# SCHIZOPHRENIA: ANOTHER EMPHASIS OF ERBB4 AND NRG1 IMPACT ON DISEASE DEVELOPMENT

by

Zeynep Sena Ağım

B.S., Molecular Biology and Genetics, Bilkent University, 2009

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Molecular Biology and Genetics

Boğaziçi University

2012

SCHIZOPHRENIA: ANOTHER EMPHASIS OF ERBB4 AND NRG1 IMPACT ON DISEASE DEVELOPMENT

APPROVED BY:

| | |
|---|---|
| Prof. A. Nazlı Başak<br>(Thesis Supervisor) | ..... |
| Assoc. Prof. Hilmi Özçelik<br>(Thesis Co-Supervisor) | ..... |
| Cem İsmail Küçükali, M.D. | ..... |
| Prof. Özlem Keskin Özkaya | ..... |
| Assoc. Prof. Müge Türet Sayar | ..... |

DATE OF APPROVAL: 04.07.2012

*To my dear parents,*

*Devran and Malik*

# ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my thesis advisors Prof. A. Nazlı Başak and Assoc. Prof. Hilmi Özçelik for their valuable criticisms and continuous support. Thanks to invaluable opportunities and strength they gave during my MSc thesis, now, I feel myself equipped to pursue my academic career.

I would like to express my special thanks to Prof. Özlem Keskin, Assoc. Prof. Müge Türet Sayar and Cem İsmail Küçükali for taking part in my thesis committee and spend their valuable time to evaluate this thesis.

At this point, I would also like to acknowledge the Suna and İnan Kıraç Foundation and Boğaziçi University Research Funds for the financial support they provided throughout my study.

I would like to express my special thanks to Özgun Uyan for his companionship during every minute of my study at NDAL. I want to thank especially to Suna Lahut, Aslıhan Özoğuz, Ceren İskender, Özgür Ömür, Güneş Birdal, Ceren Saygı, Ece Kartal, Kadir Özkan, Selda Dağdeviren, İlknur Yıldız, Aslı Gündoğdu and Irmak Şahbaz, and former NDAL members Gönenç Çobanoğlu, Pınar Deniz Aksoy and Arman Aksoy for their unforgettable friendship and support. I am grateful to every moment we spend together in three special years. I cannot forget the kind assistance and friendship of Melda Esendal during my stay at the the Samuel Lunenfeld Research Institute, Mount Sinai Hospital in Toronto.

I believe that it is impossible to be a successful scientist without a peaceful home. I feel very lucky and special for having a beloved and patient mother, Mah-ı Devran Baykal and the most intellectual and supportive father, Mümtaz Malik Ağım. Without their endless love, I would not be the person who I am, thank you…

# ABSTRACT


## SCHIZOPHRENIA: ANOTHER EMPHASIS OF ERBB4 AND NRG1 IMPACT ON DISEASE DEVELOPMENT

Schizophrenia, which is one of the most commonly reported illnesses among all psychiatric disorders, is defined as a complex disease; this implicates that it does not show a simple Mendelian pattern of inheritance. The discovery of the role of ERBB4-NRG1 axis has led to many publications, which studied the independent as well as interactive effects of ERBB4 and NRG1, associated with schizophrenia. However, most of these findings could not be validated in other studies. Also the all 21 genome-wide association studies in schizophrenia, published so far, could not identify any of *ERBB4* and *NRG1* SNPs as significant. This study exploits three publicly available genome-wide association study datasets: CATIE, GAIN and nonGAIN and aims to establish an innovative analysis methodology to highlight genetic significance of genes, such as *ERBB4* and *NRG1* that play crucial roles in molecular pathways leading to schizophrenia. In the framework of this study, novel regions of the ERBB4 and NRG1 genes associated with schizophrenia were identified in three large GWAS datasets using two different haplotype analyses, haplotype-based logistic regression and Haploview. Using the same methodology, previously associated *ERBB4* and *NRG1* blocks were validated in at least two datasets. To predict the functional effects of novel blocks, transcription factors that bind to these regions were identified. This thesis specifically and for the first time showed that the binding of particular transcription factors to intronic regions of *ERBB4* and *NRG1* might have causative or protective effects on schizophrenia. This is the first study to perform an *in silico* and bioinformatics-based functional analysis of variants located within introns in schizophrenia subjects from GWAS datasets. The methodology exploited in this thesis gives promising results that facilitate our further understanding of the functional role of intronic variants.

# ÖZET


## ŞİZOFRENİ: ERBB4 VE NRG1 ETKİSİNİN HASTALIK GELİŞİMİ ÜZERİNE YENİ BİR KANITI

En sık görülen ruhsal bozukluklardan biri olan şizofreni, kompleks bir hastalıktır; Mendel tipi kalıtım göstermez. ERBB4-NRG1 aksının rolünün bulunmasıyla, bu genlerin şizofrenideki hem bağımsız, hem de interaktif etkilerini gösteren birçok çalışma yayınlandı; fakat bu çalışmaların çoğunun sonuçları birbirini doğrulamadı. Bugüne kadar yayınlanmış olan 21 genom-çapı ilişkilendirme çalışmasının (GWAS) hiçbirinde bu iki gene ait SNP'ler anlamlı olarak gösterilemedi. Bu çalışma, üç GWAS'A ait verileri (CATIE, GAIN ve nonGAIN) kullanarak, şizofreniye neden olan moleküler yolaklarda görev alan ERBB4 ve NRG1 gibi genlerin genetik rolünü bulmak için yenilikçi bir in silico analiz yöntemi geliştirmeyi hedeflemektedir. Bu amaçla çalışma kapsamında, üç veri setinde de ERBB4 ve NRG1 genlerinin şizofreni ile ilişkilendiren bölgeleri, iki farklı haplotip-analiz yöntemi kullanılarak tanımlandı: Haplotip-tabanlı lojistik regresyon modeli ve Haploview yazılımı. Aynı yöntemler kullanılarak, daha önceki çalışmalarda, bu iki genin şizofreni ile ilişkilendirilen bölgeleri en az iki veri setinde doğrulandı. İlk defa bu çalışmada tanımlanan bölgelerin işlevsel etkilerini bulmak için, bu bölgelere bağlanan transkripsiyon faktörleri tanımlandı. ERBB4 ve NRG1 genlerinin bu bölgelerine bağlanan transkripsiyon faktörlerinin şizofreniye yatkınlık yaptığı ya da hastalıkta koruyucu etkisi olduğu gözlemlendi. Bu çalışma, ilk defa, GWAS veri setlerinde *in silico* ve biyoinformatik tabanlı işlevsel analiz yöntemleri kullanarak, intronik varyasyonların şizofreni üzerinde etkisini tanımlamaktadır. Tez kapsamında kullanılan yöntemlerin, intronik değişimlerin işlevlerini anlamak açısından yüreklendirici sonuçlar verdiği gösterilmiştir.

# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ..... iv
ABSTRACT ..... v
ÖZET ..... vi
LIST OF FIGURES ..... ix
LIST OF TABLES ..... xi
LIST OF SYMBOLS ..... xii
LIST OF ACRONYMS/ABBREVIATIONS ..... xiii
1. INTRODUCTION ..... 1
1.1. Schizophrenia ..... 2
1.1.1. Molecular Biology of Schizophrenia ..... 3
1.1.2. Genetic Architecture of Schizophrenia ..... 7
1.1.2.1. Genome-wide association studies ..... 9
1.1.2.2. Other association studies ..... 17
1.2. ERBB4-NRG1 Signalling in Schizophrenia ..... 19
1.2.1. The Schizophrenia-Associated Variants in ERBB4 ..... 21
1.2.2. The Schizophrenia-Associated Variants in NRG1 ..... 22
2. PURPOSE ..... 25
3. MATERIALS ..... 26
3.1. Subjects ..... 26
3.1.1. CATIE Study ..... 26
3.1.2. MGS Study ..... 26
3.2. Genotyping ..... 27
3.2.1. CATIE Dataset ..... 27
3.2.2. GAIN and nonGAIN Datasets ..... 28
3.3. Equipments ..... 28
3.4. Electronic Databases ..... 28
3.5. Bioinformatic Tools ..... 31

4. METHODS ..... 32
4.1. Literature Search ..... 32
4.1.1. GWAS Catalog ..... 32
4.1.2. SZGene Database ..... 33
4.1.3. Peer-reviewed Journals in PubMed ..... 33
4.2. Pathway Analysis ..... 34
4.3. Collection of Genome-wide Association Datasets ..... 34
4.4. Logistic Regression Analysis ..... 34
4.5. Haplotype Analysis ..... 35
4.6. Examination of Transcription Factor Binding Sites ..... 36
5. RESULTS ..... 37
5.1. Candidate Schizophrenia Genes in the Literature ..... 37
5.2. Pathway Analysis of Candidate Schizophrenia Proteins ..... 39
5.3. Study Group Quality Checks of CATIE, GAIN and nonGAIN datasets ..... 41
5.4. Examination of Variants and Haplotypes in Three Datasets ..... 45
5.4.1. Overview of SNPs of ERBB4 and NRG1 using GWADView ..... 45
5.4.2. Narrowing Down the Significant ERBB4 and NRG1 Regions ..... 45
5.5. Mapping Transcription Binding Sites Altered Within the Blocks ..... 53
5.5.1. Examination of Tag SNPs in the Blocks ..... 53
5.5.2. Examination of Rare SNPs in the Blocks ..... 55
5.6. Validation of Previously Identified ERBB4 and NRG1 Regions ..... 58
6. DISCUSSION ..... 60
6.1. Pathway Analysis of Candidate Genes ..... 61
6.2. Identifying Novel ERBB4 and NRG1 Regions ..... 62
6.3. Transcription Binding Sites Altered within the Blocks ..... 65
6.4. Validation of Previously Identified Regions ..... 66
REFERENCES ..... 68

# LIST OF FIGURES

Figure 1.1. Symptoms of schizophrenia and their effect in disease onset and outcome. .... 3

Figure 1.2. Proposed glutamate hypothesis of schizophrenia pathogenesis. ..... 4

Figure 1.3. Hypothetical roles of schizophrenia genes in glutamatergic, dopaminergic and neurodevelopment pathways. ..... 7

Figure 1.4. Three hypothetic models for the genetic architecture of schizophrenia. ..... 8

Figure 1.5. Binding specificities of ERBB receptor family ligand. ..... 19

Figure 1.6. The role of ERBB4-NRG1 signalling in NMDAR hypofunction. ..... 20

Figure 1.7. Schematic representation of NRG1 haplotypes associated with schizophrenia. ..... 23

Figure 4.1. Algoritm for selection and ranking of candidate schizophrenia genes from the GWAS catalog. ..... 32

Figure 4.2. Algoritm for selection and ranking of candidate schizophrenia genes from the SZGene database. ..... 33

Figure 5.1. Pathway analysis of candidate proteins in schizophrenia. ..... 41

Figure 5.2. Age distribution of cases and controls in the CATIE dataset. ..... 42

Figure 5.3. Age distribution of cases and controls in the GAIN dataset (a) before and (b) after quality checks. ..... 43

Figure 5.4. Age distribution of cases and controls in the nonGAIN dataset (a) before and (b) after quality checks. ..... 44

Figure 5.5. Visualization of the SNPs in the ERBB4 gene and its ±20 kb flanking region, according to their significance and chromosome location with the GWADView software. ..... 46

Figure 5.6. Visualization of the SNPs in the NRG1 gene and its ±20 kb flanking region according to their significance and chromosome location with the GWADView software. ..... 47

Figure 5.7. Haplotype analysis of the ERBB4 region using the Haploview software. ..... 51

Figure 5.8. Haplotype analysis of the NRG1 region using the Haploview software. ..... 52

Figure 5.9. Transcription factors that bind common and significant haplotypes of ERBB4. ..... 54

Figure 5.10. Transcription factors that bind common and significant haplotypes of NRG1. ..... 55

# LIST OF TABLES

Table 1.1. Summary of genome-wide association studies in schizophrenia. ..... 10

Table 3.1. Gender distribution of cases and controls in the MGS study. ..... 27

Table 3.2. Age distribution of cases and controls in the MGS study. ..... 27

Table 3.3. Equipment used in this study. ..... 28

Table 5.1. Candidate "Rank 1" schizophrenia genes that were collected from the GWAS Catalog, SZGene and Pubmed databases. ..... 37

Table 5.2. Top five molecular and cellular functions. ..... 39

Table 5.3. Top five physiological system development and function pathways. ..... 40

Table 5.4. Top five canonical pathways. ..... 40

Table 5.5. The significant haplotype block within ERBB4 as a result of haplotype-based logistic regression analysis. ..... 48

Table 5.6. The significant haplotype block within NRG1 as a result of haplotype-based logistic regression analysis. ..... 49

Table 5.7. Rare SNPs in six-kb ERBB4 and 25-kb NRG1 regions that change transcription binding sites. ..... 56

Table 5.8. Previously identified ERBB4 and NRG1 regions and their validations. ..... 59

# LIST OF SYMBOLS

| | |
|---|---|
| α | alpha |
| $R^2$ | R square |
| Bp | base pair |
| Mb | megabase |
| Kb | kilobase |
| P | p value |
| Chr | chromosome |

# LIST OF ACRONYMS/ABBREVIATIONS

| | |
|---|---|
| AA | African American |
| ACSM1 | acyl-CoA synthetase medium-chain family member 1 |
| ADAMTSL3 | ADAMTS-like protein 3 |
| AKT | v-akt murine thymoma viral oncogene homolog 1 |
| ALS2CL | ALS2 C-terminal like |
| AMBRA1 | autophagy/beclin-1 regulator 1 |
| ANK3 | ankyrin 3 |
| APBA2 | amyloid beta (A4) precursor protein-binding, family A, member 2 |
| ARGHAP18 | GTPase activating protein 18 |
| c/EBPa | enhancer binding protein (C/EBP), alpha |
| cAMP | Cyclic adenosine monophosphate |
| CAPON | nitric oxide synthase 1 (neuronal) adaptor protein |
| CATIE | Clinical Antipsychotic Trials of Intervention |
| CCDC60 | coiled-coil domain containing 60 |
| CDCV | Common Disease Common Variant |
| CDRV | Common Disease Rare Variant |
| CEACAM21 | carcinoembryonic antigen-related cell adhesion molecule 21 |
| CENTG2 | ArfGAP with GTPase domain, ankyrin repeat and PH domain |
| CEU | Utah residents with Northern and Western Europe ancestry |
| CNV | Copy Number Variation |
| COMT | Catechol-O-methyltransferase |
| CSF2RA | colony stimulating factor 2 receptor α |

| | |
|---|---|
| CSMD1 | CUB and Sushi multiple domains 1 |
| CYT1/2 | Cytoplasmic tail region 1/2 |
| D1/2/3 | Dopamine 1/2/3 |
| DAAO | D-amino acid oxidase |
| dbGAP | Database of Genotypes and Phenotypes |
| dbSNP | Single Nucleotide Polymorphism Database |
| DISC1 | disrupted in schizophrenia 1 |
| DOCK4 | dedicator of cytokinesis 4 |
| DSM-III/IV | Diagnostic and Statistical Manual of Mental Disorders-III/IV |
| DTNBP1 | Dysbindin |
| EA | European American |
| EFCAB2 | EF-hand calcium binding domain 2 |
| ELAVL2 | ELAV (embryonic lethal, abnormal vision, Drosophila)-like 2 |
| EML5 | echinoderm microtubule associated protein like 5 |
| ERBB4 | V-erb-a erythroblastic leukemia viral oncogene homolog 4 |
| ERK2 | mitogen-activated protein kinase 1 (MAPK1) |
| FASTSNP | Functional Analysis and Selection Tool for SNP |
| FHIT | fragile histidine triad gene |
| GABA | γ-aminobutyric acid |
| GAD67 | glutamate decarboxylase 67 |
| GAIN | The Genetic Association Information Network |
| GATA1/3 | GATA binding protein 1/3 |
| GM-CSF | Granulocyte-macrophage colony-stimulating factor |
| GRB2 | growth factor receptor-bound protein 2 |
| GRIK3 | glutamate receptor, ionotropic, kainite 3 |

| | |
|---|---|
| GRIN2B | glutamate receptor, ionotropic, N-methyl D-aspartate 2B |
| GWAS | Genome-wide Association Studies |
| HAPICE | Icelandic haplotype |
| HapMap | International Haplotype Map Project |
| HFH2 | HNF3/FH transcription factor genesis |
| HWE | Hardy-Weinberg equilibrium |
| ICCSS | Irish Case-Control Study of Schizophrenia |
| ICD | Intrecellular domain |
| IL3RA | interleukin 3 receptor α |
| IPA | Ingenuity Pathway Analysis |
| ISC | International Schizophrenia Consortium |
| JMa/b | extracellular juxtamembrane domain a/b |
| KCNIP4 | Kv channel interacting protein 4 |
| KIF26B | kinesin family member 26B |
| KPNA1 | karyopherin alpha 1 |
| LD | Linkage disequilibrium |
| LRP1 | low density lipoprotein receptor-related protein 1 |
| LSD | Lysergic acid diethylamide |
| MAF | Minor allele frequency |
| MGS | Molecular Genetics of Schizophrenia |
| MHC | Major histocompatibility complex |
| MIR137 | microRNA 137 |
| MMP16 | matrix metallopeptidase 16 |
| MSRA | methionine sulfoxide reductase A |
| NCBI | National Center for Biotechnology Information |

| | |
|---|---|
| NEK1 | NIMA (never in mitosis gene a)-related kinase 1 |
| NHGRI | National Human Genome Research Institute |
| NKX2 | NK2 homeobox |
| NMDAR | N-methyl-D-aspartate receptor |
| NOS1AP | nitric oxide synthase 1 adaptor protein (as known as CAPON) |
| NR1 | NMDA receptor subunit 1 |
| NRG1 | Neuregulin |
| NRGN | neurogranin |
| NRXN1 | neurexin 1 |
| OCT1 | organic cation transporter, member 1 |
| OPCML | opioid binding protein/cell adhesion molecule-like |
| PCGEM1 | prostate-specific transcript 1 |
| PCGEM1 | prostate-specific transcript 1 |
| PCP | phencyclidine |
| PLAA | phospholipase A2-activating protein |
| PPP3CC | protein phosphatase 3, catalytic subunit, gamma isozyme |
| PRODH | Proline dehydrogenase |
| PTBP2 | polypyrimidine tract binding protein 2 |
| PTPN21 | protein tyrosine phosphatase, non-receptor type 21 |
| RELN | Reelin |
| RGS4 | regulator of G-protein signaling 4 |
| RNA | Ribonucleic acid |
| SNAP | SNP Annotation and Proxy Search |
| SNP | Single Nucleotide Polymorphism |
| SULT6B1 | sulfotransferase family, cytosolic, 6B, member 1 |

| | |
|---|---|
| SZGENE | SchizophreniaGene |
| TAAR6 | trace amine associated receptor 6 (as known as TRAR4) |
| TCF4 | transcription factor 4 |
| TF | Transcription factor |
| UCSC | University of California Santa Cruz |
| UK | United Kingdom |
| USA | United States of America |
| USF | upstream transcription factors |
| Wt | Wild type |
| ZNF480 | zinc finger protein 480 |
| ZNF804A | zinc finger protein 804A |

# 1. INTRODUCTION

Psychiatric disorders imply for diverse conditions related to mental health and to interruption in the most crucial human characteristics, such as communication, perception, emotions, language and "sense of self" (Taylor, 2011). There are more than ten different subgroups of psychiatric disorders, but the most common ones are non-affective psychotic disorders (schizophrenia, schizoaffective disorder and delusional disorder), affective psychosis (such as bipolar disorder and major depression) and substance-induced psychotic disorder caused by addiction to alcohol or other substances (van Os and Kapur, 2009). Frustratingly, 26.2% of adults are suffering from a mental disorder in a given year in the United States, and 6% of all population have a severe mental condition (Kessler *et al.*, 2005). The statistics are terrifying: In the United Kingdom, 250000 people are admitted into psychiatric hospitals each year, and more than 4000 of those people kill themselves. In this sense, mental disorders take more lives than cancer and physical illnesses (van Os and Kapur, 2009).

The distinct line between different types of mental disorders is not very clear. Genetic and molecular biology studies show that the disease mechanisms are very close to each other and comorbidity rates among psychiatric disorders are high (Kessler *et al.*, 2005).

Schizophrenia is one of the most common illnesses among all psychiatric disorders. It was first characterized in the early $19^{th}$ century on two patients, James Tilly Matthews, and Phillipe Pinel (Heinrichs, 2003). Life time prevalence of schizophrenia increases up to 2-3% when related categories are also included (Lichtenstein *et al.*, 2009). The disease mechanism is complex and the disease risk is affected by both genetic and environmental factors.

## 1.1. Schizophrenia

Schizophrenia is a common neuropsychiatric disorder with an estimated lifetime prevalence of 1% (Lichtenstein *et al.*, 2009). It is the third-leading cause of disability with a suicide rate of approximately 5% and an increase of 50% in mortality rates (Goff *et al.*, 2005; Palmer *et al.*, 2005). The name of the disease comes from Greek roots "to split" and "mind" and was used to be characterized by "the disconnection or splitting of the psychotic functions" (Picchioni and Murray, 2007). However this name is mislead in regarding the disease symptoms. Schizophrenia is characterized by three types of symptoms: positive symptoms (delusions and hallucinations), negative symptoms (reduced emotions, interrupted speech, interest loss) and mood symptoms (Figure 1.1) (Craddock *et al.*, 2005). Most antipsychotic drugs are able to treat positive symptoms, but negative and mood symptoms are still very difficult to cure (Fleischhacker and Widschwendter, 2006). The heritability of schizophrenia was estimated as 80% in a meta-analysis of 12 twin studies (Sullivan *et al.*, 2003); however the symptoms and prognosis do not form a distinct familial subtype (Tandon *et al.*, 2008). Schizophrenia is a complex disorder and both genetic variants and environmental factors are effective in its pathophysiology (Prasad *et al.*, 2002). There are several environmental factors that might affect risk of developing schizophrenia: Prenatal famine, in utero exposure to influenza, urbanicity during upbringing, cannabis use, paternal age and migration (Pedersen and Mortensen, 2001; Selten *et al.*, 2007; Sipos *et al.*, 2004; Susser *et al.*, 1996; Takei *et al.*, 1996; Weiser and Noy, 2005).

Gender differences in schizophrenia have been identified for a long time especially regarding risk rate, age of onset and course of the disease. The risk of developing schizophrenia is 1.4 times higher in males relative to females (McGrath *et al.*, 2004). The age of onset in females is four to five years higher than males: While onset of schizophrenia peaks at age 20-25 years in males, it is 25-30 years and 45 years in females (Riecher-Rossler and Hafner, 2000). It was also reported that female patients have a more favorable disease course than males, especially in younger women, who have a better course than older women. This statistics gave rise to "estrogen hypothesis" in schizophrenia which is explained by the

protective effect of the estrogen hormone against schizophrenia between puberty and menopause when the hormone level is high (Riecher-Rossler and Seeman, 2002).

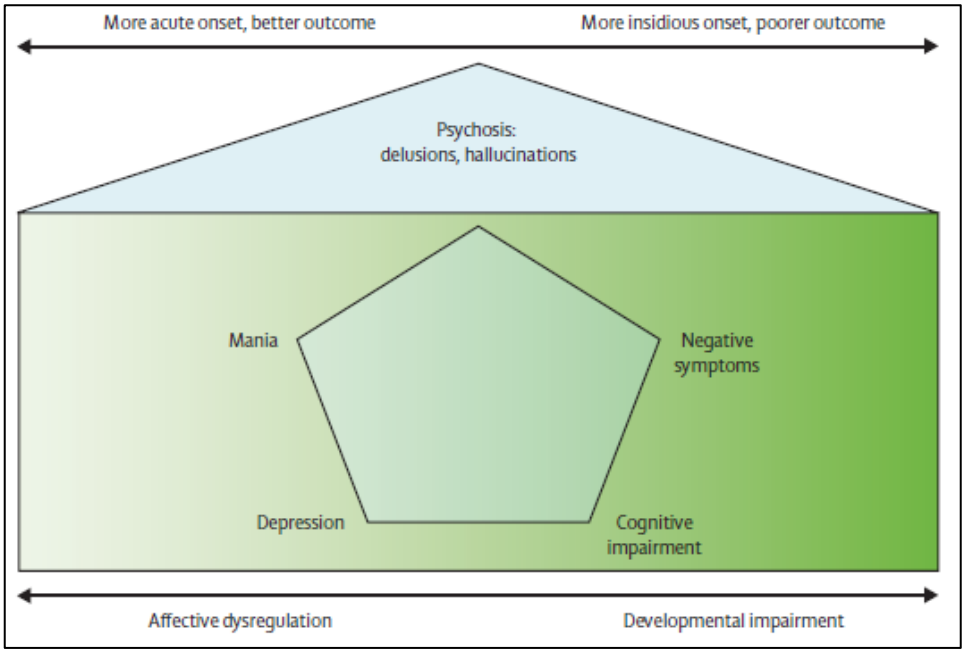



Figure 1.1. Symptoms of schizophrenia and their effect in disease onset and outcome (van Os and Kapur, 2009).

### 1.1.1. Molecular Biology of Schizophrenia

1.1.1.1. Glutamate Hypothesis. Glutamate is the major excitatory neurotransmitter in the mammalian central nervous system (Gordon, 2010). Glutamate in the post-synaptic cells binds to its receptor, *N*-methyl-D-aspartate receptor (NMDAR) and activates it (Kantrowitz and Javitt, 2010). Glutamate-induced NMDAR activity is crucial for integrity of presynaptic glutamate pathways and postsynaptic regulation and scaffolding. The role of NMDAR has been revealed two decades ago, with the observation that usage of some drugs such as phencyclidine (PCP) and ketamine causes psychosis and social withdrawal which are typical symptoms of schizophrenia (Krystal *et al.*, 2002). These two drugs are NMDAR antagonists

which lead to inhibition of receptor activity. Since then several studies have been performed to reveal glutamate hypothesis in schizophrenia.

One observation that supports the glutamate hypothesis in schizophrenia is γ-aminobutyric acid (GABA) deficits in excitatory neurons (Belforte *et al.*, 2010). GABA is the principal inhibitory neurotransmitter in the vertebrate brain (Lewis and Moghaddam, 2006). GABA is synthesized from glutamate, majorly by glutamate decarboxylase 67 (GAD67) in inhibitory neurons. Belforte *et al.* performed conditional knockdown of the NR1 subunit of the *NMDAR* in inhibitory interneurons of the hippocampus and cortex (Belforte *et al.*, 2010). Decreased NMDAR activity in inhibitory interneurons causes reduction in neuron activity leading to less GABA secretion from inhibitory interneurons to exhibitory interneurons (Figure 1.2). This feedback mechanism is essential to keep balance between inhibition and excitation for proper circuit function which is disrupted in schizophrenia. It was also shown that both RNA and protein levels of the GAD67 enzyme were lowered in the brains of patients with schizophrenia (Curley *et al.*, 2011).

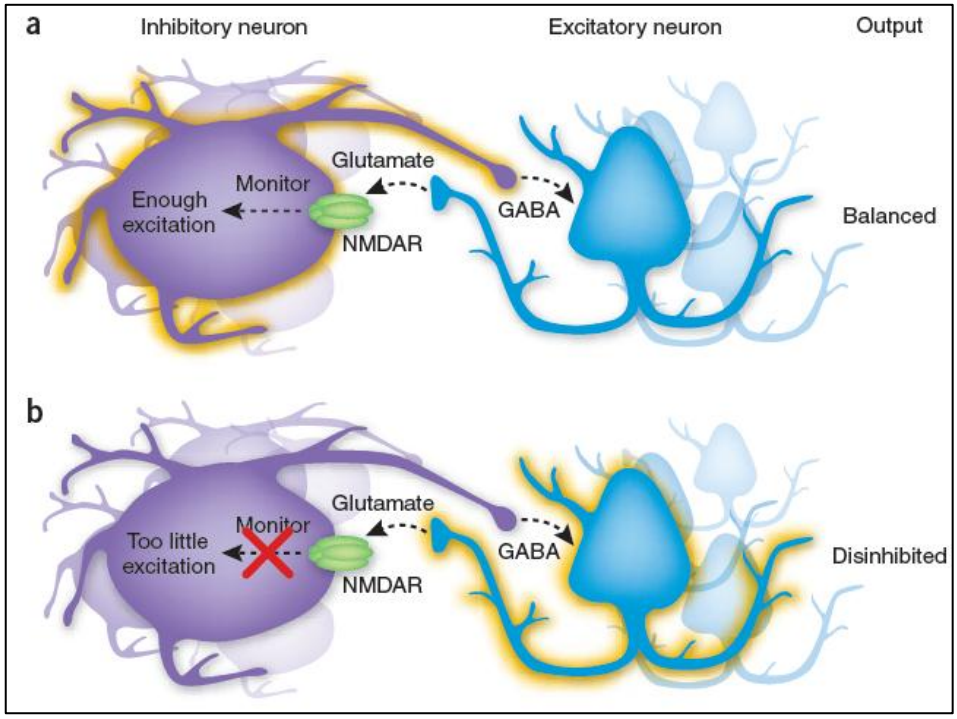



Figure 1.2. Proposed glutamate hypothesis of schizophrenia pathogenesis (Gordon, 2010).

Different NMDAR antagonists lead to different types of symptoms of schizophrenia. For example, while the effect of PCP is regarded as a negative symptom of schizophrenia, amphetamine- or LSD-based psychosis give rise to positive symptoms (Gouzoulis-Mayfrank *et al.*, 2005). These findings implicate that NMDAR hypofunction has more than one downstream pathway that act in different ways leading to schizophrenia. In spite of intensive efforts to understand glutamate hypothesis of schizophrenia, very little is known about the genetic linkage between schizophrenia and glutamate receptor subunits or other components of the signaling pathway (Gordon, 2010).

1.1.1.2. Dopamine Hypothesis. Like glutamate hypothesis, dopamine hypothesis was evolved from animal studies observing that neuroleptic drugs blocked dopamine receptors in the brain (Carlsson and Lindqvist, 1963). However three years after this observation, dopamine, a neurotransmitter, was implicated as a leading cause in schizophrenia (van Rossum, 1966). Dopamine plays important roles in motor activity, attention, executive functions, which are all disrupted in schizophrenia, in both animals and humans (Goerendt *et al.*, 2003; Nieoullon and Coquerel, 2003).

At cellular level, a direct measurement of dopamine level in humans is impossible; however some techniques were developed for the indirect determination of dopamine synthesis (Howes and Kapur, 2009). Several studies showed elevated dopamine synthesis in presynaptic neurons of schizophrenia patients (Howes *et al.*, 2007). However the results from measurement of dopamine receptor activities were not consistent. While some of the studies observed, increased D2/3 receptor activity in schizophrenia patients relative to controls, others detected no differences between both groups (Davis *et al.*, 1991). Besides D2/3 receptors, there are controversial outcomes from the measurement of D1 receptor activity that mediates dopamine transmission in the prefrontal cortex (Tamminga, 2006). While one study in this field showed increased level of D1 density in drug-free schizophrenia patients (Abi-Dargham *et al.*, 2002), two other studies observed either decreased levels or no differences (Karlsson *et al.*, 2002; Okubo *et al.*, 1997). Although after these inconsistent findings dopamine hypothesis was no longer considered as the ultimate cause of schizophrenia, many recent studies revealed that genetic variants in the genes related to the dopaminergic pathway were associated with

schizophrenia risk (Arguello and Gogos, 2008; Zheng *et al.*, 2012); thus it is not rational to reject the dopamine hypothesis (Moncrieff, 2009). Moreover there is supportive evidence for a combinatorial effect of glutamatergic and dopaminergic pathways in developing schizophrenia; such that dopaminergic hypersensitivity might be triggered by NMDAR hypofunction (Smith *et al.*, 1998).

1.1.1.3. Neurodevelopmental Hypothesis. According to neurodevelopmental hypothesis, developing schizophrenia includes pathologic processes that are caused by both genetic and environmental factors, beginning from very early stages of brain development and proceeding through the puberty (Rapoport *et al.*, 2005). This phenomena is known as "two-hit model" and implicates the role of maldevelopment in two critical time points in schizophrenia risk (Keshavan, 1999). Related to the neurodevelopmental hypothesis, enlargement of the cerebroventricular system was observed in patients with schizophrenia (Northoff *et al.*, 1999). Besides schizophrenia patients, structural abnormalities in brain were also observed in individuals at high risk for schizophrenia and in unaffected first-degree family members of schizophrenia patients (Wright *et al.*, 2000). Another support for neurodevelopmental hypothesis came from the observation of neurological soft signs in children who later develop schizophrenia (Fatemi and Folsom, 2009). Some of the genes that are involved in signal transduction, cell growth and migration, myelination and regulation of presynaptic membrane function are associated with schizophrenia: Neuregulin (*NRG1*), v-erb-a erythroblastic leukemia viral oncogene homolog 4 (*ERBB4*), dysbindin (*DTNBP1*), d-amino acid oxidase (*DAAO*), regulator of G-protein signaling 4 (*RGS4*), catechol-O-methyltransferase (*COMT*), proline dehydrogenase (*PRODH*) and reelin (*RELN*) (Chung *et al.*, 2003; Hakak *et al.*, 2001; Mirnics *et al.*, 2000; Tkachev *et al.*, 2003).

Although glutamate, dopamine and neurodevelopmental hypotheses are explained as independent pathways above, these three pathways jointly act in the risk of developing schizophrenia, as depicted in Figure 1.3.

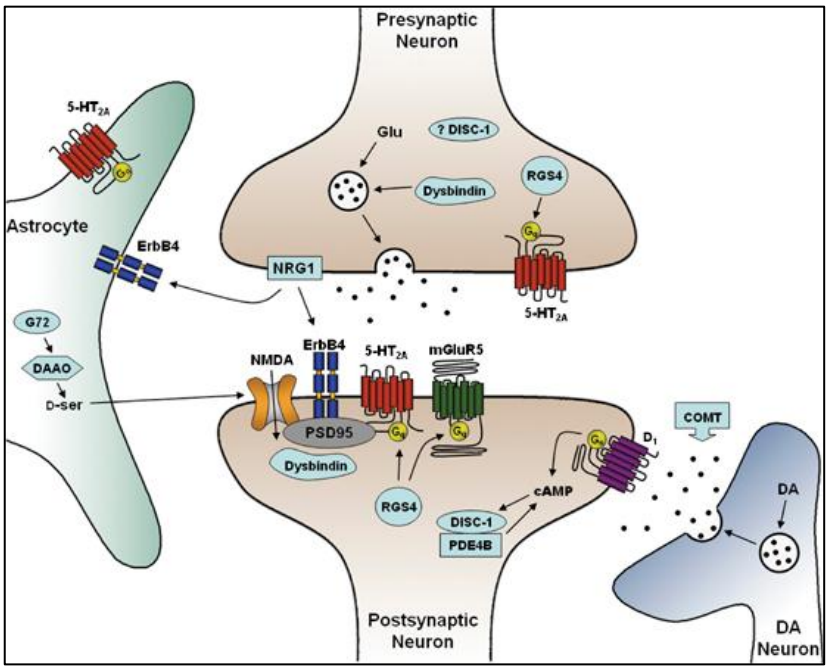



Figure 1.3. Hypothetical roles of schizophrenia genes in glutamatergic, dopaminergic and neurodevelopment pathways (Gray and Roth, 2007).

### 1.1.2. Genetic Architecture of Schizophrenia

Schizophrenia is defined as a complex disease, this implicates that it does not show a simple Mendelian pattern of inheritance (Lander and Schork, 1994). The complex and heterogeneous genetic background of schizophrenia is formulated with different hypotheses. The major two of them are "common disease-rare variant" (CDRV) and "common disease-common variant" (CDCV) models. The "common disease-rare variant" model supports the presence of rare variants with strong effect in each family with schizophrenia (Manolio *et al.*, 2009). These variants might differ from family to family and might be in the same or in different genes. This notion was supported by McClellan *et al.* due to following reasons:

- Schizophrenia is a familial disorder; the first-degree relatives of schizophrenia patients are at high risk of developing the disease.
- Paternal age is associated with increased risk of schizophrenia.
- Schizophrenia is associated with decreased fertility (McClellan *et al.*, 2007).

On the other hand, there are other strong evidences to question the CDRV model and to support the "common disease-common variant" hypothesis. The model suggests that multiple common variants with modest or weak effects jointly act in schizophrenia (Chakravarti, 1999). Contradictions to CDRV can be summarized as:

- Clinicians were not able to define a schizophrenia family with a single gene inheritance.
- So far, no single variant with a full penetrance was identified.
- Since even monozygotic twins carry 50% of schizophrenia risk, the presence of a single variant with a large effect is mathematically impossible.
- Most of the molecular genetic findings are consistent with CDCV model (Craddock *et al.*, 2007).

A recent "compound model" (Figure 1.4) assumes the presence of a rare variant modified by other common variants and environmental factors (Rodriguez-Murillo *et al.*, 2012). There are also other models that were proposed for schizophrenia which are dynamic mutations (trinucleotide expansions), genomic imprinting and mitochondrial inheritance (Ben-Shachar and Laifenfeld, 2004; Singh *et al.*, 2002). No matter what the model is, molecular genetic studies continue to reveal several variants that result in schizophrenia risk via different types of genomic approaches.

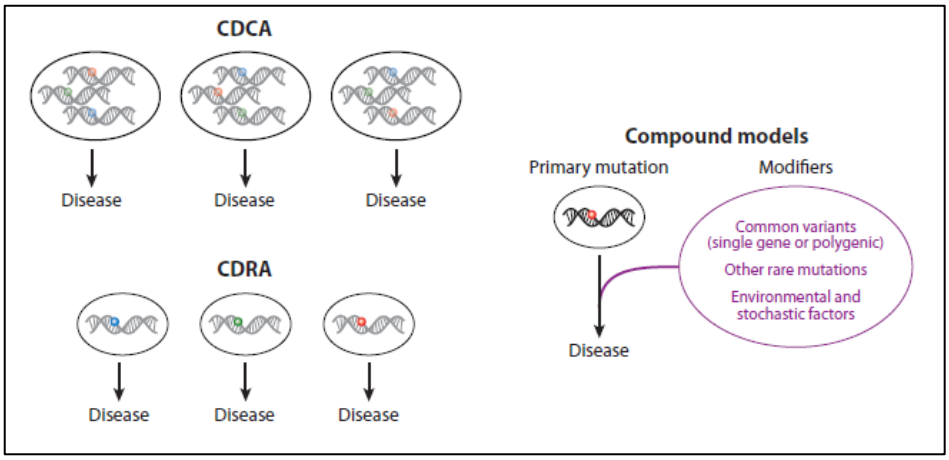



Figure 1.4. Three hypothetic models for the genetic architecture of schizophrenia (Rodriguez-Murillo *et al.*, 2012).

1.1.2.1. Genome-wide association studies. So far, a total of 21 genome-wide association studies (GWASs) in schizophrenia have been published (Table 1.1). The first paper in 2007 used a relatively small number of subjects compared to the following studies (Lencz *et al.*, 2007). In this work, 178 Caucasian cases and 144 sex- and ethnicity-matched controls were recruited to genome-wide study using two different Affymetrix chips containing ~262000 and ~238000 single nucleotide polymorphisms (SNPs). Among 439511 SNPs that passed the quality control, only one SNP (rs4129148) could achieve the genome significance of $4.2 \times 10^{-7}$. The C allele of rs4129148, that is located at the 5' region of the colony stimulating factor 2 receptor α (*CSF2RA*) gene, was shown to be over-represented in schizophrenia patients. In this study, *CSF2RA* and the neighbouring gene interleukin 3 receptor α (*IL3RA*) were subjected to haplotype analysis. One haplotype block in the intron 8 of *CSF2RA* and another haplotype block containing three variants in intron 4, 5 and 6 of *IL3RA* were significantly associated with schizophrenia in the Caucasian population. This study is the first example of a genome-wide association study in schizophrenia and for the first time reveals the role of *CSF2RA* and *IL3RA* in the disease.

The second GWAS in schizophrenia was performed using 650 cases and 2771 controls with a replication cohort of 745 cases and 759 controls from an Ashkenazi Jewish population (Shifman *et al.*, 2008). A variant in the fourth intron of the *RELN* gene, rs7341475, was correlated with schizophrenia only in women. While the G allele was overrepresented in women with the disease relative to healthy women ($p= 1.9 \times 10^{-5}$), men showed no association with this polymorphism. Results emerging from this study gave genetic support for the sex bias in the risk of schizophrenia.

Attempts to validate the findings of the first GWAS failed in the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) study that used genotype data from 738 schizophrenia patients and 733 age- and ethnicity-matched controls (Sullivan, 2008). The genotype data were produced using Affymetrix 500K "A" chipset and a custom 164K chip created by Perlegen. The significant association of *CSF2RA* and *IL3RA* with SCZ that was shown in Lencz et al (Lencz *et al.*, 2007) could not be validated. In addition, none of the 492900 SNPs reached the genome-wide significance threshold, but

Table 1.1. Summary of genome-wide association studies in schizophrenia.

| **Reference** | **Initial Sample Size (cases- controls)** | **Initial Sample Population** | **Replication Sample Size (cases-controls)** | **Replication Sample Population** | **Reported Regions** | **Reported Genes** | **Platform** |
|---|---|---|---|---|---|---|---|
| Lencz et al., 2007 | 178-144 | Caucasian | - | - | Xp22.33<br>Yp11.32 | *CSF2RA*<br>*IL3RA* | Affymetrix |
| Shifman et al., 2008 | 660- 2271 | Ashkenazi Jewish | 2274-4401 | UK, Irish, USA, Chinese | 7q22.1 | *RELN* | Affymetrix 500K array |
| Sullivan et al., 2008 | 738-733 | USA | - | - | 15q25.3<br>1q31.1<br>1p36.22<br>13q12.2<br>16p12.3<br>Xq28 | *AGBL1*<br>Intergenic<br>Intergenic<br>Intergenic<br>*ACSM1/BUCS1*<br>Intergenic | Affymetrix 500K "A" chipset |
| Walsh et al., 2008 | 150-268 | - | 83 children, 154 parents | - | - | - | Illumina 550K Microarray |
| O'Donovan, 2008 | 479-2937 | UK | 6666-9897 | UK Bipolar Samples | 2q32.1<br>11p14.1<br>16p13.12 | *ZNF804A*<br>Intergenic<br>Intergenic | Affymetrix 500K Mapping array |
| Kirov, 2009 | 574-605,<br>1,148 parents | Bulgaria | - | - | 12q24.23 | *CCDC60* | Illumina HumanHap 550 array |
| Need, 2009 | 871-863 | European | 1460-12995 | European | 15q25.2 | *ADAMTSL3* | Illumina HumanHap 300 & 550K Chip |
| Purcell, 2009 | 3322- 4692 | European | 15493- 3587 | European | 6p22.1<br>18q21.2<br>3q26.33<br>1p21.3 | *MHC*<br>*TCF4*<br>*FXR1*<br>*PTBP2* | Illumina |
| Shi, 2009 | 2681-2653,<br>1286- 973 | European<br>African-American | 5327-16424 | European | 6p22.1<br>6p21.32 | *MHC region* | Affymetrix 6.0 array |
| Stefansson, 2009 | 2663-13498 | European | 10282-21093 | European | 6p22.1<br>11q24.2<br>18q21.2<br>2p16.1<br>5q21.1<br>9q33.1 | *MHC,*<br>*NRGN*<br>*TCF4*<br>*VRK2*<br>*SLCO6A1*<br>Intergenic | Illumina HumanHap 300 & 550K Chip |

Table 1.1. Summary of genome-wide association studies in schizophrenia (cont.).

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Athanasiu, 2010 | 201-305 | Norwegian | 2663- 13780 | European | 9p21.2<br>16p12.3<br>10q21.2 | *PLAA*<br>*ACSM1*<br>*ANK3* | Affymetrix 6.0 array |
| Alkelai, 2011 | 189 individuals from 57 families | Arab-Israeli | 627-541 | European | - | - | Illumina HumanCNV 370 BeadArrays |
| Alkelai, 2011 | 331 family members | Jewish-Israeli | 189 individuals, 57 nuclear families | Arab-Israeli | 19q13.2<br><br>10p11.22 | *ATP5SL*<br>*CEACAM21*<br>*NRP1*<br>*PARD3* | Illumina HumanCNV 370 BeadArrays |
| Rietschel, 2011 | 1169-3714 | German Dutch | 7303- 26274 | European | chr6<br>11p11.2 | *ARGHAP18*<br>*AMBRA1* | Illumina HumanHap 550v3 BeadArray |
| Yamada, 2011 | 120 parent-child trios | Japanese | 506-506,<br>293 families | Japanese Chinese | 9p21.3 | *ELAVL2* | Affymetrix 100K GeneChip Mapping |
| Ma, 2011 | 98-60 | Chinese | - | - | 8p23.1 | *MSRA* | Illumina HumanHap 660 Beadchip |
| Ikeda, 2011 | 560-548 | Japanese | 1511-2451<br>479-2938 | Japanese U.K | 2p22.2 | *SULT6B1* | |
| Chen, 2011 | 1658-1655 | European | 5203- 5277<br>1875<br>1142-985 | European<br><br>Irish<br>African-American | 14q31.3<br>14q31.3 | *PTPN21*<br>*EML5* | - |
| Ripke, 2011 | 9394-12462 | European | 8442- 21397 | European | 1p21.3<br>2q32.3 8p23.2<br>8q21.3<br>10q24.32-q24.33 | *MIR137*<br>*PCGEM1*<br>*CSMD1*<br>*MMP16*<br>multiple genes | - |
| Yue, 2011 | 746- 1599 | Han Chinese | 4027 cases,<br>5603 controls | Han Chinese | 6p22.1<br>11p11.2 | *NKAPL*<br>*TSPAN18* | Illumina Human 610-Quad BeadChips |
| Shi, 2011 | 3750-6468 | Han Chinese | 4383- 4539 | Han Chinese | 8p11.23<br><br>1q24.2 | *LSM1*<br>*WHSC1L1*<br>*BRP44*<br>*DCAF6* | Affymetrix 6.0 array |

several supportive genes were proposed, such as acyl-CoA synthetase medium-chain family member 1 (*ACSM1*), kv channel interacting protein 4 (*KCNIP4*) and fragile histidine triad gene (*FHIT*).

GWASs are implied not only to identify common single nucleotide variants, but also to search for rare structural variants such as microdeletions and microduplications. In such a study, DNA samples from 150 schizophrenia cases and 268 healthy controls were assayed using the Illumina 550K microarray (Walsh *et al.*, 2008). This study showed that rare structural variants, which disrupted genes, were represented three times more in schizophrenia patients relative to healthy controls, whereas presence of variants with no effect on gene structure did not differ between case and control groups. The same study also revealed that genes disrupted by microdeletions/duplications function in neurodevelopmental pathways, such as neuregulin, axonal guidance and glutamate receptor signalling.

Overlaps in the molecular bases of neuropsychiatric disorders have been suggested previously (Schulze *et al.*, 2012). The genotyping of 362532 SNPs for 479 British cases and 2937 controls identified rs1344706 as the most significant SNP with a p value of $1.83 \times 10^{-6}$ (O'Donovan *et al.*, 2008). The significance of rs1344706 was replicated twice using two independent replication sample sets and was further reported in the British bipolar disorder cohort. Therefore the zinc finger protein 804A (*ZNF804A*), bearing rs1344706, provides a strong evidence for a common genetic basis of schizophrenia and bipolar disorder.

Parent-offspring trios in GWASs are commonly used to eliminate problems arising from population stratification. 574 unrelated schizophrenia patients, their 1148 healthy parents and 605 unrelated controls from Bulgaria were recruited to be assayed by the Illumina HumanHap550 array (Kirov *et al.*, 2009c). The most significant SNP with a p-value of $1.2 \times 10^{-6}$, rs11064768, was found to be located in the intronic region of coiled-coil domain containing 60 (*CCDC60*).

After microdeletions and microduplications were found to be correlated with the risk of schizophrenia, a large genome-wide association study was performed to analyse SNPs and

copy number variations (CNVs) in a European-ancestry population (Need *et al.*, 2009). Using a discovery cohort of 871 cases and 863 controls, the most significant SNP was found to be rs2135551 in the intron 29 of the ADAMTS-like protein 3 (*ADAMTSL3*) gene. However, the replication of association failed in an independent replication cohort that included a total of 1460 cases and 12995 healthy controls. On the other hand, previously reported schizophrenia-associated CNVs in two GWASs have been validated (Kirov *et al.*, 2009c; Walsh *et al.*, 2008). These variants were a 1.4 Mb duplication including amyloid beta (A4) precursor protein-binding family A member 2 (*APBA2*), deletions in neurexin 1 (*NRXN1*), a duplication encompassing EF-hand calcium binding domain 2 (*EFCAB2*) and kinesin family member 26B (*KIF26B*), a 1.4 Mb deletion at chromosome 1 and a 664 kb duplication at chromosome 2.

Three large genome-wide association studies of schizophrenia were published in the same volume of Nature in 2009 (Purcell *et al.*, 2009; Shi *et al.*, 2009; Stefansson *et al.*, 2009). In one of these studies, the major histocompatibility complex (MHC) region on chromosome 6 and the neighboring gene *NOTCH4* were identified as susceptibility factors for schizophrenia risk in the study of International Schizophrenia Consortium (ISC) that used 3322 European cases and 3587 controls (Purcell *et al.*, 2009). Moreover the association could be replicated in a sample collection consisting of 4692 schizophrenia patients and 15493 controls from Molecular Genetics of Schizophrenia (MGS) and SGENE project. European-ancestry samples (2681 cases and 2653 controls) and African American samples (1286 cases and 973 controls) were recruited for MGS study (Shi *et al.*, 2009). While ArfGAP with GTPase domain, ankyrin repeat and PH domain (as known as *CENTG2*) showed strong association with schizophrenia in the European-ancestry samples, *ERBB4* and its ligand *NRG1* were significantly correlated in African American samples. Meta-analysis with ISC and SGENE projects (plus 5327 cases and 16424 controls) revealed NIMA (never in mitosis gene a)-related kinase 1 (*NEK1*), polypyrimidine tract binding protein 2 (*PTBP2*), *MHC* region and transcription factor 4 (*TCF4*) as susceptible schizophrenia genes in the European population. Another study, called SGENE-plus was composed of 2663 cases and 13498 controls from England, Finland (Helsinki), Finland (Kussamo), Germany (Bonn), Germany (Munich), Iceland, Italy and Scotland (Stefansson *et al.*, 2009). Out of 314868 SNPs, the top 1500 markers were selected for replication using MGS-GAIN-EA and ISC. After replication, the

top seven SNPs were located in the MHC region, the TCF4 and the neurogranin (*NRGN*) genes were previously reported to be associated with schizophrenia risk in Portuguese males (Ruano *et al.*, 2008).

Since Scandinavians are ethnically homogenous and were not subject to immigration by other populations, they are generally considered as suitable cohorts for association studies. The TOP study includes genotype data of 459 Norwegian schizophrenia cases and 313 ethnicity-matched healthy controls by Affymetrix SNP Array 6.0 (Athanasiu *et al.*, 2010). Polymorphisms in the top genes from SZGene databases were analysed in this study and association of *DISC1*, *RELN*, *NRG1* and opioid binding protein/cell adhesion molecule-like (*OPCML*) were validated in the Norwegian population. Moreover combined analysis with SGENE-plus study (2663 European cases and 13780 controls) correlated the genes phospholipase A2-activating protein (*PLAA*), *ACSM1* that was reported in CATIE study (Sullivan, 2008), and ankyrin 3 (*ANK3*) which was previously associated with bipolar disorder (Baum *et al.*, 2008).

A GWAS of schizophrenia with a family-based design was implemented in 107 Jewish-Israeli families, consisting of a total of 331 individuals (Alkelai *et al.*, 2011b). The best SNP was rs2074127 within intron 6 of dedicator of cytokinesis 4 (*DOCK4*). The top seven SNPs were tried to be validated, using a previously published GWAS data by the same group that had 58 Arab-Israeli families with 198 individuals (Alkelai *et al.*, 2011a). In the replication cohort, the A allele of rs4803480 in the predicted intron of carcinoembryonic antigen-related cell adhesion molecule 21 (*CEACAM21*) was found to be significantly associated with schizophrenia risk.

Several genetic association studies of schizophrenia revealed that common risk alleles with very small effects have a combinatorial influence on the disease, thus it is difficult to identify reproducible results in GWASs. To overcome this problem, either very large samples must be collected or more homogeneous groups must be recruited from a particular regional area. As a model study, 1168 German and Dutch schizophrenia patients and 3714 matched controls were subjected to genotyping with Human-Hap550v3 Bead Array (Rietschel *et al.*,

2011). In this analysis, rs11154491 in the intron of *ARGHAP18* was found to be associated with schizophrenia. While the combined analysis with 2569 cases and 4108 controls for German, Dutch and Danish populations validated the association of *ARGHAP18*, four SNPs in intronic regions of the *AMBRA1* gene showed strong association. Out of the top 25 SNPs, emerging from a second replication cohort, which included 4734 patients and 18472 controls from 15 independent European populations, 11 were on chromosome 11, where *AMBRA1* and other genes related with neurodevelopmental pathways were located.

The threshold for GWASs are $5 \times 10^{-8}$ as proposed previously (McCarthy *et al.*, 2008). Although this very stringent threshold eliminates false positive results, it might sometimes cause variants with small effects to be missed. Yamada et al proposed a multistage analysis with moderate significance thresholds to cope with this problem (Yamada *et al.*, 2011). Stage I analysis was perfomed with 120 family trios from Japan assayed by Affymetrix GeneChip Mapping 100K. The significant variants here could not be replicated in Stage II analysis. In Stage II analysis, 1632 SNPs with a p value of <0.01 were assayed with Illumina Bead Array in 506 schizophrenia patients and 506 age- and sex-matched healthy Japanese controls. rs10491817 in the ELAV (embryonic lethal, abnormal vision, Drosophila)-like 2 (*ELAVL2*) gene was significantly associated with schizophrenia. In Stage III analysis, 56 tag SNPs in and around *ELAVLA2* were analysed in 293 pedigrees from the Chinese population, who are ethnically close to Japanese population (Wang *et al.*, 2000). Four of these SNPs were significant; however they could not remain after Bonferroni correction. Since *ELAVLA2* was not previously reported in schizophrenia, to implicate its role in the disease further association studies must be performed with larger number of samples.

Use of quantitative traits is a beneficial tool to strengthen the genetic association studies up to four to eight times (Wang *et al.*, 2005). A GWAS comprising 98 schizophrenia patients and 60 healthy controls from the Chinese population was performed using Illumina HumanHap 660 Beadchip. The variants were associated with fluid intelligence as an informative endophenotype of schizophrenia, and Cattell's Culture-free intelligence Test was applied to measure fluid intelligence, which was defined as a 'major measurable outcome of the influence of biological factors on intellectual development – that is heredity' (Horn and

Cattell, 1966). The study associated the *MSRA* gene with fluid intelligence in schizophrenia patients. *MSRA* plays a role in the oxidative stress pathway, which was shown to be disrupted in schizophrenia (Prabakaran *et al.*, 2007).

For a comprehensive Japanese GWAS, 575 patients-564 healthy controls and 1511 cases *versus* 2451 controls were recruited (Ikeda *et al.*, 2011). rs11895771 in the sulfotransferase family cytosolic 6B member 1 (*SULT6B1*) gene was significantly correlated with schizophrenia, as well as *NOTCH4*, that was previously associated with schizophrenia in a small UK sample (Wei and Hemmings, 2000).

After publishing large GWASs, meta-analysis using these studies was performed to analyse rich raw data. CATIE-EA (Sullivan, 2008) and MGS-GAIN-EA (Shi *et al.*, 2009) were used for initial analysis (Chen *et al.*, 2011). The SNPs with p values of $\leq 0.05$ were assigned with an empirical score as following: 2 for non-synonymous SNPs and SNPs within known genes; 1 for SNPs in evolutionary conserved regions, transcription factor binding sites, untranslated regions and synonymous SNPs; 0.5 for SNPs 2 kb up/downstream of a gene. Seven SNPs with top empirical scores were located in the protein tyrosine phosphatase, non-receptor type 21 (*PTPN21*) gene and six SNPs in the echinoderm microtubule associated protein like 5 (*EML5*) gene. These 13 SNPs were also replicated with combined analysis of MSG-nonGAIN (Shi *et al.*, 2009), IFAM and ICCSS (Chen *et al.*, 2006), ISC (Purcell *et al.*, 2009), CATIE-AA (Sullivan, 2008) and MGS-GAIN-AA (Shi *et al.*, 2009) samples. While the studies above could not validate each other, this type of meta-analysis resulted in two genes correlated with schizophrenia. There is therefore a persuasive necessity for alternate methods for mining info from GWA datasets. Another mega-analysis came in the same year, just a few months later (Ripke *et al.*, 2011). Stage I analysis with 17 published GWASs gave rise to 136 SNPs with genome-wide significance ($p< 5 \times 10^{-8}$). While 129 of them were located in MHC region, others were within either previously identified schizophrenia candidate genes (such as *TCF4* and *NRGN*) or in new regions, such as chr10q24.33 and chr8q21.3. In Stage II analysis that uses 19 independent studies, association of MHC region with schizophrenia was replicated. Moreover five new loci were identified in combined analysis of Stage I and II: chr1p21.3, 2q32.3, 8p23.2, 8q21.3 and 10q24.32-q24.33 which includes a transcript for

microRNA 137 (*MIR137*), prostate-specific transcript 1 (*PCGEM1*), CUB and Sushi multiple domains 1 (*CSMD1*), matrix metallopeptidase 16 (*MMP16*) and multiple genes, respectively.

Two most recent GWASs in schizophrenia were published simultaneously in the same journal (Shi *et al.*, 2011; Yue *et al.*, 2011). The first study used 746 Han Chinese schizophrenia patients versus 1599 healthy controls in the discovery cohort and 4027 Han Chinese cases versus 5603 controls in the replication cohort (Yue *et al.*, 2011). Two loci on chr6p21-p22.1 and chr11p11.2 were associated with schizophrenia risk in the Han Chinese population. Another study recruited 3750 patients and 6468 controls from northern, central and southern Han Chinese populations (Shi *et al.*, 2011). Association of chr8p12 and chr1q24.2 were also validated in BIOX sample (4383 cases 4539 controls) and in SGENE-plus sample (3830 patients and 14724 healthy individuals from European ancestry). Although new two loci conferred risk of schizophrenia in both Han Chinese and European populations, none of the previously identified schizophrenia candidate genes and variants could been validated, except the *MHC* region.

1.1.2.2. Other association studies. Genome-wide association studies support the emergence of the CDCV model. Besides GWAS, more than 30 linkage studies have been published (Rodriguez-Murillo *et al.*, 2012). However, reproducibility of their results was disappointing due to small number of samples and low coverage of markers (Owen *et al.*, 2005). Two meta-analyses of these linkage studies were performed (Badner and Gershon, 2002; Lewis *et al.*, 2003). The first one combined the published data that were analysed using different methods and supported involvement of chr8p, 13q and 22p loci (Badner and Gershon, 2002). The other study collected data from published and also unpublished studies that used the same analysis methods (Lewis *et al.*, 2003). While 2q was strongly associated with schizophrenia loci, 5p, 3p, 11q, 6p, 1q, 22q, 8p, 20q and 14p showed significant linkage.

Positional studies also revealed promising candidate genes for schizophrenia. Some of the strongest ones are *NRG1, DTNBP1, DAOA*, disrupted in schizophrenia 1 (*DISC1*), *RGS4*, nitric oxide synthase 1 (neuronal) adaptor protein (as known as *CAPON*), protein phosphatase 3 catalytic subunit gamma isozyme (*PPP3CC*) and trace amine associated receptor 6 (known

as *TRAR4*) (Owen *et al.*, 2005). The association of these genes with the disease also validated using neuroimaging, gene expression studies and other platforms of neurobiology (Doherty *et al.*, 2012).

Supportive evidence for CDRV model comes from studies that identify rare copy number variations. The first de novo CNV identified was a microdeletion at 22q11.2, which accounts for up to 1-2% of schizophrenia cases (Karayiorgou *et al.*, 1995). Rare deletions and duplications at 1q21.1, 15q11.2, 15q13.3, 16p11.2, 16p13.1 were identified in the following CNV study (Sebat *et al.*, 2009). These CNVs with odds ratio of 3-20 confer strong effect on schizophrenia risk (Kirov *et al.*, 2009a). Although associated CNVs span multiple genes, it is not clear which of those lead to disease development, with one exception of the *NRXN1* gene (Kirov *et al.*, 2009b). Cumulatively, all these findings show that sporadic schizophrenia cases are prone to carry rare de novo CNVs that do not affect any genes, whereas familial cases are enriched by rare inherited CNVs that are candidate to perturb gene structure and function (Xu *et al.*, 2008).

Two very recent exomic sequencing studies revealed rare point mutations in schizophrenia families. The first study was performed using 53 family trios of subjects and trios of 22 unrelated healthy controls (Xu *et al.*, 2011). 34 do novo point mutations and four de novo indel candidates were observed. While 19 out of 32 nonsynonymous de novo point mutations were evolutionarily conserved, three of four indel candidates caused protein truncations and other one single aminoacid deletions. Another study applied exomic sequencing to 14 probands with their parents and identified 15 de novo point mutations that disrupt the zinc finger protein 480 (*ZNF480*), karyopherin alpha 1 (*KPNA1*), low density lipoprotein receptor-related protein 1 (*LRP1*) and ALS2 C-terminal like (*ALS2CL*) genes (Girard *et al.*, 2011).

### 1.2. ERBB4-NRG1 Signalling in Schizophrenia

The ERBB family proteins belong to type I receptor tyrosine kinases. There are four members of this family: ERBB1 (as known as EGFR), ERBB2 that lacks the ligand binding

domain, ERBB3 that does not have an active kinase domain and ERBB4. Neuregulins are ligands for ERBB receptor, especially for ERBB4. There are four members in this cell-cell signalling proteins: NRG1, -2, -3 and -4 (Figure 1.5a). NRG ligands can only bind to ERBB3 and ERBB4 and since ERBB3 does not have a kinase domain, the signalling is conducted via ERBB4 that is homodimerized or heterodimerized with ERBB2 and ERBB3 (Figure 1.5b) (Olayioye *et al.*, 2000).

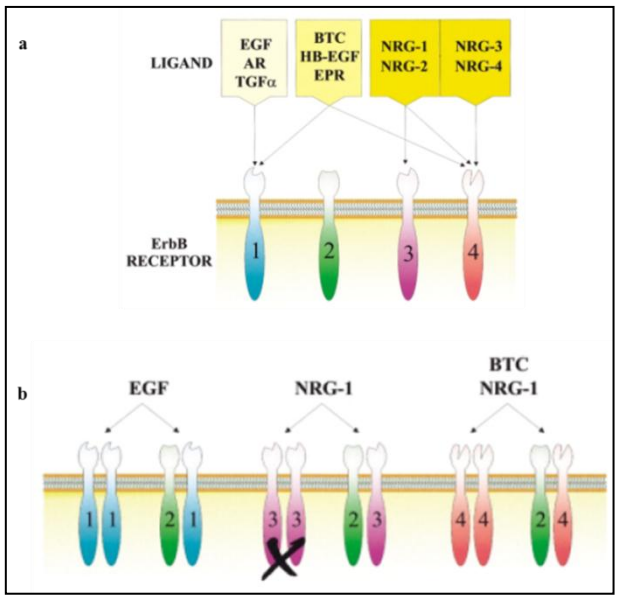



Figure 1.5. Binding specificities of ERBB receptor family ligand (adapted from (Olayioye *et al.*, 2000)).

Recently, the discovery of the role of ERBB4-NRG1 axis in schizophrenia have received a lot of scientific attention. NRG1-mediated ERBB4 signalling regulates many important cellular and molecular processes such as cellular growth, differentiation and death, in particular cell types including myelin-producing cells, glia and neurons (Falls, 2003). ERBB4 signalling is the particularly important in pathophysiology of schizophrenia due to several properties of its way of action. First, disruption of the ERBB4 signalling leads to dysfunction in neuronal migration (Anton *et al.*, 2004), NMDA hypofunction (Hahn *et al.*, 2006) and regulation of GABAergic neurotransmission (Flames *et al.*, 2004) that are also interrupted in the disease. The expression of ERBB4-NRG1 and their activity was found elevated in the

post-mortem analyses of schizophrenia patients (Law *et al.*, 2006). The hyperactivity of ERBB4-NRG1 signalling blocks phosphorylation of NMDAR by Src, leading to blockade of NMDAR-mediated synaptic current in the mouse prefrontal cortex and hippocampus which are affected in schizophrenia (Figure 1.6) (Pitcher *et al.*, 2011). Second, mice heterozygous for the *ERBB4* gene show behavioural phenotypes of schizophrenia (Stefansson *et al.*, 2002). Finally, phosphorylation of ERBB4 by NRG1 and downstream AKT and ERK2 signalling are enhanced in schizophrenia, compared to control samples (Hahn *et al.*, 2006).

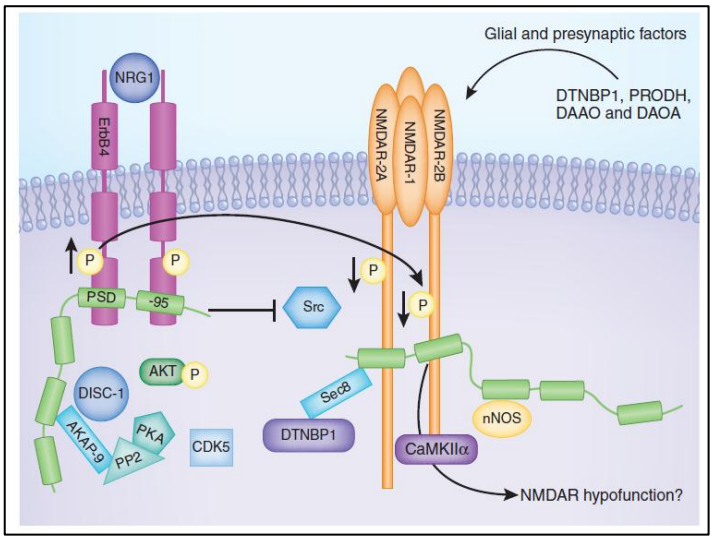



Figure 1.6. The role of ERBB4-NRG1 signalling in NMDAR hypofunction (Pitcher *et al.*, 2011).

NRG1 stimulation leads to cleavage of full-length ERBB4 by tumor necrosis factor α converting enzyme and γ-secretase and release of the intracytoplasmic domain (ERBB4-ICD). The transcriptionally active ERBB4-ICD can potentiate transcriptional activity of WT-ERα at estrogen response elements (Wong and Weickert, 2009). In post-mortem human prefrontal cortex studies, low level of ERBB4-ICD and decreased frequency of WT-ERα were observed (Chong *et al.*, 2008; Weickert *et al.*, 2008). Together these findings suggest that in schizophrenia, low levels of ERBB4-ICD cannot potentiate WT-ERα-mediated transcription, probably preventing protective effect of estrogen hormone against the disease (Wong and Weickert, 2009).

The discovery of the role of ERBB4-NRG1 axis has also led to many publications, which studied the independent as well as interactive effects of ERBB4 and NRG1, associated with schizophrenia. However, most of them were not able to be validated in other studies, due to small sample sizes and heterogeneity of the disease.

### 1.2.1. The Schizophrenia-Associated Variants in ERBB4

The *ERBB4* gene spans approximately a 1.2 Mb region on chromosome 2q34, including 28 exons. A small study of Ashkenazi Jewish schizophrenia cases and controls has shown that three SNPs (rs707284, rs839523 and rs7598440) in strong linkage disequilibrium (LD), surrounding exon 3, were associated with risk of developing schizophrenia (Silberberg *et al.*, 2006). The increased risk effect of this three-SNP haplotype on risk of schizophrenia validated by three independent family-based association studies (Nicodemus *et al.*, 2006). The group has also identified another three-SNP haplotype including rs3748962, rs2289086 and rs3791709 in the 3' end of the gene, associated with increased risk of schizophrenia. Another relatively larger Scottish study has shown that 14 out of 109 SNPs in *ERBB4*, located mostly at the two ends of the gene, have been significantly associated with schizophrenia in both allelic and genotypic tests, but a clear pattern could not be observed (Benzel *et al.*, 2007). To examine the genetic association of *ERBB4* in the Han Chinese population, 13 SNPs including those that affect the CYT1 isoforms and those previously identified in schizophrenia were selected (Lu *et al.*, 2010). Genotyping in 227 patients and 223 controls revealed the association of the ATC allele of a three-SNP haplotype, rs3791709-rs2289086-rs3748962, with increased risk of schizophrenia, as previously identified in the Caucasian population (Nicodemus *et al.*, 2006). There are also some studies that could not validate the association of *ERBB4* variants with schizophrenia. Five intronic variants in the *ERBB4* gene, rs10207288, rs2371276, rs839523, rs839511 and rs707284 were analysed in 1140 Han Chinese schizophrenia cases and 1140 controls, and no significant SNP or haplotype block was identified (Chen *et al.*, 2012).

*ERBB4* has four different isoforms created by alternative splicing (Elenius *et al.*, 1997). JM-ERBB4 isoforms differ in their extracellular juxtamembrane domain by either substitution

of exon 16 (JM-a) or exon 15 (JM-b) (Ni *et al.*, 2001). Other isoforms are CYT-1 and CYT-2 that are created by a 16 aminoacid substitution or deletion from exon 26, respectively (Junttila *et al.*, 2000). It has been shown that the A allele of rs4673628 enhanced the expression of JM-a isoform, A alleles of rs7598440, rs839523 and rs707284 induced CYT-1 isoform that were highly abundant in dorsolateral prefrontal cortex of schizophrenia patients (Law *et al.*, 2007).

### 1.2.2. Schizophrenia-Associated Variants in NRG1

The *NRG1* gene spans approximately 1.4 Mb region on chromosome 8p13, including more than 20 exons and several large introns. Beginning from 2002, single polymorphisms and haplotypes in the *NRG1* gene were associated with schizophrenia (Figure 1.7). Several linkage studies have shown chromosome 8p as a candidate risk factor in schizophrenia (Tabares-Seisdedos and Rubenstein, 2009). The risk haplotype of *NRG1*, one of the strong candidates in this region, was identified first in the Iceland population (Stefansson *et al.*, 2002). They identified $Hap_{ICE}$ haplotype in the 5' region of the gene, consisting of 5 polymorphisms (SNP8NRG122132, SNP8NRG221533, SNP8NRG241930, SNP8NRG243177, SNP8NRG422E1006) and 2 microsatellites (478B14-84, 420M91395). The association of these markers and the haplotype or other overlapping haplotypes with schizophrenia has been validated in different populations, including the Scottish (Stefansson *et al.*, 2003), British (Williams *et al.*, 2003), Han Chinese (Li *et al.*, 2004; Tang *et al.*, 2004; Yang *et al.*, 2003), Finnish (Kampman *et al.*, 2004), Japanese (Fukui *et al.*, 2006; Iwata *et al.*, 2004), Portuguese (Petryshen *et al.*, 2005), Korean (Kim *et al.*, 2006) , Bulgarian (Georgieva *et al.*, 2008), Hungarian populations (Rethelyi *et al.*, 2010), and also in meta-analysis studies (Gong *et al.*, 2009; Li *et al.*, 2006; Munafo *et al.*, 2006). Two polymorphisms from $Hap_{ICE}$ haplotype, SNP8NRG122132 and SNP8NRG243177, were also associated with increased Type I and Type IV isoforms of NRG1 in the hippocampus of schizophrenia patients, respectively (Law *et al.*, 2006).

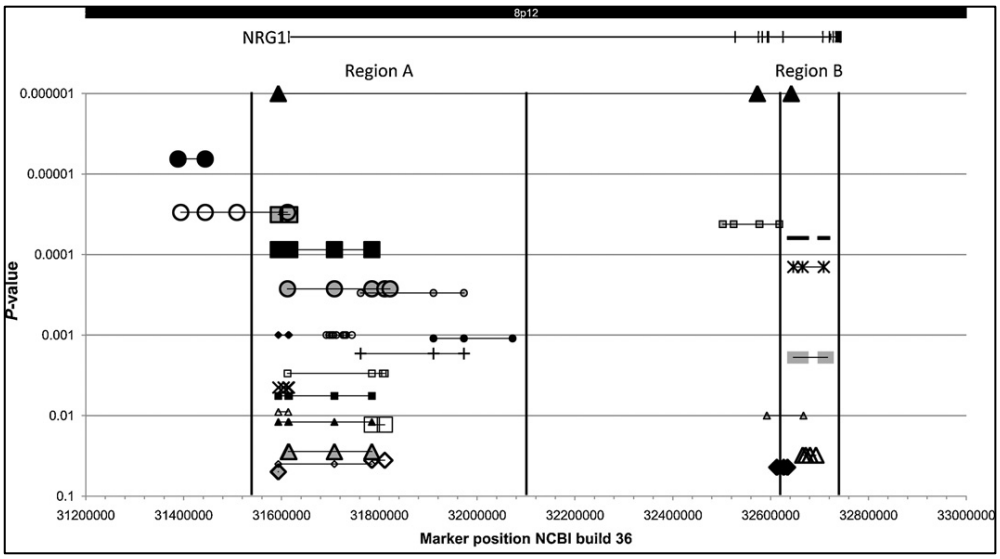



Figure 1.7. Schematic representation of *NRG1* haplotypes associated with schizophrenia (Walker *et al.*, 2010).

Other haplotypes and single markers have been also identified in several association studies. SNP8NRG221533, rs5890668, rs6150532, rs10691392 and haplotype of last three have been associated with schizophrenia in the African American population (Lachman *et al.*, 2006). A novel missense mutation from valine to leucine in exon11 has been identified in the Costa Rican population (Walss-Bass *et al.*, 2006). The same mutation could not be associated with Chinese schizophrenia patients in a recent study (Garcia-Barcelo *et al.*, 2011). Minor roles of rs2919378 and rs764059 polymorphisms of *NRG1* have been identified in the Finnish population, with p values of 0.012 and 0.048, respectively (Turunen *et al.*, 2007). While Thomson et al replicated $Hap_{ICE}$ haplotype in the Scottish population, another region in the 3' end of the gene, has been shown to be highly significant in both schizophrenia and bipolar disorder patients (Thomson *et al.*, 2007). The haplotype of two polymorphisms in this region (rs6988339 and rs3757930 with p-value 0.0037) has been validated in another study with Scottish samples (Thomson *et al.*, 2007). Alaerts et al identified five significant polymorphisms (rs7017348, rs6468061, rs7014221, rs7014410 and rs17601950) in the second intron of the gene with one protective and one disease risk-causing haplotype (p values 5 x 10-5 and 0.02) in the Northern Swedish population (Alaerts *et al.*, 2009).

Although there is a mass of positive associations of the *NRG1* haplotype, $Hap_{ICE}$, with schizophrenia, there are other studies that were not able to show the association in the Han Chinese (Hong *et al.*, 2004), Danish (Ingason *et al.*, 2006), Japanese (Ikeda *et al.*, 2008), Scandinavian (Jonsson *et al.*, 2009), Italian (Squassina *et al.*, 2010), European (Duan *et al.*, 2005), Irish populations (Thiselton *et al.*, 2004). Although a case-control study with the samples from Finland did not find significant association of SNP8NRG221533, one of the strongest single marker in the $Hap_{ICE}$, it revealed that the interaction of this variant with IL-1β-511 polymorphism causes increased risk of schizophrenia (Hanninen *et al.*, 2008). In a meta-analysis, comprising 3 family bases and 5 case-control studies, no association of SNP8NRG221533 with schizophrenia has been observed using random-effects model (p=0.61) and fixed-effects model (p=0.1) (Munafo *et al.*, 2008).

Genetic interaction analysis between *ERBB4* and *NRG1* also found significant associations. In 2006, it has been shown that the *ERBB4* IVS12-15C>T variant and the 5-kb flanking rs7424835 are associated with the *NRG1* Icelandic schizophrenia haplotype (Williams *et al.*, 2003) in the Caucasian population (Norton *et al.*, 2006). Significant pair-wise interactions between *NRG* and *ERBB* family genes have been analysed also in another study using samples from Aberdeen, UK (Benzel *et al.*, 2007). In the interaction analysis, significant interaction was observed between a recent study with a p value of 0.035 and an odds ratio 2.25 (Nicodemus *et al.*, 2010).rs2919381 in NRG1 and rs7560730 in *ERBB4* (P = 0.047, corrected) (Shiota *et al.*, 2008). Epistasis between *NRG1* (rs10503929) and its receptor *ERBB4* (rs1026882) has been associated with schizophrenia in a recent study with p value of 0.035 and odds ratio 2.25 (Nicodemus *et al.*, 2010).

# 2. PURPOSE

Schizophrenia is a complex disorder which is known to be caused from several different pathways and related genes. Although a total of 21 genome-wide association studies have revealed common polymorphisms in susceptible genes, most of them could not validate each other's findings. Moreover, variants in the causative genes, such as *ERBB4* and *NRG1*, discovered in pathway analysis, could not be associated with the disease, in GWASs, probably due to the strict significance threshold and population-specific linkage.

This study exploits three publicly available genome-wide association study datasets: CATIE, GAIN and nonGAIN, and aims to:

- establish an innovative analysis methodology to highlight genetic significance of genes, such as *ERBB4* and *NRG1*, that play crucial roles in molecular pathways leading to schizophrenia.

- find and validate genomic variants and haplotypes of the *ERBB4* and *NRG1* genes in all three datasets.

- understand the role of associated variants in the transcription of genes.

# 3. MATERIALS

## 3.1. Subjects: Study Groups

### 3.1.1. CATIE Study

The CATIE study contains a total of 738 schizophrenia patients from the United States (US) population. Cases of schizophrenia were diagnosed using Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) criteria by one or more clinicians. Patients with schizoaffective disorders and mental retardation were excluded. Other exclusion criteria included, having a history of treatment resistance, pregnancy, breastfeeding, serious mental instability and absence of written informed consent. Cases, whose ages varied from 18 to 67, self-reported themselves as Caucasian (56.3%), African-American (29.6%) and from other races (14.1%).

Controls were selected from the US population who denied having any kind of psychotic disorders including schizophrenia, schizoaffective and bipolar disorder. Ninety one percent of controls were matched with cases, according to 5-year age band, ancestry and gender. Due to insufficient African-American males in controls, African-American males in cases were matched with the females in the same age band and ancestry. Finally, 733 controls were recruited for the CATIE study.

### 3.1.2. MGS Study

The Molecular Genetics of Schizophrenia (MGS) study contains a total of 2681 European-ancestry and 1286 African American-ancestry schizophrenia patients from USA and Australia, who were diagnosed with schizophrenia and schizoaffective disorder according to DSM-III and DSM-IV criteria. Patients with schizoaffective disorders were included if only they had symptoms similar to schizophrenia for at least six months. 2653 European-ancestry

and 973 African American-ancestry controls, who self-reported their origin, were recruited for the MGS Study (Table 3.1 & 3.2)

Table 3.1. Gender distribution of cases and controls in the MGS study.

| | **EA cases** | | **EA controls** | | **AA cases** | | **AA controls** | |
|---|---|---|---|---|---|---|---|---|
| | **Count** | **%** | **Count** | **%** | **Count** | **%** | **Count** | **%** |
| **Male** | 1865 | 69.9 | 1269 | 47.8 | 803 | 62.4 | 381 | 39.2 |
| **Female** | 816 | 30.1 | 1384 | 52.2 | 483 | 37.6 | 592 | 60.8 |
| **Total** | 2681 | | 2653 | | 1286 | | 973 | |

Table 3.2. Age distribution of cases and controls in the MGS study.

| | **EA cases** | | **EA controls** | | **AA cases** | | **AA controls** | |
|---|---|---|---|---|---|---|---|---|
| | **Mean** | **SD** | **Mean** | **SD** | **Mean** | **SD** | **Mean** | **SD** |
| **at evaluation** | 43 | 11.6 | 50.6 | 16.4 | 42.7 | 10.4 | 45.6 | 13.1 |
| **at onset** | 21.4 | 7 | - | - | 20.9 | 6.8 | - | - |

## 3.2. Genotyping

### 3.2.1. CATIE Dataset

Cases and controls in the CATIE study were genotyped using three different SNP arrays: Two Affymetrix 500K ‘A’ chipsets and a custom 164K chip, created by Perlegen Sciences which had 500568 and 164871 SNPs, respectively. The genotyping was conducted by Perlegen Sciences in USA. A total of 157048 SNPs failed in quality control.

### 3.2.2. GAIN and nonGAIN Datasets

Cases and controls in the MGS study were genotyped in two phases due to funding issues, the Genetic Association Information Network (GAIN) and nonGAIN. Half of European-ancestry samples and 95% of African American samples were genotyped under the GAIN phase, whereas the remaining samples were genotyped under the nonGAIN phase after approximately six months. The Affymetrix 6.0 array including 906600 SNPs was used for both phases by Broad Institute National Center for Genotyping and Analysis. After quality controls, 696788 SNPs remained from the genotyping of European-ancestry samples, whereas 843798 SNPs passed the quality controls in African-American samples.

## 3.3. Equipment

Table 3.3. Equipment used in this study.

| Equipment | Models |
|---|---|
| Laptop Computer | Vaio VPCEB46FD, Sony, Canada |
| Desktop Computer | HP Pavilion Elite, USA |
| Desktop Computer | iMac, Apple, USA |

## 3.4. Electronic Databases

- Database of Genotypes and Phenotypes (dbGAP)
  http://www.ncbi.nlm.nih.gov/gap
  dbGAP is a public repository that provides access to genotype-phenotype files, subject characteristics and data of traits of GWASs and many other association studies of

particular genotypes with phenotypes. The database is hosted by the National Center for Biotechnology Information (NCBI).

- Single Nucleotide Polymorphism Database (dbSNP)
  http://www.ncbi.nlm.nih.gov/sites/entrez?db=snp
  dbSNP is a free online database for genetic variations including single nucleotide polymorphisms and short deletions/insertions and microsatellite markers across different species. The database is hosted by the NCBI (Bethesda, USA).

- A Functional Analysis and Selection Tool for SNP (FASTSNP)
  http://fastsnp.ibms.sinica.edu.tw/
  FASTSNP is an online tool that prioritizes SNPs according to their phenotypic and functional risk profiles. The tool gives an empirical score to a given SNP according to whether it changes the amino acid sequence and protein structure, whether it alters expression via binding to promoter or enhancer sequences and whether it has potential to disrupt splicing.

- GeneCards
  http://www.genecards.org/
  GeneCards is a database that contains inclusive information on human genes, including protein, function, expression, variants, paralogs/orthologs and association with diseases. The website is hosted by the Weizmann Institute of Science (Rehovot, Israel).

- GWAS Catalog
  http://www.genome.gov/gwastudies/
  The GWAS Catalog is an inventory for published genome-wide association studies of all diseases and traits. The catalog lists all GWASs, subjects, associated variants, platforms, studies performed for a given disease or trait. The website is hosted by the National Human Genome Research Institute (NHGRI).

- International Haplotype Map Project (HapMap)
  http://hapmap.ncbi.nlm.nih.gov/
  The HapMap data is a free-access project to describe common patterns in human genetic variations of a total of 270 individuals from the African, Asian and European-ancestry populations.

- PubMed
  http://www.ncbi.nlm.nih.gov/pubmed
  PubMed is a free resource that contains citations for biomedical literature from life sciences journals and online books and provides outsource links to relevant NCBI databases. The database is hosted by the NCBI.

- SNP Annotation and Proxy Search (SNAP)
  http://www.broadinstitute.org/mpg/snap/
  The online SNAP tool identifies proxy SNPs by calculating linkage disequilibrium and physical distances. The genotype data is based on the HapMap and 1000 Genome Pilot projects. The tool is hosted by the Broad Institute (Cambridge, USA).

- Schizophrenia Gene Database (SZGene)
  http://www.szgene.org/
  SZGene is a regularly-updated database that combines all genetic association studies and also meta-analyses of schizophrenia.

- University of California Santa Cruz (UCSC) Genome Browser
  http://genome.ucsc.edu
  UCSC Genome Browser is a comprehensive website that contains genome sequence data from several organisms and visualization and analysis tools at many levels such as genetic variants, conservation, protein and expression. The browser is hosted by the UCSC (Santa Cruz, USA).

### 3.5. Bioinformatic Tools

- GWADview

  http://www.ozceliklab.com/bioinfoProjects.asp

  The GWADview software is a visualization tool, designed by Ozcelik Lab (University of Toronto, Canada) to interpret GWAS results. The analysis is based on various algorithms such as allelic, genotypic, dominant and recessive test models. The system provides a single, integrated plot of SNP distribution according to physical location on the chromosome and p-values/odds ratio from multiple resources.

- Haploview 4.2 version

  http://www.broad.mit.edu/haploview/haploview

  The Haploview software is used to analyse and visualise linkage disequilibrium and haplotype maps. The software is developed by the Broad Institute (Cambridge, USA).

- Ingenuity Pathway Analysis (IPA)

  http://www.ingenuity.com

  The Ingenuity Pathway Analysis software is designed to establish networks for given genes/proteins and to visualize complex biological pathways based on protein-protein interactions and regulation of expression.

- PLINK

  http://pngu.mgh.harvard.edu/purcell/plink/

  PLINK is a free genome association analysis toolset that is used to manage genotype-phenotype data. gPLINK is a more advanced software, that emerges from PLINK, with a user-friendly interface. It is being developed by Massachusetts General Hospital (MGH) and the Broad Institute (Cambridge, USA).

# 4. METHODS

## 4.1. Literature Search

Genes and the encoded proteins, associated with schizophrenia, were defined via literature search. Three major resources used are: GWAS catalog, SZGene database and peer-reviewed journals from PubMed. The literature search was completed in March 2010.

### 4.1.1. GWAS Catalog

The candidate genes that were found to be associated with schizophrenia in genome-wide association studies were gathered for further examination. Each gene was searched in PubMed and was ranked according to the number of studies which found a positive and/or negative association of the gene with schizophrenia (Figure 4.1).

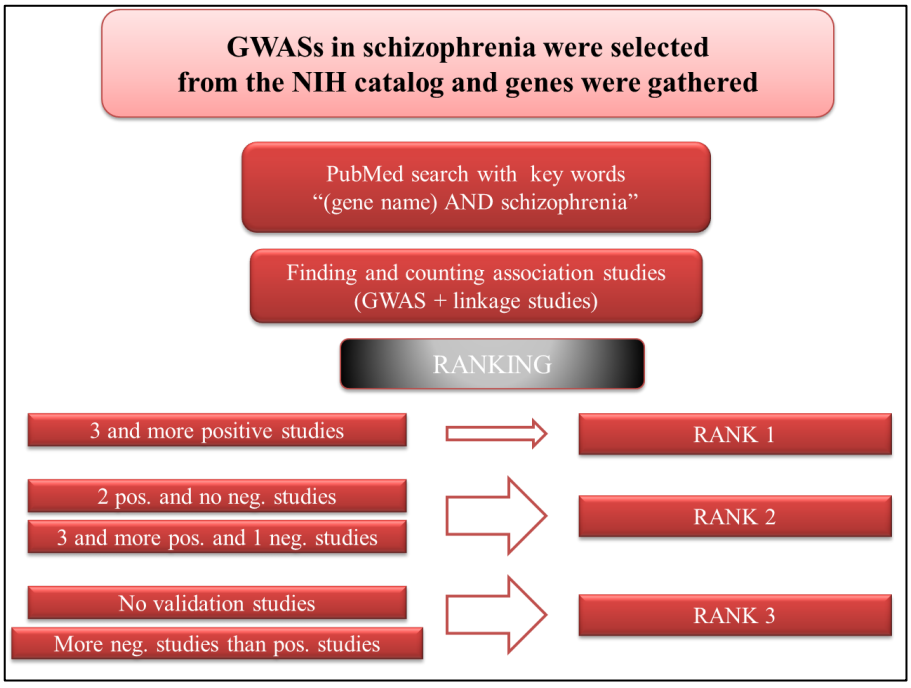



Figure 4.1. Algoritm for selection and ranking of candidate schizophrenia genes from the GWAS catalog.

### 4.1.2. SZGene Database

The SZGene database had its own ranking system that was based on HuGENet interim guidelines for the assessment of genetic association studies. According to this system, there were 39 top genes that were associated with schizophrenia. All these genes were searched for in PubMed and were ranked via a similar algorithm of the GWAS Catalog (Figure 4.2).

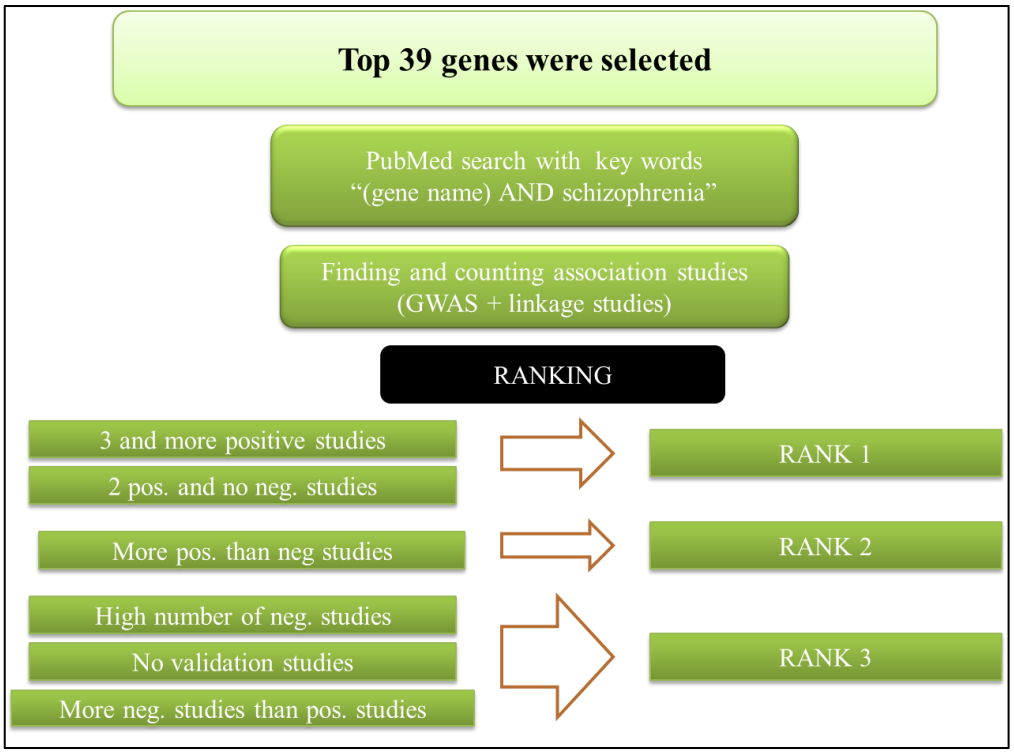



Figure 4.2. Algoritm for selection and ranking of candidate schizophrenia genes from the SZGene database.

### 4.1.3. Peer-reviewed Journals in PubMed

For mutated genes that might be missed in association studies, keywords "mutation" and "schizophrenia" were used in PubMed search. After creation of a candidate list for mutated genes, each gene was searched again in PubMed to find out how many studies could find association of that gene with schizophrenia. The genes that had three and more positive

association studies were ranked as Rank 1, the genes that were associated with schizophrenia only two times were classified as Rank 2 and the ones without validation as Rank 3.

### 4.2. Pathway Analysis

Ingenuity Pathway Analysis software was used for pathway analysis. The proteins encoded by candidate schizophrenia genes established a preliminary network. Most of the indirect interactions which were proposed by the software were eliminated for a simpler view and direct interactions were kept. Proteins were placed according to their subcellular and extracellular locations and grouped according to their functions. Top molecular/cellular functions, physiological system development/function and canonical pathways were analysed using additional features of the software.

### 4.3. Collection of Genome-wide Association Datasets

The genotype data for cases and controls was obtained from NIMH and dbGaP (dbGAP accession numbers for GAIN and nonGAIN were phs000021.v2.p1 and phs000167.v1.p1, respectively). The African-American samples were not used in further analysis. The age distribution for both female and male samples in CATIE, GAIN and nonGAIN were examined, and individuals who were older than 65 years of age were eliminated in order to match the population characteristics of the GAIN and non-GAIN datasets.

### 4.4. Logistic Regression Analysis

The chromosome locations of the ERBB4 and NRG1 genes were determined according to HapMap release 27. Using PLINK, the genotype data of these two genes were extracted from the CATIE, GAIN and nonGAIN datasets. The Perlegen and Affymetrix identities in the datasets were converted into rs numbers using Microsoft Access. Since the CATIE and MGS studies used different SNP genotyping arrays, only common SNPs were used for logistic regression analysis. Before executing the logistic model, the haplotype block structure for each gene had to be formed. Block structures were established according to the CATIE dataset to

provide uniformity. Haplotype blocks were estimated following the default procedure in Haploview. By default, pairwise LD is only calculated for SNPs within 200kb. For haplotype-based association using logistic regression model, PLINK was executed with an additional *omnibus* command which produced odds ratio values for each haplotype block. Commands are listed below:

- To extract genotype data from GWAS datasets
  *plink --file GWASdata --allow-no-sex --chr X --from-bp XXX --to-bp XXX --recode --out genedata*
- To extract common SNPs from data file
  *plink --file genedata --extract SNPlist.txt --recode --out commonSNPdata*
- To extract common SNPs from data file
  *plink --file genedata --extract SNPlist.txt --recode --out commonSNPdata*
- To generate info file for the Haploview software
  *plink --file genedata --extract SNPlist.txt --recodeHV --out commonSNPdata*
- To establish haplotype block structure
  *plink --file commonSNPdata --blocks --out blockfile*
- Haplotype-based association using logistic regression
  *plink --file commonSNPdata --hap blockfile.hlist --hap-logistic --out logisticdata*
- Additional omnibus command
  *plink --file commonSNPdata --hap blockfile.hlist --hap-logistic –hap-omnibus --out omnibusdata*

### 4.5. Haplotype Analysis

To narrow down the significant region in logistic regression model, haplotype analysis was performed using the Haploview 4.2 software. The Hardy-Weinberg p-value cut-off and minimum minor allele frequency were set to 0.05 and 0.001, respectively. Other parameters were used as default. The haplotype blocks were defined using Solid Spine of LD analysis method with D' > 0.6. The custom blocks were created when necessary.

### 4.6. Examination of Transcription Factor Binding Sites

The regions that were found to be significant in more than one dataset in our study were also subject to fine mapping for purposes of identifying transcription factor binding sites. All SNPs between candidate regions of ERBB4 and NRG1 were present in the Build 132 version of the dbSNP database. SNPs were separated into four cataegories: Common (MAF: 25-10%), rare (MAF: 9-1%), very rare (MAF: <1%) and not identified. Transcription factors that bind to major and minor allele for each polymorphism were identified using FastSNP software. Subsequently, we pursued our analysis of possible TF interactions with regions containing significant haplotypes in the CATIE, GAIN and nonGAIN datasets.

# 5. RESULTS

## 5.1. Candidate Schizophrenia Genes in the Literature

Several genes and the proteins they encode have been associated with schizophrenia so far. To list these genes, three sources were used, and the genes were classified as Rank 1, 2 and 3, as explained in the Methods section. A total of 27 candidate genes were gathered via search in the GWAS Catalog. These genes were ordered according to the number of studies that found positive and/or negative association with schizophrenia: 12 genes were ranked as Rank 1, 11 genes as Rank 2, four genes as Rank 3. Another source were the top 39 genes from the SZGene database. Eleven genes were classified as Rank 1, 7 and 21 genes as Rank 2 and 3, respectively, using the same methodology. Finally, Pubmed search was performed, not to miss any significant candidate gene in schizophrenia. In the search for mutated genes, a total of 24 genes were found. Out of 24 genes, eight, two and 14 genes were ordered as Rank 1, 2 and 3, respectively.

Out of 31 genes that were categorized as Rank 1, four were overlapping in two groups. After elimination of the overlaps, 27 schizophrenia genes, listed in Table 5.1, were taken as candidates.

Table 5.1. Candidate "Rank 1" schizophrenia genes that were collected from the GWAS Catalog, SZGene and Pubmed databases.

| Gene symbol | Gene Name | Location | Genomic Size (bp) | Source |
|---|---|---|---|---|
| *ACSM1* | acyl-CoA synthetase medium-chain family member 1 | 16p12.3 | 68020 | GWAS Catalog |
| *AHI1* | Abelson helper integration site 1 | 6q23.3 | 213759 | SZGene |
| *BTN2A2* | butyrophilin, subfamily 2, member A2 | 6p22.2 | 11745 | GWAS Catalog |

Table 5.1. Candidate "Rank 1" schizophrenia genes that were collected from the GWAS Catalog, SZGene and Pubmed databases (cont.).

| | | | | |
|---|---|---|---|---|
| *BTN3A1* | butyrophilin, subfamily 3, member A1 | 6p22.2 | 12962 | GWAS Catalog |
| *BTN3A2* | butyrophilin, subfamily 3, member A2 | 6p22.2 | 13149 | GWAS Catalog |
| *COMT* | catechol-O-methyltransferase | 22q11.21 | 27220 | PubMed |
| *DISC1* | disrupted in schizophrenia 1 | 1q42.2 | 414456 | PubMed |
| *DTNBP1* | dysbindin | 6p22.3 | 140233 | PubMed |
| *ERBB4* | erythroblastic leukemia viral oncogene homolog 4 | 2q34 | 1162911 | PubMed |
| *HIST1H2AG* | histone cluster 1, H2ag | 6p22.1 | 2250 | GWAS Catalog |
| *HIST1H2BJ* | histone cluster 1, H2bj | 6p22.1 | 480 | GWAS/SZGene |
| *NPAS3* | neuronal PAS domain protein 3 | 14q13.1 | 862493 | PubMed |
| *NRGN* | neurogranin | 11q24.2 | 7356 | SZGene |
| *OPCML* | opioid binding protein/cell adhesion molecule-like | 11q25 | 1117527 | SZGene |
| *PDE4B* | phosphodiesterase 4B | 1p31.3 | 582069 | SZGene |
| *PRSS16* | protease, serine, 16 | 6p22.1 | 8742 | SZGene |
| *RELN* | reelin | 7q22.1 | 517726 | GWAS/SZGene |
| *RGS4* | regulator of G-protein signaling 4 | 1q23.3 | 7228 | PubMed |
| *RPGRIP1L* | RPGRIP1-like | 16q12.2 | 103954 | SZGene |
| *RPP21* | ribonuclease P/MRP 21kDa subunit | 6p22.1 | 1696 | SZGene |
| *RTN4R* | reticulon 4 receptor | 22q11.21 | 26877 | PubMed |
| *SCL17A1* | solute carrier family 17 (sodium phosphate), member 1 | 6p22.2 | 47481 | GWAS Catalog |
| *SCL17A3* | solute carrier family 17 (sodium phosphate), member 3 | 6p22.2 | 37186 | GWAS Catalog |
| *SYNGR1* | synaptogyrin 1 | 22q13.1 | 35626 | PubMed |
| *TCF4* | transcription factor 4 | 18q21.2 | 360475 | GWAS/SZGene |
| *ZNF184* | zinc finger protein 184 | 6p22.1 | 22353 | GWAS Catalog |
| *ZNF804A* | zinc finger protein 804A | 2q32.1 | 341120 | GWAS/SZGene |

### 5.2. Pathway Analysis of Candidate Schizophrenia Proteins

The proteins encoded by the genes that were collected from three different sources varied greatly regarding their cellular compartments and functions. To analyse their interactions with each other, the Ingenuity Pathway Analysis software was used. Most of these proteins had molecular and cellular functions, such as cell morphology and cell-to-cell signalling (Table 5.2).

Table 5.2. Top five molecular and cellular functions.

| Name | p-value | # Molecules |
|---|---|---|
| Cell Morphology | 1.18E-08 - 5.51E-03 | 10 |
| Cell-To-Cell Signaling and Interaction | 1.57E-08 - 6.01E-03 | 14 |
| Cell Death | 2.30E-06 - 5.07E-03 | 12 |
| Cellular Movement | 8.04E-06 - 5.51E-03 | 12 |
| Carbohydrate Metabolism | 2.52E-05 - 5.51E-03 | 6 |

The candidate schizophrenia proteins were grouped according to their roles in physiological system development. While 20 out of 27 proteins function in nervous system development, 10 proteins are responsible for behaviour (Table 5.3). Finally, the software was used to reveal cellular pathways that might be crucial for disease development. The most effective pathway was shown to be the G-protein coupled receptor signaling that included six proteins from the initial candidate protein list (Table 5.4). cAMP-mediated signaling and dopamine receptor signaling ranked as second and canonical pathways with five and four candidate proteins as third.

Table 5.3. Top five physiological system development and function pathways.

| Name | p-value | # Molecules |
|---|---|---|
| Nervous System Development and Function | 1.18E-08 - 5.51E-03 | 20 |
| Behavior | 2.11E-06 - 5.51E-03 | 10 |
| Embryonic Development | 2.30E-06 - 5.51E-03 | 7 |
| Cardiovascular System Development and Function | 2.32E-06 - 5.51E-03 | 8 |
| Organismal Functions | 3.55E-06 - 8.29E-04 | 7 |

Table 5.4. Top five canonical pathways.

| Name | p-value | # Molecules |
|---|---|---|
| G-Protein Coupled Receptor Signaling | 2.32E-05 | 6 |
| cAMP-mediated Signaling | 7.33E-05 | 5 |
| Dopamine Receptor Signaling | 9.92E-05 | 4 |
| Serotonin Receptor Signaling | 2.24E-04 | 3 |
| GM-CSF Signaling | 6.93E-04 | 3 |

After analyses of molecular and cellular functions, physiological system development and canonical pathways in which candidate proteins play roles, a network of these protein and their mediator proteins was established (Figure 5.1). This complex network included several linker proteins that connected candidate proteins. Direct and indirect interactions were drawn with straight lines and dashes, respectively. All proteins were localized according to their subcellular compartments. The ones without known interactions were confined to a small area at the bottom right of the figure. This pathway analysis emphasized the importance of ERBB4-NRG1 interaction in connecting glutamate receptors, such as GRIK3 and GRIN2B, with downstream effectors (AKT1 and GRB2) and in signalling through the nucleus.

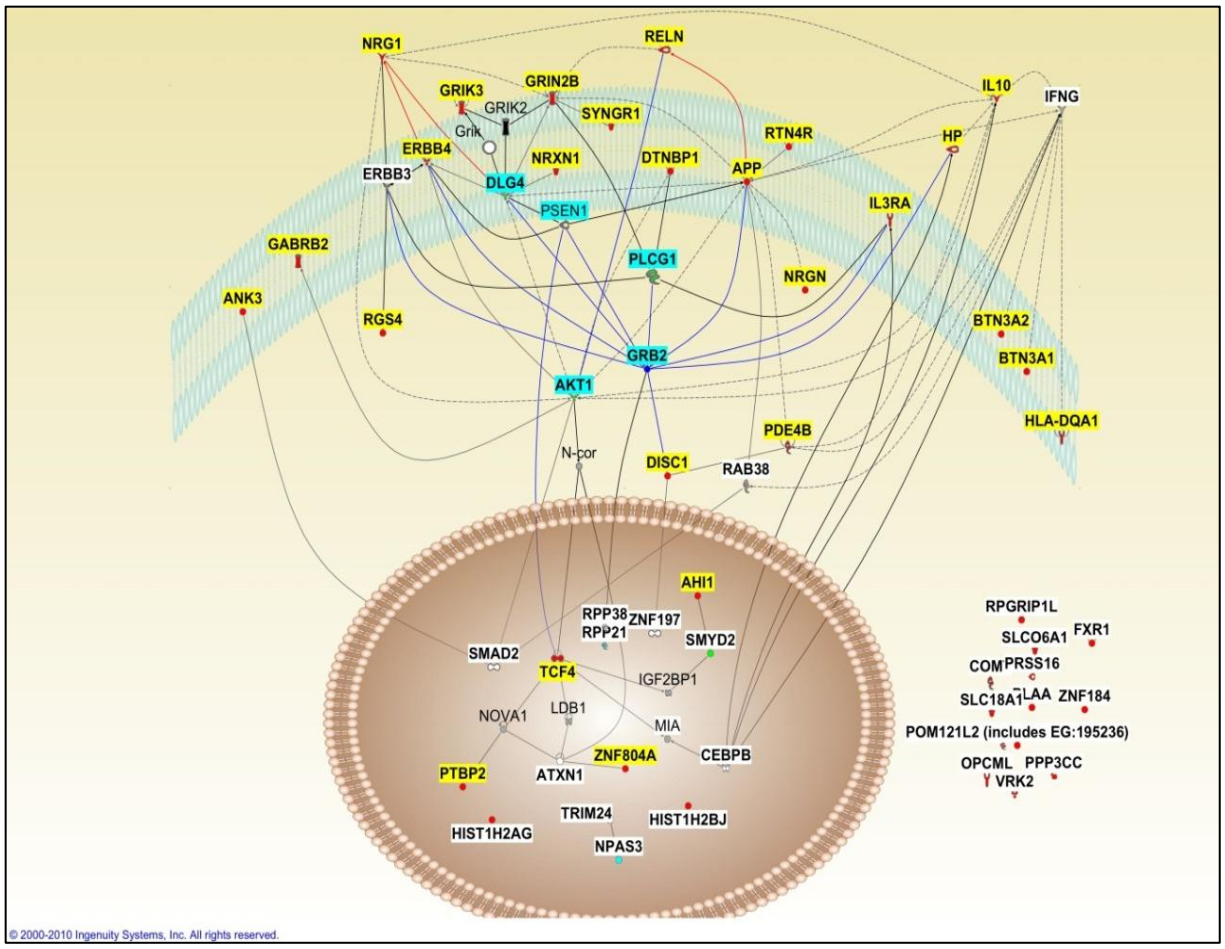



Figure 5.1. Pathway analysis of candidate proteins in schizophrenia.

### 5.3. Study Group Quality Checks of CATIE, GAIN and nonGAIN datasets

ERBB4-NRG1 signalling was shown crucial in pathway analysis. Although different SNPs and haplotypes of *ERBB4* and *NRG1* have been found to be significant in genetic studies, none of the GWASs could give significant results for these two genes. To re-analyse the raw data of three GWASs, datasets were downloaded from publicly available databases and subject groups of each were analyzed regarding their age distributions. First, cases and controls in CATIE datasets were graphed in five year age band (Figure 5.2). All controls were age-matched properly with the controls, except the 20-24 year band that had more cases than controls.

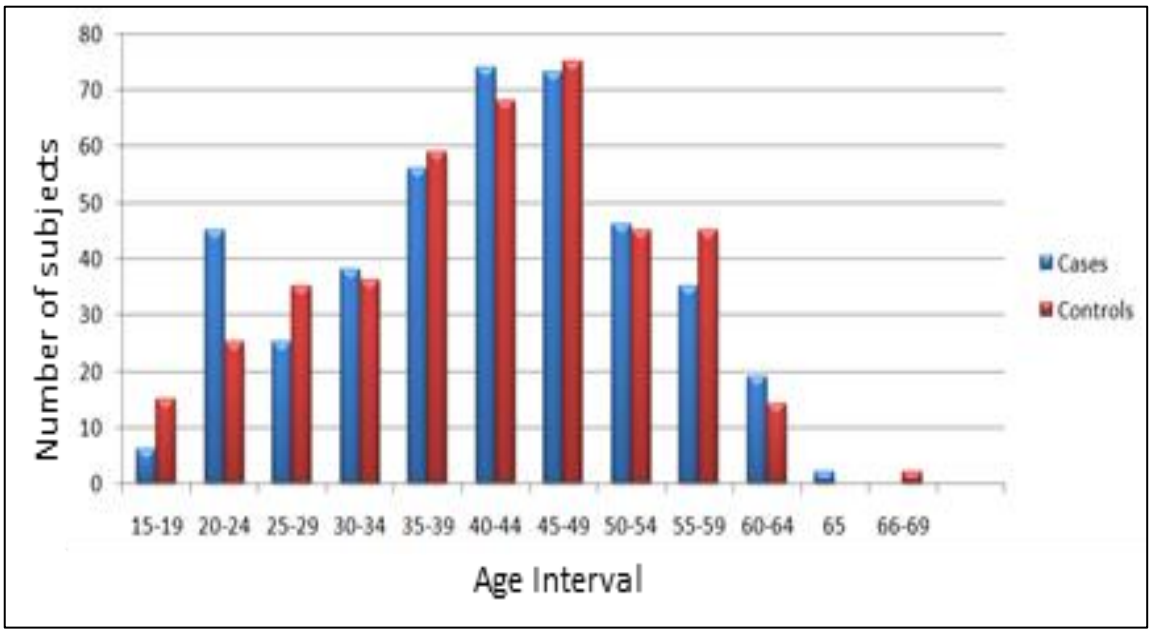



Figure 5.2. Age distribution of cases and controls in the CATIE dataset.

Subjects in the GAIN dataset were grouped according to their five year age band, as performed for the CATIE dataset. It was observed that cases and controls, older than 65 years, were not age-matched (Figure 5.3a). The number of controls over 65 was higher than the number of cases at the same age. Using the PLINK software, all subjects over 65 years were excluded and a normal age distribution was obtained (Figure 5.3b).

Like in the GAIN dataset, the nonGAIN dataset also had more controls over 65 years than cases (Figure 5.4a). To obtain a normal distribution of age in both cases and controls, all individuals over 65 years were eliminated from both groups (Figure 5.4b). Bioinformatic analyses of *ERBB4* and *NRG1* were performed using datasets after quality checks.

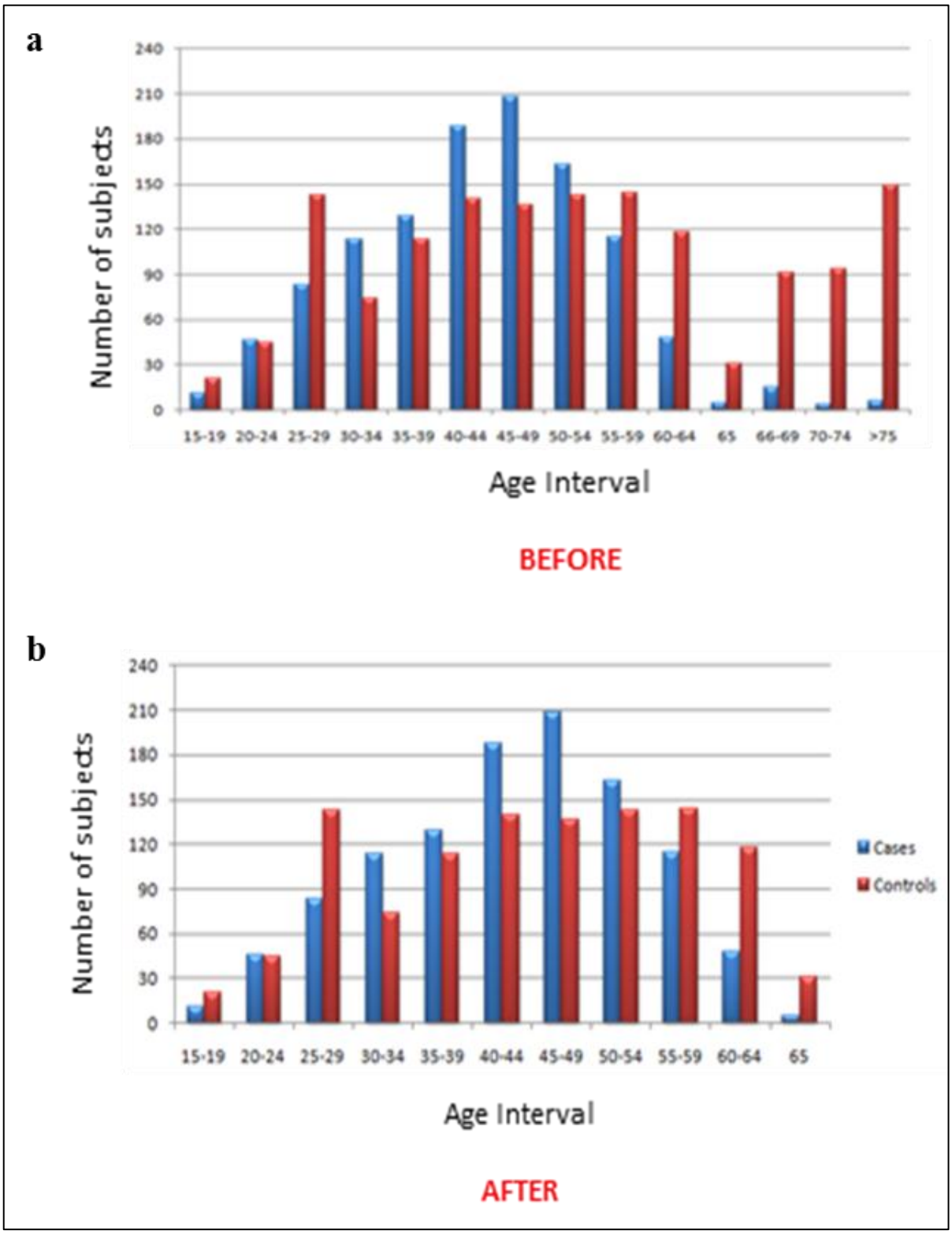



Figure 5.3. Age distribution of cases and controls in the GAIN dataset (a) before and (b) after quality checks.

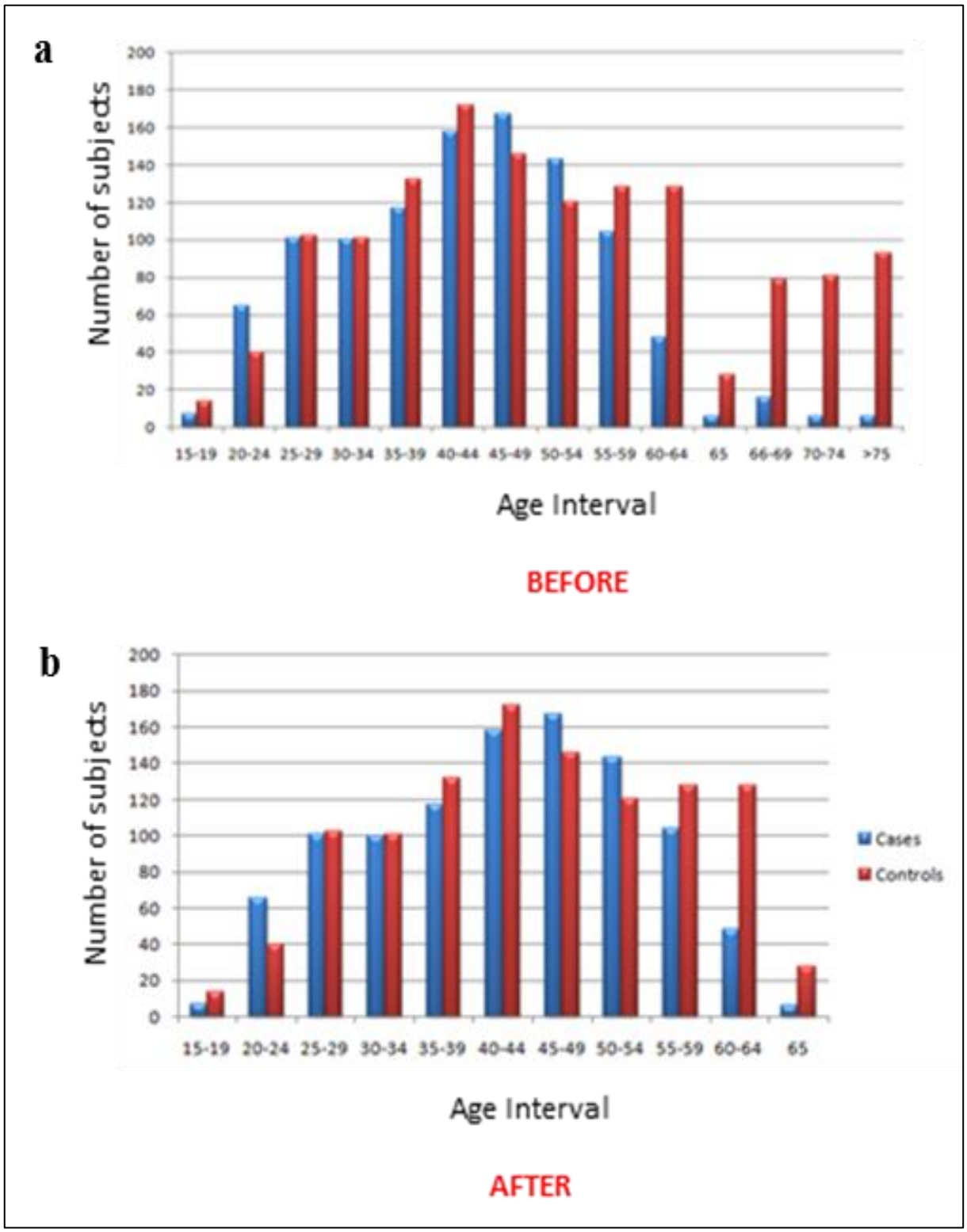



Figure 5.4. Age distribution of cases and controls in the nonGAIN dataset (a) before and (b) after quality checks.

## 5.4. Examination of Variants and Haplotypes in Three Datasets

### 5.4.1. Overview of SNPs of *ERBB4* and NRG1 using GWADView

For visualization of SNPs in *ERBB4* and *NRG1*, the GWADView software was used. SNPs in *ERBB4* and ± 20 kilobase flanking regions were extracted from each dataset and allelic test was performed to calculate p values by the software. Blue dots represent SNPs in the CATIE datasets, while red and green dots imply for SNPs in the GAIN and nonGAIN datasets, respectively. While 221 SNPs of *ERBB4* were found in the CATIE dataset, 393 in GAIN and 390 in nonGAIN were assayed (Figure 5.5). The SNPs that passed the significance threshold of 0.05 (yellow line) were shown with their rs ID in the graph. These SNPs were concentrated in the region between roughly 211950000 bp and 212230000 bp for at least two datasets.

GWADView was also employed to plot SNPs in the *NRG1* gene and its ±20 kb flanking region in three genome-wide association datasets. Blue dots represent SNPs in the CATIE datasets while red and green dots imply for SNPs in the GAIN and nonGAIN datasets, respectively. The CATIE dataset included 272 SNPs in *NRG1*. The GAIN and nonGAIN datasets had more SNPs, 496 and 497, because they had been studied with a more comprehensive SNP array (Figure 5.6). The SNPs with p values < 0.05 in more than one dataset were located within a large region between 32270000 bp and 32450000 bp.

### 5.4.2. Narrowing Down the Significant ERBB4 and NRG1 Regions

Previous analyses of single nucleotide variants showed presence of candidate regions containing significant SNPs. However these SNPs were not consistent among three datasets. To narrow down the susceptible regions in *ERBB4* and *NRG1,* and to find a smaller area that would be significant in at least two datasets, haplotype-based logistic regression analysis was performed using PLINK. For *ERBB4*, only the region between 211950000 bp and 212230000 bp, which was revealed in the previous GWADView analysis, was analysed.

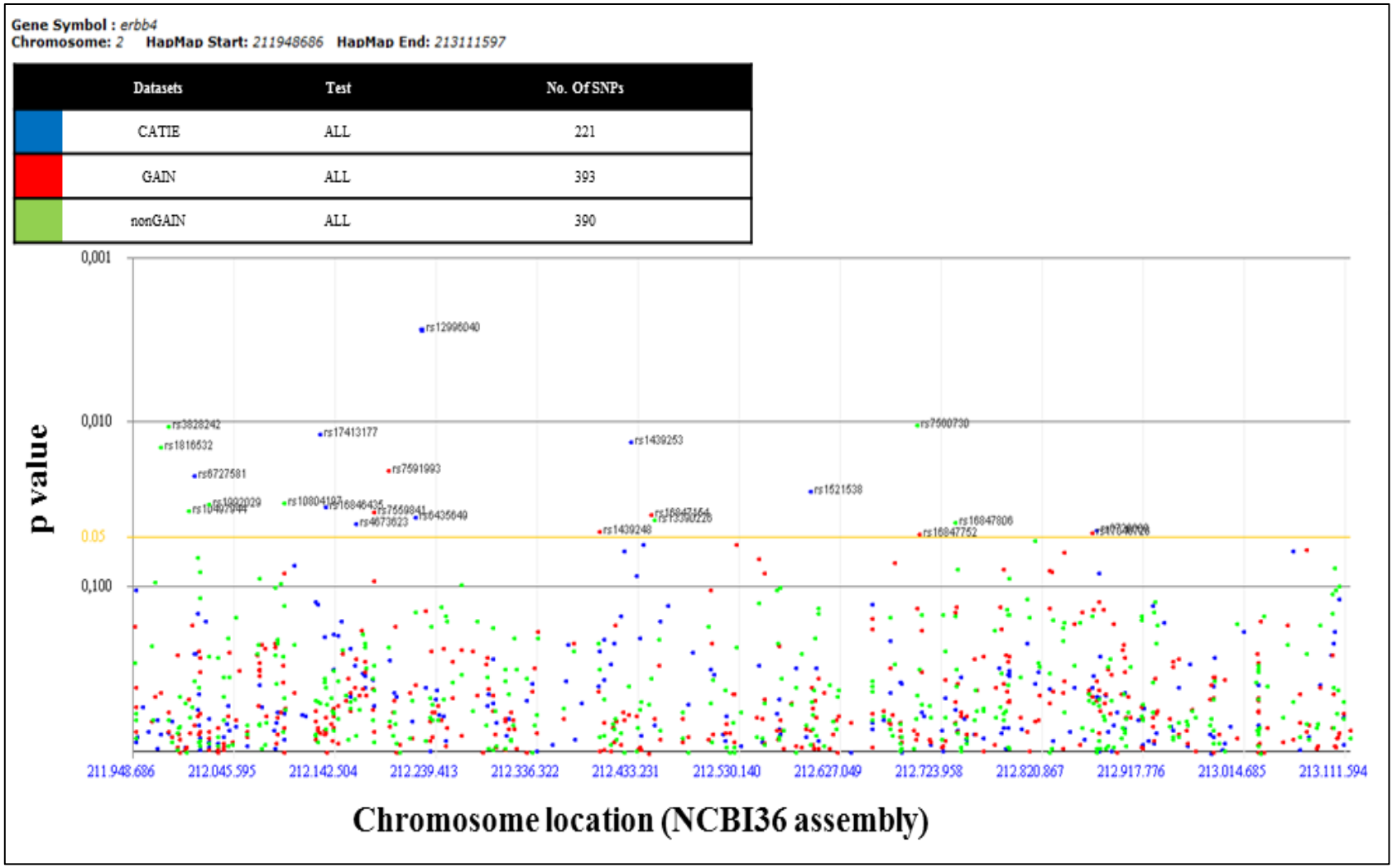



Figure 5.5. Visualization of the SNPs in the *ERBB4* gene and its ±20 kb flanking region according to their significance and chromosome location with the GWADView software.

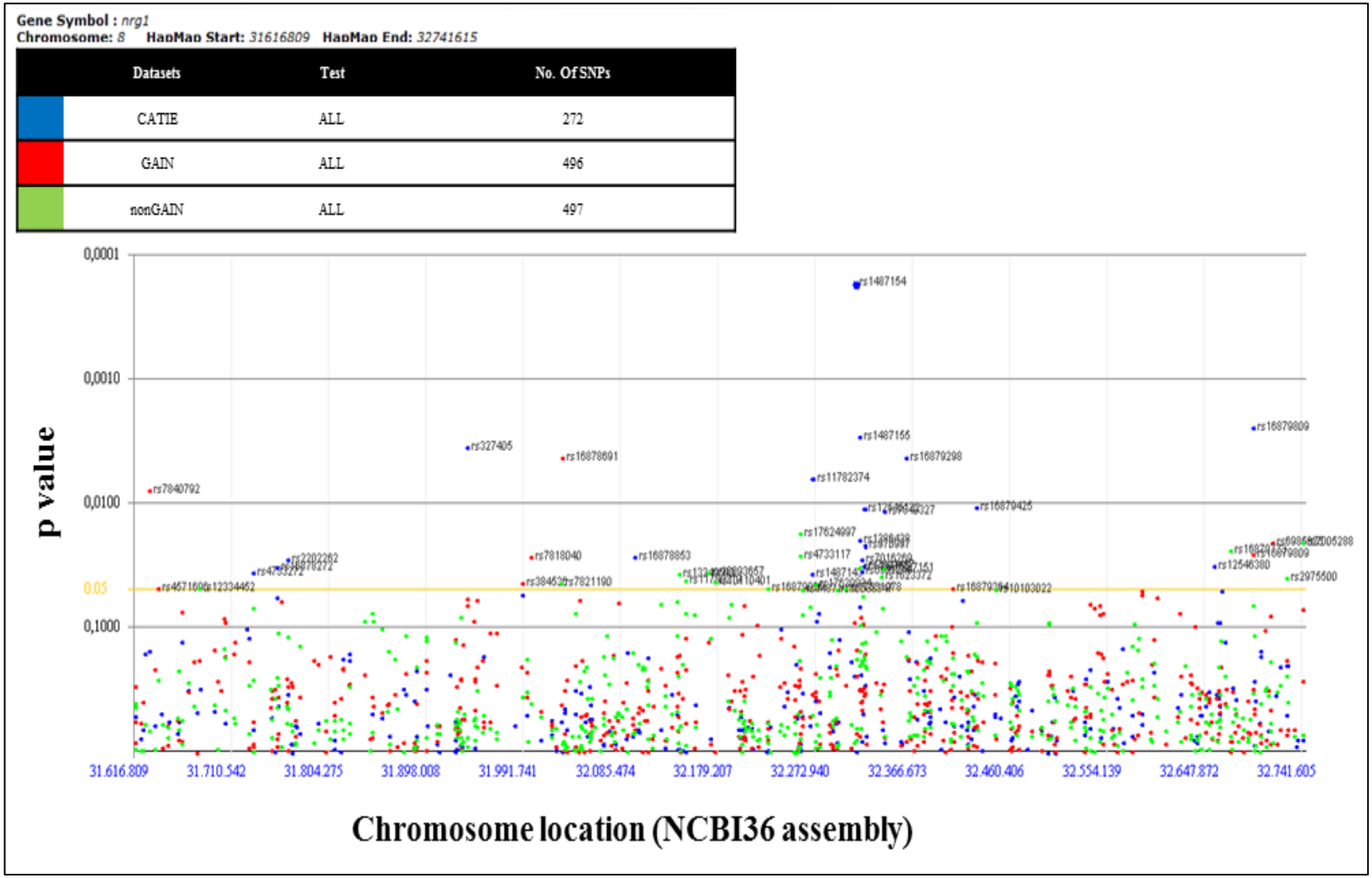



Figure 5.6. Visualization of the SNPs in the *NRG1* gene and its ±20 kb flanking region according to their significance and chromosome location with the GWADView software.

A small six-kb haplotype block from chromosome location 212156823 to 212162828 was significant in CATIE and GAIN with p values of 0.0183 and 0.00799, respectively (Table 5.5). This four-SNP haplotype block started with rs7586137 and ended with rs4673623. For the CATIE dataset, the significant haplotype was C-G-A-G with a p value of 0.0183 and an odds ratio of 0.693. For GAIN, the T-G-G-C haplotype had a significant p value (0.00799), but none of the haplotypes in this block was found to be significant in the nonGAIN datasets.

Table 5.5. The significant haplotype block within *ERBB4* as a result of haplotype-based logistic regression analysis.

| | CHR | BP1 | BP2 | SNP1 | SNP2 | Haplotype | Odds Ratio | p value |
|---|---|---|---|---|---|---|---|---|
| CATIE | | | | | | | | |
| | 2 | 212156823 | 212162828 | rs7586137 | rs4673623 | CGAG | 0,693 | 0,0183 |
| | 2 | 212156823 | 212162828 | rs7586137 | rs4673623 | CGAC | 1,95 | 0,0661 |
| | 2 | 212156823 | 212162828 | rs7586137 | rs4673623 | TAGC | 1,09 | 0,54 |
| GAIN | | | | | | | | |
| | 2 | 212156823 | 212162828 | rs7586137 | rs4673623 | CGAG | 1,09 | 0,332 |
| | 2 | 212156823 | 212162828 | rs7586137 | rs4673623 | CGAC | 0,936 | 0,79 |
| | 2 | 212156823 | 212162828 | rs7586137 | rs4673623 | TGGC | 0,468 | 0,00799 |
| | 2 | 212156823 | 212162828 | rs7586137 | rs4673623 | TAGC | 0,993 | 0,928 |
| nonGAIN | | | | | | | | |
| | 2 | 212156823 | 212162828 | rs7586137 | rs4673623 | CGAG | 1 | 0,964 |
| | 2 | 212156823 | 212162828 | rs7586137 | rs4673623 | CGAC | 1,34 | 0,218 |
| | 2 | 212156823 | 212162828 | rs7586137 | rs4673623 | TGGC | 1 | 0,991 |
| | 2 | 212156823 | 212162828 | rs7586137 | rs4673623 | TAGC | 0,926 | 0,359 |

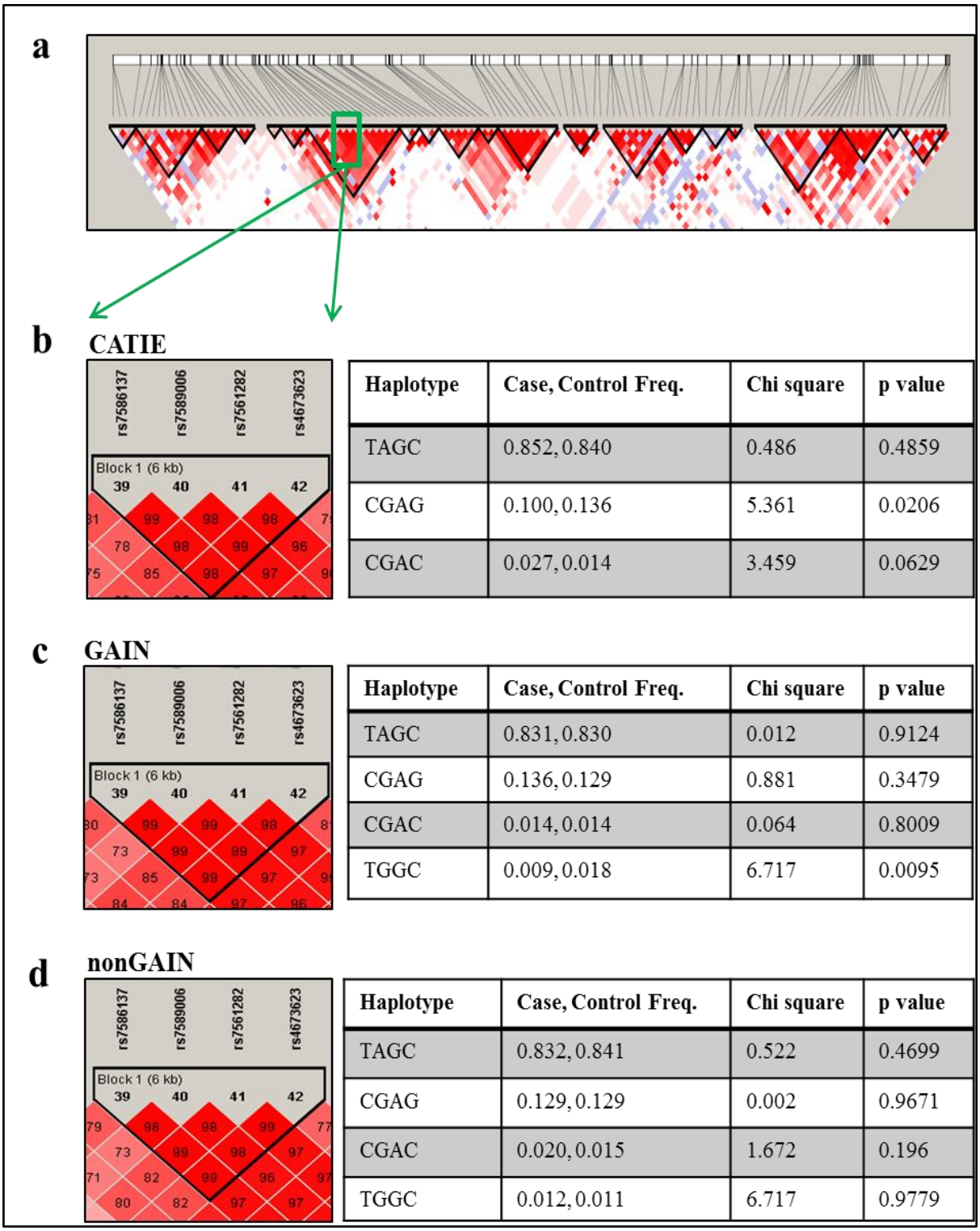
GWADView analysis showed that the *NRG1* region between 32280000 bp and 32450000 bp included significant SNPs in the CATIE and GAIN datasets. Like *ERBB4*, a smaller significant block for all three datasets in the *NRG1* gene was searched using logistic regression analysis. A 28-kb haplotype block from 32292603 bp to 32320096 bp was found to be significant in the CATIE and nonGAIN datasets (Table 5.6). The most significant haplotypes in CATIE and nonGAIN exhibited strong associations with p values of 0.0006 and 0.0199, respectively.

Table 5.6. The significant haplotype block within *NRG1* as a result of haplotype-based logistic regression analysis.

| Block | BP1 | BP2 | SNP1 | SNP2 | Haplotype | Odds Ratio | p value |
|---|---|---|---|---|---|---|---|
| CATIE | | | | | | | |
| | 32292603 | 32320096 | rs6999977 | rs970997 | GTCCGAGGAACGTACA | 0,973 | 0,834 |
| | 32292603 | 32320096 | rs6999977 | rs970997 | AACCGAGGAACGTACA | 0,899 | 0,393 |
| | 32292603 | 32320096 | rs6999977 | rs970997 | AACCAGTGAGTGCGCG | 1,75 | 0,000639 |
| | 32292603 | 32320096 | rs6999977 | rs970997 | AACCAGTGCGTGCGCG | 0,967 | 0,816 |
| | 32292603 | 32320096 | rs6999977 | rs970997 | AACTAGTGAGTGCGCG | 1,42 | 0,109 |
| | 32292603 | 32320096 | rs6999977 | rs970997 | AATCGAGGAACGTACA | 0,885 | 0,283 |
| | 32292603 | 32320096 | rs6999977 | rs970997 | AACCAATCAGTACGAG | 0,825 | 0,308 |
| GAIN | | | | | | | |
| | 32292603 | 32320096 | rs6999977 | rs970997 | AACGCTCCATCCAACA | 0,982 | 0,821 |
| | 32292603 | 32320096 | rs6999977 | rs970997 | AACGTTAGACTTGGAG | 1,06 | 0,619 |
| | 32292603 | 32320096 | rs6999977 | rs970997 | GTCGCTCCATCCAACA | 0,882 | 0,159 |
| | 32292603 | 32320096 | rs6999977 | rs970997 | AATGCTCCATCCAACA | 0,976 | 0,725 |
| | 32292603 | 32320096 | rs6999977 | rs970997 | AACGTCACACTCGGCG | 0,972 | 0,782 |
| | 32292603 | 32320096 | rs6999977 | rs970997 | AACGTCACCCTCGGCG | 0,933 | 0,433 |
| | 32292603 | 32320096 | rs6999977 | rs970997 | AACGTCACCCTTGGCG | 0,97 | 0,913 |
| | 32292603 | 32320096 | rs6999977 | rs970997 | AACATCACACTCGGCG | 1,22 | 0,101 |
| | 32292603 | 32320096 | rs6999977 | rs970997 | AACGTTAGACTCGGAG | 1,46 | 0,186 |
| GAIN | | | | | | | |
| | 32292603 | 32320096 | rs6999977 | rs970997 | AACGCTCCATCCAACA | 0,834 | 0,0278 |
| | 32292603 | 32320096 | rs6999977 | rs970997 | GTCGCTCCATCCAACA | 1,02 | 0,831 |
| | 32292603 | 32320096 | rs6999977 | rs970997 | AATGCTCCATCCAACA | 1,08 | 0,297 |
| | 32292603 | 32320096 | rs6999977 | rs970997 | AACGTTAGACTTGGAG | 0,747 | 0,0199 |
| | 32292603 | 32320096 | rs6999977 | rs970997 | AACGTCACCCTCGGCG | 1,11 | 0,246 |
| | 32292603 | 32320096 | rs6999977 | rs970997 | AACGTCACCCTTGGCG | 1,55 | 0,124 |
| | 32292603 | 32320096 | rs6999977 | rs970997 | AACGTTAGACTCGGAG | 1,4 | 0,246 |
| | 32292603 | 32320096 | rs6999977 | rs970997 | AACATCACACTCGGCG | 1,04 | 0,772 |
| | 32292603 | 32320096 | rs6999977 | rs970997 | AACGTCACACTCGGCG | 1,18 | 0,102 |

Using haplotype-based logistic regression analysis, significant haplotype blocks in two datasets were defined in the *ERBB4* and *NRG1* genes (Table 5.5 and 5.6). However, significance of these blocks had to be validated. To replicate previous results and to visualize the haplotype blocks, the Haploview software was used. The haplotype block structures of whole *ERBB4* and *NRG1* genes were visualized using "Solid Spine" analysis method, and the blocks were defined with D' greater than 0.6. SNPs that had Hardy-Weinberg Equilibrium value and minimum minor alleles less than 0.05 and 0.001, respectively, were eliminated. Then the significant haplotype blocks that emerged from logistic regression analysis were analysed, using the custom block analysis method (Figure 5.7a). The same six-kb block consisting of four SNPs (rs7586137, rs7589006, rs7561282 and rs4673623) was found to be significant in the CATIE and GAIN datasets, like in the logistic regression analysis. The significant haplotype in CATIE was C-G-A-G with a p value of 0.02 and was overrepresented in controls (case-control freq: 0.100-0.136) (Figure 5.7b). A different haplotype of the same block, T-G-G-C, was found to be significant in the GAIN dataset (p=0.0095) and again its frequency in controls was higher than in cases, 0.018 and 0.009, respectively (Figure 5.7c). In this four-SNP haplotype block, none of the haplotype blocks were defined as significant (Figure 5.7d).

For the *NRG1* gene, the same methodology was followed to replicate results of the logistic regression analysis. A 25-kb haplotype block at chr8: 32291552-32317192 was found to be significant in CATIE, GAIN and nonGAIN (Figure 5.8a). This region also included the most significant SNPs (rs10503907 and rs1487155), observed in GWADview. The most significant haplotype of this block, found in CATIE, is the A-A-C-C-G-T-G-A-T haplotype, which was overrepresented in cases over controls (case frequency=0.138, control frequemcy=0.085, p=0.0005) (Figure 5.8b). The G-C-T-C-C-T-C-C-A-C haplotype was marginally significant in GAIN (case frequency=0.117, control frequency=0.136, p=0.0589) (Figure 5.8c), while the A-A-C-G-C-T-C-C-A-C haplotype, residing in approximately 16% of cases and 19% of controls, was significant in the nonGAIN dataset (p= 0.0143) (Figure 5.8d).

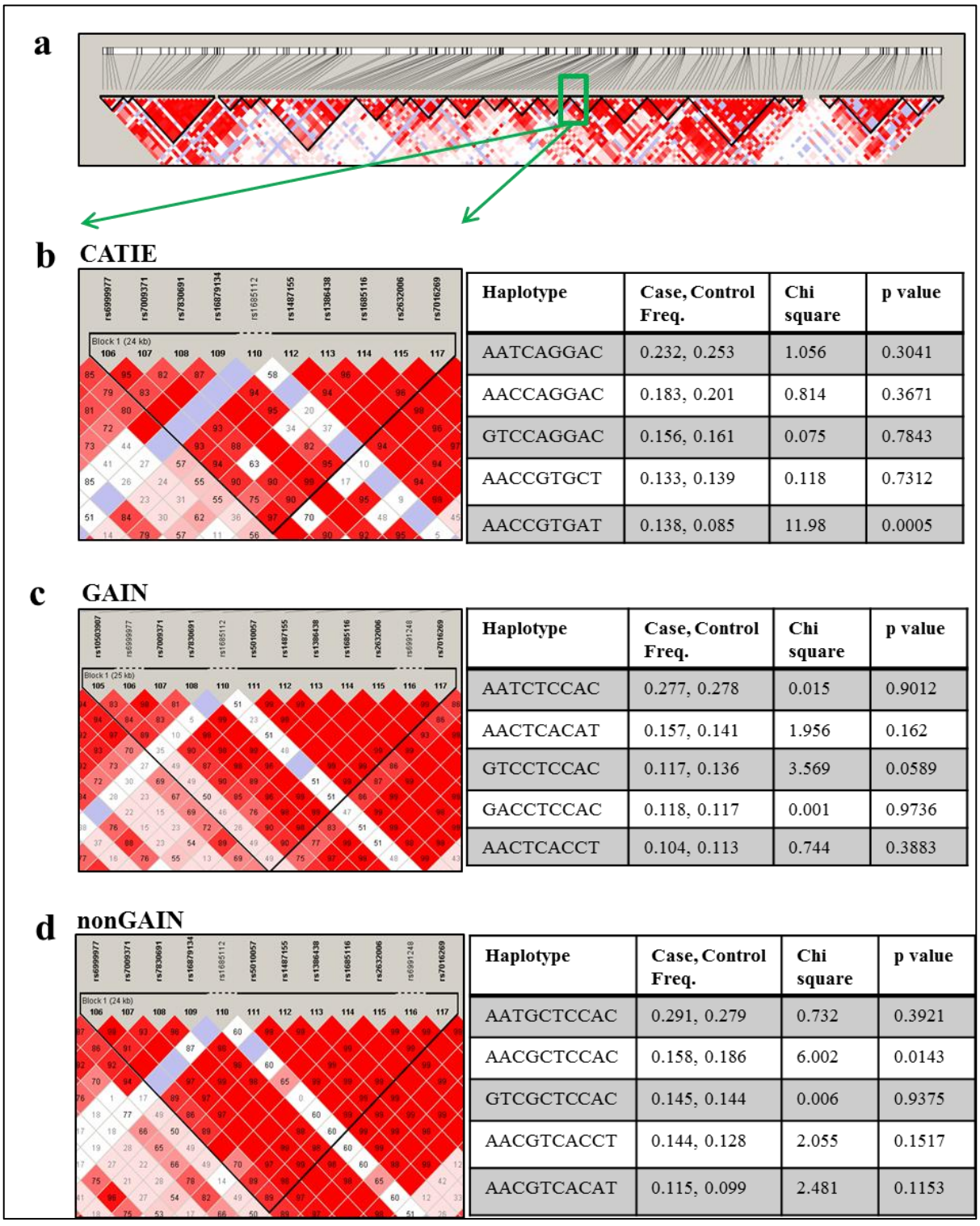



| Haplotype | Case, Control Freq. | Chi square | p value |
|---|---|---|---|
| TAGC | 0.852, 0.840 | 0.486 | 0.4859 |
| CGAG | 0.100, 0.136 | 5.361 | 0.0206 |
| CGAC | 0.027, 0.014 | 3.459 | 0.0629 |

| Haplotype | Case, Control Freq. | Chi square | p value |
|---|---|---|---|
| TAGC | 0.831, 0.830 | 0.012 | 0.9124 |
| CGAG | 0.136, 0.129 | 0.881 | 0.3479 |
| CGAC | 0.014, 0.014 | 0.064 | 0.8009 |
| TGGC | 0.009, 0.018 | 6.717 | 0.0095 |

| Haplotype | Case, Control Freq. | Chi square | p value |
|---|---|---|---|
| TAGC | 0.832, 0.841 | 0.522 | 0.4699 |
| CGAG | 0.129, 0.129 | 0.002 | 0.9671 |
| CGAC | 0.020, 0.015 | 1.672 | 0.196 |
| TGGC | 0.012, 0.011 | 6.717 | 0.9779 |

Figure 5.7. Haplotype analysis of the *ERBB4* region using the Haploview software: (a) LD structure of the whole *ERBB4* gene. Significant blocks and haplotypes in the (b) CATIE, (c) GAIN and (d) nonGAIN datasets.




| Haplotype | Case, Control Freq. | Chi square | p value |
|---|---|---|---|
| AATCAGGAC | 0.232, 0.253 | 1.056 | 0.3041 |
| AACCAGGAC | 0.183, 0.201 | 0.814 | 0.3671 |
| GTCCAGGAC | 0.156, 0.161 | 0.075 | 0.7843 |
| AACCGTGCT | 0.133, 0.139 | 0.118 | 0.7312 |
| AACCGTGAT | 0.138, 0.085 | 11.98 | 0.0005 |

| Haplotype | Case, Control Freq. | Chi square | p value |
|---|---|---|---|
| AATCTCCAC | 0.277, 0.278 | 0.015 | 0.9012 |
| AACTCACAT | 0.157, 0.141 | 1.956 | 0.162 |
| GTCCTCCAC | 0.117, 0.136 | 3.569 | 0.0589 |
| GACCTCCAC | 0.118, 0.117 | 0.001 | 0.9736 |
| AACTCACCT | 0.104, 0.113 | 0.744 | 0.3883 |

| Haplotype | Case, Control Freq. | Chi square | p value |
|---|---|---|---|
| AATGCTCCAC | 0.291, 0.279 | 0.732 | 0.3921 |
| AACGCTCCAC | 0.158, 0.186 | 6.002 | 0.0143 |
| GTCGCTCCAC | 0.145, 0.144 | 0.006 | 0.9375 |
| AACGTCACCT | 0.144, 0.128 | 2.055 | 0.1517 |
| AACGTCACAT | 0.115, 0.099 | 2.481 | 0.1153 |

Figure 5.8. Haplotype analysis of the *NRG1* region using the Haploview software: (a) LD structure of the whole *NRG1* gene. Significant blocks and haplotypes in the (b) CATIE, (c) GAIN and (d) nonGAIN datasets.

## 5.5. Mapping Transcription Binding Sites Altered Within the Blocks

### 5.5.1. Examination of Tag SNPs in the Blocks

*In silico* haplotype analyses showed that a six-kb haplotype block in the *ERBB4* and a 25-kb haplotype block in the *NRG1* genes were significantly associated with schizophrenia in all three CATIE, GAIN and nonGAIN datasets. However, it was not clear how these intronic regions might affect development of schizophrenia. Since it has been showed that transcription levels of *ERBB4* and *NRG1* differed in schizophrenia, transcription binding sites within candidate blocks were defined using the FASTSNP tool. Although the significant haplotypes of the *ERBB4* region were different in CATIE and GAIN, the most common haplotype was the same, T-A-G-C, carried by approximately 80% of all populations (Figure 5.9a). This haplotype enabled binding of NKX2 and GATA1/OCT1 transcription factors at the rs7589006 (allele A) and rs4673623 (C allele), respectively. In the significant haplotype of CATIE, C-G-A-G, conversion of A allele to G allele at rs7589006 created a new binding site for USF and deltaE transcription factors (Figure 5.9b). Other alterations were binding of C/EBPa/GATA3 and unbinding of GATA1/OCT1 transcription factors. Comparison of the significant haplotype of the GAIN dataset, T-G-G-C, with the most common haplotype, revealed that binding of USF and deltaE transcription factors instead of NKX2 at res759006 might be important in schizophrenia (Figure 5.9c). Since significant blocks in both CATIE and GAIN were found to be overrepresented in controls, binding of USF and deltaE transcription factors suggests a protective effect against schizophrenia. Since none of the haplotypes achieved the significance level ($p<0.05$) for the nonGAIN dataset, any transcription factor binding could not be concluded (Figure 5.9d).

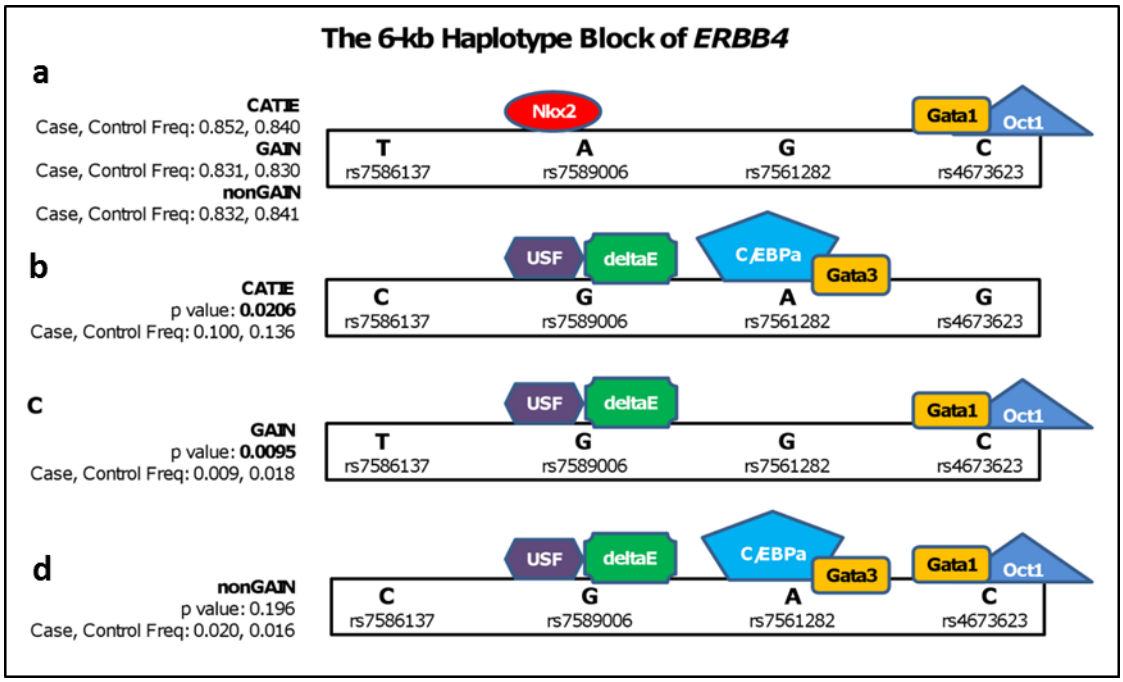



Figure 5.9. Transcription factors that bind common and significant haplotypes of *ERBB4*: (a) the most common haplotype in three datasets and the significant haplotypes of the (b) CATIE, (c) GAIN and (d) nonGAIN datasets.

FastSNP analysis revealed that the common alleles of the 11-SNP haplotype in all datasets facilitated binding of the CDXA, E2F, HFH2 complexes, when rs7009371 expressed the A allele and simultaneously the binding of the CDP CR at rs7016269, when it expressed the C allele (Figure 5.10a). The A/G-A-A-C-C-G/A-G-T-G-A-T haplotype in CATIE, which conferred a significantly increased risk, facilitated the binding sites suitable for CDXA, E2F, HFH2 (rs7009371, A allele) and S8, NKX2 and CDXA (rs7016269, T allele) (Figure 5.10b). Unlike CATIE, the significant haplotypes in GAIN and nonGAIN were found to have protective effects. The significant haplotypes in GAIN and nonGAIN showed increased binding affinity for TATA (rs7009371, T allele) and for CDP CR (rs7016269, C allele) (Figure 5.10c, d). The rs7009371 and rs7016269 resulted in binding of TATA and CDP CR, except the ones in CATIE. Mapping transcription factor binding sites in the most common and significant haplotypes of three datasets suggests that binding of CDXA, E2F, HFH2 at the position of rs7009371 (A allele), and S8, NKX2 and CDXA at the position of rs7016269 (T

allele) have a protective role in schizophrenia. On the other hand, formation of additional TATA binding sites in the candidate *NRG1* region might increase schizophrenia risk.

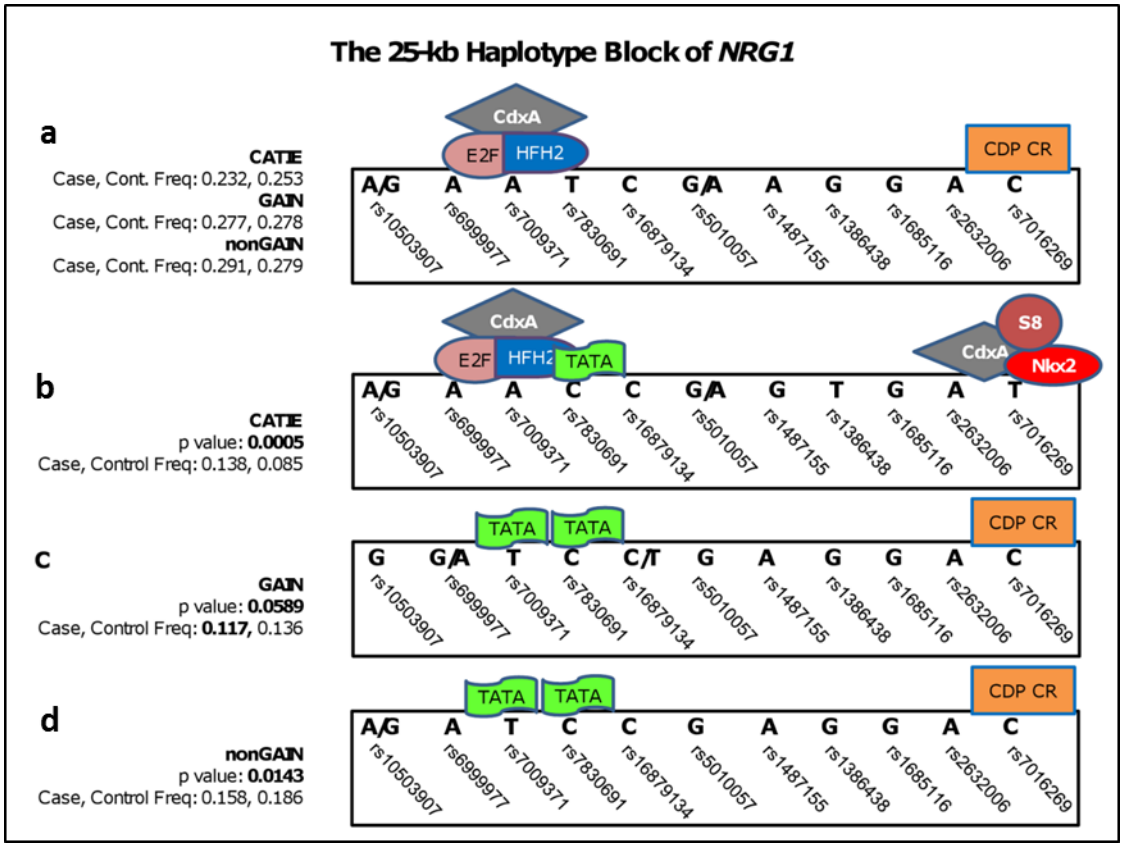



Figure 5.10. Transcription factors that bind common and significant haplotypes of *NRG1*: (a) the most common haplotype in three datasets and the significant haplotypes of the (b) CATIE, (c) GAIN and (d) nonGAIN datasets.

### 5.5.2. Examination of Rare SNPs in the Blocks

FastSNP analyses of rare SNPs that were annotated in dbSNP, yet not validated, suggested alterations of more critical transcription factor binding sites. These rare SNPs might cause three different alterations in binding affinities. For example, among the *NRG1* variants, some SNPs abolished existing binding sites such as rs13262178 [C binds NRF-2; T binds none], rs34393015 [wt binds Sox5-HFH2; 6pb insertion binds none], rs58045757 [G binds HNF-3b; delG binds none], rs34158863 [wt binds c-ETS; insG binds none] and rs34782215

[wt binds c-ETS, ELK1; insG binds none]. Other SNPs created new binding sites, such as rs12680997 [C binds none; T binds SRY], rs13278702 [G binds none; T binds GR] and rs34985716 [G binds none; T binds Sox5-SRY]. Further alterations in binding sites were replacement of an existing binding site with another, such as rs66776820 >6bp [TATA/E2F-CdxA, HFH2-Evi1], rs71832406 >6bp [S8/CdxA], rs35422231 [A binds CEBP-CDxA, C binds ARP1] and rs71512619 A/C [A binds deltaE-AML1a, and C binds MZF1]. The few rare SNPs identified within the *ERBB4* region, included rs13012759 [A binds CdxA; G binds none], rs12989265 C/T [C binds none; T binds SRY-Sox5], rs12989282 [C binds none; T binds SRY], and rs58786592 A/C [A binds c-ets; C binds Gata1-Gata2] (Table 5.7).

Table 5.7. Rare SNPs in six-kb *ERBB4* and 25-kb *NRG1* regions that change transcription binding sites.

| | | | | **FastSNP results** | | |
|---|---|---|---|---|---|---|
| **Chr. Position** | **SNP** | **Location** | **dbSNP allele** | **Risk** | **Wild type** | **Polymorphic** |
| *ERBB4* | | | | | | |
| 212449878 | rs58786592 | intron | A/C | 1-2 | c-ets | Gata1, Gata2 |
| 212451392 | rs12989265 | intron | C/T | 1-2 | - | SRY, Sox5, |
| 212451410 | rs12989282 | intron | C/T | 1-2 | - | SRY |
| 212451460 | rs13012759 | intron | A/G | 1-2 | CdxA | - |
| *NRG1* | | | | | | |
| 32172304 | rs12680997 | intron | C/T | 1-2 | - | SRY |
| 32172391 | rs13278702 | intron | G/T | 1-2 | - | GR |
| 32173481 | rs33923195 | intron | (>6bp) | 1-2 | - | Gata1, Gata2 |
| 32174315 | rs66776820 | intron | (>6bp) | 1-2 | TATA | E2F, CdxA, HFH2, Evi1 |
| 32174342 | rs3055550 | intron | (>6bp) | 1-2 | - | Evi1 |
| 32174587 | rs10614298 | intron | (>6bp) | 1-2 | - | IRF1 |
| 32176365 | rs35790889 | intron | -/A | 1-2 | Ik2 | - |
| 32177383 | rs71541814 | intron | CA/TG | 1-2 | Oct1 | - |

Table 5.7. Rare SNPs in six-kb *ERBB4* and 25-kb *NRG1* regions that change transcription binding sites (cont.).

| | | | | | | |
|---|---|---|---|---|---|---|
| 32179035 | rs10684641 | intron | -/CT | 1-2 | HFH-2 | SRY, Evi1 |
| 32179894 | rs71832406 | intron | (>6bp) | 1-2 | S8 | CdxA |
| 32180447 | rs6990964 | intron | C/T | 1-2 | - | Gata1 |
| 32180943 | rs9693341 | intron | G/T | 1-2 | MZF1 | - |
| 32181622 | rs72084148 | intron | -/AA | 1-2 | - | CdxA, HNF-3b |
| 32183447 | rs13262178 | intron | C/T | 1-2 | NRF-2 | - |
| 32184330 | rs35960536 | intron | -/T | 1-2 | C/EBPa, C/EBPb, CdxA | - |
| 32185944 | rs7825625 | intron | A/C | 1-2 | HFH-2, CdxA, HNF-3b | - |
| 32186438 | rs56657838 | intron | (>6bp) | 1-2 | - | deltaE |
| 32187323 | rs35638108 | intron | -/G | 1-2 | CdxA | C/EBP |
| 32187487 | rs35422231 | intron | A/C | 1-2 | C/EBP, CdxA | ARP1 |
| 32187503 | rs71512619 | intron | A/C | 1-2 | deltaE, AML-1a | MZF1 |
| 32187511 | rs35727240 | intron | C/T | 1-2 | - | C/ETS |
| 32188459 | rs10676449 | intron | (>6bp) | 1-2 | - | HNF-3b, HFH1 |
| 32188460 | rs34393015 | intron | (>6bp) | 1-2 | Sox5, HFH2 | - |
| 32188480 | rs58045757 | intron | G/T | 1-2 | HNF-3b | - |
| 32188605 | rs60754823 | intron | C/T | 1-2 | - | AML-1a, XFD-3 |
| 32189650 | rs72406944 | intron | -/A | 1-2 | HFH-1 | - |
| 32193522 | rs34158863 | intron | -/G | 1-2 | c-ETS | - |
| 32193545 | rs34782215 | intron | -/G | 1-2 | c-ETS, Elk1 | - |
| 32195198 | rs71208174 | intron | (>6bp) | 1-2 | Gata1 | - |
| 32195236 | rs1685114 | intron | A/G | 1-2 | - | Nkx2 |
| 32195289 | rs34464252 | intron | -/C | 1-2 | - | Gata1, MZF1 |
| 32197403 | rs34985716 | intron | G/T | 1-2 | - | Sox5, SRY |

### 5.6. Validation of Previously Identified *ERBB4* and *NRG1* Regions

Since this is the first report implicating the six-kb *ERBB4* and the 25-kb *NRG1* regions in European schizophrenia populations, previously identified *ERBB4* and *NRG1* regions were analysed using the analysis algorithm to prove the validity of the methodology. As discussed in the Introduction section, there were three major haplotype blocks in *ERBB4*. Most of the SNPs in these blocks were not assayed in all three datasets. This problem was solved by using the SNAP tool to identify SNPs which were in strong linkage disequilibrium with the previously identified SNPs. The threshold for $R^2$ has been set to 0.8 with a distance maximally confined to 500 kb in the HapMap CEU (Utah residents with Northern and Western Europe ancestry) sample. Among SNPs from the SNAP tool, the SNPs that were assayed in GWAS datasets were selected and their haplotypes were examined using the Haploview software. Two of the previously identified *ERBB4* blocks were found to be located in the third exon. These haplotypes failed to achieve significance threshold, but they implicated a borderline significance in the GAIN (p=0.0793) and nonGAIN (p=0.0527) datasets (Table 5.8). The other three-SNP haplotype, located in the region from intron 23 to exon 27, was validated in the CATIE, GAIN and nonGAIN study datasets (p=0.03, 0.048 and 0.034, respectively).

Many genetic studies identified different regions of *NRG1* as candidate schizophrenia haplotypes. These regions concentrated in two major areas: from 31550000 bp to 31850000 bp and from 32600000 to 32800000. These two regions were validated using the Haploview software to question the presence of any significant haplotype blocks in these areas. A relatively strong association of the 3' region haplotype (pCATIE=0.0008, pGAIN=0.0003 and pnonGAIN=0.0078) was identified when it was compared to the 5' $Hap_{ICE}$ haplotype (pCATIE =0.0166, pGAIN=0.0011 and pnonGAIN=0.0448) (Table 5.8). Nevertheless, the previously identified *ERBB4* and *NRG1* haplotypes in schizophrenia were also validated in the CATIE, GAIN and nonGAIN datasets, using a similar methodology that was used in the search for novel regions.

Table 5.8. Previously identified *ERBB4* and *NRG1* regions and their validations.

| | **Location**[a] | **Flanking regions** | **Associated p-value** [b] | **Proxy etc. information** [c] | **p-value CATIE** [d] | **p-value GAIN** [d] | **p-value nonGAIN** [d] |
|---|---|---|---|---|---|---|---|
| ***ERBB4*** | | | | | | | |
| 3-SNP (rs707284/rs839541-rs839523-rs7598440) haplotype | 212501443–212547291 / 212524334 | exon 3 | 0.00044 / 0.032 | $R^2$>0.8, H3, 10kb, CEU | No SNPs | 0.0793 | 0.0527 |
| 3-SNP (rs3748962-rs2289086-rs3791709) haplotype | 211960109–211993348 | intron 23-exon 27 | 0.020 | $R^2$>0.8, T, H3, 10kb, CEU | 0.03 | 0.048* | 0.034 |
| ***NRG1*** | | | | | | | |
| 5’ Region, The $Hap_{ICE}$ haplotype | 31475521-31785232 | exon 1 and 5’of intron 1 | 0.0001 – 0.05 | Haploview | 0.0166** | 0.0011** | 0.0448** |
| 3’ Region | 32600000-32800000 | several exons | 0.0001 – 0.005 | Haploview | 0.0008** | 0.0003** | 0.0078** |

a: NCBI Build 36
b: p-value of the most significant haplotype from previously published papers.
c: SNAP search criteria: R2>0.8 SNP dataset: T (1000 Genome Pilot 1), H3 (Hapmap3), Distance limit: 10 kb, Population: CEU
d: p-value of the most significant haplotype block with SNPs associated with previously identified haplotypes. P-values are calculated by Haploview with the Solid Spine analysis method (D’>0.6).
*: p-values are calculated by creating custom blocks using Haploview software.
**: p-values are calculated using the Haploview solid spine method (D’>0.6). The most significant haplotype blocks are different in CATIE, GAIN and nonGAIN.

# 6. DISCUSSION

Neuropsychiatric disorders pose a serious health problem that affects mental conditions of millions of people worldwide. Most of the diseases in this category have complex genetic structures which include a variety of common and rare variations in the genome. Like many others in this group, schizophrenia has been shown to be a complex disease with the involvement of both genetic and environmental factors. Genome-wide association studies have been a great source to unravel common variants that are responsible for the development of the disease. It has been six years since the first GWAS in schizophrenia was published, these studies have identified common variations with low penetrance, generally in noncoding regions of the genome (Mah *et al.*, 2006). While some GWASs were able to replicate the significant variants or genes, most of them could not validate results of each other. The most important reasons of this problem are use of samples from different ethnic backgrounds and application of a very stringent genome-wide significant threshold, $5 \times 10^{-8}$ (Pe'er *et al.*, 2008). The ERBB4-NRG1 signalling has been shown to be disrupted in schizophrenia in several studies. Moreover, many studies identified a genetic association of *NRG1* and *ERBB4* variants with a risk of developing schizophrenia in various ethnic groups (Silberberg *et al.*, 2006; Walker *et al.*, 2010). However, in none of the genome-wide association studies, variants in these two genes could achieve the genome-wide significance threshold. In this study, a novel approach was developed to prevent the rich GWAS raw data to miss significant risk haplotypes in schizophrenia. Novel and previously identified genetic associations of *NRG1* and *ERBB4* were analyzed using filtered European datasets from three GWASs. Having obtained promising results, this approach has been proven to be effective due to

- exploitation of multiple datasets of age and sex-matched schizophrenia patients and controls from European ethnicity,
- analyses of dense and quality SNP markers of *ERBB4/NRG1* with high genotyping call rate for SNPs and individuals,

- comprehensive genetic evaluation in which various haplotype-based association methods were systematically applied.

### 6.1. Pathway Analysis of Candidate Genes

In the first part of this study, candidate genes for schizophrenia were collected and subjected to pathway analysis. A total of 27 schizophrenia genes, which were validated in at least three different studies, were gathered using the GWAS Catalog, SZGene database and PubMed (Table 5.1). When these were subjected to pathway analysis, using the Ingenuity Pathway Analysis software, it was observed that most of the genes function in cell morphology, cell-to-cell signalling and cell death (Table 5.2). Genes responsible in cell morphology were *NRG1, TCF4, RGS4* and *RELN*. Pathway analysis also revealed the top five canonical pathways in schizophrenia: G-protein coupled receptor, cAMP-mediated, dopamine receptor, serotonin receptor and GM-CSF signalling pathways (Table 5.3). After evaluation of molecular and cellular functions of candidate schizophrenia genes, a network involving these genes was established (Figure 5.1). In this network, several new proteins that form interactions with our candidate schizophrenia proteins appeared as new candidate schizophrenia genes. For example, the phospholipase C-gamma1 protein (PLCG1) seems to have a key role in connecting ERBB4 and NRG1 with glutamate receptors and directing the signalling to a downstream effector, such as DISC1 and TCF4. *PLCG1* was not associated with schizophrenia previously, but it is a valuable candidate; because two studies identified association of exonic and intronic polymorphisms in *PLCG1* with lithium responsiveness in bipolar disorder (Ftouhi-Paquin *et al.*, 2001; Lovlie *et al.*, 2001).

Pathway analysis of candidate schizophrenia genes also re-emphasized the importance of the *ERBB4* and *NRG1* genes, because the network showed that they control the signalling from extracellular environment throughout the cell, and interact with glutamate receptors such as GRIK3 and GRIN2B. The supportive evidence comes from the observation that the hyperactivity of ERBB4-NRG1 signalling blocks NMDAR, a specific type of ionotropic glutamate receptor, via linker protein Src (Hahn *et al.*, 2006; Pitcher *et al.*, 2011).

## 6.2. Identifying Novel ERBB4 and NRG1 Regions

To analyse the association of *ERBB4* and *NRG1* with schizophrenia in large European schizophrenia cohorts, datasets from three genome-wide association studies were obtained. Access to these datasets, CATIE, GAIN and nonGAIN, was provided by the NIMH website and the dbGAP database. Prior to SNP and haplotype analyses, only European genotype data were extracted since the datasets included both European and African American samples. Then the age distributions of cases and controls were matched by eliminating samples older than 65 years in GAIN and nonGAIN, using the PLINK software.

Since *ERBB4* and *NRG1* are large genes with more than one Mb in length, it is very difficult to analyse haplotype blocks in the whole gene. The GWADView software was used to plot all SNPs in three datasets according to their chromosome locations and p values. The *ERBB4* region on chromosome 2 from approximately 211950000 bp to 212250000 bp covered seven significant SNPs from CATIE, two from GAIN and five from nonGAIN (Figure 5.5). This region included 5' UTR, exon 1 and the beginning of intron 1, but all significant SNPs were noncoding. Since none of the SNPs were very significant, all p values were between 0.05 and 0.001, the haplotypes in this area were examined using haplotype-based logistic regression test. Logistic regression is a method of modelling the effects of two susceptibility loci, influencing affection status and can be used to examine individual effects from multiple variants in linkage disequilibrium with each other (North *et al.*, 2006). Logistic regression analyses require a block structure of a given area to create haplotypes and to calculate the significance of each allele. Since CATIE, GAIN and nonGAIN had different numbers of SNPs, each had different block structures. To establish uniformity, the block structure of CATIE was used for logistic regression tests of all three datasets. The haplotype block between 212156823 bp and 212162848 bp gave the most significant haplotypes: C-G-A-G in CATIE with a p value of 0.0183 and T-G-G-C in GAIN with a p value of 0.00799 (Table 5.5).

The p values of *ERBB4* haplotypes in logistic regression tests were significant, but since they were not very powerful, another type of haplotype analysis was performed via Haploview to ascertain their significance. The same 6-kb haplotype block in *ERBB4* on chr2 between the

positions 212156823 and 212162828, which was significantly associated with risk of developing schizophrenia in the Caucasian samples of CATIE (C-G-A-G, p=0.0206) and GAIN datasets (T-G-G-C, p=0.0095) was validated (Figure 5.7). This 4-SNP haplotype block, consisting of rs7586137, rs7589006, rs7561282 and rs4673623, was significantly overrepresented in controls compared to schizophrenia patients. Although the significant haplotypes of 4-SNP block of *ERBB4* were different in CATIE and GAIN datasets, they both conferred a protective role. This novel 6-kb block, located within intron 19 of *ERBB4*, is part of a genomic region encoding the tyrosine kinase domain (known as the catalytic domain) (Uniprot ID: Q15303) that spans a region from exon 18 to exon 24. Although this haplotype does not affect the coding region directly, it may lead to altered efficiency of the splicing mechanism via altering sites for TFs and thus, resulting in modified transcription efficiency. Previously a noncoding variant in intron 12 (rs4673628) and three intronic risk SNPs (rs7598440, rs707284, rs839523) surrounding exon 3, have been shown to be strongly associated with splice variants JM-a and CYT-1 which had elevated expression levels in brain samples of schizophrenia patients (Law *et al.*, 2007). It is the first time that the six-kb haplotype block was associated with schizophrenia, further in silico and in vitro studies are needed to confirm the effect of this six-kb haplotype block on expression levels of *ERBB4* isoforms in schizophrenia.

The same methodology was applied to the *NRG1* gene. In the GWADView tool, the region between 32270000 bp and 32450000 bp on chromosome 8 has 14 significant SNPs in CATIE and nine significant SNPs in the nonGAIN datasets (Figure 5.6). One of the SNPs in the CATIE dataset, rs1487154, had a very significant p value of ~0.0001. When this large region was subjected to haplotype-based logistic regression, a novel haplotype between rs6999977 and rs970997 had the most significant haplotype ($p_{CATIE}$=0.0006, $p_{nonGAIN}$=0.0199) (Table 5.6). To ascertain the significance of these blocks, haplotype analysis was performed using Haploview. The same novel 25-kb haplotype block (11 SNPs) of *NRG1* was found to be located between the positions 32291552 to 32317192 (Figure 5.8). The significant haplotype in the CATIE dataset was overrepresented in cases, while significant haplotypes of the same block in GAIN and nonGAIN were more prevalent in controls than in cases. This block, which is located within intron 1 of *NRG1*, is likely to impact on promoter-enhancer sequences that

are part of first introns of genes. Our analysis highlights the importance of this particular *NRG1* block in schizophrenia, such that the combination of different alleles of these SNPs within the corresponding block might have an impact on the regulation of *NRG1* gene expression. Previous studies have shown that the noncoding variants in the promoter regions located in 5' UTR and the first intron were associated with the mRNA expression of a specific type of NRG1 isoform. Schizophrenia patients carrying the T allele of SNPNRG243177 (rs6994992) were shown to have a higher hippocampal mRNA expression of type IV *NRG1* isoform (Law *et al.*, 2006). The same allele was correlated with decrease in white matter size in brains of schizophrenia patients (McIntosh *et al.*, 2008) and recently was reported to be associated with reduced grey matter volume (Barnes *et al.*, 2012). Another variant in the same region, SNP8NRG222662 (rs4623364), was associated with volume reductions in superior temporal gyrus in the upper part of the temporal lobe in schizophrenia patients (Tosato *et al.*, 2012). All these evidence support the hypothesis that the single variants and/or haplotype blocks in 5' end of the *NRG1* gene might affect expression of the gene and are responsible in volume change in brains of schizophrenia patients.

Although haplotype analysis of the *ERBB4* and *NRG1* genes revealed significant alleles that were associated with schizophrenia in European GWAS datasets, the p values were not very strong. Since schizophrenia is a very complex neuropsychiatric disorder, which is under the influence of a variety of genetic factors, association of a single haplotype might not be very strong. There are several ways to strengthen the association: First, epistasis, the interaction between genes, can be examined using statistical tests (Cordell, 2002). In the case of *ERBB4* and *NRG1*, it has been shown that in 416 Japanese schizophrenia patients and 520 matched controls, none of the variants of *ERBB4* and *NRG1* were associated with the disease, whereas epistasis analysis of two variants, rs2919381 in *NRG1* and rs7560730 in *ERBB4*, significantly associated with schizophrenia, even after correction (Shiota *et al.*, 2008). Second, in particular using endophenotypes might increase the significance of the association. In a recent study, while *ERBB4* was associated with Latter-Number Sequencing Test score and California Verbal Learning Test score in European schizophrenia patients, *NRG1* was associated with p50 suppression (Greenwood *et al.*, 2012). Similarly, the SNP8NRG243177 (rs6994992) polymorphism increased the risk of psychosis in Hungarian schizophrenia

patients (Keri *et al.,* 2009). A third way is to increase sample and data size, which might be achieved by simply increasing the number of subjects in genome-wide association studies. The strengthening effect of increasing sample size has been shown for schizophrenia, as well as for other diseases (Kim *et al.,* 2011). Higher number of samples might also be achieved by meta- (combining results) and mega-analysis (combining data) (Bergen and Petryshen, 2012). In this case, one must be careful not to mix data from different ethnic backgrounds, because in this case population-specific associations will be underestimated. Moreover the datasets, that are mixed, might be from different genotype arrays and imputation may be needed.

### 6.3. Transcription Binding Sites Altered within the Blocks

Since both novel haplotype blocks were located within introns, these regions are likely to regulate the expression levels of the genes via transcription factor binding or altering splicing. Each SNP, located within candidate haplotypes, was investigated in relation to changes in transcription factor binding site affinities and pre-mRNA splicing, using the fastSNP web server. While none of the variants was found to change ERBB4 or NRG1 splicing directly, alterations in TF binding were observed between major and minor alleles. With respect to the candidate *ERBB4* region, binding of USF and deltaE, instead of NKX2, was associated with a protective role in schizophrenia, which was due in part by the rs7589006 variant within the 4-SNP haplotype (Figure 5.9). Similarly, with respect to *NRG1* haplotype block, binding of the TATA, instead of CDXA-E2F-HFH2, had a protective effect, whereas S8-CDXA-NKX2 binding to the 3' end of the block increased the risk of schizophrenia in the European population (Figure 5.10).

In addition to common SNPs investigated (as part of the SNP array data) within susceptible haplotype blocks, many candidate SNPs were in LD. However, they were not assayed in CATIE, GAIN and nonGAIN GWAS. For example, although three datasets encompassed only four SNPs in the 6-kb *ERBB4* region, there were a total of 63 variants in this region, some of which might represent a potential variant with pathogenic consequences. For *NRG1*, we observed a clear difference in transcription factor binding on block between protective and causative haplotypes. The results suggested that the binding of

CDXA/E2F/HFH2 and S8/NKX2/CDXA to the *NRG1* block increased the schizophrenia risk in European populations, and their dissociation and association of TATA and CDP CR transcription factors might be protective (Table 5.8). Some of these variants were expected to alter TFBS. Therefore, a combinatorial influence of several SNPs in gene regulation might affect schizophrenia development.

### 6.4. Validation of the Previously Identified Regions

The first 3-SNP haplotype of *ERBB4* (rs707284-rs839523-rs7598440), surrounding exon 3, was found to be previously associated with schizophrenia in the Ashkenazi population (Silberberg *et al.*, 2006). However, they were borderline-significant ($0.05<p<0.1$) in the GAIN and nonGAIN datasets (Table 5.8). The 3-SNP G-A-A haplotype (rs3748962-rs2289086-rs3791709) flanking intron 23 - exon 27 (Nicodemus *et al.*, 2006), was also validated in the three datasets used in this study: 296 families from Clinical Brain Disorders Branch/National Institute of Mental Health Sibling Study were compared with 370 healthy controls in family-based affection analyses. While this study identified a G-A-A haplotype at p-value of 0.02 significance, the similar region was significant ($0.03<p<0.048$) in the CATIE, GAIN and nonGAIN datasets. Since the same SNPs were not assayed in these GWAS datasets, we could not construct the exact haplotypes in these datasets. The same haplotype, shown to be significant in the Han Chinese case-control study (CTA haplotype, p=0.02, case vs. control= 36% vs 24%) (Lu *et al.*, 2010). rs3748962 and causing a synonymous variant in exon 27 (Val1065Val), was implicated to play a role in variable mRNA expression of maternal and paternal chromosomes in the brain (Norton *et al.*, 2006). Although the haplotype was re-validated in three independent European populations in this project, *in vivo* or *in vitro* studies are necessary to reveal its effect and possible cis-acting element in ERBB4-NRG1 signalling in the mechanism underlying schizophrenia.

The studies that have shown an association between *NRG1* and schizophrenia risk mainly focused on two genomic regions, region A ($Hap_{ICE}$) (Walker *et al.*, 2010) and region B (32600000bp-32800000bp). $Hap_{ICE}$ haplotype, located at the 5' end of the gene, covering exon 1 and the 5' of intron 1, were also validated in our study (Stefansson *et al.*, 2002). The second

region, Region B, covered most of the exons of *NRG1* which were concentrated at the 3' end of the gene as reported in Walker et al (Walker *et al.*, 2010). Several studies have identified different haplotype blocks and polymorphisms in region B as significant in various schizophrenia populations (Lachman *et al.*, 2006; Petryshen *et al.*, 2005; Thomson *et al.*, 2007). Haploview analysis in our study implicated that different blocks in region B of *NRG1* were significant in three Caucasian GWA datasets. Although the 5' end of the *NRG1* gene, including the $Hap_{ICE}$ haplotype and other haplotypes nearby, was shown to be significant in different ethnic schizophrenia populations, our study supports the role of the 3' end of the *NRG1* gene on the prevalence of schizophrenia in European populations, a finding which has been suggested in only a few studies to date (Walker *et al.*, 2010).

This is the first study to perform an *in silico* and bioinformatics-based functional analysis of variants located within introns in schizophrenia subjects from GWA datasets. This method has given promising results that facilitated our understanding of the functional role of intronic variants. However, future studies should focus on the validation of these results by *in vitro* and *in vivo* studies.

# REFERENCES

Abi-Dargham, A., O. Mawlawi, I. Lombardo, R. Gil, D. Martinez, Y. Huang, D. R. Hwang, J. Keilp, L. Kochan, R. Van Heertum, J. M. Gorman, and M. Laruelle, 2002, "Prefrontal Dopamine D1 Receptors and Working Memory in Schizophrenia", *Journal of Neuroscience*, Vol. 22, No. 9, pp. 3708-3719.

Alaerts, M., S. Ceulemans, D. Forero, L. N. Moens, S. De Zutter, L. Heyrman, A. S. Lenaerts, K. F. Norrback, P. De Rijk, L. G. Nilsson, D. Goossens, R. Adolfsson, and J. Del-Favero, 2009, "Support for Nrg1 as a Susceptibility Factor for Schizophrenia in a Northern Swedish Isolated Population", *Archives of Genetic Psychiatry*, Vol. 66, No. 8, pp. 828-837.

Alkelai, A., S. Lupoli, L. Greenbaum, I. Giegling, Y. Kohn, K. Sarner-Kanyas, E. Ben-Asher, D. Lancet, D. Rujescu, F. Macciardi, and B. Lerer, 2011a, "Identification of New Schizophrenia Susceptibility Loci in an Ethnically Homogeneous, Family-Based, Arab-Israeli Sample", *The Journal of the Fedaration of American Socities for Experimental Biology*, Vol. 25, No. 11, pp. 4011-4023.

Alkelai, A., S. Lupoli, L. Greenbaum, Y. Kohn, K. Kanyas-Sarner, E. Ben-Asher, D. Lancet, F. Macciardi, and B. Lerer, 2011b, "Dock4 and Ceacam21 as Novel Schizophrenia Candidate Genes in the Jewish Population", *International Journal of Neuropsychopharmacology*, Vol. 15, No. 4, pp. 1-11.

Anton, E. S., H. T. Ghashghaei, J. L. Weber, C. McCann, T. M. Fischer, I. D. Cheung, M. Gassmann, A. Messing, R. Klein, M. H. Schwab, K. C. Lloyd, and C. Lai, 2004, "Receptor Tyrosine Kinase Erbb4 Modulates Neuroblast Migration and Placement in the Adult Forebrain ", *Nature Neuroscience*, Vol. 7, No. 12, pp. 1319-1328.

Arguello, P. A., and J. A. Gogos, 2008, "A Signaling Pathway Akting up in Schizophrenia ", *Journal of Clinical Invesigation,* Vol. 118, No. 6, pp. 2018-2021.

Athanasiu, L., M. Mattingsdal, A. K. Kahler, A. Brown, O. Gustafsson, I. Agartz, I. Giegling, P. Muglia, S. Cichon, M. Rietschel, O. P. Pietilainen, L. Peltonen, E. Bramon, D. Collier, D. S. Clair, E. Sigurdsson, H. Petursson, D. Rujescu, I. Melle, V. M. Steen, S. Djurovic, and O. A. Andreassen, 2010, "Gene Variants Associated with Schizophrenia in a Norwegian Genome-Wide Study Are Replicated in a Large European Cohort ", *Journal of Psychiatric Research,* Vol. 44, No. 12, pp. 748-753.

Badner, J. A., and E. S. Gershon, 2002, "Meta-Analysis of Whole-Genome Linkage Scans of Bipolar Disorder and Schizophrenia ", *Moecular Psychiatry,* Vol. 7, No. 4, pp. 405-411.

Barnes, A., M. Isohanni, J. H. Barnett, O. Pietilainen, J. Veijola, J. Miettunen, T. Paunio, P. Tanskanen, K. Ridler, J. Suckling, E. T. Bullmore, P. B. Jones, and G. K. Murray, 2012, "Neuregulin-1 Genotype Is Associated with Structural Differences in the Normal Human Brain ", *Neuroimage,* Vol. 59, No. 3, pp. 2057-2061.

Baum, A. E., N. Akula, M. Cabanero, I. Cardona, W. Corona, B. Klemens, T. G. Schulze, S. Cichon, M. Rietschel, M. M. Nothen, A. Georgi, J. Schumacher, M. Schwarz, R. Abou Jamra, S. Hofels, P. Propping, J. Satagopan, S. D. Detera-Wadleigh, J. Hardy, and F. J. McMahon, 2008, "A Genome-Wide Association Study Implicates Diacylglycerol Kinase Eta (Dgkh) and Several Other Genes in the Etiology of Bipolar Disorder ", *Molecular Psychiatry,* Vol. 13, No. 2, pp. 197-207.

Belforte, J. E., V. Zsiros, E. R. Sklar, Z. Jiang, G. Yu, Y. Li, E. M. Quinlan, and K. Nakazawa, 2010, "Postnatal Nmda Receptor Ablation in Corticolimbic Interneurons Confers Schizophrenia-Like Phenotypes", *Nature Neuroscience,* Vol. 13, No. 1, pp. 76-83.

Ben-Shachar, D., and D. Laifenfeld, 2004, "Mitochondria, Synaptic Plasticity, and Schizophrenia", *International Reviews of Neurobiology,* Vol. 59, pp. 273-296.

Benzel, I., A. Bansal, B. L. Browning, N. W. Galwey, P. R. Maycox, R. McGinnis, D. Smart, D. St Clair, P. Yates, and I. Purvis, 2007, "Interactions among Genes in the Erbb-Neuregulin Signalling Network Are Associated with Increased Susceptibility to Schizophrenia", *Behavioral and Brain Functions,* Vol. 3, pp. 31.

Bergen, S. E., and T. L. Petryshen, 2012, "Genome-Wide Association Studies of Schizophrenia: Does Bigger Lead to Better Results?", *Current Opinion in Psychiatry,* Vol. 25, No. 2, pp. 76-82.

Carlsson, A., and M. Lindqvist, 1963, "Effect of Chlorpromazine or Haloperidol on Formation of 3methoxytyramine and Normetanephrine in Mouse Brain", *Acta Pharmacologica et Toxicologica (Copenh),* Vol. 20, pp. 140-144.

Chakravarti, A., 1999, "Population Genetics--Making Sense out of Sequence", *Nature Genetics,* Vol. 21, No. 1 Suppl, pp. 56-60.

Chen, J., G. Lee, A. H. Fanous, Z. Zhao, P. Jia, A. O'Neill, D. Walsh, K. S. Kendler, and X. Chen, 2011, "Two Non-Synonymous Markers in Ptpn21, Identified by Genome-Wide Association Study Data-Mining and Replication, Are Associated with Schizophrenia", *Schizophrenia Research,* Vol. 131, No. 1-3, pp. 43-51.

Chen, P., J. Chen, K. Huang, W. Ji, T. Wang, T. Li, Y. Wang, H. Wang, L. He, G. Feng, and Y. Shi, 2012, "Analysis of Association between Common Snps in Erbb4 and Bipolar Affective Disorder, Major Depressive Disorder and Schizophrenia in the Han Chinese Population", *Progress in Neuro-psychopharmacoogyl and Biological Psychiatry,* Vol. 36, No. 1, pp. 17-21.

Chen, X., X. Wang, S. Hossain, F. A. O'Neill, D. Walsh, L. Pless, K. V. Chowdari, V. L. Nimgaonkar, S. G. Schwab, D. B. Wildenauer, P. F. Sullivan, E. van den Oord, and K. S. Kendler, 2006, "Haplotypes Spanning Spec2, Pdz-Gef2 and Acsl6 Genes Are Associated with Schizophrenia", *Human Molecular Genetics,* Vol. 15, No. 22, pp. 3329-3342.

Chong, V. Z., M. Thompson, S. Beltaifa, M. J. Webster, A. J. Law, and C. S. Weickert, 2008, "Elevated Neuregulin-1 and Erbb4 Protein in the Prefrontal Cortex of Schizophrenic Patients", *Schizophrenia Research,* Vol. 100, No. 1-3, pp. 270-280.

Chung, C., T. Tallerico, and P. Seeman, 2003, "Schizophrenia Hippocampus Has Elevated Expression of Chondrex Glycoprotein Gene", *Synapse,* Vol. 50, No. 1, pp. 29-34.

Cordell, H. J., 2002, "Epistasis: What It Means, What It Doesn't Mean, and Statistical Methods to Detect It in Humans", *Human Molecular Genetics,* Vol. 11, No. 20, pp. 2463-2468.

Craddock, N., M. C. O'Donovan, and M. J. Owen, 2005, "The Genetics of Schizophrenia and Bipolar Disorder: Dissecting Psychosis", *Journal of Medical Genetics,* Vol. 42, No. 3, pp. 193-204.

Craddock, N., M. C. O'Donovan, and M. J. Owen, 2007, "Phenotypic and Genetic Complexity of Psychosis. Invited Commentary On ... Schizophrenia: A Common Disease Caused by Multiple Rare Alleles", *British Journal of Psychiatry,* Vol. 190, pp. 200-203.

Curley, A. A., D. Arion, D. W. Volk, J. K. Asafu-Adjei, A. R. Sampson, K. N. Fish, and D. A. Lewis, 2011, "Cortical Deficits of Glutamic Acid Decarboxylase 67 Expression in Schizophrenia: Clinical, Protein, and Cell Type-Specific Features", *The Americal Journal of Psychiatry,* Vol. 168, No. 9, pp. 921-929.

Davis, K. L., R. S. Kahn, G. Ko, and M. Davidson, 1991, "Dopamine in Schizophrenia: A Review and Reconceptualization", *The Americal Journal of Psychiatry*, Vol. 148, No. 11, pp. 1474-1486.

Doherty, J. L., M. C. O'Donovan, and M. J. Owen, 2012, "Recent Genomic Advances in Schizophrenia", *Clinical Genetics*, Vol. 81, No. 2, pp. 103-109.

Duan, J., M. Martinez, A. R. Sanders, C. Hou, A. J. Krasner, D. B. Schwartz, and P. V. Gejman, 2005, "Neuregulin 1 (Nrg1 ) and Schizophrenia: Analysis of a Us Family Sample and the Evidence in the Balance", *Psychological Medicine*, Vol. 35, No. 11, pp. 1599-1610.

Elenius, K., G. Corfas, S. Paul, C. J. Choi, C. Rio, G. D. Plowman, and M. Klagsbrun, 1997, "A Novel Juxtamembrane Domain Isoform of Her4/Erbb4. Isoform-Specific Tissue Distribution and Differential Processing in Response to Phorbol Ester", *The Journal of Biological Chemistry*, Vol. 272, No. 42, pp. 26761-26768.

Falls, D. L., 2003, "Neuregulins: Functions, Forms, and Signaling Strategies", *Experimental Cell Research*, Vol. 284, No. 1, pp. 14-30.

Fatemi, S. H., and T. D. Folsom, 2009, "The Neurodevelopmental Hypothesis of Schizophrenia, Revisited", *Schizophrenia Bulletin*, Vol. 35, No. 3, pp. 528-548.

Flames, N., J. E. Long, A. N. Garratt, T. M. Fischer, M. Gassmann, C. Birchmeier, C. Lai, J. L. Rubenstein, and O. Marin, 2004, "Short- and Long-Range Attraction of Cortical Gabaergic Interneurons by Neuregulin-1", *Neuron*, Vol. 44, No. 2, pp. 251-261.

Fleischhacker, W. W., and C. G. Widschwendter, 2006, "Treatment of Schizophrenia Patients: Comparing New-Generation Antipsychotics to Each Other", *Current Opinion in Psychiatry*, Vol. 19, No. 2, pp. 128-134.

Ftouhi-Paquin, N., M. Alda, P. Grof, N. Chretien, G. Rouleau, and G. Turecki, 2001, "Identification of Three Polymorphisms in the Translated Region of Plc-Gamma1 and Their Investigation in Lithium Responsive Bipolar Disorder", *American Journal of Medical Genetics*, Vol. 105, No. 3, pp. 301-305.

Fukui, N., T. Muratake, N. Kaneko, H. Amagane, and T. Someya, 2006, "Supportive Evidence for Neuregulin 1 as a Susceptibility Gene for Schizophrenia in a Japanese Population", *Neuroscience Letter*, Vol. 396, No. 2, pp. 117-120.

Garcia-Barcelo, M. M., X. Miao, C. S. Tang, H. C. So, W. Tang, T. Y. Leon, M. So, B. Yip, R. Y. Chen, E. F. Cheung, E. Y. Chen, T. Li, P. Tam, S. S. Cherny, and P. C. Sham, 2011, "No Nrg1 V266l in Chinese Patients with Schizophrenia", *Psychiatric Genetics*, Vol. 21, No. 1, pp. 47-49.

Georgieva, L., A. Dimitrova, D. Ivanov, I. Nikolov, N. M. Williams, D. Grozeva, I. Zaharieva, D. Toncheva, M. J. Owen, G. Kirov, and M. C. O'Donovan, 2008, "Support for Neuregulin 1 as a Susceptibility Gene for Bipolar Disorder and Schizophrenia", *Biological Psychiatry*, Vol. 64, No. 5, pp. 419-427.

Girard, S. L., J. Gauthier, A. Noreau, L. Xiong, S. Zhou, L. Jouan, A. Dionne-Laporte, D. Spiegelman, E. Henrion, O. Diallo, P. Thibodeau, I. Bachand, J. Y. Bao, A. H. Tong, C. H. Lin, B. Millet, N. Jaafari, R. Joober, P. A. Dion, S. Lok, M. O. Krebs, and G. A. Rouleau, 2011, "Increased Exonic De Novo Mutation Rate in Individuals with Schizophrenia", *Nature Genetics*, Vol. 43, No. 9, pp. 860-863.

Goerendt, I. K., C. Messa, A. D. Lawrence, P. M. Grasby, P. Piccini, and D. J. Brooks, 2003, "Dopamine Release During Sequential Finger Movements in Health and Parkinson's Disease: A Pet Study", *Brain*, Vol. 126, No. 2, pp. 312-325.

Goff, D. C., C. Cather, A. E. Evins, D. C. Henderson, O. Freudenreich, P. M. Copeland, M. Bierer, K. Duckworth, and F. M. Sacks, 2005, "Medical Morbidity and Mortality in

Schizophrenia: Guidelines for Psychiatrists", *Journal of Clinical Psychiatry*, Vol. 66, No. 2, pp. 183-194; quiz 147, 273-184.

Gong, Y. G., C. N. Wu, Q. H. Xing, X. Z. Zhao, J. Zhu, and L. He, 2009, "A Two-Method Meta-Analysis of Neuregulin 1(Nrg1) Association and Heterogeneity in Schizophrenia", *Schizophrenia Research*, Vol. 111, No. 1-3, pp. 109-114.

Gordon, J. A., 2010, "Testing the Glutamate Hypothesis of Schizophrenia", *Nature Neuroscience*, Vol. 13, No. 1, pp. 2-4.

Gouzoulis-Mayfrank, E., K. Heekeren, A. Neukirch, M. Stoll, C. Stock, M. Obradovic, and K. A. Kovar, 2005, "Psychological Effects of (S)-Ketamine and N,N-Dimethyltryptamine (Dmt): A Double-Blind, Cross-over Study in Healthy Volunteers", *Pharmacopsychiatry*, Vol. 38, No. 6, pp. 301-311.

Gray, J. A., and B. L. Roth, 2007, "The Pipeline and Future of Drug Development in Schizophrenia", *Molecular Psychiatry*, Vol. 12, No. 10, pp. 904-922.

Greenwood, T. A., G. A. Light, N. R. Swerdlow, A. D. Radant, and D. L. Braff, 2012, "Association Analysis of 94 Candidate Genes and Schizophrenia-Related Endophenotypes", *PLoS One*, Vol. 7, No. 1, pp. e29630.

Hahn, C. G., H. Y. Wang, D. S. Cho, K. Talbot, R. E. Gur, W. H. Berrettini, K. Bakshi, J. Kamins, K. E. Borgmann-Winter, S. J. Siegel, R. J. Gallop, and S. E. Arnold, 2006, "Altered Neuregulin 1-Erbb4 Signaling Contributes to Nmda Receptor Hypofunction in Schizophrenia", *Nature Medicine*, Vol. 12, No. 7, pp. 824-828.

Hakak, Y., J. R. Walker, C. Li, W. H. Wong, K. L. Davis, J. D. Buxbaum, V. Haroutunian, and A. A. Fienberg, 2001, "Genome-Wide Expression Analysis Reveals Dysregulation of Myelination-Related Genes in Chronic Schizophrenia", *Proceedings of the National Academy of Sciences*, Vol. 98, No. 8, pp. 4746-4751.

Hanninen, K., H. Katila, M. Saarela, R. Rontu, K. M. Mattila, M. Fan, M. Hurme, and T. Lehtimaki, 2008, "Interleukin-1 Beta Gene Polymorphism and Its Interactions with Neuregulin-1 Gene Polymorphism Are Associated with Schizophrenia", *European Archives of Psychiatry and Clinical Neuroscience,* Vol. 258, No. 1, pp. 10-15.

Heinrichs, R. W., 2003, "Historical Origins of Schizophrenia: Two Early Madmen and Their Illness", *Journal of the History of the Behavioral Sciences,* Vol. 39, No. 4, pp. 349-363.

Hong, C. J., S. J. Huo, D. L. Liao, K. Lee, J. Y. Wu, and S. J. Tsai, 2004, "Case-Control and Family-Based Association Studies between the Neuregulin 1 (Arg38gln) Polymorphism and Schizophrenia", *Neuroscience Letters,* Vol. 366, No. 2, pp. 158-161.

Horn, J. L., and R. B. Cattell, 1966, "Refinement and Test of the Theory of Fluid and Crystallized General Intelligences", *Journal of Educational Psychology,* Vol. 57, No. 5, pp. 253-270.

Howes, O. D., and S. Kapur, 2009, "The Dopamine Hypothesis of Schizophrenia: Version Iii--the Final Common Pathway", *Schizophrenia Bulletin,* Vol. 35, No. 3, pp. 549-562.

Howes, O. D., A. J. Montgomery, M. C. Asselin, R. M. Murray, P. M. Grasby, and P. K. McGuire, 2007, "Molecular Imaging Studies of the Striatal Dopaminergic System in Psychosis and Predictions for the Prodromal Phase of Psychosis", *The British Journal of Psychiatry. Supplement,* Vol. 51, pp. s13-18.

Ikeda, M., B. Aleksic, Y. Kinoshita, T. Okochi, K. Kawashima, I. Kushima, Y. Ito, Y. Nakamura, T. Kishi, T. Okumura, Y. Fukuo, H. J. Williams, M. L. Hamshere, D. Ivanov, T. Inada, M. Suzuki, R. Hashimoto, H. Ujike, M. Takeda, N. Craddock, K. Kaibuchi, M. J. Owen, N. Ozaki, M. C. O'Donovan, and N. Iwata, 2011, "Genome-

Wide Association Study of Schizophrenia in a Japanese Population", *Biological Psychiatry,* Vol. 69, No. 5, pp. 472-478.

Ikeda, M., N. Takahashi, S. Saito, B. Aleksic, Y. Watanabe, A. Nunokawa, Y. Yamanouchi, T. Kitajima, Y. Kinoshita, T. Kishi, K. Kawashima, R. Hashimoto, H. Ujike, T. Inada, T. Someya, M. Takeda, N. Ozaki, and N. Iwata, 2008, "Failure to Replicate the Association between Nrg1 and Schizophrenia Using Japanese Large Sample", *Schizophrenia Research,* Vol. 101, No. 1-3, pp. 1-8.

Ingason, A., K. Soeby, S. Timm, A. G. Wang, K. D. Jakobsen, A. Fink-Jensen, R. Hemmingsen, H. Berg Rasmussen, and T. Werge, 2006, "No Significant Association of the 5' End of Neuregulin 1 and Schizophrenia in a Large Danish Sample", *Schizophrenia Research,* Vol. 83, No. 1, pp. 1-5.

Iwata, N., T. Suzuki, M. Ikeda, T. Kitajima, Y. Yamanouchi, T. Inada, and N. Ozaki, 2004, "No Association with the Neuregulin 1 Haplotype to Japanese Schizophrenia", *Molecular Psychiatry,* Vol. 9, No. 2, pp. 126-127.

Jonsson, E. G., P. Saetre, M. Vares, D. Andreou, K. Larsson, S. Timm, H. B. Rasmussen, S. Djurovic, I. Melle, O. A. Andreassen, I. Agartz, T. Werge, H. Hall, and L. Terenius, 2009, "Dtnbp1, Nrg1, Daoa, Dao and Grm3 Polymorphisms and Schizophrenia: An Association Study", *Neuropsychobiology,* Vol. 59, No. 3, pp. 142-150.

Junttila, T. T., M. Sundvall, J. A. Maatta, and K. Elenius, 2000, "Erbb4 and Its Isoforms: Selective Regulation of Growth Factor Responses by Naturally Occurring Receptor Variants", *Trends in Cardiovascular Medicine,* Vol. 10, No. 7, pp. 304-310.

Kampman, O., S. Anttila, A. Illi, M. Saarela, R. Rontu, K. M. Mattila, E. Leinonen, and T. Lehtimaki, 2004, "Neuregulin Genotype and Medication Response in Finnish Patients with Schizophrenia", *NeuroReport,* Vol. 15, No. 16, pp. 2517-2520.

Kantrowitz, J. T., and D. C. Javitt, 2010, "N-Methyl-D-Aspartate (Nmda) Receptor Dysfunction or Dysregulation: The Final Common Pathway on the Road to Schizophrenia?", *Brain Research Bulletin,* Vol. 83, No. 3-4, pp. 108-121.

Karayiorgou, M., M. A. Morris, B. Morrow, R. J. Shprintzen, R. Goldberg, J. Borrow, A. Gos, G. Nestadt, P. S. Wolyniec, V. K. Lasseter, and et al., 1995, "Schizophrenia Susceptibility Associated with Interstitial Deletions of Chromosome 22q11", *Proceedings of the National Academy of Sciences,* Vol. 92, No. 17, pp. 7612-7616.

Karlsson, P., L. Farde, C. Halldin, and G. Sedvall, 2002, "Pet Study of D(1) Dopamine Receptor Binding in Neuroleptic-Naive Patients with Schizophrenia", *The American Journal of Psychiatry,* Vol. 159, No. 5, pp. 761-767.

Keri, S., I. Kiss, I. Seres, and O. Kelemen, 2009, "A Polymorphism of the Neuregulin 1 Gene (Snp8nrg243177/Rs6994992) Affects Reactivity to Expressed Emotion in Schizophrenia", *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics,* Vol. 150B, No. 3, pp. 418-420.

Keshavan, M. S., 1999, "Development, Disease and Degeneration in Schizophrenia: A Unitary Pathophysiological Model", *Journal of Psychiatric Research,* Vol. 33, No. 6, pp. 513-521.

Kessler, R. C., W. T. Chiu, O. Demler, K. R. Merikangas, and E. E. Walters, 2005, "Prevalence, Severity, and Comorbidity of 12-Month Dsm-Iv Disorders in the National Comorbidity Survey Replication", *Archives of General Psychiatry,* Vol. 62, No. 6, pp. 617-627.

Kim, J. W., Y. S. Lee, E. Y. Cho, Y. L. Jang, D. Y. Park, K. S. Choi, H. O. Jeun, S. H. Cho, S. Y. Jang, and K. S. Hong, 2006, "Linkage and Association of Schizophrenia with Genetic Variations in the Locus of Neuregulin 1 in Korean Population", *American*

*Journal of Medical Genetics Part B: Neuropsychiatric Genetics,* Vol. 141B, No. 3, pp. 281-286.

Kim, Y., S. Zerwas, S. E. Trace, and P. F. Sullivan, 2011, "Schizophrenia Genetics: Where Next?", *Schizophrenia Bulletin,* Vol. 37, No. 3, pp. 456-463.

Kirov, G., D. Grozeva, N. Norton, D. Ivanov, K. K. Mantripragada, P. Holmans, N. Craddock, M. J. Owen, and M. C. O'Donovan, 2009a, "Support for the Involvement of Large Copy Number Variants in the Pathogenesis of Schizophrenia", *Human Molecular Genetics,* Vol. 18, No. 8, pp. 1497-1503.

Kirov, G., D. Rujescu, A. Ingason, D. A. Collier, M. C. O'Donovan, and M. J. Owen, 2009b, "Neurexin 1 (Nrxn1) Deletions in Schizophrenia", *Schizophrenia Bulletin,* Vol. 35, No. 5, pp. 851-854.

Kirov, G., I. Zaharieva, L. Georgieva, V. Moskvina, I. Nikolov, S. Cichon, A. Hillmer, D. Toncheva, M. J. Owen, and M. C. O'Donovan, 2009c, "A Genome-Wide Association Study in 574 Schizophrenia Trios Using DNA Pooling", *Molecular Psychiatry,* Vol. 14, No. 8, pp. 796-803.

Krystal, J. H., A. Anand, and B. Moghaddam, 2002, "Effects of Nmda Receptor Antagonists: Implications for the Pathophysiology of Schizophrenia", *Archives of General Psychiatry,* Vol. 59, No. 7, pp. 663-664.

Lachman, H. M., E. Pedrosa, K. A. Nolan, M. Glass, K. Ye, and T. Saito, 2006, "Analysis of Polymorphisms in at-Rich Domains of Neuregulin 1 Gene in Schizophrenia", *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics,* Vol. 141B, No. 1, pp. 102-109.

Lander, E. S., and N. J. Schork, 1994, "Genetic Dissection of Complex Traits", *Science,* Vol. 265, No. 5181, pp. 2037-2048.

Law, A. J., J. E. Kleinman, D. R. Weinberger, and C. S. Weickert, 2007, "Disease-Associated Intronic Variants in the Erbb4 Gene Are Related to Altered Erbb4 Splice-Variant Expression in the Brain in Schizophrenia", *Human Molecular Genetics*, Vol. 16, No. 2, pp. 129-141.

Law, A. J., B. K. Lipska, C. S. Weickert, T. M. Hyde, R. E. Straub, R. Hashimoto, P. J. Harrison, J. E. Kleinman, and D. R. Weinberger, 2006, "Neuregulin 1 Transcripts Are Differentially Expressed in Schizophrenia and Regulated by 5' Snps Associated with the Disease", *Proceedings of the National Academy of Sciences*, Vol. 103, No. 17, pp. 6747-6752.

Lencz, T., T. V. Morgan, M. Athanasiou, B. Dain, C. R. Reed, J. M. Kane, R. Kucherlapati, and A. K. Malhotra, 2007, "Converging Evidence for a Pseudoautosomal Cytokine Receptor Gene Locus in Schizophrenia", *Molecular Psychiatry*, Vol. 12, No. 6, pp. 572-580.

Lewis, C. M., D. F. Levinson, L. H. Wise, L. E. DeLisi, R. E. Straub, I. Hovatta, N. M. Williams, S. G. Schwab, A. E. Pulver, S. V. Faraone, L. M. Brzustowicz, C. A. Kaufmann, D. L. Garver, H. M. Gurling, E. Lindholm, H. Coon, H. W. Moises, W. Byerley, S. H. Shaw, A. Mesen, R. Sherrington, F. A. O'Neill, D. Walsh, K. S. Kendler, J. Ekelund, T. Paunio, J. Lonnqvist, L. Peltonen, M. C. O'Donovan, M. J. Owen, D. B. Wildenauer, W. Maier, G. Nestadt, J. L. Blouin, S. E. Antonarakis, B. J. Mowry, J. M. Silverman, R. R. Crowe, C. R. Cloninger, M. T. Tsuang, D. Malaspina, J. M. Harkavy-Friedman, D. M. Svrakic, A. S. Bassett, J. Holcomb, G. Kalsi, A. McQuillin, J. Brynjolfson, T. Sigmundsson, H. Petursson, E. Jazin, T. Zoega, and T. Helgason, 2003, "Genome Scan Meta-Analysis of Schizophrenia and Bipolar Disorder, Part Ii: Schizophrenia", *The American Journal of Human Genetics*, Vol. 73, No. 1, pp. 34-48.

Lewis, D. A., and B. Moghaddam, 2006, "Cognitive Dysfunction in Schizophrenia: Convergence of Gamma-Aminobutyric Acid and Glutamate Alterations", *Archives of Neurology*, Vol. 63, No. 10, pp. 1372-1376.

Li, D., D. A. Collier, and L. He, 2006, "Meta-Analysis Shows Strong Positive Association of the Neuregulin 1 (Nrg1) Gene with Schizophrenia", *Human Molecular Genetics*, Vol. 15, No. 12, pp. 1995-2002.

Li, T., H. Stefansson, E. Gudfinnsson, G. Cai, X. Liu, R. M. Murray, V. Steinthorsdottir, D. Januel, V. G. Gudnadottir, H. Petursson, A. Ingason, J. R. Gulcher, K. Stefansson, and D. A. Collier, 2004, "Identification of a Novel Neuregulin 1 at-Risk Haplotype in Han Schizophrenia Chinese Patients, but No Association with the Icelandic/Scottish Risk Haplotype", *Molecular Psychiatry*, Vol. 9, No. 7, pp. 698-704.

Lichtenstein, P., B. H. Yip, C. Bjork, Y. Pawitan, T. D. Cannon, P. F. Sullivan, and C. M. Hultman, 2009, "Common Genetic Determinants of Schizophrenia and Bipolar Disorder in Swedish Families: A Population-Based Study", *Lancet*, Vol. 373, No. 9659, pp. 234-239.

Lovlie, R., J. O. Berle, E. Stordal, and V. M. Steen, 2001, "The Phospholipase C-Gamma1 Gene (Plcg1) and Lithium-Responsive Bipolar Disorder: Re-Examination of an Intronic Dinucleotide Repeat Polymorphism", *Psychiatric Genetics*, Vol. 11, No. 1, pp. 41-43.

Lu, C. L., Y. C. Wang, J. Y. Chen, I. C. Lai, and Y. J. Liou, 2010, "Support for the Involvement of the Erbb4 Gene in Schizophrenia: A Genetic Association Analysis", *Neuroscience Letters*, Vol. 481, No. 2, pp. 120-125.

Mah, S., M. R. Nelson, L. E. Delisi, R. H. Reneland, N. Markward, M. R. James, D. R. Nyholt, N. Hayward, H. Handoko, B. Mowry, S. Kammerer, and A. Braun, 2006,

"Identification of the Semaphorin Receptor Plxna2 as a Candidate for Susceptibility to Schizophrenia", *Molecular Psychiatry*, Vol. 11, No. 5, pp. 471-478.

Manolio, T. A., F. S. Collins, N. J. Cox, D. B. Goldstein, L. A. Hindorff, D. J. Hunter, M. I. McCarthy, E. M. Ramos, L. R. Cardon, A. Chakravarti, J. H. Cho, A. E. Guttmacher, A. Kong, L. Kruglyak, E. Mardis, C. N. Rotimi, M. Slatkin, D. Valle, A. S. Whittemore, M. Boehnke, A. G. Clark, E. E. Eichler, G. Gibson, J. L. Haines, T. F. Mackay, S. A. McCarroll, and P. M. Visscher, 2009, "Finding the Missing Heritability of Complex Diseases", *Nature*, Vol. 461, No. 7265, pp. 747-753.

McCarthy, M. I., G. R. Abecasis, L. R. Cardon, D. B. Goldstein, J. Little, J. P. Ioannidis, and J. N. Hirschhorn, 2008, "Genome-Wide Association Studies for Complex Traits: Consensus, Uncertainty and Challenges", *Nature Reviews Genetics*, Vol. 9, No. 5, pp. 356-369.

McClellan, J. M., E. Susser, and M. C. King, 2007, "Schizophrenia: A Common Disease Caused by Multiple Rare Alleles", *The British Journal of Psychiatry*, Vol. 190, pp. 194-199.

McGrath, J., S. Saha, J. Welham, O. El Saadi, C. MacCauley, and D. Chant, 2004, "A Systematic Review of the Incidence of Schizophrenia: The Distribution of Rates and the Influence of Sex, Urbanicity, Migrant Status and Methodology", *BMC Medicine*, Vol. 2, pp. 13.

McIntosh, A. M., T. W. Moorhead, D. Job, G. K. Lymer, S. Munoz Maniega, J. McKirdy, J. E. Sussmann, B. J. Baig, M. E. Bastin, D. Porteous, K. L. Evans, E. C. Johnstone, S. M. Lawrie, and J. Hall, 2008, "The Effects of a Neuregulin 1 Variant on White Matter Density and Integrity", *Molecular Psychiatry*, Vol. 13, No. 11, pp. 1054-1059.

Mirnics, K., F. A. Middleton, A. Marquez, D. A. Lewis, and P. Levitt, 2000, "Molecular Characterization of Schizophrenia Viewed by Microarray Analysis of Gene Expression in Prefrontal Cortex", *Neuron,* Vol. 28, No. 1, pp. 53-67.

Moncrieff, J., 2009, "A Critique of the Dopamine Hypothesis of Schizophrenia and Psychosis", *Harvard Review of Psychiatry,* Vol. 17, No. 3, pp. 214-225.

Munafo, M. R., A. S. Attwood, and J. Flint, 2008, "Neuregulin 1 Genotype and Schizophrenia", *Schizophrenia Bulletin,* Vol. 34, No. 1, pp. 9-12.

Munafo, M. R., D. L. Thiselton, T. G. Clark, and J. Flint, 2006, "Association of the Nrg1 Gene and Schizophrenia: A Meta-Analysis", *Molecular Psychiatry,* Vol. 11, No. 6, pp. 539-546.

Need, A. C., D. Ge, M. E. Weale, J. Maia, S. Feng, E. L. Heinzen, K. V. Shianna, W. Yoon, D. Kasperaviciute, M. Gennarelli, W. J. Strittmatter, C. Bonvicini, G. Rossi, K. Jayathilake, P. A. Cola, J. P. McEvoy, R. S. Keefe, E. M. Fisher, P. L. St Jean, I. Giegling, A. M. Hartmann, H. J. Moller, A. Ruppert, G. Fraser, C. Crombie, L. T. Middleton, D. St Clair, A. D. Roses, P. Muglia, C. Francks, D. Rujescu, H. Y. Meltzer, and D. B. Goldstein, 2009, "A Genome-Wide Investigation of Snps and Cnvs in Schizophrenia", *PLoS Genetics,* Vol. 5, No. 2, pp. e1000373.

Ni, C. Y., M. P. Murphy, T. E. Golde, and G. Carpenter, 2001, "Gamma -Secretase Cleavage and Nuclear Localization of Erbb-4 Receptor Tyrosine Kinase", *Science,* Vol. 294, No. 5549, pp. 2179-2181.

Nicodemus, K. K., A. J. Law, E. Radulescu, A. Luna, B. Kolachana, R. Vakkalanka, D. Rujescu, I. Giegling, R. E. Straub, K. McGee, B. Gold, M. Dean, P. Muglia, J. H. Callicott, H. Y. Tan, and D. R. Weinberger, 2010, "Biological Validation of Increased Schizophrenia Risk with Nrg1, Erbb4, and Akt1 Epistasis Via Functional

Neuroimaging in Healthy Controls", *Archives of General Psychiatry,* Vol. 67, No. 10, pp. 991-1001.

Nicodemus, K. K., A. Luna, R. Vakkalanka, T. Goldberg, M. Egan, R. E. Straub, and D. R. Weinberger, 2006, "Further Evidence for Association between Erbb4 and Schizophrenia and Influence on Cognitive Intermediate Phenotypes in Healthy Controls", *Molecular Psychiatry,* Vol. 11, No. 12, pp. 1062-1065.

Nieoullon, A., and A. Coquerel, 2003, "Dopamine: A Key Regulator to Adapt Action, Emotion, Motivation and Cognition", *Current Opinion in Neurology,* Vol. 16 Suppl 2, pp. S3-9.

North, B. V., P. C. Sham, J. Knight, E. R. Martin, and D. Curtis, 2006, "Investigation of the Ability of Haplotype Association and Logistic Regression to Identify Associated Susceptibility Loci", *Annals of Human Genetics,* Vol. 70, No. 6, pp. 893-906.

Northoff, G., H. Waters, I. Mooren, U. Schluter, S. Diekmann, P. Falkai, and B. Bogerts, 1999, "Cortical Sulcal Enlargement in Catatonic Schizophrenia: A Planimetric Ct Study", *Journal of Psychiatry Research,* Vol. 91, No. 1, pp. 45-54.

Norton, N., V. Moskvina, D. W. Morris, N. J. Bray, S. Zammit, N. M. Williams, H. J. Williams, A. C. Preece, S. Dwyer, J. C. Wilkinson, G. Spurlock, G. Kirov, P. Buckland, J. L. Waddington, M. Gill, A. P. Corvin, M. J. Owen, and M. C. O'Donovan, 2006, "Evidence That Interaction between Neuregulin 1 and Its Receptor Erbb4 Increases Susceptibility to Schizophrenia", *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics,* Vol. 141B, No. 1, pp. 96-101.

O'Donovan, M. C., N. Craddock, N. Norton, H. Williams, T. Peirce, V. Moskvina, I. Nikolov, M. Hamshere, L. Carroll, L. Georgieva, S. Dwyer, P. Holmans, J. L. Marchini, C. C. Spencer, B. Howie, H. T. Leung, A. M. Hartmann, H. J. Moller, D. W. Morris, Y. Shi, G. Feng, P. Hoffmann, P. Propping, C. Vasilescu, W. Maier, M. Rietschel, S. Zammit,

J. Schumacher, E. M. Quinn, T. G. Schulze, N. M. Williams, I. Giegling, N. Iwata, M. Ikeda, A. Darvasi, S. Shifman, L. He, J. Duan, A. R. Sanders, D. F. Levinson, P. V. Gejman, S. Cichon, M. M. Nothen, M. Gill, A. Corvin, D. Rujescu, G. Kirov, M. J. Owen, N. G. Buccola, B. J. Mowry, R. Freedman, F. Amin, D. W. Black, J. M. Silverman, W. F. Byerley, and C. R. Cloninger, 2008, "Identification of Loci Associated with Schizophrenia by Genome-Wide Association and Follow-Up", *Nature Genetics*, Vol. 40, No. 9, pp. 1053-1055.

Okubo, Y., T. Suhara, K. Suzuki, K. Kobayashi, O. Inoue, O. Terasaki, Y. Someya, T. Sassa, Y. Sudo, E. Matsushima, M. Iyo, Y. Tateno, and M. Toru, 1997, "Decreased Prefrontal Dopamine D1 Receptors in Schizophrenia Revealed by Pet", *Nature*, Vol. 385, No. 6617, pp. 634-636.

Olayioye, M. A., R. M. Neve, H. A. Lane, and N. E. Hynes, 2000, "The Erbb Signaling Network: Receptor Heterodimerization in Development and Cancer", *The European Molecular Biology Organization Journal*, Vol. 19, No. 13, pp. 3159-3167.

Owen, M. J., N. Craddock, and M. C. O'Donovan, 2005, "Schizophrenia: Genes at Last?", *Trends in Genetics*, Vol. 21, No. 9, pp. 518-525.

Palmer, B. A., V. S. Pankratz, and J. M. Bostwick, 2005, "The Lifetime Risk of Suicide in Schizophrenia: A Reexamination", *Archives of General Psychiatry*, Vol. 62, No. 3, pp. 247-253.

Pe'er, I., R. Yelensky, D. Altshuler, and M. J. Daly, 2008, "Estimation of the Multiple Testing Burden for Genomewide Association Studies of Nearly All Common Variants", *Genetic Epidemiology*, Vol. 32, No. 4, pp. 381-385.

Pedersen, C. B., and P. B. Mortensen, 2001, "Evidence of a Dose-Response Relationship between Urbanicity During Upbringing and Schizophrenia Risk", *Archives of General Psychiatry*, Vol. 58, No. 11, pp. 1039-1046.

Petryshen, T. L., F. A. Middleton, A. Kirby, K. A. Aldinger, S. Purcell, A. R. Tahl, C. P. Morley, L. McGann, K. L. Gentile, G. N. Rockwell, H. M. Medeiros, C. Carvalho, A. Macedo, A. Dourado, J. Valente, C. P. Ferreira, N. J. Patterson, M. H. Azevedo, M. J. Daly, C. N. Pato, M. T. Pato, and P. Sklar, 2005, "Support for Involvement of Neuregulin 1 in Schizophrenia Pathophysiology", *Molecular Psychiatry,* Vol. 10, No. 4, pp. 366-374, 328.

Picchioni, M. M., and R. M. Murray, 2007, "Schizophrenia", *British Medical Journal,* Vol. 335, No. 7610, pp. 91-95.

Pitcher, G. M., L. V. Kalia, D. Ng, N. M. Goodfellow, K. T. Yee, E. K. Lambe, and M. W. Salter, 2011, "Schizophrenia Susceptibility Pathway Neuregulin 1-Erbb4 Suppresses Src Upregulation of Nmda Receptors", *Nature Medicine,* Vol. 17, No. 4, pp. 470-478.

Prabakaran, S., M. Wengenroth, H. E. Lockstone, K. Lilley, F. M. Leweke, and S. Bahn, 2007, "2-D Dige Analysis of Liver and Red Blood Cells Provides Further Evidence for Oxidative Stress in Schizophrenia", *Journal of Proteome Research,* Vol. 6, No. 1, pp. 141-149.

Prasad, S., P. Semwal, S. Deshpande, T. Bhatia, V. L. Nimgaonkar, and B. K. Thelma, 2002, "Molecular Genetics of Schizophrenia: Past, Present and Future", *Journal of Biosciences,* Vol. 27, No. 1 Suppl 1, pp. 35-52.

Purcell, S. M., N. R. Wray, J. L. Stone, P. M. Visscher, M. C. O'Donovan, P. F. Sullivan, and P. Sklar, 2009, "Common Polygenic Variation Contributes to Risk of Schizophrenia and Bipolar Disorder", *Nature,* Vol. 460, No. 7256, pp. 748-752.

Rapoport, J. L., A. M. Addington, S. Frangou, and M. R. Psych, 2005, "The Neurodevelopmental Model of Schizophrenia: Update 2005", *Molecular Psychiatry,* Vol. 10, No. 5, pp. 434-449.

Rethelyi, J. M., S. C. Bakker, P. Polgar, P. Czobor, E. Strengman, P. I. Pasztor, R. S. Kahn, and I. Bitter, 2010, "Association Study of Nrg1, Dtnbp1, Rgs4, G72/G30, and Pip5k2a with Schizophrenia and Symptom Severity in a Hungarian Sample", *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, Vol. 153B, No. 3, pp. 792-801.

Riecher-Rossler, A., and H. Hafner, 2000, "Gender Aspects in Schizophrenia: Bridging the Border between Social and Biological Psychiatry", *Acta Psychiatrica Scandinavica*, Vol. No. 407, pp. 58-62.

Riecher-Rossler, A., and M. V. Seeman, 2002, "Oestrogens and Schizophrenia--Introduction", *Archives of Womens' Mental Health*, Vol. 5, No. 3, pp. 91-92.

Rietschel, M., M. Mattheisen, F. Degenhardt, R. S. Kahn, D. H. Linszen, J. V. Os, D. Wiersma, R. Bruggeman, W. Cahn, L. de Haan, L. Krabbendam, I. Myin-Germeys, T. W. Muhleisen, P. Kirsch, C. Esslinger, S. Herms, D. Demontis, M. Steffens, J. Strohmaier, B. Haenisch, R. Breuer, P. M. Czerski, I. Giegling, E. Strengman, C. Schmael, O. Mors, P. B. Mortensen, D. M. Hougaard, T. Orntoft, P. Kapelski, L. Priebe, F. B. Basmanav, A. J. Forstner, P. Hoffmann, S. Meier, J. Nikitopoulos, S. Moebus, M. Alexander, R. Mossner, H. E. Wichmann, S. Schreiber, F. Rivandeneira, A. Hofman, A. G. Uitterlinden, T. F. Wienker, J. Schumacher, J. Hauser, W. Maier, R. M. Cantor, S. Erk, T. G. Schulze, H. Stefansson, S. Steinberg, O. Gustafsson, E. Sigurdsson, H. Petursson, A. Kong, K. Stefansson, O. P. Pietilainen, A. Tuulio-Henriksson, T. Paunio, J. Lonnqvist, J. Suvisaari, L. Peltonen, M. Ruggeri, S. Tosato, M. Walshe, R. Murray, D. A. Collier, D. S. Clair, T. Hansen, A. Ingason, K. D. Jakobsen, L. Duong, T. Werge, I. Melle, O. A. Andreassen, S. Djurovic, I. Bitter, J. M. Rethelyi, L. Abramova, V. Kaleda, V. Golimbet, E. G. Jonsson, L. Terenius, I. Agartz, R. V. Winkel, G. Kenis, M. D. Hert, J. Veldink, C. Wiuf, M. Didriksen, N. Craddock, M. J. Owen, M. C. O'Donovan, A. D. Borglum, D. Rujescu, H. Walter, A. Meyer-Lindenberg, M. M. Nothen, et al., 2011, "Association between Genetic Variation in a

Region on Chromosome 11 and Schizophrenia in Large Samples from Europe", *Molecular Psychiatry*.

Ripke, S., A. R. Sanders, K. S. Kendler, D. F. Levinson, P. Sklar, P. A. Holmans, D. Y. Lin, J. Duan, R. A. Ophoff, O. A. Andreassen, E. Scolnick, S. Cichon, D. St Clair, A. Corvin, H. Gurling, T. Werge, D. Rujescu, D. H. Blackwood, C. N. Pato, A. K. Malhotra, S. Purcell, F. Dudbridge, B. M. Neale, L. Rossin, P. M. Visscher, D. Posthuma, D. M. Ruderfer, A. Fanous, H. Stefansson, S. Steinberg, B. J. Mowry, V. Golimbet, M. De Hert, E. G. Jonsson, I. Bitter, O. P. Pietilainen, D. A. Collier, S. Tosato, I. Agartz, M. Albus, M. Alexander, R. L. Amdur, F. Amin, N. Bass, S. E. Bergen, D. W. Black, A. D. Borglum, M. A. Brown, R. Bruggeman, N. G. Buccola, W. F. Byerley, W. Cahn, R. M. Cantor, V. J. Carr, S. V. Catts, K. Choudhury, C. R. Cloninger, P. Cormican, N. Craddock, P. A. Danoy, S. Datta, L. de Haan, D. Demontis, D. Dikeos, S. Djurovic, P. Donnelly, G. Donohoe, L. Duong, S. Dwyer, A. Fink-Jensen, R. Freedman, N. B. Freimer, M. Friedl, L. Georgieva, I. Giegling, M. Gill, B. Glenthoj, S. Godard, M. Hamshere, M. Hansen, T. Hansen, A. M. Hartmann, F. A. Henskens, D. M. Hougaard, C. M. Hultman, A. Ingason, A. V. Jablensky, K. D. Jakobsen, M. Jay, G. Jurgens, R. S. Kahn, M. C. Keller, G. Kenis, E. Kenny, Y. Kim, G. K. Kirov, H. Konnerth, B. Konte, L. Krabbendam, R. Krasucki, et al., 2011, "Genome-Wide Association Study Identifies Five New Schizophrenia Loci", *Nature Genetics*, Vol. 43, No. 10, pp. 969-976.

Rodriguez-Murillo, L., J. A. Gogos, and M. Karayiorgou, 2012, "The Genetic Architecture of Schizophrenia: New Mutations and Emerging Paradigms", *Annual Review of Medicine*, Vol. 63, pp. 63-80.

Ruano, D., Y. S. Aulchenko, A. Macedo, M. J. Soares, J. Valente, M. H. Azevedo, M. H. Hutz, C. S. Gama, M. I. Lobato, P. Belmonte-de-Abreu, A. B. Goodman, C. Pato, P. Heutink, and J. A. Palha, 2008, "Association of the Gene Encoding Neurogranin with Schizophrenia in Males", *Journal of Psychiatric Research*, Vol. 42, No. 2, pp. 125-133.

Schulze, T. G., N. Akula, R. Breuer, J. Steele, M. A. Nalls, A. B. Singleton, F. A. Degenhardt, M. M. Nothen, S. Cichon, M. Rietschel, and F. J. McMahon, 2012, "Molecular Genetic Overlap in Bipolar Disorder, Schizophrenia, and Major Depressive Disorder", *World Journal of Biological Psychiatry*.

Sebat, J., D. L. Levy, and S. E. McCarthy, 2009, "Rare Structural Variants in Schizophrenia: One Disorder, Multiple Mutations; One Mutation, Multiple Disorders", *Trends in Genetics*, Vol. 25, No. 12, pp. 528-535.

Selten, J. P., E. Cantor-Graae, and R. S. Kahn, 2007, "Migration and Schizophrenia", *Current Opinion in Psychiatry*, Vol. 20, No. 2, pp. 111-115.

Shi, J., D. F. Levinson, J. Duan, A. R. Sanders, Y. Zheng, I. Pe'er, F. Dudbridge, P. A. Holmans, A. S. Whittemore, B. J. Mowry, A. Olincy, F. Amin, C. R. Cloninger, J. M. Silverman, N. G. Buccola, W. F. Byerley, D. W. Black, R. R. Crowe, J. R. Oksenberg, D. B. Mirel, K. S. Kendler, R. Freedman, and P. V. Gejman, 2009, "Common Variants on Chromosome 6p22.1 Are Associated with Schizophrenia", *Nature*, Vol. 460, No. 7256, pp. 753-757.

Shi, Y., Z. Li, Q. Xu, T. Wang, T. Li, J. Shen, F. Zhang, J. Chen, G. Zhou, W. Ji, B. Li, Y. Xu, D. Liu, P. Wang, P. Yang, B. Liu, W. Sun, C. Wan, S. Qin, G. He, S. Steinberg, S. Cichon, T. Werge, E. Sigurdsson, S. Tosato, A. Palotie, M. M. Nothen, M. Rietschel, R. A. Ophoff, D. A. Collier, D. Rujescu, D. S. Clair, H. Stefansson, K. Stefansson, J. Ji, Q. Wang, W. Li, L. Zheng, H. Zhang, G. Feng, and L. He, 2011, "Common Variants on 8p12 and 1q24.2 Confer Risk of Schizophrenia", *Nature Genetics*, Vol. 43, No. 12, pp. 1224-1227.

Shifman, S., M. Johannesson, M. Bronstein, S. X. Chen, D. A. Collier, N. J. Craddock, K. S. Kendler, T. Li, M. O'Donovan, F. A. O'Neill, M. J. Owen, D. Walsh, D. R. Weinberger, C. Sun, J. Flint, and A. Darvasi, 2008, "Genome-Wide Association

Identifies a Common Variant in the Reelin Gene That Increases the Risk of Schizophrenia Only in Women", *PLoS Genetics*, Vol. 4, No. 2, pp. e28.

Shiota, S., M. Tochigi, H. Shimada, J. Ohashi, K. Kasai, N. Kato, K. Tokunaga, and T. Sasaki, 2008, "Association and Interaction Analyses of Nrg1 and Erbb4 Genes with Schizophrenia in a Japanese Population", *Journal of Human Genetics*, Vol. 53, No. 10, pp. 929-935.

Silberberg, G., A. Darvasi, R. Pinkas-Kramarski, and R. Navon, 2006, "The Involvement of Erbb4 with Schizophrenia: Association and Expression Studies", *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, Vol. 141B, No. 2, pp. 142-148.

Singh, S. M., B. Murphy, and R. O'Reilly, 2002, "Epigenetic Contributors to the Discordance of Monozygotic Twins", *Clinical Genetics*, Vol. 62, No. 2, pp. 97-103.

Sipos, A., F. Rasmussen, G. Harrison, P. Tynelius, G. Lewis, D. A. Leon, and D. Gunnell, 2004, "Paternal Age and Schizophrenia: A Population Based Cohort Study", *British Medical Journal*, Vol. 329, No. 7474, pp. 1070.

Smith, G. S., R. Schloesser, J. D. Brodie, S. L. Dewey, J. Logan, S. A. Vitkun, P. Simkowitz, A. Hurley, T. Cooper, N. D. Volkow, and R. Cancro, 1998, "Glutamate Modulation of Dopamine Measured in Vivo with Positron Emission Tomography (Pet) and 11c-Raclopride in Normal Human Subjects", *Neuropsychopharmacology*, Vol. 18, No. 1, pp. 18-25.

Squassina, A., P. Piccardi, M. Del Zompo, A. Rossi, A. Vita, S. Pini, A. Mucci, and S. Galderisi, 2010, "Nrg1 and Bdnf Genes in Schizophrenia: An Association Study in an Italian Case-Control Sample", *Journal of Psychiatry Research*, Vol. 176, No. 1, pp. 82-84.

Stefansson, H., R. A. Ophoff, S. Steinberg, O. A. Andreassen, S. Cichon, D. Rujescu, T. Werge, O. P. Pietilainen, O. Mors, P. B. Mortensen, E. Sigurdsson, O. Gustafsson, M. Nyegaard, A. Tuulio-Henriksson, A. Ingason, T. Hansen, J. Suvisaari, J. Lonnqvist, T. Paunio, A. D. Borglum, A. Hartmann, A. Fink-Jensen, M. Nordentoft, D. Hougaard, B. Norgaard-Pedersen, Y. Bottcher, J. Olesen, R. Breuer, H. J. Moller, I. Giegling, H. B. Rasmussen, S. Timm, M. Mattheisen, I. Bitter, J. M. Rethelyi, B. B. Magnusdottir, T. Sigmundsson, P. Olason, G. Masson, J. R. Gulcher, M. Haraldsson, R. Fossdal, T. E. Thorgeirsson, U. Thorsteinsdottir, M. Ruggeri, S. Tosato, B. Franke, E. Strengman, L. A. Kiemeney, I. Melle, S. Djurovic, L. Abramova, V. Kaleda, J. Sanjuan, R. de Frutos, E. Bramon, E. Vassos, G. Fraser, U. Ettinger, M. Picchioni, N. Walker, T. Toulopoulou, A. C. Need, D. Ge, J. L. Yoon, K. V. Shianna, N. B. Freimer, R. M. Cantor, R. Murray, A. Kong, V. Golimbet, A. Carracedo, C. Arango, J. Costas, E. G. Jonsson, L. Terenius, I. Agartz, H. Petursson, M. M. Nothen, M. Rietschel, P. M. Matthews, P. Muglia, L. Peltonen, D. St Clair, D. B. Goldstein, K. Stefansson, and D. A. Collier, 2009, "Common Variants Conferring Risk of Schizophrenia", *Nature*, Vol. 460, No. 7256, pp. 744-747.

Stefansson, H., J. Sarginson, A. Kong, P. Yates, V. Steinthorsdottir, E. Gudfinnsson, S. Gunnarsdottir, N. Walker, H. Petursson, C. Crombie, A. Ingason, J. R. Gulcher, K. Stefansson, and D. St Clair, 2003, "Association of Neuregulin 1 with Schizophrenia Confirmed in a Scottish Population", *The American Journal of Human Genetics*, Vol. 72, No. 1, pp. 83-87.

Stefansson, H., E. Sigurdsson, V. Steinthorsdottir, S. Bjornsdottir, T. Sigmundsson, S. Ghosh, J. Brynjolfsson, S. Gunnarsdottir, O. Ivarsson, T. T. Chou, O. Hjaltason, B. Birgisdottir, H. Jonsson, V. G. Gudnadottir, E. Gudmundsdottir, A. Bjornsson, B. Ingvarsson, A. Ingason, S. Sigfusson, H. Hardardottir, R. P. Harvey, D. Lai, M. Zhou, D. Brunner, V. Mutel, A. Gonzalo, G. Lemke, J. Sainz, G. Johannesson, T. Andresson, D. Gudbjartsson, A. Manolescu, M. L. Frigge, M. E. Gurney, A. Kong, J. R. Gulcher, H. Petursson, and K. Stefansson, 2002, "Neuregulin 1 and Susceptibility to

Schizophrenia", *The American Journal of Human Genetics,* Vol. 71, No. 4, pp. 877-892.

Sullivan, P. F., K. S. Kendler, and M. C. Neale, 2003, "Schizophrenia as a Complex Trait: Evidence from a Meta-Analysis of Twin Studies", *Archives of General Psychiatry,* Vol. 60, No. 12, pp. 1187-1192.

Sullivan, P. F., D. Lin, J. Y. Tzeng, E. van den Oord, D. Perkins, T. S. Stroup, M. Wagner, S. Lee, F. A. Wright, F. Zou, W. Liu, A. M. Downing, J. Lieberman, and S. L. Close, 2008, "Genomewide Association for Schizophrenia in the Catie Study: Results of Stage 1", *Molecular Psychiatry,* Vol. 13, No. 6, pp. 570-584.

Susser, E., R. Neugebauer, H. W. Hoek, A. S. Brown, S. Lin, D. Labovitz, and J. M. Gorman, 1996, "Schizophrenia after Prenatal Famine. Further Evidence", *Archives of General Psychiatry,* Vol. 53, No. 1, pp. 25-31.

Tabares-Seisdedos, R., and J. L. Rubenstein, 2009, "Chromosome 8p as a Potential Hub for Developmental Neuropsychiatric Disorders: Implications for Schizophrenia, Autism and Cancer", *Molecular Psychiatry,* Vol. 14, No. 6, pp. 563-589.

Takei, N., P. B. Mortensen, U. Klaening, R. M. Murray, P. C. Sham, E. O'Callaghan, and P. Munk-Jorgensen, 1996, "Relationship between in Utero Exposure to Influenza Epidemics and Risk of Schizophrenia in Denmark", *Biological Psychiatry,* Vol. 40, No. 9, pp. 817-824.

Tamminga, C. A., 2006, "The Neurobiology of Cognition in Schizophrenia", *Journal of Clinical Psychiatry,* Vol. 67 Suppl 9, No. pp. 9-13.

Tandon, R., M. S. Keshavan, and H. A. Nasrallah, 2008, "Schizophrenia, "Just the Facts" What We Know in 2008. 2. Epidemiology and Etiology", *Schizophrenia Research,* Vol. 102, No. 1-3, pp. 1-18.

Tang, J. X., W. Y. Chen, G. He, J. Zhou, N. F. Gu, G. Y. Feng, and L. He, 2004, "Polymorphisms within 5' End of the Neuregulin 1 Gene Are Genetically Associated with Schizophrenia in the Chinese Population", *Molecular Psychiatry,* Vol. 9, No. 1, pp. 11-12.

Taylor, J. G., 2011, "A Neural Model of the Loss of Self in Schizophrenia", *Schizophrenia Bulletin,* Vol. 37, No. 6, pp. 1229-1247.

Thiselton, D. L., B. T. Webb, B. M. Neale, R. C. Ribble, F. A. O'Neill, D. Walsh, B. P. Riley, and K. S. Kendler, 2004, "No Evidence for Linkage or Association of Neuregulin-1 (Nrg1) with Disease in the Irish Study of High-Density Schizophrenia Families (Ishdsf)", *Molecular Psychiatry,* Vol. 9, No. 8, pp. 777-783.

Thomson, P. A., A. Christoforou, S. W. Morris, E. Adie, B. S. Pickard, D. J. Porteous, W. J. Muir, D. H. Blackwood, and K. L. Evans, 2007, "Association of Neuregulin 1 with Schizophrenia and Bipolar Disorder in a Second Cohort from the Scottish Population", *Molecular Psychiatry,* Vol. 12, No. 1, pp. 94-104.

Tkachev, D., M. L. Mimmack, M. M. Ryan, M. Wayland, T. Freeman, P. B. Jones, M. Starkey, M. J. Webster, R. H. Yolken, and S. Bahn, 2003, "Oligodendrocyte Dysfunction in Schizophrenia and Bipolar Disorder", *Lancet,* Vol. 362, No. 9386, pp. 798-805.

Tosato, S., M. Bellani, C. Bonetto, M. Ruggeri, C. Perlini, A. Lasalvia, V. Marinelli, G. Rambaldelli, D. Cristofalo, M. Bertani, M. Zanoni, L. Lazzarotto, R. Cerini, R. Pozzi Mucelli, M. Tansella, P. Dazzan, M. Di Forti, R. M. Murray, D. A. Collier, and P. Brambilla, 2012, "Is Neuregulin 1 Involved in Determining Cerebral Volumes in Schizophrenia? Preliminary Results Showing a Decrease in Superior Temporal Gyrus Volume", *Neuropsychobiology,* Vol. 65, No. 3, pp. 119-125.

Turunen, J. A., J. O. Peltonen, O. P. Pietilainen, W. Hennah, A. Loukola, T. Paunio, K. Silander, J. Ekelund, T. Varilo, T. Partonen, J. Lonnqvist, and L. Peltonen, 2007, "The Role of Dtnbp1, Nrg1, and Akt1 in the Genetics of Schizophrenia in Finland", *Schizophrenia Research,* Vol. 91, No. 1-3, pp. 27-36.

van Os, J., and S. Kapur, 2009, "Schizophrenia", *Lancet,* Vol. 374, No. 9690, pp. 635-645.

van Rossum, J. M., 1966, "The Significance of Dopamine-Receptor Blockade for the Mechanism of Action of Neuroleptic Drugs", *Archives Internationales de Pharmacodynamie et de Thérapie*, Vol. 160, No. 2, pp. 492-494.

Walker, R. M., A. Christoforou, P. A. Thomson, K. A. McGhee, A. Maclean, T. W. Muhleisen, J. Strohmaier, V. Nieratschker, M. M. Nothen, M. Rietschel, S. Cichon, S. W. Morris, O. Jilani, D. Stclair, D. H. Blackwood, W. J. Muir, D. J. Porteous, and K. L. Evans, 2010, "Association Analysis of Neuregulin 1 Candidate Regions in Schizophrenia and Bipolar Disorder", *Neuroscience Letters,* Vol. 478, No. 1, pp. 9-13.

Walsh, T., J. M. McClellan, S. E. McCarthy, A. M. Addington, S. B. Pierce, G. M. Cooper, A. S. Nord, M. Kusenda, D. Malhotra, A. Bhandari, S. M. Stray, C. F. Rippey, P. Roccanova, V. Makarov, B. Lakshmi, R. L. Findling, L. Sikich, T. Stromberg, B. Merriman, N. Gogtay, P. Butler, K. Eckstrand, L. Noory, P. Gochman, R. Long, Z. Chen, S. Davis, C. Baker, E. E. Eichler, P. S. Meltzer, S. F. Nelson, A. B. Singleton, M. K. Lee, J. L. Rapoport, M. C. King, and J. Sebat, 2008, "Rare Structural Variants Disrupt Multiple Genes in Neurodevelopmental Pathways in Schizophrenia", *Science*, Vol. 320, No. 5875, pp. 539-543.

Walss-Bass, C., W. Liu, D. F. Lew, R. Villegas, P. Montero, A. Dassori, R. J. Leach, L. Almasy, M. Escamilla, and H. Raventos, 2006, "A Novel Missense Mutation in the Transmembrane Domain of Neuregulin 1 Is Associated with Schizophrenia", *Biological Psychiatry,* Vol. 60, No. 6, pp. 548-553.

Wang, L., H. Oota, N. Saitou, F. Jin, T. Matsushita, and S. Ueda, 2000, "Genetic Structure of a 2,500-Year-Old Human Population in China and Its Spatiotemporal Changes", *Molecular Biology and Evolution,* Vol. 17, No. 9, pp. 1396-1400.

Wang, W. Y., B. J. Barratt, D. G. Clayton, and J. A. Todd, 2005, "Genome-Wide Association Studies: Theoretical and Practical Concerns", *Nature Reviews Genetics,* Vol. 6, No. 2, pp. 109-118.

Wei, J., and G. P. Hemmings, 2000, "The Notch4 Locus Is Associated with Susceptibility to Schizophrenia", *Nature Genetics,* Vol. 25, No. 4, pp. 376-377.

Weickert, C. S., A. L. Miranda-Angulo, J. Wong, W. R. Perlman, S. E. Ward, V. Radhakrishna, R. E. Straub, D. R. Weinberger, and J. E. Kleinman, 2008, "Variants in the Estrogen Receptor Alpha Gene and Its Mrna Contribute to Risk for Schizophrenia", *Human Molecular Genetics,* Vol. 17, No. 15, pp. 2293-2309.

Weiser, M., and S. Noy, 2005, "Interpreting the Association between Cannabis Use and Increased Risk for Schizophrenia", *Dialogues in Clinical Neuroscience,* Vol. 7, No. 1, pp. 81-85.

Williams, N. M., A. Preece, G. Spurlock, N. Norton, H. J. Williams, S. Zammit, M. C. O'Donovan, and M. J. Owen, 2003, "Support for Genetic Variation in Neuregulin 1 and Susceptibility to Schizophrenia", *Molecular Psychiatry,* Vol. 8, No. 5, pp. 485-487.

Wong, J., and C. S. Weickert, 2009, "Transcriptional Interaction of an Estrogen Receptor Splice Variant and Erbb4 Suggests Convergence in Gene Susceptibility Pathways in Schizophrenia", *The Journal of Biological Chemistry,* Vol. 284, No. 28, pp. 18824-18832.

Wright, I. C., S. Rabe-Hesketh, P. W. Woodruff, A. S. David, R. M. Murray, and E. T. Bullmore, 2000, "Meta-Analysis of Regional Brain Volumes in Schizophrenia", *The American Journal of Psychiatry*, Vol. 157, No. 1, pp. 16-25.

Xu, B., J. L. Roos, P. Dexheimer, B. Boone, B. Plummer, S. Levy, J. A. Gogos, and M. Karayiorgou, 2011, "Exome Sequencing Supports a De Novo Mutational Paradigm for Schizophrenia", *Nature Genetics*, Vol. 43, No. 9, pp. 864-868.

Xu, B., J. L. Roos, S. Levy, E. J. van Rensburg, J. A. Gogos, and M. Karayiorgou, 2008, "Strong Association of De Novo Copy Number Mutations with Sporadic Schizophrenia", *Nature Genetics*, Vol. 40, No. 7, pp. 880-885.

Yamada, K., Y. Iwayama, E. Hattori, K. Iwamoto, T. Toyota, T. Ohnishi, H. Ohba, M. Maekawa, T. Kato, and T. Yoshikawa, 2011, "Genome-Wide Association Study of Schizophrenia in Japanese Population", *PLoS One*, Vol. 6, No. 6, pp. e20468.

Yang, J. Z., T. M. Si, Y. Ruan, Y. S. Ling, Y. H. Han, X. L. Wang, M. Zhou, H. Y. Zhang, Q. M. Kong, C. Liu, D. R. Zhang, Y. Q. Yu, S. Z. Liu, G. Z. Ju, L. Shu, D. L. Ma, and D. Zhang, 2003, "Association Study of Neuregulin 1 Gene with Schizophrenia", *Molecular Psychiatry*, Vol. 8, No. 7, pp. 706-709.

Yue, W. H., H. F. Wang, L. D. Sun, F. L. Tang, Z. H. Liu, H. X. Zhang, W. Q. Li, Y. L. Zhang, Y. Zhang, C. C. Ma, B. Du, L. F. Wang, Y. Q. Ren, Y. F. Yang, X. F. Hu, Y. Wang, W. Deng, L. W. Tan, Y. L. Tan, Q. Chen, G. M. Xu, G. G. Yang, X. B. Zuo, H. Yan, Y. Y. Ruan, T. L. Lu, X. Han, X. H. Ma, L. W. Cai, C. Jin, H. Y. Zhang, J. Yan, W. F. Mi, X. Y. Yin, W. B. Ma, Q. Liu, L. Kang, W. Sun, C. Y. Pan, M. Shuang, F. D. Yang, C. Y. Wang, J. L. Yang, K. Q. Li, X. Ma, L. J. Li, X. Yu, Q. Z. Li, X. Huang, L. X. Lv, T. Li, G. P. Zhao, W. Huang, X. J. Zhang, and D. Zhang, 2011, "Genome-Wide Association Study Identifies a Susceptibility Locus for Schizophrenia in Han Chinese at 11p11.2", *Nature Genetics*, Vol. 43, No. 12, pp. 1228-1231.

Zheng, C., Y. Shen, and Q. Xu, 2012, "Association of Intron 1 Variants of the Dopamine Transporter Gene with Schizophrenia", *Neuroscience Letters*.