การศึกษาหน้าที่ของความหลากหลายทางพันธุกรรม (*IFI16, MNDA* และ *AIM2*) ที่สัมพันธ์กับโรคเอสแอลอี

นางสาวอิงอร กิมกง

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต
สาขาวิชาจุลชีววิทยาทางการแพทย์ (สหสาขาวิชา)
บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2551

FUNCTIONAL CHARACTERIZATION OF GENETIC POLYMORPHISMS (*IFI16*, *MNDA* AND *AIM2*) ASSOCIATED WITH SLE DISEASE

Miss Ingorn Kimkong

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy Program in Medical Microbiology
(Interdisciplinary Program)
Graduate School
Chulalongkorn University
Academic Year 2008

Thesis Title	Functional characterization of genetic polymorphisms (*IFI16*, *MNDA* and *AIM2*) associated with SLE disease
By	Miss Ingorn Kimkong
Field of Study	Medical Microbiology
Thesis Principal Advisor	Associate Professor Nattiya Hirankarn, M.D. Ph.D.
Thesis Co-advisor	Assistant Professor Tewin Tencomnao, Ph.D.

Accepted by the Graduate School, Chulalongkorn University in Partial Fulfillment of the Requirements for the Doctoral Degree

..... Dean of the Graduate School
(Associate Professor Pornpote Piumsomboon, Ph.D.)

THESIS COMMITTEE

..... Chairman
(Assistant Professor Anan Chongthaleong, M.D.)

..... Thesis Principal Advisor
(Associate Professor Nattiya Hirankarn, M.D. Ph.D.)

..... Thesis Co-advisor
(Assistant Professor Tewin Tencomnao, Ph.D.)

..... External Member
(Assistant Professor Wallaya Jongjaroenprasert, M.D.)

..... Member
(Assistant Professor Jongkonnee Wongpiyabovorn, M.D. Ph.D.)

อิงอร กิมกง : การศึกษาหน้าที่ของความหลากหลายทางพันธุกรรม (*IFI16, MNDA* และ *AIM2*) ที่สัมพันธ์กับโรคเอสแอลอี (FUNCTIONAL CHARACTERIZATION OF GENETIC POLYMORPHISMS (*IFI16, MNDA* และ *AIM2*) ASSOCIATED WITH SLE DISEASE) อ. ที่ปรึกษาวิทยานิพนธ์หลัก : รศ.พญ.ดร. ณัฏฐิยา หิรัญกาญจน์, อ.ที่ปรึกษาวิทยานิพนธ์ร่วม: ผศ.ดร. เทวิน เทนคำเนาว์, 139 หน้า.

โรค SLE เป็นโรคภูมิต้านทานต่อเนื้อเยื่อตนเองที่มีผลต่อหลายอวัยวะที่สำคัญของร่างกาย งานวิจัยนี้ใช้การศึกษาแบบ case-control association study ซึ่งประกอบด้วยการศึกษาแบบ pooling genome wide association (GWA) และ candidate gene association เพื่อที่จะค้นหา SNPs ที่สัมพันธ์กับความเสี่ยงต่อการเกิดโรค SLE และ/หรือ ความรุนแรงของโรค จากผลการศึกษาของ pooling GWA ไม่สามารถระบุ SNP ได้ด้วยค่า *P* value และ odds ratio ที่แตกต่างกันเนื่องจาก power ในการศึกษาที่จำกัด อย่างไรก็ตาม ผลจากเปรียบเทียบกับ candidate genes ที่มีการศึกษาก่อนหน้านี้ บ่งชี้ได้ว่ายีน *TNFb, HLA,TNXB* และ *TNFSAIP3* เป็นยีนที่มีความสำคัญในประชากรไทยซึ่งมีค่า *P* (คำนวณโดย t-test) น้อยกว่า 0.0001 ในการศึกษานี้ได้คัดเลือกยีนที่น่าสนใจจากบริเวณ SLE major susceptibility loci ในโครโมโซม 1 เป็นตัวอย่างสำหรับการทำ individual genotyping ผลการศึกษาจาก individual genotyping พบว่ารูปแบบ allele frequency ของ SNPs ในยีน *NOS1AP, IFIX (PYHIN1), TLR5* และ *CD1D* จาก individual genotyping เหมือนกันกับรูปแบบจากวิธีการ pooling งานวิจัยนี้ได้คัดเลือก *IFIX* เพื่อศึกษาเพิ่มเติมในส่วนของการศึกษาโดย candidate gene approach นอกจากยีน *IFIX* แล้ว เรายังสนใจยีน *MNDA, IFI16* และ *AIM2* ซึ่งอยู่ในบริเวณเดียวกันและทั้งหมดเป็นยีนในกลุ่มของ IFN-inducible genes ยีนเหล่านี้เป็นยีนที่มีความเสี่ยงต่อโรค SLE ที่สำคัญเนื่องจาก 1) ข้อมูลจากการทำแผนที่ยีนในหนูที่เป็น SLE and 2) ข้อมูลจาก microarray ที่พบว่ายีนกลุ่ม IFN-inducible genes มีการแสดงออกสูงในผู้ป่วย SLE และ 3) IFI16 ได้มีการพิสูจน์ว่าเป็น autoantigen ชนิดใหม่สำหรับผู้ป่วย SLE การศึกษาส่วนนี้ได้ทำการ genotyping SNPs 10 ตำแหน่งจาก 4 ยีนนี้ ในผู้ป่วย SLE 200 รายเปรียบเทียบกับคนปกติ 200 ราย จากผลการวิเคราะห์ความสัมพันธ์ พบว่า SNP ภายใน *IFIX* (rs856084, OR = 1.37, 95%CI =1.01-1.87) และ *IFI16* (rs866484, OR = 1.37, 95%CI =1.03-1.82 และ rs1772414, OR = 1.41, 95%CI =1.06-1.88) มีความสำคัญอย่างเป็นอิสระต่อกัน เพื่อที่จะทำให้บทบาทของยีน *IFIX* และ *IFI16* ชัดเจนยิ่งขึ้น การศึกษาที่ครอบคลุมโดยการใช้ SNPs ที่ละเอียดมากขึ้นและการเพิ่มขนาดตัวอย่างจำเป็นสำหรับการศึกษาต่อไปในอนาคต และเพื่อที่จะทำให้เข้าใจหน้าที่ของยีนเหล่านี้ดียิ่งขึ้น งานวิจัยนี้ได้ศึกษาการแสดงออกของยีน *MNDA, IFIX, IFI16* และ *AIM2* จากเซลล์หลายชนิดในผู้ป่วย SLE เปรียบเทียบกับคนปกติ ซึ่งจากการศึกษาพบการแสดงออกที่เพิ่มขึ้นของ IFN-inducible genes ใน leukocytes จากผู้ป่วย SLE แต่ไม่พบใน lymphocytes และ ชิ้นเนื้อไต จากผลการศึกษาอาจบ่งชี้ถึงความสำคัญของยีนเหล่านี้ในการพัฒนาของโรค SLE โดยการแสดงออกในเม็ดเลือดขาวชนิดอื่นๆ เช่น monocytes และ granulocytes นอกจากนี้ การที่ไม่พบการแสดงออกของยีนเหล่านี้ในชิ้นเนื้อไตแสดงว่ายีนเหล่านี้อาจจะไม่มีความสำคัญในพยาธิสภาพทางไตของผู้ป่วย SLE

สาขาวิชา จุลชีววิทยาทางการแพทย์ ลายมือชื่อนิสิต.....
ปีการศึกษา 2551 ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก.....
ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์ร่วม.....

4689696220 : MAJOR MEDICAL MICROBIOLOGY

KEY WORD: SYSTEMIC LUPUS ERYTHEMATOSUS / IFI / SNP

INGORN KIMKONG : FUNCTIONAL CHARACTERIZATION OF GENETIC POLYMORPHISMS (*IFI16, MNDA* AND *AIM2*) ASSOCIATED WITH SLE DISEASE. THESIS PRINCIPAL ADVISOR : ASSOC. PROF. NATTIYA HIRANKARN, M.D. Ph.D. THESIS CO-ADVISOR : ASST. PROF. TEWIN TENCOMNAO, 139 pp.

Systemic lupus erythematosus (SLE) is an autoimmune disease that affects multiple organs. In this study, we applied case-control association study including pooling genome wide association (GWA) and candidate gene association studies to search for SNPs associated with SLE susceptibility and/or severity. We could not identify any SNPs with distinct p-value or odds ratio from our pooling GWA result due to limited power. However, the GWA result compared with known candidate genes suggested that *TNFb, HLA,TNXB and TNFSAIP3* genes are important candidate genes in Thai population giving positive association with p-value (by t-test) less than 0.0001. We selected new candidate genes from SLE major susceptibility loci in chromosome 1 as a model for individual genotyping. Our individual genotyping results showed that the allele frequency patterns of SNPs in *NOS1AP, IFIX (PYHIN1), TLR5, CD1D* genes from individual genotyping were similar to the patterns from pooling approach. We selected *IFIX* for further study in candidate gene's part. Beside *IFIX*, we also focus on *MNDA, IFI16* and *AIM2* genes which located in the same region and are all IFN-inducible genes. They are important SLE susceptibility genes due to several reasons including 1) genetic mapping from lupus murine model and 2) an upregulated IFN-inducible genes in patients with SLE from microarray studies and 3) IFI16 was identified as new autoantigen for patients with SLE. We genotyped 10 SNPs from these 4 genes in 200 patients with SLE vs. 200 normal controls. From our association results, we found that SNP within *IFIX* (rs856084, OR = 1.37, 95%CI =1.01-1.87) and *IFI16* (rs866484, OR = 1.37, 95%CI =1.03-1.82 and rs1772414, OR = 1.41, 95%CI =1.06-1.88) are independently important. To clarify the role of *IFIX* and *IFI16* gene, more extensive study using dense SNPs and increasing sample sizes are required in future study. In order to better understand functions of these genes, we studied the expression of *MNDA, IFIX, IFI16* and *AIM2* genes from various cell types in patients with SLE compared to healthy controls. We found the increasing of IFN-inducible genes in leukocytes from patients with SLE, but not in lymphocytes and kidney. These results may indicate the importance of these genes in SLE development via expression in other white blood cell subsets such as monocytes and granulocytes. The negative finding of these genes in kidney indicates that they might not be important in renal pathology.

Field of study Medical Microbiology
Academic year 2008

Student's signature.....Ingorn Kimkong.....
Principal Advisor's signature.....
Co-advisor's signature.....

ACKNOWLEDGEMENTS

I am deeply indebted to my advisor, Associate Professor Dr. Nattiya Hirankarn, for her excellent supervision, valuable advice, constructive criticism, and guidance throughout the course of this study. Her expertise in immunology improved my research skills and prepared me for future challenges. I would like to thanks my co-advisor, Assistant Professor Dr. Tewin Tencomnao for his scientific advice and great assistance.

I would like to thank the committee of Medical Microbiology Program for giving me permission to commence this dissertation in the first instance, to do the necessary research work.

I would like to thank Mr. Jeerawat Nakkuntod for his kindness in guidance throughout the laboratories and Miss Thitima Benjachat for helping in laboratories and providing the beneficial instrument for this study. Without them, this work would not be accomplished. I would also like to thank Lupus Research Unit members for fruitful discussions, constant encouragement, and great relationships.

I would like to thank Dr. Ivo Glynne Gut, Centre National de Génotypage, Evry, France, for giving me the opportunity to work in his laboratory for whole genome association study and great supporting. I would also like to thank Dr. Santitham Prom-on, Systems Biology and Bioinformatics Research Laboratory, Department of Computer Engineering, King Mongkut's University of Technology Thonburi, for helping me to analyze data in whole genome association study.

I would like to express special thank to my parents and my friends for their love, support, understanding and encouragement. Finally, I wish to give my deepest gratitude to Commission on Higher Education, Ministry of Education, Thailand, for major financial support throughout my education in Doctor of Philosophy degree according to the requirement of Kasetsart University.

CONTENTS

Page

Page

LIST OF TABLES

Table Page

Table Page

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

LIST OF ABBREVIATIONS

SLE	Systemic lupus erythematosus
dsDNA	double–standed deoxyribonucleic acid
et al	et alii
GWA	genome wide association
SNP	single nucleotide polymorphism
IFN	interferon
MNDA	Myeloid Nuclear Differentiation Antigen
IFIX	Interferon-inducible protein X
IFI16	Interferon-Inducible protein 16
AIM2	Absent In Melanoma 2
PCR	Polymerase Chain Reaction
RFLP	Restriction fragment Length Polymorphism
SSP	Sequence Specific Primer
RT-PCR	reverse transcription- Polymerase Chain Reaction
λ_s	sibling risk ratio
LD	Linkage Disequilibriam
µl	microlitter
µg	microgram
ml	milliliter
mM	millimolar
MW	molecular weight
Ng	nanogram
bp	base pair
°C	degree Celsius
EMSA	Electrophoretic Mobility Shift Assay
OR	Odd ratio
LN	lupus nephritis

CHAPTER I

INTRODUCTION

Systemic lupus erythematosus (SLE) is an autoimmune disease that affects multiple organs. SLE is more common in African-Americans, African-Caribbeans and Asians, than in Caucasians (Petri M, 2002). In addition, there are also differences in the development of disease by ethnicity. A study of Mok CC et al indicated that Asian patients with SLE have more severity in the progression of disease, especially renal involvement (Mok CC and Lau CS, 2003). The hallmark of SLE is the production of autoantibodies directed against constituents of the cell nucleus such as anti-nuclear antibodies and antibodies to double stranded DNA (ds-DNA). These autoantibodies cause end-organ damage via inflammatory response to immune complex, which result in various clinical manifestations such as glomerulonephritis, arthritis, serositis, vasculitis and neurological disorders (Nguyen C et al, 2002). The etiopathogenesis of SLE remain elusive. However, there is clearly a genetic component to SLE susceptibility shown by familial clustering (Vyse TJ and Todd JA, 1996) and twin's studies (Deapen D et al., 1992).

Two basic strategies have been employed in human SLE genetic studies: linkage and association analysis. To date, there are seven susceptibility loci in human whole-genome linkage studies have reached the threshold for significant linkage: 1q22-23, 1q41–42, 2q37, 4p16, 6p21-p11, 16q13 and 17p13 (Kelly JA et al., 2002). Several candidate genes within those loci have been studied using classical association analysis and found to be associated with SLE. Nevertheless, many of the genes associated with SLE have not been confirmed and remain controversial. Currently, with the development of genotyping platforms that permit analysis of hundreds of thousands of SNPs, it is possible to apply this principle of indirect association to the whole genome rather than just candidate genes or candidate linkage regions. Four recent genome wide association (GWA) studies in Caucasian patients with SLE from international

collaborations and only one study in Japanese patients have been reported (Harley JB et al, 2008; Hom G et al., 2008; Kozyrev S et al., 2008; Musone SL et al., 2008, Oishi T et al., 2008; Kamatani Y et al., 2008). Results from these studies confirmed known candidate genes and also discovered several new candidate genes. However, these are the studies restricted in Caucasian patients and only one study in Japanese patients, the exploration of SLE susceptibility genes in our population is still required.

In this study, we applied case-control association study including genome wide association (GWA) and candidate gene association studies.

In GWA study, an approach that balance benefit of genome-wide screening and the costs in genotyping is DNA pooling. Several groups have applied the technology from Affymetrix (microarray-based GeneChip® Mapping arrays) to assess the reliability and validity of SNP allelic frequency measurements as determined from pooling genomic DNA samples on SNP mapping arrays. They have suggested that genotyping of SNPs with DNA pooling using Affymetrix microarrays produces highly accurate results and can be used for genome-wide association studies (Butcher LM et al, 2004; Meaburn E et al., 2005; Craig DW et al., 2005; Simpson CL et al., 2005; Kirov G et al., 2006; Meaburn E et al., 2006). Several applications of DNA pooling technique for genome-wide association studies have been reported (Butcher LM et al, 2005; Johnson C et al., 2006; Steer S et al., 2007; Jongjaroenprasert W et al., 2008). In the present study, we aimed to search for SNPs associated with SLE susceptibility by genome-wide screening using Affymetrix 50K GeneChip on pooled genomic DNA.

For candidate gene study, a group of IFN-inducible genes (*MNDA, IFIX, IFI16* and *AIM2*) maps to chromosome 1q21–23 which is major SLE susceptible loci was proposed as new candidate genes for SLE susceptibility. Several studies have supported this hypothesis. First, data from lupus murine model has identified IFN-inducible protein 202 gene (ifi202) as a candidate for lupus susceptibility (Rozzo SJ et al., 2001; Choubey D et al., 2002). This gene is homologous to human IFN-inducible genes (*MNDA, IFIX, IFI16* and *AIM2*) and located on chromosome 1 similar to human. Second, IFI16 was identified as new autoantigen for patients with SLE, up to 29% of 374 SLE patients develop high titer autoantibodies to IFI16 protein (Seelig HP et al., 1994). Later study has also found that 26% of patients with SLE exhibited significantly higher

anti-IFI16 IgG antibody levels compared with control subjects. By immunohistochemical analysis, IFI16 was highly expressed in epidermis and dermal inflammatory infiltrates of skin lesions from patients with SLE (Mondini M et al., 2006). Lastly, a recent study of gene expression profiles using microarray has found 2.4 fold up-regulated IFI16 in leukocytes of patients with SLE as compared to healthy controls (Alcorta DA et al., 2007). These evidences indicate that a group of IFN-inducible genes are likely to be important factors contributing to SLE. Hence, we aimed to study the role of this gene group by searching functional single nucleotide polymorphisms (SNPs) within *MNDA, IFIX, IFI16* and *AIM2* involving in disease susceptibility and/or severity. The genotyping method for these four genes polymorphisms were conducted by Polymerase Chain Reaction-Restriction fragment Length Polymorphism (PCR-RFLP) and Polymerase Chain Reaction-Sequence Specific Primer (PCR-SSP). The genotype and allele frequencies were then compared between patients and control subjects by Chi-square test. The SNPs that found to be associated with SLE will be functional characterized. Moreover, to further understand the role of these genes in SLE, we studied the expression of *MNDA, IFIX, IFI16* and *AIM2* genes from various cell types (leukocytes, B lymphocytes, T lymphocytes and renal biopsy tissue) in patients with SLE compared to healthy controls. Real time reverse transcription-PCR (RT-PCR) was used to quantify the expression of these genes. It is likely that the difference might be limited to particular isoform. Therefore, the expression of *IFIX* and *IFI16* isoforms was also semi-quantified by conventional RT-PCR.

We hypothesized that new candidate gene from genome wide association study on pooled genomic DNA will be discovered and IFN-inducible genes (*MNDA, IFIX, IFI16* and *AIM2*) may play a role in pathogenesis of SLE from candidate gene association and expression profile studies. Our research will provide a better understanding of the disease mechanism and severity by studying the role of new candidate genes. These might lead to the development of new therapeutic strategies and prevention in the future.

CHAPTER II

OBJECTIVE

The objectives of this study were:

Part I Whole genome association study (Pooling)

To search for SNPs associated with SLE susceptibility by genome-wide screening using Affymetrix 50K GeneChip on pooled genomic DNA.

Part II Candidate gene association study

1. To search for functional SNPs of *MNDA, IFIX, IFI16* and *AIM2* genes involving in disease susceptibility and/or severity.
2. To analyze the functional role of SNPs that associated with SLE.

Part III Expression profile study

To study the expression of *MNDA, IFIX, IFI16* and *AIM2* genes from various cell types (leukocytes, B lymphocytes, T lymphocytes and renal biopsy tissue) in patients with SLE compared to healthy controls.

CHAPTER III

LITERATURE REVIEW

Systemic Lupus Erythematosus (SLE)

SLE is an autoimmune disease that affects multiple organs, including the skin, joints, kidneys, heart, lungs, and central nervous system. The hallmark of SLE is the production of autoantibodies directed against constituents of the cell nucleus such as anti-nuclear antibodies and antibodies to double stranded DNA (ds-DNA). These autoantibodies cause end-organ damage via inflammatory response to immune complex, which result in various clinical manifestations such as glomerulonephritis, arthritis, serositis, vasculitis and neurological disorders (Nguyen C et al, 2002). The diagnosis of SLE is based on the American College of Rheumatology (ACR) diagnostic criteria for classification originally developed in 1982 (Tan EM et al., 1982) and updated in 1997 (Hochberg MC et al., 1997) (Table 1). If four of the criteria are present at any time during the course of disease, a diagnosis of SLE can be made.

Table 1 Criteria for Classification of Systemic Lupus Erythematosus

Criteria	Definition
Malar rash	Fixed erythema over malar areas, sparing nasolabial folds
Discoid rash	Erythematous raised patches, keratotic scaling, follicular plugging
Photosensitivity	Skin rash after exposure to sunlight; history or physical exam
Oral ulcers	Oral or nasopharyngeal, painless, by physical exam
Nonerosive arthritis	Tenderness, swelling, effusion in 2 or more peripheral joints
Pleuritis or pericarditis	Convincing history or physical exam or ECG or other evidence
Renal disorder	>0.5g protein/d or 3+ or cellular casts
Seizures, psychosis	Not due to drugs, metabolic derangement, etc.
Hematologic disorder	Hemolytic anemia/leukopenia (<4000) or lymphopenia (<1500 twice) or thrombocytopenia (<100,000) without other causes
Immunologic disorder	Anti-dsDNA, anti-Sm, antiphospholipid antibodies, (anticardiolipin, lupus anticoagulant or false/pos test for syphilis)
Positive ANA	Not drug-induced

ECG, electrocardiogram; ANA, antinuclear antibody.

Epidemiology of SLE

Systemic lupus erythematosus (SLE) is now recognized as a common disease occurring worldwide, and the usual onset occurs between 15 and 45 years of age (Trethewey P, 2004). SLE is more prevalent in female (female:male ratio ~ 8:1). The prevalence of SLE in the general population is approximately 1 in 2000 but it varies among ethnic groups (Tsao BP, 2003). SLE is more common in African-Americans, African-Caribbeans and Asians, than in Caucasians (Petri M, 2002). The prevalence of SLE among Northern Europeans is approximately 40 cases per 100,000 persons, while among blacks have a much higher prevalence, with rate of more than 200 cases per 100,000 persons (Johnson AE et al., 1995). In Thailand, there was no official report on the incidence or prevalence of SLE. However, SLE has been reported in Hong Kong, with the estimated point prevalence is around 58.8 per 100,000 population (rates for men and women are 11.7 cases per 100000 persons and 104 cases per 100,000 persons, respectively) (Mok CC and Lau CS, 2003). In addition, the outcome of SLE may also differ among racial/ethnic groups. Mok CC and coworker have been reported that Chinese patients with SLE in Hong Kong have more serious organ manifestations than Caucasians, especially renal disease. The frequencies of renal disease were 50% in Chinese patients and 31-39% in Caucasian patients (Mok CC and Lau CS, 2003; Cervera R et al., 1993; Pistiner M et al., 1991). Renal manifestation has been reported in Thai patients with SLE with higher frequency than Chinese patients with SLE, with the reported frequency of 78% (Deesomchok U and Tumrasvin T, 1983). Furthermore, another manifestation such as neuropsychiatry was also found with higher frequency in Thai patients compare to Chinese patients (Kasitanon N et al., 2002). Therefore, from these data, it can help indicated that Asian patients with SLE, particularly Thai have more severity in the progression of disease. In the past 45 years, the 5-year survival has improved from 50% to 88-96% of patients with SLE. Furthermore, 77-85% of patients survive for 10 years and 20-year survival approaches 70% (Bongu A et al., 2002). Although mortality has hugely improved, death rate for patients with SLE remain 3.2 times greater than those in the general population (Feng PH, 2007).

Etiology

Although the precise etiopathogenesis of SLE remains elusive, both genetic and environmental factors are probably implicated.

1. Environmental Factors

Environmental triggers have been reported as a risk factor in the development of SLE comprising drugs, chemicals, ultraviolet light, dietary factors, viruses, and environmental estrogen (Mok CC and Lau CS, 2003).

Certain medications such as procainamide, hydralazine and quinidine could induce a lupus-like syndrome (D'Cruz D, 2000; Borchers AT, 2007).

Use of permanent hair dyes containing aromatic amines has also been associated with the development of SLE (Freni-Titulaer LW et al., 1989; Cooper GS et al., 2001). However, this has not been confirmed by case control or cohort studies (Petri M and Allbritton J, 1992; Sanchez-Guerrero J et al., 1996).

Ultraviolet (UV) light, particularly UVB, is a crucial trigger in patients with SLE. This light could induce apoptosis, cause inflammation and tissue damage (Bijl M et al., 2007).

Dietary factors are also important in the pathogenesis of SLE. The consumption of alfalfa sprouts that contain L-canavanine has been linked to the development of lupus-like symptoms (Prete PE, 1985).

Some infectious agent such as Epstein-Barr virus is involved in the development of SLE (James JA et al., 1997; Harley JB and James JA, 1999). EBV may induce specific responses by molecular mimicry and disturb immunoregulation (Poole BD et al., 2006).

Also, exposure to environmental estrogens through the ingestion of meat and milk products, postmenopausal estrogen and oral contraceptives may be important triggers for SLE (D'Cruz D, 2000; Mok CC and Lau CS, 2003).

2. Genetic Factors

There is clearly a genetic component to SLE susceptibility shown by familial clustering and twin's studies. The degree of familial clustering was measured by comparing the risk of a sibling with the risk in the population as a whole (λs), varies between 20 and 40 (Nath SK et al., 2004; Alarcon-Segovia D et al., 2005). In other words, the siblings of SLE patient are 20-40 times more likely to develop SLE than those who do not have an affected sibling. The evidence in twins study is higher concordance rates in monozygotic twins ranging between 24% and 58%, while the concordance rate is only 2-5% in dizygotic twins (Deapen D et al., 1992; Wong M and Tsao BP, 2006). Such 10-fold difference suggests that multiple genes shared between each pair of twins considerably influence the susceptibility to SLE.

Genetic studies in SLE

Two basic strategies have been employed in human SLE genetic studies: linkage and association analysis (Lander ES and Schork NJ, 1994). In addition, a useful alternative strategy is defining the genetics of SLE in mouse models (Kono DH and Theofilopoulos AN, 2006).

1. Human linkage analysis

The linkage study is used to seek for co-segregation of a particular genetic marker with disease in families of affected individuals. The markers themselves are not usually functional but are linked to disease causing variants according to possibility that proximal consequences on chromosome will be inherited together at meiosis. The consistent linkage of a marker with disease in several affected families therefore identifies a susceptibility region in the genome (Rhodes B and Vyse T J, 2007). However, linkage signals detected in complex diseases may result from the combined effect of a cluster of genes of modest effects. False positive results can occurred due to multiple testing since numerous numbers of genetic markers are assayed in a genome scan. Hence, Lander and Kruglyak proposed thresholds for significant linkage (*ie*, a genome-wide significance of $P < 0.05$ corresponds to a logarithm of odds score of 3.3

or 3.6, depending on the linkage method used) and suggestive linkage (a logarithm of odds score of 1.9 or 2.2) (Lander E and Kruglyak L, 1995). Only loci with highly significant threshold that can be confirmed by independent studies are believed to be true positive results. There are seven loci have reached the threshold for significant linkage to SLE (Table 2) (Kelly JA et al., 2002) and all loci (except 17p13) have been confirmed in at least one independent cohort for linkage to SLE (Nath SK et al., 2004)

Table 2 Established linkages (LOD ≥ 3.3) with Systemic Lupus Erythematosus

Chromosome	Location	Markers	LOD	collection
1	1q22-23	FcγRIIA	4.30	OMRF1B
1	1q41-42	D1s2860- D1s213	3.30	UCLA
2	2q37	D2s125	4.24	UPP 1
4	4p16	D4S2366	3.62	OMRF1B
6	6p21-p11	D6S426	4.19	UMN1 & 2
16	16q13	D16s415	3.85	UMN1 & 2
17	17p13	D17s974- D17s1298	3.64	OMRF1C

UCLA = A study conducted at University of California, Los Angeles

UMN = A study conducted at University of Minnesota

OMRF = A study conducted at Oklahoma Medical Research Foundation

USC = A study conducted at University of Southern California

UPP = A study conducted at University of Uppsala

2. Human association studies

A classical association (or candidate gene) study is the assessment of alleles or haplotypes, or any DNA polymorphisms within gene of interest in the difference of frequencies between affected patients and appropriate controls. A polymorphism observed with greater than expected frequency in affected individuals either suggests that the polymorphism being measured is the actual disease causing allele, or one located very closely to the responsible gene. Several candidate genes within confirmed susceptibility loci have been studied and found to be associated with SLE

(Table 3). In addition, there are other SLE associated genes but not confirmed in the SLE-linked region as shown in Table 4. Mostly, they are the genes involving in apoptosis such as *FAS, FASL, BCL2* and *PDCD-1*; the genes that play an important role in the regulation of immune system including *HLA, CTLA-4, TNF, TNFSF4, IL-10, STAT4* and *PTPN22*; the genes that implicate in clearance of immune complex such as *FcG Receptor, C4, MBL* and *CRP*. Since the breakthrough discovery of the central role of type I interferon in SLE pathogenesis from expression microarray, candidated genes within type I interferon (IFN) pathway have been proved to be significantly associated with SLE risk such as *TYK2* and *IRF5*. However, many of the genes implicated in SLE have not been confirmed and remain controversial.

Even if genome-wide linkage analysis using microsatellite markers has been successful in the identification of numerous complex disease loci, high-density single-nucleotide polymorphism (SNP) maps can provide greater information (Dunn G et al., 2005). In addition, most studies in candidate gene approach using a few single nucleotide polymorphisms (SNPs) which are not covered the whole gene or relevant haplotype block. With the development of genotyping platforms that permit analysis of hundreds of thousands of SNPs, it is now possible to apply this principle of indirect association to the whole genome rather than just candidate genes or candidate linkage regions. Recently, three genome wide association (GWA) studies in Caucasian patients with SLE from international collaborations have been reported (Harley JB et al, 2008; Geoffrey Hom et al., 2008; Kozyrev S et al., 2008). Results from these studies confirm known candidate genes including *HLA, FCGR, PTPN22, STAT4* and *IRF5*. Moreover, several new candidate genes have been discovered as summarized in Table 5. For instance, the discovery of *BLK* and *BANK1* genes emphasize the crucial role of B cell in pathogenesis of SLE. Another novel gene namely, *ITGAM* was also identified which is an adhesion molecule that regulates leukocyte adhesion to endothelial cells and may contribute to vasculitis in patients with SLE. From these reports of international collaborations found that odd ratios (OR) were rather low ranged between 1.25 and 2.36. Newly, a genome-wide association study has found the association of three independent SNPs in the *TNFAIP3* region with SLE (OR=1.3-2) (Musone SL et al., 2008). This data highlight the inflammatory role of TNF pathway in the pathogenesis of SLE.

Furthermore, a recent genome wide association (GWA) study in Asian population demonstrated the important role of *ITPR3* (OR=3.39) (Oish T et al., 2008) and *TNXB* (Kamatani Y et al., 2008) on chromosome 6p21 to be candidate gene susceptible to SLE in the Japanese population. Nevertheless, these are the studies restricted in Caucasian patients and only one study in Japanese, the exploration of SLE susceptibility genes in our population is still required.

In GWA study, to balance benefit of genome-wide screening and the costs in genotyping, DNA pooling approach has been proposed (Sham P et al., 2002). Several groups have applied the technology from Affymetrix (microarray-based GeneChip® Mapping arrays) to assess the reliability and validity of SNP allelic frequency measurements as determined from pooling genomic DNA samples on SNP mapping arrays. They have suggested that genotyping of SNPs with DNA pooling using Affymetrix microarrays produces highly accurate results and can be used for genome-wide association studies (Butcher LM et al, 2004; Meaburn E et al., 2005; Craig DW et al., 2005; Simpson CL et al., 2005; Kirov G et al., 2006; Meaburn E et al., 2006). Several applications of DNA pooling technique for genome-wide association studies have been reported (Butcher LM et al, 2005; Johnson C et al., 2006; Steer S et al., 2007; Jongjaroenprasert W et al., 2008). Nevertheless, although several studies have been completed, the pooling approach has the intrinsic limitations such as reduced power, loss of individual genotype data and difficulties ensuring equimolar representation of samples (McCarthy MI et al., 2008).

Table 3 Candidate genes within confirmed susceptibility regions associated with SLE

Gene	Location	Associated allele	Statistical significance[a]	References
FcGR2A	1q23	R131	1.30 (1.1-1.52), p=0.0016 2.6 (1.3-5.2), p=0.002 2.01 (1.28-3.14), p=001	Karassa FB et al., 2002 Lee HS et al., 2003 Hirankarn N et al., 2006
FcGR3A	1q23	F176	$\chi 2$=9.87, p<0.01 1.2 (1.06-1.36), p=0.006	Wu J et al., 1997 Karassa FB et al., 2003
CRP	1q23	CRP4A CRP (GT)16	$\chi 2$=3.81, p=0.05 $\chi 2$=3.82, p=0.05	Russell AI et al., 2004 Russell AI et al., 2004
FASL	1q23	–844C -844 CC	p=0.024 OR=1.53, p=0.014	Wu J et al., 2003 Chen JY et al., 2005
PDCD-1	2q37	PD-1.3A	2.6 (1.6–4.4), p=0.00001 2.73 (1.35-5.56), p=0.0019	Prokunina L et al., 2002 Velázquez-Cruz R et al., 2007
HLADR2, DR3	6p21	DR2/DR3	$\chi 2$=35.0/76.0, p<0.0005 1.75(1.4-2.19), p=0.000001 /2.02(1.44-2.83), p=0.0004	Hartung K et al., 1992 Castaño-Rodríguez N,2008
TNFa	6p21	-308A -238A -308A -308A -238A -863A allele and -863A,-308G,238G haplotype	3.7 (2.24–6.11), p<0.05 OR=3.62, p=0.02 2.3 (1.4-3.9), p=0.001 2.6 (1.77–3.83), p < 0.0001 0.04 (0-0.28), p<0.0001 1.85 (1.21-2.83), p=0.009	van der Linden MW et al., 2001 Zúñiga J et al., 2001 Parks CG et al, 2004 Correa PA et al., 2005 Correa PA et al., 2005 Hirankarn N et al., 2007
TNFb	6p21	TNFB*2	RR=3.4, p<0.0001 p<0.05	Kim TG et al., 1996 Takeuchi F et al., 2005
C4	6p21	AQ0	$p<10^{-6}$	Hartung K et al., 1992

[a] Odd ratio (OR) with 95% confidence interval, relative risk (RR) if available or x2 with p value.

Table 4 Other candidate genes associated with SLE

Gene	Location	Associated allele	Statistical significance[a]	References
TNFSF4	1q25.1	rs10912580 (G) rs844644 (A) rs2205960 (T) rs1234317 (T)	$\chi 2$=3.86, p=0.05 $\chi 2$=15.9, p=6.8 x 10^{-5} $\chi 2$=7.38, p=7 x 10^{-3} $\chi 2$=4.44, p=0.04	Graham DS et al., 2008
IL10	1q31-32	-1082A, -819T and -592A haplotype -1082G allele -2763AA genotypes -1082A, -819C and -592C haplotype	OR=1.9, p=0.02 $\chi 2$=5.85, p<0.05 P<0.05 1.47(1.02–2.12),p=0.03	Rood MJ et al., 1999 Khoa PD et al., 2005 Gibson AW et al., 2001 Hirankarn N et al., 2006
PTPN22	1p13	1858T allele (W620)	1.49(1.28–1.75), p <0.00001	Lee YH et al., 2007
STAT4	2q32	rs7574865 (T allele)	1.55(1.34-1.79), p=1.87x10^{-9}	Remmers EF et al., 2007
CTLA-4	2q33	+49G	1.72(1.2-2.4), p=0.003 1.23 (1.08–1.41), p=0.002	Ahmed S et al., 2001 Lee YH et al., 2005
IRF5	7q32	Intron -3835T (rs2004640) TInCA haplotype from rs2004640, In/Del exon6, rs2070197, rs10954213	p=4.4 x 10^{-16} $\chi 2$=24.457, p=5.72 x 10^{-6}	Graham RR et al., 2006 Kozyrev SV et al., 2007
MBL	10q11.2-q21	codon 54 B allele	1.406 (1.221-1.608), p=0.001	Lee YH et al., 2005
FAS	10q24	297C/416G -670A	RR=5, p= 0.01 p=0.004	Horiuchi T et al., 1999 Kanemitsu S et al., 2002
Bcl-2	18q21	Multiple alleles	$\chi 2$=34.95, p=0.0001	Mehrian R et al., 1998
TYK2	19p13.2	Exon (V/F) 13430 Exon (I/S) 19107	p= 5.60x10^{-5} p=1.50x10^{-4}	Sigurdsson S et al., 2005

Table 5 SLE susceptibility genes identified from genome wide association studies

Gene	Locus	SNP	P value	OR	Potential function
PTPN22 [a]	1p13.2	rs2476601	0.0000052	1.53	Inhibition of lymphocyte activation
FCGR2A [a]	1q23	rs1801274	6.78E-07	0.74	Clearance of immune complexes
STAT4 [b]	2q32.3	rs7574865	9E-14		Modulation of the production of cytokines
					in T cells and natural killer cells;
					activation of response of macrophages to interferon-α
PXK [e]	3p14.3	rs6445975	7.1E-09	1.25	Unknown effect of serine–threonine kinase
BANK1 [c]	4q24	rs10516487	0.0064		B-cell adaptor protein
HLA [a,b]	6p21.32-33	rs2187668	3E-21		Presentation of antigen
		rs1270942	1.71E-51	2.35	
		rs3131379	1.71E-52	2.36	
ITPR3 [d]	6p21.31	rs3748079	2.87×10^{-8}	3.39	Inositol 1,4,5-triphosphate receptor type 3 (ITPR3);
					plays a crucial role in the regulation of apoptosis signaling in T lymphocytes
TNXB [e]	6p21.32	rs1009382	0.00000518		tenascin XB isoform 1;
		rs3130342	0.000000930	3.11	regulator of collagen deposition by dermal fibroblasts
TNFSAIP3 [f]	6q23.3	rs13192841	5.4×10^{-6}	1.4	encodes a ubiquitin editing enzyme, A20, that restricts
		rs2230926	3×10^{-4}	2	NF-*j*B–dependent signaling and prevents inflammation
		rs6922466	1×10^{-4}	1.3	
IRF5 [a,b]	7q32.1	rs10488631	2E-11		Production of interferon-α
		rs729302	2.00E-10	0.78	
		rs10279821	6.50E-09	0.8	
		rs12537284	3.61E-19	1.54	
c8orf13-BLK [b]	8p23.1	rs2736340	0.0000004	1.37	unknown (c8orf13)-Activation of B cells (BLK)
		rs13277113	0.00000008	1.39	
KIAA1542 [e]	11p15.5	rs4963128	3.00E-10	0.78	Linkage disequilibrium with IRF7;
					production of type I interferon
ITGAM-ITGAX [b]	16p11.2	rs9937837	0.0000007	1.28	Adhesion of leukocytes to endothelial cells
		rs11574637	0.0000005	1.3	
ITGAM [e]	16p11.2	rs9888739	1.61E-23	1.62	Adhesion of leukocytes to endothelial cells
		rs1143678	8.50E-14	1.4	
		rs4548893	2.36E-12	1.34	
		rs1143679	1.70E-17	1.78	

[a] Harley JB et al, 2008, [b] Geoffrey Hom et al., 2008, [c] Kozyrev S et al., 2008, [d] Oishi T et al., 2008, [e] Kamatani Y et al., 2008 and [f] Musone SL et al., 2008

3. Genetic studies in mouse models

The studies in mouse models have been used to further understanding of lupus. There are two mouse models of lupus comprising synthetic and spontaneous mouse models (Table 6 and 7) (Nguyen C et al., 2002). The synthetic ones consist of both transgenic and knock out mouse model of various apoptosis genes implicated in SLE pathogenesis. For example, *Fas*, *FasL* and *Bcl2* were silenced or overexpressed, resulting in the deregulation of apoptosis (Watanabe-Fukunaga R et al., 1992; Takahashi T et al., 1994; Strasser A et al., 1991). In addition to synthetic mouse models, there are also classic spontaneous models, including the (NZB × NZW)F1 (or NZB/W) mouse and the congenic recombinant NZM2410 strain derived from this cross, the MRL/*lpr* mouse, and the BXSB/*yaa* mouse. These strains were subjected to linkage analysis in order to pinpoint the chromosomal regions linked with SLE susceptibility. So far, more than 50 loci have been localized. Some confirmed loci on chromosomes 1, 4, 7 and 17 were shown in Table 8. (Nguyen C et al., 2002; Kono DH and Theofilopoulos AN, 2006). Results from mouse models are useful in identifying novel genes responsible for human SLE.

Table 6 Synthetic murine models of lupus

Model	Affected function
Fas, FasL, *Bcl2*	Regulation of apoptosis
Sap, C1q, C4, DNAse	Clearance of antigen, such as apoptotic bodies and DNA
Ctla-4, p21,PD-1, Lyn, Fyn	Activation and regulation of T cells
BLyS, PD-1, Lyn, Fyn, FcγRIIB	Activation and regulation of B cells
FcγRIII, ICAM-1	Proinflammatory mechanisms

BLyS, B-lymphocyte stimulator; FasL, Fas ligand; ICAM-1, intracellular adhesion molecule-1.

Table 7 Spontaneous murine models of lupus

Model	Properties
(NZB × NZW)F1, NZM2410, (SWR × NZB)F1	Develops antinuclear autoantibodies and glomerulonephritis that resembles human lupus; exhibits a complex inheritance
MRL/*lpr*, MRL/*gld*	Contains a single-gene mutation (Fas or Fas ligand) that leads to autoimmunity when expressed in MRL background
BXSB/*yaa*	Contains the Y-linked autoimmune accelerator gene that causes a more severe disease in BXSB males

Table 8 Loci confirmed in interval-congenic mice

Locus	Chromosome	Susceptible strain	Major autoimmune trait	Reference
Sle1	1	NZM2410 (NZW)	Antinucleosome Ab	Morel L et al., 2001; Boackle SA et al., 2001; Wandstrat AE et al., 2004
FCgnz1	1	NZM2328 (NZW)	Antichromatin Ab, Glomerulonephritis (GN)	Waters ST et al, 2004
-	1	129	Antinuclear Ab, GN	Bygrave AE et al., 2004
Nba2	1	NZB	B cell hyperactivity, autoAb	Atencio S et al., 2004; Wither JE et al., 2003
Bxs1-4[a]	1	BXSB	Anti-dsDNA Ab, GN (4 overlapping intervals)	Haywood ME et al., 2004; Haywood ME et al., 2006
Sle2	4	NZM2410 (NZW, NZB)	B cell hyperactivity, incr. B1 cells	Xu Z et al., 2005
Adnz1	4	NZM2328	Glomerulonephritis	Waters ST et al., 2004
Lbw2	4	NZB	B cell hyperactivity	Haraldsson MK et al., 2005
Sle3/Sle5	7	NZM2420 (NZW)	T cell hyperactivity from hyperstimulatory antigen-presenting cells	Sobel ES et al., 2002; Zhu J et al., 2005
Nba5	7	NZB	Anti-gp70 Ab, GN	Kikuchi S et al., 2005
Sles1	17	NZW	Suppression of lupus-like disease	Subramanian S et al., 2005

[a] Four different congenic strains with chromosome 1 intervals containing one or more of the four Bxs loci;

Interferon-inducible p200 (IFI200) family

IFI200 family genes

The *IFI200* gene cluster is located on mouse and human chromosome 1q21-23 (Figure 1). The murine IFI200 family includes *IFI202 to IFI210*, while human members comprise *MNDA* (Myeloid Nuclear Differentiation Antigen) (Briggs JA et al., 1992), *IFIX* (IFN-inducible protein X) (Ding Y et al., 2004), *IFI16* (Interferon-Inducible protein 16) (Trapani JA et al., 1992), and *AIM2* (Absent In Melanoma 2) (DeYoung KL et al., 1997).

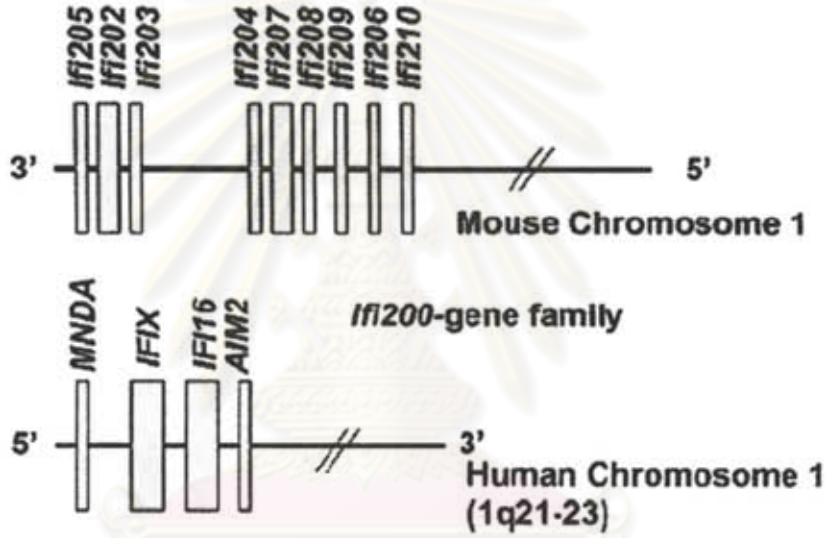


Figure 1 The relative location of IFI200-family genes (Choubey D et al., 2008)

Human *IFI200* genes and transcripts (*MNDA, IFIX, IFI16 and AIM2*)

MNDA

MNDA resides on chromosome 1q23.1. This gene consists of 7 exons and 6 introns spanning at least 18 kilobases (kb) of DNA. The 1.8 kb mRNA encode a protein of 408 amino acid (Briggs RC et al., 1994). Structure of *MNDA* gene was shown in Figure 2.

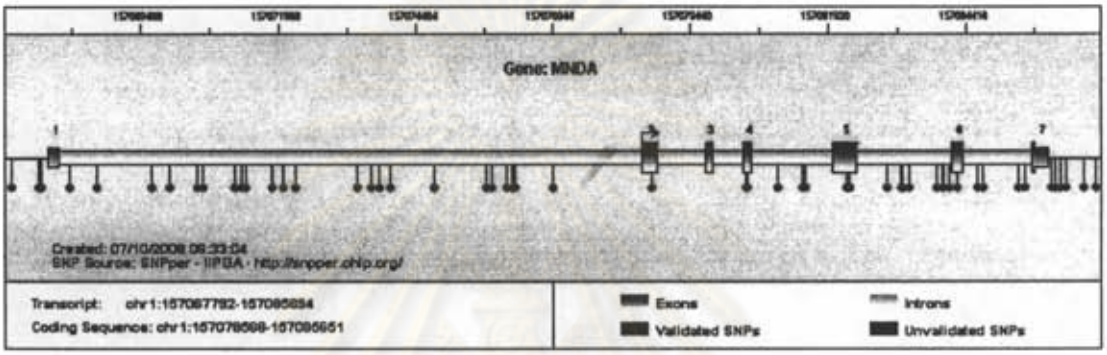


Figure 2 Structure of MNDA gene

(http://snpper.chip.org/bio/snpplet/1721/utga574229)

IFIX

IFIX is a new member of IFI200 family and locate on chromosome 1q23.1. The gene contains 9 exons and 8 interventing introns spanning approximately 45 kb of DNA. There are at least six IFIX isoforms including IFIXα1, α2, β1, β2, γ1, and γ2 encoded by alternatively spliced mRNAs (Figure 3a). The mRNAs encode each protein as follows: α1, 492 aa, α2, 483 aa, β1, 461 aa, β2, 452 aa, γ1, 246 aa, and γ2, 237 aa. The 27 base pairs (9 amino acids) of exon 3 are absent in isoforms α2, β2, and γ2 (Figure 3a and b). The C-termini of α, β, and γ isoforms are diverse due to alternative splicing (Figure 3c and d).

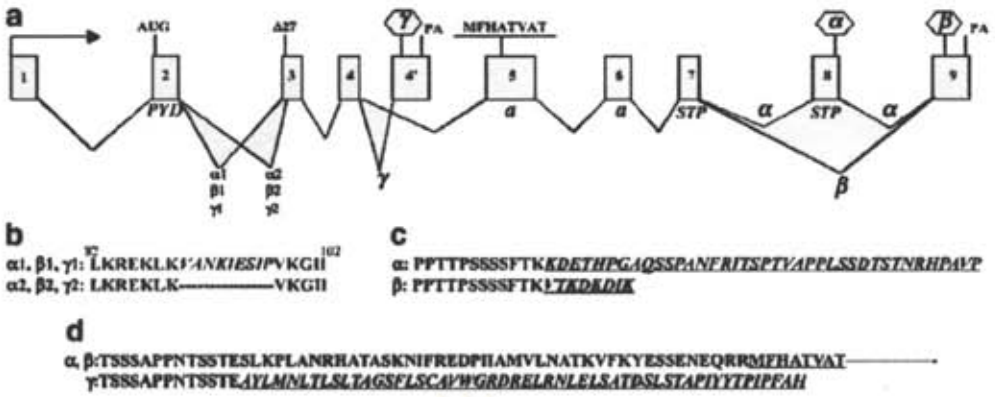


Figure 3 Structure of IFIX (a) gene and (b-d) isoforms (Ding Y et al., 2004)

IFI16

IFI16 gene maps to chromosome 1q23.1. It contains 11 exons and 10 introns spanning at least 45 kb of DNA. The 11 exons include exon 7a arisen from tandem duplication of exon 7. This gene is known to encode three isoforms including IFI16A, B and C via alternative splicing of mRNA. The longest mRNA (~2.7 kb) encodes an open reading frame of 2355 bp and generates the IFI16A isoform of 785 amino acids. The second isoform IFI16B arise from the lack of exon 7a (168 bp) to encode a protein of 729 amino acids. The smallest IFI16C isoform (2019 bp) has deleted both exon 7 and exon 7a to encode 673 amino acids. The nucleotide sequences of IFI16A, -B, and -C splice variants and the three different IFI16 mRNA isoforms were shown in Figure 4A and B (Johnstone RW et al., 1998).

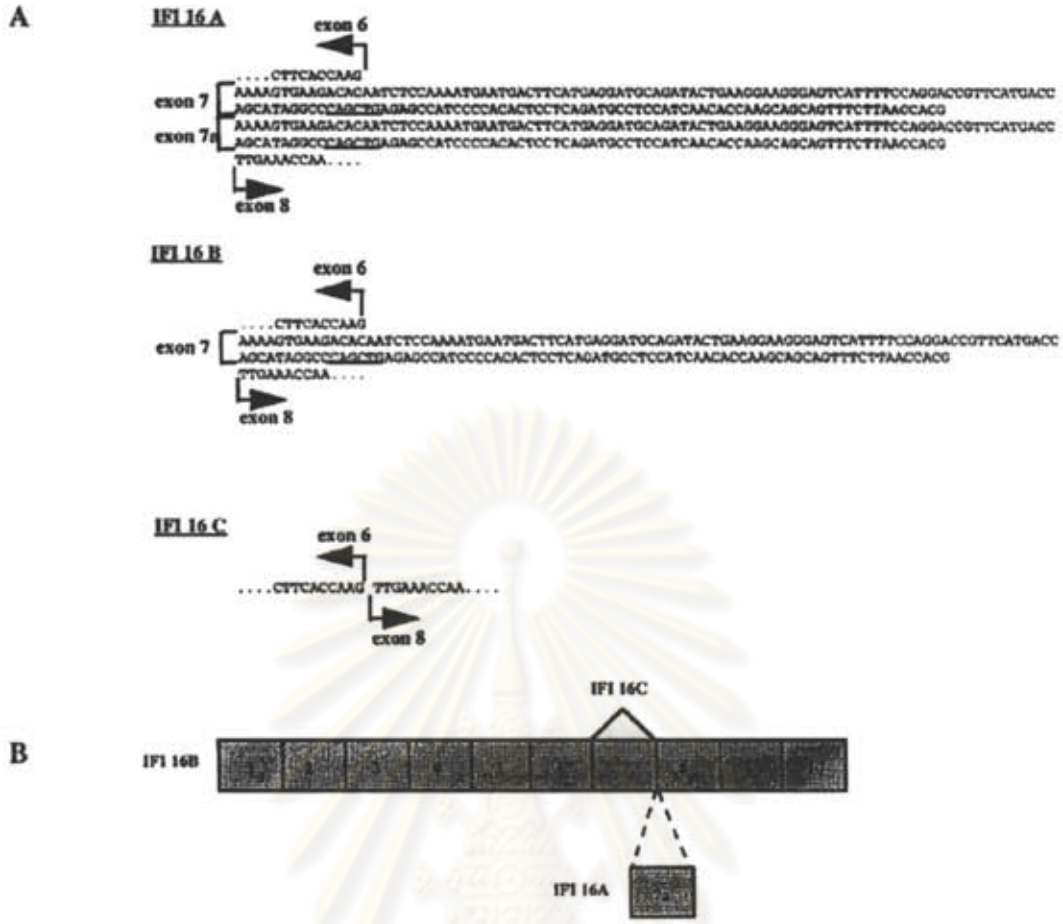


Figure 4 Nucleotide sequences of IFI 16A, -B, and -C splice variants (A) and schematic representation of the three different IFI16 mRNA isoforms (B)

AIM2

AIM2 also maps on chromosome 1q23.1. It has 6 exons and 5 interventing introns. The length of DNA is approximately 14 kb and mRNA ~1.4 kb encodes a protein of 344 amino acids. Structure of *AIM2* gene was shown in Figure 5.

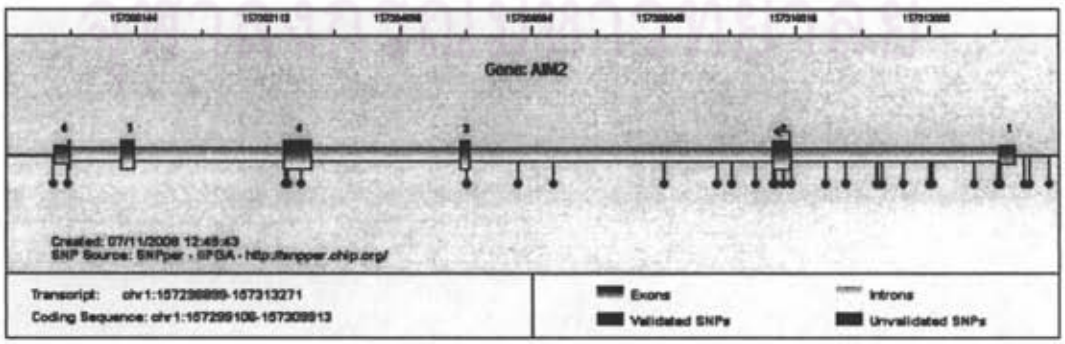


Figure 5 Structure of *AIM2* gene

(http://snpper.chip.org/bio/snpplet/1727/ihid583464)

Protein structure of IFI200 family

A structural motif found in all members (except p208 and IFIXγ) of mouse and human is a 200-amino-acid domain present either one or two copies. There are two adjacent 200-amino-acid domains (A and B) in p202 and p204, whereas the two copies of IFI16 are separated by one, two, or three copies of a highly conserved 56-amino-acid S/T/P-rich spacer region. The size of the spacer region is regulated by alternate mRNA splicing which result in three isotypic variants IFI16A, IFI16B and IFI16C (Johnstone RW et al., 1998). On the contrary, MNDA, IFIXα, IFIXβ, AIM-2, p203, p205, p206, p207, p209 and p210 contain only one 200-amino-acid domain. Within this domain contains the conserved MFHATVAT and the LXCXE pRb-binding motifs, which could mediate protein–protein interactions (Koul D et al., 1998; Choubey D and Lengyel P, 1995). The 200-amino-acid domain can be divided into three distinct subclasses A, B and C based on a broader range of motifs (Ludlow LEA et al., 2005). The sequence identity between 200-amino-acid domains ranges approximately from 24% to 88%. Nevertheless, the identity of amino acid is much greater between members of the same subclass than between members of different subclasses (Figure 6B). In addition to 200-amino-acid domains, the N-terminus of all IFI-200 proteins except p202 have PAAD/DAPIN/Pyrin domain. This domain mediates protein–protein interaction involved in apoptotic (NF-κB) and inflammatory (Caspase-1) signaling pathways (Reed JC et al., 2003). Sequence alignment of the PAAD/DAPIN/Pyrin domains in IFI200 proteins show that they could be sub-divided as 2 subclasses because the amino-acid sequences of AIM2 proteins are quite different from other IFI200 proteins. The amino-acid identity of this domain among IFI200 members ranges from 14%-100%. However, the identity within each subclass is approximately 40%-100% (Figure 6A). Schematic structural representation of IFI200 proteins was shown in Figure 7A and 7B.

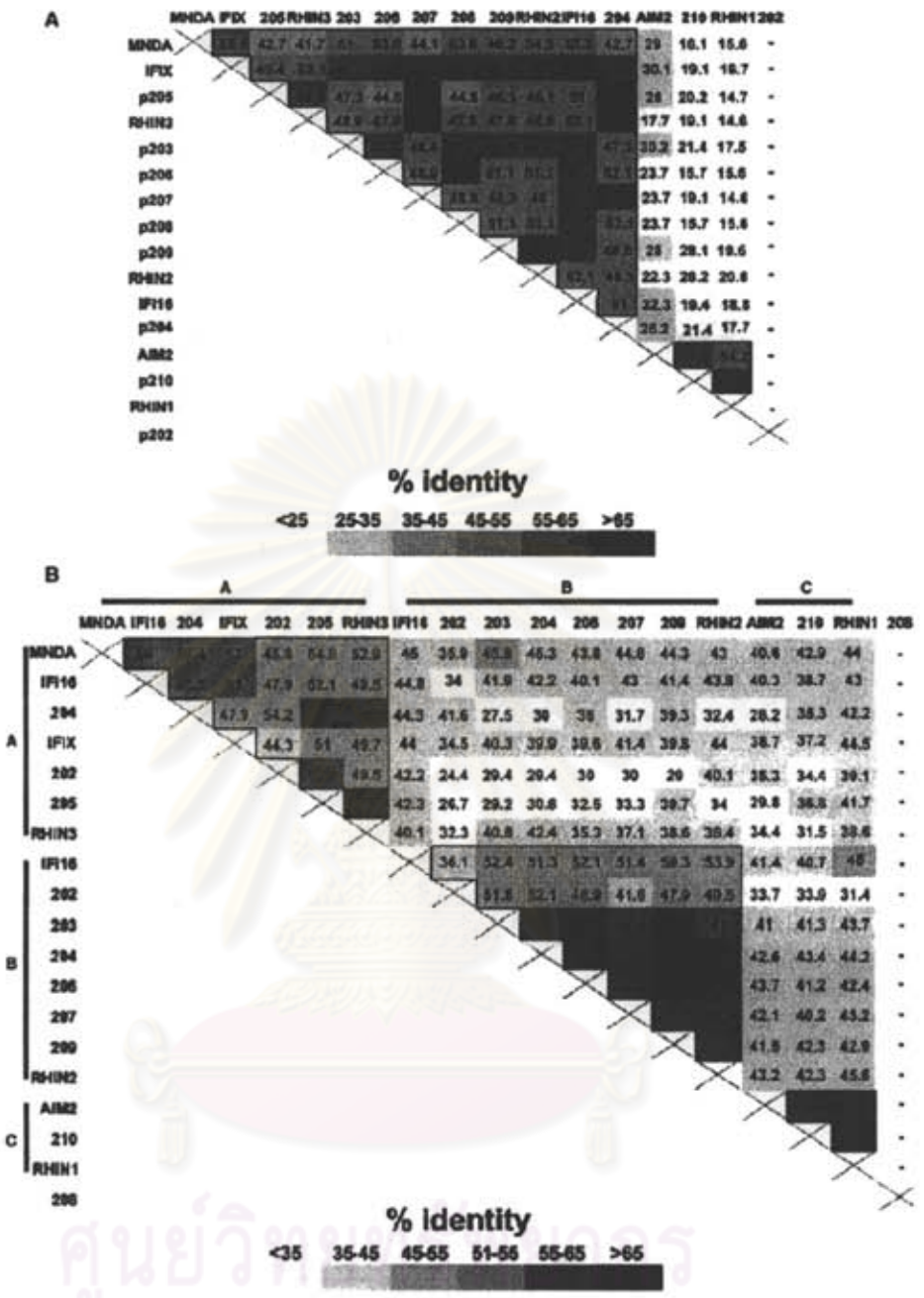


Figure 6 Percent identity of PAAD/DAPIN/Pyrin domains (A) or 200-amino-acid domains (B) among IFI200 family members (Ludlow LEA et al., 2005)

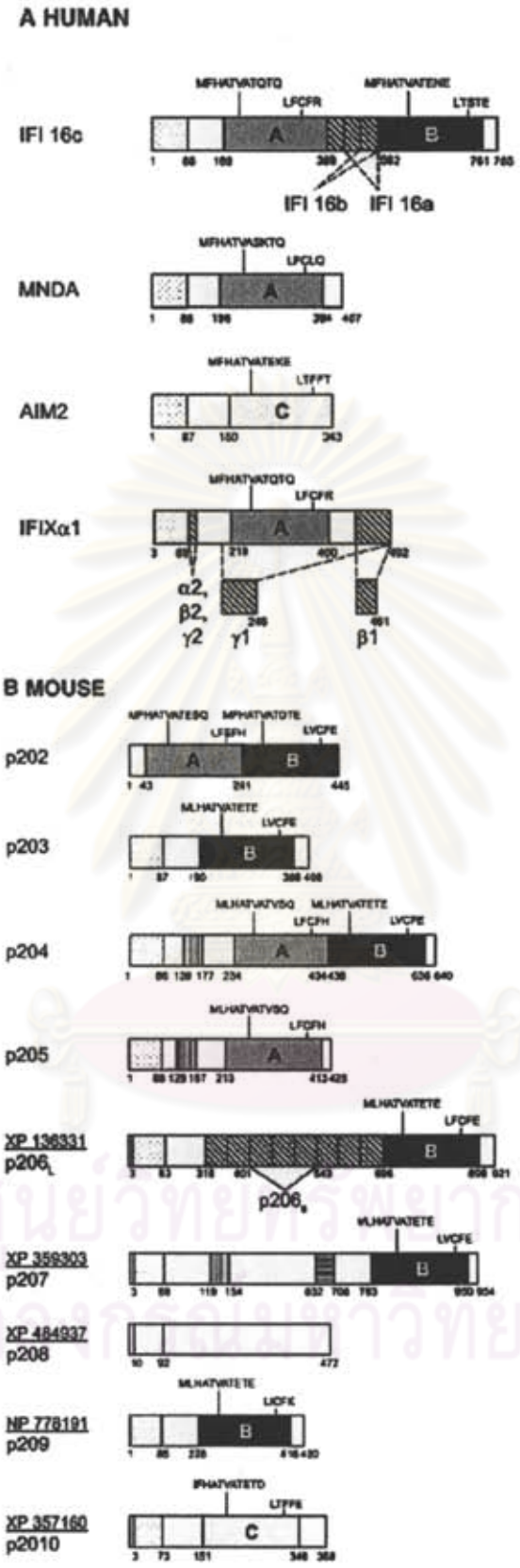


Figure 7 The structure of human (A) and murine (B) IFI200 proteins (Ludlow LEA et al., 2005)

Cellular localization IFI200 proteins

IFI200 proteins, for instance, p203, p204, p205, MNDA, IFIX and IFI 16 have been found to localize normally in the nucleus (Asefa B et al., 2003; Choubey D et al., 1992; Dermott JM et al., 2004; Duhl DM et al., 1989; Ding Y et al., 2004; Dawson MJ et al., 1995). Since they contain nuclear localization signal (NLS) to enhance nuclear import via facilitating binding to nuclear components, while p202 and AIM2 proteins do not present this motif (Johnstone RW et al., 1999; Ding Y et al., 2004). Consistent with the lack of NLS, p202 was primarily detected in cytoplasm but translocate to the nucleus after interferon treatment (Choubey D et al., 2003). AIM2 was also found in the cytoplasm of murine AKR-2B cells (Choubey D et al., 2000), but the latter study has demonstrated localization of IFN-induced AIM2 in nucleus (Cresswell KS et al., 2005). Furthermore, the newly murine p206 protein has been found the deficiency of NLS and commonly localized in the cytoplasm. Hence, p206 is the first IFI200 member identified that is exclusively expressed in the cytoplasm (Ludlow LE et al., 2008).

Expression of IFI200 family

IFI200 transcripts and proteins have been detected in both hematopoietic and non-hematopoietic origins (Table 9). Murine p202 is highly expressed in ovary, thymus, spleen and liver (Wang H et al., 1999), but low in brain, kidney, lung and testis (Wang H et al., 1999; Choubey D et al., 2000). In addition, splenic cell subsets (T, B and non-T/non-B cells) of B6.Nba2 mice highly expressed IFI202 mRNA, as compared to B6 control mice (Rozzo SJ et al., 2001). p203 is constitutively expressed in some myeloid and lymphoid tissues, namely, thymus, bone marrow, and spleen and non-hematopoietic tissue such as liver (Gribaudo G et al., 1999; Zhang Y et al., 2008). p204 and p205 are predominantly expressed in myeloid cells, particularly in monocytes or macrophages and granulocytes (Gariglio M et al., 1998; Weiler SR et al., 1999). High level of p204 expression was detected in heart and it can also be detected in skeletal muscle and kidney (Liu C et al., 2000). In contrast, p205 is low expressed in heart and skeletal muscle, but moderately in lung (Weiler SR et al., 1999). The mRNA of newly murine IFI206 was detected only in the spleen and lung of BALB/c and C57BL/6 mice

(Ludlow LE et al., 2008). In human IFI200, MNDA has only been detected in the hematopoietic origin including mature granulocytes, monocytes, activated macrophages and B cell from chronic lymphocytic leukemia (CLL) patients (Miranda RN et al., 1999; Dawson MJ et al., 1998; Joshi AD et al., 2007). IFIXα expression is restricted to secondary lymphoid organs including spleen, lymph node, and peripheral blood leukocyte. In non-hematopoietic tissues, IFIXα mRNA was not found in adult brain, heart, skeletal muscle, colon, kidney, liver, small intestine, placenta, and lung; however it could be detected in breast tissue (Ding Y et al., 2004). In contrast to MNDA, IFI16 is not restricted to hematopoietic origin but is much more widely expressed. IFI16 mRNA could be highly found in small intestine, ovary, prostate and testis, modest levels in pancreas, liver, brain and heart, and low levels in lung, placenta, colon, kidney and skeletal muscle. For hematopoietic compartment, IFI16 mRNA was abundant in spleen, thymus, peripheral blood leukocytes, CD34+ progenitor cells, mature lymphocytes and monocytes, but is absent from granulocytes, erythrocytes and megakaryocytes (Wei Wu et al., 2003; Dawson MJ et al., 1998). In addition, the study of IFI16 protein expression using anti-IFI16 monoclonal antibody has showed that IFI16 was detected in the nuclei of lymphocytes in the spleen, thymus, lymph node and palatine tonsil and also found in epithelial cells in these tissues. Yet, IFI16 protein was expressed in epithelial cells of the skin, gastrointestinal tract, urogenital tract and glands and ducts of breast tissue, but was not in heart and brain. In addition to epithelial cells, IFI16 expression is also seen in vascular endothelial cells from blood and lymph vessels (Mondini M et al., 2006; Wei Wu et al., 2003; Gariglio M et al., 2002). Another human IFI200 member, AIM2 was detected in both origins including spleen, peripheral blood leukocytes, small intestine and testis (DeYoung KL et al., 1997).

Table 9 The expression of IFI200 transcripts and/or proteins

IFI200	Hematopoietic origin	non-hematopoietic origin
IFI202	spleen, thymus, splenic B, T and non-T/non-B cells	ovary, liver, heart, brain, kidney, lung, testis
IFI203	spleen, thymus, bone marrow	liver
IFI204	spleen, thymus, bone marrow, lymph nodes, myeloid cells (monocytes, granulocytes and megakaryocytes)	heart, skeletal muscle, kidney
IFI205	spleen, myeloid cells (granulocytes and macrophages)	heart, skeletal muscle, lung
IFI206	spleen	lung
MNDA	mature granulocytes, monocytes, activated macrophages, B-cell chronic lymphocytic leukemia (CLL) cells	absent
IFIX	spleen, lymph node, peripheral blood leukocyte	breast
IFI16	spleen, thymus, peripheral blood leukocytes, PBMC, CD34+ progenitor cells, mature lymphocytes, monocytes	small intestine, ovary, prostate, testis, pancreas, liver, brain, heart, lung, placenta, colon, kidney, skeletal muscle, skin, vascular endothelial cells
AIM2	spleen, peripheral blood leukocytes	small intestine, testis

Biological functions of IFI200 family

Role of IFI200 family in cell differentiation

Several studies have reported that the IFI200 family proteins play an important role in the differentiation of certain cell types. For example, p202 hugely increased during the differentiation of cultured C2C12 myoblasts to myotubes together with decreasing of MyoD expression. Such decreasing is the consequence of inhibition of its sequence-specific binding to DNA by p202 (Datta B et al., 1998). Moreover, another family member, p204 enables the differentiation by overcoming the inhibition of the activities of MyoD, E12/E47, and other myogenic basic region helix-loop-helix (bHLH) transcription factors by Id proteins Id1, Id2, and Id3 (Liu C et al., 2002). In addition, p204 is required for the differentiation of murine P19 embryonal carcinoma stem cells to beating cardiac type myocytes (Ding B et al., 2006). In contrast to murine IFI200, human IFI16 has been reported that it was not essential for differentiation of medullary thyroid carcinoma cells induced by leukemia inhibitor factor (Kim EJ et al., 2005).

The role of IFI200 family in cell survival and apoptosis

Over-expression of IFI200 proteins in various cells is known to decrease cell proliferation and inhibit cell cycle progression at the G1–S phase transition (Lembo D et al., 1995; Yan DH et al., 1999; Lembo M et al., 1998; Dermott JM et al., 2004; Ding Y et al., 2004; Raffaella R et al., 2004). The majority of studies indicate that these proteins inhibit cell proliferation in part through the Rb/E2F and p53/p21 pathways. For instance, IFI16 expression inhibited cell cycle progression of primary human umbilical vein embryo cells (HUVECs) accompanied with up-regulation of p53, p21, and pRb, but not in HPV16 E6/E7-immortalized HUVECs (Raffaella R et al., 2004). Furthermore, knockdown of IFI16 protein in medullary thyroid carcinoma cell line resulted in up-regulation of E2F1, cyclin D1, and down-regulation of $p21^{CIP1}$ and abrogated cell cycle arrest (Kim EJ et al., 2005). In contrast, one study indicated that IFI16 is negative

regulator of p53 and $p21^{WAF1/CIP1}$. siRNA-mediated IFI16 knockdown in osteosarcoma cell line activated a G1–S checkpoint which accompanied with increasing of p53, $p21^{WAF1/CIP1}$ and decreased pRb phosphorylation. Moreover, IFI16 inhibited p53-mediated activation of the $p21^{WAF1/CIP1}$ promoter (Kwak JC et al., 2003). However, a consistent study has reported that p202 could indirectly interact with p53 through human p53-binding protein1 (53BP1) and overcome growth inhibitory activity of p202 in yeast (Datta B et al., 1996). On the other hand, p202 is a growth inhibitor in prostate cancer cells and also retard proliferation of AKR-2B fibroblasts accompanied with an increased $p21^{WAF1/CIP11}$ (Yan DH et al., 1999; Gutterman JU et al., 1999). The reasons for these incongruous results are still unclear. Another mouse IFI200 member, p204 has been reported that its growth inhibition in murine embryonic fibroblasts requires retinoblastoma protein (pRb) (Hertel L et al., 2000).

Recently, a new human IFI200 member, IFIX has been discovered by using the p202a amino-acid sequence to query human DNA databases. The expression of IFIX was associated with growth suppression, loss of tumorigenicity, and $p21^{CIP1}$ upregulation in breast cancer (Ding Y et al., 2004). Consistent with IFIX, over-expression of AIM2 retards proliferation of murine AKR-2B fibroblasts and increases the susceptibility to cell death under reduced serum conditions (Choubey D et al., 2000). Nevertheless, the mechanisms of AIM2 in apoptosis remain unclear. The biological functions of certain IFI200 family members in cell cycle regulation were shown in Figure 11.

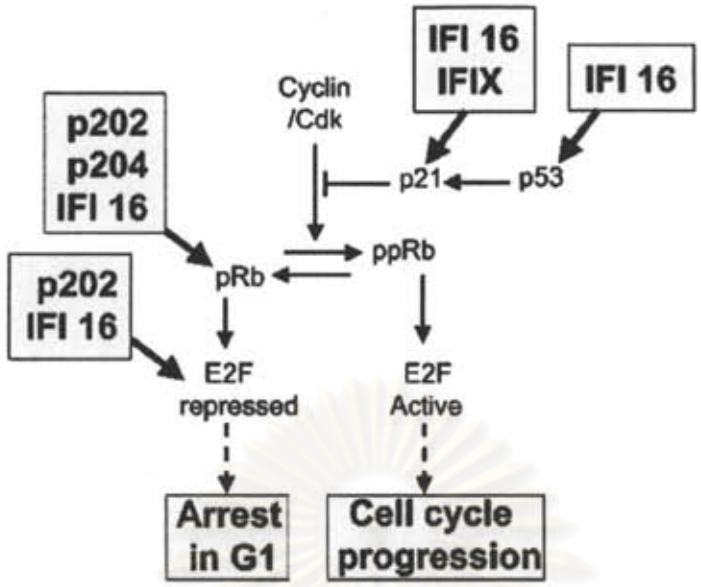


Figure 8 The functions of IFI200 family in cell cycle regulation (Ludlow LEA et al. 2005)

The role of IFI200 family in inflammation

IFI16, a member of IFI200 family is known to express in endothelial and epithelial cells such as skin in addition to hematopoietic cells (Wei W et al., 2003). Its expression is stimulated in HUVECs by oxidative stress and diverse proinflammatory cytokines (Gugliesi F et al., 2005; Mondini M et al., 2007). These data indicated that IFI16 may be implicated in the initial steps of inflammation by modulating endothelial and keratinocyte cell function. Moreover, a recent study by gene array disclosed that IFI16-overexpressing HUVECs showed an increase in expression of genes involved in immunomodulation, cell growth, and apoptosis. Consistently, IFI16 induced the expression of genes encoding adhesion molecules, such as intracellular adhesion molecule-1 (ICAM-1) and E-selectin, or chemokines, such as interleukin-8 (IL-8) and monocyte chemoattractant protein-1 (MCP-1). In addition, knockout of IFI16 using siRNA significantly inhibited the induction of ICAM-1 upon IFN-α treatment. This demonstrates that IFI16 is required for stimulating proinflammatory genes. Furthermore, the functional analysis has indicated that IFI16 regulates proinflammatory genes via NF-ΚB activation using a novel mechanism involving suppression of IΚB expression (Caposio P et al., 2007).

The role of IFI200 family members in SLE

The first evidence for the role of IFI200 family in SLE was provided using mouse model to identify SLE susceptibility loci. One region, namely Nba2 was identified as a locus responsible for mediating susceptibility to SLE (Atencio S et al., 2004; Wither JE et al., 2003). Subsequent study of Rozzo SJ and co-worker has shown that progeny from congenic mice containing the Nba2 susceptibility region on the C57Bl/6 resistant background (B6.Nba2) crossed with NZW mice (B6.Nba2xNZW) developed glomerulonephritis and IgG auto-antibodies against nuclear antigens. Additional study, they identified gene candidates within this Nba2 region with oligonucleotide microarray of wild type and B6.Nba congenic splenocytes. That study has found that only two *IFI*200 family genes were differentially expressed; *IFI202* was up-regulated while *IFI203* was down-regulated. Furthermore, they has revealed the elevated levels of *IFI202* in B cells, which showed increased resistance to apoptosis when the B cell receptor was cross-linked with IgM (Rozzo SJ et al., 2001). Second, IFI16 was identified as new autoantigen for patients with SLE, up to 29% of 374 SLE patients develop high titer autoantibodies to IFI16 protein (Seelig HP et al., 1994). Later study has also found that 26% of patients with SLE exhibited significantly higher anti-IFI16 IgG antibody levels compared with control subjects. By immunohistochemical analysis, IFI16 was highly expressed in epidermis and dermal inflammatory infiltrates of skin lesions from patients with SLE (Mondini M et al., 2006). Lastly, a recent study of gene expression profiles using microarray has found 2.4 fold up-regulated IFI16 in leukocytes of patients with SLE as compared to healthy controls (Alcorta DA et al., 2007). These evidences indicate that a group of IFN-inducible genes are likely to be important factors contributing to SLE.

CHAPTER IV

MATERIALS AND METHODS

Part I Whole genome association study (Pooling)

1. Subjects

One hundred patients with SLE from King Chulalongkorn Memorial hospital were enrolled in this study. All patients fulfilled at least 4 of the American College of Rheumatology (ACR) revised criteria for the classification of SLE (Tan EM et al., 1982). As a control group, we recruited 100 ethnically and geographically matched controls from healthy blood donor of the Thai Red Cross Society. The ethics committee of the faculty of Medicine, Chulalongkorn University, Bangkok, Thailand approved the study and the subjects gave their informed consent.

2. DNA extraction

Genomic DNA was isolated from buffy coat collected with ethylenediaminetetraacetic acid (EDTA) as anticoagulant, using a salting out method (Miller SA et al., 1988). For the genomic DNA extraction, 1 ml of red cell lysis buffer (RCLB) was added to 0.5 ml of buffy coat, vortex for 30 seconds. This solution was centrifuged at 13,000 rpm for 30 seconds and the supernatant was discarded to obtain the pellet. The pellet remaining should be white to pink. This step may be repeated if necessary. To this pellet, 200 µl nuclei lysis buffer (NLB) and 50 µl 10% SDS were added. Pellet was broken up with pipette tip and vortex to get powdery, tiny flakes. The solution, 150 µl of NLB and 10 µl of proteinase K (10 mg/ml in H2O stored frozen) were added, followed by incubation at 65 °C for 2 hours. Precipitation of proteins was obtained by adding 175 µl of 5.3 M NaCl. This solution was centrifuged at 13,000 rpm for 15 minutes in micro-centrifuge. After

centrifugation, the DNA in the supernatant was precipitated in 1 ml of cold absolute ethanol. Invert 6-10 times to precipitate DNA, it will appear as a white to translucent stringy mass. This solution was centrifuged at 13,000 rpm for 2 minutes and the supernatant was discarded to obtain the pellet. This pellet was resuspended in 1 ml of cold 70% ethanol (break pellet by tapping), followed by centrifugation 2 minutes at 13,000 rpm and the supernatant was discarded to obtain the pellet. After removal of the ethanol, the pellet was dried at 37 °C with the cap open to evaporate the ethanol. This pellet was dissolved in 100 µl of sterile distilled water, followed by incubation at 65 °C until completely dissolved.

3. DNA quantification and pool construction

Genomic DNA of each individual was quantified in duplicate using the Quant-iT™ Broad-Range DNA Assay Kit (Invitrogen) and diluted to a target concentration of 50 ng/µl. Equal amounts of each individual's DNA were randomly used to construct 10 pools in each study group. To minimize error from pipetting, all steps contained the volumes greater than 2 µl. These 10 pools of each group came from 100 patients with SLE and 100 healthy controls (i.e.10 individuals per pool).

4. SNP microarray genotyping of pooled DNA

Twenty DNA pools were genotyped using Affymetrix GeneChip Human Mapping 50K Array Xba240. The chips and reagents were obtained from Affymetrix and the assays were carried out according to the manufacturer's instructions. Briefly, 250 ng of genomic DNA were digested with XbaI and then ligated to adapters that recognize the cohesive four base-pair (bp) overhangs. A generic primer that recognizes the adapter sequence was used to amplify adapter-ligated DNA fragments with PCR conditions optimized to preferentially amplify fragments in the 250–2,000 bp size range in a GeneAmp PCR System 9700 (Applied Biosystems, Foster City, CA). After purification with a Qiagen MinElute 96 UF PCR purification system, a total of 40 µg of PCR product was fragmented and a sample of about 2.9 µg was visualized on a 4% TBE agarose gel to confirm that the average size was smaller than 180 bp. The fragmented DNA was then labeled with biotin and hybridized to

the GeneChip Mapping 50K Set for 17 hr. The arrays were washed and stained using the Affymetrix fluidics Station 450 and scanned the arrays using a GeneChip Scanner 3000 G7 (Affymetrix, Santa Clara, CA). The Affymetrix Gene- Chip1 Operating Software (GCOS) collected and extracted feature data from Affymetrix GeneChip1 Scanners. The GeneChip Genotyping analysis software (GTYPE) was used to analyze feature intensity data stored in the GCOS Database, and provided high-throughput and accurate genotyping analysis.

5. Analysis of whole genome association study (Pooling)

5.1 Data acquisition

The CEL data files were extracted from DDT archived file. Probe intensities of each SNP were extracted from the CEL data file using the R package called AffParser. Each SNP has 10 probe quartets. Each quartet probe has 4 probes, including perfect match probe of allele A (PMA), mismatch probe of allele A (MMA), match probe of allele B (PMB), mismatch probe of allele B (MMB). These probe intensities were used in the following analysis.

5.2 Quality control (QC)

In this study, there are 10 probe quartets for each SNP. Since 10 pools were generated for case and control groups, there are totally 100 probe quartets which can be detected for each SNP. The SNP level QC utilizes this information to decide whether or not the SNP is detected.

Let	Na	=	the number of detected probe quartets for that SNP (Case)
	Nu	=	the number of detected probe quartets for that SNP (Control)
	Qa	=	probe detection rate (Case), Na/100
	Qu	=	probe detection rate (Control), Nu/100
	Q	=	overall detection rate
		=	min(Qa,Qu)
	If	Q>0.95	this SNP is detected
	Else		this SNP is not detected

5.3 RAS calculation

For those SNPs that passed QC, their probe intensities were used to calculate a relative allele signal (RAS) using GenePool software. The RAS is the ratio between the intensity of the signal of A allele and the summation of the intensity of the signal from both alleles (Pearson JV et al., 2007). It is calculated as following equation.

RAS = PMA/(PMA+PMB)

Up to this point, each SNP in each group has only one RAS value representing the allele frequency of A allele.

5.4 Statistical analysis

1) Hypothesis testing (T-test)

The T-test was employed to determine SNPs that are significantly different between case and control groups. This T-test is more often used in the SNP-MaP studies (Butcher LM et al, 2005; Craig DW et al., 2005; Johnson C et al., 2006). The reason to use this T statistics is because RAS is numerical data and there are replicates in the data. The test statistics was calculated from the following equations:

$$T = (\bar{X}_a - \bar{X}_u)/S_{\bar{X}_a - \bar{X}_u}$$

$$\text{where } S_{\bar{X}_a - \bar{X}_u} = \sqrt{\frac{s_a^2}{N_a} + \frac{s_u^2}{N_u}}$$

The degree of freedom of this test statistics equals to *Na+Nu-2*. Since there are totally 20 samples, the number of degrees of freedom would be 18. From this number, we can find the p-value by reference to T-distribution table.

2) Odd ratio calculation

For each SNP, the odds ratio was calculated by pairing the data between groups for each sample. Thus, for M case pools and N control pools, there are M x N pairs. For each pair, the odds ratio was calculated as following:

OR = RAS*u* x (1-RAS*a*)/RAS*a* x (1- RAS*u*)

The subscript *a* and *u* indicate case and control groups respectively. The mean and 95% confidence interval of OR value from all pairs in each SNP were calculated.

3) Other statistical analysis

Standard errors of mean (SEM) was used to assess reliability DNA pooling approach (Johnson C et al., 2006). Data from NetAffx Analysis Centre (http://www.affymetrix.com/analysis/index.affx), the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/), International HapMap Project (http://www.hapmap.org/cgi-perl/gbrowse/hapmap_B36/) and SNPper information (http://snpper.chip.org/bio/snpper-enter) were employed in this study.

6. Selection of candidate SNPs

In this study, we focus on chromosome 1 since several regions have been confirmed as major susceptibility loci to SLE development (Wong M and Tsao BP, 2006). Only SNPs within the gene regions which passed quality control, odd ratio > 1.25 and *p*-value < 0.0001 were included. In addition, only functional SNPs in genes that are good candidates on biological reasons were selected. For example, SNPs in promoter region usually lead to different mRNA level which is the result from changing transcription factor binding site. Selection of SNPs in this region was predicted by TFSEARCH program (http://www.cbrc.jp/research/db/TFSEARCH.html) (Akiyama Y, 1998). For SNPs which are non-synonymous in exon may affect the amino acid sequence of predicted proteins where functions like DNA binding, catalytic activity and receptor– ligand contact. This region was analyzed by using SIFT program (http://blocks.fhcrc.org/sift/SIFT_dbSNP.html) (Ng PC and Henikoff S, 2006). SNPs in intron are most commonly found at the beginning and end of the donor and acceptor consensus splice sequence and may cause either exon skipping or utilization of cryptic splice sites resulting in the absence of normally spliced mRNA. The effects of such SNPs can be predicted by information theory based binding site analysis (https://splice.cmh.edu/) (Nalla VK and Rogan PK, 2005). Finally, we predicted SNPs in

3'UTR using miRNA target-gene prediction (http://cbit.snu.ac.kr/~miTarget/) (Kim SK et al., 2006).

7. Validation of candidate SNPs

PCR-RFLP (Polymerase Chain Reaction-Restriction fragment Length Polymorphism) was used to individual genotype selected SNPs. The genomic DNA of SLE patient and healthy controls in the same groups used in pooling study was amplified with specific primers under optimal condition. Amplified DNA was digested with specific restriction enzymes and digested fragments were separated on a 3% agarose gel. The resulting products were visualized under UV light by Camera Gel Doc™ MZL (BIO-LAD, USA). Primers and restriction enzymes for individual genotyping were newly designed in this study as shown in Table 10. Moreover, the results from PCR-RFLP were validated by means of direct sequencing.

Table 10 Primers and restriction enzymes used in individual genotyping of SNPs

Genes	SNPs	Primers (5'->3')	Restriction enzymes
1. *CD1D*	rs10489821	GTGGTGGGGAGAACCAGA	NspI
		GCTGGCTCACATCAAAAGG	
2. *IFIX*	rs856084	GCAACGATTGCTGACCAC	AhdI
		CCAGTGATGAGATGGGAGAA	
3. *TLR5*	rs2072493	CCGTGGAAAGAGAGAAGAGG	Tsp509I
		GGAACCAGCTCCTAGCTCCT	
4. *NOS1AP*	rs1858233	TTCCCTACCCTAGCTCCACA	AluI
		CCTGCCATCTTTCCAGTTG	

8. Statistical Analysis for individual genotyping results

Standard errors of mean (SEM) and Pearson correlation tests were used to compare DNA pooling approach to individual genotyping. Allele frequencies from individual genotyping were compared between case and control using Chi-square (χ^2) test. A *P* value corrected with Monte Carlo simulations (100000 iterations) of < 0.05 was considered significant. Odds ratios and 95% confidence intervals were also determined. Linkage disequilibrium structure (D') was generated in Haploview program (Barrett JC et al., 2005).

Part II Candidate gene association study

1. Subjects

Two hundred Thai patients from King Chulalongkorn Memorial hospital, who fulfilled at least 4 of the American College of Rheumatology (ACR) revised criteria for SLE were included in this study (Tan EM et al., 1982). As a control group, we recruited 200 ethnically and geographically matched controls from healthy blood donor of the Thai Red Cross Society. The ethics committee of the faculty of Medicine, Chulalongkorn University, Bangkok, Thailand approved the study and the subjects gave their informed consent.

2. Selection of SNP positions

The selection of SNPs in candidate gene association study was performed according to means of whole genome association study (Pooling) as previously described.

3. Genotyping study

PCR-RFLP (Polymerase Chain Reaction-Restriction fragment Length Polymorphism) was used to genotype SNPs of *MNDA* A2706G, C16432T; *IFI16* C-7217T, C6771G, A23201G, C27140T and *AIM2* G-151T. The genomic DNA of patient with SLE and healthy controls was amplified with specific primers under optimal condition as follow: an initial denaturation at 94°C for 2 minutes, followed by 30 cycles of denaturation (94°C, 20

seconds), annealing (55°C to 60 °C, 50 seconds), extension (72°C, 20 seconds) and final extension at 72°C for 7 minutes. Amplified DNA was digested with specific restriction enzymes. The digested fragments were separated on a 3% agarose gel and visualized under UV light by Gel DocTM MZL (BIO-RAD). Primers and restriction enzymes for genotyping were newly designed in this study as shown in Table 11.

PCR-SSP (Polymerase Chain Reaction-Sequence Specific Primer) was used to analyze SNPs of *IFI16* A44962T and *AIM2* C3452T. The genomic DNA of patient with SLE and healthy controls was amplified with specific primers that were newly designed as well (Table 10). Internal control primers were used to check for successful PCR amplification. Then, PCR was carried out under specific condition as follow: an initial denaturation at 94°C for 2 minutes, followed by 30 cycles of denaturation (94°C, 20 seconds), annealing (55°C and 58 °C, 50 seconds), extension (72°C, 20 seconds) and final extension at 72°C for 7 minutes. The PCR products were separated by 1.5 % agarose gel and visualized under UV light by Gel DocTM MZL (BIO-RAD). The results of these two assays were validated by means of direct sequencing.

Table 11 Primers and restriction enzymes used for analysis of SNPs

Genes	SNPs	Sequence of primers (5'→ 3')	Restriction enzyme
1. *MNDA*	rs7513873 (A2706G)	ACTCATCCTCACCAACACTG	Nla III
		GCGACAGGGCGAAACTC	
	rs2276403 (C16432 T)	GGATCCATGGATGTAGTG	Rsa I
		ATGAGGAACCTGGTGTTC	
2. *IFIX*	rs856084 (G13792T)	GCAACGATTGCTGACCAC	AhdI
		CCAGTGATGAGATGGGAGAA	
3. *IFI16*	rs4657618 (C-7217T)	GCCAGCCTGCTGAAATAGAAG	Hae III
		GTAACTCTGGCTCTTGAG	
	rs866484 (C6771G)	GCCGTTCCCCATCTCCCAAG	HpyCH4 III
		CATCAAGTATCCCTGTGAAAGC	
	rs1772414 (A23201G)	CCCATTTCCCCTTTGCTTAT	Dra III
		TCTGTGAATTGGGGCAGAAT	
	rs3754460 (C27140T)	TGTTTGCCATCTTGCTTCAG	Hae III
		AGGCAATTGGGAACTTTGAA	
	rs6940 (A44962T)	CCTGATTCAAGTATGGAAA	-
		CCTGATTCAAGTATGGAAT	
		CCCATTCCATAGGATTAAC	
4. *AIM2*	rs16841642 (G-151T)	CACTAGCAGCCACAGAAG	Dra III
		GGGTGTCGTTGGTTTTGC	
	rs2276405 (C3452T)	GCCTGTGGCAATATTAAACTC	-
		GCCTGTGGCAATATTAAACTT	
		GGCTGATCCCAAAGTTGT	
5.internal control		GCCTTCCCAACCATTCCCTTA	-
		TCACGGATTTCTGTTGTGTTTC	

4. Detection of apoptosis

Within 60 minutes after blood drawing, peripheral blood mononuclear cells (PBMCs) from 15 healthy volunteers containing different IFI16 genotypes were isolated by Ficoll-Hypaque gradient (Robbins Scientific Corporation, Sunny vale, CA) and resuspended in RPMI-1640 medium (Sigma, New York, NY, USA) with 10% fetal bovine serum (FCS) (Gibco, Karlsruhe, Germany) and 100 U/mL penicillin (Gibco, Karlsruhe, Germany) and 100 µg/ml streptomycin (Gibco, Karlsruhe, Germany) at a concentration of $5x10^6$ cells/ml. These cells were plated into 24-well plastic tissue culture plates (Greiner, Germany). PBMCs were treated with doxorubicin (Pharmacia and Upjohn, Germany) at a concentration of 10 µg/ml and then cultured at 37 °C in 5% CO_2 for 24 hr. This condition was determined to be effective to induce highest p53 expression in dose (1, 10 and 100 µg/ml) and time (8 and 24 hours) response experiments. p53 mRNA expression was studied using real time RT-PCR (Reverse Transcription-Polymerase Chain Reaction) with following primers: 5'TGGCCATCTACAAGCAGTCACA3' and 5' GCAAATTTCCTTCCACTCGGAT 3' (Yu L and Domann FE, 2006). Finally, the percentage of apoptotic cells was measured by FACScan flow cytometer (Becton Dickinson, USA) for annexin V-FITC and propidium iodide (PI) binding according to the manufactures protocol (BD Biosciences Pharmingen, San Diego, CA, USA). Positive cells for only annexin V was counted as apoptotic.

5. The binding of nuclear proteins to *AIM2* promoter at position G-151T by Electrophoretic Mobility Shift Assay (EMSA)

5.1 Screening cell lines for *AIM2* expression

Complementary DNA (cDNA) of twelve different cell lines including Hacat, Jurkat, Hela, Molt4, UACC903, HepG2, Hep2, BLCL, U937, K562, SW480 and HT29 cells were detected for *AIM2* mRNA expression by conventional RT-PCR with 5' CGTGAGGCGCTATTTACCTC 3'and 5'CAGGAGGAGAAGGAGAAAGTTG 3' primers. The PCR condition consisted of an initial denaturation at 94°C for 5 minutes, followed by 35

cycles of denaturation (94°C, 1 minute), annealing (60 °C, 1 minute), extension (72°C, 1 minute) and final extension at 72°C for 7 minutes.

5.2 Cell culture

BLCL cells and UACC903 cells (a generous gift from Dr. Ricky W. Johnstone, Gene Regulation Laboratory, Australia) were cultured in RPMI-1640 medium (Sigma, New York, NY, USA) with 100 U/mL penicillin (Gibco, Karlsruhe, Germany), 100 µg/ml streptomycin (Gibco, Karlsruhe, Germany), and supplemented with 10% FCS (Gibco, Karlsruhe, Germany). Cells were maintained at 37°C in a humidified atmosphere at 5% CO_2.

5.3 Preparation of nuclear protein extracts

Nuclear protein extracts from BLCL and UACC903 cells were prepared as previously described (Tencomnao T et al., 2004) with some modifications. Briefly, the cultured cells ($5x10^7$ to $1x10^8$) were collected by centrifugation at 250xg for 10 min. The cell pellet was resuspended in 2.5 ml of cell lysis buffer (buffer A; 10 mM HEPES, pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM DTT, 0.5 mM PMSF (phenylmethysulfonyl fluoride), 1 µg/ml aprotinin, 1 µg/ml pepstatin and 1 µg/ml leupeptin) and allowed to swell on ice for 10 min and centrifuged at 250 x g for 10 min. The cell pellet was resuspended in 1.5 ml of cell lysis buffer, Nonidet P-40 (NP40) was added to 0.05%, and cells were homogenized with about 10 strokes of a tight-fitting Dounce homogenizer to release the nuclei. The successful release of nuclei was checked by phase-contrast microscopy. After the nuclei were collected by centrifugation at 250 x g for 10 min, they were resuspended in 1 ml of nuclear extraction buffer (buffer C; 5 mM HEPES, pH 7.9, 26% glycerol, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 0.5 mM DTT, 0.5 mM PMSF, 1 µg/ml aprotinin, 1 µg/ml pepstatin and 1 µg/ml leupeptin). The total volume was measured, and NaCl was added to a final conc of 300 mM. The nuclear suspension was stirred on ice for 30 min, and then centrifuged at 24,000 x g for 20 min. The supernatant was aliquoted, snap-freezed in dry ice/ethanol, and stored in -70 °C before use. The protein concentrations were determined according to Lowry method (Lowry OH et al., 1951).

5.4 Preparation of labeled oligonucleotide probes

The oligonucleotide were synthesized and annealed to generate double-stranded oligonucleotides with overhanging sequences. A stretch of five T residues at the 5' end of the reverse oligonucleotide was added to increase the labeling efficiency. As shown in Table 12, the binding sequence for the corresponding transcription factor AP-1 was designated in bold, SNPs were underlined and mutated nucleotides were small letter.

Double-stranded oligonucleotide probes were generated as follow. Briefly, annealing reaction was performed in total volume of 50 µl using 5 µl of 10x annealing buffer (3 M KCL, 0.1 M Tris, pH 7.8, 10 mM EDTA). Equal amounts of 1,000 pmole in each oligonucleotide were added and heat to 94 °C for 2 min. The double-stranded oligonucleotides were allowed to cool slowly down to room temperature over 1 hour. Afterward, double-stranded oligonucleotides will be precipitated using 3M sodium acetate and 95% EtOH and dissolved in 50 µl water to obtain 10 pmol/µl. Then,10 pmol (1µl) of double-stranded oligonucleotide was labled by filling-in reaction at the 3' end by mixing with 1 µl (5U) of Klenow polymerase (exo-), 1 µl of [α-^{32}P]dATP (3,000 Ci/mmol), 2 µl of 5x dATP buffer containing dCTP, dGTP and dTTP (0.2 mM of each) and 5 µl of water. The mixed reaction was incubated at 37 °C for 30 min.

The labeled probes were purified by passage through sephadex G-50 column. [To prepare columns, the columns were pre-centrifuged for 1 min at 2000 g (Sephadex) or 750 g (Sephadex) upon loading of uniform slurry. The sample was applied directly to center of the shrunken gel bed and centrifuged at the same g for 2 min]. The labeled probes were finally diluted by STE buffer (100 mM Tris-HCl, pH 8.0, 1 M NaCl, 10 mM EDTA) to get the final concentration 0.1 pmol/µl.

5.5 Binding reaction and visualization

Binding reaction was performed in total volume of 30 µl by mixing the binding solution (20 mM HEPES, pH 7.9, 1 mM MgCl, 0.5 mM DTT, 0.5 mM EDTA, and 4% ficoll) with KCl to a final conc of 50 mM, 2 µg of poly (dI-dC).poly (dI-dC), 1 µg of salmon sperm DNA, 100 fmol labeled probe, and 2.5-10 µg (5 µg) of nuclear extracts form cells. In

competition experiments, a 50-fold excess of either unlabeled doubled-stranded wild type or mutant DNA probes was added to the reaction. The mixture was incubated at room temperature for 20 min. The protein-DNA complexes were separated on 5% non-denaturing polyacrylamide gel in 0.5x Tris-acetate buffer at 200 volts for 6 hours. The gel was subjected to autoradiography for 12-16 hours at -70 °C with phosphor screen. Finally, the complexes were visualized using PhosphoImager System (ImageQuaNT Software, Molecular Dynamics).

Table 12 Oligonucleotide probes used for EMSA

Probe name	Oligonucleotide probes
1. AP-1G	5' CACTTTGTGACTAAGGGGCC 3'
	3' GTGAAACACTGATTCCCCGG-TTTTT 5'
2. AP-1T	5' CACTTTTTGACTAAGGGGCC 3'
	3' GTGAAAAACTGATTCCCCGG-TTTTT 5'
3. Mutated AP-1	5' CACTTTGTGcaTAAGGGGCC 3'
	3'GTGAAACACgtATTCCCCGG-TTTTT 5'

6. Statistical Analysis

PLINK v1.03 program (Purcell S et al., 2007) was used to calculate the empirical *P* values (100,000 permutations) and odds ratios to perform the case-control association tests for single SNPs and to model the genotypic associations. In addition, haplotype analysis was also performed using this program. Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium (LD) were determined by JLIN, a Java based Linkage Disequilibrium Plotter (Carter KW et al., 2006). Moreover, the power was calculated using the CaTS Power Calculator for genetics studies (Skol AD et al., 2006).

The association between genotypes and renal pathology or apoptosis was analyzed by using non parametric test of SPSS software version 11.5 (SPSS Inc, Chicago, IL). The Mann-Whitney U test was used in case of comparing between two groups, while Kruskal-Walls H was applied in case of analysis more than two groups. A *P* value of less than 0.05 was considered significant.

Part III Expression profile study

1. Subjects

1.1 Expression study in leukocytes (buffy coat)

Twenty patients who have been diagnosed with SLE according to ACR criteria were recruited (Tan EM et al., 1982). All patients (no data in 2 patients) had high SLE disease activity index (SLEDAI) scores (mean score 14.78 ±1.21). Among 20 patients at the time of study, 2 patients were not receiving any medication. 8 out of 20 patients were receiving prednisolone alone, while 10 patients were receiving prednisolone combined with other medicines such as endoxan, cellcept and ACEI. Furthermore, 9 age and sex matched healthy controls were also included (Mean age 27.89 ± 1.35). Characteristics of SLE patients included in expression study for leukocytes were shown in Table 13.

Table 13 Characteristics of SLE patients included in expression study for leukocytes

Patients no.	sex	Age	Treatment	dose (mg/day)	SLEDAI score
1	F	37	prednisolone	5	24
2	F	30	prednisolone	15	10
3	F	44	prednisolone	7.5	12
4	F	25	prednisolone	100	14
5	F	22	-	-	16
6	F	33	prednisolone, endoxan	60, 1200	14
7	M	25	prednisolone, endoxan,ACEI	30, 1000, 10	12
8	F	38	-	-	16
9	F	27	prednisolone, cellcept, ACEI	15, 720, 20	10
10	M	45	prednisolone,endoxan	60, 800	-
11	F	30	prednisolone,ACEI	50, 2.5	16
12	F	35	prednisolone	50	12
13	F	47	prednisolone, ACEI	2.5, 10	16
14	F	22	prednisolone, cellcept	7.5, 1000	8
15	F	24	prednisolone, cellcept	10, 1500	14
16	F	25	prednisolone	12.5	30
17	F	32	prednisolone, endoxan,ACEI	40, 1000, 10	12
18	F	24	prednisolone, ACEI	50, 40	16
19	F	30	prednisolone	30	14
20	F	34	prednisolone	60	-
	Female=18; Male=2	31.45 ± 1.72 (Mean±SEM)			14.78 ±1.21 (Mean±SEM)

1.2 Expression study in B and T lymphocytes

Ten female patients from King Chulalongkorn Memorial hospital, who have been diagnosed as the American College of Rheumatology (ACR) revised criteria for SLE were enrolled in this study (Tan EM et al., 1982). SLE patients were classified into two groups according to SLE disease activity index (SLEDAI) (see Table 14). The first group included 5 inactive SLE patients receiving low doses of prednisolone at less than 5 milligrams per day within at least one month before blood collection (SLEDAI < 3). Another group composed of 5 active SLE patients receiving low to intermediate doses of corticosteroids (2.5- 15 milligrams per day of prednisolone with or without mycophenolate mofetil or azathioprine or myfortic) (SLEDAI $\geq$ 3) (Sodsai P et al., 2008). In addition, 5 age and sex matched healthy controls were also included (Mean age 26.6 ± 0.68).

Table 14 Characteristics of SLE patients included in expression study for B and T lymphocytes

Patients no.	Age	Treatment	Steroid dose (mg/day)	SLEDAI score	Stage
1	25	prednisolone	2.5	0	inactive
2	32	prednisolone	5	1	inactive
3	38	prednisolone	5	0	inactive
4	26	prednisolone	2.5	0	inactive
5	34	prednisolone	5	0	inactive
6	24	prednisolone	15	4	active
7	16	prednisolone, azathioprine	5, 75	4	active
8	30	prednisolone	5	10	active
9	29	prednisolone, mycophenolate mofetil	2.5, 1000	3	active
10	33	prednisolone	2.5	18	active
	28.7 ± 1.97 (Mean±SEM)				

1.3 Expression study in renal biopsy

Fourteen female patients underwent renal biopsy for diagnostic evaluation during 2002-2005. All patients have been diagnosed with SLE according to ACR criteria and have had at least one of the following criteria for renal involvement: a total urinary protein level of more than 0.5 gram per day, an increment of serum creatinine levels of more than 0.5 mg/dl during one month period of follow up or presence of pyuria, hematuria or urinary cast by microscopic examination. The histological types of lupus nephritis (LN) were defined by using World Health Organization (WHO) classification (Weening JJ et al., 2004). Most patients had received the same immunosuppressive treatment including prednisolone and cyclophosphamide or mycophenolate mofetil. In addition, five kidney sections from living donors or cadavers for kidney transplantation were used as controls in this study. Characteristics of SLE patients included in expression study for renal biopsy were shown in Table 15.

Table 15 Characteristics of SLE patients included in expression study for kidney biopsy

Patients no.	Age	Serum Creatinine (mg/dl)	Protienuria (g/day)	Urinary erythrocyte count	SLEDAI score	Activity score	Chronicity score	WHO Class
1	41	2.2	1.17	12	10	7	9	III
2	18	0.9	6.66	2	4	5	2	III
3	36	3.4	5.34	3	6	12	8	IV
4	34	0.6	3.6	20	18	9	1	III
5	49	0.9	1.21	40	8	0	0	V
6	30	4.3	3	1	14	6	7	IV
7	22	0.9	7.74	15	16	11	3	IV
8	31	0.8	0.53	50	10	3	0	III
9	35	0.7	4.13	5	10	6	1	IV
10	38	1.5	2	1	8	2	7	III
11	18	0.7	9.78	12	12	5	2	III
12	23	0.7	3.12	0	6	4	5	V
13	23	1	0.29	1	2	0	6	V
14	32	2.2	9.1	80	18	19	0	IV
Mean±SEM	30.71±2.43	1.49±0.31	4.12±0.84	17.29±6.32	10.14±1.33	6.36±1.37	3.64±0.87	

All subjects in the expression studies were approved by the Ethic Committee for Human Research of the Faculty of Medicine, Chulalongkorn University, and written informed consents were obtained from all patients.

2. Cell separation

Twenty milliliters (ml) heparinized blood samples were collected and processed within 2 hours of collection. Peripheral blood mononuclear cells (PBMC) were isolated using Ficoll–Hypaque reagent (Sigma, St Louis, Missouri, USA). T lymphocytes were purified using a positive selection strategy with magnetic beads (Miltenyi Biotech, Bergisch Gladbach, Germany). Monocytes were removed from non T populations by adherence to plastic at 37 °C for 1 hr (Klein SC et al., 1997). Then, non T cells without monocytes (majority B cells) were collected.

3. Total RNA Extraction

Total RNA were extracted by RNeasy mini kit (Qiagen, Chatworth, CA, USA) according to the instruction of manufacturer. The kit combines the selective binding properties of a silica-based membrane with the speed of microspin technology. Briefly, kidney tissues were first lysed and homogenized in the presence of a highly denaturing guanidine-thiocyanate containing buffer, which immediately inactivates RNases to ensure purification of intact RNA. Ethanol was added to provide appropriate binding conditions, and the sample was then applied to an RNeasy Mini spin column, where the total RNA binds to the membrane and contaminants were efficiently washed away. High-quality RNA was then eluted in 35 µl water and stored at -70 °C.

4. Complementary DNA (cDNA) synthesis

Synthesis of single-strand cDNA was carried out, as used total RNA 0.25 µg (maximum volume of RNA template not exceed 11.5 µl) for reverse-transcribed into cDNA by Taqman™ Reverse Transcriptase Reagent (Applied Biosystems, Roch Molecular Biochemical, NJ, USA). Eighteen microliters (µl) of reverse transcription mastermix

containing 3 µl of 10xRT buffer, 6.6 µl of 25mM MgCl2, 6.0 µl of 10mM 4dNTP, 1.5 ul of 50µM Random primer, 12 U of RNase inhibitor and 37.5 U of MutiScribe™ Reverse Transcriptase, 18 µl of mastermix was added into 0.25 µg (11.5 µl) RNA template and transcribe at 25 °C for 10 minutes, 48 °C for 30 minutes and 95 °C for 5 minutes. Complementary DNA was kept at -20 °C until used for measuring gene expression by real-time RT-PCR.

5. Real time RT-PCR assay

A real time RT-PCR assay was used to quantify target gene transcripts using β-actin or 18srRNA as endogenous control. All primers were newly designed in this study as shown in Table 16. PCR amplification was performed with 2x QuantiTect SYBRGreen PCR Master Mix with 0.5 µM primers, 16 ng cDNA and nuclease-free water according to the manufacturer's protocol (Qiagen). PCR conditions included an initial activation at 95°C for 15 minutes, denaturation at 95°C for 15 seconds, annealing at 55-58 °C for 30 seconds and extension at 72°C for 30 seconds followed by repeating for 40-50 cycles. The mRNA levels were measured by a Light Cycler machine (Roche Molecular Biochemicals, Indianapolis, IN). Levels of mRNA were expressed as threshold cycle (CT) and used comparative CT method for analysis. For relative quantification, the expression target genes were normalized by expression of β -actin or 18srRNA. The amount of target was calculated by $2^{-\Delta\Delta_{CT}}$.

$$\Delta CT = CT(target) - \Delta CT(reference)$$

$$\Delta\Delta CT = \Delta CT(normalized\ target) - \Delta CT(normalized\ calibrator)$$

Table 16 Primers used for Real time RT-PCR assay of *MNDA, IFIX, IFI16* and *AIM2*

Genes	Primers (5'->3')
1. *MNDA*	GGAAGAAGCATCCATTAAGG
	GTTTGTCTAGACAGGCAAC
2. *IFIX*	GAGACTGGAACCAAAAGG
	CGCGATTATTGGGTCTTC
3. *IFI16*	CTGCACCCTCCACAAG
	CCATGGCTGTGGACATG
4. *AIM2*	CAGGAGGAGAAGGAGAAAGTTG
	GTGCAGCACGTTGCTTTG
5. *Beta-actin*	ACCAACTGGGACGACATGGAGAA
	GTGGTGGTGAAGCTGTAG CC
6. 18srRNA	GCCCGAAGCGTTTACTTTGA
	TCCATTATTCCTAGCTGCGGTATC

6. Conventional reverse transcription PCR

Expression of *IFIX* and *IFI16* isoforms was semi-quantified by conventional RT-PCR. The reaction volume of the amplification reaction was 10 µl, containing 2 µl of 8.3 ng/µl cDNA, 0.2 µl of 5.0 U/µl Taq polymerase (Promega), 1 µl of 10x PCR buffer (20mM Tris-HCl pH 8.0, 100 mM KCl), 0.8 µl of 25 mM MgCl2, 0.64 µl of 10 mM dNTP, 0.2 µl (10 pmol) of each primer and 4.96 µl of sterile water. Primer sequences of IFIX (Ding Y et al., 2004) and IFI16 were shown in Table 17. The PCR step consisted of an initial denaturation at 94°C for 5 minutes, followed by 30 cycles of denaturation (94°C, 30 seconds), annealing (60°C for *IFIX* gene, or 62°C for *IFI16* gene, 30 seconds), extension (72°C, 30 seconds) and final extension at 72°C for 7 minutes. Amplification was performed in Perkin Elmer/ GeneAmp PCR system 2400. The PCR products were loaded in 1.5% Tris-acetate agarose gel and analyzed using electrophoresis containing Trisacetate buffer at 100 voltages for 40 minutes followed by staining 50 µg/ml ethidium bromide. The density of product bands was semi-quantified using software of Gel DocTM MZL (BIO-RAD).

Table 17 Primers used for conventional RT-PCR assay of *IFIX* and *IFI16* isoforms

Genes	Primers (5'->3')
1. *IFIX*	GGAACAGAGTCAGCATCC
	GTTATTTGATATCCTTGTCC
2. *IFI16*	CATCTTCGGACTCCTCAG
	GTTCAGCACCATCACTTC

7. Statistical analysis

Levels of mRNAs were expressed as log-transformation. The independent *t* test of SPSS software version 11.5 (SPSS Inc, Chicago, IL) was used to compare data between two groups. A *p*-value of <0.05 was considered statistical significance.

CHAPTER V

RESULTS

Part I Whole genome association study

1. Result summary and comparison with known candidate genes

Hybridization signal intensities were obtained from all 20 DNA pools. After the assessment of quality control, SNPs with overall detection rate more than 95% were considered. In our study, 51,215 out of 58,960 SNPs (87%) could be assigned for subsequent analyses. In 10 pools of normal control group, the averaged chip-to-chip difference (standard error of mean: SEM) in estimated allele frequency was 5.54%. Moreover, the averaged difference (SEM) between estimated allele frequency of our Thai normal control group and the published database (Asian populations) from NetAffx was 8.85%.

The 51,215 SNPs that passed quality control were analyzed to determine the difference between SLE and normal control groups. The T-test was used to analyze since RAS is numerical data and there are replicates in this data. The 3,814 SNPs were found to be associated with SLE with *P* values less than 0.0001 and odd ratio (OR) more than 1.25. The –Log10 *P* values and OR of associated SNPs were plotted according to chromosomal locations as shown in Figure 9 and 10, respectively. Top 50 SNPs ranked by *P* values and odd ratios (OR) were shown in Table A and B, respectively. The ranges of minimal and maximal values were 1.07×10^{-11} to 1.27×10^{-8} for *P* value (See in Table 18) and 2.51 to 4.32 for OR (See in Table 19). There were only 20 SNPs ranked by *P* values and 17 SNPs ranked by OR that were in the gene regions. For SNPs in the gene regions, they were not in known candidate genes and likely were not related to SLE pathogenesis. However, there were some SNPs closed to the genes that may be important to SLE. For example, SNP rs276992 ranked as 9th by *P* value near to *IRF8*, which is a gene in type I interferon pathway. This SNP is very far from *IRF8* approximately 287929 bp.

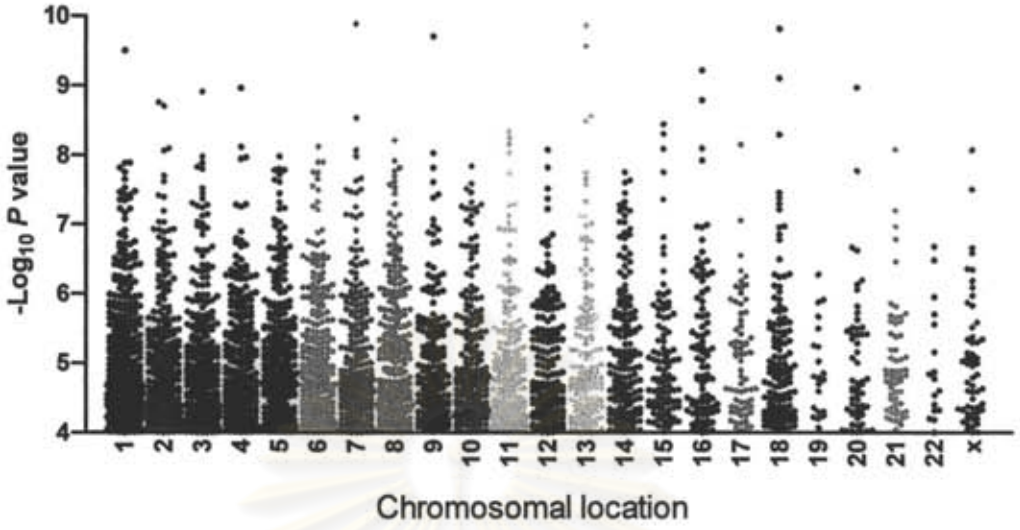


Figure 9 $-\text{Log}_{10}$ *P* values of passed QC SNPs with p values < 0.0001 and odd ratio > 1.25 according to chromosomal locations

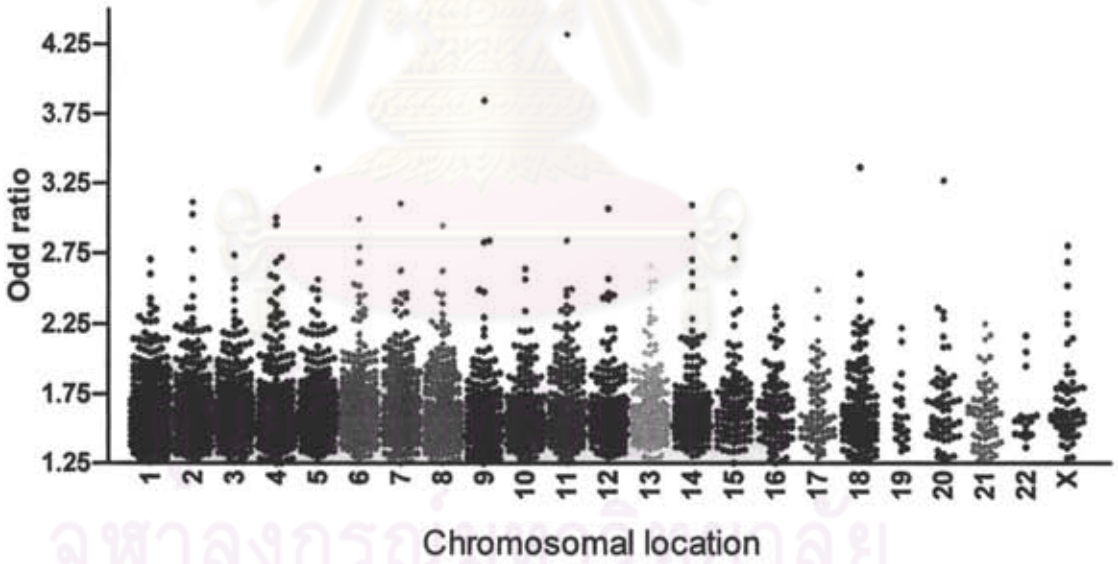


Figure 10 Odd ratios of passed QC SNPs with p values < 0.0001 and odd ratio > 1.25 according to chromosomal locations

Table 18 Top 50 SNPs ranking by P value

dbSNP RS ID	loci	OR mean	P-value	Case RAS	Control RAS	MAF	Gene symbol	Gene name or function	Gene region /distance from SNP
1.rs2914941	8q22.1	2.2260546	1.07E-011	0.46	0.64	0.4881	AC010834.19/RBM12B	AC010834.19/RNA binding motif protein 12B	16476/-14522
2.rs12081	15q15.1	2.7096148	2.98E-011	0.3	0.49	0.4286	FLJ43339	FLJ43339	Downstream
3.rs10487264	7q31.33	2.163385	1.27E-010	0.57	0.73	0.1548	SPAM1/GPR37	sperm adhesion molecule 1/G protein-coupled receptor 37	475393/-345394
4.rs9300973	13q33.2	1.6864449	1.32E-010	0.62	0.5	0.2976	SLC10A2/DAOA	solute carrier family 10, member2/D-amino acid oxidase activator	1833182/-588832
5.rs1477489	18q22.3	1.9930808	1.48E-010	0.53	0.38	0.369	NETO1/FBXO15	neuropilin (NRP) and tolloid (TLL)-like 1/F-box protein 15	344028/-981791
6.rs2252968	9q34.11	1.9314635	1.93E-010	0.58	0.43	0.3571	C9orf88	chromosome 9 open reading frame 88	Intron
7.rs9316967	13q21.1	2.0366881	2.58E-010	0.64	0.49	0.3333	PCDH17/DIAPH3	protocadherin 17/diaphanous homolog 3 (Drosophila)	380308/-1653620
8.rs2766003	1p21.3	2.1602583	3.13E-010	0.56	0.71	0.4405	CNN3/ALG14	calponin 3, acidic/asparagine-linked glycosylation 14 homolog (yeast)	63856/-21914
9.rs276992	16q24.1	2.086628	5.86E-010	0.59	0.43	0.4524	IRF8/FOXF1	interferon regulatory factor 8/forkhead box F1	287929/-323430
10.rs7228560	18q22.1	2.2276337	7.61E-010	0.29	0.45	0.369	SERPINB8/CDH7	serpin peptidase inhibitor, clade B (ovalbumin), member 8/cadherin 7	326506/-1453724
11.rs2327302	20p12.2	2.286241	1.03E-009	0.47	0.31	0.1905	JAG1/BTBD3	jagged 1 (Alagille syndrome)/BTB (POZ) domain containing 3	97070/-1156061
12.rs1388848	4q25	2.0143512	1.08E-009	0.38	0.25	0.2439	DKK2/PAPSS1	dickkopf homolog 2 (Xenopus laevis)/	271478/-420376
								3'-phosphoadenosine 5'-phosphosulfate synthase 1	
13.rs966226	3q24	2.1715929	1.20E-009	0.35	0.51	0.4762	PLSCR1/ZIC4	phospholipid scramblase 1/Zic family member 4	232408/-638461
14.rs10500349	16p13.2	2.1548719	1.58E-009	0.56	0.42	0.4881	A2BP1	ataxin 2-binding protein 1	Intron
15.rs925899	2q32.2	1.9455016	1.73E-009	0.76	0.85	0.119	DIRC1	disrupted in renal carcinoma 1	Intron
16.rs10496332	2q11.2	2.2881771	1.96E-009	0.61	0.43	0.4286	MGAT4A	mannosyl (alpha-1,3-)-glycoprotein beta-1,	Promoter
								4-N-acetylglucosaminyltransferase, isozyme A	
17.rs347405	13q14.11	1.883772	2.63E-009	0.36	0.25	0.25	DGKH	diacylglycerol kinase, eta	Intron
18.rs2189880	7q35	1.9980456	2.88E-009	0.35	0.49	0.4048	OR6B1	olfactory receptor, family 6, subfamily B, member 1	Promoter
19.rs9317014	13q21.2	2.4911385	3.06E-009	0.44	0.28	0.2381	PCDH17/DIAPH3	protocadherin 17/diaphanous homolog 3 (Drosophila)	1174363/-859565
20.rs961090	15q15.1	2.3193556	3.47E-009	0.65	0.78	0.1905	FLJ43339	FLJ43339	Downstream
21.rs2168535	11p12	1.5696301	4.44E-009	0.39	0.48	0.4405	LRRC4C/API5	leucine rich repeat containing 4C/apoptosis inhibitor 5	1687515/-1510265
22.rs10519161	15q22.2	2.0255977	4.83E-009	0.65	0.51	0.4048	NLF2	nuclear localized factor 2	Promoter
23.rs8090395	18q21.1	1.7238739	4.93E-009	0.51	0.62	0.2976	ZBTB7C/KIAA0427	zinc finger and BTB domain containing 7C/KIAA0427	347956/-163723
24.rs685428	11p13	2.016404	5.45E-009	0.29	0.41	0.4762	PAX6/RCN1	paired box gene 6 (aniridia, keratitis)/	57407/-244751
								reticulocalbin 1, EF-hand calcium binding domain	
25.rs6474087	8q12.1	1.8294323	6.05E-009	0.41	0.54	0.3095	PENK/IMPAD1	proenkephalin/inositol monophosphatase domain containing 1	472303/-49448
26.rs10502189	11q23.2	2.3233474	6.77E-009	0.45	0.3	0.2262	REXO2/FAM55A	REX2, RNA exonuclease 2 homolog (S. cerevisiae)/	33733/-48528
								family with sequence similarity 55, member A	
27.rs2323659	17p12	1.6254254	6.86E-009	0.45	0.55	0.4643	TEKT3/CDRT4	tektin 3/CMT1A duplicated region transcript 4	89808/-42402
28.rs1842129	6q22.31	1.7193189	7.43E-009	0.62	0.5	0.4167	TCBA1	T-cell lymphoma breakpoint associated target 1	Intron
29.rs3816739	4q13.3	1.9418038	7.57E-009	0.77	0.66	0.2024	APIN	APin protein	Promoter
30.rs1033043	16q12.2	2.1415044	7.82E-009	0.43	0.3	0.2262	FTO/IRX3	fatso/iroquois homeobox protein 3	479482/-99855

31.rs893371	15q26.1	1.6933301	7.95E-009	0.73	0.82	0.0119	AC091544.11/CHD2	/chromodomain helicase DNA binding protein 2	84324/-198417
32.rs10497327	2q24.3	1.8494937	7.95E-009	0.31	0.21	0.1786	XIRP2/B3GALT1	cardiomyopathy associated 3/	756822/-158367
								UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase, polypeptide 1	
33.rs2826187	21q21.1	1.706476	8.04E-009	0.48	0.6	0.4286	PRSS7/NCAM2	protease, serine, 7 (enterokinase)/neural cell adhesion molecule 2	2068258/-660685
34.rs742997	Xq27.3	2.5184443	8.16E-009	0.5	0.35	0.2619	SPANXN3/SLITRK4	SPANX family, member N3/SLIT and NTRK-like family, member 4	85392/-33987
35.rs7135979	12q15	1.88457	8.22E-009	0.31	0.43	0.4524	CAND1/DYRK2	cullin-associated and neddylation-dissociated 1/	123268/-256085
								dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 2	
36.rs7795140	7q31.32	1.9564828	8.32E-009	0.42	0.54	0.4762	FAM3C/PTPRZ1	family with sequence similarity 3, member C/	180142/-344110
								protein tyrosine phosphatase, receptor-type, Z polypeptide 1	
37.rs706551	2p16.1	1.7293454	8.51E-009	0.37	0.48	0.4762	PNPT1	polyribonucleotide nucleotidyltransferase 1	Intron
38.rs10513311	9q33.1	1.9957146	9.17E-009	0.45	0.32	0.2143	TLR4/DBC1	toll-like receptor 4/deleted in bladder cancer 1	142250/-1320187
39.rs198432	11q12.2	1.9114372	9.24E-009	0.59	0.46	0.439	C11orf11	chromosome 11 open reading frame 11	Intron
40.rs2536537	7p15.1	1.8215448	1.03E-008	0.58	0.45	0.4643	ADCYAP1R1/NEUROD6	adenylate cyclase activating polypeptide 1 (pituitary) receptor type I/	236641/-48316
								neurogenic differentiation 6	
41.rs723765	3p12.3	1.5569243	1.03E-008	0.49	0.59	0.3095	ROBO1	roundabout, axon guidance receptor,	Intron
								homolog 1 (Drosophila)	
42.rs7734499	5q34	1.8048685	1.03E-008	0.43	0.55	0.4524	MAT2B/ODZ2	methionine adenosyltransferase II, beta/	2502380/-1568368
								odd Oz/ten-m homolog 2 (Drosophila)	
43.rs2071486	4q27	1.7179562	1.07E-008	0.76	0.66	0.2738	CCNA2	cyclin A2	Intron (boundary)
44.rs10517528	4p14	2.245667	1.12E-008	0.6	0.74	0.25	HIP2	HUNTINGTIN-INTERACTING PROTEIN 2 or	Intron
								UBIQUITIN-CONJUGATING ENZYME E2-25K	
45.rs3785315	16q21	1.7340016	1.16E-008	0.56	0.43	0.3095	CDH11	cadherin 11, type 2, OB-cadherin (osteoblast)	Intron (boundary)
46.rs988031	8p12	1.6856899	1.18E-008	0.61	0.5	0.3571	NRG1	neuregulin 1	Intron
47.rs688176	6p25.2	1.8303695	1.25E-008	0.56	0.67	0.3571	C6orf146	chromosome 6 open reading frame 146	Downstream
48.rs595413	6p25.2	2.1232177	1.25E-008	0.5	0.37	0.3571	C6orf146	chromosome 6 open reading frame 146	Coding exon
49.rs4362707	3p24.3	1.7608936	1.27E-008	0.66	0.77	0.1071	ANKRD28/GALNTL2	ankyrin repeat domain 28/UDP-N-acetyl-alpha-D-galactosamine:	345488/-161925
								polypeptide N-acetylgalactosaminyltransferase-like 2	
50.rs10494410	1q23.3	1.7307293	1.27E-008	0.64	0.51	0.3571	NUF2/PBX1	Component of the evolutionarily conserved kinetochore-associated Ndc80	833972/-403365
								complex (Saccharomyces cerevisiae)/pre-B-cell leukemia transcription factor 1	

Table 19 Top 50 SNPs ranking by odd ratio (OR)

dbSNP RS ID	loci	OR mean	P-value	Case RAS	Control RAS	MAF	Gene symbol	Gene name or function	Gene region /distance from SNP
1.rs2155008	11q14.1	4.3154	3.73E-006	0.65	0.78	0.3214	NARS2/MGC33846	asparaginyl-tRNA synthetase 2 (mitochondrial)(putative)/	1616648/-2681011
								hypothetical protein MGC33846	
2.rs958842	9p23	3.8456	3.54E-005	0.55	0.42	0.2381	PTPRD/TYRP1	protein tyrosine phosphatase, receptor type, D/	1063998/-1016928
								tyrosinase-related protein 1	
3.rs1382393	18q21.32	3.3622	3.46E-008	0.55	0.38	0.4167	MC4R/CDH20	melanocortin 4 receptor/cadherin 20, type 2	312668/-806764
4.rs726336	5q34	3.3543	8.31E-007	0.68	0.52	0.4762	MAT2B/ODZ2	methionine adenosyltransferase II, beta/	1049097/-3021651
								odd Oz/ten-m homolog 2 (Drosophila)	
5.rs2326510	20p13	3.2691	4.25E-005	0.79	0.88	0.0595	PRND	prion protein 2 (dublet)	Promoter
6.rs10496865	2q22.1	3.1142	1.69E-005	0.57	0.69	0.0952	LRP1B	low density lipoprotein-related protein 1B (deleted in tumors)	Intron
7.rs10486922	7q21.11	3.1025	2.60E-006	0.7	0.56	0.5	MAGI2/GNAI1	membrane associated guanylate kinase, WW and PDZ domain containing 2/	2052680/-65086
								guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide1	
8.rs7148166	14q12	3.0916	3.77E-006	0.54	0.41	0.369	NUBPL	nucleotide binding protein-like	Intron
9.rs7138895	12q21.31	3.0659	4.79E-007	0.55	0.71	0.1786	PPFIA2/CCDC59	protein tyrosine phosphatase, receptor type, f polypeptide (PTPRF),	549924/-543650
								interacting protein (liprin),alpha2/coiled-coil domain containing 59	
10.rs768352	2q32.1	3.025	3.81E-007	0.23	0.12	0.131	NUP35/ZNF804A	nucleoporin 35kDa/zinc finger protein 804A	1340608/-133996
11.rs1112139	4q35.2	3.0012	4.11E-006	0.25	0.15	0.119	TRIML1/FRG1	tripartite motif family-like 1/FSHD region gene 1	1655586/-145790
12.rs1452882	6q21	2.9929	8.70E-006	0.61	0.49	0.369	GRIK2/HACE1	glutamate receptor, ionotropic, kainate 2/	2959547/-369729
								HECT domain and ankyrin repeat containing, E3 ubiquitin protein ligase 1	
13.rs3887103	4q13.1	2.9512	1.97E-007	0.52	0.69	0.4878	LPHN3/SRD5A2L2	latrophilin 3/steroid 5 alpha-reductase 2-like 2	1162804/-1617780
14.rs1494915	8q23.1	2.9438	2.54E-005	0.6	0.71	0.3333	ANGPT1/RSPO2	angiopoietin 1/R-spondin 2 homolog (Xenopus laevis)	610099/-39723
15.rs4981894	14q12	2.8797	3.38E-006	0.58	0.7	0.369	NUBPL	nucleotide binding protein-like	Intron
16.rs10520693	15q26.1	2.8673	1.53E-007	0.74	0.86	0.0952	CRTC3	CREB regulated transcription coactivator 3	Intron
17.rs10501899	11q22.1	2.8386	8.25E-006	0.24	0.35	0.1667	CNTN5	contactin 5; Neural recognition molecule NB-2	Intron
18.rs1024288	9q33.2	2.8385	2.83E-005	0.43	0.31	0.2738	OR1N2	olfactory receptor, family 1, subfamily N, member 2	Downstream
19.rs10511570	9p23	2.8279	6.14E-005	0.64	0.74	0.2619	PTPRD/TYRP1	protein tyrosine phosphatase, receptor type, D/tyrosinase-related protein 1	1137346/-943580
20.rs261717	Xq24	2.7989	1.01E-005	0.48	0.36	0.4762	ZCCHC12/LONRF3	zinc finger, CCHC domain containing 12/	27993/-122842
								LON peptidase N-terminal domain and ring finger 3	
21.rs10499231	6q24.2	2.7932	3.65E-005	0.34	0.23	0.0952	UTRN	utrophin (homologous to dystrophin)	Intron
22.rs892457	2p22.3	2.7746	1.20E-007	0.87	0.76	0.131	FAM98A/CRIM1	family with sequence similarity 98, member A/	858993/-191676
								cysteine rich transmembrane BMP regulator 1 (chordin-like)	
23.rs1899759	3p14.2	2.7361	3.01E-006	0.63	0.5	0.4762	C3orf67/FHIT	chromosome 3 open reading frame 67/fragile histidine triad gene	674917/-335298

24.rs763582	4q24	2.7201	5.12E-008	0.44	0.29	0.3929	TACR3/CXXC4	tachykinin receptor 3/CXXC finger 4	167550/-711288
25.rs12081	15q15.1	2.7096	2.98E-011	0.3	0.49	0.4286	FLJ43339	FLJ43339	Downstream
26.rs9324372	1p22.2	2.7046	3.76E-007	0.76	0.87	0.0854	HFM1	HFM1, ATP-dependent DNA helicase homolog (S. cerevisiae)	Intron
27.rs4981122	14q12	2.7035	7.95E-006	0.56	0.44	0.369	NUBPL	nucleotide binding protein-like	Intron
28.rs1341126	6q21	2.6871	2.08E-005	0.46	0.33	0.1707	GRIK2/HACE1	glutamate receptor, ionotropic, kainate 2/	3078483/-250793
								HECT domain and ankyrin repeat containing, E3 ubiquitin protein ligase 1	
29.rs5958198	Xq25	2.6871	2.57E-007	0.72	0.85	0.0833	GRIA3	glutamate receptor, ionotrophic, AMPA 3	Intron
30.rs1459543	4q26	2.686	1.27E-005	0.5	0.37	0.3095	TRAM1L1/NDST3	translocation associated membrane protein 1-like 1/	643613/-304420
								N-deacetylase/N-sulfotransferase (heparan glucosaminyl) 3	
31.rs2323934	13q13.3	2.6585	2.40E-006	0.45	0.58	0.2805	UFM1/FREM2	ubiquitin-fold modifier 1/FRAS1 related extracellular matrix protein 2	176221/-160964
32.rs2077946	10q23.31	2.6378	4.75E-008	0.44	0.61	0.3571	MPHOSPH1	M-phase phosphoprotein 1	Intron
33.rs6461547	7p15.3	2.6257	5.57E-008	0.53	0.69	0.2976	SP8/SP4	Sp8 transcription factor/Sp4 transcription factor	487385/-158401
34.rs2132070	8q23.1	2.6238	4.65E-006	0.5	0.39	0.3333	ANGPT1/RSPO2	angiopoietin 1/R-spondin 2 homolog (Xenopus laevis)	609379/-40443
35.rs734584	14q21.1	2.6109	6.30E-008	0.74	0.87	0.0238	FBXO33/LRFN5	F-box protein 33/leucine rich repeat and fibronectin type III domain containing 5	807472/-1402333
36.rs2198683	18q21.32	2.6036	1.50E-006	0.46	0.33	0.378	MC4R/CDH20	melanocortin 4 receptor/cadherin 20, type 2	447153/-672279
37.rs10489732	1p12	2.6033	4.15E-005	0.61	0.51	0.3452	SPAG17/TBX15	sperm associated antigen 17/T-box 15	440033/-489345
38.rs10517191	4p15.1	2.595	6.21E-005	0.51	0.4	0.381	AC024132.7/PCDH7	AC024132.7/BH-protocadherin (brain-heart)	2698874/-814036
39.rs9308004	4q32.2	2.5774	2.10E-007	0.43	0.59	0.4268	AC093700.4/FSTL5	AC093700.4/follistatin-like 5	1472970/-996058
40.rs1480491	2q14.1	2.5679	6.61E-005	0.57	0.47	0.2738	DPP10/DDX18	dipeptidyl-peptidase 10/DEAD (Asp-Glu-Ala-Asp) box polypeptide 18	966173/-1686548
41.rs1861090	12p13.31	2.5665	1.48E-006	0.4	0.55	0.4881	CD69/KLRF1	CD69 molecule/killer cell lectin-like receptor subfamily F, member 1	24593/-50398
42.rs2292623	10q26.13	2.5641	2.30E-006	0.6	0.7	0.1786	PLEKHA1	pleckstrin homology domain containing,	Intron
								family A (phosphoinositide binding specific) member 1	
43.rs16901863	5q14.3	2.5639	2.11E-007	0.38	0.24	0.4024	EDIL3/COX7C	EGF-like repeats and discoidin I-like domains 3/cytochrome c oxidase subunit VIIc	1633184/-1042474
44.rs1479959	3p14.1	2.5616	3.00E-005	0.57	0.46	0.4762	ADAMTS9/AC121493.1	ADAM metallopeptidase with thrombospondin type 1 motif, 9 /AC121493.1	809920/-28653
45.rs1413042	13q13.3	2.5566	2.18E-008	0.4	0.58	0.4405	NHLRC3/LHFP	NHL repeat containing 3/lipoma HMGIC fusion partner	90300/-214278
46.rs9315986	13q14.11	2.5509	9.30E-005	0.3	0.41	0.2073	DNAJC15	DnaJ (Hsp40) homolog, subfamily C, member 15	Intron
47.rs228857	6q23.2	2.5321	1.21E-005	0.45	0.34	0.2381	TCF21/TBPL1	transcription factor 21/TBP-like 1	40417/-23608
48.rs2326365	6p25.2	2.5186	1.12E-006	0.4	0.27	0.131	FAM50B	family with sequence similarity 50, member B	Promoter
49.rs742997	Xq27.3	2.5184	8.16E-009	0.5	0.35	0.2619	SPANXN3/SLITRK4	SPANX family, member N3/SLIT and NTRK-like family, member 4	85392/-33987
50.rs727690	14q23.1	2.5149	4.17E-006	0.42	0.53	0.3214	RTN1	reticulon 1	Intron

We compared the results from this dataset with the results of recent four genome wide association (GWA) studies in Caucasian patients with SLE, one report in Japanese patients, and with candidate gene association studies that have been previously reported according to adjacent locations (Table 20). In summary, we selected SNPs within our chips which are closest to the total of 26 known candidate genes; *PTPN22, FCGR2A, CRP, FASL, TNFSF4, IL10, STAT4, CTLA-4, PDCD-1, PXK, BANK1, TNFa, TNFb, HLA, ITPR3, TNXB, TNFSAIP3, IRF5, c8orf13, BLK, MBL, FAS, KIAA1542, ITGAM, ITGAX* and *TYK2*. Our significant SNPs do not lie within any known candidate genes when used strict criteria with $p<0.0001$. However, there were two significant SNPs that were within known candidate genes if the cutting point was $p<0.001$. The first significant SNP was rs6849308 in intron of *BANK1* ranked 5,614th ($p=0.00022$). This SNP was far 25,007 bp from significant SNP rs10516487 of previous reports. Another SNP was rs7844834 in intron of *c8orf13* which was ranked 9,055th ($p=0.00093$) and away from rs2736340 and rs13277113 of precedent WGA approximately 57,826 bp and 63,039 bp, respectively. Furthermore, the present study confirmed the association of SNPs that were proximate to *TNFb, HLA, TNXB and TNFSAIP3* when used the criteria of $p<0.0001$ and also confirmed close SNPs to *TNFa, TNFb, TNFSAIP3* and *IRF5* if the p value <0.001.

Table 20 Comparing our study to known candidate genes

Known candidate genes							Our study						
Gene	Location	SNP	Position	P value	OR	Reference	Close SNP	Postion	Distance	Rank by P value	gene (location)	OR	P value (T-test)
PTPN22	1p13.2	rs2476601	114179091	5.2E-06	1.5	Harley JB, 2008	rs3811021	114158186	2091	14584	PTPN22 (3' UTR)	1.59	0.00475
							rs1217407	114195271	-2043	19027	PTPN22 (Intron)	1.39	0.01246
FCGR2A	1q23	rs1801274	159746369	6.8E-07	0.7	Harley JB, 2008	rs10494359	159730500	15869	56866	unknown	1.05	0.89288
							rs412965	159766988	-20619	58960	HSPA6 (Downstream)	1.06	0.99996
CRP	1q23	rs1205	157948857	0.05		Russell AI et al., 2004	rs1122081	157948815	42	44684	unknown	1.19	0.373
							rs1733133	157957951	-9094	14666	PTPRN2 (Intron)	1.38	0.00482
FASL	1q23	rs763110	17089412	0.024		Wu J et al., 2003	rs10499481	17082726	6686	27579	unknown	1.39	0.053
				0.014	1.5	Chen JY et al., 2005	rs7305447	17097477	-8065	31162	unknown	1.14	0.0868
TNFSF4	1q25.1	rs1234317	171454398	0.04		Cunninghame Graham DS et al., 2008	rs10489267	171436775	-17623	20301	TNFSF4 (Intron)	1.339	0.0158
		rs2205960	171458098	0.007			rs844656	171502458	48060	55986	unknown	1.162	0.851
		rs844644	171476118	6.8E-05									
		rs10912580	171523173	0.05			rs1112468	171520353	-2820	52581	unknown	1.132	0.691
							rs10494489	171553431	30258	27762	unknown	1.152	0.0543
IL10	1q31-32	rs12024653	205015165	<0.05		Gibson AW et al., 2005	rs10494878	204982432	32733	38892	MAPKAPK2 (Downstream)	1.1	0.212
		rs1800896	205013520	3.00E-02	1.5	Hirankarn N et al., 2006	rs10494879	205018827	-3662	49933	IL10 (Promoter)	1.14	0.574
		rs1800871	205013257	2.00E-02	1.9	Rood MJ et al., 1999							
		rs1800872	205013030										
STAT4	2q32.3	rs7574865	191672878	9E-14		Geoffrey Hom, 2008	rs3821236	191611003	61875	13952	STAT4 (Intron)	1.17	0.00407
							rs1551443	191704763	-31885	27618	STAT4 (Intron)	1.37	0.05323
CTLA-4	2q33	rs231775	204440959	3.00E-03	1.7	Ahmed S et al., 2001	rs231800	204415830	25129	21449	unknown	1.45	0.0195
				2.00E-03	1.2	Lee YH et al., 2005	rs1978594	204499714	-58755	27356	unknown	1.34	0.0513
PDCD-1	2q37	rs11568821	242442585	0.00001	2.6	Prokunina L et al., 2002	rs1330695	242374397	68188	40425	unknown	1.23	0.25
				0.0019	2.7	Velázquez-Cruz R et al., 2007	rs611448	242500246	-57661	33109	unknown	1.22	0.111
PXK	3p14.3	rs6445975	58345217	7.1E-09	1.3	Harley JB, 2008	rs9311671	58166473	178744	54897	DNASE1L3 (Intron)	1.08	0.80013
							rs753819	58589433	-244216	18015	FAM3D (Downstream)	1.36	0.01022
BANK1	4q24	rs10516487	102970099	0.0064		Kozyrev S, 2008	rs4327488	102878236	91863	51399	unknown	1.08	0.63671
							rs6849308	102995106	-25007	5614	BANK1 (Intron)	1.36	0.00022
TNFa	6p21	rs1800630	31650455	0.009	1.9	Hirankarn N et al., 2007	rs10517234	31648627	1828	6953	unknown	1.59	0.00041

							rs1996886	31650546	-91	37916	unknown	1.2	0.193
		rs1800629	31651010	<0.05	3.7	van der Linden MW et al., 2001	rs1996886	31650546	464	37916	unknown	1.2	0.193
				0.001	2.3	Parks CG et al, 2004	rs10521977	31659384	-8374	36274	DMD (Intron)	1.28	0.162
				0.0001	2.6	Correa PA et al., 2005							
		rs361525	31651080	0.02	3.6	Correa PA et al., 2005	rs1996886	31650546	534	37916	unknown	1.12	0.193
							rs10521977	31659384	-8304	36274	unknown	1.28	0.162
TNFb	6p21	rs909253	31648292	<0.0001	3.4	Kim TG et al., 1996	rs2208532	31642493	5799	1424	unknown	1.51	5.6E-06
				p<0.05		Takeuchi F et al., 2005	rs10517234	31648627	-335	6953	unknown	1.6	0.00041
HLA	6p21.32-33	rs2187668	32713862	3E-21		Harley JB, 2008; Geoffrey Hom, 2008	rs2227139	32521437	192425	29205	HLADRA (Downstream)	1.13	0.06621
							rs10484561	32773398	-59536	342	unknown	2.28	2.70E-07
		rs1270942	32026839	1.71E-51	2.4		rs707937	31838993	187846	24735	MSH5 (Downstream)	1.29	0.03445
							rs9296009	32222493	-195654	41850	PPT2 (promoter)	1.11	0.28868
		rs3131379	31829012	1.71E-52	2.4		rs2248462	31554775	274237	56725	unknown	1.1	0.88721
							rs707937	31838993	-9981	24735	MSH5 (Downstream)	1.29	0.03445
ITPR3	6p21.31	rs3748079	33696125	2.87 x 10-8	3.4	Oish T et al., 2008	rs855438	33693234	2891	40860	unknown	1.114	0.262
							rs831631	33698141	-2016	47176	CD59 (Intron)	1.096	0.463
TNXB	6p21.32	rs1009382	32134085	5.18E-06		Kamatani Y et al., 2008	rs10508768	32132383	1702	13059	ARHGAP12 (Downstream)	1.432	0.00323
							rs522397	32138205	-4120	19987	CHRM5 (Intron)	1.825	0.015
		rs3130342	32188124	9.30E-07	3.1		rs2378078	32178389	9735	1218	unknown	1.568	3.8E-06
							rs4821148	32191275	-3151	27806	LARGE (Intron)	1.511	0.0547
TNFSAIP3	6q23.3	rs13192841	138008907	5.4E-08	1.4	Musone SL et al., 2008	rs10496774	138006028	-2879	3183	unknown	1.664	4.5E-05
							rs981457	138010273	1366	51467	unknown	1.093	0.639
		rs2230926	138237759	0.0003	2		rs10519416	138229507	-8252	12131	unknown	1.251	0.00249
							rs2925216	138249226	11467	7484	unknown	1.714	0.00052
		rs6922466	138486623	0.0001	1.3		rs7684854	138479174	-7449	14422	unknown	1.281	0.00457
							rs9307766	138488014	1391	4262	unknown	1.577	0.0001
IRF5	7q32.1	rs10488631	128381419	2E-11		Harley JB, 2008; Geoffrey Hom, 2008	rs7792282	128336804	44615	10604	unknown	1.32	0.00158
							rs1594423	128546309	-164890	6683	unknown	1.55	0.00036
		rs729302	128356196	2.00E-10	0.8		rs7792282	128336804	19392	10604	unknown	1.32	0.00158
							rs1594423	128546309	-190113	6683	unknown	1.55	0.00036
		rs10279821	128470783	6.50E-09	0.8		rs7792282	128336804	133979	10604	unknown	1.32	0.00158
							rs1594423	128546309	-75526	6683	unknown	1.55	0.00036

		rs12537284	128505142	3.61E-19	1.5		rs7792282	128336804	168338	10604	unknown	1.32	0.00158
							rs1594423	128546309	-41167	6683	unknown	1.55	0.00036
c8orf13-BLK	8p23.1	rs2736340	11381382	4E-07	1.4	Geoffrey Hom, 2008	rs7844834	11323556	57826	9055	C8orf13 (Intron)	1.56	0.00093
							rs2252534	11422122	-40740	21192	BLK (Intron)	1.2	0.01866
		rs13277113	11386595	8E-08	1.4		rs7844834	11323556	63039	9055	C8orf13 (Intron)	1.56	0.00093
							rs2252534	11422122	-35527	21192	BLK (Intron)	1.2	0.01866
MBL	10q11.2-q21	rs1800450	54201241	1.00E-03	1.4	Lee YH et al., 2005	rs930508	54198304	2937	21330	MBL2 (Intron (boundary))	1.38	0.0191
							rs1444161	54206225	-4984	28565	unknown	1.24	0.0606
FAS	10q24	rs1800682	90739943	0.004		Kanemitsu S et al., 2002	rs160424	90739801	142	26791	unknown	1.26	0.0471
							rs160427	90740505	-562	15685	unknown	1.38	0.00612
		rs9333296	90752838	1.00E-02	5	Horiuchi T et al., 1999	rs7920305	90750205	2633	27720	FAS (Intron)	1.32	0.054
							rs2296600	90760419	-7581	56261	SRD5A2 (Intron)	1.11	0.864
		rs3218612	90757462	1.00E-02	5	Horiuchi T et al., 1999	rs7920305	90750205	7257	27720	unknown	1.32	0.054
							rs2296600	90760419	-2957	56261	FAS (Intron (boundary))	1.11	0.864
KIAA1542	11p15.5	rs4963128	579564	3.00E-10	0.8	Harley JB, 2008	rs217238	1938894	-1359330	24115	MRPL23 (Downstream)	1.13	0.0311
ITGAM-ITGAX	16p11.2	rs9937837	31206440	7E-07	1.3	Geoffrey Hom, 2008	rs1364184	29563711	1642729	26203	BOLA2 (Intron)	1.16	0.04328
							rs2141349	31637053	-430613	42326	ZNF720 (Intron)	1.06	0.30199
		rs11574637	31276375	5E-07	1.3		rs1364184	29563711	1712664	26203	BOLA2 (Intron)	1.16	0.04328
							rs2141349	31637053	-360678	42326	ZNF720 (Intron)	1.06	0.30199
ITGAM	16p11.2	rs9888739	31220754	1.61E-23	1.6	Harley JB, 2008	rs1364184	29563711	1657043	26203	BOLA2 (Intron)	1.16	0.04328
							rs2141349	31637053	-416299	42326	ZNF720 (Intron)	1.06	0.30199
		rs1143678	31250506	8.50E-14	1.4		rs1364184	29563711	1686795	26203	BOLA2 (Intron)	1.16	0.04328
							rs2141349	31637053	-386547	42326	ZNF720 (Intron)	1.06	0.30199
		rs4548893	31271994	2.36E-12	1.3		rs1364184	29563711	1708283	26203	BOLA2 (Intron)	1.16	0.04328
							rs2141349	31637053	-365059	42326	ZNF720 (Intron)	1.06	0.30199
		rs1143679	31184312	1.70E-17	1.8		rs1364184	29563711	1620601	26203	BOLA2 (Intron)	1.16	0.04328
							rs2141349	31637053	-452741	42326	ZNF720 (Intron)	1.06	0.30199
TYK2	19p13.2	rs2304256	10336652	5.6E-05		Sigurdsson S et al., 2005	rs1413288	10336388	264	14971	unknown	1.3	0.00515
							rs206437	10338361	-1709	26749	unknown	1.18	0.0469
		rs12720356	10330975	0.00015		Sigurdsson S et al., 2005	rs1580004	10326996	3979	45799	unknown	1.16	0.411
							rs10488130	10334408	-3433	20905	unknown	1.25	0.0178

*highlight indicated significant SNPs $P<0.0001$ or $P<0.001$

2. Candidated genes in chromosome 1

We focus on chromosome 1 since several regions (1q23, 1q25-31 and 1q41-42) have been confirmed as major susceptibility loci to SLE development. In our study, 118 SNPs lying on these loci were selected for further analysis. We only interested in SNPs that were in the gene regions (promoter (5'UTR), exon, intron, and 3'UTR) based on biological reasons that these SNPs might be functional SNP. From 118 significant SNPs, 48 SNPs from 26 genes were selected (Table 20). There were 8 genes that seem to be important in SLE including *CD1A, CD1D, CD244, NOS1AP, NPHS2, PYHIN1, RGS16* and *TLR5*. However, we particularly choose putative functional SNPs according to computational predictions. There were only 4 SNPs that reached the last criteria. The first SNP was rs10489821 which was in the upstream region of *CD1D* gene and away from exon 1 approximately 5278 bp. TFSEARCH program predicted that this SNP could change transcription factor USF binding. The function of this *CD1D* gene is presenting glycolipid to T cell receptor of NKT cell and lead to NKT cell activation (Zajonc DM and Kronenberg, 2007). Several reports have suggested that *CD1D* might play an immunoregulatory role in the development of SLE. For example, a study has shown that CD1D deficiency exacerbates lupus nephritis in mice induced by the hydrocarbon oil pristane (Yang JQ et al., 2003). The next two SNPs in intron were rs1858233 of *NOS1AP* (1,083 bp away from exon 5) and rs856084 of *PYHIN1* (308 bp away from exon 7). They were predicted that the variants can affect acceptor splice site and SC35 regulatory factor binding site, respectively. *NOS1AP* is nitric oxide synthase 1 (neuronal) adaptor protein which interacts with NOS1 gene associated with end stage renal disease (ESRD) and lupus nephritis (Freedman BI et al., 2000; Vazgiourakis V et al., 2007). For *PYHIN1* or *IFIX*, it is a member of Interferon-inducible p200 (IFI200) family. The function of IFI200 family is cell cycle regulation (Ludlow LEA et al. 2005). Several evidences have supported the importance of this gene in SLE. For instance, data from lupus murine model has identified IFN-inducible protein 202 gene (ifi202) as a candidate for lupus susceptibility (Rozzo SJ et al., 2001; Choubey D et al., 2002). This gene is homologous to human IFN-inducible genes (*MNDA, IFIX, IFI16* and *AIM2*) and located on chromosome 1 similar to human. Another one was nonsynonymous SNP rs2072493 resulting in asparagine to serine alteration at residue 592 of the TLR5 protein. *TLR5* is a

member of the Toll-like receptor (TLR) family that play a key role in the activation and regulation of both innate and adaptive immune responses (Rahman AH and Eisenberg RA, 2006). The details of genes and selected SNPs on chromosome 1 were shown in Table 21 and the number of SNPs that passed each criteria was summarized in Table 22.

Table 21 The details of genes and selected SNPs on chromosome 1

dbSNP Id	loci	OR mean	P-value	Case RAS	Control RAS	MAF	Gene	function	Gene region	Functional SNP
s432611	q23.1	1.4373193	2.40E-006	0.29	0.35	0.0595	CD1A	THYMOCYTE ANTIGEN CD1A; present lipid to T cell; T cell activation	Promoter (-9161)	do not affected
s10489821	q23.1	1.5956104	1.34E-005	0.54	0.62	0.3929	CD1D	THYMOCYTE ANTIGEN CD1D; present glycolipid to NKT cell; NKT cell activation	Promoter (-5278)	C-, G: USF
s3766377	q23.3	1.4002778	4.15E-005	0.42	0.36	0.2976	CD244	NATURAL KILLER CELL RECEPTOR 2B4	3' UTR	do not affected
s10495272	q42.13	1.3669937	3.18E-005	0.75	0.8	0.1548	CDC42BPA	CDC42-BINDING PROTEIN KINASE ALPHA	Intron (boundary)	-
s10495265	q42.13	1.9619179	3.14E-006	0.33	0.23	0.2619	CDC42BPA	CDC42-BINDING PROTEIN KINASE ALPHA	Intron	-
s1022059	q23.3	1.403413	9.63E-005	0.37	0.31	0.4048	CDCA1	CELL DIVISION CYCLE-ASSOCIATED 1	Intron	-
s10495298	q42.2	1.6489747	4.53E-006	0.23	0.31	0.1071	COG2	COMPONENT OF OLIGOMERIC GOLGI COMPLEX 2	Intron	-
s6698107	q23.3	1.4489556	5.19E-005	0.34	0.28	0.1071	DDR2	DISCOIDIN DOMAIN RECEPTOR FAMILY, MEMBER 2	Intron	-
s4038287	q23.3	1.2870232	1.64E-005	0.61	0.67	0.1786	DDR2	DISCOIDIN DOMAIN RECEPTOR FAMILY, MEMBER 2	Intron	-
s499385	q42.12	1.6148655	5.03E-005	0.22	0.16	0.0952	ENAH	ENABLED, DROSOPHILA, HOMOLOG OF	Intron	-
s10489327	q25.1	1.2647954	5.75E-005	0.4	0.45	0.25	KIAA0040	-	Intron	-
s1395548	q23.3	1.445604	3.31E-005	0.4	0.34	0.2262	LMX1A	LIM HOMEOBOX TRANSCRIPTION FACTOR 1	Intron	-
s3737297	q41	2.01211	2.26E-005	0.35	0.26	0.131	MARK1	MAP/MICROTUBULE AFFINITY-REGULATING KINASE 1	Intron (boundary)	-
s4131748	q41	1.5929934	1.95E-005	0.45	0.53	0.4405	MOSC1	MOCO sulphurase C-terminal domain containing 1	Intron	-
s1330224	q25.3	1.6253811	7.98E-005	0.32	0.26	0.2857	NMNAT2	NICOTINAMIDE NUCLEOTIDE ADENYLYLTRANSFERASE 2	Intron	-
s953274	q25.3	1.5358908	2.10E-005	0.7	0.76	0.2857	NMNAT2	NICOTINAMIDE NUCLEOTIDE ADENYLYLTRANSFERASE 2	Intron	-
1932933	q23.3	1.30413	8.78E-005	0.69	0.65	0.3571	NOS1AP	NITRIC OXIDE SYNTHASE 1 (NEURONAL) ADAPTOR PROTEIN; interact with NOS1 associated with ESRD, lupus nephritis	Intron	do not affected
1858232	q23.3	1.3418209	8.05E-005	0.58	0.52	0.4881	NOS1AP	NITRIC OXIDE SYNTHASE 1 (NEURONAL) ADAPTOR PROTEIN; interact with NOS1 associated with ESRD, lupus nephritis	Intron	do not affected
347311	q23.3	1.8853574	9.73E-005	0.35	0.26	0.1548	NOS1AP	NITRIC OXIDE SYNTHASE 1 (NEURONAL) ADAPTOR PROTEIN; interact with NOS1 associated with ESRD, lupus nephritis	Intron	do not affected
1858233	q23.3	1.6470619	1.74E-008	0.53	0.43	0.4405	NOS1AP	NITRIC OXIDE SYNTHASE 1 (NEURONAL) ADAPTOR PROTEIN; interact with NOS1 associated with ESRD , lupus nephritis	Intron	C: Strengthens acceptor site
347282	q23.3	1.4600694	1.33E-005	0.48	0.41	0.2262	NOS1AP	NITRIC OXIDE SYNTHASE 1 (NEURONAL) ADAPTOR PROTEIN; interact with NOS1 associated with ESRD, lupus nephritis	Intron	do not affected
1410592	q25.2	1.3088916	1.37E-005	0.51	0.45	0.4881	NPHS2	PODOCIN; esrd, podocyte	Coding exon	A/A (Alanine/Alanine)
2087121	q42.13	1.5049364	1.36E-005	0.55	0.47	0.4524	OBSCN	OBSCURIN	Intron	-
7534667	q42.2	1.5597591	1.94E-005	0.57	0.48	0.5	PCNXL2	pecanex-like 2 (Drosophila)	Intron	-
7541396	q42.2	1.3935846	6.04E-006	0.25	0.31	0.2262	PCNXL2	pecanex-like 2 (Drosophila)	Intron	-

s7550169	42.2	1.4638134	5.93E-005	0.29	0.37	0.25	PCNXL2	ecanex-like 2 Droso hila	Intron	-
s4655345	41	2.0114131	3.73E-005	0.51	0.41	0.4762	PTPN14	PROTEIN-TYROSINE PHOSPHATASE, NONRECEPTOR-TYPE, 14	Intron	-
s10494976	41	1.6416457	6.76E-005	0.33	0.25	0.1429	PTPN14	PROTEIN-TYROSINE PHOSPHATASE, NONRECEPTOR-TYPE, 14	Intron	-
s10494979	41	1.6500708	1.94E-005	0.47	0.38	0.2857	PTPN14	PROTEIN-TYROSINE PHOSPHATASE, NONRECEPTOR-TYPE, 14	Intron	-
s856111	q23.1	1.6327233	6.90E-005	0.65	0.72	0.1429	PYHIN1	pyrin and HIN domain family, member 1; IFIX (Intereron-inducible gene)	Promoter (-9864)	do not affected
s856084	q23.1	1.3329507	9.31E-006	0.38	0.44	0.2857	PYHIN1	pyrin and HIN domain family, member 1; IFIX (Intereron-inducible gene)	Intron	T: Abolishes SC35 binding site
s10494566	25.3	1.4990793	5.82E-006	0.83	0.87	0.131	RGL1	ral uanine nucleotide dissociation stimulator-like 1	Intron	-
s10494567	25.3	1.7494743	2.11E-005	0.78	0.85	0.131	RGL1	ral guanine nucleotide dissociation stimulator-like 1	Intron	-
s10494568	25.3	1.6029457	6.42E-005	0.73	0.81	0.1071	RGL1	ral uanine nucleotide dissociation stimulator-like 1	Intron	-
s10494569	25.3	1.4973532	3.71E-005	0.26	0.2	0.0714	RGL1	ral uanine nucleotide dissociation stimulator-like 1	Intron	-
s3010062	q25.3	1.5447787	1.15E-005	0.32	0.41	0.4762	RGL1	ral guanine nucleotide dissociation stimulator-like 1	Intron	-
s10489966	q25.3	1.7488614	3.58E-005	0.45	0.36	0.1786	RGS16	regulator of G-protein signalling 16;	Promoter (-1982)	do not affected
								may regulate T lymphocyte activation		
s516567	q25.3	1.8514174	3.75E-008	0.54	0.66	0.3452	RGS16	regulator of G-protein signalling 16;	Promoter (-5641)	do not affected
								ma regulate T lymphocyte activation		
s6678820	23.3	1.3234436	6.06E-005	0.39	0.34	0.1667	RGS5	re ulator of G- rotein si nallin 5; associated with arterial wall disease	Intron	-
s4307613	41	1.7656219	8.30E-006	0.43	0.33	0.131	SPATA17	s ermato enesis associated 17 ;novel a o totic ene in s erm	Intron	-
s10495071	41	1.7613885	9.23E-006	0.33	0.24	0.122	SPATA17	s ermato enesis associated 17 ;novel a o totic ene in s erm	Intron	-
s10495073	41	1.4134948	1.42E-005	0.35	0.29	0.131	SPATA17	s ermato enesis associated 17 ;novel a o totic ene in s erm	Intron	-
s6703319	41	1.5224253	5.04E-005	0.39	0.31	0.1429	SPATA17	s ermato enesis associated 17 ;novel a o totic ene in s erm	Intron	-
s7525540	41	1.5480864	1.02E-006	0.42	0.34	0.25	SPATA17	s ermato enesis associated 17 ;novel a o totic ene in s erm	Intron	-
s2072493	q41	1.4705562	4.88E-006	0.29	0.35	0.2619	TLR5	toll-like receptor 5; activation and regulation of both innate and adaptive immunity	Coding exon	N/S (Asparagine/Serine)
s2255781	41	1.7685512	6.69E-006	0.51	0.61	0.4643	USH2A	Usher s ndrome 2A autosomal recessive, mild	Intron	-
s301748	41	1.487409	5.45E-005	0.15	0.11	0.0119	USH2A	Usher s ndrome 2A autosomal recessive, mild	Intron	-
s1323023	42.13	1.4692737	6.19E-006	0.67	0.6	0.1786	WNT3A	win less- e MMTV inte ration site famil , member 3A	Promoter	-

Table 22 The number of SNPs that passed each criteria

Number of SNPs	Criteria of selected SNPs
3814	passed QC, p<0.0001, OR>1.25
358	on chromosome 1
118	within major susceptibility loci (1q23, 1q25-31 and 1q41-42)
48	within gene regions (promoter(5'UTR), exon,intron,3'UTR)
14	8 good candidate genes related with SLE pathogenesis
4	putative functional SNPs according to bioinformatic tools

Genomic DNA from the same individuals was individually genotyped for 4 SNPs (see in Table 23). In the present study, allele frequencies from individual genotyping correlated with the pooling results with Pearson correlation coefficient of 0.991 and 0.950 for case and control, respectively. Moreover, the averaged SEM differences between two groups were small (0.0175). For association analysis, only two SNPs (rs1858233 and rs856084) of *NOS1AP* and *PYHIN1 (IFIX)* were confirmed the association with SLE, with OR (95%CI) = 1.56(1.03-2.36), p=0.035 and 1.79(1.15-2.77), p=0.009, respectively. These two SNPs ranked as 72nd and 1759th by *P* value from pooling results. The *P* value and OR were 0.0000000174 and 1.65 for rs1858233 of *NOS1AP*; 0.00000931 and 1.33 for rs856084 of *PYHIN1 (IFIX)*. For other two SNPs, no significant association was found in rs2072493 of TLR5 and rs10489821 of CD1D. However, the allele frequency patterns of both SNPs from individual genotyping were similar to the patterns from pooling approach. The *P* value and OR from pooling results were 0.00000488 and 1.47 for rs2072493 of *TLR5*; 0.0000134 and 1.60 for rs1048921 for *CD1D*. These SNPs ranked as 1351st and 2012th by *P* value from pooling data.

Table 23 Comparison between allele frequencies from pooling approach and individual genotyping

Genes	Ranking, P value, OR from pooling	SNPs	Allele	Allele frequency of case		Allele frequency of control		Odd ratio (95% CI), P value*
				pooling	individual	pooling	individual	
NOS1AP	72, 0.0000000174, 1.65	rs1858233 (C/T)	C[a]	0.53	0.52	0.43	0.41	[a]1.56(1.03-2.36), p=0.035
			T	0.47	0.48	0.57	0.59	
PYHIN1 (IFIX)	1759, 0.00000931, 1.33	rs856084 (G/T)	G	0.38	0.28	0.44	0.41	[b]1.79(1.15-2.77), p=0.009
			T[b]	0.62	0.72	0.56	0.59	
TLR5	1351, 0.00000488, 1.47	rs2072493 (C/T)	C	0.29	0.2	0.35	0.26	1.41(0.86-2.31), p=0.191
			T	0.71	0.8	0.65	0.74	
CD1D	2012, 0.0000134, 1.60	rs10489821(C/G)	C	0.46	0.44	0.38	0.42	1.09(0.7-1.71), p=0.769
			G	0.54	0.56	0.62	0.58	
Pearson correlation				0.991		0.95		p<0.0001
SEM				0.06062	0.0879	0.05701	0.06543	Averaged SEM = 0.06774
SEM differences				0.027		0.008		Averaged difference = 0.0175

[a] and [b] are comparison of allele frequencies between case and control from individual genotyping

* P value was calculated by Chi-square test

3. SNPs within PYHIN1 (IFIX) group contained on Affymetrix 50K Array Xba240

Since PYHIN1 (IFIX) showed strongest significant association with SLE from individual genotyping data, we focus on SNPs within PYHIN1 (IFIX) group which compose of *MNDA, IFIX, IFI16* and *AIM2* genes spanning approximately 300 kb. The strong association signal was in MNDA-IFIX intergenic region and within IFIX gene (Table 24). However, we cannot exclude the importance of *MNDA, IFI16* and *AIM2* genes, because there was no SNP from our chips in AIM2 and only one SNP in MNDA and IFI16 genes. In addition, we also summarized haplotype block figure (Figure 11).

Table 24 SNPs within PYHIN1 (IFIX) group contained on Affymetrix 50K Array Xba240

dbSNP RS ID	Cytoband	Gene	Physical Position	OR mean	P-value (T-test)	Case RAS	Control RAS	Rank by P value
rs2261161	q23.1	mnda	157085948	1.397085	9.86E-002	0.43	0.4	32118
rs6689517	q23.1	mnda-IFIX	157098994	1.180359	0.962	0.580	0.58	58197
rs2106092	q23.1	mnda-IFIX	157099068	1.696732	9.5E-06	0.350	0.27	1770
rs2188114	q23.1	mnda-IFIX	157100374	1.219154	0.27	0.310	0.33	41187
rs1102024	q23.1	mnda-IFIX	157149776	1.226506	0.045	0.880	0.86	26466
rs1102025	q23.1	mnda-IFIX	157150039	1.449814	5.07E-006	0.24	0.19	1369
rs856111	q23.1	IFIX	157163461	1.632723	6.90E-005	0.65	0.72	3742
rs6675945	q23.1	IFIX	157168968	1.408863	4.06E-003	0.72	0.77	13921
rs856084	q23.1	IFIX	157181765	1.332951	9.31E-006	0.38	0.44	1758
rs861319	q23.1	IFIX	157182541	1.176561	0.424	0.720	0.71	46138
rs1101991	q23.1	ifi16	157259892	1.277457	0.0783	0.530	0.5	30375

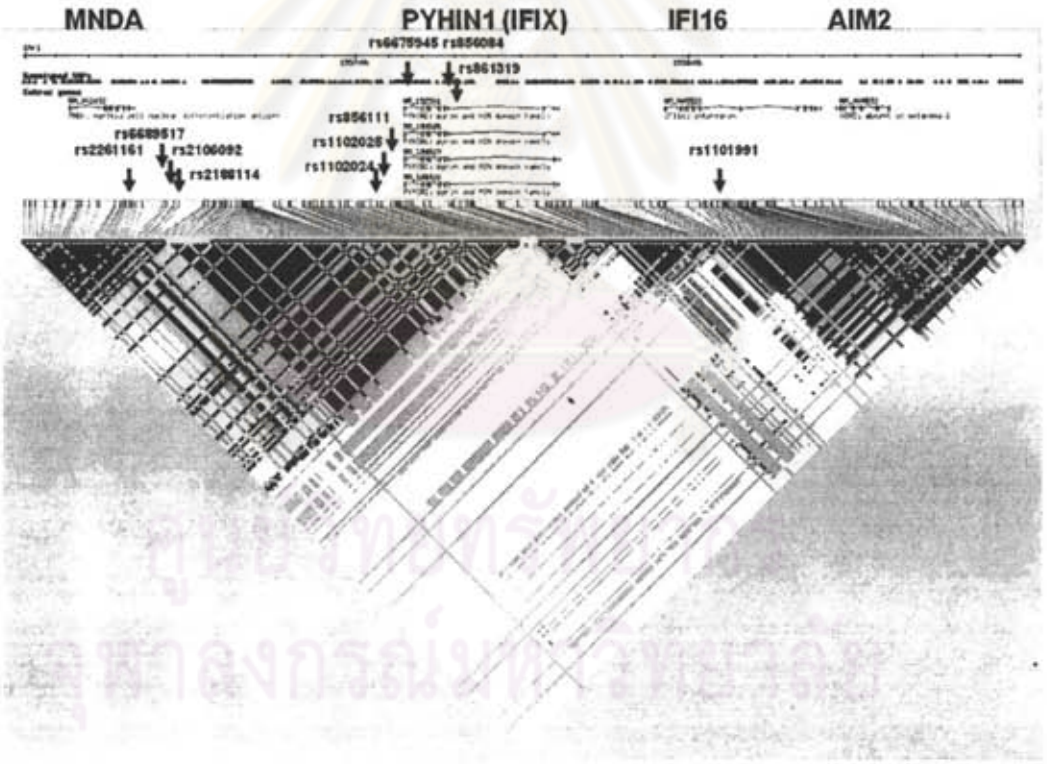


Figure 11 SNPs within IFI region contained on the Affymetrix 50K Array Xba240 were shown. Linkage disequilibrium structure (D') was generated in Haploview based on data from Chinese and Japanese populations. The darker red color stands for higher LD. Blue arrows indicate associated SNPs ($p<0.01$), while red arrows are negative SNPs.

Part II Candidate gene association study

1. Selected SNPs

We choose 10 SNPs from *MNDA*, *IFIX*, *IFI16* and *AIM2* genes according to computational prediction as shown in Table 25.

Table 25 Summary of selected SNPs in *MNDA*, *IFIX*, *IFI16* and *AIM2* genes and the predicted impacts of SNPs.

Region	Gene and SNP	Position of SNP	Predicted impact	MAF
Promoter	1. *MNDA*	-	-	
	2. *IFI16*			
	2.1 rs4657618	C-7217T	T: AML1a or RUNX1 / C: -	0.41
	3. *AIM2*			
	3.1 rs16841642	G-151T	G: AP-1 / T:-	0.22
Exon	1. *MNDA*			
	1.1 rs2276403	C16432T	H/Y (Histidine/Tyrosine)	0.06
	2. *IFI16*			
	2.1 rs866484	C6771G	T/S (Threonine/Serine)	0.33
	2.2 rs6940	A44962T	S/T (Serine/Threonine) (damaging/tolerated)	0.21
	3. *AIM2*			
	3.1 rs2276405	C3452T	E/K (Glutamic acid/Lysine) (damaging/tolerated)	0.07
Intron	1. *MNDA*			
	1.1 rs7513873	A2,706G	A:Strengthens donor site; Creates cryptic acceptor site	0.14
	2. *IFI16*			
	2.1 rs1772414	A23, 201G	A:Abolishes donor site	0.31
	2.3 rs3754460	A27,140G	A:Strengthens SF2-ASF binding site	0.21
	3. *IFIX*			
	3.1 rs856084	G13792T	T:Abolishes SC35 binding site	0.37
	3. *AIM2*	-	-	
3'UTR	-	-	-	-

2. LD analysis and comparison between Asians (Chinese and Japanese) and Caucasian from Haploview for IFI region (*MNDA, IFIX, IFI16* and *AIM2*)

LD analysis of the 238 SNPs in IFI region for Asians populations (Chinese and Japanese) and 236 SNPs for Caucasian was shown in Figure 12A and 13A, respectively. The pattern of LD was displayed as a colored plot. The darker red color represents regions of higher pairwise D'. In analyzed results, we found 17 haplotype blocks across 4 genes in Asian and only 12 haplotype blocks in Caucasian (Figure 12B and 13B). These data indicated the difference of LD patterns between two populations.

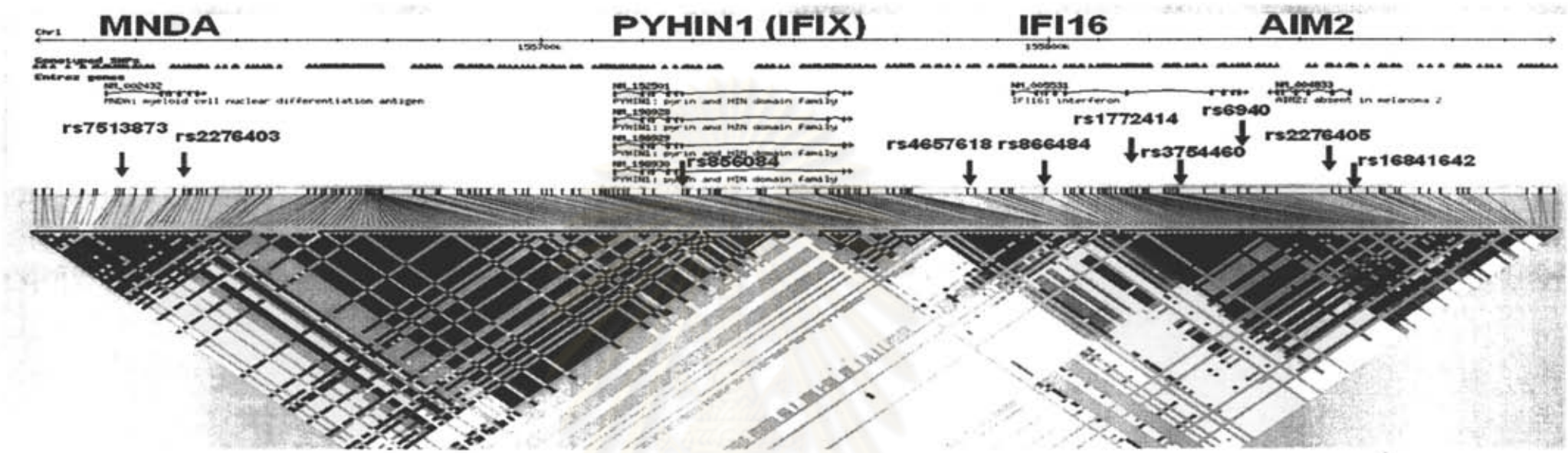


Figure 12A LD of SNPs in *MNDA, IFIX, IFI16* and *AIM2* genes for Asains (with 10 positions of selected SNPs)

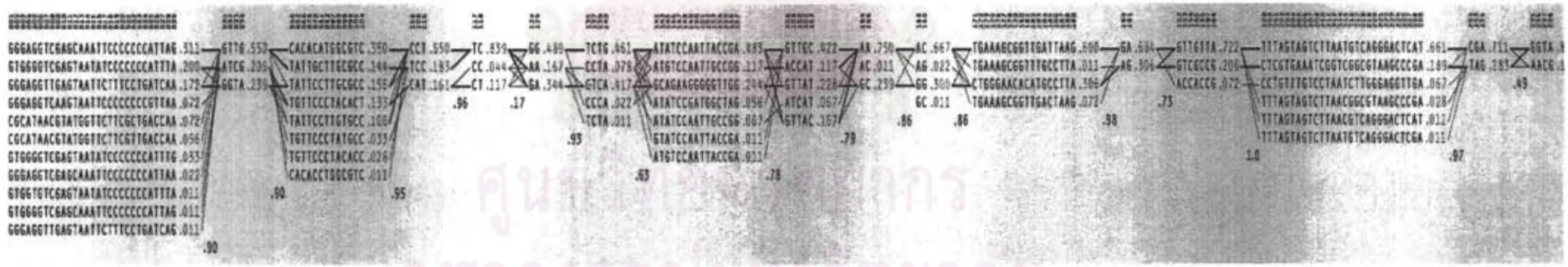


Figure 12B Haplotype blocks of *MNDA, IFIX, IFI16* and *AIM2* genes for Asians

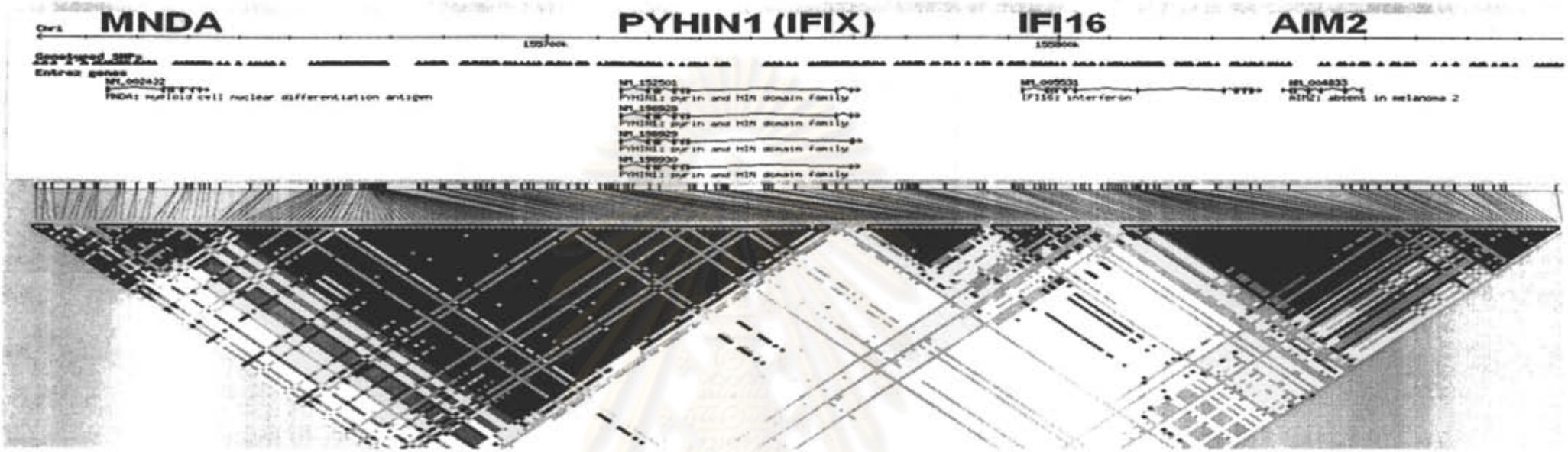


Figure 13A LD of SNPs *MNDA, IFIX, IFI16* and *AIM2* genes for Caucasian

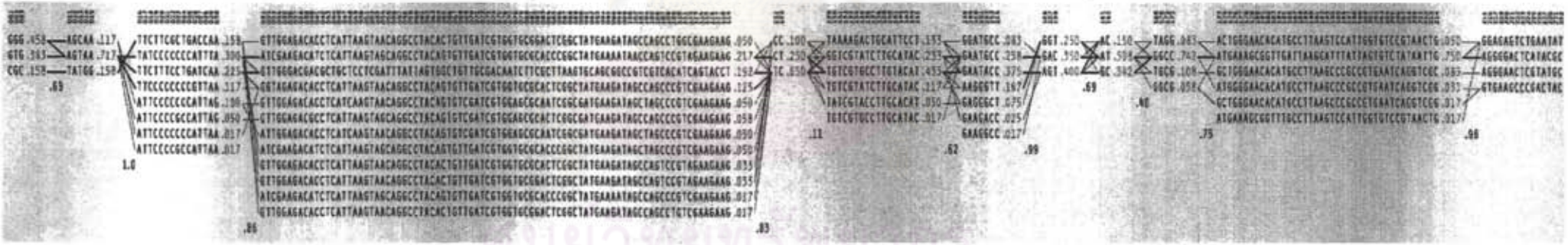


Figure 13B Haplotype blocks of *MNDA, IFIX, IFI16* and *AIM2* genes for Caucasian

3. The results of SNPs analysis

In the present study, polymorphism could be found in all positions unless rs2276405 of *AIM2* gene. In addition, two SNPs (rs6940 and rs3754460 in *IFI16* gene) were not in Hardy-Weinberg equilibrium when compared the observed and expected genotype frequencies of each SNP ($p<0.05$). Thus, these 3 SNPs were excluded from statistical analysis.

3.1 Allelic association test

In our study, there were significant associations from three SNPs including rs856084 in *IFIX* gene, rs866484 and rs1772414 in *IFI16* gene (Table 26). The G allele of *IFI16* at position C6771G (rs866484) and A23201G (rs1772414) was found to be significantly increased in SLE patients compared with healthy controls (OR = 1.37, 95% CI = 1.03–1.82, $P = 0.039$; OR = 1.41, 95% CI = 1.06–1.88, $P = 0.02$, respectively). In another significant SNP, the T allele of SNP G13792T (rs856084) in *IFIX* showed smaller significant difference with *P* value of 0.043; OR = 1.37 and 95% CI = 1.01-1.87. In this study, we did not find any allelic significant association of SNPs in *MNDA* and *AIM2* genes.

Table 26 Allelic association test for SNPs of *MNDA*, *IFIX*, *IFI16* and *AIM2* genes and SLE

Genes	SNP	Physical position	Allele	Allele frequency		Odd ratio (95% CI)	*P* value	Empirical *P* value
				SLE	Control			
MNDA	rs7513873	157070497	A	0.1256	0.1156	1.10 (0.71-1.70)	0.670	0.729
			G	0.8744	0.8844			
	rs2276403	157084223	T	0.09231	0.08333	1.12 (0.68-1.85)	0.662	0.694
			C	0.9077	0.9167			
IFI1X	rs856084	157181765	G	0.2974	0.3683	1.37 (1.01-1.87)	0.038	0.043
			T*	0.7026	0.6317			
IFI16	rs4657618	157239114	T	0.4128	0.4113	1.01 (0.75-1.34)	0.966	1
			C	0.5872	0.5887			
	rs866484	157253101	G*	0.5	0.422	1.37 (1.03-1.82)	0.031	0.039
			C	0.5	0.578			
	rs1772414	157269531	G*	0.5128	0.4274	1.41 (1.06-1.88)	0.018	0.020
			A	0.4872	0.5726			
AIM2	rs16841642	157313422	T	0.4333	0.414	1.08 (0.81-1.44)	0.589	0.608
			G	0.5667	0.586			

*** The asterisk indicated risk allele**

3.2 Modeling the pattern of inheritance

By model of inheritance analysis, the association of both SNPs (rs866484 and rs1772414) in *IFI16* gene seems to be recessive effect model. The presence of two G alleles (GG) conferred the significant *P* value of 0.009 for rs866484 and *P* value of 0.013 for rs1772414 (Table 27).

Table 27 Modeling the pattern of inheritances in SLE and controls

Gene	SNP	Minor a	Major A	SLE (N=200)			Control (N=200)			*P* value of model of inheritance	
				Aa	Aa	AA	Aa	Aa	AA	Recessive	Dominant
MNDA	rs7513873	A	G	1	50	149	1	45	154	NA	NA
	rs2276403	T	C	1	37	162	0	32	168	NA	NA
IFIX	rs856084	G	T	15	88	97	27	92	81	0.050	0.107
IFI16	rs4657618	T	C	32	101	67	34	98	68	0.788	0.916
	rs866484	G	C	57	87	56	35	103	62	0.009	0.511
	rs1772414	G	A	57	93	50	36	104	60	0.013	0.263
AIM2	rs16841642	T	G	41	94	65	35	102	63	0.444	0.830

3.3 Conditional analysis of significant SNPs

In conditional logistic analysis, we tested all three significant SNPs conditioning each other to determine which SNPs made the strongest contribution to the association (Table 28). In this study, we still found significant association of rs866484 (*P* value from 0.01271 to 0.01291) and rs1772414 (*P* value from 0.003302 to 0.006033) when conditioning on rs856084. Similarly, when we tested rs856084 conditioning on rs866484 and rs1772414, *P* values were not disappeared. From the results, it seems to be independent effects of each gene contributing the positive association.

Table 28 Conditional logistic analysis for all SNPs with single-marker allelic p<0.05

Gene	SNP	*P* value		*P* value		*P* value	
		Without conditional	conditional on rs856084	Without conditional	conditional on rs866484	Without conditional	conditional on rs1772414
IFIX	rs856084	NA	NA	0.01291	0.01271	0.006033	0.003302
IFI16	rs866484	0.01271	0.01291	NA	NA	0.7412	0.2677
	rs1772414	0.003302	0.006033	0.2677	0.7412	NA	NA

3.4 Pairwise linkage disequilibrium (LD) analysis

In this study, pairwise linkage disequilibrium (LD) among 7 SNPs was measured by different parameters, D' and r^2. There was no complete LD among all SNPs. However, strongest LD was found between two significant SNPs of *IFI16* (rs866484 and rs1772414) with D'=0.7995 and r^2 = 0.6253 (Table 29). For another significant SNP rs856084, there was small LD between rs866484 (D'=0.2078 and r^2 = 0.0345) and rs1772414 (D'=0.2974 and r^2 = 0.0690).

Table 29 Linkage disequilibrium coefficients (D'and r^2) among SNPs within *MNDA, IFIX, IFI16* and *AIM2*

		D'						
		rs7513873 (MNDA)	rs2276403 (MNDA)	rs856084 (IFIX)	rs4657618 (IFI16)	rs866484 (IFI16)	rs1772414 (IFI16)	rs16841642 (AIM2)
r^2	rs7513873	-	0.8078	0.0083	0.1759	0.4036	0.2959	0.2646
	rs2276403	0.4539	-	0.1401	0.2675	0.1899	0.0485	0.0464
	rs856084	0.00001	0.0044	-	0.4522	0.2078	0.2974	0.2563
	rs4657618	0.0040	0.0134	0.1707	-	0.3913	0.4834	0.3125
	rs866484	0.0203	0.0034	0.0345	0.1465	-	0.7995	0.6770
	rs1772414	0.0107	0.0004	0.0690	0.2187	0.6253	-	0.7794
	rs16841642	0.0090	0.0004	0.0542	0.0966	0.4434	0.5748	-

3.5 Haplotype analysis for rs856084, rs866484 and rs1772414

In the present study, we performed haplotype analysis only significant SNPs: rs856084 in *IFIX* gene, rs866484 and rs1772414 in *IFI16* gene. There were 4 common haplotypes (MHF >= 0.05) including GGG, TGG, GCA and TCA. The TCA was the mostly frequent haplotype (39.37%) followed by TGG (18.91%), GGG (18.51%) and GCA (13.11%). To test the association of haplotype and SLE development, we compared each tested haplotype with other 3 haplotypes between SLE patients and controls. Our finding demonstrated that the TGG as a risk haplotype with OR of 1.88 and $P = 0.001$, whereas GCA was displayed as protective haplotype with OR of 0.44 and P=0.004 (See in Table 30).

Table 30 Haplotype analysis for rs856084, rs866484 and rs1772414

Haplotype	Haplotype frequency		OR	P value
	SLE	Control		
GGG	0.1761	0.1851	0.92	0.710
TGG[a]	0.2845	0.1891	1.88	0.001
GCA[b]	0.07712	0.1311	0.44	0.004
TCA	0.3706	0.3937	0.90	0.490

[a] and [b] indicated risk and protective haplotypes, respectively.

Moreover, we also performed haplotype analysis of SNPs with strongest LD (rs866484 vs rs1772414). The GG and CA were the mostly frequent haplotypes as shown in Table 5. To test the association of *IFI16* haplotype and SLE development, we compared each tested haplotype with other three haplotypes between SLE patients and controls. Our finding demonstrated that the GG as a risk haplotype with OR of 1.41 and $P = 0.017$, whereas CA was displayed as protective with OR of 0.73 and P=0.032.

Table 31 Haplotype analysis for rs866484 and rs1772414 of IFI16 gene

HAPLOTYPE	Haplotype frequency		OR	P value
	SLE	Control		
GG	0.4594	0.3739	1.41	0.017
CG	0.05342	0.0535	0.99	0.996
GA	0.0406	0.04812	0.83	0.614
CA	0.4466	0.5245	0.73	0.032

3.6 Conditional haplotype-based association testing

To test whether we still see the association if we control for haplotypes of SNP. In our results, we found the decreasing of global haplotype effect when conditional on all three SNPs. This suggests that all three SNPs are important in haplotype association.

Table 32 Conditional haplotype-based association testing

Tested SNPs	*P* value by conditioning
Global	0.003715
rs856084	0.02864
rs866484	0.02748
rs1772414	0.04468

3.7 Clinical manifestation analysis

In this study, we obtained clinical data of 148 patients with SLE as shown in Table 31. There was a significant association between the G allele of rs866484 and rs1772414 with arthritis (OR = 1.77, 95% CI = 0.99-3.18, *P* = 0.039, *Pc*=0.054 and OR = 1.86, 95% CI = 1.04-3.32, *P* = 0.025, *Pc*=0.035, respectively). The effect of the G allele was similar to the autosomal dominance in which the presence of one G allele (CG compared to CC for rs866484 and AG compared to AA for rs1772414) conferred the significant OR of 4.00, 95% CI = 1.44-11.32, *P* = 0.003, *Pc*=0.005 and OR = 3.24, 95% CI = 1.20-8.85, *P* = 0.009, *Pc*=0.018, respectively. Nevertheless, significant value was slightly declined when compared CG and GG with CC genotype for rs866484 (OR = 3.16, 95% CI = 1.33-7.55, *P* = 0.004, *Pc*=0.007); and compared AG and GG with AA for rs1772414 (OR = 3.10, 95% CI = 1.27-7.59, *P* = 0.005, *Pc*=0.010). Genotype and allele frequencies of SNP rs866484 and rs1772414 of *IFI16* gene in SLE patients with and without arthritis were shown in Table 32.

Table 33 Clinical manifestation of patients with SLE (n=148)

Clinical manifestation	Number of patients with SLE (%)
Malar Rash	107 (72.30)
Discoid Rash	59 (39.86)
Photo-sensitivity	60 (40.54)
Oral/Nasal Ulcers	68 (45.95)
Arthritis	113 (76.35)
Pleurisy or Pericarditis	8 (5.41)
Renal Disorder	99 (66.89)
Neurologic Disorder	14 (9.46)
Hematologic Disorder	108 (72.97)

Table 34 Genotype and allele frequencies of SNP rs866484 and rs1772414 of *IFI16* gene in SLE patients with and without arthritis

Gene	SNP	Genotype	Arthritis (%) N=113	No arthritis (%) N=35
IFI16	rs866484	CC	26 (23)	17 (48.57)
		CG	55 (48.67)[a,b]	9 (25.71)
		GG	32 (28.32)[b]	9 (25.7)
		C	107 (47.35)	43 (61.43)
		G	119 (52.65)[c]	27 (38.57)
	rs1772414	AA	22 (19.47)	15 (42.86)
		AG	57 (50.44)[d,e]	12 (34.29)
		GG	34 (30.09)[e]	8 (22.86)
		A	101 (44.69)	42 (60)
		G	125 (55.31)[f]	28 (40)

[a] CG compared with CC genotype: OR = 4.00, 95% CI = 1.44-11.32, *P* = 0.003, *Pc*=0.005.

[b] GG+CG compared with CC genotype: OR = 3.16, 95% CI = 1.33-7.55, *P* = 0.004, *Pc*=0.007.

[c] G compared with C allele: OR = 1.77, 95% CI = 0.99-3.18, *P* = 0.039, *Pc*=0.054.

[d] AG compared with AA genotype: OR = 3.24, 95% CI = 1.20-8.85, *P* = 0.009, *Pc*=0.018.

[e] GG+AG compared with AA genotype: OR = 3.10, 95% CI = 1.27-7.59, *P* = 0.005, *Pc*=0.010.

[f] G compared with A allele: OR = 1.86, 95% CI = 1.04-3.32, *P* = 0.025, *Pc*=0.035.

3.4 Power detection

We calculated the power based on our sample size (N=200 per group), disease prevalence of 60 patients per 100,000 people in Chinese (Mok CC and Lau CS, 2003), and OR of 1.37. In this calculation, we have the power of 59-61% with α = 0.05 for a multiplicative model.

4. Functional characterization

4.1 The effects of SNP rs1772414 (A23201G) of IFI16 and rs856084 (G13792T) of IFIX on the ratio of splice isoforms

To test the effect of SNP rs1772414 genotypes, we measured the ratio of short product (B or C isoforms) to full-length product (A isoform) in leukocyte from patients with SLE containing different genotypes (AA, AG and GG). There was no significant difference between the ratio of B to A isoforms of IFI16 in each genotype (Figure 14). While comparison between the ratio of C to A isoforms found the high correlation of C isoform with the G allele, but only AG genotype reached statistical significance ($p=0.008$) when compared to AA genotype (Figure 15). For IFIX, the ratio of short product (β isoforms) to full-length product (α isoform) did not differ among genotypes (Figure 16).

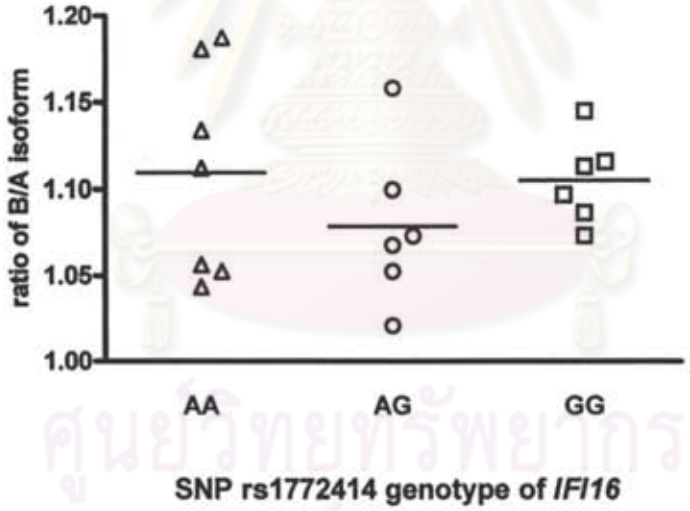


Figure 14 The ratio of short product (B isoform)/full-length product (A isoform) grouped by SNP rs1772414 genotype of *IFI16* gene

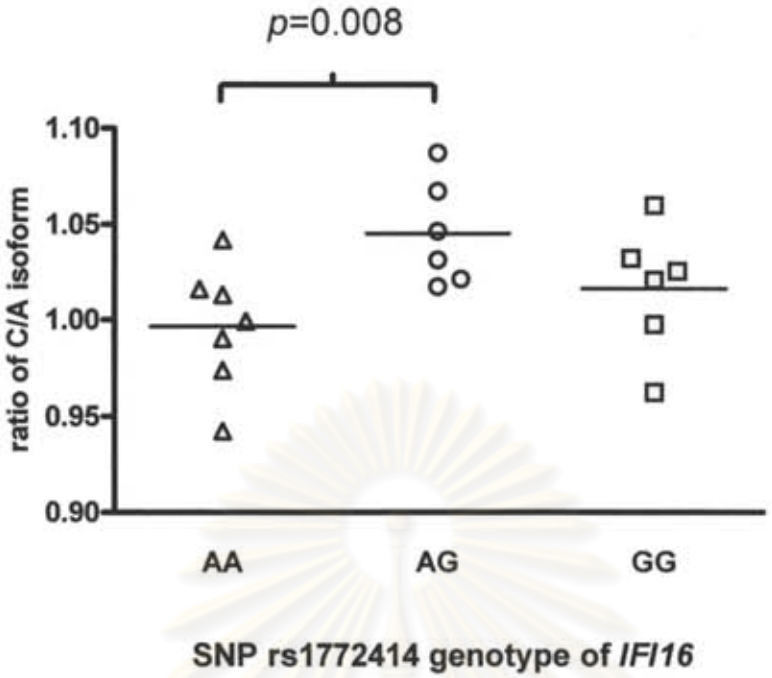


Figure 15 The ratio of short product (C isoform)/full-length product (A isoform) grouped by SNP rs1772414 genotype of *IFI16* gene

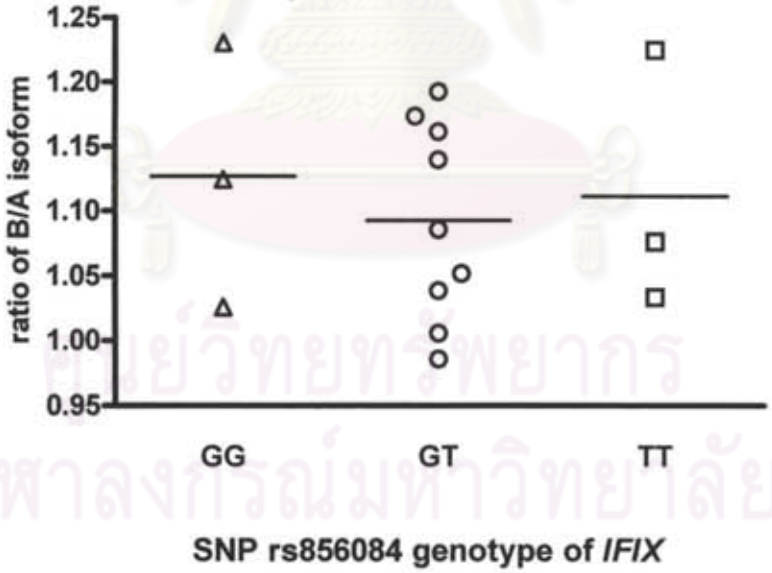


Figure 16 The ratio of short product (β isoform)/full-length product (α isoform) grouped by SNP rs856084 genotype of *IFIX* gene

4.2 The effects of non-synonymous SNP rs866484 (C6771G) of *IFI16* in apoptosis

Since SNP rs866484 of *IFI16* gene altering amino acid at position 179 from threonine to serine was found to be associated with SLE susceptibility and clinical manifestation. In addition, this SNP is located on p53 binding site of *IFI16* gene and the p53 is known to be important in the regulation of cell cycle and apoptosis (Choubey D et al., 2008). Hence, the association of SNP rs866484 and apoptosis in PBMC treated with p53 inducing agent from healthy individuals containing different genotypes was investigated.

In this study, the condition was determined to be effective to induce highest p53 expression in dose (1, 10 and 100 µg/ml) and time (8 and 24 hours) response experiments. Our finding showed that PBMCs treated with doxorubicin at a concentration of 10 µg/ml for 24 hr induced highest p53 mRNA expression (Figure 17). Moreover, the association of SNP rs866484 and apoptosis was studied using flow cytometrical analysis (Figure 18). In the present study, there was no significant difference between any *IFI16* genotypes and apoptosis (Figure 19).

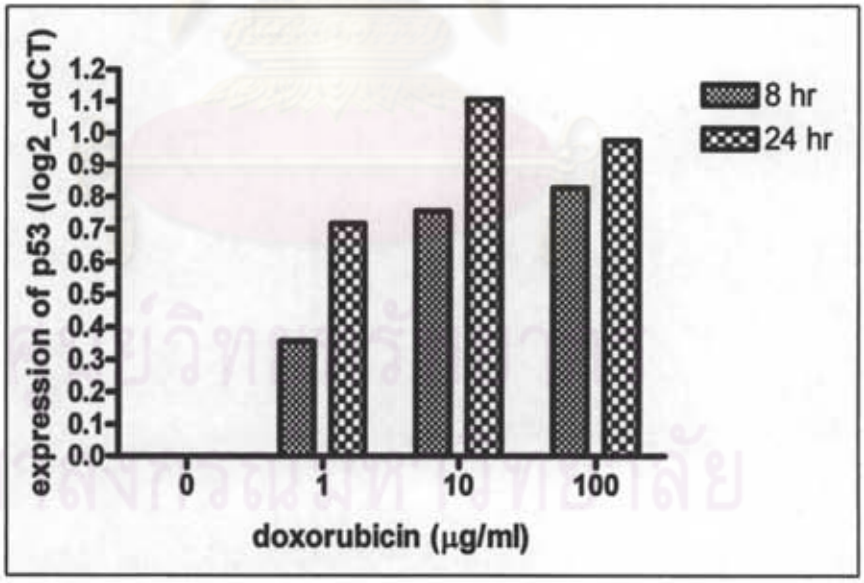


Figure 17 Relative expression of p53 mRNA (log2_ddCT) in PBMCs after 8 and 12 hrs of treatment with 0, 1, 10 and 100 µg/ml doxorubicin

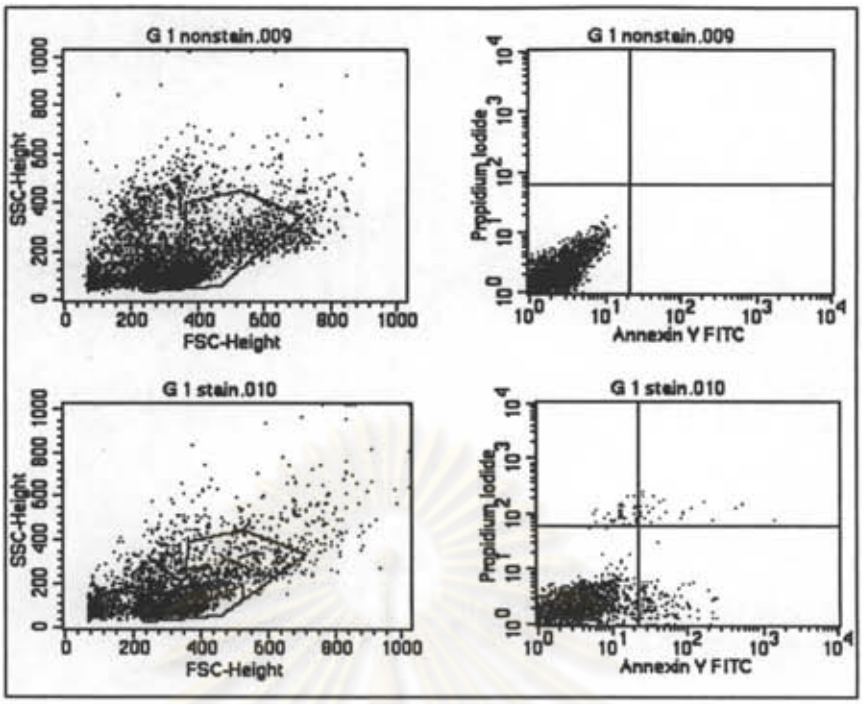


Figure 18 Flow cytometrical analysis of apoptosis as quantified by Annexin V positivity in PBMCs incubated with 10 µg/ml doxorubicin for 24 hr.

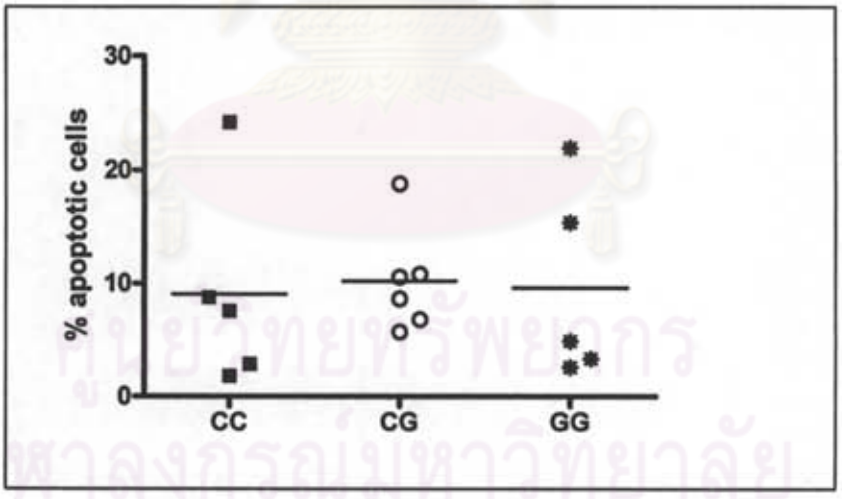


Figure 19 The effect of SNP rs866484 genotypes (CC, CG and GG; N=5, 6 and 5, respectively) on apoptotic response of healthy PBMCs treated with 10 µg/ml doxorubicin for 24 hr.

4.3 Impact of *AIM2* promoter SNP rs2276405 at position G-151T in binding of nuclear proteins and mRNA expression

The binding of nuclear proteins to *AIM2* promoter at position G-151T by EMSA

According to TFSEARCH program, it predicts that SNP of AIM2 gene at position G-151T could change transcription factor AP-1 binding as shown in Table 1. Interestingly, previous study has shown that AP-1 play an important role in regulating the expression of IFI16, which is a member of the HIN-200 family similar to AIM2 (Clarke CJP et al., 2003). Thus, our study aim to characterize this polymorphism within the promoter region of the AIM2 gene that might modulate transcription factor AP-1 binding. Although SNP G-151T of AIM2 was not associated with SLE susceptibility, it might be useful for studying in other diseases. The binding of nuclear proteins to AIM2 promoter at position G-151T was shown in Figure 20. Nuclear protein extracts could bind to both G and T alleles. Preliminary EMSA result suggested that there was different affinity between G and T alleles since the shift band can be more abolished by 50-fold by unlabelled G allele probe than unlabelled T allele probe.

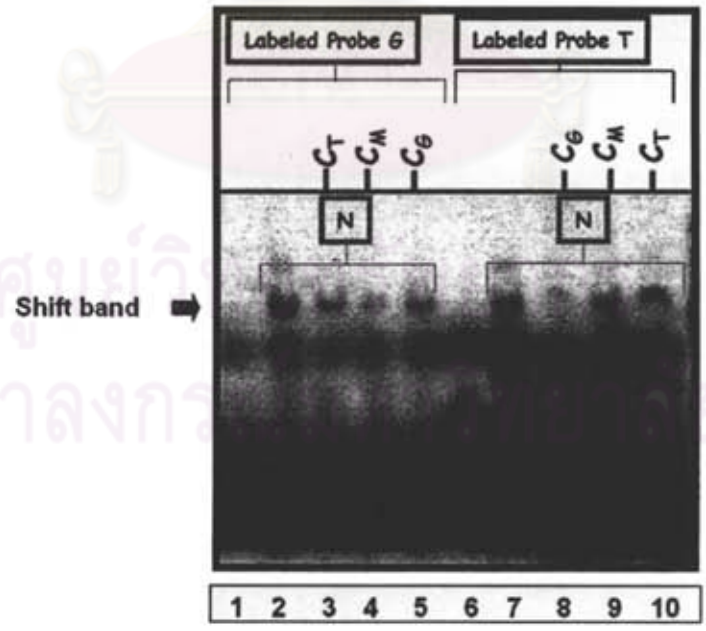


Figure 20 EMSA of SNP at position G-151T in *AIM2* promoter. Lanes 1 and 6 show the mobilities of the labeled oligonucleotides without nuclear extracts; lanes 2 and 7 present the mobilities of the labeled oligonucleotides with nuclear extracts in the absence of competitor. Specificity of nuclear protein binding is demonstrated by competition with 50-fold molar excess of unlabeled oligonucleotides containing the T allele (Lanes 3 and 10), G allele (Lanes 5 and 8) and mutant allele (Lanes 4 and 9).

The effect of SNP G-151T in *AIM2* mRNA expression

SNPs in promoter region usually led to different mRNA level which result from changing transcription factor binding site. In this study, we characterized mRNA level of *AIM2* gene in LPS-stimulated B cells from individuals with different genotype (GG, GT, or TT). It seems that B lymphocytes from individuals carrying GG and GT can be stimulated to express *AIM2* more than TT with statistical significance, $P = 0.034$ (Figure 21). However, please note that we only have 2 samples of TT genotype.

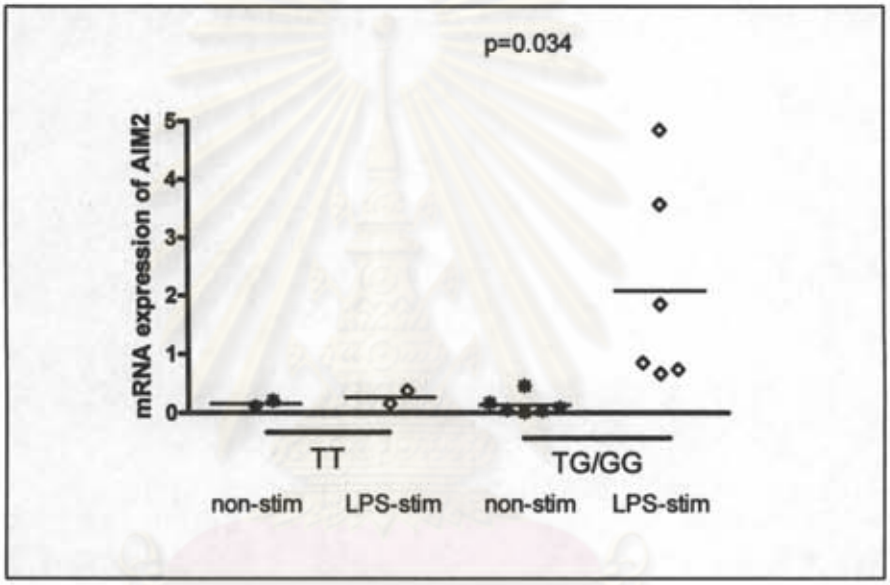


Figure 21 Effect of SNP G-151T on expression of the AIM2 gene in unstimulated B lymphocytes (non-stim) and B lymphocytes stimulated with 10 mg/ml lipopolysaccharide (LPS-stim). Stars and diamonds represent individual samples, and bars represent the mean level of expression. The y-axis shows the expression levels normalized to Beta-actin gene.

Part III Expression profile study

To further understand the role of these genes in SLE, we studied the expression of MNDA, IFIX, IFI16 and AIM2 genes from various cell types (leukocytes, B lymphocytes, T lymphocytes and renal biopsy tissue) in patients with SLE compared to healthy controls.

1. mRNA expression of *MNDA, IFIX, IFI16* and *AIM2* in leukocytes

The mRNA expression of MNDA, IFIX, IFI16 and AIM2 genes was significantly increased in leukocytes (buffy coat) from patients with SLE (N=20) as compared to normal controls (N=9) with p=0.0005, 0.011, 0.0008 and 0.0003, respectively (Figure 22-25). In addition, we also measured the relative amount of IFI16 and IFIX isoforms in leukocytes from patients with SLE (N=20) and normal controls (N=9). In our results, we found the B and C isoforms (short isoforms) of IFI16 from patients with SLE were significantly higher than normal controls with p=0.015 and p=0.049, respectively (Figure 26). For IFIX, β isoform (short isoform) was found to be decreased in patients with SLE as compared to normal controls (p=0.034) (Figure 27).

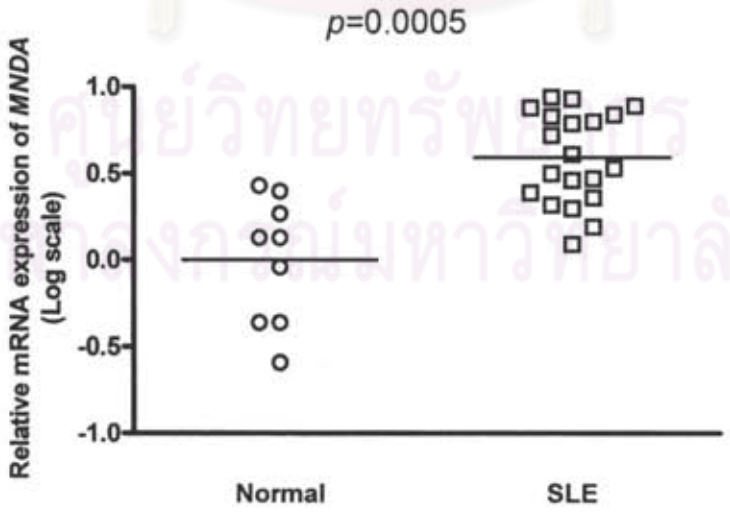


Figure 22 Level of *MNDA* mRNA in leukocytes from 20 patients with SLE and 9 normal controls.

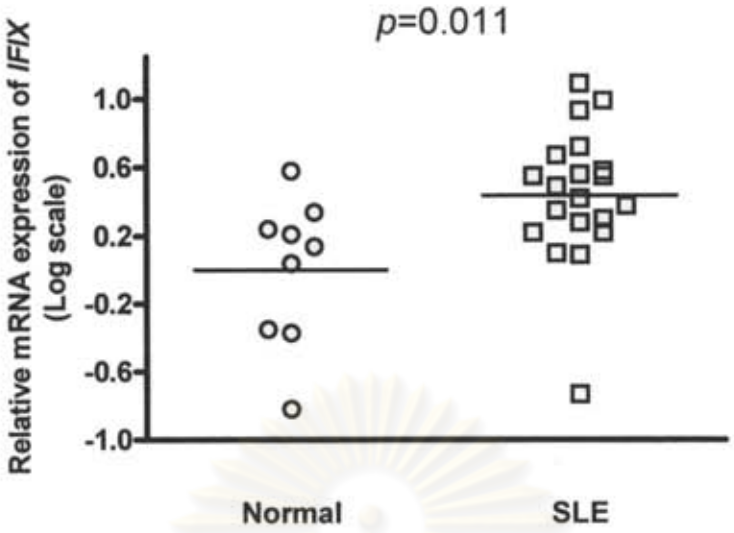


Figure 23 Level of *IFIX* mRNA in leukocytes from 20 patients with SLE and 9 normal controls.

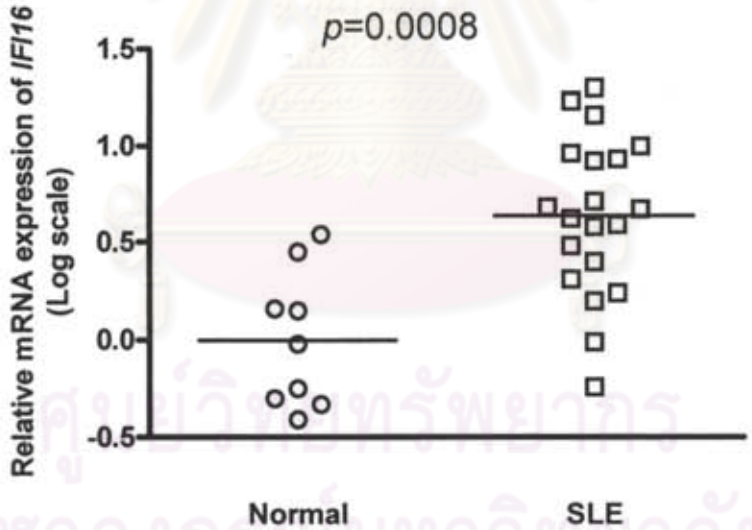


Figure 24 Level of *IFI16* mRNA in leukocytes from 20 patients with SLE and 9 normal controls.

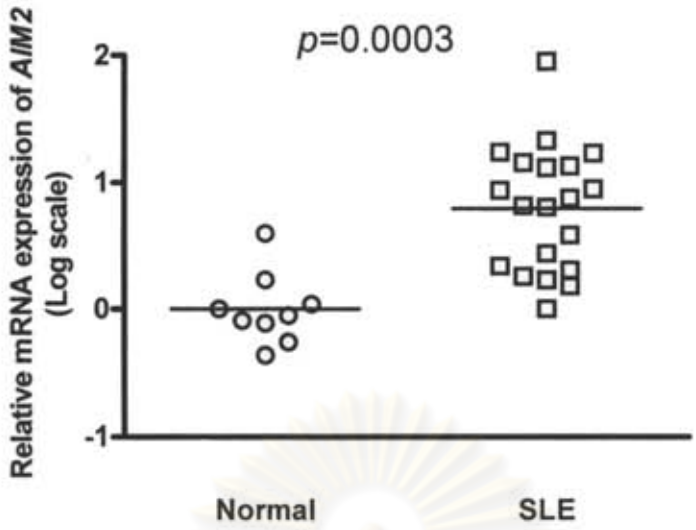


Figure 25 Level of *AIM2* mRNA in leukocytes from 20 patients with SLE and 9 normal controls.

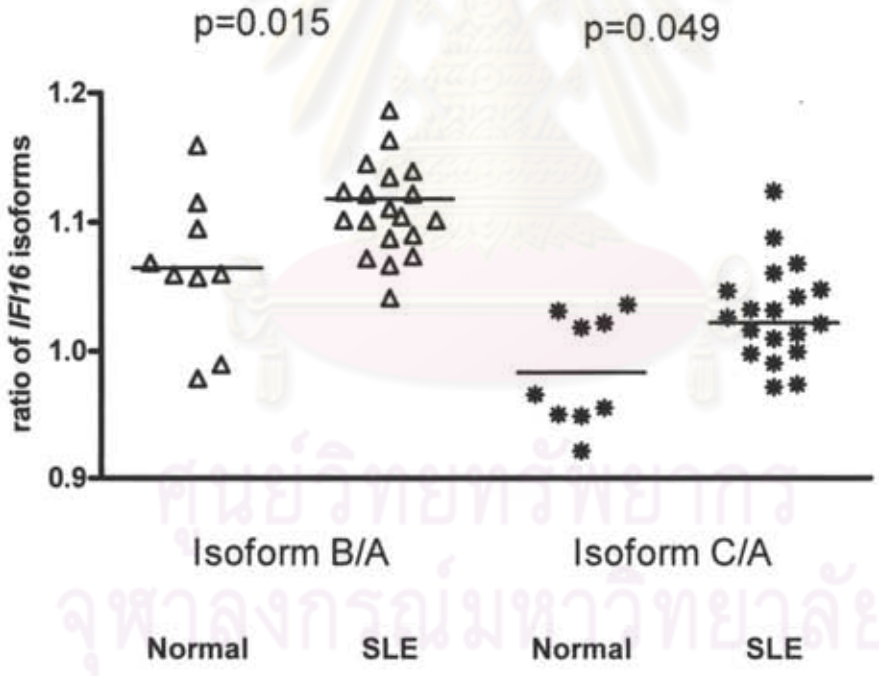


Figure 26 The ratio of short product (B and C isoform)/full-length product (A isoform) of *IFI16* gene in leukocytes from 20 patients with SLE and 9 normal controls

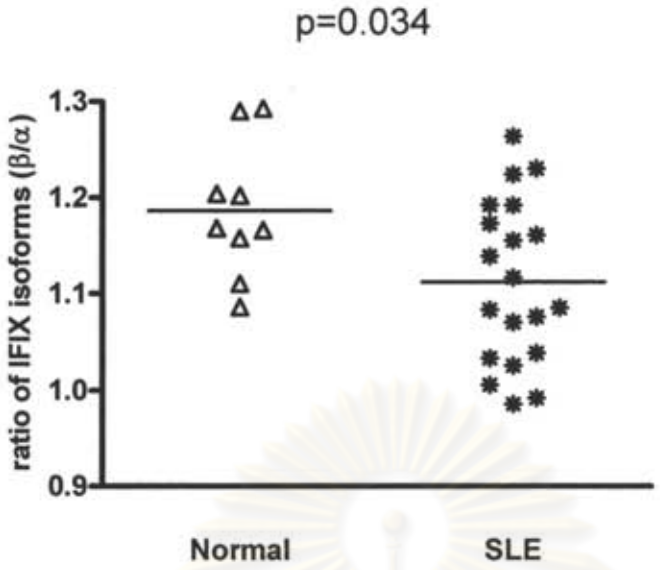


Figure 27 The ratio of short product (β isoform)/full-length product (α isoform) of *IFIX* gene in leukocytes from 20 patients with SLE and 9 normal controls

2. mRNA expression of *MNDA, IFIX, IFI16* and *AIM2* in B and T lymphocytes

The expression in B and T cells was determined using quantitative real-time RT-PCR. In this study, we did not found significant differences of all 4 genes in both subpopulations between SLE and control groups (Figure 28-31).

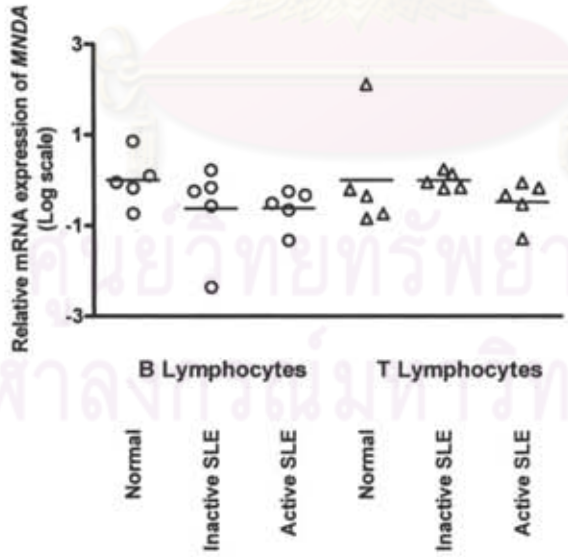


Figure 28 Level of *MNDA* mRNA in B and T lymphocytes from 10 patients with SLE (5 inactive and 5 active SLE) and 5 normal controls.

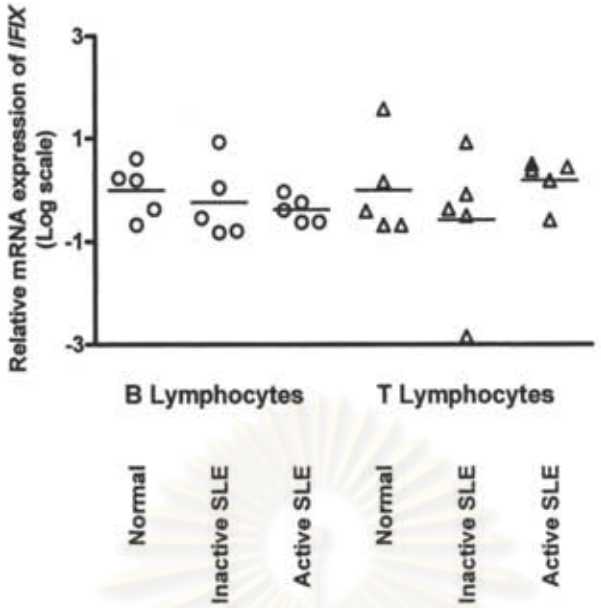


Figure 29 Level of *IFIX* mRNA in B and T lymphocytes from 10 patients with SLE (5 inactive and 5 active SLE) and 5 normal controls.

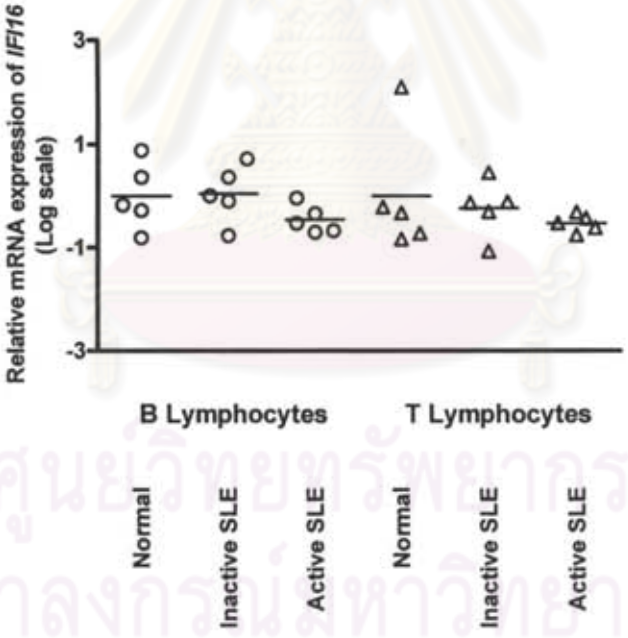


Figure 30 Level of *IFI16* mRNA in B and T lymphocytes from 10 patients with SLE (5 inactive and 5 active SLE) and 5 normal controls.

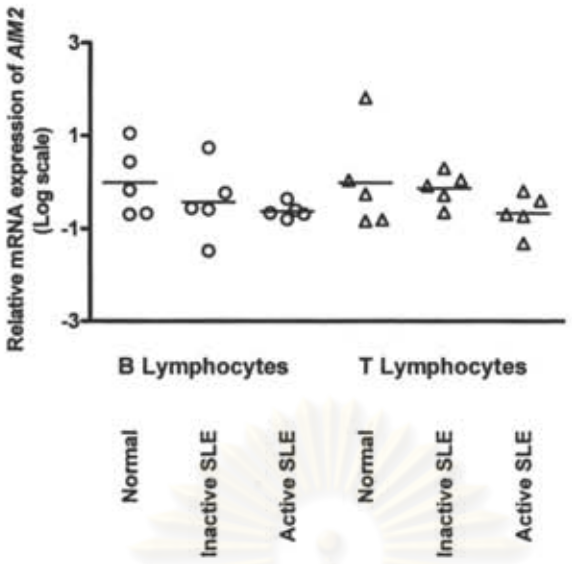


Figure 31 Level of *AIM2* mRNA in B and T lymphocytes from 10 patients with SLE (5 inactive and 5 active SLE) and 5 normal controls.

3. mRNA expression of *MNDA*, *IFIX*, *IFI16* and *AIM2* in kidney

There were no significant differences of these genes expression in renal biopsy tissues from SLE patient with renal manifestation and normal controls (Figure 32-35).

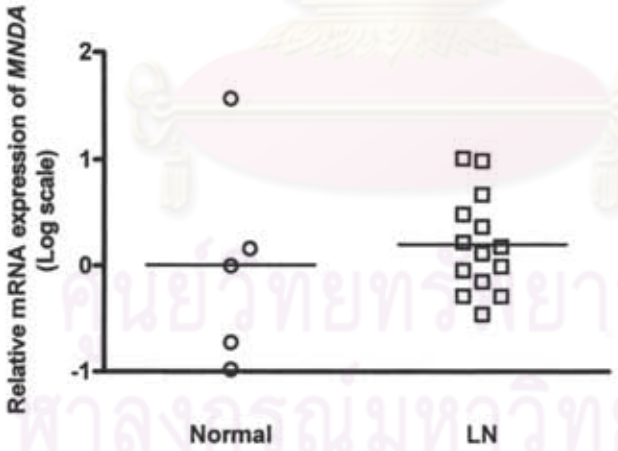


Figure 32 Level of *MNDA* mRNA in the renal biopsies from 14 lupus nephritis (LN) patients and 5 normal controls.

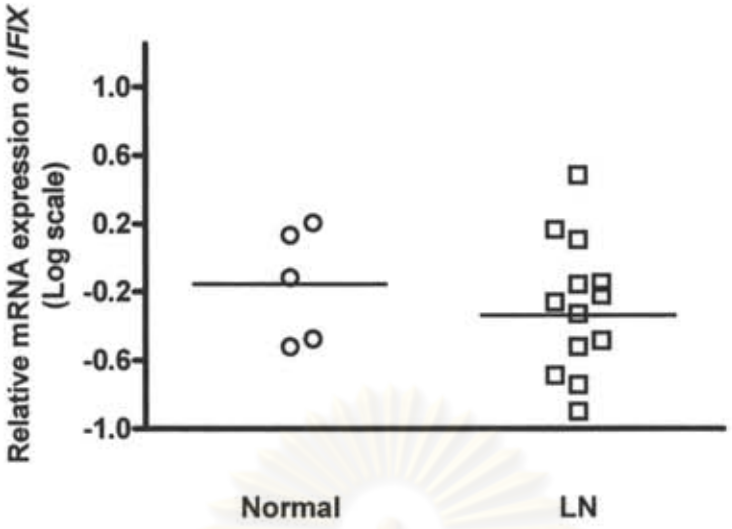


Figure 33 Level of *IFIX* mRNA in the renal biopsies from 14 lupus nephritis (LN) patients and 5 normal controls.

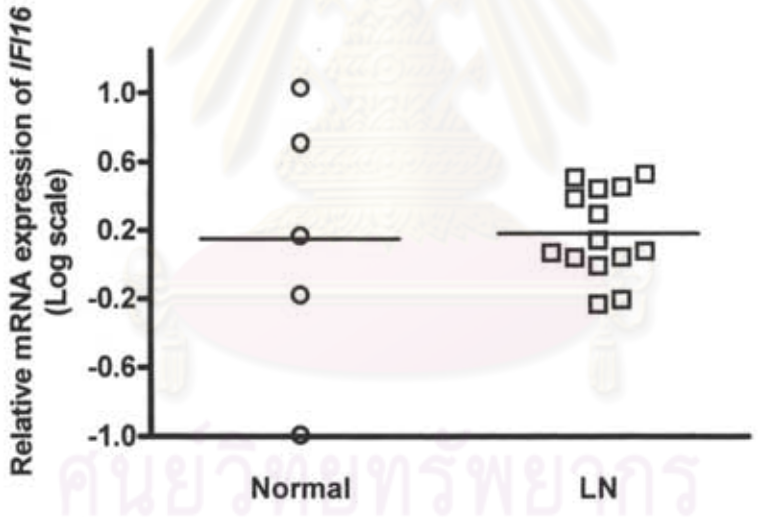


Figure 34 Level of *IFI16* mRNA in the renal biopsies from 14 lupus nephritis (LN) patients and 5 normal controls.

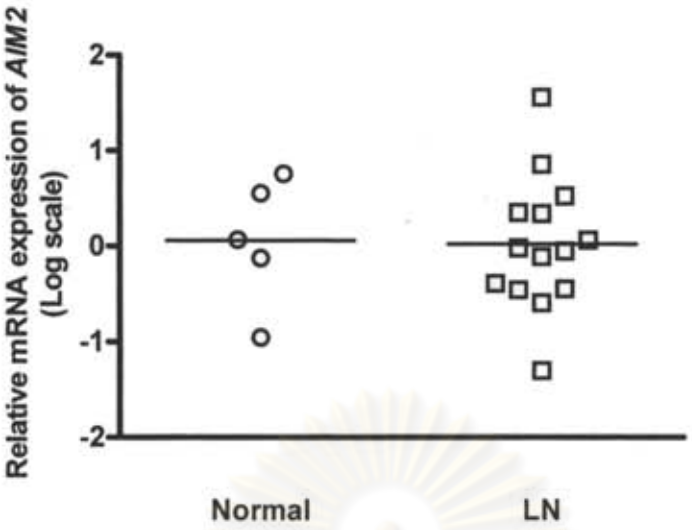


Figure 35 Level of *AIM2* mRNA in the renal biopsies from 14 lupus nephritis (LN) patients and 5 normal controls.

CHAPTER VI

DISCUSSION

Currently, genome wide association (GWA) study has ameliorated our understanding in genetics of SLE which is a complex disease. We have learnt from recent publications of GWA studies that odds ratio of SLE susceptibility genes from these studies is rather low; thus, thousands sample size is essential to obtain enough power (Harley JB et al, 2008; Geoffrey Hom et al., 2008; Kozyrev S et al., 2008; Musone SL et al., 2008). A report from the Wellcome Trust Case Control Consortium (WTCCC) that studied seven complex human diseases including bipolar disorder, coronary artery disease, Crohn's disease, hypertension, rheumatoid arthritis, type 1 diabetes, and type 2 diabetes also showed low OR and suggested that very large sample sizes are needed to increase power of detection (Wellcome Trust Case Control Consortium, 2007). In our study, we performed genome-wide screening on genomic DNA to search for SNPs associated with SLE susceptibility using pooling approach. Although this approach have been conducted in several studies (Butcher LM et al, 2005; Johnson C et al., 2006; Steer S et al., 2007; Jongjaroenprasert W et al., 2008), the limitation is that pooled DNA can be used to estimate only allele frequencies, but not genotype frequencies. This will lead to loss of haplotype and gene–gene interaction information. In addition, the limitations of this approach are the reduced power and technical difficulty to construct pools with equal amounts of samples (McCarthy MI et al., 2008). In the present study, pooled genomic DNA was genotyped on Affymetrix GeneChip Human Mapping 50K Array. This array could genotype only 58,960 SNPs which is not covered whole genome. At present, more than 1 million SNPs can be genotyped by Affymetrix or Illumina platforms which are now readily available. Another important limitation of our GWA study is the small sample sizes, which are underpowered to detect SNPs with modest effect sizes. With all these limitations, we could not identify any SNPs

with distinct p-value. The top 50 SNPs from our GWA study did not reveal any potential important SLE susceptibility genes. We then focus only on major candidate genes previously reported in the literatures.

The analysis of our pooling GWA result compared with known candidate genes suggested that *TNFb, HLA,TNXB and TNFSAIP3* genes are important candidate genes in Thai population giving positive association with p-value (by t-test) less than 0.0001. Interestingly, the best one is a SNP within the HLA region, rs10484561 ranking as 342nd with OR = 2.28 and p = 2.7 x 10^{-7} in consistent with previous observations confirming the significance of this region in most autoimmune diseases (Wellcome Trust Case Control Consortium, 2007; Burfoot RK et al., 2008). Previous HLA association with SLE in Thai population reported that *DRB1*1502 – DQB1*0501* is the major HLA haplotype that confers susceptibility to SLE (Sirikong M et al., 2002). That finding confirmed the association of *DRB1*1502 – DQB1*0501* haplotype previously reported in Taiwan Chinese (Lu LY et al., 1997). However, HLA results in Asian are different from Caucasians. Several reports from Caucasians showed that the frequencies of *HLADRB1*0301 -DQA1*0501 – DQB1*0201* and *DRB1*1501 –DQA1*0102 – DQB1*0602* haplotypes are significantly increased in SLE patients (Black CM et al., 1982; Brennan P et al., 1997; Martin-Villa JM et al., 1998; Tsuchiya N et al., 2001). This might suggest that different HLA alleles or different genes closely link with HLA might associate with SLE in Asian compared to Caucasian. Other positive genes were in the rank ranging from 1218th to 9055th suggesting that we can not use ranking as main criteria in this case. However, the SNPs within these candidate genes are not the same as the one previously reported. The distance of these SNPs from the previous ones as well as the different haplotype block structure among ethnic group are main factors contributing to association result. Therefore, negative results from our pooled GWA did not exclude the importance of some genes e.g. *STAT4* and *ITGAM*. We have investigated previously reported SNPs within these two genes in 278 patients with SLE compared to 383 normal controls. We found that SNP rs7574865 in *STAT4* showed significant association with SLE with OR (95% CI) of 1.75(1.40-2.19) and *P*=8.55 X10^{-7}. For *ITGAM*, four

SNPs were found to be associated with SLE including rs1143679, rs9888739, rs1143683, and rs1143678 with OR (95% CI) of 3.10(1.61-5.98), P=4.1X10^{-4}; 2.58(1.49-4.44), P=4.4X10^{-4}; 2.38(1.47-3.84), P=2.7 X10^{-4} and 2.84(1.65-4.90), P=9.4 X10^{-5}, respectively. In summary, we confirm positive association of *STAT4* and *ITGAM* with SLE in Thai population (Yang W, manuscripted in preparation).

Besides examining previously reported candidate genes from association studies, we also focus on genes within important susceptibility chromosomal loci that have been mapped previously. In this case, we are particularly interested in chromosome 1 as a model. We select new candidate genes based on their potential important function and from result of lupus mouse model. Based on table 21, there are at least 8 new candidate genes that can be further studied. However, due to limited budget, we only select 4 putative functional SNPs from 4 different genes (*NOS1AP*, *PYHIN1* or *IFIX*, *TLR5*, *CD1D*) for further validation by individual genotyping. Our individual genotyping results showed that the allele frequency patterns of SNPs in all 4 genes from individual genotyping were similar to the patterns from pooling approach. In our study, allele frequencies from individual genotyping correlated with the pooling results with Pearson correlation coefficient of 0.991 and 0.950 for case and control, respectively. Moreover, the averaged SEM differences between two groups were small (0.0175). Our results showed good correlation and averaged SEM differences when compared with a previously report. In that study, they showed average correlation of 0.903 and mean differences of 0.065 (Meaburn E et al, 2006). However, only SNP from *NOS1AP* and *PYHIN1* gave marginal positive association by Chi-square test (p=0.035 and P=0.009). While the other 2 SNP gave negative result. This is not surprising since these negative findings are likely due to insufficient power to detect susceptibility genes. We calculated the power based on our sample size (N=100 per group), disease prevalence of 60 patients per 100,000 people in Chinese (Mok CC and Lau CS, 2003), and OR of 1.47 for *TLR5* and 1.60 for *CD1D*. In this calculation, we only have the power of 45 and 63% with α = 0.05 for a multiplicative model, respectively. Increasing of sample sizes should be further performed to improved power of detection.

In the part of candidate gene study, we focus on IFN-inducible genes (*MNDA, IFIX, IFI16* and *AIM2*) which have been proposed as new candidate genes for SLE susceptibility by several evidences. First, data from lupus murine model has identified IFN-inducible protein 202 gene (ifi202) as a candidate for lupus susceptibility (Rozzo SJ et al., 2001; Choubey D et al., 2002). This gene is homologous to human IFN-inducible genes (*MNDA, IFIX, IFI16* and *AIM2*) and located on chromosome 1 similar to human. Second, IFI16 was identified as new autoantigen for patients with SLE, up to 29% of 374 SLE patients develop high titer autoantibodies to IFI16 protein (Seelig HP et al., 1994). Later study has also found that 26% of patients with SLE exhibited significantly higher anti-IFI16 IgG antibody levels compared with control subjects. By immunohistochemical analysis, IFI16 was highly expressed in epidermis and dermal inflammatory infiltrates of skin lesions from patients with SLE (Mondini M et al., 2006). Lastly, a recent study of gene expression profiles using microarray has found 2.4 fold up-regulated IFI16 in leukocytes of patients with SLE as compared to healthy controls (Alcorta DA et al., 2007). From our preliminary pooling GWA, we found genetic association of *IFIX* with SLE. We then select putative functional SNPs from these 4 genes for further investigation. From our association results, we could suggest that SNP within *IFIX* and *IFI16* are both important. Interestingly, a preliminary report in Caucasian patients with SLE also suggest the association signal in *IFIX-IFI16* intergenic region (Fernando MM et al., 2006 (Abstract)). However, the LD patterns of these genes between our populations and Caucasian are different (figure 12 and 13). Although our study could not identify significant association of SNP within *MNDA* and *AIM2* genes, we cannot exclude the importance of these 2 genes because we did not cover all SNPs and all haplotype block. In addition, data from our GWA study cannot define the importance of region in these two genes since there was no SNP from our chips in *AIM2* and only one SNP in *MNDA*. In summary, our approach was trying to select SNPs from candidate genes that might have function. Although this approach is the most cost-effective method and can be used to identify potential important SNPs, it did not cover the whole gene and

relevant haplotype block. Therefore, we cannot rule out the possibility of other SNPs within these genes that might influence SLE susceptibility.

As for the 3 positive SNPs in our study, since we selected them as putative functional SNPs, we try to prove our hypothesis as following. SNP rs866484 (C6771G) of IFI16 is a non-synonymous SNP altering amino acid from threonine to serine. This SNP locates on A type repeat containing p53 binding site (See figure 36). Binding of IFI16 protein to the C-terminus of p53 is reported to stimulate the transcription of p53-responsive reporter plasmids (Johnstone RW et al., 2000).

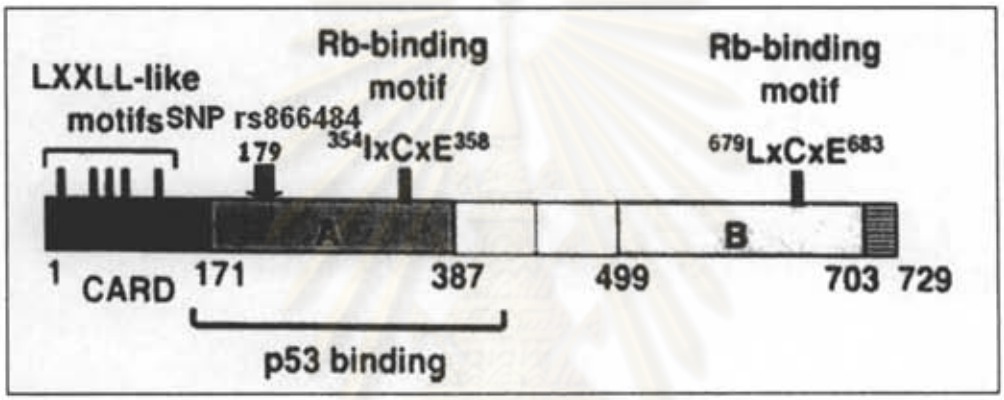


Figure 36 Position of SNP rs866484 on p53 binding site (Choubey D et al., 2008)

Consistent with one study, increased expression of IFI16 in breast cancer cell line resulted in increased transcriptional activity of p53 and led to susceptibility to apoptosis of cells (Fujiuchi N et al., 2004). In our study, we hypothesized that SNP C6771G of *IFI16* may affect p53 binding. The risk G allele may p53 bind with higher affinity and lead to more apoptosis. Consistently, lymphocytes from SLE patients were found to undergo apoptosis more than normal controls or patients with rheumatoid arthritis (Emlen W et al., 1994). Increasing of apoptotic cells might lead to the presence of numerous nuclear antigens at the extracellular tissue. These autoantigens can thus be presented to autoreactive lymphocytes and drive the development of SLE disease (Andrade F et al., 2000; Kamradt T et al., 2001). We conducted our experiment by inducing highest p53 mRNA expression with doxorubicin at a concentration of 10 µg/ml for 24 hr (Figure 17). The association of SNP rs866484 and apoptosis was studied using flow

cytometrical analysis. The limitation of our study is that we used doxorubicin to induce p53 expression, it is not a direct assay to detect IFI16 function. The present study could not reveal the association between SNP and apoptosis. This negative finding may be due to the fact that threonine and serine are in the same amino acid group. Nevertheless, further study should be perform using direct experiment such as immunoprecipitation or any protein-protein interaction methods to prove the direct role of this SNP in protein structure that affecting p53 binding or not.

Furthermore, there are other two associated SNPs in intron including SNP rs1772414 (A23201G) of *IFI16* gene (See in figure 37) and rs856084 (G13792T) of *IFIX* gene (See in figure 38). SNP rs1772414 (A23201G) in intron affect donor splice site according to computational prediction. The *IFI16* gene is known to encode three isoforms (A, B, and C) of IFI16 protein through an alternative splicing of mRNA as shown in Figure 37. The longest mRNA (~2.7 kb) encodes an open reading frame of 2355 bp and generates the IFI16A isoform of 785 amino acids. The second isoform IFI16B arise from the lack of exon 7a (168 bp) to encode a protein of 729 amino acids. The smallest IFI16C isoform (2019 bp) lacks both exon 7 and exon 7a encoding 673 amino acids protein (Johnstone RW et al., 1998). Our SNP is located in intron 6 which is upstream of exon 7, we hypothesize that this SNP may affect donors splice site and lead to the different expression of these isoforms. However, the location of this SNP is very far from intron–exon boundary approximately 16 kb.

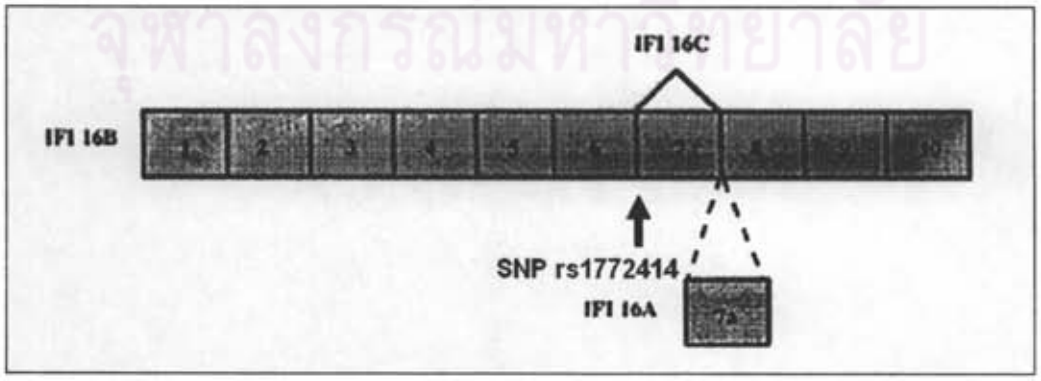


Figure 37 Position of SNP rs1772414 in intron of *IFI16* gene (Johnstone RW et al., 1998)

In addition, our mRNA expression study of *IFI16* isoforms found the B and C isoforms (short product) of *IFI16* from patients with SLE were significantly higher than normal controls. Thus, it would be interesting to explore the role of the SNP A23201G in the expression of three isoforms of *IFI16*. Our results showed the high correlation of C isoform with the G allele, but only AG genotype reached statistical significance, but not GG genotype. It is not additive effect for this SNP. It is possible that the expression of isoforms is controlled by other SNPs. There has been a study that identified SNPs modulating alternative splicing of IFI16 gene. In that study, they found high correlation of SNP rs2994824 that close to the intron–exon boundary (Hull J et al., 2007). However, this SNP could not be detected in Asian population. There may be the other mechanisms to control splice isoform pattern, for instance methylation or miRNA.

Another associated SNPs in intron, rs856084 (G13792T) of *IFIX* gene affect binding site of SC35 splicing regulatory factor according to computational prediction.

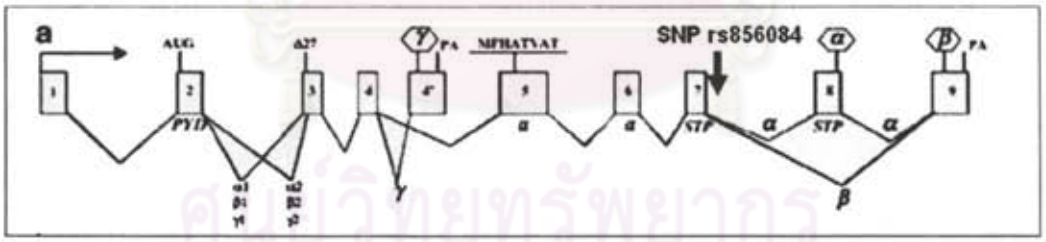


Figure 38 Position of SNP rs856084 in intron of *IFI1X* gene (Ding Y et al., 2004)

IFI1X gene is also known to encode six isoforms (α1, α2, β1, β2, γ1 and γ2) for IFIX protein through an alternative splicing of mRNA as shown in Figure 38 (Ding Y et al., 2004). The C-termini of α, β are diverse due to alternative splicing. Exon 8 is absent in isoform β. Our SNP is located in intron 7 away from exon 7 approximately 317 bp (See Figure 38). In addition, our mRNA expression study

found that β isoform (short isoform) of *IFIX* was decreased in patients with SLE as compared to normal controls. Thus, we also tested the effect of this SNP in the expression of two isoforms (α and β) of *IFIX*. In this study, we did not find the difference of isoforms among genotypes. We could suggest that this SNP does not control alternative splicing of mRNA in *IFIX*. There may be other SNPs near to the intron–exon boundary that define the expression of each isoform.

Our results could suggest that these SNPs are not likely to be functional SNPs. However, we could not exclude the possibility that these SNPs are in linkage disequilibrium (LD) with a nearby causative SNPs. In LD and haplotype analysis indicated that there may also be other SNPs playing important role in SLE disease even if they are not completely linkage. Since SLE is complex disease, combination of several functional SNPs may need to develop of SLE. To clarify the role of *IFIX* and *IFI16* gene, more extensive searching using dense SNPs and increasing of sample sizes are required in future study.

For SNP in *AIM2* promoter, we believe that SNP rs2276405 at position G-151T should change expression level based on several evidences 1) position of this SNP in proximal promoter, 2) similar regulatory transcription factor reported in homologue gene, IFI16 (Clarke CJP et al., 2003), 3) significant correlation between SNP and mRNA expression of AIM2 (figure 21), and 4) supporting evidence from our preliminary EMSA result (figure 20). However, reporter gene expression assay should be further performed to confirm our finding. Although SNP G-151T of AIM2 was not associated with SLE susceptibility, it might be useful for studying in other diseases e.g. breast cancer. There has been study found that *AIM2* suppresses human breast cancer cell proliferation *in vitro* and mammary tumor growth in a mouse model (Chen IF et al., 2006).

In this study, we also performed *MNDA, IFIX, IFI16* and *AIM2* genes expression analyses of various cell types from patients with SLE and healthy controls. We found that all 4 genes were up-regulated in leukocytes of patients

with SLE as compared to healthy controls, but not in lymphocytes and kidney. Our results confirmed a previous study of gene expression profiles using microarray in *IFI16* gene that found 2.4 fold up-regulation in leukocytes (Alcorta DA et al., 2007) from patients with SLE when compared to healthy controls. However, the different mRNA levels of *MNDA*, *IFIX* and *AIM2* genes between patients with SLE and normal controls in the study of Alcorta DA et al were not established.

The negative finding in lymphocytes may indicate the importance of these genes in SLE development via expression in other white blood cell subsets such as monocytes and granulocytes. There have been studies reported that *MNDA* is present both monocytes and granulocyte, while *IFI16* is expressed only in monocytes (Miranda RN et al., 1999; Dawson MJ et al., 1998; Wei Wu et al., 2003; Dawson MJ et al., 1998). For *IFIX* and *AIM2*, the expression of these genes in both cell types is still unclear. A study found that *MNDA* has a role regulating cell apoptosis in granulocyte-macrophage progenitor cell sensitivity to TNF-related apoptosis inducing ligand (TRAIL)–induced programmed cell death (Briggs RC et al., 2006). For *IFI16*, it plays a role in human mononuclear cell differentiation, by modulating the activity of key growth regulatory proteins (Dawson MJ et al., 1998). However, the other functions of IFI16 in myeloid cells are still unclear. In order to better understand the role of these genes in SLE, other white blood cell subsets need to be clarified in the future.

For the negative finding of these genes in kidney indicates that they might not be important in renal pathology. A study found that *IFI16* was highly expressed in epithelial and endothelial cells of skin lesions from patients with SLE (Mondini M et al., 2006). The pathogenic role of *IFI16* in epithelial cells is unknown. However, its expression in endothelial cells is suggested the implication in immunomodulation, cell growth, and apoptosis by gene array study in *IFI16*-overexpressing human umbilical vein endothelial cells (HUVECs). The functional analysis has indicated that IFI16 regulates proinflammatory genes via NF-**K**B activation using a novel mechanism involving suppression of I**K**B expression (Caposio P et al., 2007). In addition, transduction of *IFI16* into HUVECs by a herpes simplex virus (HSV)-derived replication-defective vector

efficiently suppressed the formation of capillary-like structures in vitro and induced apoptosis (Ravera R et al., 2004).

Moreover, *IFI16* and *IFI1X* gens are known to encode three isoforms (A, B and C) for IFI16 protein and six isoforms (α1, α2, β1, β2, γ1 and γ2) for IFIX protein through an alternative splicing of mRNA (Johnstone RW et al., 1998; Ding Y et al., 2004). The difference might be limited to particular isoform. Thus, we also analyzed mRNA expression of IFI16 and IFIX isoforms in leukocytes from patients with SLE compared with normal controls. The comparison of ratio between short product (B and C isoform) and full-length product (A isoform) of *IFI16* gene showed that B and C isoforms of IFI16 from patients with SLE were significantly higher than normal controls. The functional role of different isoforms is unknown. Nevertheless, a study has shown that IFI16 isoforms can dimerize via amino terminal region and these isoforms are able to interact with each other (Johnstone RW et al., 1998). The combinations of IFI16 isoforms might affect different biological functions and may lead to SLE development. For IFIX, we found that short product (β isoform) of IFIX significantly decreased in patients with SLE. In this study, we interested only α and β isoforms of IFIX since γ isoform has small identity to IFI200 family protein and does not have the characteristic 200-amino-acid signature motif of this family proteins (Ding Y et al., 2004). The decreasing of β isoforms of IFIX was associated with an increased full length mRNA (α isoform). Recently, IFIXα1 has been reported that act as tumor suppressor by interaction with HDM2 and lead to p53 stabilization and activation (Ding Y et al., 2006). Functional role of IFIX in SLE is unclear. It is possible that increased IFIXα in SLE patients destabilizes HDM2 resulting in activation of p53 and lead to cell death. Increasing of apoptotic cells might lead to the presence of numerous nuclear antigens at the extracellular tissue. These autoantigens can be presented to autoreactive lymphocytes and drive the development of SLE disease (Andrade F et al., 2000; Kamradt T et al., 2001).

CHAPTER VII

CONCLUSION

In the present study, pooled genomic DNA was genotyped on Affymetrix GeneChip Human Mapping 50K Array compared between 100 SLE patients and 100 healthy control. One important limitation of our GWA study is the small sample sizes, which are underpowered to detect SNPs with modest effect sizes. In summary, we could not identify any SNPs with distinct p-value. The top 50 SNPs from our GWA study did not reveal any potential important SLE susceptibility genes. We then focus only on major candidate genes previously reported in the literatures. The analysis of our pooling GWA result compared with known candidate genes suggested that *TNFb, HLA,TNXB and TNFSAIP3* genes are important candidate genes in Thai population giving positive association with p-value (by t-test) less than 0.0001. Interestingly, the best one is a SNP within the HLA region, rs10484561 in consistent with previous observations confirming the significance of this region in most autoimmune diseases. The distance of these SNPs from the previous reported ones as well as the different haplotype block structure among ethnic group are main factors contributing to association result. Therefore, negative results from our pooled GWA did not exclude the importance of some genes

Besides examining previously reported candidate genes from association studies, we also focus on genes within important susceptibility chromosomal loci that have been mapped previously. In this case, we are particularly interested in chromosome 1 as a model. We select new candidate genes based on their potential important function and from result of lupus mouse model. Our individual genotyping results showed that the allele frequency patterns of SNPs in *NOS1AP, IFIX (PYHIN1), TLR5, CD1D* genes from individual genotyping were similar to the patterns from pooling approach. For future study, the increasing of sample sizes should be performed to improved power of detection.

In the part of candidate gene study, we focus on IFN-inducible genes (*MNDA, IFIX, IFI16* and *AIM2*) which have been proposed as new candidate genes for SLE susceptibility. From our association results, we could suggest that SNPs within *IFIX* and *IFI16* are both important. As for the 3 positive SNPs in our study, since we selected them as putative functional SNPs, we try to prove our hypothesis by conducting some functional studies. However, our studies can not show that these SNPs are functional SNPs. Our approach was trying to select SNPs from candidate genes that might have function. Although this approach is the most cost-effective method and can be used to identify potential important SNPs, it did not cover the whole gene and relevant haplotype block. Therefore, we cannot rule out the possibility of other SNPs within these genes that might influence SLE susceptibility. Since SLE is complex disease, combination of several functional SNPs may need to develop of SLE. To clarify the role of *IFIX* and *IFI16* gene, more extensive searching using dense SNPs and increasing of sample sizes are required in future study.

Interestingly, for SNP in *AIM2* promoter, we believe that SNP rs2276405 at position G-151T should change expression level based on several evidences 1) position of this SNP in proximal promoter, 2) similar regulatory transcription factor reported in homologue gene, IFI16, 3) significant correlation between SNP and mRNA expression of AIM2, and 4) supporting evidence from our preliminary EMSA result. Although SNP G-151T of AIM2 was not associated with SLE susceptibility, it might be useful for studying in other diseases e.g. breast cancer.

Moreover, we also studied the expression of *MNDA, IFIX, IFI16* and *AIM2* genes from various cell types in patients with SLE compared to healthy controls. We found the increasing of IFN-inducible genes in leukocytes from patients with SLE, but not in lymphocytes. These results may indicate the importance of these genes in SLE development via expression in other white blood cell subsets such as monocytes and granulocytes. To further understand the role of these genes in SLE, other white blood cell subsets need to be clarified in next study.

This research provides a better understanding of the disease mechanism by studying the role of new candidate genes. These might lead to the development of new therapeutic strategies and prevention in the future.

REFERENCES

Ahmed, S., K. Ihara, et al. (2001). Association of CTLA-4 but not CD28 gene polymorphisms with systemic lupus erythematosus in the Japanese population. Rheumatology (Oxford) **40**(6): 662-7.

Alarcon-Segovia, D., M. E. Alarcon-Riquelme, et al. (2005). Familial aggregation of systemic lupus erythematosus, rheumatoid arthritis, and other autoimmune diseases in 1,177 lupus patients from the GLADEL cohort. Arthritis Rheum **52**(4): 1138-47.

Alcorta, D. A., D. A. Barnes, et al. (2007). Leukocyte gene expression signatures in antineutrophil cytoplasmic autoantibody and lupus glomerulonephritis. Kidney Int **72**(7): 853-64.

Andrade, F., L. Casciola-Rosen, et al. (2000). Apoptosis in systemic lupus erythematosus. Clinical implications. Rheum Dis Clin North Am **26**(2): 215-27, v.

Asefa, B., K. D. Klarmann, et al. (2004). The interferon-inducible p200 family of proteins: a perspective on their roles in cell cycle regulation and differentiation. Blood Cells Mol Dis **32**(1): 155-67.

Atencio, S., H. Amano, et al. (2004). Separation of the New Zealand Black genetic contribution to lupus from New Zealand Black determined expansions of marginal zone B and B1a cells. J Immunol **172**(7): 4159-66.

Barrett, J. C., B. Fry, et al. (2005). Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics **21**(2): 263-5.

Bijl, M., E. Reefman, et al. (2007). Inflammatory clearance of apoptotic cells after UVB challenge. Autoimmunity **40**(4): 244-8.

Black, C. M., K. I. Welsh, et al. (1982). HLA antigens and Bf allotypes in SLE: evidence for the association being with specific haplotypes. Tissue Antigens **19**(2): 115-20.

Boackle, S. A., V. M. Holers, et al. (2001). Cr2, a candidate gene in the murine Sle1c lupus susceptibility locus, encodes a dysfunctional protein. Immunity **15**(5): 775-85.

Bongu, A., E. Chang, et al. (2002). Can morbidity and mortality of SLE be improved? Best Pract Res Clin Rheumatol **16**(2): 313-32.

Borchers, A. T., C. L. Keen, et al. (2007). Drug-induced lupus. Ann N Y Acad Sci **1108**: 166-82.

Brennan, P., A. Hajeer, et al. (1997). Allelic markers close to prolactin are associated with HLA-DRB1 susceptibility alleles among women with rheumatoid arthritis and systemic lupus erythematosus. Arthritis Rheum **40**(8): 1383-6.

Briggs, J. A., G. R. Burrus, et al. (1992). Cloning and expression of the human myeloid cell nuclear differentiation antigen: regulation by interferon alpha. J Cell Biochem **49**(1): 82-92.

Briggs, R. C., J. A. Briggs, et al. (1994). The human myeloid cell nuclear differentiation antigen gene is one of at least two related interferon-inducible genes located on chromosome 1q that are expressed specifically in hematopoietic cells. Blood **83**(8): 2153-62.

Briggs, R. C., K. E. Shults, et al. (2006). Dysregulated human myeloid nuclear differentiation antigen expression in myelodysplastic syndromes: evidence for a role in apoptosis. Cancer Res **66**(9): 4645-51.

Burfoot, R. K., C. J. Jensen, et al. (2008). SNP mapping and candidate gene sequencing in the class I region of the HLA complex: searching for multiple sclerosis susceptibility genes in Tasmanians. Tissue Antigens **71**(1): 42-50.

Butcher, L. M., E. Meaburn, et al. (2005). SNPs, microarrays and pooled DNA: identification of four loci associated with mild mental impairment in a sample of 6000 children. Hum Mol Genet **14**(10): 1315-25.

Butcher, L. M., E. Meaburn, et al. (2004). Genotyping pooled DNA on microarrays: a systematic genome screen of thousands of SNPs in large samples to detect QTLs for complex traits. Behav Genet **34**(5): 549-55.

Bygrave, A. E., K. L. Rose, et al. (2004). Spontaneous autoimmunity in 129 and C57BL/6 mice-implications for autoimmunity described in gene-targeted mice. PLoS Biol **2**(8): E243.

Caposio, P., F. Gugliesi, et al. (2007). A novel role of the interferon-inducible protein IFI16 as inducer of proinflammatory molecules in endothelial cells. J Biol Chem **282**(46): 33515-29.

Carter, K. W., P. A. McCaskie, et al. (2006). JLIN: a java based linkage disequilibrium plotter. BMC Bioinformatics **7**: 60.

Castano-Rodriguez, N., L. M. Diaz-Gallo, et al. (2008). Meta-analysis of HLA-DRB1 and HLA-DQB1 polymorphisms in Latin American patients with systemic lupus erythematosus. Autoimmun Rev **7**(4): 322-30.

Cervera, R., M. A. Khamashta, et al. (1993). Systemic lupus erythematosus: clinical and immunologic patterns of disease expression in a cohort of 1,000 patients. The European Working Party on Systemic Lupus Erythematosus. Medicine (Baltimore) **72**(2): 113-24.

Chen, J. Y., C. M. Wang, et al. (2005). The -844C/T polymorphism in the Fas ligand promoter associates with Taiwanese SLE. Genes Immun **6**(2): 123-8.

Choubey, D., R. Deka, et al. (2008). Interferon-inducible IFI16 protein in human cancers and autoimmune diseases. Front Biosci **13**: 598-608.

Choubey, D. and B. L. Kotzin (2002). Interferon-inducible p202 in the susceptibility to systemic lupus. Front Biosci **7**: e252-62.

Choubey, D. and P. Lengyel (1992). Interferon action: nucleolar and nucleoplasmic localization of the interferon-inducible 72-kD protein that is encoded by the Ifi 204 gene from the gene 200 cluster. J Cell Biol **116**(6): 1333-41.

Choubey, D. and P. Lengyel (1995). Binding of an interferon-inducible protein (p202) to the retinoblastoma protein. J Biol Chem **270**(11): 6134-40.

Choubey, D., R. Pramanik, et al. (2003). Subcellular localization and mechanisms of nucleocytoplasmic distribution of p202, an interferon-inducible candidate for lupus susceptibility. FEBS Lett **553**(3): 245-9.

Choubey, D., S. Walter, et al. (2000). Cytoplasmic localization of the interferon-inducible protein that is encoded by the AIM2 (absent in melanoma) gene from the 200-gene family. FEBS Lett **474**(1): 38-42.

Clarke, C. J., V. Apostolidis, et al. (2003). Critical role of the transcription factor AP-1 for the constitutive and interferon-induced expression of IFI 16. J Cell Biochem **89**(1): 80-93.

Cooper, G. S., M. A. Dooley, et al. (2001). Smoking and use of hair treatments in relation to risk of developing systemic lupus erythematosus. J Rheumatol **28**(12): 2653-6.

Correa, P. A., L. M. Gomez, et al. (2005). Autoimmunity and tuberculosis. Opposite association with TNF polymorphism. J Rheumatol **32**(2): 219-24.

Craig, D. W., M. J. Huentelman, et al. (2005). Identification of disease causing loci using an array-based genotyping approach on pooled DNA. BMC Genomics **6**: 138.

Cresswell, K. S., C. J. Clarke, et al. (2005). Biochemical and growth regulatory activities of the HIN-200 family member and putative tumor suppressor protein, AIM2. Biochem Biophys Res Commun **326**(2): 417-24.

Datta, B., B. Li, et al. (1996). p202, an interferon-inducible modulator of transcription, inhibits transcriptional activation by the p53 tumor suppressor protein, and a segment from the p53-binding protein 1 that binds to p202 overcomes this inhibition. J Biol Chem **271**(44): 27544-55.

Datta, B., W. Min, et al. (1998). Increase in p202 expression during skeletal muscle differentiation: inhibition of MyoD protein expression and activity by p202. Mol Cell Biol **18**(2): 1074-83.

Dawson, M. J., N. J. Elwood, et al. (1998). The IFN-inducible nucleoprotein IFI 16 is expressed in cells of the monocyte lineage, but is rapidly and markedly down-regulated in other myeloid precursor populations. J Leukoc Biol **64**(4): 546-54.

Dawson, M. J., N. J. Elwood, et al. (1998). The IFN-inducible nucleoprotein IFI 16 is expressed in cells of the monocyte lineage, but is rapidly and markedly down-regulated in other myeloid precursor populations. J Leukoc Biol **64**(4): 546-54.

Dawson, M. J. and J. A. Trapani (1995). The interferon-inducible autoantigen, IFI 16: localization to the nucleolus and identification of a DNA-binding domain. Biochem Biophys Res Commun **214**(1): 152-62.

D'Cruz, D. (2000). Autoimmune diseases associated with drugs, chemicals and environmental factors. Toxicol Lett **112-113**: 421-32.

Deapen, D., A. Escalante, et al. (1992). A revised estimate of twin concordance in systemic lupus erythematosus. Arthritis Rheum **35**(3): 311-8.

Deesomchok, U. and T. Tumrasvin (1983). Clinical study of Thai patients with systemic lupus erythematosus. J Med Assoc Thai **66**(11): 680-7.

Dermott, J. M., J. M. Gooya, et al. (2004). Inhibition of growth by p205: a nuclear protein and putative tumor suppressor expressed during myeloid cell differentiation. Stem Cells **22**(5): 832-48.

DeYoung, K. L., M. E. Ray, et al. (1997). Cloning a novel member of the human interferon-inducible gene family associated with control of tumorigenicity in a model of human melanoma. Oncogene **15**(4): 453-7.

Ding, B., C. J. Liu, et al. (2006). p204 protein overcomes the inhibition of the differentiation of P19 murine embryonal carcinoma cells to beating cardiac myocytes by Id proteins. J Biol Chem **281**(21): 14893-906.

Ding, Y., L. Wang, et al. (2004). Antitumor activity of IFIX, a novel interferon-inducible HIN-200 gene, in breast cancer. Oncogene **23**(26): 4556-66.

Duhl, D. M., M. Gaczynski, et al. (1989). Intranuclear distribution of the human myeloid cell nuclear differentiation antigen in HL-60 cells. J Cell Physiol **141**(1): 148-53.

Dunn, G., A. L. Hinrichs, et al. (2005). Microsatellites versus single-nucleotide polymorphisms in linkage analysis for quantitative and qualitative measures. BMC Genet **6 Suppl 1**: S122.

Emlen, W., J. Niebur, et al. (1994). Accelerated in vitro apoptosis of lymphocytes from patients with systemic lupus erythematosus. J Immunol **152**(7): 3685-92.

Feng, P. H. (2007). Systemic lupus erythematosus: the face of Asia. Ann N Y Acad Sci **1108**: 114-20.

Freni-Titulaer, L. W., D. B. Kelley, et al. (1989). Connective tissue disease in southeastern Georgia: a case-control study of etiologic factors. Am J Epidemiol **130**(2): 404-9.

Fujiuchi, N., J. A. Aglipay, et al. (2004). Requirement of IFI16 for the maximal activation of p53 induced by ionizing radiation. J Biol Chem **279**(19): 20339-44.

Gariglio, M., B. Azzimonti, et al. (2002). Immunohistochemical expression analysis of the human interferon-inducible gene IFI16, a member of the HIN200 family, not restricted to hematopoietic cells. J Interferon Cytokine Res **22**(7): 815-21.

Gariglio, M., M. De Andrea, et al. (1998). The murine homolog of the HIN 200 family, Ifi 204, is constitutively expressed in myeloid cells and selectively induced in the monocyte/macrophage lineage. J Leukoc Biol **64**(5): 608-14.

Gibson, A. W., J. C. Edberg, et al. (2001). Novel single nucleotide polymorphisms in the distal IL-10 promoter affect IL-10 production and enhance the risk of systemic lupus erythematosus. J Immunol **166**(6): 3915-22.

Graham, D. S., R. R. Graham, et al. (2008). Polymorphism at the TNF superfamily gene TNFSF4 confers susceptibility to systemic lupus erythematosus. Nat Genet **40**(1): 83-9.

Graham, R. R., S. V. Kozyrev, et al. (2006). A common haplotype of interferon regulatory factor 5 (IRF5) regulates splicing and expression and is associated with increased risk of systemic lupus erythematosus. Nat Genet **38**(5): 550-5.

Gribaudo, G., L. Riera, et al. (1999). In vitro and in vivo expression analysis of the interferon-inducible 203 gene. J Interferon Cytokine Res **19**(2): 129-36.

Gugliesi, F., M. Mondini, et al. (2005). Up-regulation of the interferon-inducible IFI16 gene by oxidative stress triggers p53 transcriptional activity in endothelial cells. J Leukoc Biol **77**(5): 820-9.

Gutterman, J. U. and D. Choubey (1999). Retardation of cell proliferation after expression of p202 accompanies an increase in p21(WAF1/CIP1). Cell Growth Differ **10**(2): 93-100.

Haraldsson, M. K., N. G. dela Paz, et al. (2005). Autoimmune alterations induced by the New Zealand Black Lbw2 locus in BWF1 mice. J Immunol **174**(8): 5065-73.

Harley, J. B., M. E. Alarcon-Riquelme, et al. (2008). Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXK, KIAA1542 and other loci. Nat Genet **40**(2): 204-10.

Harley, J. B. and J. A. James (1999). Epstein-Barr virus infection may be an environmental risk factor for systemic lupus erythematosus in children and teenagers. Arthritis Rheum **42**(8): 1782-3.

Hartung, K., M. P. Baur, et al. (1992). Major histocompatibility complex haplotypes and complement C4 alleles in systemic lupus erythematosus. Results of a multicenter study. J Clin Invest **90**(4): 1346-51.

Haywood, M. E., N. J. Rogers, et al. (2004). Dissection of BXSB lupus phenotype using mice congenic for chromosome 1 demonstrates that separate intervals direct different aspects of disease. J Immunol **173**(7): 4277-85.

Haywood, M. E., S. J. Rose, et al. (2006). Overlapping BXSB congenic intervals, in combination with microarray gene expression, reveal novel lupus candidate genes. Genes Immun **7**(3): 250-63.

Hertel, L., S. Rolle, et al. (2000). The retinoblastoma protein is an essential mediator that links the interferon-inducible 204 gene to cell-cycle regulation. Oncogene **19**(32): 3598-608.

Hirankarn, N., Y. Avihingsanon, et al. (2007). Genetic susceptibility to SLE is associated with TNF-alpha gene polymorphism -863, but not -308 and -238, in Thai population. Int J Immunogenet **34**(6): 425-30.

Hirankarn, N., J. Wongpiyabovorn, et al. (2006). The synergistic effect of FC gamma receptor IIa and interleukin-10 genes on the risk to develop systemic lupus erythematosus in Thai population. Tissue Antigens **68**(5): 399-406.

Hochberg, M. C. (1997). Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum **40**(9): 1725.

Hom, G., R. R. Graham, et al. (2008). Association of systemic lupus erythematosus with C8orf13-BLK and ITGAM-ITGAX. N Engl J Med **358**(9): 900-9.

Horiuchi, T., H. Nishizaka, et al. (1999). Association of Fas/APO-1 gene polymorphism with systemic lupus erythematosus in Japanese. Rheumatology (Oxford) **38**(6): 516-20.

Hull, J., S. Campino, et al. (2007). Identification of common genetic variation that modulates alternative splicing. PLoS Genet **3**(6): e99.

James, J. A., K. M. Kaufman, et al. (1997). An increased prevalence of Epstein-Barr virus infection in young patients suggests a possible etiology for systemic lupus erythematosus. J Clin Invest **100**(12): 3019-26.

Johnson, A. E., C. Gordon, et al. (1995). The prevalence and incidence of systemic lupus erythematosus in Birmingham, England. Relationship to ethnicity and country of birth. Arthritis Rheum **38**(4): 551-8.

Johnson, C., T. Drgon, et al. (2006). Pooled association genome scanning for alcohol dependence using 104,268 SNPs: validation and use to identify alcoholism vulnerability loci in unrelated individuals from the collaborative study on the genetics of alcoholism. Am J Med Genet B Neuropsychiatr Genet **141B**(8): 844-53.

Johnstone, R. W., M. H. Kershaw, et al. (1998). Isotypic variants of the interferon-inducible transcriptional repressor IFI 16 arise through differential mRNA splicing. Biochemistry **37**(34): 11924-31.

Johnstone, R. W. and J. A. Trapani (1999). Transcription and growth regulatory functions of the HIN-200 family of proteins. Mol Cell Biol **19**(9): 5833-8.

Johnstone, R. W., W. Wei, et al. (2000). Functional interaction between p53 and the interferon-inducible nucleoprotein IFI 16. Oncogene **19**(52): 6033-42.

Jongjaroenprasert, W., S. Chanprasertyotin, et al. (2008). Association of genetic variants in GABRA3 gene and thyrotoxic hypokalaemic periodic paralysis in Thai population. Clin Endocrinol (Oxf) **68**(4): 646-51.

Joshi, A. D., G. V. Hegde, et al. (2007). ATM, CTLA4, MNDA, and HEM1 in high versus low CD38 expressing B-cell chronic lymphocytic leukemia. Clin Cancer Res **13**(18 Pt 1): 5295-304.

Kamatani, Y., K. Matsuda, et al. (2008). Identification of a significant association of a single nucleotide polymorphism in TNXB with systemic lupus erythematosus in a Japanese population. J Hum Genet **53**(1): 64-73.

Kamradt, T. and N. A. Mitchison (2001). Tolerance and autoimmunity. N Engl J Med **344**(9): 655-64.

Kanemitsu, S., K. Ihara, et al. (2002). A functional polymorphism in fas (CD95/APO-1) gene promoter associated with systemic lupus erythematosus. J Rheumatol **29**(6): 1183-8.

Karassa, F. B., T. A. Trikalinos, et al. (2002). Role of the Fcgamma receptor IIa polymorphism in susceptibility to systemic lupus erythematosus and lupus nephritis: a meta-analysis. Arthritis Rheum **46**(6): 1563-71.

Karassa, F. B., T. A. Trikalinos, et al. (2003). The Fc gamma RIIIA-F158 allele is a risk factor for the development of lupus nephritis: a meta-analysis. Kidney Int **63**(4): 1475-82.

Kasitanon, N., W. Louthrenoo, et al. (2002). Neuropsychiatric manifestations in Thai patients with systemic lupus erythematosus. Asian Pac J Allergy Immunol **20**(3): 179-85.

Kelly, J. A., K. L. Moser, et al. (2002). The genetics of systemic lupus erythematosus: putting the pieces together. Genes Immun **3 Suppl 1**: S71-85.

Khoa, P. D., T. Sugiyama, et al. (2005). Polymorphism of interleukin-10 promoter and tumor necrosis factor receptor II in Vietnamese patients with systemic lupus erythematosus. Clin Rheumatol **24**(1): 11-3.

Kikuchi, S., L. Fossati-Jimack, et al. (2005). Differential role of three major New Zealand Black-derived loci linked with Yaa-induced murine lupus nephritis. J Immunol **174**(2): 1111-7.

Kim, E. J., J. I. Park, et al. (2005). IFI16 is an essential mediator of growth inhibition, but not differentiation, induced by the leukemia inhibitory factor/JAK/STAT pathway in medullary thyroid carcinoma cells. J Biol Chem **280**(6): 4913-20.

Kim, S. K., J. W. Nam, et al. (2006). miTarget: microRNA target gene prediction using a support vector machine. BMC Bioinformatics **7**: 411.

Kim, T. G., H. Y. Kim, et al. (1996). Systemic lupus erythematosus with nephritis is strongly associated with the TNFB*2 homozygote in the Korean population. Hum Immunol **46**(1): 10-7.

Kirov, G., I. Nikolov, et al. (2006). Pooled DNA genotyping on Affymetrix SNP genotyping arrays. BMC Genomics **7**: 27.

Klein, S. C., L. H. Boer, et al. (1997). Release of cytokines and soluble cell surface molecules by PBMC after activation with the bispecific antibody CD3 x CD19. Scand J Immunol **46**(5): 452-8.

Kono, D. H. and A. N. Theofilopoulos (2006). Genetics of SLE in mice. Springer Semin Immunopathol **28**(2): 83-96.

Koul, D., N. U. Obeyesekere, et al. (1998). p202 self-associates through a sequence conserved among the members of the 200-family proteins. FEBS Lett **438**(1-2): 21-4.

Kozyrev, S. V., A. K. Abelson, et al. (2008). Functional variants in the B-cell gene BANK1 are associated with systemic lupus erythematosus. Nat Genet **40**(2): 211-6.

Kozyrev, S. V., S. Lewen, et al. (2007). Structural insertion/deletion variation in IRF5 is associated with a risk haplotype and defines the precise IRF5 isoforms expressed in systemic lupus erythematosus. Arthritis Rheum **56**(4): 1234-41.

Kwak, J. C., P. P. Ongusaha, et al. (2003). IFI16 as a negative regulator in the regulation of p53 and p21(Waf1). J Biol Chem **278**(42): 40899-904.

Lander, E. and L. Kruglyak (1995). Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat Genet **11**(3): 241-7.

Lander, E. S. and N. J. Schork (1994). Genetic dissection of complex traits. Science **265**(5181): 2037-48.

Lee, H. S., Y. H. Chung, et al. (2003). Independent association of HLA-DR and FCgamma receptor polymorphisms in Korean patients with systemic lupus erythematosus. Rheumatology (Oxford) **42**(12): 1501-7.

Lee, Y. H., J. B. Harley, et al. (2005). CTLA-4 polymorphisms and systemic lupus erythematosus (SLE): a meta-analysis. Hum Genet **116**(5): 361-7.

Lee, Y. H., Y. H. Rho, et al. (2007). The PTPN22 C1858T functional polymorphism and autoimmune diseases--a meta-analysis. Rheumatology (Oxford) **46**(1): 49-56.

Lee, Y. H., T. Witte, et al. (2005). The mannose-binding lectin gene polymorphisms and systemic lupus erythematosus: two case-control studies and a meta-analysis. Arthritis Rheum **52**(12): 3966-74.

Lembo, D., A. Angeretti, et al. (1995). Constitutive expression of the interferon-inducible protein p202 in NIH 3T3 cells affects cell cycle progression. J Biol Regul Homeost Agents **9**(2): 42-6.

Lembo, M., C. Sacchi, et al. (1998). Inhibition of cell proliferation by the interferon-inducible 204 gene, a member of the Ifi 200 cluster. Oncogene **16**(12): 1543-51.

Liu, C., H. Wang, et al. (2000). MyoD-dependent induction during myoblast differentiation of p204, a protein also inducible by interferon. Mol Cell Biol **20**(18): 7024-36.

Liu, C. J., B. Ding, et al. (2002). The MyoD-inducible p204 protein overcomes the inhibition of myoblast differentiation by Id proteins. Mol Cell Biol **22**(9): 2893-905.

Lowry, O. H., N. J. Rosebrough, et al. (1951). Protein measurement with the Folin phenol reagent. J Biol Chem **193**(1): 265-75.

Lu, L. Y., W. Z. Ding, et al. (1997). Molecular analysis of major histocompatibility complex allelic associations with systemic lupus erythematosus in Taiwan. Arthritis Rheum **40**(6): 1138-45.

Ludlow, L. E., L. L. Hii, et al. (2008). Cloning and characterisation of Ifi206: a new murine HIN-200 family member. J Cell Biochem **103**(4): 1270-82.

Ludlow, L. E., R. W. Johnstone, et al. (2005). The HIN-200 family: more than interferon-inducible genes? Exp Cell Res **308**(1): 1-17.

Martin-Villa, J. M., J. Martinez-Laso, et al. (1998). Differential contribution of HLA-DR, DQ, and TAP2 alleles to systemic lupus erythematosus susceptibility in Spanish patients: role of TAP2*01 alleles in Ro autoantibody production. Ann Rheum Dis **57**(4): 214-9.

McCarthy, M. I., G. R. Abecasis, et al. (2008). Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat Rev Genet **9**(5): 356-69.

Meaburn, E., L. M. Butcher, et al. (2005). Genotyping DNA pools on microarrays: tackling the QTL problem of large samples and large numbers of SNPs. BMC Genomics **6**(1): 52.

Meaburn, E., L. M. Butcher, et al. (2006). Genotyping pooled DNA using 100K SNP microarrays: a step towards genomewide association scans. Nucleic Acids Res **34**(4): e27.

Mehrian, R., F. P. Quismorio, Jr., et al. (1998). Synergistic effect between IL-10 and bcl-2 genotypes in determining susceptibility to systemic lupus erythematosus. Arthritis Rheum **41**(4): 596-602.

Miranda, R. N., R. C. Briggs, et al. (1999). Immunocytochemical analysis of MNDA in tissue sections and sorted normal bone marrow cells documents expression only in maturing normal and neoplastic myelomonocytic cells and a subset of normal and neoplastic B lymphocytes. Hum Pathol **30**(9): 1040-9.

Mok, C. C. and C. S. Lau (2003). Lupus in Hong Kong Chinese. Lupus **12**(9): 717-22.

Mondini, M., M. Vidali, et al. (2007). Role of the interferon-inducible gene IFI16 in the etiopathogenesis of systemic autoimmune disorders. Ann N Y Acad Sci **1110**: 47-56.

Mondini, M., M. Vidali, et al. (2006). A novel autoantigen to differentiate limited cutaneous systemic sclerosis from diffuse cutaneous systemic sclerosis: the interferon-inducible gene IFI16. Arthritis Rheum **54**(12): 3939-44.

Morel, L., K. R. Blenman, et al. (2001). The major murine systemic lupus erythematosus susceptibility locus, Sle1, is a cluster of functionally related genes. Proc Natl Acad Sci U S A **98**(4): 1787-92.

Musone, S. L., K. E. Taylor, et al. (2008). Multiple polymorphisms in the TNFAIP3 region are independently associated with systemic lupus erythematosus. Nat Genet.

Nalla, V. K. and P. K. Rogan (2005). Automated splicing mutation analysis by information theory. Hum Mutat **25**(4): 334-42.

Nath, S. K., J. Kilpatrick, et al. (2004). Genetics of human systemic lupus erythematosus: the emerging picture. Curr Opin Immunol **16**(6): 794-800.

Ng, P. C. and S. Henikoff (2006). Predicting the effects of amino acid substitutions on protein function. Annu Rev Genomics Hum Genet **7**: 61-80.

Nguyen, C., N. Limaye, et al. (2002). Susceptibility genes in the pathogenesis of murine lupus. Arthritis Res **4 Suppl 3**: S255-63.

Oishi, T., A. Iida, et al. (2008). A functional SNP in the NKX2.5-binding site of ITPR3 promoter is associated with susceptibility to systemic lupus erythematosus in Japanese population. J Hum Genet **53**(2): 151-62.

Parks, C. G., J. P. Pandey, et al. (2004). Genetic polymorphisms in tumor necrosis factor (TNF)-alpha and TNF-beta in a population-based study of systemic lupus erythematosus: associations and interaction with the interleukin-1alpha-889 C/T polymorphism. Hum Immunol **65**(6): 622-31.

Pearson, J. V., M. J. Huentelman, et al. (2007). Identification of the genetic basis for complex disorders by use of pooling-based genomewide single-nucleotide-polymorphism association studies. Am J Hum Genet **80**(1): 126-39.

Petri, M. (2002). Epidemiology of systemic lupus erythematosus. Best Pract Res Clin Rheumatol **16**(5): 847-58.

Petri, M. and J. Allbritton (1992). Hair product use in systemic lupus erythematosus. A case-control study. Arthritis Rheum **35**(6): 625-9.

Pistiner, M., D. J. Wallace, et al. (1991). Lupus erythematosus in the 1980s: a survey of 570 patients. Semin Arthritis Rheum **21**(1): 55-64.

Poole, B. D., R. H. Scofield, et al. (2006). Epstein-Barr virus and molecular mimicry in systemic lupus erythematosus. Autoimmunity **39**(1): 63-70.

Prete, P. E. (1985). The mechanism of action of L-canavanine in inducing autoimmune phenomena. Arthritis Rheum **28**(10): 1198-200.

Prokunina, L., C. Castillejo-Lopez, et al. (2002). A regulatory polymorphism in PDCD1 is associated with susceptibility to systemic lupus erythematosus in humans. Nat Genet **32**(4): 666-9.

Purcell, S., B. Neale, et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet **81**(3): 559-75.

Raffaella, R., D. Gioia, et al. (2004). The interferon-inducible IFI16 gene inhibits tube morphogenesis and proliferation of primary, but not HPV16 E6/E7-immortalized human endothelial cells. Exp Cell Res **293**(2): 331-45.

Reed, J. C., K. Doctor, et al. (2003). Comparative analysis of apoptosis and inflammation genes of mice and humans. Genome Res **13**(6B): 1376-88.

Remmers, E. F., R. M. Plenge, et al. (2007). STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. N Engl J Med **357**(10): 977-86.

Rhodes, B. and T. J. Vyse (2007). General aspects of the genetics of SLE. Autoimmunity **40**(8): 550-9.

Rood, M. J., V. Keijsers, et al. (1999). Neuropsychiatric systemic lupus erythematosus is associated with imbalance in interleukin 10 promoter haplotypes. Ann Rheum Dis **58**(2): 85-9.

Rozzo, S. J., J. D. Allard, et al. (2001). Evidence for an interferon-inducible gene, Ifi202, in the susceptibility to systemic lupus. Immunity **15**(3): 435-43.

Russell, A. I., D. S. Cunninghame Graham, et al. (2004). Polymorphism at the C-reactive protein locus influences gene expression and predisposes to systemic lupus erythematosus. Hum Mol Genet **13**(1): 137-47.

Sanchez-Guerrero, J., E. W. Karlson, et al. (1996). Hair dye use and the risk of developing systemic lupus erythematosus. Arthritis Rheum **39**(4): 657-62.

Seelig, H. P., H. Ehrfeld, et al. (1994). Interferon-gamma-inducible protein p16. A new target of antinuclear antibodies in patients with systemic lupus erythematosus. Arthritis Rheum **37**(11): 1672-83.

Sham, P., J. S. Bader, et al. (2002). DNA Pooling: a tool for large-scale association studies. Nat Rev Genet **3**(11): 862-71.

Sigurdsson, S., G. Nordmark, et al. (2005). Polymorphisms in the tyrosine kinase 2 and interferon regulatory factor 5 genes are associated with systemic lupus erythematosus. Am J Hum Genet **76**(3): 528-37.

Simpson, C. L., J. Knight, et al. (2005). A central resource for accurate allele frequency estimation from pooled DNA genotyped on DNA microarrays. Nucleic Acids Res **33**(3): e25.

Sirikong, M., N. Tsuchiya, et al. (2002). Association of HLA-DRB1*1502-DQB1*0501 haplotype with susceptibility to systemic lupus erythematosus in Thais. Tissue Antigens **59**(2): 113-7.

Skol, A. D., L. J. Scott, et al. (2006). Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. Nat Genet **38**(2): 209-13.

Sobel, E. S., L. Morel, et al. (2002). Genetic dissection of systemic lupus erythematosus pathogenesis: evidence for functional expression of Sle3/5 by non-T cells. J Immunol **169**(7): 4025-32.

Sodsai, P., N. Hirankarn, et al. (2008). Defects in Notch1 upregulation upon activation of T Cells from patients with systemic lupus erythematosus are related to lupus disease activity. Lupus **17**(7): 645-53.

Steer, S., V. Abkevich, et al. (2007). Genomic DNA pooling for whole-genome association scans in complex disease: empirical demonstration of efficacy in rheumatoid arthritis. Genes Immun **8**(1): 57-68.

Strasser, A., S. Whittingham, et al. (1991). Enforced BCL2 expression in B-lymphoid cells prolongs antibody responses and elicits autoimmune disease. Proc Natl Acad Sci U S A **88**(19): 8661-5.

Subramanian, S., Y. S. Yim, et al. (2005). Epistatic suppression of systemic lupus erythematosus: fine mapping of Sles1 to less than 1 mb. J Immunol **175**(2): 1062-72.

Takahashi, T., M. Tanaka, et al. (1994). Generalized lymphoproliferative disease in mice, caused by a point mutation in the Fas ligand. Cell **76**(6): 969-76.

Takeuchi, F., K. Nakano, et al. (2005). Genetic contribution of the tumour necrosis factor (TNF) B + 252*2/2 genotype, but not the TNFa,b microsatellite alleles, to systemic lupus erythematosus in Japanese patients. Int J Immunogenet **32**(3): 173-8.

Tan, E. M., A. S. Cohen, et al. (1982). The 1982 revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum **25**(11): 1271-7.

Tencomnao, T., D. Kapitonov, et al. (2004). Transcriptional regulation of the human UDP-galactose:ceramide galactosyltransferase (hCGT) gene expression: functional role of GC-box and CRE. Glycoconj J **20**(5): 339-51.

Trapani, J. A., K. A. Browne, et al. (1992). A novel gene constitutively expressed in human lymphoid cells is inducible with interferon-gamma in myeloid cells. Immunogenetics **36**(6): 369-76.

Trethewey, P. (2004). Systemic lupus erythematosus. Dimens Crit Care Nurs **23**(3): 111-5.

Tsao, B. P. (2003). The genetics of human systemic lupus erythematosus. Trends Immunol **24**(11): 595-602.

Tsuchiya, N., A. Kawasaki, et al. (2001). Analysis of the association of HLA-DRB1, TNFalpha promoter and TNFR2 (TNFRSF1B) polymorphisms with SLE using transmission disequilibrium test. Genes Immun **2**(6): 317-22.

van der Linden, M. W., A. R. van der Slik, et al. (2001). Six microsatellite markers on the short arm of chromosome 6 in relation to HLA-DR3 and TNF-308A in systemic lupus erythematosus. Genes Immun **2**(7): 373-80.

Velazquez-Cruz, R., L. Orozco, et al. (2007). Association of PDCD1 polymorphisms with childhood-onset systemic lupus erythematosus. Eur J Hum Genet **15**(3): 336-41.

Vyse, T. J. and J. A. Todd (1996). Genetic analysis of autoimmune disease. Cell **85**(3): 311-8.

Wandstrat, A. E., C. Nguyen, et al. (2004). Association of extensive polymorphisms in the SLAM/CD2 gene cluster with murine lupus. Immunity **21**(6): 769-80.

Wang, H., G. Chatterjee, et al. (1999). Characteristics of three homologous 202 genes (Ifi202a, Ifi202b, and Ifi202c) from the murine interferon-activatable gene 200 cluster. Genomics **60**(3): 281-94.

Watanabe-Fukunaga, R., C. I. Brannan, et al. (1992). Lymphoproliferation disorder in mice explained by defects in Fas antigen that mediates apoptosis. Nature **356**(6367): 314-7.

Waters, S. T., M. McDuffie, et al. (2004). Breaking tolerance to double stranded DNA, nucleosome, and other nuclear antigens is not required for the pathogenesis of lupus glomerulonephritis. J Exp Med **199**(2): 255-64.

Weening, J. J., V. D. D'Agati, et al. (2004). The classification of glomerulonephritis in systemic lupus erythematosus revisited. Kidney Int **65**(2): 521-30.

Wei, W., C. J. Clarke, et al. (2003). Expression of IFI 16 in epithelial cells and lymphoid tissues. Histochem Cell Biol **119**(1): 45-54.

Weiler, S. R., J. M. Gooya, et al. (1999). D3: a gene induced during myeloid cell differentiation of Linlo c-Kit+ Sca-1(+) progenitor cells. Blood **93**(2): 527-36.

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

Wither, J. E., G. Lajoie, et al. (2003). Functional dissection of lupus susceptibility loci on the New Zealand black mouse chromosome 1: evidence for independent genetic loci affecting T and B cell activation. J Immunol **171**(4): 1697-706.

Wong, M. and B. P. Tsao (2006). Current topics in human SLE genetics. Springer Semin Immunopathol **28**(2): 97-107.

Wu, J., J. C. Edberg, et al. (1997). A novel polymorphism of FcgammaRIIIa (CD16) alters receptor function and predisposes to autoimmune disease. J Clin Invest **100**(5): 1059-70.

Wu, J., C. Metz, et al. (2003). A novel polymorphic CAAT/enhancer-binding protein beta element in the FasL gene promoter alters Fas ligand expression: a candidate background gene in African American systemic lupus erythematosus patients. J Immunol **170**(1): 132-8.

Xu, Z., B. Duan, et al. (2005). Genetic dissection of the murine lupus susceptibility locus Sle2: contributions to increased peritoneal B-1a cells and lupus nephritis map to different loci. J Immunol **175**(2): 936-43.

Yan, D. H., Y. Wen, et al. (1999). Reduced growth rate and transformation phenotype of the prostate cancer cells by an interferon-inducible protein, p202. Oncogene **18**(3): 807-11.

Zhang, Y., Q. Tian, et al. (2008). Multiple splicing results in at least two p203 proteins that are expressed in the liver and down-regulated during liver regeneration. Front Biosci **13**: 2444-51.

Zhu, J., X. Liu, et al. (2005). T cell hyperactivity in lupus as a consequence of hyperstimulatory antigen-presenting cells. J Clin Invest **115**(7): 1869-78.

Zuniga, J., G. Vargas-Alarcon, et al. (2001). Tumor necrosis factor-alpha promoter polymorphisms in Mexican patients with systemic lupus erythematosus (SLE). Genes Immun **2**(7): 363-6.

APPENDICES

APPENDIX A

Reagents and equipments for SNP microarray genotyping of pooled DNA

1. Restriction Enzyme Digestion

1.1 250 ng Genomic DNA working stock: [50 ng/μL] per array

1.2 Xba I (20,000 U/mL): New England Biolab (NEB); P/N R0145L containing:

1.3 NE Buffer 2: New England Biolab (NEB); (P/N B7002S to order separately)

1.4 BSA (Bovine Serum Albumin): New England Biolab (NEB);

(P/N B9001S to order separately)

1.5 H2O (Molecular Biology Water): BioWhittaker Molecular Applications/ Cambrex; P/N 51200

1.6 96-well plate: MJ Research; P/N MLP-9601; or Applied Biosystems; P/N 403083

1.7 96-well PLT. Clear Adhesive Films: Applied Biosystems; P/N 4306311

1.8 Thermal cycler (any Pre-PCR Clean Room thermocycler)

1.9 8-Tube Strips, thin-wall (0.2 mL): MJ Research, P/N TBS-0201; Strip of 8 caps: MJ Research, P/N TCS-0801

2. Ligation

2.1 T4 DNA Ligase: New England Biolab (NEB); P/N M0202L containing:

2.2 T4 DNA Ligase Buffer: New England Biolab (NEB); P/N B0202S

2.3 H2O (Molecular Biology Water): BioWhittaker Molecular Applications/ Cambrex; P/N 51200

2.4 Adaptor Xba (5 μM); Affymetrix; P/N 900410, or Adaptor Hind (5 μM): Affymetrix; P/N 900485 available in the GeneChip® Mapping 50K Assay Kits, P/N 900520 and 900521, respectively

2.5 Thermal cycler (any Pre-PCR Clean Room thermocycler)

3. PCR

3.1 H2O (Molecular Biology Water): BioWhittaker Molecular Applications/ Cambrex; P/N 51200

3.2 dNTP (2.5 mM each): Takara; P/N 4030, or Fisher Scientific; P/N TAK 40301

3.3 PCR Primer (10 μM): Affymetrix; P/N 900409, available in the Xba and Hind GeneChip® Mapping 50K Assay Kits, P/N 900520 and 900521, respectively

3.4 Platinum Pfx DNA Polymerase (2.5 U/μL): Invitrogen Corporation, P/N 11708-039 containing:

- Pfx enzyme
- 10X Pfx Amplification Buffer
- 10X PCR Enhancer
- 50 mM MgSO4

3.5 2% TBE Gel: BMA Reliant precast (2% SeaKem Gold); P/N 54939

3.6 All Purpose Hi-Lo DNA Marker: Bionexus, Inc.; P/N BN2050, or Direct Load Wide Range DNA Marker: Sigma; P/N D7058

3.7 Gel Loading Solution: Sigma; P/N G2526

3.8 Tubes:

- Individual tubes: MJ Research; P/N TWI-0201
- 8-Tube Strips, thin-wall (0.2 mL): MJ Research, P/N TBS-0201; Strip of 8 caps: MJ Research, P/N TCS-0801

3.9 Plate:

- 96-well plate: MJ Research; P/N MLP-9601

3.10 96-well PLT. Clear Adhesive Films: Applied Biosystems; P/N 4306311

3.11 PCR Thermal Cycler (this assay has only been optimized for the following two thermal cyclers)1:

- GeneAmp PCR System 9700, Applied Biosystems, or
- DNA Engine Tetrad PTC-225, MJ Research

4. PCR Purification and Elution

4.1 Manifold - QIAvac multiwell unit: QIAGEN P/N 9014579

4.2 MinElute 96 UF PCR Purification Kit: QIAGEN P/N 28051 (four plates), or P/N 28053 (24 plates)

4.3 Buffer EB (1000 mL): QIAGEN P/N120002

4.4 Biomek Seal and Sample Aluminum Foil Lids: Beckman P/N 538619

4.5 Jitterbug 115 VAC: Boekel Scientific P/N 130000

4.6 Vacuum Regulator for use during the PCR clean up step. QIAGEN Vacuum Regulator (use with QIAvac manifolds): QIAGEN; P/N 19530*

* The QIAGEN protocol requires ~800 mb vacuum. If your lab does not have an internally regulated vacuum source, this vacuum regulator isstrongly suggested

5. Quantification of Purified Product and dilution

5.1 Nanodrop ND-1000 Spectrophotometer

5.2 EB buffer (10mM Tris-HCl, pH 8.5)

6. Fragmentation

6.1 Fragmentation Reagent (DNase I): Affymetrix, P/N 900131

6.2 10X Fragmentation Buffer: Affymetrix, P/N 900422

6.3 Molecular Biology Water: Bio Whittaker Molecular Applications/Cambrex, P/N 51200

6.4 4% TBE Gel: BMA Reliant precast (4% NuSieve 3:1 Plus Agarose); P/N 54929

6.5 All Purpose Hi-Lo DNA Marker: Bionexus, Inc.; 50 - 10000bp; P/N BN2050

6.6 Gel Loading Solution: Sigma; P/N G2526

6.7 96-well plate: MJ Research; P/N MLP-9601

6.8 96-well PLT. Clear Adhesive Films: Applied Biosystems; P/N 4306311

6.9 Thermal cycler

6.10 DNA Engine Tetrad: MJ Research, or

6.11 Gene Amp PCR System 9700

7. Labeling

7.1 GeneChip DNA Labeling Reagent (7.5 mM): Affymetrix; P/N 900484, available in the Xba I and Hind III GeneChip® Mapping 50K Assay Kits, P/N 900520 and P/N 900521, respectively

7.2 Terminal Deoxynucleotidyl Transferase (30 U/μL): Affymetrix; P/N 900426, available in the Xba I and Hind III GeneChip® Mapping 50K Assay Kits, P/N 900520 and P/N 900521, respectively

7.3 5X Terminal Deoxynucleotidyl Transferase Buffer: Affymetrix; P/N 900425, available in the Xba I and Hind III GeneChip® Mapping 50K Assay Kits, P/N 900520 and P/N 900521, respectively

8. Target Hybridization

8.1 5M TMACL (Tetramethyl Ammonium Chloride): Sigma; P/N T3411

8.2 10% Tween-20: Pierce; P/N 28320 (Surfactamps); diluted to 3% in molecular biology grade water

8.3 MES hydrate SigmaUltra: Sigma; P/N M5287

8.4 MES Sodium Salt: Sigma; P/N M5057

8.5 DMSO: Sigma; P/N D5879

8.6 EDTA: Ambion; P/N 9260G

8.7 Denhardt's Solution: Sigma; P/N D2532

8.8 HSDNA (Herring Sperm DNA): Promega; P/N D1815

8.9 Human Cot-1: Invitrogen; P/N 15279-011

8.10 Oligo Control Reagent, 0100 (OCR, 0100): Affymetrix; P/N 900541, available in the Xba and Hind GeneChip® Mapping 50K Assay Kits, P/N 900520 and P/N 900521, respectively

9. Washing, Staining and scanning

9.1 Water, Molecular Biology Grade, BioWhittaker Molecular Applications / Cambrex, P/N 51200

9.2 Distilled water, Invitrogen Life Technologies, P/N 15230147

9.3 20X SSPE (3M NaCl, 0.2M NaH2PO4, 0.02 M EDTA), BioWhittaker Molecular Applications / Cambrex, P/N 51214

9.4 Anti-streptavidin antibody (goat), biotinylated, Vector Laboratories, P/N BA-0500; reconstitute according to product instructions

9.5 R-Phycoerythrin Streptavidin, Molecular Probes, P/N S-866

9.6 10% surfact-Amps20 (Tween-20), Pierce Chemical, P/N 28320

9.7 Bleach (5.25% Sodium Hypochlorite), VWR Scientific, P/N 21899-504 (or equivalent)

9.8 Denhardt's Solution, 50X concentrate: Sigma; P/N D2532

9.9 MES hydrate, Sigma-Aldrich, P/N M5287

9.10 MES Sodium Salt, Sigma-Aldrich, P/N M5057

9.11 5M NaCl, RNase-free, DNase-free, Ambion, P/N 9760G

APPENDIX B

Reagents for DNA extraction

1. Red Cell Lysis Buffer (RCLB)

NH4Cl	1.875 g
Tris-HCl	0.25 g

Dissolve NH4Cl and Tris-HCL in 500 ml of distilled water. Adjust pH to 7.2. The solution was mixed and sterilizes by autoclaving at 121 °C for 15 min. Keep at 4 °C. Shelf life is approximately 6 months.

2. Nuclei Lysis Buffer (NLB)

1 M Tris (pH 8.0)	10 ml
5 M NaCl	0.5 ml
0.5 M EDTA (pH 8)	0.4 ml

Adjust volume to 100 ml with distilled water. Adjust pH to 7.2. Keep at 4 °C. Shelf life is approximately 6 months.

3. 1 M Tris

Tris base	12.11 g
Distilled water	100 ml

Adjust volume to 100 ml with distilled water. Adjust pH to 8.0. The solution was mixed and sterilizes by autoclaving at 121 °C for 15 min.

4. 5 M NaCl

NaCl	29.22 g
Distilled water	100 ml

Adjust volume to 100 ml with distilled water. The solution was mixed and sterilizes by autoclaving at 121 °C for 15 min.

5. EDTA

EDTA	37.22 g
Distilled water	200 ml

Adjust volume to 200 ml with distilled water. Adjust pH to 8.0. The solution was mixed and sterilizes by autoclaving at 121 °C for 15 min.

6. 5.3 M NaCl

NaCl	15.5 g
Distilled water	50 ml

Adjust volume to 50 ml with distilled water. The solution was mixed and sterilizes by autoclaving at 121°C for 15 min.

7. Proteinase K 10 mg/ml

Proteinase K	100 mg
Distilled water	10 ml

Mix the solution and store at -20 °C.

8. 10% SDS

SDS	10 g
Distilled water	100 ml

Adjust volume to 100 ml with distilled water. The solution was mixed and sterilizes by autoclaving at 121°C for 15 min.

APPENDIX C

Reagents for agarose gel electrophoresis

1. 50x Tris-acetate buffer (TAE)

Tris base	424.0 g
Glacial acetic acid	57.1 g
0.5 M EDTA pH 8.0	100 ml

Adjust volume to 1 liter with distilled water. The solution was mixed and sterilizes by autoclaving at 121°C for 15 min.

2. 10 mg/ml Ethidium bromide

Ethidium bromide	1.0 g
Distilled water	100 ml

Mix the solution and store in the dark at 4°C.

3. 1.5% Agarose gel

Agarose	0.3 g
1x TAE	20 ml

Dissolve by heating in microwave oven and occasional mix unit no granules of agarose are visible.

4. 5x Loading buffer 100 ml

Tris HCL	0.6 g
EDTA	1.68 g
SDS	0.5 g
Bromphenol Blue	0.1 g
Sucrose	40 g

Adjust volume to 100 ml with distilled water. Mix the solution, aliquot into 1.5 microtube and store at 4°C.

APPENDIX D

SNPs and haplotype analysis

PLINK v1.03 program is produced by Purcell S et al from the Center for Human Genetic Research, Massachusetts General Hospital, Boston (Purcell S et al., 2007). The software is available online at http://pngu.mgh.harvard.edu/purcell/plink/.

Input file format

Two data files were used to run PLINK program including PED and MAP files.

1. PED file

The PED file is a white-space (space or tab) delimited file: the first six columns are mandatory:

Family ID

Individual ID

Paternal ID

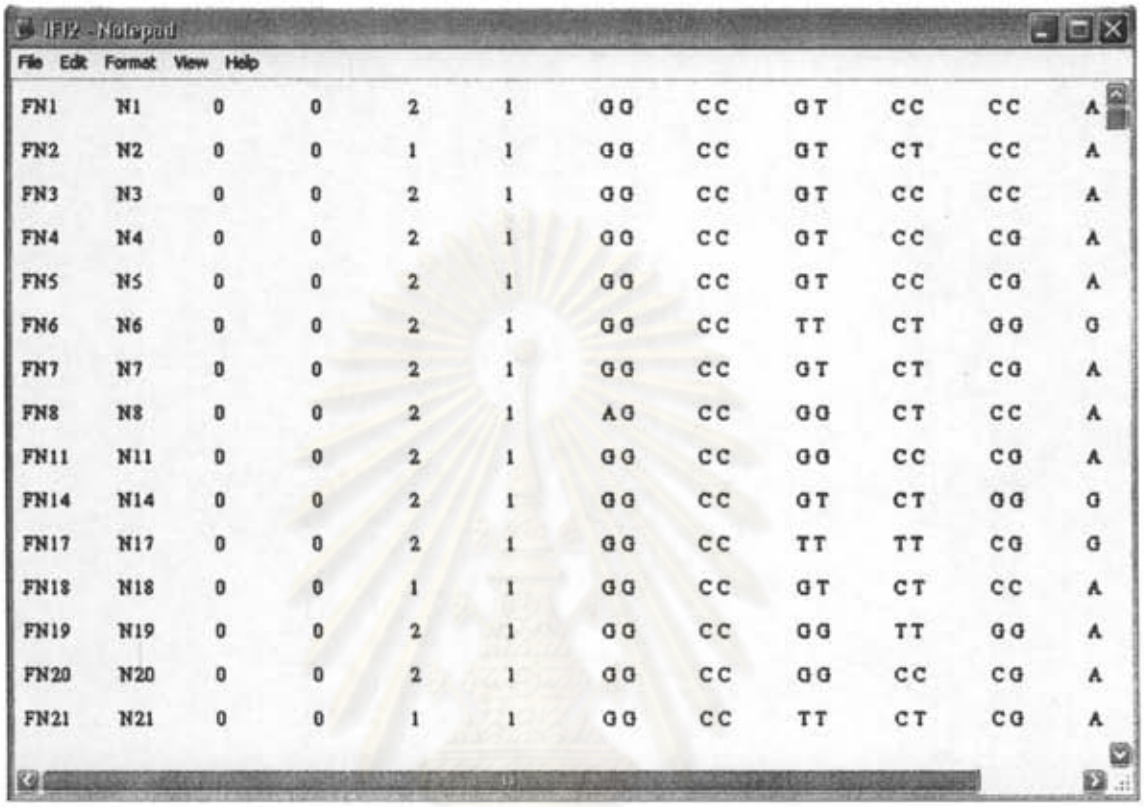
Maternal ID

Sex (1=male; 2=female)

Phenotype (normal=1; SLE=2)

For the column 7 onwards, they are genotypes which can be any character (e.g. 1,2,3,4 or A,C,G,T or anything else) except 0 which is, by default, the missing genotype character.

The example of PED file (IFI2.ped) was shown below.



IFI2 - Notepad

File Edit Format View Help

FN1	N1	0	0	2	1	G G	C C	G T	C C	C C	A
FN2	N2	0	0	1	1	G G	C C	G T	C T	C C	A
FN3	N3	0	0	2	1	G G	C C	G T	C C	C C	A
FN4	N4	0	0	2	1	G G	C C	G T	C C	C G	A
FN5	N5	0	0	2	1	G G	C C	G T	C C	C G	A
FN6	N6	0	0	2	1	G G	C C	T T	C T	G G	G
FN7	N7	0	0	2	1	G G	C C	G T	C T	C G	A
FN8	N8	0	0	2	1	A G	C C	G G	C T	C C	A
FN11	N11	0	0	2	1	G G	C C	G G	C C	C G	A
FN14	N14	0	0	2	1	G G	C C	G T	C T	G G	G
FN17	N17	0	0	2	1	G G	C C	T T	T T	C G	G
FN18	N18	0	0	1	1	G G	C C	G T	C T	C C	A
FN19	N19	0	0	2	1	G G	C C	G G	T T	G G	A
FN20	N20	0	0	2	1	G G	C C	G G	C C	C G	A
FN21	N21	0	0	1	1	G G	C C	T T	C T	C G	A

2. MAP file

By default, each line of the MAP file describes a single marker and must contain exactly 4 columns:

chromosome (1-22, X, Y or 0 if unplaced)

rs# or snp identifier

Genetic distance (morgans)

Base-pair position (bp units)

Most analyses do not require a genetic map to be specified in any case; specifying a genetic (cM) map is most crucial for a set of analyses that look for shared segments between individuals. For basic association testing, the genetic distance column can be set at 0.

The example of MAP file (IFI2.map) was shown below.

Running of PLINK program

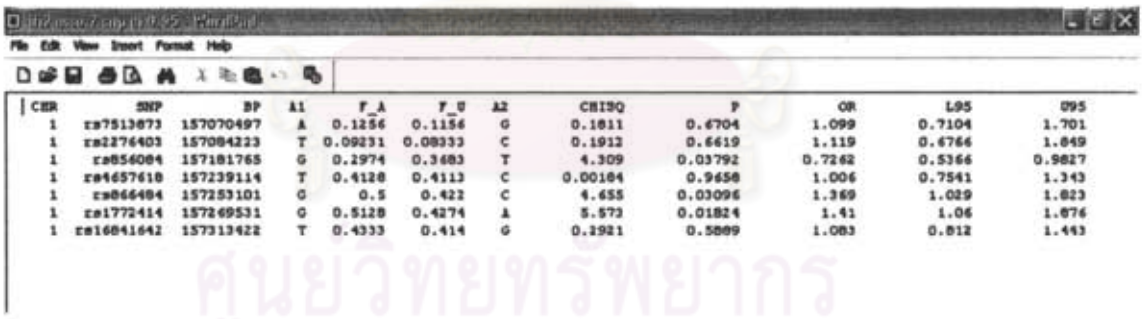
1. Allelic association test

   To perform a standard case/control association analysis, use the option as follow:

The output file contains the fields;

CHR	Chromosome
SNP	SNP ID
BP	Physical position (base-pair)
A1	Minor allele name (based on whole sample)
F_A	Frequency of this allele in cases
F_U	Frequency of this allele in controls
A2	Major allele name
P	Exact p-value for this test
OR	Estimated odds ratio (for A1)
L95	Lower bound of 95% confidence interval for odds ratio
U95	Upper bound of 95% confidence interval for odds ratio

The example of output file was shown below.



CHR	SNP	BP	A1	F_A	F_U	A2	CHISQ	P	OR	L95	U95
1	rs7513873	157070497	A	0.1256	0.1156	G	0.1811	0.6704	1.099	0.7104	1.701
1	rs2276403	157084223	T	0.09231	0.08333	C	0.1912	0.6619	1.119	0.6766	1.849
1	rs856084	157181765	G	0.2974	0.3683	T	4.309	0.03792	0.7262	0.5366	0.9827
1	rs4657618	157239114	T	0.4128	0.4113	C	0.00104	0.9658	1.006	0.7541	1.343
1	rs866484	157253101	G	0.5	0.422	C	4.655	0.03096	1.369	1.029	1.823
1	rs1772414	157269531	G	0.5128	0.4274	A	5.573	0.01824	1.41	1.06	1.876
1	rs16841642	157313422	T	0.4333	0.414	G	0.2921	0.5889	1.083	0.812	1.443

In addition, to perform permutation analysis, following option was used.

2. Model of inheritance analysis

The dominant and recessive models are tests for the minor allele (which is the minor allele can be found in the output of either the - - assoc or the - - freq commands. That is, if D is the minor allele (and d is the major allele):

Allelic:	D	versus	d	
Dominant:	(DD, Dd)	versus	dd	
Recessive:	DD	versus	(Dd, dd)	
Genotypic:	DD	versus	Dd	versus dd

As mentioned above, these tests are generated with option:

3. Conditional analysis of SNP

To condition analysis on a specific SNP, use the - - condition option and we can also specify with either

--dominant

or

--recessive

4. Haplotype analysis

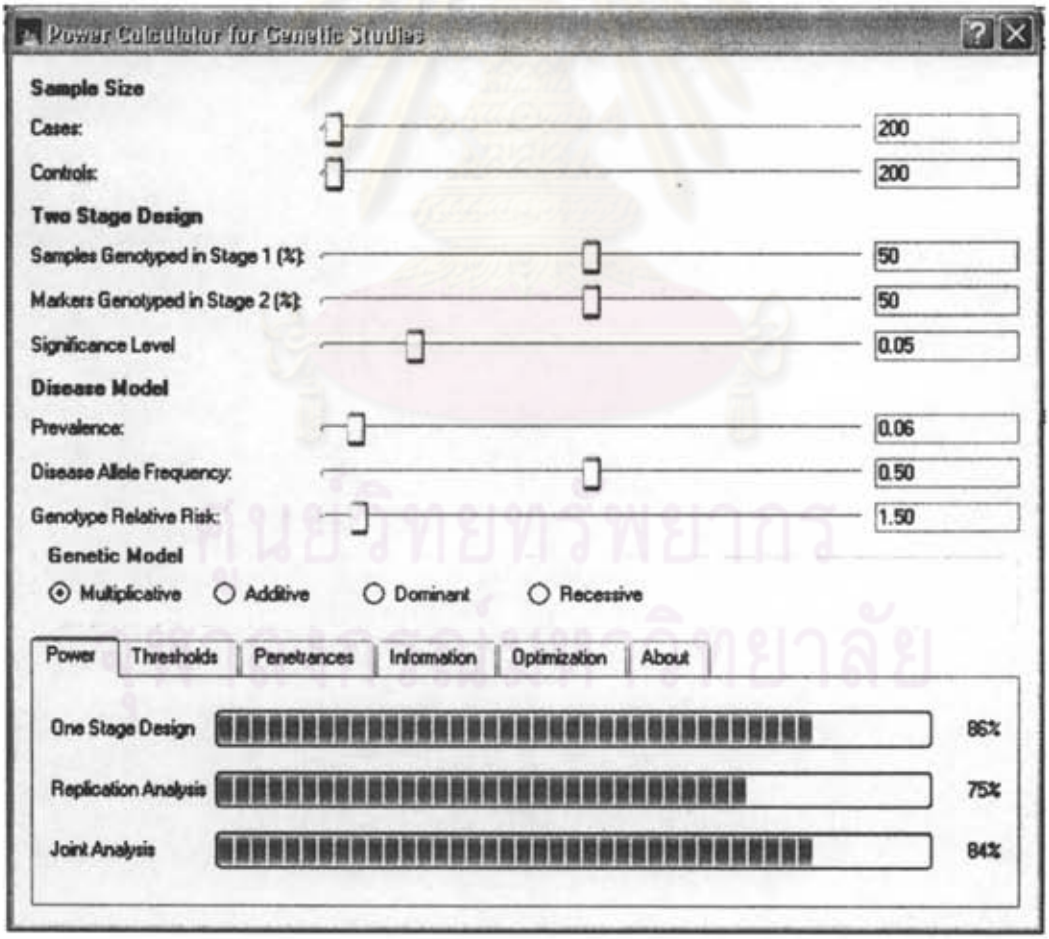
To test haplotype association, the following command was used.

```
C:\plink-1.03-dos>plink --ped IFI2.ped --map IFI2.map --hap-snps rs856084,rs8664
84,rs1772414 --hap-assoc
```

APPENDIX E

Power Calculation

The power was calculated using the CaTS Power Calculator for genetics studies (Skol AD et al., 2006; http://www.sph.umich.edu/). We calculated the power based on our sample size and disease prevalence of 60 patients per 100,000 people in Chinese (0.06) (Mok CC and Lau CS, 2003). The disease allele frequency and genotype relative risk were adjusted according to individual SNP. The significance level was set to 0.05. The interface of CaTS Power Calculator was shown below.



BIOGRAPHY

Miss Ingorn Kimkong was born on December 10, 1978 in Ratchaburi, Thailand. She graduated with the Bacherlor degree of Science in Medical Technology from Thammasat University in 2000 and received her Master degree of Science in Medical Microbiology from Chulalongkorn University in 2003. She received scholarship from The Commission on Higher Education, Ministry of Education according to the requirement of Kasetsart University, for her study in Doctor of Philosophy program in Medical Microbiology at Chulalongkorn University in 2003. This scholarship covered her research working at Centre National de Génotypage (CNG), Evry, France for 6 months in 2006.

Awards:

1. Excellent poster presentation at The Fifth Princess Chulabhorn International Science Congress: Evolving Genetics and Its Global Impact, Bangkok, Thailand, August 16-20, 2004.

2. European Commissions Marie Curie Actions scholarship from ESH EuroConference on Translational Research in Transplantation, Immunogenetics, Phamacogenomics, Proteomics and Immunobiology; July, 4-6, 2006; Paris, France

Conferences:

1. Ingorn Kimkong, Jongkonnee Wongpiyabovorn, Yingyos Avihingsanon, Nattiya Hirankarn. No Association between polymorphisms within interferon (IFN)-inducible genes (*MNDA*, *IFI16* and *AIM2*) with susceptibility to systemic lupus erythematosus. ESH EuroConference on Translational Research in Transplantation, Immunogenetics, Phamacogenomics, Proteomics and Immunobiology, Paris, France, July 4-6, 2006 (Abstract presentation).
2. Ingorn Kimkong. The role of interferon (IFN) - inducible genes (*MNDA*, *IFI16*, and *AIM2*) in SLE. The Third International Summer School Immunogenetics, Bangkok, Thailand, November 17-20, 2006 (Oral presentation).
3. Ingorn Kimkong, Yingyos Avihingsanon, Jongkonnee Wongpiyabovorn, Nattiya Hirankarn. Up-regulation of interferon (IFN)-inducible genes (*IFI16*, *MNDA*, *AIM2* and *IFIX*) in leukocyte, but not lymphocyte and kidney, of patients with systemic lupus erythematosus. The 4th Congress of the Federation of Immunology Societies of Asia-Oceania (FIMSA 2008), Taipei, Taiwan, October 17 to 20, 2008 (Oral presentation).

Publications:

1. Ingorn Kimkong, Nattiya Hirankarn. Molecular Genetics of Systemic Lupus Erythematosus. Molecular Genetics. Siriraj Medical Journal. 2008 Sep; 60(5).

2. Ingorn Kimkong, Yingyos Avihingsanon, Jongkonnee Wongpiyabovorn, Nattiya Hirankarn. Up-regulation of interferon (IFN)-inducible genes (*IFI16*, *MNDA*, *AIM2* and *IFIX*) in leukocyte, but not lymphocyte and kidney, of patients with systemic lupus erythematosus (in prep).

3. Ingorn Kimkong, Yingyos Avihingsanon, Jongkonnee Wongpiyabovorn, Nattiya Hirankarn. Association of interferon (IFN)-inducible genes (*IFI16*, *MNDA*, *AIM2* and *IFIX*) with susceptibility to SLE disease (in prep).