ASSOCIATION ANALYSES OF IMMUNE AND DNA REPAIR GENE POLYMORPHISMS WITH NON-HODGKIN LYMPHOMA RISK AND PROGNOSIS IN A MULTIETHNIC MALAYSIAN POPULATION

TEN LIK CHIN

DISSERTATION SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

INSTITUTE OF BIOLOGICAL SCIENCES
FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2018

UNIVERSITY OF MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: TEN LIK CHIN

Matric No: SGR130108

Name of Degree: MASTER OF SCIENCE (Except Mathematics & Science Philosophy

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):

ASSOCIATION ANALYSES OF IMMUNE AND DNA REPAIR GENE POLYMORPHISMS WITH NON-HODGKIN LYMPHOMA RISK AND PROGNOSIS IN A MULTIETHNIC MALAYSIAN POPULATION

Field of Study: Genetics and Molecular Biology

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date:

Subscribed and solemnly declared before,

Witness's Signature Date:

Name:

Designation:

ASSOCIATION ANALYSES OF IMMUNE AND DNA REPAIR GENE POLYMORPHISMS WITH NON-HODGKIN LYMPHOMA RISK AND PROGNOSIS IN A MULTIETHNIC MALAYSIAN POPULATION

ABSTRACT

Previous large consortia efforts and genome-wide association studies (GWASs) have implicated a number of genetic variants associated with the susceptibility and prognosis of different non-Hodgkin lymphoma (NHL) types and subtypes. However most of the published studies from the Western population and rarely in Asians. Seventeen candidate single nucleotide polymorphisms (SNPs) from two major groups of NHL-associated genes: immune function SNPs [n=10; rs3832246 (T>-), rs6773854 (T>C), rs6457327 (C>A), rs9271100 (C>T), rs2647012 (C>T), rs10484561 (T>G), rs1571011 (A>C), rs1977389 (T>G), rs10887878 (T>A) and rs4985700 (A>C)] and DNA repair SNPs [n=7; rs17655 (G>C), rs861539 (G>A), rs25487 (C>T), rs25489 (C>T), rs1799782 (G>A), rs13181 (T>G) and rs1799793 (C>T)] were genotyped. A total of 1321 subjects were successfully recruited into this study which consisted of 652 NHL cases and 669 healthy controls from three major ethnic groups in Malaysia, namely the Malay (n=612), Chinese (n=533) and Indian (n=176). Genotyping was performed using TaqMan® SNP genotyping method. Associations of SNPs with NHL risk in each ethnic group were assessed by logistic regression analysis evaluating three genetic models (additive, dominant and recessive model) and adjusted against age and gender. Meta-analysis with the random-effect model was used to estimate the effect size of SNPs in combined three ethnic groups. Two human leukocyte antigen (HLA) class II SNPs: rs2647012 (C>T) and rs10484561 (T>G) showed significant association with NHL risk after Bonferroni correction (P<0.003). SNP rs2647012 (C>T) was significantly associated with reduced risk in all NHL ($P_{\text{additive-adjusted}}$=0.0008; OR=0.54; 95% CI=0.37-0.77) in Malay group and

B-cell NHL in both pooled subjects ($P_{additive\text{-}adjusted}$=0.0008; OR=0.69; 95% CI=0.55-0.87) and Malay group ($P_{additive\text{-}adjusted}$=0.0007; OR=0.51; 95% CI=0.35-0.76). SNP rs10484561 (T>G) showed significant association with subtype FL in pooled subjects ($P_{dominant\text{-}adjusted}$=0.003; OR=2.39; 95% CI=1.35-4.22) and Malay group ($P_{dominant\text{-}adjusted}$=0.003; OR=3.50; 95% CI=1.54-7.97). Through *cis*-expression quantitative trait loci (eQTL) analysis, rs2647012 (C>T) and rs10484561 (T>G) showed potential regulatory function as it showed significant *cis*-eQTL correlation with nearby HLA class II molecule expression. Using the Kaplan-Meier method and Cox regression analysis, 130 DLBCL patients with complete clinical information were further evaluated for the association of SNPs with DLBCL prognosis. No significant association was observed between DLBCL prognosis with any of the studied SNPs. In conclusion, this study reported the association of rs2647012 (C>T) and rs10484561 (T>G) with NHL risk in the Malaysian population.

Keywords: DNA repair genes, immune function genes, non-Hodgkin lymphoma, single nucleotide polymorphisms.

ANALISIS HUBUNGAN ANTARA POLIMORFISME DALAM GEN IMUN DAN PEMBAIKPULIH DNA DENGAN RISIKO DAN PRONOSIS LIMFOMA NON-HODGKIN DI KALANGAN ETNIK MALAYSIA

ABSTRAK

Usaha konsortium dan kajian dari kumpulan genome-wide (GWASs) telah mengaitkan beberapa variasi genetik dengan risiko dan pronosis untuk mendapat jenis limfoma non-Hodgkin (NHL) yang berbeza. Walau bagaimanapun, kebanyakan laporan kajian datang dari Barat dan jarang di kalangan populasi Asia. Untuk melengkapkan usaha ini, 17 calon polimorfisme nukleotida tunggal (SNP) dari dua kumpulan utama gen: SNP dalam gen fungsi imun [n=10; rs3832246 (T>-), rs6773854 (T>C), rs6457327 (C>A), rs9271100 (C>T), rs2647012 (C>T), rs10484561 (T>G), rs1571011 (A>C), rs1977389 (T>G), rs10887878 (T>A) and rs4985700 A>C)] dan SNPs pembaikpulih DNA [n=7; rs17655 (G>C), rs861539 (G>A), rs25487 (C>T), rs25489 (C>T), rs1799782 (G>A), rs13181 (T>G) and rs1799793 C>T)] telah digenotaip. Sebanyak 1321 subjek berjaya direkrut ke dalam kajian ini yang terdiri daripada 652 kes NHL dan 669 kawalan sihat dari tiga kumpulan etnik utama di Malaysia iaitu Melayu (n = 612), Cina (n = 533) dan India (n = 176). Genotaiping dilakukan dengan menggunakan kaedah genotaiping SNP TaqMan®. Hubungan SNP dengan risiko NHL di antara setiap kumpulan etnik dianalisis menggunakan regresi logistik melalui tiga model genetik (model additif, dominan dan resesif) dan diselaraskan dengan umur dan jantina subjek. Meta-analisis dengan model kesan rawak telah diguna untuk menganggarkan hubungan SNP dengan risiko NHL dalam gabungan tiga kumpulan etnik. Dua SNP dari antigen leukocyte manusia (HLA) kelas II, rs2647012 (C>T) dan rs10484561 (T>G) menunjukkan hubungan yang signifikan dengan risiko NHL selepeas pembetulan Bonferroni (P<0.003). SNP rs2647012 (C>T) menunjukkan hubungan yang signifikan dengan semua jenis NHL di

kalangan kohort melayu ($P_{\text{additive-adjusted}}$=0.0008; OR=0.54; 95% CI=0.37-0.77) and NHL B-cell di kalangan semua subjek ($P_{\text{additive-adjusted}}$=0.0008; OR=0.69; 95% CI=0.55-0.87) and kohort Melayu ($P_{\text{additive-adjusted}}$=0.0007; OR=0.51; 95% CI=0.35-0.76). SNP rs10484561 (T>G) menunjukkan hubungan yang signifikan dengan subjenis FL di kalangan semua subjek ($P_{\text{dominant-adjusted}}$=0.003; OR=2.39; 95% CI=1.35-4.22) dan kohort Malayu ($P_{\text{dominant-adjusted}}$=0.003; OR=3.50; 95% CI=1.54-7.97). Melalui *cis*-analisis ungkapan lokus sifat kuantitatif (eQTL), rs2647012 (C>T) dan rs10484561 (T>G) berpotensi menjalani fungsi kawalan kerana ia menunjukan *cis*-eQTL yang signifikan berkorelasi dengan kemunculan beberapa molekul yang berada berdekatan di HLA kelas II. Dengan menggunakan kaedah Kaplan-Meier dan analisis regresi Cox, 130 pesakit DLBCL dengan maklumat klinikal yang lengkap dianalisis untuk hubungan SNP dengan pronosis DLBCL. Tiada hubungan yang signifikan diperhatikan antara pronosis DLCBL dengan mana-mana SNP yang dikaji. Secara kesimpulannya, hubungan antara SNP rs2647012 (C>T) dan rs10484561 (T>G) dengan risiko NHL berjaya dikaji di antara populasi Malaysia.

Kata kunci: Gen fungsi Imun, gen pembaikpulih DNA, limfoma non-Hodgkin, polimorfisme nukleotida tunggal.

ACKNOWLEDGEMENTS

First and foremost, my deepest gratitude goes to my supervisor, Associate Professor Dr. Ng Ching Ching for giving me the opportunities to carry out this research. Her advice not only helped me to resolve challenges throughout the course of my study but also sharpened my logic and critical skills. Her continued patience and valuable guidance have helped me to have a continuous progress during my candidature period.

I would like to thank University Malaya Medical Centre (UMMC) and Ampang Hospitals for the approval of samples collected for this study. My sincere thanks to the clinicians, nurses and other clinic staffs who had helped in the procurement of the samples. Not to forget all the volunteers enrolled in this study.

I thank Skim Biasiswazah Universiti Malaya for supporting my financial during my candidature period. Special thanks to High Impact Research Grant UM-MOHE UM.C/625/1/HIR/MOHE/ CHAN-02 (H-50001-A000023) from the Ministry of Higher Education, Malaysia, University of Malaya Research Grant (UMRG; RG441/12HTM) and University of Malaya Postgraduate Research Grant (PPP; PG118-2015B) for supporting this research.

I greatly appreciate the technical assistances and advice from all my labmates, especially Mr. Chin Yoon Ming, Miss Amy Khor, Miss Goh Siang Ling and Mr. Lim Chun Shen. Their help, motivation, and support throughout this journey had made me gone through all the challenges. Last but not least, I would like to thank my family members and all my friends for their unconditional love, care and support. Thank you!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

LIST OF SYMBOLS

-	:	Deletion
%	:	Percentage
<	:	Less than
>	:	More than
⩽	:	Less than and equal to
⩾	:	More than and equal to
bp	:	Base pair
D’	:	D-prime
kbp	:	Kilobase pair
min	:	Minute
ng/µl	:	Nanogram/microliter
ng/µl	:	Nanogram/microliter
ºC	:	Celcius
p	:	Chromosome short arm region
P	:	*P*-value
q	:	Chromosome long arm region
r^2	:	Pearson’s correlation
Rn	:	Normalized reporter signal
sec	:	second
α	:	Alpha
µ	:	Micro
µl	:	Microlitre
χ^2	:	Chi Square

LIST OF ABBREVIATIONS

A	:	Adenine
ABC	:	Activated peripheral-blood B-cell
ABI	:	Applied Biosystems
AIDS	:	Acquired immune deficiency syndrome
ALPS	:	Autoimmune lymphoproliferative syndrome
ASR	:	Age-standardized rate
BCL6	:	B-cell lymphoma 6
BER	:	Base excision repair
C	:	Cytosine
CD-CV	:	Common disease -common variant
CHOP	:	Cyclophosphamide, Doxorubicin, Vincristine, and Prednisone
Chr	:	Chromosome
CI	:	Confidence intervals
D	:	Aspartic acid or Asp
dbSNP	:	The Single Nucleotide Polymorphism database
dH_2O	:	Distilled water
DLBCL	:	Diffuse large B-cell lymphoma
DNA	:	Deoxyribonucleic acid
DSBR	:	Double-strand break repair
e.g.	:	For example
EBV	:	Epstein–Barr virus
ECOG	:	Eastern Cooperative Oncology Group
eQTL	:	Expression quantitative trait loci
ERCC2	:	Excision Repair Cross-complementing group 2
ERCC5	:	Excision Repair Cross-complementing group 5

et al.	:	And other people
etc.	:	And other things
FAM	:	Fluorescein amidite
FAS	:	Fas cell surface death receptor
FasL	:	Fas ligand
FDR	:	False discovery rate
FL	:	Follicular lymphoma
G	:	Guanine
GC	:	Germinal center
GCB	:	Germinal centre B-cell
GLOBOCAN	:	Global Cancer Incidence, Mortality and Prevalence
GTEx	:	Genotype-Tissue expression
GWAS	:	Genome-wide association study
H	:	Histidine or His
HapMap	:	Haplotype map
HGP	:	Human Genome Project
HCV	:	Hepatitis C virus
HIV	:	Human immunodeficiency virus
HLA	:	Human leukocyte antigen
HTLV-I	:	Human T-cell leukemia virus type I
HWE	:	Hardy-Weinberg equilibrium
Ig	:	Immunoglobulin
IPI	:	International Prognostic Index
K	:	Lysine or Lys
K-P	:	Kaplan-Meier
LD	:	Linkage disequilibrium

LDH	:	Lactate dehydrogenase
LPP	:	Lipoma-preferred partner
M	:	Methionine or Met
MAF	:	Minor allele frequency
MB	:	Megabases
MMR	:	Mismatch repair
MOH	:	Ministry of Health
N	:	Asparagine or Asn
NER	:	nucleotide excision repair
NHL	:	Non-Hodgkin lymphoma
NK	:	Natural killer
NMRR	:	National Medical Research Register
OR	:	Odd ratio
OS	:	Overall survival
PCR	:	Polymerase chain reaction
PMBCL	:	Primary mediastinal large B-cell lymphoma
Polyphen-2	:	Polymorphism phenotyping v2
Q	:	Glutamine or Gln
qPCR	:	Quantitative PCR
R	:	Arginine or Arg
RA	:	Rheumatoid arthritis
R-CHOP	:	CHOP with rituximab
RNA	:	Ribonucleic acid
ROX	:	6-Carboxyl-X-Rho- damine
SD	:	Standard deviation
SIFT	:	Sorting tolerant from intolerant.

SLE	:	Systemic lupus erythematosus
SNP	:	Single nucleotide polymorphism
SPSS	:	Statistical Package for the Social Sciences
STAT	:	Signal transducer and activator of transcription
T	:	Thiamine
T	:	Threonine or Thr
TACI	:	Transmembrane activator and CAML interactor
TNF	:	Tumour Necrosis Factor
TNFRSF13B	:	Tumour necrosis factor receptor superfamily member 13B
TSS	:	Transcription start site
UMMC	:	University Malaya Medical Centre
US	:	United States of America
UV	:	Ultraviolet
W	:	Tryptophan or Trp
WHO	:	World Health Organization
XRCC1	:	X-Ray Repair Cross Complementing 1
XRCC3	:	X-Ray Repair Cross Complementing 3

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

Non-Hodgkin's lymphomas (NHLs) are a heterogeneous group of neoplasms derived from B or T lymphocytes, which ranked among the top ten most common cancer worldwide (Ferlay *et al*., 2015). There are marked differences in the incidence patterns of NHL across a geographical area or within the population itself, where the western countries in North America and Western Europe having the higher rates (Ferlay *et al*., 2015). The overall cancer incidence rate of NHL in Asia is much lower compared to that in Europe and North America (Ferlay *et al*., 2015). In Malaysia, lymphoma ranked the sixth most common cancer, with an overall cancer incidence rate of 4.3 per 100,000 population (Zainal Ariffin & Nor Saleha, 2011). The incidence of lymphoma was slightly higher among males (age-standardized rate per 100,000 people, ASR=4.2) compared to females (ASR=3.0). Chinese were found to have a higher incidence rate compared to Malay and Indian (Zainal Ariffin & Nor Saleha, 2011).

The precise pathophysiology of NHL remains unexplained, despite its dramatic worldwide rise in incidence in recent decades. Autoimmune disorders (Landgren *et al*., 2006), immunodeficiency state of patients (Filipovich *et al*., 1992), infectious agents (Engels, 2007) such as human immunodeficiency virus (HIV) (Beral *et al*., 1991), Epstein-Barr virus (EBV; Neri *et al*., 1991), *Helicobacter pylori* (Parsonnet *et al*., 1994) and human T-cell lymphotropic/leukemia virus-1 (HTLV-1) and exposure to chemical or pharmaceutical agents such as benzene (Jiao *et al*., 2012), herbicide (Fontana *et al*., 1998), pesticides (Boccolini *et al*., 2013) are few of the reported risk factors for NHL. However, it is largely thought to be caused by a combination of the interaction between multiple genes and environmental factors.

Single nucleotide polymorphisms (SNPs) are the most common polymorphisms in the human genome and are one of the important genetic markers in genetic studies. Functional SNPs can alter the transcription and protein expression level or functions of a

gene product and therefore alter individual's susceptibility to disease including cancer. Identification of susceptibility genes to NHL is crucial for improving the understanding of the biological and etiologic mechanism involved in the development of NHL. The association between NHL risk and polymorphism in *BCL6* (Tan *et al*., 2013), HLA region (Conde *et al*., 2010; Skibola *et al*., 2009; Smedby *et al*., 2011), *FAS* (Wang *et al*., 2009), *TNFRSF13B* (Wang *et al*., 2009), *XRCC1* (Kim *et al*., 2010; Smedby *et al*., 2006), *XRCC3* (Smedby *et al*., 2006), *ERCC5, ERCC2* (Worrillow *et al*., 2009) genes had been reported, suggesting the involvement of immune function and DNA repair genes in this disease. Nevertheless, the genetic association study of NHL risk in Malaysia population is still lacking compared to other populations and only two genetic association studies had been published so far (Lim *et al*., 2014; Sujatha *et al*., 2014).

Most SNPs studied to date have yielded inconsistent association results with NHL risk between Caucasians and Asians. In addition, most studies focus on Caucasians and only a few have studied the effects of these SNPs in Asians and another ethnic origin. Importantly, there is a pronounced difference in allelic and genotype frequency of Caucasians and Asians. Furthermore, genetic and environmental heterogeneity may cause the effect of genotype on disease phenotype to vary between populations. The reasons above indicate the necessity for further investigation of the association between SNPs in candidate genes with the NHL susceptibility in different population and ethnicity.

In the present study, a case-control association study was performed to examine the NHL susceptibility SNPs in Malaysia population. By using the candidate gene approach, 17 SNPs were investigated for their associations and NHL susceptibility. The SNPs were further tested to examine the association with diffuse large B-cell lymphoma (DLBCL) patients' survival risk. To the best of our knowledge, these candidate SNPs have not been studied in the Malaysian population. Overall, this thesis is focused on the genetic susceptibility to the risk and prognosis of NHL among three ethnicities in

Malaysia.

1.1 Objectives

I. To determine the allele and genotype frequencies of selected SNPs in case (NHL patients) and control (healthy subject) groups.

II. To study the association of SNPs in candidate genes and NHL susceptibility in the Malaysian population.

III. To examine if the SNPs are associated with the prognosis of DLBCL which is the most common subtype of NHL.

CHAPTER 2: LITERATURE REVIEW

2.1 Non-Hodgkin lymphoma (NHL)

Lymphomas are hematological malignancies of a family of white blood cells known as lymphocytes which start to develop in the lymph system. Lymphoma occurs when lymphocytes grow and multiply uncontrollably. Because lymph tissue is scattered throughout the body, lymphoma can begin almost anywhere and spread throughout the body, including the lymph nodes, tonsils, thymus, spleen and other organs. The classification of human lymphoma has steadily evolved since their initial recognition by Thomas Hodgkin in 1832 (Hodgkin, 1832). There are two types of lymphoma: Hodgkin lymphoma and non-Hodgkin lymphoma (NHL). Hodgkin lymphoma is marked by the presence of Reed-Sternberg cells, which are mature B cells that enlarge and become malignant, and often seen to carry more than one nucleus. NHL, by contrast, can be derived from B or T lymphocytes.

NHL can arise in the lymph nodes as well as other organs. NHL can be classified into two major types: B-cell NHL or T-cell NHL, depends on the type of lymphocytes affected. NHL is a heterogeneous disease with more than 60 subtypes classified by the World Health Organization (WHO) (Swerdlow *et al*. 2008), where the new classification of lymphoid neoplasms was revised in 2016 (Table 2.1) (Swerdlow *et al*., 2016). NHLs are differentiated based on histopathological, immunological and molecular information of disease. Each subtype varies in presentation, survival expectation, morbidity and responses to treatment. The diagnosis is based on how the lymphoma cells look under the microscope and confirmed by additional information from other tests, including tests of genetic material within the lymphoma cells.

2.1.1 B-cell NHL

About 90% of NHL cases are B-cell NHL. Diffuse large B-cell lymphoma (DLBCL) and follicular lymphoma (FL) being the two most common B-cell NHL subtypes among Malaysian population as well as in other world population (Swerdlow *et al*., 2008). DLBCL has an annual incidence rate of over 25,000 cases, accounting for roughly 40% NHL cases (Ferlay *et al*., 2015). It is an aggressive, fast-growing lymphoma that usually affects adults but can also occur in children. DLBCL is clinically heterogeneous. Gene expression profiling using Deoxyribonucleic acid (DNA) microarrays further classified DLBCL into 3 distinct molecular subtypes: germinal center B-cell (GCB) DLBCL, activated peripheral-blood B-cell (ABC) DLBCL and primary mediastinal large B-cell lymphoma (PMBCL) (Alizadeh *et al*., 2000). FL is an indolent lymphoma, which is the second most common subtype of NHL. Other common types of B-cell NHL include mantle cell lymphoma, small lymphocytic lymphoma, marginal zone lymphoma, lymphoplasmacytic lymphoma, primary central nervous system lymphoma, Burkitt's lymphoma, lymphoblastic lymphoma and etc. (Table 2.1)

2.1.2 T-cell NHL

While there are many subtypes of T-cell NHL, they are all very rare, as this group is making up about 10% of NHL cases. A similar lymphocyte called a natural killer (NK) cell shares many features with T-cells. When NK cells become cancerous, the cancer is called NK or NK/T-cell NHL and is generally grouped together with other T-cell NHL. T-cell NHL can be aggressive (fast-growing) or indolent (slow-growing). Most published studies have demonstrated that T-cell lymphoid malignancies generally are more aggressive and confer a poorer prognosis than the corresponding B-cell lymphomas (Coiffier *et al*., 1990; Gisselbrecht *et al*., 1998).

Table 2.1: 2016 World Health Organization (WHO) classification of the mature B-cell, T-cell and NK-cell neoplasms of lymphocytes (Swerdlow *et al.*, 2016).

Mature B-cell neoplasms	Mature T-cell and NK-cell neoplasms
Chronic lymphocytic leukemia/small lymphocytic lymphoma	T-cell prolymphocytic leukemia
Monoclonal B-cell lymphocytosis*	T-cell large granular lymphocytic leukemia
B-cell prolymphocytic leukemia	Chronic lymphoproliferative disorder of NK cells
Splenic marginal zone lymphoma	Aggressive NK-cell leukemia
Hairy cell leukemia	Systemic EBV T-cell lymphoma of childhood*
Splenic B-cell lymphoma/leukemia, unclassifiable	Hydroa vacciniforme–like lymphoproliferative disorder*
Splenic diffuse red pulp small B-cell lymphoma	Adult T-cell leukemia/lymphoma
Hairy cell leukemia-variant	Extranodal NK-/T-cell lymphoma, nasal type
Lymphoplasmacytic lymphoma	Enteropathy-associated T-cell lymphoma
Waldenstrom¨ macroglobulinemia	Monomorphic epitheliotropic intestinal T-cell lymphoma*
Monoclonal gammopathy of undetermined significance (MGUS), IgM*	Indolent T-cell lymphoproliferative disorder of the GI tract*
m heavy-chain disease	Hepatosplenic T-cell lymphoma
g heavy-chain disease	Subcutaneous panniculitis-like T-cell lymphoma
a heavy-chain disease	Mycosis fungoides
Monoclonal gammopathy of undetermined significance (MGUS), IgG/A*	Sezary´ syndrome
Plasma cell myeloma	Primary cutaneous CD30 T-cell lymphoproliferative disorders
Solitary plasmacytoma of bone	Lymphomatoid papulosis
Extraosseous plasmacytoma	Primary cutaneous anaplastic large cell lymphoma
Monoclonal immunoglobulin deposition diseases*	Primary cutaneous gd T-cell lymphoma
Extranodal marginal zone lymphoma of mucosa-associated lymphoid tissue (MALT lymphoma)	Primary cutaneous CD8 aggressive epidermotropic cytotoxic T-cell lymphoma
Nodal marginal zone lymphoma	Primary cutaneous acral CD8 T-cell lymphoma*
Pediatric nodal marginal zone lymphoma	Primary cutaneous CD4[1] small/medium T-cell lymphoproliferative disorder*
Follicular lymphoma	Peripheral T-cell lymphoma, NOS
In situ follicular neoplasia*	Angioimmunoblastic T-cell lymphoma
Duodenal-type follicular lymphoma*	Follicular T-cell lymphoma*
Pediatric-type follicular lymphoma*	Nodal peripheral T-cell lymphoma with TFH phenotype*
Large B-cell lymphoma with IRF4 rearrangement*	Anaplastic large-cell lymphoma, ALK
Primary cutaneous follicle center lymphoma	Anaplastic large-cell lymphoma, ALK*
Mantle cell lymphoma	Breast implant–associated anaplastic large-cell lymphoma*
In situ mantle cell neoplasia*	
Diffuse large B-cell lymphoma (DLBCL), NOS	
Germinal center B-cell type*	
Activated B-cell type*	
T-cell/histiocyte-rich large B-cell lymphoma	
Primary DLBCL of the central nervous system (CNS)	
Primary cutaneous DLBCL, leg type	
EBV DLBCL, NOS*	
EBV mucocutaneous ulcer*	
DLBCL associated with chronic inflammation	
Lymphomatoid granulomatosis	
Primary mediastinal (thymic) large B-cell lymphoma	
Intravascular large B-cell lymphoma	
ALK large B-cell lymphoma	
Plasmablastic lymphoma	
Primary effusion lymphoma	
HHV8 DLBCL, NOS*	
Burkitt lymphoma	
Burkitt-like lymphoma with 11q aberration* High-grade B-cell lymphoma, with MYC and BCL2 and/or BCL6 rearrangements*	
High-grade B-cell lymphoma, NOS*	
B-cell lymphoma, unclassifiable, with features intermediate between DLBCL and classical Hodgkin lymphoma	

*Changes from the 2008 classification.

2.1.3 Epidemiology of NHL

Differences in the incidence rate of NHL between population group worldwide have been observed. The disease appears to be more common in higher socioeconomic groups, in urban areas, and in whites; the reasons for this pattern remain unknown.

Based on the statistics reported by Global Cancer Incidence, Mortality and Prevalence (GLOBOCAN) 2012, NHL ranks among the ten most commonly diagnosed cancer worldwide and accounts for 2.7% of all new cancers cases being diagnosed (Ferlay *et al*., 2015). The incidence rates of NHL increased 5-7% annually between 1991 and 2010; with an estimated 385,700 new cases and 199,700 deaths from NHL occurred in 2012 worldwide (Ferlay *et al*., 2015). NHL is more common in developed regions, with the highest incidence rates found in Northern America (ASR 14.6 and 10.2 per 100,000 in men and women respectively), in Australia/ New Zealand (14.3 and 10.1) and in Europe (except Central and Eastern Europe). In 2015 in the United States (U.S.), an estimated of 71,850 new cases of NHL were diagnosed, and approximately 19,790 patients died of the disease. This accounts for 4% of new cancer cases and 3% of cancer-related deaths (Siegel *et al*., 2015). The lowest rates were found in Asia and Eastern Europe (Ferlay *et al*., 2015). In general, the incidence of NHL was low in Africa with the exception of some sub-Sharan areas (particularly in East Africa) because of high incidence among children with a subtype of NHL called Burkitt lymphoma (Ferlay *et al*., 2015). The median age of patients with NHL was 60, but it may occur in all age groups (Ferlay *et al*., 2015).

In Malaysia, lymphoma ranked the sixth most common cancer among Malaysian population, with an overall cancer incidence rate of 4.3 per 100,000 population (Zainal Ariffin & Nor Saleha, 2011). The incidence of lymphoma was slightly higher among males (age-standardized rate per 100,000 people, ASR=4.2) compared to females

(ASR=3.0). Chinese were found to have a higher incidence rate compared to Malay and Indian (Zainal Ariffin & Nor Saleha, 2011).

2.1.4 Etiology of NHL

The primary known risk factor for NHL is often associated with the altered immune function. The risk of getting NHL has shown higher in autoimmune disease patients such as Sjögren syndrome, systemic lupus erythematosus (SLE) and rheumatoid arthritis (RA) (Landgren *et al*., 2006).

Organ transplant recipients who receive immunosuppressive drugs to prevent graft rejection has been associated with higher rates of NHL than in the general population (Opelz & Henderson, 1993). In addition, chronic infection with some viruses such as Epstein–Barr virus (EBV) that causes Burkitt lymphoma, human immunodeficiency virus (HIV), human T-cell leukemia virus type I (HTLV-I), human herpes virus 8, and hepatitis C virus (HCV) tend to associate with higher risk of getting NHL. Besides, types of bacteria such as *Helicobacter pylori* that cause the immune system to be continuously active are associated with gastric lymphoma (Parsonnet *et al*., 1994). NHL is classified as an acquired immune deficiency syndrome (AIDS)-defining illness among HIV-positive people, and the risk is 60 times greater among AIDS patients compared with the general population in the US (Beral *et al*., 1991).

Epidemiologic studies suggested that environmental factors may play an important role in the etiology of NHL. Agricultural work, and, in particular, frequent exposure to specific pesticides such as 2,4-dichlorophenoxyacetic acid (Boccolini *et al*., 2013) and herbicides (Fontana *et al.*, 1998) have been consistently associated with an increased risk of NHL. Exposure to certain chemicals such as benzene and chlorinated solvent have also been linked with an increased chance of getting the disease (Jiao *et al*., 2012). Working in the rubber manufacturing industry and occupational and

environmental exposure to certain chemicals (e.g., solvents such as dichloromethane) may also increase the risk for some NHL subtypes (Blair *et al*., 1998).

Like most cancers, the risk of developing NHL increases with age. In addition, blood transfusion, radiation exposure, lifestyle factors such as smoking (Willett *et al*., 2004), hair dye used (Morton *et al*., 2007), alcohol consumption (Lim *et al*., 2007), frequent sunlight exposure and dietary intake are also proved to be associated with NHL risk (Blair *et al*., 1998).

A family history of lymphoma confers an increased risk of NHL in many genetic studies (Conde *et al*., 2010; Hill *et al*., 2006; Skibola *et al*., 2009 Shen *et al*., 2007; Smedby *et al*., 2006). It is becoming increasingly clear that genetic factors, including SNPs, may play a substantial role in NHL development. However, specific loci associated with NHL have not yet been confirmed. In early years, DNA repair genes had extensively reported to be associated with NHL risk (Hill *et al*., 2006; Shen *et al*., 2007; Smedby *et al*., 2006) until recently where genome-wide association studies (GWASs) have consistently linked the HLA polymorphisms with NHL risk (Conde *et al*., 2010; Skibola *et al*., 2009; Smedby *et al*., 2011).

2.2 Human genetic polymorphisms

With the completion of international projects such as The Human Genome Project (HGP) (Sawicki *et al*., 1993), International HapMap project (International HapMap Consortium, 2003), and the most recent 1000 Genomes project (The 1000 Genomes Project Consortium, 2015) more and more data about genetic polymorphisms in various population available online as a reference, further facilitating the design of case-control association studies.

2.2.1 Single nucleotide polymorphisms (SNPs)

Molecular study of DNA variants is one of the most exciting endeavors of cancer research today. There are more than 10,000,000 SNPs found in the entire human genome (Smigielski *et al*. 2000). Even in one gene, there are many SNPs existed. Undoubtedly, genotyping all the SNPs would give us the most information, but this usually not perform as it is often a costly and time-consuming process. Researchers usually choose SNPs that tag the common variation in the region. This is an excellent choice to reduce the number of SNPs without reducing the useful information. The common disease-common variant (CD-CV) hypothesis stated that for any given common disease, the genetic risk will be due to common variants with high frequency in the population (Pritchard & Cox, 2002; Reich & Lander, 2001). Thus, choosing SNPs that represent the common haplotypes should be able to detect association with these common variants that affect the trait.

There are variations between human populations, so a SNP allele that is common in one geographical or ethnic group may be much rarer in another. The discovery of an increasing number of SNPs in perhaps all genes of an organism highlights the diversity between individuals and the potential differences in molecular responses of humans to DNA lesions. The knowledge of these differences allows the discoveries of genetic interactions and control networks in the cell, predict the response of patients to treatments with DNA damaging agents and explain individual differences in susceptibility to environmental mutagens.

2.3 Selection of NHL associated SNPs

Using the candidate gene approach, the association of 17 SNPs (Table 2.2 & Table 2.3) with NHL risk in Malaysian population was examined. These SNPs were selected from two major categories, namely immune function SNPs and DNA repair SNPs. All SNPs were selected based on previously reported findings or functional prediction from

the web database. Most SNPs study to date has yielded inconsistent association results with NHL risk between Caucasians and Asians population. In addition, most studies focus on Caucasians and only a few papers reported the association of these SNPs in Asians. Hence, this study would like to investigate the association between SNPs with the NHL risk in the Malaysian population.

2.3.1 SNPs in immune genes and NHL risk

A total of ten immune function SNPs from *BCL6*, HLA region, *FAS* and *TNFRSF13B* (Table 2.2) were selected to be examined in this study.

B-cell lymphoma 6 (*BCL6*) gene encodes for a human BCL6 protein. The protein encoded by this gene is an evolutionarily conserved zinc finger transcription factor and contains an N-terminal POZ/BTB domain. This protein acts as a sequence-specific repressor of transcription and has been shown to modulate the signal transducer and activator of transcription (STAT)-dependent interleukin-4 (*IL-4*) responses of B cells (Harris *et al*., 1999). The BCL6 protein is a transcriptional repressor that is an important regulator of lymphoid development and function. The protein is preferentially expressed in germinal center (GC) B-cells of normal lymphoid tissues as well as in a variety of B-NHL subtypes derived from germinal center B-cells irrespective of whether the BCL6 is rearranged. This gene is found to be frequently translocated and hypermutated in DLBCL and contributes to the pathogenesis of DLBCL (Cattoretti *et al*., 2005) . One published GWAS had discovered the polymorphisms in chromosome 3q27 near *BCL6* might contribute to the lymphomagenesis in Asians (Tan *et al*., 2013).

Using the advanced GWASs technology, more papers reported SNPs in the immune genes-rich region, human leukocyte antigen (HLA) region associated with NHL susceptibility (Conde *et al*., 2010; Skibola *et al*., 2009; Smedby *et al*., 2011). The HLA region maps to chromosome 6p21.3 in human. The primary function of HLA molecules

is to provide protection against pathogens. HLA class I molecules present endogenous antigens, and class II molecules present exogenous antigens to T cells, creating the "tri-molecular complex" (HLA-peptide-TCR) that initiates the immune response. The region contains more than 200 identified genes, over half of which are predicted to be expressed. Only some of the HLA region genes are involved in the immune response; the genes that encode the classical class I (A, B, and C) and class II (DR, DQ, and DP) antigens. The HLA molecule has been demonstrated to be central to physiology, protective immunity and deleterious immune reactivity. Mutations in HLA may be linked to many autoimmune disease, such as type I diabetes, ankylosing spondylitis, celiac disease, SLE, myasthenia gravis, inclusion body myositis, Sjögren syndrome, and narcolepsy (Gough & Simmonds, 2007).

FAS cell surface death receptor (*FAS*) (also known as APO-1 or CD95) is a transmembrane protein of the tumour necrosis factor (TNF) receptor family, which mediates programmed cell death or apoptosis upon trimerization induced by cross-linking to Fas ligand (FasL). *FAS* is expressed on the surface of activated T and B lymphocytes, and FAS/FasL-induced apoptosis is important for eliminating autoreactive immature T cells during ontogenesis and for maintaining peripheral lymphocyte homeostasis. Disruption of the FAS/FasL apoptotic pathway has been associated with benign lymphoproliferation, severe multisystem autoimmune disease, and hypergammaglobulinemia (Watanabe-Fukunaga *et al.*, 1992). Children who carry inherited defects in the *FAS* gene exhibit a similar pattern of phenotypes that have been collectively termed autoimmune lymphoproliferative syndrome (ALPS) (Bettinardi *et al.*, 1997). Tumour necrosis factor receptor superfamily member 13B (*TNFRSF13B*), is a transmembrane protein of the TNF receptor superfamily found predominantly on the surface of B cells, which are an important part of the immune system. It interacts with calcium-modulator and cyclophilin ligand (CAML). The protein induces activation of the

transcription factors NFAT, AP1, and NF-kappa-B and plays a crucial role in humoral immunity by interacting with a TNF ligand. Mutations in *TNFRSF13B* contribute to common variable immunodeficiency and autoimmunity in humans (Salzer *et al*., 2005)

Table 2.2: Summary of immune function SNPs examined in this study.

Gene/ region	Chromosome	SNP rs no	Position (hg 19)
BCL6	3q27	rs3832246 (T>-)	187000000
		rs6773854 (T>C)	188000000
HLA	6p21.3	rs6457327 (C>A)	31074030
		rs9271100 (C>T)	32576478
		rs2647012 (C>T)	32664458
		rs10484561 (T>G)	32665420
FAS	10q23.31	rs1571011 (A>C)	90757787
		rs1977389 (T>G)	90773494
		rs10887878 (T>A)	90779227
TNFRSF13B	17p11.2	rs4985700 (A>C)	16866075

Abbreviations: SNP: single nucleotide polymorphism.

2.3.2 SNPs in DNA repair genes and NHL risk

Studies have shown that cancer including NHL tends to occur more frequently in older people. Few experiments involved mice have confirmed the suspicion that mutation rate increases with age (Milholland *et al*., 2015). These mutations have ultimately produced a cell capable of uncontrolled growth so that older people are more vulnerable to cancer once they are exposed to carcinogens (pollution, tobacco, sun, radiation, alcohol, bad diet, and maybe even stress) longer. Their DNA repair system is not as effective as people get older. This could be because genes required for preventing or repairing DNA damage are mutated, leading to runaway DNA instability.

One of the major culprits of cancer, including NHL, is DNA damage. DNA suffers from a wide range of damage, a large proportion of DNA alterations are caused unavoidably by endogenous weak mutagens, including water, oxygen free radicals, and metabolites that can act as alkylating agents, and exogenous environmental agents such as ultraviolet (UV) light from the sun, inhaled cigarette smoke, or incompletely defined dietary factors. DNA lesions in a human cell range from 100–500 spontaneous deaminations to 20,000–40,000 single strand breaks (Hoeijmakers, 2009). Loss of function or inefficient DNA repair can lead to chromosome aberrations, a characteristic prominent in many cancers. DNA repair genes play an important role in repairing damage DNA caused by endogenous and exogenous cancer-causing agents. The repair of damage to DNA is essential to the survival of the cell and the health of the organism and over evolutionary periods the cell has developed a diverse set of defense mechanisms to deal with a wide range of DNA lesions and adducts (Friedberg *et al*., 2005) . Defects in DNA repair and genomic maintenance can cause chromosomal aberrations, and it is plausible that genetic variations in DNA repair genes influence lymphoma risk. So far there are more than 150 human DNA repair genes known (Friedberg *et al*., 2005; Wood *et al*., 2001) with five major pathways in DNA repair pathways: nucleotide excision repair (NER),

base excision repair (BER), mismatch repair (MMR), double-strand break repair (DSBR), and direct repair.

Many published data suggest that DNA repair pathways are associated with lymphoma susceptibility (Kim *et al*., 2010; Li *et al*., 2016; Smedby *et al*., 2006). In their study, Smedby *et al*. (2006) demonstrated that polymorphic variation in the *XRCC3* gene, but not in the *ERCC2* and *XRCC1* genes, may be important for susceptibility to FL. This study examined a total of seven promising functional SNPs (Table 2.3) in four key genes involved in three out of five DNA repair pathways, which enabled a broad understanding of DNA repair mechanisms involved in the pathogenesis of NHL.

Table 2.3: Summary of DNA repair SNPs examined in this study.

Gene	Type of DNA repair pathway	Chromosome	SNP rs no	Position (hg 19)	Gene Region	Amino Acid change	dbSNP functional annotation	PolyPhen-2 prediction[a]	SIFT prediction[b]
ERCC5	Nucleotide excision repair (NER)	13q33.1	rs17655 (G>C)	103528002	Exon 15	D1104H	missense	Probably damaging	Deleterious
XRCC3	Double stand break repair (DSBR)	14q32.3	rs861539 (G>A)	104165753	Exon 8	T241M	missense	Possibly damaging	Tolerated
XRCC1	Base excision repair (BER)	19q13.2	rs25487 (C>T)	44055726	Exon 10	Q399R	missense	Benign	Tolerated
			rs25489 (C>T)	44056412	Exon 9	R280H	missense	Possibly damaging	Tolerated
			rs1799782 (G>A)	44057574	Exon 6	R194W	missense	Probably damaging	Deleterious
ERCC2	Nucleotide excision repair (NER)	19q13.3	rs13181 (T>G)	45854919	Exon 23	K751Q	missense	Benign	Tolerated
			rs1799793 (C>T)	45867259	Exon 10	D312N	missense	Benign	Tolerated

Abbreviations: SNP: single nucleotide polymorphism; dbSNP: The Single Nucleotide Polymorphism database; PolyPhen: polymorphism phenotyping; SIFT: sorting tolerant from intolerant.

[a] dbSNP (https://www.ncbi.nlm.nih.gov/projects/SNP/)
[b] PolyPhen-2 prediction (http://genetics.bwh.harvard.edu/pph2/)
[c] SIFT prediction (http://sift.jcvi.org/)

2.4 Genetic association and DLBCL prognosis

A new insight into the pathogenesis of DLBCL suggested that it is a heterogeneous group of B-cell lymphomas rather than a single clinicopathologic entity (Lossos, 2005). Therefore, effective risk-adapted strategies are needed to improve the outcome of patients with DLBCL. Although the International Prognostic Index (IPI) (Shipp *et al*., 1993) or revised IPI (Sehn *et al*., 2007) have been used over years as the standard clinical tools to predict DLBCL patient outcomes; However, DLBCL patients' outcomes still differ significantly within IPI categories.

2.4.1 DLBCL prognosis

DLBCL associated with a generally poorer prognosis (median survival, about one year), and is curable in less than 50% of patients. Given the marked heterogeneity of DLBCL, a reliable prediction tool is vital for optimizing patient care. The CHOP (cyclophosphamide, doxorubicin, vincristine, and prednisone) chemotherapy regimen had been the mainstay of therapy for several decades ago. Later, the introduction of rituximab to CHOP (R-CHOP) therapy has clearly changed the prognosis of patients with DLBCL, with approximately half of the patients achieving long-term disease-free survival (Sehn *et al*., 2007). Although the adoption of R-CHOP as the new standard of care has led to improved outcomes for this curable lymphoma, patients whose lymphoma is not cured by first-line therapy continue to pose a difficult challenge.

The IPI based on pre-treatment characteristics has been the primary clinical tool used in the management of patients with DLBCL since its publication in 1993 (Shipp *et al*., 1993). Originally described in the pre-rituximab era, this model identified five factors to predict DLBCL survival: age >60, elevated serum lactate dehydrogenase (LDH), Eastern Cooperative Oncology Group (ECOG) performance status $\geqslant 2$, Ann Arbor stage III or IV, and a number of involved extranodal sites $\geqslant 2$. Four risk groups were identified

that predicted 5-year survival rates of 73%, 51%, 43%, and 26%, respectively (International Non-Hodgkin's Lymphoma Prognostic Factors Project, 1993). Early identification of poor-risk patients may allow for alternate treatment strategies to be considered. Previously, the attempt to predict the outcome of DLBCL into prognostically favorable GCB and unfavorable ABC subtype based on gene expression signatures had resulted in inconsistent results (Shipp *et al*., 2002). Therefore, effective risk-adapted strategies are needed to improve the outcome of patients with DLBCL. New biological markers such as genetic markers that reflect the heterogeneity of DLBCL must be evaluated to better determine the patient outcomes.

CHAPTER 3: MATERIAL AND METHODOLOGY

3.1 Study subjects and clinical characteristics

A total of 1321 subjects, including 652 NHL patients (304 Malay; 263 Chinese; 85 Indians) and 669 healthy controls (308 Malay; 270 Chinese; 91 Indians) were included in this study. NHL patients aged between 18-91 years old were recruited from two hospitals in Klang Valley: University Malaya Medical Centre (UMMC) and Ampang Hospital from April 2007 to July 2014. NHL types and subtypes were classified according to the WHO 2008 classification system. The inclusion criteria for cases include adult patients aged 18-year-old and above at the time of recruitment, HIV-negative NHL patients, and patients without undergoing hematopoietic stem cell transplantation. All 669 unrelated healthy controls were aged between 18-85 years old. Younger subjects were recruited from various blood donation campaign in Klang Valley during the same period of time while older control subjects were collected from an outpatient clinic in UMMC and were ensured free from any underlying malignancies. Both NHL patients and healthy controls were unrelated self-reported ethnicity of Malay, Chinese and Indian descent. Written consent was obtained from all patients and healthy controls prior to blood/ buccal samples collection. This study was approved by UMMC ethics committee (Reference no.: 565.7) and registered under the National Medical Research Register (NMRR; Research ID: 21855) from the Ministry of Health (MOH) Malaysia.

3.2 Genomic DNA extraction and quantification

Genomic DNA was extracted from peripheral blood using QIAamp DNA Blood Midi Kit or QIAamp DNA Blood Mini Kit for buccal swabs (QIAGEN Inc, Germantown, MD, USA) following manufacturer's specifications. The quality and quantity of extracted DNA were determined spectrophotometry using NanoDrop® ND-1000 spectrophotometer (Thermo Fisher Scientific Inc., Wilmington, DE, USA).

3.3 Taqman® SNP genotyping method

Genotyping was performed using TaqMan® SNP genotyping assay protocol. The TaqMan probes (Appendix A) and TaqMan® GTXpress™ Master Mix were designed and purchased from Applied Biosystems, Foster City, CA. The reaction mixture was prepared as shown in the Table 3.1.

Table 3.1: The reaction mixture of TaqMan® SNP genotyping method.

Reagent	**1x Working Reaction**
dH_2O	1.25µl
TaqMan® GTXpress™ Master Mix	2.5µl
20x TaqMan probes mixture (Appendix A)	0.25µl
5ng/ul genomic DNA	1µl
Total volume	5µl

At least one non-template controls using DNAase-free water and two DNA controls were included in each round of Taqman SNP genotyping method. The two DNA controls were successfully amplified using conventional polymerase chain reaction (PCR) method. Working reaction in the final volume of 5µl in each well was prepared in the MicroAmp Optical 384-well reaction plate. PCR amplification was performed in QuantStudio™ 12K Flex Real-Time PCR System (Applied Biosystems, Foster City, CA). The PCR conditions were described in the Table 3.2:

Table 3.2: The PCR conditions that were set in the QuantStudio™ 12K flex real-time PCR system.

Cycle Step	Temperature (°C)	Time	Cycles
Initial	95	10 min	1
Denaturation	95	15 sec	40 cycles
Annealing	58	1 min	
Extension	72	15 sec	
Final Extension	72	5 min	1

The fluorescent signal intensities and the ratios of FAM/ROX and VIC/ROX were generated at the end of PCR reaction. The endpoint allelic discrimination plate read was performed using the Applied Biosystems QuantStudio™ 12K Flex Real-Time PCR System. SNPs showing genotype call rates >95% were retained for statistical analysis. To evaluate the concordance of the genotyping calls, 5% of the samples for each SNP were randomly selected for re-genotyping.

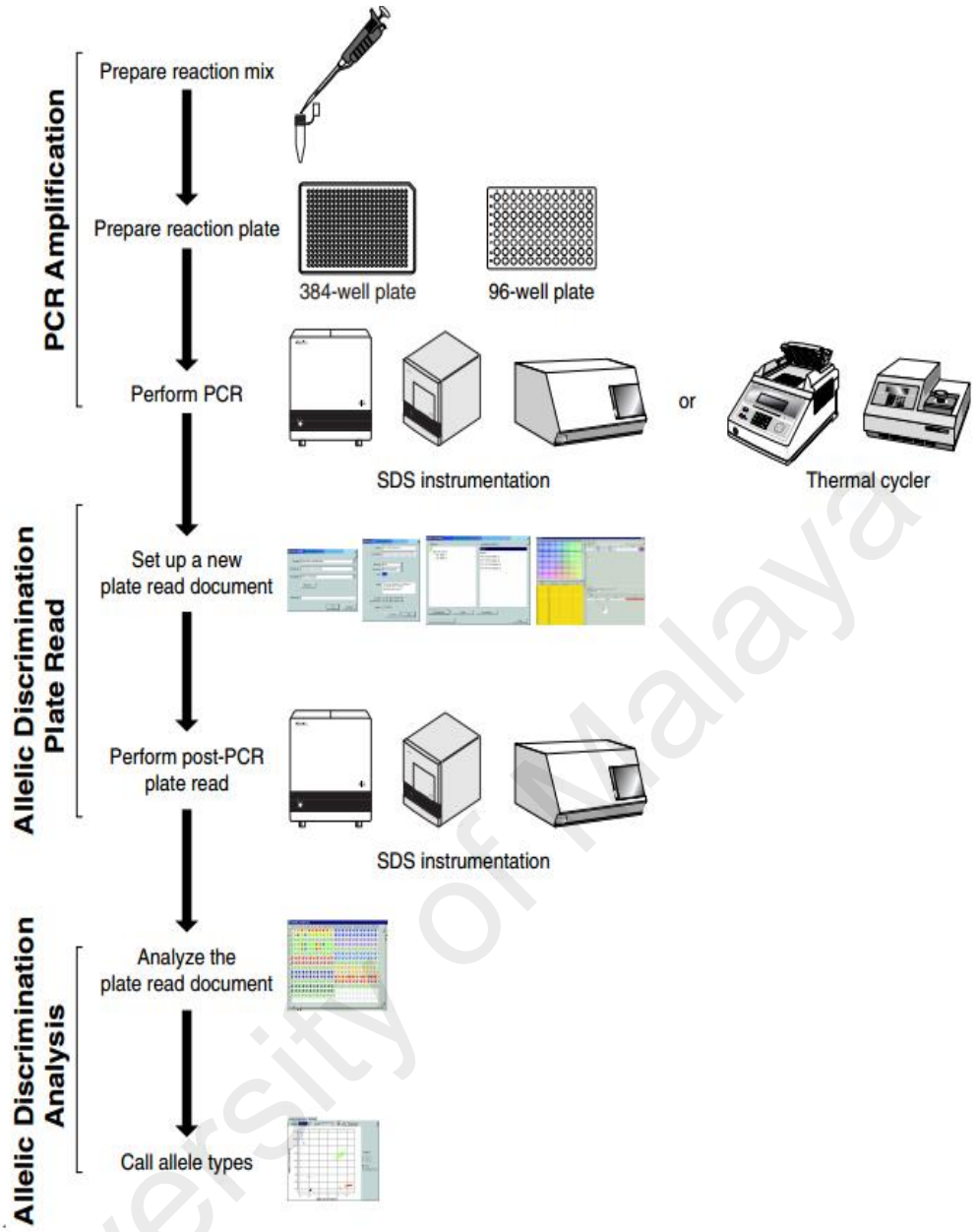


Figure 3.1: A simplified overview of the TaqMan SNP genotyping workflow. (Retrieved from TaqMan SNP genotyping assay protocol).

3.4 Statistical analyses and interpretation of genotyping data with NHL susceptibility

The SNP association analysis was performed separately for Malay, Chinese, Indian group, the results were later combined as pooled subjects. SNP association was calculated using logistic regression analysis evaluating 3 genetic models: additive, dominant and recessive model, adjusted for age and gender. Odds ratio (OR) and 95% confidence intervals (CI) were calculated according to the corresponding genetic model, with the minor allele as a reference. SNP association was calculated collectively for all NHL types as well as being stratified into types of B-cell NHL and T-cell NHL, and subtypes of DLBCL and FL. The genetic model that yielded the smallest *P*-value was considered the best-fitting model. Bonferroni threshold for multiple testing of 17 SNPs in this study was applied ($\alpha = 0.05/17 = 0.003$). Associations of SNPs with NHL were deemed significant at $P<0.003$. Deviation from Hardy–Weinberg equilibrium (HWE) was assessed using goodness-of-fit chi-square (χ^2) test on control samples. $P_{HWE}<0.05$ shows deviation from HWE equilibrium, rendering the association invalid. Breslow-Day test was performed to assess the heterogeneity of association across the Malay, Chinese and Indian groups in view of their different ancestry. All statistical analysis was performed using PLINK v1.07 unless stated otherwise (Purcell *et al*., 2007). PLINK is a free, open-source whole genome association analysis toolset, designed to perform a range of basic, large-scale analyses in a computationally efficient manner. Meta-analysis was performed using R programming software (https://www.r-project.org/) by downloading the R package Meta. This study assumes a random-effects model to estimate the combined OR and *P*-value of SNPs in the pooled subjects' analysis of Malay, Chinese and Indian groups. A random-effects model was chosen considering the different ancestry of the Malay, Chinese and Indian groups.

3.5 *Cis*-eQTL analysis and linkage disequilibrium analysis of NHL associated SNPs

The potential *cis*-eQTL function of significant SNPs was evaluated by mining data from Genotype - Tissue Expression (GTEx) V6 portal (The GTEx Consortium, 2015). GTEx project aims to cataloguing genetic variation and its influence on gene expression within and between all major tissues in the human body Considering most NHL types evaluated in this study were B-cell NHL, *cis*-eQTL associations were mined from EBV-transformed lymphocytes. Details of the *cis*-eQTL analysis have been mentioned elsewhere (The GTEx Consortium, 2015). Briefly, *cis*-eQTL was performed in Matrix eQTL (Shabalin, 2012) using linear regression, assuming an additive model, adjusting for 3 principal components and PEER factors. The *cis*-eQTL was defined as a window of +/- 1MB from the transcription start site (TSS). Storey FDR is used to correct for multiple hypotheses, with *cis*-eQTLs chosen as significant for q-values $\leqslant 0.05$. Pairwise linkage disequilibrium (LD) between SNPs in close proximity was calculated using Haploview v4.2 (Barrett *et al*., 2005) to assess the linkage or independence of the variants. LD was calculated using both Pearson's correlation (r^2) and Lewontin's D-prime (D'). Haploview is designed to simplify and expedite the process of haplotype analysis and fully compatible with data from the HapMap project and the Perlegen Genotype Browser.

3.6 DLBCL patients prognosis data

Due to the low incidence of other NHL subtypes, survival analysis was only conducted for subtype DLBCL. Clinical data of DLBCL patients were retrieved from medical records: date of diagnosis, the age of patients at diagnosis, clinical stage, nodal status, tumour size and grade, histological classification of the tumour, performance status, type of treatment, results of treatment for survival study. DLBCL patients were grouped into three risk groups (low, intermediate, and high-risk group) according to their IPI scores. All DLBCL patients were followed-up regularly (every 12 months) until the end

of the study (30 April 2015). Their survival time, date and cause of death, relapse status and last follow-up time were recorded. At the end of the study, newly diagnosed DLBCL patients that did not achieve at least 2 years follow-up time were not included in this study, except if there was an event (death/relapse/progressed) occurred in those patients.

3.7 Statistical analyses and interpretation of genotyping data with DLBCL prognosis

Outcomes of DLBCL were measured using overall survival (OS) and progression-free survival (PFS). OS is defined as an interval from the date of diagnosis until the date of death from any cause or until the last follow-up date. PFS is defined as an interval measured from the date of diagnosis until the date of progression, relapsed or death from any cause or until the last follow-up date. The Kaplan-Meier (K-P) method was used to construct survival curves for PFS and OS, and results were compared using log-rank test. The Cox regression model was used to evaluate the association of SNPs with DLBCL prognosis in case-only analysis. The homozygote of the most common allele was used as the reference group, and the heterozygote and homozygote variant genotypes were grouped together and coded as 1. Hazard ratios were adjusted against risk factors. An effect was considered statistically significant at $P<0.05$. The survival analysis was performed using the Statistical Package for the Social Sciences (SPSS) 22.0 software. SPSS is a widely used program for statistical analysis in social science.

3.8 Flow chart of experiments

Study subjects

- **Sample size:** Total 1321 (652 NHL cases & 669 healthy controls)
- **Sample source:** University Malaya Medical Centre (UMMC) & Ampang Hospital
- **Recruitment duration :** April 2007 to July 2014
- **Ethical Approval :** UMMC Ethics committee (Reference no.: 565.7) & National Medical Research Register(NMRR;Research ID: 21855)

- **Case selection:** Adult patients ≥18 years old; no mix-parentage for at least 3 generations; HIV-negative; not underwent blood transfusion.
- **Control selection**: Healthy individual ≥18 years old; no mix-parentage for at least 3 generations.

DNA extraction & quantification

- **DNA extraction:** QIAamp DNA Blood Midi Kit protocol
- **DNA quantification:** NanoDrop® ND-1000 spectrophotometer

Taqman SNPs Genotyping method

DLBCL patients prognosis outcome & follow-up

- **Study period:** April 2007- April 2015
- **Event-free survival (EFS):** an interval measured from the date of diagnosis until the date of progression, relapse or death from any cause or until the last follow-up date.
- **Overall survival (OS)** : an interval from the date of diagnosis until the date of death from any cause or until the last follow-up date.

Results analysis

- **SNP association analysis** PLINK v1.07 (http://pngu.mgh.harvard.edu/~purcell/plink/)
- **Meta analysis:** R Programming (https://www.r-project.org/)
- **eQTL analysis**: GTEx V6 (http://www.gtexportal.org/home/)
- **Linkage Disequilibrium analysis**: Haploview v4.2 (https://www.broadinstitute.org/haploview/haploview)
- **DLBCL prognosis:** SPSS 22.0 (https://www.ibm.com/analytics/data-science/predictive-analytics/spss-statistical-software)

CHAPTER 4: RESULTS

4.1 Characteristics of study population

Over the period of 7 years, this study had successfully recruited a total of 1321 subjects into this study, which included 652 NHL patients consisted of 304 (46.6%) Malays, 263 (40.3%) Chinese and 85 (13.0%) Indians and 669 healthy controls consisted of 308 (46.0%) Malays, 270 (40.4%) Chinese and 91 (13.6%) Indians. Of the 652 cases, 56% were male (age range = 18-91 years old; median = 58 years old), and 44% were female (age range = 18-87 years old; median = 55 years old). Out of 669 controls, 57% were males (age range = 18-85 years old; median=46 years old) and 43% were females (age range = 18-84 years old, median = 43.5 years old). Out of 652 NHL patients, 568 (87.1%) were B-cell NHLs and only 84 (12.9%) were T-cell NHLs. Majority of patients were DLBCL (n=334; 51.2%) followed by FL (n=84; 15.6%). Characteristics of study subjects were summarized in Table 4.1.

4.2 Genomic DNA extraction

Genomic DNA were successfully extracted from 652 NHL patients and 669 healthy controls, and approximately 10µg – 50µg DNA was obtained from 2ml – 5ml of blood. These DNA samples were then diluted to a concentration of 5-10 ng/µl for subsequent experiment. All the extracted DNA samples were relatively pure (A_{260}/A_{280} ratio ~1.70 -1.90) measured using the NanoDrop® ND-1000 spectrophotometer.

Table 4.1: Demographic characteristics of study subjects.

	Malay (n=612)		Chinese (n=533)		Indian (n=176)		Total (n=1321)	
	Case (%)	**Control (%)**	**Case (%)**	**Control (%)**	**Case (%)**	**Control (%)**	**Case (%)**	**Control (%)**
Gender								
Male	166 (54.6%)	170 (55.2%)	147 (55.9%)	156 (57.8%)	52 (61.2%)	55 (60.4%)	365 (56.0%)	381 (57.0%)
Female	138 (45.4%)	138 (44.8%)	116 (44.1%)	114 (42.2%)	33 (38.8%)	36 (39.6%)	287 (44.0%)	288 (43.0%)
Age[a]								
Mean age (SD)	51.8 (14.7)	39.7 (10.0)	57.1 (14.4)	49.1 (12.1)	57.0 (14.0)	45.2 (13.3)	54.6 (14.7)	44.2 (14.1)
Median age	54	38	60	50	58	45	57	45
Range age	18-91	19-85	18-87	18-84	19-83	18-81	18-91	18-85
Type of NHL:								
B-cell NHL	268 (88.2%)		228 (86.7%)		72 (84.7%)		568 (87.1%)	
DLBCL*	173 (56.9%)		129 (49.0%)		32 (37.6%)		334 (51.2%)	
FL*	38 (12.5%)		45 (17.1%)		19 (22.3%)		102 (15.6%)	
Other B cell[b]	57 (18.8%)		54 (20.5%)		21 (24.7%)		132 (20.2%)	
T-cell NHL[c]	36 (11.8%)		35 (13.3%)		13 (15.3%)		84 (12.9%)	
Total:	304 (100%)	308 (100%)	263 (100%)	270 (100%)	85 (100%)	91 (100%)	652 (100%)	669 (100%)

*Abbreviations: NHL, non-Hodgkin's lymphoma; DLBCL, diffuse large B-cell lymphoma; FL, follicular lymphoma

[a]Age at interviewed for cases and controls- Mean (Standard deviation)
[b]Others: Other types of B-cell NHL except DLBCL and FL
[c]T-cell NHL: Included all the subtypes of T-cell NHL

4.3 Interpretation of the Taqman SNP genotyping result

Figure 4.1 represented graphical output of a SNP with allele 1 and allele 2 as detected by the software of Applied Biosystems QuantStudio™ 12K Flex Real-Time PCR System. The genotyping results were obtained by detecting the fluorescent signal intensities of the TaqMan assay and display as a scatter plot of Allele X Rn versus Allele Y Rn (Figure 4.1).

The allelic discrimination plot contains three distinct clusters which represented three different genotypes vary along the horizontal axis (Allele X), the vertical axis (Allele Y) or diagonal (Allele X/Allele Y). The cluster of red spot signified a strong VIC/ROX signal and represented those individuals homozygous for allele 1; The cluster of green spot represented heterozygous individuals (having both allele 1 and 2), which caused increased in fluorescence signal from both reporter dyes (VIC and FAM); The cluster of blue spot indicated a strong FAM/ROX signals and represented individuals homozygous for allele 2.

Besides that, the normalized reported signal (Rn) values generated by alleles were checked to ensure the reproducibility of the result and to confirm the allelic discrimination. Allele X Rn represented the Rn for allele 1 (VIC/ROX) whereas allele Y Rn represented the Rn for allele 2 (FAM/ROX) (Figure 4.1).

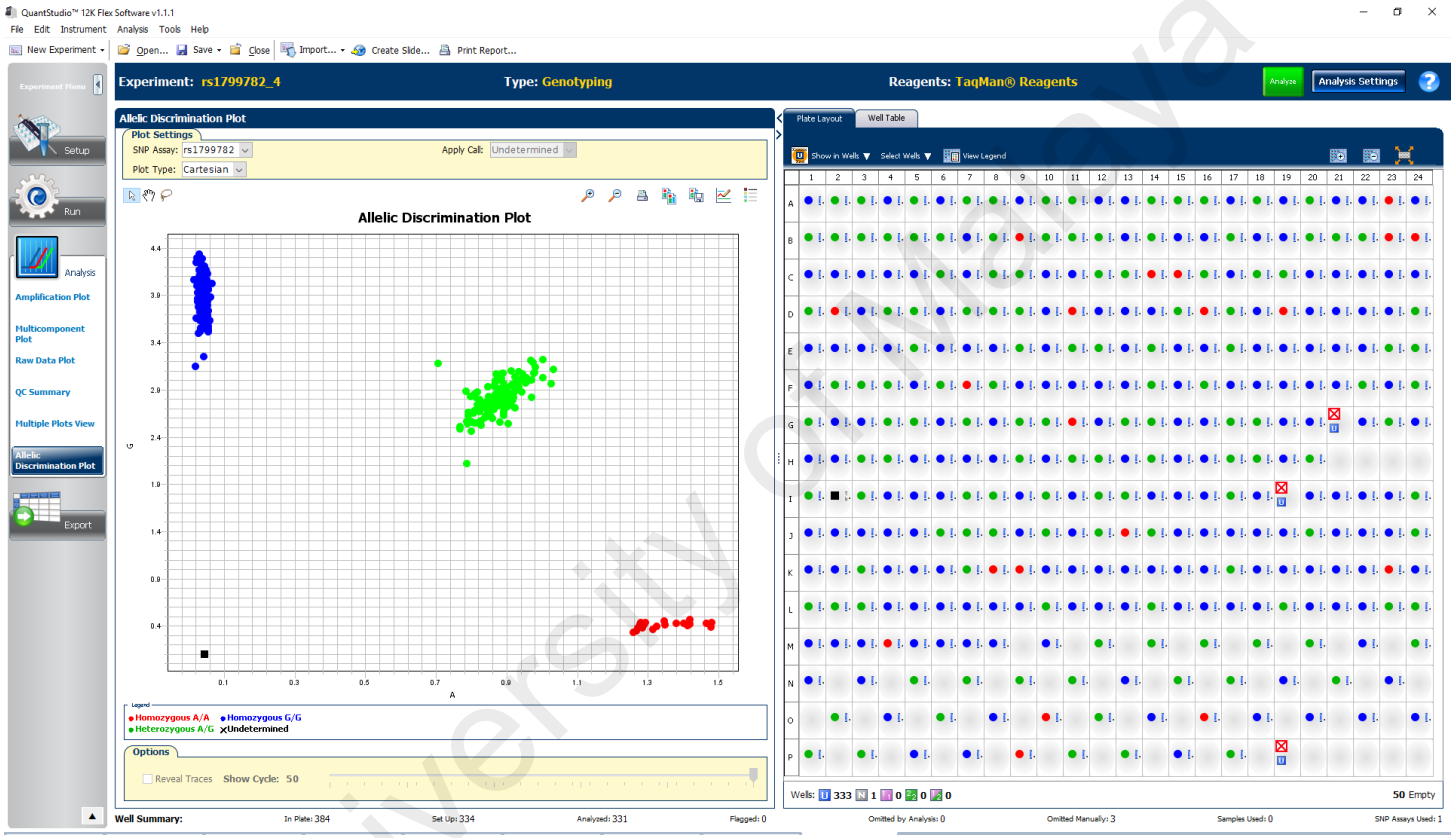


Figure 4.1: A representative graphical output of a SNP (rs1799782) detected by Applied Biosystems QuantStudio™ 12K Flex real-time PCR system. The cluster of red spot signified a strong VIC/ROX signal and represented those individuals homozygous for allele 1; The cluster of green spot represented heterozygous individuals (having both allele 1 and 2), which caused increased in fluorescence signal from both reporter dyes (VIC and FAM); The cluster of blue spot indicated a strong FAM/ROX signals and represented individuals homozygous for allele 2.

4.4 Quality control results

The genotyping call rates for the 17 SNPs were above 99%, except for rs2647012 (97.9%). To ensure the accuracy, 5% of the samples for each SNP were randomly selected to re-genotyped, and the genotype concordance rates were 100% (Table 4.2).

All the 17 SNPs were in HWE (P>0.05) in controls, in pooled subjects as well as after stratification according to ethnic groups, rendering the association is valid (Table 4.2). All the 17 SNPs had the MAF of more than 5% in control, except rs10484561, where G allele of rs10484561 had a very low prevalence in Chinese group ($MAF_{chinese\text{-}control}$ = 0.03). In overall, the allele distribution for SNPs in Chinese group was closed to the frequency distribution seen in the East Asian population based on web database 1000 Genomes Project (http://www.internationalgenome.org) Phase 3; However, the allele frequency of Malay and Indian group showed some disparity from 1000 Genomes Project Phase 3 East Asian population.

Table 4.2: Genotyping call rate, Hardy-Weinberg equilibrium (HWE) and minor allele frequencies (MAF) of SNPs examined in this study.

	Genotyping call rate (%)	Hardy-Weinberg equilibrium (HWE) *P*-value								Minor allele frequencies (MAF)								
		Pooled subjects		Malay		Chinese		Indian		Pooled subjects		Malay		Chinese		Indian		Ref [a]
		Case	Control	Case	Control	Case	Control	Case	Control	Case	Control	Case	Control	Case	Control	Case	Control	
Immune function SNPs																		
rs3832246 (T>-)	99.8	0.06	0.91	0.73	0.49	0.02	0.58	1.00	0.75	0.20	0.21	0.24	0.21	0.17	0.21	0.21	0.21	0.21
rs6773854 (T>C)	99.5	0.02	0.43	0.48	0.09	0.01	0.60	1.00	0.59	0.25	0.22	0.14	0.11	0.24	0.23	0.28	0.24	0.17
rs6457327 (C>A)	100	0.39	0.31	0.55	0.56	0.40	1.00	0.63	0.38	0.35	0.36	0.33	0.23	0.33	0.35	0.39	0.41	0.36
rs9271100 (C>T)	99.2	0.14	0.1	0.45	0.35	0.85	0.66	0.58	0.18	0.28	0.25	0.26	0.26	0.20	0.17	0.34	0.32	0.23
rs2647012 (C>T)	97.9	0.78	0.63	0.37	0.24	0.58	0.86	0.59	0.38	0.17	0.21	0.27	0.24	0.21	0.22	0.11	0.18	0.19
rs10484561 (T>G)	99.8	0.0001	0.14	0.008	0.47	1.00	1.00	0.34	0.47	0.11	0.09	0.07	0.08	0.03	0.03	0.19	0.14	0.06
rs1571011 (A>C)	99.6	0.007	0.39	0.08	0.45	0.04	0.72	0.39	0.09	0.45	0.44	0.55	0.42	0.47	0.51	0.37	0.35	0.50
rs1977389 (T>G)	99.8	0.69	0.70	0.72	0.63	0.62	1.00	0.38	0.53	0.47	0.46	0.53	0.43	0.47	0.51	0.43	0.40	0.50
rs10887878 (T>A)	99.5	0.07	1.00	0.11	1.00	0.80	0.39	0.78	0.23	0.44	0.45	0.27	0.21	0.43	0.48	0.49	0.50	0.43
rs4985700 (A>C)	99.8	0.57	0.79	0.52	0.39	0.18	1.00	0.82	0.16	0.29	0.31	0.40	0.33	0.20	0.24	0.34	0.37	0.24
DNA Repair SNPs																		
rs17655 (G>C)	99.8	0.52	0.94	0.82	0.20	0.90	0.46	0.53	0.57	0.42	0.42	0.22	0.25	0.47	0.47	0.43	0.43	0.46
rs861539 (G>A)	100	0.81	0.29	0.49	0.72	0.61	1.00	1.00	0.38	0.09	0.08	0.16	0.14	0.06	0.05	0.09	0.09	0.07
rs25487 (C>T)	99.7	0.15	0.93	0.02	0.37	0.65	0.88	0.80	0.14	0.29	0.30	0.31	0.30	0.28	0.26	0.28	0.33	0.24
rs25489 (C>T)	99.9	0.22	1.00	0.33	1.00	1.00	1.00	0.34	1.00	0.11	0.10	0.13	0.10	0.11	0.10	0.10	0.10	0.10
rs1799782 (G>A)	99.8	0.15	0.06	0.23	0.15	1.00	0.57	0.54	0.65	0.29	0.27	0.09	0.13	0.32	0.31	0.31	0.28	0.28
rs13181 (T>G)	100	0.53	0.16	0.51	0.11	1.00	0.67	0.43	0.30	0.14	0.13	0.28	0.29	0.08	0.08	0.16	0.13	0.08
rs1799793 (C>T)	100	0.008	0.23	0.25	0.05	0.09	1.00	1.00	0.17	0.11	0.09	0.30	0.25	0.05	0.05	0.11	0.08	0.05

*Abbreviations: SNP: Single nucleotide polymorphism.

[a] Reference MAF data collected from 1000 Genomes Project Phase 3 East Asian population (http://www.internationalgenome.org).

4.5 Association between the immune function SNPs and NHL susceptibility

After performing multiple corrections using Bonferroni method, logistic regression results showed that two SNPs: rs2647012(C>T) and rs10484561 (T>G) have a strong association with NHL susceptibility in the Malaysian population. SNP rs2647012 (C>T) was significantly associated with reduced risk when analysed in all NHL cases in Malay group ($P_{\text{additive-adjusted}}$=0.0008; OR=0.54; 95% CI=0.37-0.77) and B-cell NHL in both pooled subjects ($P_{\text{additive-adjusted}}$=0.0008; OR=0.69; 95% CI=0.55-0.87) and Malay group ($P_{\text{additive-adjusted}}$=0.0007; OR=0.51; 95% CI=0.35-0.76) (Table 4.3 & Table 4.4). The best association for rs2647012 (C>T) was observed for the additive model. Note of caution, the combined association in pooled subjects was largely driven by the association from the Malay group. SNP rs10484561 (T>G) showed significant association with subtype FL in pooled subjects ($P_{\text{dominant-adjusted}}$=0.003; OR=2.39; 95% CI=1.35-4.22) and Malay group ($P_{\text{dominant-adjusted}}$=0.003; OR=3.50; 95% CI=1.54-7.97) (Table 4.3 & Table 4.4). The best association for rs10484561 (T>G) was observed for the dominant model. Both association hold true as it showed no deviation from HWE (Table 4.2). Due to the small samples size especially after stratification into different NHL types and subtypes, no significant association between SNPs and NHL risk was observed in Chinese and Indian group.

No statistically significant association was observed for other eight immune function SNPs (Table 4.3, Table 4.4, Table 4.5 & Table 4.6). However, weak association (P<0.05) for a few SNPs was observed. rs3832246 (T>-) was weakly associated with all NHL ($P_{\text{dominant-adjusted}}$=0.02; OR=0.65; 95% CI=0.44-0.95), B-cell NHL ($P_{\text{dominant-adjusted}}$=0.008; OR=0.58; 95% CI=0.39-0.87), and DLBCL ($P_{\text{dominant-adjusted}}$=0.004; OR=0.49; 95% CI=0.30-0.79) in Chinese group (Table 4.5). rs6773854 (T>C) was weakly associated with B-cell NHL ($P_{\text{additive-adjusted}}$=0.02; OR=1.27; 95% CI=1.04-1.55) and DLBCL ($P_{\text{additive-adjusted}}$=0.004; OR=1.38; 95% CI=1.11-1.72) in pooled subjects

(Table 4.3) and DLBCL ($P_{\text{recessive-adjusted}}$=0.02; OR=2.65; 95% CI=1.17-60.0) in Chinese group (Table 4.5). rs9271100 (C>T) showed weakly association with T-cell NHL ($P_{\text{dominant-adjusted}}$=0.005; OR=1.98; 95% CI=1.23-3.18) in pooled subjects (Table 4.3). rs4985700 (A>C) showed weakly association with T-cell NHL ($P_{\text{additive-adjusted}}$=0.02; OR=0.46; 95% CI=0.25-0.86) in Malay group (Table 4.4). Considering large number of SNPs (n=17) examined in this study, all weak association results (P<0.05) shall not take into consideration as this might happen due to chance, as P-values did not reach the level of significance (P<0.003) set for Bonferroni correction.

Table 4.3: Association study of SNPs in immune function genes and NHL risk in pooled subjects.

	Additive Model (A vs a)		Dominant Model (AA + Aa vs aa)		Recessive Model (AA vs Aa + aa)	
	OR (95% CI)	***P*-value**	**OR (95% CI)**	***P*-value**	**OR (95% CI)**	***P*-value**
All NHL						
rs3832246 (T>-)	0.92 (0.76-1.13)	0.44	0.86 (0.68-1.10)	0.23	1.18 (0.68-2.06)	0.55
rs6773854 (T>C)	1.21 (1.00-1.46)	0.05	1.21 (0.95-1.53)	0.12	1.52 (0.95-2.45)	0.08
rs6457327 (C>A)	0.98 (0.83-1.16)	0.82	0.95 (0.75-1.20)	0.65	1.03 (0.74-1.45)	0.85
rs9271100 (C>T)	1.18 (0.98-1.42)	0.08	1.23 (0.97-1.56)	0.09	1.25 (0.81-1.92)	0.31
rs2647012 (C>T)	**0.76 (0.61-0.94)**	**0.01[d]**	**0.74 (0.58-0.95)**	**0.02[d]**	**0.60 (0.32-1.13)**	**0.11**
rs10484561 (T>G)	**1.33 (1.02-1.74)**	**0.03[d]**	**1.35 (0.99-1.83)**	**0.06**	**2.01 (0.83-4.86)**	**0.12**
rs1571011 (A>C)	0.99 (0.85-1.17)	0.94	0.93 (0.73-1.19)	0.55	1.08 (0.81-1.44)	0.59
rs1977389 (T>G)	1.04 (0.88-1.22)	0.67	1.07 (0.83-1.38)	0.61	1.02 (0.77-1.36)	0.87
rs10887878 (T>A)	0.99 (0.84-1.16)	0.87	0.91 (0.71-1.17)	0.46	1.08 (0.81-1.44)	0.59
rs4985700 (A>C)	0.95 (0.80-1.14)	0.61	0.92 (0.73-1.17)	0.51	0.99 (0.66-1.49)	0.98
B-cell NHL[a]						
rs3832246 (T>-)	0.91 (0.74-1.12)	0.37	0.84 (0.65-1.08)	0.18	1.19 (0.67-2.10)	0.56
rs6773854 (T>C)	**1.27 (1.04-1.55)**	**0.02[d]**	**1.29 (1.01-1.65)**	**0.05**	**1.63 (1.00-2.66)**	**0.05**
rs6457327 (C>A)	0.95 (0.80-1.14)	0.59	0.90 (0.70-1.15)	0.38	1.03 (0.72-1.47)	0.87
rs9271100 (C>T)	1.13 (0.93-1.37)	0.23	1.14 (0.89-1.46)	0.29	1.23 (0.78-1.93)	0.37
rs2647012 (C>T)	**0.69 (0.55-0.87)**	**0.0008[c]**	**0.65 (0.50-0.85)**	**0.002[c]**	**0.61 (0.31-1.18)**	**0.14**
rs10484561 (T>G)	**1.38 (1.04-1.82)**	**0.02[d]**	**1.39 (1.00-1.91)**	**0.05**	**2.26 (0.92-5.56)**	**0.08**
rs1571011 (A>C)	1.02 (0.86-1.20)	0.82	0.95 (0.73-1.23)	0.68	1.14 (0.85-1.53)	0.38
rs1977389 (T>G)	1.04 (0.88-1.24)	0.66	1.09 (0.84-1.43)	0.51	1.01 (0.75-1.35)	0.97
rs10887878 (T>A)	1.00 (0.84-1.19)	0.97	0.94 (0.72-1.23)	0.67	1.07 (0.79-1.44)	0.68
rs4985700 (A>C)	0.99 (0.82-1.19)	0.89	0.96 (0.75-1.22)	0.72	1.06 (0.70-1.62)	0.78
DLBCL						
rs3832246 (T>-)	0.84 (0.66-1.08)	0.17	0.75 (0.56-1.01)	0.06	1.21 (0.63-2.32)	0.56
rs6773854 (T>C)	**1.38 (1.11-1.72)**	**0.004[d]**	**1.40 (1.05-1.86)**	**0.02[d]**	**1.96 (1.16-3.31)**	**0.01[d]**
rs6457327 (C>A)	0.94 (0.77-1.15)	0.56	0.89 (0.67-1.18)	0.42	1.00 (0.66-1.50)	0.98
rs9271100 (C>T)	1.10 (0.88-1.36)	0.42	1.09 (0.82-1.44)	0.57	1.25 (0.75-2.08)	0.39
rs2647012 (C>T)	**0.72 (0.56-0.94)**	**0.02[d]**	**0.67 (0.49-0.90)**	**0.008[d]**	**0.77 (0.37-1.59)**	**0.47**
rs10484561 (T>G)	1.35 (0.99-1.84)	0.06	1.31 (0.91-1.89)	0.15	2.63 (0.99-6.93)	0.05
rs1571011 (A>C)	0.95 (0.79-1.15)	0.61	0.82 (0.61-1.10)	0.19	1.12 (0.80-1.57)	0.53
rs1977389 (T>G)	1.02 (0.84-1.24)	0.87	1.02 (0.76-1.39)	0.88	1.02 (0.73-1.43)	0.92
rs10887878 (T>A)	1.09 (0.90-1.33)	0.39	1.05 (0.77-1.43)	0.75	1.22 (0.87-1.70)	0.25
rs4985700 (A>C)	1.10 (0.88-1.36)	0.40	1.10 (0.83-1.46)	0.52	1.20 (0.75-1.92)	0.44
FL						
rs3832246 (T>-)	1.12 (0.75-1.66)	0.58	1.22 (0.77-1.95)	0.40	0.72 (0.20-2.60)	0.61
rs6773854 (T>C)	1.04 (0.70-1.55)	0.83	0.97 (0.60-1.56)	0.89	1.54 (0.59-4.02)	0.38
rs6457327 (C>A)	0.95 (0.68-1.31)	0.74	0.82 (0.52-1.29)	0.39	1.20 (0.64-2.24)	0.57
rs9271100 (C>T)	1.16 (0.81-1.66)	0.43	1.31 (0.82-2.07)	0.26	0.89 (0.35-2.21)	0.80
rs2647012 (C>T)	**0.59 (0.37-0.93)**	**0.02[d]**	**0.58 (0.35-0.96)**	**0.03[d]**	**0.26 (0.03-1.99)**	**0.19**
rs10484561 (T>G)	**2.08 (1.28-3.40)**	**0.003**	**2.39 (1.35-4.22)**	**0.003[c]**	**2.36 (0.49-11.31)**	**0.28**
rs1571011 (A>C)	1.19 (0.87-1.64)	0.28	1.44 (0.85-2.43)	0.18	1.11 (0.64-1.91)	0.72
rs1977389 (T>G)	1.15 (0.83-1.59)	0.41	1.66 (0.95-2.9)	0.07	0.86 (0.49-1.50)	0.59
rs10887878 (T>A)	0.82 (0.59-1.15)	0.25	0.84 (0.51-1.36)	0.47	0.67 (0.35-1.28)	0.23
rs4985700 (A>C)	0.78 (0.54-1.13)	0.19	0.72 (0.45-1.15)	0.17	0.77 (0.32-1.85)	0.56
T-cell NHL[b]						
rs3832246 (T>-)	1.00 (0.66-1.50)	0.99	1.01 (0.63-1.63)	0.96	0.91 (0.27-3.10)	0.88
rs6773854 (T>C)	0.95 (0.64-1.43)	0.82	0.88 (0.54-1.44)	0.62	1.27 (0.48-3.40)	0.63
rs6457327 (C>A)	1.14 (0.82-1.60)	0.43	1.35 (0.83-2.19)	0.23	0.95 (0.48-1.87)	0.88
rs9271100 (C>T)	**1.58 (1.12-2.25)**	**0.01**	**1.98 (1.23-3.18)**	**0.005[d]**	**1.40 (0.62-3.16)**	**0.41**
rs2647012 (C>T)	1.13 (0.76-1.68)	0.53	1.30 (0.81-2.07)	0.28	0.53 (0.12-2.29)	0.40
rs10484561 (T>G)	1.33 (0.78-2.25)	0.30	1.41 (0.78-2.55)	0.25	1.03 (0.12-8.69)	0.98
rs1571011 (A>C)	0.83 (0.60-1.16)	0.28	0.83 (0.51-1.35)	0.45	0.72 (0.38-1.35)	0.30
rs1977389 (T>G)	0.97 (0.70-1.33)	0.83	0.89 (0.54-1.46)	0.63	1.05 (0.60-1.82)	0.87
rs10887878 (T>A)	1.02 (0.74-1.43)	0.89	0.88 (0.53-1.45)	0.61	1.26 (0.73-2.18)	0.41
rs4985700 (A>C)	0.89 (0.62-1.29)	0.54	0.92 (0.57-1.46)	0.71	0.73 (0.30-1.76)	0.48

*Abbreviations: OR: odd ratio; CI: confidence interval; NHL: non-Hodgkin's lymphoma; DLBCL: diffuse large B-cell lymphoma; FL: follicular lymphoma; NA: not applicable.

[a]B-cell NHL: Include all the subtypes of B-cell NHL.
[b]T-cell NHL: Include all the subtypes of T-cell NHL.
[c]*P*-value that passed the Bonferroni correction ($P<0.003$).
[d]*P*-value that showed weak association ($P<0.05$).

Table 4.4: Association study of SNPs in immune function genes and NHL risk in Malay group.

	Additive Model (A vs a)		Dominant Model (AA + Aa vs aa)		Recessive Model (AA vs Aa + aa)	
	OR (95% CI)	***P*-value**	**OR (95% CI)**	***P*-value**	**OR (95% CI)**	***P*-value**
All NHL						
rs3832246 (T>-)	1.01 (0.74-1.37)	0.95	1.03 (0.71-1.49)	0.87	0.92 (0.40-2.11)	0.84
rs6773854 (T>C)	1.24 (0.94-1.63)	0.13	1.39 (0.98-1.99)	0.07	1.10 (0.58-2.07)	0.77
rs6457327 (C>A)	0.95 (0.74-1.22)	0.66	0.95 (0.66-1.37)	0.79	0.89 (0.56-1.43)	0.64
rs9271100 (C>T)	1.07 (0.83-1.39)	0.59	1.11 (0.78-1.59)	0.56	1.06 (0.61-1.83)	0.83
rs2647012 (C>T)	**0.54 (0.37-0.77)**	**0.0008[c]**	**0.52 (0.34-0.79)**	**0.002[c]**	**0.26 (0.08-0.87)**	**0.03[d]**
rs10484561 (T>G)	1.31 (0.94-1.82)	0.11	1.33 (0.90-1.96)	0.16	1.79 (0.68-4.73)	0.24
rs1571011 (A>C)	1.14 (0.88-1.46)	0.32	1.08 (0.75-1.54)	0.68	1.44 (0.87-2.40)	0.16
rs1977389 (T>G)	1.14 (0.89-1.47)	0.30	1.15 (0.80-1.67)	0.45	1.26 (0.79-2.02)	0.33
rs10887878 (T>A)	0.91 (0.71-1.17)	0.48	0.72 (0.48-1.08)	0.11	1.09 (0.73-1.64)	0.66
rs4985700 (A>C)	0.91 (0.7-1.190)	0.49	0.75 (0.52-1.08)	0.12	1.31 (0.76-2.28)	0.33
B-cell NHL[a]						
rs3832246 (T>-)	1.03 (0.75-1.42)	0.86	1.06 (0.72-1.55)	0.78	0.94 (0.40-2.23)	0.89
rs6773854 (T>C)	1.28 (0.96-1.70)	0.09	1.44 (0.99-2.09)	0.06	1.19 (0.62-2.28)	0.61
rs6457327 (C>A)	0.93 (0.72-1.21)	0.60	0.93 (0.63-1.36)	0.70	0.88 (0.54-1.45)	0.62
rs9271100 (C>T)	0.99 (0.76-1.31)	0.97	1.00 (0.69-1.46)	0.99	0.97 (0.54-1.72)	0.91
rs2647012 (C>T)	**0.51 (0.35-0.76)**	**0.0007[c]**	**0.50 (0.32-0.78)**	**0.002[c]**	**0.20 (0.05-0.76)**	**0.02[d]**
rs10484561 (T>G)	1.33 (0.94-1.87)	0.10	1.33 (0.88-2.01)	0.17	2.02 (0.75-5.43)	0.17
rs1571011 (A>C)	1.20 (0.92-1.56)	0.19	1.14 (0.78-1.66)	0.50	1.57 (0.93-2.66)	0.09
rs1977389 (T>G)	1.15 (0.88-1.51)	0.29	1.16 (0.79-1.71)	0.45	1.29 (0.79-2.11)	0.32
rs10887878 (T>A)	0.91 (0.70-1.19)	0.50	0.73 (0.48-1.13)	0.16	1.07 (0.70-1.63)	0.77
rs4985700 (A>C)	1.03 (0.78-1.36)	0.84	0.86 (0.59-1.26)	0.45	1.58 (0.89-2.80)	0.12
DLBCL						
rs3832246 (T>-)	1.01 (0.71-1.43)	0.97	0.97 (0.63-1.49)	0.90	1.20 (0.48-3.01)	0.70
rs6773854 (T>C)	1.30 (0.95-1.78)	0.10	1.47 (0.97-2.22)	0.07	1.25 (0.61-2.54)	0.54
rs6457327 (C>A)	0.90 (0.67-1.20)	0.47	0.85 (0.55-1.30)	0.45	0.90 (0.51-1.56)	0.70
rs9271100 (C>T)	0.97 (0.71-1.31)	0.82	0.94 (0.62-1.42)	0.76	1.00 (0.53-1.87)	0.99
rs2647012 (C>T)	**0.53 (0.35-0.82)**	**0.007[d]**	**0.51 (0.31-0.83)**	**0.007[d]**	**0.27 (0.07-1.10)**	**0.07**
rs10484561 (T>G)	1.22 (0.84-1.78)	0.30	1.15 (0.72-1.83)	0.55	2.26 (0.78-6.54)	0.13
rs1571011 (A>C)	1.09 (0.81-1.46)	0.58	0.98 (0.65-1.48)	0.92	1.46 (0.81-2.61)	0.21
rs1977389 (T>G)	1.12 (0.83-1.51)	0.46	1.11 (0.72-1.71)	0.64	1.25 (0.72-2.17)	0.44
rs10887878 (T>A)	1.04 (0.78-1.40)	0.77	0.91 (0.57-1.45)	0.69	1.24 (0.77-2.01)	0.38
rs4985700 (A>C)	1.17 (0.85-1.59)	0.34	1.05 (0.69-1.60)	0.84	1.68 (0.90-3.13)	0.10
FL						
rs3832246 (T>-)	1.40 (0.73-2.68)	0.32	1.67 (0.75-3.72)	0.21	0.89 (0.14-5.57)	0.90
rs6773854 (T>C)	1.04 (0.56-1.94)	0.90	0.98 (0.44-2.18)	0.96	1.34 (0.33-5.44)	0.68
rs6457327 (C>A)	1.01 (0.60-1.72)	0.96	0.75 (0.34-1.66)	0.48	1.56 (0.62-3.92)	0.35
rs9271100 (C>T)	0.98 (0.56-1.75)	0.96	1.32 (0.60-2.89)	0.49	0.42 (0.10-1.79)	0.24
rs2647012 (C>T)	0.44 (0.19-1.05)	0.06	0.48 (0.19-1.22)	0.12	NA	NA
rs10484561 (T>G)	**2.70 (1.38-5.27)**	**0.004[d]**	**3.50 (1.54-7.97)**	**0.003[c]**	**2.82 (0.44-17.94)**	**0.27**
rs1571011 (A>C)	1.57 (0.89-2.77)	0.12	1.80 (0.78-4.17)	0.17	1.84 (0.66-5.15)	0.25
rs1977389 (T>G)	1.41 (0.80-2.51)	0.24	1.88 (0.74-4.75)	0.18	1.28 (0.48-3.41)	0.62
rs10887878 (T>A)	0.73 (0.41-1.31)	0.29	0.61 (0.26-1.44)	0.26	0.75 (0.28-2.05)	0.58
rs4985700 (A>C)	0.79 (0.43-1.48)	0.47	0.72 (0.33-1.57)	0.40	0.89 (0.23-3.41)	0.86
T-cell NHL[b]						
rs3832246 (T>-)	1.00 (0.53-1.87)	0.99	1.05 (0.50-2.21)	0.89	0.68 (0.08-5.42)	0.71
rs6773854 (T>C)	1.08 (0.61-1.89)	0.80	1.22 (0.59-2.53)	0.60	0.75 (0.17-3.39)	0.71
rs6457327 (C>A)	1.04 (0.63-1.72)	0.87	1.14 (0.54-2.42)	0.73	0.94 (0.37-2.42)	0.90
rs9271100 (C>T)	1.58 (0.96-2.60)	0.07	2.18 (1.02-4.68)	0.05	1.45 (0.55-3.85)	0.46
rs2647012 (C>T)	0.61 (0.30-1.27)	0.18	0.55 (0.24-1.29)	0.17	0.55 (0.07-4.50)	0.58
rs10484561 (T>G)	1.48 (0.78-2.80)	0.23	1.68 (0.80-3.54)	0.17	1.05 (0.12-9.50)	0.97
rs1571011 (A>C)	0.82 (0.48-1.41)	0.48	0.75 (0.37-1.53)	0.44	0.85 (0.28-2.60)	0.77
rs1977389 (T>G)	1.01 (0.61-1.67)	0.96	0.92 (0.44-1.90)	0.82	1.20 (0.49-2.97)	0.69
rs10887878 (T>A)	0.98 (0.59-1.61)	0.93	0.66 (0.31-1.41)	0.28	1.46 (0.67-3.17)	0.34
rs4985700 (A>C)	**0.46 (0.25-0.86)**	**0.02[d]**	**0.43 (0.21-0.89)**	**0.02[d]**	**0.24 (0.03-1.85)**	**0.17**

*Abbreviations: OR: odd ratio; CI: confidence interval; NHL: non-Hodgkin's lymphoma; DLBCL: diffuse large B-cell lymphoma; FL: follicular lymphoma; NA: not applicable.

[a]B-cell NHL: Include all the subtypes of B-cell NHL.
[b]T-cell NHL: Include all the subtypes of T-cell NHL.
[c]*P*-value that passed the Bonferroni correction ($P<0.003$).
[d]*P*-value that showed weak association ($P<0.05$).

Table 4.5: Association study of SNPs in immune function genes and NHL risk in Chinese group.

	Additive Model (A vs a)		Dominant Model (AA + Aa vs aa)		Recessive Model (AA vs Aa + aa)	
	OR (95% CI)	***P*-value**	**OR (95% CI)**	***P*-value**	**OR (95% CI)**	***P*-value**
All NHL						
rs3832246 (T>-)	**0.77 (0.57-1.06)**	**0.11**	**0.65 (0.44-0.95)**	**0.02**[d]	**1.39 (0.58-3.32)**	**0.47**
rs6773854 (T>C)	1.03 (0.78-1.38)	0.82	0.90 (0.62-1.29)	0.56	1.92 (0.91-4.07)	0.09
rs6457327 (C>A)	0.86 (0.66-1.12)	0.27	0.75 (0.52-1.07)	0.11	1.04 (0.60-1.79)	0.90
rs9271100 (C>T)	1.28 (0.92-1.78)	0.14	1.26 (0.86-1.83)	0.24	2.09 (0.72-6.08)	0.18
rs2647012 (C>T)	0.96 (0.70-1.31)	0.79	1.00 (0.69-1.44)	0.98	0.73 (0.30-1.74)	0.48
rs10484561 (T>G)	1.14 (0.56-2.31)	0.72	1.14 (0.56-2.31)	0.72	NA	NA
rs1571011 (A>C)	0.90 (0.70-1.15)	0.39	0.81 (0.54-1.20)	0.29	0.93 (0.61-1.39)	0.71
rs1977389 (T>G)	0.87 (0.68-1.12)	0.27	0.86 (0.57-1.29)	0.45	0.80 (0.53-1.21)	0.29
rs10887878 (T>A)	0.87 (0.67-1.12)	0.27	0.81 (0.54-1.20)	0.29	0.85 (0.55-1.33)	0.48
rs4985700 (A>C)	0.85 (0.62-1.16)	0.31	0.89 (0.62-1.29)	0.55	0.51 (0.20-1.29)	0.15
B-cell NHL[a]						
rs3832246 (T>-)	**0.71 (0.51-0.99)**	**0.04**	**0.58 (0.39-0.87)**	**0.008**[d]	**1.25 (0.51-3.10)**	**0.62**
rs6773854 (T>C)	1.08 (0.80-1.46)	0.63	0.95 (0.65-1.39)	0.80	1.97 (0.90-4.29)	0.09
rs6457327 (C>A)	0.82 (0.62-1.08)	0.15	0.70 (0.48-1.02)	0.06	0.98 (0.55-1.75)	0.96
rs9271100 (C>T)	1.22 (0.86-1.72)	0.26	1.16 (0.78-1.72)	0.46	2.30 (0.76-6.95)	0.14
rs2647012 (C>T)	0.84 (0.60-1.16)	0.28	0.81 (0.55-1.20)	0.29	0.78 (0.32-1.94)	0.59
rs10484561 (T>G)	1.24 (0.59-2.58)	0.57	1.24 (0.59-2.58)	0.57	NA	NA
rs1571011 (A>C)	0.90 (0.70-1.16)	0.41	0.80 (0.53-1.21)	0.30	0.94 (0.61-1.43)	0.76
rs1977389 (T>G)	0.88 (0.68-1.14)	0.33	0.91 (0.59-1.40)	0.67	0.77 (0.50-1.19)	0.24
rs10887878 (T>A)	0.89 (0.68-1.16)	0.38	0.86 (0.57-1.30)	0.47	0.85 (0.53-1.35)	0.48
rs4985700 (A>C)	0.82 (0.59-1.14)	0.23	0.86 (0.59-1.26)	0.44	0.45 (0.16-1.29)	0.14
DLBCL						
rs3832246 (T>-)	**0.61 (0.40-0.92)**	**0.02**	**0.49 (0.30-0.79)**	**0.004**[d]	**1.06 (0.37-3.07)**	**0.91**
rs6773854 (T>C)	**1.21 (0.86-1.70)**	**0.28**	**1.03 (0.66-1.59)**	**0.90**	**2.65 (1.17-60.0)**	**0.02**[d]
rs6457327 (C>A)	0.79 (0.57-1.10)	0.16	0.71 (0.46-1.10)	0.13	0.82 (0.41-1.65)	0.58
rs9271100 (C>T)	1.17 (0.78-1.75)	0.45	1.12 (0.71-1.77)	0.64	2.03 (0.57-7.23)	0.27
rs2647012 (C>T)	0.94 (0.65-1.37)	0.75	0.93 (0.59-1.45)	0.74	0.93 (0.34-2.58)	0.89
rs10484561 (T>G)	1.29 (0.56-2.97)	0.54	1.29 (0.56-2.97)	0.54	NA	NA
rs1571011 (A>C)	0.92 (0.68-1.23)	0.57	0.81 (0.50-1.31)	0.39	0.98 (0.60-1.6)	0.94
rs1977389 (T>G)	1.12 (0.83-1.52)	0.46	1.19 (0.72-1.96)	0.50	1.15 (0.70-1.88)	0.58
rs10887878 (T>A)	0.96 (0.71-1.30)	0.79	0.84 (0.52-1.35)	0.46	1.09 (0.65-1.82)	0.75
rs4985700 (A>C)	0.85 (0.58-1.23)	0.38	0.87 (0.55-1.35)	0.53	0.58 (0.19-1.82)	0.35
FL						
rs3832246 (T>-)	0.85 (0.46-1.60)	0.62	0.87 (0.43-1.76)	0.69	0.60 (0.07-5.28)	0.65
rs6773854 (T>C)	1.03 (0.58-1.84)	0.91	0.93 (0.46-1.85)	0.82	1.83 (0.44-7.56)	0.41
rs6457327 (C>A)	0.81 (0.49-1.34)	0.41	0.66 (0.33-1.3)	0.23	1.05 (0.39-2.84)	0.92
rs9271100 (C>T)	1.47 (0.79-2.73)	0.22	1.34 (0.66-2.73)	0.42	4.29 (0.76-24.33)	0.10
rs2647012 (C>T)	0.64 (0.33-1.24)	0.19	0.60 (0.28-1.25)	0.17	0.63 (0.08-5.13)	0.66
rs10484561 (T>G)	1.46 (0.38-5.52)	0.58	1.46 (0.38-5.52)	0.58	NA	NA
rs1571011 (A>C)	0.81 (0.50-1.29)	0.37	0.91 (0.43-1.95)	0.81	0.59 (0.25-1.38)	0.22
rs1977389 (T>G)	1.25 (0.77-2.03)	0.36	2.36 (0.94-5.90)	0.07	0.85 (0.38-1.91)	0.70
rs10887878 (T>A)	0.67 (0.40-1.12)	0.12	0.79 (0.38-1.66)	0.54	0.33 (0.11-1.03)	0.06
rs4985700 (A>C)	0.67 (0.35-1.28)	0.22	0.64 (0.30-1.34)	0.24	0.49 (0.06-4.00)	0.51
T-cell NHL[b]						
rs3832246 (T>-)	1.17 (0.63-2.16)	0.62	1.12 (0.54-2.32)	0.76	1.74 (0.36-8.47)	0.50
rs6773854 (T>C)	0.82 (0.43-1.54)	0.53	0.62 (0.29-1.34)	0.23	2.03 (0.52-7.81)	0.31
rs6457327 (C>A)	1.11 (0.66-1.86)	0.71	1.16 (0.56-2.43)	0.69	1.10 (0.39-3.08)	0.86
rs9271100 (C>T)	1.86 (0.98-3.51)	0.06	2.11 (1.01-4.41)	0.05	1.57 (0.17-14.13)	0.69
rs2647012 (C>T)	1.70 (0.97-2.98)	0.07	2.54 (1.22-5.29)	0.01	0.50 (0.06-4.04)	0.52
rs10484561 (T>G)	0.86 (0.19-3.94)	0.84	0.86 (0.19-3.94)	0.84	NA	NA
rs1571011 (A>C)	1.17 (0.71-1.93)	0.55	1.37 (0.59-3.22)	0.46	1.11 (0.49-2.52)	0.81
rs1977389 (T>G)	1.25 (0.76-2.06)	0.38	1.15 (0.51-2.62)	0.73	1.56 (0.71-3.39)	0.27
rs10887878 (T>A)	0.77 (0.45-1.30)	0.32	0.63 (0.29-1.35)	0.23	0.84 (0.34-2.05)	0.71
rs4985700 (A>C)	1.19 (0.67-2.09)	0.55	1.35 (0.66-2.76)	0.42	0.88 (0.19-4.07)	0.87

*Abbreviations: OR: odd ratio; CI: confidence interval; NHL: non-Hodgkin's lymphoma; DLBCL: diffuse large B-cell lymphoma; FL: follicular lymphoma; NA: not applicable.

[a]B-cell NHL: Include all the subtypes of B-cell NHL.
[b]T-cell NHL: Include all the subtypes of T-cell NHL.
[d] *P*-value that showed weak association ($P<0.05$).

Table 4.6: Association study of SNPs in immune function genes and NHL risk in Indian group.

	Additive Model (A vs a)		Dominant Model (AA + Aa vs aa)		Recessive Model (AA vs Aa + aa)	
	OR (95% CI)	***P*-value**	**OR (95% CI)**	***P*-value**	**OR (95% CI)**	***P*-value**
All NHL						
rs3832246 (T>-)	1.25 (0.71-2.22)	0.45	1.21 (0.62-2.39)	0.58	1.92 (0.39-9.44)	0.42
rs6773854 (T>C)	1.44 (0.69-3.00)	0.34	1.39 (0.65-2.99)	0.40	NA	NA
rs6457327 (C>A)	1.38 (0.83-2.31)	0.22	1.56 (0.81-3.01)	0.18	1.32 (0.40-4.38)	0.65
rs9271100 (C>T)	0.99 (0.61-1.61)	0.97	1.04 (0.54-1.99)	0.91	0.85 (0.28-2.59)	0.77
rs2647012 (C>T)	1.20 (0.67-2.14)	0.54	1.12 (0.58-2.15)	0.75	2.46 (0.40-15.02)	0.33
rs10484561 (T>G)	0.74 (0.33-1.69)	0.48	0.73 (0.29-1.86)	0.51	0.48 (0.03-8.95)	0.63
rs1571011 (A>C)	1.55 (0.93-2.58)	0.09	1.63 (0.74-3.56)	0.22	1.90 (0.81-4.43)	0.14
rs1977389 (T>G)	1.45 (0.89-2.39)	0.14	1.77 (0.82-3.84)	0.15	1.45 (0.65-3.27)	0.37
rs10887878 (T>A)	1.26 (0.72-2.20)	0.41	1.22 (0.63-2.35)	0.55	2.06 (0.39-10.99)	0.40
rs4985700 (A>C)	1.14 (0.72-1.79)	0.58	1.39 (0.72-2.69)	0.32	0.89 (0.37-2.16)	0.80
B-cell NHL[a]						
rs3832246 (T>-)	1.36 (0.75-2.47)	0.32	1.3 (0.64-2.67)	0.47	2.40 (0.48-12.03)	0.29
rs6773854 (T>C)	1.69 (0.78-3.66)	0.18	1.65 (0.74-3.66)	0.22	NA	NA
rs6457327 (C>A)	1.30 (0.77-2.21)	0.32	1.34 (0.67-2.68)	0.41	1.65 (0.49-5.56)	0.42
rs9271100 (C>T)	0.99 (0.59-1.65)	0.96	0.99 (0.50-1.97)	0.98	0.96 (0.30-3.07)	0.94
rs2647012 (C>T)	1.14 (0.62-2.08)	0.68	1.00 (0.50-2.01)	1.00	3.09 (0.49-19.59)	0.23
rs10484561 (T>G)	0.75 (0.32-1.78)	0.51	0.73 (0.27-1.96)	0.53	0.58 (0.03-10.92)	0.72
rs1571011 (A>C)	1.38 (0.81-2.33)	0.23	1.27 (0.57-2.83)	0.55	1.84 (0.76-4.48)	0.18
rs1977389 (T>G)	1.36 (0.81-2.29)	0.24	1.57 (0.70-3.50)	0.27	1.40 (0.60-3.31)	0.44
rs10887878 (T>A)	1.21 (0.68-2.16)	0.52	1.14 (0.57-2.28)	0.71	2.14 (0.39-11.92)	0.38
rs4985700 (A>C)	1.00 (0.62-1.61)	0.99	1.17 (0.58-2.34)	0.67	0.75 (0.29-1.92)	0.55
DLBCL						
rs3832246 (T>-)	1.28 (0.59-2.78)	0.54	1.17 (0.47-2.93)	0.74	2.56 (0.34-19.36)	0.36
rs6773854 (T>C)	2.43 (0.93-6.36)	0.07	2.36 (0.85-6.53)	0.10	NA	NA
rs6457327 (C>A)	1.37 (0.70-2.68)	0.35	1.36 (0.56-3.32)	0.50	2.00 (0.46-8.68)	0.35
rs9271100 (C>T)	0.73 (0.36-1.47)	0.37	0.68 (0.27-1.69)	0.40	0.58 (0.11-3.18)	0.53
rs2647012 (C>T)	1.44 (0.66-3.12)	0.36	1.12 (0.45-2.75)	0.81	7.78 (1.00-60.71)	0.05
rs10484561 (T>G)	0.61 (0.19-2.04)	0.43	0.64 (0.18-2.31)	0.50	NA	NA
rs1571011 (A>C)	1.23 (0.62-2.44)	0.55	0.92 (0.34-2.53)	0.88	1.99 (0.65-6.12)	0.23
rs1977389 (T>G)	1.11 (0.57-2.18)	0.75	1.19 (0.43-3.31)	0.74	1.10 (0.36-3.40)	0.87
rs10887878 (T>A)	1.40 (0.65-3.04)	0.39	1.60 (0.65-3.91)	0.30	0.92 (0.10-8.52)	0.94
rs4985700 (A>C)	1.26 (0.68-2.32)	0.47	1.68 (0.65-4.29)	0.28	1.00 (0.31-3.20)	1.00
FL						
rs3832246 (T>-)	1.22 (0.46-3.27)	0.69	1.50 (0.49-4.60)	0.47	NA	NA
rs6773854 (T>C)	1.14 (0.32-4.00)	0.84	1.14 (0.32-4.00)	0.84	NA	NA
rs6457327 (C>A)	1.39 (0.61-3.16)	0.43	2.03 (0.69-5.97)	0.20	0.54 (0.05-5.24)	0.59
rs9271100 (C>T)	0.95 (0.43-2.10)	0.90	1.11 (0.39-3.12)	0.85	0.51 (0.06-4.50)	0.54
rs2647012 (C>T)	0.65 (0.23-1.82)	0.41	0.67 (0.23-1.97)	0.47	NA	NA
rs10484561 (T>G)	1.49 (0.52-4.28)	0.46	1.52 (0.41-5.56)	0.53	2.64 (0.13-53.26)	0.53
rs1571011 (A>C)	1.56 (0.67-3.66)	0.30	1.73 (0.45-6.70)	0.43	1.77 (0.47-6.68)	0.40
rs1977389 (T>G)	1.78 (0.79-4.03)	0.16	2.33 (0.58-9.34)	0.23	1.90 (0.55-6.62)	0.31
rs10887878 (T>A)	1.56 (0.68-3.57)	0.29	1.18 (0.41-3.40)	0.75	6.39 (0.93-44.07)	0.06
rs4985700 (A>C)	0.88 (0.42-1.82)	0.72	0.78 (0.27-2.23)	0.64	0.95 (0.23-3.89)	0.95
T-cell NHL[b]						
rs3832246 (T>-)	0.68 (0.20-2.27)	0.53	0.73 (0.2-2.71)	0.63	NA	NA
rs6773854 (T>C)	0.76 (0.15-3.81)	0.74	0.76 (0.15-3.81)	0.74	NA	NA
rs6457327 (C>A)	1.58 (0.62-4.02)	0.33	3.03 (0.84-10.87)	0.09	NA	NA
rs9271100 (C>T)	1.19 (0.51-2.78)	0.69	1.67 (0.49-5.71)	0.41	0.63 (0.07-5.77)	0.68
rs2647012 (C>T)	1.61 (0.54-4.78)	0.39	1.94 (0.57-6.57)	0.29	NA	NA
rs10484561 (T>G)	0.84 (0.19-3.78)	0.82	0.92 (0.18-4.71)	0.92	NA	NA
rs1571011 (A>C)	3.27 (1.08-9.94)	0.04	NA	NA	2.56 (0.60-10.79)	0.20
rs1977389 (T>G)	2.18 (0.81-5.88)	0.12	5.03 (0.59-42.60)	0.14	1.83 (0.43-7.78)	0.41
rs10887878 (T>A)	1.95 (0.70-5.43)	0.20	2.25 (0.66-7.65)	0.19	1.83 (0.12-27.18)	0.66
rs4985700 (A>C)	1.86 (0.79-4.39)	0.16	2.61 (0.65-10.51)	0.18	2.03 (0.45-9.15)	0.36

*Abbreviations: OR: odd ratio; CI: confidence interval; NHL: non-Hodgkin's lymphoma; DLBCL: diffuse large B-cell lymphoma; FL: follicular lymphoma; NA: not applicable.

[a]B-cell NHL: Include all the subtypes of B-cell NHL.
[b]T-cell NHL: Include all the subtypes of T-cell NHL.

4.6 Association between the DNA repair SNPs and NHL susceptibility

None of the 7 DNA repair SNPs examined in this study passed the Bonferroni correction threshold level at $P<0.003$. However, there are five potential DNA repair SNPs that showed weak association ($P<0.05$) with NHL susceptibility. *ERCC5* SNP rs17655 (G>C) showed weak association with subtype FL ($P_{\text{dominant-adjusted}}=0.03$; OR=0.60; 95% CI=0.37-0.95) in pooled subjects (Table 4.7); *XRCC1* SNP rs25487 (C>T) showed weak association with FL risk in Malay ethnic group ($P_{\text{additive-adjusted}}=0.03$; OR=1.87; 95% CI=1.05-3.32) and T-cell NHL ($P_{\text{additive-adjusted}}=0.008$; OR=0.41; 95% CI=0.21-0.79) in Malay group (Table 4.8) and *XRCC1* rs1799782 (G>A) showed weak association with FL risk ($P_{\text{dominant-adjusted}}=0.01$; OR=2.58; 95% CI=1.22-5.36) in Chinese group (Table 4.9). Both *ERCC2* SNPs in this study: rs13181 (T>G) and rs1799793 (C>T) showed weak association with T-cell NHL [rs rs13181 (T>G): $P_{\text{recessive-adjusted}}=0.03$, OR=4.81, 95% CI=1.13-20.49; rs1799793 (C>T): $P_{\text{recessive-adjusted}}=0.03$, OR=5.51, 95% CI=1.15-26.29] in Malay group.

Again, all weak association results ($P<0.05$) shall not take into consideration as this might happen due to chance, as *P*-values did not reach the level of significance ($P<0.003$) set for Bonferroni correction.

Table 4.7: Association study of SNPs in DNA repair genes and NHL risk examined in this study in pooled subjects.

	Additive Model (A vs a)		Dominant Model (AA + Aa vs aa)		Recessive Model (AA vs Aa + aa)	
	OR (95% CI)	***P*-value**	**OR (95% CI)**	***P*-value**	**OR (95% CI)**	***P*-value**
All NHL						
rs17655 (G>C)	1.02 (0.86-1.21)	0.82	1.00 (0.78-1.28)	1.00	1.07 (0.79-1.44)	0.68
rs861539 (G>A)	0.98 (0.73-1.33)	0.90	1.02 (0.74-1.41)	0.90	0.46 (0.12-1.82)	0.27
rs25487 (C>T)	0.98 (0.82-1.17)	0.81	0.94 (0.74-1.19)	0.60	1.08 (0.72-1.61)	0.72
rs25489 (C>T)	1.12 (0.85-1.48)	0.43	1.16 (0.86-1.55)	0.33	0.61 (0.14-2.59)	0.50
rs1799782 (G>A)	1.12 (0.93-1.34)	0.23	1.17 (0.92-1.48)	0.21	1.11 (0.74-1.67)	0.60
rs13181 (T>G)	1.19 (0.93-1.52)	0.17	1.22 (0.92-1.61)	0.16	1.19 (0.54-2.62)	0.67
rs1799793 (C>T)	1.21 (0.92-1.59)	0.18	1.19 (0.86-1.63)	0.29	1.88 (0.74-4.75)	0.18
B-cell NHL[a]						
rs17655 (G>C)	1.03 (0.87-1.23)	0.73	1.00 (0.77-1.30)	0.98	1.10 (0.80-1.51)	0.56
rs861539 (G>A)	1.04 (0.76-1.42)	0.83	1.09 (0.78-1.53)	0.60	0.39 (0.09-1.79)	0.23
rs25487 (C>T)	1.05 (0.87-1.27)	0.59	1.05 (0.82-1.34)	0.71	1.13 (0.74-1.71)	0.57
rs25489 (C>T)	1.13 (0.84-1.50)	0.42	1.16 (0.85-1.57)	0.35	0.71 (0.16-3.07)	0.64
rs1799782 (G>A)	1.08 (0.89-1.30)	0.43	1.12 (0.88-1.44)	0.36	1.05 (0.68-1.61)	0.83
rs13181 (T>G)	1.19 (0.92-1.54)	0.19	1.25 (0.93-1.67)	0.13	0.99 (0.42-2.29)	0.97
rs1799793 (C>T)	1.23 (0.92-1.65)	0.16	1.24 (0.89-1.73)	0.19	1.57 (0.60-4.15)	0.36
DLBCL						
rs17655 (G>C)	1.19 (0.97-1.46)	0.09	1.27 (0.93-1.72)	0.13	1.25 (0.88-1.78)	0.22
rs861539 (G>A)	1.15 (0.81-1.63)	0.43	1.22 (0.84-1.77)	0.30	0.52 (0.10-2.84)	0.45
rs25487 (C>T)	1.06 (0.86-1.31)	0.58	1.06 (0.80-1.39)	0.70	1.15 (0.72-1.85)	0.55
rs25489 (C>T)	1.18 (0.86-1.62)	0.31	1.20 (0.85-1.69)	0.31	1.19 (0.29-4.94)	0.81
rs1799782 (G>A)	1.07 (0.86-1.32)	0.56	1.08 (0.81-1.43)	0.61	1.12 (0.69-1.80)	0.65
rs13181 (T>G)	1.14 (0.85-1.54)	0.38	1.27 (0.91-1.76)	0.16	0.42 (0.11-1.53)	0.19
rs1799793 (C>T)	1.05 (0.74-1.48)	0.80	1.11 (0.75-1.62)	0.61	0.57 (0.14-2.36)	0.44
FL						
rs17655 (G>C)	**0.79 (0.56-1.11)**	**0.17**	**0.60 (0.37-0.95)**	**0.03**[c]	**1.07 (0.58-1.97)**	**0.83**
rs861539 (G>A)	0.51 (0.25-1.05)	0.07	0.45 (0.20-1.00)	0.05	0.72 (0.07-7.91)	0.79
rs25487 (C>T)	1.10 (0.78-1.57)	0.59	1.13 (0.71-1.78)	0.61	1.15 (0.52-2.54)	0.73
rs25489 (C>T)	0.95 (0.54-1.69)	0.87	0.99 (0.55-1.79)	0.97	NA	NA
rs1799782 (G>A)	1.13 (0.79-1.61)	0.50	1.33 (0.84-2.12)	0.23	0.76 (0.31-1.83)	0.54
rs13181 (T>G)	1.05 (0.64-1.70)	0.86	1.10 (0.63-1.93)	0.75	0.77 (0.15-3.96)	0.76
rs1799793 (C>T)	1.56 (0.95-2.54)	0.08	1.59 (0.89-2.83)	0.12	2.46 (0.60-10.11)	0.21
T-cell NHL[b]						
rs17655 (G>C)	0.92 (0.65-1.29)	0.62	0.94 (0.58-1.53)	0.79	0.82 (0.43-1.59)	0.56
rs861539 (G>A)	0.73 (0.37-1.42)	0.36	0.67 (0.32-1.40)	0.29	1.19 (0.14-10.50)	0.87
rs25487 (C>T)	**0.61 (0.41-0.91)**	**0.02**[c]	**0.50 (0.31-0.81)**	**0.005**[c]	**0.79 (0.32-1.91)**	**0.60**
rs25489 (C>T)	1.00 (0.57-1.76)	1.00	1.05 (0.58-1.89)	0.88	NA	NA
rs1799782 (G>A)	1.33 (0.94-1.87)	0.11	1.38 (0.86-2.21)	0.18	1.61 (0.79-3.27)	0.19
rs13181 (T>G)	1.13 (0.70-1.82)	0.61	1.08 (0.62-1.88)	0.79	1.81 (0.49-6.64)	0.37
rs1799793 (C>T)	1.02 (0.58-1.79)	0.95	0.85 (0.44-1.66)	0.64	3.00 (0.73-12.33)	0.13

*Abbreviations: OR: odd ratio; CI: confidence interval; NHL: non-Hodgkin's lymphoma; DLBCL: diffuse large B-cell lymphoma; FL: follicular lymphoma; NA: not applicable.

[a]B-cell NHL: Include all the subtypes of B-cell NHL.
[b]T-cell NHL: Include all the subtypes of T-cell NHL.
[c]*P*-value showed weak association ($P<0.05$).

Table 4.8: Association study of SNPs in DNA repair genes and NHL risk examined in this study in Malay group.

	Additive Model (A vs a)		Dominant Model (AA + Aa vs aa)		Recessive Model (AA vs Aa + aa)	
	OR (95% CI)	***P*-value**	**OR (95% CI)**	***P*-value**	**OR (95% CI)**	***P*-value**
All NHL						
rs17655 (G>C)	0.96 (0.73-1.26)	0.77	0.93 (0.62-1.38)	0.71	0.98 (0.60-1.60)	0.94
rs861539 (G>A)	1.01 (0.64-1.60)	0.97	1.09 (0.67-1.77)	0.73	0.17 (0.01-2.44)	0.19
rs25487 (C>T)	0.95 (0.72-1.25)	0.72	0.81 (0.56-1.18)	0.27	1.35 (0.74-2.45)	0.33
rs25489 (C>T)	1.03 (0.66-1.61)	0.90	1.07 (0.67-1.70)	0.79	0.24 (0.01-7.50)	0.41
rs1799782 (G>A)	1.04 (0.79-1.37)	0.79	1.05 (0.73-1.53)	0.78	1.04 (0.57-1.92)	0.89
rs13181 (T>G)	1.35 (0.92-1.96)	0.12	1.48 (0.96-2.27)	0.07	0.96 (0.27-3.41)	0.95
rs1799793 (C>T)	1.26 (0.82-1.94)	0.30	1.38 (0.85-2.24)	0.20	0.74 (0.15-3.71)	0.72
B-cell NHL[a]						
rs17655 (G>C)	0.96 (0.73-1.26)	0.77	0.93 (0.62-1.38)	0.71	0.98 (0.60-1.60)	0.94
rs861539 (G>A)	1.01 (0.64-1.60)	0.97	1.09 (0.67-1.77)	0.73	0.17 (0.01-2.44)	0.19
rs25487 (C>T)	0.95 (0.72-1.25)	0.72	0.81 (0.56-1.18)	0.27	1.35 (0.74-2.45)	0.33
rs25489 (C>T)	1.03 (0.66-1.61)	0.90	1.07 (0.67-1.70)	0.79	0.24 (0.01-7.50)	0.41
rs1799782 (G>A)	1.04 (0.79-1.37)	0.79	1.05 (0.73-1.53)	0.78	1.04 (0.57-1.92)	0.89
rs13181 (T>G)	1.35 (0.92-1.96)	0.12	1.48 (0.96-2.27)	0.07	0.96 (0.27-3.41)	0.95
rs1799793 (C>T)	1.26 (0.82-1.94)	0.30	1.38 (0.85-2.24)	0.20	0.74 (0.15-3.71)	0.72
DLBCL						
rs17655 (G>C)	1.11 (0.82-1.50)	0.49	1.13 (0.71-1.78)	0.61	1.18 (0.70-1.99)	0.54
rs861539 (G>A)	1.02 (0.62-1.69)	0.93	1.11 (0.65-1.89)	0.70	NA	1.00
rs25487 (C>T)	0.85 (0.62-1.17)	0.33	0.74 (0.49-1.12)	0.15	1.10 (0.55-2.20)	0.79
rs25489 (C>T)	1.18 (0.73-1.90)	0.51	1.22 (0.74-2.02)	0.43	0.37 (0.01-10.07)	0.55
rs1799782 (G>A)	1.06 (0.78-1.43)	0.71	1.02 (0.68-1.55)	0.91	1.23 (0.64-2.34)	0.54
rs13181 (T>G)	1.23 (0.80-1.88)	0.34	1.43 (0.89-2.29)	0.14	0.24 (0.03-2.15)	0.20
rs1799793 (C>T)	1.05 (0.64-1.72)	0.85	1.17 (0.68-2.02)	0.58	0.26 (0.02-2.67)	0.25
FL						
rs17655 (G>C)	0.68 (0.38-1.21)	0.19	0.55 (0.25-1.22)	0.14	0.73 (0.24-2.18)	0.57
rs861539 (G>A)	0.54 (0.18-1.58)	0.26	0.42 (0.11-1.61)	0.20	0.86 (0.04-20.16)	0.93
rs25487 (C>T)	1.87 (1.05-3.32)	0.03	1.86 (0.81-4.30)	0.15	3.16 (1.14-8.76)	0.03
rs25489 (C>T)	1.28 (0.50-3.24)	0.61	1.31 (0.50-3.42)	0.58	NA	1.00
rs1799782 (G>A)	0.52 (0.26-1.03)	0.06	0.55 (0.24-1.24)	0.15	0.15 (0.02-1.37)	0.09
rs13181 (T>G)	0.93 (0.42-2.07)	0.86	0.94 (0.36-2.41)	0.89	0.78 (0.06-9.48)	0.84
rs1799793 (C>T)	0.87 (0.35-2.20)	0.77	0.99 (0.37-2.7)	0.99	NA	1.00
T-cell NHL[b]						
rs17655 (G>C)	0.98 (0.58-1.65)	0.93	0.94 (0.44-2.00)	0.87	1.02 (0.40-2.64)	0.96
rs861539 (G>A)	0.44 (0.13-1.46)	0.18	0.44 (0.13-1.51)	0.19	NA	1.00
rs25487 (C>T)	**0.41 (0.21-0.79)**	**0.008**[c]	**0.36 (0.17-0.77)**	**0.008**[c]	**0.25 (0.03-1.95)**	**0.19**
rs25489 (C>T)	0.80 (0.30-2.12)	0.65	0.82 (0.30-2.24)	0.70	NA	1.00
rs1799782 (G>A)	1.27 (0.77-2.11)	0.35	1.58 (0.77-3.23)	0.21	1.01 (0.32-3.16)	0.99
rs13181 (T>G)	**1.51 (0.79-2.86)**	**0.21**	**1.28 (0.56-2.90)**	**0.56**	**4.81 (1.13-20.49)**	**0.03**[c]
rs1799793 (C>T)	**1.52 (0.77-3.01)**	**0.23**	**1.26 (0.51-3.11)**	**0.62**	**5.51 (1.15-26.29)**	**0.03**[c]

*Abbreviations: OR: odd ratio; CI: confidence interval; NHL: non-Hodgkin's lymphoma; DLBCL: diffuse large B-cell lymphoma; FL: follicular lymphoma; NA: not applicable.

[a]B-cell NHL: Include all the subtypes of B-cell NHL.
[b]T-cell NHL: Include all the subtypes of T-cell NHL.
[c]*P*-value showed weak association ($P<0.05$).

Table 4.9: Association study of SNPs in DNA repair genes and NHL risk examined in this study in Chinese group.

	Additive Model (A vs a)		Dominant Model (AA + Aa vs aa)		Recessive Model (AA vs Aa + aa)	
	OR (95% CI)	***P*-value**	**OR (95% CI)**	***P*-value**	**OR (95% CI)**	***P*-value**
All NHL						
rs17655 (G>C)	1.00 (0.78-1.29)	0.97	1.04 (0.70-1.53)	0.86	0.97 (0.64-1.49)	0.90
rs861539 (G>A)	1.08 (0.61-1.91)	0.80	1.05 (0.59-1.89)	0.86	NA	NA
rs25487 (C>T)	1.14 (0.86-1.52)	0.37	1.21 (0.84-1.72)	0.30	1.07 (0.53-2.13)	0.85
rs25489 (C>T)	1.11 (0.73-1.68)	0.64	1.09 (0.70-1.71)	0.70	1.6 (0.23-11.01)	0.63
rs1799782 (G>A)	1.09 (0.83-1.42)	0.55	1.14 (0.80-1.64)	0.46	1.03 (0.57-1.86)	0.93
rs13181 (T>G)	1.07 (0.67-1.71)	0.77	1.13 (0.69-1.86)	0.62	0.34 (0.03-4.12)	0.40
rs1799793 (C>T)	1.23 (0.68-2.24)	0.49	1.14 (0.60-2.16)	0.69	NA	NA
B-cell NHL[a]						
rs17655 (G>C)	1.03 (0.79-1.33)	0.85	1.05 (0.69-1.58)	0.83	1.02 (0.66-1.59)	0.93
rs861539 (G>A)	1.16 (0.65-2.09)	0.62	1.14 (0.62-2.08)	0.67	NA	NA
rs25487 (C>T)	1.22 (0.90-1.65)	0.19	1.36 (0.93-1.98)	0.11	1.01 (0.49-2.10)	0.97
rs25489 (C>T)	1.10 (0.71-1.70)	0.68	1.07 (0.67-1.71)	0.78	1.92 (0.27-13.49)	0.51
rs1799782 (G>A)	1.05 (0.78-1.40)	0.76	1.15 (0.79-1.67)	0.47	0.83 (0.44-1.59)	0.58
rs13181 (T>G)	1.06 (0.65-1.73)	0.81	1.12 (0.67-1.88)	0.67	0.38 (0.03-4.75)	0.45
rs1799793 (C>T)	1.31 (0.71-2.42)	0.39	1.21 (0.62-2.33)	0.58	NA	NA
DLBCL						
rs17655 (G>C)	1.11 (0.82-1.50)	0.51	1.23 (0.75-2.00)	0.41	1.07 (0.64-1.78)	0.80
rs861539 (G>A)	1.48 (0.78-2.80)	0.24	1.44 (0.74-2.79)	0.29	NA	NA
rs25487 (C>T)	1.35 (0.95-1.90)	0.09	1.53 (0.99-2.37)	0.06	1.18 (0.52-2.68)	0.69
rs25489 (C>T)	1.18 (0.73-1.92)	0.50	1.11 (0.65-1.89)	0.70	3.22 (0.49-21.26)	0.23
rs1799782 (G>A)	0.91 (0.65-1.28)	0.58	0.99 (0.64-1.53)	0.98	0.60 (0.26-1.39)	0.23
rs13181 (T>G)	1.29 (0.76-2.19)	0.35	1.37 (0.77-2.42)	0.28	0.73 (0.06-8.67)	0.80
rs1799793 (C>T)	1.22 (0.59-2.53)	0.59	1.12 (0.51-2.43)	0.78	NA	NA
FL						
rs17655 (G>C)	0.88 (0.55-1.41)	0.59	0.64 (0.32-1.31)	0.22	1.18 (0.53-2.62)	0.68
rs861539 (G>A)	0.37 (0.08-1.75)	0.21	0.37 (0.08-1.75)	0.21	NA	NA
rs25487 (C>T)	0.91 (0.51-1.62)	0.74	1.06 (0.53-2.12)	0.88	0.29 (0.03-2.47)	0.26
rs25489 (C>T)	0.55 (0.20-1.54)	0.26	0.56 (0.20-1.56)	0.26	NA	NA
rs1799782 (G>A)	**1.77 (1.06-2.95)**	**0.03**[c]	**2.55 (1.22-5.36)**	**0.01**[c]	**1.43 (0.51-4.04)**	**0.50**
rs13181 (T>G)	0.79 (0.30-2.06)	0.63	0.91 (0.32-2.57)	0.85	NA	NA
rs1799793 (C>T)	2.87 (1.16-7.15)	0.02	2.69 (0.99-7.26)	0.05	NA	NA
T-cell NHL[b]						
rs17655 (G>C)	0.79 (0.48-1.32)	0.37	0.76 (0.35-1.62)	0.47	0.70 (0.27-1.77)	0.45
rs861539 (G>A)	0.55 (0.12-2.43)	0.43	0.55 (0.12-2.43)	0.43	NA	NA
rs25487 (C>T)	0.76 (0.41-1.38)	0.36	0.61 (0.29-1.28)	0.19	1.25 (0.34-4.56)	0.74
rs25489 (C>T)	1.09 (0.47-2.51)	0.84	1.13 (0.48-2.67)	0.78	NA	NA
rs1799782 (G>A)	1.41 (0.84-2.36)	0.19	1.21 (0.59-2.49)	0.60	2.55 (0.99-6.59)	0.05
rs13181 (T>G)	1.09 (0.45-2.65)	0.85	1.20 (0.46-3.12)	0.71	NA	NA
rs1799793 (C>T)	0.70 (0.16-3.15)	0.64	0.70 (0.16-3.15)	0.64	NA	NA

*Abbreviations: OR: odd ratio; CI: confidence interval; NHL: non-Hodgkin's lymphoma; DLBCL: diffuse large B-cell lymphoma; FL: follicular lymphoma; NA: not applicable.

[a]B-cell NHL: Include all the subtypes of B-cell NHL.
[b]T-cell NHL: Include all the subtypes of T-cell NHL.

Table 4.10: Association study of SNPs in DNA repair genes and NHL risk examined in this study in Indian group.

	Additive Model (A vs a)		Dominant Model (AA + Aa vs aa)		Recessive Model (AA vs Aa + aa)	
	OR (95% CI)	***P*-value**	**OR (95% CI)**	***P*-value**	**OR (95% CI)**	***P*-value**
All NHL						
rs17655 (G>C)	0.98 (0.56-1.70)	0.94	0.81 (0.42-1.57)	0.53	2.56 (0.57-11.56)	0.22
rs861539 (G>A)	1.08 (0.57-2.01)	0.82	1.16 (0.56-2.39)	0.69	0.68 (0.09-4.87)	0.70
rs25487 (C>T)	0.98 (0.61-1.57)	0.92	1.03 (0.54-1.97)	0.94	0.84 (0.31-2.30)	0.74
rs25489 (C>T)	1.74 (0.79-3.80)	0.17	2.01 (0.89-4.56)	0.09	NA	NA
rs1799782 (G>A)	0.81 (0.39-1.69)	0.57	0.76 (0.34-1.70)	0.50	1.24 (0.09-16.22)	0.87
rs13181 (T>G)	1.14 (0.66-1.96)	0.65	1.11 (0.57-2.16)	0.76	1.41 (0.36-5.44)	0.62
rs1799793 (C>T)	1.26 (0.73-2.17)	0.40	1.20 (0.62-2.31)	0.59	2.06 (0.48-8.80)	0.33
B-cell NHL[a]						
rs17655 (G>C)	0.99 (0.55-1.78)	0.98	0.77 (0.39-1.56)	0.47	3.30 (0.71-15.37)	0.13
rs861539 (G>A)	1.01 (0.52-1.99)	0.97	1.11 (0.52-2.39)	0.78	0.42 (0.04-4.96)	0.49
rs25487 (C>T)	0.98 (0.60-1.62)	0.95	1.11 (0.56-2.20)	0.77	0.73 (0.25-2.13)	0.56
rs25489 (C>T)	1.74 (0.77-3.92)	0.18	2.04 (0.87-4.78)	0.10	NA	NA
rs1799782 (G>A)	0.74 (0.34-1.65)	0.46	0.66 (0.28-1.59)	0.36	1.58 (0.12-21.41)	0.73
rs13181 (T>G)	1.32 (0.74-2.37)	0.35	1.31 (0.64-2.70)	0.46	1.75 (0.45-6.84)	0.42
rs1799793 (C>T)	1.54 (0.87-2.74)	0.14	1.53 (0.76-3.08)	0.24	2.51 (0.58-10.79)	0.22
DLBCL						
rs17655 (G>C)	1.31 (0.62-2.77)	0.48	1.06 (0.43-2.60)	0.90	4.43 (0.71-27.51)	0.11
rs861539 (G>A)	1.27 (0.56-2.88)	0.57	1.41 (0.53-3.74)	0.49	0.97 (0.08-11.53)	0.98
rs25487 (C>T)	1.31 (0.70-2.43)	0.40	1.45 (0.58-3.57)	0.43	1.43 (0.42-4.89)	0.56
rs25489 (C>T)	1.18 (0.41-3.42)	0.76	1.37 (0.43-4.39)	0.59	NA	NA
rs1799782 (G>A)	0.81 (0.29-2.24)	0.68	0.63 (0.20-2.00)	0.43	3.77 (0.27-52.04)	0.32
rs13181 (T>G)	1.24 (0.56-2.74)	0.60	1.35 (0.53-3.44)	0.53	0.93 (0.1-9.03)	0.95
rs1799793 (C>T)	1.36 (0.63-2.94)	0.44	1.69 (0.68-4.16)	0.26	0.48 (0.04-5.5)	0.56
FL						
rs17655 (G>C)	0.55 (0.19-1.60)	0.27	0.40 (0.12-1.34)	0.14	2.28 (0.21-24.28)	0.49
rs861539 (G>A)	0.55 (0.15-1.96)	0.35	0.56 (0.14-2.20)	0.41	NA	NA
rs25487 (C>T)	0.66 (0.29-1.50)	0.32	0.66 (0.23-1.90)	0.45	0.37 (0.04-3.16)	0.36
rs25489 (C>T)	1.47 (0.49-4.45)	0.50	1.75 (0.52-5.88)	0.36	NA	NA
rs1799782 (G>A)	1.05 (0.35-3.10)	0.93	1.16 (0.36-3.78)	0.80	NA	NA
rs13181 (T>G)	1.36 (0.56-3.33)	0.50	1.45 (0.49-4.31)	0.50	1.37 (0.14-13.25)	0.79
rs1799793 (C>T)	1.68 (0.75-3.78)	0.21	1.49 (0.52-4.28)	0.46	4.22 (0.73-24.39)	0.11
T-cell NHL[b]						
rs17655 (G>C)	1.04 (0.35-3.12)	0.94	1.17 (0.35-3.91)	0.80	NA	NA
rs861539 (G>A)	1.44 (0.52-3.97)	0.48	1.41 (0.38-5.16)	0.60	2.60 (0.24-28.66)	0.44
rs25487 (C>T)	0.86 (0.35-2.10)	0.74	0.63 (0.18-2.19)	0.47	1.39 (0.26-7.43)	0.70
rs25489 (C>T)	1.39 (0.37-5.28)	0.63	1.64 (0.38-7.09)	0.51	NA	NA
rs1799782 (G>A)	1.11 (0.33-3.74)	0.87	1.25 (0.33-4.74)	0.75	NA	NA
rs13181 (T>G)	0.66 (0.21-2.02)	0.46	0.71 (0.21-2.41)	0.58	NA	NA
rs1799793 (C>T)	0.50 (0.15-1.61)	0.25	0.52 (0.15-1.85)	0.31	NA	NA

*Abbreviations: OR: odd ratio; CI: confidence interval; NHL: non-Hodgkin's lymphoma; DLBCL: diffuse large B-cell lymphoma; FL: follicular lymphoma; NA: not applicable.

[a]B-cell NHL: Include all the subtypes of B-cell NHL.
[b]T-cell NHL: Include all the subtypes of T-cell NHL.

4.7 *Cis*-eQTL and linkage disequilibrium analysis of NHL associated SNPs

As only HLA region SNPs showed significant association with NHL susceptibility in Malaysian population, potential regulatory function of 4 HLA SNPs was further evaluated through *cis*-eQTL analysis (Figure 4.2, Table 4.11). rs2647012 (C>T) showed significant *cis*-eQTL against *HLA-DQB1-AS1* ($P_{FDR\_cis\text{-}eQTL}$= 1.30x10^{-9}; Effect size=-0.87), *HLA-DQA2* ($P_{FDR\_cis\text{-}eQTL}$= 2.10x10^{-8}; Effect size=0.77), *HLA-DRB6* ($P_{FDR\_cis\text{-}eQTL}$=9.50x10^{-8}; Effect size=0.78) and *HLA-DQB1* ($P_{FDR\_cis\text{-}eQTL}$=1.10x10^{-7}; Effect size=-0.81) (Figure 4.2a-d). Minor allele rs2647012-T is correlated with increased expression for *HLA-DQB1-AS1* and *HLA-DQB1* (Figure 4.2a,d) and decreased expression for *HLA-DQA2* and *HLA-DRB6* (Figure 4.2b,c). rs10484561 (T>G) showed significant *cis*-eQTL with *HLA-DQB2* ($P_{FDR\_cis\text{-}eQTL}$=1.60x10^{-6}; Effect size=-0.93) (Figure 4.2e). Minor allele rs10484561-G is correlated with decreased expression for *HLA-DQB2*. *Cis*-eQTL of rs927110 in T-cell NHL was not evaluated as no suitable cell-type is available in GTEx. 4 HLA SNPs showed very weak pairwise LD in both Malay and Chinese groups (Table 4.12), and controls in pooled subjects (Figure 4.3) suggesting these variants to be independent markers.

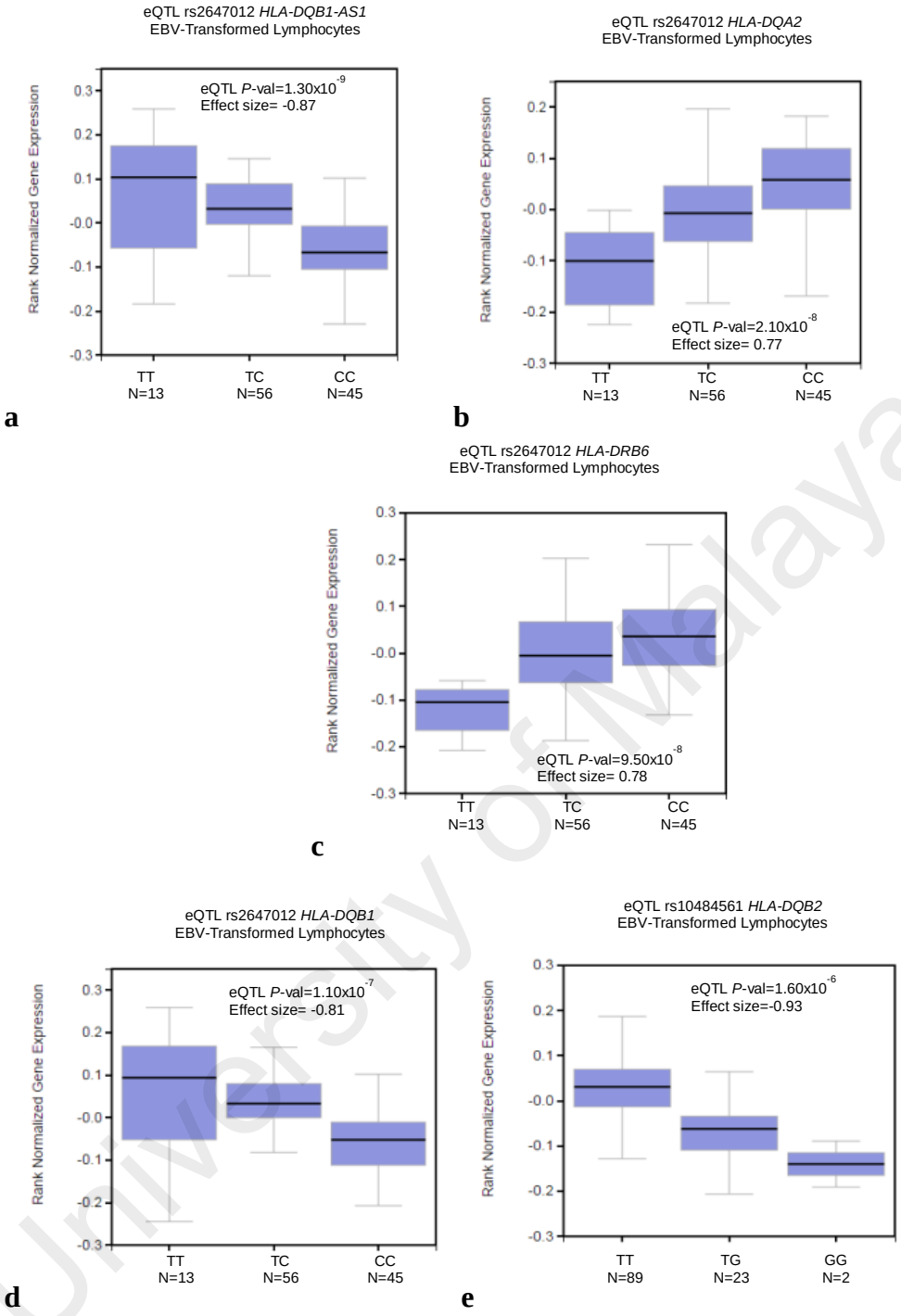


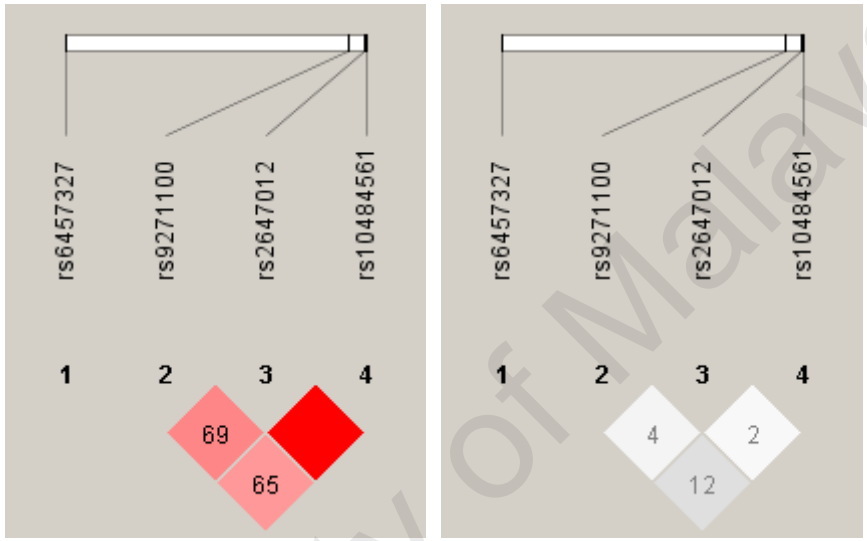
Figure 4.2: *Cis*-eQTL analysis of NHL associated SNPs in EBV-transformed lymphocytes. Correlation of **(a)** rs2647012 with *HLA-DQB1-AS1* **(b)** rs2647012 with *HLA-DQA2* **(c)** rs2647012 with *HLA-DRB6* **(d)** rs2647012 with *HLA-DQB1* **(e)** rs2647012 with *HLA-DQB2*.

Table 4.11: eQTL of 4 HLA SNPs in whole blood and EBV-transformed lymphocytes from GTEx.

Gencode Id	Gene Symbol	SNP ID	Synonym SNP	Allele	Ref/Alt allele[a]	*P*-value	Effect size[b]	Tissue
ENSG00000228789.2	*HCG22*	rs6457327	rs115316013	C>A	A/C	7.30E-15	-0.87	Cells - EBV-transformed lymphocytes
ENSG00000196126.6	*HLA-DRB1*	rs9271100	-	C>T	T/C	2.50E-15	-1	Cells - EBV-transformed lymphocytes
ENSG00000198502.5	*HLA-DRB5*	rs9271100	-	C>T	T/C	9.30E-13	-0.97	Cells - EBV-transformed lymphocytes
ENSG00000232629.4	*HLA-DQB2*	rs9271100	-	C>T	T/C	2.00E-12	0.97	Cells - EBV-transformed lymphocytes
ENSG00000237541.3	*HLA-DQA2*	rs9271100	-	C>T	T/C	1.00E-11	0.92	Cells - EBV-transformed lymphocytes
ENSG00000196735.7	*HLA-DQA1*	rs9271100	-	C>T	T/C	5.10E-11	-0.94	Cells - EBV-transformed lymphocytes
ENSG00000179344.12	*HLA-DQB1*	rs9271100	-	C>T	T/C	6.40E-09	-0.88	Cells - EBV-transformed lymphocytes
ENSG00000223534.1	*HLA-DQB1-AS1*	rs9271100	-	C>T	T/C	7.50E-09	-0.85	Cells - EBV-transformed lymphocytes
ENSG00000223534.1	*HLA-DQB1-AS1*	rs2647012	rs116393447	C>T	T/C	1.30E-09	-0.87	Cells - EBV-transformed lymphocytes
ENSG00000237541.3	*HLA-DQA2*	rs2647012	rs116393447	C>T	T/C	2.10E-08	0.77	Cells - EBV-transformed lymphocytes
ENSG00000229391.3	*HLA-DRB6*	rs2647012	rs116393447	C>T	T/C	9.50E-08	0.78	Cells - EBV-transformed lymphocytes
ENSG00000179344.12	*HLA-DQB1*	rs2647012	rs116393447	C>T	T/C	1.10E-07	-0.81	Cells - EBV-transformed lymphocytes
ENSG00000232629.4	*HLA-DQB2*	rs10484561	rs116212130	T>G	T/G	0.0000016	-0.93	Cells - EBV-transformed lymphocytes

*Abbreviations: eQTL, expression quantitative trait loci; GTEx, Genotype-Tissue expression; SNP, single nucleotide polymorphism; HLA, Human leukocyte antigen; Ref, Reference; Alt, alternative.

[a]Reference and alternative allele of the corresponding SNP
[b]Effect size calculated based on alternative allele of SNP



(a) **(b)**

Figure 4.3: Linkage disequilibrium (LD) block of 4 HLA SNPs in controls in pooled subjects. Values of D' and r^2 in percentage (%) are represented in (a) and (b), respectively. 4 HLA SNPs showed very weak pairwise LD in controls in pooled subjects.

Table 4.12 Linkage disequilibrium of 4 HLA SNPs in all NHL types for Chinese (A) and Malay (B) groups.

A: Chinese

	rs6457327	rs9271100	rs2647012	rs10484561
rs6457327		0.030	0.001	0.010
rs9271100	0.260		0.012	0.030
rs2647012	0.091	0.434		0.007
rs10484561	0.741	0.442	0.847	

B: Malay

	rs6457327	rs9271100	rs2647012	rs10484561
rs6457327		0.003	0.027	0.003
rs9271100	0.066		0.052	0.218
rs2647012	0.490	0.791		0.026
rs10484561	0.147	0.748	0.902	

Pearson's correlation, r^2

Lewontin's D-prime, D'

4.8 Association between the SNPs and DLBCL prognosis

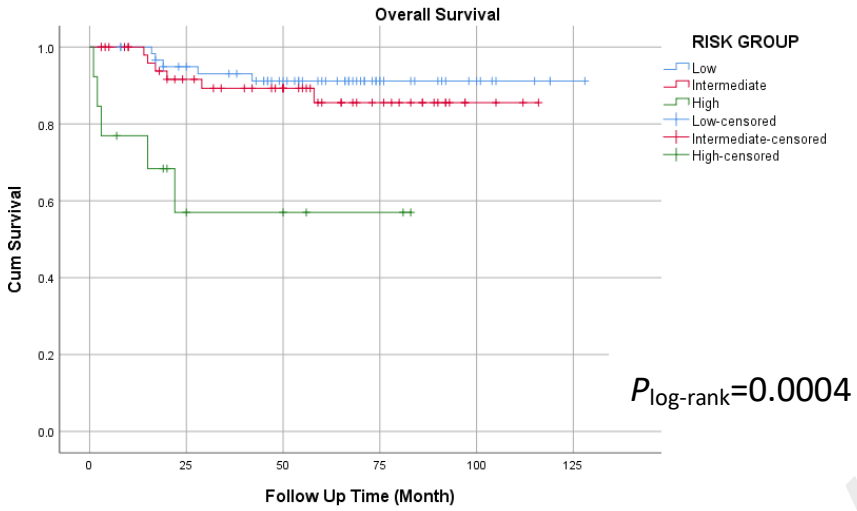
Of 334 DLBCL patients recruited in this study, 130 patients (Malay= 80; Chinese=50) with complete clinical information and meticulous follow-up record were included in the analysis. Indian group was excluded from this analysis as the number of DLBCL patients with complete clinical information was too low (n=20). Clinical characteristics of DLBCL patients are summarized in the Table 4.13. DLBCL patients were categorised into three risk groups: low, intermediate and high-risk group according to IPI score. 47.7% (n=62) patients were categorised into low risk (IPI score 0 or 1), 42.3% (n=55) intermediate-risk (IPI score 2 or 3) and 10% (n=13) high risk (IPI score of 4 or 5). With a median follow up of 52 months (range 1 to 112 months), 14 patients progressed or relapsed, and 16 patients died.

It is well established that IPI predicts the prognosis of DLBCL patients. In the present study, the classification of DLBCL patients according to three risk groups based on IPI score showed significant log-rank *P*-value in both OS ($P_{\text{log-rank}}$ =0.0004) and PFS ($P_{\text{log-rank}}$ =0.007) (Figure 4.4). Five-year OS rate for low-risk group is 90.9%, intermediate-risk group is 84.1% and high-risk group is 60.6% and the 5-year PFS rate for low-risk group is 79.4%, intermediate-risk group is 70.2% and high-risk group is 53.9%. However, univariate Cox regression analysis of 17 SNPs showed no association with survival outcomes of DLBCL patients in both OS and PFS (Tables 4.14 & 4.15).

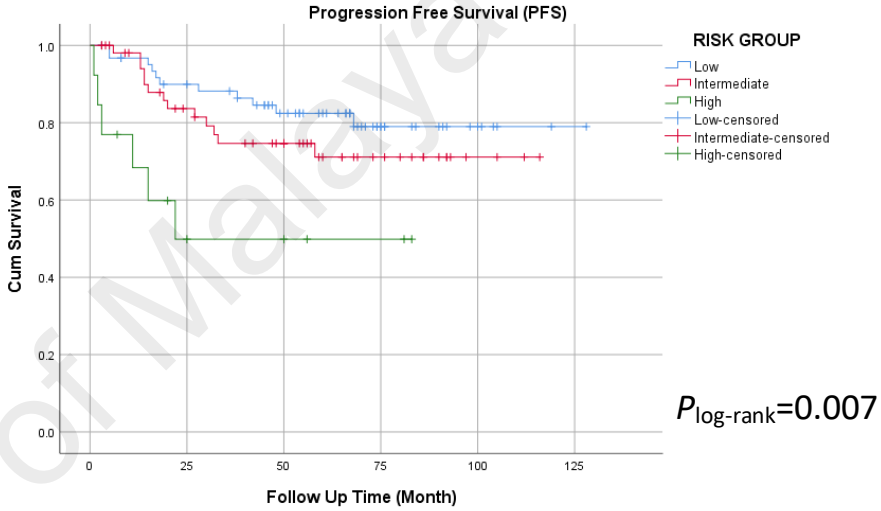
Table 4.13: Clinical characteristics of DLBCL patients.

	n	%
Race		
Malay	80	61.5
Chinese	50	38.5
Sex		
Male	59	45.4
Female	71	54.6
Age at diagnosed		
≤60	93	71.5
>60	37	28.5
Ann Arbor classification stage		
I or II	72	55.4
III or IV	57	43.8
NA	1	0.8
ECOG Performance Score		
<2	104	80
≥2	25	19.2
NA	1	0.8
Serum LDH level		
Normal	52	40
High	61	46.9
NA	17	13.1
Extranodal involvement		
≤1	96	73.8
>1	30	23.1
NA	4	3.1
Risk Group (IPI score)		
Low (0 & 1)	62	47.7
Intermediate (2 & 3)	55	42.3
High (4 & 5)	13	10
Total	130	100

*Abbreviations: ECOG, Eastern Cooperative Oncology Group; IPI, International Prognostic Index; LDH, Lactate dehydrogenase; NA, data not available.



(A)



(B)

Figure 4.4: Kaplain-Meier (K-P) curves for overall survival (OS) and progression-free survival (PFS) according to three risk groups of DLBCL patients. The classification of DLBCL patients according to three risk groups based on IPI score showed significant log-rank *P*-value in both OS ($P_{log\text{-}rank}$ =0.0004) and PFS ($P_{log\text{-}rank}$ =0.007).

*(A) OS is defined as an interval from the date of diagnosis until the date of death from any cause or until the last follow-up date. *(B) PFS is defined as an interval measured from the date of diagnosis until the date of progression, relapse or death from any cause or until the last follow-up date.

Table 4.14: Cox regression analysis of immune function SNPs and their association with overall survival (OS) and progression-free survival (PFS) of DLBCL patients.

		Overall Survival (OS)			Progression-Free Survival (PFS)		
	n	**Event[a]**	**HR (95% CI)**	***P*-value**	**Event[b]**	**HR (95% CI)**	***P*-value**
rs3832246 (T>-)							
T T	87	11	1.00 (reference)		21	1.00 (reference)	
- - + T -	43	5	0.68 (0.22-2.08)	0.50	9	0.83 (0.38-1.81)	0.64
rs6773854 (T>C)							
T T	71	12	1.00 (reference)		19	1.00 (reference)	
C C + C T	59	4	0.34 (0.11-1.05)	0.06	11	0.66 (0.31-1.38)	0.27
rs6457327 (C>A)							
C C	52	8	1.00 (reference)		15	1.00 (reference)	
A A + A C	78	8	0.59 (0.22-1.59)	0.30	15	0.59 (0.29-1.21)	0.15
rs9271100 (C>T)							
C C	67	6	1.00 (reference)		15	1.00 (reference)	
T T + C T	62	10	0.56 (0.20-1.56)	0.27	15	1.02 (0.50-2.08)	0.97
NA	1						
rs2647012 (C>T)							
C C	94	12	1.00 (reference)		23	1.00 (reference)	
T T + C T	36	4	0.56 (0.20-1.56)	0.27	7	0.81 (0.35-1.89)	0.62
rs10484561 (T>G)							
T T	98	12	1.00 (reference)		19	1.00 (reference)	
G G + G T	32	4	1.72 (0.50-5.93)	0.39	11	1.98 (0.94-4.17)	0.07
rs1571011 (A>C)							
A A	41	5	1.00 (reference)		9	1.00 (reference)	
C C + C A	89	11	0.83 (0.29-2.43)	0.74	21	1.02 (0.47-2.22)	0.96
rs1977389 (T>G)							
T T	28	4	1.00 (reference)		7	1.00 (reference)	
G G + T G	102	12	0.69 (0.21-2.10)	0.48	23	0.81 (0.35-1.88)	0.62
rs10887878 (T>A)							
T T	38	5	1.00 (reference)		12	1.00 (reference)	
A A + A T	92	11	0.92 (0.32-2.65)	0.87	18	0.62 (0.30-1.28)	0.19
rs4985700 (A>C)							
AA	68	9	1.00 (reference)		18	1.00 (reference)	
C C + C A	62	7	0.92 (0.34-2.52)	0.87	12	0.62 (0.30-1.28)	0.19
Total	130	16			30		

*Abbreviations: DLBCL, diffuse large B-cell lymphoma; HR, hazard ratio; CI, confident interval; NA, data not available.

[a] No of patient died during the follow-up period.
[b] No of patient progressed/relapsed/died during the follow-up period.

Table 4.15: Cox regression analysis of DNA repair SNPs and their association with overall survival (OS) and progression-free survival (PFS) of DLBCL patients.

		Overall Survival (OS)			Progression-Free Survival (PFS)		
	n	**Event[a]**	**HR (95% CI)**	***P*-value**	**Event[b]**	**HR (95% CI)**	***P*-value**
rs17655 (G>C)							
G G	38	6	1.00 (reference)		9	1.00 (reference)	
C C + C G	92	10	0.82 (0.28-2.36)	0.71	21	0.99 (0.46-2.17)	0.99
rs861539 (G>A)							
G G	113	12	1.00 (reference)		23	1.00 (reference)	
A A + A G	17	4	2.08 (0.67-6.49)	0.21	7	1.90 (0.81-4.42)	0.14
rs25487 (C>T)							
C C	64	8	1.00 (reference)		15	1.00 (reference)	
T T + C T	66	8	1.42 (0.48-4.24)	0.53	15	1.06 (0.52-2.17)	0.87
rs25489 (C>T)							
C C	104	10	1.00 (reference)		23	1.00 (reference)	
T T + T C	26	6	2.22 (0.80-6.21)	0.13	7	1.23 (0.53-2.87)	0.63
rs1799782 (G>A)							
G G	60	6	1.00 (reference)		14	1.00 (reference)	
A A + A G	70	10	1.21 (0.44-3.38)	0.71	16	0.96 (0.47-1.97)	0.92
rs13181 (T>G)							
T T	101	11	1.00 (reference)		21	1.00 (reference)	
G G + G T	29	5	1.92 (0.66-5.56)	0.23	9	1.63 (0.75-3.56)	0.22
rs1799793 (C>T)							
C C	108	13	1.00 (reference)		27	1.00 (reference)	
T T + C T	22	3	1.27 (0.36-4.47)	0.72	3	0.53 (1.61-1.75)	0.30
Total	130	16			30		

*Abbreviations: DLBCL, diffuse large B-cell lymphoma; HR, hazard ratio; CI, confident interval; NA, data not available.

[a] No of patient died during the follow-up period.
[b] No of patient progressed/relapsed/died during the follow-up period.

CHAPTER 5: DISCUSSION

5.1 Association between the immune function SNPs and NHL susceptibility

Out of ten immune function SNPs that examined in this study, two variants were identified: rs2647012 (C>T) and rs10484561(T>G) in HLA class II region which significantly (P<0.003) associated with NHL susceptibility in Malaysian population. SNP rs2647012 (C>T) was significantly associated with reduced risk when analysed in all NHL cases in Malay group and B-cell NHL in both pooled subjects and Malay group. SNP rs10484561 (T>G) showed significant association with subtype FL in pooled subjects and Malay group. This study further supports the findings from GWASs study that stated genetic variation in HLA class II region influenced NHL susceptibility (Conde *et al*., 2010; Skibola *et al*., 2009; Smedby *et al*., 2011). The results of this study provide supporting evidence that at least some of the genetic factors that associated with NHL risk in European ancestry could extend to the Asian population.

Interestingly, the SNP rs2647012-T allele reduced the risk for all NHL and B-cell NHL development while SNP rs10484561-G allele increased the risk of FL subtypes. This finding was supported by the study from Vijai *et al*. (2013) stated that the locus 6p21.3 act as a susceptibility locus for NHL, having both protective and risk alleles in different types and subtypes of NHL. SNP rs2647012 (C>T) is in complete LD (D'=1) with the extended HLA haplotypes DRB1*01:01- DRB1*15-DQA1*01-DQB1*06 and rs10484561(T>G) with DQA1*01:01-DQB1*05:01, respectively (Skibola *et al*., 2012). rs10484561-G allele has been associated with FL (Foo *et al*., 2013; Skibola *et al*., 2009), although it had also been implicated in DLBCL (Conde *et al*., 2010; Skibola *et al*., 2009; Smedby *et al*., 2011). This immune gene-rich region may influence NHL risk through several modes of action including effects on $CD4^+$ T-cell activation, antigen presentation of infectious or tumour-associated peptides, and HLA protein/gene expression. Further

functional studies will be needed to clarify the role of polymorphisms in HLA class II and lymphomagenesis.

This study has successfully shown that rs2647012 (C>T) and rs10484561 (T>G) act as a susceptibility locus of NHL having a potential regulatory function as both shows significant *cis*-eQTL correlation with nearby HLA class II molecule expression (Figure 4.2). The association of rs2647012 (C>T) and rs10484561 (T>G) were corroborated with RNA expression profiles from EBV-transformed lymphocytes of the GTEx database (Figure 4.2). NHL-associated SNPs against HLA class II molecules expression such as HLA-DQB1, HLA-DQB2, HLA-DQA2, HLA-DRB6 and non-coding molecule HLA-DQB1-AS1. rs2647012 confers a protective effect towards all NHL in Malay group and B-cell NHL in pooled subjects and Malay group by having a higher occurrence of the minor allele rs2647012-T in controls (Table 4.2). This is corroborated by the increased expression of HLA-DQB1 with rs2647012-T (Figure 4.2d) which was also proven by Sillé *et al*. (2014). However, rs2647012-T is correlated with lower expression of HLA-DQA2, contradicting the protective effect rendered by rs2647012-T (Figure 4.2b). In this study, rs10484561 (T>G) confers a susceptible effect towards FL subtypes with a higher occurrence of the minor allele rs10484561-G in pooled subjects and Malay group (Table 4.4). This is corroborated by a lower expression of HLA-DQB2 with rs10484561-G (Figure 4.2e). Presence of *cis*-eQTL in the EBV-transformed lymphocytes is even more relevant given that HLA class II molecules are expressed in antigen presenting cells such as B-lymphocytes, and act to present exogenous antigens to CD4+ helper T-cells. Changes in expression of these HLA class II molecules, modulated by SNPs, may compromise the efficiency of antigen presentation.

Another two studied SNPs at HLA region: rs6457327 (C>A) and rs9271100 (C>T). SNP rs6457327 (C>A), located at HLA class I region near *HLA-C* had been reported to inversely associate with risk of FL (*P*-value = 4.7×10^{-11}) (Skibola *et al*.,

2009). In the study by Wrench *et al*., (2011), it was suggested that the SNP rs6457327 (C>A) is a predictive marker for the transformation of FL to DLBCL. However, this study did not observe any association between SNP rs6457327 (C>A) with NHL risk in the Malaysian population. SNP rs9271100 (C>T) was described associated with SLE in a Chinese Han population (Han *et al*., 2009). As an autoimmune disease was often linked with NHL risk, SNP rs9271100 (C>T) was also included in this study.

Unpublished data from our group suggested that germline polymorphism in the promoter region of *BCL6*, rs3832246 (T>-), might potentially associate with NHL risk in Chinese group. The finding was further examined in larger sample size and the results showed that the deletion of one or both T alleles of the rs3832246 act as a protective role in overall NHL, B-cell NHL, and DLBCL in Chinese group (Table 4.5). The deletion of rs3832246-T allele might reduce the transcriptional activity of its co-repressor rendering its insensitive to inhibitory signalling events mediated by translocations and somatic hypermethylation, thus prevent the structural alteration of its promoter and subsequently prevent the deregulation of *BCL6* expression. Other investigators demonstrated the potential functional impact of rs3832246-T allele deletion by altering nuclear factor binding affinity using gel-shift assay (Jardin *et al*., 2003). These results suggested the potential influence of rs3832246 (T>-) in modulating *BCL6* or adjacent *LPP* expression.

SNP rs6773854 (T>C) located at chromosome 3q27, is an intergenic variant between *BCL6* and *LPP* gene. In the GWAS findings by Tan *et al*. (2013), rs6773854 reached genome-wide significance (P = 3.36 x 10^{-13}, OR= 1.44) and was found to be associated with DLBCL risk in Chinese population. In this study, rs6773854 (T>C) was observed in weak association with B-cell NHL and DLBCL risk in pooled subjects (Table 4.3) and DLBCL risk in Chinese group (Table 4.5).

In this study, we have examined three *FAS* variants: rs1571011 (A>C), rs1977389 (T>G), rs10887878 (T>A) with NHL risk in the Malaysian population. However, no

association of *FAS* variants was detected. It had been reported that two functional SNPs in the promoter of *FAS* gene: rs2234767 (−1377 G>A) and rs1800682 (−670 A>G) to influence transcriptional activities of the genes and thus causes multiple tumours susceptibility (Chen *et al*., 2015; Landowski *et al*., 1997; Zhang *et al*., 2006). Therefore, these two SNPs should be examined in the future study. *FAS* and FAS ligand (FASL) play roles in apoptotic signaling and down-regulation of this pathway may facilitate tumourigenesis. Germline mutations in the *FAS* gene have been associated with the autoimmune lymphoproliferative syndrome (ALPS), and somatic *FAS* mutations have been found in multiple myeloma (Landowski *et al*., 1997). Genetic polymorphisms in the *FAS* gene promoter may affect *FAS* gene expression and modulate apoptotic signaling, contributing to an increased risk of acute myeloid leukemia (AML) (Sibley *et al*., 2003), although the latters meta-analysis did not find any association between *FAS* gene polymorphism with leukemia risk (Chen *et al*., 2015). A study had shown that FAS-deficient mice develop elevated levels of serum autoantibodies and B cell plasmacytoid lymphomas. Patients with autoimmune lymphoproliferative syndrome or ALPS have *FAS* mutations and are at high risk for NHL.

rs4985700 (A>C), the only *TNFRSF13B* SNP examined in this study, showed weak association with T-cell NHL in Malay group (Table 4.4). Homozygous and heterozygous mutations in *TNFRSF13B*, encoding transmembrane activator and CAML interactor (TACI) are associated with common variable immunodeficiency in humans (Salzer *et al*., 2005).

5.2 Association between the DNA repair SNPs and NHL susceptibility

A total of seven SNPs in four key genes involved in four out of five DNA repair pathways were examined in this study, which enabled a broad understanding of DNA repair mechanisms involved in the pathogenesis of NHL. Although no significant association was detected after Bonferroni correction, five DNA repair SNPs: *ERCC5* rs17655 (G>C), *XRCC1* rs25487 (C>T), *XRCC1* rs1799782 (G>A), *ERCC5* rs13181 (T>G), *ERCC5* rs1799793 (C>T) showed weak association (P<0.05) with NHL susceptibility in Malaysian population.

ERCC5 SNP rs17655 (G>C) play roles in the NER pathway, and involved in excision repair of UV-induced DNA damage (Zhu *et al*., 2012). Functional effects using the computational algorithms by SIFT (Kumar *et al*., 2009) software indicated that the polymorphism in rs17655 (G>C) was likely to be "delicious" and PolyPhen-2 software (Adzhubei *et al*., 2010) indicated "probably damaging". However, such potentially functional relevance has not been validated experimentally to date. Zhu *et al*, (2012) found a null association between *ERCC5* polymorphisms with mRNA expression levels in the lymphoblastoid cell lines.

In this study, two highly reported SNPs: rs13181 (T>G) and rs1799793 (C>T) located in the *ERCC2* gene in the NER pathway were also examined. Both rs13181 (T>G) and rs1799793 (C>T) were predicted as "tolerated" by SIFT (Kumar *et al*., 2009) and "benign" by PolyPhen-2 software (Adzhubei *et al*., 2010). This mean that amino acid change caused by SNP variant can be tolerated in the protein if these positions are not involved in protein function or structure ((Kumar *et al*., 2009). In this study, both *ERCC2* SNPs rs13181 (T>G) and rs1799793 (C>T) showed a weak association with T-cell NHL in Malay group (Table 4.8). However, a meta-analysis conducted by Zhou *et al*. (2014) suggested no association between rs13181 (T>G) and rs1799793 (C>T) polymorphism and the risk of NHL.

In the present study, three highly reported SNPs in *XRCC1* which involved in BER pathway: rs25487(C>T), rs25489(C>T) and rs1799782(G>A) were examined. However, only SNP rs25487 (C>T) showed a weak association with FL risk in Malay group and T-cell NHL in Malay group and rs1799782 (G>A) showed a weak association with FL risk in Chinese group (Table 4.9). No association of rs25489(C>T) with NHL risk was observed. *XRCC1* rs25487(C>T) was predicted as "benign" by Polyphen-2 and "tolerated" using SIFT program while *XRCC1* rs1799782 (G>A) was predicted as "possibly damaging" by Polphen-2 and "tolerated" by SIFT program . A recent systematic meta-analysis conducted by Li *et al*. (2016) suggested that *XRCC1* SNP rs25487 (C>T) was associated with decreased risk for NHL and DLBCL; SNP rs1799782 (G>A) was associated with increased NHL risk and no association was detected for SNP rs25489 (C>T). The involvement of *XRCC1* and other BER genes in the processing of immunoglobulin (Ig) rearrangement intermediate during somatic hypermutation and class-switch recombination argued that BER genes could participate in early events in lymphomagenesis (Akbari *et al*., 2004).

Another SNP examined in the present study, *XRCC3* rs861539(G>A) involved in DSBR pathway. Smedby *et al*. (2006) suggested that *XRCC3* polymorphisms may contribute to FL risk among cigarette smokers. However, *XRCC3* rs861539(G>A) showed no significant association with NHL risk in Malaysian population.

The present study suggested that inherited variants in BER and NER pathway might play a role in lymphomagenesis. There are five potential DNA repair SNPs that showed weak association (P<0.05) with NHL susceptibility in Malaysian population. However, as far as our knowledge is concerned, none of the SNPs in DNA repair pathway have been identified as susceptibility locus in the published GWAS for NHL risk. Further functional studies should be undertaken to explore the mechanism underlying the NER and BER pathway and NHL development.

5.3 Association of SNPs with prognosis of DLBCL patients

IPI (Shipp *et al*., 1993) and revised IPI (Sehn *et al*., 2007) score have been used many years by clinicians to predict the DLBCL patients' clinical outcomes. The formulation of the IPI incorporated patients age at diagnosis, LDH level, performance status, clinical stage, and number of extranodal sites involved has provided generally accepted criteria to identify specific risk groups of DLBCL and to design appropriate therapies. However, due to heterogeneity of this subtypes, effective risk adapted strategies are needed to improve the outcome of patients with DLBCL. New biological markers such as genetic markers that reflect the heterogeneity of DLBCL must be evaluated to better determine the patient outcomes.

In this study, a retrospective analysis was done to evaluate the association of 17 SNPs with prognosis of DLBCL patients. The classification of DLBCL patients according to three risk group based on IPI score showed significant log-rank *P*-values in both OS ($P_{\text{log-rank}}$ =0.0004) and PFS ($P_{\text{log-rank}}$ =0.007), which supported the important role of IPI score in determination of the DLBCL patients outcomes. No significant association (P<0.05) was observed between SNPs with DLBCL patient clinical outcomes due to the small samples size of DLBCL patients with complete clinical information included in the study.

The attempt to predict outcome of DLBCL into prognostically favorable GCB and unfavorable ABC subtype based on gene expression signatures had resulted into inconsistent results (Shipp *et al*., 2002). There are many attempts to identify germline genetic variation and DLBCL patients' survival outcomes (Juszczynski *et al*., 2002; Lech-Maranda *et al*., 2004; Warzocha *et al*., 1998), however none of the findings gave a confidence results as their sample size was too small. More future studies like gene expression profiling are needed to identify the genetic markers that play role in DLBCL prognosis for future therapeutic intervention.

5.4 Strengths and limitations

The strength of this study was the results obtained from the population-based design were further supported by the *cis*-eQTL analysis. These findings are consistent and in the same trend as compared to Caucasians and thus unlikely to be biased. The potential for false positive associations still cannot fully exclude and therefore require examination in other independent and larger populations.

As many other subtypes of disease have insufficient cases (n<100) and hence it was not possible to assess risk among those smaller groups due to the limited power of study. Hence, this study can only stratify the analysis according to two major subtypes of B-cell NHL (DLBCL and FL). Some studies even suggested that further classifying DLBCL according to cell-of-origin between subtypes GCB and ABC can clearly clarify association signal of SNPs with DLBCL risk (Lossos, 2005). However, this study does not have comprehensive ABC and GCB classifications for DLBCL samples and are unable to address this question.

The Malay group in the present study showed strong and consistent association with SNP rs2647012 (C>T) and rs10484561 (T>G) and NHL risk; Chinese ethnic group showed the same direction of effect as observed in Malay group although no significant association detected. However, no association in Indians ethnic groups noted due to the very small samples size. Given that this study included three different ethnic groups in Malaysia, the presence of ethnicity dispersity in pooled results is less to influence the result because meta-analysis was used to analyse the results. Furthermore, the MAFs for the SNPs were generally similar across the three ethnic groups.

This study reported the association of HLA class II SNPs towards NHL in three major ethnicities in Malaysia, namely Malay, Chinese and Indian group. This is the first report with largest sample size to date to represent the NHL association in the Malaysian population after previous report (Lim *et al*., 2014). The inability to detect association in

the Chinese and Indian group could be due to limited sample size. Hence, future replication in other Asian population should be carried out to support this finding.

CHAPTER 6: SUMMARY AND FUTURE STUDIES

This study has successfully examined the association of 17 SNPs in various NHL risk in Malaysian population. This study has also generated the genetic data including allele and genotype frequencies of 17 SNPs for NHL patients and healthy controls in Malaysian population. In summary, this data further supported that two SNPs: rs2647012 (C>T) & rs10484561 (T>G) in HLA class II region may significantly contribute to NHL susceptibility in Malaysian population. The present study also demonstrated that rs2647012 (C>T) and rs10484561 (T>G) show potential regulatory function as both show significant *cis*-eQTL correlation with nearby HLA class II molecule expression.

This results further supported the need for in-depth analyses of genetic variants in the 6p21.3 chromosomal region. Additional studies would be needed to examine the results in this study with larger sample size or in other Asian population to confirm the findings. Besides, evaluation by meta-analysis using this results and other studies may need to perform to increase the statistical power and overcome the limitations of individual studies. Once the genetic association was confirmed, further functional studies can be carried out to investigate the role of 6p21.3 region in lymphomagenesis. Thus, this will provide a better understanding to the pathogenesis of NHL and lead to prevention, improvement of diagnosis and treatment in the future.

In view of this, future work should focus on fine mapping the HLA class II region as well genotyping the HLA class II alleles to identify potential causal variants. Future analysis is needed to include immune-related pathways of HLA class II molecules as well as the regulatory function of SNPs in modulating binding of transcription factors in promoter regions Functional assays should be carried out in future to investigate the role of these SNPs in pathogenesis of NHL.

REFERENCES

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., … Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nature Methods, 7*(4), 248–249.

Akbari, M., Otterlei, M., Peña-Diaz, J., Aas, P. A., Kavli, B., Liabakk, N. B., … Krokan, H. E. (2004). Repair of U/G and U/A in DNA by UNG2-associated repair complexes takes place predominantly by short-patch repair both in proliferating and growth-arrested cells. *Nucleic Acids Research, 32*(18), 5486–5498.

Alizadeh, A. A., Eisen, M. B., Davis, R. E., Ma, C., Lossos, I. S., Rosenwald, A., … Staudt, L. M. (2000). Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature, 403*(6769), 503–511.

Barrett, J. C., Fry, B., Maller, J., & Daly, M. J. (2005). Haploview: Analysis and visualization of LD and haplotype maps. *Bioinformatics, 21*(2), 263–265.

Beral, V., Peterman, T., Berkelman, R., & Jaffe, H. (1991). AIDS-associated non-Hodgkin lymphoma. *The Lancet, 337*(8745), 805–809.

Bettinardi, A., Brugnoni, D., Quiròs-Roldan, E., Malagoli, A., Grutta, S. L., Correra, A., & Notarangelo, L. D. (1997). Missense mutations in the Fas gene resulting in autoimmune lymphoproliferative syndrome: A molecular and immunological analysis. *Blood*, 89(3), 902-909.

Blair, A., Hartge, P., Stewart, P. A., McAdams, M., & Lubin, J. (1998). Mortality and cancer incidence of aircraft maintenance workers exposed to trichloroethylene and other organic solvents and chemicals: Extended follow up. *Occupational and Environmental Medicine, 55*(3), 161–171.

Boccolini, P. M. M., Boccolini, C. S., Chrisman, J. R., Markowitz, S. B., Koifman, S., Koifman, R. J., & Meyer, A. (2013). Pesticide use and non-Hodgkin's lymphoma mortality in Brazil. *International Journal of Hygiene and Environmental Health, 216*(4), 461–466.

Cattoretti, G., Pasqualucci, L., Ballon, G., Tam, W., Nandula, S. V., Shen, Q., … Dalla-Favera, R. (2005). Deregulated BCL6 expression recapitulates the pathogenesis of human diffuse large B cell lymphomas in mice. *Cancer Cell, 7*(5), 445–455.

Chen, Y., He, Y., Lu, X., Zeng, Z., Tang, C., Xue, T., & Li, Y. (2015). Association between Fas/FasL polymorphism and susceptibility to leukemia: A meta-analysis. *International Journal of Clinical and Experimental Medicine, 8*(3), 3817–3824.

Coiffier, B., Brousse, N., Peuchmaur, M., Berger, F., Gisselbrecht, C., Bryon, P. A., & Diebold, J. (1990). Peripheral T-cell lymphomas have a worse prognosis than B-cell lymphomas: a prospective study of 361 immunophenotyped patients treated with the LNH-84 regimen. *Annals of Oncology*, *1*(1), 45–50.

Conde, L., Halperin, E., Akers, N. K., Brown, K. M., Smedby, K. E., Rothman, N., … Skibola, C. F. (2010). Genome-wide association study of follicular lymphoma identifies a risk locus at 6p21.32. *Nature Genetics*, *42*(8), 661–664.

Engels, E. A. (2007). Infectious agents as causes of non-Hodgkin lymphoma. *Cancer Epidemiology Biomarkers & Prevention*, *16*(3), 401–404.

Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., … Bray, F. (2015). Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *International Journal of Cancer*, *136*(5), E359–E386.

Filipovich, A. H., Mathur, A., Kamat, D., & Shapiro, R. S. (1992). Primary immunodeficiencies: Genetic risk factors for lymphoma. *Cancer Research*, *52*(19 Supplement), 5465s-5467s.

Fontana, A., Picoco, C., Masala, G., Prastaro, C., & Vineis, P. (1998). Incidence rates of lymphomas and environmental measurements of phenoxy herbicides: Ecological analysis and case-control study. *Archives of Environmental Health: An International Journal*, *53*(6), 384–387.

Foo, J. N., Smedby, K. E., Akers, N. K., Berglund, M., Irwan, I. D., Jia, X., … Liu, J. (2013). Coding variants at hexa-allelic amino acid 13 of HLA-DRB1 explain independent SNP associations with follicular lymphoma risk. *The American Journal of Human Genetics*, *93*(1), 167–172.

Friedberg, E. C., Walker, G. C., Siede, W., & Wood, R. D. (2005). *DNA repair and mutagenesis*. American Society for Microbiology Press.

Gisselbrecht, C., Gaulard, P., Lepage, E., Coiffier, B., Brière, J., Haioun, C., … others. (1998). Prognostic significance of T-cell phenotype in aggressive non-Hodgkin's lymphomas. *Blood*, *92*(1), 76–82.

Gough, S. C. ., & Simmonds, M. . (2007). The HLA region and autoimmune disease: Associations and mechanisms of action. *Current Genomics*, *8*(7), 453–465.

Han, J.W., Zheng, H.F., Cui, Y., Sun, L.D., Ye, D.Q., Hu, Z., … Zhang, X.J. (2009). Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus. *Nature Genetics*, *41*(11), 1234–1237.

Harris, M. B., Chang, C.C., Berton, M. T., Danial, N. N., Zhang, J., Kuehner, D., … Rothman, P. B. (1999). Transcriptional repression of Stat6-dependent interleukin-4-induced genes by BCL-6: Specific regulation of Iε transcription and immunoglobulin E switching. *Molecular and Cellular Biology*, *19*(10), 7264–7275.

Hill, D. A., Wang, S. S., Cerhan, J. R., Davis, S., Cozen, W., Severson, R. K., … Rothman, N. (2006). Risk of non-Hodgkin lymphoma (NHL) in relation to germline variation in DNA repair and related genes. *Blood*, *108*(9), 3161–3167.

Hodgkin, D. (1832). Absorbent glands and spleen: On some morbid appearances of the absorbent glands and spleen. *The Boston Medical and Surgical Journal*, *6*(7), 101–104.

Hoeijmakers, J. H. (2009). DNA damage, aging, and cancer. *New England Journal of Medicine*, *361*(15), 1475–1485.

International HapMap Consortium. (2003). The international HapMap project. *Nature*, *426*(6968), 789–796.

International Non-Hodgkin's Lymphoma Prognostic Factors Project. (1993). A predictive model for aggressive non-Hodgkin's lymphoma. *The New England Journal of Medicine*, *329*(14), 987–994.

Jardin, F., Bastard, C., Contentin, N., Parmentier, F., Picquenot, J.M., Tilly, H., … Sahota, S. S. (2003). Intronic BCL-6 mutations are preferentially targeted to the translocated allele in t(3;14)(q27;q32) non-Hodgkin B-cell lymphoma. *Blood*, *102*(5), 1872–1876.

Jiao, J., Zheng, T., Lan, Q., Chen, Y., Deng, Q., Bi, X., … Zhang, Y. (2012). Occupational solvent exposure, genetic variation of DNA repair genes, and risk of non-Hodgkin lymphoma. *European Journal of Cancer Prevention* : *The Official Journal of the European Cancer Prevention Organisation (ECP)*, *21*(6), 580–584.

Juszczynski, P., Kalinka, E., Bienvenu, J., Woszczek, G., Borowiec, M., Robak, T., … Warzocha, K. (2002). Human leukocyte antigens class II and tumor necrosis factor genetic polymorphisms are independent predictors of non-Hodgkin lymphoma outcome. *Blood*, *100*(8), 3037–3040.

Kim, I.-S., Kim, D. C., Kim, H.-G., Eom, H.-S., Kong, S.Y., Shin, H.J., … Lee, G.W. (2010). DNA repair gene XRCC1 polymorphisms and haplotypes in diffuse large B-cell lymphoma in a Korean population. *Cancer Genetics and Cytogenetics*, *196*(1), 31–37.

Kumar, P., Henikoff, S., & Ng, P. C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature Protocols*, *4*(7), 1073–1081.

Landgren, O., Engels, E. A., Pfeiffer, R. M., Gridley, G., Mellemkjaer, L., Olsen, J. H., … Goldin, L. R. (2006). Autoimmunity and susceptibility to Hodgkin lymphoma: A population-based case–control study in Scandinavia. *Journal of the National Cancer Institute*, *98*(18), 1321–1330.

Landowski, T. H., Qu, N., Buyuksal, I., Painter, J. S., & Dalton, W. S. (1997). Mutations in the Fas antigen in patients with multiple myeloma. *Blood*, *90*(11), 4266–4270.

Lech-Maranda, E., Baseggio, L., Bienvenu, J., Charlot, C., Berger, F., Rigal, D., … Salles, G. (2004). Interleukin-10 gene promoter polymorphisms influence the clinical outcome of diffuse large B-cell lymphoma. *Blood*, *103*(9), 3529–3534.

Li, Y., Bai, O., Cui, J., & Li, W. (2016). Genetic polymorphisms in the DNA repair gene, XRCC1 associate with non-Hodgkin lymphoma susceptibility: A systematic review and meta-analysis. *European Journal of Medical Genetics*, *59*(2), 91–103.

Lim, U., Morton, L. M., Subar, A. F., Baris, D., Stolzenberg-Solomon, R., Leitzmann, M., … Hartge, P. (2007). Alcohol, smoking, and body size in relation to incident Hodgkin's and non-Hodgkin's lymphoma risk. *American Journal of Epidemiology*, *166*(6), 697–708.

Lim, Y. Y., Chin, Y. M., Tai, M. C., Fani, S., Chang, K. M., Ong, T. C., … Ng, C. C. (2014). Analysis of interleukin-10 promoter single nucleotide polymorphisms and risk of non-Hodgkin lymphoma in a Malaysian population. *Leukemia & Lymphoma*, *56*(1), 163–168.

Lossos, I. S. (2005). Molecular pathogenesis of diffuse large B-Cell lymphoma. *Journal of Clinical Oncology*, *23*(26), 6351–6357.

Milholland, B., Auton, A., Suh, Y., & Vijg, J. (2015). Age-related somatic mutations in the cancer genome. *Oncotarget*, *6*(28), 24627–24635.

Morton, L. M., Bernstein, L., Wang, S. S., Hein, D. W., Rothman, N., Colt, J. S., … Zahm, S. H. (2007). Hair dye use, genetic variation in N-acetyltransferase 1 (NAT1) and 2 (NAT2), and risk of non-Hodgkin lymphoma. *Carcinogenesis*, *28*(8), 1759–1764.

Neri, A., Barriga, F., Inghirami, G., Knowles, D. M., Neequaye, J., Magrath, I. T., & Favera, R. D.-. (1991). Epstein-Barr virus infection precedes clonal expansion in Burkitt's and acquired immunodeficiency syndrome-associated lymphoma. *Blood*, *77*(5), 1092–1095.

Opelz, G., & Henderson, R. (1993). Incidence of non-Hodgkin lymphoma in kidney and heart transplant recipients. *The Lancet*, *342*(8886), 1514–1516.

Parsonnet, J., Hansen, S., Rodriguez, L., Gelb, A. B., Warnke, R. A., Jellum, E., & Orentreich, N. (1994). *Helicobacter pylori* infection and gastric lymphoma. *The New England Journal of Medicine*, *330*(18):1267-71.

Pritchard, J. K., & Cox, N. J. (2002). The allelic architecture of human disease genes: Common disease-common variant...or not? *Human Molecular Genetics*, *11*(20), 2417–2423.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., … Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, *81*(3), 559–575.

Reich, D. E., & Lander, E. S. (2001). On the allelic spectrum of human disease. *Trends in Genetics*, *17*(9), 502–510.

Salzer, U., Chapel, H. M., Webster, A. D. B., Pan-Hammarström, Q., Schmitt-Graeff, A., Schlesier, M., … Grimbacher, B. (2005). Mutations in *TNFRSF13B* encoding TACI are associated with common variable immunodeficiency in humans. *Nature Genetics*, *37*(8), 820–828.

Sawicki, M. P., Samara, G., Hurwitz, M., & Passaro, E. (1993). Human genome project. *The American Journal of Surgery*, *165*(2), 258–264.

Sehn, L. H., Berry, B., Chhanabhai, M., Fitzgerald, C., Gill, K., Hoskins, P., … Connors, J. M. (2007). The revised International Prognostic Index (R-IPI) is a better predictor of outcome than the standard IPI for patients with diffuse large B-cell lymphoma treated with R-CHOP. *Blood*, *109*(5), 1857–1861.

Shabalin, A. A. (2012). Matrix eQTL: Ultra fast eQTL analysis via large matrix operations. *Bioinformatics*, *28*(10), 1353–1358.

Shen, M., Purdue, M. P., Kricker, A., Lan, Q., Grulich, A. E., Vajdic, C. M., … Armstrong, B. K. (2007). Polymorphisms in DNA repair genes and risk of non-Hodgkin's lymphoma in New South Wales, Australia. *Haematologica*, *92*(9), 1180–1185.

Shipp, M. A., Ross, K. N., Tamayo, P., Weng, A. P., Kutok, J. L., Aguiar, R. C. T., … Golub, T. R. (2002). Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. *Nature Medicine*, *8*(1), 68–74.

Shipp, M., Harrington, D., Anderson, J., Armitage, J., Bonadonna, G., Brittinger, G., … Yeap, B. (1993). A predictive model for aggressive non-Hodgkin's lymphoma. *New England Journal of Medicine*, *329*(14), 987–994.

Sibley, K., Rollinson, S., Allan, J. M., Smith, A. G., Law, G. R., Roddam, P. L., ... & Morgan, G. J. (2003). Functional FAS promoter polymorphisms are associated with increased risk of acute myeloid leukemia. *Cancer Research*, *63*(15), 4327-4330.

Siegel, R. L., Miller, K. D., & Jemal, A. (2015). Cancer statistics, 2015. *CA: A Cancer Journal for Clinicians*, *65*(1), 5–29.

Sillé, F. C. M., Conde, L., Zhang, J., Akers, N. K., Sanchez, S., Maltbaek, J., … Skibola, C. F. (2014). Follicular lymphoma-protective HLA class II variants correlate with increased HLA-DQB1 protein expression. *Genes and Immunity*, *15*(2), 133–136.

Skibola, C. F., Akers, N. K., Conde, L., Ladner, M., Hawbecker, S. K., Cohen, F., … Bracci, P. M. (2012). Multi-locus HLA class I and II allele and haplotype associations with follicular lymphoma. *Tissue Antigens*, *79*(4), 279–286.

Skibola, Christine F, Bracci, P. M., Halperin, E., Conde, L., Craig, D. W., Agana, L., … Brown, K. M. (2009). Genetic variants at 6p21.33 are associated with susceptibility to follicular lymphoma. *Nature Genetics*, *41*(8), 873–875.

Smedby, Karin E., Foo, J. N., Skibola, C. F., Darabi, H., Conde, L., Hjalgrim, H., … Liu, J. (2011). GWAS of follicular lymphoma reveals allelic heterogeneity at 6p21.32 and suggests shared genetic susceptibility with diffuse large B-cell lymphoma. *PLoS Genetics*, *7*(4), e1001378.

Smedby, Karin Ekström, Lindgren, C. M., Hjalgrim, H., Humphreys, K., Schöllkopf, C., Chang, E. T., … Adami, H.-O. (2006). Variation in DNA repair genes ERCC2, XRCC1, and XRCC3 and risk of follicular lymphoma. *Cancer Epidemiology and Prevention Biomarkers, 15*(2), 258-265.

Smigielski E. M., Sirotkin, K., Ward, M., Sherry, S. T. (2000). dbSNP: A database of single nucleotide polymorphisms. *Nucleic Acids Research, 28*(1), 352–355.

Sujatha, S., Gan, G. G., Zain, S. M., Haerian, B. S., Bee, P. C., Lian, L. H., … Mohamed, Z. (2014). Polymorphisms in methylenetetrahydrofolate reductase gene and risk of non-Hodgkin lymphoma in a multi-ethnic population. *Journal of Human Genetics, 59*(5), 280.

Swerdlow, S., Campo, E., Harris, N., Jaffe, E. S., Pileri, S., & Stein, H. (2008). *WHO classification of tumours of haematopoietic and lymphoid tissues (4th ed.)*. Lyon, France: IARC.

Swerdlow, S. H., Campo, E., Pileri, S. A., Harris, N. L., Stein, H., Siebert, R., … Jaffe, E. S. (2016). The 2016 revision of the World Health Organization (WHO) classification of lymphoid neoplasms. *Blood, 127*(20), 2375–2390.

Tan, D. E. K., Foo, J. N., Bei, J.-X., Chang, J., Peng, R., Zheng, X., … Liu, J. (2013). Genome-wide association study of B cell non-Hodgkin lymphoma identifies 3q27 as a susceptibility locus in the Chinese population. *Nature Genetics, 45*(7), 804–807.

The 1000 Genomes Project Consortium. (2015). A global reference for human genetic variation. *Nature, 526*(7571), 68–74.

The GTEx Consortium. (2015). The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science, 348*(6235), 648-660.

Vijai, J., Kirchhoff, T., Schrader, K. A., Brown, J., Dutra-Clarke, A. V., Manschreck, C., … Offit, K. (2013). Susceptibility loci associated with specific and shared subtypes of lymphoid malignancies. *PLoS Genetics, 9*(1), e1003220.

Wang, S. S., Purdue, M. P., Cerhan, J. R., Zheng, T., Menashe, I., Armstrong, B. K., … Rothman, N. (2009). Common gene variants in the tumor necrosis factor (TNF) and TNF receptor superfamilies and NF-kB transcription factors and non-hodgkin lymphoma risk. *PLoS ONE, 4*(4), e1003220.

Warzocha, K., Ribeiro, P., Bienvenu, J., Roy, P., Charlot, C., Rigal, D., … Salles, G. (1998). Genetic polymorphisms in the tumor necrosis factor locus influence non-Hodgkin's lymphoma outcome. *Blood*, *91*(10), 3574–3581.

Watanabe-Fukunaga, R., Brannan, C. I., Copeland, N. G., Jenkins, N. A., & Nagata, S. (1992). Lymphoproliferation disorder in mice explained by defects in Fas antigen that mediates apoptosis. *Nature*, *356*(6367), 314–317.

Willett, E. V., Smith, A. G., Dovey, G. J., Morgan, G. J., Parker, J., & Roman, E. (2004). Tobacco and alcohol consumption and the risk of non-Hodgkin lymphoma. *Cancer Causes & Control*, *15*(8), 771–780.

Wood, R. D., Mitchell, M., Sgouros, J., & Lindahl, T. (2001). Human DNA repair genes. *Science*, *291*(5507), 1284–1289.

Worrillow, L., Roman, E., Adamson, P. J., Kane, E., Allan, J. M., & Lightfoot, T. J. (2009). Polymorphisms in the nucleotide excision repair gene ERCC2/XPD and risk of non-Hodgkin lymphoma. *Cancer Epidemiology*, *33*(3–4), 257–260.

Wrench, D., Leighton, P., Skibola, C. F., Conde, L., Cazier, J.B., Matthews, J., … Fitzgibbon, J. (2011). SNP rs6457327 in the HLA region on chromosome 6p is predictive of the transformation of follicular lymphoma. *Blood*, *117*(11), 3147–3150.

Zainal Ariffin, O., & Nor Saleha, I. (2011). *National cancer registry report 2007*. Malaysia: Ministry of Health.

Zhang, Z., Wang, L.E., Sturgis, E. M., El-Naggar, A. K., Hong, W. K., Amos, C. I., … Wei, Q. (2006). Polymorphisms of FAS and FAS ligand genes involved in the death pathway and risk and progression of squamous cell carcinoma of the head and neck. *Clinical Cancer Research*, *12*(18), 5596-5602.

Zhou, J.Y., He, L.W., Liu, J., Yu, H.L., Wei, M., Ma, W.L., & Shi, R. (2014). Comprehensive assessment of associations between ERCC2 Lys751Gln/Asp312Asn polymorphisms and risk of non-Hodgkin lymphoma. *Asian Pacific Journal of Cancer Prevention*, *15*(21), 9347–9353.

Zhu, M.L., Wang, M., Cao, Z.G., He, J., Shi, T.Y., Xia, K.Q., … Wei, Q.Y. (2012). Association between the ERCC5 Asp1104His polymorphism and cancer risk: A meta-analysis. *PLoS ONE*, *7*(7), e36293.

LIST OF PUBLICATIONS AND PAPERS PRESENTED

Publication:

Ten, L. C., Chin, Y. M., Tai, M. C., Chin, E. F. M., Lim, Y. Y., Suthandiram, S., … Ng, C. C. (2017). SNP variants associated with non-Hodgkin lymphoma (NHL) correlate with human leukocyte antigen (HLA) class II expression. *Scientific Reports, 7*, 41400.

Presentations:

Ten, L. C., Gan, G. G., & Ng, C. C. (2015, August). *Association of single nucleotide polymorphisms (SNPs) in DNA repair genes and the risk of non-Hodgkin lymphoma (NHL) in Malaysia population.* Poster presented at the Annual Malaysian Society for Molecular Biology and Biotechnology (MSMBB) Scientific Meeting, Kuala Lumpur, Malaysia.

Ten, L. C., Gan, G. G., & Ng, C. C. (2015, December). *Genetic variation in human leukocyte antigen (HLA) region and the risk of non-Hodgkin lymphoma (NHL) in Malaysia population.* Poster presented at the Biological Sciences Graduate Congress (BSGC), Bangkok, Thailand.