Identification and characterisation of rare \textit{CACNG5} genetic variants in bipolar disorder and schizophrenia

Thesis submitted for the degree of

Doctor of Philosophy

UCL

Yi-Chun Lin

Molecular Psychiatry Laboratory
University College London
2010 – 2015
I, Yi-Chun Lin, confirm that all the work presented in this thesis is my own.

I confirm that the information has been derived from other sources; it has been indicated in the thesis.
Acknowledgements

This thesis would not have been completed without the help and support of colleagues and friends to whom I would now like to express my sincere thanks.

First and foremost of my thanks goes to my supervisor Dr Andrew McQuillin, who gave me this wonderful and exciting opportunity to carry out this PhD. His encouragement has provided tremendous support throughout my PhD. I am extremely grateful for his insightful advice, patience and kindness.

I would also like to express my appreciation and thanks to Professor Hugh Gurling, my second supervisor, who provided further guidelines and support to enable me to complete this PhD research.

Very special thanks to Dr. Radhika Kandaswamy and Dr. Sally Sharp, who have been guiding me with infinite patience in my experimental studies. I am indebted to them for their enthusiasm and advice, their friendship and support leading to the completion of research experiments. Also I would like to thank all my colleagues from my lab Michael, John, Alessia for their support.

Finally, I would also like to thank my family, especially my parents for their financial support and my friends for their encouragement and support.
Abstract

Schizophrenia (SCZ) and bipolar disorder (BPD) are common, highly heritability psychiatric disorders. Genome-wide association studies have found evidence of shared genetic susceptibility to both diseases. The most notable example is CACNA1C which encodes the $\alpha_1$ subunit of L-type calcium channels. Several other calcium channel genes have also been implicated in BPD and/or SCZ and together there is support for a role for these genes in both diseases.

The primary function of several $\gamma$ subunit calcium channel genes appears to be the regulation of AMPA receptor localisation and function. Collectively these are known as Transmembrane AMPA receptor Regulatory Proteins (TARPs). This thesis aimed to identify disease relevant genetic variation in one such TARP, CACNG5, and to study the effect of these variants.

CACNG5 variants in the exons and promoter region were identified in 1098 BPD, 618 SCZ, and 1087 control individuals. Four novel non-synonymous SNPs (nsSNPs) and four nsSNPs were identified. Burden analysis of nsSNPs in BPD and SCZ found evidence for association ($p=0.0022$). This association was strengthened by inclusion of data from European samples in the 1000 Genomes project ($p=0.00057$). However, combined data with the UK10K and Swedish exome sequence studies founds a weakened association signal ($p=0.0082$). Functional analyses using co-expression of AMPAR2 and CACNG5 constructs containing the eight nsSNPs were used to analyse changes in the expression and/or trafficking of $\gamma_5$ and AMPA receptors. Four of the variants were associated with decreased AMPAR2 expression as a consequence of altered trafficking to the cell surface. V146M (identified in 2 SCZ patients) overexpression increased AMPAR2 trafficking to the cell surface ($p<0.005$); conversely, T164L (identified in one SCZ patient) overexpression decreased the expression.
of AMPAR2 and its cell surface trafficking ($p<0.05$). Our results suggest a role for CACNG5 variants in SCZ and/or BPD and that this may be mediated via dysregulation of AMPARs.
# Table of Contents

Acknowledgements ........................................................................................................... 3  
Abstract ............................................................................................................................ 4  
Table of Contents ................................................................................................................. 6  
Table of Tables .................................................................................................................. 11  
Table of Figures ............................................................................................................... 14  
1 Introduction ..................................................................................................................... 28  
  1.1 Psychiatric disorders ................................................................................................. 28  
      1.1.1 Overview ............................................................................................................ 28  
      1.1.2 History of psychiatric disorders ....................................................................... 29  
      1.1.3 Diagnosis of psychiatric disorders .................................................................. 31  
      1.1.3.1 Diagnosis of bipolar disorder ....................................................................... 33  
      1.1.4 Epidemiology similarities in BPD and SCZ .................................................... 37  
      1.1.5 Treatment ......................................................................................................... 40  
      1.1.6 Pathophysiology of BPD and SCZ .................................................................. 44  
      1.1.7 Heritability of BPD and SCZ ........................................................................... 46  
1.2 Molecular Genetics Studies ....................................................................................... 49  
  1.2.1 Factors affecting genetic studies .......................................................................... 51  
  1.2.2 Genetic Markers .................................................................................................... 53  
  1.2.3 Linkage analysis .................................................................................................... 54  
  1.2.4 Association studies .............................................................................................. 58  
  1.2.5 Summary ............................................................................................................. 62  
  1.2.6 Genome-wide association studies ....................................................................... 63  
  1.2.7 Chromosomal Rearrangements and CNV ............................................................ 74  
1.3 Voltage-gated calcium channel (VGCC) .................................................................... 78  
  1.3.1 Ca\textsuperscript{2+} channel γ subunits .................................................................. 80  
1.4 Glutamate receptors .................................................................................................... 85
1.4.1 AMPA receptors ................................................................. 87
1.4.2 Trafficking of AMPARs at synapses ........................................ 95
1.4.3 TARPs as auxiliary protein of AMPARs ...................................... 96
1.4.4 Trafficking of AMPARs associated with neuro-disorders .......... 106

2 Aims of the thesis ........................................................................ 107

3 MATERIALS AND METHODS .......................................................... 108
3.1 General Materials and Methods .................................................. 108
3.1.1 Research participants – UCL Research sample ......................... 108
3.1.2 DNA Quantification ................................................................ 109
3.2 Variant detection and genotyping ............................................... 110
3.2.1 High Resolution Melting Curve Genotyping Assay ..................... 111
3.2.2 Polymerase Chain Reaction (PCR) ............................................. 119
3.2.3 Agarose gel electrophoresis ..................................................... 124
3.2.4 PCR purification ..................................................................... 125
3.2.5 DNA sequencing ..................................................................... 127
3.2.6 KASPar Genotyping Assay ....................................................... 131
3.2.7 Association study – burden analysis ......................................... 142
3.2.8 Bioinformatical analysis .......................................................... 143
3.3 Preparation of CACNG5 and AMPAR2 Constructs ....................... 144
3.3.1 Plasmid vectors ...................................................................... 146
3.3.2 Sub-cloning of CACNG5 inserts into pCMV-Tag4 vectors via traditional restriction sites ................................................................. 154
3.3.3 Sub-cloning of CACNG5 inserts into pTagRFP-C vector via PCR amplification ................................................................. 158
3.3.4 Mutating the CACNG5 gene with each variant individually ........ 161
3.3.5 Large-scale preparation of plasmid DNA .................................... 163
3.3.6 Preparation of bacterial glycerol stocks for long-term storage of clones ...... 163
3.4 Functional characterisation of CACNG5 non-synonymous variants .......... 164
3.4.1 Animal cell-culture .......................................................... 164
3.4.2 Transfection ........................................................................ 170
3.4.3 Protein isolation ................................................................. 174
3.5 Functional characterisation of CACNG5 variants on surface expression of AMPAR
........................................................................................................ 181
  3.5.1 The use of fluorescent proteins ..................................... 181
  3.5.2 Fluorescence-activated cell sorting (FACS) .................... 181
  3.5.3 Cell surface biotinylation .................................................. 183
  3.5.4 Western Blotting .............................................................. 185
  3.5.5 Antibodies ........................................................................ 195
  3.5.6 Odyssey® Imager and analytical software ..................... 198
  3.5.7 Data analysis ..................................................................... 198
4 Identification of CACNG5 variants in cases of bipolar disorder and schizophrenia .... 200
  4.1 Introduction ......................................................................... 200
  4.2 Results ................................................................................ 206
    4.2.1 CACNG5 gene variants by high resolution melting analysis (HRMA) .... 206
    4.2.2 Non-synonymous variants .............................................. 207
    4.2.3 Synonymous variants ...................................................... 214
    4.2.4 Putative CACNG5 variant promoter variants .................. 214
    4.2.5 UK10K EXOME DATA .................................................. 215
    4.2.6 Swedish exome data analysis ......................................... 217
    4.2.7 Association analysis of CACNG5 variants ....................... 218
    4.2.8 Bioinformatics analysis – prediction of the effects of variant on protein function 225
  4.3 Discussion ........................................................................... 227
5 Rare mutations identified in CACNG5 from bipolar disorder and schizophrenia
individuals exhibit impaired expression of intracellular and cell surface AMPA receptors 233
  5.1 Background ........................................................................ 233
5.2 Aims and Objective .............................................................................................................. 236
5.3 Results ............................................................................................................................... 238
  5.3.1 TARP-γ5 increased cells expressing GluR2 ............................................................... 240
  5.3.2 TARP-γ5 increases GluR2 trafficking to the cell surface .............................................. 241
  5.3.3 TARP-γ5 variants significantly increase GluR2 cell surface expression ........ ....... 242
  5.3.4 The effect of TARP-γ5 on the proportion of cells expressing GluR1 ....................... 245
  5.3.5 Mutated TARP-γ5 decreases the percentage of cells expressing intracellular and cell surface GluR1 .............................................................................................................. 247
5.4 Discussion .......................................................................................................................... 249
6 Replicated study of rare mutations identified in CACNG5 from bipolar disorder and schizophrenia individuals exhibit impaired cell surface trafficking of AMPA receptors subunit 2 ........................................................................................................................................ 253
  6.1 Background ....................................................................................................................... 253
  6.2 Results ............................................................................................................................... 256
    6.2.1 TARP-γ5 increased trafficking of GluR2 to the cell surface ...................................... 256
    6.2.2 Rare TARP-γ5 variant R69W showed no significant effect on GluR2 expression ........................................................................................................................................... 261
    6.2.3 Rare variant R71H in TARP-γ5 showed no significant effect on GluR2 expression ........................................................................................................................................... 265
    6.2.4 Rare variant R127Q in TARP-γ5 showed no significant effect on GluR2 expression ........................................................................................................................................... 269
    6.2.5 Rare variant R128M in TARP-γ5 slightly reduced GluR2 expression ................. 273
    6.2.6 Rare variant I156F in TARP-γ5 significantly reduced GluR2 expression .............. 277
    6.2.7 Rare variant T164L in TARP-γ5 significantly reduced GluR2 expression .... 281
  6.3 Discussion .......................................................................................................................... 285
7 Conclusion ............................................................................................................................... 288
8 References ............................................................................................................................. 292
### Table of Tables

Table 1-1 A summary of the symptoms of bipolar disorder, schizophrenia or other relevant disorders

Table 1-2 Genome-Wide Association Study Findings for GWA significant genes in Schizophrenia and Bipolar Disorder from the PGC SCZ (Schizophrenia Working Group of the Psychiatric Genomics Consortium. 2014) summary data for 108 genome-wide significant loci

Table 3-1 Primer sets used for HRMA

Table 3-2 HRMA assay conditions for primer optimisation

Table 3-3 Roche HRM Mix assay conditions for primer optimisation

Table 3-4 Quanta Supermix HRMA assay conditions for primer optimisation

Table 3-5 LightScanner HRM Mix assay conditions for primer optimisation

Table 3-6 HRM cycling conditions

Table 3-7 PCR primer optimisation layout

Table 3-8 Reaction mix used for Big Dye sequencing reaction

Table 3-9 Reaction mix for KASPar assay

Table 3-10 Primer sets used for Kaspar genotyping

Table 3-11 Kaspar Optimisation layout

Table 3-12 Kaspar cycling conditions: Dual colour hydrolysis probe/UPL probe program on LC480

Table 3-13 Example of ns SNV alleles counts for pooled burden analysis

Table 3-14 A table corresponed to pCMV-Tag4 vector features and its postion lists

Table 3-15 Primer sets for CACNG5 constructs site-directed mutagenesis

Table 3-16 The PCR cycling conditions for CACNG5 constructs site-directed mutagenesis

Table 3-17 Optimisation protocol using varying ratios of FuGENE® 6 Transfection reagent to DNA

Table 3-18 An example of the standard protein concentration

Table 3-19 An example of the absorbance of the BSA protein standard

Table 3-20 The composition of separation solution

Table 3-21 The composition of 5% stacking solution

Table 4-1 Summary of variant screening result

Table 4-2 Summary of non-synonymous variants genotype counts together with data from the European samples in the 1K Genomes project
Table 4-3 Summary of synonymous variant genotype counts together with data from the European samples in the 1K Genomes project. ........................................................................................................ 211
Table 4-4 Summary of putative promoter variant genotype counts together with data from the European samples in the 1K Genomes project. ........................................................................................................ 213
Table 4-5 Summary of non-synonymous variants genotype counts from the UK10K exome sequence data. ........................................................................................................................................ 215
Table 4-6 Summary of synonymous variants genotype counts from the UK10K exome sequence data. ........................................................................................................................................ 216
Table 4-7 Summary of non-synonymous variants genotype counts from the Swedish exome sequence data. ........................................................................................................................................ 217
Table 4-8 Summary of synonymous variants genotype counts from the Swedish exome sequence data. ........................................................................................................................................ 218
Table 4-9 Summary of combined total 8 non-synonymous variants genotypic count and significant association p-values from the 1K Genome Project data, UCL samples, UK10K exome sequence data and Swedish exome data. ........................................................................................................ 221
Table 4-10 Summary of genotype counts and significant association p-values for combined analysis of the 5 synonymous variants from the 1K Genome Project data, the UCL samples, the UK10K exome sequence data and the Swedish exome data. ........................................................................................................ 223
Table 4-11 Summary of putative promoter variants genotype counts and allelic association analysis........................................................................................................................................ 225
Table 4-12 Summary of bioinformatic analysis of nsSNPs.................................................................................................................................................. 226
Table 6-1 Statistical results for GluR2 expression from western blot and biotinylation assay. ........................................................................................................................................ 259
Table 6-2 Statistical results for TARP-γ5 expression in the cells with or without cell surface biotinylation assay ........................................................................................................................................ 260
Table 6-3 Statistical results for GluR2 expression from western blot and cell surface biotinylation assay. ........................................................................................................................................ 262
Table 6-4 Statistical results for TARP-γ5 expression in the cells with or without cell surface biotinylation assay. ........................................................................................................................................ 264
Table 6-5 Statistical results for GluR2 expression in the cells with or without cell surface biotinylation assay. ........................................................................................................................................ 266
Table 6-6 Statistical results for TARP-γ5 expression in the cells with or without cell surface biotinylation assay. ........................................................................................................................................ 268
Table 6-7 Statistical analysis of GluR2 expression in HEK293 cells with or without cell surface biotinylation before assay ................................................................. 270
Table 6-8 Statistical results for TARP-γ5 expression in the cells with or without cell surface biotinylation assay .................................................................................. 272
Table 6-9 Statistical results for GluR2 expression in the cells with or without cell surface biotinylation assay ................................................................. 274
Table 6-10 Statistical results for TARP-γ5 expression in the cells with or without cell surface biotinylation assay .................................................................................. 276
Table 6-11 Statistical results for GluR2 expression in the cells with or without cell surface biotinylation assay ................................................................. 278
Table 6-12 Statistical results for TARP-γ5 expression in the cells with or without cell surface biotinylation assay .................................................................................. 280
Table 6-13 Statistical results for GluR2 expression in the cells with or without cell surface biotinylation assay ................................................................. 282
Table 6-14 Statistical results for TARP-γ5 expression in the cells with or without cell surface biotinylation assay .................................................................................. 284
Table 6-15 Summary of TARP-γ5 and its variant effect results on GluR2 and TARP-γ5 expression. ................................................................................................. 285
Table 11-1 Raw data of protein expression levels ...................................................... 380
Table of Figures

Figure 1-1 Association marker rs17645023 lies 36kb from CACNG5 and 79kb from CACNG4 taken from UCSC Genome Browser ................................................................. 71
Figure 1-2 Structure of voltage-gated calcium channel α1 subunits and its subclasses. .... 79
Figure 1-3 Schematic diagrams showing predicted membrane topology and putative functional sites on γs and claudins. ........................................................................ 82
Figure 1-4 The protein structures of type I and type II TARPs. ..................................... 84
Figure 1-5 Synaptic transmission and AMPAR trafficking ........................................... 87
Figure 1-6 The structure of AMPA receptors ............................................................ 89
Figure 1-7 Structure of AMPAR subunit and its direct interacting proteins. ............... 91
Figure 1-8 A linear sequence of AMPA receptor and its formation into the cell membrane. 92
Figure 1-9 Arrangement and sequence of the alternative RNA editing to form flip and flop. 93
Figure 1-10 Ca\textsuperscript{2+} permeability of AMPARs depends on the subunit composition .... 95
Figure 1-11 TARPs generate export-competent receptors to existing AMPAR leading to synaptic trafficking ......................................................................................... 98
Figure 1-12 Model of AMPAR stabilisation via phosphorylation of TARPs by CaMKII .... 104
Figure 1-13 Lipid bilayers enhance translocation of AMPAR-TARP complexes to synapses. ................................................................................................................. 105

Figure 3-1 Examples of DNA concentration calibration from the Qubit® dsDNA assay. .. 109
Figure 3-2 An example of HRMA reveals differences between wild-type (homozygous) and variants (heterozygous) ................................................................................................. 118
Figure 3-3 DNA sequencing cycle .............................................................................. 127
Figure 3-4 Schematic showed a cycle procedure of BigDye terminator cycle sequencing.. 128
Figure 3-5 Diagram of the Applied Biosystems 3730xl DNA Analyzer run cycle steps .... 130
Figure 3-6 Diagrammatic representation of KASP reaction process ............................. 133
Figure 3-7 An example of the endpoint analysis of rs41280112 performed after KASP PCR reaction .................................................................................................................. 139
Figure 3-8 Basic procedure for DNA (gene) cloning in a plasmid vector ...................... 145
Figure 3-9 Circular map of the pCR4-TOPO vector ..................................................... 147
Figure 3-10 Circular map of the pCMV-Tag4 vector .................................................... 149
Figure 3-11 Circular map of the pCMV-Sport6 vector .................................................. 150
Figure 3-12 Circular map of the pCI vector ................................................................. 151
Figure 3-13 Circular map of the pTagRFP-C vector ..................................................... 154
Figure 3-14 Restriction digestion of both pCMV-Tag4 and pCR4-TOPO for subcloning. 155
Figure 3-15 Identification of pCMV-Tag4 clones containing 926 bp CACNG5 cDNA by colony PCR. 158
Figure 3-16 Verification of amplified fragment containing 894 bp CACNG5 insert. 160
Figure 3-17 A diagram of the improved Neubauer haemocytometer. 167
Figure 3-18 An example of the standard curve produced from the protein assay data. 177
Figure 3-19 An example of the Qubit® 2.0 Fluorometer calculates standard curve and sample concentrations. 179
Figure 3-20 Glass plate assemble for gel casting. 187
Figure 3-21 Transfer assembly for a tank transfer system. 192
Figure 4-1 LD plot of the marker rs17645023 between CACNG5 and CACNG4 genes obtained from the USCS Genome Brower website. 202
Figure 4-2 regional association of PGC BPD GWAS plot for the marker rs17645023 analysed at CACNG5 and CACNG4. 203
Figure 4-3 regional association of PGC SCZ GWAS plot for the marker rs17645023 analysed at CACNG5 and CACNG4. 204
Figure 4-4 Detection of SNPs in exon 2 of the CACNG5 gene by HRMA. 208
Figure 4-5 Validation of SNPs found in exon 2 of the CACNG5 gene by HRM analysis. 209
Figure 4-6 Evolutionary conservation of the V146 region of CACNG5. 228
Figure 4-7 Analysis of variants in the extracellular loop of CACNG5. 229
Figure 4-8 Evolutionary conservation of the R127 and T128 region of CACNG5. 230
Figure 4-9 Schematic diagrams showing predicted membrane topology and PDZ binding site motif on both types of TARP. 232
Figure 5-1 The migration of SEP-GluR2 through the compartments of the secretory pathway. 237
Figure 5-2 Typical histogram plots of GFP intensity from transfected cells generated by Kaluza software. 239
Figure 5-3 Percentage of cells expressing GluR2 was compared with cotransfected wild-type TARP-γ5 and control vector. 240
Figure 5-4 The percentage of cells expressing GluR2 on their surface was compared with co-expressed wild-type TARP-γ5 and control vector. 241
Figure 5-5 Percentage of cells expressing GluR2 was compared after co-transfection with wild-type mutated TARP-γ5. 242
Figure 5-6 Percentage of cells expressing surface GluR2 was compared after co-transfection with wild-type and mutated TARP-γ5. ................................................................. 243
Figure 5-7 The percentage of cells expressing GluR1 was compared after co-transfection with wild-type TARP-γ5 and a control vector ................................................................. 245
Figure 5-8 The percentage of cells expressing surface GluR1 was compared after co-transfection with wild-type TARP-γ5 and a control vector ................................................................. 246
Figure 5-9 The percentage of cells expressing GluR1 was compared after co-transfection with wild-type and mutated TARP-γ5. ................................................................. 247
Figure 5-10 The percentage of cells expressing GluR1 on their surface was compared after co-transfection with wild-type and mutated TARP-γ5. ................................................................. 248
Figure 6-1 Quantification of GluR2 in HEK293 when co-transfected with TARP-γ5 or a control plasmid ................................................................................................................. 258
Figure 6-2 Quantification of TARP-γ5 in HEK293 cells when co-transfected with GluR2 (n = 3) ......................................................................................................................... 260
Figure 6-3 Quantification of GluR2 in HEK293 cells when co-transfected with wild-type TARP-γ5 or mutated TARP-γ5 (n = 3). ................................................................. 261
Figure 6-4 Comparison of wild type TARP-γ5 and its carried R69W variant on their expression in HEK293 cells (n = 3) ......................................................................................... 263
Figure 6-5 Quantification of GluR2 in HEK293 cells when co-transfected with wild-type TARP-γ5 or mutated TARP-γ5 (n = 3). ................................................................. 265
Figure 6-6 Comparison of wild type TARP-γ5 and its carried R71H variant on their expression in HEK293 cells (n = 3) ......................................................................................... 267
Figure 6-7 Quantification of GluR2 in HEK293 cells when co-transfected with wild-type TARP-γ5 or mutated TARP-γ5 (n = 3). ................................................................. 269
Figure 6-8 Comparison of wild type TARP-γ5 and its carried R127Q variant on their expression in HEK293 cells (n = 3) ......................................................................................... 271
Figure 6-9 Quantification of GluR2 in HEK293 cells when co-transfected with wild-type TARP-γ5 or mutated TARP-γ5 (n = 3) ......................................................................................... 273
Figure 6-10 Comparison of wild type TARP-γ5 and its carried R128M variant on their expression in HEK293 cells (n = 3) ......................................................................................... 275
Figure 6-11 Quantification of GluR2 in HEK293 cells when co-transfected with wild-type TARP-γ5 or mutated TARP-γ5 (n = 3). ................................................................. 277
Figure 6-12 Comparison of wild type TARP-γ5 and the I156F variant on expression in HEK293 cells (n = 3). .................................................................................................................................................. 279
Figure 6-13 Quantification of GluR2 in HEK293 cells when co-transfected with wild-type TARP-γ5 or mutated TARP-γ5 (n = 3). .................................................................................................................................................. 281
Figure 6-14 Comparison of wild type TARP-γ5 and its carried T164L variant on their expression in HEK293 cells (n = 3). .................................................................................................................................................. 283
Figure 9-1 Transduction of mammalian cells using an adenoviral vector ......................... 368
Figure 9-2 Transfection of eukaryotic cells with plasmid ....................................................... 369
Figure 9-3 The best result for transfection was achieved in presence of 9 µl FuGENE6-HD. .................................................................................................................................................. 371
Figure 9-4 The best result for transfection was achieved in presence of 8 µl
Lipofectamine™2000. ................................................................................................................. 373
Figure 10-1 Genotyping plots of promoter variants in CACNG5 including rs3760263,
rs181400884, and rs75486725 are shown .................................................................................. 376
Figure 10-2 Genotyping plots of synonymous variants, R6R, C19C, A27A, and L34L found in CACNG5 are shown .......................................................................................................................... 377
Figure 10-3 Genotyping plots of non-synonymous variants, R69W, R71H, R127Q, T128M
found in CACNG5 are shown .................................................................................................... 378
Figure 10-4 Genotyping plots of non-synonymous variants, V146M, I156F, T164L, H233Y
found in CACNG5 are shown .................................................................................................... 379
<table>
<thead>
<tr>
<th>Abbreviations</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>%</td>
<td>Percentage</td>
</tr>
<tr>
<td>°C</td>
<td>Celsius</td>
</tr>
<tr>
<td>1K</td>
<td>1,000</td>
</tr>
<tr>
<td>5-HT</td>
<td>5-hydroxytryptamine</td>
</tr>
<tr>
<td>bp</td>
<td>Basepair</td>
</tr>
<tr>
<td>kb</td>
<td>Kilobase</td>
</tr>
<tr>
<td>kD</td>
<td>Kilodalton</td>
</tr>
<tr>
<td>mg</td>
<td>Milligram</td>
</tr>
<tr>
<td>ml</td>
<td>Millilitre</td>
</tr>
<tr>
<td>µl</td>
<td>Microlitre</td>
</tr>
<tr>
<td>µg</td>
<td>Microgram</td>
</tr>
<tr>
<td>ng</td>
<td>Nanogram</td>
</tr>
<tr>
<td>nM</td>
<td>Nanomolar</td>
</tr>
<tr>
<td>pg</td>
<td>Picogram</td>
</tr>
<tr>
<td>w/v</td>
<td>Weight/volume</td>
</tr>
<tr>
<td>α</td>
<td>Alpha</td>
</tr>
<tr>
<td>α2δ</td>
<td>Alpha-2/delta</td>
</tr>
<tr>
<td>X²</td>
<td>Chi square</td>
</tr>
</tbody>
</table>
A.D.  anno Domini

AD  Alzheimer’s disease

ADAR2  Adenosine deaminase acting on RNA 2

ADHD  Attention deficit hyperactivity disorder

ALSPAC  Avon Longitudinal Study of Parents and Children

AMPA  α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid

AMPAR  α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor

AMPARs  α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptors

APA  American Psychiatric Association

B.C.  Before Christ

β  Beta

BP-I  Bipolar I disorder

BP-II  Bipolar II disorder

BPAD  Bipolar affective disorder

BPD  Bipolar disorder

BP-NOS  Bipolar disorder not otherwise specified

BR  Broad rang

BSA  Bovine serum albumin

C-  Carboxyl-
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA</td>
<td><em>Cornu Ammonis</em></td>
</tr>
<tr>
<td>Ca</td>
<td>Calcium</td>
</tr>
<tr>
<td>Ca&lt;sup&gt;2+&lt;/sup&gt;</td>
<td>Calcium ion</td>
</tr>
<tr>
<td>CaMKII</td>
<td>Calcium-/calmodulin-dependent protein kinase II</td>
</tr>
<tr>
<td>cAMP</td>
<td>Cyclic adenosine monophosphate</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary DNA</td>
</tr>
<tr>
<td>cGMP</td>
<td>Cyclic guanosine monophosphate</td>
</tr>
<tr>
<td>CMV</td>
<td>Cytomegalovirus</td>
</tr>
<tr>
<td>CNIHs</td>
<td>Cornichon-like proteins</td>
</tr>
<tr>
<td>CNS</td>
<td>Central nervous system</td>
</tr>
<tr>
<td>CNV</td>
<td>Copy number variation</td>
</tr>
<tr>
<td>CNVs</td>
<td>Copy number variants</td>
</tr>
<tr>
<td>CTD</td>
<td>Carboxyl terminal domain</td>
</tr>
<tr>
<td>D2</td>
<td>Dopamine 2</td>
</tr>
<tr>
<td>DAC</td>
<td>Data access committee</td>
</tr>
<tr>
<td>DGS</td>
<td>Di George syndrome</td>
</tr>
<tr>
<td>DMEM</td>
<td>Dulbecco’s Modified Eagle’s Medium</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl sulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>ddNTP</td>
<td>dideoxynucleotides</td>
</tr>
<tr>
<td>dNTP</td>
<td>Deoxynucleoside triphosphate</td>
</tr>
<tr>
<td>dsDNA</td>
<td>double-stranded deoxyribonucleic acid</td>
</tr>
<tr>
<td>DSM</td>
<td>Diagnostic and Statistical Manual of Mental Disorders</td>
</tr>
<tr>
<td>DSM-IV</td>
<td>Diagnostic and Statistical Manual of Mental Disorders 4th edition</td>
</tr>
<tr>
<td>DSM-V</td>
<td>Diagnostic and Statistical Manual of Mental Disorders 5th edition</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>DZ</td>
<td>Dizygotic</td>
</tr>
<tr>
<td>ECACC</td>
<td>European Collection of Animal Cell Cultures</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EGA</td>
<td>European Genome-Phenotype Archive</td>
</tr>
<tr>
<td>EP</td>
<td>Ecliptic pHluorin</td>
</tr>
<tr>
<td>EPSCs</td>
<td>excitatory post-synaptic currents</td>
</tr>
<tr>
<td>ER</td>
<td>Endoplasmic reticulum</td>
</tr>
<tr>
<td>EU</td>
<td>European</td>
</tr>
<tr>
<td>EX1</td>
<td>First extracellular domain</td>
</tr>
<tr>
<td>F/F</td>
<td>Flip/flop</td>
</tr>
<tr>
<td>FACS</td>
<td>Fluorescence-activated cell sorting</td>
</tr>
<tr>
<td>FBS</td>
<td>Fetal bovine serum</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>FP</td>
<td>Fluorescent proteins</td>
</tr>
<tr>
<td>FRET</td>
<td>Forster resonance energy transfer</td>
</tr>
<tr>
<td>γ</td>
<td>Gamma</td>
</tr>
<tr>
<td>GABA</td>
<td>Gamma-aminobutyric acid</td>
</tr>
<tr>
<td>GABA&lt;sub&gt;B&lt;/sub&gt;</td>
<td>Gamma-aminobutyric acid B subunit</td>
</tr>
<tr>
<td>GFP</td>
<td>Green fluorescence protein</td>
</tr>
<tr>
<td>GK</td>
<td>Guanylate kinase-like</td>
</tr>
<tr>
<td>GluRs</td>
<td>Glutamate receptors</td>
</tr>
<tr>
<td>GQ</td>
<td>Genotype quality</td>
</tr>
<tr>
<td>GW</td>
<td>Genome-wide</td>
</tr>
<tr>
<td>GWA</td>
<td>Genome-wide association</td>
</tr>
<tr>
<td>GWAS</td>
<td>Genome-wide association study</td>
</tr>
<tr>
<td>GWASs</td>
<td>Genome-wide association studies</td>
</tr>
<tr>
<td>HCl</td>
<td>Hydrogen chloride</td>
</tr>
<tr>
<td>HEK293</td>
<td>Human embryonic kidney 293 cells</td>
</tr>
<tr>
<td>HPLC</td>
<td>High-performance liquid chromatography</td>
</tr>
<tr>
<td>HRM</td>
<td>High resolution melting</td>
</tr>
<tr>
<td>HRMA</td>
<td>High resolution melting analysis</td>
</tr>
<tr>
<td>HS</td>
<td>High sensitivity</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>HS-TK</td>
<td>Herpes simplex virus thymidine kinase</td>
</tr>
<tr>
<td>ICD</td>
<td>International Classification of Disease</td>
</tr>
<tr>
<td>ICD</td>
<td>International Classification of Disease, 10th revision</td>
</tr>
<tr>
<td>iGluRs</td>
<td>Ionotropic glutamate receptors</td>
</tr>
<tr>
<td>$K^+$</td>
<td>Potassium ion</td>
</tr>
<tr>
<td>KA</td>
<td>kainite</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani</td>
</tr>
<tr>
<td>LBD</td>
<td>Ligand binding domain</td>
</tr>
<tr>
<td>LD</td>
<td>Linkage disequilibrium</td>
</tr>
<tr>
<td>LTD</td>
<td>Long-term depression</td>
</tr>
<tr>
<td>LTP</td>
<td>Long-term potentiation</td>
</tr>
<tr>
<td>M4</td>
<td>Fourth membrane spanning segment</td>
</tr>
<tr>
<td>MAGUKs</td>
<td>Membrane associated guanylate kinases</td>
</tr>
<tr>
<td>MAOI</td>
<td>Monoamine oxidase inhibitors</td>
</tr>
<tr>
<td>MAP1-LC2</td>
<td>Microtubule associated protein 1 light chain 2</td>
</tr>
<tr>
<td>MCS</td>
<td>Multiple cloning site</td>
</tr>
<tr>
<td>mEPSCs</td>
<td>miniature excitatory post-synaptic currents</td>
</tr>
<tr>
<td>$\text{Mg}^{2+}$</td>
<td>Magnesium ion</td>
</tr>
<tr>
<td>$\text{MgCl}_2$</td>
<td>Magnesium chloride</td>
</tr>
</tbody>
</table>
mGluRs Metabotropic glutamate receptors
mRNAs Messenger ribonucleic acids
MZ Monozygotic
N- Amino-
Na+ Sodium ion
NH4+ Ammonium
NARI Noradrenaline reuptake inhibitors
Netos Neurophilin and Tolloid like protein
NCS-R National Comorbidity Survey Replication
NHGRI National Human Genome Research Institute
NHS National Health Services
NICE National Institute for Health and Clinical Excellence
NIR Near-infrared
NMDA N-methyl-D-aspartate
NMDAR N-methyl-D-aspartate receptor
nPIST Neuronal PDZ domain protein interacting specifically with TC10
NTD Amino-terminal domain
OFC Olanzapine-fluoxetine combination
OPCRIT Operational Criteria for Psychotic Illness
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBS</td>
<td>Phosphate buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PGC</td>
<td>Psychiatric Genomics Consortium</td>
</tr>
<tr>
<td>PSD</td>
<td>Postsynaptic density</td>
</tr>
<tr>
<td>Q</td>
<td>Glutamine</td>
</tr>
<tr>
<td>Q/R</td>
<td>Glutamine/arginine</td>
</tr>
<tr>
<td>R/G</td>
<td>Arginine/glycine</td>
</tr>
<tr>
<td>RDC</td>
<td>Research Diagnostic Criteria</td>
</tr>
<tr>
<td>RFP</td>
<td>Red fluorescent tag protein</td>
</tr>
<tr>
<td>RIPA</td>
<td>Radio-immunoprecipitation assay buffer</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RS</td>
<td>Arginine/serine</td>
</tr>
<tr>
<td>SADS-L</td>
<td>Schizophrenia and Affective Disorders Schedule</td>
</tr>
<tr>
<td>SCZ</td>
<td>Schizophrenia</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulphate</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium dodecyl sulphate – polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>SEP</td>
<td>Super-ecliptic pFluorin</td>
</tr>
<tr>
<td>SH3</td>
<td>src homology 3</td>
</tr>
<tr>
<td>SMR</td>
<td>Standardised mortality ratio</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>SNP</td>
<td>Single nucleotide polymorphism</td>
</tr>
<tr>
<td>SNPs</td>
<td>Single nucleotide polymorphisms</td>
</tr>
<tr>
<td>SNRI</td>
<td>Serotonin-noradrenaline reuptake inhibitors</td>
</tr>
<tr>
<td>SNV</td>
<td>Single nucleotide variant</td>
</tr>
<tr>
<td>SRP</td>
<td>Single recognition particle</td>
</tr>
<tr>
<td>ssDNA</td>
<td>Single strand deoxyribonucleic acid</td>
</tr>
<tr>
<td>SSRI</td>
<td>Selective serotonin reuptake inhibitors</td>
</tr>
<tr>
<td>SSRs</td>
<td>Simple sequence repeats</td>
</tr>
<tr>
<td>STEP-BD</td>
<td>Systematic Treatment Enhancement Program for Bipolar Disorder</td>
</tr>
<tr>
<td>STRs</td>
<td>Short tandem repeats</td>
</tr>
<tr>
<td>SV40</td>
<td>Simian virus 40</td>
</tr>
<tr>
<td>Tm</td>
<td>Temperature</td>
</tr>
<tr>
<td>TARP</td>
<td>Transmembrane AMPAR regulatory protein</td>
</tr>
<tr>
<td>TARPs</td>
<td>Transmembrane AMPAR regulatory proteins</td>
</tr>
<tr>
<td>TBE</td>
<td>Tris-borate-EDTA</td>
</tr>
<tr>
<td>TBS</td>
<td>Tris-buffered saline</td>
</tr>
<tr>
<td>TMD</td>
<td>Transmembrane domain</td>
</tr>
<tr>
<td>UCL</td>
<td>University College London</td>
</tr>
<tr>
<td>UK</td>
<td>United Kingdom</td>
</tr>
</tbody>
</table>
US: United State
UV: Ultraviolet
V: Volt
VCF: Variant call format
VCFS: Velco-cardio-facial syndrome
VDCC: Voltage-dependent calcium channel
VGCC: Voltage-gated calcium channel
VGCCs: Voltage-gated calcium channels
WHO: World Health Organisation
WIBR: Wolfson Institute for Biomedical Research
WTCCC: Wellcome Trust Case Control Consortium
Zn$^{2+}$: Zinc ion
1 Introduction

1.1 Psychiatric disorders

1.1.1 Overview

Schizophrenia (SCZ) and bipolar disorder (BPD) are serious psychiatric disorders, which are characterised by impaired thinking, emotions, and behaviours over a period of time. These disorders can cause extensive disability leading to impairments in social and ordinary life. People suffering with these illnesses are common in the population (Kessler et al. 2007) and have significantly increased morbidity and mortality due to physical illness compared to the general population (Whiteford et al. 2013). The National Comorbidity and worldwide surveys reported a lifetime prevalence of 4.5 – 4.8% for BPD (Merikangas et al. 2007, Merikangas et al. 2011). The worldwide prevalence for SCZ is approximately 0.4 – 1 % (Shivashankar et al. 2013). The combined lifetime prevalence rates of both disorders can be up to 3% with heritability rates of up to 80% (Nierenberg et al. 2010). Because of their significant impacts on the patient, family, and society, BPD and SCZ also lead to a significant economic burden (Kleinman et al. 2003, Wyatt et al. 1995). Effective management and treatment of these illnesses depends on early and correct diagnosis.

Over the decades, psychiatric classification and definition systems for the diagnosis of mental disorders have included the International Classification of Disease (ICD) produced by the World Health Organisation (WHO) since 1949 (World Health Organisation 1992a) and the Diagnostic and Statistical Manual of Mental Disorders (DSM) produced by the American Psychiatric Association (APA) since 1952 (American Psychiatric Association 2013). These classification systems have been developed and revised in order to improve the reliability of clinical diagnosis to inform treatment strategies and to guide research. They define the characteristics of SCZ as marked by positive, negative, and disorganisation symptoms;
whereas BPD is marked by episodes of depression alternating with mania and/or hypomania positive (psychotic) symptoms. The positive features of the disorders share characteristics in BPD and SCZ, and they may include mood and psychotic symptoms with varying degrees of admixture. Both are potentially severe neuropsychiatric conditions, the aetiology of which is poorly understood. Genetic studies may be particularly fruitful endeavours in such conditions, as the underlying aetiology might be illuminated by a detailed knowledge of risk alleles. Furthermore, both disorders are diagnosed solely on the basis of behavioural signs and symptoms as neither currently has a specific, clearly defined, pathophysiological foundation on which to base diagnostic tests. Because genetic studies inherently examine biological markers, they may provide clues in the development of such tests, which may be clinically useful in risk prediction.

1.1.2 History of psychiatric disorders

Early written references from Chinese, Egyptian, Hebrew, and Greek refer to individual abnormal symptoms or behaviour and attribute these to a demon or god who had taken possession of a person. If a person’s speech or behaviour appeared to have a religious or mystical significance, it was thought that he or she was possessed by a good spirit or god; whereas, if a person became excited or overactive and engaged in behaviour contrary to religious teachings, it was considered to be the work of an angry god or an evil spirit, which represented the wrath and punishment of God. It was not until the 4th century B.C, when a Greek physician Hippocrates denied the theory of deities and demons involved in the development of mental illnesses; instead he believed that the brain is the central organ of intellectual activity and injuries to the head could cause sensory and motor disorders. From clinical observation Hippocrates invented the first classification of mental disorders, dividing them into three general categories: mania, melancholia, and paranoia (Angst and Marneros 2001).
1.1.2.1 History of the concept of bipolar disorder

In the first or the second Century AD, another Greek physician Areteus of Cappadocia considered mania and melancholia as one single disorder, which was the first time that these two symptoms were linked together (Yutzy et al. 2012). During the seventeenth and eighteenth centuries, psychiatrists continued to study the association between mania and melancholia until the nineteenth century when the modern term of bipolar disorder was adopted. In 1854, a founder of German scientific psychiatrist, Wihelm Griesinger believed that the disorder was a circle of melancholia to mania with regular changes, and also further created seasonal affective disorders: melancholia usually occurs in autumn and winter, whereas mania occurs in spring. In the middle of the 19th century, Jean-Pierre Falret finally drew the conclusion that bipolar disorder is an entity of its own, based on his long-term observations. In his published statements, he named “folie circulaire”, which is characterised by a continuous cycle of melancholia, mania and the interval in between as an independent disease on its own (Angst and Marneros 2001). Later in 1899, a German psychiatrist, Emil Kraepelin characterised the major endogenous psychoses into the disease concepts of dementia praecox, which was reformulated as schizophrenia (SCZ) in 1911, and manic-depressive psychosis (Zivanovic and Nedic 2012), which has now been reconceived as bipolar disorder (BPD). These clinical definitions of Kraepelinian theory brought a revolution of psychotic characterisation in the 20th century, where a term dichotomy of psychotic disorders was introduced comprising of BPD on the one end and SCZ at the other. The term bipolar was actually invented by Leonhard in 1957, which was described for the patients who experienced both depression and mania, and it was further differentiated into bipolar I (BP-I) and bipolar II (BP-II) disorder based on DSM-IV classification (American Psychiatric Association 1980). People with BP-II (often known as hypomania) have fewer and less severe symptoms that usually do not require hospitalisation compared to those with BP-I (known as mania).
1.1.2.2 History of the concept of schizophrenia

In the middle of 19th century, a French psychiatrist, Benedict Morel used the phrase “démence-précoce” to describe a characteristic of bizarre behaviour and abnormal mental function with an early-onset deteriorating state. He believed that mental degeneration with acute episodes of madness begins in the young and this essential which is now termed SCZ. Later, Emil Kraepelin used the term “dementia praecox” to refer the “sub-acute development of mental weakness occurring at the early age”. It was not until 1911 that a Swiss psychiatrist, Eugen Bleuler effectively renamed dementia praecox as schizophrenia, and believed that the condition was characterised primarily by disorganisation of thought processes, a lack of coherence between thought and emotion, and an inward orientation away from reality. SCZ was also thought to have a split within the intellect, between intellect and emotion, and between the intellect and external reality. Bleuler believed SCZ was not a single diagnostic entity; hence the two general symptom patterns or syndromes of SCZ have been differentiated as positive- and negative syndromes (Andreasen 1985, Andreasen 1995). Although the Kraepelinian dichotomy suggested that dementia praecox and manic-depressive psychoses had specific and separate causes; Jacob Kasanin coined the term schizoaffective psychosis to refer to a disorder with mixed features of schizophrenia and affective disorder in 1933 (Kasanin 1994).

1.1.3 Diagnosis of psychiatric disorders

The clinical diagnosis of psychiatric disorders are based on the presenting symptoms, physical and mental status examinations, corroborative data obtained from interviews with the patient, family members, and other relevant sources of information as well as observations of the patient’s behaviour during hospitalisation. Laboratory or psychometric testing is warranted in certain circumstances. Currently, the definition and classification systems in use for the diagnosis of mental illness are in the tenth version of the international
statistical classification of disease and related health problems (ICD-10) (World Health Organisation 1992b) and the fifth edition of the diagnostic and statistical manual of mental disorders (DSM-V) (American Psychiatric Association 2013). A diagnostic summary of bipolar disorder and schizophrenia is described below and is also shown in Table 1-1.

Table 1-1 A summary of the symptoms of bipolar disorder, schizophrenia or other relevant disorders

<table>
<thead>
<tr>
<th>Disorder</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bipolar disorder</td>
<td>An episodic recurrent pathological disturbance in mood ranging from extreme elation or mania to severe depression, which is usually accompanied by disturbances in thinking and behaviour; psychotic features (delusions and hallucinations) often occur.</td>
</tr>
<tr>
<td>Bipolar I disorder</td>
<td>A type of bipolar disorder in which severe episodes of mania (energised state) occur.</td>
</tr>
<tr>
<td>Bipolar II disorder</td>
<td>A type of bipolar disorder in which no severe episodes of mania (energised state) occur but in which milder episodes, known as hypomania, occur.</td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>A chronic disorder in which a wide variety of clinical features can occur and which is usually characterised by deterioration in social functioning and impaired cognitive function; delusions and hallucinations are often prominent and, although mood disturbance is common, it is not the most prominent feature.</td>
</tr>
<tr>
<td>Schizophrenia disorder, bipolar type</td>
<td>An illness in which prominent features of both bipolar disorder and schizophrenia occur.</td>
</tr>
<tr>
<td>Unipolar (major depression)</td>
<td>A mood disorder in which only episodes of depressed mood occur; psychotic features can occur but are less common than for bipolar disorder.</td>
</tr>
</tbody>
</table>
1.1.3.1 Diagnosis of bipolar disorder

Bipolar disorder (BPD) is classified into bipolar I disorder (BP-I) and bipolar II disorder (BP-II) according to DSM-V classification system; whereas there is no such distinction in the ICD-10 classification system. BP-I requires at least one manic episode to be diagnosed; whereas the BP-II requires experience only milder hypomanic episodes, as well as major depressive episodes (Akiskal and Benazzi 2005). In a manic episode the patient has markedly elevated mood, and this is often interrupted by occasional outbursts of intense irritability or even violence. For a diagnosis of as BPD, the duration of a manic episode must persist for at least one week. In addition, three or more additional symptoms must occur in the same time period, such as grandiosity, decreased need for sleep, pressure of speech, flight of ideas, distractibility, hyperactivity, increased sexual activity and impulsive behaviour. Psychotic symptoms, such as delusions and hallucinations, may or may not need to be present in a manic episode. Depressive episodes are characterised by depressed mood, loss of interest, decreased sexual interest, accompanied by reduced self-confidence, decreased energy and feelings of worthlessness, which should markedly depressed for most of every day and for most days for at least 2 weeks to be diagnosed. Hypomanic episodes occur when a patient experiences abnormally elevated, expansive, or irritable mood for at least 4 days, and also has at least three other symptoms of mania but to a lesser degree (American Psychiatric Association 2013). The ICD-10 classification system defined characteristics of BPD are that the individual has to experience a manic episode along with another mood disturbance (WHO 1993). 40% of patients who experience depression and only mild or short-lived hypomanic episodes that do not quite meet the threshold criteria for a diagnosis of BP-I or BP-II are diagnosed as having unipolar depression instead. The occurrence of these episodes may last from one week to several months varying across different patients. In addition a patient with BPD may suffer from many depressive episodes before experiencing their first episode of mania (Akiskal 2005). The course of illness is generally variable among
individuals and the frequency of depressive episodes tends to be higher than manic episodes. Furthermore, researchers have also reported a seasonal pattern of episodes occurring, where manic episodes more likely to occur in the spring or early summer and depressive ones in the autumn or winter (American Psychiatric Association 2013).

Other characteristics of BPD such as cyclothymia, mixed episodes, rapid cycling, schizoaffective disorder and seasonal affective disorders are defined as bipolar disorder not otherwise specified (BP-NOS). In DSM-V, cyclothymia is defined as a less serious version of full-blown BPD because it is minus certain extreme symptoms and psychotic features such as delusions and the marked impairment caused by full-blown manic or major depressive episodes. If the person is subject to cyclical mood changes that are less severe than the mood swings seen in BPD and persist for at least a year, he or she may receive a diagnosis of cyclothymic disorder. Mixed episode is characterised by symptoms of both full-blown manic and depressive episodes nearly every day over at least one week, whether the symptoms are intermixed or alternate rapidly every few days. Additionally, psychotic features may also be present during mixed episodes (American Psychiatric Association 2013). Schizoaffective disorder is a form of recurrent mood and psychotic illness in which manic episodes occur together with schizophrenia-like psychotic symptoms. In other words, it is a boundary of schizophrenia and bipolar disorder. Rapid cycling is another condition in which the patient experiences at least four distinct episodes of mood disturbances in a year separated by periods of remission. These individuals are usually resistant to many available treatments such as lithium therapy and are therefore, clinically challenging to treat.
1.1.3.2 Diagnosis of schizophrenia

The characteristics of SCZ on the basis of DSM-V criteria includes thought echo; thought insertion or withdrawal; thought broadcasting; delusional perception and delusions of control, influence or passivity; hallucinatory voices commenting or discussing the patient in the third person; thought disorders; disorganised speech; grossly disorganised or catatonic behaviour as positive and negative symptoms (American Psychiatric Association 2013). Negative symptoms often refer to symptoms involving domains which are intact in unaffected individuals but not found in a patient with SCZ such as social impairment, lack of motivation (avolition), poverty of speech (aloria), affective blunting and inattention. By contrast, positive symptoms are those that reflect an excess or distortion in a normal repertoire of behaviour and experience such as delusions and hallucinations. The course of SCZ can be either continuous, or episodic with progressive or stable deficit, or there can be one or more episodes with complete or incomplete remission. In some cases, only one negative symptom is required if the delusions or the hallucinations consist of a voice keeping up a running commentary on the person’s behaviour our thought, or two or more voices conversing with each other. SCZ patients may also have dysfunctional work, interpersonal relations or self-care. In many cases, severe mood symptoms up to and including manic and major depressive episodes are also present in SCZ patients.

According to the classification systems, SCZ is sub-divided into five different definitions based on their clinical symptomology: paranoid, hebephrenic, catatonic, undifferentiated, and residual. The characteristic of paranoid SCZ is that it is dominated by relatively stable, often paranoid delusions and usually accompanied by hallucinations, particularly of the auditory variety and perceptual disturbances. If SCZ patients experience delusions and hallucinations that are fleeting and fragmentary, have behaviour that is irresponsible and unpredictable, have shallow and inappropriate mood, disorganised thought and have incoherent speech, he
or she is generally diagnosed as suffering from hebephrenic SCZ. The prognosis is usually poor because of the rapid development of negative symptoms, particularly flattening of affect and loss of volition. Hebephrenic SCZ should normally only be diagnosed in adolescents or young adults. Patients suffering from catatonic SCZ are required to have one or more of the following prominent catatonic behaviours such as stupor, excitement, posturing, negativism, rigidity, waxy flexibility, and command automatism for a period of at least two weeks. The catatonic phenomena may be combined with a dream-like (oneiroid) state with vivid scenic hallucinations. For diagnosis of undifferentiated SCZ there is the requirement for the general criteria for SCZ to be met, while there may be insufficient symptoms to meet the criteria of other subtypes. If the patient suffers more than four of the negative symptoms such as psychomotor slowing, blunting of affect, passivity, reduced speech throughout the previous twelve months; he or she is diagnosed as suffering from residual SCZ. This class of SCZ may also be present at a lower intensity than positive symptoms.

While there are clinical distinctions between these two nosological groups, there are no pathogenomonic signs or symptoms. Although BPD is characterised by repeated episodes of both depression and mania, no such opposite poles of SCZ exist. However to some extent the negative symptoms of SCZ do share clinical characteristics of depression, such as social withdrawal and psychomotor retardation. In addition to some extent, the positive symptoms of SCZ share characteristics of mania, a syndrome frequently characterised by delusions and hallucinations.
1.1.4 Epidemiology similarities in BPD and SCZ

Epidemiology is the study of the distribution and determinants of disease in human populations and the variations of these distributions among different groups of people. Prevalence is an epidemiological population measure, which refers to the proportion of individuals that have a disorder at a specified time or during a specified period. Lifetime prevalence is the proportion of those in the population who had a disorder at some time in their life up to their age at the time of interview. Researchers have found that BPD and SCZ tends to share several epidemiologic characteristics: (i) both of them have a similar lifetime risks of approximately 1% across the world’s populations and occur at this rate with little variation across continents (Meltzer 1999); (ii) both disorders also have similar gender and age-at-onset distributions; men and women are equally affected; they both commonly occur in young adulthood particularly between ages 15 and 25; and are relatively unusual in pre-pubertal children and after the age of 40 (Castle et al. 1993, Smith A 1992); (iii) once DSM-IV criteria for BPD or SCZ are met, the disorders tend to persists through life and become relatively chronic with a relapse-remit course, rarely showing full recovery, especially in SCZ; (iv) the two disorders are associated with high morbidity and health services use and are also characterised by increased risk of self-harm and suicidal behaviour (Guze and Robins 1970).

1.1.4.1 Epidemiology in BPD

According to the World Health Organisation, BPD affected an estimated 29.5 million individuals worldwide in 2004 (World Health Organization 2008). Recently, a study involving a combined sample of 61,392 community-dwelling individuals in 11 countries, mainly in the Americas, Europe, and Asia, found an aggregate lifetime prevalence of 2.4% in BPD, 0.6% in BP-I, 0.4% in BP-II (Merikangas et al. 2011). Of the 11 countries examined in this study, the lowest lifetime prevalence of bipolar spectrum illness was in India at 0.1%,
and the highest was in the United States at 4.4% (Merikangas et al. 2011). Interestingly, BPD prevalence decreased with increasing age and education level and was higher in unemployed/disabled individuals compared with employed individuals, or those with an income (Merikangas et al. 2011). Childhood-onset BPD prevalence showed some potentially substantive geographic variation (Merikangas et al. 2012, Stringaris et al. 2010). For example, childhood-onset BPD may be less common in Europe than in the United States (Post et al. 2008). Indeed, among the first 1,000 participants in the US-based Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD), 27.7% had an age of onset of less than 13 years, this was associated with a greater number of lifetime depressive and manic episodes, and a greater likelihood of past suicide attempt compared with adolescent onset and adult onset (Perlis et al. 2004). Childhood-onset BPD thus may entail greater genetic vulnerability, in particular the offspring of parents with childhood-onset BPD may be at higher risk of developing BPD. Furthermore BPD has been consistently associated with significant medical and psychiatric comorbidity. For instance, 94.6% of patients with BPD reported in the National Comorbidity Survey Replication (NCS-R) having at least one comorbid disorder, with a mean of 4.6 medical and/or psychiatric comorbidities reported by such individuals (Gadermann et al. 2012). Moreover, in STEP-BD, 58.8% of BPD patients had at least one medical comorbidity and the prevalence was significantly higher in those with lifetime anxiety and substance use disorders (Magalhaes et al. 2012). With regard to medical comorbidities, cardiovascular and metabolic diseases are particularly prevalent among BPD patients, who have shown increased rates of hypertension, obesity, metabolic syndrome, and diabetes (Fiedorowicz et al. 2008, Goldstein et al. 2011, McIntyre et al. 2005, Vancampfort et al. 2013). For example, a large population-based cohort study in Sweden found that BPD patients had increased mortality rates due to cardiovascular and other medical illnesses. On average they died of cardiovascular disease, 10 years earlier than the general population (Westman et al. 2013).
1.1.4.2 Epidemiology in SCZ

SCZ is devastating psychiatric disorder with a median lifetime prevalence of approximately 4 cases per a thousand people with a morbid risk of 7.2 cases per a thousand people (McGrath et al. 2008). The standardised mortality ratio (SMR; ratio of observed deaths to expected deaths) for all-cause mortality is 2.6 for patients with SCZ and dying 12 – 15 years earlier compared to the general population (McGrath et al. 2008, Saha et al. 2007), with excess deaths mainly from suicide during the early phase of the disorder and later from cardiovascular complications. The age-at-onset is typically in adolescence or early adulthood (age group 15 – 35 years) (Messias et al. 2007), onset after the age of 50s and in childhood both being rare (Girard and Simard 2008, Remschmidt and Theisen 2005). Although the prevalence for males and females is similar (McGrath et al. 2008), the course of SCZ is often more severe and with earlier onset for males (Leung and Chue 2000, Messias et al. 2007) whereas after the age of 40 it is women who are considered the most at risk (Aleman et al. 2003). The mechanism underlying the late-age-onset in women has been suggested to be the effects of oestrogen on reduction in dopamine receptors sensitivity in the central nervous system (Hafner et al. 1999). Because of this confounding information it is often accepted as a general view that there are no sex differences in the lifetime risk of developing SCZ.

There is some evidence for other epidemiological risk factors affecting localised prevalence such as family, history, seasonal or urban birth, immigration, perinatal infection and famine. In a comparison study of immigrants in the Netherlands there was evidence that not only does immigration increase the risk of psychiatric illness but that the effect is also exacerbated depending on your age. A lower age at the time of migration was associated with a higher incidence of psychosis among the immigrants (Veling et al. 2011). The effect of immigration factors on prevalence of SCZ has also been replicated on numerous occasions (Bhugra et al. 2011, Mallett et al. 2004, Saha et al. 2005).
Although substantial variations in the prevalence and incidence of SCZ across different countries and cultural groups have been reported (US Institute of Medicine 2001), these differences are reduced when stricter diagnostic criteria are applied (Jablensky 1997). The incidence of SCZ across ten countries was shown to be quite similar in a World Health Organisation (WHO) study (Jablensky et al. 1992). Replication of the WHO study across multiple countries also indicates that the clinical syndrome of SCZ is similar across a wide range of cultures and countries, including those developed and developing countries (Jablensky et al. 1992, World Health Organization 1979).

### 1.1.5 Treatment

Early pharmacological treatment are frequently used in the management of both BPD and SCZ, including antidepressants, mood stabilisers (lithium, valproic acid, carbamazapine), and benzodiazepines. Until recently, treatments of these two disorders were relatively separated. For example, mood stabilisers, such as lithium, anticonvulsants, and antidepressants, do not show substantial efficacy in SCZ. On the other hand, lithium is the most common medication for BPD treatment, which is particularly effective in patients with a family history of the disorder according to a National Health Service report (Mendlewicz 1973). In light of the evidence for overlap in susceptibility, perhaps it should be expected that some medications for one category of these disorders might be beneficial for individuals with the other category of illness.

#### 1.1.5.1 Treatment in BDP

The pharmacological treatment of BPD is based on the course of the illness, for example specific drugs for manic/hypomanic, mixed or depressive episodes (Fountoulakis and Vieta 2008). Mood stabilisers such as lithium and lamotrigine are the most significant pharmacological treatments and are used to ameliorate manic symptoms and sometimes for depression (Fountoulakis and Vieta 2008, Malhi 2009, Malhi et al. 2009). Other medications
such as anticonvulsants for epilepsy and antipsychotics for SCZ have also been used in BPD treatment. Mood episodes in BPD may relapse during continuation treatment as well as recurring during maintenance/prophylactic treatment; therefore long-term treatment to prevent mood episodes is important in the management of BPD. Since the 1960s (Abou-Saleh and Coppen 1986), lithium has been recognised to prevent mood episodes; however, it only gained approval in 1970 from the US Food and Drug Administration. Since then, lithium has become a mainstay treatment in BPD as well as in unipolar depression (Cipriani et al. 2006), where it has been proposed to have a robust neuroprotective and neurotropic action leading to the upregulation of synaptic plasticity. Recently, a study found that a 4-week administration of lithium magnified the long-term potentiating of CA1 pyramidal cells thus up-regulating the synaptic plasticity in the hippocampus (Shim et al. 2012). Conversely, lithium has a narrow therapeutic index and its toxicity can be fatal on some occasions. Alternative mood stabilisers that have been used in clinical practice are valproate, carbamazepine, gabapentine and lamotrigine.

According to clinical observation, BPD patients tend to spend more time being depressed than being manic (Fountoulakis et al. 2012), which causes greater psychosocial impairment and disability. Currently, there are a number of antidepressants have been used widely, including tricyclic antidepressants, selective serotonin reuptake inhibitors (SSRI), monoamine oxidase inhibitors (MAOI), noradrenaline reuptake inhibitors (NARI), serotonin-noradrenaline reuptake inhibitors (SNRI) and atypical antidepressents (Palucha and Pilc 2007). Among atypical antipsyhotics, quetiapine, olanzapine and fluoxetine are routinely used in monotherapy or in combination for the management of bipolar depression. Recently, lithium, lamotrigine or quetiapine monotherapy have been suggested as the first-line choice of treatment. The use of antidepressant agents in combination with antimanic agents such as lithium and valproate, as well as the use of olanzapine-fluoxetine combination (OFC) to
avoid manic switches have also been recommended (Fountoulakis et al. 2005). A similar treatment strategy has also been suggested by the NICE (National Institute for Health and Clinical Excellence). NICE also recommend adding antipsychotic drugs such as olanzapine, quetiapine, or risperidone for acute bipolar mania to the combination of antidepressant and antimanic agents (Nivoli et al. 2012). Recently, olanzapine showed efficacy among BPD patients in preventing recurrences of both mania and depression (Tohen et al. 2003).

1.1.5.2 Treatment in SCZ

Antipsychotic medication can reduce the severity of symptoms and the susceptibility to relapse of these symptoms. The rate of relapse can be worsened by substance abuse. On the other hand, psychosocial therapy can minimise the harmful effects of stress and reduce the effects of stress on vulnerability, and enhance a patient’s coping skills (Mueser and McGurk 2004). Studies of combinations of therapies have suggested that relapses were significantly reduced in those receiving high-dose antipsychotics and family therapy compared to individuals receiving neither drug treatment nor psychological intervention (Dixon et al. 2010, Goldstein et al. 1978, Huxley et al. 2000). This result suggests that complementary psychosocial treatments increase the benefits of pharmacotherapy, by enhancing social functioning and are also being associated with fewer hospitalisations.

Atypical antipsychotics represent the primary medication for schizophrenia because they improve psychotic symptoms and prevent relapse (Kane and Marder 1993). Although antipsychotics have dramatic effects on psychotic symptoms, they have more modest effect on negative symptoms and cognitive impairment (Greden and Tandon 1991).

In the early 1950s, typical or conventional antipsychotics were made available as the primary pharmacological intervention for SCZ. These antipsychotics included chlorpromazine, thioridazine, haloperidol and fluphenazine (Agid et al. 2008). Although conventional antipsychotics are relatively effective at treating psychotic symptoms, they can also produce
problematic side-effects, including Parkinsonian/extrapyramidal symptoms and tardive dyskinesia (Carlini 2004, Malhotra et al. 1993). More recently several other atypical antipsychotics were developed that have a more favourable side-effect profile. These include clozapine, risperidone, olanzapine, quetiapine and sertindole (Leucht et al. 2009). The fundamental property of this group of antipsychotics is their ability to reduce symptoms in the absence of adverse extrapyramidal side effects that patients experience with conventional antipsychotics treatment (Kapur and Mamo 2003). Despite this benefit, they confer an increased risk of metabolic side effects such as diabetes, hypercholesterolemia and weight gain (Schultz et al. 2007). All atypical antipsychotics have a relatively low binding affinity for dopamine receptors. Based on their chemical structure they are classified into dibenzodiazepines, benzisoxazoles, and thienobenzodiazepines. In addition to their low affinity for dopamine receptors, they differ from their earlier counterparts in having a higher affinity for the serotonin (5-HT) 2, receptor upon which they have an antagonistic affect (Worrel et al. 2000). A meta-analysis of comparisons between the efficacy of conventional and atypical antipsychotics showed that atypical antipsychotic treatments are more effective for negative symptoms of SCZ (Arango et al. 2013). Surprisingly, only clozapine has been found to be truly effective in the treatment of both positive and negative symptoms as well as the associated co-morbidity of depression (Kane et al. 1988). However, while clozapine has unique potency for the treatment of persistent psychotic symptoms, negative symptoms, and suicidality (Meltzer et al. 2003, Wahlbeck et al. 1999), it can cause a fatal agranulocytosis in less than 1% of patients (Alvir et al. 1995). As a result patients treated with clozapine require routine monitoring of their white blood-cell count (Cohen and Monden 2013). Despite these side effects overall clozapine has one of the lowest mortality rates in comparison to other anti-psychotics, mainly due to its effect on reducing the risk of suicide.
1.1.6 Pathophysiology of BPD and SCZ

The Kraepelinian dichotomy states that patients with dementia praecox (SCZ) and manic depression (BPD) present as two separate psychotic disorders. Accumulating evidence from clinical observations and antipsychotic treatments in SCZ and BPD recently argued against the Kraepelinian dichotomy. However there is also evidence supporting the dichotomy including data showing volumetric neuroimaging differences in brain regions such as the amygdala, hippocampus, and lateral ventricles that appear to be disorder-specific (Bellivier et al. 2013, Murray et al. 2004, Yu et al. 2010). In addition, reduction of grey matter volumes and neurocognitive neuromoter impairment occur in the early stage of SCZ, but this does not occur in BPD despite both diseases often having an onset during adolescence. Patients suffering from BPD are thought to exhibit less extensive brain morphological abnormalities and less severe neurocognitive deficits in comparison with patients suffering from SCZ (Depp et al. 2007, Ellison-Wright and Bullmore 2010, Krabbendam et al. 2005, Schretlen et al. 2007, Seidman et al. 2002). Recent findings of both shared and different genetic predispositions in BPD and SCZ do not allow firm conclusions about the nature of the Kraepelinian dichotomy (Demjaha et al. 2012, Fusar-Poli et al. 2012, Lewandowski et al. 2011, Murray et al. 2004, Napal et al. 2012, O'Donnell 2007, Young et al. 2006).

By contrast, findings in the molecular pathology of SCZ and BPD have recently challenged the Kraepelinian dichotomy. For example, the dysregulation of dopamine and serotonin have been implicated in both disorders, which suggest that antipsychotic drugs may be useful in the management of BPD as well as SCZ (Demjaha et al. 2012, Murray et al. 2004). Specific serotonin receptor knock-out mice studies consistently show an anxious and depressed phenotype and abnormal physiological responses to stress paradigms (Holmes et al. 2003, Parks et al. 1998, Richardson-Jones et al. 2011, Weisstaub et al. 2006). Selective serotonin reuptake inhibitors have also been found to be effective in treating the negative
symptoms of SCZ (Singh et al. 2010). Furthermore, some of the psychotic experiences in BPD are mediated through dopamine pathways, where antipsychotics act as dopamine D2 receptor blockers (Brugue and Vieta 2007). The depletion of monoamines does not lead to depression in all individuals and administration of monoamine-based antidepressants takes several weeks to improve core depressive symptoms, suggesting downstream neural adaption, as opposed to direct effects (Sanacora et al. 2012). Therefore, researchers have begun to focus on the hypothesis that glutamate and the N-methyl-D-aspartate receptor (NMDAR) may play an important role in the pathophysiology of BPD, SCZ and related disorders. Postmortem studies have shown increased glutamate levels in diverse brain areas in individuals with mood disorders (Hashimoto et al. 2007) and a reduction of NMDARs in BPD patients (Hashimoto et al. 2007, Rao et al. 2012). Patients suffering from SCZ showed a significant reduction in glutamate levels and a hypoactive glutamate system (Coyle 1996, Goff and Coyle 2001, Goff and Wine 1997, Jentsch and Roth 1999, Kim et al. 1980, Meador-Woodruff and Healy 2000, Olney and Farber 1995), an increased release of glutamate was shown in selected brain areas (Olney et al. 1999). Other neurotransmitter systems, such as the  \( \gamma \)-aminobutyric acid (GABA) B subunits (GABA\(_B\)) receptors, which mediate the release of a number of neurotransmitters including dopamine, serotonin, noradrenaline, somatostatin, glutamate and GABA (Nyitrai et al. 2003, Sakamaki et al. 2003, Steiniger and Kretschmer 2003, Takahashi et al. 2010, Waldmeier et al. 2008). Postmortem and genetic studies have demonstrated changes in GABA\(_B\) receptors that are thought to be involved in the pathophysiology of BPD and SCZ (Ghose et al. 2011, Ishikawa et al. 2005, Klempan et al. 2009, Mizukami et al. 2000, Zai et al. 2005).

With regards to neurocognitive functioning, several studies have reported poorer neurocognitive functioning in patients with SCZ compared to those with BPD, but there are also data suggesting that patients with SCZ are comparable with BPD patients in terms of
neurocognitive impairments (Altshuler et al. 2004, Daban et al. 2006, Depp et al. 2007, Glahn et al. 2006, Meesters et al. 2013, Sanchez-Morla et al. 2009, Simonsen et al. 2011, Varga et al. 2007). These inconsistent neurocognitive functioning findings may be due to differences in sample size, the specific neurocognitive deficits that are measured, the types of scales used, and the presence of other confounders including age and premorbid intelligence (Bora et al. 2008, Daban et al. 2006, Glahn et al. 2006, Simonsen et al. 2011).

1.1.7 Heritability of BPD and SCZ

The heritability of BPD and SCZ has been studied for almost one hundred years using twin and family studies. These have produced estimates of approximately 60 – 90% in SCZ and 60 – 80% in BPD. Co-occurrence of the two disorders has also been shown in families (Berrettini 2000, Lichtenstein et al. 2009, McGuffin et al. 2003, Sullivan et al. 2003). The lifetime risk of BPD is only 0.5 – 1.5%; however, if there is a parent or other first-degree relative with BPD, it increases to about 5 – 10%; and also if there is a more extensive family history such as both parents are affected, then this risk is multiplied several-fold (Craddock and Jones 1999). There is a ten-fold increase in the lifetime risk of SCZ in the first-degree relatives of a person suffering from SCZ (Gottesman II 1982). The genetic risk increases to nearly 50% when both parents are affected (McGuffin et al. 1995). These results strongly suggest that genetic factors are playing an important role in the aetiology of BPD and SCZ.

1.1.7.1 Twin and adoption studies

In twin studies concordance rate differences between monozygotic (MZ, identical) twins point to the role of the environment, and concordance rate differences between dizygotic (DZ, fraternal) twins point to both hereditary and environmental factors. The concordance rate is the probability that a second twin will develop a disorder if the proband (first examined) twin has the disorder, is commonly used. The concordance rates of SCZ for MZ twins have been found to be about 40 – 50% (Cardno and Gottesman 2000, Sullivan et al. 2003); where the
concordance rate for BPD is also 65% higher among MZ twins (McGuffin et al. 2003). The high heritability estimates and high monozygotic concordance rates are convincing indicators of the importance of genetic factors affecting BPD and SCZ susceptibility.

Adoption studies allow dissection of genetic from environmental contributions to a disorder in ways that twin studies cannot (Ingraham and Kety 2000). These studies have found an elevated risk for psychosis in such offspring, whether the affected parents had an onset of disease before or after adoption, and whether the rearing environment was foster parents or an institution (Heston 1966, Higgins 1976, Tienari 1991, Tienari et al. 1985, Tienari et al. 1994, Wender et al. 1974). Consistent with the risk travelling with the biological rather than the adoptive relationship, it was also shown that the risk was similar for offspring of SCZ mothers, whether they were raised by the affected biological affected parent or an adoptive un-affected parent (Higgins 1976). The offspring of mothers without SCZ also did not have an increased risk when raised by adoptive parents suffering from psychosis (Wender et al. 1974). Furthermore, adoption studies can yield some insight into gene-environment interactions, for example by comparing communication deviance in adoptive parents of high-risk adoptees (Wahlberg et al. 1997). The adoptees’ family approach starts with SCZ adoptees and matched control adoptees, and evaluates their adoptive and biological families for illness. These studies have shown elevated rates of SCZ and SCZ spectrum disorders in biological families of SCZ adoptees compared to biological families of control adoptees, coupled with low and equivalent such rates in adoptive families of both types of adoptees (Kendler et al. 1994, Kendler et al. 1981, Kety et al. 1994).
1.1.7.2 Family studies

Familial aggregation has been demonstrated repeatedly in SCZ and BPD (Lichtenstein et al. 2009), and also in schizoaffective disorder (Cardno et al. 2002). A meta-analysis of high-risk family studies recently compared offspring of parents with SCZ, BPD and major depressive disorder (Rasic et al. 2014). The study showed that the offspring of adults with SCZ, BPD or major depressive disorder had a 32% probability of developing one of these disorders themselves by adulthood and that the risks to offspring were not limited to their parent’s index disorder. For example, the offspring of patients with SCZ or BPD had an increased risk of SCZ, BPD and major depressive disorder. However, although the relative risk of SCZ and BPD was also elevated in offspring of parents with depression, this finding did not reach statistical significance. Another meta-analysis of family studies of SCZ and BPD confirmed an overlap in familial risk (Van Snellenberg and de Candia 2009). However, a Swedish population register recently showed although the cross-disorder risk is lower than the same-disorder risk, a that partial overlap exists in familial genetic susceptibility for BPD and SCZ (Lichtenstein et al. 2009). Therefore, these family studies are consistent with partial overlap in familial susceptibility of BPD and SCZ.

BPD and SCZ susceptibility gene identification has been a slow and difficult process, as with other common complex traits such as diabetes, asthma, and hypertension. This inconsistency of the findings among all these studies can be attributed to the small effect sizes of BPD or SCZ susceptibility genes, locus heterogeneity, and phenotypic heterogeneity (Schulze and McMahon 2003).
1.2 Molecular Genetics Studies

Advances in molecular genetic studies such as genome-wide association studies (GWASs), meta-analyses, mega-analyses, high-throughput DNA and RNA sequencing and gene expression arrays have detected susceptibility genes or variants that are likely to contribute to the risk diseases such as bipolar disorder (BPD) and schizophrenia (SCZ). The evidence from these studies has provided psychiatry with an unprecedented opportunity to identify the biological systems that are involved in illness and with the potential to lead to genetic medicine. Both BPD and SCZ have a complex aetiology pattern, thus they fall into the category of complex genetic disorders which may involve multiple genes either in an oligogenic model (the interaction of a small number of major genes) or a polygenic mode (minor involvement of many genes) and environmental risk factors. The clinical diagnosis of these two disorders is based on behavioural signs and symptoms from each individual. Neither disorder currently has a specific, clearly defined, pathophysiological foundation on which to base diagnostic tests. Examining biological markers by molecular genetic studies may provide clues in the development of such tests, which may be clinically useful in risk prediction. A review of molecular genetics findings in BPD and SCZ will be discussed in four different stages: linkage studies, association studies, GWAS and copy number variation (CNV) studies.

The initial genetic research focused on family linkage and allelic association methods, which have been applied to narrow the location of genes to small regions of the genome for identification and characterisation of the susceptibility gene and eventually the casual variant. Before the introduction of GWAS, molecular genetic studies can be conceptually divided into positional and candidate gene approaches. Positional linkage studies are based on chromosomal positions identification of the susceptibility genes in a few large families conferring a high degree link between the families and disease risk. This method can be
considered as a purely genetic approach that does not require prior knowledge of disease pathophysiology. In contrast, the researcher requires sufficient understanding of disease biology for the candidate gene approach to identify genes that may be associated with psychiatric disorders (Craddock and Jones 1999). With increasing understanding of common genetic variation and advances in the technological ability to genotype more people, interest shifted to association studies of larger numbers of case of BPD or SCZ and healthy people to detect population-level associations. Traditionally, association studies were focused on identification of variants within specific candidate genes of interest. With advances in mapping of the human genome and high throughput genotyping, focus shifted to GWASs. GWAS studies were made possible by advances in technology and also a reduction in genotyping costs. GWAS studies are able to detect common risk variants across the whole human genome in an unbiased way. These studies typically genotype hundreds of thousands of single nucleotide polymorphisms in thousands of people. This genome-wide analysis data can also be applied to determine rare structural chromosomal variants, also known as copy number variants (CNVs). To date, these analyses have provided robust and replicable findings in psychiatric genetics, which will be discussed in more detail later. With emerging technologies such as next generation sequencing that makes whole genome sequencing feasible in many thousands of samples. It has become possible to determine the contribution of rare variants with a modest effect on susceptibility genes. These various experimental approaches have the capacity to provide a complete picture of the genetic variation that affects disease risk.

BPD and SCZ are classified as different psychiatric disorders in modern western classification systems and historically this distinction is often referred to as the Kraeplinian dichotomy (described earlier). However, several similarities have been found, such as lifetime risks of approximately 1% across the world’s populations, an age-at-onset in young
adulthood, psychotic symptoms with a relapse-remit course, shared antipsychotic medications, high morbidity and increased risk of self-harm and suicidal behaviour, and a difficult differential diagnosis with schizoaffective disorder. These similarities have suggested possible shared common genetic susceptibilities in BPD and SCZ. This notion is supported by evidence from studies implicating common linkage regions, gene expression patterns and molecular mechanisms in BPD and SCZ (Craddock et al. 2006, Knight et al. 2009, Le-Niculescu et al. 2007, Shao and Vawter 2008). Given these clinical observations and the shared genetic findings it was widely expected that the genetic approaches described above would reveal genes that increase risk for both BPD and SCZ. In fact, BPD and SCZ susceptibility gene identification has been a slow and difficult process, as with other common complex traits such as diabetes, asthma, and hypertension. The inconsistency of the findings among all these studies can in part be attributed to lack of statistical power due to small sample size, the small effect of BPD and/or SCZ susceptibility genes, locus heterogeneity and phenotypic heterogeneity (Schulze and McMahon 2003).

1.2.1 Factors affecting genetic studies

1.2.1.1 Locus heterogeneity

The term genetic heterogeneity can be divided into two different dimensions, known as locus and allelic heterogeneity. Locus heterogeneity means that similar phenotypes may be caused by mutations at different gene loci, whereas allelic heterogeneity refers to the same phenotypic outcome at a specific gene locus for different alleles. Since linkage to the same chromosomal region is still detectable in the presence of allelic heterogeneity this does not represent a problem for linkage analysis; however, it is a problem for association analysis where tests of association generally depend on cases sharing the same allele. Conversely, the presence of several different genes with similar phenotypes generally is troublesome for linkage analysis. Although locus heterogeneity has fewer problems from association analysis,
it reduces the power of linkage and association analysis in identifying disease variants (Gershon and Goldin 1986, Risch and Botstein 1996). One of the main approaches to the problem of locus heterogeneity is by careful ascertainment. This may be achieved restricting collection of individuals or their family to geographical regions or to population isolates with the aim of reducing the number of different susceptibility genes.

1.2.1.2 Phenotypic heterogeneity

Recently, phenotype definition has become increasingly important because imprecision may result in variation in the validity of the diagnosis. A clear and distinct phenotype that captures the underlying mechanism of BPD and SCZ has yet to be defined. For instance, the identification of an affected phenotype for BP-I is highly accurate and reliable compared to the BP-II, which has a less valid and reliable diagnosis (Craddock and Jones 1999). Because of these difficulties endophenotypes have been used in order to discover clinical entities that are associated with a disease but may be closer to the underlying biology compared to the symptoms of the disease. This approach has increased the understanding of the clinical picture in SCZ (Adler et al. 1982, Gorman et al. 1990, Porjesz et al. 2002), but have not yet led to the same insights in BPD (Turecki et al. 2001). These findings suggest that the disorder spectrum phenotype for genetic analysis needs to be as broad as possible, especially for complex disorders such as BPD.

1.2.1.3 Sample sizes

The final well-known problem that has faced linkage and association studies of complex disorders such as BPD and SCZ is sample size. A sample size with sufficient statistical power is an important factor for success in genetic studies that seek to detect causal genes for complex human disease. To date many weak linkage and association findings have been reported and these have had inconsistent replication findings. Meta-analyses of genome-wide
association studies and/or meta-analyses of linkage scans may overcome some of the limitations of small sample size (Levinson et al. 2003).

1.2.1.4 Anticipation

SCZ and BPD are the two major psychotic disorders with a life-time prevalence of approximately 1%. Both disorders have a substantial genetic contribution. Several studies have shown that early onset of these two major psychoses are generally characterised by an increased number of psychotic symptoms and episodes of mania, a higher level of comorbidity with drug addiction and poor clinical outcomes (Craddock and Sklar 2009, Thibaut et al. 1995). Anticipation is one explanation for increased disease severity or decreased age of onset in succeeding generation within families. This phenomenon is known to involve either methylation effects on susceptibility genes or the expansion of trinucleotide repeat DNA sequences in some genetic illnesses, such as myotonic dystrophy, fragile X mental retardation and Huntington disease (Holmes et al. 1999, Koob et al. 1999, Ridley et al. 1988). Many recent studies have also found consistently positive results for genetic anticipation in SCZ and BPD (Vincent et al. 2000) but more studies are still required to confirm and elucidate the mechanism that underlies this phenomenon.

1.2.2 Genetic Markers

The human genome is approximately 99.5% identical and the difference between individuals is about 0.1 – 0.5%. This variation is caused by several types of structural sequence variants and polymorphisms. Genetic markers are commonly used to study the relationship between an inherited disease and its genetic cause in disease gene mapping studies. Genetic markers can be either a single nucleotide polymorphism (SNP) or comprise longer DNA sequences such as microsatellites. Microsatellites are also known as simple sequence repeats (SSRs) or short tandem repeats (STRs). They were initially introduced by Weber and May and Litt and Buty in 1989 (Litt and Luty 1989, Schlotterer 2000) and have traditionally been used in
linkage studies. Microsatellites are short tandem repeating sequences of 2-6 base pairs of DNA. They are widely distributed in the genome and are highly polymorphic among individuals (Putman and Carbone 2014). They are widely used for fingerprinting, parentage identification, genetic mapping, conservation and population genetics. Conversely, SNPs are common single nucleotide substitutions with a population frequency of approximately 1%. SNP markers have been widely used for positional fine mapping and functional candidate genes. Recently, sets of hundreds of thousands of SNPs on a single array have been used for GWAS of the whole human genome.

1.2.3 Linkage analysis

Linkage analysis utilises genetic information from families with multiple affected individuals and search for regions of the genome segregating or linked with the disorder. The presence of genetic linkage is a violation of Mendel’s law of independent assortment in which alleles at two chromosomal locations assort independently and are transmitted to offspring in random combinations.

1.2.3.1 Linkage findings susceptibility loci in common

More than 20 years of linkage studies in psychiatric genetics have produced a huge amount of data. To date linkage regions likely to harbour SCZ susceptibility genes are spread over 21 out of the 23 pairs of chromosomes (Crow 2007, Tandon et al. 2008). For BPD linkage loci have been found on every single chromosome (Serretti and Mandelli 2008). Epidemiological studies performed before linkage studies suggested an overlap in susceptibility across the boundaries between SCZ and BPD. Subsequent genetic linkage studies also demonstrated the existence of common linkage regions for these two disorders (Berrettini 2003, Hamshere et al. 2005, Shao and Vawter 2008).

Molecular linkage studies have so far revealed promising overlapping susceptibility regions for the two disorders including 18p11.2, 13q32, 22q11-3, 10p14, and 8p22 (Berrettini 2003).
Historically, the first chromosomes reported to harbour genes for mental illness was chromosome 18 (Richards et al. 1970) and an affected sibling pair study in BPD reported a susceptibility locus on 18p11 (Berrettini 1997, Berrettini et al. 1997, Berrettini et al. 1994, Detera-Wadleigh et al. 1999) that has been replicated by several investigators (MacKinnon et al. 1998, Nothen et al. 1999, Stine et al. 1995, Turecki et al. 1999). This region has also been implicated harbouring putative loci for SCZ in a study of families multiply affected by this disorder (Schwab et al. 1998b). However, it has been argued that the SCZ families might have been misdiagnosed as suffering from BPD. Other SCZ susceptibility loci were also identified in the same study including regions on 6p and 10p14 and these results indicate that the diagnosis of SCZ was likely to be correct (Schwab et al. 1995, Schwab et al. 1998a, Schwab et al. 1998b). Another positive linkage and association finding at 18p11.2 was reported to be shared between the BPD, SCZ, psychosis not otherwise specified, and schizoaffective disorder (Mukherjee et al. 2006). Later, meta-analyses were performed to solve the problem of sample size in traditional linkage studies, and also to evaluate evidence for linkage from the combined data of multiple genome scans to determine the significance of finding. A rank-based meta-analysis conducted on all the published genome scans confirmed that 18p11 is a nominally significant locus for major psychotic disorders such as BPD, BP-I, BP-II, schizoaffective disorder and recurrent unipolar disorder (Schwab et al. 1998b).

More than 44 positive findings of susceptibility loci on chromosome 22 have been reported in patients with SCZ, while only a few have been reported in patients with BPD (1996, Gill et al. 1996, Moldin 1997, Schwab and Wildenauer 1999). However, most of the positive findings were weak and do not reach the levels of significance suggested by Lander and Kruglyak (Lander and Kruglyak 1995). Furthermore, a number of studies suggest that marker D22S278 on 22q12 is possibly linked to a susceptibility locus for SCZ (Schwab and
Wildenauer 1999). Later, a genome survey of 20 North American families (Kelsoe et al. 2001) reported that the same marker had a maximum parametric lod score of 2.8 in individuals with BPD and marker D22S419 on 22q11 with a secondary lod score peak of 2.19. The location of the marker D22S419 is close to marker D22S315, which was reported to have a maximum lod score of 3.5 in a genome survey of SCZ families (Myles-Worsley et al. 1999). Deletions in chromosome 22q11 are present in 85% of individuals with velo-cardio-facial syndrome (VCFS) (Driscoll et al. 1992). VCFS is a rare congenital disease consisting of characteristic facial and palatal dysmorphology, heart disease, learning disabilities, and interestingly, a high prevalence of psychiatric symptoms including SCZ and BPD. Pulver et al. reported that nearly 80% of VCFS in their cases were given a psychiatric diagnosis, and 29% of those had SCZ (Pulver et al. 1994), while Murphy et al (Murphy et al. 1999) reported that 30% of their VCFS cases also had psychotic symptoms and nearly all of them fulfilled criteria for SCZ. These findings suggest that a small proportion of cases of SCZ may result from deletions of 22q11. Several authors have reported evidence for a BPD susceptibility locus on chromosome 22q11-13, near the VCFS locus (Detera-Wadleigh et al. 1999, Edenberg et al. 1997, Murphy et al. 1999). Furthermore, Papolos et al. (Papolos et al. 1996) reported that 16% of their VCFS sample had psychotic symptoms, and that 64% met criteria for some form of BPD. A risk chromosomal region 22q12-13 has also been identified as having an overlapping in both SCZ and BPD (Badner and Gershon 2002, Lewis et al. 2003, Liang et al. 2002, Takahashi et al. 2005).

10p14 markers have been linked to SCZ (Faraone et al. 1998, Schwab et al. 1998b, Straub et al. 1998). Although the data from these studies were statistically less significant than the other regions, the fact that three studies coincide suggests that it is a true effect. 10p14 linkage was also found in a sample of families with BPD, suggesting that this region is likely to be another locus of shared risk between BPD and SCZ (Foroud et al. 2000). Several
independent studies have revealed linkage of SCZ to a 13q32 locus (Blouin et al. 1998, Brzustowicz et al. 1999, Lin et al. 1997), while linkage of this locus to BPD was also reported in a family study (Detera-Wadleigh et al. 1999).

Other meta-analyses of linkage studies based on clinical phenotypes have identified and replicated overlapping loci including 1q32, 10p11-15, 13q32, 18p11.2 and 22q11-13 (Badner and Gershon 2002, Baron 2001, Berrettini 2000, Bramon and Sham 2001, Sklar et al. 2002). Linkage studies of psychotic BPD showed other suggestible seven regions that evidence for an overlap with SCZ including 5q33, 6q21, 8q24, 15q26, 17p12, 18q21, 20q13 (Park et al. 2004). Taken together, these studies implicate shared susceptibility loci for SCZ and BPD and this argues against traditional categorical diagnoses.

1.2.3.2 Disease specific linkage findings

Although several susceptibility loci have been identified for both SCZ and BPD, the majority of susceptibility loci identified to date are not actually shared between these two disorders. Specifically implicated in risk for SCZ are 6p22-24, 8p21-22, and 5q21 (McGuffin and Owen 1996, Williams et al. 1999); whereas unique loci for BPD are including 4p16, 12q23-24, 18q22 and 21q21 (Blackwood et al. 1996, Stine et al. 1995). In addition, several large meta-analyses have also found evidence of numerous genetic linkage of which 6p22-24, 1q21-22, and 13q32-43 are the best supported for SCZ (Lewis et al. 2003, Owen et al. 2004), and suggestive linkages have also been reported in 8p21-22, 6p22, 6q21-25, 5q21-33, and 1q42 (Baron 2001, Berrettini 2000, Lewis et al. 2003, Owen et al. 2004, Segurado et al. 2003). An international collaboration performed a linkage scan using 6,000 SNPs on over 900 pedigrees that showed the most significant evidence for linkage in a region spanning 1Mb of chromosome 8p21 (Holmans et al. 2009). This has been replicated numerous times and was validated by an independent case-control association study (Walss-Bass et al. 2006). Several other meta-analyses of BPD data sets indicated and confirmed the most promising
linkage regions were 4p16, 16q12, 18q22, 21q21 and 12q24 (Berrettini 2000, Liu et al. 2003, Segurado et al. 2003).

1.2.4 Association studies

Genetic association is a statistical statement about the co-occurrence of alleles or phenotypes. The aim of genetic association studies is to find statistically significant differences in single-locus alleles or genotype frequencies in a patient group compared with a normal control group. Several molecular genetics association studies have focused on the identification of genetic variation based on positional and/or functional evidence, which increase the risk of psychotic symptoms in psychiatric disorders. The hypothesis used to find candidate genes involved in the aetiology of BPD were those involved in neurobiological pathways such as dopamine, serotonin or in the circadian rhythm. The hypothesis considered for SCZ were genes involved in dopamine, glutamate, or in neurodevelopment.

1.2.4.1 Association findings shared across disorders

Recent reports have shown that the candidate genes originally implicated in SCZ may also influence susceptibility to BPD. Amongst the genes that have shown evidence for association are those involved in component symptom dimensions, such as psychosis or mood symptoms, in the boundary between SCZ-BPD. Association studies of psychotic BPD and subtypes such as mood-incongruent psychotic BPD have revealed modest positive results for several candidate susceptibility genes, including dysbindin (distrobrevin-binding protein 1) on 6p22.3, DISC1 (disrupted in schizophrenia 1) on 1q42, and NRG1 (neuregulin 1) on 8p12 (Goes et al. 2008). For instance, Green et al. (Green et al. 2005) found the association of NRG1 in BPD with mood-incongruent psychotic symptoms as well as with SCZ with lifetime manic episodes, suggesting that NRG1 may confer susceptibility to a phenotype with combined features of psychosis and mania. A recent study reported additional evidence for association between psychotic BPD and NRG1 (Goes et al. 2009). In addition, there is
evidence that variation in the DISC1 gene influences susceptibility to disorders of the psychosis spectrum, including SCZ, schizoaffective disorder and BPD (Owen et al. 2007).

Dysbindin has been extensively implicated in SCZ and it has also been reported to be associated with BPD with recurrent psychotic symptoms (Raybould et al. 2005). Several SNPs in the dysbindin (DTNBP1) gene have been found to be associated with psychotic major depression compared with non-psychotic major depression (Domschke et al. 2011). This finding is in accordance with the hypothesis that dysbindin crucially influences dopamine and glutamate neurotransmission and that this dysregulation is involved psychotic disorders (Papaleo and Weinberger 2011). The finding of DTNBP1 genetic association with psychotic depression (Domschke et al. 2011) substantiates the role of DTNBP1 genetic variation which had been reported to confer an increased risk of SCZ (Benson et al. 2004, Fanous et al. 2005, Pae et al. 2009, Schwab et al. 2003, Straub et al. 2002, van den Oord et al. 2003, Zuo et al. 2009), psychotic symptoms in general (Kohn et al. 2004), BPD (Breen et al. 2006, Gaysina et al. 2009, Joo et al. 2007, Pae et al. 2007, Raybould et al. 2005) and unipolar depression (Kim et al. 2008). These tentative results are consistent with the hypothesis that psychotic BPD (and psychotic unipolar depression) may represent a clinical manifestation of the effect of overlapping genes across SCZ and mood disorder syndromes.

Furthermore, a study implicated genetic variation in G72 (DAOA)/G30 (D-amino acid oxidase activator (G72)/G30) on 13q33 in susceptibility to major mood episodes across the traditional SCZ and BPD categories (Williams et al. 2006) suggesting that even though this locus was originally described as a SCZ risk gene, it may be more strongly associated with mood symptoms than with psychosis with the SCZ/BPD continuum. However not all studies support this finding (Maheshwari et al. 2009, Shi et al. 2008a).

A cross-disorder approach reported variation in exon 3 of the DRD4 (dopamine D4 receptor) gene to be associated with delusional symptoms as assessed using the OPCRIT (McGuffin et
al. 1991) across major depression, SCZ, delusional disorder, and psychotic disorder not otherwise specified (Serretti et al. 2001, Serretti et al. 1999). Association of \textit{DRD4} gene variation across affective and psychotic disorders has been corroborated in an independent sample comprising patients with SCZ, schizoaffective and unipolar affective disorders (Weiss et al. 1996). Additionally, the dopamine receptor 2 (\textit{DRD2}) S311C variant has been observed to be associated with delusional symptoms across major depression, SCZ, delusional disorder and psychotic disorder not otherwise specified (Serretti et al. 2000). \textit{COMT} (catechol-O-methyltransferase) on 22q11 (Funke et al. 2005), \textit{CACNA1C} (Green et al. 2010), and \textit{HTR5A} (Birkett et al. 2000) have also been identified as common vulnerability genes for both affective disorders and SCZ (Craddock et al. 2006).

In summary, molecular genetic studies, as well as epidemiological and family studies, have shown evidence that SCZ and BPD partly share a common genetic cause. These data challenge the current nosological dichotomy between the two types of psychosis, and are reflective of the need for reappraisal of these disorders as distinct diagnostic entities.

\textbf{1.2.4.2 Association studies identify disease specific risk genes}

The Wellcome Trust Consortium reported genes associated with BPD and in particular that there was support for genes involved in GABA neurotransmission (rs7680321 in \textit{GABRB1} encoding a ligand-gated ion channel (GABA A receptor, beta 1)), glutamate neurotransmission (rs1485171 in \textit{GRM7} (glutamate receptor, metabotropic 7)) and synaptic function (rs11089599 in \textit{SYN3} (synapsin III)). In a Japanese sample of patients with major depression, the met allele of the brain derived neurotrophic factor (BDNF) val66met polymorphism was found to be associated with psychotic features (Iga et al. 2007). Examining intermediate phenotypes in BPD has been strongly advocated, although rarely done, as a critical element in identifying informative genetic loci (Glahn et al. 2004, Lenox et al. 2002, MacQueen et al. 2005). Studies of genes controlling the circadian rhythm provide a
compelling example of how the phenotypic approach can be used to identify genetic risk factors for BPD.

Other reports suggest that CLOCK (Benedetti et al. 2007, Benedetti et al. 2003, Lamont et al. 2007, McClung 2007, Shi et al. 2008c), BmaL1 (Mansour et al. 2006, McClung 2007, Nievergelt et al. 2006), TIMELESS (Mansour et al. 2006, Shi et al. 2008c), and PERIOD1-3 (Nievergelt et al. 2006, Shi et al. 2008c) are candidate loci associated with the circadian rhythm in the BPD phenotype, although the majority of reports are preliminary and not all studies confirm these associations (Bailer et al. 2005, Nievergelt et al. 2005, Shiino et al. 2003). In a 2007 review of the genetics of BPD (Kato 2007) additional associations with TRPM2 (21q22.3), GPR50 (Xq28), Citron (12q24), CHMP1.5 (18p11.2), GCHI (14q22-24), MLC1 (22q13), GABRA5 (15q11-q13), BCR (22q11), CUX2, FLJ32356 (12q23-q24), and NAPG (18p11) looked promising but it was suggested that future replication studies were warranted.

Association studies have also identified several putative candidate genes for SCZ. Some of these are SCZ specific risk genes and they include, but are not limited to, COMT (Egan et al. 2001, Malhotra et al. 2002, Shifman et al. 2002, Wonodi et al. 2006); BDNF (brain derived neurotrophic factor) (Buckley et al. 2007, Gratacos et al. 2007, Ho et al. 2007); RGS4 (regulator of G protein signalling 4) on 1q23 (Chen et al. 2004, Chowdari et al. 2002, Morris et al. 2004, Williams et al. 2004); MTHFR, PPP3CC, GABRB2 and TP53 (Shi et al. 2008b); although the reports of successful replication of these findings vary considerably. Protective allele associations in DAOA, IL1B (interleukin-1 beta), and SLC6A4 genes were also reported in a meta-analysis (Shi et al. 2008b). By contrast, variation in the IL1B gene has been reported to be associated with the subtype of SCZ with high levels of depressive symptoms (Boks et al. 2008, Rosa et al. 2004). SCZ is a common illness with presumably multiple genes of small effects involved in the aetiology and many of the studies to date are
statistically underpowered. Geneticists suggest that tens of thousands cases and controls may be needed to find firm associations (please see the later section on GWAS studies (Abbott 2008). The development of a fundamental understanding of the effect of risk genes will undoubtedly be complex. And even though several risk genes have been implicated, the associated variants are often different in different populations and it is therefore difficult to determine the biological effect of each risk gene. Potential interactions between risk genes add to the complexity of the picture. In addition, the phenomenological heterogeneity of psychotic disorders, as well as lack of clear boundaries and biologically based definitions in the existing diagnostic categories may contribute to difficulties facing genetic studies.

1.2.5 Summary

Despite all the efforts, no single genetic variant has reliably been established as functionally implicated in the pathogenesis of SCZ or BPD. The lack of reproducibility of linkage or association findings may depend on several study-specific issues, such as variable power, differences in phenotypic models, inadequate sample sizes, differences in modelling parameters, different markers used, insufficient variation coverage, differences in linkage disequilibrium (LD) between studied populations, and most likely reflects true locus and allelic heterogeneity between and probably also within the populations studied. Although the replication of findings clearly provides an indication for the credibility of a story, there is no clear criterion for distinguishing signal from noise in single gene association studies and also lack of correlation with strong biological evidence. Thus, it is difficult to fully grasp the importance of existing findings until the causative susceptibility variants are found and functionally validated.
1.2.6 Genome-wide association studies

Although genetic linkage and single locus association studies have produced a range of evidence of association for a range of susceptibility loci and candidate genes with limited a priori evidence and/or with limited sample sizes, these have been difficult to replicate consistently. Thus, the focus has moved to large-scale genome-wide association studies (GWAS), which are geared to identify the effect of common genetic variants or candidate genes that are especially important in traits for which the biological pathway is unknown or complicated, such as psychiatric disorders. This is the first analysis of common variation across the entire human genome by genetic markers to identify positive genetic associations when there is a greater frequency in the presence of genetic variants in individuals with observable traits such as a disease or unaffected individuals. Therefore, GWAS potentially allows multiple susceptibility genes to be detected simultaneously when sufficiently large number of cases and controls are assayed (Ansorge 2009). However, a serious difficulty in evaluating the results of GWASs is the issue of multiple testing. For instance, a large number of SNPs may be tested within the same study for their association with a disease and this generates many nominally significant findings that are actually false positives. It is therefore necessary to set a stringent level for defining statistical significance, which can minimise the chance that a marker artificially exhibits evidence of phenotypic association, often requiring statistical evidence stronger than a p-value of $7.2 \times 10^{-8}$ (Dudbridge and Gusnanto 2008).

This level is dependent upon the number of SNPs analysed, and the threshold for currently available GWA chips is approximately $5 \times 10^{-8}$ (Dudbridge and Gusnanto 2008, Hoggart et al. 2008). In order to be able to reliably detect genomic regions that are genuinely influencing disease risk requires the analysis of large numbers of cases and controls for the initial study and for subsequent replication studies (Cantor et al. 2010). However, the design and execution of independent replication studies can be difficult, particularly those involving population groups that are genetically diverse from the original discovery population.
Homogenous samples that are limited to a single ancestry group such as Japanese or Ashkenazi Jewish represent an exception to this and result in the best associations. Moreover, the variants that are significantly associated in GWASs are generally at specific locations and do not implicate complete genes. Often GWAS had been found with variants that are not located close to a protein-coding gene or are within genes that were not previously considered to be candidate genes. GWASs only detect common variants (> 5%) in a population, which generally have small effect. For the successful conduct of a GWAS study it is important to have appropriate statistical methods that are necessary to reduce the risk of multiple false positive results. It is also preferable to only perform GWASs when statistical power is not limited by small sample size. Furthermore, in order to find possible casual variants, refined analysis is required to understand the biological pathway and the role that the gene variants may have in the disease process or in the condition.

Several successful GWASs have been conducted for a variety of common, complex diseases including several psychiatry disorders, and their result can be found at Catalogue of Published Genome Wide Association Studies from the National Human Genome Research Institute (NHGRI) and the European Bioinformatics Institute (EMBL-EBI) (http://www.ebi.ac.uk/gwas/).

1.2.6.1 GWASs in BPD

To date, there are several GWAS have been conducted on patients with BPD. The initial GWAS BPD studies were published in 2007 by Nick Craddock from the Welcome Trust Case-control Consortium (WTCCC), which involved 14,000 cases of seven major diseases including 1,900 BPD patients versus 3,000 shared controls. They identified a genetic association signal at chromosome 16p12 with $p = 6.3 \times 10^{-8}$, a region that contains PALB2, DCTN5, and NDUFA1 genes (Wellcome Trust Case Control Consortium 2007). Another 37 SNPs selected genotyping study by Baum et al. 2008 from the NIMH intramural program
detected the strongest association signal at a marker within the first intron of *DGKH* (diacylglycerol kinase eta; $p = 1.5 \times 10^{-8}$) with an odd ratio of 1.59 (Baum et al. 2008). These two first genome-wide association studies of BPD showed that several genes had modest effect, where implicated that BPD may be a polygenic diseases. A year later Sklar et al (Sklar et al. 2008) published a GWAS of 1,868 BPD and 2,938 controls. The strongest association signals in this study were for SNPs within the voltage-dependent calcium channel, L-type, alpha 1C subunit (*CACNA1C*) gene. Later that year, combined studies of the WTCC (Wellcome Trust Case Control Consortium 2007) data and the Sklar et al (Sklar et al. 2008) BPD data including total of 4,387 patients and 6,209 controls discovered the first genome-wide significant association signal in *ANK3* gene (encoding the protein ankyrin G) (rs10994336, $p = 9.1 \times 10^{-9}$) and further support for the previously reported *CACNA1C* gene (rs1006737, $p = 7 \times 10^{-8}$) for BPD (Ferreira et al. 2008). Since these early success BPD GWAS studies, many GWAS have been conducted with various populations, including European American, African, American, German, Japanese, Han Chinese, Norwegian, Icelandic, Bulgarian, Canadian, and UK populations (Lee et al. 2011, Hattori et al. 2009, Djurovic et al. 2010, Cichon et al. 2011, Yosifova et al. 2011, Smith et al. 2011, Smith et al. 2009), but many of these failed to detect genome-wide significant signals due to small sample size. In order to improve the power to detect BPD candidate variants with genome-wide significant findings, the Psychiatric Genomics Consortium (PGC) analysed and published a combined genome-wide association study of 7,481 BPD and 9,250 controls with a replication study of 4,496 cases and 42,422 controls to form an analysis of a total 11,974 BPD cases and 51,792 controls. This study confirmed genome-wide significant evidence of association for *CACNA1C* and also identified a new intronic variant in *ODZ4* gene (encoding teneurin-4) (11q14) (Psychiatric GWAS Consortium Bipolar Disorder Working Group1 2011). Later, Chen et al (2013) (Chen et al. 2013) conducted a GWAS meta-analysis including European and Asian populations with 7,773 cases and 9,883 controls and revealed
a significant GWAS signal near *TRANK1* gene. In the same year, another genetic variation found in neurocan gene (*NCAN*) showed genome-wide significant association with BPD in 2,411 patients and 3,613 controls (rs1064395, \( p = 3.02 \times 10^{-8} \) (odd ratio, 1.31) (Cichon et al. 2011). Last year, a large case-control study combined previously reported PGC-BPD data set of 7,481 cases and 9,250 controls with 2,266 cases and 5,028 controls from MooDs (systematic investigation of the molecular causes of major mood disorders and schizophrenia) consortium showed 56 SNPs reaching genome-wide significant association at five genomic loci including previously described risk loci, *ANK3*, *ODZ4*, and *TRANK1*, and one new risk locus *ADCY2* (5p15.31), which is a key enzyme in cAMP signalling and a region between *MIR2113* and *POU3F2* (6q16.1) (Muhleisen et al. 2014).

### 1.2.6.2 GWASs in SCZ

The early GWASs in SCZ used DNA pooling in their methodology to identify candidate genes with higher risk effect, but none of these were reached a Genome-wide (GW) significant threshold. For example, Mah et al. (2006) (Mah et al. 2006) pooled together samples from 320 European individuals with SCZ and identified a candidate locus on chromosome 1q32 (\( p = 6.0 \times 10^{-3} \); OR = 1.49), implicating the *PLXNA2* gene in SCZ. Although the association did not meet GW significance threshold, the authors supported their finding with evidence of consistent association signals in different population groups such as European Americans (OR = 1.38, \( p = 0.035 \)), Latin Americans (OR = 1.26) and Asian Americans (OR = 1.37). Later, Kirov et al (2009) (Kirov et al. 2009) also pooled their DNA samples from unrelated Bulgarian SCZ samples and identified SNP rs11064768 (\( p = 1.2 \times 10^{-6} \)), which is within the gene *CCDC60*, a coiled-coiled domain gene. Their third top SNP was rs893703 (\( p = 1.6 \times 10^{-4} \)), within *RBP1*, which had previously been reported to be a candidate gene for SCZ (Farias et al. 2005, Kalkman et al. 2006).
Later, as GWAS sample sizes increased combined SNP data from several large genome-wide scans found significant association with several markers spanning the major histocompatibility complex (MHC) region on chromosome 6p21.3-22.1, a marker located upstream of the neurogranin gene (NRGN) on 11q24.2 and a marker in intron four of transcription factor 4 (TCF4) on 18q21.2 (Stefansson et al. 2009). The results that implicated the MHC region are consistent with an immune component to SCZ risk, the association with NRGN implicates calcium signalling because NRGN may act as a calcium sensor (Zhong et al. 2009), and the TCF4 finding points to perturbation of pathways involved in brain development, memory and cognition. Another GWAS mega-analysis of SCZ, also from the PGC including over 40,000 individuals, identified one of the strongest novel genetic associations marker rs1625579 within an intron of a putative primary transcript for MIR137 gene (encoding mircoRNA 137) \( (p = 1.6 \times 10^{-11}) \) (Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium 2011), which is a key regulator of neuronal development with roles in neurogenesis and maturation and highly expressed at synapses in the cortex and hippocampus (Smrt et al. 2010, Szulwach et al. 2010, Willemsen et al. 2011).

Just recently, a larger multi-stage SCZ GWAS including 36,989 cases and 113,075 controls identified 128 independent associations spanning 108 loci all meeting genome-wide significance (Schizophrenia Working Group of the Psychiatric Genomics Consortium 2014). Of the 108 loci, 83 had never been previously reported in other GWAS before, and 75% include protein-coding genes and a further 8% are within 20 kb of a gene. Further to their study, associations were enriched among genes expressed in the brain and tissues that play a role in immunity. This second finding provides support for the previous finding that the MHC region of the immune system is associated with SCZ. This GWAS published last year highlighted approximately 35 genes that have the potential to provide entirely new insights.
into aetiology of SCZ such as the glutamatergic neurotransmission pathway, which may hold the key to future target of potential therapeutic relevance to SCZ.

1.2.6.3 Overlapping disorders finding in GWASs

Over the past years, the BPD associated genes have also been found to be a susceptible gene in SCZ suggesting a possible overlap of genetic risk for SCZ, and which may constitute common susceptibility gene for psychosis. For instance, studies of approximately 10,000 individuals revealed strong evidence for association with susceptibility to BPD at variants within two genes involved in ion channel function: ANK3 and CACNA1C (Green et al. 2010). The SNP found in CACNA1C showed maximum association with susceptibility to BPD showed similar association in UK SCZ and unipolar depression samples, indicating that variation at this locus influences susceptibility across the mood-psychosis spectrum. The study also reported analyses of combined SCZ and BPD, with association at three loci reached genome-wide significance: CACNA1C (rs4765905, \( p = 7.0 \times 10^{-9} \)), ANK3 (rs10994359, \( p = 2.5 \times 10^{-8} \)), and ITIH3-ITIH4 region (rs2239547, \( p = 7.8 \times 10^{-9} \)). This increased the levels of statistical significance in the combined analysis suggested genetic variants in these regions influencing risk of both disorders.

A separate study that used a combined dataset of around 20,000 subjects an association signal in the ZNF804A (encoding zinc finger protein 804A) gene on chromosome 2q32 (rs1344706, \( p = 1.61 \times 10^{-7} \)) with SCZ (O'Donovan et al. 2008). In order to investigate whether this association might involve a genetic variant that influenced risk of SCZ and BPD, a BPD sample was further added to the analysis. The effect size remained similar (odds ratio (OR) = 1.12), while the genome-wide significant association \( p \) value became notably more significant (\( p = 9.96 \times 10^{-9} \)). These findings have been further substantiated by a larger meta-analysis consisting of 18,945 subjects with either SCZ or schizoaffective disorder, 21,274 subjects diagnosed with SCZ plus BPD; and control samples of 38,675 and provided
evidence for association between rs1344706 that surpasses widely accepted benchmarks of significance by several orders of magnitude for both SCZ ($p = 2.5 \times 10^{-11}; \text{OR} = 1.10$) and SCZ and BPD combined ($p = 4.1 \times 10^{-13}; \text{OR} = 1.11$) (Williams et al. 2011). These results also suggested that allelic association at the ZNF804A locus to be one of the most compelling in SCZ and that the locus may also has an effect on illness susceptibility across the traditional diagnostic boundaries.

Another combined GWAS data from five psychiatric disorders including SCZ, BPD, major depressive disorder, autism spectrum disorder, and attention deficit hyperactivity disorder (ADHD) was published by the PGC Cross-Disorder Group and revealed four regions that achieved genome-wide statistical significance (Lee et al. 2013). Two of the association signals on chromosomes 3p21 and 10q24 were in regions encompassing a number of genes. The other two association signals were in CACNA1C on chromosome 12 and CACNB2 on chromosome 10. Additional pathways analysis further supported the role of channel activity genes contributing all five disorders. Further, Jan et al (Jan et al. 2014) reviewed findings from two distinct and comprehensive meta-analyses of linkage studies for BP to identify significant linkage peaks (McQueen et al. 2005, Segurado et al. 2003). Among eight selected calcium channel genes, evidence for weak association was founds for CACNA1C (rs10848635), CACNA1E (rs10848635), CACNB2 (rs11013860), and CACNG2 (rs2284018) genes were observed, and combined analysis with independent replication samples further supported the association of marker rs11013860 in CACNB2 with BPD ($p = 1 \times 10^{-6}$).

Despite the fact that powerful GWASs have revealed that common variants or genetic risk loci overlap between BPD and SCZ, clinical diagnostic and molecular distinctions between these disorders remain. An interesting BPD versus SCZ case-case comparison study carried out by our research group found association with a marker rs17645023 ($p = 6.3 \times 10^{-7}$) located between another two of the calcium channel $\gamma$ subunit genes, CACNG5 and CACNG4.
(Curtis et al. 2011). As shown in the Figure 1-1, this marker lies 36kb from CACNG5 and 79kb from CACNG4. The study revealed that the control allele frequencies for this marker were intermediate between those for BPD and SCZ cases. The results suggested that marker rs17645023 may have differential effects in the two disorders and this may help to provide an insight into the fundamental understanding of how these two disorders differ biologically possibly mediated by these two calcium channel genes. Furthermore, these comparisons of BPD and SCZ findings may also provide additional information that could someday be used in clinical diagnostics. Therefore, detailed investigation of brain calcium biology can possibly illustrate the underlying mechanism causing disorder conditions in SCZ and BPD.
Figure 1-1 Association marker rs17645023 lies 36kb from CACNG5 and 79kb from CACNG4 taken from UCSC Genome Browser.
In summary, although GWAS studies have identified genome-wide significant disease-associated variants with some insights into the genetic basis underlying SCZ and BPD, the biological effect of these variants is still unclear. GWAS only inform both conditions with common and different genes being implicated at different levels of statistical significant; thus often missing detection of candidate genes that had been previously identified by other genetic methods. Even though the greater collaborative efforts and larger sample sizes have allowed better examination and detection of common alleles of small effect sizes, there is a need to elucidate the functional significance of the extant genetic findings to better understand the human brain structural changes and biological functions or pathways that underlie the major psychoses. A summary of GWAS findings with strong evidence for association (p < 5 x 10^{-8}) that have small effect (odds ratio < 1.2) are shown in Table 1-2.
Table 1-2 Genome-Wide Association Study Findings for GWA significant genes in Schizophrenia and Bipolar Disorder from the PGC SCZ (Schizophrenia Working Group of the Psychiatric Genomics Consortium. 2014) summary data for 108 genome-wide significant loci. Abbreviations: OR = odd ratio, SCZ = schizophrenia, BPD = bipolar disorder, p-values were taken from the smallest p-value in each individual studies.

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Chromosome where marker is located</th>
<th>Nearest gene</th>
<th>OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCZ</td>
<td>1</td>
<td><em>DPYD; MIR137</em></td>
<td>1.12</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td><em>SDCCAG8</em></td>
<td>1.10</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td><em>CENTG2</em></td>
<td>0.90</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td><em>PCGEM1</em></td>
<td>1.20</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td><em>MHC</em></td>
<td>1.24</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td><em>TRIM26</em></td>
<td>1.15</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td><em>HLA-DRB9</em></td>
<td>p = 9E-14</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td><em>TSNARE1</em></td>
<td>p = 2E-10</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td><em>CNNM2</em></td>
<td>1.10</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td><em>NT5C2</em></td>
<td>1.15</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td><em>NRGN</em></td>
<td>1.15</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td><em>CCDC68</em></td>
<td>1.09</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td><em>TCF4</em></td>
<td>1.20</td>
</tr>
<tr>
<td>BPD</td>
<td>3</td>
<td><em>TRANK1</em></td>
<td>1.13</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td><em>ADCY2</em></td>
<td>1.14</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td><em>SYNE1</em></td>
<td>1.15</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td><em>ANK3</em></td>
<td>1.35</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td><em>CACNA1C</em></td>
<td>1.14</td>
</tr>
<tr>
<td>SCZ and BPD</td>
<td>2</td>
<td><em>ZNF804A</em></td>
<td>1.11</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td><em>ITIH3-ITIH4</em></td>
<td>1.12</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td><em>MAD1L1, SNX8, NUDT1, FTSJ2</em></td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td><em>NT5C2, CNNM2</em></td>
<td>1.23</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td><em>CACNA1C</em></td>
<td>1.07</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td><em>CCDC68</em></td>
<td>1.09</td>
</tr>
<tr>
<td></td>
<td>19</td>
<td><em>MAU2</em></td>
<td>p = 3E-9</td>
</tr>
</tbody>
</table>
1.2.7 Chromosomal Rearrangements and CNV

Chromosomal rearrangements can take the form of translocations between chromosomes that may be balanced or unbalanced. These translocations can lead to disruption of genes at their breakpoints and this has been implicated in disease. The study of copy number variations (CNVs) in the human genome has made substantial discoveries into the genetic aetiology of SCZ and to a lesser extent in BPD. CNVs are structural genomic variations of DNA that include micro-insertions, micro-deletions. The size range thus may vary from over one kilobase (kb) up to several megabases compared to a reference genome. Although SNPs in the genome account for the most numerous variations, CNVs have been reported to involve account for roughly 12% of phenotypic variation within normal individuals and in complex diseases, such as SCZ and BPD (Stankiewicz and Lupski 2010, Sullivan et al. 2012). In 2004, two landmark studies demonstrated a widespread distribution of submicroscopic variations (<500 kb in size) in DNA copy numbers in normal human genomes (Iafrate et al. 2004, Sebat et al. 2004). CNVs can be inherited or appear de novo. The rates of occurrence for de novo CNVs can be up to four fold greater than single nucleotide substitution rates (Lupski 2007) and this class of CNV are considered to be involved in the development of sporadic genomic disease (McCarroll 2008). In comparison to SNPs, the overall proportion of genomic nucleotides involved is high for CNVs leading to a larger functional impact per site (Malhotra and Sebat 2012). Duplication and deletion CNVs can interfere with regulatory regions or coding sequences of various genes leading to altered gene expression and biological function (Freeman et al. 2006, Hurles et al. 2008). Moreover, the effect of single nucleotide disease susceptibility variants can be amplified by the presence of CNVs, where they can cause up-regulation or down-regulation of dosage sensitive genes. The presence of deletion CNVs also has the potential to lead to variations in intermediate phenotypes in complex neuropsychiatric illnesses such as cognitive impairment or in physiological
measures (Friedman et al. 2008). CNVs have also been implicated in the biological basis of SCZ and BPD (Cook and Scherer 2008, Freeman et al. 2006, Porteous 2008).

1.2.7.1 Chromosomal rearrangements and CNV finding in BPD and SCZ

Examples of chromosomal rearrangements detected in patients led to the identification of the disrupted genes and subsequent evidence for their involvement with BPD and SCZ. Examples are translocation t(1; 11) (q42.1;q14.3) disrupting the DISC1 gene (Blackwood et al. 2001, Chubb et al. 2008, Millar et al. 2000), translocation t(1; 16) (p31.2;q21) involving PDE4B that regulate cAMP signalling (Millar et al. 2007, Millar et al. 2005), translocation t(9; 14) (q34.2;q13) disrupting the NPAS3 gene (Kamnasaran et al. 2003, Pickard et al. 2009, Pickard et al. 2006). Specific CNVs with potential roles have been investigated both in locus-specific and genome-wide studies for BPD and SCZ. Locus-specific studies have investigated CNVs in three genes in SCZ and/or BPD including a complex CNV at 15q13-14 containing a polymorphic inversion involving CHRNA7 and its fusion variant CHRFAM7A that might be implicated in BPD, SCZ, and more general psychosis (Flomen et al. 2006, Flomen et al. 2008, Freedman et al. 1997, Riley et al. 2000); a duplication at 3q13.3 disrupting the functional candidate GSK3beta (glycogen synthase kinases 3 beta) and two other genes, was shown to occur more frequently in BPD (Lachman et al. 2007); and a rare deletion at 7q34-36.1 affecting the CNTNAP2 gene was detected in three unrelated patients with both SCZ and epilepsy (Friedman et al. 2008). The 22q11.2 deletion syndrome (22q11.2DS) is a common syndrome with congenital and late-onset features and is caused by a 3Mb hemizygous microdeletion, which has also been known as a significant risk factor for SCZ, BPD, VCFS, and Di George syndrome (DGS) (Bassett and Chow 2008, Karayiorgou et al. 1995, Murphy et al. 1999).
1.2.7.2 CNV findings in SCZ

To date, CNV studies have established that rare structural variations (occurring in less than 1% of the population) are a significant risk factor for SCZ (Sebat et al. 2009). The International SCZ Consortium (International Schizophrenia Consortium 2008) reported a 1.1 – 1.5 fold enrichment of rare gene CNVs in cases compared to controls; in addition, other groups observed a 3-fold enrichment of rare CNVs in SCZ cases (Walsh et al. 2008). These findings have further been supported by several subsequent studies (Buizer-Voskamp et al. 2011, Kirov et al. 2009), and confirm that rare CNVs are collectively more common in SCZ than in normal subjects. Recent genome-wide SCZ CNV studies have found strong evidence for disease association with other loci including deletions at chr1q21.1, deletions at chr3q29, duplications of chr16p11.2, deletions at chr15q13.3, exonic deletions at chr2p16.3 (NRXN1) and duplications at chr7q36.3 (VIPR2). Early on, it was apparent that rare CNVs tended to impact genes involved in neuronal function (Walsh et al. 2008). These included functional categories related to synaptic activity and neurodevelopment (Malhotra et al. 2011, Walsh et al. 2008). De novo CNVs were significantly enriched in specific protein complexes such as N-methyl-d-aspartate receptor (NMDAR) and neuronal activity-regulated cytoskeleton-associated protein postsynaptic signalling complexes as well as other components of the postsynaptic density (Kirov et al. 2012).

1.2.7.3 CNV findings in BPD

By contrast, the literature available on CNV studies of BPD is very limited (Lachman et al. 2007), consequently the result of case-control studies has been inconsistent. While two studies have reported an enrichment of rare CNVs in bipolar disorder (Priebe et al. 2012, Zhang et al. 2009) with greater effects in subjects with an early age-at-onset, these effects were not further supported by other two studies (Grozeva et al. 2010, McQuilllin et al. 2011). Notably, de novo mutations from mood disorders were showed significantly higher
frequencies (4.3%) in BPD as compared with healthy individuals (0.09%). The rate of de novo CNVs among cases with an age at onset younger than 18 was higher still (5.6%) (Malhotra et al. 2011). The limited CNV findings in BPD have meant that it has not been possible to implicate specific genes or genomic regions in BPD; similarly pathway enrichment analyses have also been unable to illustrate clear patterns. Pathways enriched among de novo CNVs in BPD showed an enrichment of genes involved in regulation of cell shape, but no significant enrichment of genes involved in neuronal function or development (Malhotra et al. 2011).

1.2.7.4 Summary

Although CNV studies have found several variants that are more frequent in patients than in controls with odds ratios of >10 for some variations, the frequency of each individual variant in BPD or SCZ patients is low (<1%). Therefore future studies are required to determine the penetrance and mutation rate of these structural variants, as well as their phenotypic spectrum. Research has shown that some CNVs also occur more frequently in patients with other neuropsychiatric phenotypes such as autism, mental disability and epilepsy (Ben-Shachar et al. 2009, Helbig et al. 2009, Mefford et al. 2008, Miller et al. 2009). However, the mechanism underlying the phenotypic outcome still remains unclear.
1.3 Voltage-gated calcium channel (VGCC)

The voltage-gated calcium channels (VGCCs) are heteromeric proteins composed of a principle pore-forming alpha 1 (α1) subunit and auxiliary subunits: α2δ, β, and γ (Figure 1.2.a). The α1 subunit contains the voltage sensor that controls channel activation and gating. Channel activation and gating is generally regulated by second messengers, drugs, or toxins (Vacher et al. 2008). According to their physiological and pharmacological properties, three subfamilies of this subunit are known to exist: Ca\textsubscript{v}1 (L-type), Ca\textsubscript{v}2 (P/Q-, N- and R-type) and Ca\textsubscript{v}3 (T-type) as also shown in Figure 1.1.b. The function of auxiliary subunits α2δ and β are to regulate VGCC current. For instance, when co-expressed with the α\textsubscript{1} subunit, the trafficking of the channel complex to the surface membrane was increased by these auxiliary subunits and this resulted in increased current density (Dolphin 2003, Klugbauer et al. 2003). Moreover, these subunits also alter the biophysical properties by increasing activation and/or the opening probability of channel pore (Dolphin 2003, Klugbauer et al. 2003, Neely et al. 1993).
Figure 1-2 Structure of voltage-gated calcium channel α1 subunits and its subclasses. a) Structure of voltage-gated calcium channel α1 subunits; b) Three subclasses of α1 subunit according to their amino acid sequence identity (Dolphin 2012).

In expression studies, different classes of the α1 subunit and various auxiliary subunits can be found in many other tissues. The exception to this widespread expression is the γ (gamma) subunit, which was thought to be found in skeletal muscle (Klugbauer et al. 2000). When the γ subunit is co-expressed with the α1 subunit of L-type calcium channels, it appears to alter peak current and the activation and inactivation kinetics (Eberst et al. 1997, Singer et al. 1991, Wei et al. 1991). More recently, a second type of γ subunit was identified in mice with epilepsy, which was found to be due to the absence of γ subunit gene (Letts et al. 1998). The next section will discuss more about the γ subunit family.
1.3.1 \( \text{Ca}^{2+} \) channel \( \gamma \) subunits

Current literature reviews of the cellular processes influenced by the \( \gamma \) subunits suggest that they are a functionally diverse protein family. The \( \gamma \) subunits interact not just with calcium (\( \text{Ca}^{2+} \)) channels, but with other proteins as well. In fact, the recent evidence suggests that the principle cellular targets of several members of this subunit family may not be \( \text{Ca}^{2+} \) channels at all.

This first \( \gamma \) subunit described, \( \gamma_1 \), was isolated biochemically as a component of a calcium channel expressed in skeletal muscle and has been shown to alter calcium current properties in both native myocytes and in cell lines (Eberst et al. 1997, Freise et al. 2000, Held et al. 2002, Singer et al. 1991). Later, Letts (Letts et al. 1998) described a second \( \gamma \) subunit gene from stargazer mouse, Cacng2. This gene encodes a 36kD protein (stargazin/\( \gamma_2 \)) with sequence similar to the \( \gamma_1 \) subunit. Stargazer mice carry a mutant stargazin and have a complex neurological disorder including ataxia, inner-ear defects, and epilepsy (Noebels et al. 1990). Eventually, additional putative \( \gamma \) subunit genes in human, such as \( \gamma_2 - \gamma_5 \), were identified via conserved structural features of \( \gamma_1 \) and \( \gamma_2 \), where they were shown to share functional similarity with \( \gamma_2 \) (Burgess et al. 1999). Furthermore, phylogenetic analysis identified additional \( \gamma \) subunit family genes, \( \gamma_6 - \gamma_8 \): \( \gamma_6 \) might be important for \( \text{Ca}^{2+} \) channel regulation; whereas the \( \gamma_7 \) and \( \gamma_8 \) share functional similarity with \( \gamma_2 \) (Burgess et al. 2001).

Historically, the studies of co-expressed stargazin/\( \gamma_2 \) showed only a modest or even no effect on VDCC currents in heterologous cells (Green et al. 2001, Kang et al. 2001, Klugbauer et al. 2000, Letts et al. 1998). On the other hand, stargazin/\( \gamma_2 \) was discovered to prominently associate with and regulate the functions of alpha- amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptors (AMPARs) (Chen et al. 2000). Furthermore, three closely related isoforms of stargazin/\( \gamma_2 \), \( \gamma_3 \), \( \gamma_4 \), and \( \gamma_8 \), had the ability to substitute for stargazin/\( \gamma_2 \) in...
regulating AMPARs in *in vitro* expression studies. These four proteins were used to define the new family of transmembrane AMPAR regulatory proteins (TARPs) (Osten and Stern-Bach 2006, Qiao and Meng 2003). More recently, TARPs were classified into two groups on the basis of their homologous sequence alignment with type I TARPs comprising the subunits stargazin/γ2, γ3, γ4, and γ8, and the type II TARPs, γ5 and γ7. Because of differences in their sequence alignment, it is thought that the two classes of TARPs differentially modulate channel properties and/or trafficking of AMPARs in specific neuronal pathway (Cho et al. 2007, Haering et al. 2014, Kato et al. 2010, Kato et al. 2007, Tomita et al. 2005a, Tomita et al. 2003).

1.3.1.1 Distribution and cellular functions of γ subunits

The γ subunits are sub-divided into three clusters: γ1 and γ6, type I TARPs, and type II TARPs via their sequence homology. These three clusters have been shown to have a distinct distribution of their subunits in tissues, and also in their cellular functions. γ1 and γ6 are predominantly expressed in skeletal muscle, while the TARPs are differently expressed in brain regions and throughout development (Tomita et al. 2003). For example, Stargazin/γ2 has been found to mainly expressed in the cerebellum, but also occurs in the cerebral cortex and in both CA3 and dentate gyrus regions of hippocampus (Letts et al. 1998, Moss et al. 2003, Sharp et al. 2001, Tomita et al. 2003). Stargazin/γ2 is the only TARP found in cerebral granule cells, explaining the selective loss of AMPAR function in these cells in stargazer mice. γ7 is another subunit that has its highest expression level also in the cerebellum, but is localised in the somatodendritic regions of Purkinje cells and in glomeruli of the granule cell layer (Kato et al. 2007). γ3 on the other hand is found mainly in the cerebral cortex; whereas γ8 is predominantly expressed in the hippocampus (Tomita et al. 2003). By contrast, γ4 shows diffuse expression throughout the brain with local enrichment in the olfactory bulb (Tomita et al. 2003). Interestingly, within the CNS, γ4 is the only TARP found in non-
neuronal cells, particularly in glial cells. These different expression patterns of the $\gamma$ subunits indicate possible TARP-specific characteristics, which may conform to the different regulatory requirements of the different brain regions.

Although the cellular distribution is distinct between the $\gamma$ subunit, each subunit shares a common topology that consists of four transmembrane domains, with intracellular N- and C-(carboxyl) termini. The $\gamma$ subunits are also known as a tetraspanin protein superfamily members, which includes proteins called claudins that are important components of tight junctions in the epithelia (Van Itallie and Anderson 2006). These tetraspanin proteins consist of a highly conserved N-glycosylation site (GLW) and a pair of conserved cysteine residues within the first extracellular loop that may form a disulphide linkage as shown in Figure 1-3.

Figure 1-3 Schematic diagrams showing predicted membrane topology and putative functional sites on $\gamma$s and claudins (Chen et al. 2007).
1.3.1.1.1  \(\gamma_1\) and \(\gamma_6\) cluster

\(\gamma_1\) and \(\gamma_6\) are distinct from the other TARPs in that they have short C-terminal cytoplasmic regions that lack functional motifs (Figure 1-3). The \(\gamma_6\) subunit distribution is different to that of the \(\gamma_1\) subunit that only expresses in skeletal muscles. It appears to be expressed not only in skeletal muscles but also in cardiac muscles and to a lesser extent in the brain (Burgess et al. 2001, Chu et al. 2001, Fukaya et al. 2005). It is also the only member of the \(\gamma\) subunit that is also expressed as a short isoform that lacks the second and third transmembrane domains (Chu et al. 2001).

1.3.1.1.2  Type I TARPs cluster

Type I TARPs identified as regulators of AMPAR function are widely expressed in the brain and share highly conserved sequences that are different from the \(\gamma_1\) and \(\gamma_6\) cluster (Arikath and Campbell 2003, Black 2003). Notably, that the type I TARPs have the longest C-terminal cytoplasmic tails of the \(\gamma\) subunit family. These cytoplasmic tails include a number of regulatory sites including a PDZ-binding motif (TTPV) that controls AMPA receptor targeting to the synapse (Figure 1-4) (Arikath and Campbell 2003, Black 2003). In addition, this binding motif (TTPV) in stargazin/\(\gamma2\) contains a consensus threonine residue for phosphorylation by cAMP- and cGMP-dependent kinases, and the tails are also targets for binding by PDZ domain proteins (Burgess et al. 2001). This suggests phosphorylation may regulate the interaction between stargazin/\(\gamma2\) and the AMPARs.
Figure 1-4 The protein structures of type I and type II TARPs. Panel A represents type I TARPs that have a larger first extracellular domain (EX1) and a canonical PDZ-binding domain (T-T-P-V) located at their c-terminal cytoplasmic. Panel B represents type II TARPs, which have an atypical PDZ-binding domain (S/T-S-P-C). RS domain represents arginine-serine-rich domain (Kato et al. 2010).

As mentioned earlier, TARPs are differently expressed throughout development: stargazing/γ2, γ3, and γ8 are sparingly expressed in the brains of newborns and reach their highest expression levels in adult brains; by contrast, γ4 presents the opposite characteristics, showing increased expression early in development. Furthermore, γ4 are also found in embryonic epithelial cells lining the intestines (Tomita et al. 2003).

1.3.1.1.3 Type II TARPs cluster

The difference in structure between type I and type II TARPs is in their C-terminal tail, where type II TARPs lack a PDZ-binding motif (Chu et al. 2001, Moss et al. 2002). Further, these type of TARPs are encoded by five exons, whereas the other γ subunits are encoded by four exons (Chu et al. 2001, Moss et al. 2002).
1.4 Glutamate receptors

The mammalian central nervous system is comprised of an incredibly complex web of connectivity between billions of neurons that are highly specialised for the fast processing and transmission of cellular signals. Communication between neurons, each of which contains thousands of synapses, underlies all basic and higher-order information processing essential for normal brain function including learning and memory. The ability of neural circuits to strengthen or weaken their connectivity forms the molecular basis underlying the experience-dependent changes in adaptive behaviours. Glutamate is the most abundant excitatory neurotransmitter in the mammalian central nervous system (CNS) (Curtis et al. 1959, McLennan et al. 1968). Once nerve impulses are triggered to release glutamate from the pre-synaptic terminals, it binds to post-synaptic glutamate receptors (GluRs) for post-synaptic membranes depolarisation. These depolarisations may generate action potentials in the postsynaptic neuron that transmit information to subsequent neurons in the neuronal circuit (Meldrum 2000). The mechanisms of synaptic transmission are crucial factors for the proper function of neuronal communications and connections. These mechanisms are mediated not only by altering the concentration of glutamate in the synaptic cleft, but also by the number and channel properties of GluRs. Therefore, it is important to understand the physiological mechanisms that control the abundance of GluRs and their functional properties at synapses.

GluRs can be classified into two functionally distinct categories according to their role in the activation of postsynaptic current elevation. Metabotropic glutamate receptors (mGluRs) indirectly activate ion channels on the plasma membrane through a signalling cascade of G-protein second messenger systems, in order to modulate synaptic plasticity. Whereas ionotropic glutamate receptors (iGluRs) are ligand-gated ion channels that mediate fast excitatory neuro-transmission at synapses within the CNS. Binding of presynaptically
released glutamate triggers conformational changes that result in a depolarising cation current through the iGluR channel pore to allow ion flow. Based on their pharmacology and molecular characteristics, iGluRs are classified into three distinct subgroups namely: N-methyl-D-aspartate receptors (NMDAR), α-amino-3-hydroxy-5-methylisoxazole-4-propionate receptors (AMPAR), and kainite (KA) receptors (Hollmann and Heinemann 1994, Nakanishi 1992, Wisden and Seeburg 1993). In Figure 1-5, when glutamate released from presynaptic terminal and binds on both NMDARs and AMPARs, calcium (Ca$^{2+}$) and sodium (Na$^{+}$) ions flow into neurons. Ca$^{2+}$-dependent enzymes are activated by the influx of Ca$^{2+}$ at postsynaptic sites, and trigger downstream signal cascades that are involved in synaptic plasticity. The distribution of AMPARs in the postsynaptic density (PSD), as well as the number of synaptic AMPARs play an important role in modulating synaptic strength (Bredt and Nicoll 2003, Collingridge et al. 2004, Kessels and Malinow 2009, Lisman 2003, Malenka and Bear 2004, Newpher and Ehlers 2008, Shepherd and Huganir 2007). During long-term potentiation (LTP), which is a type of activity-dependent plasticity that involves neuronal learning, NMDARs mediate the influx of Ca$^{2+}$ and lead to the trafficking of AMPARs to the plasma membrane (Bredt and Nicoll 2003, Collingridge et al. 2004, Kessels and Malinow 2009, Lisman 2003, Malenka and Bear 2004, Newpher and Ehlers 2008, Shepherd and Huganir 2007). Conversely, during long-term depression (LTD), NMDARs or mGluRs mediate the influx of Ca$^{2+}$ and there is a subsequent lateral diffusion of AMPARs from the postsynaptic density (PSD) (Bredt and Nicoll 2003, Collingridge et al. 2004, Kessels and Malinow 2009, Lisman 2003, Malenka and Bear 2004, Newpher and Ehlers 2008, Shepherd and Huganir 2007). Knowledge of the mechanisms regulating AMPARs trafficking allows one to form an understanding of synaptic transmission and ultimately brain function.
Figure 1-5 Synaptic transmission and AMPAR trafficking. Glutamate released from presynaptic terminal acts on NMDARs and AMPARs at the synapse. (A) AMPARs trafficking to synapse: AMPARs at intracellular vesicles are trafficked to the cell surface by exocytosis. Surface AMPARs laterally diffuse and are stabilised at the PSD. (B) AMPARs removal from synapse: AMPARs laterally diffuse from the PSD via intracellular vesicles endocytosis.

1.4.1 AMPA receptors

AMPARs are ligand-gated ion channels of the iGluRs family that mediate the majority of fast excitatory synaptic transmission in the brain to regulate synapse stabilisation and plasticity (Malinow 2003). Since the function of AMPARs is critical for synaptic plasticity, their dysfunction leads to various neurological and psychiatric disorders, such as Alzheimer’s disease, amyotrophic lateral sclerosis, X-linked mental retardation, limbic encephalitis, ischemic brain injury, and Rasmussen’s encephalitis (Hsia et al. 1999, Hsieh et al. 2006) (Kamenetz et al. 2003, Kawahara et al. 2004, Lai et al. 2009, Lynch 2006, Rogers et al. 1994, Shepherd and Huganir 2007, Soundararapandian et al. 2005, Talos et al. 2006a, Talos et al. 2006b, Wu et al. 2007). Several characteristics have been suggested to influence the function of AMPARs, such as their subunit assembly mechanisms, conformational changes, modulation by auxiliary proteins, trafficking, and localisation. Therefore, the understanding of the cell biology and the biophysical properties of AMPARs are important to illustrate the mechanism of synaptic plasticity and assist in identifying therapeutic measures that may be
used as treatment for the diseases that are caused changes in either AMPARs or their auxiliary proteins.

1.4.1.1 Primary structure of the AMPAR subunits

AMPARs are assembled by four highly homologous subunits GluR1 to GluR4 (or GluRA-D) (Wenthold et al. 1996); recently renamed GluA1-A4, (Collingridge et al. 2009). Each subunit is comprised of four major domains, which has approximately 900 amino acids with a molecular weight of approximately 105 kDa (Granger et al. 1996). The predicted subunit topology is shown in Figure 1-6 A and B. An extracellular amino (N-) terminus forms a highly homologous N-terminal domain (NTD), which consist of about 370 amino acids. This NTD is followed by a linker sequence and connects to another extracellular domain, the ligand binding domain (LBD). The primary structure of the LBD domain can be subdivided into two separate fragments known as S1 and S2 (Armstrong et al. 1998). The LBD undergoes conformational changes resulting in channel gating upon glutamate binding. The fourth membrane spanning segment (M4) is located at the C-terminal end of the S2 fragment. The channel pore-forming transmembrane domain (TMD) consists of M1-M4 (Hollmann et al. 1994). Finally, a relatively varied C-terminal domain (CTD) extends into the cytoplasm ending with a PSD-95 discs large-zona occludens 1 (PDZ)-binding motif (Hollmann and Heinemann 1994). The PDZ-binding motif is responsible for interaction with cytosolic proteins that regulate receptor anchoring and trafficking (Malinow and Malenka 2002, Sheng and Lee 2001, Ziff 2007).
Figure 1-6 The structure of AMPA receptors. Panel A represents as primary structure of a subunit of AMPA receptor (NTD represents as N-terminal domain; CTD receptors as C-terminal domain); Panel B represents as domains of a subunit of AMPA receptor (TMD represents as transmembrane domain); Panel C represents as tetrameric structure of AMPA receptors; Panel D represents a negative stain single particle electron microscope image of an AMPA receptor. (Nakagawa 2010)

Functional AMPARs are a tetrameric complex, which are assembled from four separate subunits (Robert et al. 2001, Rosenmund et al. 1998). Assembly of tetrameric AMPARs are generally comprised of two identical heteromeric subunits dimers. The identical subunits are positioned on opposite sides of the channel pore (Figure 1-6 C and D). However, an individual AMPAR subunit is also capable of assembly into a functional homotetrameric ligand-gated ion channel, when over-expressed (Swanson et al. 1997). Several studies have revealed that the majority of AMPARs are composed of GluR2 subunits in hippocampal pyramidal neurons, such as GluR1/2 and GluR2/3 (Greger et al. 2003, Isaac et al. 2007, Lu et al. 2009). Single cell genetic approaches have suggested that GluR1/2 is the most predominant subtype, in approximately 80% of synaptic AMPARs and only 16% are GluR2/3 (Lu et al. 2009). GluR2 knock-out mice studies showed that AMPARs were not
efficiently targeted to synapses and were aberrant in their stoichiometry, suggesting a critical role for GluR2 in regulating the assembly of AMPARs (Sans et al. 2003).

Upon ligand binding AMPARs allow the influx of sodium ions (Na\(^+\)), and in the absence of the GluR2 subunit they are also permeable to calcium ions (Ca\(^{2+}\)) (see section 1.4.1.1.3). The subunit composition of AMPARs and their post-translational modification (see later) have been shown to regulate the biophysical properties of receptors including channel gating kinetics, synaptic plasticity (Cull-Candy et al. 2006, Hollmann and Heinemann 1994, Isaac et al. 2007, Jonas 2000, Seeburg 1993).

1.4.1.1.1 Long and short splice variants

The molecular diversity amongst AMPARs includes alternative splicing and RNA editing (Lomeli et al. 1994). One of the alternative splice sites is located in the C-terminal coding region leading to short and long forms of the protein. This region comprises of between 30 and 50 amino acids depending on the subunit. GluR2 and GluR4 genes are the only subunits that can undergo alternative splicing to form either long or short CTD sequences, (Dingledine et al. 1999, Hollmann and Heinemann 1994, Sobolevsky et al. 2009). To date, GluR1 has only been found to have long C-terminal sequences, which is homologous to GluR4 and the alternatively spliced form of GluR2 (GluR2L); whereas the C-terminus of GluR3 is only found as the short form, and is homologous to GluR2 and an alternatively spliced form of GluR4 (GluR4S) (Figure 1-7). Expression of these alternative isoforms is developmentally regulated and region specific; for instance, approximately 90% of GluR2 is expressed as the short form in vivo (Kohler et al. 1994). GluR4S is preferentially expressed in cerebellar granule cells and Bergmann glial cells (Gallo et al. 1992).
Figure 1-7 Structure of AMPAR subunit and its direct interacting proteins. Tetrameric AMPARs are assembled from two dimers of different subunits, such as GluR1/2 and GluR2/3. The GluR1 C-terminal domain contains type I PDZ ligand and directly interacts with SAP97, whereas the short isoform of GluR2 C-terminal domain contains type II PDZ ligand and interacts directly with PICK1 and GRIP1. In addition, GluR1 also interacts with protein 4.1N through its juxtamembrane region of the C-terminus, while GluR2 interacts with AP-2, NSF and BRAG-2 through it C-terminus in a non PDZ-dependent manner. Direct binding of AMPARs and these interacting proteins regulates various steps in AMPAR trafficking (Anggono and Huganir 2012).

Moreover, the long-tailed AMPARs are important for the activity-dependent AMPAR trafficking to synapses during synaptic strengthening, such as LTP; whereas the short-tailed AMPARs appear to constitutively recycled in and out of synapses in the absence of activity, while internalisation of both forms of AMPARs occurs during activity-dependent synaptic weakening, such as LTD (Kessels and Malinow 2009).

1.4.1.1.2 Flip/flop splice variants

The AMPAR subunit genes undergo alternative splicing resulting in flip and flop modules that differ by only 9 amino acids embedded in the region encompass the S2 sub-fragment of the LBD and the linker between the LBD and M4 (Sommer et al. 1990). Nucleotide sequences of the mRNAs of AMPAR subunits are enzymatically modified by the deaminase ADAR2 (adenosine deaminase acting on RNA 2) (Higuchi et al. 2000, Schmidt et al. 2014) introducing amino acids that are not encoded by the genome. This post-transcriptional modification known as RNA editing, occurs in two locations in the AMPAR subunits
(Sommer et al. 1991). The first is the R/G site located within the LBD immediately before the flip/flop splice site in GluR2, 3, and 4 subunits; and the other is the Q/R site only in GluR2 subunit located in the transmembrane domain of M2 region (see later) (Figure 1-8) (Higuchi et al. 1993, Lomeli et al. 1994).

Figure 1-8 A linear sequence of AMPA receptor and its formation into the cell membrane. Transmembrane domain M1 – M4 are boxed and inserted into the cell membrane. M1, M3, and M4 span through the lipid bilayer, whereas M2 enters and exits the membrane from the intracellular side. The alternatively spliced flip/flop (F/F) form in AMPAR subunits showed in a grey box, and the editing (R/G) and (Q/R) sites are indicated by lollipops (Seeburg 1996).

As shown in Figure 1-8, an arginine codon (AGA) for the amino acid of the R/G site in AMPAR subunits is adjacent to the alternative spliced exons that encode the flip/flop module. These exons begin with the adenosine in an unstable position of the preceding arginine codon, which is split by an intron (Figure 1-9). The first adenosine of the arginine codon can be altered by pre-mRNA editing to a GGA (glycine) codon. All AMPAR subunit transcripts can exist in both arginine (R) and glycine (G) versions at this site, except GluR1 that occurs only in the gene-specified arginine version. In addition, the five positions of AMPARs in the strategic segment are targeted for an arginine, where is converted to glycine via RNA editing
at the R/G site of GluR2-4. The four residues along the secondary structure of helices J and K are altered in GluR1-4 by mutually exclusive alternative splicing (Sommer et al. 1990), the default flip exon is replaced by the flop exon in a developmentally regulated manner (Figure 1-9).

Figure 1-9 Arrangement and sequence of the alternative RNA editing to form flip and flop. The upper diagram describes the position of the R/G editing site at the 3’ end of the exon 13 (red arrow head), the editing complementary sequence in intron 13 (ESC; red rectangle), and the mutually exclusive flop (green), and flip (yellow) exons. The sequence of the splice donor site is shown. Editing switches the adenosine to inosine is also show in the red letters. The lower diagram showed sequence alignment encompassing the GluR2 R/G site and flip/flop segment along helices J and K. alternative residues are highlighted in bold, the R/G site at position 743 is indicated in blue (Penn and Greger 2009).

Studies have suggested that the flip/flop modules display distinct but partially overlapping temporal and spatial expression patterns in the nervous system (Monyer et al. 1991, Sommer et al. 1990). Generally, flop variants desensitise faster, although the kinetic difference depends on the subunit and, for heteromeric channels, on subunit composition (Grosskreutz et al. 2003, Koike et al. 2000, Mosbacher et al. 1994, Sommer et al. 1990). The flip/flop splicing has an unprecedentedly strong effect on the maturation and cellular trafficking of AMPARs in several cell lines: homomeric GluR4-flop receptors accumulate in the endoplasmic reticulum (ER) whereas the corresponding GluR4-flip receptors are efficiently
transported to cell surface (Coleman et al. 2006). Different transport competence of the isoforms is independent of desensitisation and receptor activity. These differences are also manifested in secretion of the soluble ligand-binding domain, and are determined by residue 780 outside helix J. The transport block of the flop isoform is reversible and surface expression of GluR4-flop is completely rescued by co-expression with stargazin (Chen et al. 2000). These results suggest an interaction between the extracellular flip/flop region and isoform-specific ER luminal proteins as a novel step in the early trafficking of AMPARs (Coleman et al. 2006).

1.4.1.1.3 Ca²⁺ impermeability of AMPAR subunits

The presence of the GluR2 subunit has a profound impact on the biophysical property of the AMPARs. GluR2-containing AMPARs are impermeable to Ca²⁺ and Zn²⁺ with electrically linear current-voltage relationship; whereas AMPARs lacking GluR2 subunits are permeable to Ca²⁺ and Zn²⁺, also and have an inwardly rectifying current-voltage relationship (Figure 1-10) (Isaac et al. 2007). GluR2-containing AMPARs generally undergo a post-transcriptional modification by RNA edition of a glutamine (Q) to arginine (R) substitution in the M2 region resulting in Ca²⁺ impermeability (Figure 1-8 & Figure 1-10a) (Sommer et al. 1991). This substitution occurs approximately in 99% of postnatal GluR2 (Sommer et al. 1991). Therefore, a modest alteration in the level of GluR2 expression has the potential expected to have profound implications for synaptic efficacy and neuronal survival. Collectively, the alternative splicing and RNA editing adds to the molecular complexity and functional variety of AMPARs in the brain.
Figure 1.10 Ca^{2+} permeability of AMPARs depends on the subunit composition. Panel (a) presenting GluR2-lacking AMPARs are highly permeable to Ca2+. Panel (b) presenting GluR2-containing AMPARs are impermeable to Ca2+. The presence of the GluR2 in heteromeric AMPAR channels limits Ca2+ and Zn2+ influx, owing largely to the presence of positively charged R instead of a Q residue at the Q/R RNA editing site (Liu and Zukin 2007).

1.4.2 Trafficking of AMPARs at synapses

The trafficking of AMPARs at synapses has been proposed as three-step model (Opazo and Choquet 2011): (i) exocytosis of intracellular AMPARs at perisynaptic sites to the plasma membrane, (ii) lateral diffusion of surface AMPARs at synaptic sites and (iii) stabilisation of AMPARs at synapses via scaffold interactions (Figure 1-5). Activity-dependent alteration of AMPAR trafficking regulates long-term changes in synaptic efficacy that underlie physiological phenomena. An increased exocytosis of AMPARs to the plasma membrane (at the PSD sites) results in the LTP of strengthening synapses; whereas endocytosis of synaptic AMPARs leads to LTD (Opazo and Choquet 2011). Therefore, impaired AMPAR trafficking could lead to cognitive dysfunction that is associated with neurological disorders such as Alzheimer’s disease and epilepsy (Keifer and Zheng 2010). This contribution of molecular pathology has stimulated attempts to investigate the auxiliary subunits of AMPAR that modulate its trafficking and/or channel properties. Since several molecules have been recognised to be auxiliary subunits of iGluR such as TARPS, the Cornichon-like proteins (CNIHs), and the Neurophilin and Tolloid like protein (Netos), TARPs will be the only
auxiliary proteins to be described in this thesis (Diaz 2010, Jackson and Nicoll 2011a, Tomita 2010).

1.4.3 TARPs as auxiliary protein of AMPARs

Transmembrane AMPA receptors regulatory proteins (TARPs) were the first AMPA auxiliary proteins to be identified (Coombs and Cull-Candy 2009, Nicoll et al. 2006, Sager et al. 2009). TARPs are encoded by certain types of calcium channel γ subunit genes (described in section 1.3.1). There are six TARP subunits (γ-2, γ-3, γ-4, γ-5, γ-7, and γ-8), and these are classified into two types: Type I and Type II. The distinction between these two types is the C-terminal cytoplasmic tails: Type I TARPs comprised a longer C-terminal sequence contains a typical PDZ-binding motif, whereas a shorter C-terminal sequence in Type II TARPs contains atypical PDZ-binding site (Figure 1-4). The importance of these PDZ-binding sites at the end tail of TARPs is to interact with PSD-95 for efficiently transport of receptors to the postsynaptic membrane of cerebellar granule cells (Chen et al. 2000). The role of TARPs as auxiliary proteins of AMPARs has been firmly illustrated, where the channel properties of AMPARs are regulated by the stable binding of AMPARs to TARPs in the brain (Cho et al. 2007, Fukata et al. 2005, Kato et al. 2007, Milstein and Nicoll 2009, Nakagawa et al. 2005, Priel et al. 2005, Schwenk et al. 2012, Tomita et al. 2005a, Vandenberghe et al. 2005b). TARPs are non-pore-forming proteins (Tomita et al. 2004) that regulate the trafficking of AMPARs, and this is then required for synaptic AMPAR activity. Hence, the presence of TARPs is essential for the regulation of AMPARs at the synapses and lead to the regulation of synaptic plasticity.
1.4.3.1 TARPs regulate trafficking of AMPARs to the surface

The idea of TARPs regulating the surface trafficking of AMPARs has been confirmed by several research experiments. For example, the surface expression of AMPARs was increased, when the heterologous combinations of TARP isoforms were overexpressed in stargazin mice (Kato et al. 2007, Priel et al. 2005, Tomita et al. 2005a, Turetsky et al. 2005, Vandenberghe et al. 2005a, Yamazaki et al. 2004). A similar result was also obtained when TARPs were overexpressed in neurons and this led to increased surface expression of AMPARs both at the soma and in dendrites (Rouach et al. 2005, Schnell et al. 2002, Turetsky et al. 2005). Conversely, both cerebellar granule cells and hippocampal pyramidal cells of mice deficient in TARP-γ2 and TARP-γ8 respectively, had reduced ionic currents mediated by reduction of surface AMPARs compared with wild-type mice (Chen et al. 2000, Fukaya et al. 2006, Rouach et al. 2005).

1.4.3.1.1 TARPs regulate the AMPARs retention pool in the endoplasmic reticulum

Because TARPs mediate the surface trafficking of AMPAR, they may also associate with AMPAR assembly in the endoplasmic reticulum (ER), and generate a facilitating receptor complex for exit from the ER (Figure 1-11). Successive formation of dimers and tetramers of many integral membrane proteins in the ER is precondition for export from the ER; hence, the AMPAR assembly is required for escape from the ER. Although a definitive location for the first interaction of TARPs and AMPARs has not been identified, FRET studies have indicated that TARPs and AMPRs initially meet at a short distance from the ER (Bedoukian et al. 2006). Furthermore, the immature glycosylation of AMPARs were retained in the ER when the cerebellum of stargazer mice deficient TARP-γ2 (Tomita et al. 2003). These results suggest that TARPs generate export-competent receptors to facilitate AMPAR exit from the ER. Even though the mechanism of AMPAR assembly in the ER is still unclear, two basic possible steps are thought to be involved: (i) the interaction of two monomer N-termini of
each of the GluR subunits form a dimer AMPAR; (ii) then the dimerisation of these dimers occurs through interaction by the membrane-spanning regions and extracellular loops of the subunits (Ayalon and Stern-Bach 2001). This brief overview indicates that TARPs maintain an extrasynaptic pool for surface trafficking of AMPARs.

Figure 1-11 TARPs generate export-competent receptors to existing AMPAR leading to synaptic trafficking (a) AMPARs are assembled in the ER by successive formation of dimers and tetramers. TARPs associate with AMPARs in the ER and generate an ER export-competent receptor, possibly by facilitating assembly through association with subunit monomers or dimmers (not shown), displacing an ER retention factor, or by stabilising an ER export-competent receptor conformation. (B) After exiting the ER, the receptor-TARP complex traffics to the plasma membrane where it is inserted at an extrasynaptic site. The complex diffuses to the synapse, where interaction of the TARP C-terminal PDZ binding site with the synaptic scaffolding protein PSD95 anchors the complex at the synapse.
1.4.3.1.2 TARPs regulate early trafficking of AMPARs

Before AMPAR trafficking to the plasma membrane, they are required not only to exit from the ER, but they also need to travel through the Golgi apparatus. Trafficking and synaptic targeting of AMPAR involves the extreme C-terminus of TARPs (sequence, -TTPV) binding to PDZ domain-containing protein as PDZ-95 (Figure 1-4 A and 1-11 b). Another protein microtubule associated protein 1 light chain 2 (MAP1-LC2), which does not contain a PDZ-binding domain, also interacts upstream of the –TTPV sequence of TARPs in trafficking of AMPARs in cerebellar neurons before they are anchored at the synapse (Ives et al. 2004). Furthermore, the C-terminus –TTPV sequence of TARPs has also been found to bind to the protein nPIST (neuronal PDZ domain protein interacting specifically with TC10), which is also a PDZ domain containing protein involved in synaptic clustering of AMPARs. nPIST is a Golgi-enriched protein implicated in trafficking of transmembrane proteins. Overexpression nPIST increased synaptic clustering of AMPAR (Cuadra et al. 2004). Colocalisation studies suggest that both TARPs and nPIST interact and localise in the brain, which may help AMPAR complexes leave the Golgi and be targeted to the synapse (Cuadra et al. 2004). AMPARs travelling from the ER to the plasma membrane need at least three basic steps: exit from the ER, delivery through Golgi aperture, and anchoring at the PSD site; all these mechanisms require either direct or indirect interaction with various proteins, which include chaperones, vesicular delivery stabilisers, and anchoring proteins. A recent review of AMPAR interacting proteins and their function in trafficking pathways include GRIP1, KIF1A, KIF5, Rab8, Liprin-α, NEEP21, GRASP-1, Rab4, and Stx13 (Anggono and Huganir 2012).
1.4.3.2 Isoforms-specific TARPs regulate subunit-specific AMPAR trafficking

Several research studies have clearly identified that TAPRs are an essential auxiliary protein for trafficking of AMPAR, but the underlying mechanism of TARP-mediated AMPAR surface expression is still not completely understood. The mechanism may differ depending on neuronal class, TARP isoforms or AMPAR subunits. The best example is TARP-γ5, which is a type II TARP, and does not appear to have any regulatory effect on the surface expression of AMPAR (Kato et al. 2007, Tomita et al. 2004). Instead, the main effect of TARP-γ5 is to regulate the function of AMPAR subunit combined with the long-isoform, which comprises predominantly Ca$^{2+}$ permeable AMPARs (Soto et al. 2009). TARP-γ5 also appears only to regulate AMPARs in Bergmann glia, which generally acts as a negative control (Soto et al. 2009). Isoform-specific TARPs have also been suggested to regulate the retention of AMPARs in the ER (Coleman et al. 2006). The trafficking of homomeric flop isoform of GluR4 was largely blocked at the ER exit, while the flip isoform of GluR4 were trafficked to the plasma membrane at 10 times higher levels than the flop isoform. The retention of the flop isoform of GluR4 in the ER escaped when cells co-expressed either stargazin/γ2 or the flip isoform. Furthermore, although TARP-γ2 has constantly been found to be involved in an early AMPAR biosynthetic pathway in cerebellar granule cells (Vandenberghe et al. 2005a), the null mutant TARP-γ2 largely stores immature AMPARs in the cerebellum of stargazer mice (Tomita et al. 2003). These studies suggested that isoform-specific TARPs mediate specific AMPAR subunit trafficking and localisation in specific neuronal regions or cells.
1.4.3.3 Synaptic stabilisation of AMPAR-TARP complexes

The mechanism of AMPAR-TARP complex trafficking from the ER to the synapse has not been precisely identified; however, the stabilisation of these complexes by membrane-associated guanylate kinases (MAGUKs) such as PSD-95 (postsynaptic density protein 95) is still unclear, but it may involve phosphorylation of the C-terminus of TARPs. The C-terminus sequence of TARPs varies in length depending on their isoform. These are between 100 and 200 amino acids in length and are enriched for serine and threonine residues. PSD-95 is comprised of a multiple protein interactive structure, which includes: three PDZ domains, an src homology 3 (SH3) domain and a guanylate kinase-like (GK) domain. The C-terminus of AMPAR subunits cannot interact with PSD-95; whereas TARPs contain a canonical PDZ-binding motif located at the C-terminus, which allows direct binding to the PDZ domains of PSD-95 (Figure 1-4) (Chen et al. 2000, Tomita et al. 2005b). These findings suggest that the PDZ domains of PSD-95 capture and stabilise AMPAR-TARP complexes at synapses via binding to the C-terminus of TARPs. Interestingly, overexpression of PSD-95 increased synaptic clustering of AMPAR at the postsynaptic density. However, overexpression of PSD-95 did not alter the clustering of NMDARs (El-Husseini et al. 2000, Stein et al. 2003). This result suggested that overexpression of PSD-95 may enhance the retention of synaptic AMPARs by opening new slots for AMPARs at synapses and the diffusion of surface AMPARs away from the synapse was also decreased (Stein et al. 2003). In order to confirm that the C-terminus of TARPs are essential binding regions for PDS-95 to stabilise AMPARs at the synapses, mutated TARPs lacking C-terminal tail were designed. Hippocampal pyramidal cells of mutant mice harbouring TARP-γ8 lacking the C-terminal PDZ binding motif (TTPV) reduced synaptic transmission and AMPARs at a level comparable to that of TARP γ8-deficient mice (Sumioka et al. 2011). This result indicates that the C-terminal tail –TTPV sequences of TARPs stabilises AMPAR expression and regulates the abundance of synaptic AMPAR in the brain. Overexpression of TARP-γ2 lacking C-terminal sequence
(-TTPV) in hippocampal neurons reduces synaptic AMPARs activity in a dominant negative manner (Bats et al. 2007, Chen et al. 2000, Schnell et al. 2002); however, the mobility of surface AMPARs was increased at synapses (Bats et al. 2007). PSD-95 and other MAGUKs directly interact with the C-terminal PDZ binding motifs in TARPs to stabilise AMPARs at synapses (Chen et al. 2000, Coombs and Cull-Candy 2009, Dakoji et al. 2003, Diaz 2010, Jackson and Nicoll 2011a, Opazo and Choquet 2011, Tomita 2010). This interaction can also be an important mechanism for synaptic AMPAR function by AMPAR-mediated excitatory post-synaptic currents (EPSCs). This theory is observed in the reduction of the AMPAR-mediated EPSCs from stargazer mice carrying a disrupted γ2/stargazin gene (Hashimoto et al. 1999), in γ2/γ3 double knockout mice in Golgi cells (Menuz et al. 2008), in γ8 knock-out mice in hippocampal pyramidal cells (Rouach et al. 2005), and in γ2/γ7 double knock-out mice in Purkinje cells (Yamazaki et al. 2010). Furthermore, no AMPAR activity and no miniature EPSCs (mEPSCs) were observed in primary cerebellar granule cell cultures from stargazin mice (Chen et al. 2000, Hashimoto et al. 1999); however, when type I TARPs are overexpressed mEPSCs are restored (Chen et al. 2000, Tomita et al. 2003). Conversely, the decay kinetics of mEPSCs in neurons expressing Type I TARPs are faster than those in neurons expressing type II TARPs (Cho et al. 2007, Milstein et al. 2007). These results indicate that isoforms-specific TARPs modulate AMPR kinetics in the brain (Cho et al. 2007, Fukaya et al. 2005, Milstein et al. 2007, Tomita et al. 2003). Overexpression of PSD-95 in neurons also induces synaptic AMPAR potentiation (El-Husseini et al. 2000, Stein et al. 2003), which suggests that PSD-95 plays a role in AMPAR stabilisation during plasticity. However, the presence of mutated TARP-γ8 lacking C-terminus (-TTVP) sequence revealed a normal LTP induction in the hippocampus, where the interaction between PDZ-binding motif and PSD-95 was absent (Sumioka et al. 2011). This suggests that AMPAR trafficking
during activity-dependent plasticity may require more complicated interaction than the association with PSD-95 and TARPs.

Calcium-/calmodulin-dependent protein kinase II (CaMKII) is a serine/threonine-specific protein kinase, which is involved in many signalling cascades and is thought to be associated with learning and memory. There is strong evidence that activated CaMKII indirectly induces the trafficking of AMPARs to the plasma membrane and stabilised the receptor complexes at the PSD of the dendrite. In their basal state AMPARs are highly mobile at the neuronal surface, rapidly switching between extrasynaptic and synaptic sites, while the concentration of CaMKII is low in the PSD. Upon activation of NMDARs there is a Ca\textsuperscript{2+} influx that activates CaMKII and induces its postsynaptic translocation. Once CaMKII arrives at the PSD, it phosphorylates the C-terminus of TARPs to facilitate its binding to PSD95 and stabilises AMPAR-TARP complexes at the synaptic sites (Figure 1-12) (Kristensen et al. 2011, Tomita et al. 2005b).
Figure 1-12 Model of AMPAR stabilisation via phosphorylation of TARPs by CaMKII. Left panel: At the basal state, NMDAR inactivated; Right panel: activated NMDAR triggers downstream cascades (Opazo et al. 2010)

Lipid bilayers have also been shown to indirectly modulate the synaptic trafficking of AMPAR-TARP complexes (Sumioka et al. 2010). The C-terminus of TARPs binds to lipid bilayers in an electrostatic manner, this interaction further inhibits the binding of TARPs to PSD-95 at the PSD (Sumioka et al. 2010) (Figure 1-13). The positively charged arginine residue in the C-terminus of TARPs interacts with negatively charged phosphates of lipids and lipid bilayers and subsequently inhibits the binding of TARPs to PSD-95 (Roberts et al. 2011, Sumioka et al. 2010). Conversely, the phosphorylation of TARPs inhibits the interaction with lipids, which is followed by the stabilisation of AMPAR-TARP complexes at the synapse (Inamura et al. 2006, Sumioka et al. 2010, Tomita et al. 2005b). These protein interactions suggests that phosphorylation of TARPs in activity-dependent plasticity plays an important role for AMPAR stabilisation and thus control synaptic function.
Figure 1-13 Lipid bilayers enhance translocation of AMPAR-TARP complexes to synapses (Sumioka et al. 2010).

In addition to modulation of AMPAR stabilisation and trafficking, TARPs also regulate the channel properties of AMPARs (Coombs and Cull-Candy 2009, Diaz 2010, Jackson and Nicoll 2011a, Opazo and Choquet 2011, Tomita 2010). The presence of TARPs slows the kinetics of AMPAR deactivation (channel closure upon glutamate removal) and desensitisation (channel closure upon glutamate binding) (Milstein et al. 2007). Moreover, TARPs modulate the ion permeability of AMPARs. The presence of TARPs reduces the reaction of AMPARs, rendering them more calcium permeable (Cho et al. 2007, Jackson and Nicoll 2011b, Soto et al. 2007, Soto et al. 2009).
1.4.4 Trafficking of AMPARs associated with neuro-disorders

The remarkable dynamics of synaptic neurotransmitter receptor localisation and its importance to mechanisms underlying brain functioning such as learning and memory, as well as to several disease states involving cognitive dysfunction has only recently been recognised. Research studies have provided evidence for AMPAR trafficking underlying behavioural learning responses and cognitive dysfunction associated with neuro-disorder conditions such as Alzheimer’s disease (AD). A reduction in GluR1, GluR2, and GluR2/3 AMPAR subunits levels has been demonstrated in several brain areas of AD patients, such as CA1, the subiculum, and entorhinal cortex (Ikonomovic et al. 1997, Ikonomovic et al. 1995).

There is convincing evidence for the requirement of TARPs to mediate AMPAR trafficking to synapses for a variety of brain functions. While there is strong evidence of a role for TARP subunits, the mechanism linking this to pathology still requires further study. Meanwhile, human genetic studies have recently suggested that TARPs may play a role in neuropsychiatric disorders.
2 Aims of the thesis

The main aim of the thesis was to identify genetic susceptibility variants associated with bipolar disorder and schizophrenia, and characterise the contribution of these genetic variants to the development of dysfunctional cell surface AMPA receptor expression. The specific hypothesis and aims of the thesis were:

1. **Hypothesis:** rare variants contribute to disorder susceptibility and that they may be found in the same genes that harbour common disease-associated single-nucleotide polymorphisms.

   **Aims:** To scan the CACNG5 gene for functional variants and to genotype these variants in order to test for association with BPD and/or SCZ.

2. **Hypothesis:** The disease-associated rare functional CACNG5 variants may have an effect on protein interaction.

   **Aims:** To characterize the possible contribution of these CACNG5 genetic variants on the development of dysfunctional cell surface AMPA receptor expression.
3 MATERIALS AND METHODS

3.1 General Materials and Methods

3.1.1 Research participants – UCL Research sample

The UCL research samples were recruited from London and South England and consist of 1099 volunteers with bipolar disorder, 618 volunteers with schizophrenia and 986 volunteers as normal comparison subjects. The control subjects were included 614 “supernormal” – subjects were screened for an absence of psychiatric disorder and 480 unscreened British normal control subjects provided by European Collection of Animal Cell Cultures (ECACC). For all groups participants were included only both parents were of English, Irish, Welsh or Scottish descent and if three out of four grandparents were of the same descent. Furthermore, the participants were excluded if one grandparent was of Jewish or non-Caucasian European ancestry based on the EU countries before the 2004 enlargement.

All participants required to obtain UK National Health Services (NHS) multicentre with local research ethics committee approvals and signed an approved consent form. All cases subjects with bipolar disorder or schizophrenia and supernormal control subjects were interviewed by a psychiatrist using the lifetime version of the Schizophrenia and Affective Disorders Schedule (SADS-L) (Spitzer R 1977). UCL schizophrenia subjects were diagnosed according to standardised criteria in the International Classification of Diseases 10th edition (ICD-10). All cases were selected on the basis of a primary clinical diagnosis of bipolar disorder or schizophrenia, and then they were formally diagnosed again according to Research Diagnostic Criteria (RDC). The “supernormal” control subjects were strictly selected on the basis of not having a family history of schizophrenia, alcoholism or bipolar disorder and for having no past or present personal history of any RDC-defined mental disorder.
Genomic DNA was extracted by lab colleagues from frozen blood samples using standard DNA extraction techniques. DNA samples were then quantified using Qubit2.0 (Section 3.1.2 DNA quantification).

### 3.1.2 DNA Quantification

DNA concentration was quantified using Qubit® DNA assay (Invitrogen, UK) with the Qubit® 2.0 Fluorometer. This is a quantitation system relying on dyes that only fluoresce when bound to specific molecules, such as double-stranded DNA (dsDNA), single-stranded DNA (ssDNA) or RNA. The calibration of instrument was using the Qubit® dsDNA high sensitivity (HS) Assay (declared assay range between 0.2 – 100 ng; sample starting concentration between 10 pg/µl – 100 ng/µl) (Figure 3-1 A) and the Qubit® dsDNA broad rang (BR) Assay (declared assay range between 2 – 1000 ng; sample starting concentration between 100 pg/µl and 1 µg/µl) (Figure 3-1 B), according to the manufacturer’s instructions.

![Figure 3-1 Examples of DNA concentration calibration from the Qubit® dsDNA assay.](image)

The points corresponding to its DNA concentration and its corresponding fluorescence curves and plotted. A: Qubit® dsDNA HS assay kit; B: Qubit® dsDNA BR assay kit (Invitrogen, UK).
The Qubit® 2.0 protocol for calibrating the fluorescent intensity was followed according to the manufacturer’s manual. Two DNA assay kits (BR and HS) assay kits were prepared from the calibration standard provided by the manufacturer and fluorescence signal were measured using Qubit® 2.0 Fluorometer. When the DNA concentrations were higher than the calibration curves, the DNA samples were further diluted in order to put them within the range of the calibration. The dilution factor is used by Qubit 2.0 algorithm to calculate the actual concentration. After the calibration routine, the device simply displays the concentration and information about the raw fluorescence intensity is omitted. For each 50 DNA samples, a new calibration was applied to avoid maintain calibration of measurements.

**Protocol:**

Both the Qubit® dsDNA HR and BR assay required two standards provide by the manufacturer. All standards and sample tubes were prepared in labelled thin-wall 500 µl microcentrifuge tubes. The Qubit® working solution was diluted by either the Qubit® dsDNA HS or BR reagent 1:200 in Qubit® dsDNA HS or BR buffer. Then, 190 µl of Qubit® working solution was loaded into each of the tubes used for standards, and 195 µl was loaded for samples. 10 µl of each Qubit® standard to the appropriate tube was added and 5 µl was added into the samples tubes to make a final volume of 200 µl in each tubes. All samples and standards tubes were mixed by vortexing for 3 seconds, and incubated at room temperature for 2 minutes. After incubation, the standards were ready to calibrate and the samples were ready to be quantified.

### 3.2 Variant detection and genotyping

For identification of CACNG5 variants, a rapid mutation scanning of the gene was performed using high resolution melting analysis (HRMA). The exons of CACNG5 gene plus their neighbouring flanking introns and promoter regions were amplified in UCL BPD, SCZ, and control samples. The presence of potentially aetiological variants detected in cases or control
samples were further confirmed using traditional polymerase chain reaction (PCR) and followed DNA Sanger sequencing. After confirmation of validity of variants in the CACNG5 gene, the frequency of these variants in the control and cases samples was investigated using the KBiosciences competitive allele specific PCR system (KASPar) (Kbiosciences, UK). The genotypic frequencies of the variants were then analysed for departure from Hardy-Weinberg equilibrium using chi-square tests. Allele counts were tested for single marker associations and further tested by burden analysis. The possible functional effects of variants were further predicted using online bioinformatics software.

3.2.1 High Resolution Melting Curve Genotyping Assay

HRMA is a post-PCR analysis method to identify variants in disease-related genes. The principle behind it is to detect changes in the dissociation behaviour of DNA under increasing temperature (Herrmann et al. 2006, Vossen et al. 2009, Wittwer 2009). This method involves PCR in the presence of saturating dsDNA-binding dyes such as LCGreen, EvaGreen, or SyBr Green, with increasing temperature to denature the double-stranded DNA into single-stranded DNA with loss of its fluorescence signal (Reed and Wittwer 2004). Then the profile gives a specific sequence-related pattern allowing identification of wild-type sequences and homozygote-heterozygote variants (Graham et al. 2005).

3.2.1.1 Primer design for HRMA

For CACNG5 mutational analysis, 7 pairs of primers (Table 3-1) with melting temperatures of around 60°C were designed using primer 3 software following the general guidelines for primer design as described below (Method section 3.2.1.1.1). The primers were purified by standard reverse-phase cartridge purification (Sigma, UK). The primers amplified around 100 -350 basepair (bp) of genomic DNA. Using this approach the entire coding sequences plus their flanking introns and promoter regions of CACNG5 (Genebank: 27091) were amplified following the standard protocol.
### Table 3-1 Primer sets used for HRMA

<table>
<thead>
<tr>
<th>Regions</th>
<th>Primers Sequence (5’ to 3’)</th>
<th>Amplicons size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Promoter Region</td>
<td>Forward-CCCTTCATGGCAAGGATAG Reverse-TAGGGAGAGTGGGTGAGCAG</td>
<td>354</td>
</tr>
<tr>
<td>EXON 1</td>
<td>Forward-TGGCTCTCTCATCTCTTTGCTCTTT Reverse-CCTTACCTGCAAGGAGCAG</td>
<td>330</td>
</tr>
<tr>
<td>EXON 2</td>
<td>Forward-TGGGATTTCTTGCTCCTCAGAAGG</td>
<td>128</td>
</tr>
<tr>
<td>EXON 3</td>
<td>Forward-ATTCCCCTCTTGTAACGCTCT Reverse-CAGGGTGCAAGAAAAGAAA</td>
<td>153</td>
</tr>
<tr>
<td>EXON 4</td>
<td>Forward-GGCACTTTCTGAGTCATC Reverse-AGGGAGTAGGATGGTGTT</td>
<td>320</td>
</tr>
<tr>
<td>Intron 4 - 5</td>
<td>Forward-ACGTAGCCGCCGGAAGTGC Reverse-GCGGTCGCTCTTCATCAA</td>
<td>318</td>
</tr>
<tr>
<td>EXON 5</td>
<td>Forward-CGGGGTGATGTCTGTAGACGC</td>
<td>254</td>
</tr>
</tbody>
</table>

#### 3.2.1.1 Primer design guideline

The most important PCR parameter is the design of the primers. The primers must be designed to define an appropriate target with a goal to amplify a single DNA sequence. For instance, it is important to avoid repetitive DNA sequences and the following factors are needed to be considered as well:

1. Optimally about 18 – 22 nucleotides long, which are specific for sequences flanking the target sequence;
2. The forward primer sequence is complimentary to the reverse strand of DNA and the reverse primer to the leading strand.
3. The GC content should be between 40 – 60% with an even distribution of all four nucleotides;
4. No stop codons (either TTA at the 3’ end or TAG/TAA/TGA at the 5’ end);
5. The primer end with a T could possibly prone to mis-priming due to low efficiency of the *Taq* DNA polymerase

---

112
6. The melting temperature ($T_m$) of the primer is the temperature corresponding to the mid-point in the observed transition from double-stranded form to single-stranded form. The optimal $T_m$ is between 55 - 65°C. Ideally, the calculated $T_m$ values for two primers should be within 5°C of each other. The $T_m$ calculation formula as shown below:

$$T_m = [(2°C \times AT) + (4°C \times CG)] - 2$$

7. No more than 3 G’s or C’s in the last 5 bases at the 3’ terminal sequence of the primer, this increases the strength of bonding at the 3’ terminal sequence, where helps promote specific allele can be amplified;

8. No hairpins or dimerisation in primers is also essential;

9. Avoid more than 3 bases in length of inverted repeats or any self-complementary sequence.

Primer 3 software was selected to design primers (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). NetPrimer (Premier Biosoft) (http://www.premierbiosoft.com/servlet/com.pbi.crm.clientside.FreeToolLoginServlet#) is primer analysis software that to analyses the properties of primers such as $T_m$, GC%, secondary structures, 3’ terminal sequence stability. This information was used to determine the quality of the primers.

### 3.2.1.2 Optimisation of HRM assay

The accuracy of HRM scanning depends on high quality PCR, thus optimisation is critical. Amplification of $CACNG5$ primers were optimised using several HRM assays which are available from various manufacturers such as SensiMix (BIOLINE), Roche HRMA master mix, AccuMelt HRM supermix (Quanta BIOSCIENCES), and LightScanner (BioFire) HRM assays.
PCR amplifications using SensiMix (BIOLINE) were optimised with adjusted MgCl$_2$ concentrations in a final volume of 5 µl containing 10 ng dried DNA and 200 nM of each primer (Table 3-2).

**Table 3-2 HRMA assay conditions for primer optimisation**

<table>
<thead>
<tr>
<th>MgCl$_2$ Concentration conditions</th>
<th>3.5 mM (µl)</th>
<th>4.0 mM (µl)</th>
<th>4.5 mM (µl)</th>
<th>5.0 mM (µl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SensiMix HRM</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
</tr>
<tr>
<td>50 mM MgCl$_2$</td>
<td>0.05</td>
<td>0.1</td>
<td>0.15</td>
<td>0.2</td>
</tr>
<tr>
<td>EvaGreen Dye</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>Forward Primer (10 pmol/µl)</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Reverse Primer (10 pmol/µl)</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>DNA (10 ng/µl)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PCR water</td>
<td>1.05</td>
<td>1</td>
<td>0.95</td>
<td>0.9</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>5</strong></td>
<td><strong>5</strong></td>
<td><strong>5</strong></td>
<td><strong>5</strong></td>
</tr>
</tbody>
</table>

Roche HRM Mix assays were also optimised with adjusted MgCl$_2$ concentrations in a final volume of 5 µl per reaction containing 10 ng dried DNA and 200 nM of each primer (Table 3-3).

**Table 3-3 Roche HRM Mix assay conditions for primer optimisation (4 x 5 µl reactions were prepared)**

<table>
<thead>
<tr>
<th>MgCl$_2$ Concentration conditions</th>
<th>2.0 mM (µl)</th>
<th>2.5 mM (µl)</th>
<th>3 mM (µl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MasterMix</td>
<td>11.25</td>
<td>11.25</td>
<td>11.25</td>
</tr>
<tr>
<td>25 mM MgCl$_2$</td>
<td>1.8</td>
<td>2.25</td>
<td>2.7</td>
</tr>
<tr>
<td>Forward Primer (10 pmol/µl)</td>
<td>0.45</td>
<td>0.45</td>
<td>0.45</td>
</tr>
<tr>
<td>Reverse Primer (10 pmol/µl)</td>
<td>0.45</td>
<td>0.45</td>
<td>0.45</td>
</tr>
<tr>
<td>DNA (10 ng/µl)</td>
<td>4.5</td>
<td>4.5</td>
<td>4.5</td>
</tr>
<tr>
<td>PCR water</td>
<td>4.05</td>
<td>3.6</td>
<td>3.15</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>22.5</strong></td>
<td><strong>22.5</strong></td>
<td><strong>22.5</strong></td>
</tr>
</tbody>
</table>

PCR amplification using Quanta Supermix HRMA assays were optimised with adjusted primer concentrations in a final volume of 5 µl per reaction containing 10 ng dried DNA (Table 3-4).
Table 3-4 Quanta Supermix HRMA assay conditions for primer optimisation

<table>
<thead>
<tr>
<th>Primer Concentration conditions</th>
<th>200 nM (μl)</th>
<th>500 nM (μl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AccuMelt HRM SuperMix</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Forward Primer (10 pmol/μl)</td>
<td>0.1</td>
<td>0.25</td>
</tr>
<tr>
<td>Reverse Primer (10 pmol/μl)</td>
<td>0.1</td>
<td>0.25</td>
</tr>
<tr>
<td>DNA (10 ng/μl)</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PCR water</td>
<td>1.8</td>
<td>1.5</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>5</strong></td>
<td><strong>5</strong></td>
</tr>
</tbody>
</table>

LightScanner HRMA assay was the only one that do not required any adjustment of solution as shown in Table 3-5.

Table 3-5 LightScanner HRM Mix assay conditions for primer optimisation

<table>
<thead>
<tr>
<th>Primer Concentration conditions</th>
<th>200 nM (μl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MasterMix</td>
<td>2</td>
</tr>
<tr>
<td>Forward Primer (10 pmol/μl)</td>
<td>0.5</td>
</tr>
<tr>
<td>Reverse Primer (10 pmol/μl)</td>
<td>0.5</td>
</tr>
<tr>
<td>DNA (10 ng/μl)</td>
<td>1</td>
</tr>
<tr>
<td>PCR water</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>5</strong></td>
</tr>
</tbody>
</table>

Reactions using the optimum assay conditions were performed in a 384-well plate prepared using an epMotion 5075 robot (Eppendorf). Reactions were mixed using a MixMate (Eppendorf), and briefly spun in a centrifuge for a minute. The plate was loaded into the LightCycler® 480 (Roche). The amplification were run according to the optimised conditions and the cycle was applied as shown in Table 3-6. Analysis of HRM curves was carried using the LightCycler® 480 Gene Scanning software v.1.5.0 (Roche) to visualise fluorescence data using normalisation, temperature-shifting, and difference plotting, and then analysed using the automated grouping functionalities. During the high-resolution melting step that followed amplification, the software records a large number of fluorescence data points per for each change in temperature. This altered fluorescence during an experiment can be used to measure temperature- induced DNA dissociation during High Resolution Melting Analysis.
Melting. The melting profile of the amplicon depends on the GC content, sequence, length and heterozygosity, therefore, any sequence variations will result in heteroduplexes with a different melt curve shape compared to the wild-type sequence.
### Table 3-6 HRM cycling conditions

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

The LightCycler® 480 Gene Scanning software (Roche) measured the decrease in fluorescence signal during the denaturation of double-stranded amplicons with temperature increased. Then the software generated plots of fluorescence signal over temperature, where the melting curves from the variant DNA can be distinguished from the wild type samples as shown in Figure 3-2, and more apparent diagram is shown in a difference plot.

![Normalized and Temp-Shifted Melting Curves](image)

![Normalized and Temp-Shifted Difference Plot](image)

Figure 3-2 An example of HRMA reveals differences between wild-type (homozygous) and variants (heterozygous). Difference plot analysis allowed to differentiate between homozygous (blue) and heterozygous (red) samples.
3.2.2  Polymerase Chain Reaction (PCR)

PCR is a basic molecular genetics technique, where applying the ability of DNA polymerase to synthesis new strand of DNA complementary to the offered template strand. DNA polymerase synthesis of a strand complementary requires two specific designed oligonucleotide primers.

3.2.2.1  General PCR principle

The principle of PCR requires three major steps: DNA melting, primer annealing and DNA polymerase elongation. During DNA melting step, the double strands of DNA are physically separated by disrupting the hydrogen bonds between complementary bases at a high temperature (usually performed at 95ºC) for about 30 seconds to 5 minutes, yielding single-stranded DNA template. In primer annealing, temperatures are usually between 50 - 70ºC depending on the primer-template T_m. The optimal re-annealing temperature is about 5ºC lower than the primer-template allowing annealing of the primers to the single-stranded DNA template. Generally, the heat-stable DNA polymerase requires an efficient elongation temperature at about 70 - 75ºC. In the presence of heat-stable DNA polymerase and DNA precursors – the four deoxynucleoside triphosphates, dATP, dCTP, dGTP, and dTTP, the primers initiate the synthesis of new DNA strands which are complementary to the individual DNA strands of the target DNA segments. The primer concentration is another essential factor, generally the optimal concentration is between 0.1 and 0.5 µM for most PCR applications. Sufficient optimal primer concentrations increase possibility of mismatches and yielding non-specific PCR.

The electrorepulsive forces between the DNA strands are weaken by the present of cations in PCR buffers, such as K⁺ (potassium) and NH₄⁺ (ammonium) that leading to a neutralisation on the negatively charged phosphate groups of the DNA backbone. Hence, the annealing process between the primer and the templates occurs, when the repulsive forces are reduced.
And it also maintains the high ratio of specific to nonspecific primer-template binding over a wide temperature range.

Furthermore, the additional magnesium ions (Mg$^{2+}$) to the reaction mix stabilizes primer annealing, influences enzyme activity and increases the T$_m$ of dsDNA. Mg$^{2+}$ also forms soluble complexes with dNTPs in the reaction mix to produce the actual substrate that the polymerase recognizes. The optimal concentration for each deoxynucleoside triphosphate (dNTP) to minimize polymerase error is 200 µM. Other solutions such as dimethyl sulfoxide (DMSO), betaine, and glycerol are also used to improve the amplification efficiency by melting the secondary structures and decreasing non-specific products.

### 3.2.2.2 General optimisation of the PCR reaction

The primers were optimised using four common master mix conditions with altered MgCl$_2$ concentration and the inclusive or exclusive of betaine as shown in Table 3-7 on MWG-HT Primus 96 thermocycler.
Table 3-7 PCR primer optimisation layout

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume (µl)</th>
<th>2 mM MgCl₂ + Betaine</th>
<th>2 mM MgCl₂ + No Betaine</th>
<th>2.5 mM MgCl₂ + Betaine</th>
<th>2.5 mM MgCl₂ + No Betaine</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 X Buffer</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
</tr>
<tr>
<td>Betaine 5M</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>50 mM MgCl₂</td>
<td>1</td>
<td>1</td>
<td>1.25</td>
<td>1.25</td>
<td></td>
</tr>
<tr>
<td>25 mM dNTP</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td>F primer (10 pmol/µl)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>R primer (10 pmol/µl)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Taq polymerase (1 U/µl)</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>DNA 25 ng/µl</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>11.8</td>
<td>16.8</td>
<td>11.55</td>
<td>16.55</td>
<td></td>
</tr>
<tr>
<td>Total volume</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>25</td>
</tr>
</tbody>
</table>
For amplifying the PCR products, three standard PCR programs were applied such as Standard 55°C, Standard 60°C and Touchdown as shown below. In order to achieve the most efficient PCR amplification, it was essential to adjust the annealing temperature and the number of cycles in these experimental protocols.
The three standard cycling conditions were:

<table>
<thead>
<tr>
<th>Standard 55°C (STD55)</th>
<th>Standard 60°C (STD60)</th>
<th>Touch Down (MHTD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lid heated to 105 °C</td>
<td>Lid heated to 105 °C</td>
<td>Lid heated to 105 °C</td>
</tr>
<tr>
<td>Products denatured at 94 °C for 5 minutes</td>
<td>Products denatured at 94 °C for 5 minutes</td>
<td>Products denatured at 94 °C for 5 minutes</td>
</tr>
<tr>
<td>35 cycles of 94°C - 30 seconds</td>
<td>35 cycles of 94°C - 30 seconds</td>
<td>3 cycles of 94°C - 30 seconds</td>
</tr>
<tr>
<td>55°C - 30 seconds</td>
<td>60°C - 30 seconds</td>
<td>63°C - 30 seconds</td>
</tr>
<tr>
<td>72°C - 30 seconds</td>
<td>72°C - 30 seconds</td>
<td>72°C - 30 seconds</td>
</tr>
<tr>
<td>Hold at 72°C - 10 minutes</td>
<td>Hold at 72°C - 10 minutes</td>
<td>94°C - 30 seconds</td>
</tr>
<tr>
<td>Store at 4°C</td>
<td>Store at 4°C</td>
<td>60°C - 30 seconds</td>
</tr>
<tr>
<td></td>
<td></td>
<td>72°C - 30 seconds</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 cycles of</td>
</tr>
<tr>
<td></td>
<td></td>
<td>94°C - 30 seconds</td>
</tr>
<tr>
<td></td>
<td></td>
<td>57°C - 30 seconds</td>
</tr>
<tr>
<td></td>
<td></td>
<td>72°C - 30 seconds</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 cycles of</td>
</tr>
<tr>
<td></td>
<td></td>
<td>94°C - 30 seconds</td>
</tr>
<tr>
<td></td>
<td></td>
<td>54°C - 30 seconds</td>
</tr>
<tr>
<td></td>
<td></td>
<td>72°C - 30 seconds</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 cycles of</td>
</tr>
<tr>
<td></td>
<td></td>
<td>94°C - 30 seconds</td>
</tr>
<tr>
<td></td>
<td></td>
<td>51°C - 30 seconds</td>
</tr>
<tr>
<td></td>
<td></td>
<td>72°C - 30 seconds</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 cycles of</td>
</tr>
<tr>
<td></td>
<td></td>
<td>94°C - 30 seconds</td>
</tr>
<tr>
<td></td>
<td></td>
<td>48°C - 30 seconds</td>
</tr>
<tr>
<td></td>
<td></td>
<td>72°C - 30 seconds</td>
</tr>
<tr>
<td>Hold at 72°C - 10 minutes</td>
<td>Hold at 72°C - 10 minutes</td>
<td>94°C - 30 seconds</td>
</tr>
<tr>
<td>Store at 4°C</td>
<td>Store at 4°C</td>
<td>60°C - 30 seconds</td>
</tr>
</tbody>
</table>
3.2.3 Agarose gel electrophoresis

Agarose gel electrophoresis is the most effective technique to analyse the composition and quality of a nucleic acid sample. It is also used to determine the size of DNA fragments such as PCR products from 100 bases pairs (bp) to 25 kilo-bases pairs (kb) (Sambrook 2001). For this purpose, it is essential to run a standard marker containing fragments of known sizes, such as Hyperladder markers (Bioline). The principle of agarose gel electrophoresis is where the DNA is loaded into pre-cast wells in the gel and a current applied. The phosphate backbone of the DNA is negatively charged, therefore when placed in an electric field, DNA fragments will migrate to the positively charged anode. Because DNA has a uniform mass/charge ratio, DNA molecules are migrated by size within an agarose gel in a pattern such that the moving distance is inversely proportional to the logarithm of its molecular weight (Helling et al. 1974). The rate of DNA migration through a gel is determined by the following: 1) size of DNA; 2) concentration of agarose; 3) DNA conformation (Aaij and Borst 1972); 4) voltage applied; 5) presence of dye; 6) type of agarose and 7) electrophoresis buffer. Dyes such as ethidium bromide are commonly used for both detecting and quantitating DNA, where its flat ring structure has ability to stack in between the bases in nucleic acids. After migration, the DNA stained with ethidium bromide can then be detected by its fluorescence when exposed to UV (ultraviolet) light.

3.2.3.1 The principle protocol of agarose gel electrophoresis

Buffer and solutions required:

- 1X TBE running buffer: 45 mM Tris-borate, 1 mM EDTA.
- Gel staining dye: Ethidium bromide
- Loading dye: Biotaq red DNA polymerase
- Molecular size marker: Hyperladder IV or Hyperladder I (Bioline, UK)
First, an appropriate mass of agarose was weighted into a conical flask in order to prepare a w/v (weight/volume) percentage (%) solution. The concentration of agarose in a gel depends on the sizes of the DNA fragments need to be separated. Mostly 1% w/v of agarose was prepared with 1X TBE buffer in a conical flask. The mixture was then heated in a microwave oven for one or one and half minutes to allow the agarose to completely dissolve, and then cooled down to around 50 - 60ºC. For gel staining, 10 mg/ml stock solution of ethidium bromide was added to make a final concentration of 0.5 μg/ml in the gel. After that, the gel was poured into a casting tray with appropriate combs, and the agarose was allowed to set for about 20 – 30 minutes at room temperature. Once set the combs were removed and the gel was placed into the electrophoresis gel tank with enough 1X TBE in the buffer tanks to cover the gel. The DNA samples were then added with a loading dye, Biotaq red DNA polymerase, it helps to track how far the DNA sample has migrated, and also allows the sample to sink into wells of the gel. 5 µl of the DNA samples and molecular size marker were loaded into separate gel wells. Then, the electrophoresis apparatus was connected to the power supply and run at 120 V (volt) for 30 minutes. When the electrophoresis was completed, the gel was removed from the gel tank and visualised under a UV transilluminator (UVP Gel-doc-it Imaging systems, UK), and then an image was captured using the camera attached to the gel doc system.

3.2.4 PCR purification

The quickest, cheapest and most flexible PCR purification is microCLEAN, which is a DNA clean-up reagent. These reagents purify dsDNA from reaction buffers, enzymes, primers, primer dimers or dNTPs that can interfere during sequencing reactions.
3.2.4.1 The protocol of microCLEAN

Buffers and solutions required:

- MicroCLEAN (Microzone Limited, UK): 0.5 M NaCl, 1 mM Tris HCl pH 8.0, 0.1 mM EDTA, 20 % w/v PEG8000, 1.75 mM MgCl₂; All the above contents were mixed with 50 ml of PCR water (Sigma, UK) and the solution was heated gently so that it completely dissolved. The solution was then filter sterilised using a 0.45 µM filter.

- An equal volume of MicroCLEAN reagent was added to the DNA sample. The solution was mixed using a vortex for tubes or a Mixmate (Eppendorf, Germany) for plates and incubated at room temperature for 5 minutes.

MicroCLEAN protocol for tubes:

The mixture of DNA and microCLEAN solution in tubes was spun in a microcentrifuge at 13,000 rpm for 7 minutes. The supernatant was removed and the tubes were briefly spun again to remove all the liquid.

MicroCLEAN protocol for 96 well plates:

The mixture of DNA and microCLEAN solution in plates was spun in a centrifuge at 4,000 rpm for 40 minutes. The plate was placed upside down on a tissue in the centrifuge holder to remove the supernatant and spun at 1,000 rpm for 30 seconds.

The both pellet from tubes or plates was resuspended in 5 µl PCR water and then rehydrated at room temperature for 5 minutes. The purified DNA was ready for further processing.
3.2.5 DNA sequencing

DNA cycle sequencing is a fluorescence-based cycle sequencing, which requires a DNA template, a sequencing primer, a thermal stable DNA polymerase, nucleotides deoxynucleotides (dNTPs) and dideoxynucleotides (ddNTPs), and buffer. The principle of this method is that fluorescent dyes are attached to the extension product of reactions that are involve successive rounds of DNA denaturation, primer annealing, and elongation in a thermal cycler to amplify linear extension products. These amplified linear extension products are terminated by one of the four fluorescent ddNTPs as shown in Figure 3-3. The optimised ratio of deoxynucleotides to dideoxynucleotides amplifies a balanced population of long and short extension products.

Figure 3-3 DNA sequencing cycle (Taken from Big Dye Sequencing kit 3.1 manual, Applied Biosystems, UK)
3.2.5.1 BigDye terminator sequencing reaction

BigDye terminator sequencing is a reaction using fluorescent fragments that generated by dye-labelled ddNTPs (dideoxynucleotide) complexes such as terminator of extension product. Because each four ddNTP - ddATP, ddCTP, ddGTP, or ddTTP terminators carries a unique fluorescent dye. The 3’ terminal dideoxynucleotide (A, C, G or T) of the extension product is identified by the fluorescent dye that it carries (Figure 3-4).

![Figure 3-4 Schematic showed a cycle procedure of BigDye terminator cycle sequencing](Taken from Big Dye Sequencing kit 3.1 manual, Applied Biosystems, UK)

PCR products were sequenced by BigDye Terminator v3.1 sequencing kit (Applied Biosystems, Warrington, UK) according to their manufacturer's instructions. Each sample was performed with two reactions, one using designed forward primer and the other one with the reverse primer, using the following reaction mixture and the cycling conditions as shown below (Table 3-8).

**Table 3-8 Reaction mix used for Big Dye sequencing reaction**

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Final concentration</th>
<th>Volume (µl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Terminator ready reaction mix</td>
<td>1X</td>
<td>1</td>
</tr>
<tr>
<td>BigDye Sequencing buffer 5X</td>
<td>1X</td>
<td>1.5</td>
</tr>
<tr>
<td>10 µM M13 primer F or R</td>
<td>0.32 pmol</td>
<td>0.32</td>
</tr>
<tr>
<td>Template (0.5 - 1 kbp PCR product)</td>
<td>(5 - 20 ng)</td>
<td>1</td>
</tr>
<tr>
<td>Water</td>
<td>As required</td>
<td>6.18</td>
</tr>
<tr>
<td>Total volume</td>
<td>10 µl</td>
<td>10</td>
</tr>
</tbody>
</table>
Cycling conditions:

- Heat the lid to 105°C
- Initial denaturation:
  - 94°C for 5 minutes
- 25 Cycles of:
  - 96°C - 10 seconds
  - 50°C - 5 seconds
  - 60°C - 4 minutes
- Store the reactions at 4°C

3.2.5.2 DNA sequencing purification

DNA sequencing purification is an important step to completely remove the excess dye terminators and salts from previous BigDye terminator sequencing reaction before electrophoresis analysis for generating clean sequencing data. Ethanol/EDTA precipitation method was applied for DNA sequencing purification in 96-well plates. For a 10 µl precipitation reaction in 96-well plates, 2.5µl of 125 mM EDTA and 30 µl of absolute ethanol was added directly to each sample. The reaction plate was mixed using a MixMate for 15 seconds, and then left at room temperature for at least 10 minutes to precipitate the extension products. The reaction plates were then spun in a centrifuge at 4,000 rpm for 60 minutes. In order to discarded supernatant, the reaction plate was inverted onto a paper towel and spun at 1,000 rpm for a minute. Afterwards, 30 µl of 70% ethanol was added to the reaction plates and mixed with a MixMate for 15 second. The plate was then centrifuged at 4,000 rpm for 10 minutes. The supernatant was discarded again as described above by inverting the reaction plates and brief centrifugation. The reaction plate was left at room temperature for 15 minutes to evaporate any residual ethanol. The dried pellets were resuspended with 15 µl of Hi-Di™ formamide (Applied Biosystems, UK) per well, a DNA denaturing reagent and sent off for sequencing at the Centre of Comparative Genomics,
Department of Biology, UCL and GOSH, York Way, London. The samples were then run on the 3730xl DNA Analyzer (Applied Biosystems, UK) (Figure 3-5).

![Diagram of the Applied Biosystems 3730xl DNA Analyzer run cycle steps](image)

**Figure 3-5** Diagram of the Applied Biosystems 3730xl DNA Analyzer run cycle steps (Manual guide of ABI Sequencing)

### 3.2.5.3 DNA sequencing data analysis

Finch TV software was used to visualise of DNA sequence traces, the display an entire trance, linked to BLAST searching of the databases, and also has the ability to reverse complement sequences and traces. For analysis of DNA sequencing data, the Staden Package was performed with an assembly program consists of Gap4 and Pregap4. Gap4 performs assembly, contig joining, and assembly checking, repeat searching, experiment suggestion, read pair analysis and contig editing. Pregap4 provides a graphical user interface to set up the processing required to prepare trace data for assembly or analysis.
3.2.6 KASPar Genotyping Assay

The KBiosciences Competitive Allele-Specific PCR SNP genotyping system (KASPar) is the simplest and most cost effective and flexible way to determine SNP genotypes. This system relies on the discrimination power of a novel form of competitive allele specific PCR to determine the alleles at a specific locus within genomic DNA based on fluorescent resonance energy transfer (FRET). The system is comprised of two components:

1. The KASP assay mix contains three non-labelled oligonucleotides (Figure 3-6): two allele-specific forward primers (one for each allele of the SNP) and one common reverse primer. The allele-specific forward primers are labelled with a unique tail sequence at the 5' end that corresponds with a universal FRET cassette; one labelled with FAM™ dye and the other with CAL™ Fluor Orange 560, both with quenchers bound at the 3' ends.

2. The KASP Master Mix contains Taq polymerase enzyme and the universal FRET cassettes passive reference dye 5-carboxy-X-rhodamine, succinimidyl ester (ROX), and MgCl₂ in an optimised buffer solution. The composition of reaction mix is described as below
### Table 3-9 Reaction mix for KASPar assay

<table>
<thead>
<tr>
<th>Component</th>
<th>Concentration in Assay Mix (µM)</th>
<th>Volume in Assay Mix (µL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allele specific primer F1 (100 µM)</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Allele specific primer F2 (100 µM)</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Common reverse primer R (100 µM)</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Water</td>
<td></td>
<td>46</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>100</td>
</tr>
</tbody>
</table>

### Principle:

During thermal cycling of PCR, the allele-specific forward primer binds to the template and elongates, thus attaching the tail 5’ sequence to the newly synthesised strand. The complement of the allele-specific tail 5’ sequence with fluor-labelled is then generated during subsequent rounds of PCR, enabling the FRET cassette to bind to the DNA. The FRET cassette is no longer quenched and emits fluorescence. Bi-allelic discrimination is achieved through the competitive binding of the two allele-specific forward primers. For instance, the genotype at a given homozygous SNP only one of the two possible fluorescent signals generates, whereas the heterozygous genotyping SNP generates a mixed fluorescent signal (Figure 3-6).
Figure 3-6 Diagrammatic representation of KASPar reaction process (Taken from Genotyping chemistry – KASP, KBioscience LGC)
3.2.6.1 Primer design for KASPar assay

The primers for SNPs were designed using the Primer Picker software provided by KBiosciences. Approximately 25 bases on each side of SNP sequence was cut and pasted into the sequence input window and analysed using the provided software. The results consisted of two forward primers for the alternative alleles of the SNP and two reverse primers. These two reverse primers were optimised to find the best reverse primer for the reaction. A total of 16 sets of KASPar primers were designed and followed the standard protocol for analysis (Table 3-10).
<table>
<thead>
<tr>
<th>Name</th>
<th>Amino acid change</th>
<th>Primer sets sequence (5' to 3')</th>
</tr>
</thead>
</table>
| rs3760263     | N/A              | AS1-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG  
CP1-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG  
CP2-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG |
| rs181400884   | N/A              | AS1-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG  
CP1-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG  
CP2-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG |
| rs75486725    | N/A              | AS1-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG  
CP1-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG  
CP2-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG |
| rs146874664   | Cys19Cys         | AS1-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG  
CP1-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG  
CP2-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG |
| rs71379998    | Ala27Ala         | AS1-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG  
CP1-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG  
CP2-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG |
| rs11652480    | Leu34Leu         | AS1-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG  
CP1-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG  
CP2-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG |
| rs146875098   | Arg69Trp         | AS1-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG  
CP1-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG  
CP2-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG |
| rs146875105   | Arg71His         | AS1-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG  
CP1-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG  
CP2-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG |
| rs142916987   | Arg127Gln        | AS1-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG  
CP1-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG  
CP2-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG |
| rs149159754   | Thr128Met        | AS1-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG  
CP1-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG  
CP2-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG |
| rs144880644   | Val146Met        | AS1-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG  
CP1-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG  
CP2-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG |
| rs144880674   | Ile156Phe        | AS1-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG  
CP1-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG  
CP2-GAAGGTGACCAAGTTCATGCTCCCAGAGAGGACAAGGGG |

Table 3-10 Primer sets used for Kaspar genotyping.
<table>
<thead>
<tr>
<th>Name</th>
<th>Amino acid change</th>
<th>Primer sets sequence (5' to 3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>17_64880699</td>
<td>Thr164Leu</td>
<td>AS1-GAAGGTGACCAAGTTCATGCTGAAGTACGTCTCTGCAATCCTTGG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AS2-GAAGGTACGAGTCAAGGATTGAGATCCTGACATCCTTGA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CP1-GCTCTAACATCTCCAAGTCAACGTGATGCTCA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CP2-CCAGATCAGAAGTGAAGTGGCTCA</td>
</tr>
<tr>
<td>rs187075595</td>
<td>Ala182Ser</td>
<td>AS1-GAAGGTGACCAAGTTCATGCTGAAGTACGTCTCTGCAATCCTTGG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AS2-GAAGGTACGAGTCAAGGATTGAGATCCTGACATCCTTGA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CP1-CACTACAAAGTATGGGTGGTC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CP2-GAAGGTACGAGTCAAGGATTGAGATCCTGACATCCTTGA</td>
</tr>
<tr>
<td>rs41280112</td>
<td>His233Tyr</td>
<td>AS1-GAAGGTGACCAAGTTCATGCTGAAGTACGTCTCTGCAATCCTTGG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AS2-GAAGGTACGAGTCAAGGATTGAGATCCTGACATCCTTGA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CP1-CAGGATCAGAAGTGAAGTGGCTCA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CP2-TCCGATTACTCAAGCCAGTTCTCA</td>
</tr>
</tbody>
</table>

### 3.2.6.2 Optimisation of KASPar genotyping assay

The KASPar genotyping reactions of each primer sets were optimised as described in Table 3-11. The optimum conditions then carried out the PCR reaction for cases and controls on 384-well plates containing dried 10 ng of DNA. For each genotyping SNP, the master mix assay was prepared in bulk and dispensed in the 384-well plates using Eppendorf eoMotion 5075. The reactions were mixed using a Mixmate and spun in a centrifuge briefly before loading into an LC480 (Roche). The PCR cycling conditions for KASPar genotyping were used as showed in Table 3-12.

136
Table 3-11 Kaspar Optimisation layout

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume (µl)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.8 mM MgCl₂</td>
</tr>
<tr>
<td>DNA (10 ng/µl)</td>
<td>1</td>
</tr>
<tr>
<td>KASP Reaction mix (2x)</td>
<td>2</td>
</tr>
<tr>
<td>Assay mix</td>
<td>0.055</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>0</td>
</tr>
<tr>
<td>H₂O</td>
<td>0.945</td>
</tr>
<tr>
<td>DMSO</td>
<td>0</td>
</tr>
<tr>
<td>Total reaction volume</td>
<td>4</td>
</tr>
</tbody>
</table>
Table 3-12 Kaspar cycling conditions: Dual colour hydrolysis probe/UPL probe program on LC480

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

The KASP assays were run on a Light Cycler LC480 (Roche) and analysed using the Endpoint genotyping LC480 software (Figure 3-7).

Figure 3-7 An example of the endpoint analysis of rs41280112 performed after KASPPar PCR reaction. Genotyping results are clustered into four different colour dots: blue and green triangles indicate homozygous individuals for the SNP (AA/aa), red triangles indicate heterozygous (Aa) individuals, grey circles indicate negative samples.

3.2.6.4 Analysis of replication data

The diseases association of CACNG5 were further refined with additional data from the European ancestry subjects in the 1,000 genome project, UK10K and Swedish whole exome sequencing data.
3.2.6.5 1000 Genomes Projects

The 1,000 Genomes Projects (http://www.1000genomes.org) was the first sequencing project involving a large number of people to provide a comprehensive resource on human genetic variation. The subjects were recruited worldwide, where included 286 participants with East Asian ancestry, 246 participants with African ancestry, 379 participants with European ancestry, and 181 Americas ancestry. Those subjects have no associated medical or phenotype data; hence the genetic variation data was used as control samples to refine the gene association study in this project with SCZ or BPD. Data for SNPs in the region surrounding and including the gene was extracted from the European ancestry samples in 1,000 genomes database located at FTP directory /vol1/ftp/release/20110521 at ftp.1000genomes.ebi.ac.uk using a windows version of tabix. Tabix indexes a TAB-delimited genome position file in.tab.bgz from a position sorted and compressed input data file. Tabix was run with the following commands that were modified according to the gene:

```
tabix -h
ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/release/20110521/ALL.chr.phase1_release_v3.2010
1123.snps_indels_svs.genotypes.vcf.gz chr:position from-to file name.tabix
```

The extracted data was then opened in Microsoft Excel and all the letters, symbols and numbers were removed by command replacing. The genotypes were then re-coded with alleles of the SNP. The data was transposed to create a ped and a map file, and those files were merged with the SNP data from all other resources for the gene.
3.2.6.6 UK10K exome sequencing data

UK10K ([http://www.uk10k.org/](http://www.uk10k.org/)) is a large-scale sequencing project, which was a collaborative project between investigators at the Wellcome Trust Sanger Institute and clinical experts in genetic diseases. 5,500 participants were diagnosed with a variety of disease phenotypes, including SCZ as well as other disorders such as obesity, autism disorder, familial hypercholesterolemia, thyroid disorder, learning disabilities, ciliopathies, congenital heart disease, coloboma, neuromuscular disorders, and rare disorders including severe insulin resistance. Within the disease arm of the project, there are 1,392 British samples with a SCZ diagnosis. Those individuals were whole-exome sequenced to 72x depth; hence, the generated data were sufficient to discover novel rare and low-frequency variants associated with the diseases investigated in UK10K. Similarly, 4,000 participants phenotyped as control subjects were supplied from the TwinsUK ([http://www.twinsuk.ac.uk/](http://www.twinsuk.ac.uk/)) registry and the ALSPAC study (Avon Longitudinal Study of Parents and Children, Bristol University) ([http://www.bristol.ac.uk/alspac/](http://www.bristol.ac.uk/alspac/)). This data is available via controlled access approved by the specified Data Access Committee (DAC) via the European Genome-Phenotype Archive (EGA) (Study ID EGAS000001000123). The SNP of the gene from sequencing data was extracted and performed using GeneSAOcs, which is custom software designed to extract information from Variant Call Format (VCF) files. Moreover, GeneSAOcs was slightly different to PLINK/SEQ, which using appropriate filter methods to correct data by only included 98% success rate in the sequencing. For instance, if the genotype calls were not made in more than 100 samples at a location, the SNP was excluded from the analysis. Secondly, the conditional genotype quality (GQ) encoded as a phred quality- the threshold was held at 30 as well as a lower limit on read depth at each location for each sample.
3.2.6.7 Swedish schizophrenia and control exome sequencing data

The Swedish schizophrenia (Purcell et al 2014) cohort comprises 2,536 control subjects and 2,543 SCZ patients from Sweden for whom whole exome sequence data was available through controlled access via dbGap (Study ID phs00473.v1.p1). The SCZ participants were identified through the Swedish Hospital Discharge Register (Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium 2011). The criteria were included two or more hospitalisations with a discharge diagnosis of SCZ, both parents were required to be Scandinavian descent, and age at 18 or above. By contrast, exclusive participants’ criteria were included hospital register diagnosis of any disorder justifying a validity diagnosis of SCZ. The normal control participants were randomly selected from Swedish population registers with none hospitalised for SCZ or bipolar disorder, Scandinavian descent of both parents, age at 18 or above. The data was extracted and analysis performed using GeneSAOcs software as described above.

3.2.7 Association study – burden analysis

Genotype and allele counts and minor allele frequencies were calculated for each individual SNVs found in the UCL case/control samples. Chi square (X²) test of association applied on individual variants was performed. However, since several SNVs were found as rare variants (minor allele frequency (MAF) > 0.05), where the Pearson X² tests of each individual rare variants have low statistical power to detect association. Pooled burden analysis was applied in the UCL case/control sample, where group the rare variants together as assuming a dominant model to improve power over a 2 degree freedom test. For example, all nsSNVs that found in cases were grouped as a dominant model, where assumed phenotypes are affected by multiple variants in CACNG5. Table 3-13 showed the counts of nsSNVs found in CACNG5 in UCL cases, where sum up each nsSNVs heterozygotes and/or mutated homozygotes into one big variant (total).
Table 3-13 Example of ns SNV alleles counts for pooled burden analysis

<table>
<thead>
<tr>
<th></th>
<th>BP n = 1073</th>
<th>SCZ n = 603</th>
<th>Control n = 941</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>R69W</td>
<td>1071</td>
<td>0</td>
<td>599</td>
</tr>
<tr>
<td>R71H</td>
<td>1067</td>
<td>1</td>
<td>595</td>
</tr>
<tr>
<td>R127Q</td>
<td>1071</td>
<td>2</td>
<td>596</td>
</tr>
<tr>
<td>T128M</td>
<td>1069</td>
<td>3</td>
<td>599</td>
</tr>
<tr>
<td>V146M</td>
<td>1070</td>
<td>0</td>
<td>601</td>
</tr>
<tr>
<td>I156F</td>
<td>1072</td>
<td>1</td>
<td>601</td>
</tr>
<tr>
<td>T164L</td>
<td>1071</td>
<td>0</td>
<td>598</td>
</tr>
<tr>
<td>H233Y</td>
<td>1019</td>
<td>55</td>
<td>562</td>
</tr>
<tr>
<td>Total</td>
<td>1011</td>
<td>62</td>
<td>561</td>
</tr>
</tbody>
</table>

Separate analyses were performed for different classes of variant such as non-synonymous, synonymous variants, and variants that found in promoter regions. The advantage of this approach is that is achieves a dimension reduction through aggregation of multiple variants into single units of analysis. Test of association were then performed on these aggregated variants (as shown in total in Table 3-13) using $X^2$ test with merged genotypes from UCL genotyping result, the 1,000 Genome Project data, UK10K exome data, and the Swedish exome data. For these analyses a p-value less than 0.05 was regarded as significant.

3.2.8 Bioinformatical analysis

A variant was designated as “novel” if it was not listed on the SNP track from UCSC Genome Browser (assembly February 2009; http://genome.ucsc.edu/cgi-bin/hgGateway) or in the 1000 Genome Project (http://www.1000genomes.org/). Nucleotide positions correspond to genomic sequence positions (NCBI build 36) in CACNG5 cDNA (924 bases; chr17:64,831,235 – 64,881,603). All rare variants were investigated for possible functional consequences using the following tools: (i) prediction of alteration of transcription factor binding – TESS (http://www.cbil.upenn.edu/cgi-bin/tess/tess); (ii) prediction of functional effects of amino acid substitutions, i.e. benign/possibly damaging/probably damaging – PolyPhen2 (http://genetics.bwh.harvard.edu/pph2/) and possibly have a deleterious effect on

### 3.3 Preparation of CACNG5 and AMPAR2 Constructs

To achieve functional studies of those variants that were genotyped in the UCL cases/control samples it was essential to have cloned copies of the CACNG5 and AMPA receptor GluR2 subunit gene (GRIA2). Gene cloning is a fundamental part of molecular biology, where a gene of interest is transferred from one organism to a self-replicating genetic element such as a bacterial plasmid. Bacterial plasmids are self-replicating extra-chromosomal circular DNA molecules, and are often used to generate multiple copies of the same gene (Figure 3-8). There are two main ways to achieve cloning: the polymerase chain reaction (PCR) and the more traditional use of restriction enzymes and modifying enzymes to “cut and paste” the desired gene into cloning vectors, which can then be replicated using live cells, most commonly E. coli. Investigation of the effect of introducing variants into these genes then allows analysis of the possibility that they may affect their protein function and expression in animal cultured cells.
Figure 3-8 Basic procedure for DNA (gene) cloning in a plasmid vector. (Taken from Health Science Academy - http://academygenbioii.pbworks.com/w/page/37767821/Chapter%2020%20Blog%20Genetic%20Technology%20(Larissa))
3.3.1 Plasmid vectors

Plasmid vectors replicate independently in host bacterial cells and facilitate the manipulation of a newly created recombinant DNA. Other properties of vectors include that they contain a multiple cloning site that contains a number of unique restriction endonuclease cleavage sites and carrying a selectable marker to distinguish the host cells carrying the vectors from the host cells that do not contain a vector.

pCR4-TOPO cloning vector:

The pCR4-TOPO cloning vector (Figure 3-9) was provided to carry fully sequenced Homo sapiens CACNG5 complementary DNA (cDNA) cloned (IMAGE: 8069139; Source BioScience, UK). For plasmid expression in mammalian cells, the plasmid vectors requires an enhancer promoter region from either the human cytomegalovirus (CMV), the simian virus 40 (SV40), or the herpes simplex virus thymidine kinase (HS-TK) to drive transcription. The CMV promoter induces high-level constitutive expression in a variety of mammalian cell lines. In fact, pCR-TOPO vector does not contain any of these high level constitutive expression promoter regions; hence, the further sub-cloning of Homo sapiens CACNG5 cDNA into a pCMV-Tag 4 vector was required. According to manufacturer’s vector description, the cDNA of CACNG5 was reversely inserted into TOPO cloning site, where contains several restriction enzyme sites such as Pst1, EcoR1 and Not1, whereas pCMV-Tag 4 vector has a multiple cloning site region contains Pst1 and Not1 restriction enzyme sites. The process of traditional restriction enzyme sub-cloning for Homo sapiens CACNG5 cDNA is described more detailed below 3.3.2 section.
Figure 3-9 Circular map of the pCR4-TOPO vector (Taken from Invitrogen, UK)
pCMV-Tag 4 vector:

pCMV-Tag 4 vectors is an epitope tagging mammalian expression vectors with a C-terminal FLAG tagging (Invitrogen) (Figure 3-10). In addition to the features of this vector, the cytomegalovirus (CMV) promoter allows constitutive expression of the cloned DNA in a wide variety of mammalian cell lines. The neomycin-resistance gene is under control of both the prokaryotic b-lactamase promoter to provide kanamycin resistance in bacteria and the SV40 early promoter to provide G418 resistance in mammalian cells. The multiple cloning site (MCS) of the pCMV-Tag vectors allow for a variety of cloning strategies, resulting in C-terminal functions with FLAG. The relative location of the feathers for this vector is also shown in Table 3-13. Because of these features, the Homo sapiens CACNG5 cDNA from pCR4-TOPO vector was then sub-cloned into pCMV-Tag4 via their Pst1 and Not1 restriction enzyme cut and paste (as described more detail in section 3.3.2).
Figure 3-10 Circular map of the pCMV-Tag4 vector (Taken from Invitrogen, UK)

Table 3-14 A table corresponed to pCMV-Tag4 vector features and its postion lists

<table>
<thead>
<tr>
<th>Feature</th>
<th>Nucleotide Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMV promoter</td>
<td>1–602</td>
</tr>
<tr>
<td>T3 promoter and T3 primer binding site [5’ AATTACCCTCACTAAGGG 3’]</td>
<td>620–639</td>
</tr>
<tr>
<td>multiple cloning site</td>
<td>651–743</td>
</tr>
<tr>
<td>FLAG tag</td>
<td>744–767</td>
</tr>
<tr>
<td>T7 promoter and T7 primer binding site [3’ CCGGATATCCTCAGCTATTA 5’]</td>
<td>819–840</td>
</tr>
<tr>
<td>SV40 polyA signal</td>
<td>852–1235</td>
</tr>
<tr>
<td>fl origin of ss-DNA replication</td>
<td>1373–1679</td>
</tr>
<tr>
<td>bla promoter</td>
<td>1704–1828</td>
</tr>
<tr>
<td>SV40 promoter</td>
<td>1848–2186</td>
</tr>
<tr>
<td>neomycin/kanamycin resistance ORF</td>
<td>2221–3012</td>
</tr>
<tr>
<td>HSV-thymidine kinase (TK) polyA signal</td>
<td>3013–3471</td>
</tr>
</tbody>
</table>
**pCMV-Sport6 cloning vector:**

The cloned Homo sapiens glutamate receptor, AMPA receptor subunit 2 (*GRIA2*) into mammalian expression vector pCMV-Sport6 was also provided (IMAGE: 4215347; Source BioScience, UK). pCMV-Sport6 vector contained *GRIA2* was ready to express in mammalian cells, as it contains CMV and SV40 regions (Figure 3-11).

![Circular map of the pCMV-Sport6 vector](Taken from Invitrogen, UK)

**Figure 3-11 Circular map of the pCMV-Sport6 vector (Taken from Invitrogen, UK)**
pCI-SEP-GluR2 (R) cloning vector:

The cloned Rattus norvegicus glutamate receptor, AMPA receptor subunit 2 (Gria2) into mammalian expression vector pCI containing super-ecliptic pHluorin (SEP) coding sequence was provided from Addgene (Plasmid 24002). The SEP coding sequence was inserted three amino acids downstream of the predicted signal peptide cleavage site of the corresponding AMPAR subunit 2. The resultant product was then inserted into the pCI (Promega, Madison, WI) to generate pCI-SEP-GluR2 as shown in Figure 3-12.

Figure 3-12 Circular map of the pCI vector (Taken from Promega, Madison, WI)
Signal peptide in Gria2 was added to ensure that peptide sequences destined for either secretion or membrane integration cannot fold or misfold in the cytosol. Generally, cell surface membrane or secretory proteins such as AMPA receptors are directed to the endoplasmic reticulum (ER) following their translation (Higy et al. 2004). Hence, the presence of a signal peptide for targeting proteins to the ER is required, which is a small hydrophobic region of between 7 – 25 amino acids located at the N-terminus of that protein (Dev and Ray 1990). Initially, the signal peptide is translated by the ribosome, where it binds to a signal recognition particle (SRP). The binding of SRP transiently arrests translation and migrate the ribosome-SRP complex to the ER membrane via binding of SRP receptor (Gilmore et al. 1982, Walter and Blobel 1982). Once the complex arrived in the ER membrane, the signal peptide is cleaved, and the protein translation resumes (Connolly and Gilmore 1989). In this case, a predicted signal peptide sequence of 21 amino acids was located at the extreme N-terminal end of Gria2 (Everts et al. 1997)

SEP is an enhanced version of eclipitic pHluorin (EP), which is a pH-sensitive variant of green fluorescence protein (GFP). Generally, the excitation profile of native GFP possesses with a major peak at 395 nm and a minor peak at 475 nm measured with an emission maximum at 509 nm. In contrast, EP is almost invisible at pH values less than 6.0 (acidic conditions), but the fluorescence intensity increases with pH up to a maximum of 8.5 (Ashby et al. 2004, Miesenbock et al. 1998). For instance, targeted expression of EP to acidic secretory vesicles allows the real-time analysis of the exocytotic fusion of these vesicles to the plasma membrane where the pH becomes neutral, resulting in a corresponding increase in fluorescence. SEP tagging is generally applied for visualising AMPARs on the surface of neurons (Sankaranarayanan et al. 2000). The enhanced version of pH-sensitive GFP variant exhibits much brighter fluorescence compared to the EP. Thus, a chimera of SEP fused to the N-terminus of Gria2 (SEP-Gria2) acts as a marker for surface-expressed receptor.
**pTagRFP-C cloning vector:**

pTagRFP-C is a mammalian expression vector encoding red (orange) fluorescent protein TagRFP. The vector allows generation of fusions to the TagRFP C-terminus and expression of TagRFP fusion protein in mammalian cells (Figure 3-13). TagRFP codon usage is optimised for high expression in mammalian cells: Kozak consensus translation initiation site is generated upstream of the TagRFP coding sequence to increase mRNA translation efficiency. A multiple cloning site is located between TagRFP coding sequence and SV40 polyadenylation signal (SV40 polyA). Finally, the CMV promoter region provides strong, constitutive expression of TagRFP.

The vector backbone contains immediate early promoter of cytomegalovirus (pCMV-IE) for protein expression, SV40 origin for replication in mammalian cells expressing SV40 T-antigen, pUC origin of replication for propagation in E.coli, and f1 origin for single stranded DNA production. SV40 polyadenylation signals (SV40 poly A) direct proper processing of the 3’-end of the reporter mRNA. SV40 early promoter (pSV40) provides neomycin resistance gene (Neo) expression to select stably transfected eukaryotic cells using G418. Bacterial promoter (P) provides kanamycin resistance gene expression (Kan) in *E. coli*.

Homo sapience *CACNG5* cDNA was also sub-cloned into MCS of this vector. It was expressed as a fusion to the TagRFP C-terminus when inserted in the same reading frame as tagRFP and no in-frame stop codons were present (more detail is described in section 3.3.2).
3.3.2 Sub-cloning of CACNG5 inserts into pCMV-Tag4 vectors via traditional restriction sites

Upon sequence verification of the pCR4-TOPO clones containing 926 bp of CACNG5 cDNA, this inert was subcloned into pCMV-Tag4 as destination vector resulting in a wild-type of CACNG5 construct. pCMV-Tag4 vector was chosen, because of their basal transcriptional level difference between these two vectors. pCR4-TOPO vector lacks a constitutive promoter sequence result in low transcriptional level; by contrast, pCMV-Tag4 presents an SV-40 derived promoter and cytomegalovirus CMV promoter sequence to increase its transcriptional level. Two double digestion reactions using restriction enzymes (PstI and NotI) common to the pCR4-TOPO and pCMV-Tag4 vector of interest were performed. This process was to cut out the insert from the pCR4-TOPO vector and also cleave the pCMV-Tag4 vector. The restriction enzyme digest was performed in a volume of 20 µl consisting of 2 µl of 10X restriction enzyme buffer, 0.2 µl of acetylated BSA (10 µg/µl), 1 µl of plasmid DNA (1 µg/µl) and 14.8 µl of sterile deionised water, and finally a 1 µl of each restriction enzyme (10 U/µl).
The mixture in the tube was mixed gently by pipetting and spun briefly and incubated at 37°C for 4 hours. In the double restriction enzyme digestion, the most optimal buffer for the activity of both enzymes was applied according to Promega Relative activity of Restriction Enzyme Buffer Guideline. The additional acetylated BSA in the digestion reaction was to stabilise the enzymes by protecting them from proteases, non-specific adsorption and harmful environmental factors such as heat, surface tension and interfering substances. After 4 hours incubation, the digested reaction mix was analysed by 1% agarose gel electrophoresis (see above) to confirm complete cleavage of the plasmid DNA and excision of the insert from the vector. Successful digestion of the pCR4-TOPO vector showed two bands of appropriate sizes for the vector and the insert on the gel (Figure 3-14 upper panel) and one clear band for the cleaved pCMV-Tag4 vector (Figure 3-14 bottom panel). The principle of DNA fragments
migration in agarose gel confirmed the supercoiled plasmids as uncut pCMV-Tag4 migrated more rapidly than linear plasmids as cleaved pCMV-Tag (Figure 3-14 bottom panel).

After gel confirmation, the 926 bp insert and 4,319 bp cleaved pCMV-Tag4 vector were then extracted directly from the gel using QIAquick Gel Extraction Kit (Qiagen, UK) according to their manufacturer’s instructions. The extracted DNA was then quantified using NanoDrop® ND-1000 Spectrophotometer (Thermo-Scientific, UK), and ligated using T4 DNA ligase (Promega, UK).

### 3.3.2.1 Ligation reaction of insert into pCMV-Tag4 vector
Initially the sticky ends 926 bp CACNG5 digestion product was cloned into the pCMV-Tag4 sticky cloning vector following the protocol described below. A 3:1 molar ratio of insert:vector DNA was applied, and the amount of insert and vector was calculated using the formula showed as below:

\[
\text{ng of insert} = \left(\frac{\text{kb size of insert}}{\text{kb size of vector}}\right) \times \text{ng of vector} \times \frac{\text{Molar ratio of insert:vector}}{3}
\]

The ligation reactions were performed at 4°C overnight with a 10 µl volume of ligation mix consisting of 200 ng of pCMV-Tag4 vector, 35.33 ng of CACNG5 insert, 1 µl of ligase buffer, 1 µl of T4 DNA ligase (3 u/µl), and deionised water. After the ligation reaction, the ligation mix was then used to transform 50 µl of maximum efficiency DH5α/XL-10 Gold competent cells according to the standard protocol described below in section 3.3.2.2.
3.3.2.2 Transformation

50 µl of Max efficiency DH5α *E. coli* competent cells (Invitrogen) was briefly thawed on ice and transformed with 3 µl of the ligation mixture and incubated on ice for 30 minutes. After the incubation, a 42°C heat-shock was performed in a water bath for 45 seconds. The cells were immediately placed back on ice for 2 minutes. Then the cells were added to 450 µl of room temperature S.O.C medium and placed in a shaking incubator at 37°C for one hour at 225 rpm. Once the incubation time had elapsed, aliquots of 50 µl, 100 µl, and 200 µl of cells were spread on LB agar plates (LB media containing 1.5% agar under sterile conditions) containing 100 µg/ml kanamycin. Additional aliquots of 100 µl of competent cells were transformed in parallel with 50 pg of control DNA (pUC19) to determine the transformation efficiency. The plates were then inverted and incubated overnight at 37°C. Successfully transformed colonies survived on LB-amp agar plates due to the selection with kanamycin. Colony PCR of the transformed colonies was then used screen for the presence of the cloned insert.

3.3.2.3 Colony PCR for confirmation of the presence of the cloned insert

The simplest and most time efficient method for screening plasmid inserts directly from bacterial colonies is colony PCR. The transformed colony on the plate was picked up using a sterile 10 µl pipette tip and dipped in a 96-well plate containing 5 µl of PCR water. And the tip was then used to inoculate into a fresh 15 ml falcon tube containing 3 ml of LB (Luria-Bertani) medium broth with kanamycin. For amplification of the 926 bp product, the PCR master mix was prepared as described in Table 3-7, and added to a 96-well plate containing colony cells. The optimum cycling conditions were applied based on the insert of interest. After the PCR amplification, the fragments were run on 1% agarose gel to determine their specificity (Figure 3-15).
3.3.2.4 Sequence verification of plasmid DNA

The inoculated falcon tubes containing single transformed colonies carrying with the vector along with the insert were grown overnight in a shaker with 225 rpm at 37°C. On the next day, the plasmid was isolated from the cultures using QIaprep Spin Miniprep Kit (Qiagen, UK) based on the manufacturer’s protocol. The quantification of plasmid DNA was then applied using a NanoDrop® ND-1000 or NanoDrop 2000 Spectrophotometer (Thermo-Scientific, UK), and then sequenced at the Scientific Support Services at the Wolfson Institute for Biomedical Research (WIBR), UCL using CACNG5 exon 3 forward and reverse primers.

3.3.3 Sub-cloning of CACNG5 inserts into pTagRFP-C vector via PCR amplification

The advantages of applying additional red fluorescent tag protein (RFP) sequence at the N-terminus of CACNG5 sequence is to allow the visualization of CACNG5 proteins using both fixed and live cell images of fluorescent microscopy. Hence, the PCR forward primer was designed to include an ATG codon for an open reading frame site between REP sequence and CACNG5 starting coding region. In this way, the transfection efficiency of CACNG5 can be detected via fluorescence-activated cell sorting (section 3.5.2), and the appropriate antibodies can also target to RFP for western blot (section 3.5.5).
3.3.3.1 Amplification of a 894 bp CACNG5 with Xhol1 and EcoR1 restriction sites

A 894 bp CACNG5 without non-coding region PCR product with Xhol1 and EcoR1 restriction sites were amplified using the primers showed below (Sigma-Aldrich, UK)

Forward Primer = 5’ - TCAGATCTCGAGGAATGAGTGCC – 3’

Reverse Primer = 5’ – ACTGCAGAATTCTCAGCAGGGTGGAAGA – 3’

The master mix for PCR amplification consisting of 7.5 µl of 10X Pfx amplification buffer, 7.5 µl of 10X enhancer buffer, 0.5 µl of 50 mM MgSO₄, 0.3 µl of 25 mM dNTPs, 1 µl each of 10 pmol/µl forward and reverse primers, 0.4 µl of Platinum® Pfx DNA polymerase (Invitrogen, UK) and PCR water so added to a final volume of 23 µl and then dispensed into a 96-well plate containing 2 µl of 10 ng CACNG5 insert. Platinum® Pfx DNA polymerase is a highly processive polymerase that possesses proofreading 3’ to 5’ exonuclease activity with fast chain extension capability. Polymerase activity was restored after a PCR denaturation step, where providing an automatic “hot start” for increased specificity, sensitivity, and yield of the PCR product. Both the vector and insert were denatured at 94ºC for 5 minutes initially, and the PCR amplification was performed with 30 cycles of three steps cycling involving: denaturation at 95ºC for 15 seconds, annealing at 55ºC for 30 seconds, and extension at 68ºC for 5 minute and 30 second. The last extension was at 72ºC for 2 minutes and 30 second, and then maintained the reaction at 4ºC. The size of the amplified fragment was further confirmed by running a 1% agarose gel with ethidium bromide and hyperladder I as the size marker (Figure 3-16).
Upon verification, the amplified fragment band was extracted from the gel using Qiagen MinElute Gel Extraction Kit (Qiagen, UK) according to the manufacture’s instruction. The purified amplified fragment was rehydrated with PCR water, and quantified using a NanoDrop 2000 Spectrophotometer (Thermo-Scientific, UK). And the sequence verification was performed at the WIBR, UCL using CACNG5 exon 2 reverse and exon 4 forward primers.

3.3.3.2 Sub-cloning of the CACNG5 inserts into pTagRFP-C vectors

Before sub-cloning the amplified CACNG5 inserts, pTagRFP-C vectors were cleaved via traditional restriction enzyme digestion as described in section 3.3.2. Double restriction enzyme digestion was performed using both Xhol1 and EcoR1 enzyme for vectors cleavage. After gel and sequence verification of both inserts and vector, ligation reactions were carried out using T4 DNA ligase as described in section 3.3.2.1. The plasmids were propagated in DH5α/XL-10 Gold competent cells according to the protocol described in section 3.3.2.2 and isolated using Qiagen miniprep kits. The plasmid DNA was then verified by sequencing at WIBR, UCL.
3.3.4 Mutating the CACNG5 gene with each variant individually

To investigate the effect of CACNG5 variants found in UCL case/control samples, each single base of non-synonymous mutation was introduced into both wild-type CACNG5 constructs via site-directed mutagenesis using the QuikChange II XL mutagenesis kit (Strategene, UK) according to the manufacturer’s instructions. The primer for mutating CACNG5 constructs were designed using QuikChange Primer Design Program website available online at www.agilent.com/genomics/qcpd and briefly described in section 3.3.4.1.

3.3.4.1 Primer designed for mutating CACNG5 constructs

Each pair of primers contained a non-overlapping sequence at the 3’-terminus and a primer-primer complementary (overlapping) sequence at the 5’-terminus. The non-overlapping sequence was significantly larger than the overlapping to make the melting temperature of the former higher than that of the latter. The primers that were used for mutagenesis are listed in Table 3-14. These primers targeted residues R69W, R71H, R127H, T128M, V146M, T164L, and H233Y.
Table 3-15 Primer sets for CACNG5 constructs site-directed mutagenesis

<table>
<thead>
<tr>
<th>Amino Acid Change</th>
<th>Nucleotide Change</th>
<th>Primer Sequence (5' to 3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arg69Trp</td>
<td>c.265 C&gt;T</td>
<td>cacaggtgaggagtgggggcgttgctt</td>
</tr>
<tr>
<td></td>
<td></td>
<td>antisense-aagcaacgccccacctcctacactgtg</td>
</tr>
<tr>
<td>Arg71His</td>
<td>c.272 G&gt;A</td>
<td>gaggagcgggggcattgcttcaccatag</td>
</tr>
<tr>
<td></td>
<td></td>
<td>antisense-ctatggtgaagcagccccccccctcctc</td>
</tr>
<tr>
<td>Arg127Gln</td>
<td>c.440 G&gt;A</td>
<td>atccgtccccacccagacgataactgccc</td>
</tr>
<tr>
<td></td>
<td></td>
<td>antisense-gccgagtacggtgctcgcttgctctg</td>
</tr>
<tr>
<td>Arg128Met</td>
<td>c.443 C&gt;T</td>
<td>ccgctccccaccccagatctggttccttg</td>
</tr>
<tr>
<td></td>
<td></td>
<td>antisense-caaagggccagatcactcggctggagccc</td>
</tr>
<tr>
<td>Val146Met</td>
<td>c.496 A&gt;G</td>
<td>ggcctctctctctctgtggtgggcttg</td>
</tr>
<tr>
<td></td>
<td></td>
<td>antisense-cagggcccacactgagagaggaggg</td>
</tr>
<tr>
<td>Ile156Phe</td>
<td>c.526 A&gt;T</td>
<td>tgtctctatccttcagctcagccagatgagcttc</td>
</tr>
<tr>
<td></td>
<td></td>
<td>antisense-gagtctctctctcagctgtgagctctgagccacagagcagag</td>
</tr>
<tr>
<td>Thr164Leu</td>
<td>c.551 C&gt;T</td>
<td>gagatgctcaacaggtagacagatcagagagcttc</td>
</tr>
<tr>
<td></td>
<td></td>
<td>antisense-ggtctctctcagctctggtgagctctcagagagcagag</td>
</tr>
<tr>
<td>His233Tyr</td>
<td>c.757 C&gt;T</td>
<td>attactcagctcacttacacccagacgc</td>
</tr>
<tr>
<td></td>
<td></td>
<td>antisense-ggctctgggtataggaactggctctagttaat</td>
</tr>
</tbody>
</table>

3.3.4.2 Site-directed mutagenesis

Site-directed mutagenesis primers were purchased from Sigma (UK) with addition high-performance liquid chromatography (HPLC) purification. The primers were reconstituted with PCR water. The synthesis reaction of mutant strand was performed by reaction mix consisting of 5 µl of 10X reaction buffer, 10 ng of wild type CACNG5 plasmid DNA, 125 ng of forward and reverse primers, 1 µl of dNTP mix, 3 µl of QuickSolution and PCR water to a final volume of 50 µl in a 96-well PCR plate. At the end, 1 µl of Pfu Ultra HF DNA polumerase (2.5U/µl) was added to the reaction mix and mixed gently. The samples were then placed in a PCR machine using the cycling conditions described in Table 3-15.
Table 3-16 The PCR cycling conditions for CACNG5 constructs site-directed mutagenesis

<table>
<thead>
<tr>
<th>Cycles</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>95°C</td>
<td>1 minute</td>
</tr>
<tr>
<td>18</td>
<td>95°C</td>
<td>50 seconds</td>
</tr>
<tr>
<td></td>
<td>60°C</td>
<td>50 seconds</td>
</tr>
<tr>
<td></td>
<td>68°C</td>
<td>5 minute and 30 seconds</td>
</tr>
<tr>
<td>1</td>
<td>68°C</td>
<td>7 minutes</td>
</tr>
</tbody>
</table>

After the site-directed mutagenesis, the mutated plasmid DNA was transformed into E.coli as described above. The presence of the mutation was verified by DNA sequence analysis as described above.

3.3.5 Large-scale preparation of plasmid DNA

For transfection experiments, large-scale plasmid DNA preparations were performed using 25 ml LB medium containing appropriate antibiotic cultures inoculated with 50 µl of starting culture (the colonies growth culture) and grown overnight. The plasmid DNA was then extracted using QIAfilter Plasmid Midi Kit (Qiagen, UK) according to the manufacturer’s instructions.

3.3.6 Preparation of bacterial glycerol stocks for long-term storage of clones

After a single colony was inoculated in 3 ml of LB medium with appropriate antibiotic for overnight at 37°C, 500 µl of the growth culture was added to equal volume of 80% sterile glycerol in a sterile screw cap tube. The glycerol stock was then mixed well with vortex and then stored at -80°C. Meanwhile, the plasmid DNA was also isolated from 1 ml of the same growth culture using Qiagen miniprep kit, and then verified by sequence for confirmation.
3.4 Functional characterisation of CACNG5 non-synonymous variants

3.4.1 Animal cell-culture

3.4.1.1 Cell-line

Human embryonic kidney 293 cells HEK293 cells were used in this study. HEK293 cells were provided from Queen Mary University of Cancer research department. HEK293 cells were originally from the transformation of human embryonic kidney cells with sheared fragments of human adenovirus type 5 (Ad5) DNA in 1977 by Graham, Smiley, Russell, and Nairn (Graham et al. 1977). The SH-SY5Y cell line was used in the preliminary experiments. The SH-SY5Y cell line was originally isolated from a bone marrow biopsy of a neuroblastoma patient, called SK-N-SH. However, because of their low transfection efficiency, it was not used to our main functional experiments.

3.4.1.2 Sub-culturing mammalian cells

Solutions required:

- Growth medium composition: Fetal Bovine Serum (FBS; Gibco-Invitrogen, UK), 10%; 5,000 units/ml penicillin and 5,000 μg/ml streptomycin (Pen/Strep; Gibco-Invitrogen, UK), 1%; all added directly to a 500 ml bottle of high glucose (with 4,500 mg/L glucose) Dulbecco’s Modified Eagle’s Medium (DMEM; Sigma-Aldrich, Poole, UK), which was then mixed by carefully pipetting up and down several times.

- Trypsin-EDTA solution (Sigma-Aldrich, UK)

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

Cultured cells were viewed using an inverted low power microscope to assess the degree of confluency and confirm the absence of bacterial and fungal contaminants. When cells reached to 75 – 95% confluent, the cells were split or sub-cultured using the procedure shown below:

1. Culture media from the tissue culture flask was aspirated and the cell monolayer was then washed with PBS containing additional MgCl₂ and CaCl₂ to remove all serum.
2. 2 ml of 1X trypsin-EDTA solution was added to the cell culture and left for no more than 2 minutes to disaggregate cells. Additional tapping the side of the flask could help disaggregation.
3. After trypsin-EDTA treatment for 2 minutes, 10 ml of fresh culture medium was added to the flask to stop its reaction.
4. The suspended cells were aspirated into a new 15 ml falcon tube, and spun in a centrifuge at 1,500 rpm for 5 minutes.
5. The cell pellet was then re-suspended in 10 ml fresh culture medium.
6. 1 ml of the cell suspension was added into a 9 ml of fresh culture medium in a 75 cm² flask for cell seeding.
7. The flask was then gently swirled to mix and incubated at 37°C/ 5% CO₂ for approximately 3 – 4 days until cells were 75% - 95% confluent.
3.4.1.3 Mammalian cell quantification

Materials required:

- 0.4% Trypan blue solution (Sigma): Trypan blue is a vital stain that used to differentiate cell viability. The intact cell membranes of live cells are very selective in the compounds that pass through the membrane; hence, a viable cell is not absorbed cell trypan blue. By contrast, when the cell membrane is damaged, dead cells are appeared as a distinctive blue colour.

- Neubauer hemocytometer

- Microscope

Background:

A haemocytometer is used to estimate the number of viable cells under a microscope, which originally is designed for counting of blood cells. The improved Neubauer haemocytometer is a thick glass microscope side with two rectangular indentations that creates two counting chambers, and the main divisions separate each chamber into nine large squares (Figure 3-17). The gridded area of the haemocytometer consists of several 1 mm$^2$ surface area and 0.1 mm depth. With the coverslip in place, each square of the haemocytometer represents a total volume of 0.1 mm$^3$ or $10^{-4}$ cm$^3$ (1 cm$^3 = 1$ ml). For cell counting, the cell suspension is diluted in order to avoid clumps of cells. The cells are then mixed thoroughly to ensure a uniform distribution. The concentration of cell per ml was determined using the following equation:

$$\text{Cells per ml} = \text{the average count of cells per square} \times \text{the dilution factor} \times 10^4$$
Figure 3-17 A diagram of the improved Neubauer haemocytometer. Panel A left represents as top view of haemocytometer; Panel A right represents as side view of haemocytometer; Panel B represents as haemocytometer grid, each one of blue square has a 1.0 mm$^2$ surface area and a 0.1 mm$^3$ volume.

**Cell counting protocol:**

Before cell counting, both the haemocytometer and the coverslip was cleaned using 70% ethanol. The coverslip was moistened by exhaled breath on the surface, and then it was slide over the chamber back and forth using gentle downward pressure. The cell suspension was prepared using the subculturing protocol described above, and a 100 µl of cell suspension was then dispensed into a sterile microcentrifuge tube. An equal volume of Trypan Blue solution (dilution factor = 2) was added to the cell suspension in the microcentrifuge tube and mixed gently by pipetting. Both sides of the chamber were filled with mixed aliquot (10 - 20 µl) of cell suspension containing Trypan Blue solution via capillary diffusion. Finally, the haemocytometer was viewed under the light microscope using x20 magnification. Analytically, both viable and non-viable cells in the four corner squares were counted
excluding the cells lying on the outer perimeter. The concentration of viable and non-viable cells and the percentage of viable cells were calculated using the equations as shown below:

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

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

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

3.4.1.4 Preservation of mammalian cell-line

Preservation of cell cultures was used to maintain supplies of cell and to provide a backup supply of cells, in case of contamination or to have a stock of cells with a lower passage number. The preservation cells are when they are at their maximum growth rate to help in good recovery during the thawing of cells. Two ways of preservation were used in this thesis: cryogenic preservation and Bambanker™ cryopreservation.

3.4.1.4.1 Cryogenic preservation of cell-line

**Solution required:**

- Freeze medium consisting of 20% FBS, 10% DMSO and 1% PenStrep in DMEM growth medium

**Cryogenic preservation protocol:**

The freezing medium containing DMSO (Sigma, UK) is a cryoprotective reagent that reduces the freezing point of the culture medium, and also allows a slower cooling rate to reduce the risk of damaging cells from ice crystal formation. The cells were harvested using the standard subculturing protocol described above, and suspended in 2.5 ml of DMEM supplemented with 20% FBS and 1% Penstrep. After that, a 20% DMSO was added to another fresh falcon tube containing 2.5 ml of DMEM supplemented with 20% FBS and 1% Penstrep. This
medium was then slowly added to the tube containing cells drop by drop, and gently mixed. 1 ml aliquots of this cell suspension was transferred into 1.5 ml cryogenic vials (Nunc, UK) that was labelled with the cell line name, passage number, cell concentration and date. The cells containing cryogenic vials were stored in a freezer container, Nalgene®, Mr.Frosty (Sigma-Aldrich, UK) at -80°C. This freezer container is made of polycarbonate, which has a blue high-density polyethylene closure and a white high-density polyethylene via holder and foam insert to provide a critical, repeatable, 1°C/min cooling rate. This specific cooling rate is required for successful cryogenic preservation of cells. Additionally, a day before preservation procedure, the cryogenic container was filled with isopropyl alcohol overnight at -20°C. This preservation protocol can store cell-line at -80°C up to 6 months.

3.4.1.4.2 Bambanker™ cryopreservation of cell lines

The freezing medium Bambanker™ was originally developed by the Japanese company Lymphotec, where they required a suitable medium for long-term storage of highly sensitive cell lines, such as lymphocytes. To date, cell freezing medium Bambanker™ has been demonstrated in many different published articles, to allow the preservation of a high number of intact cells after thawing. The recovery rate of even sensitive cells is much higher compared to traditional media.

Solution required:

- Bambanker™ freezing medium (NIPPON Genetics EUROPE GmbH, Germany)

Bambanker™ cryopreservation protocol:

The freezing medium Bambanker™ is a serum-free medium for long-term storage of cells at -80°C. The cells were harvested using the standard subculturing protocol described above, but suspended in 1 ml of Bambanker™ and placed the cells in cryo tubes that suitable for
freezing and preservation. Then, the cells are ready to freeze and preserve in -80°C without preliminary freezing.

3.4.1.5 Resuscitation of frozen cell lines

The frozen stock cells is required a quickly thawing procedure to maximise the recovery of cells and minimise the loss of cells from the presence of DMSO in the freezing medium. Cryogenic vials were taken from -80°C freezers and immediately warmed up in a 37°C water bath for less than a minute.

Frozen cells persevered in medium containing DMSO:

The 1 ml of stock cells in cryogenic vial were transferred into a fresh 15 ml falcon tube containing 5 ml pre-warmed growth medium and mixed gently by pipetting.

Frozen cells preserved in Bambanker™ medium:

The 1 ml of stock cells in Bambanker freezing medium were suspended into a fresh 15 ml falcon tube containing 10 ml pre-warmed growth medium, and mixed gently.

The cells were then centrifuged at 1500 rpm for 3 minutes and the supernatant discarded. The cell pellet was resuspended in 1 ml of growth medium and then transferred into a 75 ml tissue culture T-flask containing 9 ml of complete growth medium. Cells were viewed under the microscope to check for their presence and then incubated in a humidified atmosphere of 5% CO₂ at 37°C until 75 – 95% confluency was reached.

3.4.2 Transfection

Transfection is the most common method for introducing new functional DNA sequences into cultured animal cells either transiently or stably. The transfected cells can then gain the ability to express one or more proteins. Transfection involves the use of chemical or physical tricks to persuade cells to take up DNA from the culture medium, the DNA eventually
finding its way to the nucleus. There are numerous different transfection methods including: chemical transfection, liposome-mediated transfection, lipofection, electroporation, and receptor-mediated endocytosis. Nowadays, different types of transfection are commercially available, and more details of transfection are described in Appendix I.

3.4.2.1 Lipofectamine® 2000 transfection

Reagents required:

- Lipofectamine® 2000 (Invitrogen Life technologies, UK)
- DMEM supplemented with 10% FBS and with or without 1% Penstrep
- Opti-MEM® I Reduced Serum Medium as serum-free culture medium for complex formation

Background:

Lipofectamine® 2000 is the most common transfection reagent used mainly in molecular and cellular biology, which is provided by Invitrogen. It is a cationic liposome based reagent that provides high transfection efficiency and high levels of transgene expression in a range of mammalian cell types. Optimum transfection efficiency and subsequent cell viability depend on a number of experimental variables such as cell density, liposome and DNA concentrations, liposome-DNA complexing time, and the presence or absence of media components such as antibiotics and serum. The important of these factors in Lipofectamine® 2000 mediated transfection is discussed in Appendix I.

Transfection protocol:

A day before transfection, HEK293 cells were seeded 2 x 10^5 cells per well in 2 ml of antibiotic-free culture medium (DMEM + 10% FBS) in 6 wells plates. The plates were then
incubated at 37°C with 5% CO₂, and were attained 90 – 95% confluence at the time of transfection.

On the transfection day, the transfection complexes (DNA-lipofectamine® 2000 complexes) were prepared for each transfection sample. First, plasmid DNA was diluted in 500 µl of Opti-MEM®I Reduced Serum Medium, and placed in a 1.5 ml of microcentrifuge tube. The stock of lipofectamine® 2000 was gently mixed before it was diluted in another 500 µl of Opti-MEM®I Medium and then incubated for 5 minutes at room temperature. After incubation, the diluted DNA was combined with diluted Lipofectamine®2000, and then mixed gently and incubated for another 20 minutes at room temperature to allow the DNA-Lipofectamine® 2000 lipoplexes to form. 1000 µl of transfection complexes were then added to each well containing cells and medium, and this was mixed gently by rocking the plate back and forth. The cells were incubated at 37°C in a humidified incubator with 5% CO₂. The medium was changed after 6 hours of incubation with complete growth medium (DMEM + 10% FBS + 1% Penstrep) and the cell were incubated for further 24 hours until they were ready for expression analysis.

3.4.2.2 FuGENE® 6 transfection

Solution required:

- FuGENE® 6 Transfection reagent (Promega, UK)
- DMEM supplemented with 10% FBS and with or without 1% Penstrep
- Opti-MEM® I Reduced Serum Medium
Background:

FuGENE® 6 transfection is also the most common reagent to be used for cellular transfection. It is a nonliposomal reagent transfects DNA into a wide variety of cell lines with high efficiency and low toxicity (more details are shown in Appendix 1)

Transfection protocol:

A day before transfection, HEK293 cells were seeded 1 x 10^5 cells per well in 500 µl of antibiotic-free culture medium (DMEM + 10% FBS) in 24 wells plates. The plates were then incubated at 37ºC with 5% CO₂, and were 90 – 95% confluent at the time of transfection.

On the transfection day, the vial of FuGENE® 6 Transfection reagent was warmed up to the room temperature and mixed by inverting the vial briefly. And the transfection optimisation was prepared using 10 µg of DNA per well at various ratios of FuGENE® 6 Transfection reagent to DNA (Table 3-17).

Table 3-17 Optimisation protocol using varying ratios of FuGENE® 6 Transfection reagent to DNA

<table>
<thead>
<tr>
<th></th>
<th>Ratio of FuGENE Transfection Reagent to DNA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6:1</td>
</tr>
<tr>
<td>A final volume of medium in a well</td>
<td>500 µl</td>
</tr>
<tr>
<td>DNA amount</td>
<td>10 µg</td>
</tr>
<tr>
<td>Volume of FuGENE 6 Transfection Reagent</td>
<td>60 µl</td>
</tr>
</tbody>
</table>

The varying amount of FuGENE® 6 transfection reagent was added into the 1.5 ml of microcentrifuge tube containing 500 µl of Opti-MEM medium (Table 3-17). This FuGENE® 6 transfection reagent/medium mixture was mixed gently by pipetting, and incubated at room temperature for 5 minutes. After this incubation, 10 µg of plasmid DNA was added to the mixture of FuGENE® 6 transfection reagent/medium, and mixed immediately and incubated again for 15 minutes at room temperature. After 15 minutes incubation, 25 µl of the
FuGENE® 6 transfection reagent/DNA mixture was added to each well of a 24-well plate containing 500 µl of cells in growth medium. The cells containing 24-well plate was mixed gently and returned to the 37°C humidified incubator with 5% CO₂ for 24 hours until they were ready to be analysed.

Live cell images of co-transfected HEK293 cells were then taken by a Zeiss AXIO Observer microscope equipped with a temperature-controlled stage at 37°C.

3.4.3 Protein isolation

Protein extraction or isolation involved an initial step of cell lysis, where the cell membrane was disrupted and the cellular contents were extracted. Historically, physical lysis was the common method for protein extraction; however, it often requires expensive, cumbersome equipment, and is not conducive for high throughput analysis. In recently years, detergent-based lysis methods have become the best choice. Different detergent-based solutions composed of particular types and concentrations of detergents, buffers, salts and reducing agents have been developed to provide the best possible protein yielding for particular types of cells and proteins. For instance, non-ionic detergent such as Triton X-100 has uncharged and hydrophilic headgroups, which are considered mild surfactants as they break protein-lipid, lipid-lipid interactions, and most of them do not denature proteins. Therefore, non-ionic detergent was preferred for the isolation of membrane proteins, such as TARP and AMPA receptor proteins.

3.4.3.1 Cell lysis

Buffers and solutions required:

- 2 M Tris-HCl buffer: Trizma® hydrochloride (Sigma, UK)
- 0.5 M NaCl buffer: sodium chloride (Sigma, UK)
- 0.5 M EDTA: ethylene diamine tetraacetic acid (Sigma, UK)
- Radio-immunoprecipitation assay buffer (RIPA): 50 mM Tris-HCl (pH 7.6), 150 mM NaCl, 1 mM EDTA (pH 8)
- Complete protease inhibitor mixture tablet (Roche, UK)
- PhosSTOP inhibitor tablet (Roche, UK)
- Triton X-100 (Sigma, UK)
- 1X PBS

Efficient solubilisation of most receptor proteins is achieved when cells are lysed in standard RIPA buffer containing both protease and phosphatase inhibitors. Before cell lysis, the standard RIPA buffer were prepared as a stock, which was filter sterilised using 0.2 µm membrane filter and stored at 4°C. After 24 hours transfection, the HEK293 cells were observed under live cell fluorescent microscopy to confirm the presence of AMPAR and TARP proteins. Then, the growth medium was removed and the cells were washed with cold 1X PBS for twice. 100 µl of RIPA lysis buffer containing 0.1% Triton X-100 and pluses both 1X protease and phosphatase inhibitors was added to each well containing transfected cells. The cells were harvested into pre-chilled 1.5 ml eppendorf tubes using cell scrapers (Nunc, UK). The lysates were then left to solubilise at 4°C with continuous agitation for 30 minutes. After solubilisation, cell and nuclear debris were removed by centrifugation at 13,000 rpm for 20 minutes. The clarified supernatant (protein extract) was then transferred to a new 1.5 ml microcentrifuge tube. Protein concentrations were determined using a Bradford reagent assay and Qubit® 2.0 protein assay as described in section 3.4.4.2.

3.4.3.2 Determination of protein concentration

Two ways of protein concentration determination, Bradford assay and Qubit® 2.0 protein assay, were applied in this thesis, and are described in next section.
3.4.3.2.1 Bradford assay

The principle of Bradford reagent assay is based on the formation of a complex between the dye, Brilliant Blue G, and proteins in solution. This protein-dye complex causes a shift in the absorption maximum of the dye from 465 to 595 nm. The amount of absorption is proportional to the protein present. The linear concentration range is 0.1 – 1.4 mg/ml of protein, using BSA (bovin serum albumin) as the standard protein.

Equipment and Solution required:

- Spectrophotometer capable of measuring absorbance in the 595 nm region
- 3 ml Disposable Plastic Cuvettes
- Bradford reagent (Sigma, UK) consists of Brilliant Blue G in phosphoric acid and methanol
- 2 mg/ml Protein standard (BSA) salutation (Sigma, UK)

Standard protein assay protocol:

The Bradford reagent was gently mixed in the bottle and warmed up to the room temperature. The protein standards were created by serially diluting the 2 mg/ml BSA protein (Table 3-18).

<table>
<thead>
<tr>
<th>Tube No.</th>
<th>Sample (µl)</th>
<th>BSA protein standard (mg/ml)</th>
<th>Bradford Reagent (ml)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50</td>
<td>0</td>
<td>1.5</td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>0.2</td>
<td>1.5</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>0.4</td>
<td>1.5</td>
</tr>
<tr>
<td>4</td>
<td>50</td>
<td>0.6</td>
<td>1.5</td>
</tr>
<tr>
<td>5</td>
<td>50</td>
<td>0.8</td>
<td>1.5</td>
</tr>
<tr>
<td>6</td>
<td>50</td>
<td>1</td>
<td>1.5</td>
</tr>
</tbody>
</table>

Table 3-18 An example of the standard protein concentration

Each standard protein samples was diluted with the cell lysis buffer; hence, any possible interference from the buffer components was compensated for in the protein standards. The unknown protein samples were also diluted in 1:10 ratio with the cell lysis. After the dilution
of protein samples, 1.5 ml of Bradford reagent was added to each cuvette containing standard protein samples, and mixed gently by pipetting. The samples were then incubated at room temperature for 25 minutes. The absorbance of each protein samples were measured at 595 nm and recorded. The unknown protein concentration was then determined by comparison of the unknown samples to the standard curve prepared using the protein standards (Table 3-19 and Figure 3-18).

Table 3-19 An example of the absorbance of the BSA protein standard

<table>
<thead>
<tr>
<th>Tube No.</th>
<th>BSA protein (mg/ml)</th>
<th>A595</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0.004</td>
</tr>
<tr>
<td>2</td>
<td>0.2</td>
<td>0.412</td>
</tr>
<tr>
<td>3</td>
<td>0.4</td>
<td>0.675</td>
</tr>
<tr>
<td>4</td>
<td>0.6</td>
<td>0.983</td>
</tr>
<tr>
<td>5</td>
<td>0.8</td>
<td>1.126</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1.325</td>
</tr>
</tbody>
</table>

Figure 3-18 An example of the standard curve produced from the protein assay data

The total concentration of protein present in the original unknown protein solution was calculated as follows:

- \((\text{The absorbance of protein samples} - 0.1074) / 1.2936 = \text{the concentration of protein sample in the cuvette.}\)
• The concentration of protein sample in cuvette x 10 = the actual protein concentration in the sample.

3.4.3.2.2 Qubit® 2.0 protein assay

The Qubit® protein assay kit is designed specifically for use with the Qubit® 2.0 Fluorometer, but it may also be used with any fluorometer or fluorescence plate reader. The principle of the Qubit® 2.0 Fluorometer is using fluorescent dyes to determine the concentration of nucleic acids or proteins in sample, which is different to the traditional UV (Ultraviolet)-absorbance method. Each dye provided from the manufacture is specific for one type of molecule, such as DNA, RNA or protein. The dye has an extremely low fluorescence until it binds to its targets (protein). Upon binding, it becomes intensely fluorescent.

Equipment and Solution required:

• Qubit® 2.0 Fluorometer (Invitrogen, UK)

• Qubit® Protein assay kits (Invitrogen, UK) provided concentrated assay reagent, dilution buffer, and pre-diluted BSA standards.

Qubit® protein assay protocol:

The Qubit® working solution was prepared by diluting the Qubit® protein reagent 1:200 in Qubit® protein buffer, and mixed by vortexing. 190 µl of Qubit® working solution was loaded into each of the thin wall microcentrifuge tubes used for standards. Then, 10 µl of each Qubit® standard was added to the appropriate tube, and mixed by vortexing 2 – 3 seconds. After that, 195 µl of Qubit® working solution was loaded into individual assay tubes and then 5 µl of each sample protein was added to make a final volume 200 µl in each sample tubes, and mixed by vortexing 2 – 3 seconds. All the tubes containing protein samples and standards were incubated at room temperature for 15 minutes. Meanwhile, the Qubit®
2.0 Fluorometer was set to the standards protein assay screen. Before determining the protein samples concentration, a standards calibration was required to set. Therefore, each of the protein standards was read in the programme according to manufacturer’s instructions. After the measurement was completed, a calibration curve was generated by the programme as shown in Figure 3-19.

![Figure 3-19 An example of the Qubit® 2.0 Fluorometer calculates standard curve and sample concentrations](image)

Since the standard calibration curve was set, the sample proteins were ready to test. The concentration of sample proteins was calculated by equation shown below:

\[
\text{Concentration of unknown protein sample} = \text{QF value} \times \frac{200}{X}
\]

Where QF value = the value given by the Qubit® 2.0 Fluorometer

\[
X = \text{the number of } \mu\text{l of sample was added to the assay tube}
\]

After the sample measurement was completed, the actual sample protein concentration was calculated by selected the dilution factor into the programme.

Once the concentration of each sample protein was determined, it was either stored at -80°C for later use or prepared for protein detection by western blot.
3.4.3.3 Preparation of sample protein for loading into SDS-PAGE

Protein detection requires a sensitive antibody to recognise the protein of interest (described in more detail in section 3.5). Generally, antibodies only recognise a small portion of the protein of interest (often referred to as the epitope), and this domain may reside within the 3-dimensional (3D) conformation of the protein. To enable access of the antibody to the specific epitope, it is necessary to unfold the protein by denaturation. For protein denaturation, a loading buffer is required with an anionic denaturing detergent, such as sodium dodecyl sulphate (SDS). The sample is then fully denatured by heat. In the presence of SDS, all of the proteins become negatively charged by their attachment to the SDS anions. The proteins were then denatured by SDS wrapping around its polypeptide backbone leading to a negative charge of polypeptide in proportion to its length.

Solution required:

- Standard Laemmli 2X buffer (pH 6.8): consisting of 4% SDS, 10% β-mercaptoethanol, 20% glycerol, 0.004% bromophenol blue, 0.125 M Tris-HCl

Protein denaturation protocol:

The sample proteins were diluted 1:1 ratio with 2X loading buffer contained the anionic denaturing detergent SDS, the disulphide bridge reducer β-mercaptoethanol, glycerol to increase the density of the sample to be loaded into the gel, and finally a small anionic dye such as bromophenol blue for visualisation (Laemmli 1970). The samples were then mixed by vortexing before boiling at 95°C for 10 minutes. After denaturing, the protein samples were mixed again by vortexing and stored at -20°C for SDS-PAGE protein separation the following day.
3.5 Functional characterisation of CACNG5 variants on surface expression of AMPAR

3.5.1 The use of fluorescent proteins

For over a decade fluorescent proteins (FP) have been widely use as tags to monitor the spatial and temporal localisation of a variety of other proteins. The advantage of introducing a FP tag protein is to visualise the interested protein under any kind of fluorescent technique, such as fluorescence microscopy, confocal, fluorescence cell cytometry. The FP tag protein can also be recognised by antibodies. The principal of using FP tag protein is that the nucleotide coding sequence is generally fused to the cDNA of the protein of interest, usually at its C- or N-terminus (Section 3.3). Then, this fusion protein is expressed in the cell of interest for further detection and investigation.

3.5.2 Fluorescence-activated cell sorting (FACS)

Fluorescence-activated cell sorting known as FACS is a specialised type of flow cytometry. It has been widely used for the analysis of cell identity and fluorescent intensity in large numbers of cells. The principle of FACS is to sort a mixture of biological cells into two or more distinct groups based upon the specific light scattering and fluorescent characteristics of the cells. It is a fast, objective and quantitative of fluorescent signals from individual cells of particular interest.
Equipment and solution required:

- BD FACSAriaII (BD Biosciences, UK)
- FACS buffer: 0.5% FBS, 2 mM EDTA, 0.1% sodium azide (Sigma, UK) in 1X PBS solution

Cell sorting protocol:

Cells were harvested by trypsinisation, washed and resuspended in FACS buffer to avoid the contamination of phenol red for cytometric analysis. Cell sorting for fluorescence protein was performed using a FACSAriaII flow cytometer equipped with a water-cooled argon laser emitting at 488nm. First, a control sample was analysed with a low forward scatter threshold to detect transfected cells while ensuring that debris and electronic noise were not captured as legitimate events. Then, 10,000 events were collected for each sample by list-mode data that consisted of side scatter, forward scatter, and fluorescence scatter. Analysis was performed using Kaluza software; Green fluorescence was (FL1) measured using a 530 + 30 nm band pass filter and red fluorescence (FL3) was determined with a 630 + 22 band pass filter. Gates were set to exclude necrotic cells and cellular debris, and compensation was used to exclude overlap between the two signals, where the fluorescence intensity of events within the gated regions was quantified. Two distinct populations of cells were visible on the flow histograms; therefore, determination of the percentage of transfected cells was based on the inclusion of only cells exhibiting high levels of fluorescence and exclusion of cells adjacent to autofluorescent, non-fluorescent protein expressed cells.
3.5.3 Cell surface biotinylation

Cell surface biotinylation assays are one of the most valuable techniques to identify cell surface proteins. The technique uses a covalent modification of cell surface receptors by the addition of biotin. The biotinylated proteins can then be isolated and identified through a process of affinity purification followed by SDS-PAGE and Western blotting.

Biotin is a water-soluble vitamin that covalently bonds to lysine groups at a low temperature to inhibit protein endocytosis. Current commercial biotin derivatives, which react with primary amines, have been developed to comprise linker regions, in order to minimise steric interference and also allow cleavage of the biotinylated molecule. Derivatives of the egg-white protein avidin that has a high affinity for biotin can be used to efficiently purify biotinylated proteins from cellular lysates. The avidin derivatives may be conjugated to silica beads to facilitated isolation of the biotinylated proteins. The addition of denaturing agents such as SDS enhances the robustness of the binding and permits purification of biotinylated proteins from complex protein solutions. An example of a biotin derivative is sulfo-NHS-biotin (N-hydroxsulfsuc-cinimidobiotin), which is commonly used in neuronal receptor trafficking experiments. The advantage of applying sulfo-NHS-biotin is that contains a negatively charged sulphate group with confers membrane impermeability, and it also has a low cytotoxicity. Biotinylation can also use to study the trafficking of multiple different receptors/membrane proteins within the same experiment. Since the proteins of interest have different molecular weights, the proteins can be separated on the same Western blot membrane and probed with antibodies that recognise their specific proteins of interest.
Material and solution required:

- 1X PBS
- 1X Tris-buffered saline (TBS) (Sigma, UK)
- Sulfo-NHS-Biotinylation kit (Pierce, UK)

Biotinylation protocol:

6-wells plates of culture cells were removed from the incubator and cooled on ice throughout all steps and using solutions at 4°C to prevent receptor endocytosis and membrane trafficking during the biotin labelling step. Culture medium was removed and cells were rinsed twice with 1 ml of ice-cold PBS. Sulfor-NHS-SS-Biotin was dissolved in 48 ml of ice-cold PBS contained 0.5 mM MgCl₂ and 1 mM CaCl₂. 2 ml of the dissolved biotin solution was added to a final concentration of 1 mg/ml and the reaction was carried out for 30 minutes at 4°C with continuous gentle agitation to ensure even coverage of all cells. After biotinylation, it was important to removing all unreacted biotin. This was achieved by adding 300 µl of Quenching solution and the plates were gently tipped back and forth to ensure even coverage. After quenching, cells were scraped into solution and transferred into a 1.5 ml microcentrifuge tube. The plates were rinsed with 1 ml of TBS and this was added to the transferred cells. The transferred cells were spun at 700 rpm for 3 minutes and the supernatant discarded. The remaining cell pellet was rinsed again with 1 ml of TBS, and the cells were twice gently pipetted up and down with a serological pipette, and then spun for 2 minutes at 700 rpm. The resulting supernatant was discarded. At this point, the cells were ready for cell lysis to identify the surface receptor pool. Quantitative western blots were then performed on both total and biotinylated (surface) proteins using antibodies again GFP and RFP (described more details in section 3.5.4).
3.5.4 Western Blotting

Western blotting is a common technique for separating proteins according to their molecular weight, where involves electrophoresis through a polyacrylamide gel in the presence of SDS (SDS polyacrylamide gel electrophoresis, or SDS-PAGE). The principle of western blotting is to detect a specific protein, and analysis of its expression level using specific antibodies in combination with SDS-PAGE. Following SDS-PAGE separation, the proteins are transferred to a membrane by running a perpendicular current through the gel into the membrane, and a specific protein detected using antibodies. These antibodies are either labelled themselves, or commonly detected by a second labelled antibody. This technique allows identification not only of the presence or absence of a protein reacted with the antigen, but also its size and an estimate of relative expression level. Hence, each transfection of wild-type and/or mutated proteins expression levels were estimated for comparison.

3.5.4.1 Polyacrylamide gel

Polyacrylamide gels are composed of long linear polyacrylamide chains cross-linked with N, N-methylenebis-acrylamide (Bis) to create a network of pores interspersed between bundles of polymer. Gel polymerisation requires the addition of ammonium persulphate along with TEMED. The structural features of a gel is neural, hydrophilic, three-dimensional networks of a long hydrocarbons cross-linked by methylene groups, made up of random distributions of solid material and pores. Thus, the movement of proteins within a gel is determined on their size and structure, relative to the pores of the gel. Large proteins migrate slower than small ones, creating distinct separation proteins within the gel.

Furthermore, polyacrylamide gels are characterised by two factors: %T, which is the weight percentage of total monomer including cross-linker and %C is the cross-linker ratio of the monomer solution. The %T indicates the relative pore size of the gel, such as increasing %T decreases the pore size. In general, the smallest pore size is composed with 5%C, any
increase or decrease in %C increases the pore size. The calculation of %T and %C are shown as the following equation.

\[
%T = \frac{\text{Grams acrylamide} + \text{Grams cross-linker}}{\text{Total volume (ml)}} \times 100
\]

\[
%C = \frac{\text{Grams cross-linker}}{\text{Grams acrylamide} + \text{Grams cross-linker}} \times 100
\]

The polyacrylamide %T is not only made as a single percentage throughout the gel, it can also be cast as a gradient of %T throughout the gel. The single percentage gels compositions are commonly from 7.5% up to 20%, or range from 4-15% to 10-20% for gradients gels

3.5.4.2 SDS-PAGE

Assembling glass plate for gel casting:

Firstly the shorter glass plate was placed on top of the longer glass plate with two spacers of equal thickness along the short edges of the rectangular plate. The glass plate sandwich was then placed into the casting frames along the shorter plate facing out, and then the four screws of the casting frames were tightened. The casting frames were placed into the casting stand so that the shorter glass faced out (Figure 3.20). Then, the sandwiches were tested for absence of leakage by pouring some water into the gap.
Figure 3-20 Glass plate assemble for gel casting

Stock gel solution required:

- Acrylamide/bis 40% (Sigma, UK)
- 1.5 M Tris-HCl, pH 8.8
- 0.5 M Tris-HCl, pH 6.8
- 10% SDS
- 5X electrode (running) buffer, pH 8.3 – Tris-glycine buffer

Gel casting:

SDS-PAGE gel in a single electrophoresis run is generally divided into stacking gel (5% acrylamide) and separating gel. The acrylamide percentage in SDS-PAGE separating gel depends on the size of the target protein in the samples. For Mini-PROTEAN® II electrophoresis cell (Bio-Rad, UK), 10 ml of separating solution was prepared for two gels. The composition of separation solution is shown in Table 3.20.
After the preparation of separation solution, it was poured into the sandwiches gap without air bubble formation, and left at least 2 cm for stacking gel solution. Immediately, the monomer solution was then covered with 2-butanol (Sigma, UK), where to prevent the solution drying out by the air. Then the gel was polymerised at the room temperature for 30 minutes. While gel was polymerised, a 5 ml of stacking gel monomer solution was prepared (Table 3.21).

Table 3-21 The composition of 5% stacking solution

<table>
<thead>
<tr>
<th>% Gel</th>
<th>4%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deionised water</td>
<td>3.12 ml</td>
</tr>
<tr>
<td>0.5 M Tris-HCl, pH 6.8</td>
<td>1.25 ml</td>
</tr>
<tr>
<td>10% SDS</td>
<td>100 µl</td>
</tr>
<tr>
<td>Arylamide/bis (40%)</td>
<td>0.5 ml</td>
</tr>
<tr>
<td>10% ammonium persulphate</td>
<td>25 µl</td>
</tr>
<tr>
<td>TEMED</td>
<td>5 µl</td>
</tr>
<tr>
<td>Total</td>
<td>5 ml</td>
</tr>
</tbody>
</table>

After separation solution polymerisation, 2-butanol layer was poured out and rinse off using deionised water. Before pouring the stacking gel, the area above the separation gel was dried using filter paper (Whatman). Once the stacking gel had been poured a comb was placed in
the gel sandwich to make sample loading wells. The stacking gel was then left at the room temperature for 30 minutes to allow polymerisation. After that, the gel sandwiches and gel comb were removed, and the gel was placed into Mini-PROTEAN® II electrophoresis cell with 1X running buffer. The gel was ready for sample loading and running step.

Before loading a protein sample, protein samples were boiled at 95°C for 10 minutes to denature protein structure. 20 µg of each protein sample was then loaded into each well using special gel loading tips. A full-range rainbow molecular weight marker (12 kDa – 225 kDa) (Amersham) was used to determine the size of the protein of interest and also to monitor the progress of electrophoresis. 5 µl of molecular weight marker was loaded in a separate well. The gel was then run at 180 V for 75 minutes, until the dye molecule reached to the bottom of the gel. After electrophoresis was completed, the gel was removed from the casting frames and was ready for western blotting.

3.5.4.3 Protein blotting

Protein blotting is the transfer of proteins to solid-phase membrane supports for visualisation and identification. The initial step is the transfer of proteins from a gel and immobilisation of those proteins on a solid membrane support. Immobilisation of proteins on a membrane makes the proteins accessible to probes for specific antibodies and enables quantitative detection. Here, wet electrophoretic transfer of proteins from a gel to a membrane was performed. The advantage of performing electrophoretic transfer of proteins is fast, efficient, and preserves the high-resolution separation of proteins by PAGE. The method of wet electrophoresis protein transfer is described below.

The principle of electrophoretic transfer:

In an electrophoretic transfer, the membrane and protein-containing gel are placed together with filter paper between two electrodes. Proteins migrate to the membrane following a
current (I) that is generated by applying a voltage (V) across the electrodes, following Ohm’s law:

\[ V = I \times R \]

Where \( R \) is the resistance generated by the materials placed between the electrodes, such as the transfer buffer, gel, membrane, and filter papers.

The electric field strength (V/cm) that is generated between the electrodes is the driving force for electrophoretic transfer. Though a number of other factors, including the size, shape, and charge of the protein and the pH, viscosity, and ionic strength of the transfer buffer and gel \(^{\%T}\) may influence, the elution of particular proteins from gels, both the applied voltage and the distance between the electrodes play a major role in governing the rate of elution of the proteins from the gel. There are practical limits on field strength, however, due to the production of heat during transfer.

The heat generated during a transfer (Joule heating) is proportional to the power consumed by the electrical elements (P), which is equal to the product of the current (I) and voltage (V).

\[ P = I \times V = I^2 \times R \]

Joule heating increases temperature and decreases resistance of the transfer buffer. Such changes in resistance may lead to inconsistent field strength and transfer, or may cause the transfer buffer to lose its buffering capacity. In addition, excessive heat may cause the gel to deteriorate and stick to the membrane. The major limitation of any electrophoretic transfer method is the ability of the chamber to dissipate heat.
Nitrocellulose membrane:

Nitrocellulose was one of the first membranes used for western blotting and is still a popular membrane for this procedure. Protein binding to nitrocellulose is instantaneous, nearly irreversible, and quantitative up to 80 to 100 ug/cm². Nitrocellulose is easily wetted in water or transfer buffer and is compatible with a wide range of protein detection systems. Membranes are commonly available in two pore sizes: the 0.45 um pore size membranes are recommended for most analytical blotting experiments, while the 0.2 um pore size membranes are most suitable for transfer of low molecular weight (< 15 kDa) proteins that might move through larger membrane pores.

Assembling the gel and membrane sandwich:

Each gel sandwich contained the gel and membrane sandwiched between pieces of blot absorbent filter paper. For each gel, one piece of membrane and two pieces of filter paper were cut to the dimensions of the gel. The gels were briefly rinse in deionised water and equilibrated in transfer buffer to remove contaminating electrophoresis buffer salt. The presence of salt will increase the conductivity of the transfer buffer and the amount of heat generated during the transfer. Also, the gel will shrink or swell to various degrees in the transfer buffer depending on the acrylamide percentage and the buffer composition. For example, low percentage gels (< 12%) will shrink in methanol buffers. Equilibration allows the gel to adjust to its final size prior to electrophoretic transfer. The membranes, filter papers, and fibre pads were also need to soak in the transfer buffer for 5 minutes before assembly in the transfer apparatus. A gel sandwich was prepared by placing one pre-wetted fibre pad on top of the black side of the cassette, then placed the pre-wetted filter paper on the top of fibre pad, followed the equilibrated gel and then placed the pre-wetted membrane on the top of gel. The gel sandwich was completed by adding a piece of filter paper on the membrane and
gently used a glass tube to roll air bubbles out (Figure 3-21). Finally, the last fibre pad was added and the cassette closed firmly.

Figure 3-21 Transfer assembly for a tank transfer system (Trans-Blot cell)
Transfer buffer

The size of SFP-AMPAR2 is 120 kDa as large protein; it tends to precipitate in the gel, hindering transfer. Thus, the suitable transfer buffer was Tris-glycine transfer buffer consisting of 48 mM base (pH 9.2), 39 mM glycine, 10% methanol and 0.1% SDS. Adding additional SDS to a final concentration of 0.1% in the transfer buffer leads to increased protein transfer from the gel, but can decrease binding of the protein to the nitrocellulose membrane. While the additional methanol tends to remove SDS from protein, so reducing the methanol percentage to 10% guard against precipitation. Lowering methanol in the transfer buffer also promotes swelling of the gel, allowing large proteins to transfer more easily. Hence, the balance of SDS and methanol in the transfer buffer, protein size, and gel percentage can affect transfer efficiency.

Electrophoresis of transfer

Once the cassette was closed and locked, inserted it into the tank with the latch side up, the black cassette plate needed to faces the black electrode plate. The transfer buffer was added to the tank until the buffer level reaches the fill line. To maintain even buffer temperature and ion distribution in the tank, magnetic stirrer bar were added to the tank and the entire tank assemble was placed on the top of a magnetic stirrer plate. The addition of an ice pack was required for the high-intensity field transfers; due to heat generation. Finally, the lid was placed on top of the cells, making sure that the colour-coded cables on the lid were attached to the electrode cards of the same colour. The cables were plugged into the power supply, and the transfer was performed at 60 volt for 90 minutes.

3.5.4.4 Coomassie gel staining

Coomassie gel staining is useful to determine if proteins have migrated uniformly and evenly. And also it can use on gels post-transfer to check the efficiency of the transfer.
**Gel staining protocol:**

After SDS-PAGE is finished, the gel were treated with a 40% distilled water, 10% acetic acid, and 50% methanol solution to allow all proteins to precipitate and prevent diffusion. In order to visualise the fixed proteins placed the gel in the same mixture of water/acetic acid/methanol with the addition of 0.25% by weight Coomassie Brilliant Blue R-250 (National Diagnostics, USA). The incubation of coomassie blue was performed for 4 hours at room temperature on a shaker. After staining, the gel had to be destained with a buffer of 67.5% distilled water, 7.5% acetic acid, and 25% methanol on a shaker. And then fresh destain buffer was required until the excess dye had been removed completely. The stain does not bind to the acrylamide, and only proteins in the gel remain stained.

**3.5.4.5 Ponceau S staining**

Ponceau S membrane staining is a method to visualise the proteins in a membrane and to check for successful of transfer.

**Staining protocol:**

After protein transfer, the membrane was washed in TBS buffer plus 0.1% Tween20 (TBST), which consist of Tris-HCl, NaCl, and Tween20 with pH 7.6. Then, the membrane was transferred into 1X Ponceau Red on a shaker for 5 minute. The 10X Ponceau Red was made of 2% Ponceau S in 30% trichloroacetic acid and 30% sulfosalicylic acid. After the staining, the membrane was then washed extensively in water until the protein bands were well-defined.
3.5.4.6 Membrane blocking

Membrane blocking is the method to prevent non-specific background binding of the primary and secondary antibodies to the membrane. Non-fat milk blocking reagent is a traditional blocking buffer, which normally consisting of 5% milk in TBST.

Membrane blocking protocol:

Upon completion of the transfer, the blotting sandwiches were dissembled and the membrane was ready for blocking. Initially, the membranes were wetted in PBS for several minutes, and then blocked with 5% dried skimmed milk blocking buffer: 2.5g of dry fat milk powder in 50 ml of TBST. While the membrane was blocked with 5% milk blocking buffer for an hour at room temperature, the primary antibodies were diluted in 1% milk buffer. The primary antibodies were incubated with the membrane at 4°C overnight, with gentle shaking.

On the next day, the pre-probed primary antibody membranes were washed three times for 10 minutes each at room temperature in 1x TBST with gentle shaking, using a generous amount of buffer. During the washing step, the fluorescently-labelled secondary antibodies were diluted in 1% milk buffer in 1x TBST. Prolonged exposure of the antibodies vial to light must be avoided. The membranes then were probed for 45 minutes at room temperature with gentle shaking and protected from light. Washes were repeated after secondary labelling, washing three times for 10 minutes in 1x TBST, then placed in normal PBS.

3.5.5 Antibodies

Antibody probing is the most important step in protein detection. The sensitivity of the antibodies determines the accuracy of the result. The selection of primary antibody depends on the antigen to be detected and their sensitivity to the target antigen. Here, three primary antibodies were used to target interested protein; anti-tRFP, anti-tGFP, and β-actin, described in more detail below. Primary antibodies are generally not directly detectable in western
blotting. Therefore, secondary antibody is usually required to detect the target antigen. The choice of secondary antibody depends upon the species of animal in which the primary antibody was raised and the detection method being used. Here, two secondary antibodies were used; goat anti-mouse antibody for targeting β-actin and goat anti-rabbit antibody for targeting GFP and RFP. Both secondary antibodies were made with fluorescence based for fluorescent detection.

3.5.5.1 Primary antibodies

- Anti-GFP (Santa Cruz, UK)

Anti-GFP was purchased from Santa Cruz, is a rabbit polyclonal antibody raised against amino acids 1 – 238 representing full length GFP of *Aequorea victoria* origin. The full length antibody has ability to also recognise GFP mutant fusion protein such as SEP. For detection of SEP-AMPAR protein expression, the antibody was used at 1:500 dilutions.

- Anti-tRFP (Evrogen, UK)

Full length anti-tRFP, a rabbit polyclonal antibody and specifically to recognise both denatured and native TagRFP was obtained from Evrogen. For detection of RFP-CACNG5 protein expression, the antibody was used at 1:1000 dilutions.

- Anti-β-actin (LI-COR Biosciences, UK)

Anti-β-actin was provided from LI-COR Biosciences, is a mouse monoclonal antibody to detect endogenous levels of β-actin protein. For measurement of protein expression, a positive control protein was required to check each lane was evenly loaded with samples, such as β-actin. It was also useful to check for even transfer from the gel to the membrane across the whole gel. Although the even loading or transfer had not occurred, β-actin bands
could also be used to quantify the protein amounts in each lane. For endogenous levels of β-actin detection, the antibody was used at 1:1500 dilution.

### 3.5.5.2 Secondary antibodies labelled with near-infrared fluorescence dyes

Traditional western blot detection uses a primary antibody directed again a protein of interest and a secondary antibody conjugated with an enzyme reporter, such as horseradish peroxidase or alkaline phosphatase. Then, chemiluminescent or colorimetric detection of the enzyme conjugate confirms the presence of the protein of interest. Fluorescent western blot detection allows quantification of the target. For example, secondary antibodies labelled with a near-infrared (NIR) fluorescent dye such as IRDye 800 CW and IRDye 680 RD provided from LI-COR Biosciences, allow direct non-enzymatic detection. The advantage of using NIR is that cells, animal tissue, plasticware, blotting membranes, and chemical compound libraries all possess intrinsic autofluorescence that can interfere with detection; whereas in the NIR spectral region 700 – 900 nm, autofluorescent background is dramatically reduced. Since the protein of interest is recognised by its primary antibody and is conjugated by the secondary antibody labelled NIR, blot is then documented with an NIR imager such as the Odyssey® Imager from LI-COR Biosciences. With this approach, fluorescent signal is directly proportional to the amount of protein of interest present.

- **Goat anti-rabbit IR-Dye 680RD (LI-COR Biosciences, UK)**

The IR-Dye 680 RD has a maximum excitation wavelength at 690 nm and emission wavelength at 694 nm, where the detection channel wavelength in Odyssey® is 700 nm. This wavelength can be detected in a combination with IR-Dye 800CW secondary antibody, known as two-colour detection. Goat anti-rabbit IR-Dye 680RD is a goat antibody to detect a rabbit primary antibody that is raised against to SEP-GluR2 and RFP-CACNG5 in our
experiment. For Odyssey® Imager detection, the antibody was optimised at 1:10,000 dilutions.

- Goat anti-mouse IR-Dye 800CW (LI-COR Biosciences, UK)

The IR-Dye 800CW has a maximum excitation wavelength at 778 nm and emission wavelength at 794 nm, where the detection channel wavelength in Odyssey® is 800 nm. Goat anti-mouse IR-Dye 800CW is a goat antibody to detect a mouse primary antibody that is raised against to endogenous levels of β-actin protein in our experiment. The optimised dilution of this antibody was used at 1:10,000 dilutions.

### 3.5.6 Odyssey® Imager and analytical software

All blots were imaged using a LICOR Odyssey® scanner and imager software to analysis and quantify the interest of protein in the cells. The IR imager in both 700 nm and 800 nm channels was set at 169 μm resolutions for all blots. Quantification of the interest of proteins was performed using Odyssey 3.0 analytical software (LI-COR Biosciences). Boxes were manually placed around each band of interest and the NIR fluorescent values of raw intensity with intra-line background subtracted was obtained. The given NIR fluorescent values were then analysed using R.

### 3.5.7 Data analysis

**3.5.7.1 Quantification analysis of cells expressing AMPAR proteins**

The data was collected from FACS where the quantification of cells expression AMPAR proteins was measured. Then, the data was further analysed using Kaluza software and presented as the mean ± SEM. The statistical analysis was performed using a two-tailed unpaired Student’s t-test in Excel. The differences were considered significant with a P-value < 0.05.

198
3.5.7.2 Quantification analysis of surface receptor proteins

The objective of the statistical analyses was to identify protein expression changes on AMPA receptor and CACNG5 between the total and surface expression. Then further analysis compared both expression between wild-type and variants. The protein expression values were given from the NIR fluorescence value for each protein of interest. Linear mixed models were used to analyse the data with the biological replicates as random effects. Statistical analysis was performed using a linear mixed model in R (R Core Team 2013). Differences of $P < 0.05$ were considered significant.
4 Identification of CACNG5 variants in cases of bipolar disorder and schizophrenia

4.1 Introduction

Bipolar disorder (BPD) and schizophrenia (SCZ) are common, severe and highly heritable psychiatric disorders. For instance, the heritability of BPD is estimated to be approximately 70 – 90%, whereas the heritability for SCZ is 80 – 85% (Cannon et al. 1998, Cardno and Gottesman 2000, Kieseppa et al. 2004). This high degree of heritability indicates that genetics has an important role in pathogenesis. Although these disorders are classified into two distinct categories, both clinical and genetic studies have shown overlapping aetiology (Berrettini 2003, Hamshere et al. 2005, Moskvina et al. 2009, Shao and Vawter 2008). Perhaps the best example of a susceptibility gene for mental illness is the CACNA1C gene which encodes the α1 subunit of the L-type voltage-gated calcium channel (VGCC). CACNA1C has been implicated in genome-wide association studies (GWASs) of BPD, SCZ, and major depressive disorder (Ferreira et al. 2008, Green et al. 2010, Psychiatric GWAS Consortium Bipolar Disorder Working Group 2011, Schizophrenia Psychiatric Genome Wide Association Study (GWAS) Consortium 2011). Recently, the Psychiatric Genomic Consortium extended previous association findings at CACNA1C, CACNB2 and CACNA1I, which encode VGCC subunits, which implicate the family members of VGCC in SCZ and other psychiatric disorders (Schizophrenia working Groups of the Psychiatric Genomic Consortium 2014). As described earlier in the GWAS section of the introduction, our research group previously found association with a marker rs17645023 (p = 6.3 x 10^-7) that lies 36kb from CACNG5 and 79kb from CACNG4 (both genes encode γ subunits of calcium channel genes) from a BPD versus SCZ case-case comparison study (Curtis et al. 2011) (Figure 1-1). The allele frequencies of this variant were 18% in SCZ, 28% in BPD, and 23%
in control, where showed that the control allele frequencies were intermediate between those of BPD and SCZ. This result implied that marker rs17645023 has significant differential effects in the two disorders and that this finding may be used to demonstrate how these two disorders differ biologically. Although this study finding brought out new insights into these two disorders, the linkage disequilibrium (LD) block analysis revealed weak linkage disequilibrium between the marker and these two genes (Figure 4-1). Both LD analysis of UCL BPD and controls and 1000 Genomes Project data revealed an LD signal shift towards CACNG4 gene region. Furthermore, the regional association plot of the Psychiatric Genomics Consortium (PGC) BPD GWAS found no associated variants in both CACNG5 and CACNG4 genes (Figure 4-2). However, the regional association of PGC SCZ shows a nominally associated variant (rs185202153, p = 4.5 x 10^{-4}) in the intronic region of CACNG4 (Figure 4-3). Even though both PGC GWAS data and LD analysis did not strengthen the finding of disease susceptibility variants in both genes, an interesting 17q24.2 microdeletion study suggested that deletion of CACNG5 and CACNG4 regions might be responsible for speech delay and intellectual disability observed in four patients and mood swings, hallucinations and seizures in two patients (Vergult et al. 2012). These findings suggest that GWAS may not detect all genetic susceptibility genes or variants and therefore if studies only rely on GWAS findings or LD score analyses they may fail to detect hidden casual variants or genes. The finding of a nominal association with the CACNG4 SNP rs185202153 in the PGC SCZ data, may suggest that CACNG4 is involved more in SCZ than in BPD. However, none of PGC GWASs had found any association variants in CACNG5. This begs an interesting question as to whether CACNG5 is a gene with overlapping or distinct susceptibilities for these two conditions. Since the genetic evidence implicates the importance of calcium channel signalling in the path-physiology of psychiatric disorders. Therefore, the genetic
sequence and genotyping of *CACNG5* may illustrate the distinct biological patterns or mechanisms that underlie SCZ and/or BPD.

**Figure 4-1** LD plot of the marker rs17645023 between *CACNG5* and *CACNG4* genes obtained from the USCS Genome Brower website. The analysis was applied by using the UCL and 1000 Genome Project data by PLINK.
Figure 4-2 regional association of PGC BPD GWAS plot for the marker rs17645023 analysed at CACNG5 and CACNG4.
Here, we explored the hypothesis that genetic causes of BPD and SCZ may be found by identifying genetic variants that overlap in both disorders, as well as by variants that distinguish between these disorders. Previous observations have indicated that CACNG5 may be a candidate gene for these disorders with an involvement of calcium channel signalling in both BPD and SCZ (Curtis et al. 2011). In fact, according to literature reviews CACNG5 encodes TARPs (transmembrane AMPA receptor regulatory proteins), a protein intimately involved in the trafficking of glutamatergic AMPA (α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid) receptors (Burgess et al. 1999, Burgess et al. 2001). Moreover, the γ5 subunit (CACNG5) was also shown to contribute to the regulation of calcium permeability via AMPA receptor assembly (Soto et al. 2009). The regulation of intracellular calcium ions and AMPA receptor trafficking and gating are important mechanisms that mediate the...
excitatory synaptic transmission at the postsynaptic density. Thus dysfunctional control mechanisms may lead to disorders, such as BPD and SCZ (Du et al. 2004a, Du et al. 2004b). Despite these results, the role of CACNG5 polymorphisms in these disorders has still not been completely investigated. This is likely to reflect the limitation of GWAs studies, which are only designed to identify common variants, possibly of small effect, and may therefore fail to detect associations with rare susceptibility variants (Maher 2008). We therefore focused our attention on identifying CACNG5 polymorphisms in subjects with either BPD or SCZ.

To assess the presence of functional variants within CACNG5, we initially screened 1,099 patients with BPD, 618 patients with SCZ and 986 control samples (this included 614 supernormal controls). The screened regions included the entire coding regions of the gene (exon 1 to exon 5) plus their neighbouring flanking introns and promoter regions. We chose to use high resolution melting analysis (HRMA) to perform rapid variant scanning of the CACNG5 in DNA samples from both patients and controls. The HRMA variant scanning principle is based on the dissociation behaviour of DNA under increasing temperature. Melting analysis is generated by using a saturating double-stranded DNA-binding dye with increasing temperature to denature double-stranded into single-stranded DNA with the loss of its fluorescence (Reed and Wittwer 2004). The melting profile gives a specific sequence-related pattern allowing differentiation between wild-type sequences and homozygote heterozygote variants (Graham et al. 2005). Because HRMA is a simple PCR-based method, with high accuracy, low cost, and yields rapid results, it is an attractive choice for the detection of disease-associated variants. Although the accuracy of HRM is very high, several factors have to be taken into consideration when applying this technique into clinical practice, such as PCR specificity, length of the amplicon, percentage of GC content, dye, instrument and melting analysis software. For example, as melting analysis is performed directly after
PCR, different heterozygotes variants may reveal similar melting curves that cannot be distinguished from each other, even though they are clearly distinct from homozygous wild-type variants. Therefore careful design of primers and optimised temperature cycling is required to improve the specificity of amplification of the target regions (Er et al. 2012). Secondly, studies have suggested that shortening the length of the amplicons can also improve melting curve differentiation between mutant and wild-type alleles (Gundry et al. 2003, Montgomery et al. 2010). As shown in Reed GH and Wittwer CT (2004) study, (Reed and Wittwer 2004), when PCR products below or equal to 300 bp, both sensitivity and specificity were 100%; when PCR products were between 300 bp and 1000 bp, sensitivity and specificity decreased to 96.1% and 99.4% respectively. Third, studies have also suggested that there was better sensitivity with GC mismatches, i.e. there were fewer missed variants (false negative) in these circumstances compared to AA and TT mismatches (Reed and Wittwer 2004, Krypuy et al. 2006). Finally, the different dye used in HRM, work with different efficiencies and also the use of different instruments has an impact on the sensitivity and specificity of HRM. In order to overcome these factors, any sample that was detected with a suspected variant was directly sequenced to confirm the results from HRMA. The frequency of any change then was assessed in healthy subjects compared to our patient samples. Genotyping was validated by the Kaspar genotyping method. There was 100% concordance between the HRMA and Kasp genotyping data for the variants described.

4.2 Results

4.2.1 CACNG5 gene variants by high resolution melting analysis (HRMA)

We detected a total of 12 single nucleotide exchange variants within exonic regions and 3 extra variants were in promoter regions. These results are summarised in the Table 4.1, the rare SNPs are named by chromosomal location. For validation of these variants, we genotyped all cases and controls samples by the Kaspar as described in the Methods section
and the analysed results of each of the variants can be found in Appendix II. The validated variants include 8 predicted amino acid substitution variants (non-synonymous SNPs, nsSNPs), 4 synonymous SNPs (sSNPs) and 3 single nucleotide exchange polymorphisms in promoter regions.

Table 4-1 Summary of variant screening result

<table>
<thead>
<tr>
<th>Name</th>
<th>Nucleotide Change</th>
<th>Predicted Amino Acid Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs3760263</td>
<td>G&gt;A</td>
<td>N/A</td>
</tr>
<tr>
<td>rs181400884</td>
<td>G&gt;T</td>
<td>N/A</td>
</tr>
<tr>
<td>rs75486725</td>
<td>C&gt;T</td>
<td>N/A</td>
</tr>
<tr>
<td>17_64873468</td>
<td>c.77 G&gt;A</td>
<td>Arg6Arg</td>
</tr>
<tr>
<td>rs146874664</td>
<td>c.117 T&gt;C</td>
<td>Cys19Cys</td>
</tr>
<tr>
<td>rs71379998</td>
<td>c.141 G&gt;A</td>
<td>Ala27Ala</td>
</tr>
<tr>
<td>rs11652480</td>
<td>c.162 G&gt;A</td>
<td>Leu34Leu</td>
</tr>
<tr>
<td>17_64875098</td>
<td>c.265 C&gt;T</td>
<td>Arg69Trp</td>
</tr>
<tr>
<td>17_64875105</td>
<td>c.272 G&gt;A</td>
<td>Arg71His</td>
</tr>
<tr>
<td>rs142916987</td>
<td>c.440 G&gt;A</td>
<td>Arg127Gln</td>
</tr>
<tr>
<td>rs149159754</td>
<td>c.443 C&gt;T</td>
<td>Thr128Met</td>
</tr>
<tr>
<td>17_64880644</td>
<td>c.496 A&gt;G</td>
<td>Val146Met</td>
</tr>
<tr>
<td>17_64880674</td>
<td>c.526 A&gt;T</td>
<td>Ile156Phe</td>
</tr>
<tr>
<td>17_64880699</td>
<td>c.551 C&gt;T</td>
<td>Thr164Leu</td>
</tr>
<tr>
<td>rs41280112</td>
<td>c.757 C&gt;T</td>
<td>His233Tyr</td>
</tr>
</tbody>
</table>

4.2.2 Non-synonymous variants

HRMA clearly showed an aberrant curve with a base change from cytosine to thymine at cDNA position 265 (c.265 C>T) within exon 2 (Figure 4-4 and Figure 4-5 panel A). Direct sequencing indicated that the underlying variation corresponded to a novel variant that was predicted to cause a non-synonymous p.Arg69Trp (R69W) amino acid change. Further genotyping confirmed this change as a rare polymorphism that occurred in only one SCZ sample and not in any of the BPD or control samples (Figure 4-4 and Figure 4-5 panel A). A
second novel nsSNP with a predicted amino acid change of p.Arg71His (R71H) with a guanine to adenine base change at position 272 (c.272 G>A) was only found in one BPD sample and not in the SCZ or control samples (Figure 4-4 and Figure 4-5 panel B). Both of these nsSNPs were absent from the European samples in the 1K Genome project. However, R69W was found in one individual from Africa (NA19175) in the 1K Genome project.

Figure 4-4 Detection of SNPs in exon 2 of the CACNG5 gene by HRMA. p.Arg69Trp is shown in panel A and p.Arg71His is shown in panel B. The upper panels show the detection of the novel changes obtained by comparing melting-curve shapes after signal normalisation. Samples that contain heterozygote polymorphisms (red or green) form heteroduplexes that are denatured at lower temperatures compared to the wild-type samples (blue). The lower panels indicate differences in fluorescence plotted against temperature for the sample curves. This leads to clear separation between samples with polymorphisms (red or green) and wild-type samples (blue).
Figure 4-5 Validation of SNPs found in exon 2 of the CACNG5 gene by HRM analysis. The upper panels of A and B show clear genotype groups; wild-type homozygous individuals are shown in blue, heterozygous individuals are shown in red, and negative samples are shown in grey. The lower panel of A and B shows DNA sequence traces for the two samples (the heterozygous polymorphic bases are shown with an arrow.

HRMA also detected another two nsSNPs within exon 3. A base change from guanine to adenine at position 440 (c.440 G>A) was predicted to result in Arg127Gln (R127Q). Genotyping results revealed that this SNP occurred in two BPD samples and did not occur in the control or SCZ samples. However, this SNP was found in one of the 379 European samples from the 1K Genomes project (Table 4-2). Another exon 3 nsSNP was detected which was a cytosine to thymine base change at position 443 (c.443 C>T). This SNP was predicted to result in p.Thr128Met (T128M) and was found in three BPD samples plus two control samples and not in the SCZ samples. A different single nucleotide change that was predicted to result in a different amino acid substitution, p.Thr128Lys (T128K) was found in only one Asian sample (NA18610) and not in any European sample from the 1K Genome project.
A Further three nsSNPs were detected by HRMA; the predicted amino acid changes that these SNPs caused were p.Val146Met (V146M), p.Ile156Phe (I156F) and p.Thr164Leu (T164L). According to our genotyping data, only one BPD patient was identified who carried the I156F SNP, whereas two SCZ patients with V146M were identified and only one SCZ patient was detected with T164L. None of these SNPs were reported in the 1K Genomes project database. A more common nsSNP, rs41280112 was also detected in the UCL samples. This SNP was predicted to lead to the His233Tyr (H233Y) amino acid substitution. Interestingly, the allele frequencies (AF) for this SNP is higher in both BPD and SCZ samples compared to UCL controls and in the European samples from the 1K Genomes project; 2.89% in BPD, 3.48% in SCZ, 1.70% in our control, and 1.85% in European samples from the 1K Genomes project.

Furthermore, a nsSNP with the predicted amino acid substitution of Ala182Ser (A182S) was observed in a single sample of European origin in the 1K Genomes project. Additional genotyping of this variant in the UCL BPD, SCZ and control samples was performed in order to identify whether it was a false negative finding from our HRMA. Our results showed that A182S did not occur in the UCL samples and also suggests that our HRMA had high sensitivity and specificity.
Table 4-2 Summary of non-synonymous variants genotype counts together with data from the European samples in the 1K Genomes project. The sum of total wild-type homozygous individuals genotyped in each SNP is showed in 0|0 column; the sum of total heterozygous individuals genotyped in each SNP is showed in 1|0 column. Percentage of allele frequencies (AF %) was calculated for each variant. The numbers in total row represents an aggregation of all non-synonymous variants found in UCL cases.

<table>
<thead>
<tr>
<th></th>
<th>1K EURO n = 379</th>
<th>BP n = 1073</th>
<th>SCZ n = 603</th>
<th>Control n = 941</th>
<th>Aberdeen Control n = 518</th>
<th>Aberdeen SCZ n = 643</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>00</td>
<td>10</td>
<td>AF %</td>
<td>00</td>
<td>10</td>
<td>AF %</td>
</tr>
<tr>
<td>R69W</td>
<td>379</td>
<td>0</td>
<td>0</td>
<td>1071</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>R71H</td>
<td>379</td>
<td>0</td>
<td>0</td>
<td>1067</td>
<td>1</td>
<td>0.05</td>
</tr>
<tr>
<td>R127Q</td>
<td>378</td>
<td>1</td>
<td>0.13</td>
<td>1071</td>
<td>2</td>
<td>0.09</td>
</tr>
<tr>
<td>T128M</td>
<td>379</td>
<td>0</td>
<td>0</td>
<td>1069</td>
<td>3</td>
<td>0.14</td>
</tr>
<tr>
<td>V146M</td>
<td>379</td>
<td>0</td>
<td>0</td>
<td>1070</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>H156F</td>
<td>379</td>
<td>0</td>
<td>0</td>
<td>1072</td>
<td>1</td>
<td>0.05</td>
</tr>
<tr>
<td>T164L</td>
<td>379</td>
<td>0</td>
<td>0</td>
<td>1071</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>H233Y</td>
<td>367</td>
<td>12</td>
<td>1.58</td>
<td>1019</td>
<td>55</td>
<td>2.56</td>
</tr>
<tr>
<td>Total</td>
<td>365</td>
<td>14</td>
<td>1.85</td>
<td>1011</td>
<td>62</td>
<td>2.89</td>
</tr>
</tbody>
</table>

Aberdeen SCZ n = 643
Control n = 941
1K EURO n = 379
BP n = 1073
SCZ n = 603
Table 4-3 Summary of synonymous variant genotype counts together with data from the European samples in the 1K Genomes project. The sum of total wild-type homozygous individuals genotyped in each SNP is showed in 0|0 column; the sum of total heterozygous individuals genotyped in each SNP is showed in 1|0 column; the sum of total mutated homozygous individuals genotyped in each SNP is showed in 1|1 column. Percentage of allele frequencies (AF %) was calculated for each variant. The numbers in total row represents an aggregation of all synonymous variants found in UCL cases.

<table>
<thead>
<tr>
<th></th>
<th>1K EURO n = 379</th>
<th>BP n = 1071</th>
<th>SCZ n = 600</th>
<th>Control n = 942</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>R6R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C19C</td>
<td>374</td>
<td>5</td>
<td>0</td>
<td>0.66</td>
</tr>
<tr>
<td>A27A</td>
<td>367</td>
<td>12</td>
<td>0</td>
<td>1.58</td>
</tr>
<tr>
<td>L34L</td>
<td>314</td>
<td>58</td>
<td>7</td>
<td>9.5</td>
</tr>
<tr>
<td>S87S</td>
<td>378</td>
<td>1</td>
<td>0</td>
<td>0.13</td>
</tr>
<tr>
<td>Total</td>
<td>296</td>
<td>76</td>
<td>7</td>
<td>11.87</td>
</tr>
</tbody>
</table>
Table 4-4 Summary of putative promoter variant genotype counts together with data from the European samples in the 1K Genomes project. The sum of total wild-type homozygous individuals genotyped in each SNP is showed in 0|0 column; the sum of total heterozygous individuals genotyped in each SNP is showed in 1|0 column; the sum of total mutated homozygous individuals genotyped in each SNP is showed in 1|1 column. Percentage of allele frequencies (AF %) was calculated for each variant. The numbers in total row represents an aggregation of all putative promoter variants found in UCL cases.

<table>
<thead>
<tr>
<th></th>
<th>1K EURO n = 379</th>
<th>BP n = 1057</th>
<th>SCZ n = 598</th>
<th>Control n = 937</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>rs3760263</td>
<td>345</td>
<td>32</td>
<td>2</td>
<td>4.75</td>
</tr>
<tr>
<td>rs181400884</td>
<td>377</td>
<td>2</td>
<td>0</td>
<td>0.26</td>
</tr>
<tr>
<td>rs75486725</td>
<td>379</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>343</td>
<td>34</td>
<td>2</td>
<td>5.01</td>
</tr>
</tbody>
</table>
4.2.3 Synonymous variants

Exon 1 was difficult to screen for novel polymorphisms using only a single set of primer pairs because it contains the two common database (DB) SNPs, rs71379998 and rs11652480. This is because HRMA detects all polymorphisms in the region that is amplified and this makes it difficult to distinguish novel changes amongst a background of common variants. In order to resolve this problem several samples from each melting curve cluster were sequenced directly from the HRM product. A novel guanine to adenine sSNP at cDNA sequence position 77 from the start codon (c.77 G>A; Arg6Arg (R6R)) was only found in one SCZ sample and not in control or BPD samples. A 1K Genomes project thymine to cytosine sSNP at position 117 (c.117 T>C; Cys19Cys (C19C)) was found in the UCL sample with minor AF of 0.42% in BPD, 0.75% in SCZ, and 0.27% in the control sample (Table 4-3). The common SNPs rs71379998 (Ala27Ala (A27A)) and rs11652480 (Leu34Leu (L34L)) have similar AFs in the BPD, SCZ and control samples to the European samples in the 1K Genomes project.

4.2.4 Putative CACNG5 variant promoter variants

The putative promoter region of CACNG5 presented 3 aberrant curves in both BPD and SCZ samples. SNP rs3760263 was found in the UCL samples and also in the European samples from the 1K Genomes project. The AF the UCL SCZ sample was 5.10%, 4.59% in the UCL BPD sample, 3.79% in the UCL control sample and 4.75% in the European samples from the 1K Genomes project (Table 4-4). Furthermore, no homozygous individuals were identified in the UCL control sample, while two homozygous individuals were detected in the European samples from the 1K Genomes project. Interestingly, SNP rs181400884 was found in UCL BPD and control samples and also in two European samples from the 1K Genomes project, but was not found in the UCL SCZ samples. Finally SNP rs75486725 was not detected in the European samples from the 1K Genomes project but this SNP was detected in the UCL
samples albeit at a low AF; 0.17% in the SCZ samples, 0.38% in the BPD samples and 0.43% in the control samples.

### 4.2.5 UK10K EXOME DATA

Table 4-5 Summary of non-synonymous variants genotype counts from the UK10K exome sequence data. Percentage of allele frequencies (AF %) was calculated for each variant. The sum of total wild-type homozygous individuals genotyped in each SNP is showed in 0|0 column; the sum of total heterozygous individuals genotyped in each SNP is showed in 1|0 column; the sum of total mutated homozygous individuals genotyped in each SNP is showed in 1|1 column. The numbers in total row represents an aggregation of all non-synonymous variants found in UCL cases.

<table>
<thead>
<tr>
<th>Variant</th>
<th>UK10K Control n = 982</th>
<th>UK10K SCZ n = 1392</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>0 1</td>
</tr>
<tr>
<td>R69W</td>
<td>982 0 0.00</td>
<td>1391 1 0 0.04</td>
</tr>
<tr>
<td>R71H</td>
<td>982 0 0.00</td>
<td>1390 2 0 0.07</td>
</tr>
<tr>
<td>R127Q</td>
<td>982 0 0.00</td>
<td>1391 1 0 0.04</td>
</tr>
<tr>
<td>T128M</td>
<td>980 2 0.10</td>
<td>1392 0 0 0.00</td>
</tr>
<tr>
<td>V146M</td>
<td>982 0 0.00</td>
<td>1392 0 0 0.00</td>
</tr>
<tr>
<td>I156F</td>
<td>981 1 0.05</td>
<td>1391 1 0 0.04</td>
</tr>
<tr>
<td>T164L</td>
<td>982 0 0.00</td>
<td>1392 0 0 0.00</td>
</tr>
<tr>
<td>H233Y</td>
<td>952 30 1.53</td>
<td>1346 44 1 1.65</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>949 33 1.68</td>
<td>1342 49 1 1.83</td>
</tr>
</tbody>
</table>

The UK10K study examined 982 UK sample with chronic obesity which we used on controls for our study and 1,392 SCZ patients. Five out of our total eight rare nsSNPs were found in UK10K exome data. These were R69W, R71H, R127Q, T128M, and I156F. R69W and R71H were found one and two SCZ cases, respectively and were absent in the controls. Conversely, T128M was found one control subject but was absent in the cases. Interestingly, variant I156F was found once in control and case, where was not found in both 1K Genomes project data and UCL control subjects. Thus, the minor AF of this variant was almost equal in both control and case samples of UK10K exome sequence data. The common variant H233Y was also found in both control and case subject with increased minor AF in cases, but this was dramatically decreased compared with UCL BPD and SCZ subjects (from AF = 2.56%
and 3.15%, respectively, to 1.65% in UK10K cases). The additional UK10K exome sequence data did not strengthen the findings for the frequencies of all nsSNP variants combined in case versus control subject comparison. Possible reasons could be the two rare variants were absent in the UK10K data and also the common variant appears to be occurring more frequently.

Table 4-6 Summary of synonymous variants genotype counts from the UK10K exome sequence data. Percentage of allele frequencies (AF %) was calculated for each variant. The sum of total wild-type homozygous individuals genotyped in each SNP is showed in 0|0 column; the sum of total heterozygous individuals genotyped in each SNP is showed in 1|0 column; the sum of total mutated homozygous individuals genotyped in each SNP is showed in 1|1 column. The numbers in total row represents an aggregation of all synonymous variants found in UCL cases.

<table>
<thead>
<tr>
<th></th>
<th>UK10K Control n = 982</th>
<th></th>
<th>UK10K SCZ n = 1392</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>R6R</td>
<td>982 0 0 0 1391 1 0</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>C19C</td>
<td>963 19 0 0.97 1374 18 0</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>A27A</td>
<td>966 16 0 0.81 1360 31 1</td>
<td>1.19</td>
<td></td>
</tr>
<tr>
<td>L34L</td>
<td>827 151 4 8.10 1160 221 11</td>
<td>8.73</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>792 186 4 9.88 1109 271 12</td>
<td>10.60</td>
<td></td>
</tr>
</tbody>
</table>

Nevertheless, all sSNP variants found in the UCL sample were also found in UK10K exome data with similar minor AF. Interestingly, the R6R variant was found in one case not in any of the controls. This is similar to the findings in the UCL samples and 1K Genome project data. C19C and A27A variants were detected at different AFs the controls and cases. A27A was more common in the SCZ samples compared to controls but the difference was not significant. L34L common variant showed modest evidence for association in UK10K cases (p = 0.042). Burden analysis of combined sSNP variants showed significantly associated (p = 0.0195) as shown in Table 4-6.
4.2.6 Swedish exome data analysis

Table 4-7 Summary of non-synonymous variants genotype counts from the Swedish exome sequence data. Percentage of allele frequencies (AF %) was calculated for each variant. The sum of total wild-type homozygous individuals genotyped in each SNP is showed in 0|0 column; the sum of total heterozygous individuals genotyped in each SNP is showed in 1|0 column. The numbers in total row represents an aggregation of all non-synonymous variants found in UCL cases.

<table>
<thead>
<tr>
<th></th>
<th>Swedish Control n = 2545</th>
<th></th>
<th>Swedish SCZ n = 2545</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>R69W</td>
<td>2545</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>R71H</td>
<td>2545</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>R127Q</td>
<td>2538</td>
<td>7</td>
<td>0.14</td>
</tr>
<tr>
<td>T128M</td>
<td>2543</td>
<td>2</td>
<td>0.04</td>
</tr>
<tr>
<td>V146M</td>
<td>2545</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>I156F</td>
<td>2544</td>
<td>1</td>
<td>0.02</td>
</tr>
<tr>
<td>T164L</td>
<td>2545</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>H233Y</td>
<td>2484</td>
<td>60</td>
<td>1.18</td>
</tr>
<tr>
<td>Total</td>
<td>2475</td>
<td>70</td>
<td>1.38</td>
</tr>
</tbody>
</table>

Swedish SCZ exome sequencing data for 2,500 SCZ and 2,500 controls was downloaded from dbGAP. Only four out of eight UCL genotyping nsSNP variants were detected and these included R127Q, T128M, I156F, and H233Y. Interestingly, the minor AF of R127Q was increased in both control and case samples (AF = 0.14%) of Swedish exome data. This SNP variant was not found in control samples from the UK10K exome sequencing data and UCL control subjects (Table 4-7). Conversely, the minor AF of T128M was decreased in controls and increased in case. Moreover, I156F was only found one in control subject and was absent in the cases. This was an opposite result to the previous sample. The minor AF of H233Y was quite similar between control and case samples, 2 individual subjects with SCZ were found to be homozygous for H233Y. Overall, the burden analysis of the variants showed almost equal minor AF in cases and controls. One reason for this is that only half the variants were found in Swedish data, and these were more common in controls compared to cases.
Table 4-8 Summary of synonymous variants genotype counts from the Swedish exome sequence data. Percentage of allele frequencies (AF %) was calculated for each variant. The sum of total wild-type homozygous individuals genotyped in each SNP is showed in 0|0 column; the sum of total heterozygous individuals genotyped in each SNP is showed in 1|0 column; the sum of total mutated homozygous individuals genotyped in each SNP is showed in 1|1 column. The numbers in total row represents an aggregation of all synonymous variants found in UCL cases.

<table>
<thead>
<tr>
<th>Variant</th>
<th>Swedish Control n = 2545</th>
<th>Swedish SCZ n = 2545</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>R6R</td>
<td>2545</td>
<td>0</td>
</tr>
<tr>
<td>C19C</td>
<td>2499</td>
<td>46</td>
</tr>
<tr>
<td>A27A</td>
<td>2472</td>
<td>73</td>
</tr>
<tr>
<td>L34L</td>
<td>2162</td>
<td>366</td>
</tr>
<tr>
<td>Total</td>
<td>2043</td>
<td>485</td>
</tr>
</tbody>
</table>

Three out of four UCL sSNP variants were found in the Swedish data (Table 4-8). Interestingly, the rare variant R6R was not found in the Swedish exome sequencing data, and C19C variant showed an increase in minor AF in cases compared to controls. However, both A27A and L34L variants showed higher minor AF in controls than in cases, but this was not significant. Burden analysis of combined total allele counts of three variants showed a slight significant associated in controls \((p = 0.0434)\).

4.2.7 Association analysis of CACNG5 variants

Burden analysis was performed using data for the combined allele counts of the 8 nsSNVs detected in CACNG5. The case versus control comparison study using the UCL control samples, were significant for SCZ \((p = 0.0016)\) and BPD \((p = 0.0127)\), and for the two diseases combined \((p = 0.0022)\) (Table 4-9 Upper). These findings were supported by the use of the data from the European samples in the 1K Genomes project as controls. In these analyses there was evidence for association with SCZ \((p = 0.0213)\), but was not significant in BPD \((p = 0.0814)\) and also for both diseases combined \((p = 0.0389)\) (data not shown). The evidence for association with both diseases seems to be decreased by BPD, and also UCL
controls seems to give better results than 1 K. A good explanation is because the UCL study includes healthy supercontrols. Using a combined set of control data that comprised genotypes from the UCL control samples and the data from the European samples in the 1K Genomes project compared with the BPD and SCZ sample data showed increased evidence for association with SCZ (\( p = 5.89 \times 10^{-4} \)) and with BPD (\( p = 5.91 \times 10^{-3} \)), and for both diseases combined (\( p = 5.69 \times 10^{-4} \)). Hence, increasing the sample size for controls (and cases) appeared to increase the evidence for association (Table 4-9 Lower). Furthermore, using the combined set of control data that comprised genotypes from the UCL control samples, the data from the 1K Genomes projects, and the data from the UK10K control exome sequencing compared with the BPD and SCZ samples data showed increased evidence for association with SCZ (\( p = 1.14 \times 10^{-5} \)) and with BPD (\( p = 1.71 \times 10^{-4} \)), and for both diseases combined (\( p = 2.17 \times 10^{-3} \)). However, when both diseases were combined with additional UK10K SCZ data and this was compared with the combined set of control data there was decreased evidence for association (\( p = 3.44 \times 10^{-3} \)). Similar significant findings were observed when the combined set of control data comprised additional genotypes from the Swedish control exome sequencing, when compared with the BPD and SCZ samples data showed increased evidence for association with SCZ (\( p = 9.45 \times 10^{-7} \)) and with BPD (\( p = 1.51 \times 10^{-5} \)), for both diseases combined (\( p = 2.04 \times 10^{-8} \)), for both diseases and UK10K SCZ combined (\( p = 9.98 \times 10^{-6} \)), and for both diseases combined with UK10K SCZ and Swedish SCZ (\( p = 5.12 \times 10^{-3} \)).

The results showed that the additional data from UK10K and Swedish exome sequencing data did not strengthen the evidence for association. Perhaps, some individual variants such as R69W, R71H, V146M, and T164L were still interesting to be further investigated with functional test. This is particularly the case for V146M and T164L which were not found in both UK10K and Swedish controls and cases data, and R69W and R71H were only not found in both sets of control data. According to the previous GWAS by the International
Schizophrenia Consortium (PSC) and CNV study, there was no evidence of major population stratification within our samples {Purcell, 2009 #917} {Consortium, 2008 #918}. As the UCL case-control samples were only recruited if both parents were English, Scottish or Welsh, and with at least three grandparents having the same origin. Furthermore, the data was only included if the fourth grandparent was of another white European origin, but if one grandparent was of Jewish or non-European Union (EU) then the subject was excluded. In the Aberdeen samples, all participants self-identified as born in the British Isles and 95% were in Scotland {Purcell, 2009 #919}. Therefore, despite the finding that both V146M and T164L were absent in the UK10K, it is unlikely not the presence of these variants in the cohorts studies here represents the effect of population stratification. Similar analyses of data for the five sSNPs revealed no evidence for association with either disorder. The results of the nsSNP and sSNP analyses are summarised in Table 4-9 and Table 4-10
Table 4-9 Summary of combined total 8 non-synonymous variants genotypic count and significant association p-values from the 1K Genome Project data, UCL samples, UK10K exome sequence data and Swedish exome data. The upper table represents genotype counts of rare non-synonymous variants found in CACNG5 from each cohort. The lower table represents CACNG5 a pooled burden analysis of rare non-synonymous variants in each of the cohorts and in a combined analysis of the cohorts. 1K = 1K genome project data; UCL CONT = UCL control subjects; UCL BP = UCL bipolar disorder subjects; UCL_SCZ = UCL schizophrenia subjects; UK10K = UK10K exome data control and schizophrenia subjects; Swedish exome data = Swedish exome data control and schizophrenia; 1K + UCL CONT = combined set of 1K and UCL CONT. 0|0 column represents the sum of total wild-type homozygous individuals genotyped in each data set; 1|0 column represents the total of heterozygous individuals genotype; 1|1 column represents the total mutated homozygous individuals genotyped.

<table>
<thead>
<tr>
<th>Genotypic Count</th>
<th>Control</th>
<th>BPD</th>
<th>SCZ</th>
<th>Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>UCL</td>
<td>909</td>
<td>32</td>
<td>0</td>
<td>1011</td>
</tr>
<tr>
<td>UCL + 1K</td>
<td>1274</td>
<td>45</td>
<td>0</td>
<td>1011</td>
</tr>
<tr>
<td>UK10K</td>
<td>949</td>
<td>33</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>UCL + 1K + UK10K</td>
<td>2223</td>
<td>78</td>
<td>0</td>
<td>1011</td>
</tr>
<tr>
<td>Swedish</td>
<td>2475</td>
<td>70</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>UCL + 1K + UK10K + Swedish</td>
<td>4698</td>
<td>148</td>
<td>0</td>
<td>1011</td>
</tr>
<tr>
<td>Association analysis (p-value)</td>
<td>UCL BPD</td>
<td>UCL SCZ</td>
<td>UCL Cases</td>
<td>UCL + UK10K SCZ</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>---------</td>
<td>---------</td>
<td>-----------</td>
<td>----------------</td>
</tr>
<tr>
<td>Controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCL</td>
<td>1.27E-02</td>
<td>1.58E-03</td>
<td>2.21E-03</td>
<td>1.18E-01</td>
</tr>
<tr>
<td>UCL + 1K</td>
<td>5.91E-03</td>
<td>5.89E-04</td>
<td>5.69E-04</td>
<td>8.11E-02</td>
</tr>
<tr>
<td>UCL + UK10K</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCL + Swedish</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCL + 1K + UK10K</td>
<td>1.71E-04</td>
<td>1.13503E-05</td>
<td>2.17E-06</td>
<td>3.53E-02</td>
</tr>
<tr>
<td>All Controls</td>
<td>1.52464E-05</td>
<td>9.92111E-07</td>
<td>1.11E-08</td>
<td>1.16E-03</td>
</tr>
</tbody>
</table>
Table 4-10 Summary of genotype counts and significant association p-values for combined analysis of the 5 synonymous variants from the 1K Genome Project data, the UCL samples, the UK10K exome sequence data and the Swedish exome data. The upper table represents genotype counts of synonymous variants found in \textit{CACNG5} from each cohort. The lower table represents \textit{CACNG5} pooled burden analysis association test result for the synonymous variants in each cohort or combined analysis of these cohorts. 1K = 1K genome project data; UCL CONT = UCL control subjects; UCL BP = UCL bipolar disorder subjects; UCL_SCZ = UCL schizophrenia subjects; UK10K = UK10K exome data control and schizophrenia subjects; Swedish exome data = Swedish exome data control and schizophrenia; 1K + UCL CONT = combined set of 1K and UCL CONT. 0|0 column represents the sum of total wild-type homozygous individuals genotyped in each data set; 1|0 column represents the total of heterozygous individuals genotype; 1|1 column represents the total mutated homozygous individuals genotyped.

<table>
<thead>
<tr>
<th>Genotypic Count</th>
<th>Control</th>
<th>BPD</th>
<th>SCZ</th>
<th>Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>UCL</td>
<td>775</td>
<td>158</td>
<td>9</td>
<td>890</td>
</tr>
<tr>
<td>UCL + 1K</td>
<td>1071</td>
<td>234</td>
<td>16</td>
<td>890</td>
</tr>
<tr>
<td>UK10K</td>
<td>792</td>
<td>186</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>UCL + 1K + UK10K</td>
<td>1863</td>
<td>420</td>
<td>20</td>
<td>890</td>
</tr>
<tr>
<td>Swedish</td>
<td>2043</td>
<td>485</td>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td>UCL + 1K + UK10K + Swedish</td>
<td>3906</td>
<td>905</td>
<td>37</td>
<td>890</td>
</tr>
<tr>
<td>Controls</td>
<td>Association analysis (p-value)</td>
<td>UCL BPD</td>
<td>UCL SCZ</td>
<td>UCL Cases</td>
</tr>
<tr>
<td>----------</td>
<td>--------------------------------</td>
<td>---------</td>
<td>---------</td>
<td>-----------</td>
</tr>
<tr>
<td>UCL</td>
<td>7.55E-01</td>
<td>2.02E-01</td>
<td>7.03E-01</td>
<td></td>
</tr>
<tr>
<td>UCL + 1K</td>
<td>2.38E-01</td>
<td>5.19E-01</td>
<td>6.03E-01</td>
<td></td>
</tr>
<tr>
<td>UCL + UK10K</td>
<td></td>
<td></td>
<td></td>
<td>3.19E-01</td>
</tr>
<tr>
<td>UCL + Swedish</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCL + 1K + Swedish</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The allelic association of the three promoter region SNPs between patients and healthy controls are summarised in Table 4-11. As shown in the table, there was no evidence for association with SCZ, BPD, or in a combined analysis. Only SNP 17_64873175 showed a significant association with UCL controls (p = 0.0236) and with a combined set of controls (p = 0.0327) when compared with UCL SCZ samples, no carrier found.

Table 4-11 Summary of putative promoter variants genotype counts and allelic association analysis

<table>
<thead>
<tr>
<th>rs3760263</th>
<th>Genotype Counts</th>
<th>p-value</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1K EURO</td>
<td>345</td>
<td>32</td>
<td>2</td>
<td>0.857</td>
<td>0.709</td>
</tr>
<tr>
<td>UCL CONT</td>
<td>866</td>
<td>71</td>
<td>0</td>
<td>0.210</td>
<td>0.076</td>
</tr>
<tr>
<td>1K + UCL CONT</td>
<td>1211</td>
<td>103</td>
<td>2</td>
<td>0.377</td>
<td>0.139</td>
</tr>
<tr>
<td>UCL BPD</td>
<td>963</td>
<td>91</td>
<td>3</td>
<td>0.534</td>
<td>0.076</td>
</tr>
<tr>
<td>UCL SCZ</td>
<td>538</td>
<td>53</td>
<td>4</td>
<td>0.997</td>
<td>0.024</td>
</tr>
<tr>
<td>UCL BPD + SCZ</td>
<td>1501</td>
<td>144</td>
<td>7</td>
<td>0.802</td>
<td>0.033</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>rs37687317B</th>
<th>Genotype Counts</th>
<th>p-value</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>rs75486725</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs75486725</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs75486725</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.2.8 Bioinformatics analysis – prediction of the effects of variant on protein function

To assess the possibility that the CACNG5 nsSNPs variants have functional effects on the protein, we performed bioinformatic analysis of the variants with the widely used programmes SIFT, PolyPhen2, and Mutation Taster. These programmes predict the functional effects of variants on
a protein. The results of these analyses are summarised in Table 4-12. The variant that was predicted to have the strongest effect on protein function variant was p.Val146Met and this variant had the highest scores in all tests with the prediction of the protein as probably functionally affected. Variants p.Arg69Trp and p.His233Tyr were predicted to be possibly damaging; H233Y is a relatively common SNP and is widely reported in the genome databases, whilst R69W was found in a single individual of African origin in the 1K Genomes project data.

Three variants, p.Thr128Met, p.Ile156Phe, and p.Thr164Ile were predicted to be probably damaging by PolyPhen2; conversely they were predicted to be tolerated by SIFT. Two further variants p.Arg127Gln and p.Ala182Ser, which were detected in the European sample from the 1K genome databases and were predicted to be benign.

Table 4-12 Summary of bioinformatic analysis of nsSNPs

<table>
<thead>
<tr>
<th>Amino acid change</th>
<th>PolyPhen2</th>
<th>SIFT</th>
<th>Mutation Taster</th>
</tr>
</thead>
<tbody>
<tr>
<td>R69W</td>
<td>Possibly damaging</td>
<td>Affect protein function</td>
<td>Protein affected</td>
</tr>
<tr>
<td>R71H</td>
<td>Probably damaging</td>
<td>Tolerated</td>
<td>Disease causing</td>
</tr>
<tr>
<td>R127Q</td>
<td>Benign</td>
<td>Tolerated</td>
<td>Protein affected</td>
</tr>
<tr>
<td>T128M</td>
<td>Probably damaging</td>
<td>Tolerated</td>
<td>Disease causing</td>
</tr>
<tr>
<td>V146M</td>
<td>Probably damaging</td>
<td>Affect protein function</td>
<td>Protein affected</td>
</tr>
<tr>
<td>I156F</td>
<td>Probably damaging</td>
<td>Tolerated</td>
<td>Protein affected</td>
</tr>
<tr>
<td>T164L</td>
<td>Probably damaging</td>
<td>Tolerated</td>
<td>Protein affected</td>
</tr>
<tr>
<td>A182S</td>
<td>Benign</td>
<td>Tolerated</td>
<td>Disease causing</td>
</tr>
<tr>
<td>H233Y</td>
<td>Possibly damaging</td>
<td>Affect protein function</td>
<td>Disease causing</td>
</tr>
</tbody>
</table>

PolyPhen2: [http://genetics.bwh.harvard.edu/pph2/index.shtml](http://genetics.bwh.harvard.edu/pph2/index.shtml)

SIFT: [http://sift.jcvi.org/](http://sift.jcvi.org/)

Mutation Taster: [http://www.mutationtaster.org/](http://www.mutationtaster.org/)
4.3 Discussion

The aim of our study was to identify functional risk variants for psychiatric disorders in the calcium channel subunit gene CACNG5. We chose HRMA as a fast and cost effective technique to analyse the complete coding and putative promoter regions of the CACNG5 gene. Amplicons that were likely to contain variants were then selected for Sanger sequencing. All variants successfully validated by Sanger sequencing were then genotyped in all available BPD, SCZ and controls samples. Our results do not unequivocally implicate the rare variants of the CACNG5 gene in the aetiology of BPD and SCZ. They also do not provide a clear answer to the question of whether certain variants are involved in distinct or overlapping susceptibility. They may however, go some way to explaining some of the association finding reported by Curtis et al., 2011 (Curtis et al. 2011).

One of the most promising variants detected was p.Val146Met in exon 4 of CACNG5. This variant was first identified in a UCL SCZ patient but was absent in the UCL BPD and control samples and in the European samples from the 1K Genomes project. Additionally, this variant was also detected in one SCZ patient from a replication case (n = 643) control (n = 518) cohort from Aberdeen (Table 4.2.2). Furthermore, it was also not found in both UK10K and Swedish exome sequencing data. Therefore p.Val146Met could be considered to be a rare SCZ relevant polymorphic variant. Bioinformatic analysis and localisation of the variant have also suggested that it is highly likely to have a functional effect on the protein. Interestingly, a recent functional study reported a de novo variant p.Val143Leu in CACNG2 (also known as stargazin), which is another calcium channel γ subunit gene, was found in a patient with intellectual disability. This variant significantly decreased stargazin’s ability to bind to GluR1 or GluR2 AMPAR subunits in HEK293 cells (Hamdan et al. 2011). It also showed a decrease in GluR1
cell surface expression and excitatory postsynaptic current in transfected hippocampal neurons. Amino acid alignments with the other TARP subunits show that the V146 and V143 residues are conserved within TARP subgroups (Figure 4-6).

<table>
<thead>
<tr>
<th>Residue</th>
<th>Sequence</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>γ4</td>
<td>VLSAGILFVAAGLSNIIGIIV</td>
<td>153</td>
</tr>
<tr>
<td>γ8</td>
<td>ILGAGILFVAAGLSNIIGVIV</td>
<td>168</td>
</tr>
<tr>
<td>γ2</td>
<td>ILSAGIFFVSAGLSNIIGIIV</td>
<td>155</td>
</tr>
<tr>
<td>γ3</td>
<td>ILSAGIFFVSAGLSNIIGIIV</td>
<td>155</td>
</tr>
<tr>
<td>γ5</td>
<td>AFVSGIFFILSGLSLVVGLVL</td>
<td>151</td>
</tr>
<tr>
<td>γ7</td>
<td>AFVSGIFFILSGLSLVVGLVL</td>
<td>151</td>
</tr>
</tbody>
</table>

**Homo sapiens**  NP_665810.1   ILSGLSLVVGLVLY 153  
**Mus musculus**  NP_542375.1   ILSGLSLVVGLVLY 153  
**Rattus norvegicus**  NP_542424.1   ILSGLSLVVGLVLY 153  
**Gallus gallus**  XP_415681.2   ILSGLSLVVGLVLY 153  
**Danio rerio**  XP_001337872.1   ILSGLSLVVGLVLY 153  

**Figure 4-6 Evolutionary conservation of the V146 region of CACNG5.** A. Protein sequence alignments of all human calcium channel γ subunits surrounding the mutations. B. Cross species Amino acid conservation of CACNG5 residues affected by the variant (Hamdan et al. 2011).
Figure 4-7 Analysis of variants in the extracellular loop of CACNG5. A. Schematic diagrams showing predicted membrane topology and putative functional sites on TARP subunits and claudins. B. Protein sequence alignments of the first large extracellular loop of all human TARP subunits (Chen et al. 2007).

The molecular dissection of TARP function has begun to elucidate the domains that mediate differential regulation of AMPA receptor trafficking and function (Tomita et al. 2005a). Structural and functional studies suggest that the long extracellular loop is essential for the modulation of the channel properties and also a binding site for AMPAR ligand binding domain (Tomita et al. 2005a, Turetsky et al. 2005). However, the protein alignments of TARPs show important differences between subgroups. Previous studies showed that the substitution of this domain of stargzin (γ2) with that from γ4 increased glutamate affinity and reduced the kinetics of deactivation and desensitisation (Cho et al. 2007, Milstein et al. 2007). Thus, the sequence differences in this domain may confer altered receptor pharmacology and channel gating. In addition, this domain has a close sequence alignment with Claudin proteins particularly at N-glycosylation sites and at two cysteine residues. Interestingly, two of our variants R69W and R71H, are located in the first extracellular between the two cysteine residues (Figure 4-7). Bioinformatic prediction revealed that R69W possibly effects protein function, while R71H was
only predicted to be probably damaging by one analysis. This is likely to be a reflection of
difference in the conservation of the two residues. Maher et al (Maher et al. 2011) showed that
residue of R69W has similar amino acids present in >50% of sequences in other calcium channel
γ subunits. Whereas, Chu et al (Chu et al. 2001) showed that this residue was only common to γ2
through γ8, excluding γ6. Conversely, R71H polymorphism is not a conserved residue, but the
alternative codon histidine is identical to the equivalent γ4 residue. Even though there is not
strong support for this finding, there is still the possibility that the physical properties of this
substitution could have an effect on the functioning of neighbouring amino acid(s) that in turn
alter protein affinity.

Figure 4-8 Evolutionary conservation of the R127 and T128 region of CACNG5. A. Protein
sequence alignments of all human γ subunits surrounding the R127Q and T128M variant
sites. B. Cross species amino acid conservation of CACNG5 residues affected by the
variants.
Bioinformatic analysis of the R127Q amino acid substitution revealed that this change is likely to have a benign effect on protein function. Similar analysis of T128M with Polyphen 2, SIFT produced a prediction of probably damaging (albeit with a damaging score of 0.997; the most damaging score being 1) with Polyphen 2 whilst SIFT predicted that this amino acid change would be tolerated. According to their amino acid position, R127Q and T128M are likely to be located in the intracellular loop of the protein, close to the third transmembrane domain. Cross species amino acid conservation analysis showed that both R127Q and T128M SNPs are highly conserved within the species analysed (Figure 4-8 pane B). However, the amino acid alignment of all human γ subunits revealed a low conservation in T128M, while R127Q has approximately 50% identical residue to other γ-subunits (Figure 4-8 pane A). Hence, these two amino acid substitutions might be predicted to have a low possibility of effecting protein function.

Two novel nsSNPs that were predicted to lead to I156F and T164I amino acid substitutions were predicted to be probably damaging in PolyPhen2 analysis but were predicted to be tolerated in SIFT. Both residues are located in the second extracellular loop, which does not appear to be involved in AMPA receptor binding or trafficking. However, amino acid substitutions could cause alterations in the secondary structure of the protein. Hence, there is still a possibility that these SNPs could still have an effect on protein function.
Bioinformatic analysis of the H233Y substitution caused by the more common nsSNP rs41280112 was also been predicted to affect protein function. Type I TARPs contain a typical PDZ domain binding motif (-TTPV) at their C-terminus (Figure 4-9). This binding motif has showed to regulate the synaptic localisation of AMPARs. In contrast, Type II TARPs contain an atypical PDZ-binding motif (-S/TTPC) at their C-terminus. H233Y is located in the C-terminus and close to the atypical PDZ-binding motif (Figure 4-9). Hence, the amino acid substitution, which changes the physical properties of the residue, could affect the binding affinity of the domain.

In summary, screening of CACNG5 for variants by HRMA enabled identification of rare variants of that may have a role in susceptibilities to SCZ and/or BPD. However, bioinformatic analysis only provides a prediction of the effect of these substitutions without giving evidence for how these changes may affect the normal functioning of the protein. To find mechanistic evidence for the impact of the candidate variants on CACNG5 function, further functional studies must be conducted. Thus, the full characterisation of molecular aetiology of variants in the CACNG5 gene may facilitate possible therapeutic intervention.
5 Rare mutations identified in CACNG5 from bipolar disorder and schizophrenia individuals exhibit impaired expression of intracellular and cell surface AMPA receptors

5.1 Background

Bipolar disorder and schizophrenia are important causes of disability as well as financial and emotional burden worldwide. Over the past decade, the glutamatergic system has been implicated in the pathophysiology and possible treatment of major depressive disorder, bipolar disorder, and schizophrenia; however, the molecular changes that underlie this pathology still remain poorly understood (Auer et al. 2000, Hashimoto et al. 2007, Mauri et al. 1998, Mitani et al. 2006). Accordingly, ionotropic glutamate receptors (iGluRs) which mediate fast excitatory neuro-transmission at synapses within the central nervous system are promising therapeutic targets for the treatment of neuropsychiatric disorders. Previous studies have suggested that the iGluR N-methyl-D-aspartate (NMDA) receptor may be involved in the pathophysiology of mood disorders (Javitt 2004, Petrie et al. 2000, Skolnick et al. 2009). NMDA receptor antagonists such as phencyclidine (PCP) and ketamine induce schizophrenia-like symptoms in normal subjects and that there are even worse in schizophrenia patients (Allen and Young 1978, Barbon et al. 2007, Coyle 1996, Coyle et al. 2003, Ellison 1995, Lahti et al. 1995, Meador-Woodruff and Healy 2000). Recently, a review of ketamine suggests that a low-dose ketamine effects on the glutamatergic system and abnormalities in this neurotransmitter system are present in depression (Naughton et al. 2014). One mechanism that may explain the abnormal regulation of NMDA is that this may be due to abnormal expression and localisation of AMPA glutamate receptors. There is evidence that alteration of AMPA receptor activity at the postsynaptic density either by their dysregulation of receptor expression or cell surface trafficking could decrease activity of
NMDA receptors (NMDAR), resulting in the appearance of psychotic symptoms (Coyle et al. 2003, Meador-Woodruff and Healy 2000). Some studies have also found increased AMPA receptor (AMPAR) binding in the cortex of schizophrenia patients (Noga et al. 2001) (Zavitsanou et al. 2002). These findings suggest that the dysregulation of AMPA receptors (AMPARs) may lead to psychotic illness.

A potential mechanism underlying AMPAR disturbances in neuropsychiatric conditions is abnormal expression of AMPAR auxiliary proteins that regulate AMPAR function, localisation, and trafficking (Beneyto and Meador-Woodruff 2006, Dracheva et al. 2005, Hammond et al. 2010, Malinow and Malenka 2002, Mirnics et al. 2000, Song and Huganir 2002, Toyooka et al. 2002, Whiteheart and Matveeva 2004). Transmembrane AMPAR regulatory protein gamma (γ) subunit 2 (TARP-γ2/stargazin), was the first such protein found to interact with AMPARs (Chen et al. 2000, Diaz 2010, Nakagawa and Sheng 2000, Tomita 2010, Vandenberghhe et al. 2005b). γ2 was initially identified from the natural occurring mutation that is found in the so called stargazer mice, result in absence epilepsy, cerebellar ataxia, and a characteristic abnormal motor syndrome (Khan et al. 2004). The stargazin mutation is associated with selective loss of AMPARs function in cerebellar granule cells (Chen et al. 2000, Hashimoto et al. 1999). The gene encoding γ2 is known as voltage-dependent calcium channel gamma subunit 2 (CACNG2) due to its sequence homology to the skeletal muscle tetraspanning calcium channel γ1 subunit (Jay et al. 1990). Currently, eight TARPs have been identified, each subunit seems to have a variety of roles in AMPAR cell surface trafficking and its biophysical properties (Chen et al. 2003, Chen et al. 2007, Coombs and Cull-Candy 2009, Diaz 2010, Jackson and Nicoll 2011a, Kato et al. 2007, Klugbauer et al. 2000, Tomita et al. 2003). For instance, γ4 shows an effect on the slow gating of AMPARs (Cho et al. 2007, Korber et al. 2007, Kott et al. 2007, Milstein et al. 2007); whereas γ7
greatly enhance glutamate-evoked currents from GluR1 (Kato et al. 2007). Specifically, γ5 has been widely used as a negative control in experiments for identifying TARPs (Tomita et al. 2005a, Tomita et al. 2004, Turetsky et al. 2005), because it does not restore functional AMPARs on stargazer cerebellar granule cells (Kato et al. 2007). Recently, γ5 was found to selectively regulate the functional properties of a splice variant of the AMPAR GluR2 subunit and the homomeric AMPAR of GluR1 receptor (Soto et al. 2009). In addition to trafficking of AMPAR, γ5 significantly decreased the cell surface expression of homomeric GluR2(Q) and long form GluR2(Q) protein; in contrast, it led to a slight increase in the cell surface expression of homomeric GluR2(R) protein (Soto et al. 2009).

Although several studies have reported the regulation of AMPAR by TARPs, little is known about the expression of TARP-γ5. Several recent studies on inflammatory bowel disease (Rivas et al. 2011), multiple sclerosis (Zhuang et al. 2015), type 2 diabetes (Bonnefond et al. 2012), schizophrenia (Gulsuner et al. 2013, Kenny et al. 2014) have demonstrated that susceptibility to common disease may mediated by rare variants. With this disease mechanism in mind, we characterised eight non-synonymous SNPs (R69W; R71H; rs142916987, R127Q; rs149159754, T128M; V146M; I156F; T174L; rs41280112, H233Y) in human CACNG5, using a combination of in vitro co-transfection and flow cytometry (fluorescence activated cell sorting; FACS). Functional characterisation of such variants may prove useful for the translation of genetic findings to disease prediction, pharmacogenetic applications and novel drug development.
5.2 Aims and Objective

The previous genotyping analysis found evidence that nsSNP CACNG5 variants were associated with psychiatric illness. Bioinformatic analysis of the potential consequence that the variant may have a damaging effect on protein structure/function. It is therefore possible that the nsSNPs may have an effect upon the regulation of protein expression and/or AMPAR surface expression, which in turn may be associated with psychosis illness. Additionally, in vitro mutated CACNG5 constructs should provide opportunities to understand how TARP-γ5 function is affected.

To determine the effects of these variants, the previous genotyping results were used to design mutated CACNG5 expression constructs that were subsequently co-transfected into HEK293 cells alone with GluR1 or GluR2 homotetramer subunits. The subsequent effect of the CACNG5 variants on GluR1 or GluR2 homotetramer subunit expression was monitored by the presence of a pH-sensitive green fluorescent protein (GFP) fusion protein on the cell membrane. This pH-sensitivity GFP tag protein is a mutated GFP with pH sensor, known as super-ecliptic pHluorin (SEP), which was tagged on AMPAR subunits and acts as a marker for surface-expressed receptors (Ashby et al. 2004, Miesenbock et al. 1998, Sankaranarayanan et al. 2000). The construct is a chimera of SEP fused to the N-terminus of the AMPAR subunit (i.e. SEP-GluR2). When the N-terminal domain of receptor tagged with SEP migrates to the pH neutral outside of the cell from the relatively acidic internal vesicle compartments there is a dramatic (100 fold) increase in fluorescence (Figure 5-1) (Ashby et al. 2004). The green fluorescent intensity signal from each cells were captured by FACS. This cell sorting technique allowed discrimination of cells by the intensity of their green fluorescent signal. The actual numbers of cells with different levels of fluorescence determined automatically by the software. FACS has been widely used for
Figure 5-1 The migration of SEP-GluR2 through the compartments of the secretory pathway. ER: endoplasmic reticulum (Ashby et al. 2004).

Since the disease-associated rare variants found in our previous study were predicted to have a possible effect on protein structure or binding affinity, it is a highly interesting investigation on their effect. To test this, SEP-GluR1 or SEP-GluR2 alone with wild-type TARP-γ5 or mutated TARP-γ5 were co-transfected in HEK293 cells. 24 hours after transfection, the cells were collected and sorted using FACS (BD FACSariaII) to quantify the number of cells and the intensity of green fluorescence in each cell. Discrimination the location of AMPAR subunits expressed between intracellular and surface was set by the given green fluorescent intensity from
the membrane AMPAR proteins, where it showed a 100-fold brighter green fluorescent signal. Data were analysed using Kaluza software and presented as the mean ± SEM. The statistical analysis was performed using a two-tailed unpaired Student’s t-test in Excel. The differences were considered significant with a P-value < 0.05.

5.3 Results
To quantify the number of GluR2 expressing cells, a control experiment was run from with and without fluorescent tag proteins. The data from these experiments were used to determine baseline parameters for the Kaluza analysis software. Each sample of cells was sorted at a rate of 7500 – 8500 events ensuring that necrotic cells and cellular debris were not captured as legitimate events. Co-transfection of non-tagged AMPAR and TARP-γ5 in HEK293 cells revealed zero green fluorescent intensity as shown in region C of figure 5-2 A; this represented the distribution of unstained cells. In contrast, SEP tagged AMPAR co-expressed with TARP-γ5 showed a 10-fold increase in green fluorescent intensity and this represented the intracellular expression of the AMPAR fusion protein in the cells (Figure 5-2 B; area D and Y). As mentioned previously SEP tagged GluR2 constructs expressed a super-ecliptic pHluorin protein, which is known as an enhanced version of pHluorin. The trafficking of the AMPAR to the cell surface would be expected to show a 100-fold brighter green light from cells. Cells with these high levels of fluorescence would represent cells expressing cell surface trafficked AMPAR (Figure 5-2 B; area Y).
Figure 5-2 Typical histogram plots of GFP intensity from transfected cells generated by Kaluza software. The figure illustrates how regions were set to discriminate between the regions of SEP-AMPAR expression in the data. Areas E+D+Y represent all AMPAR expressing cells; areas C or E represent non-green fluorescent protein expression cells; area D represents intracellular AMPAR expressing cells; area Y represents membrane AMPAR expressing cells.

<table>
<thead>
<tr>
<th>Area</th>
<th>Cell counts</th>
<th>% Total</th>
<th>% Area</th>
<th>Medium</th>
</tr>
</thead>
<tbody>
<tr>
<td>E+D+Y</td>
<td>7,890</td>
<td>78.9</td>
<td>100</td>
<td>5.59</td>
</tr>
<tr>
<td>D</td>
<td>4,814</td>
<td>48.14</td>
<td>61.01</td>
<td>14.92</td>
</tr>
<tr>
<td>Y</td>
<td>19</td>
<td>0.19</td>
<td>0.24</td>
<td>132.31</td>
</tr>
</tbody>
</table>
5.3.1 TARP-γ5 increased cells expressing GluR2

Figure 5-3 Percentage of cells expressing GluR2 was compared with cotransfected wild-type TARP-γ5 and control vector. Unstained represents cells co-expressing non-fluorescent control constructs; GluR2 represents as cells co-expressing SEP-GluR2 and a control vector; TARP-γ5 represents cells co-expressing SEP-GluR2 and the wild-type TARP-γ5 construct.

The discrimination parameter utilised by the Kaluza software showed that the percentage of cells expressing GluR2 was increased when the cells were co-transfected with wild-type TARP-γ5 from 58.06% to 67.05% of cells (Figure 5-3). Only two replicates of these experiments were performed and it was not possible to statistically analyse these data. However, the findings were in agreement with the previous literature (Tomita et al. 2005a, Tomita et al. 2004, Turetsky et al. 2005). If wild-type TARP-γ5 enhanced GluR2 expression in the cells, does TARP-γ5 also regulate the trafficking of GluR2 to the cell surface?
5.3.2 TARP-γ5 increases GluR2 trafficking to the cell surface

![Graph showing percentage of cells expressing surface GluR2](image)

Figure 5-4 The percentage of cells expressing GluR2 on their surface was compared with co-expressed wild-type TARP-γ5 and control vector. Unstained represents cells co-expressing non-fluorescent constructs; GluR2 represents cells co-expressing SEP-GluR2 and control vector constructs; TARP-γ5 represents cells co-expressing SEP-GluR2 and the wild-type TARP-γ5 construct.

To identify whether TARP-γ5 regulates the trafficking of GluR2 in the cells, the surface and intracellular expression of GluR2 were discriminated via the intensity of the green fluorescence (Figure 5-2 B; area Y). The results from these experiments revealed that wild-type TARP-γ5 co-transfection led to increased surface trafficking of GluR2, from 0.21% to 0.78% (Figure 5-4). A similar result has been published which showed that the co-expression of wild-type TARP-γ5 increased surface trafficking of edited GluR2, known as Ca^{2+} impermeable subunit (Soto et al. 2009). Taken together, the results suggested that wild-type TARP-γ5 regulates specific Ca^{2+} impermeable GluR2 subunit expression and surface trafficking in the cells.
5.3.3 TARP-γ5 variants significantly increase GluR2 cell surface expression

![Graph showing percentage of cells expressing GluR2](image)

Figure 5-5 Percentage of cells expressing GluR2 was compared after co-transfection with wild-type mutated TARP-γ5. WT represents cells co-expressing wild-type TARP-γ5 constructs; SNP represents cells co-expressing mutated TARP-γ5 constructs. Data were presented as mean ± SEM (n = 3 or 4).

To characterise the rare variants (nsSNPs) found in our schizophrenia and bipolar disorder patients (Table 4-2), each variant was introduced into the CACNG5 gene, and then co-transfected with GluR2 in HEK293 cells. The cells expressing GluR2 were not significantly affected by the co-transfection of mutated CACNG5 (Figure 5-5).
Figure 5-6 Percentage of cells expressing surface GluR2 was compared after co-transfection with wild-type and mutated TARP-γ5. WT represents cells co-expressing wild-type TARP-γ5 constructs; SNP represents cells co-expressing mutated TARP-γ5 constructs. Data are presented as mean ± SEM (n = 3 or 4), * = p<0.05; ** = p<0.005; *** = p<0.001.

By contrast cell surface GluR2 expression was significantly affected by co-transfection with specific TARP-γ5 variants (Figure 5-6). The most promising result is that the cells co-transfected with an amino acid substitution at p.Val146Met of TARP-γ5 gene significantly increased cells expressing surface GluR2, from 0.78% ± 0.40 cells to 18.89% ± 1.22 cells, p = 3.12 x 10^{-5} (Figure 5-6 B). Based on the observed changes in cell surface GluR2 expression, it may be assumed that this rare allele may have a role in psychiatric conditions such as schizophrenia (characterised by excess of GluR2 trafficking). Another functional study of the de novo variant p.Val143Leu in TARP-γ5 found in an individual with intellectual disabilities showed significantly decreased binding affinity to GluR2 in HEK293 cells (Hamdan et al. 2011). These results suggest that rare variants in Ca^{2+} channel γ subunit genes have the potential to affect regulation of GluR2.
Another variant, p.Ile156Phe also located in second extracellular loop of TARP-γ5 protein showed a significant increase in the percentage of cells expressing surface GluR2 from 0.78% ± 0.40 cells to 10.15% ± 3.23 cells (p = 0.019; Figure 5-6 B).

The variant p.His233Tyr located at the C-terminus and close to the atypical PDZ-binding motif of TARP-γ5, also showed a significantly increase in the percentage of cells expressing GluR2 on their surface (from 0.78% ± 0.40 cells to 11.38% ± 2.20 cells, p = 0.0026; Figure 5-6 B). The atypical PDZ domain binding motif (-TTPV) of type I TARPs at their C-terminus has been shown to regulate the synaptic localisation of AMPARs (Itakura et al. 2014). Our results indicate that the C-terminus of TARP-γ5 may play a role in regulating cell surface trafficking of Ca^{2+} impermeable GluR2 subunits.
5.3.4 The effect of TARP-γ5 on the proportion of cells expressing GluR1

Figure 5-7 The percentage of cells expressing GluR1 was compared after co-transfection with wild-type TARP-γ5 and a control vector. Unstained represents cells co-expressing unstained constructs; GluR1 represents cells co-expressing SEP-GluR1 and control vector constructs; TARP-γ5 represents cells co-expressing SEP-GluR1 and the wild-type TARP-γ5 construct. Data are presented as mean ± SEM (n = 4)

Several of the Ca\(^{2+}\) channel γ subunit gene family members interact with GluR1 subunits of AMPAR. Therefore, TARP-γ5 could conceivably facilitate the elevation of intracellular or cell surface GluR1 expression and trafficking. Our result showed that there was a modest trend for an increase in the percentage of cells expressing GluR1 following by co-transfection with wild-type TARP-γ5 (from 55.91% ± 3.41 cells to 61.43% ± 4.77 cells, \(p = 0.14\); Figure 5-7).
Figure 5-8 The percentage of cells expressing surface GluR1 was compared after co-transfection with wild-type TARP-γ5 and a control vector. Unstained represents cells co-expressing unstained constructs; GluR1 represents cells co-expressing SEP-GluR1 and control vector constructs; TARP-γ5 represents cells co-expressing SEP-GluR1 and the wild-type TARP-γ5 construct. Data are presented as mean ± SEM (n = 4).

There was also a non-significant trend for an increase in the percentage of cells expressing GluR1 on their cell surface after co-transfection with wild-type TARP-γ5 (Figure 5-8). One explanation for the lack of significance was that the number of cells expressing intracellular and cell surface GluR1 were variable. Although we did not produce strong evidence of TARP-γ5 regulating the percentage of cells expressing GluR1, it is still interesting to understand the regulation of GluR1 by TARP-γ5 in the cells.
5.3.5 Mutated TARP-γ5 decreases the percentage of cells expressing intracellular and cell surface GluR1

![Graph showing % cells expressing GluR1](image)

**Figure 5-9** The percentage of cells expressing GluR1 was compared after co-transfection with wild-type and mutated TARP-γ5. WT represents cells co-expressing wild-type TARP-γ5 constructs; SNP represents cells co-expressing mutated TARP-γ5 constructs. Data are presented as mean ± SEM (n = 4), * = p<0.05; ** = p<0.005.

To identify whether each rare variant in TARP-γ5 affects regulation of the percentage of GluR1 expressing cells, co-transfection of GluR1 and mutated TARP-γ5 cells were compared with co-transfection with wild-type TARP-γ5. Both p.Arg69Trp and p.Arg71His located in the long extracellular loop of the TARP-γ5 protein showed a significant reduction in the percentage of cells expressing GluR1 (Figure 5-9). Structural and functional studies have suggested that the long extracellular loop of TARP-γ5 plays an important role in the modulation of AMPAR channel properties. This loop also harbours the AMPAR ligand binding domain (Tomita et al. 2005a, Turetsky et al. 2005). The substitution of this domain of TARP-γ2 with the domain from TARP-γ4 increases glutamate affinity (Cho et al. 2007, Milstein et al. 2007). Therefore, the reduction in the percentage of cells expressing GluR1 that resulted from the presence of these two variants may also affect channel properties and binding affinity.
The most significant changes in GluR1 expression were with the p.Thr128Met and p.Thr164Leu variants. Both showed a decrease in cells expressing GluR1 ($p = 0.0029$ and $p = 0.0041$, respectively; Figure 5-9). p.Thr164Leu is also located in the extracellular loop of the TARP-$\gamma$5 protein, which is known to mediate the binding affinity of AMPAR (Turetsky et al. 2005); whereas p.Thr128Met is located in the intracellular loop, close to the third transmembrane domain. To date, there is no evidence to suggest a functional role for the intracellular loop of the TARP-$\gamma$5 protein.

Figure 5-10 The percentage of cells expressing GluR1 on their surface was compared after co-transfection with wild-type and mutated TARP-$\gamma$5. WT represents cells co-expressing wild-type TARP-$\gamma$5 constructs; SNP represents cells co-expressing mutated TARP-$\gamma$5 constructs. Data are presented as mean ± SEM (n = 3 or 4), * = p<0.05.

The four CACNG5 variants that were found to reduce the percentage of cells expressing GluR1 also showed significant decreases in the percentage of cells expressing GluR1 on their surface (Figure 5-10). This significant reduction also included p.Arg127Gln one of the variants that was also located in the intracellular loop of the TARP-$\gamma$5 protein. These results suggest that TARP-$\gamma$5
variants may have a potential effect on the expression of both GluR1 and GluR2 subunits and particularly on their trafficking to the cell surface.

5.4 Discussion

TARPs are important regulatory proteins that indirectly mediate glutamatergic synapses. Several variants in this family of proteins have been shown to be associated with different neuropsychiatric disorders, such as schizophrenia, epilepsy, and intellectual disability, (Drummond et al. 2013, Hamdan et al. 2011, Letts et al. 1998). However, gaps remain in the in depth understanding of functional characterisation of all TARPs. From our previous genetic study, the rare variants found in TARP-γ5 gene, known as CACNG5, showed an association in bipolar disorder and schizophrenia. Such rare variants may contribute to the aetiology of these complex disorders and therefore the functional characterisation of their effects is important (Bamshad et al. 2011, Gibson 2011).

Quantitative and functional analysis of eight variants in the CACNG5 gene in comparison to the wild-type have provided insights into which of them may affect functionally important amino acids. Of the eight CACNG5 variants analysed, p.Val146Met and p.Ile156Phe which are located in the second extracellular loop region of TARP-γ5, significantly increased cell surface expression of GluR2 (Figure 5-6 B). Conversely, these two variants led to non-significant decreases in the percentage of cells expressing GluR1 on their surface (Figure 5-7). The analysis presented here suggests that this TARP-γ5 domain may differentially interact with specific AMPAR subunits.

The p.Thr164Leu variant is also located in the second extracellular loop region of TARP-γ5, and showed a significant decrease in the percentage of cells expressing GluR1 on their surface
(Figure 5-8). There was also a non-significant reduction in cells expressing surface GluR2 was also reduced when co-transfected with this in CACNG5 variant (Figure 5-6 A). These results suggest that this variant may reduce the expression and trafficking of AMPAR subunits.

Four rare variants, p.Arg69Trp, p.Arg71His, p.Arg127Gln, and p.Thr128Met are located in the first extracellular loop and intracellular loop regions of TARP-γ5 and led to significant decreases in the percentage of cells expressing GluR1 on their surface (Figure 5-9). These variants also led to non-significant decreases in the percentage of cells expressing surface GluR2 on their surface (Figure 5-6 A). Functional and structural studies have suggested that the first extracellular loop of TARP-γ5 controls AMPAR channel properties and its binding affinity (Tomita et al. 2005a, Turetsky et al. 2005). Our results are in agreement with the previous findings and suggest that p.Arg69Trp and p.Arg71His variants reverse the normal regulation of AMPAR subunits in these cells. By contrast, there is an unexpected considering that p.Arg127Gln, and p.Thr128Met variants fall in the region involved in trafficking of AMPAR subunits. To date this intracellular loop region of TARPs has not been functionally characterised. Therefore, our results are the first evidence that variants embedded in this intracellular loop region of TARP-γ5 affect expression and trafficking of AMPAR subunits.

Functional analysis of the p.His233Tyr variant showed an increase in the percentage of cells expressing GluR2 on their surface, but a decrease in the percentage of cells expressing GluR1 on their surface. Both TARP-γ5 and its close relative TARP-γ7 have unusually short cytosolic tails (C-tails), and consequently lack two, of the ten, phosphorylation sites present in the C-tails of other TARP members. This shorter PDZ binding domain differs from those of other members suggesting that they may interact with distinct protein partners, such as Ca^{2+} impermeable GluR2
subunits (Bats et al. 2007). Moreover, as these sites are thought to be involved in regulation of AMPAR trafficking (Choi et al. 2002, Tomita et al. 2005b), it seems likely these differences between TARPs could enable differential receptor regulation.

Since the regulation role of TARP-γ5 may be functionally different from other TARPs, the functional analysis of each amino acid substitution revealed that the specific functional domains of TARP-γ5 appear to differentially regulate specific AMPAR subunits. For instance, the variants embedded in the second extracellular loop and C-terminus region showed an increase in cells expressing GluR2 on their surface, but a decrease in cells expressing GluR1 on their surface. As a result, TARP-γ5 appears to act as a unique functional regulator of AMPARs compared to other members of calcium channel γ subunit gene family.

Variants in the extracellular loop and C-terminus of TARP-γ5 (p.Arg69Trp, p.Arg71His, p.Val146Met, p.Ile156Phe, p.Thr164Leu and p.His233Tyr) appear to be involved in GluR1 and/or GluR2 trafficking. These variants appear to affect the binding of TARP-γ5 to AMPARs leading to altered trafficking rate. These variants provide a substrate that could be used to develop and understanding of the functional relationship between TARP-γ5 and this important neurotransmitter receptor which is also a high affinity antipsychotic target.

From the genetic perspective the burden analysis suggested that in aggregate all of the variants studied here were associated (Chapter 4.0); however, the analysis presented here did not functionally implicate all of the variants. For the variants that do appear to alter AMPAR trafficking to the cell surface, these may begin to explain some of the complexity in the clinical features, etiology, family history and treatment response of bipolar disorder and schizophrenia. Psychostimulants that mimic altered AMPAR subunit trafficking in the manner that appear to be
caused by some of the CACNG5 variants described here, may be of therapeutic value for bipolar disorder, schizophrenia, and similar neuropsychiatric disorders.

Continued functional investigation of genetic variants could hold promise for improved understanding of the underlying mechanisms that lead to disease and also in the development of new therapeutics. For example, a recent review on this topic highlighted of a rare homozygous mutation in the PCSK9 gene that reduces cholesterol levels and the potential for the translation of this finding into a blockbuster drug (Hall 2013). Our results based on rare amino acid substitution variants and their possible functional analysis provide a further level of evidence that some rare variants (V146M, I156F, and T164L) may disrupt cellular or molecular mechanisms and that these may lead to complex neuropsychiatric conditions.
6 Replicated study of rare mutations identified in CACNG5 from bipolar disorder and schizophrenia individuals exhibit impaired cell surface trafficking of AMPA receptors subunit 2

6.1 Background

The pathophysiology of major depressive disorder, bipolar disorder, and schizophrenia are thought to associate with the abnormal regulation of glutamate receptor signalling; however, the molecular changes that underpin this genetic susceptibility remain unclear (Auer et al. 2000, Hashimoto et al. 2007, Mauri et al. 1998, Mitani et al. 2006). The principle glutamate receptors are classified as ionotrophic and metabotropic receptors. The ionotropic receptor family comprises NMDA, AMPA, and kainate receptors. The AMPA receptor (AMPAR) subfamily is particularly important in diverse sensory, behavioural and cognitive processes involving fast synaptic transmission, including learning and memory. By contrast, excessive AMPAR activity at the postsynaptic density mediated by dysregulation of receptor expression or cell surface trafficking, could cause central nervous system disorders ranging from stroke to epilepsy (Addae et al. 2007). The development of AMPAR antagonists has been suggested as a way to improve abnormal glutamate neurotransmission, and would possibly to provide treatments for neuropsychiatric disorders such as schizophrenia, depression, and bipolar disorder.

In the past decade, transmembrane AMPAR regulatory proteins (TARPs) were identified as AMPAR auxiliary subunits. They are thought to control receptor trafficking, gating, and pharmacology (Chen et al. 2000, Hashimoto et al. 1999, Vandenberghe et al. 2005a). Stargazin increases AMPAR glutamate affinity, enhances single-channel conductance, slows deactivation and desensitisation, and reduces the extent of desensitisation (Bedoukian et al. 2006, Priel et al.
2005, Tomita et al. 2005a, Turetsky et al. 2005). Mutations of TARP-γ2 (CACNG2; known as stargazing) cause absence epilepsy and cerebellar ataxia, and are thought to be associated with selective loss of AMPAR function (Chen et al. 2000, Hashimoto et al. 1999). Later, a family of related TARPs were identified, including γ3, γ4, γ5, γ7, and γ8, and these regulate AMPARs in distinct cell types throughout the brain (Tomita et al. 2003). Molecular analyses have found that abnormal expression of TARPs is also associated with neurological illness. Post-mortem studies using brains from people suffering from schizophrenia and major depressive disorder showed increased and decreased stargazin mRNA expression, respectively (Beneyto and Meador-Woodruff 2006). Additionally, the allelic polymorphisms found in the gene that encodes stargazin, CACNG2 are associated with increased response to lithium, the classical treatment for bipolar disorder (Silberberg et al. 2008). CACNG2 deletion copy number variants have also been found in patients with bipolar disorder and schizophrenia (Wilson et al. 2006). Taken together, there is good evidence to suggest that dysregulation of TARPs may associate with these neuropsychiatric illnesses. The results of our FACS analysis presented in chapter 5 suggested that transient co-expression of GluR2 and TARP-γ5 showed a modest increase of GluR2 expression in HEK293 cells and that this may have been on the cell surface. Moreover, a statistically significant increase in GluR2 expression was shown when the cells were co-transfected with GluR2 and certain TARP-γ5 variants, such as V146M, I156F, and H233Y; whereas the other variants led to decreased GluR2 expression. In addition experiments using the GluR1 subunit gave similar results but there were not statistically significant. These results suggested that TARP-γ5 may regulate the expression of GluR2, and that some of the rare variants detected in our studies may have specific effects on GluR2 expression and trafficking. TARP-γ5 may therefore be an effective tool for modulating the plasticity of excitatory synapses.
Although the increased green fluorescence intensity of GluR2 indicated that co-transfection with TARP-γ5 was having an effect in the cells, the actual mechanism by which these proteins interact are still unclear. For instance, the green fluorescence intensity of the SEP fusion protein in the cells could be altered by its environment pH level, and this could provide spurious results. Also, the intensity of green fluorescence was not a quantitative measure of the amount of GluR2 protein expressed in the cells or trafficked to the cell surface. In addition to TARP-γ5 expression, the possible effect of the TARP-γ5 variants on TARP-γ5 expression or localisation was not measured in the previous experiments. Therefore, in order to determine whether TARP-γ5 variants alter the trafficking or expression of GluR2 in the cells, cell surface biotinylation assays with an additional red fluorescent tag protein on TARP-γ5, RFP-γ5 wild type or variant constructs were applied (see Methods section).

Cell surface biotinylation is a method to quantify the total amount of protein on the surface of a cell (Fairfax et al. 2004). This is an useful method for comparing treated versus untreated groups of cells, and also can be used to determine the proportion of the total receptor pool that resides at the cell surface (Fairfax et al. 2004). Cell surface biotinylation has been used to investigate the relative increase in the proportion of GluR1 receptor subunits that reside at the neuronal surface with increasing maturity in spinal cord neurons (Mammen et al. 1997). It is therefore a suitable method for identifying small differences in the amount of glutamate receptor trafficked to the cell surface between TARP-γ5 wild type and its variants, which might not have been apparent with FACS analysis. Additionally, this method can also be used to demonstrate that whether TARP-γ5 is trafficked to the cell surface with GluR2, or is retained intracellularly.
The aim of this chapter was to quantify the total amount of GluR2 and TARP-γ5 at the cell surface and confirms the changes in its trafficking by the co-transfection of TARP-γ5 variants. To quantify the amount of GluR2 and TARP-γ5 at the cell surface, SEP-GluR2 and wild type RFP-TARP or mutated RFP-TARP were co-expressed in HEK293 cells. 24 hours after transfection, the cells were separated into two aliquots. The first was directly lysed and this was used to measure the total cellular amount of GluR2 and TARP-γ5. The second aliquot of transfected cells were used to measure the amount of GluR2 and TARP-γ5 on the cell surface. The cells were subjected to biotinylation prior to lysis which meant that proteins on the surface of the cell could then be purified away from cytoplasmic proteins using streptavidin. The quantification of fluorescent intensity from each protein expression band in each experiment and in the biological replicates was variable. Thus, the data was analysed using a linear mixed model to account for systematic variation. This model helps avoid potentially misleading results due to very different expression of proteins in different experiment or different biological replicates. Our findings may help to implicate the pathogenic importance TARP/AMPA interaction and pave the way for further investigations into the disruption of the excitatory synaptic transmission at the post-synapse in BPD and/or SCZ.

6.2 Results

6.2.1 TARP-γ5 increased trafficking of GluR2 to the cell surface

To determine whether TARP-γ5 regulates cell surface trafficking of GluR2, cell surface biotinylation assays with a red fluorescent protein (RFP)-CACNG5 and pH sensitivity GFP fluorescent (SEP)-GluR2 constructs was performed (See Methods section). After biotin cleavage and western blotting, antibodies recognising RFP and SEP were able to target in GluR2 and
TARP-γ5. This method allowed quantification of both total and cell-surface expression of both GluR2 and TARP-γ5 proteins. Cells were co-transfected with either SEP-GluR2 and control vector or SEP-GluR2 and RFP-TARP-γ5. Co-transfected SEP-GluR2 and control vector cells showed a faint green fluoresce signal on both live cell image and western blot gel analysis (Figure 6-1 A); whereas co-transfected SEP-GluR2 and RFP-TARP-γ5 cells showed both a yellow signal which was as combination of both green and red fluorescent signals (Figure 6-1 B).

When GluR2 was co-expressed with TARP-γ5 protein, there was a 3 fold increase expression (from 260.02 ± 90.65 to 962.5 ± 128.2; (n = 3, p = 0.0029) as shown in Table 6-1. However, this increase was not significant for the cell surface of GluR2 expression (from 335 ± 861.5 to 2788.3 ± 1218.3; (n = 3, p = 0.084). A similar result was reported by Soto et al. (Soto et al. 2009). In this study, the surface expression of GluR2(R) was increased by the presence of TARP-γ5 in HEK293 cells but this change was not statistically significant. Our result suggests that TARP-γ5 may regulate the expression of GluR2 in the HEK293 cells, and may also have a role in influencing cell surface trafficking of GluR2.
Figure 6-1 Quantification of GluR2 in HEK293 when co-transfected with TARP-γ5 or a control plasmid (n = 3). Panel A is live cell imaging of HEK293 cells co-transfected with SEP-GluR2 and control vector, where only green fluorescence appears this demonstrates expression of GluR2 in the cell. Panel B is live cell imaging of HEK293 cells co-transfected with SEP-GluR2 and RFP-TARP-γ5, two cells are shown with red fluorescence that represents TARP-γ5 expression. The cell with the yellow signal represents overlapping expression of TARP-γ5 and GluR2. Lower panel: western blot of cell surface GluR2 isolated after cell surface biotinylation and of total GluR2. Panel C: box plot of western blot results from total cell’s GluR2 expression with and without TARP-γ5 expression (n = 3). Panel D: cell surface biotinylation result of surface GluR2 expression with and with TARP-γ5 expression analysed by linear model regression (n = 3). * P < 0.05; ** P < 0.005.
Table 6-1 Statistical results for GluR2 expression from western blot and biotinylation assay. Total represents total GluR2 in the cells; surface represents biotinylated GluR2 as surface GluR2. Medium represents the mean fluorescent intensity of expression; SEM = standard error of medium; R-square = linear regression of GluR2 expression; p-value is a probability under the condition that the null hypothesis is false, where the null hypothesis is “TARP-γ5 has no effect on GluR2 expression.

<table>
<thead>
<tr>
<th>Co-transfected</th>
<th>Total</th>
<th>Surface</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Medium</td>
<td>SEM</td>
</tr>
<tr>
<td>R2</td>
<td>260.02</td>
<td>90.65</td>
</tr>
<tr>
<td>R2 + γ5</td>
<td>962.5</td>
<td>128.2</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.9137</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.002878</td>
<td></td>
</tr>
</tbody>
</table>

The expression of GluR2 and its trafficking in the cells appeared to be regulated by the presence of TARP-γ5; however, the underlying mechanism was still unclear. For instance, the question of whether TARP-γ5 traffics together with GluR2 or only assists it intracellularly was still unanswered. To address this question the expression of TARP-γ5 in the cells and on the cell surface was quantified via western blot with the combination of cell surface biotinylation assays (Figure 6-2). The result showed that more than half of TARP-γ5 in the cells was trafficked to the cell surface (see Table 6-2). This finding suggests that approximately half of the TARP-γ5 is trafficked to the cell surface along with GluR2, but that some is also retained intracellularly.
Figure 6-2 Quantification of TARP-γ5 in HEK293 cells when co-transfected with GluR2 (n = 3). Panel A: Total TARP-γ5 in HEK293 cells was quantified via fluorescence intensity (R2 + g5 represents co-transfected GluR2 and TARP-γ5; surface represents only surface proteins in the cells; whereas total represents as total proteins in the cells). Panel B: western blot of cell surface TARP-γ5 isolated by biotinylation and of total TARP-γ5.

Table 6-2 Statistical results for TARP-γ5 expression in the cells with or without cell surface biotinylation assay (Total represents all cellular proteins; surface represents cell surface proteins detected using the cell surface biotinylation assay). Medium represents as mean fluorescent intensity of expression; SEM = standard error of medium; R2 = linear regression of TARP-γ5 expression; p-value is a probability under the condition that the null hypothesis is false, where the null hypothesis is “GluR2 has no effect on trafficking TARP-γ5 to cell surface.

<table>
<thead>
<tr>
<th>Co-transfected</th>
<th>R2 + γ5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Medium</td>
</tr>
<tr>
<td>Total</td>
<td>1174.4</td>
</tr>
<tr>
<td>Surface</td>
<td>688.7</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.2556</td>
</tr>
<tr>
<td>p-value</td>
<td>0.3063</td>
</tr>
</tbody>
</table>
6.2.2 Rare TARP-γ5 variant R69W showed no significant effect on GluR2 expression

Figure 6-3 Quantification of GluR2 in HEK293 cells when co-transfected with wild-type TARP-γ5 or mutated TARP-γ5 (n = 3). Panel A is live cell imaging of co-transfected HEK293 cells with GluR2 and p.Arg69Trp mutated TARP-γ5 (R69W), the red – orange fluorescent signal indicates that there is little GluR2 expression on the cell surface. Panel B is live cell imaging of co-transfected HEK293 cells with GluR2 and wild type TARP-γ5. The cells expressed both red fluorescent signal and yellow signal, which indicates that the cells express both GluR2 and TARP-γ5 proteins on their surface. Panel C: western blot of cell surface GluR2 isolated by biotinylation and of total GluR2. Panel D: box plot from western blot and cell surface biotinylation assays for GluR2 expression (n = 3).
Table 6-3 Statistical results for GluR2 expression from western blot and cell surface biotinylation assay. Total represents total GluR2 in the cells; surface represents cell surface biotinylated GluR2. Medium represents as mean fluorescent intensity of expression; SEM = standard error of medium; R2 = linear regression of GluR2 expression, p-value is a probability under the condition that the null hypothesis is false, where the null hypothesis is “R69W variant has no effect GluR2 expression in the total or surface cells.

<table>
<thead>
<tr>
<th>Co-transfected</th>
<th>AMPA-R2 Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
</tr>
<tr>
<td>Wild-type</td>
<td>1003</td>
</tr>
<tr>
<td>R69W</td>
<td>1525.7</td>
</tr>
<tr>
<td>Adjusted R2</td>
<td>0.126</td>
</tr>
<tr>
<td>p-value</td>
<td>0.49</td>
</tr>
</tbody>
</table>

The Arg69Trp (R69W) variant was identified in one SCZ patient (Table 4-2). It is located in the first extracellular loop of the TARP-γ5 protein. To investigate whether this variant affects surface trafficking or expression of GluR2, the variant was introduced into a wild-type TARP-γ5 and then co-transfected with GluR2 in HEK293 cells. The western blot studies showed that R69W was no significant effect on the expression of GluR2 in the cells; however, a slightly decrease in the cell surface expression of GluR2 was observed (Figure 6-3 D). Live cell imaging also demonstrated reduced GFP signal in the R69W transfected compared to the wild-type transfected cells (Figure 6-3 A and B). This result could implicate the first extracellular loop of TARP-γ5 as a possible regulatory domain for GluR2 trafficking to the cell surface.
Figure 6-4 Comparison of wild type TARP-γ5 and its carried R69W variant on the their expression in HEK293 cells (n = 3). Panel A: Total wild type TARP-γ5 and its carried R69W variant in HEK293 cells were quantified via fluorescence intensity (SNP69 represents co-transfected GluR2 and TARP-γ5 carried R69W variant; wild-type represents as wild-type TARP-γ5; surface represents only surface TARP-γ5 in the cells; total represents as total proteins in the cells). Panel B: western blot of cell surface TARP-γ5 isolated by biotinylation and of total TARP-γ5 in either co-transfected with GluR2 and wild-type TARP-γ5 or TARP-γ5 with R69W variant.
Table 6-4 Statistical results for TARP-\(\gamma 5\) expression in the cells with or without cell surface biotinylation assay (Total represents as without surface biotinylation assay; surface represents as with biotinylation assay). Medium represents as mean fluorescent intensity of expression; SEM = standard error of medium; \(R2\) = linear regression of TARP-\(\gamma 5\) expression; \(p\)-value is a probability under the condition that the null hypothesis is false, where the null hypothesis “R69W variant has no effect total or surface expression of TARP-\(\gamma 5\).

<table>
<thead>
<tr>
<th>Co-transfected</th>
<th>TARP-(\gamma 5) Expression</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Surface</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>SEM</td>
<td>Medium</td>
</tr>
<tr>
<td>Wild-type</td>
<td>1308.7</td>
<td>512.8</td>
<td>836.3</td>
</tr>
<tr>
<td>R69W</td>
<td>1174.4</td>
<td>725.3</td>
<td>984</td>
</tr>
<tr>
<td>Adjusted R2</td>
<td>0.008504</td>
<td></td>
<td>0.03978</td>
</tr>
<tr>
<td>(p)-value</td>
<td>0.8621</td>
<td></td>
<td>0.7048</td>
</tr>
</tbody>
</table>

Furthermore, both TARP-\(\gamma 5\) surface and total expression showed no effect in the present of R69W variant (Figure 6-4). Hence, R69W variant does not affect expression of TARP-\(\gamma 5\) in the cells, but slightly reduced the trafficking of GluR2 to the cell surface (Figure 6-4).
6.2.3 Rare variant R71H in TARP-γ5 showed no significant effect on GluR2 expression

Figure 6-5 Quantification of GluR2 in HEK293 cells when co-transfected with wild-type TARP-γ5 or mutated TARP-γ5 (n = 3). Panel A is live cell imaging of co-transfected HEK293 cells with GluR2 and wild type TARP-γ5, green fluorescent represents GluR2 protein, red fluorescent represents TARP-γ5, and yellow represents overlapping these two proteins. Panel B is live cell imaging of co-transfected HEK293 cells with GluR2 and p.Arg71His mutated TARP-γ5 (R71H). Panel C: western blot of cell surface GluR2 isolated by biotinylation and of total GluR2. Panel D: pooled data from western blot and biotinylation assay for GluR2 expression (n = 3).
Table 6-5 Statistical results for GluR2 expression in the cells with or without cell surface biotinylation assay (Total represents as without surface biotinylation assay; surface represents as with biotinylation assay). Medium represents as mean fluorescent intensity of expression; $\text{SEM} = \text{standard error of medium}; R^2 = \text{linear regression of GluR2 expression}; \text{p-value is a probability under the condition that the null hypothesis is false, where the null hypothesis is “R71H has no effect on the total or surface expression of GluR2.}$

<table>
<thead>
<tr>
<th>Co-transfected</th>
<th>AMPA-R2 Expression</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Surface</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>$\text{SEM}$</td>
<td>Medium</td>
</tr>
<tr>
<td>Wild-type</td>
<td>869.7</td>
<td>287.1</td>
<td>1940</td>
</tr>
<tr>
<td>R71H</td>
<td>735.7</td>
<td>203</td>
<td>756.7</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.05164</td>
<td>0.175</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.665</td>
<td>0.4091</td>
<td></td>
</tr>
</tbody>
</table>

p.Arg71His (R71H) was identified in one bipolar disorder patient (Table 4-2). Our functional studies result showed that there was no significantly decrease in total or cell surface of GluR2 expression (Figure 6-5). Both R69W and R71H variants are located at the first long extracellular loop of TARP-γ5 protein. Structural and functional studies of TARP-γ2/stargazin suggest that this long extracellular loop is an essential functional domain for the channel properties modulation and also the binding site for AMPAR ligand binding domain (Tomita et al. 2005a) (Turetsky et al. 2005). The substitution of this domain of stargzin/γ2 with that from TARP-γ4 increased glutamate affinity and reduced kinetics of deactivation and desensitisation (Cho et al. 2007, Milstein et al. 2007). Although these variants had no effect on the trafficking or expression of GluR2 in the cells, it could possibly effect on the kinetics of receptor and binding affinity of glutamate.
Figure 6-6 Comparison of wild type TARP-γ5 and its carried R71H variant on their expression in HEK293 cells (n = 3). Panel A: Total wild type TARP-γ5 and the R71H variant in HEK293 cells were quantified via fluorescence intensity (SNP71 represents cells co-transfected GluR2 and TARP-γ5 with the R71H variant; wild-type represents wild-type TARP-γ5; surface represents cell surface TARP-γ5; total represents as total proteins in the cells). Panel B: western blot of cell surface TARP-γ5 isolated by biotinylation and of total TARP-γ5 in either co-transfected with GluR2 and wild-type TARP-γ5 or TARP-γ5 with R71H variant.
Table 6-6 Statistical results for TARP-γ5 expression in the cells with or without cell surface biotinylation assay (Total represents as without surface biotinylation assay; surface represents as with biotinylation assay). Medium represents as mean fluorescent intensity of expression; SEM = standard error of medium; R2 = linear regression of TARP-γ5 expression; p-value is a probability under the condition that the null hypothesis is false, where the null hypothesis is “R71H variant has no effect on the total or surface expression of TARP-γ5.

<table>
<thead>
<tr>
<th>Co-transfected</th>
<th>TARP-γ5 Expression</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Surface</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>SEM</td>
<td>Medium</td>
</tr>
<tr>
<td>Wild-type</td>
<td>1174.3</td>
<td>535.2</td>
<td>511.3</td>
</tr>
<tr>
<td>R71H</td>
<td>1500</td>
<td>378.4</td>
<td>688.7</td>
</tr>
<tr>
<td>Adjusted R2</td>
<td>0.08474</td>
<td></td>
<td>0.08883</td>
</tr>
<tr>
<td>p-value</td>
<td>0.5757</td>
<td></td>
<td>0.5662</td>
</tr>
</tbody>
</table>
6.2.4 Rare variant R127Q in TARP-γ5 showed no significant effect on GluR2 expression

Figure 6-7 Quantification of GluR2 in HEK293 cells when co-transfected with wild-type TARP-γ5 or mutated TARP-γ5 (n = 3). Panel A is live cell imaging of HEK293 cells with co-transfected GluR2 and wild type TARP-γ5, green fluorescent represents GluR2 protein, red fluorescent represents TARP-γ5, and yellow represents overlap between these two proteins. Panel B is live cell imaging of co-transfected HEK293 cells with GluR2 and TARP-γ5 p.Arg127Gln mutated (R7127Q). Panel C: western blot of cell surface GluR2 isolated by biotinylation and of total GluR2. Panel D: pooled data from western blot and cell surface biotinylation assay for GluR2 expression (n = 3).
Table 6-7 Statistical analysis of GluR2 expression in HEK293 cells with or without cell surface biotinylation before assay (Total represents as without surface biotinylation assay; surface represents as with biotinylation assay). Medium represents as mean fluorescent intensity of expression; SEM = standard error of medium;; R2 = linear regression of GluR2 expression; p-value is a probability under the condition that the null hypothesis is false, where the null hypothesis is “R127H variant has no effect the total or surface expression of GluR2.

<table>
<thead>
<tr>
<th>Co-transfected</th>
<th>AMPA-R2 Expression</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>Surface</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Medium</td>
<td>SEM</td>
<td>Medium</td>
<td>SEM</td>
<td>Medium</td>
<td>SEM</td>
</tr>
<tr>
<td>Wild-type</td>
<td>916.7</td>
<td>163.898</td>
<td>2045.3</td>
<td>1230.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R127H</td>
<td>1002</td>
<td>115.893</td>
<td>1078</td>
<td>870.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted R2</td>
<td>2.585 x 10^-5</td>
<td>0.4085</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.9924</td>
<td>0.1718</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The p.Arg127Gln (R127Q) variant was found in two bipolar disorder patients and in one person from European 1000 Genome project control (1K Genome project) (Table 4-2). According to our previous bioinformatic analysis, R127Q was predicted to have a benign effect on protein function. The green fluorescent intensity from our FACS analysis showed a reduction of both surface and total percentage of cells expressed GluR2 (Figure 5-3 and 5-4). A similar result also revealed from our biotinylation and western blot assay, where the GluR2 expression in the cells was showed no effect, but the surface GluR2 expression showed reduced in the present of R127Q variant (Figure 6-7 D). From the live cell image, the visible green fluorescent signal from the R127Q variant co-transfected cells was lower than the wild-type co-transfected cells (Figure 6-7 A and B). Although the statistical analysis showed that the R127Q variant has effect, it is still possible that it may have an effect on the trafficking of GluR2. Further experiments may reveal the specific effect on this variant.
Figure 6-8 Comparison of wild type TARP-γ5 and its carried R127Q variant on the their expression in HEK293 cells (n = 3). Panel A: Total wild type TARP-γ5 and its carried R127Q variant in HEK293 cells were quantified via fluorescence intensity (SNP127 represents co-transfected GluR2 and TARP-γ5 carried R127Q variant; wild-type represents as wild-type TARP-γ5; surface represents only surface TARP-γ5 in the cells; total represents as total proteins in the cells). Panel B: western blot of cell surface TARP-γ5 isolated by biotinylation and of total TARP-γ5 in either co-transfected with GluR2 and wild-type TARP-γ5 or TARP-γ5 with R127Q variant.
Table 6-8 Statistical results for TARP-γ5 expression in the cells with or without cell surface biotinylation assay (Total represents as without surface biotinylation assay; surface represents as with biotinylation assay). Medium represents as mean fluorescent intensity of expression; SEM = standard error of medium; R2 = linear regression of TARP-γ5 expression; p-value is a probability under the condition that the null hypothesis is false, where the null hypothesis is “R127 has no effect on the total or surface TARP-γ5 expression”.

<table>
<thead>
<tr>
<th>Co-transfected</th>
<th>TARP-γ5 Expression</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Surface</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>SEM</td>
<td>Medium</td>
</tr>
<tr>
<td>Wild-type</td>
<td>1174.3</td>
<td>285.2</td>
<td>688</td>
</tr>
<tr>
<td>R127H</td>
<td>898</td>
<td>208.7</td>
<td>594.7</td>
</tr>
<tr>
<td>Adjusted R2</td>
<td>0.1797</td>
<td></td>
<td>0.0171</td>
</tr>
<tr>
<td>p-value</td>
<td>0.4022</td>
<td></td>
<td>0.805</td>
</tr>
</tbody>
</table>

The R127Q variant only appeared to have a slight effect on GluR2 cell surface expression and therefore it was of interest to know whether to this variant could have an effect on its own expression or trafficking. Our result revealed that this variant did not have significant effect on its own protein expression, or variant located in one of the loops of TARp-γ5 only have a modest effect on GluR2 trafficking to the cell surface (Table 6-8). Thus, this intracellular variant only has a possibility effect on GluR2 trafficking.
6.2.5 Rare variant R128M in TARP-γ5 slightly reduced GluR2 expression

Figure 6-9 Quantification of GluR2 in HEK293 cells when co-transfected with wild-type TARP-γ5 or mutated TARP-γ5 (n = 3). Panel A is live cell imaging of co-transfected HEK293 cells with GluR2 and wild type TARP-γ5, green fluorescent represents GluR2 protein, red fluorescent represents TARP-γ5, and yellow represents overlapping these two proteins. Panel B is live cell imaging of co-transfected HEK293 cells with GluR2 and p.Thr128Met mutated TARP-γ5 (R128M). Panel C: western blot of cell surface GluR2 isolated by biotinylation and of total GluR2. Panel D: pooled data from western blot and biotinylation assay for GluR2 expression (n = 3).
Table 6-9 Statistical results for GluR2 expression in the cells with or without cell surface biotinylation assay (Total represents as without surface biotinylation assay; surface represents as with biotinylation assay). Medium represents as mean fluorescent intensity of expression; SEM = standard error of medium; R2 = linear regression of GluR2 expression; p-value is a probability under the condition that the null hypothesis is false, where the null hypothesis is “T128M has no effect on the total or surface expression of GluR2”.

<table>
<thead>
<tr>
<th>Co-transfected</th>
<th>AMPA-R2 Expression</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>Surface</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Medium</td>
<td>SEM</td>
</tr>
<tr>
<td>Wild-type</td>
<td>1003.7</td>
<td>200.6</td>
<td>2215.3</td>
</tr>
<tr>
<td>T128M</td>
<td>512</td>
<td>141.9</td>
<td>908</td>
</tr>
<tr>
<td>Adjusted R2</td>
<td>0.6002</td>
<td></td>
<td>0.4543</td>
</tr>
<tr>
<td>p-value</td>
<td>0.074</td>
<td></td>
<td>0.1421</td>
</tr>
</tbody>
</table>

We then investigated whether the p.Thr128Met variant, which lies adjacent to the R127Q variant in the first intracellular loop of TARP-γ5 affects GluR2 trafficking and its expression in co-transfected HEK293 cells. This variant was found in three bipolar disorder patients and two in UCL control samples (Table 4-2). Previous Polyphen2 bioinformatic analysis predicted it was probably damaging to protein function; whereas the SIFT analysis revealed this variant to be tolerated. The percentage of cells expressing surface GluR2 in co-transfected variant cells was reduced compared to wild-type TARP-γ5 expressed cells, and so does the total GluR2 expressed in the cells, but without statistical significant (Figure 5-3 and 5-4). In addition, the western blot and cell surface analysis showed decreased total GluR2 expression in the cells and on the cell surface when co-transfected with R128M variant. These results were not statistically significant, but the result with total GluR2 expression was so close to $p < 0.05$ (Table 6-9).
Figure 6-10 Comparison of wild type TARP-γ5 and its carried R128M variant on the their expression in HEK293 cells (n = 3). Panel A: Total wild type TARP-γ5 and its carried R128M variant in HEK293 cells were quantified via fluorescence intensity (SNP128 represents co-transfected GluR2 and TARP-γ5 carried R128M variant; wild-type represents as wild-type TARP-γ5; surface represents only surface TARP-γ5 in the cells; total represents as total proteins in the cells). Panel B: western blot of cell surface TARP-γ5 isolated by biotinylation and of total TARP-γ5 in either co-transfected with GluR2 and wild-type TARP-γ5 or TARP-γ5 with R128M variant.
Table 6-10 Statistical results for TARP-γ5 expression in the cells with or without cell surface biotinylation assay (Total represents as without surface biotinylation assay; surface represents as with biotinylation assay). Medium represents as mean fluorescent intensity of expression; SEM = standard error of medium; R2 = linear regression of TARP-γ5 expression; p-value is a probability under the condition that the null hypothesis is false, where the null hypothesis is “T128M has no effect on the total or surface expression of TARP-γ5”

<table>
<thead>
<tr>
<th>Co-transfected</th>
<th>TARP-γ5</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td></td>
<td>Surface</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>SEM</td>
<td>Medium</td>
<td>SEM</td>
<td></td>
</tr>
<tr>
<td>Wild-type</td>
<td>1174.3</td>
<td>244.1</td>
<td>688.6</td>
<td>356.6</td>
<td></td>
</tr>
<tr>
<td>T128M</td>
<td>633</td>
<td>172.6</td>
<td>390.3</td>
<td>252.1</td>
<td></td>
</tr>
<tr>
<td>Adjusted R2</td>
<td>0.5515</td>
<td></td>
<td>0.1489</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.0982</td>
<td></td>
<td>0.4498</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The presence of the T128M variant in the co-transfected cells appeared decrease to lead to a modest in the expression and cell surface trafficking of GluR2 in the cells. It was therefore important to know whether this variant affects its own expression and cell surface trafficking. Interestingly, the biotinylation assays showed that the presence of this variant actually also affect on its own expression, and the statistical result was much closed to the p < 0.05. By contrast, the trafficking of TARP-γ5 was not significantly much affected by this variant. One of the possible explanation of reducing expression and trafficking of GluR2 in the cells is due to the reduction of TARP-γ5 expression in the cells. The reduction of TARP-γ5 expression decreased the expression of GluR2; it suggests that TARP-γ5 may have a regulation role in intracellular GluR2 assembly.
6.2.6 Rare variant I156F in TARP-γ5 significantly reduced GluR2 expression

Figure 6-11 Quantification of GluR2 in HEK293 cells when co-transfected with wild-type TARP-γ5 or mutated TARP-γ5 (n = 3). Panel A is live cell imaging of co-transfected HEK293 cells with GluR2 and wild type TARP-γ5, green fluorescent represents GluR2 protein, red fluorescent represents TARP-γ5, and yellow represents overlapping these two proteins. Panel B is live cell imaging of co-transfected HEK293 cells with GluR2 and p.Ile156Phe mutated TARP-γ5 (I156F). Panel C: western blot of cell surface GluR2 isolated by biotinylation and of total GluR2. Panel D: pooled data from western blot and biotinylation assay for GluR2 expression (n = 3). * P < 0.05.
Table 6-11 Statistical results for GluR2 expression in the cells with or without cell surface biotinylation assay (Total represents as without surface biotinylation assay; surface represents as with biotinylation assay). Medium represents as mean fluorescent intensity of expression; SEM = standard error of medium;; R2 = linear regression of GluR2 expression; p-value is a probability under the condition that the null hypothesis is false, where the null hypothesis is “I156F variant has no effect on the cell total or surface expression of GluR2”.

<table>
<thead>
<tr>
<th>Co-transfected</th>
<th>AMPA-R2 Expression</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Surface</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>SEM</td>
<td>Medium</td>
</tr>
<tr>
<td>Wild-type</td>
<td>1003.7</td>
<td>176.2</td>
<td>3123.7</td>
</tr>
<tr>
<td>I156F</td>
<td>501</td>
<td>124.6</td>
<td>375</td>
</tr>
<tr>
<td>Adjusted R2</td>
<td>0.6704</td>
<td>0.5625</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.04627</td>
<td>0.08593</td>
<td></td>
</tr>
</tbody>
</table>

Variant p.Ile156Phe (I156F) is one of the more interesting variants. This variant that was discovered found in one bipolar disorder patient. The I156F variant showed significantly increased surface expression of GluR2 compared to the wild type in our FACS analysis (Figure 5-6). However, the biotinylation assay revealed an opposite effect, where the expression and trafficking of GluR2 were decreased by the presence of this variant (Figure 6-11); specifically, the expression of GluR2 in cells was decreased in the presence of the I156F variant (from 1003.7 ± 176.2 to 501 ± 124.6, n = 3, p = 0.046) (Table 6-11). Both live cell image and western blot analysis showed a lower signal of GluR2 compared to the wild type. Although the different assays were not in argument, it still showed a statistical significant effect on the expression of GluR2 by the presence of I156F variant. Therefore, a single amino acid substitution in second extracellular loop of TARP-γ5 may affect its functional domain and the binding affinity to GluR2 protein; however, further experiments are required to explain this mechanism.
Figure 6-12 Comparison of wild type TARP-γ5 and the I156F variant on expression in HEK293 cells (n = 3). Panel A: Total wild type TARP-γ5 and the TARP-γ5 I156F variant in HEK293 cells were quantified via fluorescence intensity (SNP156 represents co-transfected GluR2 and TARP-γ5 carried I156F variant; wild-type represents as wild-type TARP-γ5; surface represents surface TARP-γ5 expression; total represents total TARP-γ5 in the cells). Panel B: western blot of cell surface TARP-γ5 isolated by biotinylation and of total TARP-γ5 in either co-transfected with GluR2 and wild-type TARP-γ5 or TARP-γ5 with the I156F variant.
Table 6-12 Statistical results for TARP-γ5 expression in the cells with or without cell surface biotinylation assay (Total represents as without surface biotinylation assay; surface represents as with biotinylation assay). Medium represents as mean fluorescent intensity of expression; SEM = standard error of medium; R2 = linear regression of TARP-γ5 expression; p-value is a probability under the condition that the null hypothesis is false, where the null hypothesis is “I156F has no effect on the total or surface expression of TARP-γ5”.

<table>
<thead>
<tr>
<th>Co-transfected</th>
<th>TARP-γ5 Expression</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>Surface</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Medium</td>
<td>SEM</td>
<td>Medium</td>
<td>SEM</td>
</tr>
<tr>
<td>Wild-type</td>
<td>1174</td>
<td>301.2</td>
<td>688.7</td>
<td>365.3</td>
<td></td>
</tr>
<tr>
<td>I156F</td>
<td>475.3</td>
<td>213</td>
<td>271</td>
<td>258.3</td>
<td></td>
</tr>
<tr>
<td>Adjusted R2</td>
<td>0.5739</td>
<td></td>
<td>0.2463</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.08105</td>
<td></td>
<td>0.3167</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The result of biotinylation assay on TARP-γ5 expression showed that in the presence of the I156F variant, the expression and trafficking of TARP-γ5 may also have been reduced. However this effect was not statistically significant. This reduction of TARP-γ5 expression and trafficking could actually affect on the rate of GluR2 assembly in the cells, and also their cell surface trafficking.
6.2.7 Rare variant T164L in TARP-γ5 significantly reduced GluR2 expression

Figure 6-13 Quantification of GluR2 in HEK293 cells when co-transfected with wild-type TARP-γ5 or mutated TARP-γ5 (n = 3). Panel A is live cell imaging of co-transfected HEK293 cells with GluR2 and wild type TARP-γ5, green fluorescent represents GluR2 protein, red fluorescent represents TARP-γ5, and yellow represents overlapping these two proteins. Panel B is live cell imaging of co-transfected HEK293 cells with GluR2 and p.Thr164Leu mutated TARP-γ5 (T164L). Panel C: western blot of cell surface GluR2 isolated by biotinylation and of total GluR2. Panel D: pooled data from western blot and biotinylation assay for GluR2 expression (n = 3). * P < 0.05.
Table 6-13 Statistical results for GluR2 expression in the cells with or without cell surface biotinylation assay (Total represents as without surface biotinylation assay; surface represents as with biotinylation assay). Medium represents as mean fluorescent intensity of expression; SEM = standard error of medium; R2 = linear regression of GluR2 expression; p-value is a probability under the condition that the null hypothesis is false, where the null hypothesis is “T164L has no effect on the total or surface expression of GluR2”.

<table>
<thead>
<tr>
<th>Co-transfected</th>
<th>AMPA-R2 Expression</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Surface</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>SEM</td>
<td>Medium</td>
</tr>
<tr>
<td>Wild-type</td>
<td>1003.67</td>
<td>132.14</td>
<td>2638.7</td>
</tr>
<tr>
<td>T164L</td>
<td>285</td>
<td>93.44</td>
<td>484.7</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.8809</td>
<td></td>
<td>0.5431</td>
</tr>
<tr>
<td>p-value</td>
<td>0.005548</td>
<td></td>
<td>0.0947</td>
</tr>
</tbody>
</table>

TARP-γ5 variant p.Thr164Leu was found in one schizophrenia patient (Table 4-2). According to our FACS analysis, the presence of this variant in the TARP-γ5 protein led to a non-significant reduction in GluR2 expression in the cells. A further study with western blot analysis combined with cell surface biotinylation assays also showed a decreased expression of GluR2 in the cells from $1003.67 \pm 132.14$ to $285 \pm 93.44$ (n = 3, $p = 0.0055$) (Table 6-13). There was also a trend for decreased GluR2 trafficking to the cell surface was also showed when the cells were co-transfected with this variant (Table 6-13). This variant was the most interesting finding from our experiments, and it showed a reduction of GluR2 expression and trafficking with two different experimental approaches. Additionally, the live cell image also showed red signal in most of the co-transfected cells (Figure 6-13 B).
Figure 6-14 Comparison of wild type TARP-γ5 and its carried T164L variant on the their expression in HEK293 cells (n = 3). Panel A: Total wild type TARP-γ5 and its carried T164L variant in HEK293 cells were quantified via fluorescence intensity (SNP164 represents co-transfected GluR2 and TARP-γ5 carried T164L variant; wild-type represents as wild-type TARP-γ5; surface represents only surface TARP-γ5 in the cells; total represents as total proteins in the cells). Panel B: western blot of cell surface TARP-γ5 isolated by biotinylation and of total TARP-γ5 in either co-transfected with GluR2 and wild-type TARP-γ5 or TARP-γ5 with T164L variant.
Table 6-14 Statistical results for TARP-\(\gamma\)5 expression in the cells with or without cell surface biotinylation assay (Total represents as without surface biotinylation assay; surface represents as with biotinylation assay). Medium represents as mean fluorescent intensity of expression; SEM = standard error of medium; R2 = linear regression of TARP-\(\gamma\)5 expression; p-value is a probability under the condition that the null hypothesis is false, where the null hypothesis is “T164L has no effect on the total or surface expression of TARP-\(\gamma\)5”.

<table>
<thead>
<tr>
<th>Co-transfected</th>
<th>TARP-(\gamma)5 Expression</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Surface</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>SEM</td>
<td>Medium</td>
</tr>
<tr>
<td>Wild-type</td>
<td>1174.3</td>
<td>293.5</td>
<td>688.7</td>
</tr>
<tr>
<td>T164L</td>
<td>663.3</td>
<td>207.5</td>
<td>435</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.4312</td>
<td></td>
<td>0.116</td>
</tr>
<tr>
<td>p-value</td>
<td>0.1566</td>
<td></td>
<td>0.5089</td>
</tr>
</tbody>
</table>

Although the expression of GluR2 in the cells were significantly reduced in the co-transfected T164L variant of TARP-\(\gamma\)5, the expression of TARP-\(\gamma\)5 was not significantly influenced by the presence of this variant. In particular, the surface expression of TARP-\(\gamma\)5 in the presence of this variant led to approximated equal expression to the wild-type (Table 6-14). Hence, approximately 1.5 fold reduction on the surface expression of GluR2 was not changed by altered TARP-\(\gamma\)5 in the cells. This result suggests that an amino acid substitution at the second extracellular loop of TARP-\(\gamma\)5 would possibly affect the binding affinity of GluR2, where leading to the reduction of their trafficking.
6.3 Discussion

Table 6-15 Summary of TARP-γ5 and its variant effect results on GluR2 and TARP-γ5 expression. – represents as no effect; ↑ represents modest increased expression; ↓ represents modest decreased expression; ↓↓ represents significantly decreased expression, p < 0.05; ↓↓↓ represents significantly decreased expression, p < 0.005.

<table>
<thead>
<tr>
<th></th>
<th>GluR2</th>
<th>TARP-γ5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Surface</td>
</tr>
<tr>
<td>TARP-γ5</td>
<td>↑↑↑</td>
<td>↑</td>
</tr>
<tr>
<td>R69W</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>R71H</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>R127Q</td>
<td>-</td>
<td>↓</td>
</tr>
<tr>
<td>T128M</td>
<td>↓</td>
<td>↓</td>
</tr>
<tr>
<td>I156F</td>
<td>↓↓</td>
<td>↓</td>
</tr>
<tr>
<td>T164L</td>
<td>↓↓</td>
<td>↓</td>
</tr>
</tbody>
</table>

The data presented here provide evidence that TARP-γ5 mediates AMPAR2 trafficking and its expression. This confirms and extends results from one recent study reporting that TARP-γ5 decreases both cell surface trafficking of unedited GluR2 short and long isoform, and increased the trafficking of edited GluR2 receptor, known as Ca^{2+} impermeable receptors (Soto et al. 2009). Although the statistical significant alteration was not revealed in the trafficking of edited GluR2 receptor, the similar result trend has also been showed in our preliminary experiment (Figure 5-3). One of the explanations is that the three biological replications and each expression were variable.

R69W and R71H are variants that located at the first extracellular loop of TARP-γ5, and were found in one schizophrenia and one bipolar disorder patient, respectively (Table 4-2). Molecular dissection of stargazin/γ2 identified that the first extracellular loop controls channel properties and glutamate affinity (Cho et al. 2007, Milstein et al. 2007, Osler et al. 2005, Tomita et al. 2005a). In addition, our biotinylation assay showed no effect on both expression and trafficking.
of GluR2 when cells were co-expressed with R69W and R71H individually. Furthermore, the green fluorescent intensity from surface sensitivity tag protein revealed a decreased level, but without statistical significant. This result suggests that the first extracellular loop domain does mediate GluR2 trafficking; hence, further electrophysiological experiments are required to investigate of this variant.

R127H and T128M were found only in bipolar disorder patients, and also in control subjects (Table 4-2). These two variants are located at the short intracellular loop and close to transmembrane domain III. Bioinformatic analysis of these two variants has revealed a benign effect on protein function; the green fluorescent intensity from surface GluR2 showed a trend for decreased trend. Here, our biotinylation assay showed that both expression and surface GluR2 were decreased when co-expressed with these two variants individually, but also without any statistical significant. Although the explanation could be variable on each expression level, the expression of TARP-γ5 was reduced when co-expressed with T128M variant. Therefore, one possible reduction reason could be due to TARP-γ5 expression reduced in the cells. Furthermore, literature has suggested any function of this region to mediate AMPAR channel properties or trafficking. Hence, these two variants are also confirmed to have no effect on protein function.

Another two rare variants I156F and T164L are found in one bipolar disorder and one schizophrenia, respectively. Both residues are located in the second extracellular loop, which is thought to mediate changes in AMPAR channel properties and its binding affinity (Tomita et al. 2005a). Our FACS data has showed that co-expression I156F variant in the HEK293 cells, surface GluR2 expression was significantly enhanced; conversely, the biotinylated surface GluR2 expression showed a decreased trend, while the TARP-γ5 expression in the cells was also reduced. These results suggest that the expression level in the cells and on the cell surface is
regulated by the amount of TARP-γ5 expressed in the cells. The most promising characteristic variant is T164L, which has not occurred in any other cohort. Previously green fluorescent intensity result revealed a reduction of surface GluR2 expression when cells were co-expressed with T164L variant. Here, both total GluR2 and biotinylated GluR2 were decreased when cells were co-transfected with T164L variant, particular significantly reduced in the total GluR2 expression (Table 6-15). This reduction from the present of T164L was not affected by the expression levels of TARP-γ5.

Analysis of mutants as a single amino acid substitution variants shows that alters the total expression of AMPAR2 and also its surface expression. Co-expressed wild-type TARP-γ5 increased the expression of AMPAR2; whereas, the specific amino acid substitution, T164L decreased AMPAR2 expression. This evidence suggests that a single rare variants found in TARP-γ5 have an effect on AMPAR2 assembly or trafficking in the cells. In term this dysregulation of AMPAR2 function may increase susceptibility to a variant of neuropsychiatric conditions.
7 Conclusion

According to the Kraepelinian dichotomy bipolar disorder (BPD) and schizophrenia (SCZ) are traditionally regarded as two separate disorders. However, there is increasing evidence to support the notion that they share common neurobiological abnormalities (Domjan et al. 2012). Clinical genetic studies imply a familiar aggregation and considerable heritability of BPD shared with SCZ (Lichtenstein et al. 2009). As such, it would be natural to hypothesize that multiple risk variants may be involved in the aetiology of both diseases. Despite advancing study design and technology, the underlying genetic and pathophysiological components of the disorders remain unclear. Intensive research in the field of shared variants in BPD and SCZ, including numerous family linkage studies, locus specific association studies, and GWAS, has demonstrated susceptibility loci and candidate genes that may confer risk to both BPD and SCZ.

In this study, the CACNG5 gene was screened for all variants that may be associated with BPD and/or SCZ in the UCL BPD and/or SCZ patient and control sample. This was followed by genotyping the identified CACNG5 variants to confirm allele frequencies from the gene scanning method. These variants were then tested for association with BPD and/or SCZ. The genetic study revealed four novel rare non-synonymous (SNPs) (nsSNPs), three rare nsSNPs, and one more common nsSNP. It provided evidence for the efficiency of the gene scanning method and demonstrated that HRMA has a high sensitivity and specificity.

Burden analysis of all nsSNPs in UCL BPD versus UCL controls found evidence for weak association ($p = 0.0127$); whereas UCL SCZ versus UCL controls found moderately associated ($p = 0.00158$). Combined both UCL cases versus UCL controls did not strengthen the association signal ($p = 0.0022$). Combined analysis including the European samples from the 1000 Genomes
project strengthened the association signal (p = 0.00057). The larger combined analysis that included the UK10K data reduced the association signal to (p = 0.0034); moreover combined analysis that included the Swedish exome sequence data also did not strengthen the association signal (p = 0.0051). Analysis of data from all controls versus all cases found a weakened association signal (p = 0.0082). These results have revealed that the association signal was stronger in UCL samples, perhaps due to the healthy super-normal controls that were included.

In fact, our genetic study showed a different result compared to the original case-case study (Curtis, 2011 #810), where the association result implicated that the rs17645023 SNP is a disorder specific variant, but our current result showed multiple coding variants moderately associated with SCZ and even less in BPD. It is possible to speculate that CACNG5 may be more involved in susceptibility to SCZ gene and that CACNG4 may be more involved in susceptibility to BPD and that this may be the explanation for the finding with SNP rs17645023, that was found to be associated in case versus case analysis, which is located between these two genes. Because of this possible implication, one of our colleagues has been working on sequencing and genotyping of CACNG4 gene to identify it is a possible explanation. Furthermore, as our association signal was modest, this study is probably underpowered to detect significant association between individual CACNG5 variants and SCZ. Thus, much larger samples than available in the current study are needed to obtain statistical power to address the role of CACNG5 mutations in SCZ and/or BPD, or other complex phenotypes. Moreover, a further family study may strength the genetic association analysis and to understand whether these rare variants are affected in specific family

The most promising rare variant detected in this thesis was p.Val146Met in exon 4 of CACNG5, which was identified in a UCL SCZ patient but was absent in other UCL cohorts. The variant
was also detected in one SCZ patient in a replication cohort from Aberdeen. The limitations of genetic analysis mean that it was not possible to unequivocally implicate the rare CACNG5 gene variants in the aetiology of SCZ and/or BPD. We were also unable to obtain clear evidence that individual variants were involved in distinct or overlapping susceptibility. One way to obtain further evidence to support the genetic association finding is to sequence this gene in further large cohorts or family samples in order to further ascertain their involvement in mental health illness.

Missense variants that are associated with Mendelian disorders typically interfere with protein stability, folding solubility or cellular processing {Kleppe, 2001 #920}. Although SIFT and Polyphen can predict whether individual variants are likely to have a benign or deleterious effect on a protein, these results are often discrepant between the different analyses. This is probably due to the overall limitations of analysis and the limitations encountered, for instance, due to lack of structural or functional data for many proteins. Functional studies of likely susceptibility variants may produce evidence that supports their likely pathogenic role in BPD and SCZ. This type of evidence may also highlight the importance of the glutamatergic pathways in these disorders. Recent sequencing of candidate glutamatergic genes in nonsyndromic intellectual disability and SCZ also reported an excess of de novo deleterious mutations in these diseases {Hamdan et al. 2011}. These observations further support the notion that BPD and SCZ or other psychiatric disorders may share a common causal synaptic component.

The most promising findings from our functional studies is the characterisation of both V146M and T164L variants in CACNG5; overexpressed V146M significantly increased the percentage of cells surface expressed GluR2 AMPA receptor (AMPAR) subunits, and overexpressed T164L decreased cell surface expression of GluR2. Indeed, TARP-γ5 has been shown to regulate
AMPAR trafficking and also the permeability to calcium in the synapse. These significant alterations in the presence of the variants may implicate dysfunctional trafficking of AMPAR to the synapse resulting in unbalanced intracellular ion concentrations. Abnormal AMPAR trafficking and/or activity has been linked to several neuropathologies, as outlined in section 1.4. This suggests a possible pathogenic mechanism for these variants. For instance, mutations in GRIA3, which encodes an AMPAR subunit and in several genes known to affect AMPAR trafficking (such as CACNG2) have been previously shown to cause forms of neurophathology (Wu et al. 2007) (Hamdan et al. 2011). Thus, our results suggest a potential causative role for rare CACNG5 variants in SCZ and/or BPD, and emphasise that CACNG5 gene mutations may contribute to dysregulation of AMPAR leading to a variety of neuropsychiatric disorders. Although our functional study has shown that two of the rare variants identified from SCZ samples altered the trafficking of AMPA receptors to the cell surface, how the localisation of individual variants affect to the trafficking of AMPAR receptors was not characterised. A further protein binding affinity study could be applied to each individual variant to detect whether each variant led to a loss of protein binding affinity; thus reducing the surface trafficking of AMPA receptors. Furthermore 3 dimensional protein structural analyses may also explain whether each individual variant has an effect on protein structure including altering protein folding or stability. Phosphorylation of TARP-?5 is also an essential mechanism which regulates AMPAR trafficking, raising the possibility that deleterious mutations at critical sites may reduce protein binding affinity. Additional work is needed to investigate whether this is the case for these variants. Finally, further functional studies of disease-associated variants may identify new drug targets and therapeutic pathways.
8 References


Castle DJ, Wessely S, Murray RM. Sex and schizophrenia: effects of diagnostic stringency, and associations with and premorbid variables. The British journal of psychiatry : the journal of mental science. 1993;162:658-64. Epub 1993/05/01.


Coombs ID, Cull-Candy SG. Transmembrane AMPA receptor regulatory proteins and AMPA receptor function in the cerebellum. Neuroscience. 2009;162(3):656-65. Epub 2009/02/03.


Held B, Freise D, Freichel M, Hoth M, Flockerzi V. Skeletal muscle L-type Ca(2+) current modulation in gamma1-deficient and wildtype murine myotubes by the gamma1 subunit and cAMP. The Journal of physiology. 2002;539(Pt 2):459-68. Epub 2002/03/08.


Jentsch JD, Roth RH. The neuropsychopharmacology of phencyclidine: from NMDA receptor hypofunction to the dopamine hypothesis of schizophrenia. Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology. 1999;20(3):201-25. Epub 1999/03/04.


Porteous D. Genetic causality in schizophrenia and bipolar disorder: out with the old and in with the new. Current opinion in genetics & development. 2008;18(3):229-34. Epub 2008/08/05.


Walsh T, McClellan JM, McCarthy SE, Addington AM, Pierce SB, Cooper GM, et al. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. Science. 2008;320(5875):539-43. Epub 2008/03/29.


Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007;447(7145):661-78. Epub 2007/06/08.


9 Appendix I

9.1 Transient transfection efficiency of CACNG5 & GIRA2

9.1.1 Introduction

The most essential step for studying mutated gene function and regulation and further protein function is to deliver genetic materials into cell cultures to produce genetically modified cells (Nikcevic et al. 2003, Ohama et al. 2005). Over the past decades, scientists have been desperately finding suitable gene delivery methods to overcome the understanding of human genetic disorders. The principle of gene delivery methods is classified into viral and non-viral ways. Viral techniques are most commonly used in clinical research also known as transduction (Pfeifer and Verma 2001) (Figure 9-1). Transductions normally use several classes of viruses as a tool for gene delivery such as retrovirus, adenovirus, adeno-associated virus, and herpes simplex virus. Viral vector are generally most effective and easy to achieve transgene expression in vivo, but they present harmful side-effects such as immunogenicity and cytotoxicity (Hacein-Bey-Abina et al. 2002, Roesler et al. 2002, Woods et al. 2003). Thus, much effect has been made to develop non-viral transfection and become the most prevalent methods in contemporary research (Glover 2012, Hart 2010).
Figure 9-1 Transduction of mammalian cells using an adenoviral vector (http://ibidi.com/applications/transfection-transduction-and-proteofection/transduction/)

In contrast, non-viral vectors are generally less effective in gene delivery and expression, when compared with viral vectors especially in vivo application (Piskin 2005). However, scientists have developed new non-viral vectors to overcome these limitations, such as increasing genetic material carrying capacity and producing easily in large quantities (Luo and Saltzman 2000, Schmidt-Wolf and Schmidt-Wolf 2003). Non-viral vectors are classified into physical or chemical based transfections. The physical methods include direct micro injection, biolistic particle delivery, electroporation, and laser-based transfection (Mehier-Humbert and Guy 2005). Those methods either require expensive instruments, or often cause cell death and physically damage to the samples (Kim and Eberwine 2010). Conversely, the chemical methods are much cheaper and easier to handle, such as cationic polymer transfection, calcium phosphate precipitation, and cationic lipid based transfection (Holmen et al. 1995, Schenborn and Goiffon 2000, Washbourne and McAllister 2002). Their principle methods are similar; positively charged chemicals combine nucleic acid to form chemical complexes. Thus, the positively charged
complexes bind to the negatively charged cell membrane through ionic interaction and enter into cells by endocytosis, without significant cytotoxicity (Dincer et al. 2005) (Figure 9-2).

**Figure 9-2 Transfection of eukaryotic cells with plasmid**

Although many commercial non-viral vectors hold promises to achieve high transfection efficiencies, there are still several factors need to be considered such as nucleic acid/chemical ratio, medium conditions, cell types and cell density on the time of transfection (Nikcevic et al. 2003). Thus, the optimised transfection conditions should be required before the studies are implemented. The aim of this study was optimising and improving selected non-viral gene transfection commercially available into HEK293 (human embryonic kidney) cells. We tested the hypothesis that, under the manufacture recommended optimised conditions, these transfection are capable of yielding high-transfection efficiencies for gene function and expression studies. We chose two of the most prevalence commercial reagents; FuGENE6, a lipid with other component, and Lipofectamine2000, a cationic lipid. For these studies, two
different plasmid DNAs encoding green and red fluorescent tag fusion proteins were determined by flow cytometric analysis of the resulting shift in fluorescence patterns on dot plots.

9.1.2 Results of FuGene6-HD transfection efficiency

To optimise the best FuGENE6-HD mediated gene delivery into HEK293 cells, the ratios of plasmid DNA/ FuGENE6-HD were prepared as 1:0.75, 1:1.5, 1:2, 1:3, 1:4, and 1:6, and then added to $10^5$ or $20^5$ cells in 6 well plates. The best transfection rate was conferred up to 55% by using 9 µl of FuGENE6-HD (1:1.5 ratios) with $20^5$ cells in each well (Fig 9-3B). Moreover, more than 70% of HEK293 cells were viable when FuGENE6-HD reagent was increased up to 36 µl (6:1 ratio). However, when FuGENE6-HD ratio was increased to 3:1 and further, the rate of gene delivery was decreased dramatically (Figure 9-3).
Figure 9-3 The best result for transfection was achieved in presence of 9 µl FuGENE6-HD. Increasing the volume of FuGENE6-HD did not cause any cell toxicity, but the transfection efficiency was decreased. A: 1 x 10^5 HEK293 cells were plated 24 hours before transfection time; B: 2 x 10^5 HEK293 cells were plated.
**9.1.3 Lipofectamine2000 transfection efficiency**

Although the rate of transfection into HEK293 cells in presence of Lipofectamine\textsuperscript{TM}2000 investigated in different condition such as amount of Lipofectamine\textsuperscript{TM}2000. The best transfection rate was conferred up to 73% by using 8 ul of Lipofectamine\textsuperscript{TM}2000 (2:1 ratio) with 2 x $10^5$ HEK293 cells in each well (Fig 9-4B). HEK293 cells were transfected at the average of 75% by increased amount of Lipofectamine\textsuperscript{TM}2000 with 1 x $10^5$ of HEK293 cells per well (Fig 9-4A). However, increased amount of Lipofectamine\textsuperscript{TM}2000 decreased the cell viability because of exacerbating cell toxicity.
The best result for transfection was achieved in presence of 8 ul Lipofectamine™2000. Increasing the volume of LipofectamineTM2000 increased the transfection efficiency rate, but the cell toxicity was also increased. A: 1 x 10^5 HEK293 cells were plated 24 hours before transfection time; B: 2 x 10^5 HEK293 cells were plated.
9.1.4 Discussion

Having an acceptable gene delivery level into eukaryotic cells is important to determine gene expression and function, protein production for recombinant genes and mutated gene studies (Nikcevic et al. 2003). Significant differences in amounts of transfected cells were measured depending on the transfection reagent, the reagent/DNA ratio, and amount of plasmid per well. Currently, there are few data that have been published on Lipofectamine™2000 based on transient expression in human cell line such as HEK293 cells. In fact, Lipofectamine™2000 is not common as much as FuGENE6-HD. FuGene6-HD, the most effective transfection reagent across literature, however was not transfected very well in our studies. It only allowed transfected up to 40%, whereas Lipofectamine2000 can transfected up to 73%. No remarkable differences were found between the obtained transfection results with increased FuGene6-HD. Another similar result from Wiesenhofer and Humpel has shown only 16.3% and 5.1% transfection efficiency with the optimised FuGENE-HD in C6 glioma cells and in primary glial cells, respectively (Wiesenhofer and Humpel 2000).

According to our findings, cationic lipid gene delivery based on using Lipofectamine™2000 is an appropriate procedure to mediate DNA transfer into HEK293T cells because of their high transfection in comparison with FuGENE6-HD, although it also had a higher cytotoxicity. Taken together, the significant factors that can influence gene transfection including cell conditions such as cell number and amount of transfection reagents and the type of cell lines should be considered for gene delivery experiments, as reported before and shown here (Felgner et al. 1987, Nikcevic et al. 2003, Wiesenhofer and Humpel 2000).

Furthermore, our finding pointed that to achieve the highest gene delivery into HEK293 cell line; one cannot rely on the transfection reagents suggested by manufacturers. As a matter of fact,
reagents such as FuGENE6-HD despite being highly recommended by manufacturers for gene transfer, they were not efficient as much as the reagent Lipofectamine™2000 for HEK293 cell line. Finally, it can be suggested that Lipofectamine™2000 is an appropriate reagent to transfer HEK293 cells.
10 Appendix II

10.1 Genotyping plots are shown for 15 variants found in CACNG5 gene for the UCL cases and control samples using Karspar analysis.

![Genotyping plots](image)

Figure 10-1 Genotyping plots of promoter variants in CACNG5 including rs3760263, rs181400884, and rs75486725 are shown. The SNP genotypes have been assigned based on cluster formation in scatter plots of normalised allele intensities X and Y. Each circle represents one individual’s genotype. Blue and purple colours indicate homozygote genotypes for the SNP (AA/aa), red colour indicates heterozygote (Aa) and green colour indicates negative sample such as water. Four distinct, tight clusters exhibited by all three representative SNPs indicate good discrimination of the three genotypes.
Figure 10. Genotyping plots of synonymous variants, R6R, C19C, A27A, and L34L found in CACNG5 are shown. The SNP genotypes have been assigned based on cluster formation in scatter plots of normalised allele intensities X and Y. Each circle represents one individual’s genotype. Blue and purple colours indicate homozygote genotypes for the SNP (AA/aa), red colour indicates heterozygote (Aa) and green colour indicates negative sample such as water. Four distinct, tight clusters exhibited by all three representative SNPs indicate good discrimination of the three genotypes.
Figure 10. 3 Genotyping plots of non-synonymous variants, R69W, R71H, R127Q, T128M found in CACNG5 are shown. The SNP genotypes have been assigned based on cluster formation in scatter plots of normalised allele intensities X and Y. Each circle represents one individual’s genotype. Blue colour indicates homozygote genotype for the SNP (AA), red colour indicates heterozygote (Aa) and green colour indicates negative sample such as water. Three distinct, tight clusters exhibited by all three representative SNPs indicate good discrimination of the three genotypes.
Figure 10. 4 Genotyping plots of non-synonymous variants, V146M, I156F, T164L, H233Y found in CACNG5 are shown. The SNP genotypes have been assigned based on cluster formation in scatter plots of normalised allele intensities X and Y. Each circle represents one individual’s genotype. Blue colour indicates homozygote genotype for the SNP (AA), red colour indicates heterozygote (Aa) and green colour indicates negative sample such as water. Three distinct, tight clusters exhibited by all three representative SNPs indicate good discrimination of the three genotypes.
11 Appendix III

11.1 Raw data of protein expression levels

<table>
<thead>
<tr>
<th>Samples</th>
<th>Replication 1</th>
<th>Replication 2</th>
<th>Replication 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total Expression Level</td>
<td>Surface Expression Level</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AMPA</td>
<td>TARPS</td>
<td>AMPA</td>
</tr>
<tr>
<td>GluR2</td>
<td>33.1</td>
<td>0</td>
<td>694</td>
</tr>
<tr>
<td>TARP5</td>
<td>911</td>
<td>1,650</td>
<td>1,640</td>
</tr>
<tr>
<td>SNP 69</td>
<td>764</td>
<td>410</td>
<td>1,120</td>
</tr>
<tr>
<td>SNP 71</td>
<td>704</td>
<td>2,300</td>
<td>2,830</td>
</tr>
<tr>
<td>SNP 127</td>
<td>795</td>
<td>1,200</td>
<td>1,150</td>
</tr>
<tr>
<td>SNP 128</td>
<td>359</td>
<td>642</td>
<td>741</td>
</tr>
<tr>
<td>SNP 156</td>
<td>232</td>
<td>201</td>
<td>307</td>
</tr>
<tr>
<td>SNP 164</td>
<td>266</td>
<td>727</td>
<td>510</td>
</tr>
</tbody>
</table>

Table 11-2 Raw data of protein expression levels