TARGETED NEXT GENERATION SEQUENCING OF *ZNF804A*, *DTNBP1*, *AKT1* AND *TCF4* TO IDENTIFY RISK VARIANTS IN SCHIZOPHRENIA MULTIPLEX FAMILIES

CHOW TZE JEN

DOCTOR OF PHILOSOPHY IN SCIENCE

LEE KONG CHIAN
FACULTY OF ENGINEERING AND SCIENCE
UNIVERSITI TUNKU ABDUL RAHMAN
APRIL 2018

TARGETED NEXT GENERATION SEQUENCING OF *ZNF804A*, *DTNBP1*, *AKT1* AND *TCF4* TO IDENTIFY RISK VARIANTS IN SCHIZOPHRENIA MULTIPLEX FAMILIES

By

CHOW TZE JEN

A dissertation submitted to the Department of
Mechatronics and Biomedical Engineering,
Lee Kong Chian Faculty of Engineering and Science,
Universiti Tunku Abdul Rahman,
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Science
April 2018

To my parents,

my eternal cheerleaders

May this be a blessing for all who are affected by schizophrenia

be it directly or indirectly

ABSTRACT

TARGETED NEXT GENERATION SEQUENCING OF *ZNF804A*, *DTNBP1*, *AKT1* AND *TCF4* TO IDENTIFY RISK VARIANTS IN SCHIZOPHRENIA MULTIPLEX FAMILIES

Chow Tze Jen

Cognitive impairment is one of the core features of schizophrenia and is extensively associated to genetic influences. First-degree relatives of patients with schizophrenia often showed similar cognitive deficit as seen in patients. Until now, most of the case-control and family-based neurocognition studies have reported on candidate variants in the intron regions. Therefore, the current study utilised a multiplex family-based case-control sample cohort to investigate the association of functional regions of *ZNF804A*, *DTNBP1*, *AKT1*, and *TCF4* with schizophrenia and cognitive deficits. This study involved 12 multiplex families (n = 26), comprising of 12 probands and their affected first-degree relatives, together with 11 control families (n = 22). Cognitive ability of participants was assessed using Trail Making Test (TMT). Variants were identified using Taqman® Real-time PCR and Ion Torrent AmpliSeq™ sequencing. On average, the cognitive performances of the patients fell below normal standard. Significantly poorer scores in patients were found in TMT-A (p = 4.08 x 10^{-5}), TMT-B (p = 1.96 x 10^{-4}), and B-A (p = 1.68 x 10^{-3}). A total of 117 variants were detected. For *ZNF804A*, a haploblock was identified in the 5' UTR. Together with exonic rs3731834, 5' UTR of *ZNF804A* may be associated with schizophrenia. In *DTNBP1*, 5' UTR variant rs11558324 was

only found in patients and predicted to regulate protein coding of exon 1, while 3' UTR rs1047631 was predicted to inhibit mRNA expression. For *AKT1*, the 3' UTR spanning seven variants (739 bp) was shown to be highly associated to schizophrenia. They were only found in patients, in high LD (D' = 1.00), and were predicted to inhibit gene expression. In the 3' UTR region of *TCF4*, risk allele of rs369444573 was found higher in patients and was predicted to promote *TCF4* expression. In conclusion, the genetic pathways that underlie cognitive deficits in schizophrenia may lie within the 5' and 3' regulatory regions of *ZNF804A*, *DTNBP1*, *AKT1*, and *TCF4*.

ACKNOWLEDGEMENTS

This work would not have been possible without the invaluable contributions and support from numerous people. First and foremost, I wish to express heartfelt gratitude to my supervisor, Dr. Tang Pek Yee and co-supervisor, Dr. Tee Shiau Foon, for sharing their knowledge and insights, and whose tutelage has guided me through all aspects of my PhD journey.

I am especially grateful to Kuok Foundation Berhad for the PhD scholarship. Because of it, I was able to realise many dreams and tick off bucket lists. I am forever indebted to Prof. Emeritus Yong Hoi Sen, who has made available his support in so many ways. I thank Dr Loh Siew Yim and staffs from Hospital Permai Johor Bahru for their contributions in sample collection. Special thanks to Dr Song Sze Looi for her patience to a NGS newbie. I also wish to acknowledge funding sources for this work: Ministry of Higher Education (HIR-H-50001-00-A000025 and UM.C/62SH/HIR/MOHE/MED/26), UTAR (IPSR/RMC/UTARRF/2014C1/T07) and Malaysia Toray Science Foundation.

Palms together to my guru Ven. Khen Rinpoche Jei Samten Drodul Dorjee, whose wisdom and compassion inspires me to persevere during the hardest moments. Not forgetting a big hug to my friends Hui Lane and Yong Hui for their encouragement and camaraderie. *Fortitudine Vincimus!*

Above all, I owe my deepest gratitude to my family for their love, patience, and endless support. A deep bow and thank you!

Tze Jen
August 2017

APPROVAL SHEET

This dissertation entitled "**TARGETED NEXT GENERATION SEQUENCING OF *ZNF804A*, *DTNBP1*, *AKT1* AND *TCF4* TO IDENTIFY RISK VARIANTS IN SCHIZOPHRENIA MULTIPLEX FAMILIES**" was prepared by CHOW TZE JEN and submitted as partial fulfillment of the requirements for the degree of Doctor of Philosophy in Science at Universiti Tunku Abdul Rahman.

Approved by:

_____

(Assoc. Prof. Dr. TANG PEK YEE) Date:……..
Supervisor
Department of Mechatronics and BioMedical Engineering
Lee Kong Chian Faculty of Engineering and Science
Universiti Tunku Abdul Rahman

_____

(Asst. Prof. Dr. TEE SHIAU FOON) Date:……..
Co-supervisor
Department of Chemical Engineering
Lee Kong Chian Faculty of Engineering and Science
Universiti Tunku Abdul Rahman

LEE KONG CHIAN FACULTY OF ENGINEERING AND SCIENCE
UNIVERSITI TUNKU ABDUL RAHMAN

Date: _____

SUMMISSION OF DISSERTATION

It is hereby certified that **CHOW TZE JEN** (ID No: **11UED06773**) has completed this dissertation entitled "TARGETED NEXT GENERATION SEQUENCING OF *ZNF804A*, *DTNBP1*, *AKT1* AND *TCF4* TO IDENTIFY RISK VARIANTS IN SCHIZOPHRENIA MULTIPLEX FAMILIES" under the supervision of Dr. Tang Pek Yee (Supervisor) from the Department of Mechatronics and BioMedical Engineering, Lee Kong Chian Faculty of Engineering and Science, and Dr. Tee Shiau Foon (Co-Supervisor) from the Department of Chemical Engineering, Lee Kong Chian Faculty of Engineering and Science.

I understand that the University will upload softcopy of my dissertation in pdf format into UTAR Institutional Repository, which may be made accessible to UTAR community and public.

Yours truly,

_____

(CHOW TZE JEN)

DECLARATION

I hereby declare that the dissertation is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at UTAR or other institutions.

Name _____

Date _____

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviations

5-HTT	Serotonin Transporter gene
AAO	Age at Onset
AIDS	Acquired Immune Deficiency Syndrome
AKT1	V-Akt Murine Thymoma Viral Oncogene Homolog 1 gene
ANKRD1	Ankyrin Repeat Domain 1 gene
ANOVA	Analysis of Variance
APOE	Apolipoprotein E gene
BDNF	Brain Derived Neurotrophic Factor gene
bp	Base pair
CHRM2	Cholinergic Receptor Muscarinic 2 gene
CHRNA7	Cholinergic Receptor Nicotinic Alpha 7 subunit
CMOS	Complementary Metal-oxide-semiconductor
CNV	Copy Number Variation
COMT	Catechol-O-methyltransferase gene
DAOA	D-Amino Acid Oxidase Activator gene
DEL	Deletion
DNA	Deoxyribonucleic Acid
DRD2	Dopamine Receptor D2 gene
DTNBP1	Dystrobrevin-Binding Protein 1 gene
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders (fourth edition)
EDTA	Ethylene-diamine-tetra-acetic Acid
EOS	Early-onset Schizophrenia

FDR	False Discovery Rate
FU	Fluorescence Unit
GSK-3	Glycogen Synthase Kinase-3
GWAS	Genome Wide Association Study
Hoxc8	Homeobox C8 gene
IN	Insertion
ISFET	Ion-sensitive Field-effect Transistor
ISP	Ion Sphere Particle
IQ	Intelligence Quotient
kb	Kilo base
KCNE	Potassium voltage-gated channel subfamily E
kDa	Kilo Dalton
KIBRA	Kidney and Brain expressed gene
LD	Linkage Disequilibrium
Max	Maximum
Mb	Mega base
Min	Minimum
MNV	Multi Nucleotide Variant
miRNA	MicroRNA
mRNA	Messenger RNA
n	Number of Subjects
NCBI	National Center for Biotechnology Information
NF-1C	Nuclear Factor 1C
NF-Y	Nuclear Transcription Factors Y
ng	Nanogram

nm	Nanometer
nM	Nanomolar
NRG1	Neuregulin 1 gene
NGS	Next Generation Sequencing
PacBio	Pacific Biosciences
PGM	Personal Genome Machine
PI3K	Phosphoinositide-3 Kinase
PRSS16	Protease, Serine 16 gene
p	Probability
PCR	Polymerase Chain Reaction
rpm	Revolutions per Minute
rs	Reference SNP
SD	Standard Deviation
SDS	Sodium Dodecyl Sulphate
SNP	Single Nucleotide Polymorphism
SNV	Single Nucleotide Variant
SPSS	Statistical Package for Social Sciences
Ta	Annealing Temperature
TCF4	Transcription Factor 4 gene
TGF-β	Transforming Growth Factor β
TMT	Trail Making Test
µl	Microliter
UTR	Untranslated Region
WHO	World Health Organisation
ZNF804A	Zinc-Finger Protein 804A gene

CHAPTER 1

INTRODUCTION

1.1 Schizophrenia

Schizophrenia is a complex and debilitating psychiatric disorder that impairs an individual's perception, thought, and behaviour (Ministry of Health Malaysia, 2009). It involves a complex array of neurophysiology, neurochemical, and psychological disturbance (Laviolette, 2007) that results from multifaceted genetic interactions with environmental factors (Oliver, 2011). Schizophrenia affects approximately 1% of the population worldwide (Tandon *et al.*, 2008), whereby life expectancy is reduced by approximately 10 years, mainly as a result of suicide (Rossler *et al.*, 2005). In Malaysia, the reported prevalence of mental health problem escalated from 10.7% in 1996 to 29.2% in 2015 (Institute for Public Health, 2015).

Clinical symptoms of schizophrenia usually begin in late adolescence or early adulthood and seldom occur after the age of 45. The symptoms are presented in three main categories of clinical features: positive, negative, and cognitive symptoms (National Institute of Mental Health, 2009). Positive, also known as psychotic symptoms, are portrayed by delusions (false beliefs), hallucinations (false perceptions), and paranoid, with a 90% incidence in all patients (Barbato, 1998). On the other hand, negative symptoms reflect a loss

of normal psychomotor, emotional and social abilities, showing social isolation, low emotional expression, and poverty of affect and speech (Cohen and Docherty, 2004). They are less obvious but often persist even after the resolution of positive symptoms (Ministry of Health Malaysia, 2009).

Cognitive symptoms are related to concentration, executive functions, and memory impairment, characterised by chaotic speech and slow thinking (Peralta and Cuesta, 2001). Cognitive deficits might occur as prodromal symptoms that are present before the onset of clinical symptoms and persist to be central to the functional disability of schizophrenia (Lewis and Lieberman, 2000). Similar to negative symptoms, cognitive symptoms may be hard to recognise and are often only detected with tests (National Institute of Mental Health, 2009). Patients often experience overlapped symptoms and may develop their own unique combination of symptoms (Ministry of Health Malaysia, 2009).

There is currently no cure for schizophrenia, and researches focus on developing ways for early detection to increase the chances of early treatment. Treatment methods rely heavily on antipsychotic medication that suppress dopamine activity and biochemical imbalance that causes schizophrenia (Freedman, 2003). They are mostly effective in reducing positive symptoms but exert limited effect on negative and cognitive symptoms (Erhart *et al.*, 2006; Keefe *et al.*, 2007), which give the main impact on patients' quality of life (Hyman and Fenton, 2003).

The increasing expenditure on antipsychotic medications every year reflected the economic burden of mental illness. Malaysian government has been subsidizing heavily on antipsychotic medications and spent nearly 2.3 billion Malaysian Ringgit on drug expenditure in 2005 (Chee, 2009). The rising trend was evident when the expenditure for psycholeptics was ranked top 5 in the country's total medicines expenditure in 2010 but was absent in the top 10 list in previous years (Ministry of Health Malaysia, 2014). Aside from that, there are also indirect costs including loss of productivity, disability, and legal problems such as violence (Rossler *et al.*, 2005) which contributed to the staggering economic burden of mental illness to the society.

1.2 Cognitive Impairment as Phenotype of Schizophrenia

Cognitive impairment is one of the core features of schizophrenia and is found apparent in almost all patients and is consistent across illness stages (Lesh *et al.*, 2011; Keefe and Harvey, 2012). It usually appears even before signs of the first onset of psychosis emerge (Bilder *et al.*, 2011). The most important domains of cognitive deficits in schizophrenia are working memory (Mark and Toulopoulou, 2016), attention and executive functions including reasoning, problem solving, and speed of processing (Green *et al.*, 2004).

Working memory is defined as a temporary storage of information that is no longer present in external environment. Functional working memory is essential for a variety of executive functions (Andre *et al.*, 2016). Attention,

sometimes refer to as vigilance, is the ability to maintain concentration over time. Impairment in attention leads to difficulties in following instructions and conversations (Keefe and Harvey, 2012). Executive functions are hypothesised to be associated to dorsolateral prefrontal cortex dysfunction which leads to the limitation of the ability for response and performance (Lesh *et al.*, 2011). Patients with executive functions impairment often lose the ability to adapt to changes (Keefe and Harvey, 2012).

Recently, neurocognitive ability is increasingly being studied as a heritable phenotype for schizophrenia, while some categorised it as an intermediate phenotype or endophenotype (Toulopoulou *et al.*, 2010). Patients with family history would be expected to have increased genetic risk burden. Study reported that first-degree relatives of patients with schizophrenia showed similar cognitive impairment as seen in patients (Harvey, 2008). This proposes that cognitive impairment may be manifest in individuals who have family history of schizophrenia.

1.3 Genetics Effects between Schizophrenia and Cognition

Schizophrenia is known to be a highly polygenic disorder, presumably caused by a number of multifaceted genotypic networks related to distinct clinical symptoms (Lee *et al.*, 2012b). The aetiology of schizophrenia is elusive but strong evidences indicate that genetic inheritance plays the most important role, accounting for approximately 80% of all cases (Tsuang *et al.*,

2000; Tandon *et al.*, 2008). Family-based and twin-based studies showed that risk of developing schizophrenia is higher amongst those with family history. Monozygotic twin have the highest risk, followed by children with two affected parents, dizygotic twin, children, siblings, and lastly parents (Nasrallah and Smeltzer, 2002). Nearly 21.6% of patients with schizophrenia in Malaysia had a family history of mental illness (Aziz *et al.*, 2008). A family where an individual diagnosed with schizophrenia has a 1st- or 2nd-degree relative with similar diagnosis is known as schizophrenia multiplex family. Previous large family studies of schizophrenia observed that multiplex families are expected to have higher prevalence compared to the general unrelated population (DeLisi, 2016).

Cognitive impairment correlates to heritable deterioration in brain areas especially in the frontal and temporal lobe, hippocampus and subcortical (Kuperberg and Heckers, 2000), which are often evident even before the onset of symptoms (Reichenberg *et al.*, 2002). Cognitive performances such as attention, executive functioning, verbal fluency, and memory, are associated with frontal and temporal lobe functioning, and they contribute to the pathophysiology of cognitive impairments in schizophrenia (Lesh *et al.*, 2011). Twins and relatives-study found that healthy counterparts of the probands experience slight abnormalities when compared to healthy controls (Cannon *et al.*, 2000; Snitz *et al.*, 2006), suggesting a shared genetic liability (Toulopoulou *et al.*, 2010).

Evidence showed that cognitive deficits are extensively associated to genetic influences (Toulopoulou *et al.,* 2007). Recent work found significant correlation between cognitive phenotype and schizophrenia, with up to 90% of the cause due to shared genetic effects (Toulopoulou *et al.,* 2010; Owens *et al.,* 2011). In a genome-wide association study (GWAS) study, schizophrenia susceptible gene variants were identified to be association with both brain development and cognition (Stefansson *et al.,* 2009). Besides that, early onset schizophrenia appears to be associated with cognitive impairments and reflects a genetic predisposition to the disorder. Genetic studies have suggested several genes such as *DAOA* (Addington *et al.*, 2004), *DTNBP1* (Gornick et al., 2005), and *NRG1* (Addington *et al.,* 2007) to be linked with childhood onset schizophrenia. Interestingly, these genes are also associated with impaired cognition (Jansen *et al.,* 2009; Baek *et al.,* 2012).

Several genes have been found to contain variants associated with the individual phenotypes of cognitive impairments. Among genes investigated, *CHRM2* and dopaminergic *DRD2* are found related specifically to intelligence, while serotonergic *5-HTT* and *KIBRA* are linked to memory. Four other genes: *DTNBP1*, *COMT*, *BDNF*, *APOE* have associations to both intelligence and memory (Bilder *et al.*, 2011). Apart from that, *ZNF804A* (Lencz *et al.*, 2010), *DTNBP1* (Wolf *et al.*, 2011), *NRG1* (Yokley *et al.*, 2012), *AKT1* (Tan *et al.*, 2008), and *TCF4* gene (Albanna *et al.*, 2014) also had reported links to both schizophrenia and cognitive deficit. Several rare variants such as *KCNE1*/*KCNE2* and *CHRNA7* have also been highlighted for possible shared genetic liability between schizophrenia and cognition (Tam *et al.*, 2010).

These findings provide evidence for the role of these genes as contributors to the structural changes in the brain, resulting in cognitive impairments shown in schizophrenia (Lesh *et al.*, 2011). This underlines the need to understand the genetic basis underlying cognitive deficits in schizophrenia for better understanding of disease pathophysiology and effective treatments.

In particular, schizophrenia susceptible genes *ZNF804A*, *DTNBP1*, *AKT1*, and *TCF4* are often found in between neurotransmitter pathways and are reported to play a role in pathway regulation (Flora *et al*., 2007; Lai *et al*., 2006; Guo *et al*., 2009; Esslinger *et al*., 2011). Previous study on *AKT1* and *TCF4* also yield significant results that support the association of these genes with early onset schizophrenia and cognition (Chow *et al*., 2016). Thus they might serve as good candidates to investigate the association of these genes with schizophrenia and cognition.

1.4 Objectives of Study

With the alarming high prevalence of schizophrenia and the complications of cognitive deficits, there is a need to detect genetic association for schizophrenia and cognitive phenotypes. Until now, most of the case-control and family-based neurocognition studies have relied on single gene or candidate variants located in the intron regions and no causative genes have yet to be identified (Schreiber *et al.*, 2013). Functional regions, such as exons, untranslated regions (UTRs), and promoters are known to regulate gene and

protein expressions, thus they serve as a better candidate to explore for the association of neurocognition and schizophrenia.

Besides, disease-specific family cohort design, where subjects are enriched for genetic and familial risk, serves as an advantage for studying gene-gene interactions and to adjust for unmeasured environmental factors (Hopper, 2011). Therefore, the current study utilised a multiplex family-based case-control sample cohort to investigate the functional regions of genes involved in the genetic pathway of neurocognition with schizophrenia and cognitive deficits. This will contribute to the understanding of specific genetic and molecular expression pathways that underlie cognitive deficits in schizophrenia, facilitating disease prevention study and treatment.

Thus, the objectives of this study are: (1) to study the association of *ZNF804A*, *DTNBP1*, *AKT1*, and *TCF4* gene variants with schizophrenia; and (2) to identify the risk variants of these genes specific to cognition.

CHAPTER 2

LITERATURE REVIEW

2.1 Schizophrenia

The multiple facets of schizophrenia can be divided into five subtypes, namely paranoid, undifferentiated, disorganised, residual, and catatonic subtype. The paranoid subtype is characterised by hallucination and delusion but thought disorder, disorganised behaviour, and affective flattening are absent from the symptoms. On the other hand, undifferentiated patients have psychotic symptoms without the criteria for paranoid, disorganised, and catatonic. For the disorganised subtype, both thought disorder and flat affect can be detected, while the residual subtype portrays positive symptoms at a low intensity. Less frequently observed today, the catatonic subtype is characterised by psychomotor disturbances together with symptoms such as catatonic stupor and waxy flexibility (Bilder, 2006).

Schizophrenia may be diagnosed at any time in one's lifetime and the severity of illness differs from mild to severe. Age at onset (AAO) is the age at diagnosis for schizophrenia and is fundamental to understand the illness. The typical AAO is in late adolescence or early twenties. Males tend to have an earlier AAO and higher illness severity compared to females (Bergen *et al.*, 2014). The onset of schizophrenia peaks at age of 18 to 25 years for men, while in women, the mean AAO is 25 to 35 (Ochoa *et al.*, 2012). Patients of

early-onset schizophrenia (EOS) develop the psychotic symptoms during childhood or adolescence up to the age of 17 years. The proportion of patients with early onset is only around 4% of schizophrenia patients (Cannon *et al.*, 1999). Those with EOS show insidious onset, more severe premorbid neurodevelopmental abnormalities and cognitive performance, poorer response to treatment, more frequent terrifying visual hallucinations, and higher rate of familial psychopathology (Hoff *et al.*, 1996; Margari *et al.*, 2008; Rajji *et al.*, 2009).

2.2 Demographics of Schizophrenia

The incidence of schizophrenia was reported to be higher in migrants and minority ethnic populations, mainly due to greater exposure to social adversity (Morgan *et al.*, 2010). Besides that, the AAO was also found to differ across countries and migration rate (Versola-Russo, 2006). Patients in developing countries tend to have earlier onset than developed countries, while immigrants have earlier onset than non-immigrants. This may be due to differences in environment and social factors (Rabinowitz and Fennig, 2002). On the other hand, the 2015 Malaysian National Health and Morbidity Survey showed that individuals that received up till secondary education have the highest prevalence on mental health, followed closely by those with primary and tertiary education, and lastly, individuals with no formal education have the lowest prevalence (Institute for Public Health, 2015).

Malaysia National Mental Health Registry indicated that the registered schizophrenia patients in Malaysia were mostly Malays (54%), followed by Chinese (28%), Indians (9%), and others (9%) (Aziz, 2008). This ratio is consistent with the Malaysian ethnic group distribution, though the prevalence is similar among the three major ethnics (Malays: 28.2%, Chinese: 24.2%, Indians: 28.9%) (Institute for Public Health, 2015).

Gender differences for the cumulative lifetime risk are generally similar between the male and female (Rossler *et al.*, 2005), but males scored a higher rate, especially in the age group younger than 40 years (McGrath *et al.*, 2004). Of note, more than 60% of schizophrenia cases in Malaysia were males (Aziz *et al.*, 2008). Females tend to develop the disorder later in life, with approximate onset of 3 to 6 years later than male (Thara and Kamath, 2015). It is believed that the gonadal hormone, estrogen, might play a neuro-protective role in schizophrenia development through interactions with neurotransmitters (Boerma *et al.*, 2010). Although schizophrenia is one of the leading causes of disease burden for women below 40 years (WHO, 2004), females tend to have a better outcome after treatment compared to males (Thara and Kamath, 2015).

2.3 Aetiology of Schizophrenia

The aetiology of schizophrenia remains enigmatic. There is strong evidence indicating that both genetic and environmental factors may institute the disorder (Maric and Svrakic, 2012). Besides the influence of familial

inheritance, epidemiology studies also suggested that genetic risk can be attributed to the polygenic nature of schizophrenia itself. Multiple susceptibility genes, each with small to moderate effect, combined and contributed to cause the disorder (Corvin, 2011). Although the heterogeneity of the aetiology is well accepted, yet, only a small portion of the factors have been identified (Brown and McGrath, 2011).

Besides genetic factors, there are numerous environmental risk factors associated with schizophrenia, including: urbanisation (Pedersen and Mortensen, 2001), migration history (McGrath *et al.*, 2004), malnutrition (Tandon *et al.*, 2008), cannabis abusers (Zammit *et al.*, 2002), and prenatal virus infection (Koponen *et al.*, 2004). Despite many speculations on environmental risk factors, aetiology of the disorder remains poorly understood (Schwab and Wildenauer, 2008). Recent advancement in epigenetics studied on molecular mechanisms that portray environmental influences on gene expression (Dick, 2011).

2.4 Current Methods in Schizophrenia Genetics

The study of genes in schizophrenia employs two main approaches: linkage analyses (ancestral study) and association studies (case-control study) (Dick, 2011). Linkage analyses provide information on the genome position of the schizophrenia risk genes, while association studies describe specific gene variation that is involved in disease pathogenesis. Besides that, there is a

diversity of genetic variation, including duplication, insertions, deletions, Single Nucleotide Polymorphisms (SNPs), and copy number variations (CNVs). In human genome that consist of three billion nucleotides, approximately 0.2% of the nucleotides are polymorphic (Tishkoff and Verrelli, 2003; Jorde *et al.*, 2001). This provides a platform to study the foundations of phenotypic variation.

Schizophrenia is a polygenic disorder with multiple SNPs contributing to certain phenotype and the search for selective signatures across the genome has been continuing for decades. With rapid advances in technology developments, many powerful approaches have emerged. Introduction of GWAS and Next Generation Sequencing (NGS) have made bioinformatics analyses of genomic DNA sequence data and frequencies of polymorphic alleles a common approach in search for susceptible genes for schizophrenia (Need *et al.*, 2009; Stefansson *et al.*, 2009). GWAS is powered to identify common polymorphisms in large sample cohorts, but most findings were intragenic and unlikely to be functional. As such, NGS can serve as a complementary approach (Schreiber *et al.*, 2013).

NGS, also known as massively parallel sequencing, is an advanced sequencing method that allows for qualitative and quantitative information of genes from single base to large copy number polymorphism (Reis-Filho, 2009). Unlike conventional Sanger sequencing, it is able to sequence multiple genes and process millions of sequence reads in parallel using minimal DNA input in a shorter turnaround time (Mardis, 2008).

There are a few types of NGS platforms available in the market (Table 2.1). Of note, Illumina, Ion Torrent Personal Genome Machine (PGM), Pacific Biosciences (PacBio) RS, and Roche 454 platforms are among the few that are commonly used, depending on study design, sample size and cost-effectiveness. Illumina and Ion Torrent PGM are widely used for diagnostic purposes, the latter more common for small sample studies due to its cost-effectiveness (de Leng *et al.*, 2016). Ion Torrent adopts the sequencing by synthesis method and detects hydrogen ions that are released during DNA polymerisation. The resulting pH change is detected by an ion sensor known as ion-sensitive field-effect transistor (ISFET) and integrated complementary metal-oxide-semiconductor (CMOS) chip. It has limited accuracy in measuring homopolymers as the resulting pH change is not proportional to the number of nucleotides detected (Goodwin *et al.*, 2016). Although the read length is much shorter compared to Roche 454 and PacBio, but this can be compensated by the relatively lower costs and shorter run time.

Different approaches including whole-genome, whole-exome, targeted exome and hotspot sequencing can be studied using these platforms. For human, sequencing the entire genome is possible but not financially feasible, especially for a large number of subjects (Schuster, 2008). Also, the data processing and bioinformatics systems involved requires huge amount of time and work. Recent advances in NGS focused on exome sequencing, which contains most functional variation and can facilitate the identification of rare genetic variants, especially in familial study.

Table 2.1: Comparison of major NGS platforms (Goodwin *et al.*, 2016; Mardis, 2017)

NGS platforms	Sequencing reaction	Read length	Major advantages	Major disadvantages
Roche 454	Pyrosequencing	400 bp	Long read length, short run time	High cost per Mb, high error rate in homopolymer
Illumina	Reverse terminator	150 – 300 bp (paired end)	Versatile, scalable, high-throughput per cost	Long run time, higher cost per Mb, requires high sample concentration
ABI SOLiD	Ligation-based sequencing	50 – 75 bp	Low instrument cost and high accuracy	Short reads, long run time, error in palindromic sequences
PacBio	Real-time single molecule sequencing	Up to 40 kb	Real time sequencing, longest read length, short run time	High capital cost, high error rate
Ion Torrent PGM	Semiconductor sequencing	200 – 400 bp	Short run time, simple method, low instrument cost	Error in homopolymer
Oxford Nanopore	Single molecule sequencing	Depends on library	Very long reads, portable	Low throughput
Qiagen GeneReader	Fluorescence-based sequencing by synthesis	107 bp	All-in-one platform from sample preparation to analysis	Long run time, limited applications

2.5 Cognition-Related Genes in Schizophrenia

2.5.1 Zinc-Finger Protein 804A (*ZNF804A*)

Zinc-finger protein 804A (*ZNF804A*) is a 341 kb gene located in chromosome 2q32.1. It consists of four exons that transcribe a 137 kDa protein with a zinc finger domain at the N-terminal that enables DNA binding and thus regulate multiple genes expression (Walters *et al.*, 2010). *ZNF804A* is expressed predominantly in human brain, particularly in the developing cortex, hippocampus, and adult cerebellum (Johnson *et al.*, 2009; Lencz *et al.*, 2010).

Its function is largely unknown and much work is needed to study the role of this gene (Donohoe *et al.*, 2010). However, identification of the mouse homologue of *ZNF804A*, *zfp804a*, as a downstream target for *Hoxc8*, a transcription factor essential in promoting nerve growth and expression, suggested a role for *ZNF804A* in regulating early neurodevelopment (Chung *et al.*, 2010). Besides that, it is also associated with abnormal brain white matter microstructure (Mallas *et al.*, 2016) and schizophrenia-associated genes such as *COMT*, *DRD2*, and *PRSS16* (Girgenti *et al.*, 2012), hence suggesting *ZNF804A* to be a susceptibility gene for schizophrenia.

ZNF804A was first reported to be associated with schizophrenia in 2008 through a GWAS report and an intron 2 variant, rs1344706, was the first to achieve genome-wide significance for the disorder (O'Donovan *et al.*, 2008). Since then, genetic studies of *ZNF804A* in schizophrenia reported predominantly on rs1344706. Many replication studies were done on

independent GWAS, case-control cohorts, and meta-analysis to show evidence of association between rs1344706 and schizophrenia (Riley *et al.*, 2010; Williams *et al.*, 2011b; Zhang *et al.*, 2012). Although not all results reported positive association (Li *et al.*, 2011), most studies, especially in the European cohorts, showed consistent evidences for schizophrenia association (Schwab *et al.*, 2013; Mallas *et al.*, 2016).

Rs1344706 is often associated with schizophrenia in a range of brain structure and neurocognitive phenotypes. It alters neural connectivity of the prefrontal cortex and hippocampus region, contributing to the cognitive impairments explicitly in the working memory domain (Esslinger *et al.*, 2011). In patients with schizophrenia, rs1344706 was linked to abnormal sleep phenotypes, which leads to poorer memory consolidation during sleep (Hellmich *et al.*, 2015). Besides that, interactions between *ZNF804A* variants and IQ level give rise to significant effects on working memory and executive functions. It was suggested that individual IQ level modulates the role of rs1344706 in both cognitive deficits and schizophrenia, whereby the risk allele A is linked to lower cognitive performance in patients with low IQ but the effect is reverse in those with high IQ (Chen *et al.*, 2012). Evidence also showed that individual carrying the rs1344706 risk allele A had relatively spared memory function (Walters *et al.*, 2010) and relatively larger grey matter volumes (Donohoe *et al.*, 2011). These findings suggest that this risk allele A might pose a putative protective role to schizophrenia neurocognition (Schultz *et al.*, 2014) but much work is to be done as there might be more unidentified *ZNF804A* variants involved.

2.5.2 Dystrobrevin-Binding Protein 1 (*DTNBP1*)

Dystrobrevin-binding protein 1 (*DTNBP1*) gene lies on chromosome 6p22.3, spanning 140 kb with 10 exons. It codes for the protein dysbindin, which interacts with α- and β-dystrobrevin in the brain and skeletal muscle cells, forming part of the dystrophin-associated protein complex as a scaffold for signalling proteins (Guo *et al.*, 2009). As such, *DTNBP1* has been greatly implicated in signalling pathways, neuronal glutamate regulation and axonal growth (Numakawa, 2004; Fallgatter *et al.*, 2010). It is associated with bipolar disorder (Gaysina *et al.*, 2009), brain tumor (Correa *et al.*, 2016), and is known as one of the most promising genes in schizophrenia susceptibility (Bray *et al.*, 2005).

Dysbindin was first discovered in mouse, mostly expressed in the axonal terminals of the mouse brain (Benson *et al.*, 2001), hence the human ortholog of mouse dysbindin, *DTNBP1*, was suggested to have a possible role in neuronal transmission. Straub *et al.* (2002) first identified *DTNBP1* as a schizophrenia susceptibility gene by showing significant association between schizophrenia and *DTNBP1* variants in a large pedigree-based family study (Straub *et al.*, 2002). Independent replication studies across different populations later provided further evidence for the association of *DTNBP1* with schizophrenia (Tang *et al.*, 2003; Numakawa, 2004; Pae *et al.*, 2009; Voisey *et al.*, 2010), but several others failed to replicate the result (Liu *et al.*, 2007; Turunen *et al.*, 2007). Inconsistency in the association patterns might be due to considerable allelic heterogeneity of *DTNBP1* (Schwab *et al.*, 2003).

Interestingly, *DTNBP1* is located in chromosome 6, which is identified as a highly susceptibility region for schizophrenia (Straub *et al.,* 2002). Besides that, overexpression of *DTNBP1* is often associated with increased activity of *AKT1*, a schizophrenia susceptible gene, hence suggesting the involvement of *DTNBP1* in the AKT signalling pathway (Numakawa *et al.*, 2004; Riley and Kendler, 2006). Post-mortem study also reported reduced expression of presynaptic *DTNBP1* in hippocampus and prefrontal cortex, both brain regions which are known to control cognitive functions critical to schizophrenia (Weickert *et al.*, 2004; Trost *et al.*, 2013; Zhang *et al.*, 2017).

Supporting this notion, association of *DTNBP1* with cognitive functioning has been reported in various studies, where the gene appears to play an important role in intellectual functioning, attention, processing speed, spatial working memory, verbal memory and visual memory (Burdick *et al.*, 2007; Varela-Gomez *et al.*, 2015; Bakanidze *et al.*, 2016). *DTNBP1* has been implicated in the decline of intelligence portrayed in schizophrenia patients, suggesting that genetic variations of *DTNBP1* may increase the risk of schizophrenia by disrupting intellectual functioning (Zinkstok *et al.*, 2007). Besides that, working memory was found to be related to altered dopaminergic and glutamatergic neurotransmission caused by synaptic effects of *DTNBP1* (Wolf *et al.*, 2011), hence the association between *DTNBP1* and schizophrenia may possibly arise from the influence of the gene on cognition (Baek *et al.*, 2012). Despite accumulating evidences supporting *DTBNP1* as a schizophrenia susceptible gene, its functional variants have yet to be identified (Zinkstok *et al.*, 2007).

2.5.3 V-Akt Murine Thymoma Viral Oncogene Homolog 1 (*AKT1*)

V-akt murine thymoma viral oncogene homolog (AKT) was termed after it was initially cloned as a cellular homolog of v-act oncogene (Hallmayer, 2004). Also known as protein kinase B, AKT is a serine/threonine-specific protein kinase that functions in multiple cellular functions for instance, glucose metabolism, apoptosis, cell proliferation, cell migration (Woodgett, 2005) and neurodevelopmental pathways such as cell signalling and signal transduction (Lai *et al.*, 2006; Arguello and Gogos, 2008).

The AKT family consists of three isoforms: AKT1, AKT2, and AKT3, each encoded by separate genes and have specific functions (Beaulieu, 2012). AKT1 is the most extensively studied of this family and is often related to metabolic syndrome, mediating two major pathways: muscle atrophy regulation and glucose transport pathways (Harmon *et al.*, 2010). It is also part of an important signal transduction pathway that regulates cellular functions for example survival, growth, transcriptional regulation and metabolism (Thiselton *et al.*, 2008). AKT1 is activated in response to stimulation by the plasma membrane receptor, phosphoinositide-3 kinase (PI3K). When triggered by dopamine, PI3K signals the recruitment of AKT1 to neuronal cell surface, where AKT1 will then negatively regulates the constitutively active glycogen synthase kinase-3 (GSK-3) through phosphorylation. Together with AKT1, GSK-3 is often associated with mood disorder and anxiety, whereby inhibition of GSK-3 is demonstrated in the action of the mood stabilizer lithium (Freyberg *et al.*, 2010; Beaulieu, 2012).

AKT plays an essential role in modulating synaptic plasticity and is involved in intracellular trafficking of dopamine and norepinephrine in the brain (Freyberg *et al.*, 2010). It is a key signalling intermediate that is located downstream of dopamine receptor D2 (DRD2), a well-established target for antipsychotic drugs, hence AKT serves as a suitable candidate to study the dopaminergic transmission and dopamine-associated behaviours (Thiselton *et al.*, 2008). Due to its role with DRD2, AKT has been widely implicated in pathogenesis of neurodegenerative disorders such as Parkinson's disease (Ran *et al.*, 2011), bipolar disorder, and schizophrenia (Karege *et al.*, 2012).

The association of *AKT1* with schizophrenia was first demonstrated in a large European familial study (Emamian *et al.*, 2004). Since then, genetic contributions of *AKT1* to schizophrenia were widely reported in case-control and familial association studies (Thiselton *et al.*, 2008; Pietilainen *et al.*, 2009; Liu *et al.*, 2016). Deficiency in *AKT1* expression was found to cause abnormal prefrontal cortical structure leading to cognitive deficiencies such as impaired working memory in schizophrenia (Lai *et al.*, 2006). Evidence also showed *AKT1* variant rs2494732 to be associated with lower memory and attention performance in patients (Ohi *et al.*, 2013). Although some studies reported negative results (Sanders *et al.*, 2008; Lee *et al.*, 2010) and the association of *AKT1* seems inconclusive across different populations. Yet, post-mortem brains of patients showed reduced AKT1 levels in hippocampus and frontal cortex (Hallmayer, 2004), further highlighting *AKT1* as a candidate gene for schizophrenia susceptibility. Nevertheless, the influence of ethnic variation in gene association is worth examined.

2.5.4 Transcription Factor 4 (*TCF4*)

Transcription factor 4 (*TCF4*) is a basic helix–loop–helix transcription factor gene located at chromosome 18q21.1. The encoded protein regulates transcription by recognising the E-box (5'-CANNTG-3') binding site of certain genes, such as those involved in cell differentiation (Forrest *et al.*, 2013). *TCF4* is required in the development of lymphoid progenitors in immune system, hence it is sometimes referred to as immunoglobulin transcription factor 2 (ITF2). In addition, *TCF4* also plays an essential role in early neurodevelopment (Flora *et al.*, 2007) and grey matter thickness during cerebral aging (Albanna *et al.*, 2014).

Mutations in *TCF4* are associated with risk of many diseases. Haploinsufficiency of *TCF4* may cause Pitt-Hopkins Syndrome, a disorder associated with mental retardation (Blake *et al.*, 2010); while intronic CTG-trinucleotide repeat expansion causes a cornea degenerative disease known as Fuchs endothelial corneal dystrophy (Wieben *et al.*, 2014). Besides that, evidence also suggests that common variants in or near *TCF4* gene are prominent to confer risk of schizophrenia (Navarrete *et al.*, 2012).

Following the initial report of *TCF4* variant rs9960767 in achieving genome-wide significance to schizophrenia susceptibility (Stefansson *et al.*, 2009), many GWAS replication studies, coupled with meta-analyses, associated other *TCF4* variants such as rs4309482, rs12966547, and rs2958182 with increased risk of schizophrenia (Li *et al.*, 2010; Ripke *et al.*,

2011; Steinberg *et al.*, 2011). Among the identified variants, rs9960767 was the most studied (Lennertz *et al.*, 2011). Post-mortem analysis revealed that this variant is not functional and does not alter mRNA expression as measured in the adult brain (Williams *et al.*, 2011a). Nonetheless, *TCF4* mRNA level in peripheral blood was found significantly higher in patients with schizophrenia and bipolar disorder compared to controls (Wirgenes *et al.*, 2012). Of note, these genetic association studies only identified regions with disease risk loci but are unable to determine the actual functional variants (Forrest *et al.*, 2014).

On another note, these schizophrenia-associated *TCF4* risk variants have shown great influence on cognitive functions. In animal study, overexpression of *TCF4* in mouse brain was shown to reduce prepulse inhibition, a core phenotype portrayed in human schizophrenia (Brzozka *et al.*, 2010). Interestingly, the risk allele of rs9960767 is correlated with dysfunctional sensorimotor gating, which is measured by prepulse inhibition in patients, hence mirroring an established schizophrenia endophenotypes (Lennertz *et al.*, 2011; Quednow *et al.*, 2011). Besides that, *TCF4* gene is also implicated in cognitive functions such as reasoning and problem-solving (Albanna *et al.*, 2014), memory, language performance (Hui *et al.*, 2015), intellectual disability (Hill *et al.*, 2014; Maduro *et al.*, 2016), verbal learning, and executive functioning (Wirgenes *et al.*, 2012). These studies highlight the role of *TCF4* in information processing pathways that contributes to cognitive deficits observed in schizophrenia.

CHAPTER 3

MATERIAL AND METHODS

3.1 Subject Selection

This study involved 12 multiplex families (n = 26), comprising of 12 probands and their affected first-degree relatives, recruited from Hospital Permai, Johor Bahru, Malaysia. All patients diagnosed with schizophrenia met the DSM-IV criteria. Those with co-morbidity and current substance abuse were excluded. Besides that, 11 control families (n = 22) were also recruited (Table 3.1). Controls were selected based on the following exclusion criteria: (i) no personal nor family history of mental disorders, (ii) no history of major illness such as cancer, stroke and AIDS, (iii) not on drug therapy or substance abuse (excluding nicotine), and (iv) no learning disabilities.

Education levels of patients were matched to controls as possible. Written consent was obtained for all participants (Appendix A). All subjects were self-identified as being of Malay, Chinese and Indian descent. All subjects had normal or corrected-to-normal vision, 18 years or older and able to understand the instructions of neurocognitive assessment. The study was approved by the University of Malaya Medical Center Medical Research Ethics Committee Institutional Review Board, Malaysia (approval number: UMMC/MECIRB926.63).

Table 3.1: Demographic table

	Patients	Controls
Number of families	12	11
Number of subjects	26	22
Mean age ± standard deviation	40.20 ± 9.73	38.25 ± 9.63
Gender (male / female)	17 / 9	12 / 10
Ethnic (Malay / Chinese / Indian)	12 / 10 / 4	8 /12 / 2

3.2 Clinical and Neurocognitive Assessment

Cognitive ability of all participants was assessed using Trail Making Test (TMT) according to instructions (Reitan, 1958). The test consists of two parts: in TMT-A (Appendix B), participants were required to connect 25 randomly distributed numbers in an ascending order (i.e. 1-2-3-4-etc.); while in TMT-B (Appendix C), participants were required to connect randomly distributed numbers and letters in an alternate manner (i.e. 1-A-2-B-3-C-etc.). An example was shown to participants prior to the test. Participants were instructed to complete the test as fast and accurately as possible.

The total time (in seconds) required to complete each test were recorded as dependent variable (score). TMT-A was recorded for assessing processing speed, and TMT-B for executive functioning, including reasoning and problem-solving (Perianez *et al.*, 2007). TMT-A relies mainly on visuoperceptual abilities, while TMT-B reflects working memory and task-

required to complete TMT-A and TMT-B are 29 seconds and 75 seconds respectively. Individuals will be considered as deficient if their completion time exceeds 78 seconds for TMT-A and 273 seconds for TMT-B.

From the direct scores, four derived scores: difference score (B-A), ratio score (B:A), log B:A and proportion score (B–A):A were calculated as additional indexes for better description of the cognitive tests (Perianez *et al.*, 2007). B-A score minimizes speed demands from the test evaluations and represents a relative indicator of executive control function; B:A score correlates with task-switching ability; log B:A score generalize results across different groups; while (B–A):A score reflects an index for prefrontal cortex functions (Perianez *et al.*, 2007; Sanchez-Cubillo *et al.*, 2009).

3.3 DNA Isolation

DNA was extracted from 200 ml peripheral blood samples of each participants using QIAamp® DNA Blood Mini Kit (Qiagen, CA, USA) following the manufacturer's instructions with minor modification: DNA was eluted using 50 µl of nuclease-free deionised water instead of Tris-EDTA buffer. This is to avoid the interference of ethylene-diamine-tetra-acetic acid (EDTA) to enzymatic activities in NGS, causing reduced final yield, failure of library prep, or low coverage. DNA products were run on 0.7% agarose gel with 1 kb marker ladder (Vivantis, Malaysia).

3.4 DNA Quantification and Quality Control

The purified DNA was quantified using Qubit® dsDNA High Sensitivity Assay Kit (Thermo Fisher Scientific, USA). The Qubit® working solution was prepared by diluting Qubit® dsDNA HS Reagent 1:200 in Qubit® dsDNA HS Buffer. In each sample tube, 190 µl of Qubit® working solution was added to 10 µl of sample and mixed by vortex for 3 seconds. Mixtures were incubated 2 minutes at room temperature before quantification using Qubit® 2.0 Fluorometer (Thermo Fisher Scientific, USA). The minimum concentration required for each sample was 20 ng/µl.

The quality of DNA samples was ascertained using Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific, USA) by measuring the A260/280 nm and A260/230 nm purity ratios. Nuclease-free deionised water water was used as blank and 1 µl of DNA sample was used for each measurement read. The minimum standard requirement ranges from 1.8 to 2.1 for A260/280 nm ratio and 1.6 to 2.2 for A260/230 nm ratio.

3.5 Genotyping

3.5.1 Taqman® Real-time Polymerase Chain Reaction (PCR)

Sixteen variants within the intron regions from four genes (*ZNF804A, DTNBP1, AKT1* and *TCF4*) were selected from NCBI-GENBANK. The selection criteria included: (i) widely reported in GWAS or case-control

studies; and (ii) reported with high percentage of positive association. Taqman® genotyping was used for previously-reported variants to serve as a marker for identification of causal SNPs in functional regions which is in LD with the marker. The chromosomal position and alleles of the variants were shown in Table 3.2. Genotyping was performed with Taqman® Assays-by-Design SM SNP Genotyping-based assays using ABI PRISM®7900 Sequence Detection System (Thermo Fisher Scientific, USA).

Table 3.2: Selected intronic variants for Taqman® genotyping

Gene	SNP	Position	Major Allele	Minor Allele
ZNF804A	rs12613195	185,489,221	C	G
	rs7597593	185,533,580	A	G
	rs3931790	185,763,376	A	C
	rs1344706	185,778,428	A	C
	rs1583048	185,783,141	C	T
DTNBP1	rs9370822	15,544,736	A	C
	rs9370823	15,550,658	A	G
AKT1	rs3803304	105,239,146	C	G
	rs2494732	105,239,192	T	C
	rs3730358	105,246,407	G	A
	rs1130214	105,259,734	C	A
TCF4	rs12966547	52,752,017	A	G
	rs2958182	53,149,021	A	T
	rs9960767	53,155,002	A	C
	rs10401120	53,192,498	C	T
	rs17512836	53,194,961	C	T

The Taqman® Allele Discrimination Assay (Thermo Fisher Scientific, USA) employs the 5' nuclease activity of *Taq* polymerase during polymerase chain reaction (PCR) to detect a fluorescent reporter signal. The assay includes allele-specific oligonucleotides labelled with different fluorescent probes. During PCR, the 5' nuclease activity cleaves one or both probes to produce a fluorescent reporter signal respective to its associated allele. Both alleles were detected simultaneously and genotypes were determined by the probe ratio.

PCR reactions for each SNP include: 10ng genomic DNA, 900 nM forward and reverse primers, 200 nM fluorescent probes (major allele FAM-labelled; minor allele VIC-labelled) (Thermo Fisher Scientific, USA) and Taqman® Universal PCR Master Mix without AmpErase® UNG (Thermo Fisher Scientific, USA) in a final volume of 10 µl. Amplification was done using QuantStudio™ 12K Flex real-time PCR system (Thermo Fisher Scientific, USA), with 10 minutes of denaturation at 95 ℃; 44 cycles of 15 seconds at 92 ℃ and 1 minute at annealing temperature of 60 ℃.

3.5.2 Targeted Next Generation Sequencing

3.5.2.1 Custom Panel Design

Targeted regions on four schizophrenia and cognition related candidate genes: *ZNF804A*, *AKT1*, *DTNBP1* and *TCF4* were sequenced using Ion AmpliSeq™ Targeted Sequencing (Thermo Fisher Scientific, USA). For the design of the sequencing panel content, the primary aim was to allow for the

detection of possible variations in the functional regions of *ZNF804A*, *AKT1*, *DTNBP1* and *TCF4* genes. Target regions of exons, 3' and 5' UTR (with ± 200 bp) of *ZNF804A*, *AKT1*, *DTNBP1* and *TCF4* genes were selected based on the NCBI (hG19/GRCh37) coordinates. The probes for this targeted custom panel were designed using Ion AmpliSeq™ Designer (Thermo Fisher Scientific, USA) by the manufacturer and ready-to-use prepooled, multiplexed primers were delivered. The custom panel size was 32.55 kb, consisted of 122 amplicons, with a coverage depth of 100X. The amplicon sizes ranged from 125 bp to 375 bp, with 98.59% coverage.

The amount of DNA used was a minimum of 3,000 copies (10 ng) of DNA for each primer pool. Two primer pools were used for this custom panel to cover large target regions. For a 2X Ion AmpliSeq™ primer pool concentration without using Ion AmpliSeq™ Sample ID Panel (Thermo Fisher Scientific, USA), the maximum volume and concentration of DNA per reaction was set to be 6 µl and 1.67 ng/µl respectively.

3.5.2.2 Ion Torrent PGM Library Preparation

An Ion Torrent barcode adapter-ligated library was generated using 1 µg of genomic DNA using Ion AmpliSeq™ Library Kit 2.0 (Thermo Fisher Scientific, USA). For each sample DNA target amplification using a 2X primer pool, 4 µl of 5X Ion AmpliSeq™ HiFi Mix (Thermo Fisher Scientific, USA), 10 µl of 2X Ion AmpliSeq™ Primer Pool (Thermo Fisher Scientific,

USA), and 10 ng of DNA were mixed to a final volume of 20 µl. The reaction mixture was vortexed thoroughly and spun down before proceeding to amplification using Veriti™ Thermal Cycler (Thermo Fisher Scientific, USA), following the program: holding at 99 ℃ for 2 minutes for enzyme activation; denaturation at 99 ℃ for 15 seconds and 4 minutes annealing and extension at 60 ℃ for 20 cycles. After amplification, 2 µl of FuPa Reagent (Thermo Fisher Scientific, USA) was added to each amplified sample, giving a total volume of 22 µl. The reaction mixture was vortexed thoroughly and span down before partial digestion by thermal cycler (Thermo Fisher Scientific, USA): 50 ℃ for 10 minutes; 55 ℃ for 10 minutes; followed by 60 ℃ for 20 minutes.

The next step in library preparation was adapter ligation to the amplicons. A barcode adapter mix was prepared prior to the ligation. The barcode adapter mix for two reactions was prepared using: 1 µl of Ion P1 Adapter (Thermo Fisher Scientific, USA), 1 µl of the chosen Ion Xpress™ Barcode X (Thermo Fisher Scientific, USA) and 2 µl of nuclease-free water to a total volume of 4 µl. For each reaction, 2 µl of the barcode adapter mix; 4 µl of Switch Solution (Thermo Fisher Scientific, USA); and 2 µl of DNA ligase (Thermo Fisher Scientific, USA) were added to the 22 µl previously digested amplicon, giving a total volume of 30 µl. The reaction mixture was vortexed thoroughly and span down before running in thermal cycler (Thermo Fisher Scientific, USA) for 22 ℃ for 30 minutes followed by 72 ℃ for 10 minutes.

3.5.2.3 Library Quantification and Quality Control

Size estimation and quality of the library were determined using Agilent 2100 Bioanalyzer (Agilent Technologies, USA) together with Agilent High Sensitivity DNA Kit (Agilent Technologies, USA). Firstly, the High Sensitivity DNA dye concentrate (Agilent Technologies, USA) and High Sensitivity DNA gel matrix (Agilent Technologies, USA) were equilibrated to room temperature for 30 minutes. A total of 15 µl High Sensitivity DNA dye concentrate was added to the High Sensitivity DNA gel matrix vial to form a gel-dye mixture. The mixture was vortexed for 10 seconds and spun down before transferred to a spin filter. It was then centrifuged at 2,240 xg for 10 minutes. The mixture was protected from light and stored at 4 ℃.

The gel-dye mixture was equilibrated 30 minutes at room temperature before loading. A High Sensitivity DNA chip (Figure 3.1) (Agilent Technologies, USA) was positioned on the chip priming station. A total of 9 µl gel-dye mix was pipetted at the bottom of the black 'G' well. The plunger was positioned at 1 ml, pressed until held by the clip and after 60 seconds, the plunger and clip was released. After another 5 seconds, the plunger was then pulled back to the 1 ml position and the process was repeated in other wells.

Then, 5 µl of High Sensitivity DNA marker (Agilent Technologies, USA) was pipetted into the ladder well and all sample wells. This was followed by loading 1 µl of High Sensitivity DNA ladder (Agilent Technologies, USA) in the ladder well. Then, 1 µl of sample was pipetted in

all of the sample wells. The chip was inserted horizontally in the adapter of an IKA MS3 vortex mixer (IKA®, Germany) and vortexed for 1 min at 2400 rpm. The chip was inserted and processed in the Agilent 2100 Bioanalyzer (Agilent Technologies, USA) within 5 minutes. Results were analysed using 2100 Expert Software (Agilent Technologies, USA).

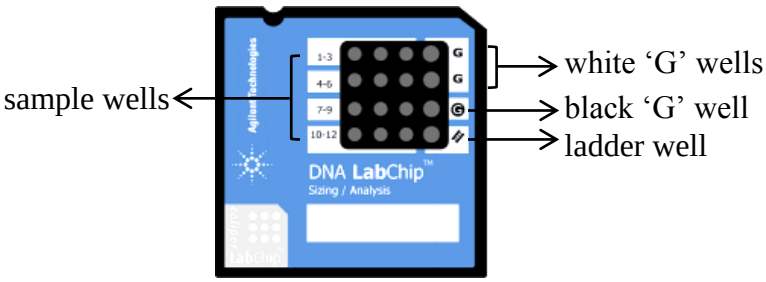


Figure 3.1: High Sensitivity DNA Bioanalyzer chip

3.5.2.4 Ion Torrent PGM Sequencing

The DNA library from all 26 patients and 22 controls were normalised to 100 pM using deionized water before combined in a pool with the same amount of DNA. A total of 50 ng pooled amplicons were ligated using Ion Torrent adapters P1 and A (Thermo Fisher Scientific, USA), followed by purification using AMPure beads (Beckman Coulter, USA). After PCR amplification for 10 cycles, the adapter-ligated library was purified again using AMPure beads. The concentration and quality of the amplicon was determined using Qubit DNA HS Assay Kit (Thermo Fisher Scientific, USA) and Agilent BioAnalyzer DNA High-Sensitivity LabChip (Agilent Technologies, USA), according to the manufacturer's instructions.

Then, the amplicons were pooled in equimolar concentrations and Ion PGM™ Template OT2 400 Kit (Thermo Fisher Scientific, USA) was used for sample emulsion PCR and enrichment. The input concentration of one DNA template copy per Ion Sphere Particles (ISPs) (Thermo Fisher Scientific, USA) was mixed to the emulsion PCR master mix (Thermo Fisher Scientific). After ISPs recovery, template-positive ISPs were enriched using Dynabeads MyOne Streptavidin C1 beads (Thermo Fisher Scientific, USA) and their quantity confirmed using the Qubit® 2.0 Fluorometer (Thermo Fisher Scientific, USA).

One hour before performing sequencing run, the Ion PGM system was cleaned and initialised according to manufacturer's instructions. Meanwhile, Ion PGM™ Sequencing 400 Kit (Thermo Fisher Scientific, USA) was used to prepare the sample for sequencing. Firstly, 15 µl of ISPs were annealed with 12 µl Sequencing Primers (Thermo Fisher Scientific, USA) using thermal cycler (Thermo Fisher Scientific, USA) for 95 ℃ for 2 minutes followed by 37 ℃ for 2 minutes. Meanwhile, Chip Check was done on an Ion 316^{TM} chip (Figure 3.2) (Thermo Fisher Scientific, USA) prior to sample loading following manufacturer's instructions.

The ISPs was then removed from the thermal cycler and 3 µl of Ion PGM™ Sequencing 400 Polymerase (Thermo Fisher Scientific, USA) was added to the ISPs, for a total final volume of 30 µl. The sample was mixed and incubated at room temperature for 5 minutes before gently and slowly loaded into the Ion 316^{TM} chip to avoid introduction of bubbles into the chip. The chip was centrifuged for 30 seconds using MiniFuge (Thermo Fisher Scientific,

USA) and the sample in the chip was mixed by means of pipetting the sample in and out three times without removing the tip. The chip centrifuge and mixing steps were repeated twice followed by removing all liquid from the chip. After that, the chip was sequenced on the Ion Torrent PGM sequencing system (Thermo Fisher Scientific, USA) following manufacturer's instructions.

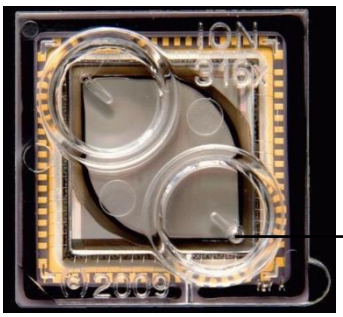


Figure 3.2: Ion 316^{TM} chip

3.5.2.5 Sequence Analysis

Raw sequence reads were extracted from the PGM system in FASTQ format. Data were analysed on the Ion Torrent platform-specific pipeline software Torrent Suite v1.3.1 (Thermo Fisher Scientific, USA) to generate sequence reads, trim adapters, and remove poor signal-profile reads. The trimmed sequences were aligned against the genomic reference sequence from NCBI (hG19/GRCh37) and targeted variants were detected. Ion Torrent Server v4.2 (Thermo Fisher Scientific, USA), Variant Caller and Coverage Analysis plugins (Thermo Fisher Scientific, USA) were used for variants annotation and Integrative Genomics Viewer (Robinson *et al.*, 2011) were used for reads visualization. Several filters were applied to identify high-

quality variant calling, such as a minimum coverage of 20x and at least 5% of reads must be detected before variant identification.

3.6 Validation

Selected NGS variant calls were validated against Sanger sequencing. Variants selected were based on the following self-designed rule-of-thumb selection criteria: (i) no reference SNP cluster ID (rs number); (ii) variant occurrence in more than two subjects; (iii) only found in patient or control; or (iv) having a significant association between patient and control.

Primer pairs sequences flanking each variant were obtained from Ion Reporter server (Thermo Fisher Scientific, USA) or designed using Primer3 software version 4.0.0 (http://primer3.sourceforge.net/) (Untergasser *et al.*, 2012). For Primer3, the conditions for primer design were: primer length of 18 bp to 22 bp, melting temperature between 57 ℃ to 60 ℃, and GC% of 40% to 60%. Sequences that best fit the selection criteria were selected (Table 3.3).

Variants were amplified using PCR method. PCR reaction for each variant included 20 ng genomic DNA, 1 U *Taq* polymerase (Vivantis, Malaysia), 1X PCR buffer (Vivantis, Malaysia), 1 µM each of forward and reverse primers (First BASE Laboratories, Malaysia), 0.2 mM dNTP mix (Vivantis, Malaysia), 1.5 mM $MgCl_2$ (Vivantis, Malaysia) and topped up with distilled water to a final volume of 25 µl. The amplification was performed

using Veriti™ Thermal Cycler (Thermo Fisher Scientific, USA) following the protocol: 5 minutes of initial denaturation at 94 ℃; followed by 35 cycles of (i) denaturation at 94 ℃ for 30 seconds, (ii) annealing at primer-specific annealing temperature for 30 seconds, and (iii) extension at 72 ℃ for 30 seconds; and lastly a final extension of 10 minutes at 72 ℃.

Following amplification, size of the PCR product was verified by running on 2% agarose gel with VC 100 bp plus ladder (Vivantis, Malaysia) or MassRuler™ low range DNA ladder (Thermo Fisher Scientific, USA) as reference. Products with a single clean band and proper size were outsourced for Sanger sequencing at First BASE Laboratories, Malaysia, using BigDye® Terminator v3.1 (Thermo Fisher Scientific, USA) cycle sequencing kit chemistry.

Table 3.3: Primer sequences for validation

Position	Size (bp)	Ta (^{o}C)	Forward	Reverse
***ZNF804A*:**				
185,801,596	231	56	TGAGCCATTTGTACCTGTCCT	TCTGAAAGTGGTCCCATGCA
185,802,542	156	57	CGAAAACGTAGACAACATTCACA	GGGGTGGAATCTGCCTCTTT
***DTNBP1*:**				
15,522,879	326	57	ACAGGAAGTGGCATTTTCCACT	CTGGAATCCAGTTTTGGCTGTATG
15,522,988	326	57	ACAGGAAGTGGCATTTTCCACT	CTGGAATCCAGTTTTGGCTGTATG
15,523,101	326	57	ACAGGAAGTGGCATTTTCCACT	CTGGAATCCAGTTTTGGCTGTATG
15,638,037	184	56	ACTTGCACAGTCTTTGGCTC	CCTTGCAACATGTCCCCATT

(Table 3.3 continued)

***AKT1*:**				
105,235,818	244	58	CACTTGACCTTTTCGACGCT	CCCAGCAAAGTGAGCCAAAA
105,235,824	244	58	CACTTGACCTTTTCGACGCT	CCCAGCAAAGTGAGCCAAAA
105,235,825	244	58	CACTTGACCTTTTCGACGCT	CCCAGCAAAGTGAGCCAAAA
105,235,860	244	58	CACTTGACCTTTTCGACGCT	CCCAGCAAAGTGAGCCAAAA
105,236,287	320	57	CAATAACAAATTTAAACCTTGCTCCTCTGT	AATCCGATTCACGTAGGGAAATGT
105,236,557	208	58	CGCATAGCTGCAGAAGTCCTTA	GCCTCGTGCCATGATCTGTATTT
***TCF4*:**				
52,893,255	280	56	CAGTTTTTCAACACACACACAGCTA	CTGTTCTCCTGTTTTGCCAGTTATATAGTA
52,893,680	291	56	TGCCAAGAATTGTGTGGCTGTA	GCCAATATTCAATCATCATGCAGCATTATA
53,252,388	305	56	ATCGCACAAGATAACATTTTGTGTAAGT	ACCCTTAGCATAAACTCTAAGCTGTTTG

* bp: base pair; Ta: annealing temperature

3.7 Statistical Analyses

Scores of trail making tests were compared between patients and controls using independent *t*-test. Allelic and genotype frequency differences between patient and control families were analysed using Fisher's exact test in Statistical Package for the Social Sciences (SPSS) for Windows (SPSS Inc., Chicago). Linkage disequilibrium (LD) was assessed using Haploview (http://www.broad.mit.edu/mpg/haploview). Significant *p*-value was set at $p < 0.05$ for all statistical tests.

All results were corrected using false discovery rate (FDR) (Benjamini and Hochberg, 1995) to minimize false positive rate, and only *p*-value that is significant after correction were highlighted. In FDR correction, an excel datasheet with pre-set FDR formula was used to enter the *p*-values obtained by multiple tests of the same data together with the desired level of significance, in this case $p = 0.05$. The calculation will be done based on the *p*-values entered to generate a 'corrected' significance level, which is used as the threshold for result significance.

3.8 *In-Silico* Variant Functional Effect Prediction

Functional effect prediction for variants that are significantly different between patients and controls; patient-specific; or have high potential as schizophrenia risk variant were done using Variant Effect Predictor

(http://grch37.ensembl.org/Homo_sapiens/Tools/VEP) and SNP Nexus (http://snp-nexus.org/) while 3' UTR variants functions were predicted using MirSNP (http://bioinfo.bjmu.edu.cn/mirsnp/search/). Phyre2 (http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index) was used for protein homolog and folding prediction. SIFT dbSNP (http://sift.jcvi.org/www/SIFT_dbSNP.html) was used for protein prediction of non-synonymous variant. GOR IV (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_gor4.html) was used for protein secondary structures prediction.

CHAPTER 4

RESULTS

4.1 Cognitive Performance

On average, the cognitive performances of the patients fell below normal standard (Table 4.1). To complete TMT-A, patients spent an average of 51.07 ± 23.34 seconds, falling in the range between the standard average of 29 seconds and deficient average of 78 seconds. Similarly, they obtained a below average score of 135.60 ± 92.13 seconds (standard: 75 seconds) in TMT-B. There were significant differences between healthy controls and patients on TMT-A ($p = 4.08 \times 10^{-5}$) and TMT-B ($p = 1.96 \times 10^{-4}$) scores. The mean results showed that the cognitive performance of the patients was below par, although they were not considered as cognitive deficit.

From the derived scores, B-A difference score is the relative indicator of executive control function omitting the influence of speed factor. Controls fared a significantly ($p = 1.68 \times 10^{-3}$) better B-A score compared to patient, hence better executive control function compared to patients. Controls average B–A score remained similar to their TMT-A score, while patients' average B–A score was 65% higher than their TMT-A score. On the other hand, both B:A ($p = 0.046$) and (B-A):A ($p = 0.046$) scores between patients and controls were not significant after FDR correction. The slightly higher B:A score of patients

showed poor task-switching skill while high (B-A):A reflected abnormality in prefrontal cortex functions. No significant difference was found in log B:A score (p = 0.093) between patient and control.

4.2 DNA Isolation

The product size of extracted DNA was approximately 20 kb on a 0.7% agarose gel electrophoresis (Figure 4.1). The concentration of DNA obtained from Qubit® 2.0 Fluorometer (Thermo Fisher Scientific, USA) reading ranged between 22 to 76 ng/µl.

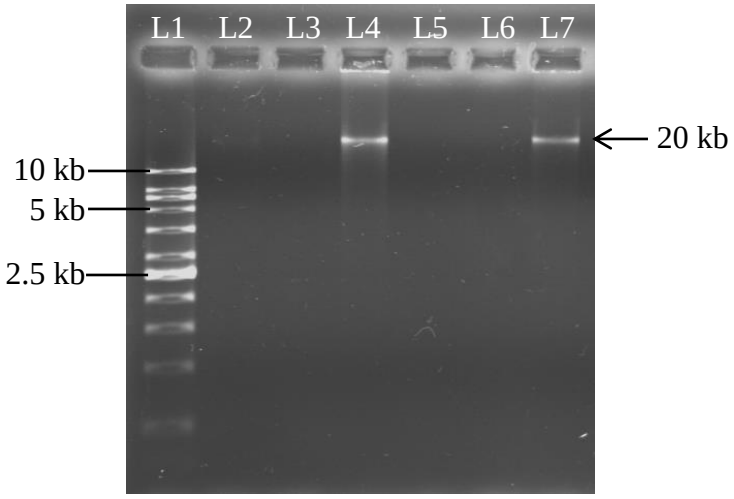


Figure 4.1: Genomic DNA extraction product. Lane 1: VC 1 kb DNA ladder (Vivantis, Malaysia); Lane 4 and 7: genomic DNA

Table 4.1: Descriptive statistics of direct (TMT-A and TMT-B) and derived (B-A, B:A, log B:A and (B-A):A) TMT scores

Subject	TMT scores					
	TMT-A	**TMT-B**	**B-A**	**B:A**	**log B:A**	**(B-A):A**
Patient						
Mean	51.07	135.60	84.54	2.713	0.387	1.713
SD	23.34	92.13	79.93	1.285	0.206	1.286
Min-Max	22 – 115	43 – 430	3 – 349	1.04 – 5.31	0.02 – 0.73	0.04 – 4.31
Control						
Mean	27.48	55.59	28.11	2.105	0.300	1.104
SD	9.63	23.70	20.81	0.726	0.144	0.725
Min-Max	13 – 50	25 – 114	3 -78	1.09 – 3.89	0.04 – 0.59	0.09 – 2.89
***p*-value**	**4.08×10^{-5}***	**1.96×10^{-4}***	**1.68×10^{-3}***	0.046	0.093	0.046

* Significant after FDR correction (SD: standard deviation; Min: minimum score; Max: maximum score)

4.3 Library Quantification and Quality Control

The quality of the barcode adapter-ligated library was determined using Agilent 2100 Bioanalyzer (Agilent Technologies, USA). Individual electropherogram was obtained for each sample. Example of an electropherogram is shown in Figure 4.2.

As seen in the Figure 4.2, both lower and upper markers showed a single peak at 43bp and 11.3kb respectively. All of the bands were detected within the markers range. Both lower and upper markers were used for the alignment or sizing of the bands, while the upper marker was used for band quantitation. The electropherogram profiles were the similar for all of the samples with little or no noise at the base line.

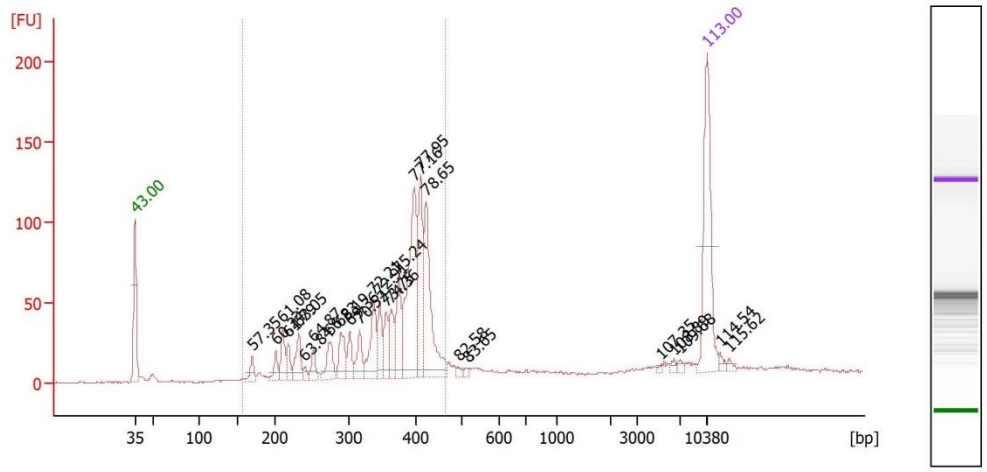


Figure 4.2: Electropherogram of NGS library, showing peak intensity in fluorescence unit (FU) against DNA size in base pair (bp). Lower marker peak indicated at 43 bp and upper marker peak at 11.3 kb

4.4 Association Analyses

For NGS sequencing, the coverage depth of the 48 samples ranged between 312x to 1007x, which is 3x to 10x above expected coverage. A total of 117 significant variants were identified from Taqman® and NGS analysis. Among them, 33 variants were observed in patients only (1 *ZNF804A* variant, 9 *DTNBP1* variants, 16 *AKT1* variants, and 7 *TCF4* variants); 22 variants observed in controls only (6 *ZNF804A* variants, 4 *DTNBP1* variants, 2 *AKT1* variants, and 10 *TCF4* variants); while 62 variants were found in both patients and controls (12 *ZNF804A* variants, 7 *DTNBP1* variants, 16 *AKT1* variants, and 27 *TCF4* variants). For the latter, association test of genotype and allele frequencies showed that variants from *ZNF804A* (rs12613195, rs7597593) and *TCF4* (rs2958182) were significantly different between patients and controls after FDR corrections (data shown in respective gene subsections).

In the current study, risk allele is defined as the allele which is found significantly higher in patients compared to controls. Major alleles are ancestral alleles which are common among similar population, while minor alleles are the second most common allele that occurs in a given population, most of the time serves as the risk allele.

4.4.1 *ZNF804A*

Table 4.2 shows the allelic and genotypic frequencies of variants detected in *ZNF804A*. Rs12613195 ($p = 1.52 \times 10^{-7}$) and rs7597593 ($p = 0.002$) were found to have significantly different genotype frequencies in patients compared to controls after FDR correction. Figure 4.3 shows the LD structure of the *ZNF804A* variants detected in patients. Three haploblocks were detected: one among the 5' variants and two in the intron regions. The 5' haploblock consisted of four variants, among which two of the variants marked statistical significance, indicating that all variants in the block are in linkage disequilibrium with each other and are inherited together. The other two intronic haploblock did not have statistically significant variants.

Figure 4.4 shows PCR products of rs75132823 and rs78607324, which were selected for validation using Sanger sequencing. The sequencing chromatograms were shown in Table 4.3. In functional predictions, rs359895 was predicted to be a transcription factor binding site and rs5836928 as a location where peptide shift occur. Nine variants: rs12476147, rs35676856, rs61739290, rs4667001, rs1366842, rs79776875, rs12477430, rs78607324, and rs3731834 were predicted to be nonsynonymous variants; while another three exonic variants: rs75132823, rs4667002, and rs728534 were synonymous substitutions (Table 4.4).

Table 4.2: Variants detected in *ZNF804A*

No.	Position	SNP ID	Location	Type	ref/variant (1/2)	Genotype frequency (11/12/22)		*p*-value	Allele frequency (1/2)		*p*-value
						Patient	Control		Patient	Control	
1	185,463,120	rs2279266	5' UTR	SNV	C/T	61.54/38.46/0.00	40.91/40.91/18.18	0.059	80.77/19.23	61.36/38.64	**0.042**
2	185,463,185	rs359895	5' UTR	SNV	T/A	57.69/42.31/0.00	77.27/13.64/9.09	**0.024**	78.85/21.15	84.09/15.91	0.604
3	185,463,430	rs370931594	5' UTR	SNV	C/T	84.62/15.38/0.00	100.00/0.00/0.00	-	92.31/7.69	100.00/0.00	-
4	185,463,535	rs73041379	5' UTR	SNV	G/A	53.85/46.15/0.00	77.27/18.18/4.55	0.065	76.92/23.08	86.36/13.64	0.299
5	185,489,221	rs12613195	intron	SNV	C/G	7.69/76.92/15.38	36.36/40.91/22.73	**1.52 x 10^{-7}***	46.15/53.85	56.82/43.18	0.157
6	185,533,580	rs7597593	intron	SNV	A/G	65.38/26.92/7.69	81.82/18.18/0.00	**0.002***	78.85/21.15	90.91/9.09	**0.028**
7	185,731,054	rs530610969	intron	SNV	G/T	100.00/0.00/0.00	95.45/4.55/0.00	-	100.00/0.00	97.73/2.27	-
8	185,763,376	rs3931790	intron	SNV	A/C	50.00/34.62/15.38	50.00/40.91/9.09	0.392	67.31/32.69	70.45/29.55	0.761
9	185,778,428	rs1344706	intron	SNV	C/A	23.08/61.54/15.38	27.27/45.45/27.27	**0.039**	53.85/46.15	50.00/50.00	0.671
10	185,783,141	rs1583048	intron	SNV	C/T	15.38/34.62/50.00	9.09/40.91/50.00	0.392	32.69/67.31	29.55/70.45	0.761
11	185,798,504	rs4667000	intron	SNV	A/C	0.00/38.46/61.54	0.00/22.73/77.27	0.351	19.23/80.77	11.36/88.64	0.400
12	185,800,905	rs12476147	exon	SNV	A/T	0.00/34.62/65.38	0.00/13.64/86.36	0.180	17.31/82.69	6.82/93.18	0.214
13	185,801,559	rs35676856	exon	SNV	A/G	100.00/0.00/0.00	95.45/4.55/0.00	-	100.00/0.00	97.73/2.27	-
14	185,801,596	rs75132823	exon	SNV	C/T	100.00/0.00/0.00	81.82/18.18/0.00	-	100.00/0.00	90.91/9.09	-
15	185,801,597	rs61739290	exon	SNV	A/G	100.00/0.00/0.00	95.45/4.55/0.00	-	100.00/0.00	97.73/2.27	-
16	185,801,747	rs4667001	exon	SNV	G/A	0.00/34.62/65.38	4.55/13.64/81.82	0.138	17.31/82.69	11.36/88.64	0.564

(Table 4.2 continued)

17	185,801,755	rs4667002	exon	SNV	A/G	0.00/34.62/65.38	4.55/9.09/86.36	0.059	17.31/82.69	9.09/90.91	0.370
18	185,801,917	rs728534	exon	SNV	A/G	53.85/42.31/3.85	68.18/22.73/9.09	0.404	75.00/25.00	79.55/20.45	0.634
19	185,802,211	rs5836928	exon	IN	-/ACA	0.00/34.62/65.38	4.55/13.64/81.82	0.138	17.31/82.69	11.36/88.64	0.564
20	185,802,243	rs1366842	exon	SNV	C/A	0.00/34.62/65.38	4.55/13.64/81.82	0.138	17.31/82.69	11.36/88.64	0.564
21	185,802,335	rs79776875	exon	SNV	A/G	100.00/0.00/0.00	95.45/4.55/0.00	-	100.00/0.00	97.73/2.27	-
22	185,802,363	rs12477430	exon	SNV	A/G	0.00/34.62/65.38	4.55/9.09/86.36	0.059	17.31/82.69	9.09/90.91	0.370
23	185,802,542	rs78607324	exon	SNV	A/C	100.00/0.00/0.00	90.91/9.09/0/00	-	100.00/0.00	95.45/4.55	-
24	185,803,364	rs3731834	exon	SNV	C/G	57.69/34.62/7.69	81.82/4.55/13.64	0.034	75.00/25.00	84.09/15.91	0.321

* Significant after FDR correction (ref: reference; 1: allele 1 (reference allele); 2: allele 2 (variant); SNV: single nucleotide variant; IN: insertion)

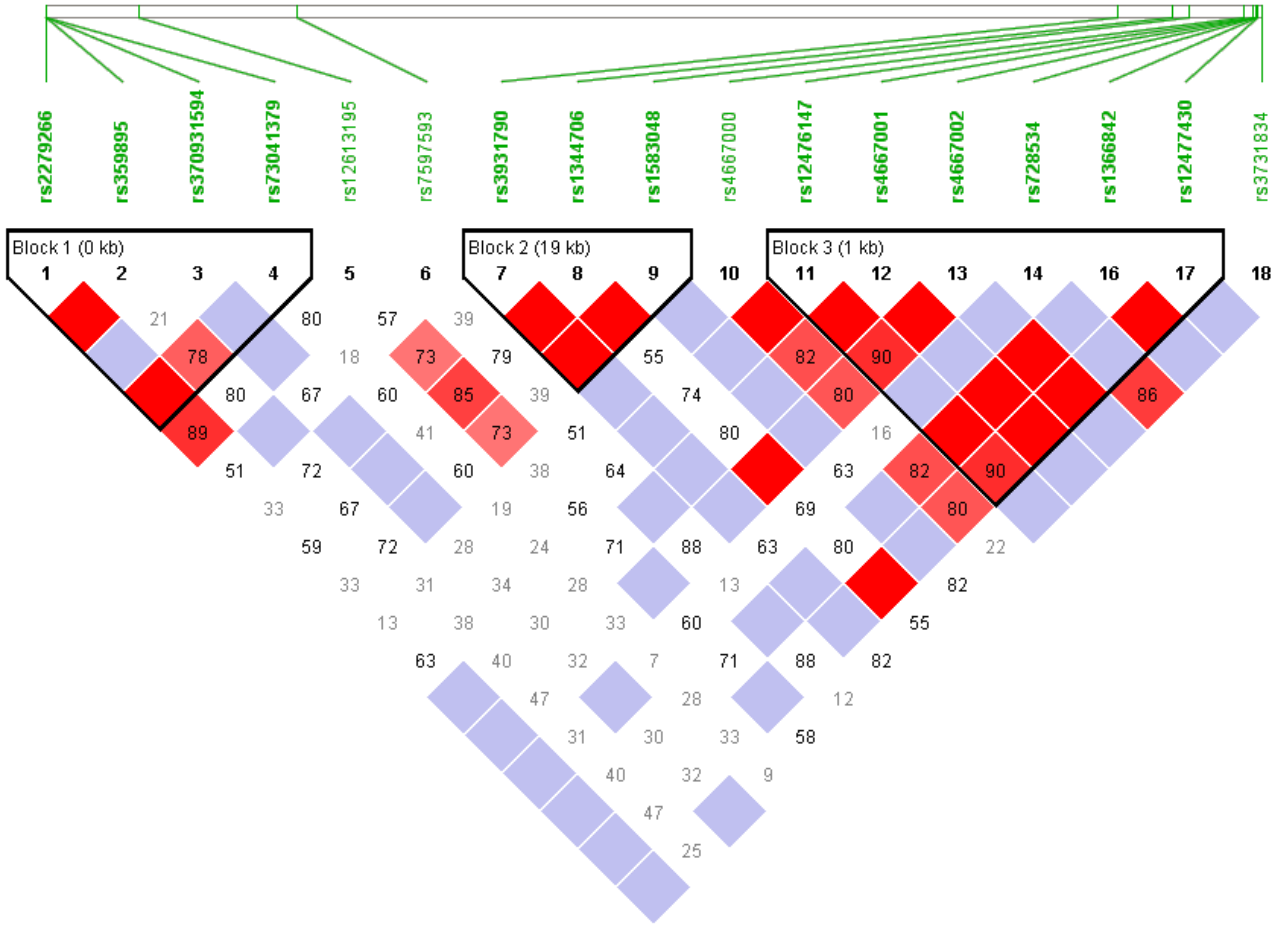


Figure 4.3: LD structure of *ZNF804A* variants detected in patients, showing three haploblocks

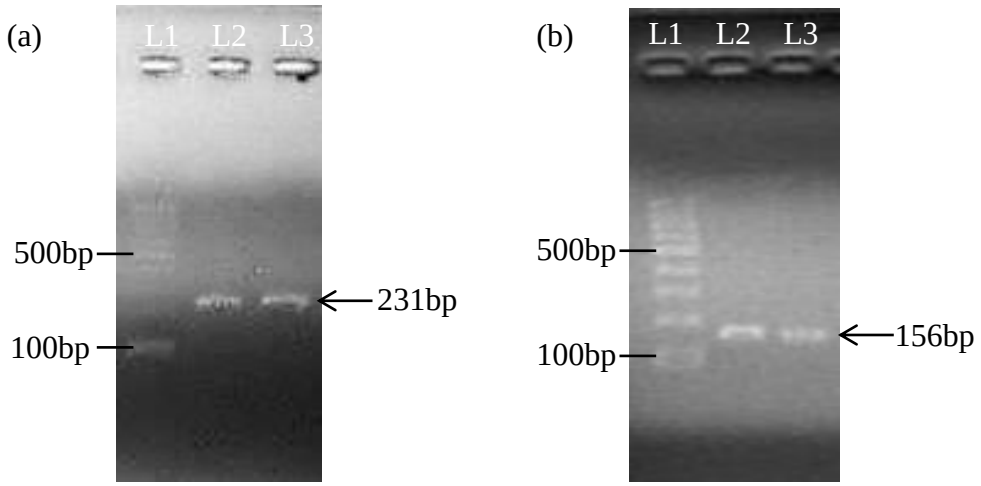


Figure 4.4: Agarose gel electrophoresis of selected *ZNF804A* variants. (a) Rs75132823, Lane 1: VC 100 bp Plus DNA ladder (Vivantis, Malaysia), Lane 2 to 3: PCR products; (b) rs78607324, Lane 1: VC 100 bp DNA ladder (Vivantis, Malaysia), Lane 2 to 3: PCR products

Table 4.3: Validation of selected variants in *ZNF804A*

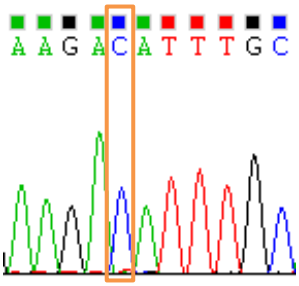
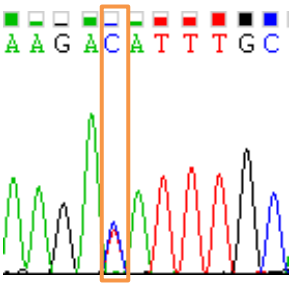
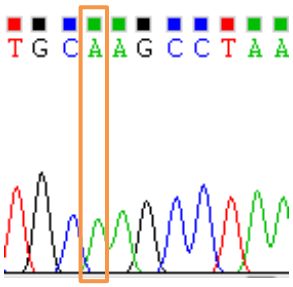
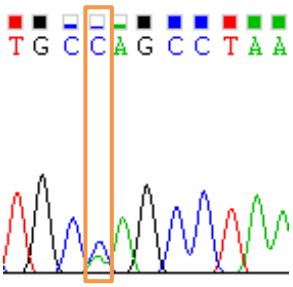
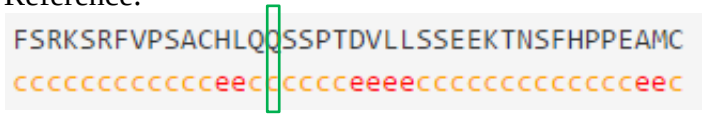
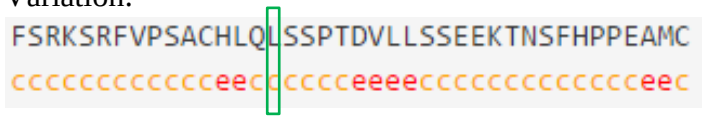
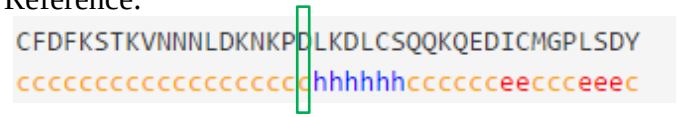
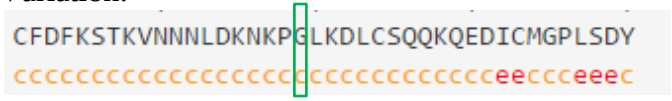
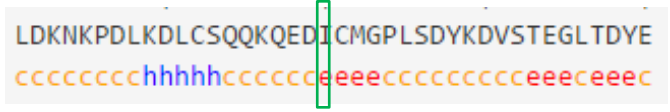
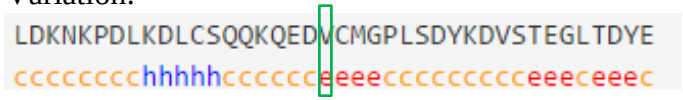
Variant position	Allele	Reference	Variation
185,801,596 (rs75132823)	C>C/T		
185,802,542 (rs78607324)	A>A/C		

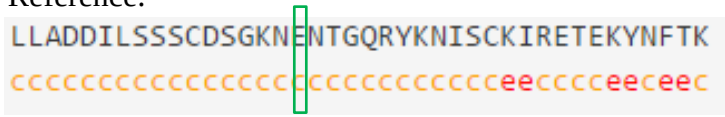
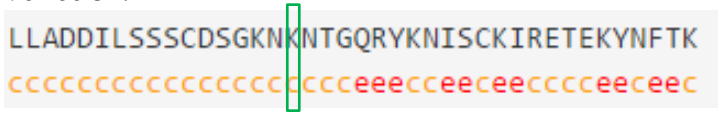
Table 4.4: *In-silico* variant functional prediction of *ZNF804A* variants

Position (SNP ID)	Variant Type	Ref	Function	Effect*
185,463,185 (rs359895)	SNV (5' UTR)	T	Transcription factor binding site for nuclear factor 1 (NF-1) (P08651) and nuclear transcription factors: CP1A (P25208), CP1C (Q13952), NF-YA (P23511)	unknown
185,800,905 (rs12476147)	SNV (exon)	A	Nonsynonymous nucleotide mutation. Amino acid change from Glutamine (Q) to Leucine (L) at position 261	Different protein family but tolerated mutation Reference:  Variation: 

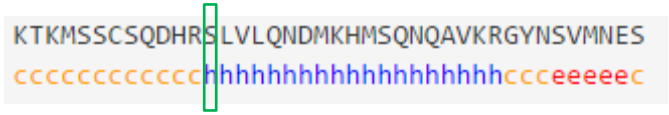
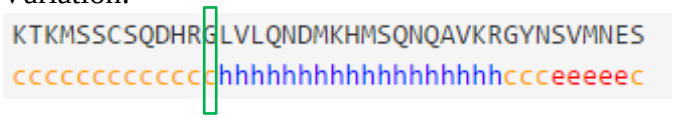
(Table 4.4 continued)

185,801,559 (rs35676856)	SNV (exon)	A	Nonsynonymous nucleotide mutation. Amino acid change from Aspartic acid (D) to Glycine (G) at position 479	Protein folding to change from alpha helix to supercoil Reference:  Variation: 
185,801,596 (rs75132823)	SNV (exon)	C	Synonymous nucleotide mutation. Amino acid remain as Aspartic acid (D) at position 491	
185,801,597 (rs61739290)	SNV (exon)	A	Nonsynonymous nucleotide mutation. Amino acid change from Isoleucine (I) to Valine (V) at position 492	Tolerated mutation. No protein folding change within same amino acid group Reference:  Variation: 

(Table 4.4 continued)

185,801,747 (rs4667001)	SNV (exon)	G	Nonsynonymous nucleotide mutation. Amino acid change from Glutamic acid (E) to Lysine (K) at position 542	Same protein family but tolerated mutation. Protein folding to change from supercoil to extended strand Reference:  Variation: 
185,801,755 (rs4667002)	SNV (exon)	A	Synonymous nucleotide mutation. Amino acid remain as Threonine (T) at position 544	
185,801,917 (rs728534)	SNV (exon)	A	Synonymous nucleotide mutation. Amino acid remain as Lysine (K) at position 598	
185,802,211 (rs5836928)	IN (exon)		Amino acid Threonine (T) added at position 697	Peptide shift occurs

(Table 4.4 continued)

185,802,243 (rs1366842)	SNV (exon)	C	Nonsynonymous nucleotide mutation. Amino acid change from Threonine (T) to Lysine (K) at position 707	Different protein family but tolerated mutation
185,802,335 (rs79776875)	SNV (exon)	A	Nonsynonymous nucleotide mutation. Amino acid change from Serine (S) to Glycine (G) at position 738	Protein folding to change from alpha helix to supercoil Reference:  Variation: 
185,802,363 (rs12477430)	SNV (exon)	A	Nonsynonymous nucleotide mutation. Amino acid change from Histidine (H) to Arginine (R) at position 747	Same amino acid family, tolerated mutation

(Table 4.4 continued)

185,802,542 (rs78607324)	SNV (exon)	A	Nonsynonymous nucleotide mutation. Amino acid change from Lysine (K) to Glutamine (Q) at position 807	No protein folding change Reference: LNRQNHLPEEFLRPPSTSVAPCKPKKKRRRKRGRFHPGFETLELKE cccccccccccccccccccccccccccccccccccccccceeeeeec Variation: LNRQNHLPEEFLRPPSTSVAPCQPKKKRRRKRGRFHPGFETLELKE cccccccccccccccccccccccccccccccccccccccceeeeeec
185,803,364 (rs3731834)	SNV (exon)	C	Nonsynonymous nucleotide mutation. Amino acid change from Leucine (L) to Valine (V) at position 1081	Same family, tolerated mutation Reference: LQPNMLANKVKFTFPPAALPPPSTPLQPLPLQQSLCSTSVTTIHHTVLQQ cccccccccccccccccccccccccccccccccccccccceeeeceeeecc Variation: LQPNMLANKVKFTFPPAALPPPSTPVQPLPLQQSLCSTSVTTIHHTVLQQ cccccccccccccccccccccccccccccccccccccccceeeeceeeecc

* Protein prediction using GOR IV (h: alpha helix; e: extended strand; c: random coil; Ref: reference allele; SNV: single nucleotide variant; IN: insertion)

4.4.2 *DTNBP1*

Table 4.5 shows the allelic and genotypic frequencies of variants detected in *DTNBP1*, and none scored statistical significance after FDR correction. Figure 4.5 shows PCR products of four variants (KX619615, rs200695686, rs1047631, and rs140124163) selected for validation using Sanger sequencing and Table 4.6 shows the sequencing chromatograms. The LD structure of variants detected in patients shows two haploblocks in the intron region (Figure 4.6). Functional predictions predicted rs1047631 to be a microRNA (miRNA)-messenger RNAs (mRNA) binding site, and rs144524387 as a location for stop-gained mutation. Rs140124163 was predicted as a 5' splice variant for intron 3 and rs11558324 was predicted to affect protein coding of exon 1 (Table 4.7).

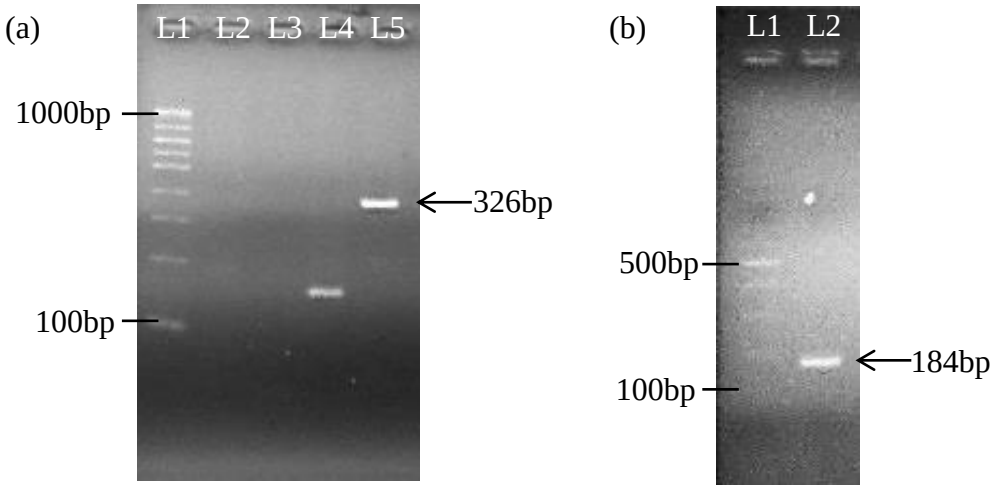


Figure 4.5: Agarose gel electrophoresis of selected *DTNBP1* variants: (a) KX619615, rs200695686, and rs1047631, Lane 1: VC 100 bp DNA ladder (Vivantis, Malaysia), Lane 5: PCR product; (b) rs140124163, Lane 1: VC 100 bp Plus DNA ladder (Vivantis, Malaysia), Lane 2: PCR product

Table 4.5: Variants detected in *DTNBP1*

No.	Position	SNP ID	Location	Type	ref/variant (1/2)	Genotype frequency (11/12/22)		*p*-value	Allele frequency (1/2)		*p*-value
						Patient	Control		Patient	Control	
1	15,522,879	KX619615*	downstream	SNV	C/T	100.00/0.00/0.00	90.91/9.09/0/00	-	100.00/0.00	95.45/4.55	-
2	15,522,988	rs200695686	downstream	DEL	AT/-	96.15/3.85/0.00	100.00/0.00/0.00	-	98.08/1.92	100.00/0.00	-
3	15,523,101	rs1047631	3' UTR	SNV	T/C	84.62/15.38/0.00	100.00/0.00/0.00	-	92.31/7.69	100.00/0.00	-
4	15,523,806	rs34782642	intron	MNV	TT/AA	100.00/0.00/0.00	95.45/4.55/0.00	-	100.00/0.00	97.73/2.27	-
5	15,523,807	rs55944541	intron	SNV	T/A	80.77/19.23/0.00	72.73/27.27/0.00	0.732	90.38/9.62	86.36/13.64	0.749
6	15,524,064	rs766437357	intron	SNV	C/T	92.31/7.69/0.00	100.00/0.00/0.00	-	96.15/3.85	100.00/0.00	-
7	15,524,151	rs374546569	intron	SNV	G/A	96.15/3.85/0.00	100.00/0.00/0.00	-	98.08/1.92	100.00/0.00	-
8	15,524,480	rs742106	intron	SNV	G/A	7.69/50.00/42.31	9.09/50.00/40.91	1.000	32.69/67.31	34.09/65.91	1.000
9	15,524,599	rs117610176	intron	SNV	A/G	92.31/7.69/0.00	100.00/0.00/0.00	-	96.15/3.85	100.00/0.00	-
10	15,524,715	rs144524387	intron	SNV	G/A	96.15/3.85/0.00	100.00/0.00/0.00	-	98.08/1.92	100.00/0.00	-
11	15,533,464	rs534257582	intron	SNV	C/G	96.15/3.85/0.00	100.00/0.00/0.00	-	98.08/1.92	100.00/0.00	-
12	15,544,736	rs9370822	intron	SNV	A/C	15.38/50.00/34.62	9.09/40.91/50.00	0.079	40.38/59.62	29.55/70.45	0.182
13	15,550,658	rs9370823	intron	SNV	A/G	19.23/57.69/23.08	9.09/59.09/31.82	0.082	48.08/51.92	38.64/61.36	0.254
14	15,593,240	rs7758659	intron	SNV	T/C	0.00/3.85/96.15	0.00/0.00/100.00	1.000	1.92/98.08	0.00/100	1.000
15	15,593,399	rs552812050	intron	SNV	T/C	100.00/0.00/0.00	95.45/4.55/0.00	-	100.00/0.00	97.73/2.27	-
16	15,615,637	rs3829893	intron	SNV	G/A	53.85/34.62/11.54	45.45/54.55/0.00	0.202	71.1/28.85	72.73/27.27	1.000

(Table 4.5 continued)

17	15,627,553	rs16876738	intron	SNV	C/G	53.85/34.62/11.54	45.45/54.55/0.00	0.202	71.1/28.85	72.73/27.27	1.000
18	15,627,771	rs12525702	intron	SNV	C/T	92.31/7.69/0.00	81.82/18.18/0.00	0.392	96.15/3.85	90.91/9.09	0.408
19	15,638,037	rs140124163	5’ splicesite	IN	-/A	100.00/0.00/0.00	95.45/4.55/0.00	-	100.00/0.00	97.73/2.27	-
20	15,638,161	rs372222238	intron	SNV	T/A	92.30/7.69/0.00	100.00/0.00/0.00	-	96.15/3.85	100.00/0.00	-
21	15,652,452	rs2619526	intron	SNV	G/A	84.62/11.54/3.85	86.36/13.64/0.00	1.000	90.38/9.62	93.18/6.82	0.723
22	15,663,118	rs11558324	5’ UTR	SNV	T/C	96.15/3.85/0.00	100.00/0.00/0.00	-	98.08/1.92	100.00/0.00	-

* Genbank accession number for novel DNA sequence

(ref: reference; SNV: single nucleotide variant; 1: allele 1 (reference allele); 2: allele 2 (variant); MNV: multi nucleotide variant; DEL: deletion; IN: insertion)

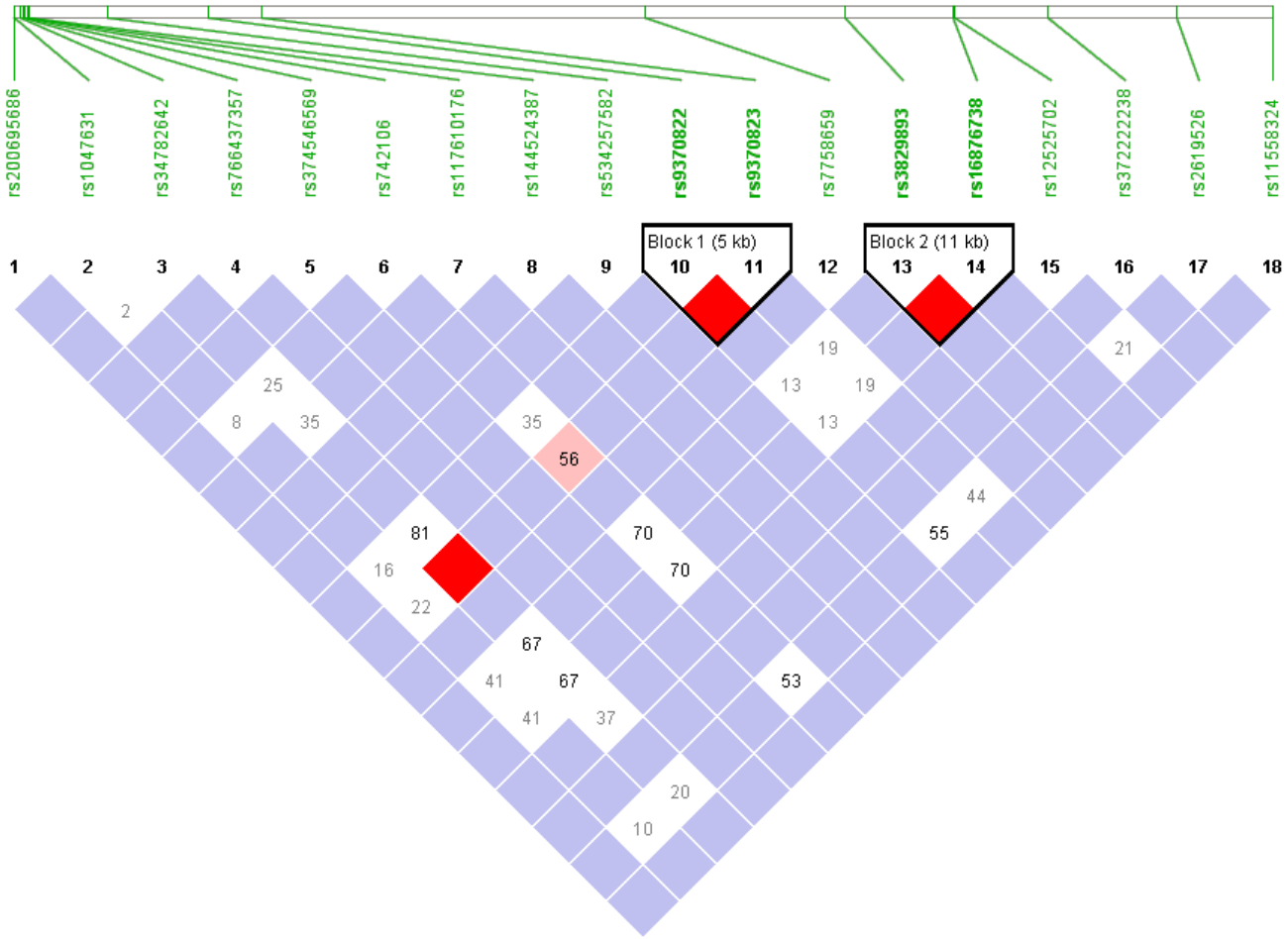


Figure 4.6: LD structure of *DTNBP1* variants detected in patients, showing two haploblocks

Table 4.6: Validation of selected variants in *DTNBP1*

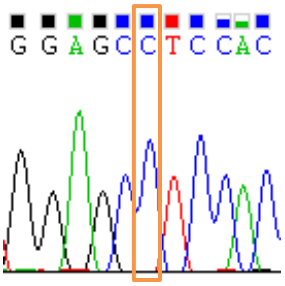
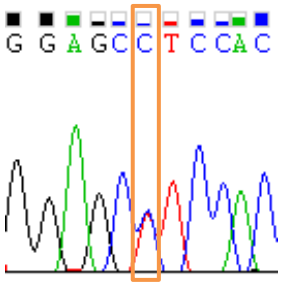
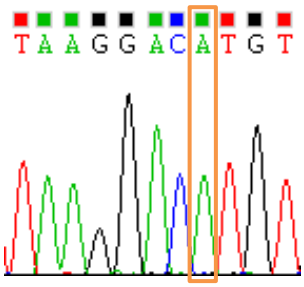
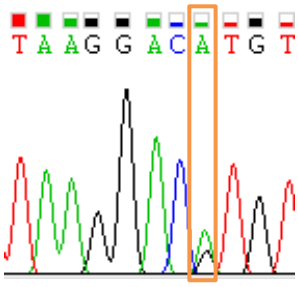
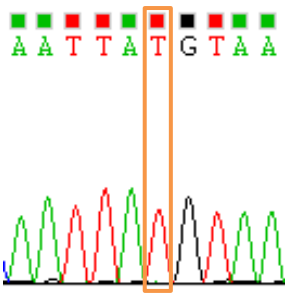
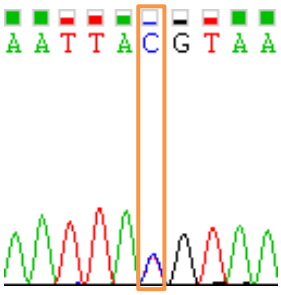
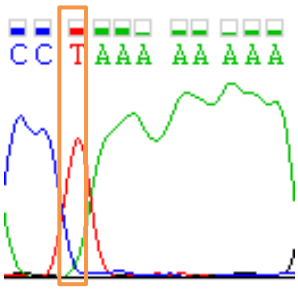
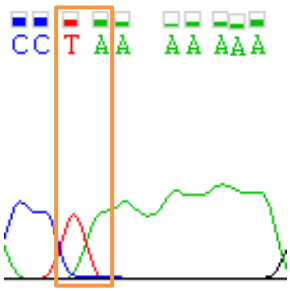
Variant position	Allele	Reference	Variation
15,522,879 (KX619615)	C>C/T		
15,522,988 (rs200695686)	A>A/G		
15,523,101 (rs1047631)	T>T/C		
15,638,037 (rs140124163)	T>T/A		

Table 4.7: *In-silico* variant functional prediction of *DTNBP1* variants

Position (SNP ID)	Variant Type	Ref/ variant	Function	Effect
15,523,101 (rs1047631)	SNV (3' UTR)	T/C	Binding site for hsa-miR-377-3p	When allele changes from C to T, miRNA-mRNA binding will be broken. Hence inhibition of gene expression by risk allele
15,524,715 (rs144524387)	SNV (intron)	G/A	Stop-gained mutation. Amino acid change from Arginine (R) to stop codon at position 285 in dysbindin isoform b	dysbindin isoform
15,638,037 (rs140124163)	IN (5'splicesite)	-/A	Splice variant that changes the 3' end of intron 3	Affect protein coding and might cause protein truncation or loss of function
15,663,118 (rs11558324)	SNV (5' UTR)	T/C	5' UTR with benign allele	Affect protein coding of exon 1

(Ref: reference allele; SNV: single nucleotide variant; IN: insertion)

4.4.3 *AKT1*

Table 4.8 shows the allelic and genotypic frequencies of variants detected in *AKT1*. Among the 33 variants detected, all were intronic, except seven 3' UTR variants and two synonymous variants. Association test showed that genotype (p = 0.031) and allele (p = 0.041) frequencies of rs2494738 were different between patients and controls, but the results did not survive FDR correction. Figure 4.7 shows PCR products of the six variants (rs17846829, rs58565216, rs17846828, rs3803305, rs35416681, and rs17846826) selected for validation using Sanger sequencing and Table 4.9 shows the sequencing chromatograms.

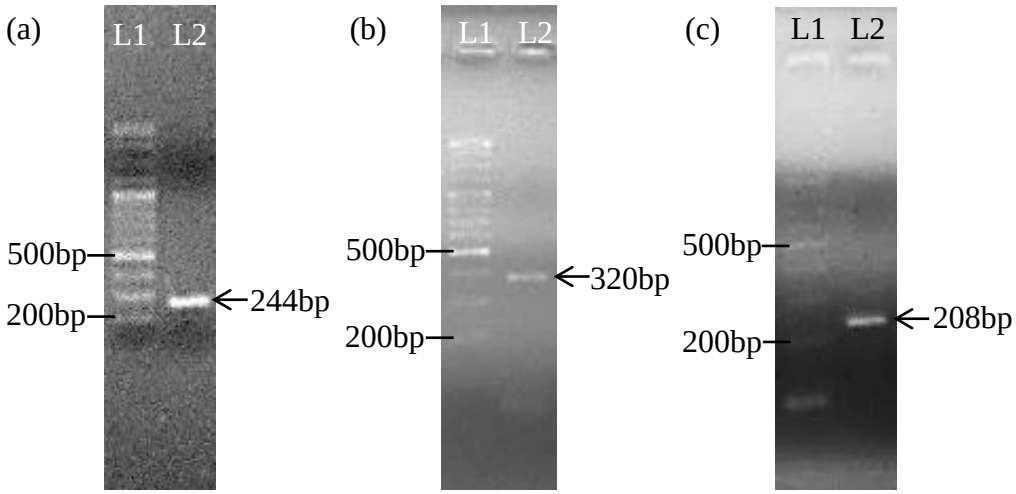


Figure 4.7: Agarose gel electrophoresis of selected *AKT1* variants: of (a) rs17846829, rs58565216, rs17846828 and rs3803305; (b) rs35416681; and (c) rs17846826. Lane 1: VC 100 bp Plus DNA ladder (Vivantis, Malaysia); Lane 2: PCR products

Table 4.8: Variants detected in *AKT1*

No.	Position	SNP ID	Location	Type	ref/variant (1/2)	Genotype frequency (11/12/22)		*p*-value	Allele frequency (1/2)		*p*-value
						Patient	Control		Patient	Control	
1	105,235,558	rs2498801	downstream	SNV	T/C	15.38/46.15/38.46	22.73/36.36/40.91	0.806	38.46/61.54	40.91/59.09	0.837
2	105,235,818	rs17846829	3' UTR	SNV	C/T	92.31/7.69/0.00	100.00/0.00/0.00	-	96.15/3.85	100.00/0.00	-
3	105,235,824	rs58565216	3' UTR	SNV	G/C	88.46/11.54/0.00	100.00/0.00/0.00	-	94.23/5.77	100.00/0.00	-
4	105,235,825	rs17846828	3' UTR	SNV	G/A	92.31/7.69/0.00	100.00/0.00/0.00	-	96.15/3.85	100.00/0.00	-
5	105,235,860	rs3803305	3' UTR	SNV	G/A	92.31/7.69/0.00	100.00/0.00/0.00	-	96.15/3.85	100.00/0.00	-
6	105,235,897	rs3840005	3' UTR	DEL	AAGGTC/-	92.31/7.69/0.00	100.00/0.00/0.00	-	96.15/3.85	100.00/0.00	-
7	105,236,287	rs35416681	3' UTR	SNV	A/C	88.46/11.54/0.00	100.00/0.00/0.00	-	94.23/5.77	100.00/0.00	-
8	105,236,557	rs17846826	3' UTR	SNV	G/A	92.31/7.69/0.00	100.00/0.00/0.00	-	96.15/3.85	100.00/0.00	-
9	105,236,690	rs200213561	exon	SNV	G/A	100.00/0.00/0.00	95.45/4.55/0.00	-	100.00/0.00	97.73/2.27	-
10	105,237,026	rs150070451	intron	DEL	CTCT/-	92.31/7.69/0.00	100.00/0.00/0.00	-	96.15/3.85	100.00/0.00	-
11	105,238,591	rs55839843	intron	DEL	CTGAG/-	23.08/38.46/38.46	22.73/36.36/40.91	1.000	42.31/57.69	40.91/59.09	1.000
12	105,238,604	rs2498800	intron	SNV	C/T	23.08/38.46/38.46	22.73/36.36/40.91	1.000	42.31/57.69	40.91/59.09	1.000
13	105,238,636	rs17846832	intron	SNV	G/A	80.77/19.23/0.00	100.00/0.00/0.00	-	90.38/9.62	100.00/0.00	-
14	105,238,670	rs8192700	intron	SNV	C/T	88.46/11.54/0.00	100.00/0.00/0.00	-	94.23/5.77	100.00/0.00	-
15	105,239,146	rs3803304	intron	SNV	C/G	80.77/19.23/0.00	90.91/9.09/0.00	0.065	90.38/9.62	95.45/4.55	0.283
16	105,239,192	rs2494732	intron	SNV	T/C	3.85/46.15/50.00	13.64/45.45/40.91	0.621	26.92/73.08	36.36/63.64	0.380

(Table 4.8 continued)

17	105,239,894	rs1130233	exon	SNV	C/T	34.62/26.92/38.46	22.73/36.36/40.91	0.712	48.08/51.92	40.91/59.09	0.540
18	105,240,111	rs3730362	intron	SNV	G/A	88.46/11.54/0.00	100.00/0.00/0.00	-	94.23/5.77	100.00/0.00	-
19	105,241,576	rs3730344	intron	SNV	G/A	88.46/11.54/0.00	100.00/0.00/0.00	-	94.23/5.77	100.00/0.00	-
20	105,242,831	rs3001371	intron	SNV	C/T	23.08/38.46/38.46	22.73/36.36/40.91	1.000	42.31/57.7	40.91/59.09	1.000
21	105,242,926	rs17846822	intron	SNV	G/A	92.31/7.69/0.00	100.00/0.00/0.00	-	96.15/3.85	100.00/0.00	-
22	105,242,966	rs2494735	intron	SNV	T/C	11.54/42.31/46.15	22.73/36.36/40.91	0.638	32.69/67.31	40.91/59.09	0.524
23	105,243,220	rs2498796	intron	SNV	G/A	30.77/30.77/38.46	36.36/22.73/40.91	0.873	46.15/53.85	47.73/52.27	1.000
24	105,246,325	rs2494737	intron	SNV	T/A	19.23/46.15/34.62	4.55/45.45/50.00	0.311	42.31/57.69	27.27/72.73	0.140
25	105,246,407	rs3730358	intron	SNV	G/A	92.31/7.69/0.00	95.45/4.55/0.00	1.000	96.15/3.85	97.73/2.27	1.000
26	105,246,565	rs17846816	intron	SNV	C/T	96.15/3.85/0.00	95.45/4.55/0.00	1.000	98.08/1.92	97.73/2.27	1.000
27	105,246,681	rs35635404	intron	SNV	G/A	92.31/7.69/0.00	100.00/0.00/0.00	-	96.15/3.85	100.00/0.00	-
28	105,246,686	rs2494738	intron	SNV	G/A	42.31/26.92/30.77	9.09/50.00/40.91	**0.031**	55.77/44.23	34.09/65.91	**0.041**
29	105,246,692	rs12588965	intron	SNV	T/A	92.31/7.69/0.00	100.00/0.00/0.00	-	96.15/3.85	100.00/0.00	-
30	105,258,892	rs2494748	intron	MNV	CA/TG	69.23/26.92/3.85	86.36/13.64/0.00	0.379	82.69/17.31	93.18/6.82	0.214
31	105,258,893	rs2494749	intron	SNV	A/G	69.23/23.08/7.69	68.18/31.82/0.00	0.585	80.77/19.23	84.09/15.91	0.791
32	105,259,706	rs10138227	intron	SNV	C/T	84.62/15.38/0.00	81.82/18.18/0.00	1.000	92.31/7.69	90.91/9.09	1.000
33	105,259,734	rs1130214	intron	SNV	C/A	84.62/15.38/0.00	81.82/18.18/0.00	1.000	92.31/7.69	90.91/9.09	1.000

(ref: reference; SNV: single nucleotide variant; 1: allele 1 (reference allele); 2: allele 2 (variant); MNV: multi nucleotide variant; DEL: deletion)

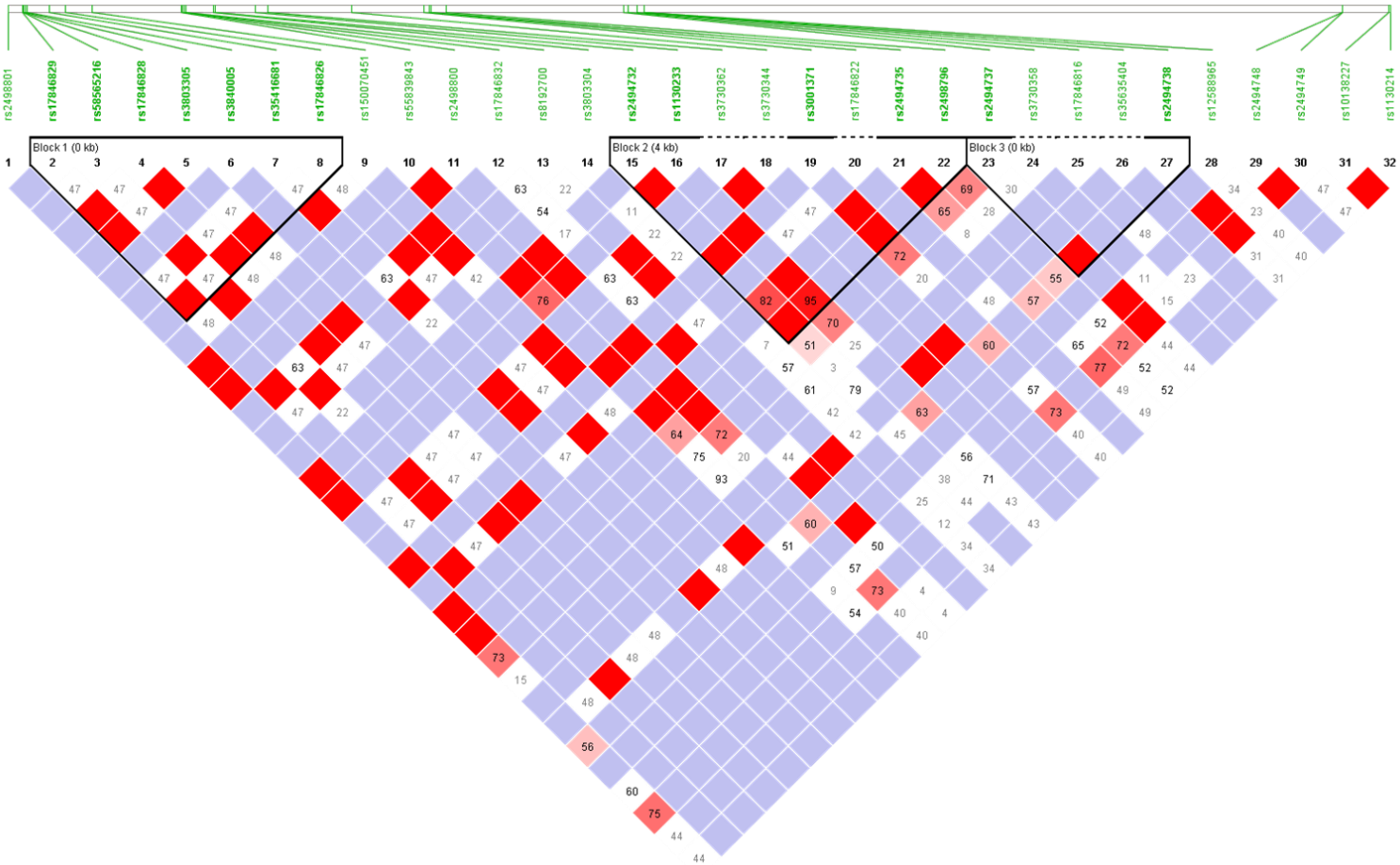
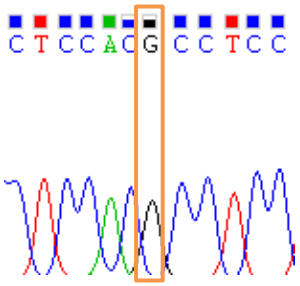
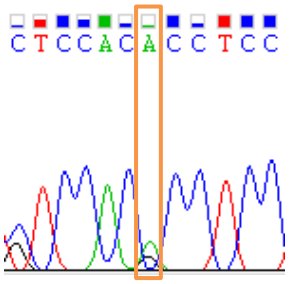
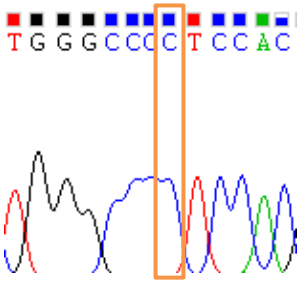
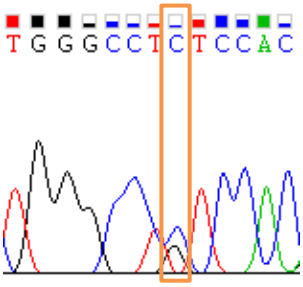


Figure 4.8: LD structure of *AKT1* variants detected in patients, showing three haploblocks

The LD structure of variants detected in patients shows three haploblocks (Figure 4.8). The 3' UTR region showed a haploblock with D' = 1.0 and consisted of seven patient-specific variants which are in LD with each other and are inherited together. Functional predictions predicted six 3'UTR variants: rs17846829, rs58565216, rs17846828, rs3803305, rs35416681, and rs17846826 as miRNA-mRNAs binding sites. Two exonic variants, rs200213561 and rs1130233, were predicted to have synonymous nucleotide substitutions (Table 4.10). Four of the 3' UTR variants were predicted to promote miRNA-mRNA binding and reduce gene expression of *AKT1*, while the remaining two were predicted to break miRNA-mRNA binding.

Table 4.9: Validation of selected variants in *AKT1*

Variant position	Allele	Reference	Variation
105,235,818 (rs17846829)	G>G/A		
105,235,824 (rs58565216)	C>C/G		

(Table 4.9 continued)

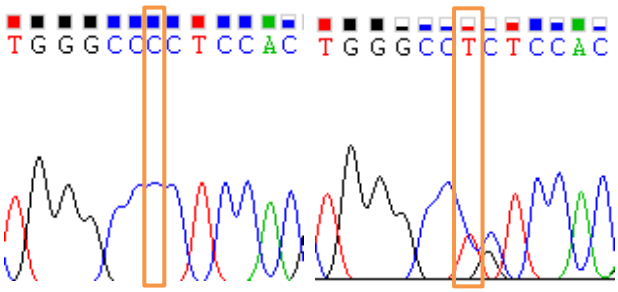
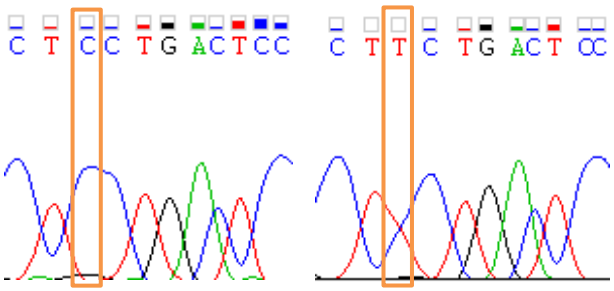
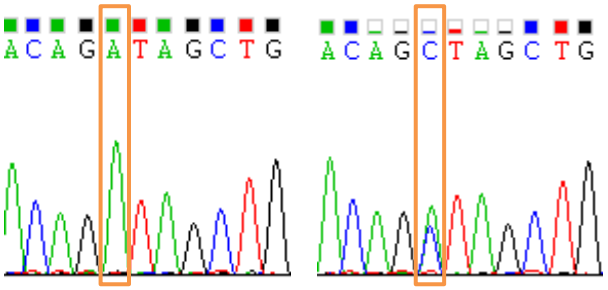
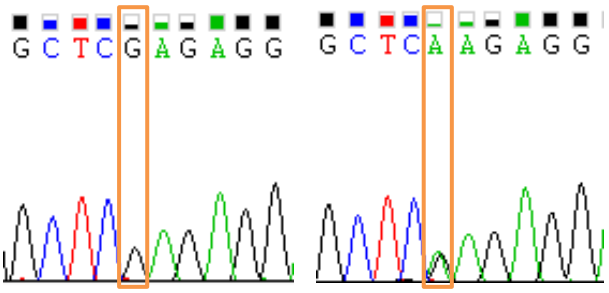
105,235,825 (rs17846828)	C>T	
105,235,860 (rs3803305)	C>T	
105,236,287 (rs35416681)	A>A/C	
105,236,557 (rs17846826)	G> G/A	

Table 4.10: *In silico* variant functional prediction of *AKT1* variants

SNP ID	ref/ variant	Function	Effect	Inhibition of *AKT1* expression by allele	
				reference	variant
3' UTR:					
rs17846829	C/T	Binding site for hsa-miR-665	When allele changes from T to C, new miRNA-mRNA binding will be created	Yes	No
rs58565216	G/C	Binding site for hsa-miR-3184-5p	When allele changes from G to C, new miRNA-mRNA binding will be created	No	Yes
rs17846828	G/A	Binding site for hsa-miR-3184-5p	When allele changes from A to G, miRNA-mRNA binding will be broken	No	Yes
rs3803305	G/A	Binding site for hsa-miR-149-3p	When allele changes from A to G, miRNA-mRNA binding will be broken	No	Yes
rs35416681	A/C	Binding site for hsa-miR-143-3p	When allele changes from A to C, miRNA-mRNA binding will be broken	Yes	No
rs17846826	G/A	Binding site for hsa-miR-4674	When allele changes from A to G, miRNA-mRNA binding will be broken	No	Yes

(Table 4.10 continued)

Exon:			
rs200213561	G/A	Synonymous nucleotide substitution	Amino acid remain as Serine at position 477
rs1130233	C/T	Synonymous nucleotide substitution	Amino acid remain as Glutamic acid at position 242

4.4.4 *TCF4*

Table 4.11 shows the allelic and genotypic frequencies of variants detected in *TCF4*. Both genotype ($p = 3.10 \times 10^{-29}$) and allele ($p = 2.53 \times 10^{-12}$) frequencies of rs2958182 were found significantly different in patient after FDR correction. Figure 4.9 shows PCR products of rs141970461, rs867452925, and rs599550, which were selected for validation using Sanger sequencing. The sequencing chromatograms were shown in Table 4.12. LD structure of variants detected in patients is shown in Figure 4.10. In functional predictions, three 3' UTR variants: rs369444573, rs76640061, and rs180957404 were predicted to be miRNA-mRNAs binding sites, while exonic rs8766 was predicted to be a synonymous variant (Table 4.13).

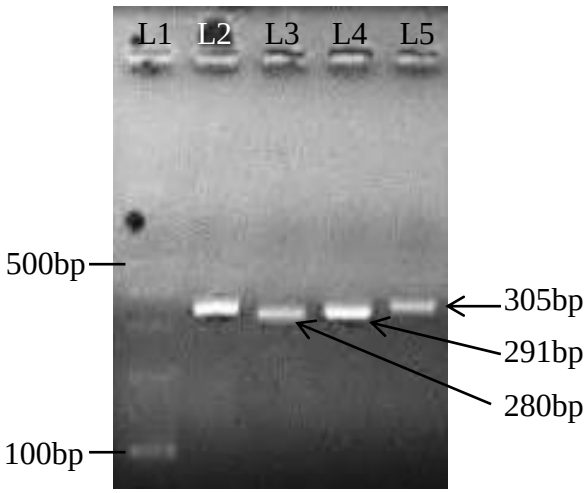


Figure 4.9: PCR products of selected *TCF4* variants. Lane 1: VC 100 bp DNA ladder (Vivantis, Malaysia); Lane 2 and 5: rs599550; Lane 3: rs141970461; Lane 4: rs867452925

Table 4.11: Variants detected in *TCF4*

No.	Position	SNP ID	Location	Type	ref/variant (1/2)	Genotype frequency (11/12/22)		*p*-value	Allele frequency (1/2)		*p*-value
						Patient	Control		Patient	Control	
1	52,752,017	rs12966547	3' UTR	SNV	A/G	38.46/38.46/23.08	22.73/50.00/27.27	0.060	57.69/42.31	47.73/52.27	0.202
2	52,889,967	rs1261085	3' UTR	SNV	T/C	46.15/34.62/19.23	27.27/59.09/13.64	0.269	63.46/36.54	56.82/43.18	0.536
3	52,890,073	rs541061758	3' UTR	DEL	GTCA/-	100.00/0.00/0.00	95.45/4.55/0.00	-	100.00/0.00	97.73/2.27	-
4	52,890,160	rs1261084	3' UTR	SNV	A/G	46.15/34.62/19.23	27.27/59.09/13.64	0.269	63.46/36.54	56.82/43.18	0.536
5	52,890,543	rs767178570	3' UTR	IN	T/-	100.00/0.00/0.00	90.91/9.09/0.00	-	100.00/0.00	95.45/4.55	-
6	52,890,940	rs369444573	3' UTR	IN	-/C	88.46/7.69/3.85	86.36/13.64/0.00	0.812	92.31/7.69	93.18/6.82	1.000
7	52,891,204	rs140440103	3' UTR	IN	-/T	100.00/0.00/0.00	95.45/4.55/0.00	-	100.00/0.00	97.73/2.27	-
8	52,892,315	rs76640061	3' UTR	SNV	T/C	100.00/0.00/0.00	95.45/4.55/0.00	-	100.00/0.00	97.73/2.27	-
9	52,893,255	rs141970461	3' UTR	IN	-/A	73.08/0.00/26.92	95.45/0.00/4.55	0.055	73.08/26.92	95.45/4.55	**0.005**
10	52,893,265	rs201139617	3' UTR	SNV	C/A	96.15/3.85/0.00	95.45/4.55/0.00	1.000	98.08/1.92	97.73/2.27	1.000
11	52,893,377	rs180957404	3' UTR	SNV	A/C	80.77/19.23/0.00	81.82/18.18/0.00	1.000	90.38/9.62	90.91/9.09	1.000
12	52,893,680	rs867452925	3' UTR	SNV	C/T	92.31/7.69/0.00	100.00/0.00/0.00	-	96.15/3.85	100.00/0.00	-
13	52,894,540	rs781564405	3' UTR	IN	G/-	96.00/0.00/4.00	100.00/0.00/0.00	-	96.00/4.00	100.00/0.00	-
14	52,895,418	rs6567210	intron	SNV	G/A	84.62/15.38/0.00	90.91/9.09/0.00	0.674	92.31/7.69	95.45/4.55	0.684
15	52,895,531	rs8766	exon	SNV	T/C	34.62/30.77/34.62	31.82/50.00/18.18	0.335	50.00/50.00	56.82/43.18	0.543
16	52,895,657	rs10221362	intron	SNV	C/T	42.31/30.77/26.92	40.91/45.45/13.64	0.491	57.69/42.31	63.64/36.36	0.676

(Table 4.11 continued)

17	52,896,000	rs8087569	intron	SNV	A/G	84.62/15.38/0.00	86.36/13.64/0.00	1.000	92.31/7.69	93.18/6.82	1.000
18	52,896,003	rs8087570	intron	SNV	A/G	84.62/15.38/0.00	86.36/13.64/0.00	1.000	92.31/7.69	93.18/6.82	1.000
19	52,921,606	rs1539950	intron	SNV	C/T	46.15/38.46/15.38	50.00/36.36/13.64	1.000	65.38/34.62	68.18/31.82	0.830
20	52,921,673	rs1539951	intron	SNV	T/C	46.15/38.46/15.38	50.00/36.36/13.64	1.000	65.38/34.62	68.18/31.82	0.830
21	52,922,000	rs62092442	intron	SNV	T/C	38.46/38.46/23.08	40.91/40.91/18.18	1.000	57.69/42.31	61.36/38.64	0.835
22	52,922,007	rs62092443	intron	SNV	C/T	96.15/3.85/0.00	100.00/0.00/0.00	-	98.08/1.92	100.00/0.00	-
23	52,927,275	rs138389096	intron	IN	A/-	96.15/3.85/0.00	100.00/0.00/0.00	-	98.08/1.92	100.00/0.00	-
24	52,942,827	rs1788027	intron	SNV	G/A	15.38/46.15/38.46	22.73/45.45/31.82	0.865	38.46/61.54	45.45/54.55	0.537
25	53,017,488	rs117303010	intron	SNV	C/T	100.00/0.00/0.00	90.91/9.09/0.00	-	100.00/0.00	95.45/4.55	-
26	53,070,764	rs771654348	intron	SNV	G/A	100.00/0.00/0.00	90.91/9.09/0.00	-	100.00/0.00	95.45/4.55	-
27	53,070,914	rs17522826	intron	SNV	G/A	65.38/26.92/7.69	40.91/50.00/9.09	0.248	78.85/21.15	65.91/34.09	0.174
28	53,089,352	rs376676806	intron	DEL	-/A	96.15/3.85/0.00	100.00/0.00/0.00	-	98.08/1.92	100.00/0.00	-
29	53,089,361	rs796552500	intron	SNV	A/C	96.15/3.85/0.00	100.00/0.00/0.00	-	98.08/1.92	100.00/0.00	-
30	53,131,206	rs111953577	intron	SNV	T/A	100.00/0.00/0.00	90.91/9.09/0.00	-	100.00/0.00	95.45/4.55	-
31	53,131,208	rs9945479	intron	SNV	T/A	96.15/3.85/0.00	86.36/4.55/9.09	0.449	98.08/1.92	88.64/11.36	0.090
32	53,131,210	rs11152407	intron	SNV	A/T	53.85/23.08/23.08	40.91/22.73/36.36	0.564	65.38/34.62	52.27/47.73	0.216
33	53,131,212	rs147574299	intron	SNV	A/T	53.85/46.15/0.00	68.18/31.82/0.00	0.382	76.92/23.08	84.09/15.91	0.447
34	53,131,213	rs549984245	intron	SNV	G/C	96.15/3.85/0.00	100.00/0.00/0.00	-	98.08/1.92	100.00/0.00	-

(Table 4.11 continued)

35	53,131,214	rs748458759	intron	SNV	A/T	100.00/0.00/0.00	95.45/4.55/0.00	-	100.00/0.00	97.73/2.27	-
36	53,131,388	rs758306015	intron	SNV	T/C	100.00/0.00/0.00	95.45/4.55/0.00	-	100.00/0.00	97.73/2.27	-
37	53,149,021	rs2958182	intron	SNV	A/T	7.69/23.08/69.23	36.36/63.64/0.00	**3.10×10^{-29}***	19.23/80.77	68.18/31.82	**2.53×10^{-12}***
38	53,155,002	rs9960767	intron	SNV	A/C	76.92/23.08/0.00	100.00/0.00/0.00	-	88.46/11.54	100/0.00	-
39	53,177,670	rs9957668	intron	SNV	T/C	23.08/50.00/26.92	18.18/40.91/40.91	0.695	48.08/51.92	38.64/61.36	0.412
40	53,177,679	rs144036944	intron	SNV	A/C	92.31/7.69/0.00	86.36/13.64/0.00	0.649	96.15/3.85	93.18/6.82	0.658
41	53,177,740	rs9954890	intron	SNV	A/T	23.08/50.00/26.92	18.18/40.91/40.91	0.695	48.08/51.92	38.64/61.36	0.412
42	53,177,774	rs9964328	intron	SNV	G/C	23.08/50.00/26.92	18.18/40.91/40.91	0.695	48.08/51.92	38.64/61.36	0.412
43	53,178,061	rs67387556	intron	SNV	G/A	80.77/11.54/7.69	90.91/9.09/0.00	0.674	86.54/13.46	95.45/4.55	0.173
44	53,192,498	rs10401120	intron	SNV	C/T	76.92/23.08/0/00	86.36/13.64/0/00	0.144	88.46/11.54	93.18/6.82	0.335
45	53,194,961	rs17512836	intron	SNV	C/T	0.00/1.38/84.62	0.00/0/00/100.00	-	7.69/92.31	0.00/100	-
46	53,252,388	rs599550	intron	SNV	G/A	0.00/19.23/80.77	0.00/0.00/100.00	0.054	9.62/90.38	0.00/100	0.060
47	53,255,335	rs533414588	intron	SNV	C/T	100.00/0.00/0.00	95.45/0.00/4.55	-	100.00/0.00	95.45/4.55	-
48	53,255,630	rs536947841	intron	DEL	G/-	3.85/0.00/96.15	18.18/0.00/81.82	0.165	3.85/96.15	18.18/81.82	**0.040**

*significant after FDR correction (ref: reference; SNV: single nucleotide variant; 1: allele 1 (reference allele); 2: allele 2 (variant); DEL: deletion; IN: insertion)

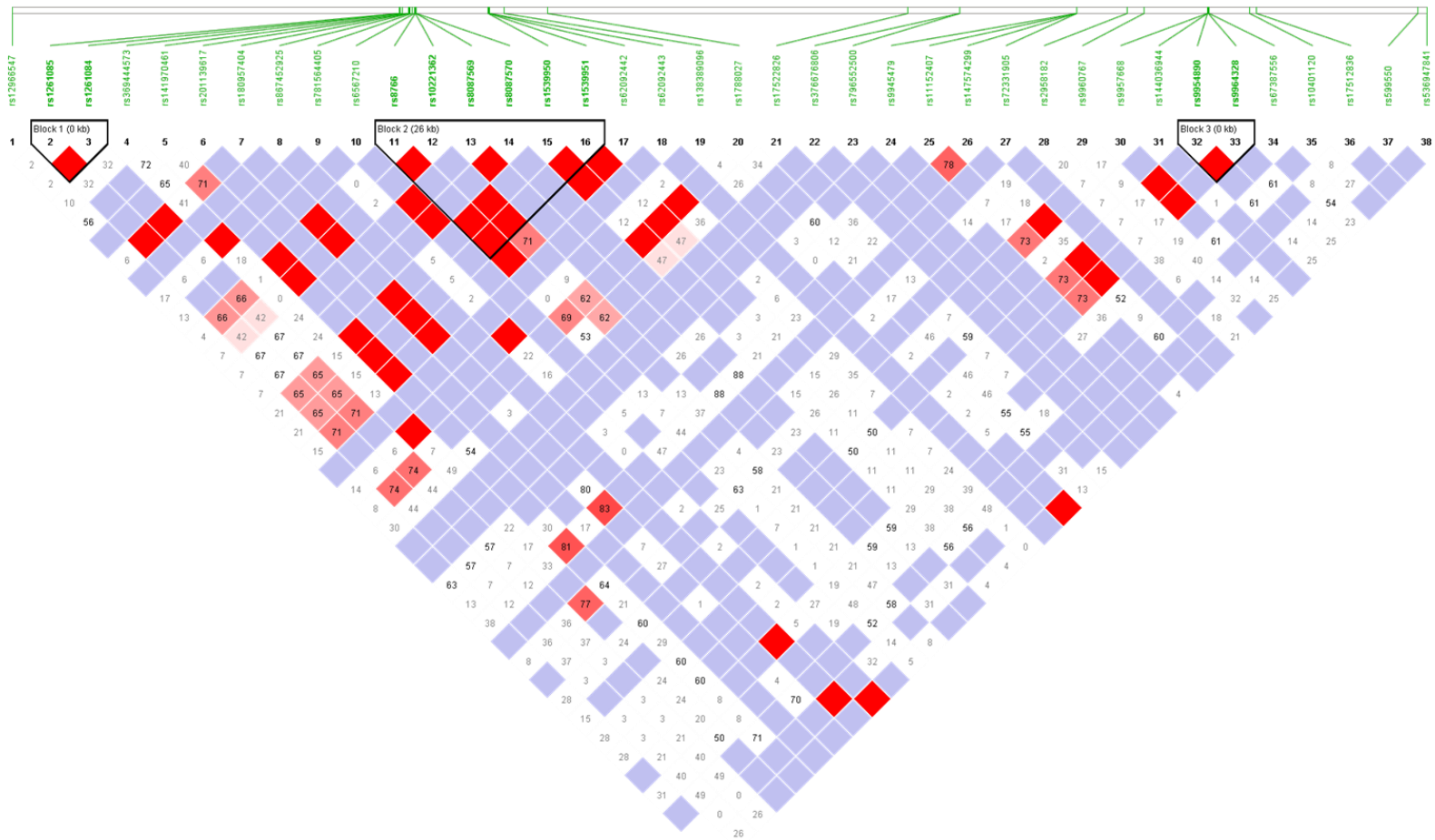


Figure 4.10: LD structure of variants detected in *TCF4*, showing three haploblocks

Table 4.12: Validation of selected variants in *TCF4*

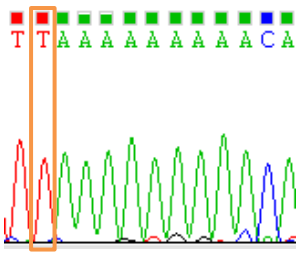
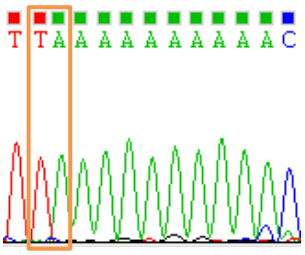
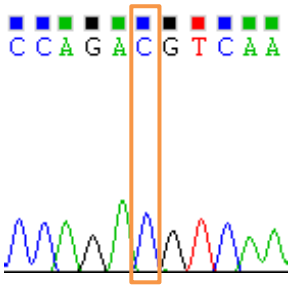
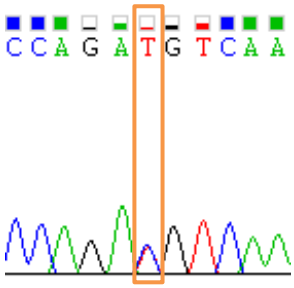
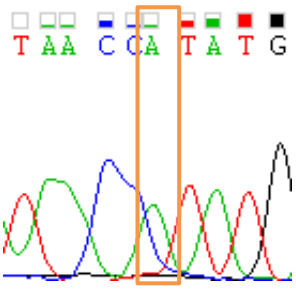
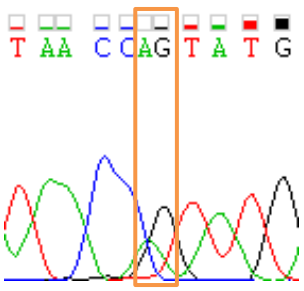
Variant position	Allele	Reference	Variation
52,893,255 (rs141970461)	T>TA		
52,893,680 (rs867452925)	C>CT		
53,252,388 (rs599550)	A>A/G		

Table 4.13: *In-silico* variant functional prediction of *TCF4* variants

Position (SNP ID)	Variant Type	Ref	Function	Effect
52,890,940 (rs369444573)	IN (3' UTR)	-/C	Binding site for miR-204/211	miRNA-mRNA binding broken with presence of allele C, promoting gene expression
52,892,315 (rs76640061)	SNV (3' UTR)	T/C	Binding site for hsa-miR-204-5p	When allele changes from T to C, miRNA-mRNA binding will be broken, promoting gene expression
52,893,377 (rs180957404)	SNV (3' UTR)	A/C	Transcription factor binding site for C/EBPalpha (P49715) and Hlf (Q16534)	
52,895,531 (rs8766)	SNV (exon)	T/C	Synonymous nucleotide mutation. Amino acid remains as Serine (S) at position 643	

(Ref: reference allele; IN: insertion; SNV: single nucleotide variant)

CHAPTER 5

DISCUSSION

5.1 Cognitive Performance in Schizophrenia

The current study examined neurocognitive functions in schizophrenia. Results showed differences between patients and controls in both direct and derived scores where patients underperformed in TMT-A, TMT-B, and B-A scores. TMT-A demonstrates processing speed and visuoperceptual abilities, while TMT-B reflects primarily on working memory and executive functioning such as reasoning, problem-solving, and task-switching ability (Perianez *et al.*, 2007; Sanchez-Cubillo *et al.*, 2009). Significant result in the B-A score highlighted the accuracy of test evaluations for relative executive control, free from the influence of speed factor.

However, the result of B:A and (B–A):A indicated that there were no difference in task-switching ability and prefrontal cortex functions between patients and controls. The negative results might be due to the small sample size which greatly decreases the statistical power when analysing the derived scores. Similarly, due to the small sample size, the present study only involved two groups: case and control, hence the log B:A result is not indicative to the current sample population. Overall TMT results for the current sample

population indicate that patient exhibited greater deficiency in processing speed and overall executive performance without the influence of speed.

The degree of cognitive impairment is associated with familial genetic loading for the illness, where higher genetic loading leads to greater level of impairment (Gur *et al.,* 2007). In cognition analysis, TMT is a widely used neuropsychological assessment tool for schizophrenia. Evidences from past TMT studies reported a positive linear relationship between brain injury and TMT performance, which is especially demonstrated in patients with frontal lobe damage (Demakis, 2004; Lange *et al.,* 2005). On a side note, it is important to obtain TMT results from current healthy controls to serve as a norm data for Malaysian population which is useful for comparative studies across different populations (Perianez *et al.,* 2007).

5.2 Association of *ZNF804A*, *DTNBP1*, *AKT1*, and *TCF4* with Schizophrenia and Cognition

Accumulating evidence suggests that the risk of schizophrenia and cognitive deficits are extensively associated with genetic variations and they harbour certain amounts of shared genetic effects (Owens *et al.,* 2011). A plausible strategy to explore this association is to study the functional regions of genes associated to both schizophrenia and neurocognition in a familial cohort setting, since such approach would capture functional variants that have direct impact on gene expressions. The current results suggest that cognitive

deficits is related to variations within cognition-associated genes *ZNF804A*, *DTNBP1*, *AKT1*, and *TCF4*, leading to increased risk for schizophrenia.

The interactions between these four genes are shown in Figure 5.1. *ZNF804A* is a GWAS identified schizophrenia risk gene which functions to regulate DNA binding and transcription of various other schizophrenia-associated genes such as *COMT*, *PRSS16*, *DRD2*, and *ANKRD1* (Girgenti *et al.*, 2012; Umeda-Yano *et al.*, 2013). Its function in regulating gene expression is associated to transforming growth factor β (TGF-β) signalling, which is also often implicated in schizophrenia and where elevated levels of TGF-β were found in hippocampus of patients (Umeda-Yano *et al.*, 2013). Overexpression of *ZNF804A* correlates to increased TGF-β level, and the synergic interaction between *ZNF804A* and TGF-β was found to increase risk of schizophrenia and cognitive deficits (Umeda-Yano *et al.*, 2013; Borovcanin *et al.*, 2016).

Meanwhile, kinase activity of TGF-β will activate PI3-K, which plays an important role to phosphorylate and activates *AKT1* (Zhang, 2009). *AKT1* is a schizophrenia susceptible gene and disruption in the AKT signaling pathway plays a significant role in the pathogenesis of schizophrenia (Emamian, 2012). Decrease level of *AKT1* is often seen in schizophrenia and is associated with poor cognitive performance (Lai *et al.*, 2006). Besides that, *AKT1* is important to negatively regulate GSK-3. Elevated level of GSK-3 is implicated in mood disorder and anxiety (Freyberg *et al.*, 2010), besides causing the degradation of β-catenin, an important substrate in *TCF4* transcription regulation pathway (Fang *et al.*, 2007).

Besides that, up-regulation of *DTNBP1* will also lead to increased phosphorylation and activity of *AKT1* (Numakawa *et al.*, 2004; Riley and Kendler, 2006). *DTNBP1* is implicated as schizophrenia susceptible gene and low levels of the gene have been observed in postmortem brain of patients (Weickert *et al.*, 2004). Besides that, decreased *DTNBP1* level is also associated with cognitive deficits such as low intellectual functions (Zinkstok *et al.*, 2007). The activated *AKT1* will phosphorylate β-catenin and stimulate uptake of *TCF4* to form a complex important in regulating gene transcription (Fang *et al.*, 2007). *TCF4* is a schizophrenia susceptible gene where overexpression of the gene can lead to cognitive impairments such as poor executive functioning (Wirgenes *et al.*, 2012; Albanna *et al.*, 2014). Altogether, *ZNF804A*, *DTNBP1*, *AKT1*, and *TCF4* gene may exert individual effect and also play a part in this network to giving rise to the accumulating effects on the pathophysiology of schizophrenia and neurocognition.

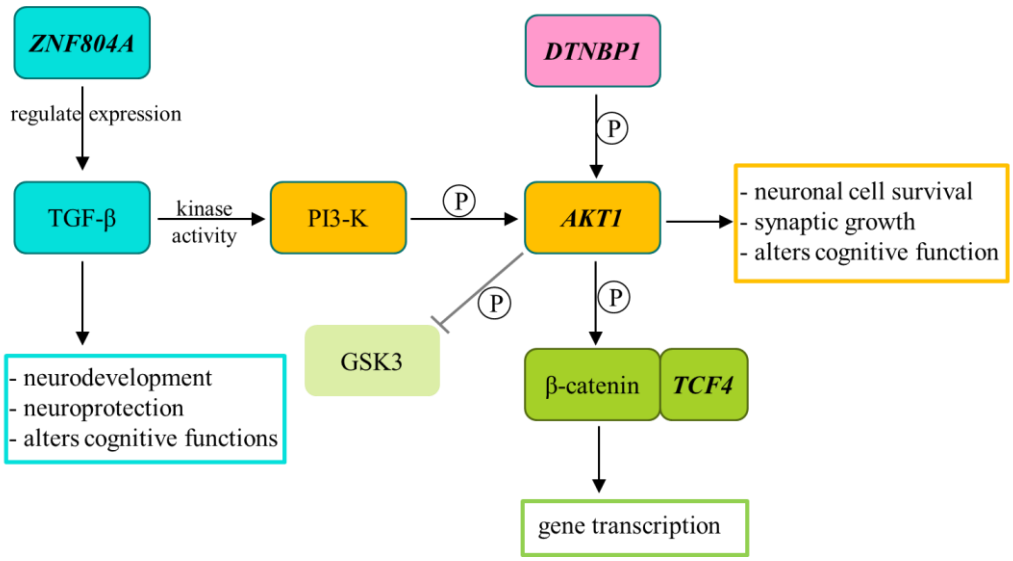


Figure 5.1: Association of *ZNF804A*, *DTNBP1*, *AKT1*, and *TCF4* with schizophrenia and cognition (Ⓟ: phosphorylate)

5.2.1 *ZNF804A*

In upstream 5' UTR regions of *ZNF804A*, four variants (rs227926, rs359895, rs370931594, and rs73041379) were shown to be in LD with each other in the present study. Among them, rs227926 and rs359895 portrayed marginal association with schizophrenia while rs370931594 was only detected in patients. Interestingly, the present result shows that all patients do not carry the homozygous risk genotype (variant/variant) in any of these 5' variants. This finding proposes a protective role for the homozygous risk genotypes in 5' UTR against the risk of schizophrenia and neurocognition.

The functional role for rs227926 is unknown but it is assumed to be related with rs359895, since both of them were only 65 bp apart and in linkage with each other. The risk allele at rs359895 is reported to influence the expression of *ZNF804A* by promoting promoter activity, hence suggesting its potential to be a functional variant (Li *et al.*, 2011). Rs359895 resides in the core promoter region approximately 500 bp from the transcription start site of *ZNF804A* and serves as a transcription factor binding site for nuclear factor 1C (NF-1C) and nuclear transcription factors Y (NF-Y) subunits. NF-1C is capable of activating transcription and replication, while NF-Y subunits can function as both activator and repressor, depending on their interacting cofactors. Studies have shown NF-1C protein to be associated with Alzheimer's disease, mental retardation, anxiety, and depression; while NF-Y subunits were linked to astrocytoma and central nervous system cancer (Chaudhry *et al.*, 1997; Lamani *et al.*, 2009; Lee *et al.*, 2012a; Jun *et al.*, 2017).

Besides that, rs359895 was also found to alter the binding affinity of transcription factor Sp1, where risk allele T was shown to have higher binding affinity and promoter activity (Li *et al.*, 2011). However, there was a dispute on determining whether the reported risk allele T corresponds to the minor or the major allele, as this variant has a canonical substitution (Sand, 2013). To further support this, NCBI Entrez database explicitly showed that the A allele as the minor allele, which is similar to the current study where A allele is the risk allele and T allele serving as the ancestral allele. A total of 77% of controls were identified as TT genotype carriers whereas only 57% of the patients carry the same genotype. This showed a significant genotype effect and further strengthens the role of A allele as the risk allele in schizophrenia.

Meanwhile, non-coding variants of *ZNF804A* have often been associated with schizophrenia. The current study found significant association of rs12613195 and rs7597593 with schizophrenia. Not much is known about these two variants, but located 244 kb downstream of rs7597593 is rs1344706, which was marginally significant in patients after FDR correction. Rs1344706 was the first *ZNF804A* risk variant achieving genome-wide significance in GWAS studies for schizophrenia (O'Donovan *et al.*, 2008) and was replicated in independent studies (Zhang *et al.*, 2012). It was suggested that rs1344706 risk allele A serves a protective role to schizophrenia neurocognition (Schultz *et al.*, 2014) where individual carrying this allele had spared memory function (Walters *et al.*, 2010) and relatively larger grey matter volumes (Donohoe *et al.*, 2011). This is consistent with the current result where patients have lower allele A frequency compared to controls. It was suggested that *ZNF804A* is a

common risk gene for European populations but there were other causal risk variants such as rs359895 in the Chinese population (Li *et al.*, 2011). Furthermore, it was found that although rs1344706 is associated with higher *ZNF804A* expression, it is not responsible for it (Williams *et al.*, 2011b).

Most of the exonic variants detected in this study lies within the vicinity of exon 4 but they showed no significant associations to schizophrenia and were mostly found in controls. However, a risk haplotype was identified encompassing the exon 4 variants from rs12476147 to rs12477430. Approximately 1kb apart, rs3731834 (Leu1081Val) showed marginal association with schizophrenia. Meta-analysis suggested association of rs3731834 with schizophrenia, not as a functional locus, but rather contributing as a genetic marker that is linked to the causal variant (Zhang *et al.*, 2012). Indeed, current functional prediction result of rs3731834 indicated a tolerated non-synonymous substitution, most likely having a benign effect on protein structure and function. Rs61739290 (Ile492Val) and rs1366842 (Thr707Lys) were reported to be in LD with rs3731834 in the Chinese population (Zhang *et al.*, 2012), but not in the present population. Besides that, GWAS study showed that rs4667001 (Glu542Lys) and rs728534 (Lys598Lys), which also lies in exon 4, are associated with schizophrenia (O'Donovan *et al.*, 2008). The G allele of rs4667001 was associated with higher *ZNF804A* expression and showed strong linkage with rs1344706 (Williams *et al.*, 2011b). The apparent discrepancies between the results might be due to the small sample size in the present study. However, the accumulation of these missense

variants, shown to be in LD with each other, suggests this exon 4 region may encompass susceptibility variants to schizophrenia (Williams *et al.*, 2011b).

The *ZNF804A* protein is known to have pleiotropic effects on schizophrenia where over- or under-expression of *ZNF804A* will lead to different phenotypic symptoms, predominantly affecting brain connectivity (Bray *et al.*, 2010). *ZNF804A* knockdown was shown to alter expression of cell adhesion-related genes, hence reflecting influence on neurodevelopmental functions such as synapse development, neural migration, and neurite outgrowth (Hill *et al.*, 2011). Although the current information about the function and expression of *ZNF804A* is still insufficient, but it appears that it may predominantly affect brain connectivity, reflecting neurodevelopmental abnormalities in patients (Bray *et al.*, 2010).

In the present study, patients who carried the *ZNF804A* risk alleles were found to have below average functioning of both processing speed and executive controls. In addition, patients having greater amount of detected risk variants seemed to perform poorer in the speed domain. This is consistent with the findings that *ZNF804A* risk alleles alter processing speed (Nicodemus *et al.*, 2014) and cause impairments on a timed visuomotor performance task (Lencz *et al.*, 2010). Besides that, evidence also showed the association of *ZNF804A* variant rs1344706 with the executive control of attention in schizophrenia (Balog *et al.*, 2011). Given that poor processing speed and executive deficit is a prominent phenotype in schizophrenia and that *ZNF804A* is a schizophrenia susceptibility gene, the present results supports the role of

ZNF804A in neurocognitive phenotypes reflecting schizophrenia susceptibility. Taken together, the deteriorating effects of *ZNF804A* on schizophrenia and neurocognitive may lies within the 5'UTR and exon 4 regions of the gene.

5.2.2 *DTNBP1*

DTNBP1 is a large gene with extensive intron and comparatively smaller coding regions that codes for a short transcript with high 5'UTR and 3'UTR variation (De Luca *et al.,* 2005). This explains that the majority of the *DTNBP1* variants reported in the current study were found in introns, having only detected a single 5' UTR variant and three downstream variants. Of note, most of these functional variants were only found in patients. LD analysis showed weak linkage among the introns detected in the present study and association analysis between case and control yield no significance. Although most of the genetic studies on *DTNBP1* mainly reported on intronic regions (Voisey *et al.,* 2010; Tan *et al.,* 2015), the current study showed more promising result in the functional regions which could provide a better understanding to the pathophysiology of schizophrenia.

In the present study, the 5' UTR variant, rs11558324 can only be detected in patients and was predicted to affect protein coding of exon 1. This is in accordance with a large Korean cohort study whereby polymorphism of this variant was shown to influence cortical mRNA levels through protein coding regulation of exon 1 (Pae *et al.,* 2009). Reduced *DTNBP1* expression

in the prefrontal cortex and glutamatergic terminals of the hippocampal formation were often reported in schizophrenia (Weickert *et al.*, 2004). Evidence also suggests that this negative effect of glutamatergic terminals might be related to cognitive deficit or negative symptoms of schizophrenia (Pae *et al.*, 2009). Hence the association between *DTNBP1* and schizophrenia may possibly arise from the influence of the gene on cognition

Among the three downstream variants identified in the present study, rs1047631 in the 3' UTR was reported as a prominent candidate for schizophrenia aetiology (Voisey *et al.*, 2010). It is located within a miRNA binding site, thus polymorphism in this region will affect mRNA and gene regulation (Luciano *et al.*, 2009). This is consistent with the current result where this variant was only detected in patients and functional prediction based on the free energy of miRNA-mRNA duplex and the miRanda score predicted rs1047631 risk allele C to inhibit *DTNBP1* expression through miRNA-mRNA binding. Studies reported that this allele was associated with reduced mRNA expression in cortical and prefrontal cortex of schizophrenia patients (Weickert *et al.*, 2004; Bray *et al.*, 2005). The disturbance of *DTNBP1* expression in these brain areas might affect general cognitive ability and intelligence (Burdick *et al.*, 2006).

For the other two downstream variants, the present study was the first to report occurrence of these variants in schizophrenia susceptibility. The function of rs200695686 is unknown and the single nucleotide variant located in position Chromosome 6: 15,522,879 was found to be a novel *DTNBP1*

variant (Genbank accession number: KX619615) although it were only detected in controls. A larger sample size might yield a more promising result. Although no single *DTNBP1* variant has yet to provide direct evidence for functional effects (Burdick *et al.*, 2006), it was hypothesised that variants in the 3' UTR region might contribute pathologically to the disorder due to the modification on miRNA binding (Duan *et al.*, 2007; Gong *et al.*, 2013).

Current results on neurocognitive analysis of *DTNBP1* variants showed that patients carrying the risk variants performed below average in processing speed and executive functions, but not to the extent of deficiency. Similar findings were reported where the negative influences of *DTNBP1* on executive functions such as attention and vigilance have been regarded as phenotypes of schizophrenia (Baek *et al.*, 2012). *DTNBP1* is more prone to exert effects in old age and might have stronger influence on specific cognitive phenotype such as memory, executive function and speed tasks, compared to other general cognitive deficiencies (Luciano *et al.*, 2009). Taking into consideration of the current association and neurocognitive results of *DTNBP1* variants, the 5' and 3' regulatory regions of *DTNBP1* are identified to confer greater risk to schizophrenia susceptibility while playing a milder role in influence neurocognitive phenotypes, particularly in the domain of processing speed and executive functions of the disorder.

5.2.3 *AKT1*

In this study, the 3' region in *AKT1* was shown to be highly associated to schizophrenia. This association was observed in a 739 bp gene region located between rs1784682 and rs17846829, spanning a total seven variants found to be in LD with each other. Of note, all of these 3' UTR variations were only found in patients, suggesting the 3' end of *AKT1* as a hot spot region for schizophrenia susceptibility. On a similar note, a recent study on the *AKT1* isoform, *AKT3*, identified a putative functional variant (rs14403) in the 3' UTR region as genome-wide associated to schizophrenia and might impact the development of prefrontal cortical-mediated cognitive function (Howell *et al.*, 2017).

Generally, 3' UTRs of mRNAs interact with specific miRNAs, which negatively regulate protein-coding gene expression by translational inhibition (Mellios and Sur, 2012). In the current study, functional prediction analysis based on the free energy of miRNA-mRNA duplex and the miRanda score showed that four of the 3' UTR variants were predicted to inhibit gene expression through miRNA-mRNA binding. Not much is known for these *AKT1* downstream variants and it is difficult to trace their individual functional influence (Beveridge and Cairns, 2012). However, it was implied that rather than individual miRNA expression, accumulation of miRNA expressions as a whole might provide a more convergent effect that influence the magnitude of risk to schizophrenia (Beveridge *et al.*, 2010). Although 3' UTR variants are not common in psychiatric diseases except in autism

(Shibayama *et al.*, 2004), schizophrenia studies reported that dysregulation of certain miRNAs might be linked to the pathophysiology of the disorder (Hansen *et al.*, 2007; Beveridge and Cairns, 2012).

AKT1 has 16 exons and in the current study, two exonic variants, rs20021361 (Ser477Ser) and rs1130233 (Glu242Glu), were found to have mutation but none marked statistical significance. The functional role of rs20021361 was not reported elsewhere, but it is possible that this exonic variant might be linked to the 3' UTR hot spot located merely 133 bp apart and contribute to the risk of schizophrenia and neurocognition. Rs1130233, on the other hand, was found to reduce the expression of *AKT1* hence altering the function of neuronal cells vital for cognitive functions (Arguello and Gogos, 2008). In neuroimaging genetics, rs1130233 was found to have strong influence on brain structure abnormalities associated to schizophrenia. Variation in rs1130233 expression alters the frontal lobe functions, thus affecting IQ and executive functions. Reduced gray matter volumes in the basal ganglia and prefrontal cortex were also found in patients, conferring risk in schizophrenia (Tan *et al.*, 2008). Besides that, rs1130233 was reported to exert some effect on facial emotion perception, which is an important phenotype in schizophrenia social cognition (Chuang *et al.*, 2014).

Among the *AKT1* variants found in patients and controls, rs2494738 showed marginal significance. Little is known about this variant except that it lies in the intronic region. It is often reported together with its downstream neighbour rs3730358, being 279 bp apart and reported to have strong linkage

(D'=1) with each other (Norton *et al.*, 2007; Zai *et al.*, 2008). In the current sample population, rs3730358 was reported as a low-frequency variant (minor allele frequency < 5%) but did not show significant association or linkage with rs2494738. The negative results might be due to the small familial sample size in the current study. Nonetheless, the results may serve as an indicator as both variants portrayed diverse effect according to different medical conditions, but generally they serve as a marker complementing for each other. For instance, both rs2494738 and rs3730358 showed promising association with Tardive Dyskinesia in African-American populations, although replication with a larger sample size was required (Zai *et al.*, 2008). However, in thyroid cancer, rs2494738 mutation showed no association with tumour invasion while alteration in the adjacent rs3730358 was related to malignant tumour (Silva, 2012). Similarly, rs3730358 was reported to confer higher risk to develop late-onset depression but not rs2494738 (Pereira *et al.*, 2014). This supports that the two variants were disease-specific while being closely related to each other. The mechanism on how these two variants work with each other is still unknown but it was believed that both variants have a small contributing effect to the risk of schizophrenia.

Approximately 45% of the variants detected in the current study were only found in patients. Overall results showed that patients having minor alleles of these variants appear to performed poorer in TMT-A compared to TMT-B, although both were below average. This finding proposed the influence of *AKT1* polymorphism on both processing speed and executive performance, which is consistent with the implication of *AKT1* on cognitive

factors such as tasks processing speed and executive control with relation to brain injury in prefrontal-striatal structure (Tan *et al.*, 2008). In mice, *AKT1* deficiency was associated with abnormal prefrontal cortex structure and functions, which resulted in poor cognitive performances (Lai *et al.*, 2006). Moreover, evidence also showed that genetic factor contributes to the cognitive performances tested in TMT and it represents the phenotypic link between schizophrenia and cognition (Owens *et al.*, 2011). Due to the supportive fact that *AKT1* was identified as a schizophrenia risk gene, this implicates that the *AKT1* gene might play an important role in influencing neurocognitive deficits in schizophrenia. Also, the current results where the majority of the *AKT1* functional variants detected were found in 3' UTR implicate evidence that this region contributes to the risk of schizophrenia and neurocognition, especially in the domain of processing speed.

5.2.4 *TCF4*

The *TCF4* variants detected in the current study mainly lies in the 3' UTR and intron region, with a single exonic variant. Similarly, a recent study also reported that most schizophrenia-risk variants in *TCF4* were located in introns and intragenic downstream region (Badowska, 2015). In the downstream region, thirteen 3' UTR variants were identified where 15% were only found in patients. Apart from that, only rs141970461 showed marginal significant difference in allele frequency between patient and control. This was the first case where rs141970461 was reported to be associated to

schizophrenia but its function is yet to be known. Approximately 2kb apart, an allele C insertion in rs369444573 was predicted to inhibit miRNA-mRNA binding and promote gene expression of *TCF4*. This risk allele C showed higher frequencies in patients compared to controls. This suggests that increased *TCF4* level in patients might be associated with schizophrenia. In supporting this notion, elevated peripheral *TCF4* mRNA levels was also found in blood (Wirgenes *et al.*, 2012) and postmortem brain tissues (Guella *et al.*, 2013) of schizophrenic patients. Thus, the identification of these 3' variants may serve to influence the susceptibility of the disorder.

The non-coding variant, rs2958182 was found to be significantly associated to schizophrenia in the current study. It is often reported as a prominent schizophrenia risk variant (Li *et al.*, 2010; Steinberg *et al.*, 2011) and shown to influence cognitive functions, especially intelligence and attention functions (Zhu *et al.*, 2013). Located in intron 3, rs2958182 lies approximately 6 kb apart from the robustly studied schizophrenia risk variant rs9960767. Both variants were in high linkage disequilibrium in the Chinese population and were often reported together as interrelated variants susceptible for risk of schizophrenia (Li *et al.*, 2010). The current sample population did not show linkage between them. In addition, the current study reported rs2958182 risk allele to be the T allele, and not the A allele, which was linked to the risk C allele of rs9960767 in European population (Zhu *et al.*, 2013; Hui *et al.*, 2015). Similarly, Li *et al.* (2010) reported T allele as risk allele and suggested rs2958182 as a population-specific risk variant (Li *et al.*, 2010). Population stratification might lead to the considerable differences in genetic

distribution. Although the specific risk alleles are different, rs2958182 appears to be relevant for schizophrenia in both populations. Results demonstrated association of rs2958182 with lower processing speed and executive functioning in schizophrenia patients and their affected family members. Besides that, rs2958182 was shown to have interaction effects with cognitive intelligence and attention functions, but no effect on working memory in patients (Zhu *et al.*, 2013).

Current study found an exonic variation in rs8766 (Ser643Ser), but it appeared not to be significantly associated to schizophrenia. In accordance with this, a recent report showed that rs8766 was not associated to schizophrenia susceptibility (Basmanav *et al.*, 2015). Located within the last coding exon of *TCF4*, rs8766 is a synonymous coding variant, where overexpression of the C allele was found in the post mortem brain of patients with neurodegenerative disorder (Buonocore *et al.*, 2010). It was suggested that rs8766 does not exerts its effect through regulation of *TCF4* transcript, thus the functional mechanism responsible for the association remains unknown (Williams *et al.*, 2011a).

In the current study, patients with the *TCF4* risk alleles exhibited below average performance in both processing speed and executive controls. *TCF4* is often implicated in executive functioning such as reasoning and problem-solving deficits as seen in schizophrenia (Albanna *et al.*, 2014). Besides that, *TCF4* was also associated with earlier schizophrenia onset which is correlated with more severe cognitive impairments (Chow *et al.*, 2016).

Previous studies have shown that both supressed and overexpressed *TCF4* level will lead to impaired cognitive functions. For instance, under-expression of the gene causes mental retardation (Flora *et al.*, 2007), while overexpression of *TCF4* gene in mice resulted in reduced memory (Brzozka *et al.*, 2010). The current study implicated increased level of *TCF4* in schizophrenia, proposing the influence of *TCF4* on processing speed and executive performance impairments observed in patients.

The strength of this study is the disease-specific family cohort design. The subjects consisted of first-degree familial samples, hence strong genetic inheritance and familial risk effect could be observed. This serves as an advantage for studying gene-gene interactions and limits the effect of the unmeasured environmental influences (Hopper, 2011).

The limitation to address is the small sample size that limits the possibility to draw a projection on the effects of genetic variants and cognition. Some seriously impaired patients failed to complete one or more of the test in this study and had to be excluded from the analysis, hence the small familial sample size. Some study reported that family-based association studies are relatively rare in psychiatric genetics, mainly due to difficulties in obtaining samples (Alkelai *et al.*, 2011). Although much effort have been made in recruiting a larger cohort for the current study, there are very limited resources and data history in Malaysia (Chee and Abdul Aziz, 2014), especially for familial samples. The different ethnic background of subjects also serves as a

limitation as it increases the level of heterogeneity. In this study, patients and controls were matched in terms of ethnicity in order to reduce this effect.

Besides that, the positive detecting rate of the genes as seen in the results is very high. This is mainly contributed by the strong genetic load observed in familial patient cohort and the small sample size limitation. Although NGS was employed, the current study only targeted specific regions of four genes, hence it is difficult to detect rare mutations. Future study using NGS exome sequencing will yield a better chance for rare mutation discovery.

CHAPTER 6

CONCLUSION

Overall findings implicated evidence for association of *ZNF804A*, *DTNBP1*, *AKT1*, and *TCF4* variants with the risk of schizophrenia and neurocognition. The 5'UTR and exon 4 regions of *ZNF804A*, the 5' and 3' regulatory regions of *DTNBP1*, and the 3' UTR of *AKT1* and *TCF4* were identified as contributors to schizophrenia susceptibility and impairments in neurocognitive phenotypes of processing speed and executive functions. In conclusion, the genetic expression pathways that underlie cognitive deficits in schizophrenia may lie within the regulatory regions of these genes. This supports the hypothesis of the current study that *ZNF804A*, *DTNBP1*, *AKT1*, and *TCF4* variants might confer risk to schizophrenia and cognitive deficits. As impaired cognitive functions are being recognised increasingly as core features of schizophrenia, these findings may serve as additional support to better understand neurocognition in schizophrenia, hence facilitating disease prevention and treatment study.

REFERENCES

Addington, A.M., et al., 2004. Polymorphisms in the 13q33.2 gene G72/G30 are associated with childhood-onset schizophrenia and psychosis not otherwise specified. *Biological Psychiatry*, 55 (10), pp. 976-980.

Addington, A.M., et al., 2007. Neuregulin 1 (8p12) and childhood-onset schizophrenia: susceptibility haplotypes for diagnosis and brain developmental trajectories. *Molecular Psychiatry*, 12 (2), pp. 195-205.

Albanna, A., et al., 2014. TCF4 gene polymorphism and cognitive performance in patients with first episode psychosis. *Schizophrenia Research*, 152 (1), pp. 124-129.

Alkelai, A., et al., 2011. Identification of new schizophrenia susceptibility loci in an ethnically homogeneous, family-based, Arab-Israeli sample. *FASEB Journal*, 25 (11), pp. 4011-4023.

Andre, J., Picchioni, M., Zhang, R. and Toulopoulou, T., 2016. Working memory circuit as a function of increasing age in healthy adolescence: A systematic review and meta-analyses. *Neuroimage Clinical*, 12, pp. 940-948.

Arguello, P.A. and Gogos, J.A., 2008. A signaling pathway AKTing up in schizophrenia. *Journal of Clinical Investigation*, 118 (6), pp. 2018-2021.

Aziz, A.A., et al., 2008. The National Mental Health Registry (NMHR). *The Medical Journal of Malaysia*, 63 Suppl C, pp. 15-17.

Badowska, D., 2015. Schiophrenia risk factor TCF4 and gene x environment interaction in mice. PhD thesis, University of Göttingen.

Baek, J.H., et al., 2012. Association of genetic variations in DTNBP1 with cognitive function in schizophrenia patients and healthy subjects. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 159B (7), pp. 841-849.

Bakanidze, G., et al., 2016. Association of Dystrobrevin-Binding Protein 1 Polymorphisms with Sustained Attention and Set-Shifting in Schizophrenia Patients. *Neuropsychobiology*, 74 (1), pp. 41-47.

Balog, Z., Kiss, I. and Keri, S., 2011. ZNF804A may be associated with executive control of attention. *Genes, Brain and Behavior*, 10 (2), pp. 223-227.

Barbato, A., 1998. *Schizophrenia and public health*. (WHO Publication No. WHO/MSA/NAM/97.6). Geneva: World Health Organization.

Basmanav, F.B., et al., 2015. Investigation of the role of TCF4 rare sequence variants in schizophrenia. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 168B (5), pp. 354-362.

Beaulieu, J.M., 2012. A role for Akt and glycogen synthase kinase-3 as integrators of dopamine and serotonin neurotransmission in mental health. *Journal of Psychiatry & Neuroscience*, 37 (1), pp. 7-16.

Benjamini, Y. and Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57 (1), pp. 289-300.

Benson, M.A., Newey, S.E., Martin-Rendon, E., Hawkes, R. and Blake, D.J., 2001. Dysbindin, a Novel Coiled-coil-containing Protein That Interacts with the Dystrobrevins in Muscle and Brain. *Journal of Biological Chemistry*, 276 (26), pp. 24232-24241.

Bergen, S.E., et al., 2014. Genetic modifiers and subtypes in schizophrenia: investigations of age at onset, severity, sex and family history. *Schizophrenia Research*, 154 (1-3), pp. 48-53.

Beveridge, N.J. and Cairns, M.J., 2012. MicroRNA dysregulation in schizophrenia. *Neurobiology of Disease*, 46 (2), pp. 263-271.

Beveridge, N.J., Gardiner, E., Carroll, A.P., Tooney, P.A. and Cairns, M.J., 2010. Schizophrenia is associated with an increase in cortical microRNA biogenesis. *Molecular Psychiatry*, 15 (12), pp. 1176-1189.

Bilder, R.M., 2006. 'Schizophrenia' in P.J., Snyder, P.D., Nussbum, and D.L., Robins, (eds.), *Clinical Neuropsychology: A Pocket Handbook for Assessment*, pp. 398-414. Washington: American Psychological Association.

Bilder, R.M., Howe, A., Novak, N., Sabb, F.W. and Parker, D.S., 2011. The genetics of cognitive impairment in schizophrenia: a phenomic perspective. *Trends in Cognitive Sciences*, 15 (9), pp. 428-435.

Blake, D.J., et al., 2010. TCF4, schizophrenia, and Pitt-Hopkins Syndrome. *Schizophrenia Bulletin*, 36 (3), pp. 443-447.

Boerma, M.A., Van Der Stel, J.C., Van Amelsvoort, T., Linszen, D.H. and De Haan, L., 2010. Women, schizophrenia and oestrogen; neurobiological hypotheses and hormonetherapy studies. *Tijdschrift Voor Psychiatrie*, 52 (4), pp. 235-244.

Borovcanin, M., et al., 2016. Possible Role of TGF – B Pathways in Schizophrenia / Moguća Uloga TTGF - B Signalnih Puteva U Shizofreniji. *Serbian Journal of Experimental and Clinical Research*, 17 (1), pp. 3-8.

Bray, N.J., Leweke, F.M., Kapur, S. and Meyer-Lindenberg, A., 2010. The neurobiology of schizophrenia: new leads and avenues for treatment. *Current Opinion in Neurobiology*, 20 (6), pp. 810-815.

Bray, N.J., et al., 2005. Haplotypes at the dystrobrevin binding protein 1 (DTNBP1) gene locus mediate risk for schizophrenia through reduced DTNBP1 expression. *Human molecular genetics*, 14 (14), pp. 1947-1954.

Brown, A.S. and Mcgrath, J.J., 2011. The prevention of schizophrenia. *Schizophrenia Bulletin*, 37 (2), pp. 257-261.

Brzozka, M.M., Radyushkin, K., Wichert, S.P., Ehrenreich, H. and Rossner, M.J., 2010. Cognitive and sensorimotor gating impairments in transgenic mice overexpressing the schizophrenia susceptibility gene Tcf4 in the brain. *Biological Psychiatry*, 68 (1), pp. 33-40.

Buonocore, F., et al., 2010. Effects of cis-regulatory variation differ across regions of the adult human brain. *Human Molecular Genetics*, 19 (22), pp. 4490-4496.

Burdick, K.E., et al., 2007. DTNBP1 genotype influences cognitive decline in schizophrenia. *Schizophrenia Research*, 89 (1-3), pp. 169-172.

Burdick, K.E., et al., 2006. Genetic variation in DTNBP1 influences general cognitive ability. *Human Molecular Genetics*, 15 (10), pp. 1563-1568.

Cannon, M., et al., 1999. School Performance in Finnish Children and Later Development of Schizophrenia. *Archives of General Psychiatry*, 56 (5), pp. 457.

Cannon, T.D., et al., 2000. Childhood cognitive functioning in schizophrenia patients and their unaffected siblings: a prospective cohort study. *Schizophrenia Bulletin*, 26 (2), pp. 379-393.

Chaudhry, A.Z., Lyons, G.E. and Gronostajski, R.M., 1997. Expression patterns of the four nuclear factor I genes during mouse embryogenesis indicate a potential role in development. *Developmental Dynamics*, 208 (3), pp. 313-325.

Chee, K.Y., 2009. Outcome study of first-episode schizophrenia in a developing country: quality of life and antipsychotics. *Social Psychiatry and Psychiatric Epidemiology*, 44 (2), pp. 143-150.

Chee, K.Y. and Abdul Aziz, S., 2014. A review of schizophrenia research in malaysia. *The Medical Journal of Malaysia*, 69 Suppl A, pp. 46-54.

Chen, M., et al., 2012. Evidence of IQ-Modulated Association Between ZNF804A Gene Polymorphism and Cognitive Function in Schizophrenia Patients. *Neuropsychopharmacology*, 37 (7), pp. 1572-1578.

Chow, T.J., Tee, S.F., Yong, H.S. and Tang, P.Y., 2016. Genetic Association of TCF4 and AKT1 Gene Variants with the Age at Onset of Schizophrenia. *Neuropsychobiology*, 73 (4), pp. 233-240.

Chuang, L.Y., et al., 2014. Identification of SNP barcode biomarkers for genes associated with facial emotion perception using particle swarm optimization algorithm. *Annals of General Psychiatry*, 13, pp. 15.

Chung, H.J., et al., 2010. Mouse Homologue of the Schizophrenia Susceptibility Gene ZNF804A as a Target of Hoxc8. *Journal of Biomedicine and Biotechnology*, 2010, pp. 231708.

Cohen, A.S. and Docherty, N.M., 2004. Deficit Versus Negative Syndrome in Schizophrenia: Prediction of Attentional Impairment. *Schizophrenia Bulletin*, 30 (4), pp. 827-835.

Correa, D.D., et al., 2016. COMT,BDNF, andDTNBP1polymorphisms and cognitive functions in patients with brain tumors. *Neuro-Oncology*, 18 (10), pp. 1425-1433.

Corvin, A.P., 2011. Two patients walk into a clinic...a genomics perspective on the future of schizophrenia. *BMC Biology*, 9 (1), pp. 77.

De Leng, W.W., et al., 2016. Targeted Next Generation Sequencing as a Reliable Diagnostic Assay for the Detection of Somatic Mutations in Tumours Using Minimal DNA Amounts from Formalin Fixed Paraffin Embedded Material. *PLoS One*, 11 (2), pp. e0149405.

De Luca, V., Voineskos, D., Shinkai, T., Wong, G. and Kennedy, J.L., 2005. Untranslated region haplotype in dysbindin gene: analysis in schizophrenia. *Journal of Neural Transmission*, 112 (9), pp. 1263-1267.

DeLisi, L.E., 2016. A case for returning to multiplex families for further understanding the heritability of schizophrenia: A psychiatrist's perspective. *Molecular Neuropsychiatry*, 2(1), pp. 15-19.

Demakis, G.J., 2004. Frontal lobe damage and tests of executive processing: a meta-analysis of the category test, stroop test, and trail-making test. *Journal of Clinical and Experimental Neuropsychology*, 26 (3), pp. 441-450.

Dick, D.M., 2011. Gene-Environment Interaction in Psychological Traits and Disorders. *Annual Review of Clinical Psychology*, 7, pp. 383-409.

Donohoe, G., Morris, D.W. and Corvin, A., 2010. The Psychosis Susceptibility Gene ZNF804A: Associations, Functions, and Phenotypes. *Schizophrenia Bulletin*, 36 (5), pp. 904-909.

Donohoe, G., et al., 2011. ZNF804A risk allele is associated with relatively intact gray matter volume in patients with schizophrenia. *NeuroImage*, 54 (3), pp. 2132-2137.

Duan, J., et al., 2007. DTNBP1 (Dystrobrevin Binding Protein 1) and Schizophrenia: Association Evidence in the 3′ End of the Gene. *Human Heredity*, 64 (2), pp. 97-106.

Emamian, E.S., 2012. AKT/GSK3 signaling pathway and schizophrenia. *Frontiers in Molecular Neuroscience*, 5, pp. 33.

Emamian, E.S., Hall, D., Birnbaum, M.J., Karayiorgou, M. and Gogos, J.A., 2004. Convergent evidence for impaired AKT1-GSK3beta signaling in schizophrenia. *Nature Genetics*, 36 (2), pp. 131-137.

Erhart, S.M., Marder, S.R. and Carpenter, W.T., 2006. Treatment of Schizophrenia Negative Symptoms: Future Prospects. *Schizophrenia Bulletin*, 32 (2), pp. 234-237.

Esslinger, C., et al., 2011. Cognitive state and connectivity effects of the genome-wide significant psychosis variant in ZNF804A. *NeuroImage*, 54 (3), pp. 2514-2523.

Fallgatter, A.J., et al., 2010. DTNBP1 (dysbindin) gene variants modulate prefrontal brain function in schizophrenic patients - support for the glutamate hypothesis of schizophrenias. *Genes, Brain and Behavior*, pp. 1-9.

Fang, D., et al., 2007. Phosphorylation of beta-catenin by AKT promotes beta-catenin transcriptional activity. *The Journal of Biological Chemistry*, 282 (15), pp. 11221-11229.

Flora, A., Garcia, J.J., Thaller, C. and Zoghbi, H.Y., 2007. The E-protein Tcf4 interacts with Math1 to regulate differentiation of a specific subset of neuronal progenitors. *Proceedings of the National Academy of Sciences USA*, 104 (39), pp. 15382-15387.

Forrest, M.P., Hill, M.J., Quantock, A.J., Martin-Rendon, E. and Blake, D.J., 2014. The emerging roles of TCF4 in disease and development. *Trends in Molecular Medicine*, 20 (6), pp. 322-331.

Forrest, M.P., Waite, A.J., Martin-Rendon, E. and Blake, D.J., 2013. Knockdown of human TCF4 affects multiple signaling pathways involved in cell survival, epithelial to mesenchymal transition and neuronal differentiation. *PLoS One*, 8 (8), pp. e73169.

Freedman, R., 2003. Schizophrenia. *The New England Journal of Medicine*, 349 (18), pp. 1738-1749.

Freyberg, Z., Ferrando, S.J. and Javitch, J.A., 2010. Roles of the Akt/GSK-3 and Wnt signaling pathways in schizophrenia and antipsychotic drug action. *The American Journal of Psychiatry*, 167 (4), pp. 388-396.

Gaysina, D., et al., 2009. Association of the dystrobrevin binding protein 1 gene (DTNBP1) in a bipolar case-control study (BACCS). *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 150B (6), pp. 836-844.

Girgenti, M.J., Loturco, J.J. and Maher, B.J., 2012. ZNF804a regulates expression of the schizophrenia-associated genes PRSS16, COMT, PDE4B, and DRD2. *PLoS One*, 7 (2), pp. e32404.

Gong, Y., et al., 2013. Polymorphisms in microRNA target sites influence susceptibility to schizophrenia by altering the binding of miRNAs to their targets. *European Neuropsychopharmacology*, 23 (10), pp. 1182-1189.

Goodwin, S., Mcpherson, J.D. and Mccombie, W.R., 2016. Coming of age: ten years of next-generation sequencing technologies. *Nature Review Genetics*, 17 (6), pp. 333-351.

Gornick, M.C., et al., 2005. Dysbindin (DTNBP1, 6p22.3) is associated with childhood-onset psychosis and endophenotypes measured by the Premorbid Adjustment Scale (PAS). *Journal of Autism and Developmental Disorders*, 35 (6), pp. 831-838.

Green, M.F., et al., 2004. Approaching a consensus cognitive battery for clinical trials in schizophrenia: the NIMH-MATRICS conference to select cognitive domains and test criteria. *Biological Psychiatry*, 56 (5), pp. 301-307.

Guella, I., et al., 2013. Analysis of miR-137 expression and rs1625579 in dorsolateral prefrontal cortex. *Journal of Psychiatry Research*, 47 (9), pp. 1215-1221.

Guo, A.Y., et al., 2009. The dystrobrevin-binding protein 1 gene: features and networks. *Molecular Psychiatry*, 14 (1), pp. 18-29.

Gur, R.E., et al., 2007. The Consortium on the Genetics of Schizophrenia: neurocognitive endophenotypes. *Schizophrenia Bulletin*, 33 (1), pp. 49-68.

Hallmayer, J., 2004. Getting our AKT together in schizophrenia? *Nature Genetics*, 36 (2), pp. 115-116.

Hansen, T., et al., 2007. Brain expressed microRNAs implicated in schizophrenia etiology. *PLoS One*, 2 (9), pp. e873.

Harmon, B.T., et al., 2010. Functional characterization of a haplotype in the AKT1 gene associated with glucose homeostasis and metabolic syndrome. *Human Genetics*, 128 (6), pp. 635-645.

Harvey, P.D., 2008. The Genetics of Cognitive Impairment in Schizophrenia. *Psychiatry (Edgmont)*, 5 (6), pp. 65-67.

Hellmich, C., et al., 2015. Genetics, sleep and memory: a recall-by-genotype study of ZNF804A variants and sleep neurophysiology. *BMC Medical Genetics*, 16, pp. 96.

Hill, M.J., Forrest, M.P., Martin-Rendon, E. and Blake, D.J., 2014. Association of Transcription Factor 4 (TCF4) variants with schizophrenia and intellectual disability. *Current Behavioral Neuroscience Reports*, 1 (4), pp. 206-214.

Hill, M.J., Jeffries, A.R., Dobson, R.J.B., Price, J. and Bray, N.J., 2011. Knockdown of the psychosis susceptibility gene ZNF804A alters expression

of genes involved in cell adhesion. *Human Molecular Genetics*, 21 (5), pp. 1018-1024.

Hoff, A.L., et al., 1996. A neuropsychological study of early onset schizophrenia. *Schizophrenia Research*, 20 (1), pp. 21-28.

Hopper, J.L., 2011. Disease-specific prospective family study cohorts enriched for familial risk. *Epidemiologic Perspectives & Innovations*, 8 (1), pp. 2.

Howell, K.R., Floyd, K. and Law, A.J., 2017. PKBgamma/AKT3 loss-of-function causes learning and memory deficits and deregulation of AKT/mTORC2 signaling: Relevance for schizophrenia. *PLoS One*, 12 (5), pp. e0175993.

Hui, L., et al., 2015. TCF4 gene polymorphism is associated with cognition in patients with schizophrenia and healthy controls. *Journal of Psychiatry Research*, 69, pp. 95-101.

Hyman, S.E. and Fenton, W.S., 2003. Medicine. What are the right targets for psychopharmacology? *Science*, 299 (5605), pp. 350-351.

Institute for Public Health, 2015. *National Health and Morbidity Survey 2015. Vol. II: Non-Communicable Diseases, Risk Factors & Other Health Problems*. (MOH Publication No. MOH/S/IKU/52.15). Kuala Lumpur: Ministry of Health

Jansen, A., et al., 2009. Effect of the G72 (DAOA) putative risk haplotype on cognitive functions in healthy subjects. *BMC Psychiatry*, 9, pp. 60.

Johnson, M.B., et al., 2009. Functional and evolutionary insights into human brain development through global transcriptome analysis. *Neuron*, 62 (4), pp. 494-509.

Jorde, L.B., Watkins, W.S. and Bamshad, M.J., 2001. Population genomics: a bridge from evolutionary history to genetic medicine. *Human Molecular Genetics*, 10 (20), pp. 2199-2207.

Jun, G.R., et al., 2017. Transethnic genome-wide scan identifies novel Alzheimer's disease loci. *Alzheimers & Dementia*, 13 (7), pp. 727-738.

Karege, F., et al., 2012. Genetic overlap between schizophrenia and bipolar disorder: a study with AKT1 gene variants and clinical phenotypes. *Schizophrenia Research*, 135 (1-3), pp. 8-14.

Keefe, R.S., et al., 2007. Neurocognitive effects of antipsychotic medications in patients with chronic schizophrenia in the CATIE Trial. *Archives of General Psychiatry*, 64 (6), pp. 633-647.

Keefe, R.S. and Harvey, P.D., 2012. Cognitive impairment in schizophrenia. *Handbook of Experimental Pharmacology*, (213), pp. 11-37.

Koponen, H., et al., 2004. Childhood central nervous system infections and risk for schizophrenia. *European Archives of Psychiatry and Clinical Neuroscience*, 254 (1), pp. 9-13.

Kuperberg, G. and Heckers, S., 2000. Schizophrenia and cognitive function. *Current Opinion in Neurobiology*, 10 (2), pp. 205-210.

Lai, W.S., et al., 2006. Akt1 deficiency affects neuronal morphology and predisposes to abnormalities in prefrontal cortex functioning. *Proceedings of the National Academy of Sciences USA*, 103 (45), pp. 16906-16911.

Lamani, E., et al., 2009. Tissue- and cell-specific alternative splicing of NFIC. *Cells Tissues Organs*, 189 (1-4), pp. 105-110.

Lange, R.T., Iverson, G.L., Zakrzewski, M.J., Ethel-King, P.E. and Franzen, M.D., 2005. Interpreting the trail making test following traumatic brain injury: comparison of traditional time scores and derived indices. *Journal of Clinical and Experimental Neuropsychology*, 27 (7), pp. 897-906.

Laviolette, S.R., 2007. Dopamine Modulation of Emotional Processing in Cortical and Subcortical Neural Circuits: Evidence for a Final Common Pathway in Schizophrenia? *Schizophrenia Bulletin*, 33 (4), pp. 971-981.

Lee, K.Y., et al., 2010. No association between AKT1 polymorphism and schizophrenia: a case-control study in a Korean population and a meta-analysis. *Neuroscience Research*, 66 (3), pp. 238-245.

Lee, L.C., et al., 2012a. Role of the CCAAT-binding protein NFY in SCA17 pathogenesis. *PLoS One*, 7 (4), pp. e35302.

Lee, S.H., et al., 2012b. Estimating the proportion of variation in susceptibility to schizophrenia captured by common SNPs. *Nat Genet*, 44 (3), pp. 247-250.

Lencz, T., et al., 2010. A Schizophrenia Risk Gene, ZNF804A, Influences Neuroanatomical and Neurocognitive Phenotypes. *Neuropsychopharmacology*, 35 (11), pp. 2284-2291.

Lennertz, L., et al., 2011. Impact of TCF4 on the genetics of schizophrenia. *European Archives of Psychiatry and Clinical Neuroscience*, 261 Suppl 2, pp. S161-165.

Lesh, T.A., Niendam, T.A., Minzenberg, M.J. and Carter, C.S., 2011. Cognitive control deficits in schizophrenia: mechanisms and meaning. *Neuropsychopharmacology*, 36 (1), pp. 316-338.

Lewis, D.A. and Lieberman, J.A., 2000. Catching up on schizophrenia: natural history and neurobiology. *Neuron*, 28 (2), pp. 325-334.

Li, M., et al., 2011. Allelic differences between Han Chinese and Europeans for functional variants in ZNF804A and their association with schizophrenia. *American Journal of Psychiatry*, 168 (12), pp. 1318-1325.

Li, T., et al., 2010. Common variants in major histocompatibility complex region and TCF4 gene are significantly associated with schizophrenia in Han Chinese. *Biological Psychiatry*, 68 (7), pp. 671-673.

Liu, C.-M., et al., 2007. No association evidence between schizophrenia and dystrobrevin-binding protein 1 (DTNBP1) in Taiwanese families. *Schizophrenia Research*, 93 (1-3), pp. 391-398.

Liu, L., et al., 2016. The mRNA expression of DRD2, PI3KCB, and AKT1 in the blood of acute schizophrenia patients. *Psychiatry Research*, 243, pp. 397-402.

Luciano, M., et al., 2009. Variation in the dysbindin gene and normal cognitive function in three independent population samples. *Genes, Brain and Behavior*, 8 (2), pp. 218-227.

Maduro, V., et al., 2016. Complex translocation disrupting TCF4 and altering TCF4 isoform expression segregates as mild autosomal dominant intellectual disability. *Orphanet Journal of Rare Diseases*, 11 (1), pp. 62.

Mallas, E.-J., et al., 2016. Genome-wide discovered psychosis-risk gene ZNF804A impacts on white matter microstructure in health, schizophrenia and bipolar disorder. *PeerJ*, 4, pp. e1570.

Mardis, E.R., 2008. The impact of next-generation sequencing technology on genetics. *Trends in Genetics*, 24 (3), pp. 133-141.

Mardis, E.R., 2017. DNA sequencing technologies: 2006-2016. *Nature Protocols*, 12 (2), pp. 213-218.

Margari, F., et al., 2008. Very early onset and greater vulnerability in schizophrenia: A clinical and neuroimaging study. *Neuropsychiatric Disease and Treatment*, 4 (4), pp. 825-830.

Maric, N.P. and Svrakic, D.M., 2012. Why schizophrenia genetics needs epigenetics: a review. *Psychiatria Danubina*, 24 (1), pp. 2-18.

Mark, W. and Toulopoulou, T., 2016. Cognitive intermediate phenotype and genetic risk for psychosis. *Current Opinion in Neurobiology*, 36, pp. 23-30.

Mcgrath, J., et al., 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*, 2 (1), pp. 13.

Mellios, N. and Sur, M., 2012. Emerging Role of microRNAs in schizophrenia and autism spectrum disorders. *Frontiers in Psychiatry*, 3, pp. 39.

Ministry of Health, 2009. *Clinical practice guidelines*. (MOH Publication No. MOH/P/PAK/185.09). Putrajaya: Ministry of Health Malaysia

Ministry of Health, 2014. *Malaysian Statistics on Medicine 2009 & 2010*. Kuala Lumpur: Pharmaceutical Services Division and Clinical Research Centre.

Morgan, C., Charalambides, M., Hutchinson, G. and Murray, R.M., 2010. Migration, ethnicity, and psychosis: toward a sociodevelopmental model. *Schizophrenia Bulletin*, 36 (4), pp. 655-664.

Nasrallah, H.A. and Smeltzer, D.J. 2002. *Contemporary Diagnosis and Management of the Patient with Schizophrenia*, Handbooks in Health Care Company.

National Institute of Mental Health, 2009. *Schizophrenia*. (NIH Publication No. 09-3517). Bethesda, MD: U.S. Government Printing Office.

Navarrete, K., et al., 2012. TCF4 (e2-2; ITF2): a schizophrenia-associated gene with pleiotropic effects on human disease. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 162B (1), pp. 1-16.

Need, A.C., et al., 2009. A genome-wide investigation of SNPs and CNVs in schizophrenia. *PLoS Genetics*, 5 (2), pp. e1000373.

Nicodemus, K.K., et al., 2014. Category fluency, latent semantic analysis and schizophrenia: a candidate gene approach. *Cortex*, 55, pp. 182-191.

Norton, N., et al., 2007. Association analysis of AKT1 and schizophrenia in a UK case control sample. *Schizophrenia Research*, 93 (1-3), pp. 58-65.

Numakawa, T., 2004. Evidence of novel neuronal functions of dysbindin, a susceptibility gene for schizophrenia. *Human Molecular Genetics*, 13 (21), pp. 2699-2708.

Numakawa, T., et al., 2004. Evidence of novel neuronal functions of dysbindin, a susceptibility gene for schizophrenia. *Human Molecular Genetics*, 13 (21), pp. 2699-2708.

O'donovan, M.C., et al., 2008. Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nature Genetic*, 40 (9), pp. 1053-1055.

Ochoa, S., Usall, J., Cobo, J., Labad, X. and Kulkarni, J., 2012. Gender differences in schizophrenia and first-episode psychosis: a comprehensive literature review. *Schizophrenia Research and Treatment*, 2012, pp. 916198.

Ohi, K., et al., 2013. The AKT1 gene is associated with attention and brain morphology in schizophrenia. *World Journal of Biological Psychiatry*, 14 (2), pp. 100-113.

Oliver, P.L., 2011. Challenges of analysing gene-environment interactions in mouse models of schizophrenia. *Scientific World Journal*, 11, pp. 1411-1420.

Owens, S.F., et al., 2011. Genetic overlap between schizophrenia and selective components of executive function. *Schizophrenia Research*, 127 (1-3), pp. 181-187.

Pae, C.-U., et al., 2009. Dysbindin gene (DTNBP1) and schizophrenia in Korean population. *European Archives of Psychiatry and Clinical Neuroscience*, 259 (3), pp. 137-142.

Pedersen, C.B. and Mortensen, P.B., 2001. Evidence of a dose-response relationship between urbanicity during upbringing and schizophrenia risk. *Archives of General Psychiatry*, 58, pp.

Peralta, V. and Cuesta, M.J., 2001. How many and which are the psychopathological dimensions in schizophrenia? Issues influencing their ascertainment. *Schizophrenia Research*, 49 (3), pp. 269-285.

Pereira, P.A., et al., 2014. Genetic variant of AKT1 and AKTIP associated with late-onset depression in a Brazilian population. *International Journal of Geriatric Psychiatry*, 29 (4), pp. 399-405.

Perianez, J.A., et al., 2007. Trail Making Test in traumatic brain injury, schizophrenia, and normal ageing: sample comparisons and normative data. *Archives of Clinical Neuropsychology*, 22 (4), pp. 433-447.

Pietilainen, O.P., et al., 2009. Association of AKT1 with verbal learning, verbal memory, and regional cortical gray matter density in twins. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 150B (5), pp. 683-692.

Quednow, B.B., et al., 2011. The schizophrenia risk allele C of the TCF4 rs9960767 polymorphism disrupts sensorimotor gating in schizophrenia spectrum and healthy volunteers. *Journal of Neuroscience*, 31 (18), pp. 6684-6691.

Rabinowitz, J. and Fennig, S., 2002. Differences in age of first hospitalization for schizophrenia among immigrants and nonimmigrant in a national case registry. *Schizophrenia Bulletin*, 28 (3), pp. 491-499.

Rajji, T.K., Ismail, Z. and Mulsant, B.H., 2009. Age at onset and cognition in schizophrenia: meta-analysis. *British Journal of Psychiatry*, 195 (4), pp. 286-293.

Ran, C., et al., 2011. Genetic studies of the protein kinase AKT1 in Parkinson's disease. *Neuroscience Letters*, 501 (1), pp. 41-44.

Reichenberg, A., et al., 2002. A population-based cohort study of premorbid intellectual, language, and behavioral functioning in patients with schizophrenia, schizoaffective disorder, and nonpsychotic bipolar disorder. *American Journal of Psychiatry*, 159 (12), pp. 2027-2035.

Reis-Filho, J.S., 2009. Next-generation sequencing. *Breast Cancer Research*, 11 Suppl 3, pp. S12.

Reitan, R.M., 1958. Validity of the Trail Making Test as an Indicator of Organic Brain Damage. *Perceptual and Motor Skills*, 8 (3), pp. 271-276.

Riley, B. and Kendler, K.S., 2006. Molecular genetic studies of schizophrenia. *European Journal of Human Genetics*, 14 (6), pp. 669-680.

Riley, B., et al., 2010. Replication of association between schizophrenia and ZNF804A in the Irish Case–Control Study of Schizophrenia sample. *Molecular Psychiatry*, 15 (1), pp. 29-37.

Ripke, S., et al., 2011. Genome-wide association study identifies five new schizophrenia loci. *Nature Genetics*, 43 (10), pp. 969-976.

Robinson, J.T., et al., 2011. Integrative Genomics Viewer. *Nature Biotechnology*, 29 (1), pp. 24-26.

Rossler, W., Salize, H.J., Van Os, J. and Riecher-Rossler, A., 2005. Size of burden of schizophrenia and psychotic disorders. *European Neuropsychopharmacology*, 15 (4), pp. 399-409.

Sanchez-Cubillo, I., et al., 2009. Construct validity of the Trail Making Test: role of task-switching, working memory, inhibition/interference control, and visuomotor abilities. *Journal of the International Neuropsychological Society*, 15 (3), pp. 438-450.

Sand, P.G., 2013. ZNF804Aand schizophrenia: An open peer commentary. *American Journal of Medical Genetics Part B*, 162 (1), pp. 71-72.

Sanders, A.R., et al., 2008. No significant association of 14 candidate genes with schizophrenia in a large European ancestry sample: implications for psychiatric genetics. *American Journal of Psychiatry*, 165 (4), pp. 497-506.

Schreiber, M., Dorschner, M. and Tsuang, D., 2013. Next-generation sequencing in schizophrenia and other neuropsychiatric disorders. *American Journal of Medical Genetics*, 162B (7), pp. 671-678.

Schultz, C.C., et al., 2014. ZNF804A and cortical structure in schizophrenia: in vivo and postmortem studies. *Schizophrenia Bulletin*, 40 (3), pp. 532-541.

Schuster, S.C., 2008. Next-generation sequencing transforms today's biology. *Nature Methods*, 5 (1), pp. 16-18.

Schwab, S.G., et al., 2003. Support for Association of Schizophrenia with Genetic Variation in the 6p22.3 Gene, Dysbindin, in Sib-Pair Families with Linkage and in an Additional Sample of Triad Families. *American Journal of Human Genetics*, 72 (1), pp. 185-190.

Schwab, S.G., et al., 2013. Association of rs1344706 in the ZNF804A gene with schizophrenia in a case/control sample from Indonesia. *Schizophrenia Research*, 147 (1), pp. 46-52.

Schwab, S.G. and Wildenauer, D.B., 2008. Research on causes for schizophrenia: are we close? *Schizophrenia Research*, 102 (1-3), pp. 29-30.

Shibayama, A., et al., 2004. MECP2 structural and 3'-UTR variants in schizophrenia, autism and other psychiatric diseases: a possible association with autism. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 128B (1), pp. 50-53.

Silva, A. C. N., 2012. *Study of clinical utility in the identification of genetic alternations in PI3K/AKT and MAPK pathways in diagnosis of thyroid nodules and in prediction of outcome in patients with differentiated thyroid carcinoma*, PhD thesis, University of Campinas.

Snitz, B.E., Macdonald, A.W., 3rd and Carter, C.S., 2006. Cognitive deficits in unaffected first-degree relatives of schizophrenia patients: a meta-analytic review of putative endophenotypes. *Schizophrenia Bulletin*, 32(1), pp.179-194.

Stefansson, H., et al., 2009. Common variants conferring risk of schizophrenia. *Nature*, 460 (7256), pp. 744-747.

Steinberg, S., et al., 2011. Common variants at VRK2 and TCF4 conferring risk of schizophrenia. *Human Molecular Genetics*, 20 (20), pp. 4076-4081.

Straub, R.E., et al., 2002. Genetic Variation in the 6p22.3 Gene DTNBP1, the Human Ortholog of the Mouse Dysbindin Gene, Is Associated with Schizophrenia. *The American Journal of Human Genetics*, 71 (2), pp. 337-348.

Tam, G.W., et al., 2010. Confirmed rare copy number variants implicate novel genes in schizophrenia. *Biochemical Society Transactions*, 38 (2), pp. 445-451.

Tan, G.K.N., Tee, S.F. and Tang, P.Y., 2015. Genetic association of single nucleotide polymorphisms in dystrobrevin binding protein 1 gene with schizophrenia in a Malaysian population. *Genetics and Molecular Biology*, 38 (2), pp. 138-146.

Tan, H.Y., et al., 2008. Genetic variation in AKT1 is linked to dopamine-associated prefrontal cortical structure and function in humans. *Journal of Clinical Investigation*, 118 (6), pp. 2200-2208.

Tandon, R., Keshavan, M.S. and Nasrallah, H.A., 2008. Schizophrenia, "just the facts" what we know in 2008. 2. Epidemiology and etiology. *Schizophrenia Research*, 102 (1-3), pp. 1-18.

Tang, J.X., et al., 2003. Family-based association study of DTNBP1 in 6p22.3 and schizophrenia. *Molecular Psychiatry*, 8 (8), pp. 717-718.

Thara, R. and Kamath, S., 2015. Women and schizophrenia. *Indian Journal of Psychiatry*, 57 (Suppl 2), pp. S246-S251.

Thiselton, D.L., et al., 2008. AKT1 is associated with schizophrenia across multiple symptom dimensions in the Irish study of high density schizophrenia families. *Biological Psychiatry*, 63 (5), pp. 449-457.

Tishkoff, S.A. and Verrelli, B.C., 2003. Role of evolutionary history on haplotype block structure in the human genome: implications for disease mapping. *Current Opinion in Genetics & Development*, 13 (6), pp. 569-575.

Toulopoulou, T., et al., 2010. Impaired intellect and memory: a missing link between genetic risk and schizophrenia? *Archives of General Psychiatry*, 67 (9), pp. 905-913.

Toulopoulou, T., et al., 2007. Substantial genetic overlap between neurocognition and schizophrenia: genetic modeling in twin samples. *Archives of General Psychiatry*, 64 (12), pp. 1348-1355.

Trost, S., et al., 2013. The DTNBP1 (dysbindin-1) gene variant rs2619522 is associated with variation of hippocampal and prefrontal grey matter volumes in humans. *European Archives of Psychiatry and Clinical Neuroscience*, 263 (1), pp. 53-63.

Tsuang, M.T., Stone, W.S. and Faraone, S.V., 2000. Schizophrenia: vulnerability versus disease. *Dialogues in Clinical Neuroscience*, 2 (3), pp. 257-266.

Turunen, J.A., et al., 2007. The role of DTNBP1, NRG1 and AKT1 in genetics of schizophrenia in Finland. *Schizophrenia Research*, 91 (1-3), pp. 27-36.

Umeda-Yano, S., et al., 2013. The regulation of gene expression involved in TGF-beta signaling by ZNF804A, a risk gene for schizophrenia. *Schizophrenia Research*, 146 (1-3), pp. 273-278.

Untergasser, A., et al., 2012. Primer3—new capabilities and interfaces. *Nucleic Acids Research*, 40 (15), pp. e115-e115.

Varela-Gomez, N., et al., 2015. Dysbindin gene variability is associated with cognitive abnormalities in first-episode non-affective psychosis. *Cognitive Neuropsychiatry*, 20 (2), pp. 144-156.

Versola-Russo, J., 2006. Cultural and Demographic Factors of Schizophrenia. *International Journal of Psychosocial Rehabilitation*, 10 (2), pp. 89-103.

Voisey, J., et al., 2010. Analysis of HapMap tag-SNPs in dysbindin (DTNBP1) reveals evidence of consistent association with schizophrenia. *European Psychiatry*, 25 (6), pp. 314-319.

Walters, J.T., et al., 2010. Psychosis susceptibility gene ZNF804A and cognitive performance in schizophrenia. *Archives of General Psychiatry*, 67 (7), pp. 692-700.

Weickert, C.S., et al., 2004. Human dysbindin (DTNBP1) gene expression in normal brain and in schizophrenic prefrontal cortex and midbrain. *Archives of General Psychiatry*, 61 (6), pp. 544-555.

Wieben, E.D., et al., 2014. Comprehensive assessment of genetic variants within TCF4 in Fuchs' endothelial corneal dystrophy. *Investigative Ophthalmology & Visual Science*, 55 (9), pp. 6101-6107.

Williams, H.J., et al., 2011a. Association between TCF4 and schizophrenia does not exert its effect by common nonsynonymous variation or by influencing cis-acting regulation of mRNA expression in adult human brain. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 156B (7), pp. 781-784.

Williams, H.J., et al., 2011b. Fine mapping of ZNF804A and genome-wide significant evidence for its involvement in schizophrenia and bipolar disorder. *Molecular Psychiatry*, 16 (4), pp. 429-441.

Wirgenes, K.V., et al., 2012. TCF4 sequence variants and mRNA levels are associated with neurodevelopmental characteristics in psychotic disorders. *Translational Psychiatry*, 2, pp. e112.

Wolf, C., Jackson, M.C., Kissling, C., Thome, J. and Linden, D.E.J., 2011. Dysbindin-1 genotype effects on emotional working memory. *Molecular Psychiatry*, 16 (2), pp. 145-155.

Woodgett, J.R., 2005. Recent advances in the protein kinase B signaling pathway. *Current Opinion in Cell Biology*, 17 (2), pp. 150-157.

Yokley, J.L., et al., 2012. Genetic associations between neuregulin-1 SNPs and neurocognitive function in multigenerational, multiplex schizophrenia families. *Psychiatric Genetics*, 22 (2), pp. 70-81.

Zai, C.C., et al., 2008. Genetic study of eight AKT1 gene polymorphisms and their interaction with DRD2 gene polymorphisms in tardive dyskinesia. *Schizophrenia Research*, 106 (2-3), pp. 248-252.

Zammit, S., Allebeck, P., Andreasson, S., Lundberg, I. and Lewis, G., 2002. Self reported cannabis use as a risk factor for schizophrenia in Swedish conscripts of 1969: historical cohort study. *British Medical Journal*, 325 (7374), pp. 1199.

Zhang, R., et al., 2012. Further evidence for the association of genetic variants of ZNF804A with schizophrenia and a meta-analysis for genome-wide significance variant rs1344706. *Schizophrenia Research*, 141 (1), pp. 40-47.

Zhang, W., et al., 2017. BDNF rescues prefrontal dysfunction elicited by pyramidal neuron-specific DTNBP1 deletion in vivo. *Journal of Molecular Cell Biology*, 9 (2), pp. 117-131.

Zhang, Y.E., 2009. Non-Smad pathways in TGF-beta signaling. *Cell Research*, 19 (1), pp. 128-139.

Zhu, X., et al., 2013. Associations between TCF4 gene polymorphism and cognitive functions in schizophrenia patients and healthy controls. *Neuropsychopharmacology*, 38 (4), pp. 683-689.

Zinkstok, J.R., et al., 2007. Association between the DTNBP1 gene and intelligence: a case-control study in young patients with schizophrenia and related disorders and unaffected siblings. *Behavioral and Brain Functions,* 3 (1), pp. 19.

APPENDIX A

SAMPLE OF RESEARCH CONSENT FORM

RESEARCH CONSENT FORM 1a

(Patient)

Schizophrenia Health Study
University Malaya-MOHE-Tampoi Hospital
Blood Draw Consent Form *[Borang Persetujuan Penagmbilan Darah]*

ID Number:

Particular of Volunteer *[Maklumat Subjek]*:

Name *[Nama]:* …… Age *[Umur]*: ……. Ethnic *[Bangsa]:* ……

NRIC *[Nombor IC]*: …… Education level *[Tahap Pelajaran]*: ……..

Telephone *[Telepon]*: …… Email *[Emel]*: ……

Purpose *[Tujuan]*:

This is a genetic study on schizophrenia involving patients (and their relative with schizophrenia), and healthy subjects (and their healthy relatives). The purpose of the study is to identify genetic marker for Schizophrenia (single nucleotide polymorphisms). You are invited as part of the study involving patients with schizophrenia.

[Ini adalah kajian genetik yang melibatkan pesakit skizophrenia (dan anggota keluarga yang juga menghidapi skizophrenia), dengan subjek normal yang sihat (dan anggota keluarganya yang sihat). Tujuan kajian ini adalah untuk mengkaji marker genetik untuk Skizophrenia (Single Nucleotide Polymorphism). Anda dijemput untuk kajian bahagian ini]

Procedure *[Prosedur]*:

A small amount of blood (2ml) will be taken from you.

[Sebanyak 2ml darah akan diambil dari anda]

Blood Drawing Statement *[Kenyataan Pengambilan Darah]*:

[ ] I certify that I have been explained about the purpose of the study.
[Saya sahkan bahawa saya telah diberitahu tentang tujuan kajian ini]

[ ] I certify that I have had the opportunity to ask question, and it was explained to my satisfaction.
[Saya sahkan bahawa saya ada peluang untuk menanya soalan tentang kajian ini dan pertanyaan saya telah dijawab dengan memuaskan]

[ ] I voluntarily participate in this study.
[Saya secara sukarela melibatkan diri dalam kajian ini]

Signature/Name *[Tandatangan/Nama]*: ……

Date *[Tarikh]*: ……

RESEARCH CONSENT FORM 1b

(Healthy subject)

Schizophrenia Health Study
University Malaya-MOHE-Tampoi Hospital
Blood Draw Consent Form *[Borang Persetujuan Penagmbilan Darah]*

ID Number:

Particular of Volunteer *[Maklumat Subjek]*:

Name *[Nama]:* …… Age *[Umur]*: ……. Ethnic *[Bangsa]:* ……

NRIC *[Nombor IC]*: …… Education level *[Tahap Pelajaran]*: ……..

Telephone *[Telepon]*: …… Email *[Emel]*: ……

Purpose *[Tujuan]*:

This is a genetic study on schizophrenia involving patients (and their relative with schizophrenia), and healthy subjects (and their healthy relatives). The purpose of the study is to identify genetic marker for Schizophrenia (single nucleotide polymorphisms). You are invited to this second part of the study involving healthy subjects.

[Ini adalah kajian genetik yang melibatkan pesakit skizophrenia (dan anggota keluarga yang juga menghidapi skizophrenia), dengan subjek normal yang sihat (dan anggota keluarganya yang sihat). Tujuan kajian ini adalah untuk mengkaji marker genetik untuk Skizophrenia (Single Nucleotide Polymorphism). Anda dijemput untuk kajian bahagian ini]

Procedure *[Prosedur]*:

A small amount of blood (2ml) will be taken from you.

[Sebanyak 2ml darah akan diambil dari anda]

Blood Drawing Statement *[Kenyataan Pengambilan Darah]*:

[ ] I certify that I have been explained about the purpose of the study.
[Saya sahkan bahawa saya telah diberitahu tentang tujuan kajian ini]

[ ] I certify that I have had the opportunity to ask question, and it was explained to my satisfaction.
[Saya sahkan bahawa saya ada peluang untuk menanya soalan tentang kajian ini dan pertanyaan saya telah dijawab dengan memuaskan]

[ ] I voluntarily participate in this study.
[Saya secara sukarela melibatkan diri dalam kajian ini]

Signature/Name *[Tandatangan/Nama]*: ……

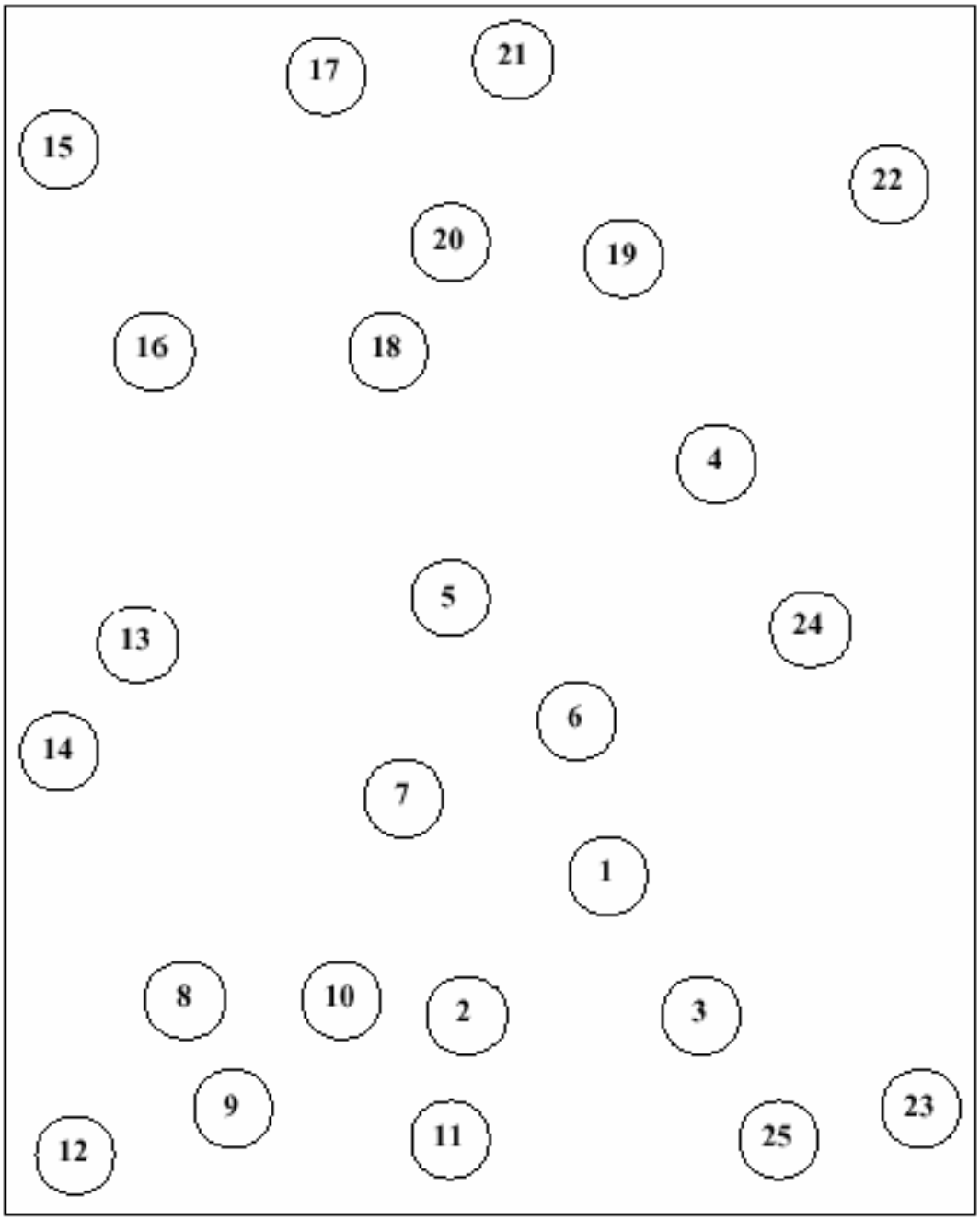
Date *[Tarikh]*: ……

APPENDIX B

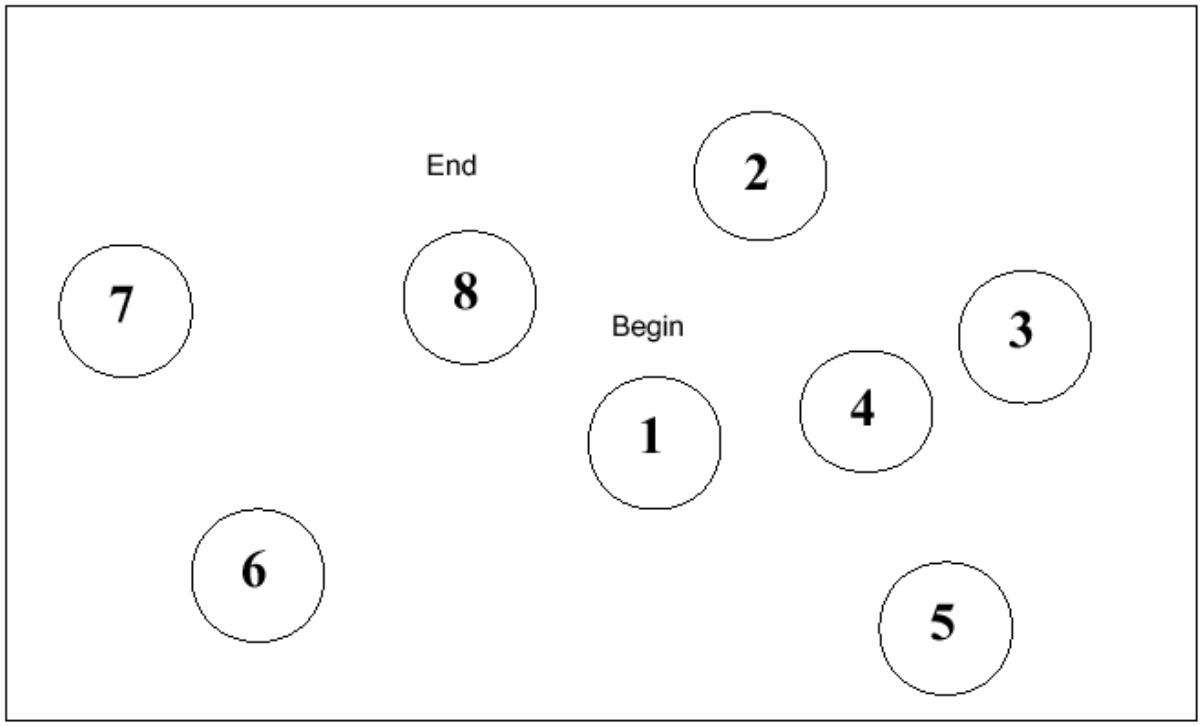
TRAIL MAKING TEST-A

Trail Making Test Part A

Patient's Name: _____ Date: _____



Trail Making Test Part A – *SAMPLE*

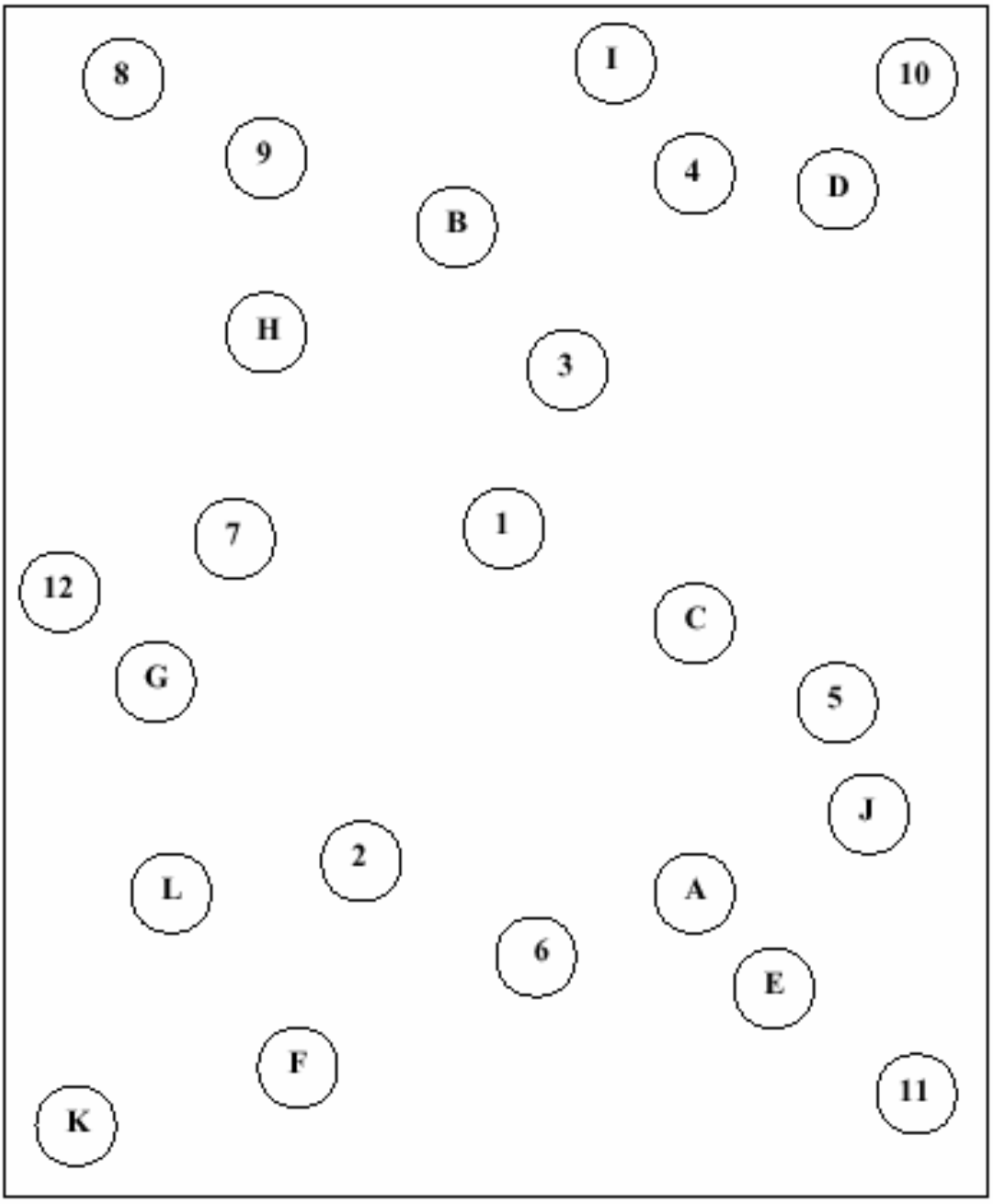


APPENDIX C

TRAIL MAKING TEST-B

Trail Making Test Part B

Patient's Name: _____ Date: _____



Trail Making Test Part B – *SAMPLE*