



Copy Number Variation and Schizophrenia

by

Gloria Wing Chi TAM

Wellcome Trust Sanger Institute
Trinity College, University of Cambridge

This dissertation is submitted for the degree of Doctor of Philosophy

Nov 2008



Preface

This dissertation reports the work carried out at the Wellcome Trust Sanger Institute, between April 2005 and October 2008. It is submitted for the degree of Doctor of Philosophy, contains 248 pages (excluding bibliography and appendices), 58 figures and 20 tables and does not exceed the limit set by the Degree Committee.

This dissertation is my own work and contains nothing that is the outcome of work done in collaboration with others, except as specified in the text and below.

In section 3, oligonucleotide array hybridization experiments (for the *ABCA13* deletion) were performed in collaboration with Tomas Fitzgerald (Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, UK).

In section 4, whole-genome tiling path array hybridization experiments were performed in collaboration with Karen Porter (Wellcome Trust Sanger Institute), who performed approximately one third of the hybridization experiments. Affymetrix SNP array experiments and analysis were performed by the International Schizophrenia Consortium (Broad Institute, Boston, Cambridge MA) and the data was provided by Prof. Douglas Blackwood (Department of Psychiatric Genetics, University of Edinburgh, Edinburgh). Perl scripts for CNV genotyping were developed by Dr Richard Redon (Wellcome Trust Sanger Institute).

I hereby declare that I am the sole author of this dissertation and no part of the work contained in this dissertation has been previously submitted for any other degree.

Abstract

Schizophrenia is a debilitating psychiatric illness affecting 1% of the population worldwide. The aetiology of schizophrenia is largely unknown, and deciphering schizophrenia genetics has remained a major challenge during the past decades in psychiatric research. In the past, visible alterations of the genome have been recognized as the underlying causes in a number of cognitive or behavioural defects. Structural chromosomal abnormalities, such as the 22q11 microdeletion and the Disrupted in Schizophrenia 1 (*DISC1*) translocation, were demonstrated to play a role in a proportion of schizophrenia cases. Furthermore, recent studies have identified a number of novel recurrent submicroscopic copy number changes significantly associated with schizophrenia (ISC 2008; Stefansson et al. 2008). This thesis describes a multi-faceted investigation to identify schizophrenia-related copy number variations (CNVs), defined as deletions and duplications larger than 1 kb in the genome.

As a first approach I performed array CGH on the whole-genome tiling path (WGTP) platform to screen for CNVs in three familial cases. Each pedigree consists of multiple patients affected with schizophrenia and other psychiatric illnesses. I identified a duplication on chromosome 1p36 common to all four affected members in one family, which was not identified in the normal HapMap controls (n=269). The CNV extends from the gene *H6PD* (Hexose-6-phosphate dehydrogenase precursor) to *SPSB1* (SPRY domain-containing SOCS box protein SSB-1). Using quantitative PCR, long range PCR and Fiber-FISH, I sequenced

the duplication breakpoint and delineated the structure of this potential pathogenic variant. Next, in a candidate-gene targeted approach I screened a multiplex schizophrenia family for CNVs in the gene *ABCA13* (ATP Binding Cassette Gene 13) at 7p12. I demonstrated the segregation of an intronic deletion with disease status.

Complementary to the family-based approach, I designed a population-based CNV study in schizophrenia versus matched control cohorts. Still using WGTP arrays, I performed a genome-wide screen for CNVs in 91 Scottish schizophrenia patients and 92 Lothian Birth Control DNA samples. In the WGTP dataset I identified a previously established schizophrenia-associated deletion at 15q11.2 (Stefansson et al. 2008) in a schizophrenia patient, near the *CYFIP1* (cytoplasmic FMR1 interacting protein 1 isoform) gene. I also identified a number of rare variants overlapping genes that are linked to various psychiatric diseases, including *SGCE* (sarcoglycan, epsilon), *OXTR* (Oxytocin) and *RCAN1* (Down Syndrome Critical Region 1). My results are consistent with recent reports demonstrating the role of rare CNVs in schizophrenia (ISC 2008, Walsh et al. 2008). In terms of common copy number variations, I genotyped 577 common CNVs using the WGTP data, and identified 31 candidates with putative bias in genotype distributions in cases versus controls. Two of these candidates, one at 3p26 and another at 15q13, were genotyped in an extended case-control cohort. Neither of them showed significant association with disease in the extended cohort.

The last approach was based on the hypothesis that CNVs could be linked to variations in learning, memory and brain function, both in the normal population and in psychiatric patients. The strategy involved a CNV screen on a set of proteins with important neuronal and synaptic functions. The NMDA receptor complex (NRC/MASC) was selected due to known roles of its components in cognitive and behavioural traits. Out of 186 NRC/MASC proteins, 20 of them showed CNVs in normal HapMap individuals. Four of these were linked to components of the core synaptic machinery, including a common CNV at *DLG1* (Discs, Large homolog 1 (Drosophila)). In addition, I investigated the multi-allelic variant at 17q21 near the gene N-ethylmaleimide-sensitive factor (*NSF*). I identified two major CNV blocks with interesting population bias, and identified for the first time a European-specific haplotype in an allelic variant known as H1.

Acknowledgements

Firstly, I would like to express my deepest gratitude towards my supervisors, Professor Seth Grant and Dr Nigel Carter. Their continuous support, advice and guidance have made my research life a lot more rewarding than it would have been. I am also particularly grateful to Dr Richard Redon, who has frequently inspired me on scientific ideas and technical knowledge, as well as communication skills and thought processes. My PhD studies were greatly enhanced by the generous support and patience from these three mentors.

I would also like to thank all members of the labs in Team 32 and Team 70, in particular Karen Porter and Tomas Fitzgerald, who have contributed to parts of the research work described in this thesis, as well as being supportive and joyful companions in the lab. I am also grateful to everyone who has provided me with advice, scientific discussions and technical support, including Dr Louie Van de Lagemaat, Dr Norboru Komiyama, Dr Fengtang Yang, Dr Jianxiang Chi (University of Oxford), Diane Rigler, Diana Rajan, Dr Douglas Strathdee, Kathryn Elsegood, Ellie Tuck, Dr Lianne Stanford and Charles Pettit.

I would like to express my appreciation to members of my thesis committee, Professor Nabeel Affara (Department of Pathology, University of Cambridge) and Professor Allan Bradley, for their scientific insights and constructive comments during my PhD reviews.

I would like to acknowledge my research collaborators at the University of Edinburgh, including Professor Douglas Blackwood, Professor Ian Deary, Dr. Walter Muir, Dr. Ben Pickard, Mary Malloy and Margaret Van Beck, for the provision of DNA samples, CNV data and the useful discussions on the clinical aspects of the project.

Special thanks go to the Wellcome Trust and Trinity College (University of Cambridge) for providing financial support. Last but not least, I am grateful to my family and all my friends for their endless love and support, and finally to Anson Ma for the wonderful years we have spent in Cambridge.

** Affiliation: Wellcome Trust Sanger Institute, unless otherwise specified in brackets.*

Gloria Tam

November 2008.

List of Figures

Figure 1.1 Types of genetic variants and their relative sizes.....	6
Figure 1.2 Comparative Genome Hybridization applied on a cancer cell line revealed amplification of the <i>myc</i> locus.....	10
Figure 1.3 Schematics of an array comparative genome hybridization experiment....	13
Figure 1.4 Sensitivity and throughput of various CNV detection techniques.....	24
Figure 1.5 Non-allelic homologous recombination by low copy repeat.....	29
Figure 1.6 Phenotypic effects of CNVs.....	31
Figure 1.7 Lifetime risk of developing schizophrenia based on relationship.....	53
Figure 1.8 Genomic loci of the three novel recurrent deletions associated with schizophrenia.....	67
Figure 1.9 Schematic diagram of the synapse displaying known involvement of synaptic proteins in schizophrenia CNV loci.....	72
Figure 2.1 Amplified DNA samples after BioPrime Labelling Procedure.....	82
Figure 2.2 PCR genotyping of a deletion at 3p26 5' upstream of Close Homolog of L1 (<i>CHL1</i>)	92
Figure 2.3 Long Range PCR gel electrophoresis analysis of a tandem duplication at Chr1p36.....	94
Figure 2.4 Dissociation curves analysis for different primer sets.....	97
Figure 2.5 A typical standard curve of the control primer.....	97
Figure 2.6 Labelled DNA probes for Fluorescent In Situ Hybridisation.....	102
Figure 3.1 Three families analysed by whole-genome CNV screen.....	108
Figure 3.2 qPCR validation of duplication 1p36 in Pedigree F-29.....	109
Figure 3.3 Delineation of 1p36.22 duplication structure by Fiber-FISH.....	112
Figure 3.4 Sequencing breakpoints of 1p36.22 revealed repeat structures.....	113
Figure 3.5 Array CGH detection of the 7p12 deletion for Patient 4398.....	119
Figure 3.6 Array CGH profiles for all available members in Family 340.....	120

Figure 4.1 Case and control cohorts and CNV detection platforms used in our study.....	128
Figure 4.2 Normalization and filtering steps applied to WGTP hybridization data before CNV analysis.....	130
Figure 4.3 WGTP array quality control indicators.....	131
Figure 4.4 Frequency and types of CNVs and CNVRs in SCZ and LBC.....	133
Figure 4.5 Size and frequency distributions of CNVRs in Schizophrenia (SCZ) and Lothian Birth Control (LBC) Cohorts detected using the WGTP platform.....	134
Figure 4.6 CNVR gene content in the SCZ and LBC cohorts.....	135
Figure 4.7 Correlation of CNV discovery rate with data quality in SCZ and LBC.....	136
Figure 4.8 WGTP data compared with 4 known schizophrenia CNV loci.....	138
Figure 4.9 WGTP data detected a deletion at 15q11.2 in one patient.....	140
Figure 4.10 WGTP CNV dataset validation strategies.....	142
Figure 4.11 Schizophrenia cohort-specific CNVRs containing brain-related or neuronal-related genes.....	149
Figure 4.12 Recurrent SCZ-specific CNVR regions detected by both WGTP and Affymetrix platforms.....	152
Figure 4.13 Vst scores to identify clones showing SCZ and LBC differentiation in the WGTP analysis.....	159
Figure 4.14 Examples of bivariate clustering based on log2ratio of consecutive clones.....	162
Figure 4.15 Three regions with significant difference of genotype distributions between SCZ and LBC as detected by bivariate clustering.....	164
Figure 5.1 A copy number polymorphism at chromosome 3p26.3 was detected by three BAC clones.....	174
Figure 5.2 <i>CHL1</i> CNV genotypes from bivariate clustering.....	175
Figure 5.3 Genomic location of the <i>CHL1</i> 5' CNV.....	176
Figure 5.4 <i>CHL1</i> expression level against BAC clone log2ratio in HapMap samples..	178
Figure 5.5 Genotyping of the <i>CHL1</i> 5' deletion polymorphism in HapMap samples...	179
Figure 5.6 <i>CHL1</i> gene expression against copy number of the <i>CHL1</i> 5' CNV in HapMap CEU samples.....	180

Figure 5.7 The 15q13-14 genomic locus.....	184
Figure 5.8 The three existing polymorphic structures of <i>CHRFAM7A</i> allele at 15q13-14.....	185
Figure 5.9 Copy number polymorphism at 15q13.2-13.3 spanning the gene <i>CHRFAM7A</i>	194
Figure 5.10 <i>CHRFAM7A</i> CNV genotype distributions in cases versus controls.....	195
Figure 5.11 Analysis of a typical Taqman qPCR experiment to determine <i>CHRFAM7A</i> genotypes. In each 96-well plate 40 samples (+2 controls) were genotyped.....	197
Figure 6.1 Schematic diagram of a glutamatergic excitatory synapse.....	204
Figure 6.2 Log2ratio distributions for WGTP clones reporting CNV among 269 HapMap samples in 20 MASC regions.....	207
Figure 6.3 CNVs affecting core components of the NRC/MASC signalling complex..	212
Figure 6.4 Schematic representation of the 17q21 locus.....	216
Figure 6.5 Array CGH genomic profiles of the 3 HapMap ethnic groups at chr17q21.	219
Figure 6.6 Population differentiation at two major CNV blocks at 17q21.....	220
Figure 6.7 High resolution oligo array CGH profiles for 20 individuals at chr17 41.5 Mb-42.2 Mb.....	222
Figure 6.8 Quantitative PCR validation of CNV _{KIAA1267} and CNV _{NSF}	224
Figure 6.9 Fiber-FISH experiment to visualize copy number of CNV _{NSF}	224
Figure 6.10 Comparing SCZ and LBC samples at CNV _{KIAA1267} and CNV _{NSF}	227
Figure 6.11 Locations of derived and ancestral loci of segmental duplications at 17q21.....	228
Figure 7.1 Identification of CNVs as disease risk loci for further characterization.....	247

List of Tables

Table 1.1 Two most influential diagnostic guidelines for schizophrenia the DSM-IV and ICD-10.....	49
Table 1.2 Schizophrenia candidate genes with evidence from linkage and association analysis.....	56
Table 1.3 Three novel recurrent deletions associated with schizophrenia.....	68
Table 2.1 Cycling protocol for general PCR amplification.....	90
Table 2.2 Cycling protocol for long-range PCR amplification.....	93
Table 2.3 Reagent mixture for each Taqman quantitative PCR assay.....	95
Table 2.4 Cycling protocol for Quantitative PCR amplification.....	96
Table 3.1 Known genomic rearrangement at 1p36.22.....	115
Table 3.2 Schizophrenia patients with DNA analysed on custom arrays.....	119
Table 4.1 Validated SCZ-specific rare variants.....	143
Table 4.2 SCZ-specific rare variants with genes associated with psychiatric disorders.	146
Table 4.3 Recurrent SCZ-specific variants as detected in 297 SCZ samples.....	151
Table 4.4 Validating 5 top Vst-regions using custom-design oligonucleotide array.....	160
Table 5.1 <i>CHL1</i> 5' CNV distribution determined by the original genome-wide array CGH data.....	182
Table 5.2 <i>CHL1</i> 5' CNV distribution in an extended case-control cohort determined by PCR genotyping.....	182
Table 5.3 Previous reports associating the 15q13-14 locus to schizophrenia and related psychiatric disorders.....	189
Table 5.4 <i>CHRFAM7A</i> CNV genotype distributions in the extended case and control cohorts.....	198
Table 6.1 CNVs detected at 20 genes encoding NRC/MASC signalling complex components.....	206
Table 6.2 CNV genotypes for all unrelated HapMap individuals at CNV _{KIAA1267} and CNV _{NSF}	225
Table 6.3 SCZ and LBC genotype counts at CNV _{KIAA1267} and CNV _{NSF}	227

List of Abbreviations

ADEOAD	Autosomal Dominant Form of Early-Onset Alzheimer Disease
Alu	a family of repeat elements named after the <i>AluI</i> restriction site
AS	Angelman Syndrome
ASD	Autism Spectrum Disorder
BAC	Bacterial Artificial Chromosome
bp	Base Pairs
CAA	Cerebral Amyloid Angiopathy
CD-CV	Common-Disease Common-Variant
CD-RV	Common-Disease Rare-Variant
CEU	HapMap DNA: Utah samples with European ancestry
CGH	Comparitive Genome Hybridisation
CHB	HapMap DNA: Chinese samples with Asian ancestry
CNP	Copy Number Polymorphism
CNV	Copy Number Variation
CNVR	Copy Number Variation Region
COS	Childhood-Onset
cR	combined ratio
Cy3	Indocarbocyanine
Cy5	Indodicarbocyanine
DECIPHER	Database of Chromosomal Imbalance
DGS	DiGeorge Syndrome
DGV	Database of Genomic Variants
DISC1	Disrupted In Schizophrenia 1
DNA	Deoxyribonucleic Acid

dNTP	Deoxynucleoside Triphosphate
DOP-PCR	Degenerate Oligonucleotide Primed Polymerase Chain Reaction
DSBs	Double Stranded Breaks
DSM-IV	Diagnostic and Statistic Manual of Mental Disorder- 4th Edition
ECS	Electroconvulsive Shocks
ESP	Clone-End Sequence-Pair
EtOH	Ethanol
FISH	Fluorescent In Situ Hybridisation
GAD	Genetic Association Database
G-banding	Giemsa-banding
GWAS	Genome-Wide Association Studies
ICD-10	International Statistical Classification of Diseases and Related Health Problems 10th Revision
INDEL	Insertions and Deletions in a chromosome
ISC	International Schizophrenia Consortium
JPT	HapMap DNA: Japanese samples with Asian ancestry
kb	Kilo Base (One Thousand Base Pairs)
LBC	Lothian Birth Control Cohort
LCR	Low Copy Repeat
LD	Linkage Disequilibrium
LINE	Long Interspersed Nuclear Element
LOD	Logarithm of Odds
LOH	Loss of Heterozygosity
LTP	Long-Term Potentiation
M	Molar
MAGUK	Membrane-Associated Guanylate Kinase

MAPH	Multiplex Amplifiable Probe Hybridisation
MAQ	Multiplex Amplicon Quantification
MASC	MAGUK Associated Signaling Complex
Mb	Mega Base
min	Minute
MLPA	Multiplex Ligation-Dependent Probe Amplification
MR	Mental Retardation
MRI	Magnetic Resonance Imaging
mRNA	Messenger RNA
NAHR	Non-Allelic Homologous Recombination
NHEJ	Non-Homologous End Joining
NMDA	N-Methyl-D-Aspartate
NRC	NMDA Receptor Complex
nt	Nucleotide
OMIM	Online Mendelian Inheritance In Man
OR	Odds Ratio
PBS	Phosphate-Buffered Saline
PCP	Phencyclidine
PCR	Polymerase Chain Reaction
PEM	Paired-End Mapping
PET	Positron Emission Tomography
PFGE	Pulsefield Gel Electrophoresis
PPI	Prepulse Inhibition
PSD	Postsynaptic Density
PWS	Prader Willi Syndrome

qPCR	Quantitative PCR
RACE	Rapid Amplification of cDNA Ends
rcf	Relative Centrifugal Force
RNA	Ribonucleic Acid
rpm	Revolutions Per Minute
SCZ	Schizophrenia Cohort
SD	Segmental Duplication
SDe	Variability Measure for Array CGH Experiments
SINE	Short Interspersed Nuclear Element
SKY	Spectral Karyotype
SNP	Single Nucleotide Polymorphism
Tm	Melting Temperature
VCFS	Velo-Cardio-Facial Syndrome
VNTR	Variable Number Of Tandem Repeat
Vst	A variance-based measure to compare quantitative data from different cohorts
WGTP	Whole Genome Tiling Path
YRI	HapMap DNA: Yoruba samples with African ancestry

CHAPTER 1 INTRODUCTION

1.1 COPY NUMBER VARIATION (CNV) AS A SOURCE OF GENETIC DIVERSITY	2
1.2 DETECTION OF COPY NUMBER VARIATION	7
1.2.1 Classical Cytogenetic Techniques for the Detection of Structural Variations.....	7
1.2.2 Array-based Comparative Genome Hybridization	11
1.2.2.1 BAC Array CGH	12
1.2.2.2 Oligonucleotide Array CGH	14
1.2.2.3 Choice of Reference DNA for Array Hybridization.....	15
1.2.3 Genotyping CNVs using Single Nucleotide Polymorphisms	16
1.2.3.1 Linkage Disequilibrium Based Tag-SNP Approach	16
1.2.3.2 CNV Genotyping Using SNP Arrays	17
1.2.4 Validation and Detection of CNV at Targeted Loci	19
1.2.4.1 Quantitative Fluorescent Real-time PCR (qPCR).....	19
1.2.4.2 Multiplex Quantitative Fluorescent Real-time PCR	20
1.2.4.3 Other Methods.....	20
1.2.5 Genome Sequencing and CNV Detection.....	22
1.2.5 Genome Sequencing and CNV Detection.....	22
1.2.5.1 Clone End Mapping and Sequencing.....	22
1.2.5.2 Novel Sequencing Technology and Genome Comparison.....	23
1.3 MECHANISMS OF COPY NUMBER VARIATION GENERATION.....	25
1.3.1 Non-allelic Homologous Recombination	25
1.3.2 Non-homologous End Joining.....	26
1.3.3 Other Mechanisms.....	27
1.3.4 Insights from Breakpoint Mapping.....	27
1.4 BIOLOGICAL IMPACT OF COPY NUMBER VARIATION	30
1.4.1 Phenotypic Effect of CNVs.....	30
1.4.1.1 CNV and Expression	32
1.4.1.2 CNV and Gene Disruption	34
1.4.2 Phenotypic Variations and Evolution.....	35
1.4.2.1 CNV and Human Traits	35
1.4.2.2 Positive and Negative Selections on CNV	36
1.4.3 CNV and Disease	38
1.4.3.1 Genome Disorder	38
1.4.3.2 Rare CNVs in Mendelian Disease Traits.....	40
1.4.3.3 Common CNVs in Multifactorial or Complex Disease	41
1.5 SCHIZOPHRENIA.....	45
1.5.1 The Concept of Schizophrenia.....	45
1.5.2 Phenotypes and Diagnosis	47

1.5.2.1	Positive and Negative Symptoms	47
1.5.2.2	Endophenotypes	47
1.5.2.3	Course of Illness.....	48
1.5.2.4	Standardised Diagnostic Methods	48
1.5.3	Aetiology and Neurobiology	50
1.5.3.1	Dysfunction of Neurotransmitter Systems.....	50
1.5.3.2	Neurodevelopment and Neuropathology	52
1.5.4	The Genetics of Schizophrenia.....	53
1.5.4.1	Evidence of Genetic Contribution	53
1.5.4.2	A Search for Candidate Genes.....	54
1.5.4.3	Genes and Environment.....	57
1.6	CNV IN SCHIZOPHRENIA AND OTHER PSYCHIATRIC DISEASES	58
1.6.1	Early Studies on Chromosomal Abnormalities in Schizophrenia	58
1.6.1.1	DISC1 and Breakpoint Study in Schizophrenia.....	58
1.6.1.1	22q11 Microdeletion and Schizophrenia.....	60
1.6.2	Large Scale CNV Screen in Schizophrenia Patients.....	63
1.6.2.1	Summary of CNV Findings	63
1.6.2.1	Identification of Large Recurrent Schizophrenia Loci.....	65
1.6.2.2	Increased Mutation Burden of CNV in Schizophrenia Patients	69
1.6.2.3	Rare Variants Converging into Neurodevelopmental Pathways.....	70
1.7	SCOPE OF THESIS	77

CHAPTER 2 MATERIALS AND METHODS

2.1	ARRAY COMPARATIVE GENOME HYBRIDIZATION (ARRAY CGH)	80
2.1.1	Patient and Control DNA Samples.....	80
2.1.2	DNA Labelling.....	81
2.1.3	Sample Precipitation and Preparation.....	82
2.1.4	Array Hybridisation	83
2.1.5	Image Acquisition and Data Analysis.....	83
2.2	AGILENT OLIGO CUSTOM-DESIGNED ARRAY CGH	86
2.2.1	Custom Array Designs.....	86
2.2.2	DNA Labelling.....	86
2.2.3	Preparation of Labelled Genomic DNA.....	87
2.2.4	Array Hybridisation	87
2.2.5	Slide Washing.....	88
2.2.6	Image Acquisition and Data Analysis.....	89
2.3	POLYMERASE CHAIN REACTION (PCR).....	90
2.3.1	PCR genotyping of the 3p26 Deletion near <i>CHL1</i>	90

2.3.2	Long Range Polymerase Chain Reaction (LR-PCR)	92
2.3.3	Quantitative Real-time Polymerase Chain Reaction (qRT-PCR)	94
2.3.3.1	SYBR Green Method	94
2.3.3.2	Taqman Method with MGB Probes	95
2.3.3.3	Thermal Cycler and Reaction Condition	95
2.3.3.4	Standard Curve Generation	96
2.3.3.5	DNA Quantification and Data Analysis	98
2.4	FLUORESCENCE IN-SITU HYBRIDISATION (FISH)	99
2.4.1	Growing Cell lines	99
2.4.2	Preparation of Extended Chromatin Fibre Slides	99
2.4.3	Preparation of Fosmid Clone Insert DNA	100
2.4.4	Amplification and Labeling of DNA probes	101
2.4.4.1	GenomePlex® Whole Genome Amplification	101
2.4.4.2	DNA Fluorescent Labelling	101
2.4.4.3	Probe Fragmentation	102
2.4.5	Immunofluorescence and Image Acquisition	103

CHAPTER 3 FAMILIAL STUDY IN SCHIZOPHRENIA

3.1	WHOLE-GENOME CNV SCREEN IN FAMILIES	106
3.1.1	Whole-genome Array CGH Screen in Three Familial Cases	106
3.1.2	Characterization of the Rare Familial Duplication at 1p36.22	110
3.1.3	Known Genomic Rearrangements Near the 1p36.22 duplication	114
3.2	DELETION AT <i>ABCA13</i> IN AN EXTENDED FAMILY WITH SCHIZOPHRENIA	116
3.2.1	Evidence of Functional Mutations of <i>ABCA13</i> at 7p12.3	116
3.2.2	Oligo Array CNV Screen in an Extended Pedigree	118
3.3	CHAPTER SUMMARY AND DISCUSSION	122

CHAPTER 4 POPULATION-BASED CNV STUDY IN SCHIZOPHRENIA

4.1. EXPERIMENTAL DESIGN AND ARRAY DATA QUALITY CONTROL ..	127
4.1.1 Case-Control CNV Screen Experimental Design	127
4.1.2 WGTP Array Data Quality Control	129
4.2 COPY NUMBER VARIATION DETECTION ON THE WGTP ARRAY	132
4.2.1 Distribution of CNVs and CNV Regions in Case and Control Cohorts.....	132
4.2.2 Bias of CNV Discovery Rate in SCZ Versus LBC	135
4.3 COMPARING WGTP DATA WITH KNOWN SCHIZOPHRENIA CNV REGIONS.....	137
4.4 RARE VARIANTS SPECIFIC TO THE SCHIZOPHRENIA COHORT	141
4.4.1 Rare Variant Detection Using Consecutive Clone Calling Criteria.....	141
4.4.2 Validation of SCZ-Specific Rare Variants	142
4.4.3 Rare Variants in SCZ with Genes Involved in Psychiatric Disorders	144
4.5 RECURRENT SCZ-SPECIFIC VARIANTS IN EXTENDED COHORT	150
4.5.1 A CNVR at Down Syndrome Critical Region 1 (RCAN1/DSCR1).....	153
4.5.2 A Variant Near Olfactomedin1 and other Recurrent CNVRs	154
4.6 FREQUENT COPY NUMBER VARIATIONS IN SCZ AND LBC	156
4.6.1 Variance-based Clone-by-Clone Cohort Comparison	156
4.6.2 CNV Genotyping with Bivariate Clustering.....	161
4.7 CHAPTER SUMMARY AND DISCUSSION.....	165

CHAPTER 5 GENOTYPING TWO SCHIZOPHRENIA CNVS IN AN EXTENDED CASE CONTROL COHORT

5.1 CANDIDATE I: <i>CHL1</i> (CLOSE HOMOLOG OF L1) AT 3P26	170
5.1.1 Functional Significance of <i>CHL1</i>	170
5.1.2 Structure of the <i>CHL1</i> 5' Deletion Polymorphism	173
5.1.3 Correlation of <i>CHL1</i> CNV to mRNA Expression.....	177
5.1.4 Genotyping <i>CHL1</i> CNV in a Larger Case-Control Cohort.....	181
5.2 CANDIDATE II: <i>CHRFAM7A</i> (CHRNA7-FAM7A FUSION GENE) AT 15Q13	183

5.2.1	Genomic Architecture at chromosome 15q13-14	183
5.2.2	Molecular Genetic Studies linking 15q13-14 to Schizophrenia	187
5.2.3	Functional Significance of <i>CHRNA7</i> and <i>CHRFAM7A</i>	191
5.2.4	Copy Number Polymorphism at the <i>CHRFAM7A</i> Region	192
5.3	CHAPTER SUMMARY AND DISCUSSION	199

CHAPTER 6 CNVS AND THE NMDA RECEPTOR COMPLEX

6.1	NMDA RECEPTOR COMPLEX, SCHIZOPHRENIA AND COGNITION..	203
6.2	COPY NUMBER VARIATION AND THE NMDA RECEPTOR COMPLEX	205
6.3	CNV AT 17Q21 NEAR N-ETHYLMALEAMIDE-SENSITIVE FACTOR (NSF)	213
6.3.1	Introduction to the 17q21 locus near <i>NSF</i>	213
6.3.2	Known Genomic Structure of 17q21	215
6.3.3	Array CGH Data Reveals Two Major CNVs at 17q21	217
6.3.3.1	CNV _{NSF} : Copy number Variant at 5' end of <i>NSF</i>	217
6.3.3.2	CNV _{KIAA1267} : Copy number Variant at <i>KIAA1267</i>	218
6.3.4	Resolving CNVKIAA1267 with SNP and High Resolution Oligo Array Data	221
6.3.5	Validating CNVNSF and CNVKIAA1267 by qPCR and FISH	223
6.3.6	Genotyping HapMap Individuals for 17q21 Structural Variants	225
6.3.7	CNVNSF and CNVKIAA1267 in Schizophrenia Versus Control	226
6.3.8	Evolutionary History of CNVNSF and CNVKIAA1267	228
6.4	CHAPTER SUMMARY AND DISCUSSION	229

CHAPTER 7 GENERAL DISCUSSION

7.1	DESIGN OF CNV DISCOVERY AND ASSOCIATION STUDY.....	232
7.1.1	Multiple Approaches of CNV Study	232
7.1.2	Future Large-Scale CNV Studies	234
7.1.3	Enrichment of Schizophrenia Subtypes	235
7.2	UNDERSTANDING THE GENETIC MODEL OF SCHIZOPHRENIA.....	237
7.2.1	Rare Variants Versus Common Variants	237
7.2.2	Incomplete Penetrance and Expressivity	239
7.2.3	Aetiological Overlap of Schizophrenia with Other Psychiatric Diseases	239

7.3	CLINICAL RELEVANCE OF CNVS IN SCHIZOPHRENIA.....	241
7.3.1	CNV Findings Translating into Disease Classification.....	241
7.3.2	Genetic Counselling and Therapeutic Potential	242
7.4	THESIS SUMMARY	244
7.5	FUTURE DIRECTION	246
	REFERENCE.....	249
	APPENDICES.....	278