LEABHARLANN CHOLÁISTE NA TRÍONÓIDE, BAILE ÁTHA CLIATH | TRINITY COLLEGE LIBRARY DUBLIN
Ollscoil Átha Cliath | The University of Dublin

Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

ON THE MOLECULAR GENETICS OF BIPOLAR DISORDER



A thesis submitted to the University of Dublin for the degree of
Doctor of Philosophy
by

Carmel Kealey M.Sc.

Trinity College,
University of Dublin.

February, 2003

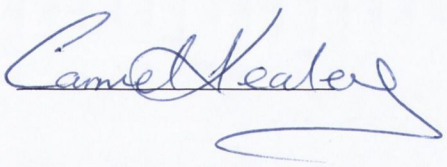
TRINITY COLLEGE
13 OCT 2003
LIBRARY DUBLIN
THESIS
~~7498~~
7374

This thesis is dedicated to my wonderful parents.
Their unwavering support, encouragement
and belief in me has enabled the
completion of this work.
Thank you both.

DECLARATION

I hereby certify that this thesis submitted for the degree of Doctor of Philosophy to the University of Dublin, has not been previously submitted to this or any other university for examination for a higher degree. The work presented here is entirely my own except where noted.

This thesis may be made available for consultation within the university library. It may be photocopied or lent to other libraries for the purpose of consultation.



Carmel Kealey M.Sc.

February, 2003

ABSTRACT

A modified genome scan, concentrating primarily on gene rich regions of the genome was employed by this group to search for regions harboring susceptibility loci for bipolar disorder in a panel of 48 Irish sib-pair families (Murphy *et al.,* 2000). Results from this analysis revealed six independent regions on chromosomes 17, 16, 14, 5, 1 and 9 that produced linkage findings considered worthy of further investigation. Based on the strength of the finding and evidence from previously conducted linkage studies, regions implicated on 17, 14 and 5 were selected for further investigation in this study. These regions were saturated with appropriate microsatellite markers and non-parametric linkage analysis was performed using GENEHUNTER v 2.0. These analyses revealed that the incorporation of flanking markers for the implicated regions on chromosomes 17q11-q12 and 14q22-q24 further supported the original findings thereby reducing the possibility of the original findings being false positives. The addition of flanking markers for the suggested region on 5q, however, resulted in the disappearance of support for a susceptibility region on chromosome 5 in this sample, possibly indicating the presence of a false positive in the original finding. In addition to the traditional non-parametric sib-pair linkage analysis, it was decided to perform a two-locus analysis whereby the evidence for linkage on chromosome 14 was used to reassess the evidence for linkage on chromosome 17. It is hypothesised that some of the difficulties encountered in replicating results obtained in genome scans for complex disorders, may be alleviated by conducting analyses to identify potential interactions. It is further proposed that analyses that allow for the contributions of multiple susceptibility loci may be critical to the successful positional cloning of genes for complex disorders.

To further investigate the suggestive linkage findings on 17q and 14q, it was decided to test for association using the candidate gene approach. As there is limited published data suggesting plausible candidate genes for psychiatric disorders on 14q, various web-based tools were employed to search the implicated region for genes that were either directly or indirectly involved in neurotransmitter activities and displayed significant expression in the brain. The GTP cyclohydrolase I (GCHI) gene was identified as a feasible candidate for application to association studies. Failing to find suitable SNPs

for use in these studies, the coding regions, 5'UTR and approximately 1000bp of the promoter region were screened for the presence of polymorphisms in a sample of 30 unrelated individuals using denaturing high performance liquid chromatography technology. Four variants were detected, three in the promoter region and one in the first intron of the gene. Upon elucidation of the polymorphisms by automated sequencing and further examination, it was found that the variant G/A at position –959bp in the promoter region was suitable for employment in a family based association study between the GCHI gene and bipolar disorder. Preferential transmission of the variant A allele was detected.

The region implicated in chromosome 17q harbors the serotonin transporter, a long hypothesised functional candidate for involvement in bipolar disorder. This gene possesses three published polymorphisms, extensively examined by independent groups for association with bipolar disorder. This study investigated each of the three polymorphisms in both a family based association study and a case-control study. The family trio study revealed evidence of an association between bipolar disorder and the short allele of the promoter polymorphism, known to have reduced transcriptional activity. There was also a tendency for excess transmission of the T allele of the 3'UTR G/T polymorphism in BPI cases compared to the controls. Multimarker haplotypes were also constructed with the most significant association observed being for a two-marker combination of 5-HTTLPR and the 3'UTR G/T. Excess transmission of the haplotype s-T was detected, reduced transmission of haplotype 1-G, and approximately equal transmission/nontransmission of 1-T and s-G. Attempts to replicate these findings in a case-control sample did not suggest that any single marker was associated with bipolar disorder nor was there any indication of associated haplotypes.

An additional functional candidate gene, COMT, was investigated within the context of furthering a previous association detected between bipolar disorder and female probands within the Irish population. Association analyses were performed on four single nucleotide polymorphisms in addition to the previously examined functional polymorphisms in an extended trio and combined trio and case-control cohort. Haplotype analysis was also presented as well as SDT analysis, which enabled the inclusion of families of less than ideal structure. Observations from this study continued to support a tentative association between the functional val/met polymorphism and

bipolar disorder among the BPI female probands. Furthermore, results were indicative of an association between the exon 6 insertion/deletion polymorphism and bipolar disorder within this same cohort. There was also evidence to suggest that a bipolar haplotype involving the exon 6 polymorphism and the functional polymorphism was identifiable within this region.

The candidate gene approach was also applied to the estrogen receptor beta (ERβ) gene using a family trio cohort. Due to suggestive preliminary linkage data for 14q22-24, this gene, which is located at 14q22-q24, was considered a plausible functional candidate for bipolar disorder. The TDT test was employed to test for association in the sample as a whole and also on the basis of the gender of the proband. However, this study failed to provide evidence that genetic variation within ERβ is involved in conferring susceptibility to bipolar disorder within the Irish population.

PUBLICATIONS

Kealey, C., Reynolds, A., Mynett-Johnson, L., Claffey, E., and McKeon, P. (2001). No evidence to support an association between the oestrogen receptor beta gene and bipolar disorder. *Psychiatric Genet* 11(4):223-6.

Mynett-Johnson, L., Kealey, C., Claffey, E., Curtis, D., Bouchier-hayes, L., Powell, C., and McKeon, P. (2000). Multimarkerhaplotypes within the serotonin transporter gene suggest evidence of an association with bipolar disorder. *Am J Med Genet* 96:845-849.

Murphy, V., Mynett-Johnson, L., Claffey, E., Bergin, E., McAuliffe, M., Kealey, C., and McKeon, P. (2000). Search for bipolar disorder susceptibility loci: the application of a modified genome scan concentrating on gene-rich regions. *Am J Med Genet* 96:728-732.

ACKNOWLEDGEMENTS

I would like to sincerely thank my supervisor, Lesley, for her constant guidance and support throughout this thesis and in particular for the many hours of proof reading of this manuscript. Special thanks to Prof. Patrick McKeon, who especially in the past two years, has provided me with every support to complete this work. It is much appreciated. Also, a big thank you to Pete, for affording me with lab space and the necessary facilities to conduct my work as well as for being there if needed. I would like to thank all the members of the Ocular Genetics lab for creating a congenial, supportive working environment which I've thoroughly enjoyed being part of: Sophie, Sophia, Paul, Avril, Aileen, Niamh, Marian, Alex, Arpad, Mary, Brian, Therese, Danny, Fiona, Helena, Alison, Philip, Sarah, Sylvie, Gearóid, Denis, Mandy and Jane. Thanks also to Aisling, Alison and Connor who've worked with me in the lab. Special thanks to Dave, Paul and Brenda for all their help in the ordering department. Thanks to Simon, Kevin and Andrew for helping with the computer stuff. And to all members of the Genetics Department, it's been a pleasure to work and party alongside you all!!

I would like to thank in a special way my family and all my friends for their constant support and encouragement. You've helped make this long road considerably easier and I want you all to know how much it's appreciated. To Catriona and Dave whom I've been lucky to share this Ph.D. experience with; it was great to have you both there to moan and whine to when things weren't going to plan and then to laugh and joke about it all later!!! To all the gang from U.C.G. days and the extended circle since then, especially Aileen and David, thanks for everything. Those weekends in Galway, Kerry, Cork and Kildare provided the perfect escape many times!!!

To my parents, Pat and Monica, who've wholeheartedly supported me throughout my many, many years of study, thanks again. I'm definitely finished now!!!! And to my sister Eimear and my brothers Damien and Noel, I couldn't have done it without you. Thanks.

Last but not least I would like to thank Aware Defeat Depression for their funding and support. We are also forever indebted to all those families who have participated in these studies.

TABLE OF CONTENTS

Table of Contents i
List of Tables v
List of Figures vii
Abbreviations viii

CHAPTER ONE: GENERAL INTRODUCTION – BIPOLAR DISORDER

1.1 Bipolar Disorder - origin and concept..... 1
1.2 Diagnosis of bipolar disorder..... 3
1.3 Bipolar Disorder - Family, Twin and Adoption Studies..... 5
1.4 Bipolar Disorder - mode of inheritance..... 7
1.5 Investigations into the genetic component of Bipolar Disorder..... 8
1.6 Parametric Linkage Analysis..... 10
1.7 Non-Parametric Linkage Analysis..... 12
1.7.1 Sib-Pair Analysis..... 12
1.7.2 Affected Sib-Pair Analysis..... 13
1.7.3 Association Studies - Candidate gene approach..... 15
1.7.3.1 Principles of the association paradigm..... 16
1.7.3.2 Methods for analysing case-control studies..... 16
1.7.3.3 Methods for analysing family based studies..... 17
1.8 Further statistical considerations..... 20
1.8.1 Reporting linkage findings..... 20
1.8.2 Multiple testing..... 21
1.8.3 Power and sample size..... 22
1.9 Genome scans and complex disease..... 23
1.10 Whole genome scans and bipolar disorder..... 25
1.11 Follow-up chromosomal regions..... 26
1.11.1 Chromosome 18..... 26
1.11.2 Chromosome 21q..... 27
1.11.3 Chromosome 12q..... 27
1.11.4 Chromosome 4p..... 28
1.12 Common susceptibility regions and psychosis genes..... 28
1.12.1 Chromosome 10p..... 29
1.12.2 Chromosome 13q..... 29
1.12.3 Chromosome 22q11-13..... 30
1.12.4 Chromosome 18p..... 30
1.13 Modified genome scan..... 31
1.14 Candidate gene studies and bipolar disorder..... 33
1.14.1 Dopaminergic system..... 33
1.14.1.1 Dopamine D1 receptor (DRD1)..... 33

1.14.1.2 Dopamine D2 receptor (DRD2)..... 34
1.14.1.3 Dopamine D3 receptor (DRD3)..... 35
1.14.1.4 Dopamine D4 receptor (DRD4)..... 35
1.14.1.5 Dopamine D5 receptor (DRD5)..... 36
1.14.1.6 Dopamine transporter (DAT)..... 36
1.14.2 Serotonergic system..... 37
1.14.2.1 Serotonin transporter (5-HTT)..... 37
1.14.2.2 Serotonin receptors..... 38
1.14.3 Tyrosine hydroxylase..... 40
1.14.4 Monoamine oxidase A (MAOA)..... 41
1.14.5 Catechol-O-methyltransferase (COMT)..... 41
1.15 The future of complex genetics..... 42
1.15.1 Human genome project and its application to complex disease..... 42
1.15.2 Whole genome association (LD) studies..... 44
1.15.3 Population isolates..... 46
1.15.4 Meta-analysis..... 47
1.15.5 Endophenotypes..... 48
1.15.6 Animal models and Bipolar Disorder..... 50
1.16 Study Aim..... 51

CHAPTER TWO: MATERIALS AND METHODS

2.1 Patient sample and ascertainment procedure..... 52
2.1.1 Diagnosis..... 52
2.1.2 Irish Bipolar Patient Collections..... 52
2.1.3 Sib-Pair Family Collection..... 52
2.1.4 Case-Control Association Collection..... 53
2.1.5 Family-based Association Collection..... 53
2.2 DNA Extraction..... 54
2.3 Non-radioactive polymerase chain reaction (PCR)..... 55
2.4 Radioactive PCR..... 55
2.5 Polyacrylamide Gel Electrophoresis..... 56
2.5.1 Denaturing Acrylamide Gel Electrophoresis..... 56
2.5.2 Non-Denaturing Acrylamide Gel Electrophoresis..... 57
2.6 Agarose Gel Electrophoresis..... 57
2.7 Restriction Fragment Length Polymorphism..... 57
2.8 Serotonin transporter VNTR polymorphism..... 58
2.9 Serotonin transporter promoter polymorphism..... 58
2.10 Denaturing High Performance Liquid Chromatography (DHPLC)..... 59
2.10.1 Temperature Modulated Heteroduplex Analysis (TMHA)..... 60
2.11 PCR purification for sequencing..... 63
2.12 Automated sequencing..... 63

2.13 Statistical and Linkage Packages 63
2.13.1 Genehunter Version 2.0 64
2.13.2 Two Locus Analysis 64
2.13.3 NNLD 65
2.13.4 Transmission Disequilibrium Test (TDT) 65
2.13.5 HHRR, GHRR and MGRR 66
2.13.6 Extended Transmission Disequilibrium Test (ETDT) 66
2.13.7 Sibship Disequilibrium Test (SDT) 67
2.13.8 EH 67

CHAPTER THREE: FOLLOWING UP ON PREVIOUS LINKAGE FINDINGS FROM A MODIFIED GENOME SCAN PERFORMED USING AN IRISH COHORT

3.1 Introduction 69
3.1.1 Linkage Analysis 69
3.1.2 Affected Sib-Pair Analysis 69
3.1.3 Regions of interest 70
3.1.4 Two-locus analysis 74
3.2 Results 76
3.2.1 Chromosome 17 76
3.2.2 Chromosome 14 78
3.2.3 Chromosome 5 79
3.2.4 Two-locus analysis 80
3.3 Discussion 82
3.3.1 Chromosome 17 82
3.3.2 Chromosome 14 83
3.3.3 Chromosome 5 84
3.3.4 Two-locus analysis 85

CHAPTER FOUR: INVESTIGATIONS OF GTP CYCLOHYDROLASE I (GCHI) – A PLAUSIBLE CANDIDATE GENE FOR BIPOLAR DISORDER?

4.1 Introduction 87
4.1.1 Previous linkage analysis 87
4.1.2 Screening region of interest for potential candidate genes 87
4.1.3 GTP Cyclohydrolase I (GCHI) - A functional candidate for bipolar disorder? 88
4.1.4 Detecting DNA sequence variation 91
4.1.5 Detection of polymorphisms using Denaturing High Performance Liquid Chromatography (DHPLC) 92
4.2 Results 94
4.2.1 Searching for candidate genes on 14q22-24 using Web based tools 94

4.2.2 Investigation of GCHI using Web based tools..... 96
4.2.3 WAVE analysis of the GCHI gene to detect novel polymorphisms..... 98
4.2.4 RFLP analysis of novel C/G polymorphism at position +393bp (Intron1)..... 101
4.2.5 RFLP analysis of novel G/A polymorphism at position -959bp (Promoter)..... 102
4.2.6 TDT based association study using novel SNP..... 103
4.3 Discussion..... 104

CHAPTER FIVE: FURTHER STUDIES TO EXAMINE THE INVOLVEMENT OF THREE WELL-RECOGNISED CANDIDATE GENES FOR BIPOLAR DISORDER, IN THE IRISH POPULATION

5.1 Introduction..... 107
5.1.1 Candidate gene association approach..... 107
5.1.2 Serotonin transporter (5-HTT)..... 109
5.1.3 Catechol-O-methyltransferase (COMT)..... 112
5.1.4 Estrogen receptor beta (ERβ)..... 114
5.2 Results..... 117
5.2.1 Serotonin transporter (5-HTT)..... 117
5.2.2 Catechol-O-methyltransferase (COMT)..... 121
5.2.3 Estrogen receptor beta (ERβ)..... 126
5.3 Discussion..... 129
5.3.1 Serotonin transporter (5-HTT)..... 129
5.3.2 Catechol-O-methyltransferase (COMT)..... 131
5.3.3 Estrogen receptor beta (ERβ)..... 135

CHAPTER SIX: CONCLUSIONS AND FUTURE PROSPECTS

6.1 Conclusions..... 136
6.2 Future prospects..... 139

BIBLIOGRAPHY 142

APPENDIX A Stock solutions..... 171
APPENDIX B B.1 Chromosome 17q microsatellite markers..... 175
B.2 Chromosome 14q microsatellite markers..... 176
B.3 Chromosome 5q microsatellite markers..... 176
B.4 Gene polymorphism primer sequences (Association studies)..... 177
B.5 Primer sequences for DHPLC and sequence analysis..... 178

LIST OF TABLES

Table 1.1 Classification of Bipolar Disorder..... * 2
Table 1.2 DSM IV criteria that must be reached for a patient to be diagnosed with a depressive or manic episode..... * 4
Table 1.3 Approximate life-time rates of mood disorder in various classes of relative of bipolar probands..... 6
Table 1.4 Published whole genome scans for bipolar disorder..... * 25
Table 1.5 Généthon poly-CA markers which selectively screen EST rich regions of the human genome..... 32
Table 1.6 Results for six chromosomal regions showing positive NPL scores from a modified genome scan (Murphy *et al.*, 2000)..... * 32
Table 2.1 Family composition and diagnostic details of the sib-pair families under the broad and narrow affection models..... * 52
Table 2.2 Details of Restriction Fragment Length Polymorphisms..... * 57
Table 2.3 Column temperatures and % buffer B selected for each region of the GCHI gene analysed to ensure optimal conditions for mutation detection..... 62
Table 3.1 NPL scores and associated p values obtained for chromosome 17 using GENEHUNTER V2 under the broad and narrow affection models..... 77
Table 3.2 NPL scores and associated p values obtained for chromosome 14 using GENEHUNTER V2 under the broad and narrow affection models..... 78
Table 3.3 NPL scores and associated p values obtained for chromosome 5 using GENEHUNTER V2 under the broad and narrow affection models..... 79
Table 3.4 Family NPL trend across 14q22-24 under the narrow model only..... 80
Table 3.5 2 point NPL score for each family at D14S281 and the weighting applied under both models of epistasis and heterogeneity..... 81
Table 3.6 Original and weighted 2 point NPL scores for each of the chromosome 17 markers based on the evidence for linkage at D14S281..... * 81
Table 4.1 Known genes found in the region of interest 14q21.1-q24.3..... * 95
Table 4.2 TDT result for novel SNP (G/A, at position -959), detected in the promoter region of GCHI, in a sample of 93 Irish BPI trios..... 103
Table 5.1 Pairwise analysis for each of the hSERT polymorphisms..... 118
Table 5.2 ETDT analysis of multimarker haplotypes for hSERT..... * 118
Table 5.3 Pairwise SDT analysis of individual hSERT polymorphisms..... 119
Table 5.4 Output of NNLD programme for case-control hSERT sample..... 120
Table 5.5 Genotypic and allelic frequencies for the combined data set (all trio and case-control hSERT data; n=200)..... * 120
Table 5.6 Details of SNPs found within the COMT gene..... 121
Table 5.7 χ^2 tests for genotype and allele frequencies (transmitted vs non-transmitted) for COMT polymorphisms..... 121

*** Indicates that the relevant Table is located on the following page**

Table 5.8 Extended TDT analysis - Transmission of haplotypes from heterozygous parents to probands for all possible COMT SNP combinations..... 123
Table 5.9 TDT results for expanded female COMT trio sample..... 123
Table 5.10 Genotypic and allelic frequencies for the combined COMT female data set (n = 82)..... 124
Table 5.11 Pairwise SDT analysis for all COMT family data (n = 73)..... 125
Table 5.12 Linkage disequilibrium analysis - χ^2 tests of linkage disequilibrium between each of the single nucleotide polymorphisms within the COMT gene..... 126
Table 5.13 TDT results for the ERβ gene from the sample as a whole and from the sample divided by proband gender..... * 127
Table 5.14 Transmission of alleles in 13 multiplex families to affected offspring, from parents heterozygous for the A/G polymorphism within the ERβ gene..... 128

*** Indicates that the relevant Table is located on the following page**

LIST OF FIGURES

Figure 2.1 Microsatellite D14S271 (dinucleotide repeat)..... * 56
Figure 2.2 Microsatellite D17S1299 (tetranucleotide repeat)..... * 56
Figure 2.3 Wavemaker graphical output illustrating the melting behaviour of a fragment..... 60
Figure 2.4 Wavemaker graphical output illustrating the recommended temperatures for the different melting domains of the fragment..... 61
Figure 3.1 NPL multipoint analysis for chromosome 17q..... * 77
Figure 3.2 NPL multipoint analysis for chromosome 14q..... * 78
Figure 3.3 NPL multipoint analysis for chromosome 5q..... * 79
Figure 4.1 Schematic presentation of the biosynthesis of BH_4 and its use as a cofactor for the hydroxylation of L-Tyrosine and L-Tryptophan in the biosynthesis of the neurotransmitters L-dopa and serotonin..... * 88
Figure 4.2 Formation of homoduplex and heteroduplex DNA molecules following denaturation and reannealing of DNA strands..... 92
Figure 4.3 Ensembl *contigview* of 14q22.2-22.3, highlighting the microsatellite markers and the Ensembl known and predicted genes for this position on chromosome 14..... 94
Figure 4.4 Diagrammatic output from the human genome browser for GCHI..... 96
Figure 4.5 Structure of GTP Cyclohydrolase I gene..... * 97
Figure 4.6 Nucleotide sequence of the 5' region of the human GCHI gene..... * 97
Figure 4.7 Regions of the GCHI gene screened for novel polymorphisms using DHPLC..... 98
Figure 4.8 Wave chromatograms obtained for Exon 1 (507bp fragment) of the GCHI gene and the corresponding sequence patterns..... * 99
Figure 4.9 Wave chromatograms obtained for the promoter (region 3) of the GCHI gene and the corresponding sequence patterns..... * 99
Figure 4.10 Wave chromatograms obtained for the promoter (region 4) of the GCHI gene and the corresponding sequence patterns..... * 99
Figure 4.11 C/G SNP, at +393bp in Intron 1 (GCHI gene), resolved using RFLP analysis..... 101
Figure 4.12 G/A SNP, at –959bp in the promoter region (GCHI gene), resolved using RFLP analysis..... 102
Figure 5.1 A/G SNP at bp 1730 in 3'UTR of exon 8 (ERβ gene), resolved using RFLP analysis..... 126
Figure 5.2 Correction formula for non-independence of sibs..... 128

*** Indicates that the relevant Figure is located on the following page**

ABBREVIATIONS

APM	Affected pedigree member
APOE	Apolipoprotein E
ASM	Allele sharing method
ASP	Affected sib pair
ATG	Methionine
BDNF	Brain-derived neutrophic factor
BH4	Tetrahydrobiopterin
BP	Bipolar Disorder
bp	Base pair
BPI	Bipolar Disorder I
BPII	Bipolar Disorder II
BRCA 1	Breast cancer 1
BRCA 2	Breast cancer 2
cDNA	Complementary DNA
CEPH	Centre d'Etude du Polymorphisme Humain
cM	Centimorgan
COMT	Catechol-O-methyltransferase
CYP19	Aromatase
DAT	Dopamine transporter
DHPLC	Denaturing High Performance Liquid Chromatography
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
DRD	Dopa-responsive dystonia
DRD2	Dopamine receptor D2
DRD3	Dopamine receptor D3
DRD4	Dopamine receptor D4
DRD5	Dopamine receptor D5
DRDI	Dopamine receptor D1
DSM-IV	Diagnostic and Statistical Manual, 4th Edition
ERα	Estrogen receptor alpha
ERβ	Estrogen receptor beta
EST	Expressed sequence tag
ETDT	Extended transmission disequilibrium test
FTP	File transfer programme
GCHI	GTP Cyclohydrolase I
GHRR	Genotype haplotype relative risk
G-olfα	Guanosine triphosphate (GTP)-binding protein alpha
GRR	Genotype relative risk
GTP	Guanosine triphosphate

HGMP	Human Genome Mapping Project
HHRR	Haplotype haplotype relative risk
HPLC	High Performance Liquid Chromatography
HRR	Haplotype relative risk
HTGS	High Throughput Genomic Sequence
IBD	Identical by descent
IBS	Identical by state
ICD10	International Statistical Classification of Diseases-10^{th} revision
IMPA2	Myoinositol monophosphatase
LD	Linkage disequilibrium
LOD	Log of odds
MAOA	Monoamine oxidase A
MDD	Major depressive disorder
MFP	Multifactorial Polygenic Model
MGRR	Matched genotype relative risk
MLS	Maximum lod score
MSP	Multiple Scan Probability
NIDDM1	Non-insulin dependent diabetes mellitus
NIMH	National Institute of Mental Health
NNLD	Artificial neural network
NOS	Not otherwise specified
NPL	Non-parametric lod score
NT	Non-transmitted
OCD	Obsessive compulsive disorder
OMIM	Online Mendelian Inheritance in Man
OPCRIT	Operational Criteria Checklist for Psychotic Illness
PCR	Polymerase chain reaction
PFKL	Phosphofructokinase locus
PTPS	6-pyruvoyl-tetrahydrobiopterin synthase
QTL	Quantitative Trait Loci
RDC	Research Diagnostic Criteria
RDS	Retinal degeneration slow
RFLP	Restriction fragment length polymorphism
rpm	Revolutions per minute
RT-PCR	Reverse transcriptase polymerase chain reaction
RUD	Recurrent unipolar disorder
SA	Schizoaffective
SADS	Schedule of Affective Disorders and Schizophrenia
SAGE	Serial analysis of gene expression
SDS	Sodium dodecyl sulphate
SDT	Sibship disequilibrium test

SNP	Single nucleotide polymorphism
SR	Sepiapterin reductase
SSCPE	Single stranded conformational polymorphism electrophoresis
SZP	Schizophrenia
TDT	Transmission disequilibrium test
TH	Tyrosine hydroxylase
TMHA	Temperature Modulated Heteroduplex Analysis
TPH	Tryptophan hydroxylase
UTR	Untranslated region
VCFS	Velocardiofacial syndrome
VNTR	Variable number tandem repeat
w/v	Weight per volume
θ	Recombination rate
χ^2	Chi-Square
5-HT	Serotonin
5-HT1D (β)	Serotonin receptor
5-HT2A	Serotonin receptor
5-HT2C	Serotonin receptor
5-HT3A	Serotonin receptor
5-HT5A	Serotonin receptor
5-HT6	Serotonin receptor
5-HTT	Serotonin transporter
5-HTTLPR	Serotonin transporter promoter polymorphism

CHAPTER ONE

GENERAL INTRODUCTION
BIPOLAR DISORDER

1.1 Bipolar Disorder – origin and concept

The origin and concept of bipolar disorder is rooted in the work and views of the Greek physicians of the classical period. Hippocrates (460-337 BC) formulated the first classification of mental disorders, namely melancholia, mania and paranoia. The classical perceptions of melancholia and mania were broader than modern concepts and were comprised of a spectrum that included mixed states, schizoaffective disorders, some types of schizophrenia, some types of acute organic psychosis and 'atypical' psychosis in addition to melancholia and mania (Angst and Maneros, 2001). Aretaeus of Cappadocia (50–130AD) was the first to explicitly link melancholia and mania and therefore conceive the idea of bipolarity. In 1851, Falret described for the first time a separate entity of mental disorder termed 'Folie Circulaire', which was characterised by a continuous cycle of depression, mania and free intervals of varying length. While Falret's concept was met with critical opposition, by 1884 it had finally found general acceptance. Emil Kraeplin (late 1800's – early 1900's), known as the 'Father of Modern Psychiatry', contributed enormously to the understanding, diagnosis and prognosis of manic-depressive illness. His unification of all affective disorders within the concept of manic-depressive illness, however, caused strong opposition. Wernicke, Kleist and Leonhard were all vehement supporters of subdividing and differentiating the different kinds of affective syndromes. One of the most important aspects of their system, namely the unipolar/bipolar distinction, remained largely unrecognised by international psychiatry and Kraeplin's unitary concept was adopted worldwide. In 1966, two independent publications by Angst and Perris presented support for the differentiation between unipolar and bipolar disorders. As a result of their work the concepts of Falret and Baillarger were replicated, completed and developed, whilst also incorporating essential aspects of the work of Wernicke, Kleist and Leonhard, thus renewing the idea of bipolar disorders.

Today, bipolar mood disorder (BP), also known as bipolar affective disorder or more informally as manic-depression, is thought to be one of the most common, present-day psychiatric illnesses with a lifetime prevalence in the general population of approximately 1%. The disorder has been defined by the American Psychiatric Association: 'Bipolar disorder (Manic-depression psychosis) is a common serious illness characterised by two types of relapse, mania (elation with disinhibition, over active behaviour) and depression'. It is a complex disorder in which the core feature is

pathological disturbance in mood, ranging from extreme elation or mania to severe depression usually accompanied by disturbances in thinking and behaviour, which may include psychotic symptoms, such as delusions and hallucinations (Craddock and Jones, 1999). A notable feature of bipolar disorder is that it is typically an episodic illness with full recovery between episodes. All modern classifications of bipolar disorder such as ICD-10 and DSM IV (table 1.1) require that an episode of mania has to occur at some time during the course of the illness thus distinguishing bipolar disorder from unipolar disorder in which subjects suffer one or more episodes of depression without ever experiencing episodes of pathologically raised mood. In DSM IV, bipolar disorder is sub-classified into BPI disorder, which features either one or more manic or mixed episodes and at least one major depressive episode, BPII disorder, which is characterised by one or more episodes of major depression and at least one hypomanic (a less severe elation) episode, cyclothymic disorder and bipolar not otherwise specified (NOS). Schizoaffective disorder is categorized in a separate group (Marneros, 2001). ICD-10 defines two very global groups of affective disorders, namely the bipolar affective disorders and the bipolar schizoaffective disorders, defined according to the current episode.

Table 1.1: Classification of Bipolar Disorder.

Panel 1: DSM IV classification of bipolar disorder

DSM IV category	Criteria	Course specifiers and examples
Bipolar I disorder	One or more manic or mixed episodes, Usually accompanied by one or more Major depressive episodes	To describe current or more recent episode: mild, moderate, severe without psychotic features Severe with psychotic features In partial or full remission With catatonic features With postpartum onset To describe current (or most recent) major depressive episode: Chronic With melancholic features With atypical features To describe pattern of episodes: With or without full interepisode recovery With seasonal pattern With rapid cycling (≥4 episodes in previous 12 months)
Bipolar II disorder	Recurrent major depressive episodes with one or more hypomanic (milder than manic) episodes	To describe current (or more recent) episode Hypomanic Depressed To describe current (or more recent) major depressive episode and pattern of episodes: see bipolar I disorder
Cyclothymic disorder	Chronic (>2yrs), fluctuating mood disturbance, involving numerous periods of mild hypomanic and depressive symptoms that do not meet criteria for a major depressive episode.	Over 2 years any symptom free intervals last no longer than 2 months
Bipolar disorder not otherwise specified	Disorders with bipolar features that do not meet criteria for any specific bipolar disorder	Examples: Very rapid cycling (over days) Recurrent hypomanias without depressive symptoms Indeterminate whether primary or secondary (due to a general medical condition or substance abuse)

DSM IV = Diagnostic and Statistical Manual of Mental Disorders, 4th edn.

Adapted from Muller-Oerlinghausen *et al.*, 2002

Bipolar disorder has been found to occur nearly equally in men and women (Smith and Weissman, 1992) and it is also noted to be more frequent in creative people, such as artists and their relatives, than in normal controls. Scores of influential 18th and 19th century poets, artists and composers were thought to have suffered from bipolar disorder: Alfred Lord Tennyson, William Blake, Vincent Van Gogh, Gustav Mahler, Virginia Woolf and Edgar Allen Poe to name but a few. Recent studies have found that a considerable number of established artists, far more than would be expected by chance, meet the DSM IV criteria for manic depression or major depression (Jamison, 1995). An alarming element of this disease is that approximately 10-20% of individuals with bipolar disorder take their own life, and nearly one-third of patients admit to at least one suicide attempt (Muller-Oerlinghaisen *et al.,* 2002). It is therefore quite clear that bipolar disorder is an extremely debilitating and destructive disease both for patients and their families. There are at present effective therapies for manias and depressions. However, these drugs can often dampen a person's general intellect and limit his/her emotional and perceptual range. For this reason, many patients with manic depression can stop taking the drugs often with very serious consequences. It is therefore imperative that a better understanding of the complex biological basis for mood disorders is achieved so that more specific and less problematic therapies can be developed.

1.2 Diagnosis of bipolar disorder

Ideally for research purposes a balance is required that is often quite difficult to accomplish: well-characterised diagnostic entities that are narrow enough to be reliably associated with specific susceptibility genes and yet broad enough to occur at a frequency suitable for genetic studies. It has been proposed that misclassification could result in a reduction in the evidence for linkage. Typically, genetic studies have attempted to define a phenotypic spectrum that minimizes false positives and it appears that a spectrum limited to schizoaffective-bipolar, BPI and BPII and recurrent unipolar depressive disorders is the most appropriate, while studies using BPI only or those employing a more liberal spectrum generally result in a lowering of LOD scores (logarithm of odds).

As previously mentioned standardised methods or tools exist for diagnosing psychiatric disorders. Recognised standards include DSM IV (The Diagnostic and Statistical Manual of Mental Disorders, 4th edition), and ICD-10. These manuals provide information on the criteria that must be met in order for a patient to be diagnosed with a certain disorder (table 1.2). As diagnostic uncertainty has plagued collaborative studies and the replication and comparison of results, these methods of diagnosis have assisted, at least from a diagnostic perspective, in ensuring standardization amongst most studies. However, problems still remain. As already mentioned, because of the nature of bipolar disorder in addition to the variable age of onset, an individual who is initially diagnosed as unaffected may subsequently change in diagnosis to affected. Given that a change in diagnosis can have a dramatic effect on the genetic linkage results (Egeland *et al,* 1987; Kelsoe *et al,* 1989), it is therefore essential that families are regularly assessed for possible alterations in diagnosis.

Since the 1970's there has been an ever-increasing propensity to broaden the concept of bipolar disorder. Hypomanic episodes, for example, have been reported to occur quite commonly in the general population and BPII has been discovered in many patients originally diagnosed as having recurrent depressions (Hantouche *et al,* 1998). Pioneering research performed by Akiskal and colleagues has provided evidence for the identification of soft spectrum forms of bipolar disorder in the general population and within bipolar families (Akiskal *et al,* 2000). Undoubtedly, the recognition of a broader spectrum of bipolar disorders is paramount to improving case recognition and to the provision of the most effective treatment. However, this expansion of the classification of bipolar disorder may have serious consequences for genetic studies. Broadening the spectrum would select for increased genetic heterogeneity and complexity, the net effect being that more treatment resistant cases would be identified and consequently more heterogeneous patient samples would be included in research studies. Therefore, there is justification for upholding the integrity of the bipolar concept (Baldessarini, 1970) and for maintaining stringent diagnostic boundaries of bipolar disorder in research.

Table 1.2: DSM IV criteria that must be reached for a patient to be diagnosed with a depressive or manic episode (adapted from Goodwin and Guze, 1995).

Criteria for Major Depressive Episode	Criteria for Manic Episodes
A Five or more of the following symptoms have been present during the same 2-week period and represent a change from the previous functioning (i) Depressed mood most of the day, nearly every day, and indicated by either subjective report or observation made by others (ii) Marked diminished interest or pleasure in all. or almost all, activities of the day, nearly every day (iii) Significant weight loss when not dieting or weight gain, or decrease or increase in appetite nearly every day (iv) Insomnia or hypersomnia nearly every day (v) Psychomotor agitation or retardation nearly every day (vi) Fatigue or loss of energy nearly every day (vii) Feelings of worthlessness or excessive or inappropriate guilt nearly every day (viii) Diminished ability to think or concentrate or indecisiveness, nearly every day (ix) Recurrent thoughts of death, recurrent suicidal ideation without a specific plan, or a suicide attempt or a specific plan for committing suicide **B** The symptoms cause clinically significant distress or impairment in social, occupational, or other important areas of functioning **C** The symptoms are not due to the direct physiological effects of a substance or general medical condition **D** The symptoms are not better accounted for by bereavement, the symptoms persist for longer than 2 months and are characterised by marked functional impairment, morbid preoccupation with worthlessness, suicidal ideation, psychotic symptoms or psychomotor retardation	**A** A distinct period of abnormally and persistently elevated, expansive or irritable mood , lasting at least one week **B** During the period of mood disturbance, three (or more) of the following symptoms have persisted and have been present to a significant degree; (i) Inflated self-esteem or grandiosity (ii) Decreased need for sleep (iii) More talkative than usual or pressure to keep talking (iv) Flight of ideas or subjective experience that thoughts are racing (v) Distractability (i.e. attention too easily drawn to unimportant or irrelevant external stimuli) (vi) Increase in goal directed activity or psychomotor agitation (vii) Excessive involvement in pleasurable activities that have a high potential for painful consequences **C** The mood disturbance is sufficiently severe to cause marked impairment in occupational functioning or in usual social activities or relationships with others, or to necessitate hospitalisation to prevent harm to self or others, or there are psychotic features **D** The symptoms are not due to the direct physiological effects of a substance or general medical condition

1.3 Bipolar Disorder – Family, Twin and Adoption Studies

There has always been an element of controversy when considering the concept of there being a genetic component to psychiatric conditions such as bipolar disorder. The genetic hypothesis predicts that relatives of ill probands will be at a greater risk of developing the disorder than the relatives of the control probands and that the risk to relatives of ill probands will diminish as the amount of genes they have in common with the proband decreases. The relative risk is defined as the ratio of risk of bipolar disorder in the first-degree relatives of bipolar probands to the risk in first-degree relatives of controls or in studies where a control group was not included, to the assumed general population baseline risk of 1%.

Many of the early studies of mood disorders, a review of which can be found in Tsuang and Faraone (1990), are not particularly accurate in their assessment of relative risk as they do not employ a clear description of the clinical features associated with the diagnostic categories used. However, over the last 30 years, about 20 studies have been conducted which have clearly employed the modern understanding of bipolar illness. Qualitatively, all these studies indicate a marked increase in life-time risk of bipolar disorder in the first degree relatives of the proband. Quantitative results, while varying among studies consistent with methodology and method of ascertainment employed, estimate a risk of 5-10 times that of the general population (Craddock and Jones, 1999). This finding is comparable to the relative risk to first degree relatives established for other complex diseases: for multiple sclerosis the accepted relative risk to first degree relatives is 3-5% and for rheumatoid arthritis it is approximately 4%. Family studies have also consistently found that first-degree relatives of bipolar probands are also at a significantly increased risk of developing unipolar major depression and are at an increased risk of developing BPII if the proband suffers from BPI. The sex of probands or relatives does not appear to influence lifetime risk, but early age of onset and the number of ill relatives do seem to have an effect on the life-time risk to relatives. At present, there are no adequate data to provide meaningful estimates of the rates of affective disorders in second-degree relatives of bipolar probands, but available evidence would suggest that the rates lie between those for first-degree relatives and those for the general population (Craddock and Jones, 1999).

The well-established, elevated relative risk to relatives of bipolar probands clearly reflects the familial nature of both bipolar disorder and unipolar affective disorders. However, family studies cannot definitively substantiate the involvement of a genetic component in the manifestation of bipolar disorder, as generally family members share a common environment. Twin and adoption studies are therefore required to delineate the effects of nature and nurture. Early twin studies, which are reviewed in Tsuang and Faraone (1990), did not distinguish between bipolar disorder and unipolar disorder but they did support the involvement of genes in broadly defined mood disorders. Six twin studies, all using the modern concept of bipolar disorder, have shown an increased risk in monozygotic (identical) co-twins compared with dizygotic (fraternal) co-twins of a proband with bipolar disorder (Kringlen, 1967; Allen *et al.,* 1974; Bertelsen *et al.,* 1977; Torgersen, 1986; Kendler *et al.*, 1993; Cardno *et al.,* 1999) (table 1.3).

Table 1.3: Approximate life-time rates of mood disorder in various classes of relative of bipolar probands (adapted from Craddock and Jones, 1999).

Degree of relationship to bipolar proband	Risk of bipolar disorder (%)	Additional risk of unipolar depression (%)
Monozygotic co-twin	40 – 70	15 – 25
First degree relative	5 – 10	10 – 20
General population (i.e. unrelated)	0.5 – 1.5	5 - 10

Again most risk studies estimate the risk to monozygotic co-twins to be 45-75 times that of the general population. Moreover, twin studies indicate that very few identical co-twins of an individual with bipolar disorder escape all affective morbidity; if not developing BPI, they will have some form of affective disorder during their lifetime thereby providing robust evidence that genes play an important role in bipolar disorder. However, the fact that not all monozygotic twins are concordant for narrowly defined bipolar disorder indicates that factors other than the genetic component, namely environmental factors, are important in the manifestation of this disease.

Heritability is a measure of phenotypic variance that can be attributed to genetic influences and concordance rates for monozygotic twins infer a heritability of 80% for bipolar disorder (McGuffin & Katz, 1986). A later study performed by McGuffin and

Katz (1989) implied that the contribution of genetic factors was greater than 80%, in comparison to 7% for shared and 7% for non-shared environmental factors.

Adoption studies can be employed to investigate further the contribution of environmental factors to the disorder. Adoption studies, by virtue of their nature are particularly difficult to perform and to date only two adoption studies have been carried out which employ the modern concept of bipolar disorder. Mendlewicz and Rainer (1977) investigated the genetic element in bipolar disorder using a study group that consisted of the biological and adoptive parents of 29 bipolar and 22 normal adoptees and the biological parents of 31 bipolar non-adoptees. They found that psychopathology in the biological parents was in excess of that found in the adoptive parents of the same manic-depressive offspring, and when focussing on the affective spectrum disorders, found that the difference was significant at the level of $p<0.025$ ($\chi^2 = 5.10$). They also eatablished, that the risk in the biological relatives of bipolar adoptees was comparable to that in the biological relatives of bipolar non-adoptees. A smaller study carried out by Wender *et al.,* (1986) also indicated that there was an increased risk for biological relatives of probands as compared with adoptive relatives.

1.4 Bipolar Disorder - mode of inheritance

Family, twin and adoption studies undoubtedly provide an impressive and reliable body of evidence supporting the existence of genes determining predisposition to bipolar disorder. However, it is obvious that this disease does not conform to Mendelian patterns of inheritance i.e either dominant or recessive. When a disorder displays dominant inheritance, affected individuals will have at least one affected parent and the progeny of an affected individual will have a 50% chance of inheriting the disorder. In the case of recessive inheritance, affected individuals usually have two phenotypically normal parents who will be carriers of the mutant allele and therefore will have a 25% chance of producing an affected offspring. Non-conformity to Mendelian inheritance is also corroborated by the fact that psychiatric disorders are far more common than a typical Mendelian disease and also the risk to first-degree relatives compared to population frequency are far below Mendelian rates. In order to investigate the possible mode of inheritance of bipolar disorder, various models of transmission have been proposed and tested.

The first possible model is that of a Single Major Locus (SML) being solely responsible for the manifestation of the disease. The second model proposed is the Multifactorial Polygenic (MFP) model, or more commonly termed Quantitative Trait Loci (QTL), which hypothesizes that the illness is due to a combination of many small genes and environmental factors which additively or interactively contribute towards the development of the disease. The final model is termed the mixed model and is essentially a combination of the first and second models, where a major gene produces its effects against a polygenic background. These models have been assessed by either of two techniques: Prevalence analysis or Segregation analysis (discussed in Mynett-Johnson, 1996). Although some segregation analyses advocate a role for single locus inheritance (Spence *et al.,* 1995), it seems very unlikely that a single major locus accounts for the transmission of affective disorders. Results in general are not consistent with this mode of inheritance, and the relatively high prevalence of affective disorders implies the involvement of more than one locus contributing to disease susceptibility. The rapid drop off in recurrence risk from monozygotic co-twins to first-degree relatives and back to the general population (e.g. the λ values for sibs is 5-10%) is also more consistent with several interacting genes. Of course this interaction could involve one gene of relatively major effect and several with a more minor effect, or several genes all having a more modest effect or even a large number of genes of small effect. In addition to the genetic effects, there is a definite environmental influence, which further adds to the complexity of the disorder. It has also been suggested that mechanisms such as dynamic mutations, imprinting and mitochondrial inheritance could be involved in bipolar disorder and therefore should be considered in analyses (Owen and Craddock, 1996). In summation, bipolar mood disorder is a complex disease with the majority of cases probably resulting from the interaction of several genes of comparatively small effect in addition to a distinct environmental contribution of unknown magnitude.

1.5 Investigations into the genetic component of Bipolar Disorder

There are many problems inherent in attempting to identify genetic factors contributing to a complex disease such as bipolar disorder. As previously mentioned, the patterns of inheritance are unspecified and most probably multiple. There may exist unknown degrees of penetrance (*Full penetrance* is where each genotype unequivocally

determines a phenotype. However, many dominant alleles can display *incomplete penetrance* where there is an absence of symptoms in individuals known to be carrying the mutant allele), and unknown prevalence of predisposition genes in the general population. Phenocopies may exist i.e. persons presenting with bipolar disorder who do not have a genetic predisposition, which further complicates efforts to identify predisposing genes. Biological endophenotypes are heritable disease related traits, which are also present in unaffected relatives. It is possible that the present clinical phenotype of bipolar disorder may not accurately encapsulate the genetic endophenotypes that underlay this disease. Finally, it is possible that this mood disorder is most likely heterogenous and that the disease develops as a result of cumulative effects of mutations in several distinct genes, most probably interacting with one another and with other internal and external environmental risk factors.

Despite the difficulty in distinguishing the various factors contributing to the manifestation of bipolar disorder, a genetic basis has clearly been established for this disease. The next obvious step is to decide on the most appropriate means of investigating this genetic component. Essentially, there are two independent approaches being applied in genetic studies of bipolar disorder at present; the candidate gene approach and the screen for genetic linkage, both of which will be discussed in more detail (sections 1.7.3 and 1.9 respectively). The former is based either on the limited knowledge of the biochemical process involved in the aetiology of bipolar disorder or on existing evidence from linkage studies. Because of the nature of this disorder, numerous genes involved in neurotransmission and other neurophysiological processes have been proposed as candidate genes. Furthermore, genes that produce proteins that are known to be affected by the various drugs used in the treatment of this disorder, in addition to genes located in chromosomal regions implicated by classical linkage studies, are all considered as plausible candidates to warrant investigation. The latter approach involves performing genome-wide scans for evidence of genetic linkage or even linkage disequilibrium. This process involves genotyping a sample of individuals through genetic markers regularly positioned along the genome and performing linkage analysis in anticipation of finding a region that may harbour a susceptibility locus (for example, 10cM intervals screen approximately 300-400 markers). If such a region is implicated, the next step is to attempt to fine map the region and localise candidate genes. Both of these approaches are at present being extensively applied to bipolar

disorder as well as a wide variety of other complex genetic diseases including schizophrenia, Alzheimers and diabetes. Finally, other considerations include the statistical approach to be employed for the analysis and the nature of the patient sample, considerations which are linked.

1.6 Parametric Linkage Analysis

Traditionally, parametric type analyses have been employed to map Mendelian disease genes, and have met with considerable success e.g. Huntington's Chorea, Cystic Fibrosis. The logic of this type of linkage analysis is straightforward: if a marker is closely linked i.e. physically near the mutation, affected individuals are expected to have inherited the same allele from an affected parent whereas unaffected members will receive different alleles from their parents. The further away the marker locus is from the disease causing mutation, the greater the likelihood that it will be separated from the disease-causing mutation during meiotic recombination. Therefore, the rate of recombination is a measure of distance on the chromosome between the marker and the disease-causing mutation. The likelihood of obtaining the specific assembly of marker alleles and affected status in the family is compared under two hypotheses; the test hypothesis that the marker is linked to the disease at a specific distance (that reflects the recombination rate, θ), and the null hypothesis, which assumes that there is no linkage between the marker and the disease (Burmeister, 1999). Dividing these two likelihoods produces a quotient known as the likelihood ratio. The log to the base 10 of this ratio generates what is commonly referred to as the LOD score, which can be calculated by modelling in the test hypothesis any distance, θ, of the marker to the disease locus. Traditionally a LOD score of 3, corresponding to an odds ratio of 1000:1, is accepted as evidence for linkage. To increase the power and certainty with which linkage is detected multipoint analysis can be performed; several marker loci, known to be located close to each other, are simultaneously checked for linkage with the disease thereby generating a linked haplotype.

Parametric analysis has conventionally employed large multiply affected pedigrees and, as previously stated, this type of approach has proven very successful for Mendelian disorders. While many still advocate the utilisation of such densely affected pedigrees in psychiatric genetics, there are certain problems inherent in this type of sample. Linkage

results in large families can be sensitive to genotyping errors or changes in diagnosis of single individuals as illustrated by the dramatic change in LOD score for affective disorder in the well-known Amish study that resulted from previously well individuals becoming ill for the first time (Kelsoe *et al.,* 1989). This is likely to be a persistent problem in psychiatric diseases where the appearance of the first episode can extend to middle and even old age. A further problem is that large pedigrees typically display a broad spectrum of phenotypes making it extremely difficult to classify individuals on a simple affected/unaffected basis. Moreover, parametric analysis itself is problematic as it is model based i.e. it requires that the mode of inheritance is specified, but as previously discussed, segregation analysis has not identified the mode of transmission for bipolar disorder. However, there are various options available to minimise some of the inherent difficulties in applying this methodology to complex disease. Firstly, both dominant and recessive modes of inheritance are examined and various analytical tools such as affected only analysis and multiple affection models (broad and narrow definitions of affection status) are employed. Penetrance models are also applied to the pedigree and marker data, taking into account the particular age of onset of the disease and the age of each individual. To incorporate an age of onset function in the LINKAGE programme, which performs the analysis, liability classes are required in conjunction with the affection status locus type (i.e. 2 = affected, 1 = unaffected, 0 = unknown). The number of liability classes is directly related to the number of different penetrance classes required to specify the penetrance for each individual. Two-point linkage analysis is generally performed using MLINK and ILINK from the LINKAGE package version 5.1 (http://www.hgmp.mrc.ac.uk; http://linkage.rockefellar.edu).

MLINK computes LOD scores at a set of recombination fractions defined by the user, generally starting at $\theta = 0$ and calculating LOD scores in steps of 0.1 until $\theta = 0.5$ is reached (loci far apart). The object of linkage analysis is to estimate θ and then test if θ is less than 0.5, i.e. whether an observed deviation from 50% recombination is statistically significant. ILINK is also employed for two-point analysis but it does not produce LOD scores at predefined points, but rather attempts to maximise the likelihood numerically. It returns the likelihoods at the maximum likelihood estimate of the iterated parameter, in this case the recombination fraction. Incorporation of the aforementioned modifications into the analysis enables linkage to be tested under several different 'hypothetical models' and those giving the highest LOD scores are reported. A final point in regard to this type of analysis is that corrections for multiple

tests have to be made and therefore the thresholds for claiming linkage are required to be significantly higher than the traditional limits (Lander and Kruglyak, 1995).

In conclusion, large single families offer certain advantages to studying complex diseases such as bipolar disorder. The disease is likely to be caused by a more limited number of interacting genes and the penetrance and mode of inheritance can often be more clearly defined in a single pedigree. There have been several important linkage findings reported using single large families and parametric analysis (Adams *et al.*, 1998; Blackwood *et al.*, 1996; Ewald *et al.*, 1995), and as continuing advances are made in statistical analysis, it is evident that this type of approach will continue to have a significant role in identifying the genes contributing to bipolar disorder and other complex diseases.

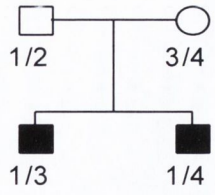
1.7 Non-Parametric Linkage Analysis

An alternative approach for the study of affective disorders and one that has gained much attention in recent times, is the use of non-parametric, or model free, methods. These methods can essentially be categorised into two main types of analysis: (i) sib pair analysis and affected pedigree member method (APM) (ii) association studies.

1.7.1 Sib-pair analysis

Sib-pair analyses are based on the idea that independent of whether the disease is dominant, recessive, or more complex, if there is a disease causing mutation in a specific chromosomal region in a high proportion of families, two individuals from the same family would be expected to share an allele of a marker locus more often than expected by chance alone (50% for siblings). The fundamental idea is that sibs are expected to share one allele IBD (*Identity by descent*: one allele is inherited from a common ancestral source) When sibs are both affected with a given disease and the analysis is done with a marker tightly linked to the disease-predisposing gene, then siblings are expected to share more than one allele IBD. For example, consider the following pedigree:

Pedigree I



It is clear that these two children share one allele IBD, allele 1. If there is no association between the disease and the marker locus, then each of the possible genotype combinations is equally likely. There are four combinations with two alleles shared IBD, 8 combinations with one allele shared IBD and 4 combinations with no alleles shared IBD. Therefore, the expected average number of alleles shared IBD is (4x2 + 8x1) / 16 = 1. However, if there is complete linkage at $\theta = 0$, and the disease is dominant, then assuming that the disease comes from the father, and is on the same haplotype with the 1 allele, the only possibilities are cases where both sons carry the 1 allele and the expected number of alleles IBD is 1.5, which can be detected by sib pair analysis of moderate sample size. Similarly, if the disease was recessive, one combination with two alleles shared IBD would be expected, which again would be detected by this type of analysis. For complex diseases however, the true model is often unknown and consequently most real situations are somewhere between these two extremes. Sib-pair analysis examines allele sharing for a large number of sib pairs, it does not require the mode of inheritance to be specified and it can detect linkage in the presence of heterogeneity (Burmeister, 1999). This method also has the advantage of being able to avoid unclear phenotypes by designing the study to include only pairs that possess a specific diagnosis.

1.7.2 Affected pedigree member method (APM)

In extended pedigrees, where some individuals are distantly related, it has been proposed that IBS (*Identity by state:* any two copies of an allele at a given locus are considered identical by state) rather than IBD relationships be examined, because for the aforementioned individuals it is not always possible to determine the inheritance status of the alleles. This nonparametric method of linkage analysis is also based on

sharing of alleles among affected pedigree members but it involves more complex mathematical calculations, details of which can be found in Weeks and Lange (1988). In addition, the reliability and robustness of the method is strongly dependent on good gene frequency estimates. Despite these difficulties, it remains a popular approach for nonparametric analysis in pedigrees with structures that are not conducive to affective sib-pair analysis.

Currently the trend has been to study smaller nuclear families, which many believe may be more useful for the investigation of complex diseases. Genotyping a panel of affected sib pairs throughout the genome in the hope of detecting susceptibility loci for complex diseases such as bipolar disorder has proved to be a powerful and efficient tool. Numerous genome scans have been undertaken in recent times using affected sib pairs and have generated interesting findings for possible susceptibility loci for bipolar disorder (details can be found in the chromosome workshop reports from the 5^{th} and 6^{th} World Congress of Psychiatric Genetics). It has been proposed that a sample size of approximately 400 sib pairs is required to achieve adequate power to detect linkage (Hauser *et al.,* 1996), which at present far outnumbers the size of the majority of studies. Despite this possible constraint, the affected sib pair method of analysis does offer several advantages over large affected pedigrees, including (i) cases are more representative of typical disease, (ii) it is possible to focus effort on collecting affected subjects who satisfy narrow and stable definitions of illness, (iii) large samples are readily obtained and (iv) linkage results are less sensitive to errors.

Genehunter (Kruglyak *et al.,* 1996), one of the common programmes applied to sib pair allele sharing data, is a powerful linkage analysis tool, capable of performing parametric and non-parametric linkage analysis, information content mapping and haplotype re-construction. The programme can compute LOD scores for moderately sized pedigrees under particular models of inheritance, allowing the user to specify allele frequencies at the disease locus and penetrances for liability classes. It also permits the user to test for linkage under genetic heterogeneity by using an admixture model (Terwilliger and Ott, 1994) to estimate the proportion of linked families, α. Alternatively, the admixture parameter, α, can be specified. For model free analysis, an NPL score is generated with its associated p value, indicating the level of the IBD sharing among the affected relatives, both for the individual pedigrees and for the sample as a whole on the basis of

information obtained from multiple markers and all genotyped relatives. Using the available information the programme can also report the most likely haplotype even in pedigrees with missing data, marking crossovers and highlighting double crossovers, potentially arising from genotyping errors. Finally, the information content mapping feature provides a measure of the fraction of the total inheritance information extracted from the marker data. The Genehunter programme facilitates efficient multipoint analysis to be performed, extracting inheritance information from all markers and all individuals, thereby substantially increasing the power to detect linkage.

1.7.3 Association studies – Candidate gene approach

The other type of non-parametric analysis is association, or linkage disequilibrium studies, whose primary application to date has been in the context of candidate gene studies. A large number of such studies using DNA markers have been published since the 1980s. Candidate gene studies involve the direct examination of one or more polymorphisms within a candidate gene. An important advantage of this approach is that the polymorphisms studied are likely to lie very close to the disease causing variation, or indeed, may be the pathogenic variant. However, for many common complex disorders and especially in the case of psychiatric illness, selection of appropriate candidate genes is constrained by a relatively poor understanding of the pathophysiology of the disease. Also linkage studies for bipolar disorder have been inconsistently replicated (Gershon *et al.,* 1998) and have not convincingly suggested any one area of the genome that merits intensive investigation with association studies. However, in the case of bipolar disorder, numerous studies have focused on neurotransmitter systems that appear to be influenced by medications employed in the management of the disorder, e.g. dopamine, serotonin, noradrenaline, for the selection of candidate genes (Alda, 2001). While no findings have unequivocally been replicated in all samples, a number of interesting findings exist which await robust validation.

Association studies together with linkage have succeeded in leading to the fine mapping of the genes involved in several simple genetic disorders such as cystic fibrosis (Rommens *et al.,* 1989) and myotonic dystrophy (Harley *et al.,* 1991), as well as implicating specific genetic loci in the pathogenesis of complex disorders e.g. the APOE in late onset Alzheimer's disease (Strittmatter *et al.,* 1993). As theories of mood

disorder biology become more sophisticated and as additional information from linkage and association scans become available, it is likely that candidate gene studies are less likely to focus on neurotransmitter-related gene sequence and more likely to concentrate on regulatory mechanisms and signal transduction pathways, such as secondary messengers.

1.7.3.1 Principles of the association paradigm

The goal of allelic association studies of disease is to demonstrate a significantly different distribution of allelic variants in affected and unaffected individuals. Association may arise due to either (a) direct involvement of the polymorphism in the pathogenesis of the disease or (b) linkage disequilibrium due to the close physical proximity between the marker and the disease locus.

The traditional approach uses a control group of unrelated individuals who are well matched to the proband group for appropriate variables such as ethnicity, age and sex. However, significant associations between marker alleles and disease may be produced artifactually by inappropriate matching between proband and control groups mostly because of unsuspecting ethnic differences (population stratification). To counteract this possibility, family based association studies have gained popularity, as they are robust to population stratification. Several methodologies have been, and continue to be, developed for this type of approach.

1.7.3.2 Methods for analysing case-control studies

Artificial Neural Network (NNLD)

Because of the presumed multigenic nature of complex traits, it may be required to identify association between disease and a group of tightly linked biallelic markers. The pattern recognition properties of artificial neural networks represent a type of statistical approach suitable to this type of analysis. The method uses a case-control sample of affected and unaffected individuals together with the multi-locus biallelic genotypes of each subject and based on the genotype data provided, the occurrence of haplotypes or partial haplotypes that confer susceptibility to disease are detected. If association exists between a haplotype and the disease, it may be possible to distinguish the affection

status of the individual based on marker genotypes and consequently output scores will tend to be significantly larger for the affected cases than for the controls. The success of the programme in discriminating case and control genotypes is also calculated by the generation of a p value that compares the output scores of the case and control genotypes.

1.7.3.3 Methods for analysing family based studies

Haplotype relative risk (HRR)

The choice of well-matched controls for genetic association studies is complicated by population stratification where different subgroups may display considerable variation in marker and disease frequency. To avoid these difficulties, Falk and Rubinstein (1987) proposed the use of the parents of affected probands, as an alternative type of control. For the design of this method, which has become known as the 'haplotype relative risk' (HRR) method, the two alleles of the parents that are not transmitted to the affected child are combined to create a pseudocontrol. The association of the disease with particular marker alleles is then assessed using traditional case control methods (Schaid, 1998). Thus the frequencies of particular marker alleles among cases and the pseudocontrols are used to derive an odds ratio (known as the 'HRR'when applied to the pseudocontrols). The HRR method is not intended as a test for linkage and standard statistical theory implies that it is only a valid test for association when population structure has been eliminated (Speilman and Ewens, 1996).

HHRR (Haplotpe Haplotype Relative Risk)

This test is a variation of the Falk and Rubenstein (1987) HRR approach. It extracts information about allelic associations by analysing haplotype-based data rather than genotypic data. The original HRR method merges together the parental homozygotes and heterozygotes. However, under the null hypothesis the two parental genotypes are independent, implying that it is possible to treat each parent as an independent observation and merely look at the fate of each parental marker allele. As HHRR uses more of the information present in the family it has been found to be more powerful than the original HRR approach (Terwilliger, 1992).

Transmission disequilibrium test (TDT)

Presently, one of the most popular methods of analysis for family based studies is the transmission disequilibrium test (TDT) (Speilman *et al.,* 1993). This method was originally designed as a test for linkage, but is also valid as a test for association in simplex families, and, unlike conventional case-control studies, it is not sensitive to population stratification. If a marker locus is considered with two alleles M and N, TDT tests for equal transmission of M and N from heterozygous parents to the affected offspring. The TDT statistic is $(b-c)^2 / (b+c)$ and the resulting statistic is treated as a χ^2 with 1 degree of freedom (df). If linkage exists between a marker and the disease, as well as allelic association, b and c will tend to differ in value and a significant difference provides evidence that the marker is linked to the disease locus (Speilman and Ewens, 1996).

Genotype analysis (MGRR and GHRR)

The frequencies of genotypes can also be compared between the case and the control groups. Almost all biological mechanisms, which would produce association with a genotype, especially a heterozygous genotype, would also produce detectable allelic association. Genotypic association is tested for using either genotype haplotype relative risk (GHRR) or matched genotype relative risk (MGRR). However, one of the major limitations of these approaches is that there is no distinction between a case (or a control) possessing one or two copies of M, and hence homozygote MM and heterozygote MN cases are given equal weight.

For MGRR analysis each case, possessing the two transmitted parental alleles are matched with a control, representing the two non-transmitted parental alleles. They are classified according to those cases where M is present in the case but absent in the control and where M is absent in the case but present in the control. When there is no population stratification, the transmitted and non-transmitted alleles are independent under the null hypothesis of no association. Therefore, the cases and controls are not required to be analysed as a matched design; rather they can be analysed as an unmatched design such as GHRR analysis. This methodology uses more information than the matched analysis thereby giving a more powerful statistical test.

Extended Transmission Disequilibrium Test (ETDT)

The ETDT package, an extension of the original TDT, performs the transmission disequilibrium test for a marker with multiple alleles as opposed to a bi-allelic marker. This allele-wise analysis attempts to establish a pattern of preferential transmission of certain alleles across genotypes, and thus implements a plausible model to test the hypothesis that different alleles may vary in the extent to which they are preferentially transmitted to affected offspring (Sham and Curtis, 1995).

Sibship disequilibrium test (SDT)

For many complex diseases, particularly those with a late age of onset, data from parents may be difficult or impossible to obtain. SDT is a programme which implements the 'sibship disequilibrium test' of Horvath and Laird (1998) for testing sibships containing affected and unaffected siblings for linkage and association with a multiallelic marker, and combines it with a TDT test based on families with at least one affected sibling and at least one heterozygous parent (Curtis *et al.,* 1999). This approach facilitates the incorporation of families of less than ideal structure into a TDT analysis, thereby increasing the data set and thus the power of the sample. The test involves examining each allele for the presence of linkage and association by using a sign test which basically entails comparing the observed number of sibships with increased frequency of the allele among affecteds, or parents demonstrating preferential transmission of the allele, to that which would be expected by chance i.e. half the number of informative sibships or parents. The outputs from the SDT programme takes the form of three independent statistics, the first relating to discordant-sibling pairs, the second is based on TDT analyses for heterozygous parents and the final output presents a statistic for the combined test comprising of results from both types of family and therefore providing an overall test for linkage and association (Curtis *et al.,* 1999).

The relative robustness to genetic heterogeneity and the ability of association studies to detect genes of modest to minor effect using feasible sample sizes ensure that they will continue to play an increasingly important role in the search for susceptibility loci for bipolar disorder. Although, in general non-parametric analyses are less powerful than parametric analysis when the correct model of inheritance is specified, it remains a powerful tool for any investigation into the genetic factors contributing to complex diseases such as bipolar disorder.

1.8 Further statistical considerations

1.8.1 Reporting linkage findings

Genetic mapping of any trait - simple or complex - essentially involves scanning the entire genome with a dense collection of genetic, or polymorphic, markers in order to determine by means of linkage analysis, those chromosomal regions that tend to be shared among affected relatives. Conceptually, this appears quite straightforward, but inherent in this type of analysis is the fact that linkage to any region will vary just by chance across an entire genome scan. It is therefore necessary to determine what constitutes a significant deviation from what would be expected under independent assortment and what threshold should be required for declaring linkage. It is apparent that if too moderate a standard is adopted the literature will be inundated with false positive claims of linkage, which will undoubtedly result in non-replication of findings thereby undermining the integrity of the field. On the other hand, if too high a standard is employed for reporting results, the likelihood of missing a true positive finding is quite high as initial genetic analyses may fall short of a strict threshold for statistical significance but may nevertheless indicate important regions worthy of further rigorous investigation. It is therefore essential that a correct balance be reached when deciding upon the criteria necessary for reporting linkage findings, thereby ensuring that a flood of false positive findings is avoided and, at the same time, true hints of linkage are not overlooked.

Geneticists have traditionally reported linkage results in one of three essentially equivalent ways: Z score, a LOD score, or a p value. The Z score represents the number of standard deviations by which allele sharing exceeds its null hypothesis of 0.50. The LOD score is the logarithm of the likelihood of the observed data to arise under a specified hypothesis of linkage than under the null hypothesis of independent assortment. The p value reflects the point wise chance of generating a statistic as large or larger than the one observed under independent assortment. Conventionally, a LOD score of 3 (odds = 1000:1) is considered to be proof of linkage and a LOD score of –2 as proof that two loci are not linked at the recombination fraction value that leads to that score. These thresholds work well for classical two point linkage studies of simple Mendelian traits. For reporting linkage findings of complex traits, Lander and Kruglyak (1995) have proposed to maintain the chance of encountering a false positive at no more than 5%. To attain this standard a threshold of $Z \geq 4.1$, $LOD \geq 3.6$ or $p = 2 \times 10^{-5}$ must

be imposed. If the threshold were any less stringent there would be a substantial chance (>5%) of reporting false linkages. They have recommended a classification for reporting linkage based on the number of times that one would expect to see a result at random in a dense, complete genome scan.

Suggestive linkage – statistical evidence that would be expected to occur one time at random in a genome scan. In the case of a sib pair study this would correspond to a point-wise significance level of 7×10^{-4} or a LOD score of 2.2.

Significant linkage – statistical evidence expected to occur 0.05 times in a genome scan (i.e. probability 5%), corresponding to a p value of 2×10^{-5} or a LOD score of 3.6 in a sib pair study.

Highly significant linkage – statistical evidence expected to occur 0.001 times in a genome scan. Again in the case of a sib pair study this would be consistent with a p value of 3×10^{-7} and a LOD score of 5.4.

Confirmed linkage – significant linkage from one or a combination of initial studies that has subsequently been validated in a further sample, preferably by an independent group of investigators for which a nominal p value of 0.01 should be required.

They also suggest that all regions with a nominal p value of $p = 0.05$ encountered in a complete genome scan be reported, but without any claims of linkage. Lander and Kruglyak's suggested criteria for reporting linkage findings has been widely adopted and has significantly facilitated the communication of results.

1.8.2 Multiple testing

Each time that a sample is tested, the chance of generating a false positive result is increased. For example, if a null hypothesis, which is in fact true, is tested using 0.05 as the critical significance level, the probability of coming to a 'not significant' (correct) conclusion is 0.95. However, if twenty such independent hypotheses are tested the probability that none will be significant will be $(0.95)^{20} = 0.36$ which generates a probability of 0.64 (1- 0.36) of getting at least one significant result i.e. increasing the risk of false positives. The easiest way to circumvent this problem is to avoid where possible performing multiple tests on the same sample. In some cases however, multiple testing is deemed necessary and worthwhile. Application of the Bonferroni correction is therefore encouraged to allow for the effect of multiple testing. The Bonferroni adjustment deflates the type I error rate, α, applied to each test, so the study–wide error

rate remains at 0.05. A true Bonferroni correction can be made using the formula $1-(1-\alpha)^{1/n}$, where n is the number of tests performed. An approximate correction can be made by multiplying the observed p value generated by the number of tests, *n*, with any *n*p less than 0.05 being considered significant. The primary advantages of the Bonferroni correction include its simplicity and its applicability to different types of data and different test statistics. However, this correction is often considered over conservative resulting in low power. It is advised that Bonferroni's corrections should only be made in situations where the number of tests is quite small.

1.8.3 Power and sample size

Risch and Merkingas (1996), evaluated the affected sib pair linkage analysis (Blackwelder and Elston, 1985) and the transmission disequilibrium test (TDT; Speilman *et al.,* 1993) with respect to the power of these methods to detect genes of modest effect. They assumed a multiplicative relationship for the genotypic relative risk (GRR), which is the increased chance that the individual with a particular genotype is affected, and calculated the sample sizes required to obtain 80% power of rejecting the null hypothesis when it is false. They estimated that when using the linkage approach and a disease gene with a GRR of 4 or greater, the number of affected sibling pairs necessary to detect linkage is 185 or 297, depending upon the frequency of the disease allele. The aforementioned numbers are sufficient provided that the allele frequency is between 5 and 75%. However, for a gene with a GRR of 2 or less the sample sizes are generally beyond reach (well over 2000), implying that genes of minor effect cannot be identified using this approach. In contrast, they calculated that the required sample sizes for the association test are vastly less than for linkage; with a GRR of 4, sample sizes of 150 or 103 are recommended and even for a GRR of 2 the proposed sample sizes are an achievable 695 or 340, again depending on the disease allele frequency. Knapp (1999) also performed power approximations for TDT and proposed that the sample size necessary to gain 80% power for TDT with singletons employing a gene with a GRR of 4 under the multiplicative mode of inheritance was 101 and under an additive mode of inheritance was 218 and for a gene with a GRR of 2, a sample size of 338 was suggested under both modes of inheritance. These calculations are in concordance with the suggestions of Risch and Merkingas and therefore have been adopted as guidelines for the numerous linkage and association studies in progress for complex traits.

Two of the largest linkage studies undertaken to date include the Wellcome Trust UK-Irish bipolar affective disorder sibling-pair genome screen (Bennet *et al.,* 2002), which contains 258 affected sib pairs (ASP), and a bipolar linkage study performed by Cichon *et al,* (2002) with 194 ASP. While these large studies attempt to meet the suggested criteria for sample size, it is recognised that the linkage study group of 83 sib pairs employed in this thesis falls short. However, this data set has been collected from a relatively homogeneous population by a single clinical team, it is continuously being expanded and it is comparable with previously published linkage data (Esterling *et al.,* 1998; Kelsoe *et al.,* 1996).

The association study data sets: case control (108) and family trio (92), are comparable to studies that have recently been published (case-control: Ospina-Duque *et al.,* 2000; Ho *et al.,* 2000; Family trio studies: Muglia *et al.,* 2002; Kirov *et al.,* 1999). These collections are also continuously being enlarged as it is widely acknowledged that greater power will be achieved in detecting genes of modest effect with a larger sample (Long and Langely *et al.,* 1999).

1.9 Genome scans and complex disease

The traditional approach to identifying disease susceptibility loci is to first identify the genetic regions within which one or more disease-predisposing genes lie and, once these have been found, to localise the genes and determine their function and biological role in the disease. The development of highly informative microsatellite markers across the genome has greatly facilitated the localisation of disease loci in Mendelian and complex diseases via both linkage and association (linkage disequilibrium) studies. Traditionally, genome wide linkage analyses would involve the testing of about 300-400 highly polymorphic markers, usually microsatellites, with an average spacing of 10cM along the genome using large multigenerational pedigrees where LOD score linkage analysis is a powerful technique for localising disease genes. This technique for genetic linkage between polymorphic markers and human diseases was first proposed by Lander and Botstein (Lander and Botstein, 1986) and has been very successful when applied to a number of Mendelian traits e.g. cystic fibrosis (Rommens *et al.,* 1989) and Huntingtons Chorea (Gusella *et al.,* 1983) and also for subsets of complex traits that show simple Mendelian inheritance – for example, early onset Alzheimer's disease (St George-

Hyslop *et al.,* 1987; Chartier-Harlin *et al.,* 1991), and the familial breast/ovarian cancer and breast cancer genes BRCA1 and BRCA2 (Miki *et al.,* 1994; Wooster *et al.,* 1995). For complex diseases, however, the involvement of many genes and the strong influence of environmental factors means that large multigeneration pedigrees are rarely if ever seen. Linkage analysis has thus tended to be performed using the less powerful affected sibpair families. Such genome-wide linkage scans are being conducted on many complex diseases, including typel and type 2 diabetes, multiple sclerosis, rheumatoid arthritis, schizophrenia, as well as bipolar disorder and some success has been reported. Two studies of type 1 diabetes, one study of rheumatoid arthritis, one study of inflammatory bowel disease and two studies of psoriasis all showed one or more highly significant linkages and two showed confirmed linkage (Davies *et al.,* 1994; Field *et al.,* 1994; Shiozawa *et al.,* 1998; Satsangi *et al.,* 1996; Tomforde *et al.,* 1994; Trembath *et al.,* 1997).

In the case of bipolar disorder a total of 23 genome scans have been completed to date with a number of studies still in progress. Despite the diversity of the ascertainment procedures used, some consistency in findings has begun to emerge. Of the completed genome scans, 'suggestive' and 'significant' linkages as proposed by Lander and Kruglyak have been detected. No one locus was detected at a 'significant' threshold by more than one study, but linkage near 4p12-13 was detected at the significant threshold in one study and the suggestive level in another (Ewald *et al.,* 1998; Blackwood *et al.,* 1996). Linkage near 13q31-33 was detected at the 'suggestive' level in three studies (Kelsoe *et al.,* 2001; Detera-Wadleigh *et al,* 1999; Stine *et al.,* 1997). However in view of the time, effort and cost that whole genome scans generate, it is perhaps disappointing that no two studies implicate the same region at or above the 'significant' level of genome-wide significance.

There may be several reasons for this. Firstly, in general, mapping of human susceptibility loci for complex human disease can be complicated by any or all of the following: high population frequency, incomplete penetrance, phenocopies, possible epistasis, and pleitrophy (Weeks and Lathrop, 1995). Also bipolar disorder like most complex diseases is likely to be genetically heterogeneous i.e. mutations in one of several independent genes may produce a similar clinical phenotype and different mutations in the same gene may cause a variety of related symptoms. The lack of

consistent findings from genome-wide linkage studies in bipolar disorder may in part be a result of the reduced power of the linkage approach for diseases that show genetic heterogeneity (Terwilliger and Weiss, 1998). Study design also has an important part to play in the replication of findings and it is now recognised that the marker densities and the sample sizes used in the published whole genome scans for bipolar disorder may be insufficient for detecting loci of modest effect (Risch and Merikangas, 1996). In addition, the observable differences in the diagnostic and ascertainment methods used by different groups make the replication of positive findings from whole genome scans extremely difficult and thus relatively limited.

1.10 Whole genome scans and bipolar disorder

Table 1.4 presents a summary of the 23 whole genome scans conducted to date for bipolar disorder. There are several regions of the human genome that have been repeatedly implicated by independent investigators. These include 4p16, 12q24, 18q22, 18p11, 10p, 13q, 21q21 and 22q11. Four of these regions (10p, 13q, 18p11 and 22q11) have also been implicated in genome scans for schizophrenia, suggesting that these two distinct categories of psychological disorder may share certain genetic susceptibility loci. While it is apparent that considerable progress has been made in the mapping of bipolar disorder susceptibility loci to the human genome, none of the confirmed linkage loci have as yet resulted in an identified gene. Advances in multipoint and nonparametric analysis over the past decade continue to increase the power of linkage methods to detect loci of modest effect. However, there remains a continuing need for novel methods, that will help define genetically more homogeneous phenotypes, take into account interactions between multiple susceptibility loci and accommodate additional complexities such as parent of origin effects, in the ongoing search for linkage.

Table 1.4: Published whole genome scans for bipolar disorder.

Study First author Year	Sample	Pedi-grees	Ascertainment	Affected Subjects	Number of Markers	Analysis performed	Regions to follow up	Comments
Coon, 1993	Utah	8	Proband with BPI<30 + >2 1st degree relatives with BPI, BPII or RUP	51	328	Pairwise lod score; positive findings followed up by APM	5q	Dominant model- 3 markers on 5q and 1 on 11 give lod >1. APM, D5S62 p<0.01
Straub, 1994	Columbia/ hadassah	47	12 / pedigree; at least 2 BPI; only 1 affected parent by history	490	5-153 marker / chromosome	Lod score analysis and model free APM and ESPA	21q22 (PFKL)	Large lod score seen in 1 family but APM yields p<0.0003
Blackwood, 1996	Edinburgh	1	Large family and more than one member with BPI	27	135, 87 extra on 4p	Lod score Mlink, affected only, linkmap, ERPA	4p12-13	D4S394, 2point lod score of 4.1 under dominant model. Support from other analyses
McInnes, 1996	Costa Rican	2	BPI probands in extended multiplex pedigrees	31	473	2 point linkage analysis	18q22-23 16q, 11p13-14	Only BPI considered as affected
Polymeropou-lus, 1996	Old Order Amish	1	5 BPI, 2 BPII	10	1772	2 point lod score. dominant model	6pter, 4pter 9q11-13, 19p13-q13	
Ginns, 1996	Old Order Amish	1	31 BPI, 50 other psychiatric diagnoses	81	≈400	Sib pair tests,ASP,APM, TDT, lod score analyses with 16 different models	6pter-p24 13q13 15q11-qter	Parametric and nonpa-rametric analyses provide evidence
NIMH initiative – Genome split into 4 ↓	NIMH	97	BPI probands with a BPI/SABP 1st degree relative with 2 living sibs + > 1 parent unaffected by BPI/SABP	392	318			
Rice, 1997	1,6,8,10, 12	As above	2 additional affected relatives with SABP/BPI/BPII/RUP	392	65	ASP, multipoint ASP Pedigree analysis-MOD	D1S1648 D6S1009 D10S188	Isolated findings need denser map

Study First author Year	Sample	Pedi-grees	Ascertainment	Affected Subjects	Number of Markers	Analysis performed	Regions to follow up	Comments
Stine, 1997	2,11,13,14,X	As above	As above	392	53	ASP, and pedigree methods (MOD)	13q14-32 Xp22 Xq26-28	Support from flanking markers for chrs. 13 and Xq
Edenberg, 1997	3,5,15,16,17,22	As above	As above	392	74	ASP, using SIBPAL and MOD analysis	16p, 5q, 3, D15S217, D17S961, D22S533	Support from flanking markers for 16, 5 and 3. 15, 17 and 22 are isolated results
Detera-Wadle-Igh, 1997	4,7,9,18,19,20, 21	As above	As above	392	126	ASP	4p, 4q, 7p 7q, 9q, 20p 20q, 21q	Only 21q supported by data from flanking markers.
Adams, 1998	Australian	1	Minimum of 3 affected individuals at least 2 of whom are BPI	11	214	Lod score analysis and APM method	4q35	Analysis of a further 10 families support the finding for 4q35
Ginns, 1998	Old Order Amish	4	Systematic ascertainment of bipolar probands	68	980	Model free ASP using sibpal and multipoint using genehunter-plus	4p12-13 4q, suggestive	Mental health wellness phenotype
Detera-Wadle-igh, 1999	NIMH-CNG + right extension of Amish 880	22	BPI/II proband with at least 6 living relatives and only 1 source of BP	159	607	ASP and multipoint analysis and parametric lod score; dominant and recessive models	13q32, 1q31-32 18p11.2	Also identified 7q31 and 22q11-13 as interesting . confirmed findings for 4p16, 12q23-24, 21q22
Morissette, 1999	Sanguenay-lac St john	1	Lithium responsive BPI probands + family history of affective disorders	53	332	2 point lod score, Sim APM and SimIBD genehunter,	12q23-24	Some support for two other loci on chromosomes 5 and 21
Friddle, 2000	JHU-Dana	50	BPI proband with 2/more affected sibs and only one affected parent	236	267	Parametric lod score analysis	8q24 4q 7p	All produced Hlod scores >2

Study First author Year	Sample	Pedi-grees	Ascertainment	Affected Subjects	Number of Markers	Analysis performed	Regions to follow up	Comments
Foroud, 2000	NIMH	97	As above 1997	392	318	Model free affected relative pair	10p	Lod score 2.5 near D10S1423, support previous finding on 10p
Kelsoe, 2001	SanDiego/BC	20	Proband with BPI/BPII + at least 2 relatives with BPI/BPII/SA/RUD	76	443	Parametric lod score analysis only	22q13, 3p21, 3q27, 5p15, 10q26, 13q31-33, 21q22	Chr 22 finding is significant. Other regions all have suggestive evidence for linkage
Badenhop, 2001	Australian of British descent	1	Minimum of 3 affected, 2 of whom have BPI	6	400	Lod score, multipoint and nonparametric analysis	13q14 3q21-25 5q37, 9q33	13q14 strong result under all analyses and haplotype analyses narrowed region to 6cM
UK-Irish Wellcome Trust, 2002	UK + Republic of Ireland	151	Proband BPI and at least 1 other sibling with BPI/BPII.SABP/BPNOS /MDD(R)	509	398	Mapmaker sibs calculate MLS lod score	2,3,4,6,7,9,10, 12,17,18,X	All reach nominal pointwise significance of 5% under narrow model. All will be followed up in 2nd stage.
Ekhlom, 2002	Finland	1	8BPI, 3BPII/BPNOS /cyclothymia/ 6 RMDD	18	386	Lod score	Xq25, 4q32, 12q15, 18p11, 19q13, 7q31 5q31, 10q22	Strongest finding for Xq25. DXS1047, lod score 1.76. other regions have lod score >1
Cichon, 2002	German/Israeli /Italian	75	Proband with BPI, secondary affected sib with BPI/BPII/ SABP/ UPR, availability of both parents or at least two more sibs	275	382	Parametric and nonparametric Genehunter imprinting	8q24, 10q25-26, 1p33-36, 2q21-23, 3p14, 3q26-27, 6q21-22, 8p21, 13q11, 14q12-13	Strongest evidence for 8q24. Putative paternally imprinted loci at 2p24-21 + 2q31-32. Maternally imprinted loci on 14q32 and 16q21-33
Ewald, 2002	Danish/all from Aarhus county	2	Narrow: BP, single manic episode or SA manic /depressed type Broad: narrow +modest depressive episodes		613	Two-point parametric, multipoint non-parametric linkage analysis	12q24.3, 1p22-p21, 4p16, 6q14-q22, 10q26, 16p13.3	12q24.3 'significant' and 1p22-p21, evidence from both parametric and NPL analysis.

1.11 Follow-up chromosomal regions

Independent studies have repeatedly implicated several distinct chromosomal regions, as possibly harbouring susceptibility loci for bipolar disorder. Although many of these regions have been confirmed according to the recommended criteria of Lander and Krugylak, individual susceptibility genes have as yet remained elusive.

1.11.1 Chromosome 18

Several independent studies have been performed on chromosome 18 and findings to date suggest that this chromosome may harbour at least two distinct regions of interest for bipolar disorder. Berrettini *et al.,* (1994), initially reported evidence for a bipolar disorder susceptibility locus in the pericentromeric region of chromosome 18 (18p11.2), employing 22 Caucasian kindreds of European ancestry and model free methods of analyses. Additional studies have since narrowed the original broad region of interest to about 11-15 Mb from the p-telomere where a LOD score of 2.32 was obtained (Detera-Wadleigh *et al.,* 1999). Numerous studies have independently confirmed this finding (Stine *et al.,* 1995; Turecki *et al.,* 1999; Bennett *et al.,* 2002). Interestingly, McInnes *et al.*, (2001) have conducted fine scale mapping of a 331kb region on 18p11.3 and their results suggest a candidate region for bipolar disorder containing six genes. Also of note is the fact that linkage appears to be more prominent in those families with paternal transmission of the illness (Nothen *et al.,* 1999; Gershon *et al.,* 1996). In addition to the 18p11.2 region, Stine and colleagues (1995) also reported evidence for linkage to a distinct and separate region of chromosome 18, 18q21-2. This study also displayed a parent-of–origin effect with excess sharing of the paternally transmitted alleles. This finding was followed up in an entirely new set of 30 families and results obtained substantiated the initial linkage found at 18q21 (McMahon *et al.,* 1997). Other groups have also found evidence for linkage at 18q, although in varying regions, including q12 (Ewald *et al.,* 1997), q22 (de Bruyn *et al.,* 1996) and q23 (Freimer *et al.,* 1996; Coon *et al.,* 1996; Nothen *et al.,* 1999). Although there is considerable evidence for the presence of bipolar susceptibility loci on chromosome 18, there have also been several negative reports (Bowen *et al.,* 1999; Knowles *et al.,* 1998; Kalsi *et al.,* 1997; Mynett- Johnson *et al.,* 1997; Detera-Wadleigh *et al.,* 1997). Despite the fact that caution must be exerted in accepting linkage to either 18p11 or 18q21 as confirmed, these regions remain potentially promising and are certainly worthy of further investigation.

1.11.2 Chromosome 21q

Straub *et al.,* (1994) detected in one of 47 bipolar disorder families, a robust LOD score of 3.41 at the phosphofructokinase (PFKL) locus on chromosome 21q22.3. Recently, this same group reported evidence of linkage in the same region about 5cM proximal to PFKL in 40 additional families (Aita *et al.,* 1999), and in a further follow-up linkage study employing all 56 pedigrees a LOD score of 3.56 was obtained with D21S1260 at 21q22 (Liu *et al.,* 2001), thereby further corroborating the original finding. Several independent studies have reported positive results for 21q; Detera-Wadleigh and colleagues reported evidence for excess allele sharing at 21q in two unrelated data sets (Detera-Wadleigh *et al.,* 1997, 1996) and Kwok *et al.,* (1999) reported support for the suggestive linkage of a susceptibility locus for bipolar disorder on 21q22.3. Also of interest is the study conducted by Smyth *et al.,* (1997), which produced a maximum LOD score of 1.29 with markers PFKL and D21S171, but when a two locus admixture model was employed incorporating the 21q locus and the tyrosine hydroxylase locus on 11p15, the LOD score increased to 3.87. There have, however, been several reports that have failed to find satisfactory evidence for linkage to 21q (McInnes *et al.,* 1996; Ewald *et al.,* 1996; Byerley *et al.,* 1995).

1.11.3 Chromosome 12q

A report indicating the co-segregation of Darier's disease (a dominantly inherited dermatological disease) and bipolar disorder in a single family raised interest in 12q24, the region to which the Darier's disease gene has been localised (Craddock *et al.,* 1994). Recently, two groups have reported suggestive positive linkage findings on 12q: Ewald *et al.,* (1998) reported a significant two point LOD score of 3.37 at D12S1639 located in the region 12q22-24 in two Danish families; Morisette *et al.,* (1999), using a homogeneous population in the province of Quebec, pointed to a locus of major effect for bipolar disorder on 12q23-24. Further evidence for a susceptibility locus for bipolar disorder on 12q24 comes from a study conducted by Degn *et al.,* (2001), which investigated haplotype sharing among bipolar patients from the Faroe Islands. To date, several candidate genes located in this region of chromosome 12 have been investigated, but as yet no evidence has been found for the direct involvement of these genes in producing susceptibility to bipolar disorder.

1.11.4 Chromosome 4p

A linkage study conducted by Blackwood *et al.,* (1996), using a single large Scottish family revealed a maximum LOD score of 4.8 at 4p16. Results from a study performed by Ewald *et al.,* (1998), of Danish families, supported this finding by generating a LOD score of 2.00 for the same marker, D4S394, that had displayed the tightest linkage in the Scottish family study. Further evidence for a locus at 4p was provided by a group in Germany, where they interestingly noted highly significant allele sharing at D4S394 in a subset of their families when they restricted their analysis to families showing paternal inheritance to bipolar disorder (Nothen *et al.,* 1997). Asherson *et al.,* (1998) also observed a LOD score of 1.84 with markers in this region of chromosome 4 for a family with multiple members affected by schizoaffective disorders. Interestingly, Ginns *et al.,* reported linkage to this region for a locus that potentially protects against bipolar disorder in four Amish families (Ginns *et al.,* 1998). Since the initial finding, the Scottish group has constructed an accurate, high-resolution physical map of 4p15.3-p16.1, which includes the critical region as an essential preliminary step to the identification of candidate genes that predispose to bipolar disorder (Evans *et al.,* 2001).

1.12 Common susceptibility regions and Psychosis genes

Classical Kraepelinian dichotomy classifies schizophrenia and bipolar disorder as separate diagnostic categories. However, currently there is increasing evidence to suggest that the two disorders may be more closely related than was originally envisaged. It is generally accepted that bipolar disorder and schizophrenia display considerable epidemiological overlap. They each have a population incidence of 1%, are lifelong conditions with a similar age of onset and both syndromes affect men and women equally. Furthermore, family studies considering segregation and phenotypic variation have demonstrated the existence of both phenotypes in the same kindred (Maier *et al.,* 1993), and the presence of intermediate phenotypes such as schizoaffective disorders in families with schizophrenia and bipolar disorder (Taylor and Amir, 1994). The presence of affective disorder within a family has also been found to predispose to schizophrenia in later generations (Coryell *et al.,* 1985). Of immense interest is the observation that epidemiological data appears to be supported by recent genome and chromosomal scans where overlapping susceptibility regions have been found for both schizophrenia and bipolar disorder. Importantly, these studies were

conducted in independently ascertained families where the index proband was diagnosed with either schizophrenia or bipolar disorder. The emergence of common susceptibility regions for the aforementioned psychiatric disorders increases the possibility of the existence of shared genes that contribute to the molecular basis of functional psychosis.

1.12.1 Chromosome 10p

Evidence for a schizophrenia susceptibility locus on 10p, was recently reported by three independent groups studying kindreds that were of general European ancestry. All three groups detected excess allele sharing at the marker D10S1423 or at other markers in close proximity (Faraone *et al.,* 1998; Straub *et al.,* 1998; Schwab *et al.,* 1998). Furthermore, a genome wide scan conducted by DeLisi *et al.,* (2002) produced an impressive LOD score of 3.60 at marker D10S189, located on 10p13-p15. The NIMH Initiative for Bipolar Disorder also revealed suggestive linkage for D10S1423 (Rice *et al.,* 1997) and additional analysis of these pedigrees again produced a peak LOD score of 2.5 near D10S1423 (Foroud *et al.,* 2000). It appears that strong linkage signals have emerged for both disorders within a narrow region on 10p13-12. Recent studies have also alluded to there being a second common susceptibility locus on 10q. A genome scan conducted by Levinson *et al.,* (1998) for schizophrenia highlighted 10q23 as possibly harbouring a susceptibility locus and this finding was further supported in the second stage of this genome scan (Mowry *et al.,* 2000). Linkage studies conducted by two other groups (Bailer *et al.,* 2000; Williams *et al.,* 1999) have also implicated 10q23 as a region of interest for schizophrenia. In regard to bipolar disorder, genome scans performed by Cichon *et al.,* (2002) and Kelsoe *et al.,* (2001) suggested a possible susceptibility locus for bipolar disorder on 10q25-26.

1.12.2 Chromosome 13q

There are several independent reports for a susceptibility locus for schizophrenia on 13q14-32. Lin *et al.,* (1995) found suggestive evidence of linkage in this region, a finding, which they subsequently followed up and substantiated by reporting a LOD score of 2.58. Further support for a susceptibility locus for schizophrenia on 13q14-32 was obtained in at least four unrelated schizophrenia studies (Blouin *et al.,* 1998;

Brzustowicz *et al.,* 1999; Camp *et al.,* 2001; Cardno *et al.,* 2001). Substantial evidence has also been obtained for a bipolar susceptibility locus at 13q. A genome scan conducted by Detera-Wadleigh *et al.,* (1999) revealed a LOD score of 3.5 at 13q32. Kelsoe *et al.,* (2001) also reported linkage of bipolar disorder to 13q31-q34 markers and recently a genome screen conducted by Badenhop *et al.,* (2001) also provided evidence in support of a susceptibility locus on 13q14-32.

1.12.3 Chromosome 22q11-13

In addition to the physical abnormalities characteristic of velocardiofacial syndrome (VCFS), a substantial number of patients have also been found to display psychiatric manifestations including bipolar disorder and schizophrenia. Pulver *et al.,* (1994) first described evidence for schizophrenia linkage to 22q11-13, the region associated with VCFS. Since then there have been a number of studies implicating this region of chromosome 22 as possibly harbouring a risk locus for schizophrenia (Gill *et al.,* 1996; Vallada *et al.,* 1995; Polymeropoulos *et al.,* 1994). Moreover, several studies have reported evidence for the presence of a bipolar susceptibility locus at 22q11-13 [Lachman *et al.,* (1996); Edenberg *et al.*, (1997); Kelsoe *et al.,* (2001); Detera-Wadleigh *et al.,* (1999)]. Thus for the 22q11-13 region, it appears that there are several reports of linkage to bipolar disorder and schizophrenia.

1.12.4 Chromsomome 18p

Chromosome 18p has previously been discussed in detail in relation to linkage findings for bipolar disorder (see section 1.9.2.1). Interestingly, Schwab and colleagues (1998) conducted a linkage analysis of 59 multiplex German and Israeli schizophrenia pedigrees and produced a maximum LOD score of 3.1 at D18S53, located at 18p11.2, thereby overlapping the region of interest for bipolar disorder. Two plausible functional candidates for psychiatric disorder have been localised to 18p: G-olf$_{\alpha}$, a G protein that primarily functions in transmembrane signalling involving neurotransmitters, and myoinositol monophosphatase (IMPA2), a signalling protein inhibited by lithium which is frequently administered in the treatment of bipolar disorder. Schwab and colleagues (1998) reported modest LD between the 124bp allele of the intronic CA-repeat within the G-olf$_{\alpha}$ gene and schizophrenia and since then the region surrounding this gene has

been subjected to further analysis. Currently, several recently identified novel cDNA clones await additional investigation (Rojas *et al.,* 2000).

A review of molecular linkage studies to date has disclosed evidence for at least four shared chromosomal regions proposed to be involved in genetic predisposition to bipolar disorder and schizophrenia: 18p11.2, 22q11-13, 13q32 and 10p14. Partial shared genetic risk is further endorsed by family studies and the acknowledged existence of independent susceptibility loci for both disorders. It is likely therefore that a number of genes may predispose individuals to psychosis, with the distinct clinical form of the disorder emerging due to epistatic or additive effects in addition to environmental contributions.

1.13 Modified genome scan

It was perceived by McKusick (1991) while maintaining a catalogue of human genes and phenotypes, that gene density varied considerably among chromosomes. It was also observed by Langlois *et al.,* (1982) that not all chromosomes have the same G+C versus A+T composition. Chromosomes 16, 17, 19, 20 and 22 were determined to be the most GC-rich while chromosomes 13, 18 and Y were the least. Saccone *et al.,* (1992) showed that even within chromosomes there was uneven distribution of GC-rich regions, which were found to be generally located near the telomeres of chromosomes. Antonarakis (1994) suggested that instead of using the systematic genome wide screen employing approximately 300 markers to cover the entire genome, it would be more efficient to use markers strategically located in the middle of the chromosomal regions that are GC rich and therefore likely to be gene rich. He concluded that there would be a good chance of localising at least 70% of the elusive genes with much less effort and expense. Inglehern (1997) selected two sets of Généthon markers (table 1.5), to screen expressed sequence tag (EST)-rich regions of the genome. He proposed that by employing just 65 markers, one could search 17% of the genome and thus screen at least 37% of the localised ESTs. Because of the estimated variation in EST density, several chromosomes were omitted from Inglehern's panel of markers, namely 13, 18, 21 and X, all of which are considered to be relatively gene poor.

Table 1.5: Généthon poly-CA markers, which selectively screen EST rich regions of the human genome (adapted from Inglehearn, (1997).

Set A					
D1S513	D1S255	D1S447	D1S252	D1S305	D1S212
D1S510	D1S213	D2S291	D2S113	D2S152	D3S1581
D3S1290	D5S658	D6S273	D7S632	D7S479	D8S531
D8S270	D9S176	D9S282	D10S198	D11S1346	D11S1313
D11S987	D11S916	D11S1317	D11S908	D11S924	D12S361
D12S90	D12S84	D15S211	D16S496	D17S933	D17S806
D19S226	D19S223	D20S174	D22S282		
Set B					
D1S464	D1S236	D1S248	D2S390	D3S1267	D4S1558
D4S395	D4S414	D5S668	D7S653	D7S480	D7S512
D9S165	D9S289	D10S537	D14S281	D14S271	D15S132
D15S155	D15S125	D17S938	D17S841	D17S748	D20S119
D22S275					

In a study performed by Murphy *et al.,* (2000), 48 Irish families were genotyped with Inglehern's panel of markers (1997). For each marker, genotypic data were analysed using Genehunter to obtain a two point nonparametric LOD score and associated p value. Heterogeneity LOD scores were also determined using three different genetic models and two affection models; broad and narrow definition of affection status. This study failed to reach the accepted stringent criteria for establishing suggestive or significant linkage. However, there were a number of chromosomal regions that generated positive results worthy of further consideration.

NPL scores >1 were observed for 6 individual chromosomal markers; D1S305, D5S563, D9S282, D14S271, D16S496, D17S841, D17S933 (table 1.6). Because of the nature of this study, close flanking markers (<10cM) did not exist for the positive findings on 5, 9, 14 or 16. On chromosome 17 however, NPL scores that approached significance were found for two adjacent markers (D17S841 – 7cM – D17S933). This result was obtained under the broad affection model but it also remained positive under the narrow model, although the significance of the result for D17S841 was considerably reduced.

Table 1.6: Results for six chromosomal regions showing positive NPL scores from modified genome scan (Murphy *et al.*, 2000).

	BROAD (BPI, BPII, SA, and RUP)					NARROW (BPI only)				
Marker	**NPL**	**pvalue**	**Hlod[1]**	**Hlod[2]**	**Hlod[3]**	**NPL**	**pvalue**	**Hlod[1]**	**Hlod[2]**	**Hlod[3]**
D1S305	1.192	0.11	0.35	0.27	0.17	0.995	0.06	0.41	0.26	0.83
D5S658	1.264	0.07	0.36	0.28	0.82	0.74	0.11	0.23	0.20	1.05
D9S282	1.652	0.04	1.25	0.93	0.24	0.485	0.21	0.37	0.53	0.14
D14S271	1.388	0.07	0.74	0.41	0.46	0.985	0.06	0.36	0.15	1.34
D16S496	1.330	0.07	1.00	1.37	0.47	0.331	0.30	0.51	0.84	0.03
D17S841	1.588	0.06	0.67	0.31	0.12	0.734	0.14	0.33	0.19	0.21
D17S933	1.490	0.07	0.76	0.78	0.00	1.15	0.05	1.25	1.35	0.11

All NPL scores quoted refer to use of Z_{All} statistic. Heterogeneity Lod scores (Hlod) are given for genetic models 1, 2, 3.

1.14 Candidate gene studies and bipolar disorder

To date, the pathophysiology of bipolar disorder remains poorly understood and as a consequence association studies have concentrated on perhaps the most obvious candidate genes i.e. the neurotransmitters themselves and those genes involved in the various aspects of neurotransmission regulation. Association studies and candidate genes are discussed in more detail in sections 1.7.3 and 1.7.3.1, but there have been a number of interesting findings for selected candidate genes and these are discussed in detail in the following section.

1.14.1 Dopaminergic system

There is considerable pharmacological evidence to suggest a possible role for dopamine, one of the central neurotransmitters, in the manifestation of bipolar disorder. Dopamine agonists such as L-dopa, amphetamine and bromocriptine have been found to induce mania (Gerner *et al.,* 1976), and a number of the dopamine reuptake inhibitors such as α – amineptine, display antidepressant effects (Paes de Sousa & Tropa, 1989). Dopamine antagonists are particularly effective in blocking an elated response (Wald *et al.,* 1978) and antipsychotics, which bind dopamine receptors, are routinely used in the treatment of mania (Cookson *et al.,* 1981). In addition, anti-manic drugs such as lithium and carbamazepine have significant effects on central dopamine activity (Avissar and Schreiber, 1992). These observations would suggest that dopamine activity is decreased during a depressive episode and increased during a manic episode thereby alluding to its involvement in bipolar disorder. A number of studies have examined dopamine receptor genes for association with bipolar disorder.

1.14.1.1 Dopamine D1 receptor (DRD1)

The dopamine D1 (DRD1) receptor gene is located at chromosome 5q35.1, a region that has produced notable linkage results in four whole genome screens on bipolar disorder (Coon *et al.,* 1993; Edenberg *et al.,* 1997; Badenhop *et al.,* 2001; Ekhlom *et al.,* 2002). At present there is no established relationship between any dopaminergic therapy and this receptor, however, several polymorphisms have been detected within this gene (Cichon *et al.,* 1996) and consequently this receptor has been the subject of a number of association studies. To date no evidence of association or linkage has been reported for the DRD1 receptor and bipolar disorder (Nothen *et al.,* 1992; Cichon *et al.,* 1996) and

several studies have also reported a lack of single strand conformation polymorphisms in patients with bipolar disorder that differed from controls (Shah *et al.,* 1995; Cichon *et al.,* 1994). Polymorphisms within this receptor have also been examined in other disorders such as Tourette's syndrome, Attention Deficit Hyperactivity Disorder (ADHD), autism and alcoholism. Generally the findings have been negative but an association between homozygosity at the D*de* I polymorphism in the 5'-untranslated region of DRD1 and subjects with compulsive, addictive behaviours has been reported (Comings *et al.,* 1997).

1.14.1.2 Dopamine D2 receptor (DRD2)

The human dopamine D2 (DRD2) receptor gene is located on chromosome 11q22-23. Whilst genome scans have not implicated this region of chromosome 11, the clinical effectiveness of neuroleptics, which are used in the treatment of bipolar disorder and schizophrenia, have been found to depend upon DRD2 receptor affinity. Consequently, several studies of these disorders have examined the dopamine D2 receptor gene for evidence of association. Generally, results have been inconsistent with both positive and negative findings emerging. A recent European multi-centre case – control study, which consisted of 358 bipolar patients and 358 matched controls, revealed evidence for a significant association between bipolar disorder and a DRD2 microsatellite ($p= 0.002$ for allele 5), suggesting possible LD with a nearby bipolar susceptibility variant (Massat *et al.,* 2002). Li *et al.,* (1999) also detected association between two different DRD2 variants and bipolar disorder in a Chinese cohort, but due to non-replication in a Caucasian sample, the authors suggested that this might have been a false positive.

There have been numerous negative reports for the involvement of DRD2 in the aetiology of bipolar disorder. Recently, parametric and non-parametric linkage methods were employed to analyse data from 23 mood disorder pedigrees. Results from this study excluded linkage to DRD2 (Serretti *et al.,* 2000), which concurs with findings from two previous linkage studies (De Bruyn *et al.,* 1996; Ewald *et al.,* 1995). Several association studies have also reported negative findings for DRD2 and bipolar disorder (Trio studies: Kirov *et al,* 1999; Bocchetta *et al.,* 1999; case control studies: Savoye *et al.,* 1998; Furlong *et al.,* 1998; Manki *et al.,* 1996).

1.14.1.3 Dopamine D3 receptor (DRD3)

The dopamine DRD3 receptor gene is of particular interest because of its expression pattern in brain structures controlling various aspects of behaviour, cognition and emotions (Civelli *et al.,* 1993). Moreover, it encodes a receptor protein that is a target for psychotropic drugs, which are efficient in the treatment of affective disorders. Hence, this gene has attracted considerable interest as a possible candidate for involvement in psychiatric disorders. Crocq *et al.,* (1992) reported increased homozygosity for the *Bal*I polymorphsim in schizophrenics as compared to controls and subsequently several studies were undertaken including a meta-analysis, which provided modest evidence in support of this finding in schizophrenia (Williams *et al.,* 1998). This polymorphism has also been extensively examined in bipolar disorder but with inconsistent findings. A meta-analysis of previous association studies undertaken by Elvidge *et al.,* (2001) revealed no significant difference in allele distributions between bipolar patients and controls for the *Bal*I polymorphism in ethnically homogeneous samples (OR, = 1.04; p = 0.60; 95% CI = 0.89-1.20). However, a study performed by Parsian *et al.,* (1995) did find an association between bipolar disorder and DRD3 using the haplotype relative risk method of analysis, and also of interest is a study carried out by Chiaroni *et al.,* (2000), where they observed a gender distribution difference for the *Bal*I polymorphism, with bipolar females being preferentially heterozygous and males homozygous.

1.14.1.4 Dopamine D4 receptor (DRD4)

The dopamine D4 (DRD4) receptor gene is located on chromosome 11p15.5, a region that has produced conflicting linkage data (see sections 1.2 and 1.14.3) but nonetheless has been the subject of several association studies. There is a hypervariable region within this gene and variants in this polymorphism have been employed for association studies with bipolar disorder. Results from such studies are inconsistent, with some studies providing no support for DRD4 (Oruc *et al.,* 1997; Bochetta *et al.,* 1999) and others displaying significant association between the dopamine D4 receptor gene and mood disorders (Manki *et al.,* 1996). Serretti *et al.,* (2001) have consistently found an association between the DRD4 exon 3 long allele variant and delusional symtomatology and have proposed that these variants may constitute a liability factor for the

development of delusional symptomatology in patients with major psychoses (Serretti, 1999).

1.14.1.5 Dopamine D5 receptor (DRD5)

The dopamine D5 receptor gene (DRD5) is located within the region of 4p14-16 where suggestive linkage findings for bipolar disorder have been noted (Blackwood *et al.,* 1996; Ewald *et al.,* 1998). Accordingly, Muir *et al.,* (2001) conducted a study looking for linkage disequilibrium between a series of microsatellite markers from this region and major psychoses including schizophrenia and bipolar disorder. The findings from this study were consistent with an association between markers close to the D5 dopamine receptor and schizophrenia but not bipolar disorder. Studies performed by Kirov *et al.,* (1999) and Asherson *et al.,* (1998) also failed to provide evidence to suggest that mutations in DRD5 are involved in the aetiology of bipolar disorder.

1.14.1.6 Dopamine Transporter (DAT)

The dopamine transporter (DAT) is a plausible candidate for bipolar disorder because of its important role in the regulation of dopaminergic neurotransmission by mediating the active reuptake of synaptic dopamine. The DAT gene has been mapped to chromosome 5p15.3, a region where previous linkage results have suggested the presence of a susceptibility locus for bipolar disorder (Kelsoe *et al.,* 1996). Generally, association studies, both family and case control, have reported negative findings for involvement of this gene in susceptibility to bipolar disorder (Georgieva *et al.,* 2002; Bochetta *et al.,* 1999). As a functional variant has not been identified in the DAT gene to date, studies rely on linkage disequilibrium to assess the genetic contribution of DAT to disorders such as bipolar disorder. Greenwood *et al.,* (2002) conducted a study to investigate the complex structure of linkage disequilibrium across the DAT gene and they found that while there was a high degree of LD at the 5' and 3' regions of the DAT gene, there was very little significant LD between them. The complex pattern of LD variation observed for the DAT gene implies that association studies using markers chosen at random from polymorphic sites even within the gene may not be reliable for the detection of a nearby causal variation. Therefore, this study suggests that knowledge of the LD structure and

intermarker relationships is essential to the design and efficacy of future association studies.

1.14.2 Serotonergic system

There is reasonable evidence to suggest that serotonin, a key neurotransmitter in the central and peripheral nervous system, is involved in the control of a variety of complex behaviours inclusive of psychiatric disorders. Although the role of serotonin in bipolar disorder has not been as extensively researched as for unipolar depression, post-mortem studies, cerebrospinal fluid and neuroendocrine studies have all found significantly reduced serotonergic activity in bipolar depressed patients. In the manic phase results are less consistent, although decreased serotonin responsiveness has been reported in some studies (Mahmood and Silverstone, 2001).

1.14.2.1 Serotonin transporter (5-HTT)

The serotonin transporter (5-HTT) is a gene that has attracted considerable interest as a possible candidate for bipolar disorder. Its candidacy is based on the fact that in addition to regulating extracellular serotonin levels, the serotonin transporter is also the site of action of the selective serotonin re-uptake inhibitors, which are a major class of antidepressants, effective in the treatment of bipolar depression.

Two polymorphisms within the serotonin transporter have been studied extensively. The first is a variable number of tandem repeats (VNTR) in the second intron (three alleles Stin 2.12, Stin 2.10, Stin 2.12). Functionally, Stin 2.10 and Stin 2.12 have been found to act as positive transcriptional regulators of serotonin with allele dependent differential enhancer-like properties in mouse embryonic rostral hind-brain (Mackenzie & Quinn, 1999). There have been various conflicting reports regarding association between the VNTR polymorphism and bipolar disorder. A family based study (Kirov *et al.,* 1999) and several case control studies (Collier *et al.,* 1996; Kunugi *et al.,* 1997; Rees *et al.,* 1997) have provided evidence supporting an association between the allele 12 variant of the VNTR with susceptibility to bipolar disorder. In contrast, there have been several negative findings for the VNTR polymorphism (Saleem *et al.,* 2000; Liu *et al.,* 1999; Gutierrez *et al.,* 1998). A meta-analysis conducted by Furlong *et al.,* (1998) consisting

of over 375 bipolar individuals of European Caucasian origin also observed a negative finding for the VNTR polymorphism. Recently, studies have been conducted to investigate serotonin uptake in human blood platelets. Platelets, which have receptors and serotonin uptake mechanisms similar to those of presynaptic serotonergic neurons (Sneddon, 1973), have been widely used as a peripheral model of central serotonergic activity. A study carried out by Kaiser *et al.,* (2002) observed that individuals who were homozygous for the allele 12 variant of the VNTR polymorphism appeared to have a lower affinity for serotonin uptake than individuals heterozygous for the other alleles.

The second polymorphism within the serotonin transporter is a functional deletion/insertion polymorphism in the promoter (5-HTTLPR), which displays two allelic forms, the long variant (l) and the short variant (s). The transcriptional activity of the promoter is restricted by the short variant, resulting in low functional expression of the serotonin transporter (Collier *et al.,* 1996). Serotonin uptake is consequently altered, which *in vivo* may have physiological consequences. Collier *et al.,* (1996) found a significant association between bipolar disorder and the low transcriptional activity variant in a large European case-control study. This finding was replicated in a meta-analysis conducted by Furlong *et al.,* (1998). Conversely, a number of studies have resulted in negative findings for the promoter polymorphism, most notably a meta-analysis of 11 case-control studies involving 1149 bipolar individuals and 1693 controls, which showed no evidence to support a role for the promoter (Craddock *et al.,* 2001). Evidence from these studies would suggest that the serotonin transporter is not a gene of major effect and further studies are required to determine the exact nature of the contribution of this gene to the aetiology of bipolar disorder.

1.14.2.2 Serotonin receptors

The serotonin receptors, because of their role in the neuronal action of serotonin, are plausible candidates for involvement in the aetiology of bipolar disorder. There are several serotonin receptor subtypes, some of which have gained more attention than others. The serotonin 5-HT2A has been examined in several studies, mainly due to the fact that this gene is located on chromosome 13q, a region that has been implicated in bipolar disorder through linkage studies (Badenhop *et al.,* 2001). Findings are varied, with some groups reporting a positive association (Chee *et al.,* 2001; Bonnier *et al.,*

2002; Vincent *et al.,* 1999) and others presenting negative results (Murphy *et al.,* 2001; Massat *et al.,* 2000). A meta-analysis of 8 case-control studies, recently conducted by Craddock *et al.,* (2001), revealed no evidence that variation at the T102C polymorphism influenced susceptibility to bipolar disorder. When 5HT-2A gene expression was examined from skin biopsies of retinoblastoma patients with chromosomal deletions and re-arrangements, it was observed that only the maternally inherited allele, which displayed methylation, was expressed, thereby suggesting the possibility that 5HT-2A may be subjected to genomic imprinting (Kato *et al.,* 1996). The involvement of parent of origin effects in bipolar disorder have been investigated and the outcomes of these studies have been varied. Some clinical studies suggest a predominance of maternal transmission patterns (McMahon *et al.,* 1995; Gershon *et al.,* 1996), some point towards paternal transmission (Kornberg *et al.,* 2000), while other analyses indicate comparable rates of maternal and paternal transmission (Grigoroiu-Serbanescu *et al.,* 1995; Kato *et al.,* 1996; Murphy *et al.,* 2001). It has also been shown that 5HT-2A exhibits a phenomenon known as polymorphic imprinting in the human adult brain (Bunzel *et al.,* 1998), with some individuals displaying monoallelic expression while others show biallelic expression. One of the advantages of family based association studies is that they offer the ability to observe the path of parental transmission of alleles, which is vital when considering the possibility of genomic imprinting. Future studies investigating the role of 5HT-2A in bipolar disorder, would be advised to apply parental allele specific association tests to large family based studies in order to ensure that the possibility of genomic imprinting is accounted for.

Another serotonin receptor that has received some interest is the 5-HT2C receptor, which is located on chromosome Xq24, a region that has been previously implicated through linkage studies (Ekhlom, 2002). Lerer *et al.,* (2001), found evidence to support a possible role for genetically based structural variation in 5-HT2C receptor in the pathogenesis of bipolar disorder. However, there have also been negative findings for this receptor; Gutierrez *et al.,* (2001) found no association between DNA polymorphisms in the 5' regulatory region of this gene, a finding replicated by Meyer *et al.,* (2002). Other receptor subtypes that have been investigated for possible association with bipolar disorder are: 5-HT1D (β)- where a family based association study excluded direct involvement of this receptor in the pathogenesis of bipolar disorder (Mundo *et al.,* 2001); 5-HT3A- where Niesler *et al.,* (2001) detected 2 novel mutations in the 5'region

of this receptor and found a significant association between the functional variant C178T and bipolar disorder; 5-HT5A- where a study performed by Arias *et al.,* (2001), examined two polymorphisms within this gene and found no genotype, allele or haplotype association with bipolar disorder. In contrast to this finding, Birkett *et al.,* (2000) did find allelic association between one of these polymorphisms (19G/C) and bipolar disorder; 5-HT6- both positive and negative results have also been observed for this receptor (Vogt *et al.,* 2000; Hong *et al.,* 1999).

While the serotonin receptors remain plausible candidates for involvement in the aetiology of bipolar disorder, conclusive findings have not been generated from studies conducted to date.

1.14.3 Tyrosine hydroxylase

Catecholamine dysfunction has long been hypothesised to be involved in a number of psychiatric and behavioural disorders. Tyrosine hydroxylase (TH) is the rate-limiting enzyme that catalyses the first step in the synthesis of catecholamines and, therefore, it has been proposed as a plausible functional candidate for genetic studies of bipolar disorder. The tyrosine hydroxylase gene is located on chromosome 11p15, which is the region alluded to in the initial Old Order Amish positive linkage study (Egeland *et al.,* 1987). This study was later extended and reanalysed (Kelsoe *et al.,* 1989) and the results failed to find support for the involvement of this region of chromosome 11 in bipolar disorder. This negative finding for 11p15 has been replicated by many other subsequent linkage studies, thereby casting doubt on the original finding. However, interest in this region was revived when studies carried out by Leboyer *et al.,* (1990) and Meloni *et al.,* (1995) reported association with two RFLPs at the tyrosine hydroxylase gene and a microsatellite marker in intron two of the gene, in a small French case-control sample. These polymorphisms have been the subject of several subsequent studies but findings have proven inconsistent. Serretti *et al.,* (1998) found that the tyrosine hydroxylase gene was associated with depressive symptomatology in mood disorder and Malafosse *et al.,* (1997) produced data to suggest that risk factors for bipolar disorder were found in the 11p15 region, with tyrosine hydroxylase being the most likely candidate gene. However, there have been numerous studies that refute these findings (McQuillin *et al.,* 1999; Burgert *et al.,* 1998; Oruc *et al.,* 1997). Also meta-analyses preformed by Furlong

et al., (1999) and Turecki *et al.,* (1997) failed to reveal evidence that variation within this gene influenced susceptibility to bipolar disorder. It is noteworthy however, that the sample sizes in these analyses have only modest power to detect susceptibility genes for complex diseases (Bellivier *et al.,* 1998).

1.14.4 Monoamine oxidase A (MAOA)

Monoamine oxidases play a critical role in the deamination of biogenic amines including neurotransmitters such as noradrenaline, serotonin and dopamine and consequently are possible candidate genes for bipolar disorder. Like the previously mentioned candidate genes, results for MAOA have been inconsistent with positive and negative findings being reported (Lim *et al.,* 1995; Kirov *et al.,* 1999; Syagailo *et al.,* 2001). However, two recently performed meta-analyses have provided support for the association of MAOA polymorphisms and bipolar disorder. Due to marked differences in population allele frequencies, Furlong *et al.,* (1999), carried out separate meta-analyses for Caucasian and Japanese samples. Data from all published studies were included and the analyses employed two of the common polymorphisms found within the MAOA gene. They reported associations between bipolar disorder and the MAOA-CA marker in both populations. Preisig *et al.,* (2000) replicated the positive finding of Lim and colleagues by finding an association between bipolar disorder and the MAOA-CA marker in female subjects of a French-Swiss sample. They also carried out a meta-analysis of the available data obtained in the five studies that have examined the association between MAOA markers and bipolar disorder in Caucasians, and they reported confirmation of the association between the MAOA-CA polymorphism and bipolar disorder in women and in the overall sample. It is unlikely that the microsatellite itself is functionally responsible for susceptibility to bipolar disorder but it may be close to the contributing genetic variant. It appears that the MAOA gene may be a gene of modest effect in bipolar disorder, particularly for female probands.

1.14.5 Catechol-O-methyltransferase (COMT)

Catechol-O-methyltransferase (COMT) catalyses the methylation and hence the inactivation of catecholamines and it therefore plays a crucial role in the metabolism of such neurotransmitters as L-dopa, noradrenaline, adrenaline and dopamine. It exists as

common high and low activity alleles in the population, which result from a valine to methionine substitution at codon 158 of the protein. Several case-control and family trio studies have been undertaken to investigate this functional polymorphism in bipolar disorder. Results from these studies have been varied but noteworthy is a meta-analyses of 13 case-control samples conducted by Craddock *et al.,* (2001), which reported evidence that the low activity allele may increase susceptibility to bipolar disorder. A small family-based association study performed by Mynett-Johnson *et al.,* (1998) suggested the possibility of a gender effect by finding an association between bipolar disorder in female probands and the low activity allele of the COMT gene. Also of interest is a report by Lachman *et al.,* (1996) where they found that there was an apparent association between the low activity allele and the development of bipolar spectrum disorder, in particular the rapid cycling form, in a population of patients with velo-cardio-facial-syndrome (VCFS). This finding was supported by a study performed by Kirov *et al.,* (1998) in a larger sample of bipolar patients without VCFS. Papolos *et al.,* (1998) reported evidence that the low activity allele predicted ultra-ultra rapid cycling but not standard rapid cycling. These studies suggest that the low activity allele of COMT could represent a modifying gene (i.e. a gene that influences clinical features and treatment response but does not affect liability) that predisposes to rapid cycling in patients with bipolar disorder.

1.15 The future for complex genetics

1.15.1 Human genome project and its application to complex disease

The human genome project is an international effort, begun in 1990, to determine the sequence of the approximately 3.2 billion subunits that make up human DNA, the chemical code that instructs cells to produce the proteins that not only produce physical attributes but also contribute to less tangible traits, such as behaviour, learning and predisposition to disease. For the past 12 years, genome centres all over the world have generated raw sequence information by determining the sequence of the base sequence of small fragments of human DNA and at present a comprehensive working draft of the human genome (90% of which is fully sequenced) is freely available to the public domain (http://genome.ucsc.edu/; http://www.ncbi.nlm.gov/).

This immense project is expected to be complete by the end 2003. The generated data will unquestionably underpin human biology and medicine in the next century, providing a single, essential reference to all genetic information. Investigations of common complex disorders including psychiatric genetics are significantly improved as a result of the vast amount of essential information arising from the availability of the human genome sequence.

Complete and accurate maps of each chromosomal region establish the order of linked genetic markers, define the size of the critical region and aid the identification of positional candidate genes. Contigs are a source for the isolation of new polymorphic markers and they facilitate high-resolution mapping of genes relative to linked or associated markers, and provide a physical template for transcript identification procedures. The human genome sequence will also provide a complete description of all expressed genes. While annotation of the raw DNA sequence is no easy task, it is possible via a combination of computational analysis and experimental confirmation by cDNA sequencing. Homologies between newly defined gene products and proteins of known function can be detected, and this information will help to postulate biochemical functions for these novel genes, which can then be tested experimentally. As our knowledge of gene location and function expands, it will be possible to pinpoint more accurately the genes and proteins that are involved in common complex disorders (http://www.ensembl.org/).

Comparative genomics is also an area of key importance for interpretation of the human genome sequence. By comparing corresponding genomic sequences in different species, regions that have been highly conserved during evolution can be identified. Many of these regions will reflect sequences of conserved functions such as those that control gene expression (http://www.ensembl.org/).

The human genome sequence also provides a starting point for the identification and the characterisation of individual sequence variation, including variation that confers susceptibility to common complex disorders. DNA sequence variation in the form of single nucleotide polymorphisms (SNPs) are remarkably common, occurring on average once every 1000 bp between any two individuals picked at random from the population (Roses *et al.*, 2001). The SNP consortium (http://snp.cshl.org/), a unique collaboration

between academic institutions and industry, and the analysis of clone overlaps by the International Human Genome Sequencing Consortium (ftp://ftp.ncbi.nih.gov/snp/) have succeeded in providing a map of 1.42 million SNPs distributed throughout the human genome, with an average density on available sequence of one SNP every 1.9 kb (Sachidanandam *et al.,* 2001). This high density SNP map provides a public resource that will facilitate efficient genome-wide association studies, replication studies and meta-analyses and ultimately should help in the identification of biomedically important genes for diagnosis and therapy (Evans *et al.,* 2001).

1.15.2 Whole genome association (LD) studies

With the advent of new molecular tools, in particular high-density SNP maps (Wang *et al.*, 1998), genome-wide large-scale linkage disequilibrium mapping using tens of thousands of candidate SNPs appears possible. It has been proposed that association mapping in many cases may be more efficient than linkage analyses in detecting genetic regions involved in disease (Risch and Merkingas, 1996). Numerous genetic effects, too weak to be detected by linkage, could be identified by genomic association studies. There are advantages to SNP genome-wide association mapping which can be exploited in the search for complex disease genes. Firstly, the abundance of SNPs, which are freely available in the public databases, allow for high-density gene mapping studies on a scale larger and more intensive than previous efforts thereby increasing the power of resolution of gene mapping (Johnson *et al.,* 2001). Secondly, the presence of SNPs throughout the genome, including exons, introns, promoters, and enhancers, imply that SNPs could have direct relevance to protein function, impact splicing or influence gene expression. Thirdly, the mutation rate of SNPs is lower than for other polymorphisms, which allows for greater reliability in assessing co-segregation and locus associations across generations. Finally, SNPs have a simple structure and therefore they lend themselves to rapid and efficient genotyping. There are, however, uncertainties associated with genome-wide linkage disequilibrium studies, which should not be overlooked.

Linkage disequilibrium is found to be highly variable across the genome (Boehnke, 2000), with variability occurring even between adjacent regions of a chromosome (Taillon-Miller *et al.,* 2000). Experimental data have produced widely varying estimates

of LD from 5 kb to several hundred kb (Taillon-Miller, 2000; Abecasis *et al.,* 2001). LD has also been found to vary greatly in different populations (Reich *et al.,* 2001), and the extent of LD in a population will predict the number of SNPs that need to be tested when performing genome-wide screens for susceptibility genes. If LD extends 60kb on average then as few as 30,000 SNPs should be sufficient for a genome-wide scan, however, if as predicted from computer simulations (Kruglyak, 1999), LD does not extend beyond 3kb, then around 500,000 SNPs would be required for a genome scan. Genotyping effort could be reduced by the use of DNA pooling and new technologies such as DNA chips, or DNA microarrays could speed up the process and lower the overall cost. However, at the present time, the effort and cost of such a project would need to improve considerably for genome-wide association studies to become feasible.

There are also other issues concerning high-density SNP mapping which also need to be considered. Many disease genes might be overlooked simply because a negative linkage disequilibrium result does not rule out a significant gene effect in a particular genomic region. There may also be uncertainties about the accurate interpretation of SNP allele frequencies in various populations, which would have an impact on the value of SNPs for association mapping studies (Goddard *et al.,* 2000).

Finally, there are statistical issues that may hamper the interpretation of SNP association studies, such as appropriate p-value criteria to avoid false positives due to multiple test effects. The power of such a study could also be compromised by the possibility of the presence of allelic heterogeneity or population heterogeneity (Kruklyak, 1997). High recombination rates that can disrupt the correlation between SNPs and disease-related variants, and the presence of multiple mutations (per gene) that predispose an individual to disease, may further hinder the detection of gene effects in an SNP analysis (Pennisi, 1998).

The traditional linkage genome scan remains a powerful tool for detecting genes of moderate to large effect but unfortunately it has generally been unsuccessful in detecting genetic loci contributing to susceptibility to complex diseases. Genome-wide association studies target the entire complement of genes including genes of small effect but currently are constrained by genomic, statistical and logistical complexities. The ever-improving map of the human genome together with continued advances in the

areas of study design, statistical methodologies and recent technological developments imply that while linkage analyses will continue to play an important role, the future of the genetics of complex diseases is likely to also require large-scale testing by association analysis (Risch and Merkingas, 1996).

1.15.3 Population Isolates

The utilization of population isolates, i.e. populations that are more genetically homogeneous due to historical reasons, has gained a lot of attention over the past number of years as they are informative for a number of different types of analyses such as classical linkage analysis, LD studies and fine mapping. The theory for using such a sample is quite straightforward. If there are several genes that when mutated can cause the same phenotype in the human population as a whole, then affected individuals from an isolated or homogeneous genetic pool, may have fewer of those predisposing genes segregating within that population than those that were present in the founding ancestors. If there are less predisposing genes, then current linkage methods should be even more successful at detecting susceptibility genes than they would be if applied to a more heterogeneous population (Escamilla, 2001). Certain population isolates, namely those with a small number of founders, coupled to substantial growth for 10-20 generations, and quite a large current population, have been proposed to be considerably more appropriate for association-based or LD based mapping at the genome screening stage than outbred populations. In population isolates of recent founding, the chance of detecting an association between the founding disease causing mutation and a segment of DNA inherited IBD is much greater than if the same exercise was performed in an outbred population. Allelic heterogeneity reduces the power of the association analysis in the outbred population. Given that linkage disequilibrium is eroded over time by recombination, population isolates, especially those with an older history such as Finnish and Sardinian populations are very useful for fine mapping, as LD is thought only to extend over distances of 1cM in these populations. Thus the number of markers that would be required for fine mapping a region of 20-30 cM would be reasonably manageable and yet at the same time, LD when detected, would allow for tight localisation of the susceptibility gene. There have been several studies of genetic mapping of bipolar disorder carried out in population isolates including Finland, Iceland, the Central Valley of Costa Rica, the Old Order Amish, Quebec and Scotland.

To date most of these studies have focussed on large multiply affected pedigrees using classical linkage analyses. There has been some success in identifying interesting chromosomal regions using population isolates e.g. 4p, Blackwood *et al.*, (1996) using a large Scottish pedigree; 12q, Morissette *et al,* (1999) using a homogenous population in Quebec; 18q, Freimer *et al,* (1996) using a Costa Rican pedigree; chromosomes 6, 13 & 15, Ginns *et al,* (1996) using Amish pedigrees. Complete genome scans have only been reported for studies in Costa Rica, Amish population and Quebec. Despite the potential for performing LD analyses in these populations, LD studies have thus far only been performed in the Amish population (Ginns *et al.,* 1996) and in Costa Rica (Escamilla *et al,* 1996; Ophoff *et al,* 2002). LD fine mapping is at present underway in at least three populations (Costa Rica, Scotland, Finland) where population sizes are large enough to enable such studies to be carried out.

Of course there are potential problems with population isolates. Not all population isolates might be as homogeneous as one would like to think. If the founders of a particular region were genetically diverse, then it is possible that the population deriving from those founders could maintain a high level of heterogeneity at the genetic level. However, it is evident that population isolates will continue to play a central role in unravelling the molecular genetics of bipolar disorder.

1.15.4 Meta-analysis

Although significant progress has been made in defining the genetic basis of complex human diseases, it is acknowledged that even relatively large studies are likely to have low power to map, genes of modest effect by linkage (Risch 2000; Terwilliger and Goring, 2000). One potential solution to the recognised limitations of linkage studies is to perform a meta-analysis on multiple linkage scans of the genetic trait under consideration. While this concept is not new, it can be difficult to apply to published data given the lack of consistency in the presentation of linkage results. Badner and Gershon (2002) presented a meta-analysis technique, called Multiple Scan Probability (MSP), which can readily be applied to published genome scans. This technique allows for the fact that in the majority of reported studies, only regional p values may be available. The MSP result is calculated by taking the minimum p value over *n* cM in each study, correcting the observed p value by estimating the probability of it occurring

within *n* cM and then summing the logs of the corrected p values. The null hypothesis is that the distribution of p values is uniform and if this hypothesis is rejected, the implication is that a true linkage is present in at least one of the studies. There are limitations to this type of analysis; for example the power of the meta-analysis to detect a particular susceptibility locus is dependent upon the power of the original studies, which could result in false negatives, different meta-analysis on the same data could generate different results depending on the criteria chosen for the selection of appropriate studies. However, it has been shown by means of simulation that the MSP technique is both more conservative and more powerful than requiring that one of several genome scans meets Lander and Kruglyak's criteria for significant linkage. The replication MSP (an analysis that excludes the most significant study), which is routinely employed in this approach, offers a means of determining if the MSP result is primarily due to one significant result or if at least one other study is contributing, thereby affording further support for a positive finding. This method is also found to be robust to a considerable amount of study heterogeneity and it also tends to be more inclusive than collaborative projects as it requires no effort on the original investigators other than publishing their findings. MSP meta-analysis has been applied to all the published genome scans for bipolar disorder and schizophrenia (Badner *et al.,* 2002) with results showing strong evidence for susceptibility loci for both bipolar disorder and schizophrenia on 13q and 22q. There was also substantial evidence for a susceptibility locus for schizophrenia on 8q. There are further meta-analyses in progress for bipolar disorder and schizophrenia as it is now widely acknowledged that the combination of evidence from multiple studies may prove to be critical to the localisation of genes of modest effect in common complex human disease.

1.15.5 Endophenotypes

Failure to replicate much of the linkage findings for bipolar disorder has led to a growing consensus that utilisation of additional approaches and methodologies is essential for the identification of susceptibility genes in a disease as complex as bipolar disorder. Leboyer *et al.,* (1998) have proposed that identifying more homogeneous forms of diseases through a 'candidate symptom approach' among affected subjects and employing an endophenotype methodology that identifies sub-clinical traits among non-affected relatives might produce better results. The former approach identifies narrow

clinical characteristics that are most likely to be associated with the disease genotype and that show a simple pattern of inheritance. Identification of specific subforms of the disorder would then lead to the identification of homogeneous families that are more appropriate for linkage studies. There are several examples where clinical characteristics such as disease severity, age of onset and family history have succeeded in narrowing the phenotype and facilitated the identification of the relevant genes. For example, in breast cancer research, restricting the gene search to families with early onset disease led to the identification of the BRCA1 and BRCA2 genes. In psychiatry, early onset has been found to be associated with increased familial risk in schizophrenia and bipolar disorder (Leboyer *et al.,* 1998). It has also been suggested that BPII, which in studies at John Hopkins, constitute the largest group of affected relatives, could have an important role as a subtype (Simpson *et al.,* 1993). McMahon *et al.,* (2001) identified a chromosomal region, which potentially harbours a susceptibility gene in the BPII rich families but not in families where few members have BPII.

The latter strategy is based on the identification of subclinical associated traits 'endophenotypes', which are believed to represent the genetic liability of the disorder among the non-affected subjects. Endophenotypes are thought to have simpler genetic underpinnings and may have a considerably higher incidence in the population than the disorder. There are several examples of somatic diseases where the endophenotype level has helped to define the molecular genetic basis of the illness. Epilepsy is a common disease, displaying analogous difficulties in nosology and phenotypic definition as those found in psychiatric diseases. A genetic susceptibility factor in juvenile myoclonic epilepsy was identified when investigators chose a subclinical trait as an endophenotype in non-affected relatives to identify linkage to chromosome 6 (Greenberg *et al.,* 1988). There are several proposed endophenotypes for bipolar disorder, for example, abnormal regulation of circadian cycling, a manic response to sleep deprivation, a response to psychostimulants, all of which are documented in literature to differ between patients with bipolar disorder and control subjects (Lenox *et al.,* 2002). These potential endophenotypes for bipolar disorder have yet to be proven true, but there is the hope that endophenotypes can potentially narrow down the genetic determinants contributing to the disorder, thereby increasing the available power to identify the genes responsible for complex behavioural disorders.

1.15.6 Animal models and Bipolar Disorder

Animal models have an intrinsic role to play in delineating a fundamental understanding of the genes that are in involved in the manifestation of mood disorders and also in identifying the changes in the brain that underlie the diverse symptoms of these illnesses. The major difficulty in devising animal models for mood disorders is defining a relevant rodent phenotype. Bipolar disorder is further complicated by the oscillating pattern of the illness, which presents a further difficulty in the development of an appropriate model for the overall syndrome. Presently, the majority of the animal models for depression are based on reproducing symptoms of depression in the animals. By using a variety of stressful conditions, symptoms that are inferred to be depressive-like can be induced in the animals. While these models can be used to study the biological mechanisms underlying the symptoms and to develop new treatments that alleviate the symptoms, they are limited by the possibility that they may poorly reflect the mechanisms involved in the human. In the case of mania, traditional animal models have focussed on modelling a single manic episode; by for example, concentrating on the hyperactivity aspect of the disorder or on the effects of sleep deprivation, but as yet there is no bona fide rodent model for mania.

It has been proposed that animal models for depression and bipolar disorder could be improved by using a variety of mutational techniques in mice (Nestler *et al.,* 2002). Random mutations, using chemical mutagens, could be introduced into the animals, which could then be screened for certain behavioural patterns, the disrupted gene could be identified and related to the behavioural phenotype. Such large scale mutagenesis studies are now getting underway in the mood disorders field. Mutations could also involve altering a particular protein that has been implicated in other animal models of depression or antidepressant treatment and the resulting phenotype characterised. Although such models may not replicate the pathophysiology of human depression, they do offer insights into the neural and biochemical pathways involved in regulation of mood. Also endophenotypes (section 1.15.5) may have a major role to play, as it should be more straightforward to model these in laboratory animals.

While the development of an unambiguous animal model for bipolar disorder seems unlikely, animal research will undoubtedly play an essential role because once human disease genes are discovered, these genes will have to be placed in laboratory animals to

understand, at the molecular, cellular, anatomical and neural system level, how the genetic abnormality leads to the disease. This is a fundamental step in translating genetic discoveries into improved treatments.

1.16 Study Aim

The aim of this project was to investigate a number of chromosomal regions to determine if they continue to provide evidence of linkage and/or linkage disequilibrium to bipolar disorder within the Irish population. This was accomplished through a combination of additional sib-pair based linkage studies to further localise the regions displaying linkage and linkage disequilibrium studies using functional and/or positional candidate genes, a number of which were identified by *in silico* methodology.

CHAPTER TWO

MATERIALS AND METHODS

2.1 Patient sample and ascertainment procedure

2.1.1 Diagnosis

All family members were interviewed by trained psychiatric nurses using the Schedule for Affective Disorders and Schizophrenia (Lifetime Version; SADS-LB) The resulting information together with information from previous hospitalisations and clinical records was collated and run through Operational Crtiteria Checklist for Psychotic Illness (OPCRIT) (McGuffin *et al.*, 1991), a computerised checklist, which produces diagnoses under differing classifications such as DSMIII R, RDC and ICD-10. Diagnoses were reviewed by a consultant psychiatrist, and all met DSM IV criteria (Diagnostic and Statistical Manual, American Psychiatric Association, 1994).

2.1.2 Irish Bipolar Patient Collections

The collections have been established throughout the life of the genetic susceptibility project and numbers included in this section relate to the current samples examined in the majority of studies detailed in this thesis.

2.1.3 Sib-Pair Family Collection

Families with at least one affected sib-pair were identified at the Mood Disorder Clinic at St. Patrick's Hospital, Dublin and other similar clinics throughout Ireland. They were diagnosed as previously described in section 2.1.1. In order to qualify as an affected sib-pair, one sib is required to be diagnosed with BPI and the additional sib must be diagnosed with an affective disorder such as BPI, BPII, schizoaffective disorder (SA) or recurrent unipolar depression (RUP). Parents were also interviewed and blood samples were obtained. If either of the parents were unavailable, unaffected sibs were interviewed and blood samples were obtained. Sib-pair analysis was performed under two affection models; a broad affection model where BPI, BPII, schizoaffective disorder, recurrent unipolar depression were all classed as affected, and a narrow model where only BPI individuals were classed as affected (individuals with other psychiatric diagnoses were classed as unknown). Table 2.1 shows family composition and diagnostic details of sib-pair families under both affection models.

Table 2.1: Family composition and diagnostic details of the sib-pair families under the broad and narrow affection models.

		Broad	Narrow
Number of families	with 2 affected sibs	41	23
	with 3 affected sibs	5	1
	with 4 affected sibs	2	0
	with 5 affected sibs	0	0
	with 6 affected sibs	1	0
Total number of families		**49**	**24**
Maximum number of sibpairs		83	26
Number of families	with both parents	21	12
	with one parent	25	11
	with no parents	3	1
Total number of families		**49**	**24**
Diagnostic details			
Total number of affected individuals		133	57
Bipolar I (BPI)		84	57
Bipolar II (BPII)		30	0
Schizoaffective disorder		2	0
Recurrent unipolar depression		17	0

2.1.4 Case-Control Association Collection

108 BPI probands were identified at the Mood Disorder Clinic at St. Patrick's Hospital, Dublin. These patients were diagnosed as described in section 2.1.1. Through the cooperation of the Blood Transfusion Service Board, a large sample of control blood samples was obtained. A subset of these controls was appropriately matched to BPI probands under the criteria of age, sex and county of origin/residence. If it was not possible to match to the county itself, a match was made to the nearest county.

2.1.5 Family-based Association Collection

A blood sample was obtained from 92 BPI probands who were diagnosed as described in section 2.1.1 at either the Mood Disorder Clinic at St. Patrick's Hospital, Dublin or at a similar clinic throughout Ireland. Blood samples were collected from both their parents. 78 of these 92 families had only one BPI child, whereas 14 families had two BPI children. Although family based association studies based on standard methodology such as the TDT require genotypic data from both parents, alternate methodology to allow for missing samples was also used in some studies. An additional sample of 29 families, lacking one (n=27) or both (n=2) parents, but including BPI and unaffected siblings, were included on these occasions.

2.2 DNA Extraction

50ml of blood was extracted from selected individuals in all studies. This was divided into two lots of 25ml each. One was used for DNA extraction, while the second was stored as a backup.

5 ml of blood was mixed with 20 ml of buffer A. Samples were spun at 3.5K for 20 minutes in a Sorvall RC-5B Refrigerated Superspeed Centrifuge. The pellet was resuspended in 1.25ml of buffer B. 25 µl of proteinase K (10mg/ml) was added, followed by 62.5 µl of 10% (w/v) SDS. The contents were mixed thoroughly by inversion and incubated overnight at 37°C. An equal volume of saturated phenol was added to each of the samples, which were then mixed thoroughly and spun for 30 minutes at 3000rpm in a bench top centrifuge. The aqueous layer of clear fluid was removed and extracted twice more. Two volumes of ice-cold ethanol (100% v/v) were added to each of the samples and the samples were then placed in the freezer at -20°C for 30 minutes to allow the DNA to precipitate out. Following removal using a glass 'hook', made using a pasteur pipette, the DNA was washed 3 times in 70% ethanol. The DNA was then dissolved in 200 µl of sterile water and stored in appropriately labelled eppendorfs at -20°C (stock DNA). Stock DNA plates for linkage analysis were also prepared using 96 well microtitre plates ('Concord' Biosciences); each well represented a specific DNA sample, which was recorded on master templates for gel loading and genotype scoring. 10 µl of each family members DNA was added to its corresponding well. 40 µl of cresol red dye was then added to each well to enhance visualisation. Finally, all wells were overlayed with 50 µl of mineral oil. The plates were labelled appropriately and stored at -20°C.

***All relevant solutions used in DNA extractions, PCR, staining methods, electrophoresis and DHPLC are located in Appendix A.**

2.3 Non-radioactive polymerase chain reaction (PCR)

In a standard PCR, 50 ng of DNA was added to a master mix which consisted of 2.5 µl of $MgCl_2$ PCR buffer of appropriate concentration, 4 µl of cold dNTP mix (200pM dATP, dCTP, dTTP, dGTP), 20 pM of each primer and 0.625U Taq polymerase in a total volume of 25 µl. After an initial denaturation (94°C for 5 min), 35 cycles of PCR amplification were performed, with each cycle consisting of denaturation (94°C for 1min), primer annealing (typically 56°C for 1 min) and amplification (72°C for 1 min) and then a final extension (72°C for 5 min). The appropriate magnesium concentration for a particular PCR reaction was determined by performing a magnesium curve; replica PCR reactions were set up with a range of magnesium concentrations typically from 10mM to 20mM $MgCl_2$. Other conditions such as annealing temperature, concentration of reagents and the number of cycles performed were modified to optimise amplification with particular primer pairs.

All PCR reactions were carried out on a 96 well Programmable Thermal Controller PTC-100. All primer sequences are given in Appendix B.3, B.4 and B.5. All fragments generated were analysed according to the assay for that polymorphism.

2.4 Radioactive PCR

Microsatellite markers were analysed by radioactive PCRs, performed in 96-well microtitre plates (Anachem). 1 µl of stock DNA was transferred from the stock plate to the new microtitre plate using a multichannel pipette (Hamilton) so that the new plate was a replica of the stock plate. A master mix containing all PCR reagents was prepared. The final concentration of the PCR reagents, except for the dNTP mix, were the same as previously described for non-radioactive PCRs. dATP, dTTP and dGTP were at a final concentration of 250pM while dCTP was at a final concentration of 3.125pM. Additional dCTP was provided by the inclusion in the reaction master mix of deoxycytidine 5' triphosphate, α-^{32}P (NEN) at a final concentration of 0.05 µCi. 9 µl of the master mix was added to each well using a multipipettor (Gilson) and all wells were overlayed with 50 µl of mineral oil. Inclusion of cresol red in the DNA sample plates aids the process of setting up PCR reactions since samples change colour once the master reaction mix has been added; any errors in adding the mix are easily identifiable.

Each plate was covered with a foil seal to ensure good heat transfer. All radioactive PCRs were carried out on a 96-well programmable controller PTC-100 and cycling conditions were similar to those previously described for non-radioactive PCRs, again with occasional modifications of reagent concentrations and cycling conditions to ensure optimal amplification. All primer sequences are given in Appendix B.1, B.2 and B.3.

Microsatellite markers were selected from the CHLC and Généthon marker maps (http://www.chlc.org/ChlcMaps.html/, Dib *et al.*, 1996). The marker order and distance were determined according to the published maps at Marshfield Medical Research Foundation (www.marshfieldclinic.org/research/genetics/Map_Markers/mapmaker/mapformframes.html).

2.5 Polyacrylamide Gel Electrophoresis

2.5.1 Denaturing Acrylamide Gel Electrophoresis

Radioactive PCRs were separated and visualised on 8% (w/v) acrylamide gels. Gels were pre-run for 1 hour in 1xTBE buffer to reach optimum temperature. 2µl of formamide loading dye was then added to each 10µl sample. The samples were denatured for 2 minutes at 94°C and 4-6µl of each sample was loaded on the gel. Electrophoresis was performed at a maximum constant current of 40mA and a maximum voltage of 2000V. After running the gel for the appropriate length of time (dependent on the published size of the fragment amplified) the gels were transferred to Whatman paper and were dried at 80°C for 1 hour on a speed Gel^{TM}SG 200 drier. The gel was subsequently exposed to AGFA CURTIX RP1 X-ray film for visualisation. To ensure consistency in the designation of alleles, CEPH reference individuals were included on all gels and also representative alleles were selected from each gel and at least two cross reference gels, depending on the clarity of the alleles, were additionally examined for each marker. Examples of a tetra- and di- nucleotide microsatellite gel are given in figures 2.1 and 2.2.

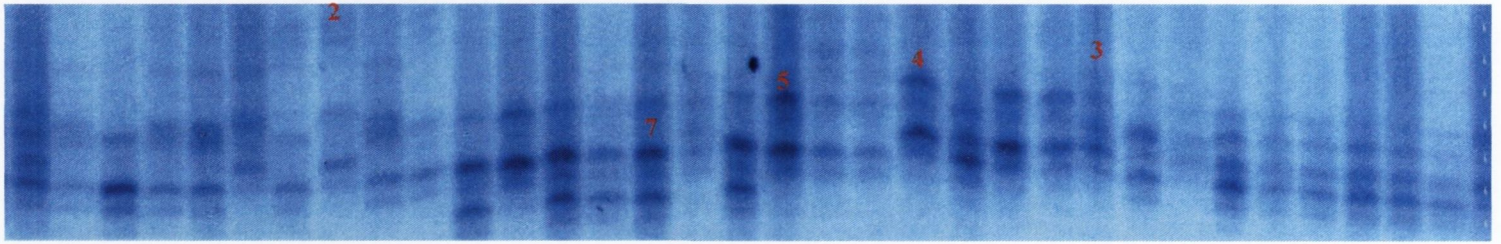


Figure 2.1: Microsatellite D14S271 (dinucleotide repeat). Allele designation is indicated by numbers. All alleles are not visible on this gel. Allele designation is finally decided upon cross-referencing across all gels for each marker.

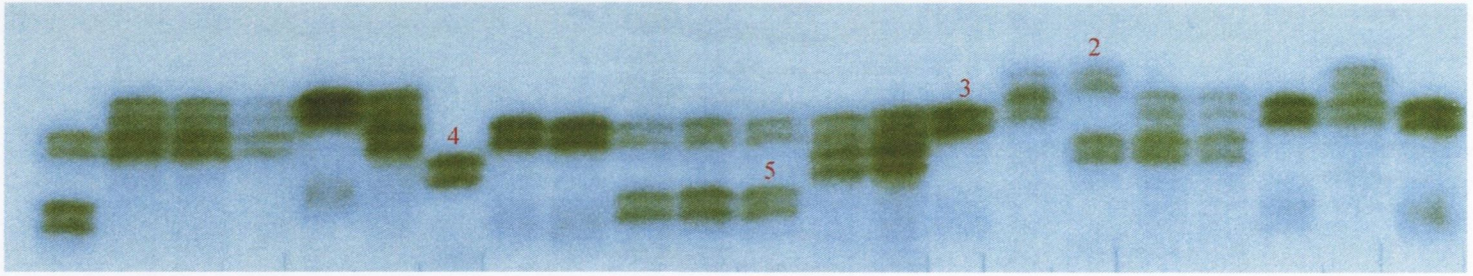


Figure 2.2: Microsatellite D17S1299 (tetranucleotide repeat). Allele designation is indicated by numbers. Allele 1 does not appear on this gel. Allele designation is finally decided upon cross-referencing across all gels for each marker.

2.5.2 Non-denaturing Acrylamide Gel Electrophoresis

Non-denaturing polyacrylamide gels were used in the assaying of certain polymorphisms such as the serotonin transporter promoter (5-HTTLPR) using a 6% polyacrylamide gel. 2-5 μl of the sample, together with 1 μl of loading dye, was loaded in each well and the gel was run in 1XTBE buffer at a constant voltage of 100V. The gel was visualised using the silver staining technique. Firstly, the gel was fixed by incubating in 250 mls of solution 1 (AppendixA) for 3 minutes. The solution was discarded and the gel was re-incubated in 250 mls of solution 1 for a further 3 minutes. Solution 1 was discarded and 500 mls of solution 2 (Appendix A) was poured over the gel and left for 15 minutes. The gel was then rinsed quickly with 2 changes of distilled water. Solution 3 (Appendix A) was the poured over the gel and it was incubated for 20 mins. This solution was then discarded and the gel was fixed for 10 mins in solution 4 (Appendix A). During all previous steps the gel was gently shaken on a shaking platform and was kept submerged in the solution. The bands were visualised using a light box and the wet gels were photographed using a Polaroid camera.

2.6 Agarose Gel Electrophoresis

Non-radioactive PCR fragments were run on 2 or 3% (w/v) agarose gels, which were stained using ethidium bromide at a concentration of 0.02mg/ml. The gels were run in 1XTBE buffer at 100V. These gels were viewed on an ultra violet illuminator and photographed using a Polaroid camera or Gene Genius bio-imaging system.

2.7 Restriction Fragment Length Polymorphisms

Aliquots of the PCR products were digested with the appropriate restriction endonuclease and buffer supplied by the manufacturer (New England Biolabs). The details of the polymorphisms analysed, and the relevant assays and digests, are given in table 2.2. The restriction digest protocol for the serotonin transporter 3'UTR, and COMT polymorphisms were previously published (Battersby *et al.,* 1999; Lachman *et al.,* 1996). The appropriate restriction endonucleases for the ERβ and GCH1 polymorphisms were determined using the known sequence from the draft of the human genome (http://genome.ucsc.edu/) and WEBCUTTER 2.0 (http://www.firstmarket.com/cutter/cut2.html).

Table 2.2: Details of Restriction Fragment Length Polymorphisms.

Gene Name	Polymorphism Location	Polymorphism	Enzyme (U/µl)	Incubation Temperature	Fragment Size	Resolution	Reference
Serotonin transporter	SNP at bp 689 in 3'UTR	G → T	*Mse* I (0.25U/µl)	Overnight @ 37°C	Uncut: 741bp Cut: 689+52bp	2% agarose gel	Battersby *et al.*, (1999) J Neurochem *72:1384-88*
COMT	SNP Codon 108/158 Exon 4	G → A	*Nla* III (0.25U/µl)	Overnight @ 37°C	Uncut: 26+29+114 bp Cut: 26+29+18+96 bp	2% agarose gel	Lachman *et al.*, (1996) Pharmacogenetics 6:243-250
	SNP at bp –287 in promoter	A → G	*Hind* III (5U/µl)	Overnight @37°C	Uncut: 350 bp Cut: 276 + 74 bp	2% agarose gel	Levinson *et al.*, (1998) *Psych Genet* 8 115-220
	SNP Codon 62, Exon 3	C → T	*Pml* I	Overnight @ 37°C	Uncut: 237 bp Cut: 148 + 89 bp	2% agarose gel	Karayiorgou *et al.*, (1998) *Biol Psych* 43: 425-431
	SNP Codon 136, Exon 4	C → G	*Bcl* I (5U/µl)	Overnight @ 50°C	Uncut: 169 bp Cut: 114 + 55 bp	2% agarose gel	Hoda *et al.*, (1996) *BBRC* 228:780-784
	3'UTR, Exon 6	Insertion C	*Bgl* I (5U/µl)	Overnight @ 37°C	Uncut: 280 bp Cut: 165 + 115 bp	2% agarose gel	Chen *et al.*, (1996) *Am J Med Genet* 67:556-9
ERβ	SNP at bp 1730 in 3'UTR of exon 8	A → G	*Alu* I (1U/µl)	Overnight @ 37°C	Uncut: 234 bp Cut: 185 + 69 bp	3% agarose gel	Rosenkranz *et al.*, (1998) J Clin Endocrinol Metab *83:4524-7*
GCHI	SNP at bp -959 in promoter	G → A	*BsrD*I (3.75U/µl)	Overnight @ 65C	Uncut: 351 bp Cut: 191 + 160 bp	3% agarose gel	Novel SNP
GCHI	SNP at bp +353 in intron 1	C → G	*Mnl* I (2U/µl)	Overnight @ 37°C	Uncut: 131 bp Cut: 79 + 52 bp	3% agarose gel	Novel SNP

2.8 Serotonin Transporter VNTR polymorphism

A standard 25 μl non-radioactive PCR reaction was performed (section 2.4) using the appropriate forward and reverse primers for the VNTR polymorphism (Appendix B.3). 35 cycles of PCR amplification were carried out, with each cycle consisting of denaturation (94°C for 1 min), primer annealing (62°C for 1 min) and amplification (72°C for 5 min). PCR fragments were run on 2% (w/v) agarose gels, visualised on an ultra violet illuminator and photographed using a Polaroid camera (section 2.6).

2.9 Serotonin Transporter Promoter polymorphism

GC rich regions, such as the promoter region of the serotonin transporter (5-HTTLPR) were amplified using a slightly modified PCR protocol; 50 ng of DNA was added to a master mix which consisted of 2.5 μl of 15 mM $MgCl_2$ PCR buffer, 4 μl of cold dNTP mix (200 pM dATP, dCTP, dTTP), 1.25 μl DMSO, 0.5 μl of 7-deaza dGTP (200 μM), 20 pM of each primer (Appendix B.3), 0.625U Taq polymerase in a total volume of 25 μl. Touchdown PCR cycling conditions were used which consisted of two cycles of 94°C for 30 secs, 63°C for 30 secs, and 72°C for 60 secs, followed by two cycles with an annealing temperature of 62°C and finally 34 cycles in which the annealing temperature was 61°C. The PCR fragments were run and visualised using non-denaturing acrylamide gel electrophoresis (section 2.5.2).

2.10 Denaturing High Performance Liquid Chromatography (DHPLC)

PCR primers were selected using the Primer3 programme (http://www-genome.wi.mit.edu/cgo-bin/primer/primer3_www.cgi) and are shown in Appendix B.4. Reliable DHPLC analysis requires amplified fragments of quality similar to that suitable for direct conventional sequencing or subcloning analysis; PCR products should ideally be between 150bp and 600bp in length, although mutations in 1.5kb fragments have been detected by DHPLC. The specificity of the primer sequences were first validated by blasting the chosen sequences against all sequences held in the non-redundant and high throughput genomic sequence (HTGS) database (GenBank) maintained by the National Centre of Bioinformatics using the BLASTN programme. Additionally, all primers chosen were required to have a G/C content of between 40-60% and a difference in annealing temperature of no more than 2°C within each primer pair. The region adjacent to and spanned by the primer is difficult to analyse, therefore to ensure that the entire region of interest was included, primers were chosen at least 40 bp between the end of the primer sequence and the start of the region of interest. To further minimise the problems of non-specific PCR amplification, 'hot-start' approaches were employed and no more than 30 PCR cycles were performed. Also, it was generally found that the quality of the PCR was significantly improved by the use of 'HotStar' Taq DNA polymerase (Qiagen) at a concentration of 1.25U/µl. HotStar taq is inactive at ambient temperature, which prevents the formation of misprimed products and primer-dimers at low temperatures. The cycling conditions were therefore slightly different to those used for the standard PCR protocol. Firstly, there was an initial activation step of 95°C for 15 minutes, followed by 25-30 cycles of 94°C for 45 secs, 61°C for 45 secs and 72°C for 1minute, and finally an extension step of 72°C for 10 minutes. 10µl of Q solution, provided with the HotStar taq polymerase, was included in the majority of 50µl PCR amplifications. This solution facilitates amplification of difficult templates by modifying the melting behaviour of DNA. All PCR products were checked on 2% (w/v) agarose gels stained with ethidium bromide, prior to DHPLC analysis, to ensure that the product was of sufficient quality.

Following PCR amplification, the double stranded DNA was heated to 95°C for 5 minutes and then cooled to 25°C at a rate of 1.5°C per minute (heteroduplex step). During this process the DNA strands separate, then randomly re-anneal, thereby ensuring heteroduplex formation between wild-type and mutant DNA. Samples were

transferred to 0.2ml thin walled eppendorfs (Treff, AG) and stored at 4°C or -20°C until DHPLC analysis was performed. Approximately 10µl of the heteroduplexed PCR reaction was required for each temperature at which the sample was analysed (see below for details).

2.10.1 Temperature Modulated Heteroduplex Analysis (TMHA)

The sensitivity of TMHA analysis is dependent on the temperature at which samples are passed through the WAVE™ DNASEP column. The column temperature is selected to ensure partial (approximately 75%) strand denaturation in the presence of a mutation, thus guaranteeing that the heteroduplex species is eluted from the column prior to the homoduplex species. Two different methods are generally employed to select the temperature of the column. Firstly, the melting behaviour of the duplex PCR amplicon sequence is analysed using WAVEMAKER software version 4.0 (Transgenomic). The primary melting point is determined using Fixman and Freire's implementation (Fixman and Friere, 1977) of Poland's algorithm (Poland, 1974). This software also displays, by means of a graphical output, the predicted temperature range at which the fragment will separate from double stranded DNA to single stranded DNA (helical fraction versus temperature), in addition to the pattern of denaturation along the fragment at any given temperature (helical fraction versus base position). For fragments with multiple melting domains, additional temperatures may be manually chosen to ensure that the entire fragment displays a helical fraction of 75% or greater (figure 2.3 and figure 2.4).

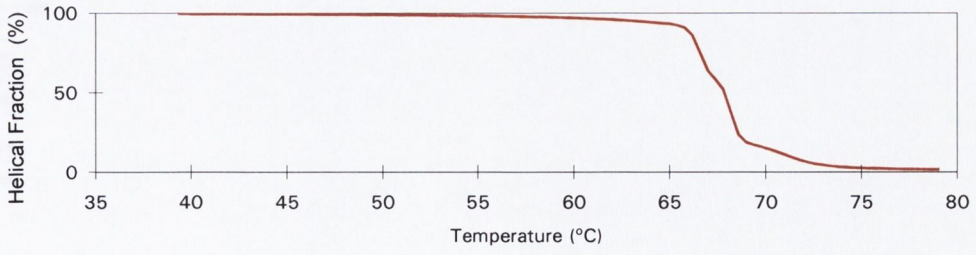


Figure 2.3: Wavemaker graphical output illustrating the melting behaviour of the fragment.

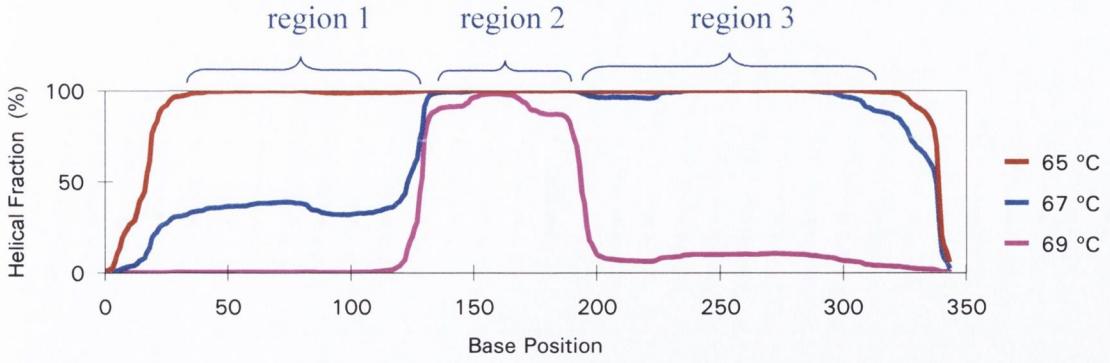


Figure 2.4: WAVEMAKER graphical output illustrating the recommended temperatures for the different melting domains of the fragment.

The analysis for mutational analysis should be set at or above 75% 'not melted' or still helical for each of the melting domains.

Region 1: recommended temperature of 65°C

Region 2: recommended temperature of 69°C

Region 3: recommended temperature of 67°C

Secondly, the melting behaviour of the sequence can be predicted using DHPLCMelt software (http://www.insertion.stanford.edu/melt.html/) (Jones *et al.,* 1999). This programme employs the Wartell and Benight version of Poland's algorithm (Wartell, 1985) to determine the melting point of the double stranded DNA. To ensure maximum sensitivity, it has been suggested that in addition to the temperature suggested by the software (n °C), each fragment must also be analysed at n + 2 °C (Jones *et al.,* 1999). It is essential that the sample is eluted from the column 2.5 – 5 minutes following sample injection. For approximately every 1°C deviation from the primary melting temperature automatically selected by the WAVEMAKER software, there is a 30 second shift in elution time. For example, if the primary melting temperature selected for a specific DNA sequence is 56°C, and one of the manually chosen temperatures is 58°C, then the time shift column on the sample table, which is created individually for each fragment, will be changed to +1 to account for the difference in temperature. Importantly, correct adjustment of the time shift, automatically produces the correct change in the percentage of Buffer B (Appendix A) flowing through the column, and therefore ensures that the fragment is eluted at the correct time. Temperatures chosen here, for the analysis of each DNA fragment, include recommendations made by both WAVEMAKER and DHPLCMelt software (table 2.3). The samples were run through

the column at the specified temperatures using a combination of buffers A and B (Appendix A) at a flow rate of 0.9ml/minute. A wave mutation standard, supplied by Transgenomic, was analysed at the beginning of each a project, to verify the quality of the analysis.

Table 2.3: Column temperatures and % of Buffer B selected for each region of the GCHI gene analysed to ensure optimal conditions for mutation detection.

GCH1 region	Column Temp	%Buffer B	GCH1 region	Column Temp	% Buffer B
Promoter: region 1 (389bp)	62	59	**Exon 1 (344bp)**	65	59
	66	58		67	58
	68	56		69	57
	70	56			
			5'UTR (265)	65	56
Promoter: region 2 (447bp)	60	61		67	54
	62	60		69	53
	65	58			
	67	55	**Exon 2 (315bp)**	51	60
				55	56
Promoter: region 3 (427bp)	56	60		57	54
	59	57			
	62	56	**Exon 3 (260bp)**	54	55
	64	55		58	53
				60	54
Promoter: region 4 (418bp)	56	62			
	58	61	**Exon 4 (313bp)**	54	56
	61	57		57	55
	63	54		63	54
Promoter: region 5 (428bp)	52	62	**Exon 5 (352bp)**	56	59
	56	59		58	56
	58	57		60	55
	61	55		62	55
Exon 1 (507)	66	60	**Exon 6 (310bp)**	58	57
	69	61		60	55
	71	63		62	54

Following WAVE fragment analysis, DNA samples that produced a deviation from the standard peak profile for each respective region, required verification that a mutation was present. The original DNA sample was re-amplified by PCR and sequenced as described in sections 2.11 and 2.12.

2.11 PCR purification for sequencing

PCRs were purified prior to sequencing to remove proteins and primers using QIA quick PCR purification kit (Qiagen) and the recommended protocol. The concentration of the purified product was estimated by eye following electrophoresis on a 2% (w/v) agarose gel.

2.12 Automated sequencing

250ng of purified PCR product and 5pM of sequencing primer was added to 8µl reaction mix from the Big Dye ^{TM} Terminator V 3.0 Cycle Sequencing Ready Reaction (ABI Prism, Applied Biosystems). Sequencing reactions were performed on an MJ Research DNA Engine. Cycling conditions were as follows; 96°C for 2 mins, 26 cycles of 96°C for 10 secs, 50°C for 5 secs and 60°C for 4 minutes. The reactions were precipitated with 100% ethanol and 0.3M sodium acetate and then washed with 70% ethanol. They were then re-suspended in 25µl of Template Suppression Reagent (ABI Prism, Applied Biosystems) and transferred to sequencing tubes, heated to 95°C for 2 minutes and then chilled on ice. The sequencing reactions were resolved on an ABI Prism^{TM} 310 Genetic Analyser according to the manufacturers instructions. The data was processed and analysed using ABI software.

2.13 Statistical and Linkage Packages.

To ensure consistency in the designation of alleles across gels and therefore assigning genotypes to individuals, various methods were employed. As previously mentioned, CEPH individuals, for whom genotypes were available, were included on all gels. Additionally, representative alleles were selected from each gel and individual cross reference gels were examined to ensure uniformity across all gels. As a further means of reducing error, the predictive power of multipoint analysis was used to generate haplotypes and therefore multiple crossovers, possibly indicative of errors in genotyping, were easily detected. Following the assignment of a genotype to each individual, data analysis was either performed by hand (marker based case-control and family association studies) or pedigree and genotypic data were entered into one of a

number of computer programmes. Types of analyses performed and the relevant programmes are described below.

2.13.1 Genehunter Version 2.0

Genehunter V 2.0, an updated version of Genehunter (Kruglyak, 1996), was used to perform two-point and multipoint non-parametric linkage analyses on genotypic data from the Irish sib-pair collection. This type of analysis requires two input files; the '.pre' file contains information regarding the family structure and the genotypic information, the '.dat' file contains information specific to the markers such as the allele frequencies, the order of the markers used and the distance between them.

The programme was run on an in house Unix machine (Oscar) and two-point and multipoint NPL scores were generated. In addition to these calculations, the multipoint inheritance information enabled maximum-likelihood haplotypes for all individuals in the pedigree to be constructed. These haplotypes were examined for multiple cross-overs, which might indicate potential genotypic errors; those found were subject to scrutiny and re-genotyping if necessary. Genehunter also measures the fraction of the total inheritance information extracted from the marker data; this information content mapping informs future analysis by indicating regions of low information content.

2.13.2 Two Locus Analysis

Two-locus analysis enables multipoint allele sharing analysis to take the evidence for linkage in one region into account in assessing the evidence for linkage in a second region. Non-parametric analysis using Genehunter-Plus (v1.1) was first performed. The Genehunter-Plus 'scan' command produces two output files, 'nullprobs.dat' and 'probs.dat', containing information about the distribution of the test statistic; one under the null hypothesis and the other conditional on the position and marker data for each kindred. These files are required to be present in the directory where the ASM programme, which performs the two-locus analysis, is run. The ASM programme also requires that a 'weights.dat' file be present. The 'weights.dat' file contains the family name and the appropriate weight, depending on the model of locus interaction, assigned to that family. Two 'weight.dat' files were created. The first, which modelled epistasis,

was created by assigning weight 0 to families with 0 or negative NPL linkage scores, and weight 1 to families with positive linkage scores. The second 'weight.dat file', which modelled heterogeneity, was generated by assigning weight 1 to families with negative linkage scores, and weight 0 to families with 0 or positive linkage scores. The ASM programme produced an output file, 'asm.out', which contained the weighted NPL score for each marker.

2.13.3 NNLD

NNLD is a programme designed to use an artificial neural network to identify association between disease status and a group of tightly linked biallelic markers using a case-control sample group. The programme requires two input files in order to run. The first is a control file, which specifies various input parameters and options required by the programme. The second is an input data file, consisting of genotype data of cases and controls. NNLD then produces two output files. The first is a result file, which contains p-values reflecting the subject's status as a case or control. The second output file represents all possible genotypes together with the score associated with each one. These scores indicate which genotypes tend to be particularly associated with susceptibility to disease.

2.13.4 Transmission Disequilibrium Test (TDT)

Family based association studies were tabulated by hand using the TDT statistic, $(b-c)^2 / (b+c)$, and p-values were calculated using the Chiprob program (Linkage Utilities Package, HGMP). When using multiplex families (i.e. families with more than one affected sib), Martin's correction was applied (Martin *et al.,* 1997), which corrects for the non-independence of multiple sibs.

2.13.5 HHRR, GHRR and MGRR

HHRR is an unmatched χ^2 analysis (test for allelic association) based on transmitted and non-transmitted parental alleles from family based association data. It is given by

HHRR = $\dfrac{(wz - xy)^2 4n}{(w + y)(x + z)4n^2}$

where: w = the number of transmissions of allele M
x = the number of transmissions of allele N
y = the number of non-transmissions of allele M
z = the number of non-transmissions of allele N
n = the number of trios in the sample

GHRR is a non-matched χ^2 analysis for case-control genotypes given by

GHRR = $\dfrac{(wz - xy)^2 2n}{(w + y)(x + z)n^2}$

where: w = the number of cases where M is present
x = the number of cases where M is absent
y = the number of controls where M is present
z = the number of controls where M is absent
n = the number of trios in the study.

MGRR is a matched analysis for case-control genotypes, giving a McNemar χ^2 statistic

MGRR = $\dfrac{(B - C)^2}{(B + C)}$

where: B = the number of times M is present in the cases and absent in the controls.
C = the number of times M is absent in the cases and present in the controls

2.13.6 Extended Transmission Disequilibrium Test (ETDT)

The ETDT package, is available at http://www.hgmp.mrc.ac.uk and all ETDT analyses were performed remotely at that site, with input files being transferred from local Unix storage using the file transfer programme (FTP). This analysis requires three separate programmes. The first, called PED2TDT takes a linkage pedigree file as input and outputs a file consisting of the genotypes of all affected subjects and their parents; alternatively an equivalent output file can be created by hand from raw genotypic data.

The second, called CALCTDT, reads in this file and for each parental genotype counts the number of times it occurs and the number of times each allele of the genotype is passed to an affected offspring. It then uses these counts to construct a matrix for the logistic regression analysis and outputs this file. The third program, called LRTDT, reads in this file containing the data matrix, performs the logistic regression and outputs chi-squared statistics.

2.13.7 Sibship Disequilibrium Test (SDT)

The SDT programme is provided along with a utility called PED2SDT, which converts LINKAGE data files into those suitable for SDT. The PED2SDT programme then goes through the LINKAGE pedigree file, breaks it up into nuclear families containing at least one affected child and then outputs an SDT datafile containing the relevant information for each family. The first output relates to the discordant-sibling analyses using the SDT method. The difference between the number of sibships in which each allele is commoner among affected siblings and the number of sibships in which it is commoner among unaffecteds. Below these differences the total number of informative sibships for each allele is output. Below the total for each allele is output a p-value for the sign test determining whether the difference is greater than would be expected by chance. Finally, the Tsdt, is output together with the p-value assuming that all alleles are informative. The p-value given for the Tsdt statistic assumes that under the null hypothesis Tsdt is distributed as a chi-squared statistic with M-1 degrees of freedom, M being the number of marker alleles. However, if only M' alleles are informative then the Tsdt statistic should be referred to as a chi-squared distribution with M'-1 degrees of freedom.

The second set of output files consist of exactly analogous results based on TDT analyses for heterozygous parents. Finally, an output is provided for the combined test comprising results from both kinds of family. The relevant subtotals are summed together and T_{both} statistic will then provide an overall test for linkage and association (Curtis *et al.*, 1999)

2.13.8 EH

EH programme is available at http://www.hgmp.mrc.ac.uk. EH is a linkage utility programme to test and estimate linkage disequilibrium between different markers/polymorphisms. The data consists of a number of individuals for each genotype that are taken to be collected at random from a population. Based on these sample data the programme estimates allele frequencies for each marker. Haplotype frequencies are estimated with allelic association (H_1) and without (H_0). The programme also provides log likelihood, chi-square and the number of degrees of freedom under hypotheses H_0 and H_1.

CHAPTER THREE

FOLLOWING UP ON PREVIOUS LINKAGE FINDINGS FROM A MODIFIED GENOME SCAN PERFORMED USING AN IRISH COHORT

Introduction

3.1.1 Linkage analysis

Linkage analysis tests for the co-segregation of a gene marker and disease phenotype within a family, to ascertain if the marker and the disease are physically linked. Presently, there are two principal approaches to performing linkage analysis. The first is parametric linkage analysis, which has proven very successful in detecting genes of moderate to large effect, as in the case of simple Mendelian disorders. It employs traditional LOD score methodology to analyse data from large multiply affected pedigrees, requires that the mode of inheritance be specified, and that the phenotype be classified on a simple affected/unaffected basis; although variable disease liability can be factored into the analysis. However, applying such model based linkage analysis to complex diseases, such as bipolar disorder, is complicated by the intrinsic characteristics of such diseases; several genes of modest effect are thought to contribute to susceptibility to complex diseases, the disease model is unknown and frequently a broad spectrum of phenotypes exist causing difficulties in classification. Although there are various options available to minimise some of the inherent difficulties in applying parametric linkage analysis to complex diseases, such as examining both dominant and recessive modes of inheritance, employing 'affecteds only analysis', multiple affection models, and penetrance models, an alternative approach of non-parametric analysis or 'model free' analysis is generally advocated for complex diseases. This methodology, which employs smaller nuclear families can essentially be categorised into two main types of analysis (i) sib pair analysis and affected pedigree member method (APM) and (ii) association studies (discussed in detail in chapter 5).

3.1.2 Affected sib pair analysis

Affected sib pair analysis infers that if a given marker is co-segregating with a disease-predisposing allele, then affected siblings of affected probands are more likely to receive the same allele identical by descent (IBD) at a closely linked marker locus than if the marker locus was segregating independently (i.e. unlinked) of the disease predisposing allele. Essentially, it is expected that two siblings would share one allele IBD. However, when the sibs are both affected with a given disease and the analysis is performed with a marker tightly linked to the disease predisposing gene, it is anticipated that the siblings would share more than one allele IBD. Non-parametric analysis of

genotypic data from sib-pair collections is performed using computer programmes such as Genehunter (see 2.13.1). The Z statistic (NPL_{all}, all affected relatives) is employed in the analysis because it has been determined that if the same allele is found IBD in many affected relatives, it is more powerful than considering one relative pair at a time. IBD sharing is also assessed on the basis of information from multiple markers and all genotyped relatives, thereby implying that this methodology is far less likely to be misled by chance sharing of alleles and is less sensitive to specification of allele frequencies. An additional essential feature of NPL analysis, especially in the case of complex disorders, is that it requires neither consideration of multiple models of inheritance, thereby avoiding corrections for multiple testing, nor advance knowledge of the correct model of inheritance, thus averting problems of mis-specification. The non-parametric method of sib-pair analysis is therefore an acknowledged powerful linkage tool. It is rapid, easy to apply and interpret, robust and most importantly Z_{all} appears to provide a nonparametric pedigree-analysis method that loses relatively little power compared to the traditional LOD score method using the correct model of inheritance. In general, for complex disorders, it is advocated that sib-pair non-parametric analysis be employed either *in lieu* of or in addition to parametric analyses, to ensure confidence in the observed result.

3.1.3 Regions of interest

Initial linkage findings from a modified genome scan for bipolar disorder within the Irish population, performed by this group (Murphy *et al.*, 2000), suggested six regions worthy of further investigation on chromosomes 17, 16, 14, 5, 1 and 9. These regions were prioritised based on the strength of the finding and evidence from previous linkage studies conducted by other independent groups. Consequently, the regions implicated on chromosomes 17, 14 and 5 were further investigated by saturating the regions with microsatellite markers and performing the appropriate non-parametric linkage analysis.

Chromosome 17

Whole genome scans to date have failed to reveal a major susceptibility locus for bipolar disorder on chromosome 17. However, there have been a few isolated reports implicating its involvement in psychiatric disorders such as bipolar disorder. In 1997, as part of the four-centre NIMH genetics initiative on bipolar disorder, a genome scan was

performed on a set of 97 families, enriched for affected relative pairs and parents where available (Edenberg *et al.,* 1997). This study presented one marker on chromosome 17 with significantly increased allele sharing (D17S969, p = 0.022). However, upon further inspection, it appears that this finding requires cautious interpretation as it was only produced for one of the three models used in the analysis, and furthermore, adjacent markers did not show similar increases in allele sharing. Recently, a genome-wide linkage study of age of onset in schizophrenia revealed a maximum likelihood LOD score of 2.54 with genome-wide p = 0.27, fulfilling Lander and Kruglyak's criteria for suggestive linkage in that it would be expected to occur once or less (0.3 times) per genome scan (Cardno *et al.,* 2001). However, this finding appeared to be almost exclusively accounted for by excess allele sharing at one marker (D17S787), with virtually no evidence of linkage over flanking markers.

A modified genome scan concentrating on gene rich regions previously performed by this group observed positive NPL scores that approached significance for two adjacent markers on 17q11-q12 (D17S841; NPL = 1.588, p = 0.06 - 6.97cM - D17S933; NPL = 1.490, p = 0.07). The gene encoding the serotonin transporter, long considered a plausible candidate for involvement in bipolar disorder, is in fact flanked by D17S841 and D17S933. As a result, this region of chromosome 17 has been the subject of numerous candidate based association studies (discussed in detail in chapter 5). Although studies to date do not produce convincing evidence for the serotonin transporter being a gene of major effect, results are consistent with a modest influence on susceptibility to bipolar disorder. Considering the presence of an excellent functional candidate for bipolar disorder in this region of chromosome 17 for which numerous positive findings have previously been obtained, and moreover, the fact that support for this region was obtained from two adjacent markers, the initial finding was followed up with an appropriate linkage study.

Chromosome 14

To date, chromosome 14 has not been a focus for candidate gene or candidate region studies for psychiatric disorders. The majority of studies providing data in regard to chromosome 14 have been undertaken within the context of genome screens for linkage. Although there is no compelling evidence for a region of interest, modestly interesting

findings for bipolar disorder and schizophrenia have been observed, and appear to fall into two distinct clusters; the first on proximal 14q and the second region on 14q22-q24. Genome scans conducted by Detera-Wadleigh *et al.,* (1998) and more recently, Cichon *et al.,* (2002) provide suggestive evidence for linkage at 14q12-q13. Further support for this region comes from a balanced translocation between proximal 14q and chromosome 18 in a patient displaying psychiatric illness that includes bipolar disorder (Overhauser *et al.,* 1998). Support for the second region of interest on 14q22-q24 comes from genome scans conducted by Detera-Wadleigh *et al.,* (1998) and McInnis *et al.,* (1998) for bipolar disorder and Faraone *et al.,* (1998) and Shaw *et al.,* (1998) for schizophrenia. Additionally, a balanced translocation between this region of chromosome 14 and chromosome 18 has been identified in schizophrenics (Goodman, 1994). However, many of the genome scans conducted for bipolar disorder and schizophrenia have failed to identify regions of interest on chromosome 14 (Adams *et al.,* 1998; Levinson *et al.,* 1998). Imprinted loci have previously been suggested to be involved in the development of bipolar disorder. Accordingly, a recent finding by Cichon *et al.,* (2002) who found evidence for putative maternally imprinted loci for bipolar disorder at 14q32 is of note, and suggests that future genome scan data be analysed using methods that adequately take a parent-of-origin effect into account.

Data from the modified genome scan revealed excess allele sharing at D14S271 (NPL= 1.388, p=0.07 under the broad model of affection and NPL=0.985, p = 0.06 under the narrow model of affection). Although a close marker, D14S281, was included in the screen, a review of the genotypic data indicated that technical problems with that marker resulted in a very low rate of genotypic data, possibly contributing to the observation that this marker did not reveal any evidence of excess allele sharing. It is noteworthy, that D14S271 is located within 14q22-q24, the region previously highlighted in the aforementioned linkage studies for bipolar disorder and schizophrenia. This observation, in addition to the actual finding itself, substantiates further investigation of this region of chromosome 14.

Chromosome 5

Independent groups have detected significant findings for schizophrenia and bipolar disorder, implicating two distinct regions on chromosome 5. The first is 5p14.1-p13.1, which was initially identified as a region of interest for schizophrenia by Silverman *et*

al., (1996). This finding was later supported by Garver *et al.,* (1998), who reported an NPL score of 2.03 for the D5S111 locus in addition to support from flanking markers. Although these studies are consistent with linkage in this region, five multiplex schizophrenia pedigrees from eastern Canada failed to produce evidence for linkage to 5p14.1-13.1 (King *et al.,* 1997). Furthermore, recent linkage studies involving schizophrenia and 5p have also reported negative findings (Delisi *et al.,* 2002; Maziade *et al.,* 2001). Using a set of 20 North American pedigrees, Kelsoe *et al.,* (1996) indicated suggestive evidence of linkage to 5p15 for bipolar disorder, the region where the dopamine transporter has been localised. Recently, a genome survey conducted by Kelsoe *et al.,* (2001) provided additional support for their previous finding. The second region of interest on chromosome 5, namely 5q22-33, emerged from independent schizophrenia studies (Schwab *et al.,* 1997; Straub *et al.,* 1997). The involvement of this region in schizophrenia was further supported by findings obtained by Paunio *et al.,* (2001) and Gurling *et al.,* (2001). A genome scan of five genetically related pedigrees from the Saguenay-Lac-St. John region in Quebec, conducted by Shink *et al.,* (1998), presented a suggestive finding for 5q31.3 – q35.1 and bipolar disorder. Support for this region has come from several independent studies (Edenberg *et al.,* 1997; Garner *et al.,* 2001). Detera-Wadleigh *et al.,* (1998) presented a finding from a genome scan for bipolar disorder that revealed excess allele sharing in a region located at 5q34. However there have been numerous negative reports for the involvement of loci on chromosome 5 in either schizophrenia or bipolar disorder. These include Faraone *et al.,* (1998) and Levinson *et al.,* (1998) for schizophrenia, and Raeymaekers *et al.,* (1997) and Berrettini *et al.,* (1997) for bipolar disorder. It is intriguing however, that the regions on chromosome 5 implicated in schizophrenia are also alluded to in bipolar disorder. This observation raises the possibility that common susceptibility genes may be involved and that the relationship between the two disorders may be more complex than previously thought.

The modified genome scan performed on the Irish cohort of sib-pair families revealed suggestive evidence for linkage at D5S658 (NPL=1.264 p = 0.07 under the broad model of affection), which is located within the previously proposed region of 5q31.1-q35.1 for harbouring a susceptibility locus for bipolar disorder. Unfortunately, the marker set on which this study was based did not include nearby flanking markers to assess support for the finding. Nonetheless, considering the evidence indicating the possibility of a

susceptibility locus within this region of chromosome 5 for bipolar disorder, it was essential to perform further investigations to determine the accuracy of the original finding.

3.1.4 Two-locus analysis

As previously mentioned, transmission of complex diseases appears to exhibit composite actions and interactions between multiple genetic and environmental factors. It has been proposed that genetic analyses that allow for the simultaneous consideration of susceptibility from multiple regions may enhance the ability to map genes for complex disorders (Schork *et al.,* 1993). Cox *et al.,* (1999) suggested that when evidence for a statistical interaction between regions has been obtained from preliminary studies, it is possible by weighting families according to their evidence for linkage, to incorporate linkage evidence from one region in assessing evidence for linkage in a second region, modelling both epistasis and heterogeneity. The traditional multipoint allele sharing approach, initially described by Kruglyak *et al.,* (1996) and later extended by Kong and Cox (1997), usually weighted families (or pairs) equally. Although equal weighting of families is appropriate for many situations, especially where information is limited, it has been advocated that when such information as pedigree structure, number of affecteds and/or their evidence for linkage at a particular location is available, that it be incorporated in the analysis by altering the weights for families. Basically this approach involves modelling positive interactions (such as epistasis) by assigning weight 0 to families with 0 or negative linkage scores at a particular locus and weight 1 to families with positive linkage scores. Heterogeneity is modelled by designating weight 1 to families with negative linkage scores and weight 0 to families with 0 or positive linkage scores. A modified version of the Genehunter programme has been developed to perform such two-locus analysis. Non-parametric analysis is first conducted on the actual data using GENEHUNTER-PLUS (v1.1 or later) and then many replicate weighting files are generated randomly and used to calculate the final weighted NPL scores. As this analysis involves multiple comparisons, Bonferroni adjustment is required. However, even with this adjustment such secondary analysis may increase the overall false positive rate because they are designed to strengthen the support for regions that do not themselves meet genome-wide criteria for significance. Therefore, more stringent criteria for determining

significance is employed, where p values must be lower than 0.01 instead of 0.05. Employing this methodology in the search for susceptibility loci for diabetes, Cox *et al.*, (1999) demonstrated that the CYP19 region on chromosome 15, which in a previous genome scan conducted on a Mexican American cohort did not meet genome-wide criteria for significance, displayed significant correlations between their family NPL scores and the family NPL scores at NIDDM1 on chromosome 2, even after Bonferroni correction for the number of correlations was employed. They proposed that in collections of Mexican American families similar in size to that of the original genome scan (408 individuals), the linkage signals from both regions might be comparable and only modest unless the interaction is properly accounted for. Thus, they concluded that conducting analyses to identify potential interactions could possibly assist in resolving some of the difficulties encountered in replicating results obtained in genome scans for bipolar disorder.

3.2 Results

The inclusion of genotypic errors in linkage analyses can have a significant effect on the statistical result. Consequently, various measures were undertaken for all conducted analyses to ensure consistency in allele designation. CEPH individuals, for whom published genotypes were available, were included for each marker, thereby enabling blind comparison of the designated genotype with the standard genotype for the CEPH individual. Representative alleles were selected from each gel and cross-referencing was performed to ensure uniformity across the gels. Also, the predictive power of multipoint analysis was employed to infer haplotypes for each chromosomal region, enabling the identification of multiple crossovers, which may be indicative of genotyping errors. Incorporation of such processes into each analysis ensured that the data generated was robust and reliable. Also, the population allele frequency for each marker was determined using the parental genotypes, rather than the standard 1/n, and thus are representative of the frequency of the alleles in the Irish population.

3.2.1 Chromosome 17

A collection of 49 Irish families comprising of 210 individuals, (details in 2.1.3) were genotyped for 10 microsatellite markers spanning a total distance of 67.87cM on chromosome 17q. The original two markers employed in the previously described modified genome scan, D17S933 and D17S841, were re-genotyped for all individuals, thus leading to a more complete data set given that genotypes were not available for a number of key individuals in the original screen. The present cohort includes an additional family previously not available for inclusion. Genotypic data from an additional marker formerly genotyped in a separate analysis was also incorporated in the analysis performed here. A further marker, D17S1294, was also genotyped for all individuals but was not included in the multipoint analysis due to an inconsistency in determining its exact location. Table 3.1 summarises the two-point NPL results obtained for each of the microsatellite markers under the broad and narrow definitions of the disease with their associated p values. Also included in this table are the genetic distances between each of the markers, as produced by the Marshfield map (http://research.marshfieldclinic.org/genetics/Map_Markers/mapmaker/MapFormFrames.html).

Table 3.1: NPL scores and associated p values obtained for chromosome 17 using Genehunter V2 under the broad and narrow affection models.

Marker	Distance (cM)*	Broad NPL	Broad P value	Narrow NPL	Narrow P value
GATA158H04		0.3936	0.3457	0.9969	0.1602
	8.87				
D17S1303		-0.2109	0.5815	1.2027	0.1132
	8.4				
ATA78D02		-0.0544	0.5198	0.6919	0.2438
	12.66				
D17S2196		-0.2319	0.5896	0.4304	0.3332
	6.12				
D17S841		**1.6207**	**0.0533**	1.2087	0.1132
	2.67				
D17S798		1.2807	0.1007	**1.4808**	**0.0703**
	4.3				
D17S933		**1.5388**	**0.0627**	**1.7876**	**0.0367**
	4.3				
D17S1299		-0.2915	0.6128	0.7855	0.2189
	4.64				
D17S806		-0.0219	0.5068	-0.1655	0.5642
	15.71				
D17S948		0.4641	0.3202	0.4791	0.3146
	34.30				
D17S784		0.3022	0.3797	-0.4278	0.6642

* **The distance given in cM refers to the genetic distance between each marker and the subsequent marker in the list.**

Figure 3.1 depicts the NPL multipoint analysis. The Z_{all} results for 17q, inclusive of all 11 markers under the broad and narrow affection models, are shown graphically. In contrast to the conventional two-point analysis, which tests for the cosegregation of the disease and a single marker, multipoint analysis uses all the available information from multiple markers which are in close proximity to each other. This may enable greater precision in determining the most likely position of the disease locus, or it may permit the exclusion of the disease locus from a particular region on the chromosome.

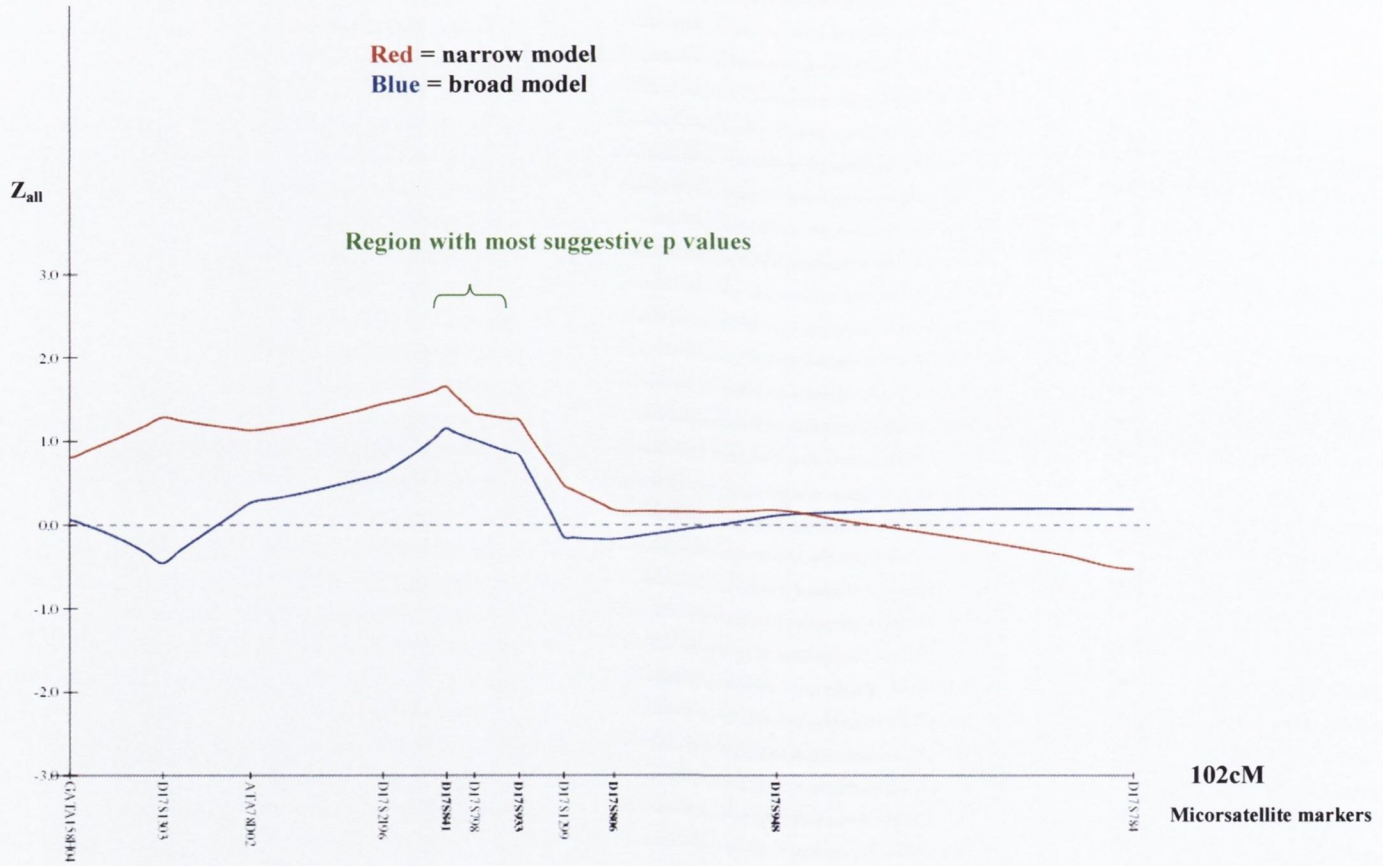


Figure 3.1: NPL multipoint analysis for chromosome 17q. The region with the most suggestive p values is indicated by a green brace from D17S841 to D17S933. The markers employed in the original genome scan are highlighted in bold.

3.2.2 Chromosome 14

This study was performed using the same, slightly expanded Irish cohort, as previously described for the chromosome 17 analysis. A total distance of 85.43cM on 14q11.2-q24.3 was investigated using genotypic data from 13 microsatellite markers, which included the original two markers from the modified genome scan, D14S281 and D14S271. Five of these markers were located distally to D14S281; three proximally to D14S271 and three were evenly spaced between D14S271 and D14S281. Table 3.2 summarises the two-point NPL results obtained for each of the microsatellite markers under the broad and narrow definitions of the disease and their associated p values. The genetic distances between each of the markers, as produced by the Marshfield map, are also given in table 3.2. The NPL multipoint analysis inclusive of all 13 markers under the broad and narrow affection models are displayed graphically in figure 3.2.

Table 3.2: NPL scores and associated p values obtained for chromosome 14 using Genehunter V2 under the broad and narrow affection models.

Marker	Distance (cM)*	Broad NPL	Broad P value	Narrow NPL	Narrow P value
D14S1023		0.444	0.327	0.448	0.326
	13.23				
D14S972		1.16	0.121	**1.88**	**0.031**
	10.2				
ATA76C02		1.14	0.127	1.06	0.144
	8.9				
D14S599		0.229	0.408	0.635	0.264
	9.2				
D14S1027		0.322	0.372	0.759	0.226
	5.9				
D14S281		1.351	0.089	**2.486**	**0.006**
	4.6				
D14S980		0.926	0.177	**2.219**	**0.013**
	2.8				
D14S274		**1.552**	**0.061**	**1.76**	**0.039**
	3.6				
D14S592		0.969	0.166	**2.112**	**0.017**
	2.4				
D14S271		1.047	0.148	**1.845**	**0.033**
	6.4				
D14S588		0.993	0.160	0.308	0.379
	7.0				
D14S71		1.41	0.08	**1.817**	**0.034**
	11.2				
D14S974		-0.243	0.594	-0.034	0.515

*** The distance in cM refers to the genetic distance between each marker and the subsequent marker in the list.**

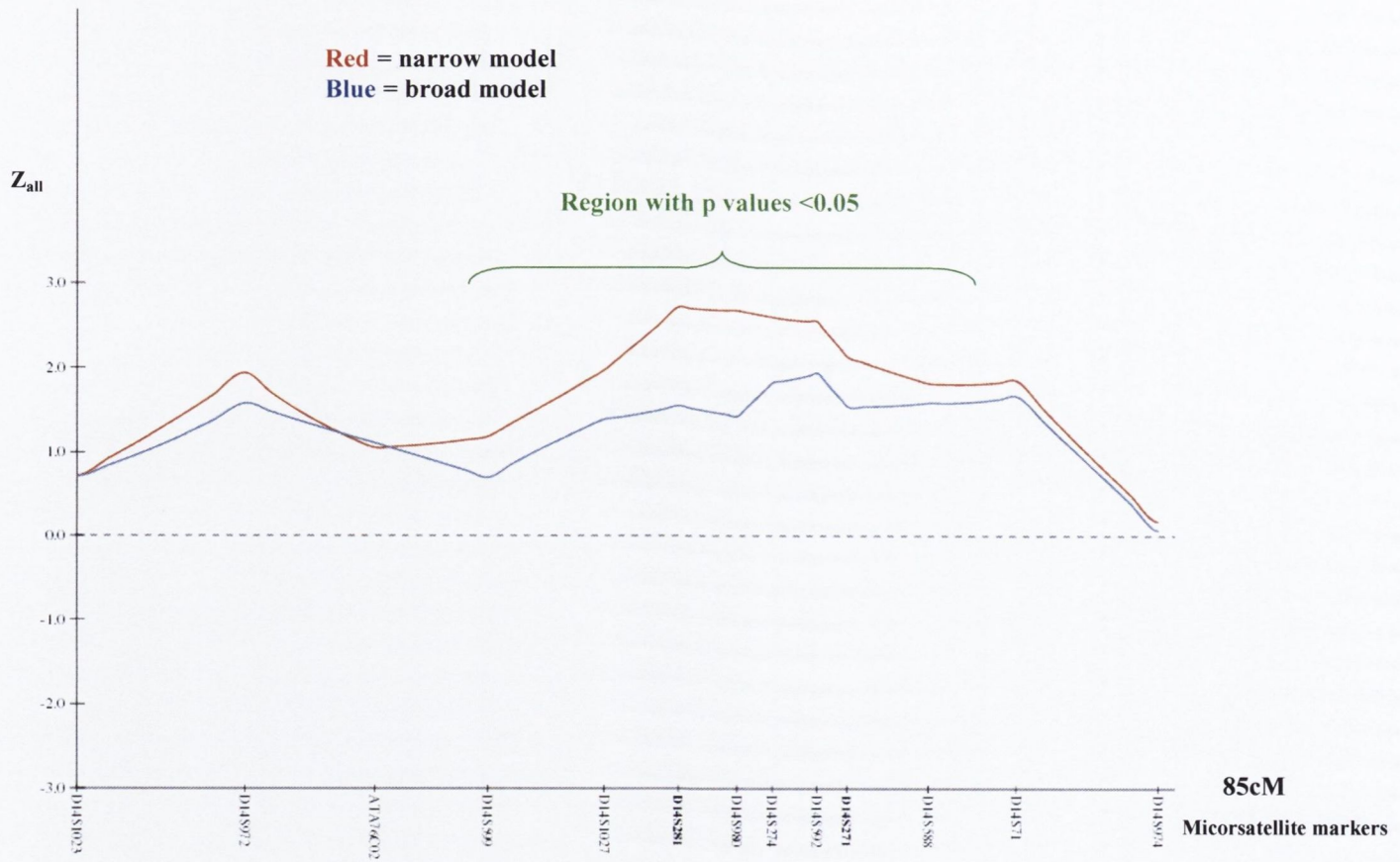


Figure 3.2: NPL multipoint analysis for chromosome 14q. The extensive region with p values <0.05 is indicated by a green brace from D14S599 to D14S71. The markers used in the original genome scan analysis are highlighted in bold.

3.2.3 Chromosome 5

Again, using the same sib-pair cohort as previously employed for the chromosome 17 and chromosome 14 analyses, all individuals were genotyped for five microsatellite markers spanning a total distance of 26.1cM on 5q. The original marker, D5S658, employed for the modified genome scan, was re-genotyped, in addition to four markers evenly spaced either side of D5S658. Table 3.3 illustrates the two-point NPL scores obtained for these markers under the broad and narrow affection models, and figure 3.3 illustrates the results of the multipoint analysis for this interval.

Table 3.3: NPL scores and associated p values obtained for chromosome 5 using GenehunterV2 under the broad and narrow affection models.

Marker	Distance (cM)*	Broad NPL	Broad P value	Narrow NPL	Narrow P value
D5S1505		-0.016	0.505	0.445	0.326
	9.5				
D5S816		0.353	0.361	-0.571	0.713
	3.6				
D5S658		-0.229	0.587	-0.043	0.515
	4.6				
D5S436		-0.214	0.582	-0.531	0.699
	8.4				
D5S673		0.368	0.355	-0.007	0.506

*** The distance in cM refers to the genetic distance between each marker and the subsequent marker on the list.**

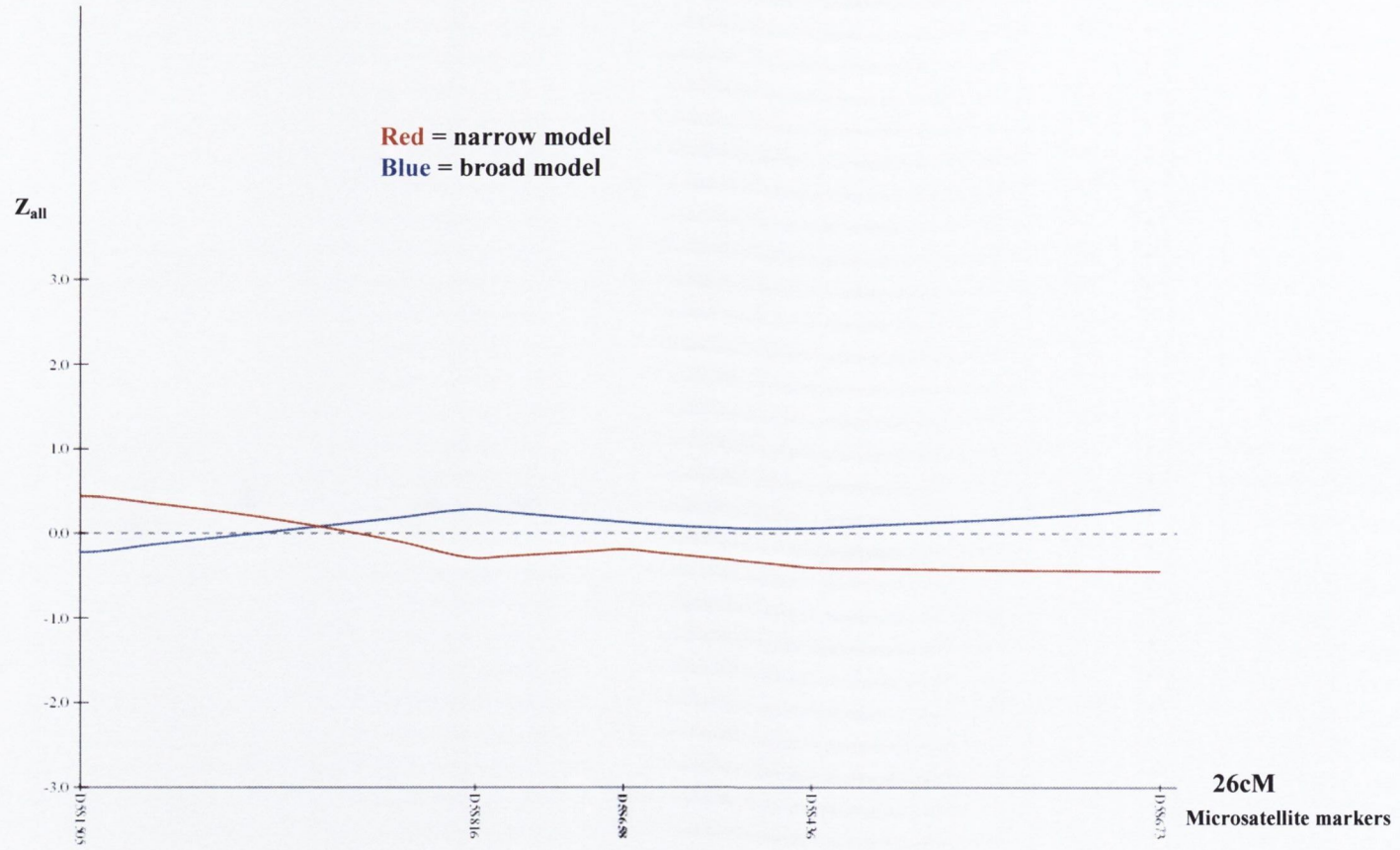


Figure 3.3: NPL multipoint analysis for chromosome 5q. The marker employed in the original genome scan analysis is highlighted in bold.

3.2.4 Two-locus analysis

It was decided to reanalyse the data for chromosome 17 taking into account the linkage data for chromosome 14. An examination of the region revealed that marker D14S281 provided the strongest evidence for linkage in the region of interest on 14q; therefore each of the 24 families with a BPI-BPI sib pair, and hence included in the narrow model of affection, were assigned a weighting based on the Genehunter NPL scores obtained for this marker. Epistasis was modelled by assigning weight 0 to families with 0 or negative linkage scores at D14S281 and weight 1 to families with positive linkage scores ($weight_{0-1}$). Heterogeneity was modelled by assigning weight 1 to families with negative linkage scores at D14S281 and weight 0 to families with 0 or positive linkage scores ($weight_{1-0}$). Table 3.4 shows the family NPL trends across the region examined, and table 3.5 presents the two-point NPL scores for each family at D14S281, and the weighting under epistasis and heterogeneity.

Table 3.4: Family NPL trend across the region examined under the narrow model only - green = positive NPL score, red = negative NPL score, yellow = 0 NPL score, blue = cannot be determined. Numbers indicate family code.

Marker	D14S599	D14S1027	D14S281	D14S980	D14S274	D14S592	D14S271	D14S588	D14S71	D14S974
Family										
1										
2										
3										
5										
6										
8										
10										
12										
14										
16										
19										
21										
22										
24										
25										
27										
33										
34										
36										
37										
38										
39										
41										
48										

Table 3.5: Two-point NPL score for each family at D14S281 and the weighting applied under both models of epistasis and heterogeneity.

Family	D14S281 NPL score	$Weighting_{(0-1)}$ Epistasis	$Weighting_{(1-0)}$ Heterogeneity
1	0.4472	1	0
2	0.8441	1	0
3	-0.7071	0	1
5	-0.4472	0	1
6	1.414	1	0
8	0.4472	1	0
10	-0.6257	0	1
12	0.6249	1	0
14	0.1369	1	0
16	0.7071	1	0
19	0.5399	1	0
21	0.0000	0	0
22	0.0000	0	0
24	1.3416	1	0
25	1.0909	1	0
27	0.6191	1	0
33	1.4142	1	0
34	0.7829	1	0
36	0.0000	0	0
37	0.4589	1	0
38	1.6808	1	0
39	0.4472	1	0
41	0.2589	1	0
48	0.7071	1	0

Before using the ASM programme, which performs the two-locus analysis, non-parametric analysis was first performed using Genehunter-Plus. This analysis produced two output files, 'nullprobs.dat' and 'probs.dat', which contained information about the distribution of the test statistic; one under the null hypothesis and the other conditional on the position and marker data for each kindred. In the presence of these two files, ASM used the weighting files for chromosome 14 (D14S281) to reanalyse the chromosome 17 data under the narrow definition of affected sib-pairs (BPI-BPI).

Table 3.6 presents the original chromosome 17 two-point NPL scores, both from the modified genome scan and from the extended linkage anaysis, in addition to the weighted NPL scores under both models of epistasis and heterogeneity for each of the chromosome 17 markers, based on the evidence for linkage at D14S281.

Table 3.6: Original (from modified genome scan and extended linkage analysis) and weighted two point NPL scores for each of the chromosome 17 markers based on the evidence for linkage at D14S281.

Modified genome scan markers	NPL score from genome scan	Extended linkage analysis markers	NPL score from extended analysis	Weight $_{(0\text{-}1)}$ NPL score (Epistasis)	Weight $_{(1\text{-}0)}$ NPL score (Heterogeneity)
		GATA158H04	0.9969	1.205	0.2616
		D17S1303	1.2027	1.017	0.0933
		ATA78D02	0.6919	0.5104	0.0456
		D17S2196	0.4304	0.9747	-1.0367
D17S841	1.588	**D17S841**	1.2087	1.0676	0.7628
		D17S798	1.4808	1.1438	0.9246
D17S933	1.490	**D17S933**	1.7876	1.2347	1.2153
		D17S1299	0.7855	0.6703	0.5164
D17S806	-0.036	**D17S806**	-0.1655	-0.8248	0.7746
D17S948	0.709	**D17S948**	0.4791	0.3549	0.4858
		D17S784	-0.4278	0.4292	-2.1338

3.3 Discussion

3.3.1 Chromosome 17

The original modified genome scan included four markers located on 17q, spanning a region of 31.82cM. Of the investigated microsatellite markers, two produced results that approached significance under the broad definition of the disease, (similar to that used in the present analysis): D17S841, NPL= 1.588, p value = 0.06; D17S933, NPL= 1.490, pvalue = 0.07. Under the narrow definition of the disease (for scan analysis only BPI individuals were classed as affected, individuals with other psychiatric diagnoses were classed as unknown), positive NPL scores were produced for both markers: D17S933, NPL= 1.15, p value= 0.05; D17S841, NPL= 0.734, p value= 0.14. However, the significance of the result for D17S841 was considerably reduced. In this present study, the original four markers were re-genotyped, in addition to a further seven markers, expanding the total region under investigation to 102.17cM. It is apparent from table 3.1 and figure 3.1, that under the broad definition of the disease, D17S841 and D17S933, as well as an intermediate marker D17S798, all generated positive NPL scores comparable to those produced in the original analysis, with D17S841 again yielding the most significant result: D17S841, NPL = 1.62, p value = 0.053. However, under the narrow definition of the disease, which for this analysis includes only those families with a BPI-BPI sib-pair, there is a notable difference in the NPL scores produced for D17S841 and D17S933 to those obtained in the original analysis: D17S841, NPL= 1.21, pvalue = 0.113; D17S933, NPL= 1.79, p value = 0.037. The disparity in the results obtained in the present analysis as compared to those generated in the original scan analysis most likely reflects the inclusion of the genotypic data from an extra family, as well as the availability of more genotypic data overall (all markers included in this analysis were typed to greater than 95% completion).

While this study did not produce compelling evidence for a susceptibility locus for bipolar disorder, the most significant positive NPL scores continued to be produced for markers within a 10cM region that is known to harbour the serotonin transporter gene, an acknowledged potential candidate for bipolar disorder. The results obtained here justify candidate gene studies with the serotonin transporter gene, and also possibly saturating the region with genetic markers and employing linkage disequilibrium studies to fine map this region of interest on chromosome 17q.

3.3.2 Chromosome 14

The original modified genome scan employed only two microsatellite markers for examination of chromosome 14, D14S281 and D14S271. Linkage analysis generated a result for D14S271worthy of further investigation: D14S271, NPL= 1.388, p value = 0.07 under the broad model and under the narrow model of affection, NPL= 0.985, p value = 0.06.

The data in table 3.2 for the present analysis indicates that under the broad model of affection, no one marker produced a two-point NPL score that reached significance, although NPL scores obtained for D14S281, D14S274 and D14S71 produced associated p values that approached significance. However, when multipoint analysis was applied, figure 3.2, it was found that the region that spans from D14S980 to D14S71, a distance of approximately 15.17cM, produced a p value of <0.05. Under the narrow definition of the disease, it was observed that seven of the microsatellite markers produced two-point NPL scores with p values considerably less than 0.05 (table 3.2). Employing multipoint analysis, the region that produced a p value of <0.05 extended from D14S599 to D14S71 which spans approximately 34.92cM. The most significant NPL score that was obtained was 2.714, with an associated p value of 0.003, corresponding to D14S281. Comparison of the results from this analysis, to those of the modified genome scan, revealed considerable disparities. These differences are most probably due to retyping and inclusion of missing genotypes, and the addition of an extra family to the analysis. When evaluating the results from the present analysis, it is encouraging to find that support from flanking markers is observed for the marker that produces the most significant NPL score, thereby reducing the possibility of it being a false positive result.

It was anticipated that the inclusion of additional flanking markers at approximately 10cM intervals either side of the extensive region of interest would bring down the outlaying multipoint NPLs, and thus reduce the size of the final region where p<0.05. However, it has been observed that there has been only modest change to the multipoint NPL scores upon inclusion of additional markers. Therefore, it appears from the consistency and supporting nature of the results of this analysis, that chromosome 14q22-q24 may indeed harbour a susceptibility locus for bipolar disorder, and hence continued examination of this region is justified. Published data relating to this region

of chromosome 14, does not reveal the presence of an obvious candidate gene for bipolar disorder. It is therefore proposed, that further investigation of this region of interest would involve screening the region for potential candidates using various web based tools, which possess extensive information relating to all annotated genes located in this region of chromosome 14q.

3.3.3 Chromosome 5

The original modified genome scan utilised microsatellite markers, D5S658 and D5S668, separated by 72.8cM, to examine the evidence for linkage on chromosome 5. D5S668 did not produce any significant result either under the broad or narrow affection models: under the broad model, NPL= -0.3445, p value = 0.64 and under the narrow model, NPL= 0.459, p value = 0.21. However, D5S658 did generate an interesting result, especially under the broad model of affection, where an NPL score of 1.264 with an associated p value of 0.07 was obtained. Under the narrow model the NPL score while still positive did not approach significance: NPL= 0.74, p value = 0.11. Although this was a modest finding for chromosome 5q, the observation for D5S658 occurred 4.5 cM proximal to the 24cM interval (D5S436-D5S415) where previous positive LOD scores were observed (discussed in 3.1.3). Simulation studies imply that a true susceptibility locus is flanked by a region of markers displaying significant ($p<0.05$) allele sharing (Goldin *et al.,* 1995; Terwilliger *et al.,* 1997). However, the nature of the modified genome scan was such that makers flanking D5S658 were not included in the analysis. Nonetheless, given the positive NPL score, albeit modest, and the location of this marker in relation to previous findings, it was decided to investigate this finding further. What is immediately apparent upon examination of the data in table 3.3 and figure 3.3, is that the original positive NPL score that was obtained under the broad affection model in the scan analysis has completely disappeared under both models of affection, and the flanking markers appear to support this new finding. As the definition of the broad model of affection status remained the same for this present analysis, it can be concluded that the loss of the original finding was due to the inclusion of significantly more genotypic data, from which the analysis was performed. It is likely that the original finding was a false positive, and that there is no linkage evidence for a region of interest for bipolar disorder on chromosome 5q in this Irish cohort.

3.3.4 Two Locus Analysis

It has long been acknowledged that composite actions and interactions between multiple genetic and environmental factors are involved in complex disorders, and consequently genetic analyses that simultaneously consider susceptibility from multiple regions have been advocated as a means of enhancing the ability to map genes for these disorders. Kong and Cox (1997) proposed an approach that assesses the evidence for statistical interactions between unlinked regions. It enables multipoint allele sharing analysis to take the evidence for linkage in one region into account in assessing the evidence for linkage in a second region. In the analysis performed here, the evidence for linkage at D14S281, the microsatellite marker yielding the most significant NPL scores, was used to reassess the evidence for linkage at 17q under two possible models of loci interaction – epistasis which refers to gene-gene interactions between loci, and heterogeneity which alludes to the possibility that mutations at two or more loci each result in similar phenotypes. Given that the evidence for linkage for chromosome 14q22-q24 was strongest under the narrow model of affection, the 24 families comprising this cohort were weighted according to their evidence for linkage at D14S281 (as described in 3.2.4). Based on this weighting, the ASM programme (galton.uchicago.edu on the /pub/kong directory) was employed to examine the evidence for linkage at each of the chromosome 17 markers independently under both models of epistasis and heterogeneity. The results of the two-point analysis are presented in table 3.5. Comparison of the original NPL values to those observed under both the epistasis model and the heterogeneity model indicate that weighting the analysis based on linkage on chromosome 14 has not provided evidence for linkage on chromosome 17, with NPL values actually decreasing. Given this, simulation studies to provide p values for the weighted NPL scores were not undertaken. Examination of the multipoint figures (not presented here) also revealed similar findings under both models. Conclusions of the analysis performed here would therefore suggest that there is no evidence implicating interactions between loci on chromosome 14q and 17q as contributory factors to bipolar disorder.

Cox *et al.,* (1999) using this approach, showed that the interaction of genes on chromosome 2 (NIDDM1) and 15 (near CYP19), contributed to susceptibility to type II diabetes in Mexican Americans. In the baseline analysis, the LOD in the CYP19 region was 1.3 but increased to 4.0 when families were weighted by their evidence for linkage

at NIDDM1, using the epistasis model. They suggested that the linkage signals from both regions might be comparable and only modest unless the interaction is properly accounted for. They also implied that many of the difficulties encountered in replicating findings obtained in genome scans for complex disorders could be considerably reduced by conducting analyses to identify potential interactions.

CHAPTER FOUR

INVESTIGATIONS OF GTP CYCLOHYDROLASE I (GCHI) – A PLAUSIBLE CANDIDATE GENE FOR BIPOLAR DISORDER?

4.1 Introduction

4.1.1 Previous Linkage Analysis

Initial findings from a modified genome scan, previously undertaken by this group, revealed excess allele sharing on chromosome 14q, at the marker D14S271. Due to the nature of this genome scan (discussed in section 1.13), flanking markers were not available to determine if there was support for this positive finding. Further linkage analysis was performed in this region of chromosome 14, which initially involved re-genotyping 49 Irish sib-pair families, including some previously untyped individuals (described in section 2.1.3) through the original two markers, D14S271 and D14S281. In addition, a further eleven microsatellite markers positioned either proximal, distal or between the original two markers, were incorporated into the analysis thereby expanding the original study to cover, in detail, a total distance spanning 84cM. Single point and multipoint analyses were performed using Genehunter and are described in detail elsewhere (see chapter 3). In effect, p values <0.05 were observed for a region extending over 34.92cM on chromosome 14q22-24, with a peak NPL score of 2.714 and an associated p value of 0.003 being obtained for the marker D14S281. Support for this finding was also obtained from flanking markers, thus reducing the possibility of the finding being a false positive. Linkage analysis results would therefore appear to indicate that the region 14q22-24 might possibly harbour a susceptibility locus for bipolar disorder within the Irish population.

4.1.2 Screening region of interest for potential candidate genes

Unfortunately, there is limited published data to suggest that 14q22-24 may harbour susceptibility genes contributing to psychiatric disorders. However, since the launch of the human genome project, thousands of genes distributed across the whole genome have been identified. The vast array of information relating to these genes, for example, their structure, expression patterns, functional profiles etc, has been collated and made freely available to the public by means of various databases (http://ensembl.org/; http://genome.ucsc.edu/; http://www.ncbi.nlm.nih.gov/). This central resource, which is continuously being expanded as new genes are identified, provides an initial, highly valuable, starting point for the screening of a putative chromosomal region for genes that may contribute to the disease in question. A pivotal step in the identification of disease causing genes is the ability to detect polymorphisms or even disease causing

mutations. Previously, techniques such as single stand conformational polymorphism electrophoresis (SSCPE), heteroduplex analysis, denaturing gradient electrophoresis and direct sequencing were employed for the investigation of candidate genes for a particular disease. For example in the case of retinitis pigmentosa, SSCPE and direct sequencing enabled the identification of a three base pair deletion at codon 118 or 119, resulting in the loss of one of a pair of highly conserved cysteine residues in the predicted third transmembrane domain of the peripherin-RDS protein, a protein proposed to be involved in photoreceptor morphogenesis (Farrar *et al.,* 1991). While these methodologies enabled the detection of numerous genetic variants, many remained labour intensive, time consuming, lacked sensitivity and were expensive. However, significant developments have been made in the area of variant detection, and presently there are methodologies available, which are >98% sensitive, are almost fully automated and relatively inexpensive to maintain (discussed in detail in 4.1.4).

4.1.3 GTP Cyclohydrolase I (GCHI) – A functional candidate for bipolar disorder?

By consolidating the information obtained from the various databases and using the aforementioned criteria for the selection of potential candidates, the GTP cyclohydrolase I (GCHI) gene was selected for further investigation (see 4.2.1).

GCHI catalyses the conversion of GTP to dihydroneopterin triphosphate, the first and rate limiting step in the biosyntheses of tetrahydrobiopterin (BH_4) (Thőny *et al.,* 2000). The biosynthesis of BH_4 is then completed by the subsequent action of 6-pyruvoyl-tetrahydropterin synthase and sepiapterin reductase (figure 4.1).

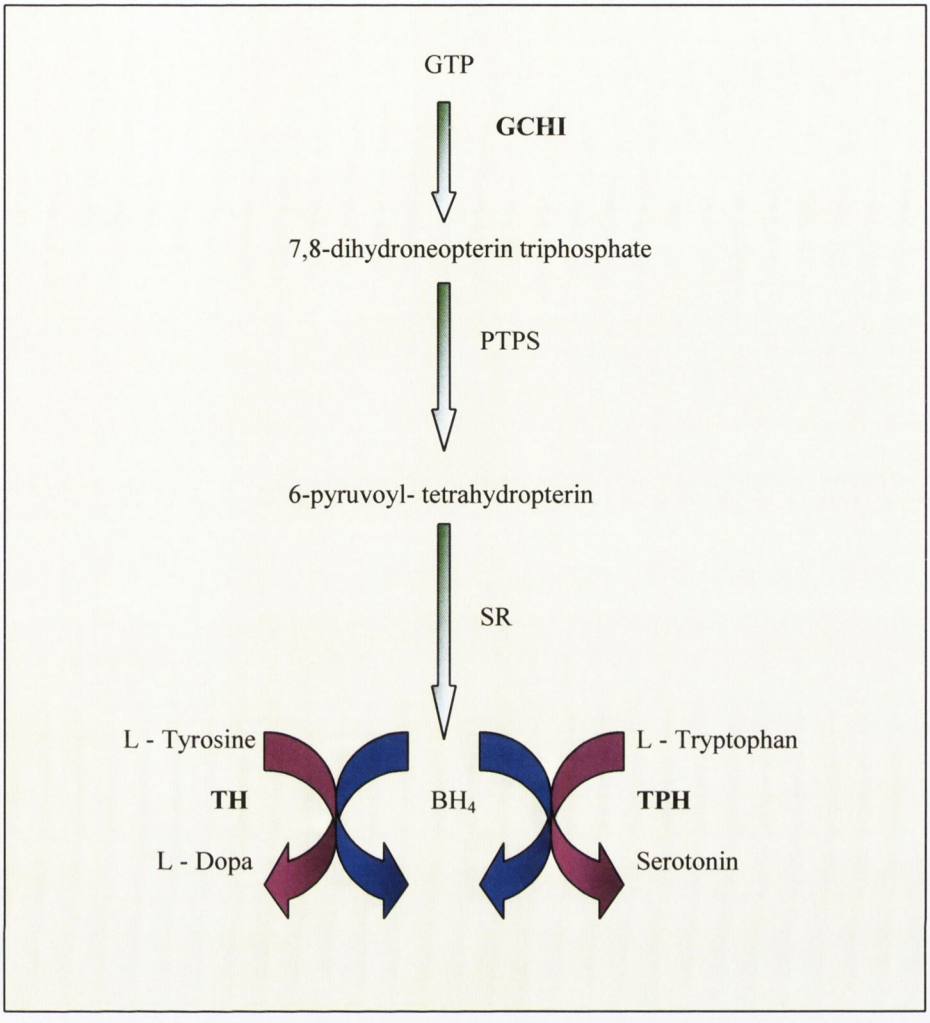


Figure 4.1: Schematic presentation of the biosynthesis of BH_4 and its use as a cofactor for the hydroxylation of L-Tyrosine and L-Tryptophan in the biosynthesis of the neurotransmitters L-Dopa and serotonin.

Enzyme abbreviations:

GCHI: GTP cyclohydrolase I (EC 3.5.4.16)

PTPS: 6 – pyruvoyl –tetrahydropterin synthase (EC 4.6.1.10)

SR: sepiapterin reductase (EC 1.1.1.153)

TH: tyrosine hydroxylase (EC 1.14.16.2)

TPH: tryptophan hydroxylase (EC 1.14.16.4)

The rate limiting enzymes, GCHI, TH, and TPH are written in boldface.

Tetrahydrobipoterin (BH_4), a natural cofactor, is present almost ubiquitously in higher organisms and is required for various enzyme activities. It acts as an essential cofactor for three amino-acid monooxygenases i.e. phenylalanine, tyrosine, and tryptophan hydroxylases (Kaufman, 1995). These enzymes are essential for synthesising hormones and neurotransmitters such as dopamine, noradrenaline, adrenaline and serotonin, all of which have displayed evidence for involvement in psychiatric disorders. It has been hypothesised that diminished transport of the transmitter into the presynaptic neuron and/or the synaptic vesicles, may result in a deficit or overflow of the respective transmitter, which cannot be counterbalanced, thus contributing to the manifestation of depression and/or mania (Ackenheil, 2001). It has also been shown that BH_4 acts as a co-factor for the generation of nitric oxide. Nitric oxide is a short-lived neurotransmitter or modulator, which at a central level has been described to play a role in memory, convulsions, cerebral ischemia, learning, behaviour, schizophrenia and bipolar disorder (Karatinos *et al.,* 1995). BH_4 deficiency is predominantly associated with a rare variant of hyperphenylalaninaemia (Scriver *et al.,* 1994), which affects either all organs, including the central nervous system, or only the peripheral hepatic phenylalanine hydroxylating system. It is caused by mutations in genes encoding the enzymes involved in BH_4 biosynthesis or regeneration. It is inherited in an autosomal recessive manner and phenotypically presents with a deficit of the neurotransmitters dopamine and serotonin and progressive neurological disorders. However, two forms of BH_4 deficiency may occur without hyperphenylalaninaemia. Dopa-responsive dystonia (DRD), which is inherited as autosomal dominant mutations of the GCHI gene (Ichinose, 1994) and an apparent central nervous system-localised form of dihydropteridine reductase deficiency (Blau *et al.,* 1998). DRD patients display a severity of symptoms that vary greatly from generalised or focal dystonia, abnormal gait, to subtle signs such as a slight tremor or abnormal results on a writing test. A mutation in exon 2 of the GCHI gene was recently demonstrated in a large Texan family, who in addition to displaying the recognised clinical features of DRD, also exhibited significant psychiatric dysfunction, including depression and anxiety (Hahn *et al.*, 2001). Decreased levels of BH_4 in the cerebrospinal fluid have also been documented in other neurological diseases presenting phenotypically without hyperphenylalaninaemia, such as Parkinson's disease (Curtius *et al.,* 1984), autism (Tani *et al.,* 1994), depression (Bottiglieri *et al.,* 1992) and Alzheimer's disease (Barford *et al.*, 1984).

Structural analysis of isolated genomic DNA clones for human and mouse GCHI revealed that the gene is encoded by a single copy, and is composed of six exons spanning approximately 30 kb (Ichinose *et al.,* 1995). Prior to the large scale sequencing of the Human Genome project, the GCHI gene had been broadly localised to 14q22.1-22.2 (Ichinose *et al.,* 1994). Witter *et al.,* (1996) identified an active promoter region for the GCHI gene, which was found to contain CCAAT and TATA boxes. They also detected several Sp-1 responsive elements in the GC rich region close to the transcription start point, which were required for maximum promoter activity.

To date, approximately 60 different mutations (missense, nonsense and frameshift mutations) in the coding region or in the exon-intron junctions of the GCHI gene have been reported in patients with autosomal dominant DRD all over the world (Nishiyama *et al.,* 2000) with the majority being established as familial (Steinberger *et al.,* 2000).

4.1.4 Detecting DNA sequence variation

A plethora of techniques currently exist for the detection of sequence variation, which is fundamental to our understanding of phenotypic variability including susceptibility to disease and the significant disparity observed in response to treatment. These diverse approaches for allelic discrimination include enzymatic assays, such as restriction fragment length polymorphism (Shi *et al.,* 1999); electrophoretic methods such as single strand conformation polymorphism (SSCP) (Orita *et al.,* 1989) and heteroduplex analysis (Rubocki *et al.,* 1992); direct sequencing (Edwards *et al.,* 1989); high-throughput technologies, for example fluorescence resonance energy transfer (FRET) (Clegg *et al.,* 1992), Taqman (Livak, 1995), Invader assay (Lyamichev, 1999), DNA microarray genotyping (Chee *et al.,* 1996), denaturing high performance liquid chromatography (Underhill *et al.,* 1997) and finally *in silico* methods, such as high-throughput analysis of expressed sequence tag (EST) data (http://genome.wustl.edu/gsc/polybayes) (Shi , 2001). Some of these approaches are of limited efficacy due to insufficient sensitivity (e.g. SSCP), others (e.g. direct sequencing) are labour intensive and time consuming and the more automated high throughput technologies, whilst displaying efficiency and accuracy, tend to be expensive to maintain.

Denaturing High Performance Liquid Chromatography (DHPLC) represents an attractive compromise as it combines high sensitivity, speed (through semi-automisation) and relative inexpense. A number of studies have compared the sensitivity of DHPLC to standard methodology for sequence detection. Recently, DHPLC was compared to direct sequencing on fragments of the BRCA1 gene from women with hereditary cancers (Arnold *et al.,* 1999), and it was revealed that this method succeeded in detecting 100% of the DNA alterations observed by direct sequencing; in contrast SSCP only detected 94%. Furthermore, a blind analysis of DHPLC as a tool for the detection of novel mutations, conducted by O'Donovan *et al.,* (1998) also reported 100% sensitivity and specificity.

4.1.5 Detection of polymorphisms using Denaturing High Performance Liquid Chromatography (DHPLC)

Individuals who are heterozygous for a polymorphism have a 1:1 ratio of wild type to mutant DNA. When the PCR product from the region surrounding the sequence change in such an individual is heated to 95°C and allowed to cool slowly, the DNA strands separate and randomly re-anneal to form a mixture of three distinct species: a wild type homoduplex, a mutant homoduplex and heteroduplices (figure 4.2).

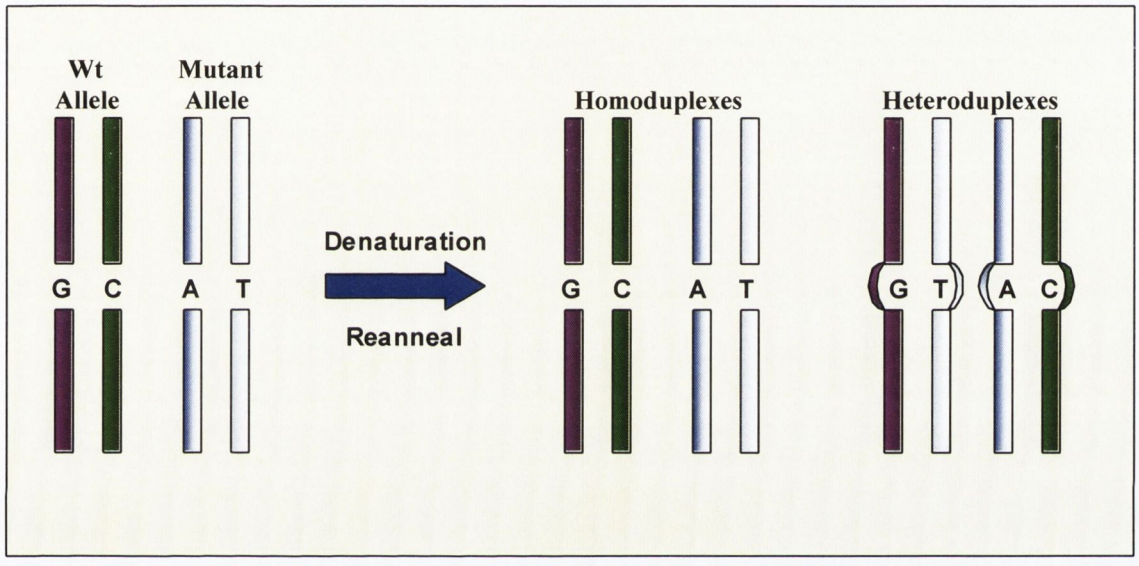


Figure 4.2: Formation of homoduplex and heteroduplex DNA molecules following denaturation and reannealing of DNA strands (Wt: wild type).

If an individual is homozygous for a mutation the same species can be generated by first mixing the homozygous DNA with wild type DNA, and then allowing the denaturing and hybridisation steps to occur (Underhill *et al.,* 1996). Material from a homozygous wild-type individual, however, will only ever form one species, the wild-type homoduplex.

Temperature modulated heteroduplex analysis (TMHA) used for the detection of novel polymorphisms, relies upon the physical changes that occur in DNA molecules induced by mismatched heteroduplex formation during re-annealing of wild type and mutant DNA. During the formation of the heteroduplex, a 'bubble' is formed at the region of base pair mismatches resulting in a short linear region of single stranded DNA being produced, thereby physically altering the molecule. The heteroduplex and homoduplex

species are separated on the WAVE^{TM} DNA Fragment Analysis System using stationary phasealkylated (polystyrene/divinylbenezene) particles, which are converted into a dynamic anion exchanger by an ion-pairing reagent (Kuklin *et al.,* 1997). The double stranded homoduplexes form stronger interactions with the column matrix than the heteroduplex species, which contains single stranded DNA, and therefore, the former remains attached to the matrix for a longer period, resulting in the separation of the two species. TMHA technology employs a rapid and robust scanning method for identifying the presence of mutations, and because wave peak samples can be directly captured for further analysis such as automated sequencing, this technique saves time and labour in addition to eliminating the errors often associated with alternative methodologies.

To date, numerous mutations, many of which are sporadic, have been detected in the GCHI gene, primarily in the coding regions and the intron-exon boundaries. Variants found in the coding region, the intron-exon boundaries and/or the promoter region, which contains the regulatory elements for the gene, may affect protein structure or expression. Hence, for this study the coding regions, the 5'untranslated region and approximately 1000bp upstream of the start (ATG) codon of the putative promoter region, were screened for the presence of mutations using DHPLC technology. DHPLC analysis indicated the presence of four distinct polymorphisms, which were elucidated using automated sequencing. Upon further investigation, it was found that only one of the detected polymorphisms, located in the promoter region of the GCHI gene, was suitably polymorphic to facilitate an association study between bipolar disorder and the GCHI gene.

4.2 Results

4.2.1 Searching for candidate genes on 14q22-24 using Web based tools

Prior linkage studies suggested a region of interest on chromosome 14q22-24 (see chapter 3). Using web-based tools, this 34.92cM region extending from D14S599 to D14S71, was extensively searched for the presence of plausible candidate genes for bipolar disorder. Initially the Ensembl database (http://www.ensembl.org/), was employed to enable the identification of genes present in the specified region. Using markers as boundary tags, sections of the chromosome varying in size from 3 cM to 10cM depending on gene density, were explored. Ensembl *contigview* Web pages feature the ability to scroll along the chromosome, whilst viewing the features within the elected region in detail (figure 4.3).

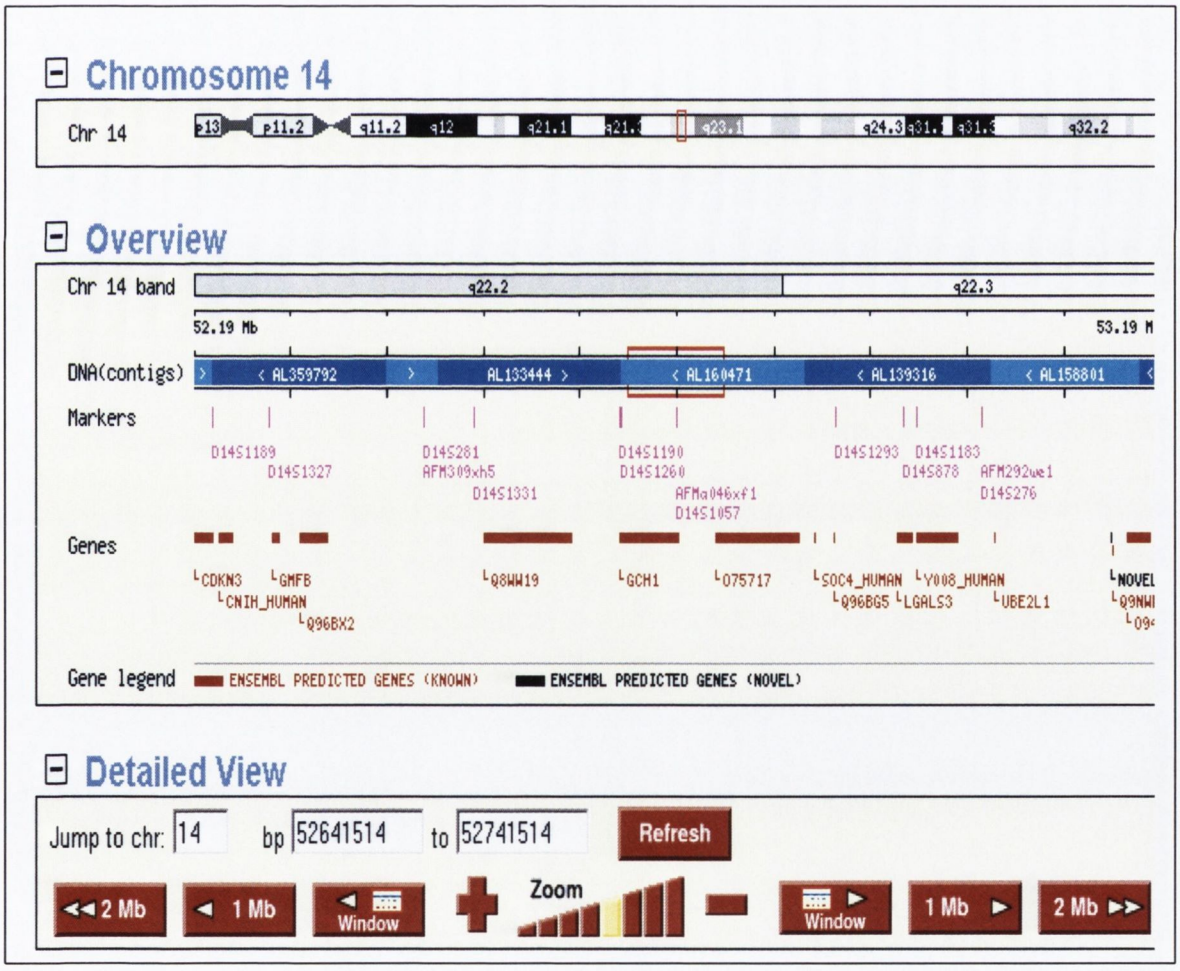


Figure 4.3: Ensembl *contigview* of 14q22.2-22.3, highlighting the microsatellite markers and the Ensembl known and predicted genes for this position on chromosome 14.

By selecting individual genes one can gain access to the Ensembl gene report, where, in addition to providing information about the structure of the gene, direct links to other databases are also provided. All documented information relating to the gene, it's protein and it's possible biological function can be accessed through these links. Using Ensembl, at the time of analysis, a total of 186 known genes were identified within this 34.92cM region of chromosome 14q; all known genes, their location, description, expression profiles and protein function are presented in table 4.1. By consolidating the information obtained from the various databases, and adhering to the criteria that plausible candidate genes must be either directly or indirectly involved in neurotransmitter activities and display significant expression in the brain, the GTP cyclohydrolase I gene (GCHI; protein 58 in table 4.1) was chosen as the best possible candidate and was therefore employed for additional studies.

Table 4.1: Known genes found in the region of interest 14q21.2-q24.3. Last updated 28-07-02

D14S1027 (AFMB340vf1) → D14S71 (AFM197xf12)
14q21.2 → q24.3 spanning a region from 40.8Mb → 73.7Mb

Gene Name	Location: Ensembl	Gene Description	Protein Function	Gene Expression
1. Q9HOJ3	40.9Mb	Hypothetical 75.1kDa protein		Testis
2. Q9NXS3	41.4Mb	CDNA, clone COLO3242	Hypothetical protein	Kidney, testis, uterus
3. Q9Y4F4	41.4Mb	KIAA0423 protein fragment		
4. FKBP3	41.5Mb	Rapamycin selective immunophilin	May bind immunosuppressants Fk506 and rapamycin	Everywhere including brain
5. Q9NWR1	41.5Mb	CDNA, clone KAIA608	Hypothetical protein, KLJ20006	Most areas + brain, CNS
6. Q9HCH6	41.6Mb	KIAA1596 protein (fragment)		
7. Q96PY4	43.4Mb	KIAA1903 protein (fragment)	DNA binding?	
8. RPLI0L	44.8Mb	Ribosomal protein L-10 like	Functional ribosomal protein?	Testis and brain
9. MGAT2	46.6Mb	Alpha-1,6-mannosyl-glycoprotein β-1,2N-acetyl glucosaminyltransferase	Catalyses a step in conversion of oligomann complex N-glycans	Most areas + brain, CNS
10. RPL36A	46.6Mb	60S ribosomal protein L44 (L36A)		Most areas + brain
11. RPS29	46.6Mb	40S ribosomal protein S29	Component of small 40S ribosomal unit, has zinc finger	Everywhere including brain
12. ARF6	46.8Mb	ADP ribosylation factor 6	GDP binding protein, involved in endocytosis and stimulates cholera toxin activity	Everywhere including brain and CNS
13. Q9NWG1	46.8Mb	Hypothetical protein (fragment)		
14. Q9P160	46.8Mb	PRO2610		
15. Q9H867	47.1Mb	CDNA, flji3920	Hypothetical protein	Most areas + brain
16. SOS2	47.1Mb	Son of sevenless protein homolog 2	ras guanine nucleotide exchange factor, binds receptor bound GRB2	Everywhere
17. Q9H9P8	47.2Mb	Similar hypothetical protein flji2618		
18. CDKL1	47.3Mb	Serine/threonine protein kinase	May regulate cell cycle like CDK2	Uterus, leiomios
19. Atpw_human	47.3Mb	ATP synthase coupling factor B	Hypothetical protein	Brian, colon, kidney
20. MAP4K5	47.4Mb	KSH1	Protein kinase, activates jun N terminal kinase	Most areas including brain and CNS
21. Q9H4B6	47.6Mb	WW45 protein	WW domain containing gene	Most areas + brain
22. Q95890	47.6Mb	Unknown	Hypothetical protein	Most areas + brain
23. NIN	47.7Mb	Ninein-lm isoform	Similar to murine centromere binding protein NIN	Most areas + brain

Gene Name	Location: Ensembl	Gene Description	Protein Function	Gene Expression
24. Q95l50	47.8Mb	4-1BB-mediated signalling molecule		
25. RPL36AL	47.8Mb	60S ribosomal protein L44	Ribosomal protein	Ubiquitous
26. Q9NVR5	47.8Mb	Ambiguous		Most areas
27. POLE2	47.8Mb	DNA polymerase episilon subunit B	Stabilise DNA polymerase complex	Most areas
28. Q8WYI1	47.9Mb	MSTP025		
29. Q9Y2U9	47.9Mb	Host cell factor homolog	Intracellular protein transport?	Ubiquitous
30. SDCCAG1	47.9Mb	Serologically defined colon cancer antigen 1		Ubiquitous
31. PYGL	47.9Mb	Glycogen phosphorylase (liver form)	Liver glycogen phosphorylase	Most areas
32. Q92557	48.0Mb		Tripartite motif protein	
33. Q8WXQ3	48.2Mb	No description		
34. Q9BRR1	48.4Mb	No description	Hypothetical protein FLJ12618	Most areas
35. SPG3A1	48.7Mb	ATLASTIN	Spastic Paraplegia 3A	Ubiquitous
36. NID2	48.9Mb	Nidogen2 precursor (osteonidogen)	Basement membrane protein	Most areas +brain
37. TRIM9	49.1Mb	Tripartite motif protein 9	May be involved in development and cell growth	Ubiquitous
38. PD2R_human	49.2Mb	Prostaglandin D2 receptor		Brain
39. PTGER2	49.2Mb	Prostaglandin E2 receptor	G protein coupled receptor	Kidney, uterus, marrow
40. Q9P2K2	49.3Mb	KIAA1344 protein (fragment)		
41. TXNDC	49.4Mb	Thioredoxin-related transmembrane protein	Thioredoxins are small enzymes that participate in redox reactions	Ubiquitous
42. Q9H3A2	49.4Mb	PRO2425		
43. PSMC6	49.6Mb	26S protease regulatory subunit	ATPase subunit of 26S proteasome	Most areas + brain
44. Q9P109	49.6Mb	ERO1 like protein	Maybe formation of disulphide bonds in ER protein folding	Most areas + brain
45. Q96NE9	49.8Mb	CDNa FLJ30983	Weakly similar to H.S. PTPL1	
46. Q14840	49.8Mb	Mig-2 protein (fragment)		
47. GNG2	50.1Mb	PRO1107		
48. CG99_human	50.2Mb	No description	Function unknown	Ubiquitous
49. Q9H3A2	50.6Mb	PRO2425		
50. Q8WUJ0	50.9Mb	Protein tyrosine phosphatase		
51. Q96EK6	50.9Mb	Similar to glucosamine-phosphate N-acetyl transferase		
52. BMP4	50.9Mb	Bone morphogenetic protein 4 precursor	TGF-beta family, like murine bpm4 which specifies tooth identity	Most areas + brain
53. Q8WVH3	51.2Mb	KIAA1705 protein (fragment)	Phosphatidic acid preferring phospholipase (mouse)	

Gene Name	Location: Ensembl	Gene Description	Protein Function	Gene Expression
54. CDKN3	51.3Mb	Cyclin dependent kinase inhibitor 3	Dual specificity phosphatase that interacts with CDK kinases	Most areas + brain
55. CNIH_human	51.3Mb	Cornichon homolog	May play signalling role during T cell activation	Most areas + brain
56. Q99675	51.4Mb	Cell growth regulator CGR19	Cell growth regulator	Most areas + brain, CNS
57. Q9UPU9	51.6Mb	KIAA1053 protein (fragment)		
58. GCH1	51.8Mb	GTP cyclohydrolase I	Catalyses 1st step in BH_4 synthesis, cofactor in aromatic hydroxylation and nitric oxide synthesis	Many tissues inclusive of the brain
59. And1_human	51.9Mb	Acidic nucleoplasmic DNA binding protein	Contains WD repeat and HMG box motifs	Many tissues except brain
60. LGALS3	52.1Mb	Galectin-3	Lactose binding lectin involved in cell growth regulation	Everywhere and brain
61. Y008_human	52.1Mb	Hypothetical protein KIAA0008		High in testis
62. UBE2L3	52.1Mb	Ubiquitin-conjugating enzyme	Ubiquitination of p53, c-fos?	Everywhere
63. Q94290	52.3Mb	KIAA0831		Most areas + brain, CNS
64. Q9NWN3	52.3Mb	CDNA FLJ20725 fis	Hypothetical protein	Most areas + brain
65. KIN1	52.5Mb	CG1 protein	MAP, large ER membrane protein	Everywhere
66. GMFB	52.6Mb	Glia maturation factor beta	Differentiation and regeneration of the nervous system	Ubiquitous
67. Q96BX2	52.7Mb	Cell growth regulatory with ring finger domain	Stress response? Negative control of cell proliferation?	Ubiquitous
68. PEL12	53.2Mb	Pellino 2	Pellino (drosophilia) homolog 2	Most areas + brain
69. Q9H716	53.2Mb	CDNA FLJ21555		
70. SOC4_human	53.2Mb	Suppressor of cytokine signalling 4	Cytokine-inducible negative regulators of cytokine signalling	Testis, uterus, kidney, hippocampus
71. Q96BG5	53.3Mb	Ambiguous	Galectin-3(mouse?)	
72. Q9NX78	53.5Mb	CDNA FLJ20392	Hypothetical protein	Most areas + brain
73. OTX2	53.7Mb	Homeobox protein OTX2	Orthodentical (drosophilia) homolog	Brain, eye, embryo
74. Q9UHT0	54.1Mb	PRO1912 protein		
75. SEC10L1	54.1Mb	Brain secretory protein HSEC10p	Homolog of *S.Cerevisciae* Sec10p needed for pose golgi traffic	Most areas + brain and CNS
76. Q9NWY2	55.1Mb	Uncharacterised hypothalamus protein	Member of actin family	Most areas + brain
77. PSMA3	55.2Mb	Proteasome subunit alpha type 3	Multicatalytic protease complex	Most areas + brain, CNS
78. Q9P150	55.2Mb	PRO1617	Hypothetical protein	Kidney, marrow, placenta
79. RBBP1	55.2Mb	Retinoblastoma binding protein 1	Transcriptional regulation	Most areas + brain
80. TIMM9	55.3Mb	Mitochondrial import inner membrane translocase subunit	Member of family of TIM related proteins	Most areas + brain

Gene Name	Location: Ensembl	Gene Description	Protein Function	Gene Expression
81. Q96DX3	55.4Mb	Ambiguous	Hypothetical protein FLJ10813	Ubiquitous
82. Q9P055	55.4Mb	HSPC213	Hypothetical protein	Most areas + brain
83. Q9NVL8	55.6Mb	Ambiguous	Hypothetical protein FLJ10650	Kidney, colon, stomach, prostate, uterus, pancreas
84. Q9NYF0	55.6Mb	Hepatocellular carcinoma novel gene 3 protein		Most areas including brain and CNS
85. Q9YD41	56.2Mb	KIAA0666 protein (fragment)		
86. Q9UHU2	56.4Mb	PRO1693		Human fetal liver
87. Y586_human	56.6Mb	No description	Novel prediction (mouse)	Ubiquitous
88. Q9UKU2	57.1Mb	Retinal short chain dehydrogenase reductase (RETSDR4)	CGI-86 protein	Ubiquitous
89. Q9H9F2	57.1Mb	Weekly similar to Pecanexprotein	Hypothetical protein FLJ12799	Most areas + brain
90. PPM1A	57.2Mb	Protein phosphatase 2C isoform	Mg dependent protein phosphatase	Most areas + brain
91. DAAM1	57.4Mb	Dishevelled-associated activator of morphogenesis	Belong to FH protein family implicated in cell polarity	Ubiquitous
92. SIX 6	57.4Mb	Homeobox protein SIX6	SIX 6 family of homeodomain transcription factors	Brain only
93. Q96LJ5	57.6Mb	cDNA FLJ25436	Similar to riken cDNA(mouse)	
94. SIX1	57.6Mb	Homeobox protein SIX1	Drosophilia homolog, muscle development	Some areas + brain
95. MNAT1	57.7Mb	CDK activating kinase assembly factor MAT1	Ménage a trios CDK assembly factor	Most areas + brain, CNS
96. SIX4	57.7Mb	Homeobox protein SIX4	May be involved in embryogenesis	Breat, colon, prostate
97. RTN1	57.8Mb	Reticulon 1	Euroendocrine specific protein	Most areas + brain
98. Q9P2F4	57.9Mb	KIAA1393 protein (fragment)		
99. PRKCH	58.3Mb	Protein kinase C eta type	Binds phorbol esters	Some areas + brain
100.Q9Y394	58.4Mb	CGI-86 protein	Dehydrogenase? (predicted)	Ubiquitous
101.Q9BVQ0	58.5Mb	Unknown	Hypothetical protein	Some areas not brain
102. HIF1A	58.7Mb	Hypoxia-inducible factor 1 alpha	Basic helix-loop-helix transcription factor	Ubiquitous
103.Q9BQR8	60.3Mb	Synaptotagmin fragment	Synaptomagin (mouse)	
104. Q9H189	60.7Mb	Sphingosine-1 phosphate phosphatase		
105. ARHJ	61.4Mb	TC10-like RHO GTPase	GTP binding?	Most areas + brain
106. ESR2	61.4Mb	Steroid hormone receptor	Transcriptional activator involved in reproduction	Germ cell, lung, brain, testis, cell line
107. MTHFD1	61.5Mb	C-tetrahydrofolate synthase	Protein catalyses 1 of 3 sequential reactions in the interconversion of 1 carbon derivatives of BH_4, substrates for methionine, thymidylate, purines	Most areas including brain

Gene Name	Location: Ensembl	Gene Description	Protein Function	Gene Expression
108. PPP2R5E	61.6Mb	Serine/threonine phosphatase 2A	Regulate cell growth + division, muscle contraction + transcription	Highly expressed in brain
109. AKAP5	61.6Mb	A-kinase anchor protein	Anchors c-AMP dependent protein kinase to post synaptic densities	Only found in CNS
110. ZNF4	61.6 Mb	Zinc finger protein 46	Kruppel zinc finger protein 46	Bone, skin
111. HSPA 2	61.7Mb	Heat shock related 70kDa protein2	Similar to murine Hsp70-2	Most areas+ brain
112. Q9Y2K1	61.7Mb	KIAA0997 protein		Most areas + brain, CNS
113. Q6FK6	61.8Mb	Ambiguous	Similar to Riken cDNA	Most areas + brain
114. SPTB	61.9Mb	Spectrin beta chain, erythrocyte	May crosslink actin proteins of the membrane associated cytoskeleton	Brain, colon, heart, neck, lung, spleen, marrow
115. Q9Y2LE	62.1Mb	NESPRIN –2 gamma		
116. FUT8	62.2Mb	n-acetyl-beta-d-glucosaminide	Functions in asparigine linked glyco protein oligosaccharide synthesis	Most areas + brain
117. Q96LQ0	62.8Mb	cDNA FLJ25315	Novel prediction	
118. MAX	62.9Mb	MAX protein	Forms heterodimers with myc, mad and regulated gene expression	Everywhere + brain
119. FNTB	62.9Mb	Farensyltransferase β subunit	Transfers farensyl; groups to protein	
120. Q92BT82	63.0Mb	No description	Novel prediction	
121. GPX2	63.0Mb	Glutathione peroxidase	Catalyses the reduction of hydrogen peroxide, hydroperoxide+ lipid peroxides by reduced glutothione	Highly expressed in several libraries
122. Q8WUH1	63.1Mb	CHURCHILL	Novel prediction	
123. FOS	64.1Mb	Murine osteosarcoma viral oncogene homolog	Can dimerize with JUM proteins to form transcription factor complex	
124. NM_006827	64.2Mb	Transmembrane trafficking protein	Protein transport from ER to golgi	Ubiquitous
125.VATD_human	64.9Mb	Vacuolar ATP synthase subunit	Vacuolar protonδpolypeptide	Everywhere
126. Q9ULM0	64.9Mb	KIAA1200 protein (fragment)		
127. GPHN	65.0Mb	GEPHYRIN	Postsynaptic localisation of inhibitory neurotransmiiter receptors	Ubiquitous
128.EIF2S1	65.0Mb	Eukaryotic translation initiation factor 2 subunit 1	Binds gtp and assists with met-trna binding to 40S ribosomal subunit	Most areas + brain
129. Q9H357	65.0Mb	PRO1751		
130 Q9NYT0	65.0Mb	Pleckstrin 2 (mouse homolog)	Involved in actin reorganisation?	Not in brain
131 ARG2	65.2Mb	ArginaseII precursor	Hydrolyses arginine to urea and ornithine	Everywhere inclusive of brain
132 PIGH	65.2Mb	Phosphotidylinositol glycan, class h	Catalyses transfer of GlcNAc to PI	Everywhere + brain
133. O15035	65.3Mb	KIAA0321		
134 VTL1_human	65.3Mb	Vesicle transport vsnare protein	Vesicle transport pathways?	Most areas + brain

Gene Name	Location: Ensembl	Gene Description	Protein Function	Gene Expression
135. MPP5	65.8Mb	Guanylate kinase	Signalling	Most areas
136. RAD51L1	66.4Mb	DNA repair protein	Sensing DNA damage?	Most areas + brain
137. BRF1	66.4Mb	Butyrate response factor 1	Putative zinc finger nuclear transcription factor	Most areas + brain
138. ACTN1	66.5Mb	Alpha-actinin 1	Cytoskeletal proteins	Most areas + brain
139. Q9NVH0	66.8Mb	CDNA FLJI0738	Hypothetical protein	Brain, CNS + others
140. Q9ULT9	66.9Mb	KIAA1130 protein (fragment)		
141. Q9NUM3	67.0Mb	CDNA FLJI1274	Hypothetical protein	Most areas + brain
142. ERH	67.0Mb	Enhancer of rudimentary homolog	Drosophilia, pyrimidine pathway	Most areas + brain
143. Y247_human	67.3Mb	Hyopthetical protein	Unknown	Most areas + brain
144. Q8WY65	67.4Mb	No description	Human zinc finger protein	Colon
145. Q9NQR7	67.4Mb	Myelin proteolipid protein like protein		
146. 64117	67.5Mb	Unknown	Unknown	Only in placenta
147. SFRS5	67.5Mb	Splicing factor arginine/serine rich	Regulator of pre-mRNA splicing	Everywhere +brain
148. SLC10A1	67.5Mb	Sodium/bile acid co-transporter	Extraction of bile acids from portal blood plasma by liver	Liver
149. Q9Y4F8	67.6Mb	Secreted modular calcium binding protein precursor		Most areas including brain
150. SLC8A3	67.9Mb	Na/Ca exchanger 3 precursor		
151. O00368	68.2Mb	Putative p150		
152. Q9UHL1	68.2Mb	RNA polymerase II transcriptional regulational mediator	Roles in transcriptional activation and repression	Some areas + brain
153. ADAM20	68.2Mb	ADAM20 precursor	Membrane anchored proteins	Testis, tongue
154. Q8WYG5	68.3Mb	No description		
155. Q92623	68.3Mb	Myleoblast KIAA1227 (fragment)		
156. Q9H2N5	68.4Mb	Mixed lineage kinase MLK1		
157. Q9Y2J9	68.6Mb	KIAA0995 protein (fragment)	Unknown	
158. Q9POS2	68.9Mb	Ambiguous	Hypothetical protein LOC51241	Ubiquitous
159. OM25_human	68.9Mb	Mitochondrial outer membrane protein	Membrane targeted signalling	Ubiquitous
160. Q9H7N3	68.9Mb	FLJ00035 protein (fragment)	Unknown	
161. ADAM 21	69.0Mb	Disintegrin and metalloprotinease domain 21	Membrane anchored proteins	Testis
162. LRE1	69.1Mb	ORF2 contains reverse transcriptase domain		
163. M3K9_human	69.3Mb	Mitogen activated protein kinase		
164. Q94897	69.4Mb	Pecanex-like protein 1	Mouse similarity	
165. Q9P167	69.7Mb	PRO2372	Unknown	
166. Q9PIN6	69.7Mb	PRO1158	Unknown	

Gene Name	Location: Ensembl	Gene Description	Protein Function	Gene Expression
167. Q9H367	69.7Mb	PRO1953	Unknown	
168. Q9H365	69.8Mb	PRO2459	Unknown	
169. Q9UNU4	70.1Mb	KIAA0440 protein	Unknown	Ubiquitous
170. RGS6	70.5Mb	Regulator of G-protein signalling	Activated GTPase activity of alpha subunits	Brian, lung testis, liver, spleen, neck,
171. O95248	70.9Mb	Hypothetical 133.3kDa protein		
172. PSEN1	70.8Mb	Presenilin 1	Role in processing amyloid beta precursor protein	Highly expressed in several libraries
173. Numb_human	71.0Mb	Numb protein homolog (drosophilia)	Generate asymmetric cell division during neurogenesis	Most areas + brain
174. REQC_human	71.1Mb	Zinc-finger protein CER-D4	Maybe DNA binding protein, involved in cell death	Most areas + brain
175. E2Bb_human	71.4Mb	Translation initiation factor subunit	Translation initiation factor	Most areas +brain
176. ZNFN12A1	71.5Mb	KIAA1589	Zinc binding?	
177. Q96K22	71.5Mb	CDNA FLJ 14839	Hypothetical KDA protein (mouse)	
178. Q9H6W3	72.0Mb	No description	Unknown	Most areas + brain
179. Q96N88	72.2Mb	Ambiguous	Hypothetical protein MGC12435	Ubiquitous
180. Q96MY7	72.5Mb	CDNA FLJ31697	Weakly similar to Plectin	
181. PTE2_human	72.5Mb	Peroximal acetyl co-enzyme thioester hydrolase 2	Peroxisomal thioesterase	Most areas + brain
182. Q96DE2	72.6 Mb	No description	Novel prediction	
183. PNMA1	72.7Mb	Paraneoplastic neuronal antigen		Most areas + brain, CNS
184. C0Q6_human	72.9Mb	Putative ubiquitone biosynthesis monooxygenase	CGI-10 protein	Most areas + brain, CNS
185. LDH6A1	73.0Mb	Methylmaldonate-semialdehyde dehydrogenase	Same	Some areas + brain
186. Q9H5P8	73.0Mb	CDNA FLJ23189	Hypothetical protein	Some areas + brain

4.2.2 Investigation of GCHI using WEB based tools

The GTP cyclohydrolase I gene was investigated using the August 2001 release of the human genome browser (http://genome.ucsc.edu/).

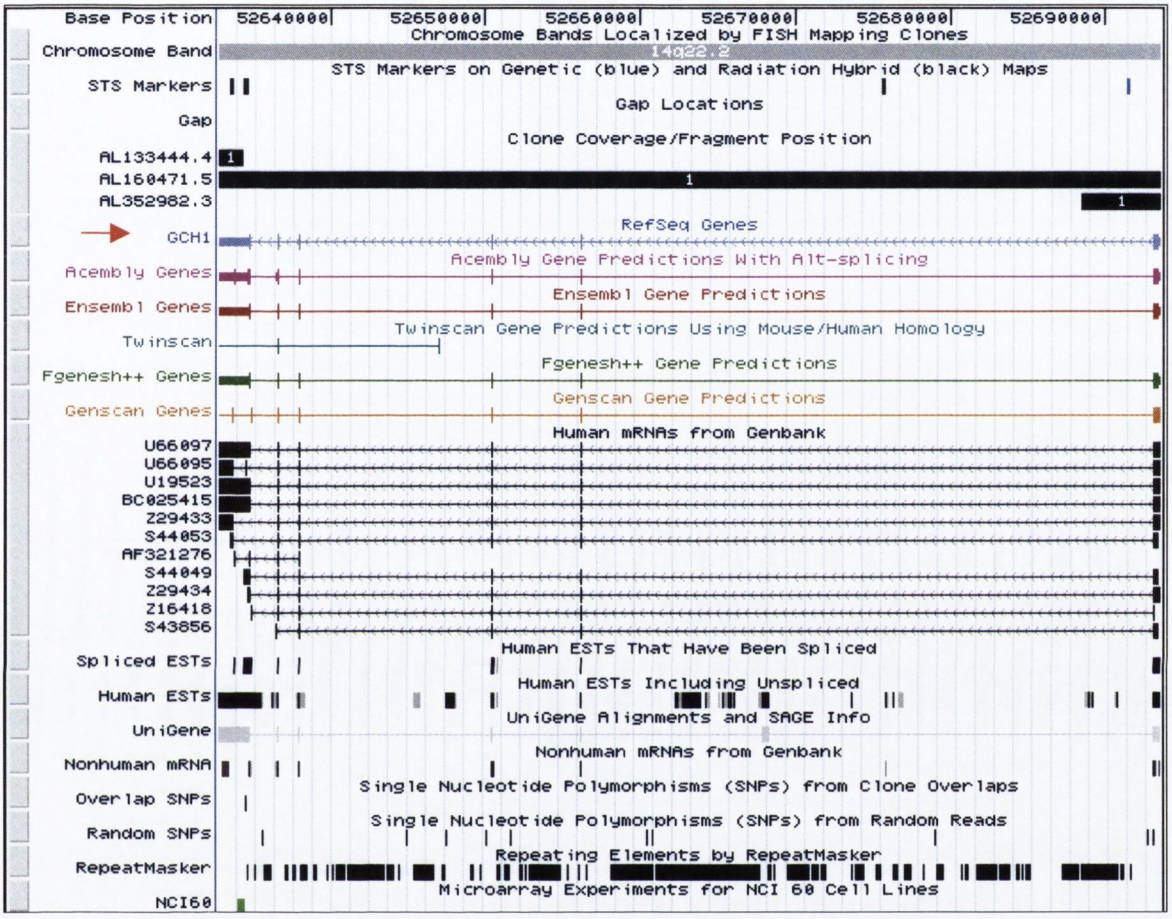


Figure 4.4: Diagrammatic output from the human genome browser for GCHI gene.

Upon examination of the data displayed in figure 4.4, it is observed that the GCHI gene is located on chromosomal band 14q22.2. The gene is fully annotated and several databases including GenBank, Human Genome browser, Ensembl gene predictions, and Assembly gene predictions, all concur that this gene contains six coding regions. The expression profile of the gene can be accessed via the Unigene alignments and the Serial Analysis of Gene Expression (SAGE) link. Also, the known SNPs are displayed individually and it is possible to directly access the SNP database (www.ncbi.nlm.nih.gov/SNP/), where additional information regarding location, type, heterozygosity etc., of the SNP can be acquired. By clicking on the blue GCHI link (red arrow), contact can be made with the PUBMED and OMIM databases, where textual information about the gene and genetic disorders associated with its malfunction as well as citations and references from biomedical literature can be obtained. Also, through this link it is possible to access the raw genomic and mRNA sequences. This data is presented in a user-friendly environment, where the coding regions, 3' and 5' untranslated regions, and the promoter sequence can be easily identified. The structure of the GCHI gene is presented in figure 4.5. The 5' nucleotide sequence of the gene, inclusive of part of the putative promoter region, is illustrated in figure 4.6.

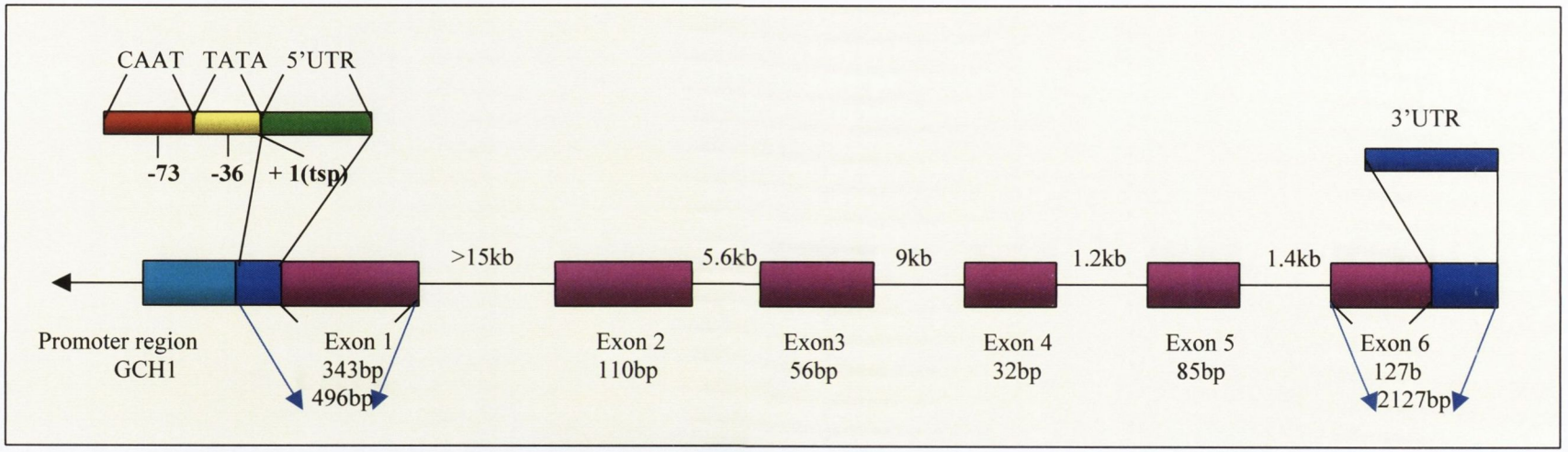


Figure 4.5: Structure of GTP Cyclohydrolase I gene.



Figure 4.6 : Nucleotide sequence of the 5' region of the human GCHI gene (adapted from Witter *et al.*, 1996). The putative CCAAT and TATA boxes are indicated. Putative binding sites for the following transcription factors were identified by computer based sequence analysis: upstream stimulatory factor (USF), serum response factor (SRF), nuclear factor y (NF-Y), glucocorticoid-response element (GRE), nuclear factor κB (NFκB) and the SP1 transcription factor. The SNPs detected using DHPLC analysis are also indicated by a red arrow: G/A at position –959bp, C/T at position –904bp, G/A at position -740bp and C/G at position +393bp. [tsl (translation start site) is numbered +1]

4.2.3 WAVE Analysis of the GCHI gene to detect novel polymorphisms

The GTP cyclohydrolase I gene (coding region, 5'UTR and 1000bp upstream of the start (ATG) codon of the putative promoter region) was amplified using non-radioactive PCR in 50µl volumes (section 2.5), prior to TMHA analysis. The promoter region was amplified in five overlapping fragments of size varying between 380bp and 447bp in length. Exon I was initially amplified and analysed as a 507 bp fragment. However, results were inconclusive and consequently the 5' and 3' ends of exon I, were re-analysed separately (figure 4.7).

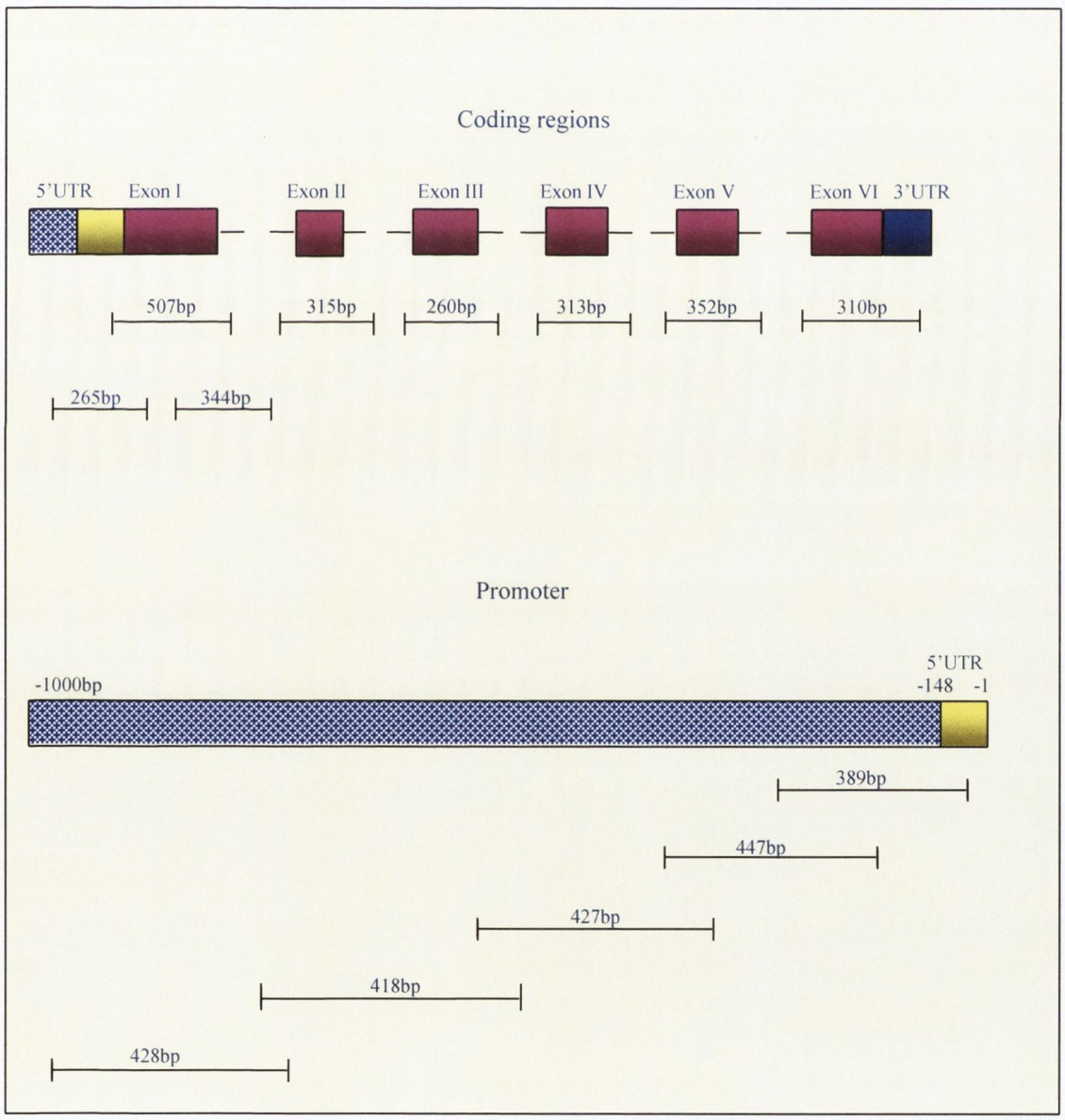


Figure 4.7: Regions of the GCHI gene screened for novel polymorphisms using DHPLC.

Power calculations have shown that to detect a genetic variant, which occurs at a frequency of 10% in the general population, 28 individuals must be screened to assure 95% detection power. Therefore, a sample cohort of 15 unrelated BPI individuals and 15 unrelated control individuals were employed in this analysis. Unaffected individuals were included in the analysis to avoid false negatives. In the unlikely event that all affected individuals contain an identical functional polymorphism, all chromatographic outputs would be the same, thereby producing a situation where the polymorphism could be mistaken as a PCR artefact.

Although DHPLC does not require any special PCR primers, reagents, or post-PCR sample pre-treatment, reliable DHPLC analysis does require high quality PCR products. This is achieved by careful primer design, elimination of pre-PCR artefacts and avoidance of excessive cycling. Ideally PCR products should be between 150bp and 600bp in length, although mutations have been detected in fragments as large as 1.5kb. It has been found that the quality of the water can influence the efficiency of the PCR reactions and can also affect the column performance. Consequently, HPLC grade water is used at all times. In addition, several reagents such as DMSO, Triton and mineral oil, routinely used in standard PCR protocols, are detrimental to the column and are therefore not used.

Each PCR protocol, to amplify the segments of the gene to be screened, was optimised prior to DHPLC analysis and checked on 2% (w/v) agarose gels to ensure that amplification was specific. Following heteroduplex formation, each sample was analysed using the most appropriate pre-specified column temperatures. PCR heteroduplexes were resolved from homoduplexes via differential elution profiles under partially denaturing conditions. In the absence of a mutation, the chromatogram displayed only a single homoduplex peak, whereas in the presence of a mutation two peaks were produced, the first peak representing the heteroduplex and the second peak corresponding to the homoduplex. Figures 4.8, 4.9 and 4.10 depict the characteristic WAVE chromatograms produced for three of the regions where variants were detected; namely C/G base change at position +393bp, C/T base change at position –740bp, G/A base change at position –959bp and C/T base change at positon –904bp.

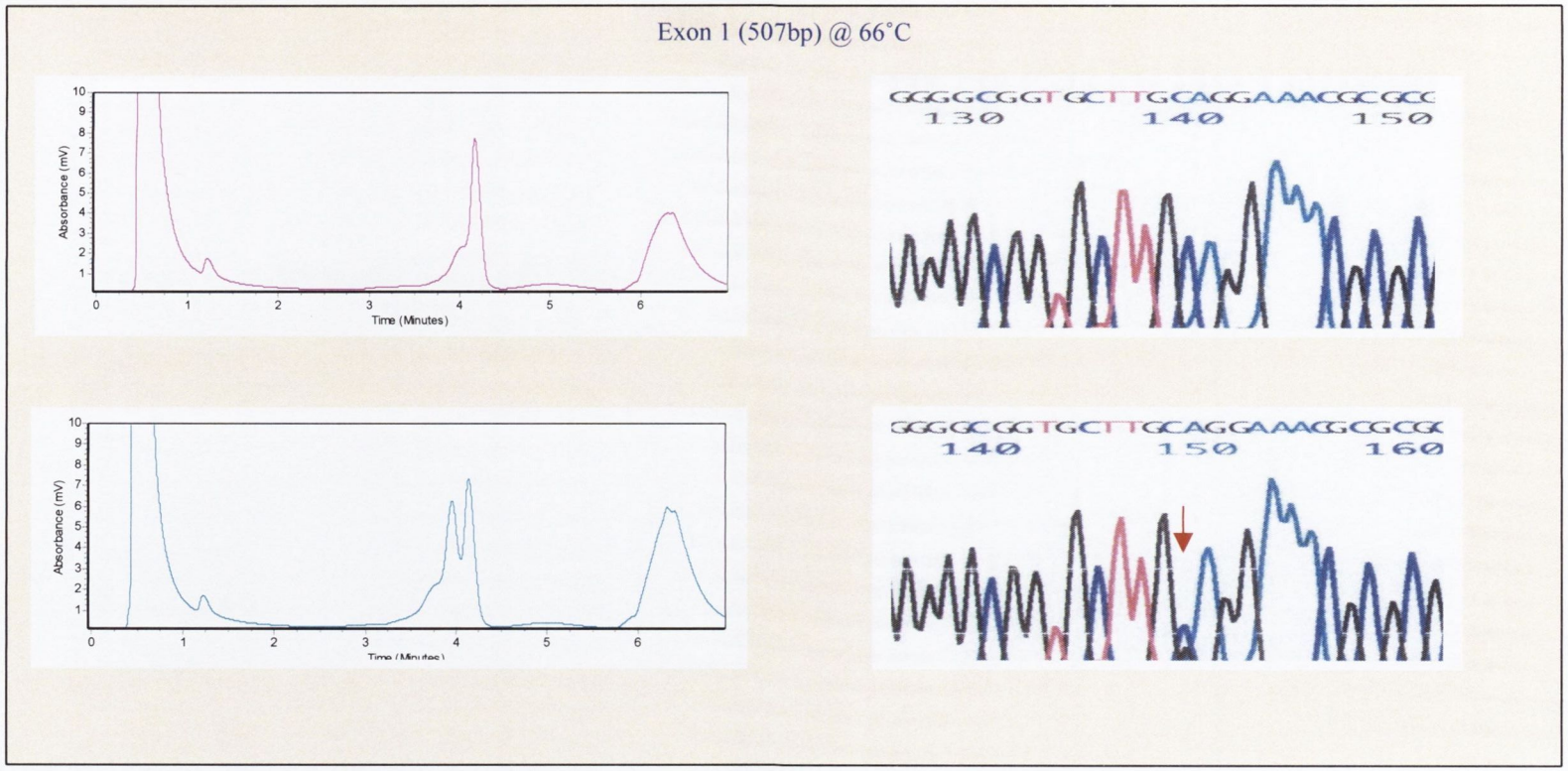


Figure 4.8: Wave chromatograms obtained for GCHI gene Exon I (507bp) (L.H.S). The upper chromatogram illustrates a homogeneous sample and the lower chromatogram depicts a heterogeneous sample. The corresponding sequence patterns for the region containing the SNP are also shown (R.H.S). The C/G base change at position +393bp is clearly marked.

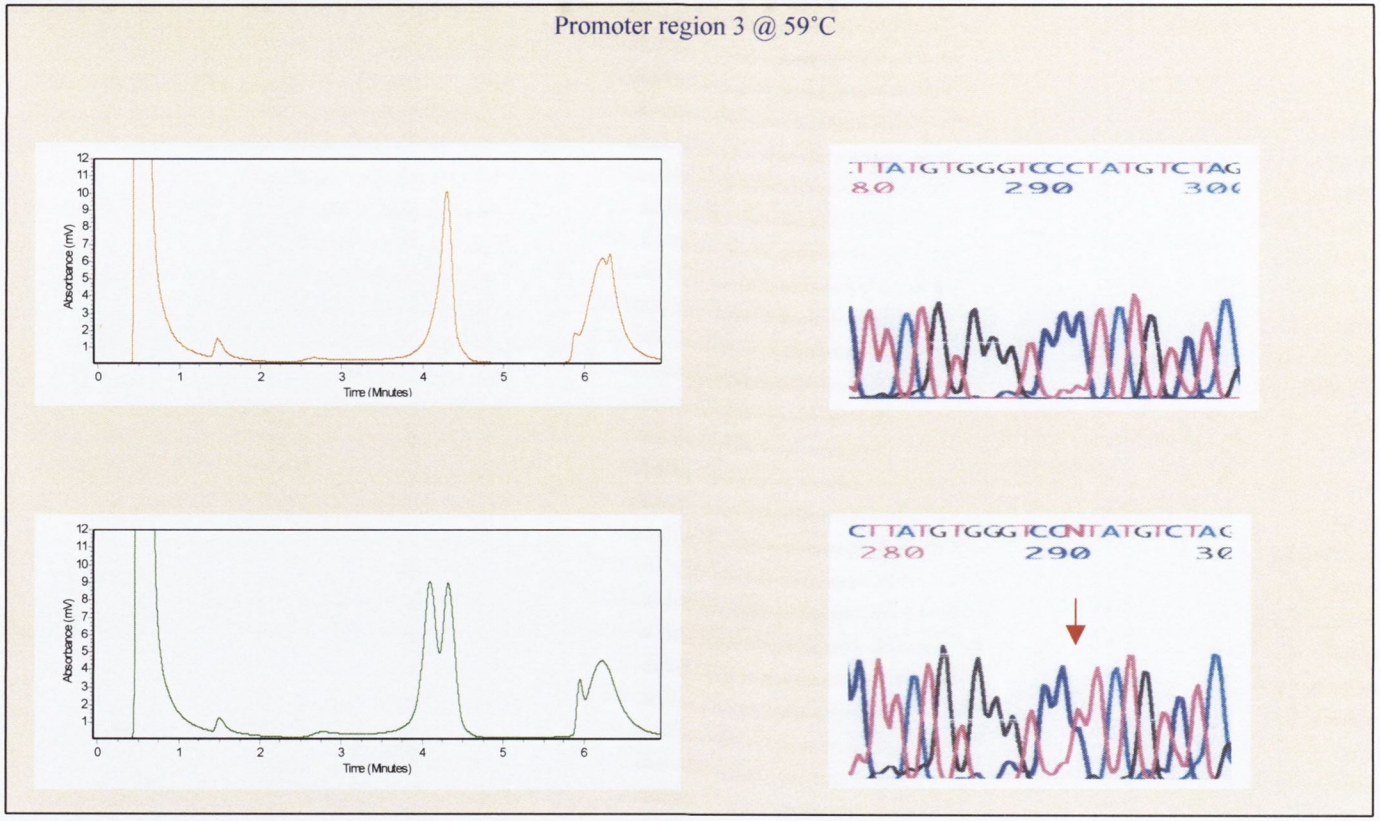


Figure 4.9: WAVE chromatograms obtained for the promoter (region 3) of the GCHI gene (LHS). The upper chromatogram illustrates a homogeneous sample and the lower chromatogram depicts a heterogeneous sample. The corresponding sequence pattern for the region containing the SNP are shown (RHS). The C/T base change at position –740bp is clearly marked.

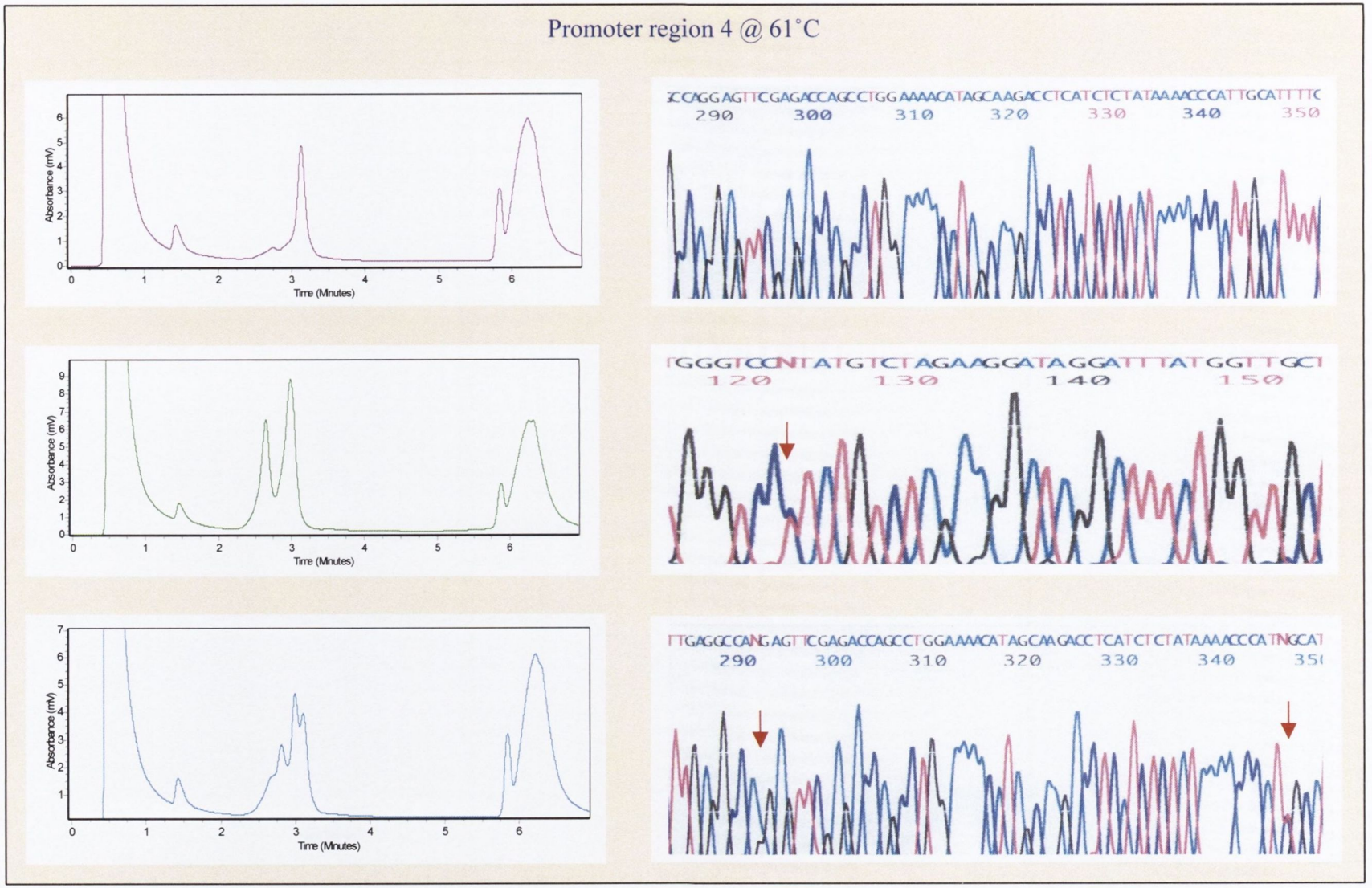


Figure 4. 10: Wave chromatograms for the promoter (region 4) of the GCHI gene (L.H.S). The upper sample is homogeneous, the middle sample contains one SNP and the lower sample, with 3 peaks contains two SNPs. The corresponding sequence patterns are displayed on the R.H.S and the SNPs are clearly marked. There is a C/T at position –740bp in the middle chromatogram and two SNPs, G/A at position –959bp and C/T at position –904bp, in the lower chromatogram.

DHPLC analysis indicated the presence of four independent SNPs. The fragments containing the SNPs were re-amplified and the SNPs elucidated by automated sequencing (sections 2.11 and 2.12). Exon I was initially amplified as a 507 bp fragment, with the primers located in intron 1 and the 5'UTR. When analysed using DHPLC, a heteroduplex was clearly visible in 13% of the samples. However, when these samples were sequenced the SNP was not detected. This was probably due to the fact that the SNP is located only 29bp upstream of the reverse primer in the original fragment and the corresponding sequence of this region was not of sufficient quality to detect the SNP. This fragment was then reanalysed in two separate fragments, incorporating more of the intronic sequence and the 5'UTR sequence. The SNP, a C/G base change, was found to be located + 393 bp from the ATG codon, in intron 1. The other five coding regions and the 5'UTR were screened in their entirety but the chromatograms produced for each region did not display any deviations, thereby implying that SNPs were not present in these regions of the GCH1 gene. Approximately 1000bp of the putative promoter region was analysed in five distinct fragments. Because of the nature of the sequence in this region, it was difficult to design appropriate primers. Consequently, there was considerable overlap between these fragments (see section 4.2.3). Region 3, displayed the characteristic pattern for the presence of a heteroduplex in 13% of the samples, which when sequenced was found to correspond to a C/T base change at position –740bp. This SNP was also detected in region 4 of the promoter, which was consistent with the overlap between the fragments. Two additional SNPs were also detected in this fragment, G/A at position –959bp and C/T at position –904bp in 33% of the samples.

To determine if any of these polymorphisms were suitable for application in an association study with bipolar disorder, their heterozygosity had to be established. Two of the SNPs, at positions –740bp and –904bp were previously revealed as neutral polymorphisms (Steinberger *et al.,* 2000). The published heterozygosity of these polymorphisms was 0.2 and 0.3 respectively and therefore they were not considered to be suitable for further studies. For the two novel polymorphisms identified, G/A at position –959bp and C/G at position +393bp, suitable RFLP assays were designed and performed on a sample of 30 individuals to determine heterozygosity.

4.2.4 RFLP analysis of novel C/G polymorphism at position +393 (Intron 1)

Analysis of the sequence and the polymorphism using WEBCUTTER (http://firstmarket.com/cutter/cut2.html) revealed that the *Mnl*I restriction site was affected by this polymorphism. The assay was performed as described in section 2.7. The characteristic band pattern for this RFLP following gel electrophoresis is shown in figure 4.11.

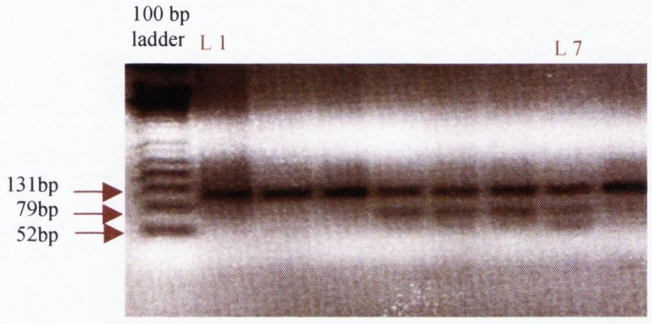


Figure 4.11: SNP at +393bp in intron 1 resolved using RFLP analysis. Presence of the mutant G allele creates a recognition site for the restriction enzyme *Mnl*I, visualised on the gel by the presence of the bands 79bp and 52bp. L1, L2 and L3, corresponding to samples D58, D59 and D60, represent the uncut fragment (common C allele is present). L4, L5, L6 and L7, corresponding to samples D61, D62, D63 and D64, are heterozygotes displaying bands of size 131, 79 and 52bp.

A sample of 30 unrelated individuals was investigated using this assay to determine the heterozygosity of this polymorphism. However, the heterozygosity was determined to be 0.16 and as a result it was eliminated as a possible candidate polymorphism for an association study.

4.2.5 RFLP analysis of novel G/A polymorphism at position –959 bp (Promoter)

Analysis of the sequence and the polymorphism using WEBCUTTER revealed that only one endonuclease, the restriction enzyme *BsrD*I was affected by this polymorphism. An internal positive control site was not available for inclusion in this assay. The assay was performed as described in section 2.7. The characteristic band pattern for this RFLP following gel electrophoresis is shown in figure 4.12.

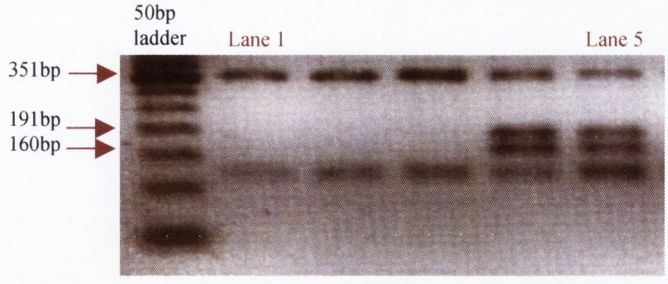


Figure 4.12: SNP at –959bp in the promoter region resolved using RFLP analysis. Presence of the mutant A allele creates a recognition site for the restriction enzyme *BsrD*I visualised on the gel by the presence of the bands 191bp and 160bp. Lanes 1, 2 and 3, corresponding to samples D1, D2 and D3, represent the uncut fragment (351bp). Lanes 4 and 5, corresponding to samples D4 and A6, are heterozygotes producing bands of sizes 351bp, 191bp and 160 bp. In all lanes there is a band of approximate size of 120bp. This represents a spurious PCR product that does not interfere with the RFLP analysis.

Using a sample of 30 unrelated individuals it was established that the heterozygosity for the G/A polymorphism at position –959bp was 0.46, making it an appropriate SNP for further association analysis.

4.2.6 TDT based association study using novel SNP

Ninety-three families, with both parents available for genotyping, were included in a TDT based association study. The genotypic data for the G/A single nucleotide polymorphism at positon –959bp was analysed by hand using TDT, in a total of 93 family trios where the proband was diagnosed as BPI. The result of this analysis is presented in table 4.2

Table 4.2: TDT result for novel SNP (G/A, at position –959bp), detected in the promoter region of GCHI and a sample of 93 Irish BPI trios.

Trio Sample	Number of heterozygous parents	G/A transmission counts	Chi-square statistic	p value
All Probands	92	35-57	5.26	0.0218

Upon examination of the alleles transmitted from 92 heterozygous parents it was revealed that there was preferential transmission of the variant A allele (57A transmissions versus 35G transmissions, $\chi^2 = 5.26$, $p = 0.0218$).

4.3 Discussion

Chromosome 14q22-24 was previously highlighted by this group as a region of interest, based on the results of a modified genome scan (Murphy *et al.*, 2000; section 1.13). Further linkage analysis supported the initial finding, and resulting from these studies a 34.92cM region on chromosome 14q22-24 was suggested as harbouring a susceptibility locus for bipolar disorder in the Irish population. This region was then subjected to an extensive search using web-based tools for the presence of plausible candidate genes with which to conduct association studies for bipolar disorder. To date 186 known genes have been localised to this region. Following rigorous examination of these genes and based on the criteria that the proposed gene should be implicated in the neurotransmission pathways and also display significant expression in the brain, it was decided that the GTP cyclohydrolase gene (GCHI) presented the best possible candidate gene for further studies.

Serotonin and catecholamines are central to neurotransmission pathways and therefore have long been proposed to be involved in the aetiology of neuropsychiatric disorders such as bipolar disorder. The biosynthesis of these neurotransmitters is regulated by tryptophan and tyrosine hydroxylase respectively. The activity of these enzymes is controlled by the concentration of the cofactor tetrahydrobiopterin (BH_4), which in turn is regulated by the rate-limiting enzyme GTP cyclohydrolase I (GCHI). It is therefore possible that GCHI is indirectly involved in controlling tyrosine hydroxylase and tryptophan hydroxylase activity and consequently influences catecholamine and serotonin levels. To date, approximately 60 different muations have been reported in the GCHI gene in patients affected by the autosomal dominant disease, dopa-responsive dystonia. However, the majority of these mutations have been established as familial. To investigate the presence of polymorphisms within the Irish case/control bipolar disorder cohort, a sample of 30 unrelated individuals were screened using DHPLC technology, which ensures > 95% power to detect a genetic variant occurring at a frequency of 10% in the general population. The six coding regions of GCHI, together with the 5'UTR and a portion of the putative promoter region, were amplified and examined using TMHA at appropriate column temperatures, and the data was analysed using the WAVE system in conjunction with Transgenomic software. Three independent fragments, two of which were in the promoter region and the third incorporating 103 bp of the 3'end of exon1 and the first 203bp of intron 1, displayed the

characteristic chromatographic patterns indicating the presence of a single nucleotide polymorphism (i.e. a double peak profile). The wild type profile (i.e. single peak) was observed for all other fragments analysed. The samples that displayed a deviation from the wild type peak profile were reamplified and sequenced in both the 5' and 3' directions.

A C/G base change was located +393bp from the ATG codon, in intron 1. This SNP is not located in the exon-intron junction and it does not appear to affect any splicing regulatory sites. Three additional polymorphisms were detected in the promoter region of the GCHI gene; a C/T base change was detected at position –740bp from the ATG codon, a C/T base change at position –904bp and a G/A base change at position –959bp. These polymorphisms do not appear to directly affect any obvious or predicted transcription regulatory sites in the promoter. The polymorphisms at –740bp and –904bp were previously revealed as neutral polymorphisms with insufficient heterozygosity rates for use in further studies (Steinberger *et al.*, 2000). The novel SNPs at +393bp in intron 1, and –959bp in the promoter, were examined using RFLP analysis in a sample of initially 30 unrelated individuals to determine the heteozygosity of the polymorphisms. It was established that only the G/A polymorphism at position –959bp, with a heterozygosity rate of 0.46 was suitable for further analysis.

The transmission disequilibrium test was applied to the genotypic data from a total of 93 family trios, since such family based studies nullify the problem of false positive associations that may arise from ethnic stratification of case-control samples (Schaid, 1998). The test revealed a statistically significant distortion in the transmission of allele A to BPI probands ($\chi^2 = 5.26$, $p = 0.0218$). The transition from G to A in the promoter region is unlikely to be functional as it does not appear to directly alter known transcriptional elements in the promoter region (figure 4.6). A definitive answer, however, would require functional expression studies such as those performed on the serotonin transporter promoter (Collier *et al.*, 1996). These studies displayed the effects of the polymorphism on restricting the transcriptional activity of the serotonin promoter resulting in low functional expression of the serotonin transporter. There is also the possibility that this polymorphism may be in linkage disequilibrium with another as yet unidentified functional mutation that is contributing to bipolar disorder. Further studies could also involve attempts to replicate this finding in an independent sample of trios. It

may not be beneficial to perform an association study using the case-control sample based on evidence from studies previously conducted on the serotonin gene, where it was found that this sample was unable to replicate suggestive findings in the trio sample.

CHAPTER FIVE

FURTHER STUDIES TO EXAMINE THE INVOLVEMENT OF THREE WELL-RECOGNISED CANDIDATE GENES FOR BIPOLAR DISORDER, IN THE IRISH POPULATION

5.1 Introduction

5.1.1 Candidate gene association approach

The candidate gene association study approach has and continues to make a considerable contribution to the genetic analysis of complex disorders. This methodology involves investigating whether there is an excessive presence of a particular polymorphism in individuals with the disorder as compared to individuals who are unaffected. Currently, association studies are typically performed in one of two ways (i) case control study; where affected individuals are matched to unaffected, unrelated controls based on a number of criteria such as age, sex, area of origin. The allelic distribution is then compared between the affected group and the matched unaffected control group, with a significant difference being indicative of an association between the polymorphism and the disease, (ii) family based approach; where the sample is comprised of a collection of family trios consisting of an affected proband and his/her parents. The non-transmitted alleles from the parents represent a random sample of alleles from the population from which the proband was sampled and therefore they can be used to create a well-matched control sample (Falk and Rubenstein, 1987).

Traditionally, the case-control design has been employed for association studies as it is widely acknowledged that there are several advantages associated with this approach. Suitable participants are relatively easy and inexpensive to recruit for case-control studies. A representative sample of subjects is more easily achieved. Issues of sample recruitment might result in certain biases with regard to family based studies; generally both parents are required to be alive for family based studies, thereby possibly imposing bias towards early onset forms of a disease. Also it is possible that parents without a psychiatric disorder could be over-represented in trio studies as it has been observed that the mortality rate is usually increased for disorders such as bipolar disorder. This situation might result in the under-representation of dominant-like risk genes and/or the over-representation of phenocopies (Bruun and Ewald, 1999). Genotype-environment interactions can be examined in case-control studies (Khoury, 1994) and it has been noted that the power to detect pathogenetically relevant associations can be substantially greater under some plausible models with this methodology (Long and Langley, 1999; Craddock *et al.*, 2001).

However, despite the significant advantages associated with the case-control design, it is acknowledged that the power of this methodology is considerably compromised by the possibility of spurious associations that can arise from unsuspected population stratification. The alternative family based approach is robust to such stratification and offers the additional advantage of being able to easily detect genotyping inconsistencies, but this approach is limited in its applicability for diseases with onset in adulthood, such as bipolar disorder, as parents are frequently unavailable for participation in these studies. It is obvious that there are strengths and weaknesses associated with both the case-control and the family based designs, but undoubtedly each have an important role to play in the investigation of any complex disease including bipolar disorder. The current understanding of the pathophysiology of bipolar disorder is relatively poor, hence to date candidate gene studies have focused mainly on the neuro-transmitter systems, as these appear to be the prime targets for the medications used in the clinical management of this disorder. Numerous studies of mood disorders have examined polymorphisms within genes encoding receptors, transporters and enzymes involved in the action, re-uptake or metabolism, of dopamine, serotonin and noradrenaline and while no study has produced the conventional levels of statistical significance, there have been some positive findings which await validation.

The candidate gene association studies performed during this study, were selected primarily based on findings obtained from a modified genome scan, undertaken by our group (Murphy *et al.*, 2000), for bipolar disorder within the Irish population (details in 1.13); in contrast, the studies involving the COMT gene were aimed at furthering previous association study findings. The region highlighted on chromosome 17q was found to include the serotonin transporter gene (5-HTT), which is generally recognised as a plausible functional candidate gene for bipolar disorder. Well-characterised polymorphisms are associated with this gene and have been extensively investigated in bipolar disorder by several independent groups. This gene was therefore the obvious candidate for further investigation of this region using association studies. The catechol-O-methyltransferase gene (COMT) is also an acknowledged possible functional candidate gene for involvement in bipolar disorder and association studies with this gene were performed in the context of furthering previous findings within a subset of the extended cohort presently available for analysis. Early linkage studies conducted on chromosome 14q, suggested a region of interest that was known at a relatively early

stage of linkage analysis to be in the region of the estrogen receptor beta (ERβ) gene. Because of the involvement of this gene in mediating the many functions of the hormone estrogen, and in addition, the recognised possibility that steroid hormones may be involved in the regulation of mood disorders, the ER β gene was therefore chosen as a functional candidate for association studies with bipolar disorder.

Family trio and/or case-control samples (see sections 2.1.4 and 2.1.5) were used for all association studies performed here and appropriate methods of analysis were employed to test for association.

5.1.2 Serotonin Transporter (5-HTT)

Serotonin, a major neurotransmitter in the central and peripheral nervous system, has been extensively implicated in the control of mood, cognition, sleep, and a variety of complex behaviours and psychiatric disorders. Further to it's significant role as a neurotransmitter, various studies have suggested that serotonin acts as a regulator of morphogenic activities such as cell proliferation, migration, and differentiation during the early development of the central nervous system (Lauder, 1993). The serotonin transporter (5-HTT), which regulates the extracellular serotonin level by recycling serotonin back into the presynaptic neuron, has been proposed as a candidate gene for bipolar disorder. In addition to its critical involvement in serotonin metabolism, it also represents the site of action of many antidepressant drugs. Furthermore, reduced binding to platelet 5-HTT has been suggested as an endophenotype/trait marker for bipolar disorder (Leboyer *et al.*, 1999).

Two polymorphisms within the serotonin transporter gene have been studied extensively; the first is a variable number of tandem repeats (VNTR) in the second intron (three alleles: Stin2.9, Stin2.10, Stin2.12). Stin 2.12 and Stin2.10 have been found to act as consistent positive transcriptional regulatory elements of 5-HTT in mouse embryonic rostral hind brain, with Stin 2.12 being significantly stronger (Mackenzie and Quinn, 1999). The second polymorphism is a 44 base pair insertion/deletion within the promoter region that displays predominantly two allelic forms, the long variant (l; 16 repeat alleles) and the short variant (s; 14 repeat alleles). The short variant of the 5-HTT linked polymorphic region, was found to restrict the

transcriptional activity of the promoter, leading to low functional expression of the serotonin transporter (Collier *et al.*, 1996), which may have physiological consequences in vivo by altering the uptake of serotonin. However, a total of 14 different variants have been detected in the 5-HTTLPR with the long and short alleles being divided into four and six kinds of allelic variants respectively. Significant ethnic differences in allele distribution and the genotype distribution have been identified. In a study carried out by Nakamura *et al.*, (2000), six types of alleles were detected (14-A, 14-D, 16A, 16-D, 16-E and 16-F) in a Caucasian sample whereas in a Japanese sample eleven types of alleles were found. While most of the newly reported variants seem to have low frequencies, it has been suggested that ambiguous associations between diseases and 5-HTTLPR may be due to analysis discriminating only the s and l types and that maybe a more complete subdivision of 5-HTTLPR alleles could be warranted in future studies (Nakamura *et al.*, 2000).

There have been various conflicting reports regarding association between bipolar disorder and 5-HTTLPR and VNTR. Collier *et al.* (1996) carried out a case control study and reported an allelic association between allele 12 of the VNTR and bipolar mood disorder. This finding was replicated by several groups (Kunugi *et al.,* 1997; Rees *et al.*, 1997; Kirov *et al.*, 1999). In contrast however, many case-control samples failed to reproduce the association with allele 12 (Bellivier *et al.,* 1997; Furlong *et al.*, 1998; Gutiérrez *et al.*, 1998; Hoehe *et al.*, 1998; Saleem *et al.*, 2000). A similar situation arises in regard to the promoter polymorphism. In a large European collaborative case-control study, Collier *et al.*, (1996) found significant association between bipolar disorder and the low transcriptional activity variant of 5-HTTLPR, a finding reproduced by Furlong *et al.*, (1998). However, again subsequent studies have resulted in negative findings (Rees *et al.*, 1997; Esterling *et al.*, 1998; Vincent *et al.*, 1999; Mundo *et al.*, 2000; Ospina-Duque *et al.*, 2000). The majority of these reports with the exception of Esterling *et al.*, (1998), Kirov *et al.*, (1999), Bocchetta *et al.*, (1999) and Mundo *et al.*, (2000) were all case control studies. As previously mentioned such studies are subject to population stratification and hence the production of false positive results is a real possibility. An additional polymorphism within a putative polyadenylation signal for one of the commonly used polyadenylation sites was identified (3'UTR G/T) (Battersby *et al.*, 1999). Although a case-control study showed no significant difference in abundance of the two alleles in bipolar disorder, this polymorphism represents another

useful genetic marker in the evaluation of the role of the serotonin transporter in mood disorders.

The modified genome scan undertaken by our group and previously discussed in chapter one (section 1.13), revealed excess allele sharing at markers flanking the gene encoding the serotonin transporter ([D17S841, NPL 1.59, p = 0.06] - 7cM - [D17S933, NPL 1.49, p = 0.07]; broad affection model). This finding therefore warranted further investigation, even though previous linkage studies using classic parametric analysis did not support the presence of a susceptibility locus for bipolar disorder near the serotonin transporter (Kelsoe *et al.*, 1996; Ewald *et al.*, 1997). Thus, it was decided to perform linkage disequilibrium studies on the serotonin transporter using a combination of family-based association tests and multimarker haplotypes. Traditionally, the tendency has been to examine polymorphisms individually but it is hypothesised that the extensive use of haplotypes of many markers within and around each gene may be a more powerful approach. In support of this proposition, it is observed that a disease causing mutation will be in linkage disequilibrium with an entire haplotype of surrounding alleles. Therefore, the ability to investigate the entire region of the gene must provide considerably more information than examination of a single polymorphism. Moreover, if multiple mutations exist within a gene, which have a similar effect on susceptibility to disease, the use of haplotypes may be better suited to revealing the association, as examination of single polymorphisms could result in the complex pattern of associations cancelling each other out (Collier and Sham, 1998). Haplotype methodology has been extensively employed in association studies examining the possible involvement of various candidate genes in bipolar disorder. Recently, haplotype analysis revealed evidence to suggest that a DNA variant in the brain-derived neutrophic factor (BDNF) locus may confer susceptibility to bipolar disorder (Neves-Pereira *et al.*, 2002), and in a separate study, this methodology indicated that variations within the serotonin 4 receptor (5-HT4) gene might be involved in the aetiology of mood disorders (Ohtsuki *et al.*, 2002). For this study, findings from family based association tests and haplotype analysis on the serotonin transporter gene were followed up with a case-control study.

51.3 Catechol-O-methyltransferase (COMT)

Catecholamines, such as dopamine, adrenaline and noradrenaline, have long been hypothesised to be involved in the aetiology of affective disorders (Schildkraut, 1965). There are several lines of evidence suggesting that the catechol-O-methyltransferase gene (COMT) is a plausible candidate gene for involvement in bipolar disorder. Firstly, COMT, a ubiquitous enzyme, catalyses the O-methylation of the catecholamine neurotransmitters in addition to other xenobiotic catechols, resulting in their inactivation. COMT thereby represents one of the major degradative pathways of endogenous catecholamine neurotransmitters (Kopin, 1994). Secondly, the COMT gene is located on 22q11, which is microdeleted in velo-cardio-facial syndrome (VCFS). In addition to the characteristic presentation of this syndrome, there appears to be a marked prevalence of psychiatric illness among affected VCFS individuals; an observation first reported by Shprintzen *et al.,* (1992). This would suggest that individuals, hemizygous for a gene/genes located in this region experience haploinsufficiency effects, possibly contributing to the manifestation of psychiatric illness (Scambler *et al.,* 1992). Attempts to characterise these psychiatric manifestations initially revealed associations between deletions of 22q11 and psychotic illness, specifically schizophrenia (Pulver *et al.,* 1994) and more recently bipolar disorder (Papolos *et al.,* 1996). Further evidence implicating chromosome 22q11 in bipolar disorder has come from a linkage study performed by Lachman *et al.,* (1997). Three microsatellite markers spanning 13cM around the VCFS region were examined in 13 bipolar pedigrees. However, as the marker providing the strongest evidence in this study lies 5cM distal to the region deleted in VCFS, it has been suggested that this region of chromosome 22 may harbour two susceptibility loci for bipolar disorder. The COMT gene presents in two forms (soluble and membrane bound), which differ in the presence or absence of a 50 amino acid hydrophobic N-terminal region; both forms being encoded by a single gene. A functional polymorphism has been characterised that is responsible for the substantial variability in COMT enzymatic activity within the population. A relatively low activity allele is found to be associated with a methionine residue at position 108/158 (soluble/membrane bound) whereas the high activity variant has a valine residue at this position (Lotta *et al.,* 1995). Lachman *et al.,* (1996) developed a straightforward PCR assay to distinguish the polymorphic genetic variants within the COMT gene. Several association studies have since been undertaken to examine the possible involvement of this functional polymorphism in the aetiology of bipolar disorder. Results are conflicting with positive

findings being reported by some groups for an association between bipolar disorder and the low activity allele of COMT (Ohara *et al.,* 1998; Li *et al.,* 1997) and negative findings being reported by others (Kirov *et al.,* 1999; Biomed European Bipolar Collaborative Group, 1997; Kunugi *et al.,* 1997; Gutierrez *et al.,* 1997). However, a recent meta-analysis of 13 case-control samples including 910 individuals suffering from bipolar disorder and 1069 matched controls provides evidence that the low activity allele may increase susceptibility to the illness, although the estimated effect size is quite modest [pooled estimate of OR 1.18 (CI 1.02-1.35)] (Craddock *et al.,* 2001).

An intriguing finding has been the discovery of a sexual dimorphism with respect to the low activity allele, which was first reported in association with another psychiatric disease, namely obsessive compulsive disorder (OCD) (Karayiorgou *et al.,* 1997). Employing a case-control sample of 42 male and 31 female affected individuals and 148 ethnically matched controls, they showed that the low activity allele was significantly associated in a recessive manner with susceptibility to OCD, particularly in male patients (allelic association males χ^2 =13.68, p = 0.002; females χ^2 = 0.01, p= 0.9097). A follow-up study conducted by Gogos *et al.,* (1998), provided evidence of an important sex- and region-specific contribution of COMT in the maintenance of steady-state levels of catecholamines in the brain and suggested a role for COMT in some aspects of emotional and social behaviour in mice. Specifically, homozygous female (but not male) mice displayed impairment in emotional reactivity in the light/dark exploratory model of anxiety and heterozygous COMT-deficient male mice exhibited increased aggressive behaviour. Mutant mice also demonstrated region-specific changes of dopamine levels, notably in the frontal cortex, a region of the brain that has previously been implicated in bipolar disorder (Strakowski *et al.,* 2000). To provide a possible biological basis to the observed sex-specific association, Karayiorgou and colleagues (1997) suggested that oestrogens might epigenetically reduce COMT activity. Recently, the low activity allele has also been found to be associated with female BPI probands (Mynett-Johnson *et al.,* 1998), which fits the hypothesis that 'epigenetically challenged' females carrying the low activity allele may be more susceptible to bipolar disorder. Of further interest is the finding by several independent groups that the low activity allele could represent a modifying gene that predisposes to rapid-cycling in patients with bipolar disorder (Kirov *et al.,* 1998; Papolos *et al.,* 1998; Lachman *et al.,* 1996). This condition is more commonly found in female patients and the possibility that the low

activity allele may be associated with rapid cycling is consistent with the tendency for anti-depressants to induce rapid cycling; both resulting in an increase of the availability of catecholamines at the neuronal synapses.

The studies described here were performed within the context of furthering a previous association detected between bipolar disorder and female probands within the Irish population. Four single nucleotide polymorphisms in addition to the previously examined functional polymorphism, were investigated using appropriate statistical methods to enable the detection of association, in an extended trio and combined trio and case-control female cohort. Haplotype analysis was also performed, the advantages of which have been previously discussed. In a further effort to increase the power of the study, SDT analysis was performed to enable the inclusion of families of less than ideal structure into the study.

5.1.4 Estrogen Receptor Beta (ERβ)

In addition to their well-established involvement in reproductive functions, estrogens have long been recognised as exerting a variety of actions on many regions of the nervous system thereby influencing a broad spectrum of behavioural and physiological functions. Recent evidence from clinical and experimental studies supports a protective role for estrogen in diseases such as Alzheimer's disease and schizophrenia (Garcia-Segura *et al.*, 2001). Moreover, estrogen mediated mechanisms are continuously being established as having a role in alterations to pathways involved in neurotransmission, such as the dopaminergic and serotonergic pathways, thereby implicating the possible involvement of estrogens in mood disorders. Some of the variations in mood, mental state, and cognition, in both males and females, are thought to be mediated by the action of estrogen on the 5-hydroxytryptamine$_{2A}$ receptor (5-HT$_{2A}$R) and the serotonin transporter (SERT) in the brain (Fink *et al.*, 1998). Several of the actions of testosterone appear to depend on the conversion of testosterone to estradiol, by the enzyme aromatase, and therefore the action of testosterone on these two serotonergic mechanisms may be similar to that of estrogen (Fink *et al.,* 1999). Of particular note is the observation that there is a significantly higher prevalence of major depression in women as compared to men (Weissman *et al.*, 1988). Periods of hormonal fluctuations or estrogen instability (i.e. premenstrual, postpartum, perimenopausal) have been

associated with increased vulnerability to depression among susceptible women (Halbreich and Kahn, 2001). There is evidence to suggest that depression may be rapidly reduced in patients with postpartum symptoms who have documented estradiol deficiency with the administration of 17beta-estradiol (Ahokas *et al.*, 2001; Gregoire *et al.*, 1996) and in some studies hormome replacement therapy has been found to improve mood in symptomatic peri- and postmenopausal women with mild depressive symptoms (Ditkoff *et al.*, 1991; Sherwin, 1988). In males it has been reported that a decrease in levels of testosterone is accompanied by changes in mood and mental state (Vermeulen, 1993) and that testosterone replacement therapy has been known to improve mood in hypogonadal men (Wang, 1996). The specific multifaceted interactions of estrogen with the central nervous system and the endocrine system have yet to be fully elucidated but it appears that estrogen, or its absence, is strongly implicated in the regulation of mood and behaviour, as well as the pathobiology of mood disorders, especially in women.

The effects of estrogen in both males and females are most likely mediated through the estrogen receptors. To date two estrogen receptors have been identified, ER α and ER β. Both are members of the superfamily of nuclear receptors involved in the transduction of extracellular signals into transcriptional responses. ER α was first identified by Green *et al.*, (1986), and subsequently a subtype of this estrogen receptor, now known as ER β, was identified and characterised by Mosselman *et al.*, (1996). Recent advances in brain imaging techniques have enabled the identification of regions of the brain that may be involved in mood disorders. To date the prefrontal cortex, medial temporal lobe, striatum and cerebellum, which participate in two inter-related brain circuits, have been implicated in the pathophysiology of mood disorders (Soares, 2002) and whilst the hypothalamus and the pituitary are the primary sites for the estrogen receptors, nerve cells containing ER α and ER β have recently been identified in these regions (McEwen, 2001).

Most published linkage studies providing data about chromosome 14 have been undertaken within the context of whole genome screens for linkage (discussed in Craddock and Lendon, 1999). While current data does not provide compelling evidence for a region of interest on chromosome 14, there are modestly interesting findings for both bipolar disorder and schizophrenia and they fall into two clusters. The first is

located on proximal 14q, where genome scans conducted by Detera-Wadleigh *et al.*, (1998) and more recently, Cichon *et al.*, (2002) revealed evidence for suggestive linkage at 14q12-q13 for bipolar disorder. A balanced translocation has also been identified between proximal 14q and chromosome 18 [t(14;18)(q11.2;q22.1)] in a woman displaying psychiatric illness that includes bipolar disorder (Overhauser *et al.*, 1998). Genome scans carried out by Detera-Wadleigh *et al.*, (1998), McInnis *et al.*, (1998) for bipolar disorder and Faraone *et al.*, (1998) and Shaw *et al.*, (1998) for schizophrenia provided modest evidence for linkage to the second region of interest on chromosome 14, 14q22-q24. Furthermore a balanced translocation between this region of chromosome 14 and chromosome 18 [t (14;18)(q22;q21)] has been identified in schizophrenics (Goodman, 1994). However, it should be noted that many other genome scans for bipolar disorder and schizophrenia have not identified regions of interest on chromosome 14 (Berrettini *et al.*, 1997 [BP]; Adams *et al.*, 1998 [BP]; Karayiorgou *et al.*, 1994 [SZP]; Levinson *et al.*, 1998 [SZP]). Also worthy of mention is the recent finding by Cichon *et al.*, (2002), who found evidence for putative maternally imprinted loci for bipolar disorder at 14q32. This region has previously been proposed as a candidate for containing imprinted genes (Martin *et al.*, 1999) and as it has previously been suggested that imprinted loci may play a role in the development of bipolar disorder (Gershon *et al.*, 1996) this region may be worthy of further investigation.

Evidence from linkage studies reported here in this study, indicated 14q21-q24, as possibly harbouring a susceptibility locus for bipolar disorder. The gene encoding ERβ is localised to 14q22-q24 (Enmark *et al.*, 1997) and in light of the previously discussed possible involvement of estrogen in mood disorders, this gene was chosen as a plausible candidate for association studies with bipolar disorder. The ER β gene is composed of 8 exons spanning approximately 41 kb (Enmark, 1997). Previously Rosenkranz *et al.*, (1998) screened the coding region and part of the 5′ and 3′ untranslated regions of the ER β gene for polymorphisms. Examining a sample of obese individuals as well as patients with anorexia nervosa and bulimia nervosa, a total of 5 different sequence variants were identified. Of these, an A1730G mutation in the 3′ untranslated region of exon 8 was selected for use in the present study (see table, 2.2). Although A1730G is a silent mutation and therefore unlikely to contribute directly to the disease, it was found to be highly polymorphic within the population originally studied (Rosenkranz, 1998).

5.2 Results

5.2.1 Serotonin Transporter (5-HTT)

Ninety- two families, with both parents available for genotyping were included in this study; 78 of these families had only one BPI child, whereas 14 families had 2 BPI children. Twenty of the 92 families had previously been included in the genome screen (Murphy *et al.,* 2000). Single marker and multimarker haplotype analyses were performed using both (i) 92 independent trios (ii) 106 total trios (92 trios plus the second trio from 14 families). However, given the prior evidence of linkage to this region, albeit modest, the use of 106 trios could only be considered as a test for linkage, not association. An additional sample of 29 families, lacking one (n=27) or both (n=2) parents but including BPI and unaffected siblings, was also incorporated in the study using SDT (Curtis *et al.,* 1999); SDT is a valid test of association even in the presence of non-independent trios.

PCR reactions were performed for each of the three polymorphisms, VNTR, 5-HTTLPR, 3'UTR (see 2.3, 2.7, 2.8 and 2.9). The transmission of alleles from heterozygous parents (TDT; Spielman *et al.,* 1993) for each of the hSERT polymorphisms were examined and the results are given in table 5.1. Results for 92 independent trios indicated that only 5-HTTLPR showed a statistically significant distortion in allele transmission, with the short allele (deletion) being preferentially transmitted to affected individuals (56 transmissions [T] versus 37 nontransmissions [NT], $\chi^2 = 3.88$, df 1, $p = 0.049$). Although the VNTR displayed virtually equal allele transmissions, the results for 3'UTR G/T approached significance in the test for linkage (104 trios, 61 T versus 42 NT, $\chi^2 = 3.51$, df 1, $p = 0.06$).

Table 5.1: Pairwise TDT Analysis of Individual Polymorphisms*

	Alleles Transmitted	χ^2	df	p value
5-HTTLPR	Insertion (l)-Deletion (s)			
Independent trios (N = 92)	37-56	3.88	1	0.049
All trios (N = 106)	44-60	2.18	1	0.140
VNTR[a]	12-10			
Independent trios	51-44	0.52	1	0.471
All trios (N = 106)	59-50	0.74	1	0.389
3'UTR G/T	G-T			
Independent trios (N = 91)	39-53	2.13	1	0.144
All trios (N = 104)	42-61	3.51	1	0.061

* Based on allele transmissions from all heterozygous parents in 92 families. Total trio numbers (N = 106) are based on one trio per family (78 families) plus two trios per family (14 families).

[a] the rare VNTR allele 9 was pooled with allele 10.

The analysis of the three polymorphisms within hSERT coupled to the availability of complete parental data (n = 92 families), enabled the transmission of multimarker haplotypes to be examined. Two and three marker haplotypes were constructed by hand and haplotype transmission was examined using the extended transmission disequilibrium test (ETDT; Sham and Curtis, 1995). Empirical p values were obtained using MCETDT and 1000 replicates (Zhao *et al.,* 1999). The results, given in table 5.2, were based solely on those families in which haplotypes could be unequivocally assigned to all family members. All of the two-marker combinations, with the exception of 5-HTTLPR-VNTR, showed some evidence of association and/or linkage ($p< 0.05$). Moreover, combining the data of 5-HTTLPR and 3' UTR G/T together strengthened the association observed for 5-HTTLPR alone; the association being due to an excess transmission of haplotype s-T (24 T to 9 NT, independent trios, $p = 0.005$). Consideration of all three polymorphisms at once also gave statistically significant evidence of an association ($\chi^2 = 17.65$, df 8, $p = 0.04$) with the distortion in haplotype transmission occurring for haplotype s-12-T (15 T vs 3 NT).

Table 5.2: ETDT Analysis of Multimarker Haplotypes*

Marker Combination	Independent trios		All trios	
And Haplotype	**T**	**NT**	**T**	**NT**
5-HTTLPR-VNTR-3'UTR G/T[a]				
1-9-T	3	3	3	3
1-10-T	24	29	27	34
1-10-G	3	6	3	7
1-12-T	16	17	24	18
1-12-G	9	23	11	25
s-10-T	11	8	12	9
s-10-G	0	2	0	2
s-12-T	15	3	16	4
s-12-G	26	16	28	22
	$\chi^3 = 17.65$, df 8, p = 0.04		$\chi^2 = 16.60$, df 8, p = 0.049	
5-HTTLPR-VNTR[b]				
1-9	3	3	3	3
1-10	25	23	28	39
1-12	22	36	31	40
s-10	11	10	12	11
s-12	41	20	44	25
	$\chi^2 = 8.89$, df 4, p = 0.076		$\chi^2 = 6.34$, df 4, p = 0.186	
5-HTTLPR-3'UTR G/T[c]				
1-G	12	27	14	31
1-T	26	33	32	35
s-G	26	19	28	24
s-T	24	9	27	11
	$\chi^2 = 12.91$, df 3 , p = 0.005		$\chi^2 = 12.14$, df 3, p = 0.005	
VNTR-3'UTR G/T[d]				
9-T	4	3	4	3
10-G	3	8	3	9
10-T	33	35	37	41
12-G	31	40	34	47
12-T	37	22	45	23
	$\chi^2 = 6.01$, df 4, p = 0.217		$\chi^2 = 9.82$, df 4, p = 0.050	

*The statistic given refers to the global allelewise test (ETDT) calculated for fully informative families only. Empirical p values were obtained using MCETDT (1,000 replicates). The rare allele 9 (VNTR) was observed only on haplotype 1-9-T.

[a]n = 67 families. [c]n = 68 families.

[b]n = 72 families. [d]n = 79 families.

In an attempt to increase the power of the study, an additional 29 families were genotyped at each of the three polymorphisms. However, since these families had incomplete parental data, it was only possible to incorporate this data into the individual marker analysis, as definitive haplotypes could not be constructed. The results (using SDT; Curtis *et al.,* 1999) are given in table 5.3. All family data (n=121) was included in a single SDT input file, and the programme split families into the most appropriate analysis. Discordant sibling analysis alone failed to show evidence of association at any of the three polymorphisms. Only the TDT analysis revealed evidence of an association to 5-HTTLPR (T_{tdt}, $p = 0.022$), although the evidence was reduced slightly once data from the discordant siblings was combined in a joint SDT + TDT (all families) analysis (T_{both}, $p = 0.033$).

Table 5.3: Pairwise SDT Analysis of Individual Polymorphisms*

	5-HTTLPR			VNTR			3'UTR G/T		
	χ^2	df	p	χ^2	df	p	χ^2	df	p
T_{sdt}	0.059	1	0.808	0.333	1	0.564	0.333	1	0.564
T_{tdt}	5.226	1	0.022	1.382	2	0.501	3.082	1	0.079
T_{both}	4.570	1	0.033	1.094	1	0.579	1.988	1	0.158

* Based on the combined data from 92 families with complete parental data and 29 families with incomplete parental data. Data was analysed using SDT for discordant sibling analysis (T_{sdt}) alone, TDT analysis alone (T_{tdt}), and combined SDT and TDT (T_{both}).

Linkage disequilibrium between pairwise polymorphisms was assessed using parental genotypic data alone (n=92 families) and Associate (version 2.31). The predicted haplotype frequencies showed significant evidence of linkage disequilibrium between all pairwise combinations of the three polymorphisms. However, as a proportion of the maximum disequilibrium that could exist between the pairs of loci, the D' values were 0.35 (5-HTTLPR-VNTR), 0.30 (5-HTTLPR-3'UTR G/T) and 0.81 (VNTR-3'UTR G/T). Thus with the exception of the VNTR-3'UTR G/T combination, linkage disequilibrium had only a modest effect on the inheritance pattern of the loci.

It was decided to investigate if the association to the s-T haplotype could be replicated in the case-control sample. The cohort for this study consisted of a total of 108 BPI cases and matched controls (see 2.1.4), typed for all three polymorphisms. Given that it is not possible to generate definitive haplotypes for random individuals (cases and controls), the data generated was analysed by a neural network methodology, NNLD (Curtis *et al.*, 2001); the results are presented in table 5.4. The results from this analysis did not suggest that any single marker was associated with bipolar disorder and the network failed to find any indication of associated haplotypes.

Table 5.4: Output of NNLD programme for case-control sample.

	Single marker p values
Single marker p-value for locus 1	0.556886
Single marker p-value for locus 2	0.334953
Single marker p-value for locus 3	0.433459
Smallest single marker p-value	0.334953
Smallest single marker p-value (corrected)	1.004860
Neural network scores – Wilcoxon test	0.381782
Smallest p-value obtained	0.334953
Smallest p-value obtained (corrected)	1.339813

In an attempt to increase the power of the study all trio data was combined with the case-control data to produce a combined cohort of n=200 and genotypic and allelic χ^2 analyses were performed for each of the three polymorphisms. Artificial controls can be constructed for the trio samples from the parental untransmitted alleles; allele and genotype counts for these NT controls can then simply be combined with the control allele and genotype counts from the case-control sample. The results of this analysis are presented in table 5.5. The pooled test of all data remained significant for the preferential transmission of the short allele (deletion) of the promoter polymorphism to affected individuals.

5.2.2 Catechol-O-Methyltransferase (COMT)

As a follow up to the previously obtained tentative association (Codon 108/158, p = 0.043, uncorrected) between the low activity allele of the COMT gene and bipolar disorder in females, four single nucleotide polymorphisms (SNPs) within the COMT gene in addition to the functional polymorphism in exon 4, were examined in a group of 33 female BPI trios the details of which are presented in table 5.6

Table 5.6 Details of SNPs found within the COMT gene.

Polymorphism	Position	Nucleotide Change	Restriction Enzyme	Reference
Promoter	-287	A – G	*Hind* III	Levinson *et al.*, (1998) *Psych Genet* 8: 115-120
Exon 3	Codon 62	C – T	*Pml* I	Karayiorgou *et al.*, (1998). *Biol Psych.* 43: 425-431
Exon 4	Codon 136	C – G	*Bcl* I	Hoda et al., (1996) *BBRC* 228: 780-784
Exon 4	Codon 108/158	G – A	*Nla* III	Lachman et al., (1996) *Pharmacogenetics* 6:243-250
Exon 6	3'UTR	Insertion C	*Bgl* I	Chen et al., (1996) *Am. J. Med. Genet.* 67:556-9

Standard PCR protocol and cycling conditions were employed for each of the polymorphisms (details in section 2.3), the only variation being in the primer annealing temperatures; promoter 1 and both exon 4 polymorphisms @ 57°C and exon 3 and exon 6 polymorphisms @ 64°C followed by appropriate RFLP analysis (details in section 2.7 and table 2.2).

Several statistical methods have been proposed for the analysis of family data, which have been discussed in detail in section 1.7.3.3. The data generated in this study was analysed using pooled GHRR and HHRR tests as well as TDT; the results are presented in table 5.7

Table 5.7: χ^2 tests for genotype and allele frequencies (transmitted vs non-transmitted).

Polymorphism	GHRR						HHRR			TDT		
	Allele 1			Allele 2								
	χ^2	df	p value	χ^2	df	p value	χ^2	df	p value	χ^2	df	p exact
Promoter 1	0.159	1	0.69	0	1	0	0.032	1	0.86	0	1	1
Exon 3	1.991	1	0.158	0.31	1	0.578	1.11	1	0.292	1.5	1	0.3075
Exon 4 (*BclI*)	1.153	1	0.283	2.905	1	0.088	3.285	1	0.069	2.94	1	0.1214
Exon 6	4.25	1	0.039	2.885	1	0.089	4.492	1	0.034	5.76	1	0.0266

Genotype frequencies for the insertion/deletion polymorphism in the 3'UTR of exon 6 indicated that the 1,1 genotype was more frequent in the cases, a finding reflected by the statistically significant GHRR values for allele 1 (χ^2 = 4.25, df =1, p value =0.039). Allele frequencies were also found to be significant (HHRR, p value = 0.034) and the TDT test, which examines only the alleles transmitted to probands from heterozygous parents, also displayed a statistically significant association (χ^2 = 4.76, df =1, p_{exact} =0.0266). There also appeared to be a trend towards an increased presence of the 1,1 genotype in the cases (transmitted) as compared to the controls (non-transmitted) for the exon 4 (*Bcl*I) polymorphism. However, despite the observed tendency in the data, all statistical tests failed to reach the conventional levels of significance. The results for the promoter 1 polymorphism and exon 3 polymorphism did not reveal significant p values for any of the statistical tests employed in this analysis.

The availability of the parental genotypes enabled the absolute designation of haplotypes. Using the genotype data from the association studies for each of the four single nucleotide polymorphisms in conjunction with previously published data for the functional val-108-met polymorphism (Mynett-Johnson *et al.,* 1998), pairwise haplotypes were constructed by hand for all possible 2 marker combinations. ETDT analysis (discussed in 1.7.3.3 and 2.13.6) was employed to obtain allele-wise and genotype-wise measurements for the preferential transmission of paired haplotypes from parents to probands. The results are presented in table 5.8

Table 5.8: Extended TDT analysis - Transmission of haplotypes from heterozygous parents to probands (for all possible combinations of the 5 SNPs).

Polymorphsim Pair	Allele-wise p value	Haplotype transmitted
Promoter – Exon 3	0.61	
Promoter – Exon 4 (*Bcl* I)	0.47	
Promoter – Exon 4 (*Nla* III)	0.13	
Promoter – Exon 6	0.63	
Exon 3 – Exon 4 (*Bcl* I)	0.10	
Exon 3 – Exon 4 (*Nla* III)	0.17	
Exon 3 – Exon 6	0.21	
Exon 4 (*Bcl* I) – Exon 4 (*Nla* III)	0.05	1 – 2*
Exon 4 (*Bcl* I) – Exon 6	0.04	1 – 1*
Exon 4 (*Nla* III) – Exon 6	0.03	2 – 1*

*** For those Pairwise combinations giving p<0.05, the haplotype showing preferential transmission is shown.**

Two haplotype combinations produced results that either approached or achieved significant results for both allele-wise and genotype-wise p-values; exon 4 (*Bcl* I)/ exon 4 (*Nla* III) and exon 4 (*Nla* III) / exon 6.

To expand this study, an additional 15 female trios were genotyped for the functional polymorphism and the TDT analysis was repeated. The results of this analysis are presented in table 5.9

Table 5.9: TDT results for expanded female trio sample.

Trio Sample	Transmissions	Chi-Square Statistic	p value
Simplex trios (n=40)	17-23	$\chi^2 = 0.09$	P > 0.05
Independent trios (n=44)	18-26	$\chi^2 = 1.45$	P > 0.05
Total trios (n=47)	22-26	$\chi^2 = 0.3$	P > 0.05

In an attempt to increase the power of the study, all female trios (excluding 2nd sib) data was combined with the original case-control data from the female subgroup (n=37) to produce a combined cohort of n=82 and traditional χ^2 analysis was performed. For the trio sample artificial controls are constructed from the alleles not transmitted to the parents; allele and genotype counts for these NT controls can then simply be combined with the control allele and genotype counts from the case-control sample. The results of this analysis are shown in table 5.10.

Table 5.10: Genotypic and allelic frequencies for the combined female data set (n=82).

Genotype	Cases	Controls	Allele	Cases	Controls
1 1	14	30	1	75	95
1 2	47	35	2	89	69
2 2	21	17			
$\chi^2 = 7.995$	df =2	p<0.02	$X^2 = 4.885$	df=1	p <0.03

The pooled test of all data remained significant for the preferential transmission of the low activity allele of COMT to female probands.

An additional sample of 29 families, lacking one (n=27) or both parents (n=2) but including BPI and unaffected siblings was also incorporated into the study using SDT analysis (see section 1.7.3.3) (Curtis *et al.,* 1999). SDT is a valid test of association in the presence of non-independent trios. All family data (n=73) was included in a single SDT input file, and the programme split the families into the most appropriate analysis. The results of this analysis are presented in table 5.11

Table 5.11: Pairwise SDT analysis*

	Exon 4 – Codon 108/158		
	χ^2	df	p value
T_{sdt}	0.2	1	0.655
T_{tdtt}	0.64	1	0.423
T_{both}	0.82	1	0.366

*Based on combined data from 44 families with complete parental data and 29 families with incomplete parental data. Data was analysed using SDT for discordant sibling analysis (T_{sdt}) alone, TDT analysis alone (T_{tdt}), and combined SDT and TDT (T_{both}).

All three analyses, discordant sibling, TDT and combined SDT + TDT, failed to reveal evidence of association ($p > 0.05$).

Attempts were made to investigate whether any or all of the five polymorphic loci investigated were in linkage disequilibrium with one another. Parental genotypes for each of the ten pairwise combinations were tabulated by hand and entered into the EH programme (see section 1.7.3.3). Observed haplotype frequencies were then compared to those expected under the condition of no association, for evidence of linkage disequilibrium. χ^2 values were calculated by the EH programme, while the associated p-values at one degree of freedom were computed by the CHIPROB programme (part of linkage utilities package available at //www.hgmp.mrc.ac.uk). The results are presented in table 5.12

Table 5.12: Linkage disequilibrium analysis - χ^2 tests of linkage disequilibrium between each of the single nucleotide polymorphisms within the COMT gene.

Markers	X^2	d.f.	p-value
Prmoter 1/Exon 3	6.69	1	0.0097
Promoter 1/Exon 4 (*Bcl* I)	11.38	1	0.0007
Promoter 1/Exon 4 (*Nla* III)	7.77	1	0.0053
Promoter 1/Exon 6	2.06	1	0.1519
Exon 3/Exon 4 (*Bcl* I)	15.43	1	0.00008
Exon 3/Exon 4 (*Nla* III)	44.10	1	0.00000
Exon 3/Exon6	12.75	1	0.00036
Exon 4 *(Bcl* I)/Exon 4 (*Nla* III)	67.98	1	0.00000
Exon 4 (*Bcl* I)/Exon 6	6.00	1	0.01433
Exon 4 (*Nla* III)/Exon 6	11.34	1	0.00076

The results revealed that all five polymorphisms are in strong linkage disequilibrium with each other. An exception to this was the failure to detect linkage disequilibrium between the promoter 1 and exon 6 polymorphisms.

5.2.3 Estrogen Receptor Beta (ERβ)

Analysis of the sequence and the polymorphism using GENEJOCKEY revealed that an *Alu*I restriction site was affected by the A/G polymorphism. The assay was performed as described in section 2.7. The characteristic band pattern for this RFLP following gel electrophoresis is shown in figure 5.1.

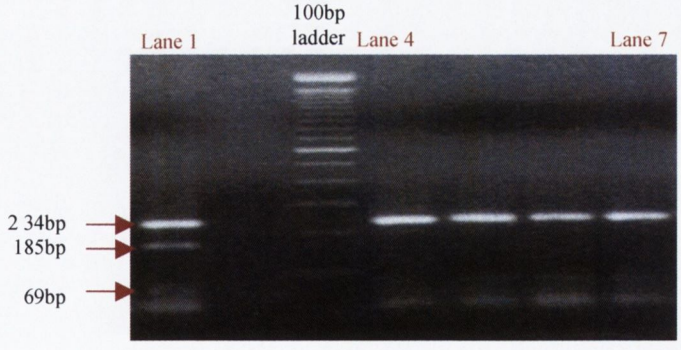


Figure 5.1 SNP at bp 1730 in 3'UTR of exon 8 resolved using RFLP analysis. Presence of the mutant G allele creates a recognition site for the restriction enzyme *Alu*I, producing bands 185bp and 69 bp. The uncut fragment is visualised as a 234bp band in lanes 4, 5, 6 and 7, corresponding to samples D127, D128, D155 and D156. Lane 1, corresponding to sample D30, represents a heterozygote with the bands 234bp and 185 bp clearly visible. The lower band of 69bp is present although it is not clearly visible, as it tended to run with a primer dimer.

Eighty-nine families, with both parents available for genotyping, were included in the family based association study; 76 of these families had only one BPI child, whereas 13 families had two BPI children. The transmission disequilibrium test (TDT), which tests for linkage in the presence of association, was employed to analyse the genotypic data for a single nucleotide polymorphism within the ER β gene (3'UTR, A1730G) in a total of 102 trios where the proband had been diagnosed as BPI. The trios were also divided on the basis of the gender of the proband (53 male, 49 female) and the TDT test was applied to each subset: all results are presented in table 5.13. Given that the sample was composed of both simplex and multiplex families, allele transmissions were examined for (i) simplex families only (n = 76), (ii) one trio from each family (n = 89 trios), (iii) all trios (n = 102), applying a correction (Martin *et al.,* 1997), for non-independence of multiple sibs in some families. The allele transmissions from heterozygous parents to their affected offspring, was determined for 13 multiplex families available for this study, and the correction for non-independence (figure 5.2) was applied to the sample as a whole and also to the sample divided according to gender. Conditional on the parental genotypes, the set of transmissions from one parent to his or her affected offspring is independent of the set of transmissions from the other parent, if there is no linkage disequilibrium. In this way the independence property is maintained and valid χ^2 tests are obtained.

Table 5.14: Transmission of alleles to affected offspring from heterozygous parents for 13 multiplex families.

Transmitted Alleles	Number of parents (Whole)	Number of parents (Males only)	Number of parents (Females only)
S_{AA}	3	1	2
S_{GG}	4	1	3
S_{AG}	5	4	1

Only heterozygous parents from the 13 multiplex families were examined to determine whether they transmitted the same allele (S_{AA} or S_{GG}) or different alleles (S_{AG}) to both affected sibs. These counts shown in table 5.14 were appropriately applied to the calculation shown below (figure 5.2) in conjunction with the required number A/G transmissions in simplex families, to correct for non-independence. The final corrected statistics are given in table 5.13.

$$\chi^2 = \frac{m-1}{m} \sum_{i=1}^{m} \left(\sqrt{\frac{2h_{xij} + h_{yij}}{\frac{h^*x}{hx}\ 4h_{xij} + h_{yij}}}\ 2s_{xii,jj} + s_{yi,j} - 2s_{xjj,ii} - s_{yj,i} \right)^2 \Big/ \ 2h_{xi} + h_{yi}$$

Figure 5.2: Correction formula for non-independence of sibs (Martin *et al.*, 1997)

$S_{xii,jj}$ = number of heterozygous parents from multiplex families who donate A allele to both affected children (s_{AA})

$S_{xjj,ii}$ = number of heterozygous parents from multiplex families who donate G allele to both affected children (s_{GG})

$S_{yi,j}$ = total number of A transmissions in simplex families only

$S_{yj,i}$ = total number of G transmissions in simplex families only

hx = total number of heterozygous parents who have two affected children ($S_{AA} + S_{GG} + S_{AG}$)

hx^* = total number of heterozygous parents who transmit the same allele to both affected children ($S_{AA} + S_{GG}$)

h_{xij} = heterozygous parents with two affected children ($S_{AA} + S_{GG} + S_{AG}$)

h_{yij} = heterozygous parents with single affected child (number of simplex families)

The results of the overall trios and subdividing by proband gender provided no evidence of preferential allele transmission. For the complete data set, an examination of the alleles transmitted from heterozygous parents revealed that there was virtually equal transmission of the A and G alleles (50 G transmissions versus 47 A transmissions, $\chi^2 = 0.093$, $p = 0.7604$ for all probands). Results for all other trio structures gave virtually identical p values. When the trios were subdivided according to the gender of the BPI proband, there was, once again, no significant difference in the transmission of the A allele versus the G allele to either male or female probands (male, $\chi^2 = 0.0909$, $p = 0.763$, $\chi^2 = 0.4714$, $p = 0.4922$).

5.3 Discussion

A variety of methodology has been used to extend the power of association studies to either investigate new candidate genes (ERβ) or to extend previous association study findings (COMT; hSERT) from the cohorts of Irish bipolar patients. These have included the construction of haplotypes across genes, the combining of family based and case control samples and the inclusion of more complex family structures, in the case of family based association studies.

5.3.1 Serotonin Transporter (5-HTT)

Disease association studies have in the last few years, undergone a transformation, in part due to the constant development and refinement of methodology designed to increase the reliability and informativeness of such studies. Consequently, the use of family based approaches has been advocated (Schaid, 1998) because they negate the problem of spurious false positives arising from poor matching of case control samples. Furthermore, the consideration of many markers, within and around the candidate gene has also been promoted (Collier and Sham, 1998), because this permits the examination of haplotype transmissions, which as previously discussed is likely to bring about greater power than the study of single polymorphisms alone. Since initial reports of association between bipolar disorder and polymorphisms within hSERT (5-HTTLPR: Collier *et al.,* 1996; VNTR: Collier *et al.,* 1996; 3'UTR: Battersby *et al.,* 1999), there have been numerous further association studies (Cusin *et al.,* 2001; Oliveira *et al.,* 2000; Liu *et al.,* 1999). However, to date only five of these association studies have been

family based (Dimitrova *et al.,* 2002; Piccardi *et al.,* 2002; Bocchetta *et al.,* 1999; Kirov *et al.,* 1999; Esterling *et al.,* 1998) and of these, only the study conducted by Kirov *et al.,* (1999) considered more than a single hSERT polymorphism.

This study revealed evidence of an association between bipolar disorder and 5-HTTLPR (TDT and SDT analysis), thus supporting other findings (Bellivier *et al.,* 1998; Collier *et al.,* 1996; Furlong *et al.,* 1998) and as in the previous studies it was the deleted/short promoter allele, known to have reduced transcriptional activity, that showed evidence of association. Although this study failed to show evidence of an association to the VNTR, the tendency for excess transmission of allele T at the 3'UTR G/T polymorphism is in agreement with the findings of Battersby *et al.,* (1999); allele T was more frequent albeit not significantly, in BPI cases compared with controls. Although a number of tests designed to detect evidence of association and/or linkage, have been applied to the data from three polymorphic markers, it should be noted that no correction for multiple testing was applied, because not all tests were independent of each other.

An examination of single polymorphisms and multimarker haplotypes provided some evidence of both association and linkage between bipolar disorder and variation within the serotonin transporter gene. Furthermore, results indicated that, at least for this study, consideration of multimarker haplotypes provided slightly greater power. The most significant association observed was for a two-marker combination of 5-HTTLPR and 3'UTR G/T (χ^2 = 12.91, df 3, p = 0.005) with haplotype s-T showing excess transmission, haplotype l-G reduced tranmission, and l-T and s-G approximately equal transmission/nontransmission. Although such data could be consistent with the two alleles acting in an additive manner with respect to disease susceptibility, alternative conclusions are probably more likely. Thus, the evidence obtained from this study is such that either one or both variants may be contributing to the disease or are in linkage disequilibrium with an as yet undefined susceptibility locus. This study suggests that the incorporation of the 3'UTR polymorphism into other family-based association analyses of the serotonin transporter would perhaps be helpful towards resolving the possible role of the serotonin transporter in bipolar disorder.

Further to these findings, it was decided to investigate if the association to the s-T haplotype could be replicated in the case-control sample, which included 108 BPI cases

and appropriately matched controls. This data was analysed by means of a new methodology, NNLD (Curtis *et al.,* 2001), which is designed to employ the pattern recognition ability of the neural network to detect, based on the genotype data provided, the occurrence of haplotypes or partial haplotypes that confer susceptibility to disease. This analysis did not suggest that any single marker was associated with bipolar disorder and the network failed to find any indication of associated haplotypes. Although the results of the case-control study did not replicate the findings in the trio study, it is believed the finding from the family based trio study remains worthy of further investigation as the relatively small size of the case-control study may have been a limiting factor in replicating the finding. The trio sample data was combined with the case-control data, using the non-transmitted parental alleles as artificial controls, and traditional case-control χ^2 performed. It was found that the pooled test of all data continued to show evidence of association between bipolar disorder and the deleted/short allele of the promoter. The VNTR and the 3'UTR did not reveal any evidence of association.

To address a possible functional hypothesis for this haplotype, future studies could include *in vivo* and *in vitro* expression analysis; reporter gene constructs/assays, competitive RT-PCR or real time PCR could be used to evaluate and quantify expression in comparison to suitable controls. The alternative hypothesis that the haplotype is in linkage disequilibrium with an uncharacterised susceptibility allele, could be investigated by screening the serotonin transporter gene coding and control regions for novel polymorphisms using DHPLC technology (discussed in detail in chapter 4). Such a study could be done on a sample of just 15 unrelated BPI probands; this would be sufficient to give 95% power to detect a genetic variant with a frequency of 5% in the general population (Speight *et al.,* 2000).

5.3.2 Catechol-O-methyltransferase (COMT)

A previous study conducted on a collection of 60 Irish BPI probands and their parents by Dr. Lesley Mynett-Johnson did not reveal evidence of either allelic or genotypic association between the COMT gene and bipolar disorder. However, when the data was re-analysed based on proband gender, it was found that the female group (n=30) did show evidence of association with the low activity allele of COMT (HHRR, $p = 0.043$)

although the significance of this result was lost when a Bonferroni correction for multiple testing was applied (p=0.084). The TDT test also displayed increased transmission of the low activity allele to female probands but it just failed to reach significance (TDT; χ^2 =4.17, df=1, p=0.06). The genotypic analysis also approached statistical significance (GHRR; allele 1, p=0.052, allele 2, p=0.21). Based on these results, a previously performed case-control study, which did approach the conventional level of statistical significance, was reanalysed according to proband gender. Again the female sample (n=37) did show some tendency towards association, with the genotypic test revealing a p value of 0.014, although the allelic test failed to reach significance (p=0.068). In an attempt to increase the power of this study, all data from the case-control study and the 60 nuclear families were pooled to produce a combined set of 147 BPI probands and 147 matched controls divided according to proband gender; females n=67, males n=80. The data was analysed using traditional case control methods resulting in a statistically significant finding for the female cohort (allelic, p=0.0072 and genotypic, p=0.012), although the significance was lost following a conservative correction.

These studies were extended here and initially, a cohort of 33 female trios, were examined for possible association between bipolar disorder and four independent polymorphisms found within the COMT gene. A comparison of the genotypes of parents and their offspring enabled 'control' genotypes to be constructed from the alleles that were not transmitted to the proband. The genotype and allele frequencies of the cases and 'controls' were investigated for statistically significant associations using GHRR and HHRR pooled tests. The transmission disequilibrium test (TDT), which tests for the preferential transmission of a particular allele from heterozygous parents to their affected offspring, was also applied to the data. Only the insertion/deletion polymorphism in exon 6 generated statistically significant results for an association between allele 1 and bipolar disorder in female probands (GHRR, p value = 0.039 for allele 1; HHRR, p value = 0.034; TDT, p value = 0.0266). Although the exon 4 (*Bcl* I) polymorphism displayed a trend towards the preferential transmission of allele 1 to the BPI proband, all statistical tests performed failed to reach the conventional levels of statistical significance. Likewise, all statistical analysis performed on the data generated for the promoter 1 polymorphism and the exon 3 polymorphism resulted in insignificant findings. The availability of parental genotypes also enabled the absolute designation of

haplotypes. It has been previously suggested that the construction and examination of haplotypes within and around a candidate gene may be a more powerful method to detect disease associations than studying a single polymorphism. The rationale being that inspection of haplotypes will lead to an increase of informativeness as a result of the haplotype being present at a lower frequency on non-disease chromosomes. Pairwise haplotypes were constructed and ETDT analysis was performed to determine if there was preferential transmission of paired haplotypes from parents to probands. Two haplotype combinations generated allele-wise and genotype-wide p values that either approached or achieved significant results (see table 5.8), pointing towards the possible identification of marker haplotypes that are associated with bipolar I disorder in females of Irish origin (i.e. 2/1 haplotype observed for the exon 4(*Nla* III)/exon 6 combination). However, no paired haplotype was found to exhibit an increase in significance over that displayed by individual markers suggesting that presentation of a particular haplotype did not imply a greater risk than the individual alleles. It is possible that the failure to strengthen the weak association observed in the female probands by the analysis of haplotypes may be a reflection of the sample size used in the study.

Attempts were also made to determine if any or all of the five polymorphic loci were in linkage disequilibrium (LD) with each other. Evidence of LD was established by comparing the observed haplotype frequencies with those expected under the condition of no association, using the EH programme (see section 1.7.3.3). Results from this analysis revealed that all five polymorphisms were in strong linkage disequilibrium with each other, with the exception of the promoter 1 polymorphism and exon 6 polymorphism. Variable mutation rates and/or large distances are the conventional explanations for failure to detect LD but in this case a more plausible explanation may be that the polymorphisms have had separate evolutionary histories.

In an attempt to increase the power of the study, 15 additional female trios were typed for the original marker examined (i.e. the functional polymorphism; Val 108 Met) and the TDT analysis was repeated. Increasing the size of the trio cohort, however, led to a loss in significance. The expanded female trio sample was then combined with the original case-control female sample and traditional case-control χ^2 analysis was performed. It was found that the pooled test of all data remained significant for the preferential transmission of the low activity allele of COMT to female BPI probands. A

further attempt at increasing the size of the data set and thus the power of the study was achieved by the incorporation of families of less than ideal structure into the TDT analysis. A sample of 29 families lacking one or both parents, but including BPI and unaffected siblings, was included and SDT analysis (details in 1.7.3.3) was performed. This analysis, however, did not reveal any evidence of association.

Observations from this study continue to support a tentative association between the functional val/met polymorphism and bipolar disorder among BPI female probands within the Irish population. In addition, results are also indicative of a preliminary association between the exon 6 insertion/deletion polymorphism and bipolar disorder within this same cohort. Furthermore, there is evidence to suggest that a bipolar haplotype involving the exon 6 polymorphism and val-108-met functional polymorphism, may be identifiable within this region. As previously discussed, a sexually dimorphic association has been reported for the functional polymorphism and obsessive compulsive disorder, whereby it has been proposed that estrogens could epigenetically reduce COMT activity (Karayiorgou *et al.,* 1997). In support of this theory it was recently found that estrogen can specifically down-regulate human COMT gene transcription (Xie *et al.,* 1999). These observations could provide an attractive explanation for the observed association in female BPI patients such that epigenetically challenged females possessing the low activity allele may be more susceptible to bipolar disorder. Future studies could involve further expansion of the cohort and also replication of the study with male patients to investigate if the association with the exon 6 polymorphism represents another sexual dimorphism within the COMT gene. The possible function of the proposed haplotype could also be investigated by means of expression analysis, reporter gene constructs/assays, competitive RT-PCR or real time PCR to evaluate and quantify expression in comparison to controls.

5.3.3 Estrogen Receptor Beta (ERβ)

It has been hypothesised that estrogens have a role to play in alterations to pathways involved in neurotransmission (McEwen and Alves, 1999) and therefore may be involved in the aetiology of neuropsychiatric disorders such as bipolar disorder.

The ERβ is one of two estrogen receptors that mediate the effects of estrogen, in both males and females. The gene encoding ERβ is located in a region of chromosome 14 (14q22-24), previously highlighted by our group as a region of interest based on the results of a modified genome scan (Murphy *et al.,* 2000). Five novel polymorphisms have been detected within the ERβ gene (Rosenkranz *et al.,* 1998) and of these, a silent A1730G mutation in the 3' untranslated region of exon 8 was selected for use in the present study. The transmission/disequilibrium test (TDT) was applied to the genotypic data from a total of 102 family trios, since such family based approaches negate the problem of false positive associations arising from ethnic stratification of case-control samples (Schaid, 1998). Even though the data includes multiple sibs from certain families, TDT remains a valid test of association since a correction for non-independence (Martin *et al.,* 1997) was incorporated into the analysis.

The data was also analysed on the basis of the gender of the BPI proband. It is plausible that females may represent a more homogeneous sample with regard to studying any possible role of ERβ in bipolar disorder, since in males the production of estrogen is dependant upon the conversion of testosterone to estradiol by the enzyme aromatase. It is thus possible that in males the effects of disease causing variations in ERβ are masked, in terms of detecting disease associations, by this conversion process. However, the transmission/disequilibrium test did not reveal evidence of preferential transmission of either allele to BPI probands in the sample as a whole or when the sample was divided on the basis of proband gender. Thus this present study has failed to provide evidence that genetic variation within ERβ plays a role in susceptibility to bipolar disorder. However, given the silent nature of the polymorphism examined, it is not possible to definitively exclude a role for ERβ in the aetiology of bipolar disorder.

6.1 Conclusions

Over the past few years this research group has continued to ascertain samples of DNA from Irish bipolar patients and their families with the aim of applying modern molecular genetical analysis to these collections. These and similar studies, conducted throughout the world, are aimed at furthering the knowledge of the fundamental causes not only of bipolar disorder, but of a whole array of neuropsychiatric diseases. The main body of research presented here was performed primarily within the context of following up on results previously obtained from a modified genome scan (Murphy *et al.,* 2000). An exception to this, were the association studies examining polymorphisms within the catechol-O-methyltransferase (COMT) gene; these studies were performed to further previous association study findings (Murphy *et al.,* 2001).

It is important to note that whilst, the Irish collections used in this research are relatively small in comparison to other collections (certainly this is true for the sib-pair sample used for non-parametric linkage analysis), the samples were ascertained by a small coordinated clinical team, thereby negating many of the standardisation problems associated with large multi-centre collaborative projects.

The previously conducted modified genome scan highlighted a number of chromosomal regions worthy of further investigation, which were then prioritised for additional study based on the strength of the finding and supportive evidence from studies conducted by other independent groups. The chromosomal regions thus chosen for further investigation were, 17q11-q12, 5q31.1-q35.1 and 14q22-q24.

Using microsatellite markers extensive linkage analysis was performed on 17q11-q12. Modest evidence continued to be found for a susceptibility locus within a 10cM region that is known to harbour the serotonin transporter gene, a long acknowledged candidate for bipolar disorder. To further examine the possible role of the serotonin transporter in bipolar disorder within the Irish cohort, family based and case-control association studies were undertaken. Analyses revealed evidence of an association between bipolar disorder and the deleted/short promoter allele, thereby supporting a number of other independent findings with this polymorphism and bipolar disorder. It was also observed that there was a tendency for the excess transmission of allele T at the 3'UTR G/T polymorphism. Examination of multimarker haplotypes, previously suggested to be a

more powerful approach than the examination of individual polymorphisms in isolation, revealed a significant association for a two marker combination of 5-HTTLPR and 3'UTR G/T with haplotype s-T showing excess transmission. As a follow-up to this finding, it was decided to investigate if the association with the s-T haplotype could be replicated in the case-control sample. The applied analysis did not suggest that any single marker was associated with bipolar disorder, nor was there any indication of associated haplotypes. It is recognised however, that the relatively small size of the case-control sample may have been a limiting factor in replicating the finding. Future studies might address the possible functional hypothesis for the s-T haplotype. Such studies could include *in vivo* and *in vitro* expression analyses, reporter gene constructs/assays, competitive RT-PCR or real time PCR, to evaluate and quantify expression in comparison to suitable controls. An alternative hypothesis that the haplotype is in linkage disequilibrium with an as yet unidentified susceptibility allele could be investigated by screening the serotonin transporter gene coding and control regions for novel polymorphisms using mutation detection technology such as denaturing high performance liquid chromatography (DHPLC).

The nature of the modified genome scan was such that the panel of microsatellite markers used did not include markers flanking the original marker showing interesting results for the region 5q31.1-q35.1. Thus, it was not possible to assess the strength of this finding based on information from close markers. However, as this region was alluded to in previous studies as possibly harbouring a susceptibility locus for bipolar disorder it was therefore considered worthy of further investigation. Upon incorporation of additional flanking markers, it was established that there was no evidence to support the presence of a potential susceptibility locus for bipolar disorder in this region of chromosome 5, thereby implying that the original result was most likely a false positive.

By far the most promising results from the extensive follow-up linkage analyses performed were obtained for the putative region on chromosome 14q22-q24; with the region of follow up extending to 32.9cM. Under the narrow definition of the disease, it was observed that seven consecutive microsatellite markers produced two-point NPL scores with associated p values < 0.05. Multipoint analysis also supported this finding, revealing a region of interest of approximately 34.92cM, where the p values are <0.05. Initial linkage studies suggested a region that was known at a relatively early stage to be

in the vicinity of the estrogen receptor beta (ERβ) gene. Due to the recognised possibility that steroid hormones may be involved in mood disorders, the ERβ gene was chosen as a functional candidate for association studies with bipolar disorder. The transmission disequilibrium test employed to assess evidence for association did not reveal preferential transmission of either allele of the silent A/G polymorphism in the 3'untranslated region of exon 8 to BPI probands in the sample as a whole or when the sample was divided on the basis of proband gender. Further linkage analysis extended the region of interest, and as current published data did not indicate an additional obvious candidate for bipolar disorder in this region of 14q, web based tools were employed to conduct an extensive search of the region in anticipation of locating a plausible candidate gene on which to perform association studies. The GTP cyclohydrolase I gene (GCHI), which may indirectly influence catecholamine and serotonin levels was chosen as an appropriate candidate gene for further studies. The coding region and putative promoter region of the gene were screened for the presence of novel polymorphisms using DHPLC. Four polymorphisms were obtained, three of which were located in the promoter region of the gene and the fourth was located in the first intron. Of the four polymorphisms, the novel G/A polymorphism at position –959bp was determined to be suitable for family based association studies. Analyses revealed that there was a statistically significant distortion in the transmission of allele A to BPI probands. As this polymorphism does not appear to affect known regulatory sites within the promoter, it is unlikely to be functional, although a definitive answer would require functional expression studies. There is the possibility that the polymorphism is in linkage disequilibrium with an as yet unidentified functional mutation that is contributing to the disease. Future work might entail refined localisation of the region of interest with a view to identifying further possible candidate genes for bipolar disorder.

Support for the presence of a susceptibility locus on 14q also comes from a recently conducted meta–analysis (Seguardo *et al.,* personal comm.) employing data from a total of 18 completed genome scans (excluding the data set used in this study). The method employed in this investigation was Genome Scan Meta-Analysis (Wise *et al.,* 1999). Essentially it involved defining 'bins' of fixed width across the genome. The marker set used in each genome scan was split into the bins and the maximum linkage score in each bin was noted. The regions of the genome were then ranked according to the

linkage score, with the subsequent summation of the ranks across all scans participating in the analysis. Computer simulations were conducted to derive a distribution for the summed rank under the null hypothesis of no underlying susceptibility gene. Significance was determined in the form of an empirical bin-wise (BW) p value from the distribution of simulations, representing the probability that any bin would achieve a summed rank higher than the observed value by chance. A total of five disease models were included in the analysis, but the two disease models considered for primary analysis were the most stringent disease model which incorporated BPI and schizoaffective disorder-bipolar subtype (SAB) (Model 1) and the intermediate model which, in addition to BPI and SAB, also included BPII (Model 2).

Although genome wide significance was not observed for any region in this meta-analysis, interesting trends were observed on chromosomes 8, 14 and 18. Under Model 2, 14q24.1-q32.12, attained a significance level which could be considered 'suggestive' (BW p-value = 0.003), although this threshold corresponds to that expected at least once per GSMA, by chance. However, there was cross diagnostic support for this region and corroborative evidence was also observed from adjacent bins (14q13.1-q24.1) thereby affording further reliability to the finding. It therefore appears from this meta-analysis, that 14q13.1-q32.12 represents a plausible location for the presence of a susceptibility locus common to many studies, thus providing further support for the linkage findings in this present work.

6.2 Future prospects

Replication studies for the GCHI positive association finding are essential to confirm the significance of this result. The presently available case-control sample may not provide the most suitable sample for such a study. Previously, the positive finding for the hSERT trio study was not replicated in this cohort, the limited size of the case-control sample possibly being a contributory factor to this non-replication. However, the association study could perhaps be replicated in another independent Irish family based sample.

Linkage studies for bipolar disorder continue to be plagued with non-replication of results, which may in part be due to the lack of consensus regarding the proper

definition of the affected phenotype and the questionable homogeneity of bipolar illness. Age of onset has frequently been successfully used as a key indicator in delineating disorder subtypes leading to gene identification e.g. Breast cancer, Alzheimer's disease, Diabetes. Genetic association and linkage studies also suggest differences between early and late onset bipolar affective disorder e.g. an association has been reported between the apolipoprotein gene polymorphism and early onset bipolar disorder (Bellivier *et al.,* 1997). Also recent segregation analysis indicated that the disease is transmitted differently in early and late onset groups (Grigoriu-Serbanescu *et al.,* 2001). Early onset BP I could possibly be a genetically homogeneous subtype of manic- depressive illness. The sib pair and the association cohorts employed in this research are continuously being expanded in order to increase the power of future studies. It may be possible to obtain an early onset BPI study group and incorporate such a group into future studies, thereby attempting to deal with the possibility of heterogeneity in the sample. Presently, trios with a proband age of onset of less than twenty years from our cohort will be employed in a multicentre collaborative study which will attempt to peform a genome scan aimed at identifying candidate regions using a large number of early-onset bipolar affected sibling pairs, and subsequently to perform fine mapping and candidate gene analysis in the regions of interest in anticipation of detecting candidate genes for bipolar disorder.

Reyes *et al.,* (2002) investigated a region on 18p11.2, which previously had revealed evidence for linkage to bipolar disorder (Detera-Wadleigh *et al.,* 1999) by constructing an array of overlapping BAC and PAC clones spanning an 8cM interval. Evans *et al.,* (2001) has adopted a similar approach as an initial step to further examining the region of interest on 4p. Reyes and colleagues generated a total of 29 new STSs and these were used to facilitate chromosome walking and ordering of clones. Their data in conjuction with the data available from the first draft of the human genome sequence permits the determination of the position and order of candidate genes and new STSs. They have succeeded in identifying genes worthy of further investigation, which are in close proximity to markers that have previously yielded suggestive linkage to bipolar disorder. This region of interest on 14q22-24, particularly in light of the positive findings for this region from the recently conducted meta-analysis for bipolar disorder (Seguardo *et al.,* personal comm.), demands further investigation to identify additional candidate genes. This would involve using the various database resources available and

possibly creating a physical map to obtain a more reliable, accurate order of the genes located in this region. It is proposed to concentrate further work initially on the region in close proximity to D14S281, which displays the strongest evidence for linkage in this study and which also happens to be in close proximity to the GTP cyclohydrolase I gene (GCHI).

BIBLIOGRAPHY

Abecasis, G.R., Noguchi, E., Heinzmann, A., Traherne, J.A., Bhattacharyya, S., Leaves, N.I., Anderson, G.G., Zhang, Y., Lench, N.J., Carey, A., Cardon, L.R., Moffatt, M.F. and Cookson, W.O. (2001) Extent and distribution of linkage disequilibrium in three genomic regions. *Am J Hum Genet*, **68**, 191-197.

Ackenheil, M. (2001) Neurotransmitters and signal transduction processes in bipolar affective disorders: a synopsis. *J Affect Disord*, **62**, 101-111.

Adams, L.J., Mitchell, P.B., Fielder, S.L., Rosso, A., Donald, J.A. and Schofield, P.R. (1998) A susceptibility locus for bipolar affective disorder on chromosome 4q35. *Am J Hum Genet*, **62**, 1084-1091.

Ahokas, A., Kaukoranta, J., Wahlbeck, K. and Aito, M. (2001) Estrogen deficiency in severe postpartum depression: successful treatment with sublingual physiologic 17beta-estradiol: a preliminary study. *J Clin Psychiatry*, **62**, 332-336.

Aita, V.M., Liu, J., Knowles, J.A., Terwilliger, J.D., Baltazar, R., Grunn, A., Loth, J.E., Kanyas, K., Lerer, B., Endicott, J., Wang, Z., Penchaszadeh, G., Gilliam, T.C. and Baron, M. (1999) A comprehensive linkage analysis of chromosome 21q22 supports prior evidence for a putative bipolar affective disorder locus. *Am J Hum Genet*, **64**, 210-217.

Akiskal, H.S., Bourgeois, M.L., Angst, J., Post, R., Moller, H. and Hirschfeld, R. (2000) Re-evaluating the prevalence of and diagnostic composition within the broad clinical spectrum of bipolar disorders. *J Affect Disord*, **59 Suppl 1**, S5-S30.

Alda, M. (2001) Genetic factors and treatment of mood disorders. *Bipolar Disord*, **3**, 318-324.

Allen, M.G., Cohen, S., Pollin, W. and Greenspan, S.I. (1974) Affective illness in veteran twins: a diagnostic review. *Am J Psychiatry*, **131**, 1234-1239.

American, A.P. (1994) *Diagnostic and statistical manual of mental disorders, 4th edition.* American Psychiatric Press, Washington, DC.

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

Antonarakis, S.E. (1994) Genome linkage scanning: systematic or intelligent? *Nat Genet*, **8**, 211-212.

Arias, B., Collier, D.A., Gasto, C., Pintor, L., Gutierrez, B., Valles, V. and Fananas, L. (2001) Genetic variation in the 5-HT5A receptor gene in patients with bipolar disorder and major depression. *Neurosci Lett*, **303**, 111-114.

Arnold, N., Gross, E., Schwarz-Boeger, U., Pfisterer, J., Jonat, W. and Kiechle, M. (1999) A highly sensitive, fast, and economical technique for mutation analysis in hereditary breast and ovarian cancers. *Hum Mutat*, **14**, 333-339.

Asherson, P., Mant, R., Williams, N., Cardno, A., Jones, L., Murphy, K., Collier, D.A., Nanko, S., Craddock, N., Morris, S., Muir, W., Blackwood, B., McGuffin, P. and Owen, M.J. (1998) A study of chromosome 4p markers and dopamine D5 receptor gene in schizophrenia and bipolar disorder. *Mol Psychiatry*, **3**, 310-320.

Avissar, S. and Schreiber, G. (1992) Interaction of antibipolar and antidepressant treatments with receptor-coupled G proteins. *Pharmacopsychiatry*, **25**, 44-50.

Badenhop, R.F., Moses, M.J., Scimone, A., Mitchell, P.B., Ewen, K.R., Rosso, A., Donald, J.A., Adams, L.J. and Schofield, P.R. (2001) A genome screen of a large bipolar affective disorder pedigree supports evidence for a susceptibility locus on chromosome 13q. *Mol Psychiatry*, **6**, 396-403.

Badner, J.A. and Gershon, E.S. (2002) Meta-analysis of whole-genome linkage scans of bipolar disorder and schizophrenia. *Mol Psychiatry*, **7**, 405-411.

Badner, J.A. and Gershon, E.S. (2002) Regional meta-analysis of published data supports linkage of autism with markers on chromosome 7. *Mol Psychiatry*, **7**, 56-66.

Bailer, U., Leisch, F., Meszaros, K., Lenzinger, E., Willinger, U., Strobl, R., Gebhardt, C., Gerhard, E., Fuchs, K., Sieghart, W., Kasper, S., Hornik, K. and Aschauer, H.N. (2000) Genome scan for susceptibility loci for schizophrenia. *Neuropsychobiology*, **42**, 175-182.

Baldessarini, R.J. (1970) Frequency of diagnoses of schizophrenia versus affective disorders from 1944 to 1968. *Am J Psychiatry*, **127**, 759-763.

Barford, P.A., Blair, J.A., Eggar, C., Hamon, C., Morar, C. and Whitburn, S.B. (1984) Tetrahydrobiopterin metabolism in the temporal lobe of patients dying with senile dementia of Alzheimer type. *J Neurol Neurosurg Psychiatry*, **47**, 736-738.

Battersby, S., Ogilvie, A.D., Blackwood, D.H., Shen, S., Muqit, M.M., Muir, W.J., Teague, P., Goodwin, G.M. and Harmar, A.J. (1999) Presence of multiple functional polyadenylation signals and a single nucleotide polymorphism in the 3' untranslated region of the human serotonin transporter gene. *J Neurochem*, **72**, 1384-1388.

Bellivier, F., Laplanche, J.L., Leboyer, M., Feingold, J., Bottos, C., Allilaire, J.F. and Launay, J.M. (1997) Serotonin transporter gene and manic depressive illness: an association study. *Biol Psychiatry*, **41**, 750-752.

Bellivier, F., Laplanche, J.L., Schurhoff, F., Feingold, J., Feline, A., Jouvent, R., Launay, J.M. and Leboyer, M. (1997) Apolipoprotein E gene polymorphism in early and late onset bipolar patients. *Neurosci Lett*, **233**, 45-8.

Bellivier, F., Schurhoff, F., Nosten-Bertrand, M., Mallet, J., Feingold, J. and Leboyer, M. (1998) Methodological problems in meta-analysis of association studies between bipolar affective disorders and the tyrosine hydroxylase gene. *Am J Med Genet*, **81**, 349-352.

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

Berrettini, W.H., Ferraro, T.N., Goldin, L.R., Detera-Wadleigh, S.D., Choi, H., Muniec, D., Guroff, J.J., Kazuba, D.M., Nurnberger, J.I., Jr., Hsieh, W.T., Hoehe, M.R. and Gershon, E.S. (1997) A linkage study of bipolar illness. *Arch Gen Psychiatry*, **54**, 27-35.

Berrettini, W.H., Ferraro, T.N., Goldin, L.R., Weeks, D.E., Detera-Wadleigh, S., Nurnberger, J.I., Jr. and Gershon, E.S. (1994) Chromosome 18 DNA markers and manic-depressive illness: evidence for a susceptibility gene. *Proc Natl Acad Sci U S A*, **91**, 5918-5921.

Bertelsen, A., Harvald, B. and Hauge, M. (1977) A Danish twin study of manic-depressive disorders. *Br J Psychiatry*, **130**, 330-351.

Biomed, E.B.C.G. (1997) No association between bipolar disorder and alleles at a functional polymorphism in the COMT gene. Biomed European Bipolar Collaborative Group. *Br J Psychiatry*, **170**, 526-528.

Birkett, J.T., Arranz, M.J., Munro, J., Osbourn, S., Kerwin, R.W. and Collier, D.A. (2000) Association analysis of the 5-HT5A gene in depression, psychosis and antipsychotic response. *Neuroreport*, **11**, 2017-2020.

Blackwelder, W.C. and Elston, R.C. (1985) A comparison of sib-pair linkage tests for disease susceptibility loci. *Genet Epidemiol*, **2**, 85-97.

Blackwood, D.H., He, L., Morris, S.W., McLean, A., Whitton, C., Thomson, M., Walker, M.T., Woodburn, K., Sharp, C.M., Wright, A.F., Shibasaki, Y., St Clair, D.M., Porteous, D.J. and Muir, W.J. (1996) A locus for bipolar affective disorder on chromosome 4p. *Nat Genet*, **12**, 427-430.

Blau, N., Thony, B., Renneberg, A., Arnold, L.A. and Hyland, K. (1998) Dihydropteridine reductase deficiency localized to the central nervous system. *J Inherit Metab Dis*, **21**, 433-434.

Blouin, J.L., Dombroski, B.A., Nath, S.K., Lasseter, V.K., Wolyniec, P.S., Nestadt, G., Thornquist, M., Ullrich, G., McGrath, J., Kasch, L., Lamacz, M., Thomas, M.G., Gehrig, C., Radhakrishna, U., Snyder, S.E., Balk, K.G., Neufeld, K., Swartz, K.L., DeMarchi, N., Papadimitriou, G.N., Dikeos, D.G., Stefanis, C.N., Chakravarti, A., Childs, B., Pulver, A.E. and et al. (1998) Schizophrenia susceptibility loci on chromosomes 13q32 and 8p21. *Nat Genet*, **20**, 70-73.

Bocchetta, A., Piccardi, M.P., Palmas, M.A., Chillotti, C., Oi, A. and Del Zompo, M. (1999) Family-based association study between bipolar disorder and DRD2, DRD4, DAT, and SERT in Sardinia. *Am J Med Genet*, **88**, 522-526.

Boehnke, M. (2000) A look at linkage disequilibrium. *Nat Genet*, **25**, 246-247.

Bonnier, B., Gorwood, P., Hamon, M., Sarfati, Y., Boni, C. and Hardy-Bayle, M.C. (2002) Association of 5-HT(2A) receptor gene polymorphism with major affective disorders: the case of a subgroup of bipolar disorder with low suicide risk. *Biol Psychiatry*, **51**, 762-765.

Bottiglieri, T., Hyland, K., Laundy, M., Godfrey, P., Carney, M.W., Toone, B.K. and Reynolds, E.H. (1992) Folate deficiency, biopterin and monoamine metabolism in depression. *Psychol Med*, **22**, 871-876.

Bowen, T., Kirov, G., Gill, M., Spurlock, G., Vallada, H.P., Murray, R.M., McGuffin, P., Collier, D.A., Owen, M.J. and Craddock, N. (1999) Linkage studies of bipolar disorder with chromosome 18 markers. *Am J Med Genet*, **88**, 503-509.

Brunn, T.G. and Ewald, H. (1999) Selection bias of susceptibility genes possible when using parent-offspring trios in genetic association studies. *Molecular Psychiatry*, **4**, 415-416.

Brzustowicz, L.M., Honer, W.G., Chow, E.W., Little, D., Hogan, J., Hodgkinson, K. and Bassett, A.S. (1999) Linkage of familial schizophrenia to chromosome 13q32. *Am J Hum Genet*, **65**, 1096-1103.

Bunzel, R., Blumcke, I., Cichon, S., Normann, S., Schramm, J., Propping, P. and Nothen, M.M. (1998) Polymorphic imprinting of the serotonin-2A (5-HT2A) receptor gene in human adult brain. *Brain Res Mol Brain Res*, **59**, 90-92.

Burgert, E., Crocq, M.A., Bausch, E., Macher, J.P. and Morris-Rosendahl, D.J. (1998) No association between the tyrosine hydroxylase microsatellite marker HUMTH01 and schizophrenia or bipolar I disorder. *Psychiatr Genet*, **8**, 45-48.

Burmeister, M. (1999) Basic concepts in the study of diseases with complex genetics. *Biol Psychiatry*, **45**, 522-532.

Byerley, W., Holik, J., Hoff, M. and Coon, H. (1995) Search for a gene predisposing to manic-depression on chromosome 21. *Am J Med Genet*, **60**, 231-233.

Camp, N.J., Neuhausen, S.L., Tiobech, J., Polloi, A., Coon, H. and Myles-Worsley, M. (2001) Genomewide multipoint linkage analysis of seven extended Palauan pedigrees with schizophrenia, by a Markov-chain Monte Carlo method. *Am J Hum Genet*, **69**, 1278-1289.

Cardno, A.G., Holmans, P.A., Rees, M.I., Jones, L.A., McCarthy, G.M., Hamshere, M.L., Williams, N.M., Norton, N., Williams, H.J., Fenton, I., Murphy, K.C., Sanders, R.D., Gray, M.Y., O'Donovan, M.C., McGuffin, P. and Owen, M.J. (2001) A genomewide linkage study of age at onset in schizophrenia. *Am J Med Genet*, **105**, 439-445.

Cardno, A.G., Marshall, E.J., Coid, B., Macdonald, A.M., Ribchester, T.R., Davies, N.J., Venturi, P., Jones, L.A., Lewis, S.W., Sham, P.C., Gottesman, II, Farmer, A.E., McGuffin, P., Reveley, A.M. and Murray, R.M. (1999) Heritability estimates for psychotic disorders: the Maudsley twin psychosis series. *Arch Gen Psychiatry*, **56**, 162-168.

Chartier-Harlin, M.C., Crawford, F., Houlden, H., Warren, A., Hughes, D., Fidani, L., Goate, A., Rossor, M., Roques, P., Hardy, J. and et al. (1991) Early-onset Alzheimer's disease caused by mutations at codon 717 of the beta-amyloid precursor protein gene. *Nature*, **353**, 844-846.

Chee, I.S., Lee, S.W., Kim, J.L., Wang, S.K., Shin, Y.O., Shin, S.C., Lee, Y.H., Hwang, H.M. and Lim, M.R. (2001) 5-HT2A receptor gene promoter polymorphism -1438A/G and bipolar disorder. *Psychiatr Genet*, **11**, 111-114.

Chee, M., Yang, R., Hubbell, E., Berno, A., Huang, X.C., Stern, D., Winkler, J., Lockhart, D.J., Morris, M.S. and Fodor, S.P. (1996) Accessing genetic information with high-density DNA arrays. *Science*, **274**, 610-614.

Chen, C.H., Lee, Y.R., Liu, M.Y., Wei, F.C., Koong, F.J., Hwu, H.G. and Hsiao, K.J. (1996) Identification of a BglI polymorphism of catechol-O-methyltransferase (COMT) gene, and association study with schizophrenia. *Am J Med Genet*, **67**, 556-559.

Chiaroni, P., Azorin, J.M., Dassa, D., Henry, J.M., Giudicelli, S., Malthiery, Y. and Planells, R. (2000) Possible involvement of the dopamine D3 receptor locus in subtypes of bipolar affective disorder. *Psychiatr Genet*, **10**, 43-49.

Cichon, S., Nothen, M.M., Rietschel, M., Korner, J. and Propping, P. (1994) Single-strand conformation analysis (SSCA) of the dopamine D1 receptor gene (DRD1) reveals no significant mutation in patients with schizophrenia and manic depression. *Biol Psychiatry*, **36**, 850-853.

Cichon, S., Nothen, M.M., Stober, G., Schroers, R., Albus, M., Maier, W., Rietschel, M., Korner, J., Weigelt, B., Franzek, E., Wildenauer, D., Fimmers, R. and Propping, P. (1996) Systematic screening for mutations in the 5'-regulatory region of the human dopamine D1 receptor (DRD1) gene in patients with schizophrenia and bipolar affective disorder. *Am J Med Genet*, **67**, 424-428.

Cichon, S., Schumacher, J., Muller, D.J., Hurter, M., Windemuth, C., Strauch, K., Hemmer, S., Schulze, T.G., Schmidt-Wolf, G., Albus, M., Borrmann-Hassenbach, M., Franzek, E., Lanczik, M., Fritze, J., Kreiner, R., Reuner, U., Weigelt, B., Minges, J., Lichtermann, D., Lerer, B., Kanyas, K., Baur, M.P., Wienker, T.F., Maier, W., Rietschel, M., Propping, P. and Nothen, M.M. (2002) A genome screen for genes predisposing to bipolar affective disorder detects a new susceptibility locus on 8q. *Hum Mol Genet*, **10**, 2933-2944.

Civelli, O., Bunzow, J.R. and Grandy, D.K. (1993) Molecular diversity of the dopamine receptors. *Annu Rev Pharmacol Toxicol*, **33**, 281-307.

Clegg, R.M., Murchie, A.I., Zechel, A., Carlberg, C., Diekmann, S. and Lilley, D.M. (1992) Fluorescence resonance energy transfer analysis of the structure of the four-way DNA junction. *Biochemistry*, **31**, 4846-4856.

Clerget-Darpoux, F. (2001) Extension of the LOD score: the mod score. *Adv Genet*, **42**, 115-124.

Collier, D.A., Arranz, M.J., Sham, P., Battersby, S., Vallada, H., Gill, P., Aitchison, K.J., Sodhi, M., Li, T., Roberts, G.W., Smith, B., Morton, J., Murray, R.M., Smith, D. and Kirov, G. (1996) The serotonin transporter is a potential susceptibility factor for bipolar affective disorder. *Neuroreport*, **7**, 1675-1679.

Collier, D.A. and Sham, P.C. (1998) The usual suspects: tyrosine hydroxylase and the serotonin transporter in affective disorders. *Mol Psychiatry*, **3**, 103-105.

Collier, D.A., Stober, G., Li, T., Heils, A., Catalano, M., Di Bella, D., Arranz, M.J., Murray, R.M., Vallada, H.P., Bengel, D., Muller, C.R., Roberts, G.W., Smeraldi, E., Kirov, G., Sham, P. and Lesch, K.P. (1996) A novel functional polymorphism within the promoter of the serotonin transporter gene: possible role in susceptibility to affective disorders. *Mol Psychiatry*, **1**, 453-460.

Comings, D.E., Gade, R., Wu, S., Chiu, C., Dietz, G., Muhleman, D., Saucier, G., Ferry, L., Rosenthal, R.J., Lesieur, H.R., Rugle, L.J. and MacMurray, P. (1997) Studies of the potential role of the dopamine D1 receptor gene in addictive behaviors. *Mol Psychiatry*, **2**, 44-56.

Cookson, J., Silverstone, T. and Wells, B. (1981) Double-blind comparative clinical trial of pimozide and chlorpromazine in mania. A test of the dopamine hypothesis. *Acta Psychiatr Scand*, **64**, 381-397.

Coon, H., Hoff, M., Holik, J., Hadley, D., Fang, N., Reimherr, F., Wender, P. and Byerley, W. (1996) Analysis of chromosome 18 DNA markers in multiplex pedigrees with manic depression. *Biol Psychiatry*, **39**, 689-696.

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

Coryell, W., Endicott, J., Keller, M. and Andreasen, N.C. (1985) Phenomenology and family history in DSM-III psychotic depression. *J Affect Disord*, **9**, 13-18.

Cox, N.J., Frigge, M., Nicolae, D.L., Concannon, P., Hanis, C.L., Bell, G.I. and Kong, A. (1999) Loci on chromosomes 2 (NIDDM1) and 15 interact to increase susceptibility to diabetes in Mexican Americans. *Nat Genet*, **21**, 213-215.

Craddock, N., Dave, S. and Greening, J. (2001) Association studies of bipolar disorder. *Bipolar Disord*, **3**, 284-298.

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

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

Craddock, N., McGuffin, P. and Owen, M. (1994) Darier's disease cosegregating with affective disorder. *Br J Psychiatry*, **165**, 272.

Crocq, M.A., Mant, R., Asherson, P., Williams, J., Hode, Y., Mayerova, A., Collier, D., Lannfelt, L., Sokoloff, P., Schwartz, J.C. and et al. (1992) Association between schizophrenia and homozygosity at the dopamine D3 receptor gene. *J Med Genet*, **29**, 858-860.

Curtis, D., Miller, M.B. and Sham, P.C. (1999) Combining the sibling disequilibrium test and transmission/disequilibrium test for multiallelic markers. *Am J Hum Genet*, **64**, 1785-1786.

Curtis, D., North, B.V. and Sham, P.C. (2001) Use of an artificial neural network to detect association between a disease and multiple marker genotypes. *Ann Hum Genet*, **65**, 95-107.

Curtius, H.C., Niederwieser, A., Levine, R. and Muldner, H. (1984) Therapeutic efficacy of tetrahydrobiopterin in Parkinson's disease. *Adv Neurol*, **40**, 463-466.

Cusin, C., Serretti, A., Lattuada, E., Lilli, R., Lorenzi, C., Mandelli, L., Pisati, E. and Smeraldi, E. (2001) Influence of 5-HTTLPR and TPH variants on illness time course in mood disorders. *J Psychiatr Res*, **35**, 217-223.

Davies, J.L., Kawaguchi, Y., Bennett, S.T., Copeman, J.B., Cordell, H.J., Pritchard, L.E., Reed, P.W., Gough, S.C., Jenkins, S.C., Palmer, S.M. and et al. (1994) A genome-wide search for human type 1 diabetes susceptibility genes. *Nature*, **371**, 130-136.

De bruyn, A., Souery, D., Mendelbaum, K., Mendlewicz, J. and Van Broeckhoven, C. (1996) A linkage study between bipolar disorder and genes involved in dopaminergic and GABAergic neurotransmission. *Psychiatr Genet*, **6**, 67-73.

Degn, B., Lundorf, M.D., Wang, A., Vang, M., Mors, O., Kruse, T.A. and Ewald, H. (2001) Further evidence for a bipolar risk gene on chromosome 12q24 suggested by investigation of haplotype sharing and allelic association in patients from the Faroe Islands. *Mol Psychiatry*, **6**, 450-455.

DeLisi, L.E., Shaw, S.H., Crow, T.J., Shields, G., Smith, A.B., Larach, V.W., Wellman, N., Loftus, J., Nanthakumar, B., Razi, K., Stewart, J., Comazzi, M., Vita, A., Heffner, T. and Sherrington, R. (2002) A genome-wide scan for linkage to chromosomal regions in 382 sibling pairs with schizophrenia or schizoaffective disorder. *Am J Psychiatry*, **159**, 803-812.

Detera-Wadleigh, S., Yoshikawa, T., Padigaru, M., Berrettini, W.H., Badner, J.A., Sanders, A.R. and al., e. (1998) Genome screen and candidate gene analysis in bipolar disorder. *Am J Med Genet*, **81**, 463.

Detera-Wadleigh, S.D., Badner, J.A., Berrettini, W.H., Yoshikawa, T., Goldin, L.R., Turner, G., Rollins, D.Y., Moses, T., Sanders, A.R., Karkera, J.D., Esterling, L.E., Zeng, J., Ferraro, T.N., Guroff, J.J., Kazuba, D., Maxwell, M.E., Nurnberger, J.I., Jr. and Gershon, E.S. (1999) A high-density genome scan detects evidence for a bipolar-disorder susceptibility locus on 13q32 and other potential loci on 1q32 and 18p11.2. *Proc Natl Acad Sci U S A*, **96**, 5604-5609.

Detera-Wadleigh, S.D., Badner, J.A., Goldin, L.R., Berrettini, W.H., Sanders, A.R., Rollins, D.Y., Turner, G., Moses, T., Haerian, H., Muniec, D., Nurnberger, J.I., Jr. and Gershon, E.S. (1996) Affected-sib-pair analyses reveal support of prior evidence for a susceptibility locus for bipolar disorder, on 21q. *Am J Hum Genet*, **58**, 1279-1285.

Detera-Wadleigh, S.D., Badner, J.A., Yoshikawa, T., Sanders, A.R., Goldin, L.R., Turner, G., Rollins, D.Y., Moses, T., Guroff, J.J., Kazuba, D., Maxwell, M.E., Edenberg, H.J., Foroud, T., Lahiri, D., Nurnberger, J.I., Jr., Stine, O.C., McMahon, F., Meyers, D.A., MacKinnon, D., Simpson, S., McInnis, M., DePaulo, J.R., Rice, J., Goate, A., Gershon, E.S. and et al. (1997) Initial genome scan of the NIMH genetics initiative bipolar pedigrees: chromosomes 4, 7, 9, 18, 19, 20, and 21q. *Am J Med Genet*, **74**, 254-262.

Dimitrova, A., Georgieva, L., Nikolov, I., Poriazova, N., Krastev, S., Toncheva, D., Owen, M.J. and Kirov, G. (2002) Major psychiatric disorders and the serotonin transporter gene (SLC6A4): family-based association studies. *Psychiatr Genet*, **12**, 137-141.

Ditkoff, E.C., Crary, W.G., Cristo, M. and Lobo, R.A. (1991) Estrogen improves psychological function in asymptomatic postmenopausal women. *Obstet Gynecol*, **78**, 991-995.

Edenberg, H.J., Foroud, T., Conneally, P.M., Sorbel, J.J., Carr, K., Crose, C., Willig, C., Zhao, J., Miller, M., Bowman, E., Mayeda, A., Rau, N.L., Smiley, C., Rice, J.P., Goate, A., Reich, T., Stine, O.C., McMahon, F., DePaulo, J.R., Meyers, D., Detera-Wadleigh, S.D., Goldin, L.R., Gershon, E.S., Blehar, M.C. and Nurnberger, J.I., Jr. (1997) Initial genomic scan of the NIMH genetics initiative bipolar pedigrees: chromosomes 3, 5, 15, 16, 17, and 22. *Am J Med Genet*, **74**, 238-246.

Edwards, A., Gibbs, R.A., Nguyen, P.N., Ansorge, W. and Caskey, C.T. (1989) Automated DNA sequencing methods for detection and analysis of mutations: applications to the Lesch-Nyhan syndrome. *Trans Assoc Am Physicians*, **102**, 185-194.

Egeland, J.A., Gerhard, D.S., Pauls, D.L., Sussex, J.N., Kidd, K.K., Allen, C.R., Hostetter, A.M. and Housman, D.E. (1987) Bipolar affective disorders linked to DNA markers on chromosome 11. *Nature*, **325**, 783-787.

Ekholm, J.M., Pekkarinen, P., Pajukanta, P., Kieseppa, T., Partonen, T., Paunio, T., Varilo, T., Perola, M., Lonnqvist, J. and Peltonen, L. (2002) Bipolar disorder susceptibility region on Xq24-q27.1 in Finnish families. *Mol Psychiatry*, **7**, 453-459.

Elvidge, G., Jones, I., McCandless, F., Asherson, P., Owen, M.J. and Craddock, N. (2001) Allelic variation of a BalI polymorphism in the DRD3 gene does not influence susceptibility to bipolar disorder: results of analysis and meta-analysis. *Am J Med Genet*, **105**, 307-311.

Enmark, E., Pelto-Huikko, M., Grandien, K., Lagercrantz, S., Lagercrantz, J., Fried, G., Nordenskjold, M. and Gustafsson, J.A. (1997) Human estrogen receptor beta-gene structure, chromosomal localization, and expression pattern. *J Clin Endocrinol Metab*, **82**, 4258-4265.

Escamilla, M.A. (2001) Population isolates: their special value for locating genes for bipolar disorder. *Bipolar Disord*, **3**, 299-317.

Escamilla, M.A., Spesny, M., Reus, V.I., Gallegos, A., Meza, L., Molina, J., Sandkuijl, L.A., Fournier, E., Leon, P.E., Smith, L.B. and Freimer, N.B. (1996) Use of linkage disequilibrium approaches to map genes for bipolar disorder in the Costa Rican population. *Am J Med Genet*, **67**, 244-253.

Esterling, L.E., Yoshikawa, T., Turner, G., Badner, J.A., Bengel, D., Gershon, E.S., Berrettini, W.H. and Detera-Wadleigh, S.D. (1998) Serotonin transporter (5-HTT) gene and bipolar affective disorder. *Am J Med Genet*, **81**, 37-40.

Evans, K.L., Le Hellard, S., Morris, S.W., Lawson, D., Whitton, C., Semple, C.A., Fantes, J.A., Torrance, H.S., Malloy, M.P., Maule, J.C., Humphray, S.J., Ross, M.T., Bentley, D.R., Muir, W.J., Blackwood, D.H. and Porteous, D.J. (2001) A 6.9-Mb high-resolution BAC/PAC contig of human 4p15.3-p16.1, a candidate region for bipolar affective disorder. *Genomics*, **71**, 315-323.

Evans, K.L., Muir, W.J., Blackwood, D.H. and Porteous, D.J. (2001) Nuts and bolts of psychiatric genetics: building on the Human Genome Project. *Trends Genet*, **17**, 35-40.

Ewald, H., Degn, B., Mors, O. and Kruse, T.A. (1998) Significant linkage between bipolar affective disorder and chromosome 12q24. *Psychiatr Genet*, **8**, 131-140.

Ewald, H., Degn, B., Mors, O. and Kruse, T.A. (1998) Support for the possible locus on chromosome 4p16 for bipolar affective disorder. *Mol Psychiatry*, **3**, 442-448.

Ewald, H., Eiberg, H., Mors, O., Flint, T. and Kruse, T.A. (1996) Linkage study between manic-depressive illness and chromosome 21. *Am J Med Genet*, **67**, 218-224.

Ewald, H., Flint, T., Kruse, T.A. and Mors, O. (2002) A genome-wide scan shows significant linkage between bipolar disorder and chromosome 12q24.3 and suggestive linkage to chromosomes 1p22-21, 4p16, 6q14-22, 10q26 and 16p13.3. *Mol Psychiatry*, **7**, 734-744.

Ewald, H., Mors, O., Flint, T., Friedrich, U., Eiberg, H. and Kruse, T.A. (1995) Linkage analysis between manic-depressive illness and markers on the long arm of chromosome 11. *Am J Med Genet*, **60**, 386-392.

Ewald, H., Mors, O., Koed, K., Eiberg, H. and Kruse, T.A. (1997) Susceptibility loci for bipolar affective disorder on chromosome 18? A review and a study of Danish families. *Psychiatr Genet*, **7**, 1-12.

Falk, C.T. and Rubinstein, P. (1987) Haplotype relative risks: an easy reliable way to construct a proper control sample for risk calculations. *Ann Hum Genet*, **51 (Pt 3)**, 227-233.

Faraone, S.V., Matise, T., Svrakic, D., Pepple, J., Malaspina, D., Suarez, B., Hampe, C., Zambuto, C.T., Schmitt, K., Meyer, J., Markel, P., Lee, H., Harkavy Friedman, J., Kaufmann, C., Cloninger, C.R. and Tsuang, M.T. (1998) Genome scan of European-American schizophrenia pedigrees: results of the NIMH Genetics Initiative and Millennium Consortium. *Am J Med Genet*, **81**, 290-295.

Farrar, G.J., Kenna, P., Jordan, S.A., Kumar-Singh, R., Humphries, M.M., Sharp, E.M., Sheils, D.M. and Humphries, P. (1991) A three-base-pair deletion in the peripherin-RDS gene in one form of retinitis pigmentosa. *Nature*, **354**, 478-480.

Field, L.L., Tobias, R. and Magnus, T. (1994) A locus on chromosome 15q26 (IDDM3) produces susceptibility to insulin-dependent diabetes mellitus. *Nat Genet*, **8**, 189-194.

Fink, G., Sumner, B., Rosie, R., Wilson, H. and McQueen, J. (1999) Androgen actions on central serotonin neurotransmission: relevance for mood, mental state and memory. *Behav Brain Res*, **105**, 53-68.

Fink, G., Sumner, B.E., McQueen, J.K., Wilson, H. and Rosie, R. (1998) Sex steroid control of mood, mental state and memory. *Clin Exp Pharmacol Physiol*, **25**, 764-775.

Fixman, M. and Freire, J.J. (1977) Theory of DNA melting curves. *Biopolymers*, **16**, 2693-2704.

Foroud, T., Castelluccio, P.F., Koller, D.L., Edenberg, H.J., Miller, M., Bowman, E., Rau, N.L., Smiley, C., Rice, J.P., Goate, A., Armstrong, C., Bierut, L.J., Reich, T., Detera-Wadleigh, S.D., Goldin, L.R., Badner, J.A., Guroff, J.J., Gershon, E.S., McMahon, F.J., Simpson, S., MacKinnon, D., McInnis, M., Stine, O.C., DePaulo, J.R., Blehar, M.C. and Nurnberger, J.I., Jr. (2000) Suggestive evidence of a locus on chromosome 10p using the NIMH genetics initiative bipolar affective disorder pedigrees. *Am J Med Genet*, **96**, 18-23.

Freimer, N.B., Reus, V.I., Escamilla, M., Spesny, M., Smith, L., Service, S., Gallegos, A., Meza, L., Batki, S., Vinogradov, S., Leon, P. and Sandkuijl, L.A. (1996) An approach to investigating linkage for bipolar disorder using large Costa Rican pedigrees. *Am J Med Genet*, **67**, 254-263.

Freimer, N.B., Reus, V.I., Escamilla, M.A., McInnes, L.A., Spesny, M., Leon, P., Service, S.K., Smith, L.B., Silva, S., Rojas, E., Gallegos, A., Meza, L., Fournier, E., Baharloo, S., Blankenship, K., Tyler, D.J., Batki, S., Vinogradov, S., Weissenbach, J., Barondes, S.H. and Sandkuijl, L.A. (1996) Genetic mapping using haplotype, association and linkage methods suggests a locus for severe bipolar disorder (BPI) at 18q22-q23. *Nat Genet*, **12**, 436-441.

Friddle, C., Koskela, R., Ranade, K., Hebert, J., Cargill, M., Clark, C.D., McInnis, M., Simpson, S., McMahon, F., Stine, O.C., Meyers, D., Xu, J., MacKinnon, D., Swift-Scanlan, T., Jamison, K., Folstein, S., Daly, M., Kruglyak, L., Marr, T., DePaulo, J.R. and Botstein, D. (2000) Full-genome scan for linkage in 50 families segregating the bipolar affective disease phenotype. *Am J Hum Genet*, **66**, 205-215.

Furlong, R.A., Coleman, T.A., Ho, L., Rubinsztein, J.S., Walsh, C., Paykel, E.S. and Rubinsztein, D.C. (1998) No association of a functional polymorphism in the dopamine D2 receptor promoter region with bipolar or unipolar affective disorders. *Am J Med Genet*, **81**, 385-387.

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

Furlong, R.A., Ho, L., Walsh, C., Rubinsztein, J.S., Jain, S., Paykel, E.S., Easton, D.F. and Rubinsztein, D.C. (1998) Analysis and meta-analysis of two serotonin transporter gene polymorphisms in bipolar and unipolar affective disorders. *Am J Med Genet*, **81**, 58-63.

Furlong, R.A., Rubinsztein, J.S., Ho, L., Walsh, C., Coleman, T.A., Muir, W.J., Paykel, E.S., Blackwood, D.H. and Rubinsztein, D.C. (1999) Analysis and metaanalysis of two polymorphisms within the tyrosine hydroxylase gene in bipolar and unipolar affective disorders. *Am J Med Genet*, **88**, 88-94.

Garcia-Segura, L.M., Azcoitia, I. and DonCarlos, L.L. (2001) Neuroprotection by estradiol. *Prog Neurobiol*, **63**, 29-60.

Garner, C., McInnes, L.A., Service, S.K., Spesny, M., Fournier, E., Leon, P. and Freimer, N.B. (2001) Linkage analysis of a complex pedigree with severe bipolar disorder, using a Markov chain Monte Carlo method. *Am J Hum Genet*, **68**, 1061-1064.

Garver, D., Barnes, R., Holcomb, J., Filbey, F., Wilson, R. and Bowcock, A. (1998) Genome-wide scan and schizophrenia in African Americans. *Am J Hum Genet*, **6**, 454-455.

Georgieva, L., Dimitrova, A., Nikolov, I., Koleva, S., Tsvetkova, R., Owen, M.J., Toncheva, D. and Kirov, G. (2002) Dopamine transporter gene (DAT1) VNTR polymorphism in major psychiatric disorders: family-based association study in the Bulgarian population. *Acta Psychiatr Scand*, **105**, 396-399.

Gerner, R.H., Post, R.M. and Bunney, W.E., Jr. (1976) A dopaminergic mechanism in mania. *Am J Psychiatry*, **133**, 1177-1180.

Gershon, E.S., Badner, J.A., Detera-Wadleigh, S.D., Ferraro, T.N. and Berrettini, W.H. (1996) Maternal inheritance and chromosome 18 allele sharing in unilineal bipolar illness pedigrees. *Am J Med Genet*, **67**, 202-207.

Gershon, E.S., Badner, J.A., Goldin, L.R., Sanders, A.R., Cravchik, A. and Detera-Wadleigh, S.D. (1998) Closing in on genes for manic-depressive illness and schizophrenia. *Neuropsychopharmacology*, **18**, 233-242.

Gill, M., Vallada, H., Collier, D., Sham, P., Holmans, P., Murray, R., McGuffin, P., Nanko, S., Owen, M., Antonarakis, S., Housman, D., Kazazian, H., Nestadt, G., Pulver, A.E., Straub, R.E., MacLean, C.J., Walsh, D., Kendler, K.S., DeLisi, L., Polymeropoulos, M., Coon, H., Byerley, W., Lofthouse, R., Gershon, E., Read, C.M. and et al. (1996) A combined analysis of D22S278 marker alleles in affected sib-pairs: support for a susceptibility locus for schizophrenia at chromosome 22q12. Schizophrenia Collaborative Linkage Group (Chromosome 22). *Am J Med Genet*, **67**, 40-45.

Ginns, E.I., Ott, J., Egeland, J.A., Allen, C.R., Fann, C.S., Pauls, D.L., Weissenbachoff, J., Carulli, J.P., Falls, K.M., Keith, T.P. and Paul, S.M. (1996) A genome-wide search for chromosomal loci linked to bipolar affective disorder in the Old Order Amish. *Nat Genet*, **12**, 431-435.

Ginns, E.I., St Jean, P., Philibert, R.A., Galdzicka, M., Damschroder-Williams, P., Thiel, B., Long, R.T., Ingraham, L.J., Dalwaldi, H., Murray, M.A., Ehlert, M., Paul, S., Remortel, B.G., Patel, A.P., Anderson, M.C., Shaio, C., Lau, E., Dymarskaia, I., Martin, B.M., Stubblefield, B., Falls, K.M., Carulli, J.P., Keith, T.P., Fann, C.S., Paul, S.M. and et al. (1998) A genome-wide search for chromosomal loci linked to mental health wellness in relatives at high risk for bipolar affective disorder among the Old Order Amish. *Proc Natl Acad Sci U S A*, **95**, 15531-15536.

Goddard, K.A., Hopkins, P.J., Hall, J.M. and Witte, J.S. (2000) Linkage disequilibrium and allele-frequency distributions for 114 single-nucleotide polymorphisms in five populations. *Am J Hum Genet*, **66**, 216-234.

Gogos, J.A., Morgan, M., Luine, V., Santha, M., Ogawa, S., Pfaff, D. and Karayiorgou, M. (1998) Catechol-O-methyltransferase-deficient mice exhibit sexually dimorphic changes in catecholamine levels and behavior. *Proc Natl Acad Sci U S A*, **95**, 9991-9996.

Goldin, L.R., Chase, G.A., King, T.M., Badner, J.A. and Gershon, E.S. (1995) Use of exact and adjusted liability scores to detect genes affecting common traits. *Genet Epidemiol*, **12**, 765-769.

Goodman, A.B. (1994) Elevated risks for amyotrophic lateral sclerosis and blood disorders in Ashkenazi schizophrenic pedigrees suggest new candidate genes in schizophrenia. *Am J Med Genet*, **54**, 271-278.

Goodwin, S. and Guze, A.D. (1995) *Psychiatric Diagnosis, 5th edition*. Oxford University Press.

Green, S., Walter, P., Kumar, V., Krust, A., Bornert, J.M., Argos, P. and Chambon, P. (1986) Human oestrogen receptor cDNA: sequence, expression and homology to v-erb-A. *Nature*, **320**, 134-139.

Greenberg, D.A., Delgado-Escueta, A.V., Widelitz, H., Sparkes, R.S., Treiman, L., Maldonado, H.M., Park, M.S. and Terasaki, P.I. (1988) Juvenile myoclonic epilepsy (JME) may be linked to the BF and HLA loci on human chromosome 6. *Am J Med Genet*, **31**, 185-192.

Greenwood, T.A., Alexander, M., Keck, P.E., McElroy, S., Sadovnick, A.D., Remick, R.A., Shaw, S.H. and Kelsoe, J.R. (2002) Segmental linkage disequilibrium within the dopamine transporter gene. *Mol Psychiatry*, **7**, 165-173.

Gregoire, A.J., Kumar, R., Everitt, B., Henderson, A.F. and Studd, J.W. (1996) Transdermal oestrogen for treatment of severe postnatal depression. *Lancet*, **347**, 930-933.

Grigoroiu-Serbanescu, M., Martinez, M., Nothen, M., Grinberg, M., Sima, D., Propping, P., Marinescu, E. and Hrestic, M. (2001) Different familial transmission patterns in bipolar I disorder with onset before and after 25. *Am J Med Genet*, **105**, 765-773.

Grigoroiu-Serbanescu, M., Nothen, M., Propping, P., Poustka, F., Magureanu, S., Vasilescu, R., Marinescu, E. and Ardelean, V. (1995) Clinical evidence for genomic imprinting in bipolar I disorder. *Acta Psychiatr Scand*, **92**, 365-370.

Gurling, H.M., Kalsi, G., Brynjolfson, J., Sigmundsson, T., Sherrington, R., Mankoo, B.S., Read, T., Murphy, P., Blaveri, E., McQuillin, A., Petursson, H. and Curtis, D. (2001) Genomewide genetic linkage analysis confirms the presence of susceptibility loci for schizophrenia, on chromosomes 1q32.2, 5q33.2, and 8p21-22 and provides support for linkage to schizophrenia, on chromosomes 11q23.3-24 and 20q12.1-11.23. *Am J Hum Genet*, **68**, 661-673.

Gusella, J.F., Wexler, N.S., Conneally, P.M., Naylor, S.L., Anderson, M.A., Tanzi, R.E., Watkins, P.C., Ottina, K., Wallace, M.R., Sakaguchi, A.Y. and et al. (1983) A polymorphic DNA marker genetically linked to Huntington's disease. *Nature*, **306**, 234-238.

Gutierrez, B., Arias, B., Papiol, S., Rosa, A. and Fananas, L. (2001) Association study between novel promoter variants at the 5-HT2C receptor gene and human patients with bipolar affective disorder. *Neurosci Lett*, **309**, 135-137.

Gutierrez, B., Arranz, M.J., Collier, D.A., Valles, V., Guillamat, R., Bertranpetit, J., Murray, R.M. and Fanas, L. (1998) Serotonin transporter gene and risk for bipolar affective disorder: an association study in Spanish population. *Biol Psychiatry*, **43**, 843-847.

Gutierrez, B., Bertranpetit, J., Guillamat, R., Valles, V., Arranz, M.J., Kerwin, R. and Fananas, L. (1997) Association analysis of the catechol O-methyltransferase gene and bipolar affective disorder. *Am J Psychiatry*, **154**, 113-115.

Hahn, H., Trant, M.R., Brownstein, M.J., Harper, R.A., Milstien, S. and Butler, I.J. (2001) Neurologic and psychiatric manifestations in a family with a mutation in exon 2 of the guanosine triphosphate-cyclohydrolase gene. *Arch Neurol*, **58**, 749-755.

Halbreich, U. and Kahn, L.S. (2001) Role of estrogen in the aetiology and treatment of mood disorders. *CNS Drugs*, **15**, 797-817.

Hantouche, E.G., Akiskal, H.S., Lancrenon, S., Allilaire, J.F., Sechter, D., Azorin, J.M., Bourgeois, M., Fraud, J.P. and Chatenet-Duchene, L. (1998) Systematic clinical methodology for validating bipolar-II disorder: data in mid-stream from a French national multi-site study (EPIDEP). *J Affect Disord*, **50**, 163-173.

Harley, H.G., Walsh, K.V., Rundle, S., Brook, J.D., Sarfarazi, M., Koch, M.C., Floyd, J.L., Harper, P.S. and Shaw, D.J. (1991) Localisation of the myotonic dystrophy locus to 19q13.2-19q13.3 and its relationship to twelve polymorphic loci on 19q. *Hum Genet*, **87**, 73-80.

Hauser, E.R., Boehnke, M., Guo, S.W. and Risch, N. (1996) Affected-sib-pair interval mapping and exclusion for complex genetic traits: sampling considerations. *Genet Epidemiol*, **13**, 117-137.

Ho, L.W., Furlong, R.A., Rubinsztein, J.S., Walsh, C., Paykel, E.S. and Rubinsztein, D.C. (2000) Genetic associations with clinical characteristics in bipolar affective disorder and recurrent unipolar depressive disorder. *Am J Med Genet*, **96**, 36-42.

Hoda, F., Nicholl, D., Bennett, P., Arranz, M., Aitchison, K.J., al-Chalabi, A., Kunugi, H., Vallada, H., Leigh, P.N., Chaudhuri, K.R. and Collier, D.A. (1996) No association between Parkinson's disease and low-activity alleles of catechol O-methyltransferase. *Biochem Biophys Res Commun*, **228**, 780-784.

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

Hong, C.J., Tsai, S.J., Cheng, C.Y., Liao, W.Y., Song, H.L. and Lai, H.C. (1999) Association analysis of the 5-HT(6) receptor polymorphism (C267T) in mood disorders. *Am J Med Genet*, **88**, 601-602.

Horvath, S. and Laird, N.M. (1998) A discordant-sibship test for disequilibrium and linkage: no need for parental data. *Am J Hum Genet*, **63**, 1886-1897.

Ichinose, H., Ohye, T., Matsuda, Y., Hori, T., Blau, N., Burlina, A., Rouse, B., Matalon, R., Fujita, K. and Nagatsu, T. (1995) Characterization of mouse and human GTP cyclohydrolase I genes. Mutations in patients with GTP cyclohydrolase I deficiency. *J Biol Chem*, **270**, 10062-10071.

Ichinose, H., Ohye, T., Takahashi, E., Seki, N., Hori, T., Segawa, M., Nomura, Y., Endo, K., Tanaka, H., Tsuji, S. and et al. (1994) Hereditary progressive dystonia with marked diurnal fluctuation caused by mutations in the GTP cyclohydrolase I gene. *Nat Genet*, **8**, 236-242.

Inglehearn, C.F. (1997) Intelligent linkage analysis using gene density estimates. *Nat Genet*, **16**, 15.

Jamison, K.R. (1995) Manic-depressive illness and creativity. *Sci Am*, **272**, 62-67.

Johnson, G.C., Esposito, L., Barratt, B.J., Smith, A.N., Heward, J., Di Genova, G., Ueda, H., Cordell, H.J., Eaves, I.A., Dudbridge, F., Twells, R.C., Payne, F., Hughes, W., Nutland, S., Stevens, H., Carr, P., Tuomilehto-Wolf, E., Tuomilehto, J., Gough, S.C., Clayton, D.G. and Todd, J.A. (2001) Haplotype tagging for the identification of common disease genes. *Nat Genet*, **29**, 233-237.

Jones, A.C., Austin, J., Hansen, N., Hoogendoorn, B., Oefner, P.J., Cheadle, J.P. and O'Donovan, M.C. (1999) Optimal temperature selection for mutation detection by denaturing

HPLC and comparison to single-stranded conformation polymorphism and heteroduplex analysis. *Clin Chem*, **45**, 1133-1140.

Kaiser, R., Muller-Oerlinghausen, B., Filler, D., Tremblay, P.B., Berghofer, A., Roots, I. and Brockmoller, J. (2002) Correlation between serotonin uptake in human blood platelets with the 44-bp polymorphism and the 17-bp variable number of tandem repeat of the serotonin transporter. *Am J Med Genet*, **114**, 323-328.

Kalsi, G., Smyth, C., Brynjolfsson, J., Sherrington, R.S., O'Neill, J., Curtis, D., Rifkin, L., Murphy, P., Petursson, H. and Gurling, H.M. (1997) Linkage analysis of manic depression (bipolar affective disorder) in Icelandic and British kindreds using markers on the short arm of chromosome 18. *Hum Hered*, **47**, 268-278.

Karatinos, J., Rosse, R.B. and Deutsch, S.I. (1995) The nitric oxide pathway: potential implications for treatment of neuropsychiatric disorders. *Clin Neuropharmacol*, **18**, 482-499.

Karayiorgou, M., Altemus, M., Galke, B.L., Goldman, D., Murphy, D.L., Ott, J. and Gogos, J.A. (1997) Genotype determining low catechol-O-methyltransferase activity as a risk factor for obsessive-compulsive disorder. *Proc Natl Acad Sci U S A*, **94**, 4572-4575.

Karayiorgou, M., Gogos, J.A., Galke, B.L., Wolyniec, P.S., Nestadt, G., Antonarakis, S.E., Kazazian, H.H., Housman, D.E. and Pulver, A.E. (1998) Identification of sequence variants and analysis of the role of the catechol-O-methyl-transferase gene in schizophrenia susceptibility. *Biol Psychiatry*, **43**, 425-431.

Karayiorgou, M., Kasch, L., Lasseter, V.K., Hwang, J., Elango, R., Bernardini, D.J., Kimberland, M., Babb, R., Francomano, C.A., Wolyniec, P.S. and et al. (1994) Report from the Maryland Epidemiology Schizophrenia Linkage Study: no evidence for linkage between schizophrenia and a number of candidate and other genomic regions using a complex dominant model. *Am J Med Genet*, **54**, 345-353.

Kato, M.V., Shimizu, T., Nagayoshi, M., Kaneko, A., Sasaki, M.S. and Ikawa, Y. (1996) Genomic imprinting of the human serotonin-receptor (HTR2) gene involved in development of retinoblastoma. *Am J Hum Genet*, **59**, 1084-1090.

Kato, T., Winokur, G., Coryell, W., Keller, M.B., Endicott, J. and Rice, J. (1996) Parent-of-origin effect in transmission of bipolar disorder. *Am J Med Genet*, **67**, 546-550.

Kaufman, S. (1995) Tyrosine hydroxylase. *Adv Enzymol Relat Areas Mol Biol*, **70**, 103-220.

Kelsoe, J.R., Ginns, E.I., Egeland, J.A., Gerhard, D.S., Goldstein, A.M., Bale, S.J., Pauls, D.L., Long, R.T., Kidd, K.K., Conte, G. and et al. (1989) Re-evaluation of the linkage relationship between chromosome 11p loci and the gene for bipolar affective disorder in the Old Order Amish. *Nature*, **342**, 238-243.

Kelsoe, J.R., Sadovnick, A.D., Kristbjarnarson, H., Bergesch, P., Mroczkowski-Parker, Z., Drennan, M., Rapaport, M.H., FLODman, P., Spence, M.A. and Remick, R.A. (1996) Possible locus for bipolar disorder near the dopamine transporter on chromosome 5. *Am J Med Genet*, **67**, 533-540.

Kelsoe, J.R., Spence, M.A., Loetscher, E., Foguet, M., Sadovnick, A.D., Remick, R.A., FLODman, P., Khristich, J., Mroczkowski-Parker, Z., Brown, J.L., Masser, D., Ungerleider, S., Rapaport, M.H., Wishart, W.L. and Luebbert, H. (2001) A genome survey indicates a possible

susceptibility locus for bipolar disorder on chromosome 22. *Proc Natl Acad Sci U S A*, **98**, 585-590.

Kendler, K.S. (1993) Twin studies of psychiatric illness. Current status and future directions. *Arch Gen Psychiatry*, **50**, 905-915.

Khoury, M.J. (1994) Case-parental control method in the search for disease-susceptibility genes. *Am J Hum Genet*, **55**, 414-415.

King, N., Bassett, A.S., Honer, W.G., Barsellis, M. and Kennedy, J.L. (1997) Absence of linkage for schizophrenia on the short arm of chromosome 5 in multiplex Canadian families. *Am J Hum Genet*, **74**, 472-474.

Kirov, G., Jones, I., McCandless, F., Craddock, N. and Owen, M.J. (1999) Family-based association studies of bipolar disorder with candidate genes involved in dopamine neurotransmission: DBH, DAT1, COMT, DRD2, DRD3 and DRD5. *Mol Psychiatry*, **4**, 558-565.

Kirov, G., Murphy, K.C., Arranz, M.J., Jones, I., McCandles, F., Kunugi, H., Murray, R.M., McGuffin, P., Collier, D.A., Owen, M.J. and Craddock, N. (1998) Low activity allele of catechol-O-methyltransferase gene associated with rapid cycling bipolar disorder. *Mol Psychiatry*, **3**, 342-345.

Kirov, G., Norton, N., Jones, I., McCandless, F., Craddock, N. and Owen, M.J. (1999) A functional polymorphism in the promoter of monoamine oxidase A gene and bipolar affective disorder. *Int J Neuropsychopharmcol*, **2**, 293-298.

Kirov, G., Rees, M., Jones, I., MacCandless, F., Owen, M.J. and Craddock, N. (1999) Bipolar disorder and the serotonin transporter gene: a family-based association study. *Psychol Med*, **29**, 1249-1254.

Knapp, M. (1999) A note on power approximations for the transmission/disequilibrium test. *Am J Hum Genet*, **64**, 1177-1185.

Knowles, J.A., Rao, P.A., Cox-Matise, T., Loth, J.E., de Jesus, G.M., Levine, L., Das, K., Penchaszadeh, G.K., Alexander, J.R., Lerer, B., Endicott, J., Ott, J., Gilliam, T.C. and Baron, M. (1998) No evidence for significant linkage between bipolar affective disorder and chromosome 18 pericentromeric markers in a large series of multiplex extended pedigrees. *Am J Hum Genet*, **62**, 916-924.

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

Kopin, I.J. (1994) Monoamine oxidase and catecholamine metabolism. *J Neural Transm Suppl*, **41**, 57-67.

Kornberg, J.R., Brown, J.L., Sadovnick, A.D., Remick, R.A., Keck, P.E., Jr., McElroy, S.L., Rapaport, M.H., Thompson, P.M., Kaul, J.B., Vrabel, C.M., Schommer, S.C., Wilson, T., Pizzuco, D., Jameson, S., Schibuk, L. and Kelsoe, J.R. (2000) Evaluating the parent-of-origin effect in bipolar affective disorder. Is a more penetrant subtype transmitted paternally? *J Affect Disord*, **59**, 183-192.

Kringlen, E. (1967) *Heredity and Environment in the Functional Psychoses*. London: Heinemann.

Kruglyak, L. (1997) What is significant in whole-genome linkage disequilibrium studies? *Am J Hum Genet*, **61**, 810-812.

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

Kruglyak, L., Daly, M.J., Reeve-Daly, M.P. and Lander, E.S. (1996) Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet*, **58**, 1347-1363.

Kuklin, A., Munson, K., Gjerde, D., Haefele, R. and Taylor, P. (1997) Detection of single-nucleotide polymorphisms with the WAVE DNA fragment analysis system. *Genet Test*, **1**, 201-206.

Kunugi, H., Hattori, M., Kato, T., Tatsumi, M., Sakai, T., Sasaki, T., Hirose, T. and Nanko, S. (1997) Serotonin transporter gene polymorphisms: ethnic difference and possible association with bipolar affective disorder. *Mol Psychiatry*, **2**, 457-462.

Kunugi, H., Vallada, H.P., Hoda, F., Kirov, G., Gill, M., Aitchison, K.J., Ball, D., Arranz, M.J., Murray, R.M. and Collier, D.A. (1997) No evidence for an association of affective disorders with high- or low-activity allele of catechol-o-methyltransferase gene. *Biol Psychiatry*, **42**, 282-285.

Kwok, J.B., Adams, L.J., Salmon, J.A., Donald, J.A., Mitchell, P.B. and Schofield, P.R. (1999) Nonparametric simulation-based statistical analyses for bipolar affective disorder locus on chromosome 21q22.3. *Am J Med Genet*, **88**, 99-102.

Lachman, H.M., Kelsoe, J.R., Remick, R.A., Sadovnick, A.D., Rapaport, M.H., Lin, M., Pazur, B.A., Roe, A.M., Saito, T. and Papolos, D.F. (1997) Linkage studies suggest a possible locus for bipolar disorder near the velo-cardio-facial syndrome region on chromosome 22. *Am J Med Genet*, **74**, 121-128.

Lachman, H.M., Morrow, B., Shprintzen, R., Veit, S., Parsia, S.S., Faedda, G., Goldberg, R., Kucherlapati, R. and Papolos, D.F. (1996) Association of codon 108/158 catechol-O-methyltransferase gene polymorphism with the psychiatric manifestations of velo-cardio-facial syndrome. *Am J Med Genet*, **67**, 468-472.

Lachman, H.M., Papolos, D.F., Saito, T., Yu, Y.M., Szumlanski, C.L. and Weinshilboum, R.M. (1996) Human catechol-O-methyltransferase pharmacogenetics: description of a functional polymorphism and its potential application to neuropsychiatric disorders. *Pharmacogenetics*, **6**, 243-250.

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

Lander, E.S. and Botstein, D. (1986) Mapping complex genetic traits in humans: new methods using a complete RFLP linkage map. *Cold Spring Harb Symp Quant Biol*, **51 Pt 1**, 49-62.

Langlois, R.G., Yu, L.C., Gray, J.W. and Carrano, A.V. (1982) Quantitative karyotyping of human chromosomes by dual beam flow cytometry. *Proc Natl Acad Sci U S A*, **79**, 7876-7880.

Lauder, J.M. (1993) Neurotransmitters as growth regulatory signals: role of receptors and second messengers. *Trends Neurosci*, **16**, 233-240.

Leboyer, M., Bellivier, F., Nosten-Bertrand, M., Jouvent, R., Pauls, D. and Mallet, J. (1998) Psychiatric genetics: search for phenotypes. *Trends Neurosci*, **21**, 102-105.

Leboyer, M., Malafosse, A., Boularand, S., Campion, D., Gheysen, F., Samolyk, D., Henriksson, B., Denise, E., des Lauriers, A. and Lepine, J.P. (1990) Tyrosine hydroxylase polymorphisms associated with manic-depressive illness. *Lancet*, **335**, 1219.

Leboyer, M., Quintin, P., Manivet, P., Varoquaux, O., Allilaire, J.F. and Launay, J.M. (1999) Decreased serotonin transporter binding in unaffected relatives of manic depressive patients. *Biol Psychiatry*, **46**, 1703-1706.

Lenox, R.H., Gould, T.D. and Manji, H.K. (2002) Endophenotypes in bipolar disorder. *Am J Med Genet*, **114**, 391-406.

Lerer, B., Macciardi, F., Segman, R.H., Adolfsson, R., Blackwood, D., Blairy, S., Del Favero, J., Dikeos, D.G., Kaneva, R., Lilli, R., Massat, I., Milanova, V., Muir, W., Noethen, M., Oruc, L., Petrova, T., Papadimitriou, G.N., Rietschel, M., Serretti, A., Souery, D., Van Gestel, S., Van Broeckhoven, C. and Mendlewicz, J. (2001) Variability of 5-HT2C receptor cys23ser polymorphism among European populations and vulnerability to affective disorder. *Mol Psychiatry*, **6**, 579-585.

Levinson, D.F. and Coon, H. (1998) Chromosome 22 workshop. *Psychiatr Genet*, **8**, 115-120.

Levinson, D.F., Mahtani, M.M., Nancarrow, D.J., Brown, D.M., Kruglyak, L., Kirby, A., Hayward, N.K., Crowe, R.R., Andreasen, N.C., Black, D.W., Silverman, J.M., Endicott, J., Sharpe, L., Mohs, R.C., Siever, L.J., Walters, M.K., Lennon, D.P., Jones, H.L., Nertney, D.A., Daly, M.J., Gladis, M. and Mowry, B.J. (1998) Genome scan of schizophrenia. *Am J Psychiatry*, **155**, 741-750.

Li, T., Liu, X., Sham, P.C., Aitchison, K.J., Cai, G., Arranz, M.J., Deng, H., Liu, J., Kirov, G., Murray, R.M. and Collier, D.A. (1999) Association analysis between dopamine receptor genes and bipolar affective disorder. *Psychiatry Res*, **86**, 193-201.

Li, T., Vallada, H., Curtis, D., Arranz, M., Xu, K., Cai, G., Deng, H., Liu, J., Murray, R., Liu, X. and Collier, D.A. (1997) Catechol-O-methyltransferase Val158Met polymorphism: frequency analysis in Han Chinese subjects and allelic association of the low activity allele with bipolar affective disorder. *Pharmacogenetics*, **7**, 349-353.

Lim, L.C., Powell, J., Sham, P., Castle, D., Hunt, N., Murray, R. and Gill, M. (1995) Evidence for a genetic association between alleles of monoamine oxidase A gene and bipolar affective disorder. *Am J Med Genet*, **60**, 325-331.

Lin, M.W., Curtis, D., Williams, N., Arranz, M., Nanko, S., Collier, D., McGuffin, P., Murray, R., Owen, M., Gill, M. and et al. (1995) Suggestive evidence for linkage of schizophrenia to markers on chromosome 13q14.1-q32. *Psychiatr Genet*, **5**, 117-126.

Liu, J., Juo, S.H., Terwilliger, J.D., Grunn, A., Tong, X., Brito, M., Loth, J.E., Kanyas, K., Lerer, B., Endicott, J., Penchaszadeh, G., Gilliam, T.C. and Baron, M. (2001) A follow-up linkage study supports evidence for a bipolar affective disorder locus on chromosome 21q22. *Am J Med Genet*, **105**, 189-194.

Liu, W., Gu, N., Feng, G., Li, S., Bai, S., Zhang, J., Shen, T., Xue, H., Breen, G., St Clair, D. and He, L. (1999) Tentative association of the serotonin transporter with schizophrenia and unipolar depression but not with bipolar disorder in Han Chinese. *Pharmacogenetics*, **9**, 491-495.

Livak, K.J., Flood, S.J., Marmaro, J., Giusti, W. and Deetz, K. (1995) Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. *PCR Methods Appl*, **4**, 357-362.

Long, A.D. and Langley, C.H. (1999) The power of association studies to detect the contribution of candidate genetic loci to variation in complex traits. *Genome Res*, **9**, 720-731.

Lotta, T., Vidgren, J., Tilgmann, C., Ulmanen, I., Melen, K., Julkunen, I. and Taskinen, J. (1995) Kinetics of human soluble and membrane-bound catechol O-methyltransferase: a revised mechanism and description of the thermolabile variant of the enzyme. *Biochemistry*, **34**, 4202-4210.

Lyamichev, V., Mast, A.L., Hall, J.G., Prudent, J.R., Kaiser, M.W., Takova, T., Kwiatkowski, R.W., Sander, T.J., de Arruda, M., Arco, D.A., Neri, B.P. and Brow, M.A. (1999) Polymorphism identification and quantitative detection of genomic DNA by invasive cleavage of oligonucleotide probes. *Nat Biotechnol*, **17**, 292-296.

MacKenzie, A. and Quinn, J. (1999) A serotonin transporter gene intron 2 polymorphic region, correlated with affective disorders, has allele-dependent differential enhancer-like properties in the mouse embryo. *Proc Natl Acad Sci U S A*, **96**, 15251-15255.

Mahmood, T. and Silverstone, T. (2001) Serotonin and bipolar disorder. *J Affect Disord*, **66**, 1-11.

Maier, W., Lichtermann, D., Minges, J., Hallmayer, J., Heun, R., Benkert, O. and Levinson, D.F. (1993) Continuity and discontinuity of affective disorders and schizophrenia. Results of a controlled family study. *Arch Gen Psychiatry*, **50**, 871-883.

Malafosse, A., Leboyer, M., d'Amato, T., Amadeo, S., Abbar, M., Campion, D., Canseil, O., Castelnau, D., Gheysen, F., Granger, B., Henrikson, B., Poirier, M.F., Sabate, O., Samolyk, D., Feingold, J. and Mallet, J. (1997) Manic depressive illness and tyrosine hydroxylase gene: linkage heterogeneity and association. *Neurobiol Dis*, **4**, 337-349.

Manki, H., Kanba, S., Muramatsu, T., Higuchi, S., Suzuki, E., Matsushita, S., Ono, Y., Chiba, H., Shintani, F., Nakamura, M., Yagi, G. and Asai, M. (1996) Dopamine D2, D3 and D4 receptor and transporter gene polymorphisms and mood disorders. *J Affect Disord*, **40**, 7-13.

Marneros, A. (2001) Expanding the group of bipolar disorders. *J Affect Disord*, **62**, 39-44.

Martin, E.R., Kaplan, N.L. and Weir, B.S. (1997) Tests for linkage and association in nuclear families. *Am J Hum Genet*, **61**, 439-448.

Martin, R.A., Sabol, D.W. and Rogan, P.K. (1999) Maternal uniparental disomy of chromosome 14 confined to an interstitial segment (14q23-14q24.2). *J Med Genet*, **36**, 633-636.

Martinez, M., Khlat, M., Leboyer, M. and Clerget-Darpoux, F. (1989) Performance of linkage analysis under misclassification error when the genetic model is unknown. *Genet Epidemiol*, **6**, 253-258.

Massat, I., Souery, D., Del-Favero, J., Van Gestel, S., Serretti, A., Macciardi, F., Smeraldi, E., Kaneva, R., Adolfsson, R., Nylander, P.O., Blackwood, D., Muir, W., Papadimitriou, G.N., Dikeos, D., Oruc, L., Segman, R.H., Ivezic, S., Aschauer, H., Ackenheil, M., Fuchshuber, S., Dam, H., Jakovljevic, M., Peltonen, L., Hilger, C., Hentges, F., Staner, L., Milanova, V., Jazin, E., Lerer, B., Van Broeckhoven, C. and Mendlewicz, J. (2002) Positive association of dopamine

D2 receptor polymorphism with bipolar affective disorder in a European Multicenter Association Study of affective disorders. *Am J Med Genet*, **114**, 177-185.

Massat, I., Souery, D., Lipp, O., Blairy, S., Papadimitriou, G., Dikeos, D., Ackenheil, M., Fuchshuber, S., Hilger, C., Kaneva, R., Milanova, V., Verheyen, G., Raeymaekers, P., Staner, L., Oruc, L., Jakovljevic, M., Serretti, A., Macciardi, F., Van Broeckhoven, C. and Mendlewicz, J. (2000) A European multicenter association study of HTR2A receptor polymorphism in bipolar affective disorder. *Am J Med Genet*, **96**, 136-140.

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

McEwen, B.S. (2001) Invited review: Estrogens effects on the brain: multiple sites and molecular mechanisms. *J Appl Physiol*, **91**, 2785-2801.

McEwen, B.S. and Alves, S.E. (1999) Estrogen actions in the central nervous system. *Endocr Rev*, **20**, 279-307.

McGuffin, P., Farmer, A. and Harvey, I. (1991) A polydiagnostic application of operational criteria in studies of psychotic illness. Development and reliability of the OPCRIT system. *Arch Gen Psychiatry*, **48**, 764-770.

McGuffin, P., Katz R. (1986) Nature, nurture and affective disorder. In *The biology of affective disorders*. London: Royal College of Psychiatrists, Gaskell Press, pp. 26-51.

McGuffin, P. and Katz, R. (1989) The genetics of depression and manic-depressive disorder. *Br J Psychiatry*, **155**, 294-304.

McInnes, L.A., Escamilla, M.A., Service, S.K., Reus, V.I., Leon, P., Silva, S., Rojas, E., Spesny, M., Baharloo, S., Blankenship, K., Peterson, A., Tyler, D., Shimayoshi, N., Tobey, C., Batki, S., Vinogradov, S., Meza, L., Gallegos, A., Fournier, E., Smith, L.B., Barondes, S.H., Sandkuijl, L.A. and Freimer, N.B. (1996) A complete genome screen for genes predisposing to severe bipolar disorder in two Costa Rican pedigrees. *Proc Natl Acad Sci U S A*, **93**, 13060-13065.

McInnes, L.A., Service, S.K., Reus, V.I., Barnes, G., Charlat, O., Jawahar, S., Lewitzky, S., Yang, Q., Duong, Q., Spesny, M., Araya, C., Araya, X., Gallegos, A., Meza, L., Molina, J., Ramirez, R., Mendez, R., Silva, S., Fournier, E., Batki, S.L., Mathews, C.A., Neylan, T., Glatt, C.E., Escamilla, M.A., Luo, D., Gajiwala, P., Song, T., Crook, S., Nguyen, J.B., Roche, E., Meyer, J.M., Leon, P., Sandkuijl, L.A., Freimer, N.B. and Chen, H. (2001) Fine-scale mapping of a locus for severe bipolar mood disorder on chromosome 18p11.3 in the Costa Rican population. *Proc Natl Acad Sci U S A*, **98**, 11485-11490.

McInnis, M., Mackinnon, D., McMahon, F., Foroud, T., Edenberg, H.J., Goate, A., Detera-Wadleigh, S., Stine, O., Rice, J., Blehar, M., Reich, T., Gershon, E., Nurnberger, J., Simpson, S. and JR, D.P. (1998) Evidence for a susceptibility locus for bipolar disorder on chromosome 11p11.5. *Am J Hum Genet*, **81**, 463.

McKusick, V.A. (1991) Current trends in mapping human genes. *Faseb J*, **5**, 12-20.

McMahon, F.J., Hopkins, P.J., Xu, J., McInnis, M.G., Shaw, S., Cardon, L., Simpson, S.G., MacKinnon, D.F., Stine, O.C., Sherrington, R., Meyers, D.A. and DePaulo, J.R. (1997) Linkage

of bipolar affective disorder to chromosome 18 markers in a new pedigree series. *Am J Hum Genet*, **61**, 1397-1404.

McMahon, F.J., Simpson, S.G., McInnis, M.G., Badner, J.A., MacKinnon, D.F. and DePaulo, J.R. (2001) Linkage of bipolar disorder to chromosome 18q and the validity of bipolar II disorder. *Arch Gen Psychiatry*, **58**, 1025-1031.

McMahon, F.J., Stine, O.C., Meyers, D.A., Simpson, S.G. and DePaulo, J.R. (1995) Patterns of maternal transmission in bipolar affective disorder. *Am J Hum Genet*, **56**, 1277-1286.

McQuillin, A., Lawrence, J., Curtis, D., Kalsi, G., Smyth, C., Hannesdottir, S. and Gurling, H. (1999) Adjacent genetic markers on chromosome 11p15.5 at or near the tyrosine hydroxylase locus that show population linkage disequilibrium with each other do not show allelic association with bipolar affective disorder. *Psychol Med*, **29**, 1449-1454.

Meloni, R., Leboyer, M., Bellivier, F., Barbe, B., Samolyk, D., Allilaire, J.F. and Mallet, J. (1995) Association of manic-depressive illness with tyrosine hydroxylase microsatellite marker. *Lancet*, **345**, 932.

Mendlewicz, J. and Rainer, J.D. (1977) Adoption study supporting genetic transmission in manic--depressive illness. *Nature*, **268**, 327-329.

Meyer, J., Saam, W., Mossner, R., Cangir, O., Ortega, G.R., Tatschner, T., Riederer, P., Wienker, T.F. and Lesch, K.P. (2002) Evolutionary conserved microsatellites in the promoter region of the 5-hydroxytryptamine receptor 2C gene (HTR2C) are not associated with bipolar disorder in females. *J Neural Transm*, **109**, 939-946.

Miki, Y., Swensen, J., Shattuck-Eidens, D., Futreal, P.A., Harshman, K., Tavtigian, S., Liu, Q., Cochran, C., Bennett, L.M., Ding, W. and et al. (1994) A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science*, **266**, 66-71.

Morissette, J., Villeneuve, A., Bordeleau, L., Rochette, D., Laberge, C., Gagne, B., Laprise, C., Bouchard, G., Plante, M., Gobeil, L., Shink, E., Weissenbach, J. and Barden, N. (1999) Genome-wide search for linkage of bipolar affective disorders in a very large pedigree derived from a homogeneous population in quebec points to a locus of major effect on chromosome 12q23-q24. *Am J Med Genet*, **88**, 567-587.

Mosselman, S., Polman, J. and Dijkema, R. (1996) ER beta: identification and characterization of a novel human estrogen receptor. *FEBS Lett*, **392**, 49-53.

Mowry, B.J., Ewen, K.R., Nancarrow, D.J., Lennon, D.P., Nertney, D.A., Jones, H.L., O'Brien, M.S., Thornley, C.E., Walters, M.K., Crowe, R.R., Silverman, J.M., Endicott, J., Sharpe, L., Hayward, N.K., Gladis, M.M., Foote, S.J. and Levinson, D.F. (2000) Second stage of a genome scan of schizophrenia: study of five positive regions in an expanded sample. *Am J Med Genet*, **96**, 864-869.

Muglia, P., Petronis, A., Mundo, E., Lander, S., Cate, T. and Kennedy, J.L. (2002) Dopamine D4 receptor and tyrosine hydroxylase genes in bipolar disorder: evidence for a role of DRD4. *Mol Psychiatry*, **7**, 860-866.

Muir, W.J., Thomson, M.L., McKeon, P., Mynett-Johnson, L., Whitton, C., Evans, K.L., Porteous, D.J. and Blackwood, D.H. (2001) Markers close to the dopamine D5 receptor gene (DRD5) show significant association with schizophrenia but not bipolar disorder. *Am J Med Genet*, **105**, 152-158.

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

Mundo, E., Walker, M., Tims, H., Macciardi, F. and Kennedy, J.L. (2000) Lack of linkage disequilibrium between serotonin transporter protein gene (SLC6A4) and bipolar disorder. *Am J Med Genet*, **96**, 379-383.

Mundo, E., Zai, G., Lee, L., Parikh, S.V. and Kennedy, J.L. (2001) The 5HT1Dbeta receptor gene in bipolar disorder: a family-based association study. *Neuropsychopharmacology*, **25**, 608-613.

Murphy, V.E., Mynett-Johnson, L.A., Claffey, E., Bergin, P., McAuliffe, M., Kealey, C. and McKeon, P. (2000) Search for bipolar disorder susceptibility loci: the application of a modified genome scan concentrating on gene-rich regions. *Am J Med Genet*, **96**, 728-732.

Murphy, V.E., Mynett-Johnson, L.A., Claffey, E., Shields, D.C. and McKeon, P. (2001) No association between 5HT-2A and bipolar disorder irrespective of genomic imprinting. *Am J Med Genet*, **105**, 422-425.

Mynett-Johnson, L., Kealey, C., Claffey, E., Curtis, D., Bouchier-Hayes, L., Powell, C. and McKeon, P. (2000) Multimarkerhaplotypes within the serotonin transporter gene suggest evidence of an association with bipolar disorder. *Am J Med Genet*, **96**, 845-849.

Mynett-Johnson, L.A., Murphy, V.E., Claffey, E., Shields, D.C. and McKeon, P. (1998) Preliminary evidence of an association between bipolar disorder in females and the catechol-O-methyltransferase gene. *Psychiatr Genet*, **8**, 221-225.

Mynett-Johnson, L.A., Murphy, V.E., Manley, P., Shields, D.C. and McKeon, P. (1997) Lack of evidence for a major locus for bipolar disorder in the pericentromeric region of chromosome 18 in Irish pedigrees. *Biol Psychiatry*, **42**, 486-494.

Mynett-Johnson, L.M.P. (1996) The molecular genetics of affective disorders: An overview. *Irish Journal of Psychological Medicine*, **13**, 155-161.

Nakamura, M., Ueno, S., Sano, A. and Tanabe, H. (2000) The human serotonin transporter gene linked polymorphism (5-HTTLPR) shows ten novel allelic variants. *Mol Psychiatry*, **5**, 32-38.

Nestler, E., Gould, E. and Manji, H. (2002) Preclinical models: status of basic research in depression. *Biol Psychiatry*, **52**, 503.

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

Niesler, B., Weiss, B., Fischer, C., Nothen, M.M., Propping, P., Bondy, B., Rietschel, M., Maier, W., Albus, M., Franzek, E. and Rappold, G.A. (2001) Serotonin receptor gene HTR3A variants in schizophrenic and bipolar affective patients. *Pharmacogenetics*, **11**, 21-27.

Nishiyama, N., Yukishita, S., Hagiwara, H., Kakimoto, S., Nomura, Y. and Segawa, M. (2000) Gene mutation in hereditary progressive dystonia with marked diurnal fluctuation (HPD), strictly defined dopa-responsive dystonia. *Brain Dev*, **22 Suppl 1**, S102-106.

Nothen, M.M., Cichon, S., Rohleder, H., Hemmer, S., Franzek, E., Fritze, J., Albus, M., Borrmann-Hassenbach, M., Kreiner, R., Weigelt, B., Minges, J., Lichtermann, D., Maier, W., Craddock, N., Fimmers, R., Holler, T., Baur, M.P., Rietschel, M. and Propping, P. (1999)

Evaluation of linkage of bipolar affective disorder to chromosome 18 in a sample of 57 German families. *Mol Psychiatry*, **4**, 76-84.

Nothen, M.M., Erdmann, J., Korner, J., Lanczik, M., Fritze, J., Fimmers, R., Grandy, D.K., O'Dowd, B. and Propping, P. (1992) Lack of association between dopamine D1 and D2 receptor genes and bipolar affective disorder. *Am J Psychiatry*, **149**, 199-201.

O'Donovan, M.C., Oefner, P.J., Roberts, S.C., Austin, J., Hoogendoorn, B., Guy, C., Speight, G., Upadhyaya, M., Sommer, S.S. and McGuffin, P. (1998) Blind analysis of denaturing high-performance liquid chromatography as a tool for mutation detection. *Genomics*, **52**, 44-49.

Ohara, K., Nagai, M. and Suzuki, Y. (1998) Low activity allele of catechol-o-methyltransferase gene and Japanese unipolar depression. *Neuroreport*, **9**, 1305-1308.

Ohtsuki, T., Ishiguro, H., Detera-Wadleigh, S.D., Toyota, T., Shimizu, H., Yamada, K., Yoshitsugu, K., Hattori, E., Yoshikawa, T. and Arinami, T. (2002) Association between serotonin 4 receptor gene polymorphisms and bipolar disorder in Japanese case-control samples and the NIMH Genetics Initiative Bipolar Pedigrees. *Mol Psychiatry*, **7**, 954-961.

Oliveira, J.R., Carvalho, D.R., Pontual, D., Gallindo, R.M., Sougey, E.B., Gentil, V., Lafer, B., Maia, L.G., Morais, M.A., Jr., Matioli, S., Vallada, H., Moreno, R.A., Nishimura, A., Otto, P.A., Passos-Bueno, M.R. and Zatz, M. (2000) Analysis of the serotonin transporter polymorphism (5-HTTLPR) in Brazilian patients affected by dysthymia, major depression and bipolar disorder. *Mol Psychiatry*, **5**, 348-349.

Ophoff, R.A., Escamilla, M.A., Service, S.K., Spesny, M., Meshi, D.B., Poon, W., Molina, J., Fournier, E., Gallegos, A., Mathews, C., Neylan, T., Batki, S.L., Roche, E., Ramirez, M., Silva, S., De Mille, M.C., Dong, P., Leon, P.E., Reus, V.I., Sandkuijl, L.A. and Freimer, N.B. (2002) Genomewide linkage disequilibrium mapping of severe bipolar disorder in a population isolate. *Am J Hum Genet*, **71**, 565-574.

Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K. and Sekiya, T. (1989) Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proc Natl Acad Sci U S A*, **86**, 2766-2770.

Oruc, L., Verheyen, G.R., Furac, I., Jakovljevic, M., Ivezic, S., Raeymaekers, P. and Van Broeckhoven, C. (1997) Analysis of the tyrosine hydroxylase and dopamine D4 receptor genes in a Croatian sample of bipolar I and unipolar patients. *Am J Med Genet*, **74**, 176-178.

Ospina-Duque, J., Duque, C., Carvajal-Carmona, L., Ortiz-Barrientos, D., Soto, I., Pineda, N., Cuartas, M., Calle, J., Lopez, C., Ochoa, L., Garcia, J., Gomez, J., Agudelo, A., Lozano, M., Montoya, G., Ospina, A., Lopez, M., Gallo, A., Miranda, A., Serna, L., Montoya, P., Palacio, C., Bedoya, G., McCarthy, M., Reus, V., Freimer, N. and Ruiz-Linares, A. (2000) An association study of bipolar mood disorder (type I) with the 5-HTTLPR serotonin transporter polymorphism in a human population isolate from Colombia. *Neurosci Lett*, **292**, 199-202.

Overhauser, J., Berrettini, W.H. and Rojas, K. (1998) Affective disorder associated with a balanced translocation involving chromosome 14 and 18. *Psychiatr Genet*, **8**, 53-56.

Owen, M.J. and Craddock, N. (1996) Modern molecular genetic approaches to complex traits: implications for psychiatric disorders. *Mol Psychiatry*, **1**, 21-26.

Paes de Sousa, M. and Tropa, J. (1989) Evaluation of the efficacy of amineptine in a population of 1,229 depressed patients: results of a multicenter study carried out by 135 general practitioners. *Clin Neuropharmacol*, **12 Suppl 2**, S77-86.

Papolos, D.F., Faedda, G.L., Veit, S., Goldberg, R., Morrow, B., Kucherlapati, R. and Shprintzen, R.J. (1996) Bipolar spectrum disorders in patients diagnosed with velo-cardio-facial syndrome: does a hemizygous deletion of chromosome 22q11 result in bipolar affective disorder? *Am J Psychiatry*, **153**, 1541-1547.

Papolos, D.F., Veit, S., Faedda, G.L., Saito, T. and Lachman, H.M. (1998) Ultra-ultra rapid cycling bipolar disorder is associated with the low activity catecholamine-O-methyltransferase allele. *Mol Psychiatry*, **3**, 346-349.

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

Paunio, T., Ekelund, J., Varilo, T., Parker, A., Hovatta, I., Turunen, J.A., Rinard, K., Foti, A., Terwilliger, J.D., Juvonen, H., Suvisaari, J., Arajarvi, R., Suokas, J., Partonen, T., Lonnqvist, J., Meyer, J. and Peltonen, L. (2001) Genome-wide scan in a nationwide study sample of schizophrenia families in Finland reveals susceptibility loci on chromosomes 2q and 5q. *Hum Mol Genet*, **10**, 3037-3048.

Pennisi, E. (1998) A closer look at SNPs suggests difficulties. *Science*, **281**, 1787-1789.

Piccardi, M.P., Ardau, R., Chillotti, C., Deleuze, J.F., Mallet, J., Meloni, R., Oi, A., Severino, G., Congiu, D., Bayorek, M. and Del Zompo, M. (2002) Manic-depressive illness: an association study with the inositol polyphosphate 1-phosphatase and serotonin transporter genes. *Psychiatr Genet*, **12**, 23-27.

Poland, D. (1974) Recursion relation generation of probability profiles for specific-sequence macromolecules with long-range correlations. *Biopolymers*, **13**, 1859-1871.

Polymeropoulos, M.H., Coon, H., Byerley, W., Gershon, E.S., Goldin, L., Crow, T.J., Rubenstein, J., Hoff, M., Holik, J., Smith, A.M. and et al. (1994) Search for a schizophrenia susceptibility locus on human chromosome 22. *Am J Med Genet*, **54**, 93-99.

Polymeropoulos, M.H. and Schaffer, A.A. (1996) Scanning the genome with 1772 microsatellite markers in search of a bipolar disorder susceptibility gene. *Mol Psychiatry*, **1**, 404-407.

Preisig, M., Bellivier, F., Fenton, B.T., Baud, P., Berney, A., Courtet, P., Hardy, P., Golaz, J., Leboyer, M., Mallet, J., Matthey, M.L., Mouthon, D., Neidhart, E., Nosten-Bertrand, M., Stadelmann-Dubuis, E., Guimon, J., Ferrero, F., Buresi, C. and Malafosse, A. (2000) Association between bipolar disorder and monoamine oxidase A gene polymorphisms: results of a multicenter study. *Am J Psychiatry*, **157**, 948-955.

Pulver, A.E., Karayiorgou, M., Wolyniec, P.S., Lasseter, V.K., Kasch, L., Nestadt, G., Antonarakis, S., Housman, D., Kazazian, H.H., Meyers, D. and et al. (1994) Sequential strategy to identify a susceptibility gene for schizophrenia: report of potential linkage on chromosome 22q12-q13.1: Part 1. *Am J Med Genet*, **54**, 36-43.

Raeymaekers, P., Verheyen, G., Van Zand, K., Souery, D., Cales, S. and Van Broeckhoven, C. (1997) Linkage studies with chromosomal candidate regions in bipolar families. *Am J Hum Genet*, **74**, 587.

Rees, M., Norton, N., Jones, I., McCandless, F., Scourfield, J., Holmans, P., Moorhead, S., Feldman, E., Sadler, S., Cole, T., Redman, K., Farmer, A., McGuffin, P., Owen, M.J. and Craddock, N. (1997) Association studies of bipolar disorder at the human serotonin transporter gene (hSERT; 5HTT). *Mol Psychiatry*, **2**, 398-402.

Reich, D.E., Cargill, M., Bolk, S., Ireland, J., Sabeti, P.C., Richter, D.J., Lavery, T., Kouyoumjian, R., Farhadian, S.F., Ward, R. and Lander, E.S. (2001) Linkage disequilibrium in the human genome. *Nature*, **411**, 199-204.

Reyes, G.D., Esterling, L.E., Corona, W., Ferraren, D., Rollins, D.Y., Padigaru, M., Yoshikawa, T., Monje, V.D. and Detera-Wadleigh, S.D. (2002) Map of candidate genes and STSs on 18p11.2, a bipolar disorder and schizophrenia susceptibility region. *Mol Psychiatry*, **7**, 337-339.

Rice, J.P., Goate, A., Williams, J.T., Bierut, L., Dorr, D., Wu, W., Shears, S., Gopalakrishnan, G., Edenberg, H.J., Foroud, T., Nurnberger, J., Jr., Gershon, E.S., Detera-Wadleigh, S.D., Goldin, L.R., Guroff, J.J., McMahon, F.J., Simpson, S., MacKinnon, D., McInnis, M., Stine, O.C., DePaulo, J.R., Blehar, M.C. and Reich, T. (1997) Initial genome scan of the NIMH genetics initiative bipolar pedigrees: chromosomes 1, 6, 8, 10, and 12. *Am J Med Genet*, **74**, 247-253.

Risch, N. and Merikangas, K. (1996) The future of genetic studies of complex human diseases. *Science*, **273**, 1516-1517.

Risch, N.J. (2000) Searching for genetic determinants in the new millennium. *Nature*, **405**, 847-856.

Rojas, K., Liang, L., Johnson, E.I., Berrettini, W.H. and Overhauser, J. (2000) Identification of candidate genes for psychiatric disorders on 18p11. *Mol Psychiatry*, **5**, 389-395.

Rommens, J.M., Iannuzzi, M.C., Kerem, B., Drumm, M.L., Melmer, G., Dean, M., Rozmahel, R., Cole, J.L., Kennedy, D., Hidaka, N. and et al. (1989) Identification of the cystic fibrosis gene: chromosome walking and jumping. *Science*, **245**, 1059-1065.

Rosenkranz, K., Hinney, A., Ziegler, A., Hermann, H., Fichter, M., Mayer, H., Siegfried, W., Young, J.K., Remschmidt, H. and Hebebrand, J. (1998) Systematic mutation screening of the estrogen receptor beta gene in probands of different weight extremes: identification of several genetic variants. *J Clin Endocrinol Metab*, **83**, 4524-4527.

Roses, A.D. (2001) Pharmacogenetics. *Hum Mol Genet*, **10**, 2261-2267.

Rubocki, R.J., Wisecarver, J.L., Hook, D.D., Cox, S.M. and Beisel, K.W. (1992) Histocompatibility screening by molecular techniques: use of polymerase chain reaction products and heteroduplex formation. *J Clin Lab Anal*, **6**, 337-341.

Sachidanandam, R., Weissman, D., Schmidt, S.C., Kakol, J.M., Stein, L.D., Marth, G., Sherry, S., Mullikin, J.C., Mortimore, B.J., Willey, D.L., Hunt, S.E., Cole, C.G., Coggill, P.C., Rice, C.M., Ning, Z., Rogers, J., Bentley, D.R., Kwok, P.Y., Mardis, E.R., Yeh, R.T., Schultz, B., Cook, L., Davenport, R., Dante, M., Fulton, L., Hillier, L., Waterston, R.H., McPherson, J.D., Gilman, B., Schaffner, S., Van Etten, W.J., Reich, D., Higgins, J., Daly, M.J., Blumenstiel, B., Baldwin, J., Stange-Thomann, N., Zody, M.C., Linton, L., Lander, E.S. and Altshuler, D. (2001) A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature*, **409**, 928-933.

Saleem, Q., Ganesh, S., Vijaykumar, M., Reddy, Y.C., Brahmachari, S.K. and Jain, S. (2000) Association analysis of 5HT transporter gene in bipolar disorder in the Indian population. *Am J Med Genet*, **96**, 170-172.

Satsangi, J., Parkes, M., Louis, E., Hashimoto, L., Kato, N., Welsh, K., Terwilliger, J.D., Lathrop, G.M., Bell, J.I. and Jewell, D.P. (1996) Two stage genome-wide search in

inflammatory bowel disease provides evidence for susceptibility loci on chromosomes 3, 7 and 12. *Nat Genet*, **14**, 199-202.

Savoye, C., Laurent, C., Amadeo, S., Gheysen, F., Leboyer, M., Lejeune, J., Zarifian, E. and Mallet, J. (1998) No association between dopamine D1, D2, and D3 receptor genes and manic-depressive illness. *Biol Psychiatry*, **44**, 644-647.

Scambler, P.J., Kelly, D., Lindsay, E., Williamson, R., Goldberg, R., Shprintzen, R., Wilson, D.I., Goodship, J.A., Cross, I.E. and Burn, J. (1992) Velo-cardio-facial syndrome associated with chromosome 22 deletions encompassing the DiGeorge locus. *Lancet*, **339**, 1138-1139.

Schaid, D.J. (1998) Transmission disequilibrium, family controls, and great expectations. *Am J Hum Genet*, **63**, 935-941.

Schaid, D.J. and Sommer, S.S. (1994) Comparison of statistics for candidate-gene association studies using cases and parents. *Am J Hum Genet*, **55**, 402-409.

Schildkraut, J.J., Gordon, E.K. and Durell, J. (1965) Catecholamine metabolism in affective disorders. I. Normetanephrine and VMA excretion in depressed patients treated with imipramine. *J Psychiatr Res*, **3**, 213-228.

Schork, N.J., Boehnke, M., Terwilliger, J.D. and Ott, J. (1993) Two-trait-locus linkage analysis: a powerful strategy for mapping complex genetic traits. *Am J Hum Genet*, **53**, 1127-1136.

Schwab, S.G., Eckstein, G.N., Hallmayer, J., Lerer, B., Albus, M., Borrmann, M., Lichtermann, D., Ertl, M.A., Maier, W. and Wildenauer, D.B. (1997) Evidence suggestive of a locus on chromosome 5q31 contributing to susceptibility for schizophrenia in German and Israeli families by multipoint affected sib-pair linkage analysis. *Mol Psychiatry*, **2**, 156-160.

Schwab, S.G., Hallmayer, J., Albus, M., Lerer, B., Hanses, C., Kanyas, K., Segman, R., Borrman, M., Dreikorn, B., Lichtermann, D., Rietschel, M., Trixler, M., Maier, W. and Wildenauer, D.B. (1998) Further evidence for a susceptibility locus on chromosome 10p14-p11 in 72 families with schizophrenia by nonparametric linkage analysis. *Am J Med Genet*, **81**, 302-307.

Schwab, S.G., Hallmayer, J., Lerer, B., Albus, M., Borrmann, M., Honig, S., Strauss, M., Segman, R., Lichtermann, D., Knapp, M., Trixler, M., Maier, W. and Wildenauer, D.B. (1998) Support for a chromosome 18p locus conferring susceptibility to functional psychoses in families with schizophrenia, by association and linkage analysis. *Am J Hum Genet*, **63**, 1139-1152.

Scriver, C.R., Eisensmith, R.C., Woo, S.L. and Kaufman, S. (1994) The hyperphenylalaninemias of man and mouse. *Annu Rev Genet*, **28**, 141-165.

Serretti, A., Lilli, R., Lorenzi, C., Lattuada, E. and Smeraldi, E. (2001) DRD4 exon 3 variants associated with delusional symptomatology in major psychoses: a study on 2,011 affected subjects. *Am J Med Genet*, **105**, 283-290.

Serretti, A., Macciardi, F., Catalano, M., BelLODi, L. and Smeraldi, E. (1999) Genetic variants of dopamine receptor D4 and psychopathology. *Schizophr Bull*, **25**, 609-618.

Serretti, A., Macciardi, F., Cusin, C., Lattuada, E., Souery, D., Lipp, O., Mahieu, B., Van Broeckhoven, C., Blackwood, D., Muir, W., Aschauer, H.N., Heiden, A.M., Ackenheil, M., Fuchshuber, S., Raeymaekers, P., Verheyen, G., Kaneva, R., Jablensky, A., Papadimitriou,

G.N., Dikeos, D.G., Stefanis, C.N., Smeraldi, E. and Mendlewicz, J. (2000) Linkage of mood disorders with D2, D3 and TH genes: a multicenter study. *J Affect Disord*, **58**, 51-61.

Serretti, A., Macciardi, F., Verga, M., Cusin, C., Pedrini, S. and Smeraldi, E. (1998) Tyrosine hydroxylase gene associated with depressive symptomatology in mood disorder. *Am J Med Genet*, **81**, 127-130.

Shah, M., Coon, H., Holik, J., Hoff, M., Helmer, V., Panos, P. and Byerley, W. (1995) Mutation scan of the D1 dopamine receptor gene in 22 cases of bipolar I disorder. *Am J Med Genet*, **60**, 150-153.

Sham, P.C. and Curtis, D. (1995) An extended transmission/disequilibrium test (TDT) for multi-allele marker loci. *Ann Hum Genet*, **59 (Pt 3)**, 323-336.

Shaw, S.H., Kelly, M., Smith, A.B., Shields, G., Hopkins, P.J., Loftus, J., Laval, S.H., Vita, A., De Hert, M., Cardon, L.R., Crow, T.J., Sherrington, R. and DeLisi, L.E. (1998) A genome-wide search for schizophrenia susceptibility genes. *Am J Med Genet*, **81**, 364-376.

Sherwin, B.B. (1988) Affective changes with estrogen and androgen replacement therapy in surgically menopausal women. *J Affect Disord*, **14**, 177-187.

Shi, M.M. (2001) Enabling large-scale pharmacogenetic studies by high-throughput mutation detection and genotyping technologies. *Clin Chem*, **47**, 164-172.

Shi, M.M., Bleavins, M.R. and de la iglesia, F.A. (1999) Technologies for detecting genetic polymorphisms in pharmacogenomics. *Mol Diagn*, **4**, 343-351.

Shink, E., Morissette, J., Rochette, D., Bordeleau, L., Plante, M., Villeneuve, A. and Barden, N. (1998) Bipolar affective disorder susceptibility loci on chromosomes 5 and 21: heterogeneity in a homogeneous population in Quebec. *Am J Hum Genet*, **6**, 541-542.

Shiozawa, S., Hayashi, S., Tsukamoto, Y., Goko, H., Kawasaki, H., Wada, T., Shimizu, K., Yasuda, N., Kamatani, N., Takasugi, K., Tanaka, Y., Shiozawa, K. and Imura, S. (1998) Identification of the gene loci that predispose to rheumatoid arthritis. *Int Immunol*, **10**, 1891-1895.

Shprintzen, R.J., Goldberg, R., Golding-Kushner, K.J. and Marion, R.W. (1992) Late-onset psychosis in the velo-cardio-facial syndrome. *Am J Med Genet*, **42**, 141-142.

Silverman, J.M., Greenberg, D.A., Altstiel, L.D., Siever, L.J., Mohs, R.C., Smith, C.J., Zhou, G., Hollander, T.E., Yang, X.P., Kedache, M., Li, G., Zaccario, M.L. and Davis, K.L. (1996) Evidence of a locus for schizophrenia and related disorders on the short arm of chromosome 5 in a large pedigree. *Am J Med Genet*, **67**, 162-171.

Simpson, S.G., Folstein, S.E., Meyers, D.A., McMahon, F.J., Brusco, D.M. and DePaulo, J.R., Jr. (1993) Bipolar II: the most common bipolar phenotype? *Am J Psychiatry*, **150**, 901-903.

Smith, A.L. and Weissman, M.M. (1992) Epidemiology. Handbook of affective disorders. In. Edinburgh: Churchill Livingstone, pp. 111-129.

Smyth, C., Kalsi, G., Curtis, D., Brynjolfsson, J., O'Neill, J., Rifkin, L., Moloney, E., Murphy, P., Petursson, H. and Gurling, H. (1997) Two-locus admixture linkage analysis of bipolar and unipolar affective disorder supports the presence of susceptibility loci on chromosomes 11p15 and 21q22. *Genomics*, **39**, 271-278.

Sneddon, J.M. (1973) Blood platelets as a model for monoamine-containing neurones. *Prog Neurobiol*, **1**, 151-198.

Soares, J.C. (2002) Can brain-imaging studies provide a 'mood stabilizer signature?' *Mol Psychiatry*, **7 Suppl 1**, S64-70.

Speight, G., Turic, D., Austin, J., Hoogendoorn, B., Cardno, A.G., Jones, L., Murphy, K.C., Sanders, R., McCarthy, G., Jones, I., McCandless, F., McGuffin, P., Craddock, N., Owen, M.J., Buckland, P. and O'Donovan, M.C. (2000) Comparative sequencing and association studies of aromatic L-amino acid decarboxylase in schizophrenia and bipolar disorder. *Mol Psychiatry*, **5**, 327-331.

Spence, M.A., FLODman, P.L., Sadovnick, A.D., Bailey-Wilson, J.E., Ameli, H. and Remick, R.A. (1995) Bipolar disorder: evidence for a major locus. *Am J Med Genet*, **60**, 370-376.

Spielman, R.S. and Ewens, W.J. (1996) The TDT and other family-based tests for linkage disequilibrium and association. *Am J Hum Genet*, **59**, 983-989.

Spielman, R.S., McGinnis, R.E. and Ewens, W.J. (1993) Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet*, **52**, 506-516.

St George-Hyslop, P.H., Tanzi, R.E., Polinsky, R.J., Haines, J.L., Nee, L., Watkins, P.C., Myers, R.H., Feldman, R.G., Pollen, D., Drachman, D. and et al. (1987) The genetic defect causing familial Alzheimer's disease maps on chromosome 21. *Science*, **235**, 885-890.

Steinberger, D., Korinthenberg, R., Topka, H., Berghauser, M., Wedde, R. and Muller, U. (2000) Dopa-responsive dystonia: mutation analysis of GCH1 and analysis of therapeutic doses of L-dopa. German Dystonia Study Group. *Neurology*, **55**, 1735-1737.

Stine, O.C., McMahon, F.J., Chen, L., Xu, J., Meyers, D.A., MacKinnon, D.F., Simpson, S., McInnis, M.G., Rice, J.P., Goate, A., Reich, T., Edenberg, H.J., Foroud, T., Nurnberger, J.I., Jr., Detera-Wadleigh, S.D., Goldin, L.R., Guroff, J., Gershon, E.S., Blehar, M.C. and DePaulo, J.R. (1997) Initial genome screen for bipolar disorder in the NIMH genetics initiative pedigrees: chromosomes 2, 11, 13, 14, and X. *Am J Med Genet*, **74**, 263-269.

Stine, O.C., Xu, J., Koskela, R., McMahon, F.J., Gschwend, M., Friddle, C., Clark, C.D., McInnis, M.G., Simpson, S.G., Breschel, T.S. and et al. (1995) Evidence for linkage of bipolar disorder to chromosome 18 with a parent-of-origin effect. *Am J Hum Genet*, **57**, 1384-1394.

Strakowski, S.M., DelBello, M.P., Adler, C., Cecil, D.M. and Sax, K.W. (2000) Neuroimaging in bipolar disorder. *Bipolar Disord*, **2**, 148-164.

Straub, R.E., Lehner, T., Luo, Y., Loth, J.E., Shao, W., Sharpe, L., Alexander, J.R., Das, K., Simon, R., Fieve, R.R. and et al. (1994) A possible vulnerability locus for bipolar affective disorder on chromosome 21q22.3. *Nat Genet*, **8**, 291-296.

Straub, R.E., MacLean, C.J., Martin, R.B., Ma, Y., Myakishev, M.V., Harris-Kerr, C., Webb, B.T., O'Neill, F.A., Walsh, D. and Kendler, K.S. (1998) A schizophrenia locus may be located in region 10p15-p11. *Am J Med Genet*, **81**, 296-301.

Straub, R.E., MacLean, C.J., O'Neill, F.A., Walsh, D. and Kendler, K.S. (1997) Support for a possible schizophrenia vulnerability locus in region 5q22-31 in Irish families. *Mol Psychiatry*, **2**, 148-155.

Strittmatter, W.J., Saunders, A.M., Schmechel, D., Pericak-Vance, M., Enghild, J., Salvesen, G.S. and Roses, A.D. (1993) Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. *Proc Natl Acad Sci U S A*, **90**, 1977-1981.

Syagailo, Y.V., Stober, G., Grassle, M., Reimer, E., Knapp, M., Jungkunz, G., Okladnova, O., Meyer, J. and Lesch, K.P. (2001) Association analysis of the functional monoamine oxidase A gene promoter polymorphism in psychiatric disorders. *Am J Med Genet*, **105**, 168-171.

Taillon-Miller, P., Bauer-Sardina, I., Saccone, N.L., Putzel, J., Laitinen, T., Cao, A., Kere, J., Pilia, G., Rice, J.P. and Kwok, P.Y. (2000) Juxtaposed regions of extensive and minimal linkage disequilibrium in human Xq25 and Xq28. *Nat Genet*, **25**, 324-328.

Tani, Y., Fernell, E., Watanabe, Y., Kanai, T. and Langstrom, B. (1994) Decrease in 6R-5,6,7,8-tetrahydrobiopterin content in cerebrospinal fluid of autistic patients. *Neurosci Lett*, **181**, 169-172.

Taylor, M.A. and Amir, N. (1994) Are schizophrenia and affective disorder related?: the problem of schizoaffective disorder and the discrimination of the psychoses by signs and symptoms. *Compr Psychiatry*, **35**, 420-429.

Terwilliger, J. and Ott, J. (1994) *Handbook of human genetic linkage.* John Hopkins University Press, Baltimore.

Terwilliger, J.D. and Goring, H.H. (2000) Gene mapping in the 20th and 21st centuries: statistical methods, data analysis, and experimental design. *Hum Biol*, **72**, 63-132.

Terwilliger, J.D. and Ott, J. (1992) A haplotype-based 'haplotype relative risk' approach to detecting allelic associations. *Hum Hered*, **42**, 337-346.

Terwilliger, J.D., Shannon, W.D., Lathrop, G.M., Nolan, J.P., Goldin, L.R., Chase, G.A. and Weeks, D.E. (1997) True and false positive peaks in genomewide scans: applications of length-biased sampling to linkage mapping. *Am J Hum Genet*, **61**, 430-438.

Terwilliger, J.D. and Weiss, K.M. (1998) Linkage disequilibrium mapping of complex disease: fantasy or reality? *Curr Opin Biotechnol*, **9**, 578-594.

Thony, B., Auerbach, G. and Blau, N. (2000) Tetrahydrobiopterin biosynthesis, regeneration and functions. *Biochem J*, **347 Pt 1**, 1-16.

Tomfohrde, J., Silverman, A., Barnes, R., Fernandez-Vina, M.A., Young, M., Lory, D., Morris, L., Wuepper, K.D., Stastny, P., Menter, A. and et al. (1994) Gene for familial psoriasis susceptibility mapped to the distal end of human chromosome 17q. *Science*, **264**, 1141-1145.

Torgersen, S. (1986) Genetic factors in moderately severe and mild affective disorders. *Arch Gen Psychiatry*, **43**, 222-226.

Trembath, R.C., Clough, R.L., Rosbotham, J.L., Jones, A.B., Camp, R.D., Frodsham, A., Browne, J., Barber, R., Terwilliger, J., Lathrop, G.M. and Barker, J.N. (1997) Identification of a major susceptibility locus on chromosome 6p and evidence for further disease loci revealed by a two stage genome-wide search in psoriasis. *Hum Mol Genet*, **6**, 813-820.

Tsuang, M., Faranone, SV. (1990) *The Genetics of Mood Disorders*. The John Hopkins University Press.

Turecki, G., Grof, P., Cavazzoni, P., Duffy, A., Grof, E., Martin, R., Joober, R., Rouleau, G.A. and Alda, M. (1999) Lithium responsive bipolar disorder, unilineality, and chromosome 18: A linkage study. *Am J Med Genet*, **88**, 411-415.

Turecki, G., Rouleau, G.A., Mari, J., Joober, R. and Morgan, K. (1997) Lack of association between bipolar disorder and tyrosine hydroxylase: a meta-analysis. *Am J Med Genet*, **74**, 348-352.

Underhill, P.A., Jin, L., Lin, A.A., Mehdi, S.Q., Jenkins, T., Vollrath, D., Davis, R.W., Cavalli-Sforza, L.L. and Oefner, P.J. (1997) Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. *Genome Res*, **7**, 996-1005.

Underhill, P.A., Jin, L., Zemans, R., Oefner, P.J. and Cavalli-Sforza, L.L. (1996) A pre-Columbian Y chromosome-specific transition and its implications for human evolutionary history. *Proc Natl Acad Sci U S A*, **93**, 196-200.

Vallada, H.P., Gill, M., Sham, P., Lim, L.C., Nanko, S., Asherson, P., Murray, R.M., McGuffin, P., Owen, M. and Collier, D. (1995) Linkage studies on chromosome 22 in familial schizophrenia. *Am J Med Genet*, **60**, 139-146.

Vermeulen, A. (1993) The male climacterium. *Ann Med*, **25**, 531-534.

Vincent, J.B., Masellis, M., Lawrence, J., Choi, V., Gurling, H.M., Parikh, S.V. and Kennedy, J.L. (1999) Genetic association analysis of serotonin system genes in bipolar affective disorder. *Am J Psychiatry*, **156**, 136-138.

Vogt, I.R., Shimron-Abarbanell, D., Neidt, H., Erdmann, J., Cichon, S., Schulze, T.G., Muller, D.J., Maier, W., Albus, M., Borrmann-Hassenbach, M., Knapp, M., Rietschel, M., Propping, P. and Nothen, M.M. (2000) Investigation of the human serotonin 6 [5-HT6] receptor gene in bipolar affective disorder and schizophrenia. *Am J Med Genet*, **96**, 217-221.

Wald, D., Ebstein, R.P. and Belmaker, R.H. (1978) Haloperidol and lithium blocking of the mood response to intravenous methylphenidate. *Psychopharmacology (Berl)*, **57**, 83-87.

Wang, C., Alexander, G., Berman, N., Salehian, B., Davidson, T., McDonald, V., Steiner, B., Hull, L., Callegari, C. and Swerdloff, R.S. (1996) Testosterone replacement therapy improves mood in hypogonadal men--a clinical research center study. *J Clin Endocrinol Metab*, **81**, 3578-3583.

Wang, D.G., Fan, J.B., Siao, C.J., Berno, A., Young, P., Sapolsky, R., Ghandour, G., Perkins, N., Winchester, E., Spencer, J., Kruglyak, L., Stein, L., Hsie, L., Topaloglou, T., Hubbell, E., Robinson, E., Mittmann, M., Morris, M.S., Shen, N., Kilburn, D., Rioux, J., Nusbaum, C., Rozen, S., Hudson, T.J., Lander, E.S. and et al. (1998) Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science*, **280**, 1077-1082.

Wartell, R.M. and Benight, A.S. (1985) Thermal-denaturation of DNA-molecules: a comparison of theory with experiment. *Phys Rep*, **126**, 67-107.

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

Weeks, D.E. and Lathrop, G.M. (1995) Polygenic disease: methods for mapping complex disease traits. *Trends Genet*, **11**, 513-519.

Weissman, M.M., Leaf, P.J., Bruce, M.L. and Florio, L. (1988) The epidemiology of dysthymia in five communities: rates, risks, comorbidity, and treatment. *Am J Psychiatry*, **145**, 815-819.

Wender, P.H., Kety, S.S., Rosenthal, D., Schulsinger, F., Ortmann, J. and Lunde, I. (1986) Psychiatric disorders in the biological and adoptive families of adopted individuals with affective disorders. *Arch Gen Psychiatry*, **43**, 923-929.

Williams, J., Spurlock, G., Holmans, P., Mant, R., Murphy, K., Jones, L., Cardno, A., Asherson, P., Blackwood, D., Muir, W., Meszaros, K., Aschauer, H., Mallet, J., Laurent, C., Pekkarinen, P., Seppala, J., Stefanis, C.N., Papadimitriou, G.N., Macciardi, F., Verga, M., Pato, C., Azevedo, H., Crocq, M.A., Gurling, H., Owen, M.J. and et al. (1998) A meta-analysis and transmission disequilibrium study of association between the dopamine D3 receptor gene and schizophrenia. *Mol Psychiatry*, **3**, 141-149.

Williams, N.M., Rees, M.I., Holmans, P., Norton, N., Cardno, A.G., Jones, L.A., Murphy, K.C., Sanders, R.D., McCarthy, G., Gray, M.Y., Fenton, I., McGuffin, P. and Owen, M.J. (1999) A two-stage genome scan for schizophrenia susceptibility genes in 196 affected sibling pairs. *Hum Mol Genet*, **8**, 1729-1739.

Wise, L.H. and Lewis, C.M. (1999) A method for meta-analysis of genome searches: application to simulated data. *Genet Epidemiol*, **17**, S767-71.

Witter, K., Werner, T., Blusch, J.H., Schneider, E.M., Riess, O., Ziegler, I., Rodl, W., Bacher, A. and Gutlich, M. (1996) Cloning, sequencing and functional studies of the gene encoding human GTP cyclohydrolase I. *Gene*, **171**, 285-290.

Wooster, R., Bignell, G., Lancaster, J., Swift, S., Seal, S., Mangion, J., Collins, N., Gregory, S., Gumbs, C. and Micklem, G. (1995) Identification of the breast cancer susceptibility gene BRCA2. *Nature*, **378**, 789-792.

Xie, T., Ho, S.L. and Ramsden, D. (1999) Characterization and implications of estrogenic down-regulation of human catechol-O-methyltransferase gene transcription. *Mol Pharmacol*, **56**, 31-38.

Zhao, J.H., Sham, P.C. and Curtis, D. (1999) A program for the monte carlo evaluation of significance of the extended transmission/disequilibrium test. *Am J Hum Genet*, **64**, 1484-1485.

APPENDIX

Appendix A: Stock Solutions
All chemical reagents were obtained from Sigma, BDH or GIBCO BRL.

DNA Extractions (section 2.2)

Buffer A: (1 Litre)

1M TRIS (pH 7.5)	10ml
1M $MgCl_2$	5ml
100% Triton-X100	10ml
0.32M Sucrose	100.5g
Water	up to 1L

Buffer B: (200ml)

0.5 M NaCl	30ml
0.5 M EDTA	7.5ml
Water	up to 200ml

PCR Reagents (section 2.3)

$MgCl_2$ Buffer	10mM/15mM/20mM
1M KCl	2.5ml
1M TRIS (pH 9.0)	0.5ml
1M $MgCl_2$	50µl/75µl/100µl
1% Gelatin	0.5ml
100% Triton-X100	50µl
Water	up to 5ml

Cold dNTP mix		***Hot dNTP mix***	
dTTP (0.1µm/µl)	10µl	dTTP (0.1µm/µl)	10µl
dATP (0.1µm/µl)	10µl	dATP (0.1µm/µl)	10 µl
dGTP (0.1µm/µl)	10µl	dGTP (0.1µm/µl)	10 µl
dCTP (0.1µm/µl)	10µl	dCTP 1/100 dil (0.1µm/µl)	12.5 µl
Water	760µl	Water	717.5 µl

Denaturing Gel Electrophoresis (section 2.5.1)

10X TBE (1Litre)

Boric Acid	55g
Trizma Base	108g
0.5M EDTA, pH 8.0	40ml
Water	up to 1L

7.75M *Urea Solution* (1Litre)

Urea	233.5g
10X TBE	50ml
Water	up to 1L

20% (w/v) Denaturing Acrylamide Solution (500ml)

Acrylamide	96.5g
Bis-acrylamide	3.35g
Urea	233.5g
10X TBE	50ml
Water	up to 500ml

8% (w/v) Denaturing Acrylamide Gel

20% Denaturing Acrylamide Solution	30ml
7.75M Urea Solution	45ml
Ammonium Persulphate 10% (w/v)	600µl
TEMED	75µl

Formamide Loading Dye

Bromophenol Blue	50mg
Xylene Cyanol	50mg
0.5M EDTA (pH 8.0)	1ml
Formamide (80%)	up to 50ml

Cresol Red Dye

Cresol Red	0.125g
1 X TBE	50 ml

Non-Denaturing Gel Electrophoresis (section 2.5.2)

6% (w/v) Non-Denaturing Gel

40% Acrylamide Mix (Sequagel)	6ml
10X TBE	4ml
Water	30ml
Ammonium Persulphate 10% (w/v)	280µl
TEMED	14µl

Agarose Loading Dye

Bromophenol Blue	12.5mg
Xylene Cyanol	12.5mg
Glycerol	15ml
Water	up to 50 ml

Silver Staining Solutions

Solution 1:	Ethanol 10% (v/v)	50ml
	Acetic Acid 0.5% (v/v)	2.5ml
	Water	up to 500ml
Solution 2:	Silver Nitrate ($AgNO_3$) 0.1%	50ml
	Water	up to 500ml
Solution 3:	Sodium Hydroxide	7.5g (final concentration 0.1% (w/v))
	Formaldehyde 0.1%	3ml
	Water	up to 500ml
Solution 4:	Sodium Carbonate 0.75% (w/v)	3.75g
	Water	up to 500ml

Agarose Gels

Gels of 1.5-3% were made by the addition of agarose to 1XTBE. The agarose was solubilised by microwaving. 5µl of ethidium bromide was added to the gel mix before pouring. 1XTBE was used as the running buffer.

Denaturing High Performance Liquid Chromatography (DHPLC) (section 2.10)

All buffer solutions for the Wave were made to analytical chemistry standards, using quality assured glassware.

Buffer A

2M TEAA	50ml
Acetonitrile (HPLC grade)	250µl
Water	up to 1L

Buffer B

2M TEAA	50ml
Acetonitrile (HPLC grade)	250ml
Water	up to 1L

Buffer C

Acetonitrile (HPLC grade)	750ml
Water	up to 1L

Buffer D

Acetonitrile (HPLC grade)	80ml
Water	up to 1L

Buffers A, B and D were made weekly and Buffer C was made once every 3 weeks.

APPENDIX B: MICROSATELLITE MARKERS - PRIMER SEQUENCES

Primers were supplied by Sigma-Genosys and Research Genetics.

All primers are written in 5'-3' direction

B.1 Chromosome 17q microsatellite markers

Marker	Forward Primer	Reverse Primer
GATA158H04	ACATACAGCACAGGCCAAAT	CCAGTCTTCCGTCACTATGC
D17S1303	CTCTCCAAGGCTCACTCAAA	TGGTCTTTTTCCATTCCAAA
ATA78D02	GGCAAGAGTTTTCCCATTTT	GGCCAAGAGAACTCACGTAA
D17S2196	CCAACATCTAGAATTAATCAGAATC	ATATTTCAATATTGTAACCAGTCCC
D17S841	TGGACTTTCTTACATGGCAG	AGGTTAGTAGTCTATGTCACAGCG
D17S1294	TGGCATGCAATTGTAGTCTC	TTCTTCCTTACTAAGTTGAGAACG
D17S798	CCATGAGAAAGTTGTTTAGTAGTC	TGTTCTTGGGAGTGCAG
D17S933	ACTCACTGGGGTCCTGG	TGTGGTTTCCTTATAGACTGTAGA
D17S1299	TAGCACTTGAGCACACATGG	GTGCATTATGGGGACCATTA
D17S806	AGCTTTACTGCACTAGCCCC	GGAGATTGGGATTTGGTGAC
D17S948	GTCTCTGTCCTTAGGGAGTTTA	TATCTTTGCCCATTTCTTG
D17S784	GAGTCTCCTAAATGCTGGG	AGCTCCTGCACAGTTCTTAAATA

B.2 Chromosome 14q microsatellite markers

Marker	Forward Primer	Reverse Primer
D14S1023	TGCATTTCCCGTAGACATT	GACTCTTGTAGTTCTTTGAAGCC
D14S972	TTAACGCATAACAGCCAAGA	TCTGACTGCCTCCATGA
ATA76C02	TGCCATGTCACTATGTCTGG	TGCTGTAAGCCATGATCTCA
D14S599	AACTTACTGGCGCAGATGAC	ACTCAGTGCAATATTTGCTGC
D14S1027	TCCCTTGATTTTTCAACG	AGTCAGTTGGGTTTGGC
D14S281	GAGCTAGGGATGCCTGCTACT	TGGCAGAATGTTAGACTGCTTCT
D14S980	CTGGGCAACAAGAGTGG	GAAGCGGGACAATTCTCTAAG
D14S274	TGAACTTTGGGCACCCT	TCTGACAAACCAGCAAATGA
D14S592	TTCCAGAGTATTTGCTTAAGAGG	GCATTGTGGGATGAGGTATG
D14S271	GGCAGTATTTTAAAAATTGTGGT	GAGANTCATCTGAGCCTGG
D14S588	GCCGAAAGAAAGAAAAAAGG	CGAATGCATACTTGCTGTTG
D14S71	TGCACCAATGCCTCCT	CCCGGCCAGAAATGCT
D14S974	CACACACATGCACATGC	GGCTAAAGTCTACGCCAC

B.3 Chromosome 5q Microsatellite Markers

Marker	Forward Primer	Reverse Primer
D5S1505	TAAGTGCCAGAGTCTCCCAC	TAAGGCATGTCTCGGAGCTA
D5S816	TTGCCACTGAAAATCATATCC	TCTACTTGGCATCCCTGATG
D5S658	TTTGAAGGGCACTACGAAGATCCTC	ATTCTCTATTGGAGCCAAGCCAAG
D5S436	TATGTGCCAGGCATTACGCT	GTCTCCACCCACAGCAGG
D5S673	GGTCTGAGTTCCACATCAT	TGGAGAGATAGTGTCCCG

B.3 Gene Polymorphism Primer Sequences (Association Studies)

Polymorphism	Forward Primer	Reverse Primer
Serotonin Transporter		
VNTR 5-HTTLPR 3'UTR	GTCAGTATCACAGGCTGCGAG GGCGTTGCCGCTCTGAATGC CCGCTTGAATGCTGTGTAACACAC	TGTTCCTAGTCTTACGCCAGTG GAGGGACTGAGCTGGACAACCAC GTACCCTTCCAATAATAACCTCC
COMT		
SNP in codon 108/158, exon4	ACTGTGGCTACTCAGCTGTG	CCTTTTTTCCAGGTCTGACAA
SNP at –287 bp in promoter	TAGTAACAGACTGGGCACGAA	GTTCAAAGGACATTTATCATG
SNP in codon 62, exon 3	TTGCTGGGCCTGGTGCTGCT	ACGTTCATGGCCC-ACTCCTCT
SNP in codon 136 exon 4	ACTGTGGCTACTCAGCTGTG	CCTTTTTCCAGGTCTGACAA
Ins/del, 3'UTR, exon 6	GACAACGTGATCTGCCCAGG	GAGGTGTGCTTTGCATTTAG
ERβ		
SNP at 1730bp in 3'UTR of exon 8	GACCTGCTGCTGGAGATGCT	AATGAGGGACCACACAGCA
GCH1		
SNP at –959bp in promoter	TCTTCACCAAGGGCTACCAG	AGTGAGGCAACTCCGGAAAC
SNP at +393bp in intron 1	TCAAGTGAGGAAAAAGGTCCA	TGATTTGCAAACGAAGATAGGA

B.4 Primer sequences for DHPLC and Sequence Analysis

Marker	Forward Primer	Reverse Primer
Exon1 (a)*	GGAGTTTAGCCGCAGACC	GTGAGGCAACTCCGGAAAC
Exon1 (b)	CCTACTCGTCCATCCTGAGC	GAGTTCCGGGGCTGAGAG
Exon2	TTTTAGGGAGGATTAACGTTCG	TCAGCAATTGGCAGCTAAAA
Exon3	AACAGTTCCAGATGTTTTCAAGG	CAGCAGATGAGGGCAGGT
Exon4	CCACTTGCTTCAACAATGTCC	CCTGGGTGACAGAGCAAGAC
Exon5	AGCTGGTGTGTCTTGGCTCT	AGGCTCASGGGATGGAAATCT
Exon6	CGATATACTTGGTAACTGTGAGC	CAAACAAGACCGGACAGACA
5'UTR (264bp)	CCGGGCCATAAAAAGGAG	GAACCCATTGCTGCACCT
Promoter 1	CTGGACGACTGCCTCTTTTC	GAGGTCTGCGGCTAAACTCC
Promoter 2	GCCTTGAAAGCTTAAGAGTTGG	CGCGCCTCCTTTTTATGG
Promoter 3*	ACCGACGGAAATGGATTTTA	GAAAAGAGGCAGTCGTCCAG
Promoter 4*	GGCATGGTGATGCACTCTTA	AGCCCACTTGCTTCAACAAT
Promoter 5	GAGGCAGGAGAATCGTTTGA	GTGGCTCACGCCTGTACTC

All primers for DHPLC analysis were designed using the web site
(www.genome.wi.mit.edu/cgi bin/primer/primer3_www.cgi)

An asterix (*) denotes primers used in sequencing reactions in addition to DHPLC analysis.