SINGLE NUCLEOTIDE POLYMORPHISM (SNPS) STUDY ON X-CHROMOSOME IN THAI SLE POPULATIONS

Miss Krisana Jaiwan

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Molecular Science of Medical Microbiology and Immunology
Department of Transfusion Medicine and Clinical Microbiology
FACULTY OF ALLIED HEALTH SCIENCES
Chulalongkorn University
Academic Year 2020

การศึกษา Single nucleotide polymorphisms (SNPs) ใน X-chromosome ของกลุ่มผู้ป่วยโรค SLE ประเทศไทย

น.ส.กฤษณา ใจวัน

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิทยาศาสตร์ระดับโมเลกุลทางจุลชีววิทยาทางการแพทย์และวิทยาภูมิคุ้มกัน ภาควิชาเวช
ศาสตร์การธนาคารเลือดและจุลชีววิทยาคลินิก
คณะสหเวชศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2563

Thesis Title	SINGLE NUCLEOTIDE POLYMORPHISM (SNPS) STUDY ON X-CHROMOSOME IN THAI SLE POPULATIONS
By	Miss Krisana Jaiwan
Field of Study	Molecular Science of Medical Microbiology and Immunology
Thesis Advisor	Pattarin Tangtanatakul, Ph.D.
Thesis Co Advisor	Wang Yong Fei, Ph.D.

Accepted by the FACULTY OF ALLIED HEALTH SCIENCES, Chulalongkorn University in Partial Fulfillment of the Requirement for the Master of Science

..... Dean of the FACULTY OF ALLIED HEALTH SCIENCES
(Associate Professor PALANEE AMMARANOND, Ph.D.)

THESIS COMMITTEE

..... Chairman
(Assistant Professor TEWIN TENCOMNAO, Ph.D.)

..... Thesis Advisor
(Pattarin Tangtanatakul, Ph.D.)

..... Thesis Co-Advisor
(Wang Yong Fei, Ph.D.)

..... Examiner
(Assistant Professor Tewarit Sarachana, Ph.D.)

..... Examiner
(Pumipat Tongyoo, Ph.D.)

กฤษณา ใจวัน : การศึกษา Single nucleotide polymorphisms (SNPs) ใน X-chromosome ของกลุ่มผู้ป่วยโรค SLE ประเทศไทย. (SINGLE NUCLEOTIDE POLYMORPHISM (SNPS) STUDY ON X-CHROMOSOME IN THAI SLE POPULATIONS) อ.ที่ปรึกษาหลัก : อ. ดร.ภัทริน ตั้งธนตระกูล, อ.ที่ปรึกษาร่วม : Wang Yong Fei

โรคพุ่มพวงเป็นโรคภูมิแพ้ตนเองที่สำคัญของประเทศไทย โรคพุ่มพวงเกิดในเพศหญิงมากกว่าผู้ชาย ในอัตราส่วน 9 ต่อ 1 จากการศึกษาก่อนหน้านี้แสดงให้เห็นว่าความหลากหลายทางพันธุกรรมบนโครโมโซมเอกซ์ เป็นปัจจัยหนึ่งของการเกิดโรคพุ่มพวง อย่างไรก็ดี ยังไม่มีรายงานความหลากหลายทางพันธุกรรมบนโครโมโซมเอกซ์ในผู้ป่วยโรคพุ่มพวงในประชากรไทย ดังนั้น วัตถุประสงค์ของโครงการวิจัย คือ การศึกษาความหลากหลายทางพันธุกรรมบนโครโมโซมเอกซ์ที่เกี่ยวข้องกับการเกิดโรคพุ่มพวง โดยใช้ข้อมูลจากกลุ่มประชากรชาวไทย ข้อมูล Genotyping จากการศึกษาก่อนหน้านี้ จะนำมาแบ่งเป็นข้อมูลที่ใช้ศึกษาในขั้นต้น (กลุ่มควบคุมสุขภาพปกติ จำนวน1,683 คน และ กลุ่มผู้ป่วยโรคพุ่มพวง จำนวน 487 คน) และข้อมูลที่ใช้ยืนยันผล (กลุ่มควบคุมที่มีโรคประจำตัวที่ไม่เกี่ยวข้องกับโรคพุ่มพวงจำนวน 1,711 คน และ กลุ่มผู้ป่วยโรคพุ่มพวง จำนวน 455 คน) โดยวิธี meta-analysis และ imputation กับ 1 KGP จากผลการวิเคราะห์ผล พบว่า rs1059702 บริเวณยีนส์ *IRAK1- MECP2- TMEM187* (p-value = 1.82 x 10^{-7}; OR = 0.68) rs3853839 (p-value = 2.03 x 10^{-4}; OR = 0.74) บริเวณยีนส์ Toll-like receptor 7 (*TLR7),* X:9165034 บริเวณยีนส์ Family With Sequence Similarity 9 Member B (*FAM9B*) (p-value = 1.14 x 10^{-5}; OR=1.3), และ rs12398129 บริเวณยีนส์ *CXorf61* (p-value = 1.72 x 10^{-4}; OR = 1.39) ความหลากหลายทางพันธุกรรมดังกล่าว มีรายงานในประชากรเชื้อชาติอื่นๆ มาแล้ว นอกจากนี้เรายังพบความหลากหลายทางพันธุกรรมที่อาจมีความจำเพาะในเชื้อชาติไทย ดังนี้ rs6528443 บริเวณยีนส์ *GPR101* (p-value = 8.71 x 10^{-6} ;OR = 3.55) และ rs7052503 บริเวณยีนส์ *MIR891A* (p-value = 2.46 x 10^{-4}; OR = 0.77) โดยความหลากหลายทางพันธุกรรม เหล่านี้มีความเกี่ยวข้องกับการทำงานของระบบภูมิคุ้มกันซึ่งอาจมีผลต่อการเกิดโรคพุ่มพวงในประชากรไทย อย่างไรก็ดี การทดสอบโดยวิธีอื่นๆ จะช่วยยืนยันผลการทดสอบ

สาขาวิชา	วิทยาศาสตร์ระดับโมเลกุลทางจุลชีววิทยาทางการแพทย์และวิทยาภูมิคุ้มกัน	ลายมือชื่อนิสิต .....
ปีการศึกษา	2563	ลายมือชื่อ อ.ที่ปรึกษาหลัก .....
		ลายมือชื่อ อ.ที่ปรึกษาร่วม .....

6176757837 : MAJOR MOLECULAR SCIENCE OF MEDICAL MICROBIOLOGY AND IMMUNOLOGY
KEYWORD: Systemic Lupus Erythematosus (SLE) X chromosome Genome-Wide Association Study (GWAS) Thai populations

Krisana Jaiwan : SINGLE NUCLEOTIDE POLYMORPHISM (SNPS) STUDY ON X-CHROMOSOME IN THAI SLE POPULATIONS. Advisor: Pattarin Tangtanatakul, Ph.D. Co-advisor: Wang Yong Fei, Ph.D.

Systemic Lupus Erythematosus (SLE) is common autoimmune disease in Thailand which dominantly in females in ratio 9:1 of patients. Previous study has shown that the genetic components especially in X chromosome contributed a lot to the disease development. However, the susceptibility loci in Thai population have not been fully examined. Here, we conducted genome-wide association study (GWAS) on X chromosome using the data from two independent cohorts: primary dataset (controls = 1,683, SLE = 487) and secondary dataset (controls = 1,711, SLE = 455). Through meta-analyzing and imputation base on 1 KGP the two data set, SNP rs1059702 in *IRAK1- MECP2- TMEM187* (*p*-value = 1.82 x 10^{-7}; OR = 0.68), rs3853839 (*p*-value = 2.03 x 10^{-4}; OR = 0.74) in Toll-like receptor 7 (*TLR7)*, X:9165034 (p-value = 1.14 x 10^{-5}; OR=1.3) closet *FAM9B* (Family With Sequence Similarity 9 Member B) and rs12398129 (*p*-value = 1.72 x 10^{-4}; OR = 1.39) on *CXorf61* was successfully replicated in other populations. In addition, we also identified a number of loci specific with Thai population such as rs6528443 (*p*-value = 8.71 x $10^{-6;\ OR}$ = 3.55) on *GPR101* and rs7052503 (*p*-value = 2.46 x 10^{-4}; OR = 0.77) nearly *MIR891A*. These loci are involved immune systems may affect to Thai SLE patients, which worth to be further investigated.

Field of Study:	Molecular Science of Medical Microbiology and Immunology	Student's Signature .....
Academic Year:	2020	Advisor's Signature .....
		Co-advisor's Signature .....

ACKNOWLEDGEMENTS

This thesis cannot be successful without the hand of these people. First, my advisor Pattarin Tangtanatrakul, Ph.D. is the first one who makes me know about coding and science. I never thought I would be able to code programs for analysis. Still, she encourages me and supports me in any way, such as the material storage server for running the program, the suggestion from the beginning, sacrificing personal time to recheck my work, and giving me a chance to connect with experienced people. Next is my co-advisor Wang Yong Fei, Ph.D. the important one who makes me understand the program and the genetic statistic. He suggested with his kind and cheerful. I am grateful to work with my co-adviser. I have to thank you for your kind and sacrifice. I want to thank you, Mr. Manon Boonbangyang, who help me in a tough time with the impute programming. Without his hand, I could not finish that step. In addition have to mention to Mr. Lei Yao, who was helpful and advised me anytime. I received many experiences from this project, from study with myself, learn from my mistakes. In the future, I will be better off your suggestions. Thank you a lot for your help.

Krisana Jaiwan

TABLE OF CONTENTS

Page

ABSTRACT (THAI) iii
ABSTRACT (ENGLISH) iv
ACKNOWLEDGEMENTS v
TABLE OF CONTENTS vi
LIST OF TABLES viii
LIST OF FIGURES ix
CHAPTER I INTRODUCTION 1
Rational and background 1
Objective 2
Conceptual framework 3
CHAPTER II LITERATURE REVIEW 4
Systemic lupus erythematosus (SLE) 4
Genetic variation 5
Linkage disequilibrium (LD) 6
Genotyping 7
Genotyping format files output 9
Genetic Susceptibility in SLE on X-chromosome 11
Genome-wide associated study (GWAS) 17
CHAPTER III MATERIALS AND METHODS 19
Data collection 19
Quality Control data filtration 22
Pre-phasing and imputation data 23
Association analysis 24
Function annotation 25
CHAPTER IV RESULTS AND DISCUSSION 26

Quality control and Impute result 26
Association result 28
Function annotation GTEx result 33
Discussion 39
CHAPTER V CONCLUSION 43
Conclusions 43
Future Perspective 43
APPENDIX 45
REFERENCES 55
VITA 62

LIST OF TABLES

Page

Table 1 SNPs on X chromosome associated with SLE in Anhui China.....13

Table 2 SLE patients' characteristics from Pattarin Tangtanatakul, Chisanu Thumarat et al., 2020.....21

Table 3 Sample number in primary data and secondary data after quality controls cut-off and imputation.....26

Table 4 List of SNPs on X chromosome association significant at $p < 1 \times 10^{-5}$ from Meta-analysis of the X chromosome in females.....28

Table 5 List of SNPs on X chromosome association significant at $p < 1 \times 10^{-5}$ from Meta-analysis of the X chromosome in males.....29

Table 6 List of know SNPs on X chromosome association significant at $p < 1 \times 10^{-3}$ from Meta-analysis of X chromosome in primary dataset and secondary dataset. 30

Table 7 List of novel SNPs on X chromosome association significant at $p < 1 \times 10^{-3}$ from Meta-analysis of X chromosome in primary dataset and secondary dataset. 30

Table 8 Results of eQTL analyses for novel loci in multiple human tissues.....33

Table 9 List of SNPs association significant at $p < 1 \times 10^{-3}$ from Meta-analysis of the X chromosome from primary data and secondary data. 45

LIST OF FIGURES

Page

Figure 1 Pathogenesis of SLE....5

Figure 2 Allele....6

Figure 3 Genotype and Phenotype....6

Figure 4 The suitability of detection of SNPs by different genotypes....8

Figure 5 The Infinium HTS Workflow–The Infinium HTS format provides a rapid 3-day workflow with minimal hands-on time....9

Figure 6 Overview of various commonly used PLINK files....10

Figure 7 The online single nucleotide polymorphism (SNP) databases....11

Figure 8 Odds ratios for SLE risk loci based on the immunological pathways they affect....12

Figure 9 Single nucleotide polymorphisms (SNPs) [43]....18

Figure 10 Genome-wide association studies (GWAS) pipeline....18

Figure 11 Sample collection: Inclusion and Exclusion criteria....19

Figure 12 1982-Revised American College of Rheumatology (ACR) Criteria for Diagnosis of SLE. The patient fulfils 4 of 11 criterion will classify to SLE....20

Figure 13 Genotyping call rate per SNPs....23

Figure 14 Pre-phasing and Imputing....24

Figure 15 Flow process chart....25

Figure 16 Quantile- Quantile plot (Q-Q plot) of P values from the X chromosome-wide association study (blue line). −log10 P values were plotted against the expected null distribution (red line) with inflation factor value (λ) in two datasets and meta of two datasets....27

Figure 17 Manhattan plot of meta-analysis of the X chromosome in females from the primary dataset and secondary dataset. The cut-off p-value is 1×10^{-5} (red line). According to statistical power analysis. The significant locus in the SLE is labeled in orange....29

Figure 18 Manhattan plot of meta-analysis of the X chromosome from the primary dataset and secondary dataset. The cut-off p-value is 1×10^{-3} (red line). According to

statistical power analysis. The novel significant locus in SLE is labeled in green. The known significant locus is labeled in yellow. 31

Figure 19 The Locus zoom plots of 8 novel SNPs (A-H) show the location and SNPs surrounding LD (r2). 33

CHAPTER I
INTRODUCTION

Rational and background

Systemic lupus erythematosus (SLE) is an autoimmune disease that has a high incidence and prevalence in Asia around 2.5-8.6 per 100,000 [1] [2] [3]. Since the severity of SLE is heterogeneous, a number of observations have suggested that it was possibly due to the nationality background and environmental exposures [1]. Previous studies have shown that African, Hispanic, and Asian have a higher severity when compared to Caucasian nationality [1]. Correspondingly, the vital organs affected, such as neurological and renal systems, are often found in Thai SLE patients [4]. These indicated that the Thai population's genetic background might contain specific SLE susceptible variants leading to higher disease severity.

SLE has 66% of heritability, mainly in young adolescent females [5]. There has been hypothesized that estrogens and double X-chromosomes in females are predisposing factors contributing to SLE [1]. Defective epigenetic silencing (methylation or histone modification) or X-Chromosome Inactivation (XCI) has been suggested to involve in SLE development [6]. Moreover, evidence in Klinefelter syndrome patients, an extra X chromosome (47, XXY) abnormality, showed an increased risk of developing SLE by 14-fold [7, 8]. In the same manner, a previous study found that 23%[9]. Thus, a genetic variation on X-chromosome might be associated with SLE development, especially in females.

Single Nucleotide Polymorphisms (SNPs) is genetic polymorphisms expressed greater than 1 % in the populations [10]. It can localize on exons or splice sites, thereby interfere gene expression [11]. The SNPs have a huge influence on human phenotype, drug response and disease development [11]. Genome-wide association study (GWAS), a high throughput technique to identify SNPs, becomes a popular approach to characterize the susceptible loci-associated with disease [12]. In SLE, a number of GWAS studies have been reported susceptible SNPs on X-chromosome in many

different population [13]. For examples, in Chinese background has been reported rs4830478 (p-value = 1.9 x 10^{-9}; OR = 1.33) on TLR*7*(Toll-like receptor 7), rs17422 (p-value = 1.4 x 10^{-14}; OR = 1.26) on *TMEM187* (Transmembrane protein 187), rs1059702 (p-value = 1.8 x 10^{-16}; OR = 0.76) on *IRAK1* (Interleukin-1 receptor-associated kinases) and rs887369 (p-value = 9.2 x 10^{-7}; OR = 1.16) on *CXorf21* (Chromosome X open reading frame 21) [13, 14] [15].

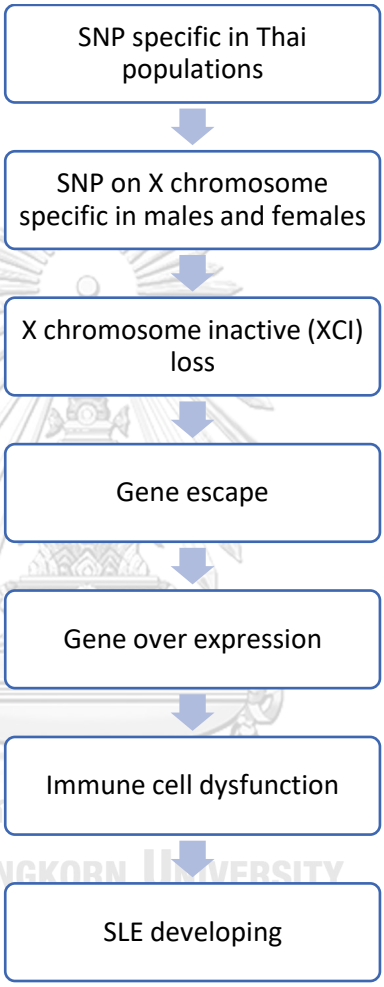
Interestingly, functional annotation of some reported SLE susceptible alleles is to drive X-chromosome silencing failure and promotes immune cell activation, leading to autoimmune susceptibility [13, 16]. For example, the rs2734647 in methyl-CpG binding protein 2 (*MECP*) allele, the highest and most consistent SLE susceptible alleles on the X chromosome has been demonstrated to affect epigenetic regulation, especially on the X chromosome [17] [7]. The rs1059702 on *IRAK1* is identified as a risk allele in Chinese, Japanese and Korean [18]. This SNP influences the coding sequence at position 196 in *IRAK1* gene, resulting in a missense mutation from serine to phenylalanine. This non-synonymous variation increases NF-κB transcription activity, thereby enhance inflammatory response in SLE patients [19]. Moreover, the rs5914778 within *LINC01420* (Long intergenic noncoding RNA 1420); a specific expression in females [17], this SNPs disrupt DNase I hypersensitive site that controls X inactivation [7]. The rs13440883 in *GPR173* (G protein-coupled receptors 173 are risk loci in Chinese, European, and Thai SLE patients [13]. The *GPR173* gene encodes the G protein-coupled receptor 1 family in the T-cell surface as extracellular ligands affect T-cell activation and morphology [20].

Although there is a specific allele study of X-chromosome polymorphisms in Thais, the whole genome analysis of SNPs on X-chromosome in the Thai population is still absent. Our study is the first study that characterizes the SNPs on X-chromosome in the Thai populations. This could open the windows for novel targeted therapy, which is one of Thailand's future expected treatment policies.

Objective

To identify entire SNPs on X-chromosome associated with Thai SLE patients.

Conceptual framework



CHAPTER II
LITERATURE REVIEW

Systemic lupus erythematosus (SLE)

SLE is a complex systemic inflammation resulting from auto antibody complexes with self-antigen [21]. SLE is one of the global health problems because of its chronic and relapsing-remitting disease course. In addition, standard treatment currently inhibits the immune systems, but it cannot cure the disease. SLE patients with more than 10-years of immune-suppressive drug treatment are usually ended up with mortality from severe infection or vital organ failure [22].

Females have a higher disease frequency more than males, in the ratio 9:1. This disease typically developed from teenage until late adulthood, approximately 15 – 44 years old [1, 23]. Because of the complication of disease pathogenesis, it is difficult to identify the virtual cause of disease. However, the risk factors associated with SLE development has been proposed including 1) environmental triggers such as UV, tobacco, infections, silica, and solvent [24], 2) Drug-induced lupus (DIL) such as hydralazine and procainamide [25], 3) hormonal factors such as progesterone, estrogen [1], and 4) genetic background. These factors induce gene alteration and start autoimmune disorders. Their immune system loses self-tolerance and produces autoantibody to stimulating inflammatory systemic (Figure 1).

It is well established that disease severity can be varied according to their nationality background [22]. Previous studies highlighted that Black, Asians, and Hispanics have a high incidence rate and more severity when compare with white people [26]. In Thailand, the prevalence of vital organs affected is high, especially renal disorders [27]. We hypothesize that specific genetic variants associated with SLE in the Thai population may be observed regarding this finding.

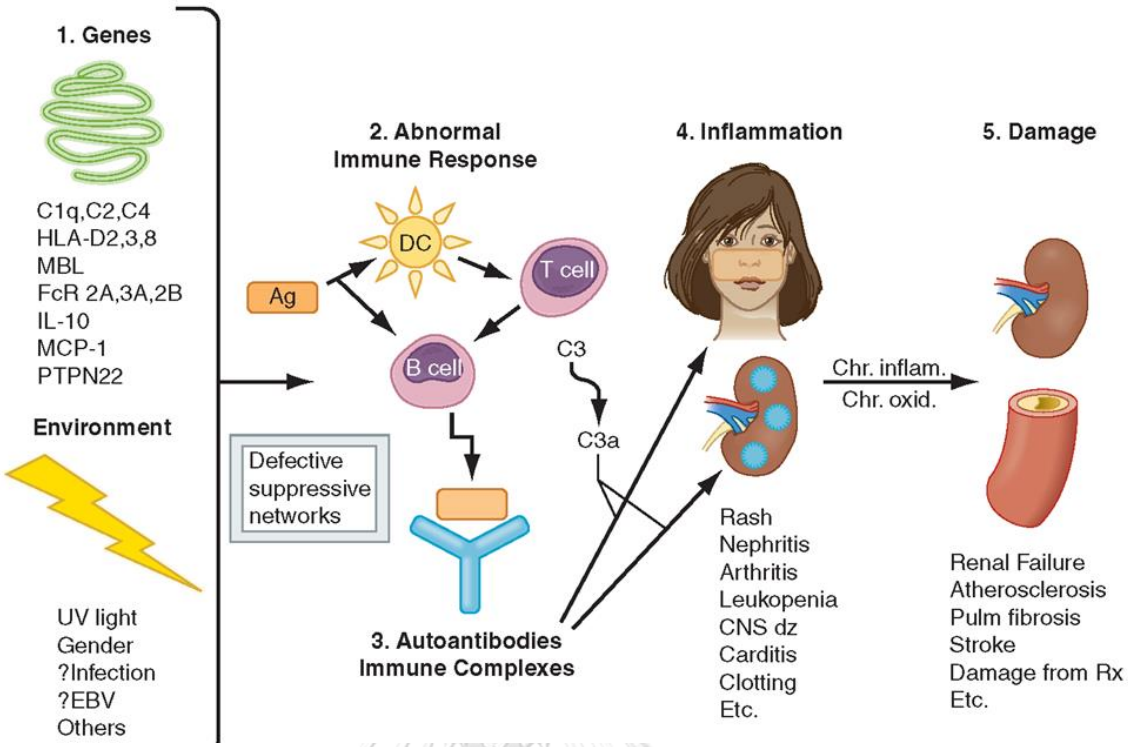


Figure 1 Pathogenesis of SLE (10)

Genetic variation

Firstly, we would like to review basic genetics and important terminology using in this study. Unit for storage genetic code is call genes. We call the location of genes on a chromosome is locus. Each locus contains a sequence for protein-coding and non-coding regions the sequence includes 4 base types, or we know as allele A, T, C, and G. This allele always has pairs at the same loci because we received alleles from parents. There are two types of allele homozygous allele and heterozygous allele. Though the term allele is used initially to describe variation among genes and non-coding regions. Characteristic of allele influence on genotypes in an individual that affects phenotype expression. Therefore, allele frequency is necessary to analyze phenotype in populations [28].

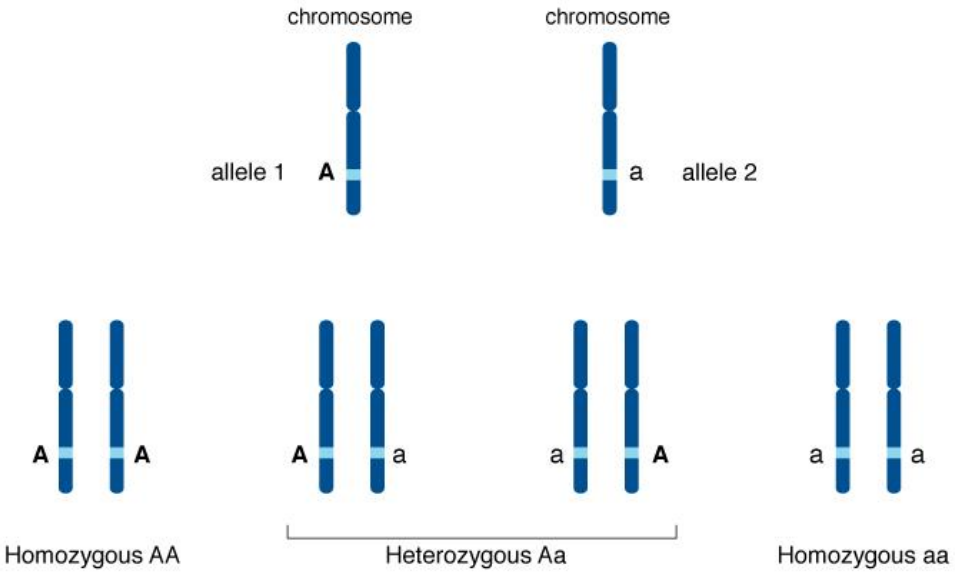


Figure 2 Allele [29]

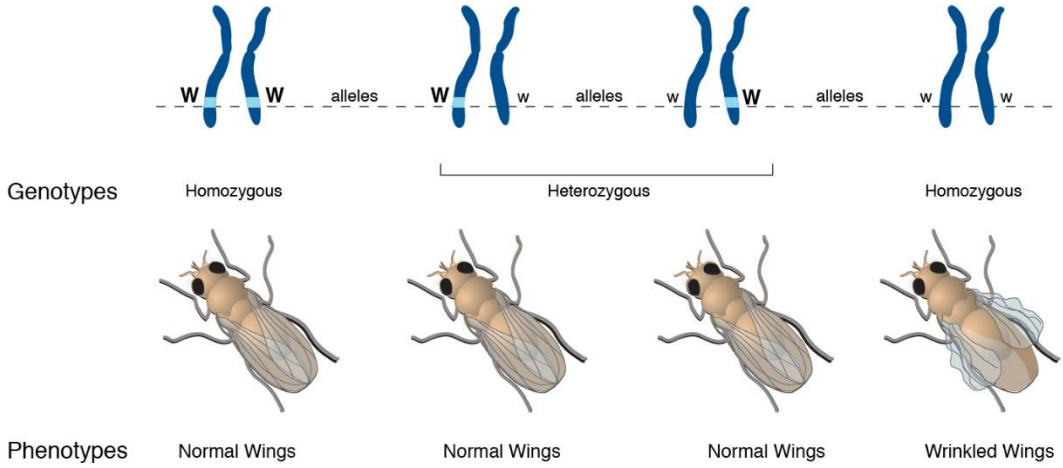


Figure 3 Genotype and Phenotype [29]

Linkage disequilibrium (LD) is used to measure alleles the non-random association of alleles at two or more positions that can be coinherited. LD measure by r^2 calculation [28].

Example calculation LD (r^2)

SNP locus A: A1=T, A2=C SNP

locus B: A1=1, A2=G

Haplotype Symbol Frequency

Haplotype	Symbol	Frequency
A1B1	x11	0.6
A1B2	x12	0.1
A2B1	x21	0.2
A2B2	x22	0.1

Calculated allelic frequency

Allele	Symbol	Frequency
A1	p1	0.7
A2	p1	0.3
B1	q1	0.8
B2	q2	0.2

D = x11 – p1q1: D = 0.6 – (0.7) (0.8) = 0.04
or
D = (x_{11}) (x_{22}) – (x_{12}) (x_{21}): *D* = (0.6) (0.1) – (0.1) (0.2) = 0.04
r^2 Calculating

$$r^2 = \frac{D^2}{p1q1p2q2}$$

Cut off $r^2 \geq 0.2$ are interpreted to SNPs have LD, if $r^2 = 1$ is SNPs in complete LD.

Genotyping

The method for determining differences in a person's genetic (genotype) by determining their DNA sequence and comparing to another person sequence or a reference sequence. It explains the alleles an individual has inherited from their parents. Polymerase chain reaction (PCR) is a commonly used genotyping technique that must prepare a primer-pair and target-specific fluorescent probe. It takes a long time for a sensitive and specific way to detect SNPs [28]. The Genotyping method for GWAS must be appropriate with high-throughput sample processing to deliver high-quality, genome-wide information, high-density oligonucleotide SNP arrays. Amounts of thousands of probes are arrayed on a small chip, recognizing many SNPs to be interrogated together because SNP alleles are only unlike in a single base and is a complication to optimal hybridization conditions for all probes on the array. Infinium

Asian Screening Array-24 v1.0 Bead Chip performs a genome-wide assay that can genotype over 500,000 human SNPs [30]. The Infinium Global Screening Array-24 v1.0 (GSA) Bead Chip is an advanced genotyping array that supports high-throughput processing of thousands of samples per week for population-scale studies medicine research. Robust, High-Quality Assay maintains the same data quality of Illumina genotyping arrays with call rates >99% and reproducibility >99.9% [30].

1 - 10	• TaqMan • LightTyper • Pyrosequencing
1 - 500	• SNaPshot • SNPlex • Sequenom MassARRAY • Illumina Golden Gate with BeadXpress readout
384 – 3,072	• Illumina Golden Gate with iScan readout
6,000 – 70,000	• Illumina Infinium iSelec Custom Beadchip
500,000 – 4,800,000	• Illumina Omni Whole-Genome Array • Affymetrix 6.0 Array

Figure 4 The suitability of detection of SNPs by different genotypes [30]

Figure 5 The Infinium HTS Workflow–The Infinium HTS format provides a rapid 3-day workflow with minimal hands-on time [30].

Genotyping format files output

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; other=unknown)

Phenotype (-9 missing, 0 missing, 1 unaffected, 2 affected)

MAP file describes a single marker and must contain exactly 4 columns (23):

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

rs# or SNP identifier

Genetic distance (Morgans)

Base-pair position (bp units)

Binary PED files including 3 file FAM fila, BED file, BIM file. These files are often to analysis in the plink tool [31].

*.fam

FID	IID	PID	MID	Sex	P
1	1	0	0	2	1
2	2	0	0	1	0
3	3	0	0	1	1

*.bed

Contains binary version of the SNP info of the *.ped file. (not in a format readable for humans)

*.bim

Chr	SNP	GD	BPP	Allele 1	Allele 2
1	rs1	0	870000	C	T
1	rs2	0	880000	A	G
1	rs3	0	890000	A	C

Legend			
FID	Family ID	rs{x}	Alleles per subject per SNP
IID	Individual ID	Chr	Chromosome
PID	Paternal ID	SNP	SNP name
MID	Maternal ID	GD	Genetic distance (morgans)
Sex	Sex of subject	BPP	Base-pair position (bp units)
P	Phenotype	C{x}	Covariates (e.g., Multidimensional Scaling (MDS) components)

Figure 6 Overview of various commonly used PLINK files [31].

Researchers can currently use database SNPs to identify rs number as an accession number to refer to specific SNPs. It stands for Reference SNP cluster ID. Many SNPs database including the International HapMap Project aims to develop a haplotype map or HapMap of the human genome to explain the patterns of human genetic variation. To find variants affecting traits, and drugs response, and other factors. The project is freely available for the researcher [32]. However, the researcher is available to access many databases, including the National Center for Biotechnology Information (NCBI), National Human Genome Research Institute (NHGRI), SNPedia [31]. 1000

Genome Project (1KGP): The project started in 2008 and finished in 2015 become the largest resource of human genetic variation or called single nucleotide polymorphism (SNPs) with frequencies of at least 1% in the populations. They identify around 40 million SNPs of the population from every region and assemble for one project.

Database	Host organization	Gateway URL for initiating SNP data searches
dbSNP	NCBI	http://www.ncbi.nlm.nih.gov/sites/entrez?db=snp
HapMap	The HapMap Consortium	http://www.hapmap.org/cgi-perl/gbrowse/
Ensembl	EMBL-EBI/Sanger Center	http://www.ensembl.org/Homo_sapiens/index.html
Santa Cruz	University of California, Santa Cruz	http://genome.ucsc.edu/cgi-bin/hgGateway
Perlegen	Perlegen Sciences	http://genome.perlegen.com/browser/index_v2.html
Assays-on-Demand	Applera (Applied Biosystems)	https://products.appliedbiosystems.com/ab/en/US/adirect/ab?cmd=ABGTKeywordSearch&catID=600769
SeattleSNPs	US NHLBI (PGA)	http://gvs.gs.washington.edu/GVS/

NCBI National Center for Biotechnology Information, *NHLBI*, National Heart, Lung, and Blood Institute, *PGA* Program for Genomic Applications

Figure 7 The online single nucleotide polymorphism (SNP) databases [33].

Genetic Susceptibility in SLE on X-chromosome

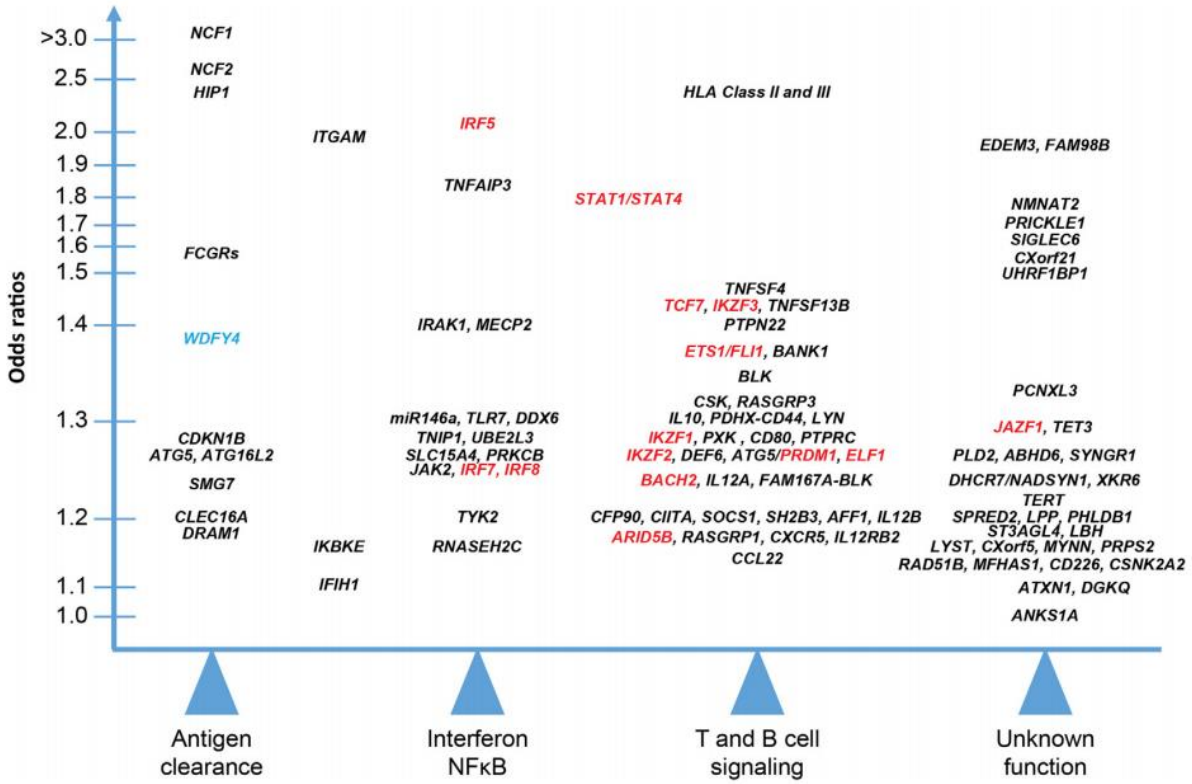


Figure 8 Odds ratios for SLE risk loci based on the immunological pathways they affect [16].

Genetic is one acritical key factor contributing to SLE [16]. A previous study reported various genetic susceptibility of SLE which are involved mainly in the immune cell. The X chromosome are playing a crucial role in SLE, especially in females. Typically, one copy of the X chromosome in females are being silent through epigenetic mechanisms [34]. This process is called X-chromosome inactivation. The silencing is essential to keep specific gene expression in control and allow specific genes to express, specialize to immune genes coding on the X chromosome. Despite 23% of genes on the X chromosome are escaped from the silencing mechanism, resulting in females bias found in SLE [9]. In addition, many genes are escaping from the silencing process too. For example, *TLR7*, *TMEM187*, *IRAK1*, and *CXorf21* are reported to escape from the silencing process [34-36]. The *CXorf21* is functioned to induce interferon IFN-γ and IFN-α in monocytes and B-cell [37], as well as *TLR7*, is an innate pattern recognition receptor that targeted single-stranded RNA, resulting in stimulating IFN-type I response.

Several studies identified SNPs on the X chromosome are SLE susceptibility specific with ethnicity. Example SNPs in Chinese SLE from GWAS are shown in Table 1. Taken together, this confirmed that the X chromosome may add a significant gender bias in female SLE patients.

Table 1 SNPs on X chromosome associated with SLE in Anhui China [17].

SNP	Gene	Minor Allele	OR	SE	L95	U95	P-value
rs1059702	IRAK1	G	0.72	0.09	0.61	0.85	1.34E-04
rs2239464	MECP2	G	0.72	0.09	0.6	0.85	1.74E-04
rs2734647	MECP2	G	0.73	0.09	0.62	0.86	2.18E-04
rs6631753	DMD	A	0.78	0.07	0.69	0.9	3.85E-04
rs5956251	-	A	0.76	0.08	0.65	0.89	4.06E-04
rs5972178	-	C	0.8	0.07	0.7	0.91	5.68E-04
rs10218247	-	A	0.77	0.08	0.66	0.89	5.71E-04
rs2536576	-	G	1.36	0.09	1.14	1.61	6.41E-04
rs1860995	ATP1B4	G	1.26	0.07	1.1	1.44	6.53E-04
rs1860814	-	G	1.35	0.09	1.13	1.61	7.45E-04
rs5914638	-	G	1.28	0.07	1.11	1.49	7.95E-04
rs2516036	FAM120C	G	0.68	0.12	0.54	0.85	8.90E-04
rs2266888	TMEM187	G	0.77	0.08	0.66	0.9	9.57E-04
rs2495794	FAM120C	A	0.68	0.12	0.54	0.86	1.09E-03
rs2806010	MIR548AE1	G	1.24	0.07	1.09	1.4	1.13E-03
rs17422	HCFC1	A	0.77	0.08	0.66	0.9	1.17E-03
rs3761622	TLR8-AS1	C	0.74	0.09	0.61	0.89	1.21E-03
rs942273	MIR548AE1	C	1.23	0.07	1.09	1.4	1.23E-03
rs1408095	MIR548AE1	A	1.23	0.07	1.08	1.4	1.37E-03
rs2495782	FAM120C	A	0.69	0.12	0.55	0.87	1.59E-03
rs17326228	MORC4	G	1.23	0.07	1.08	1.41	1.61E-03
rs5960060	-	A	1.23	0.07	1.08	1.4	1.66E-03
rs5960395	PHF8	A	0.68	0.12	0.54	0.87	1.69E-03
rs17329976	-	G	0.77	0.09	0.65	0.91	1.75E-03
rs12556165	-	C	0.82	0.07	0.72	0.93	1.81E-03
rs12688561	FAM120C	A	0.69	0.12	0.55	0.87	1.82E-03
rs5909765	-	A	0.74	0.1	0.61	0.89	1.92E-03
rs7062536	PRPS2	A	0.8	0.07	0.7	0.92	1.92E-03
rs4288493	-	G	1.24	0.07	1.08	1.43	1.93E-03
rs6612662	-	G	1.24	0.07	1.08	1.43	1.96E-03
rs4907832	-	A	0.76	0.09	0.63	0.9	1.98E-03
rs4535870	-	C	1.24	0.07	1.08	1.43	1.98E-03
rs5914778	LINC01420	A	1.25	0.07	1.09	1.44	2.00E-03
rs5914860	-	C	1.24	0.07	1.08	1.43	2.09E-03
rs5936901	-	G	1.23	0.07	1.08	1.41	2.11E-03

rs17267184	RPS6KA6	A	0.74	0.1	0.62	0.9	2.11E-03
rs1343096	-	A	0.74	0.1	0.62	0.9	2.15E-03
rs1560514	FAAH2	A	1.24	0.07	1.08	1.42	2.16E-03
rs1323751	MIR548AE1	A	1.22	0.07	1.07	1.38	2.18E-03
rs17281143	-	G	1.26	0.08	1.09	1.47	2.22E-03
rs6418619	-	A	0.8179	0.06579	0.7189	0.9305	2.25E-03
rs6612720	-	A	1.24	0.07048	1.08	1.424	2.25E-03
rs2026622	LINC01420	A	1.24	0.07049	1.08	1.424	2.27E-03
rs5914806	-	G	1.24	0.07049	1.08	1.424	2.27E-03
rs6616617	MORC4	A	1.225	0.06666	1.075	1.396	2.34E-03
rs2532869	-	C	0.819	0.0658	0.7199	0.9318	2.41E-03
rs5960810	-	A	1.239	0.07051	1.079	1.422	2.41E-03
rs5970959	PTCHD1-AS	G	1.26	0.08	1.09	1.47	2.42E-03
rs6638625	-	A	1.244	0.07201	1.08	1.433	2.43E-03
rs6521788	-	G	0.6944	0.1205	0.5483	0.8793	2.46E-03
rs5960612	PHF8	A	0.6924	0.1216	0.5456	0.8788	2.51E-03
rs5961058	-	A	0.8215	0.06511	0.7231	0.9333	2.53E-03
rs4379572	-	G	1.236	0.07041	1.077	1.419	2.60E-03
rs5914776	-	A	1.244	0.07272	1.079	1.435	2.66E-03
rs5922916	RPS6KA6	A	0.7489	0.09626	0.6201	0.9044	2.66E-03
rs2411864	-	G	0.77	0.09	0.64	0.91	2.67E-03
rs3764880	TLR8	A	0.7549	0.09364	0.6283	0.9069	2.67E-03
rs6611574	-	A	1.235	0.07046	1.076	1.418	2.75E-03
rs1527803	-	A	0.82	0.07	0.72	0.93	2.84E-03
rs6529663	-	G	1.24	0.07	1.08	1.43	3.03E-03
rs4826508	LINC01420	G	1.233	0.0708	1.073	1.417	3.06E-03
rs2335517	-	A	1.301	0.08897	1.093	1.549	3.09E-03
rs5914037	-	A	1.231	0.07035	1.073	1.413	3.12E-03
rs4843993	-	G	1.295	0.08755	1.091	1.538	3.14E-03
rs5915082	-	A	1.23	0.07	1.07	1.42	3.21E-03
rs5936343	-	A	1.294	0.08755	1.09	1.536	3.24E-03
rs5944365	-	A	1.286	0.0856	1.087	1.521	3.32E-03
rs3788935	TLR8	A	0.7606	0.09322	0.6336	0.913	3.33E-03
rs11094927	-	A	1.214	0.066	1.066	1.381	3.34E-03
rs995154	-	A	1.21	0.07	1.07	1.38	3.39E-03
rs5960307	-	G	0.8266	0.06505	0.7277	0.939	3.43E-03
rs5914785	LINC01420	A	1.229	0.07064	1.071	1.412	3.45E-03
rs5960235	SPIN3	G	1.228	0.07034	1.07	1.41	3.48E-03
rs12835268	-	A	0.8238	0.06637	0.7233	0.9382	3.50E-03
rs5914795	LINC01420	A	1.229	0.07079	1.07	1.412	3.59E-03
rs5913993	LINC01420	A	1.227	0.07063	1.069	1.41	3.73E-03
rs726441	-	A	0.8299	0.06438	0.7315	0.9415	3.78E-03
rs5933907	-	A	1.268	0.08218	1.08	1.49	3.81E-03
rs6641214	-	A	1.282	0.08586	1.083	1.517	3.81E-03

rs11094877	-	A	0.8262	0.06599	0.726	0.9403	3.82E-03
rs11091412	-	A	1.22	0.06888	1.066	1.396	3.88E-03
rs6612746	SPIN3	A	1.225	0.07033	1.067	1.406	3.90E-03
rs5978593	TLR8_AS1	G	0.7525	0.0986	0.6203	0.913	3.94E-03
rs2269368	ARHGAP4	G	0.817	0.07042	0.7117	0.938	4.11E-03
rs3810757	-	A	1.269	0.08315	1.078	1.494	4.12E-03
rs10854983	-	G	0.83	0.07	0.73	0.94	4.16E-03
rs6612721	-	A	1.225	0.07073	1.066	1.407	4.16E-03
rs5925798	-	A	0.82	0.07	0.71	0.94	4.17E-03
rs7884579	-	G	1.213	0.06736	1.063	1.384	4.18E-03
rs6571303	TMEM187	G	0.8068	0.07537	0.696	0.9352	4.39E-03
rs2056918	-	G	0.7275	0.1118	0.5844	0.9058	4.44E-03
rs5966868	-	A	0.7437	0.1041	0.6064	0.912	4.44E-03
rs7883778	-	G	0.6869	0.132	0.5303	0.8897	4.44E-03
rs5925786	-	A	0.8207	0.0696	0.7161	0.9407	4.53E-03
rs9306569	-	G	0.83	0.06572	0.7297	0.9441	4.58E-03
rs7065919	DMD	G	0.8288	0.06625	0.7279	0.9438	4.61E-03
rs5914036	SPIN3	A	1.22	0.07032	1.063	1.4	4.69E-03
rs5960936	-	A	0.8294	0.06617	0.7285	0.9442	4.70E-03
rs5963635	LOC286442	A	1.25	0.08	1.07	1.45	4.76E-03
rs6617836	-	A	0.7956	0.081	0.6788	0.9325	4.76E-03
rs1342219	-	A	0.7278	0.1128	0.5835	0.9079	4.85E-03
rs6622208	-	G	1.206	0.0665	1.059	1.374	4.85E-03
rs5918209	CASK	G	0.7832	0.08709	0.6603	0.9289	5.01E-03
rs6523960	-	G	1.208	0.06739	1.058	1.378	5.08E-03
rs5924847	-	C	1.251	0.08007	1.07	1.464	5.10E-03
rs6617830	-	A	0.7973	0.08095	0.6803	0.9344	5.13E-03
rs1323757	-	A	0.8303	0.0665	0.7288	0.9458	5.15E-03
rs5914893	-	G	1.22	0.07	1.06	1.4	5.17E-03
rs3859913	-	G	1.204	0.06631	1.057	1.371	5.20E-03
rs2890089	-	C	1.218	0.07064	1.061	1.399	5.21E-03
rs5936206	-	C	1.279	0.08811	1.076	1.52	5.25E-03
rs5913850	-	A	0.8319	0.06613	0.7308	0.947	5.38E-03
rs1937249	-	A	1.207	0.06752	1.057	1.377	5.39E-03
rs5916449	-	G	1.215	0.07017	1.059	1.395	5.43E-03
rs12013552	-	C	1.215	0.07019	1.059	1.395	5.46E-03
rs5921138	-	A	0.7492	0.1042	0.6109	0.9189	5.57E-03
rs2982249	-	A	1.217	0.07074	1.059	1.398	5.58E-03
rs11795541	-	G	1.237	0.07672	1.064	1.438	5.59E-03
rs5925802	-	A	0.8248	0.06955	0.7197	0.9452	5.60E-03
rs5977894	-	A	1.207	0.06809	1.057	1.38	5.63E-03
rs2188615	-	A	1.196	0.06468	1.053	1.358	5.69E-03
rs2188616	-	A	1.196	0.06468	1.053	1.358	5.69E-03
rs1925926	GDPD2	C	0.8097	0.07638	0.6971	0.9405	5.72E-03

rs7055735	DACH2	A	0.8101	0.07623	0.6977	0.9407	5.75E-03
rs5936524	EDA	C	1.216	0.07088	1.058	1.397	5.78E-03
rs2214563	-	G	1.234	0.07617	1.063	1.432	5.83E-03
rs7877755	-	A	0.7225	0.1179	0.5734	0.9103	5.83E-03
rs5975417	-	A	1.206	0.06801	1.055	1.378	5.89E-03
rs4403537	-	A	0.7515	0.1039	0.6131	0.9212	5.96E-03
rs9699111	-	A	0.8354	0.06549	0.7348	0.9498	6.02E-03
rs17344059	-	G	0.8153	0.07438	0.7047	0.9432	6.04E-03
rs5926470	-	A	1.202	0.06693	1.054	1.37	6.04E-03
rs6654792	-	G	1.259	0.08376	1.068	1.483	6.04E-03
rs2285563	ARX	C	1.195	0.06493	1.052	1.357	6.06E-03
rs5914994	FAAH2	A	1.212	0.06996	1.056	1.39	6.09E-03
rs5914700	-	G	1.217	0.07176	1.058	1.401	6.10E-03
rs5986613	-	G	1.262	0.08477	1.069	1.49	6.10E-03
rs697664	-	G	1.212	0.07009	1.056	1.39	6.15E-03
rs1467342	-	G	1.212	0.07022	1.056	1.391	6.17E-03
rs5959353	-	G	0.7367	0.1117	0.5919	0.9169	6.21E-03
rs5953534	-	A	1.327	0.1035	1.083	1.625	6.28E-03
rs765076	-	G	0.8179	0.07362	0.708	0.9449	6.33E-03
rs707346	SPIN2B	A	1.211	0.07017	1.055	1.39	6.35E-03
rs5928345	IL1RAPL1	C	1.272	0.08823	1.07	1.512	6.48E-03
rs5961051	-	A	0.8357	0.06593	0.7344	0.951	6.48E-03
rs2808725	-	A	1.192	0.06468	1.05	1.353	6.53E-03
rs5914902	-	A	1.21	0.07	1.05	1.39	6.61E-03
rs5978005	-	G	1.201	0.06746	1.052	1.371	6.69E-03
rs4826580	-	A	1.211	0.07053	1.054	1.39	6.73E-03
rs5908660	-	G	1.317	0.1024	1.078	1.61	7.14E-03
rs5977810	-	G	1.201	0.06814	1.051	1.373	7.14E-03
rs859603	SASH3	G	1.259	0.08565	1.064	1.489	7.20E-03
rs5960434	-	A	1.196	0.06686	1.049	1.364	7.32E-03
rs209764	NDP	A	0.8377	0.0661	0.7359	0.9535	7.36E-03
rs7059234	-	G	1.19	0.06489	1.048	1.351	7.37E-03
rs6610903	EFHC2	A	1.322	0.1045	1.077	1.623	7.51E-03
rs1266322	-	A	1.21	0.0714	1.052	1.392	7.57E-03
rs5986629	-	G	1.243	0.08142	1.059	1.458	7.61E-03
rs5911059	-	G	0.8408	0.06498	0.7403	0.955	7.62E-03
rs512119	-	G	1.203	0.06939	1.05	1.379	7.67E-03
rs6628425	IL1RAPL1	G	1.226	0.07657	1.055	1.424	7.83E-03
rs6627929	-	A	0.8312	0.06956	0.7253	0.9526	7.85E-03
rs2280964	CXCR3	A	1.198	0.06807	1.048	1.369	7.93E-03
rs6527265	DMD	A	0.838	0.0666	0.7355	0.9549	7.98E-03
rs5923562	DACH2	A	1.204	0.06999	1.05	1.381	8.03E-03
rs5914734	-	A	1.209	0.07152	1.05	1.39	8.08E-03
rs723556	ARAF	G	0.84	0.06586	0.7383	0.9557	8.10E-03

rs6617168	-	G	1.192	0.06661	1.046	1.358	8.39E-03
rs5928201	DMD	A	0.7415	0.1136	0.5935	0.9265	8.49E-03
rs2071251	ZNF185	A	0.8348	0.06871	0.7296	0.9552	8.60E-03
rs6521411	-	C	1.206	0.0713	1.049	1.387	8.60E-03
rs7876155	FRMPD4	A	0.7902	0.08973	0.6628	0.9422	8.69E-03
rs5955456	-	G	1.237	0.08124	1.055	1.451	8.75E-03
rs5942373	-	C	0.8139	0.07877	0.6975	0.9498	8.96E-03
rs12009868	-	A	1.193	0.06744	1.045	1.361	8.97E-03
rs1489965	-	G	1.205	0.07143	1.048	1.386	9.05E-03
rs5975460	-	A	1.192	0.06753	1.045	1.361	9.14E-03
rs4830593	KAL1	A	1.207	0.0723	1.048	1.391	9.16E-03
rs5924783	-	G	0.8361	0.06867	0.7308	0.9566	9.17E-03
rs5924779	ZNF185	G	0.8361	0.06871	0.7308	0.9567	9.20E-03
rs5911011	-	A	0.7329	0.1196	0.5798	0.9266	9.38E-03
rs1860012	-	G	0.8413	0.0666	0.7383	0.9586	9.46E-03
rs5975439	-	A	1.192	0.06755	1.044	1.36	9.49E-03
rs5930628	-	A	1.186	0.06585	1.043	1.35	9.54E-03
rs2813808	-	G	1.186	0.06598	1.042	1.35	9.66E-03
rs5935409	-	C	1.239	0.08297	1.053	1.458	9.73E-03
rs5923542	DACH2	A	1.2	0.07063	1.045	1.378	9.76E-03
rs5962817	-	A	1.189	0.06727	1.042	1.357	9.98E-03
rs5933555	KDM5C	A	0.8124	0.08067	0.6936	0.9515	9.99E-03

Genome-wide associated study (GWAS)

The X chromosome has become popular to find SNPs susceptibility with the autoimmune disorder by Genome-wide association study (GWAS). GWAS is a high throughput technology for genotyping single nucleotide polymorphisms (SNP) from an interesting group applying statistical use. Single nucleotide polymorphisms are single base variation patterns which components of adenine (A), thymine (T), cytosine (C), or guanine (G). Stereotypes vary from person to person, ethnicity, and genetic background. That is why humans exhibit different phenotypes. The single-point base change of the SNP differs from the point of the mutation because the frequency of SNPs is greater than 1% in the population [38]. SNPs impact on the gene has reported 50% on non-coding regions, 25% are missense mutations, and the 25% left are silent mutations. The synonymous or nonsynonymous SNPs influence individuals' diseases exposed, drug response, and genome evolution [39].

The aim of genome-wide association studies (GWAS) is to identify single nucleotide polymorphisms (SNPs) of which the allele frequencies vary systematically as a function of phenotypic trait values. Identification of trait-associated SNPs may subsequently reveal new insights into the biological mechanisms underlying these phenotypes. Technological advancements allow investigation of the impact of large numbers of SNPs distributed throughout the genome [40-42].

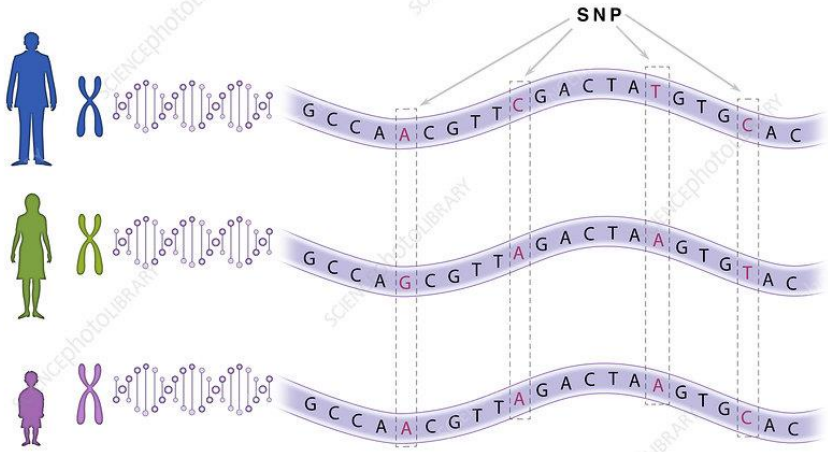


Figure 9 Single nucleotide polymorphisms (SNPs) [43]

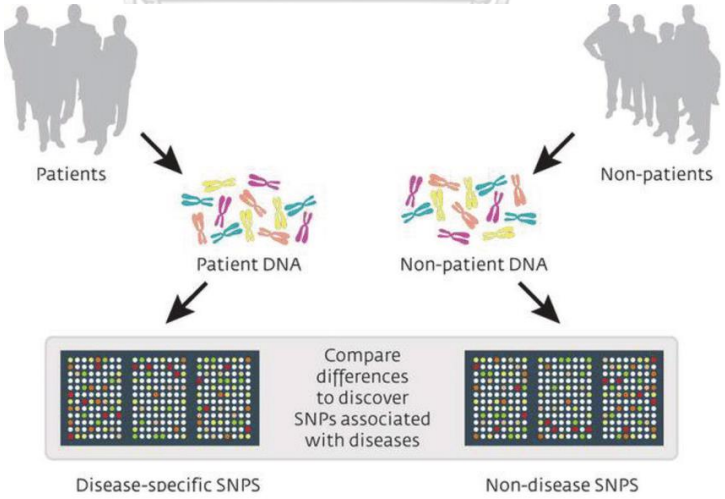


Figure 10 Genome-wide association studies (GWAS) pipeline [40].

CHAPTER III
MATERIALS AND METHODS

Data collection

We use previous publish data from Pattarin Tangtanatakul, Chisanu Thumarat et al., 2020

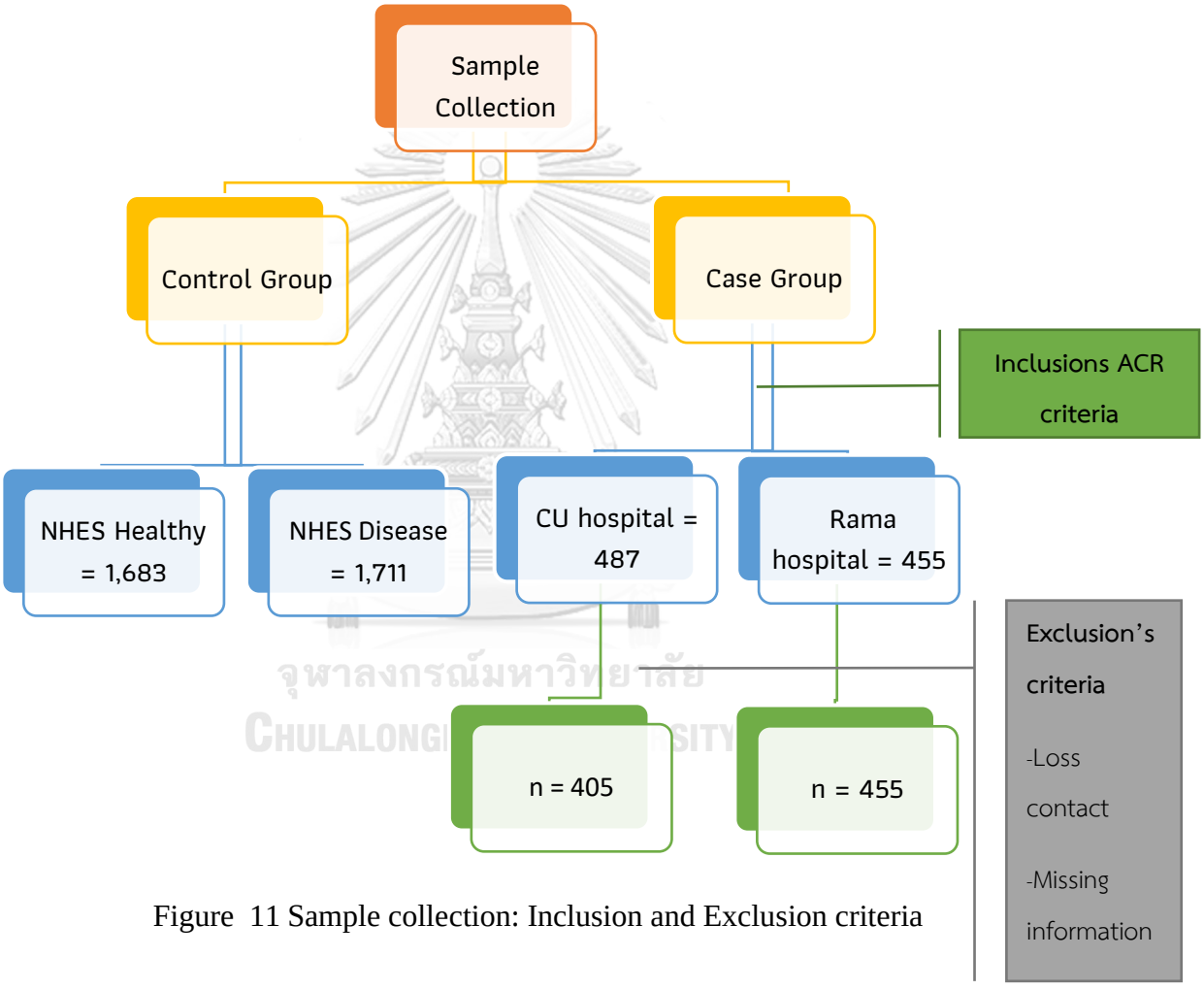


Figure 11 Sample collection: Inclusion and Exclusion criteria

Criterion	Definition
1. Malar rash	Fixed erythema, flat or raised, over the malar eminences, tending to spare the nasolabial folds
2. Discoid rash	Erythematous raised patches with adherent keratotic scaling and follicular plugging; atrophic scarring may occur in older lesions
3. Photosensitivity	Skin rash as a result of unusual reaction to sunlight, by patient history or physician observation
4. Oral ulcers	Oral or nasopharyngeal ulceration, usually painless, observed by physician
5. Non-erosive arthritis	Involving two or more peripheral joints, characterised by tenderness, swelling or effusion
6. Pleuritis or pericarditis	a. Pleuritis—convincing history of pleuritic pain or rubbing heard by a physician or evidence of pleural effusion OR b. Pericarditis—documented by electrocardiogram or rub or evidence of pericardial effusion
7. Renal disorder	a. Persistent proteinuria > 0.5 g/d or > than 3+ if quantisation not performed OR b. Cellular casts—may be red cell, haemoglobin, granular, tubular or mixed
8. Neurological disorder	a. Seizures—in the absence of offending drugs or known metabolic derangements; e.g. uraemia, ketoacidosis or electrolyte imbalance OR b. Psychosis—in the absence of offending drugs or known metabolic derangements; e.g. uraemia, ketoacidosis or electrolyte imbalance
9. Haematological disorder	a. Haemolytic anaemia—with reticulocytosis OR b. Leucopaenia—<4000/mm^3 on ≥ 2 occasions OR c. Lymphopenia—<1500/mm^3 on ≥ 2 occasions OR d. Thrombocytopaenia—<100,000/mm^3 in the absence of offending drugs
10. Immunological disorder	a. Anti-DNA: antibody to native DNA in abnormal titre OR b. Anti-Sm: presence of antibody to Sm nuclear antigen OR c. Positive finding of antiphospholipid antibodies on 1. An abnormal serum level of IgG or IgM anticardiolipin antibodies 2. A positive test result for lupus anticoagulant using a standard method, or 3. A false-positive test result for at least 6 months confirmed by Treponema pallidum immobilisation or fluorescent treponemal antibody absorption test
11. Positive anti-nuclear antibody	An abnormal titre of anti nuclear antibody by immunofluorescence or an equivalent assay at any point in time and in the absence of drugs

Figure 12 1982-Revised American College of Rheumatology (ACR) Criteria for Diagnosis of SLE. The patient fulfils 4 of 11 criterion will classify to SLE [44, 45].

Table 2 SLE patients' characteristics from Pattarin Tangtanatakul, Chisanu Thumarat et al., 2020

Patients' characteristics	Clinical cases			
	Observatory cohort n = 455[a]		Replication cohort n = 371[a]	
	n	(%)	n	(%)
Age of onset (mean ± SD)	30.38	± 13.68	30.39	± 11.43
Sex				
Female	425	(93.41%)[b]	337	(90.84%)[c]
Male	26	(5.71%)[b]	27	(7.28%)[c]
Clinical aspects				
Hematologic disorders	243	(53.41%)[b]	136	(36.66%)[c]
Neurological disorders	62	(13.63%)[b]	33	(8.89%)[c]
Ulcer	115	(25.27%)[b]	52	(14.02%)c
Discoid rash	161	(35.38%)[b]	49	(13.21%)[c]
Malar rash	142	(31.21%)[b]	82	(22%)[c]
Arthritis	133	(29.23%)[b]	148	(39.89%)[c]
Renal disorders	284	(62.42%)[b]	149	(40.16%)[c]
ANA	350	(76.92%)[b]	214	(57.68%)[c]

[a]The sample number after quality control processes

[b]The percentages of unknown clinical data (n/a) in the observatory dataset are listed here. Sex = 0.88%, hematologic disorder = 1.76%, neurological disorder = 2.20%, ulcer = 4.18%, discoid rash = 3.96%, malar rash = 5.71%, arthritis = 4.18%, renal disorders = 1.76%, and ANA = 9.89%

[c]The percentages of unknown clinical data (n/a) in the replication dataset are listed here. Sex = 0.00%, hematologic disorder = 36.93%, neurological disorder = 37.2%, ulcer = 37.4%, discoid rash = 37.2%, malar rash = 37.47%, arthritis = 37.2%, renal disorders = 37.74%, and ANA = 36.93%

Quality Control data filtration

This process is important. We must exclude poor data before the association test. We divide two-part of QC consist of individual filter and SNP filter.

Individual QC

In the individual part, we removed the gender discordant sample. This rule base on the heterozygosity rate on the X chromosome. Normally, females have a higher heterozygosity rate on the X chromosome than males that P-value can identify. Therefore, the P-value of the male is more than 0.8; the female is less than 0.2. If samples have a P-value discordant to those criteria, the program will remove those samples automatically [12]. Thus, the sample with a low heterozygosity rate of less than 95 % was removed after checking the allele on any locus of the individual sample. The sample that has high homologous and low heterozygosity must be deleted to reduce false positives. Heterozygosity rate identical by inbreeding coefficient is the level of genetically related mating between ancestry [12, 31]. Therefore, the sample with a high inbreeding coefficient means a low heterozygosity rate. The next step is identity by descent (IBD) or Pi-hat. This step helps to measure the pair of the individual sample who has genetic or allele from the same family identified by IBD calculation, called Pi-hat (PI). The Pi-hat are interpreting with 4 degrees including IBD = 1 for duplicates or monozygotic twins, IBD = 0.5 for first-degree relatives, IBD = 0.25 for second-degree relatives, IBD = 0.125 for third-degree relatives [12, 41]. After calculated the sample with Pi-hat, samples that are more than 0.125 must be removed to prevent false-positive error. The last step of individual QC is the genotype rate per individual check. We removed the sample with common SNPs expression less than 95 % [12, 31].

SNPs QC

In part of SNPs, we filter inferior quality SNPs from the study. Starting with test missing, we identify SNPs missing between case and control group. SNPs that are missing a difference p-value less than 1×10^{-4} must be removed. Next, we continue with genotyping call rate per SNPs check. In this step, we excluded SNPs with a low expression of less than 95 % of total SNPs [12, 31].

SNP Call Rate/Proportion

	SNP1	SNP2	SNP3	SNP4	SNP5
Sample1	**00**	AG	GG	GA	**00**
Sample2	**00**	GG	GG	AA	CC
Sample3	AC	**00**	GG	AA	CC
Sample4	AA	AG	GC	AA	CC
Sample5	AC	AA	**00**	AA	CA
SNP Call Rate	**60%**	80%	80%	100%	80%

ptember 2013 Entebbe, Uganda 23

Figure 13 Genotyping call rate per SNPs. (40)

Minor allele frequency (MAF) is the second or minor allele frequency beside the major allele and can be inherited together. MAF is used to estimate major allele distribution if MAF is less than 95%, meaning that SNPs have no significance with the disease and must be excluded [12, 31]. Hardy-Weinberg Equilibrium (HWE) is the role for life heredity, including 1. no mutation 2. random mating 3. no gene flow 4. infinite population size, and 5. no selection [12, 31, 46]. It is the principle for maintaining race. In GWAS, the HWE is used to find SNPs with over-genotype error (Statistical methods for GWAS), and we exclude SNPs genotype error at p-value < 1 x 10^{-4}.

Pre-phasing and imputation data

Data pass filtered overrun to pre-phasing to estimate of haplotype from genotype data by SHAPEIT software tool. The process compares study data with database resource 1000 genome project (1KP) or hg19 to evaluate the possible missing allele by the statistic applying. The pre-phasing step is to prepare data for impute step [47, 48].

Imputation genotyping data from the phasing step stage allow imputing data for boots variant SNPs from original data. The IMUPE2 software compares sample data with database references such as the 1000 Genomes project and HapMap Project, then

calculates statistics and replaces the possible variants within the sample sequence [49].

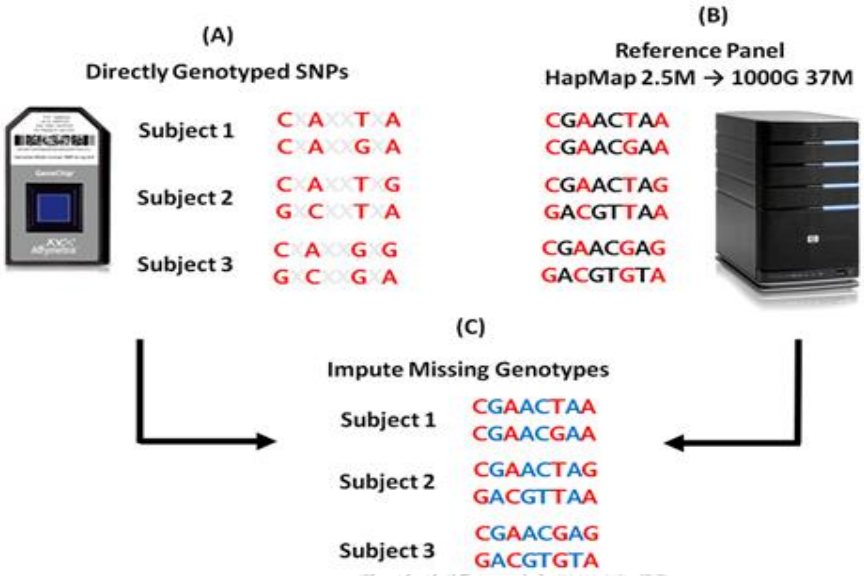


Figure 14 Pre-phasing and Imputing [50].

Association analysis

Testing for any SNP shows that it is significant to the trail of interest by statistical proof. [51]. In other words, this is a method for case and control test in which each SNP was tested for association with disease or whether trails of interest. The statistical for GWAS is dependent on sample conditional such as single locus, control of population stratification, generalized linear models for covariate control sample then select a model for proof hypotheses such as odd ratio (OR), Relative risk (RR), and Logistic regression. PLINK is a program for the analysis of single SNP associations in genome-wide studies. The tests implemented include [41]. We chose logistic regression for this GWAS because it is a task like linear regression. This model is appropriate for a binary test (case-control) with multiple variable analysis (age, gender, genetic variants) and predictable risk allele outputs. Logistic regression is standard apply in GWAS [13, 52].

After association analysis, data were obtained to quantile-quantile plot (Q-Q plot) for performing graphical study p-value (y-axis) against expected values (x-axis) under the

null hypothesis of no association from a theoretical χ^2-distribution [53]. To look up the SNPs' significant-high p-value from the average population.

Function annotation

The SNPs with significant association levels are investigated their function by using Haploreg v.4.1 algorithms. This program is used to identified susceptibility loci and integrates expression quantitative trait locus (eQTL) variants and their tissue-specific target genes from The Genotype-Tissue Expression (GTEx) project [54].

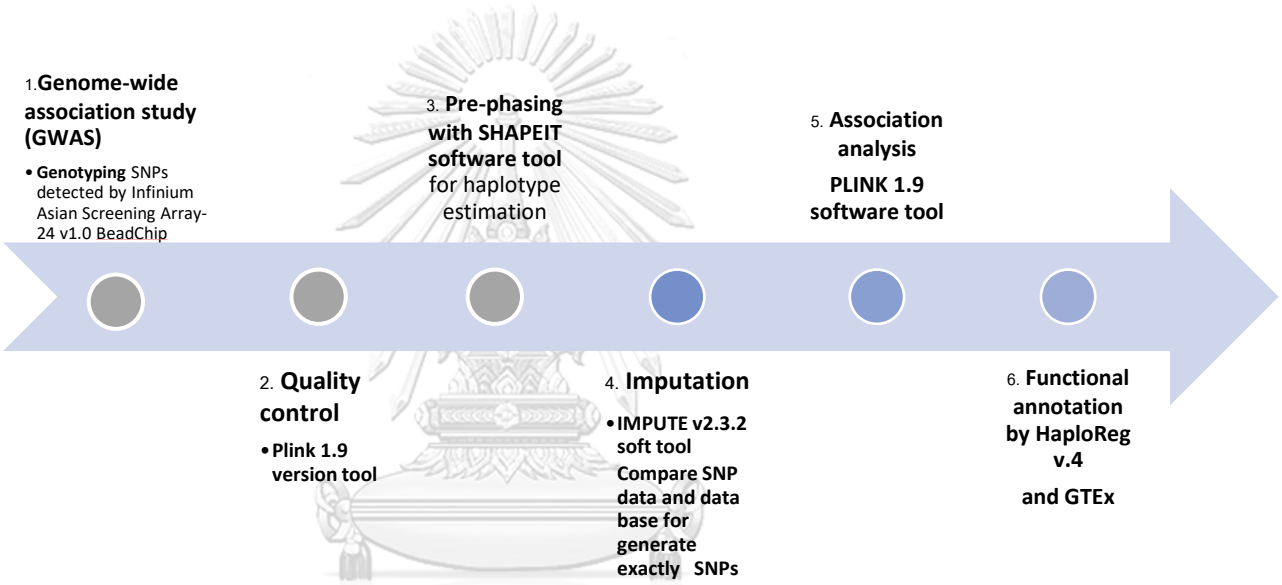


Figure 15 Flow process chart

CHAPTER IV
RESULTS AND DISCUSSION

Quality control and Impute result

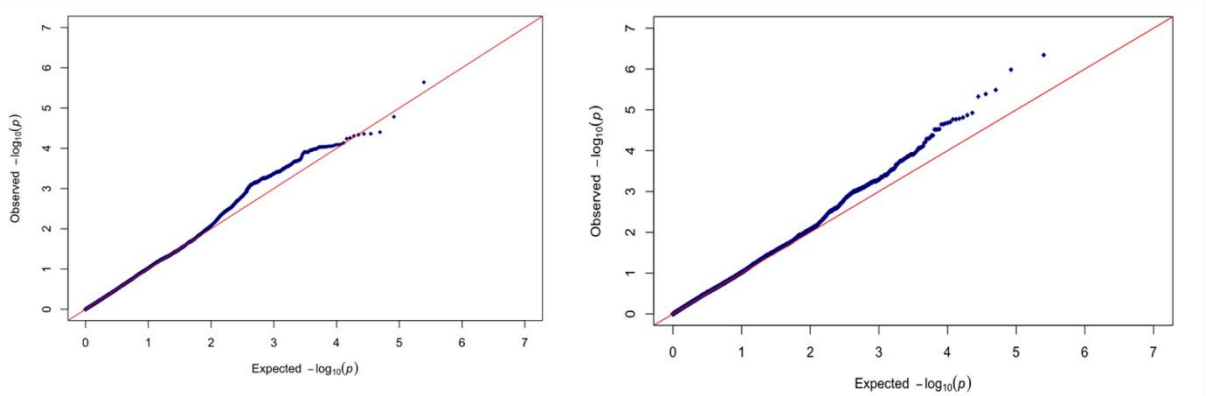
The low-quality samples were filtered out according to the criteria mentioned in the materials and methods session, such as heterozygosity rate, gender discordant, and missing genotyping rate. After the QC processing, primary dataset has females (n=1,232), males (n=809) (Table 1). Simultaneously, females (n=1,149) and males (n=783) are remaining in secondary dataset. The inflation factors of primary and secondary datasets are 0.988 and 1.09799, and the two-dataset merge is 1.09749 (Figure 14). These suggested that our population stratification has normal distribution when compared with the expected p-value. These confirmed that false-positive results could be devoid of the analysis.

Inflation factor value (λ) calculated from a median p-value of study divided by median p-value of theoretical distribution (normal distribution), which is 0.4549, the λ value not over 1.1 that is meaning the study p-value are normal distribution [55]. Add inflation factor calculation from our data and mention figure 14A, figure 14B, figure 14C.

Table 3 Sample number in primary data and secondary data after quality controls cut-off and imputation.

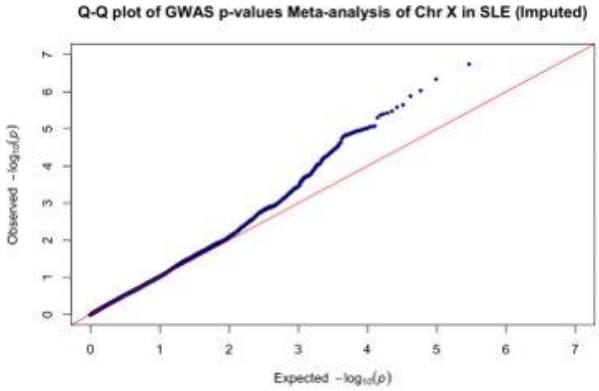
Primary dataset					
Gender	Raw data		QC Pass filter		After Imputation
	Sample (n)	Variants	Sample (n)	Variants	Variants
Females	1,319	21,510	1,232	15,163	132,144
Males	851	20,667	809	14,938	150,611
Total	2,170	42,177	2,041	30,101	282,755
Secondary dataset					
Gender	Raw data		QC Pass filter		After Imputation

	Sample (n)	Variants	Sample (n)	Variants	Variants
Females	1,226	21,380	1,149	15,096	132,180
Males	890	19,694	783	14,307	147,110
Total	2,116	41,074	1,932	29,403	279,290



(A) Primary dataset λ = 0.9883898

(B) Secondary dataset λ = 1.09799



C. Meta-two dataset λ = 1.09749

Figure 16 Quantile- Quantile plot (Q-Q plot) of P values from the X chromosome-wide association study (blue line). −log10 P values were plotted against the expected null distribution (red line) with inflation factor value (λ) in two datasets and meta of two datasets.

1)

Association result

Meta-analysis of females

SNPs on X chromosome specific in females at significant *p*-value < 1 x 10^{-5} was shown in Table 4. Unfortunately, we found only one locus which is rs1059702 (p-value = 4.54 x 10^{-7}, OR = 0.68) in *IRAK1*. However, linkage disequilibrium (LD) of that SNPs showed rs1734791 (p-value = 1.04 x 10^{-6}, OR = 0.69) in *MECP2* and rs6643656 (p-value = 4.74E x 10^{-6}, OR = 0.68) in *TMEM187* which are also significant associated with Thai female SLE patients. The Manhattan plot showing significant SNPs are presented in Figure 15.

Table 4 List of SNPs on X chromosome association significant at p < *1* x *10^{-5}* from Meta-analysis of the X chromosome in females.

dbSNP	BP	A1	A2	Locus	Annotation	MAF	OR	P
rs1059702	153284192	A	G	IRAK1	missense	0.2	0.69	4.54E-07
rs1734791	153330920	A	T	MECP2	intronic	0.81	0.69	1.04E-06
rs2734647	153292180	T	C	MECP2	3'-UTR	0.2	0.71	3.26E-06
rs2075596	153297392	A	G	MECP2	intronic	0.2	0.71	4.10E-06
rs6643656	153254605	C	G	TMEM187	synonymous	0.82	0.69	4.74E-06

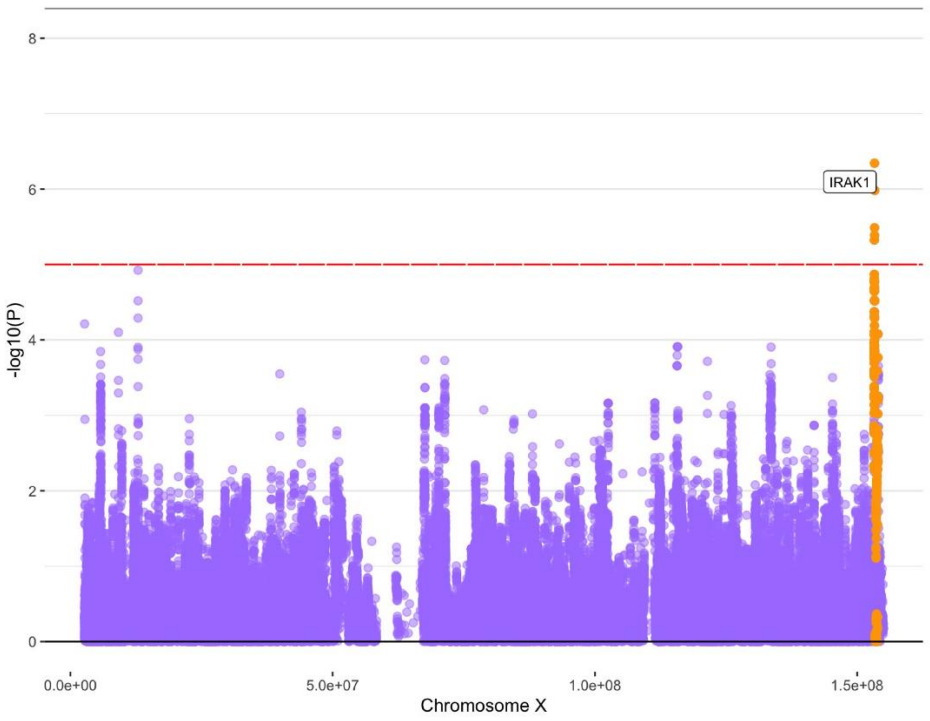


Figure 17 Manhattan plot of meta-analysis of the X chromosome in females from the primary dataset and secondary dataset. The cut-off p-value is 1 x 10^{-5} (red line). According to statistical power analysis. The significant locus in the SLE is labeled in orange.

Meta-analysis of males

Interesting, SNPs on X chromosome specific in males at significant p-value < 1 x 10^{-5} was show 4 loci repeated significantly in *GPR101* region including rs6528443 (p-value = 8.71 x 10^{-6}, OR = 3.55), rs1413644 (p-value = 8.85 x 10^{-6}, OR = 3.54), rs4829611 (p-value = 9. x 10^{-6}, OR = 3.53) and rs5929811 (p-value = 9.76 x 10^{-6}, OR = 3.52) (Table 5). According to literature review, these SNPs has not been identified in other population. The Manhattan plot showing significant SNPs associated with Thai male SLE patients was show in Figure 16.

Table 5 List of SNPs on X chromosome association significant at p < *1* x *10^{-5}* from Meta-analysis of the X chromosome in males.

dbSNP	BP	A1	A2	Locus	Annotation	MAF	OR	P
rs6528443	136353993	T	C	GPR101	none	0.65	3.55	8.71E-06
rs1413644	136355984	G	A	GPR101	none	0.65	3.55	8.85E-06
rs4829611	136355922	T	C	GPR101	none	0.65	3.53	9.60E-06
rs5929811	136356727	T	G	GPR101	none	0.65	3.53	9.76E-06

Meta-analysis

The meta-analysis using primary dataset and secondary dataset all susceptibility SNPs were are reported in Table 6. The loci significantly at p-value 1 x 10-3 was show 214 loci seeing in Manhattan plot (Figure 17).

The highest known SLE susceptibility loci is IRAK1- MECP2- TMEM187 loci, and we found in Thai SLE population including rs1059702 (p-value = 1.82 x 10-7; OR = 0.68). Next, we identified on IRAK1- MECP2- TMEM187 and the risk loci of TLR region express rs3853839 (p-value = 2.03 x 10-4; OR = 0.74) on TLR7and X:9165034 (p-value = 1.14 x 10-5; OR=1.3) on FAM9B (Family With Sequence Similarity 9 Member B) and rs12398129 (p-value = 1.72 x 10-4 ;OR = 1.39) on CXorf61. The results are consistent with the previous reports [12, 13, 54] (Table 7).

Strikingly, meta-analysis of two datasets also identifies several novel susceptible loci significantly associated with Thai SLE patients. First, rs5961374 (p-value = 2.50 x 10-4; OR = 0.61) on NLGN4X (Neuroligin 4 X-Linked) , rs11282724 (p-value = 4.31

x 10-4; OR = 0.79) on EFHC2 (EF-Hand Domain Containing 2), rs138858396 (p-value = 2.42 x 10-4; OR = 3.09) on WAS (Wiskott-Aldrich syndrome), rs75079700 (p-value = 8.21 x 10-4; OR = 0.77) on SNX12 (Sorting Nexin 12) , rs11094246 (p-value = 1.87 x 10-4; OR = 0.60) on NHSL2 (NHS like 2), rs3861732 (p-value = 2.68 x 10-4; OR = 1.33) on TCEAL5 (Transcription Elongation Factor A Like 5), rs6528443 (p-value = 8.71 x 10-6; OR = 3.55) GPR101 (G protein-coupled receptor 101 or GPCR101) and rs7052503 (p-value = 2.46 x 10-4; OR = 0.77) on MIR891A (MicroRNA 891a) (Table 8).

Table 6 List of know SNPs on X chromosome association significant at $p < 1 \times 10^{-3}$ from Meta-analysis of X chromosome in primary dataset and secondary dataset.

dbSNP	**BP**	**A1**	**A2**	**Locus**	**Annotation**	**MAF**	**OR**	**P**
rs1059702	153284192	A	G	IRAK1	missense	0.2	0.68	1.82E-07
X:9165034	9165034	A	AAAAAT	FAM9B	none	0.35	1.3	1.14E-05
rs12398129	115778625	C	A	CXorf61	none	0.14	1.39	1.72E-04
rs3853839	12907658	C	G	TLR7	3'-UTR	0.76	0.74	2.03E-04

Table 7 List of novel SNPs on X chromosome association significant at $p < 1 \times 10^{-3}$ from Meta-analysis of X chromosome in primary dataset and secondary dataset.

dbSNP	**BP**	**A1**	**A2**	**Locus**	**Annotation**	**MAF**	**OR**	**P**
rs6528443	136353993	T	C	GPR101	none	0.65	3.55	8.71E-06
rs11094246	71361418	C	T	NHSL2	none	0.04	0.6	1.87E-04
rs138858396	48518415	G	A	WAS	none	0.07	3.09	2.42E-04
rs7052503	145297247	C	T	MIR891A	none	0.33	0.77	2.46E-04
rs5961374	5751679	C	T	NLGN4X	none	0.61	0.78	2.50E-04
rs3861732	102519802	G	A	TCEAL5	none	0.11	1.33	2.68E-04
rs11282724	44016996	C	CCCGCCA	EFHC2	none	0	0.79	4.31E-04
rs75079700	70271108	C	CA	SNX12	none	0.16	0.77	8.21E-04

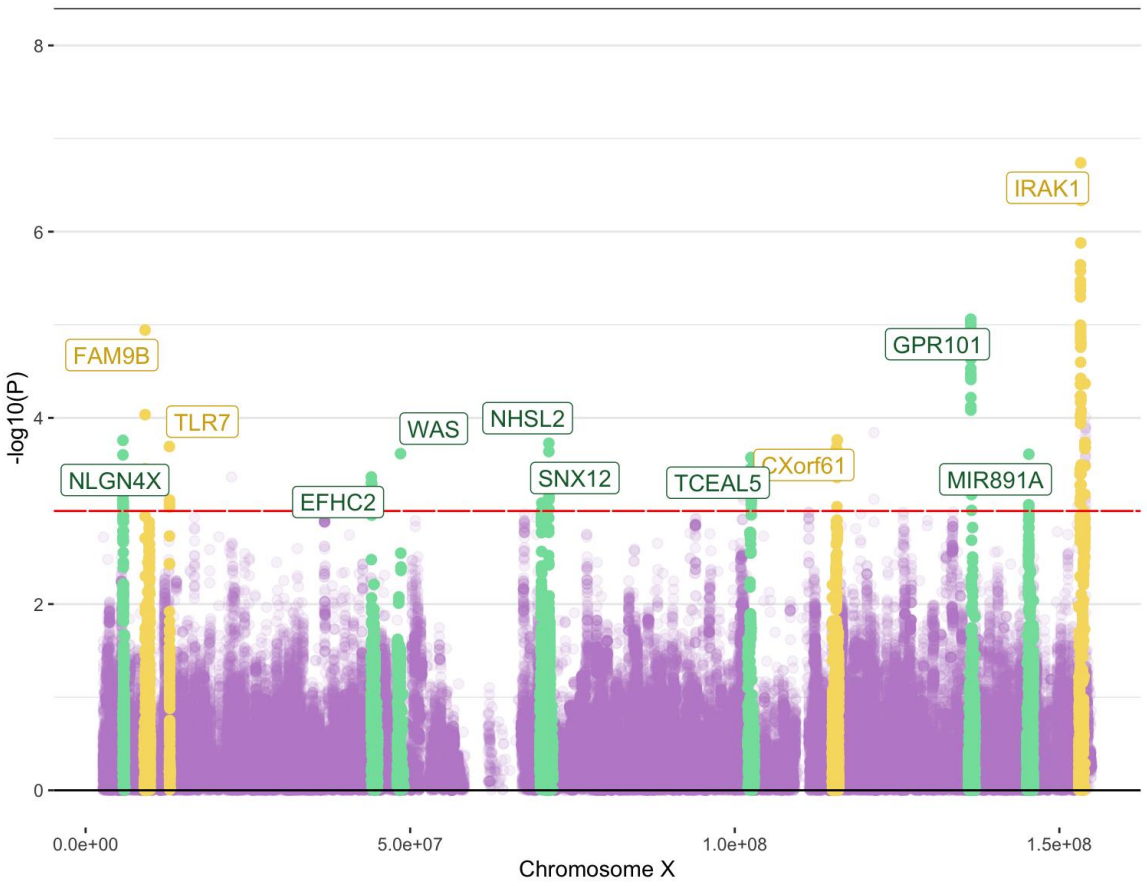
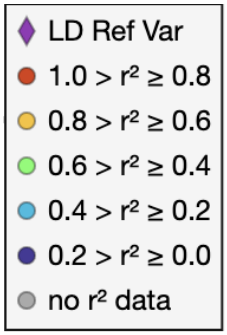
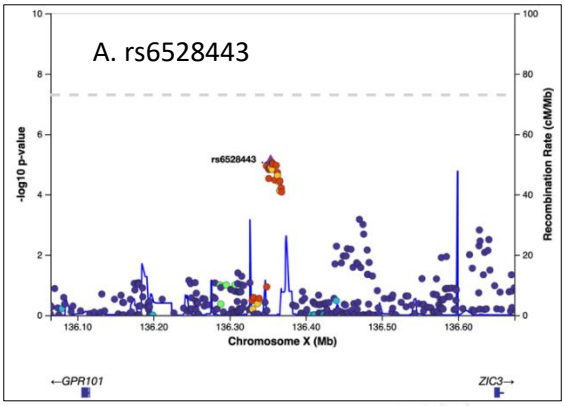
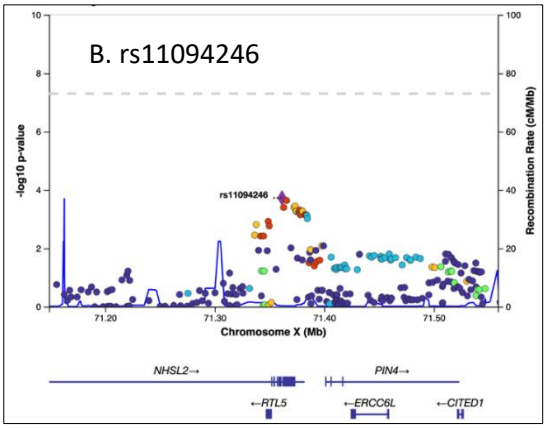


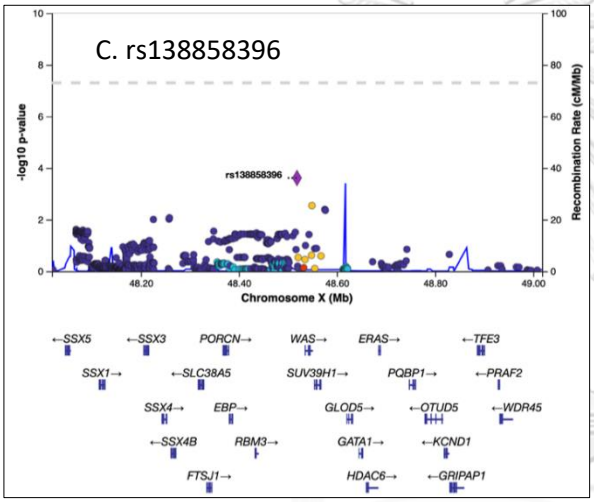
Figure 18 Manhattan plot of meta-analysis of the X chromosome from the primary dataset and secondary dataset. The cut-off p-value is 1 x 10^{-3} (red line). According to statistical power analysis. The novel significant locus in SLE is labeled in green. The known significant locus is labeled in yellow.

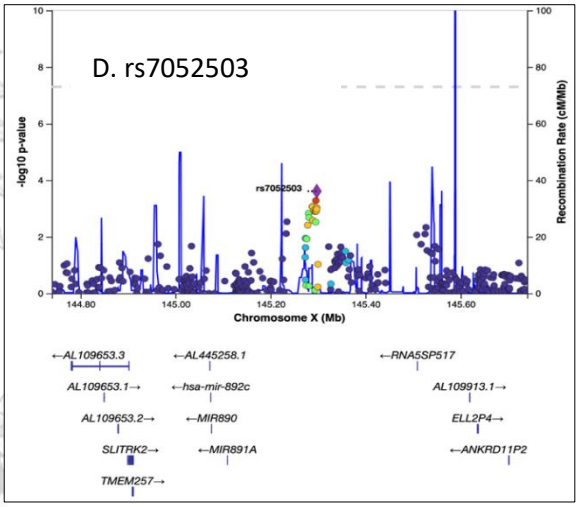
LD (r^2) score of SNPs around novel loci was shown in Locus zoom plot (Figure 18 A-H). SNPs around of novel location express $r^2 > 0.2$ that confirm novel SNPs are heritable with non-random association.

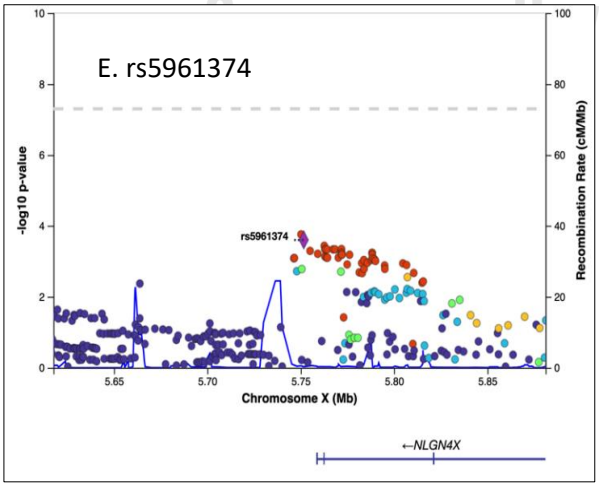


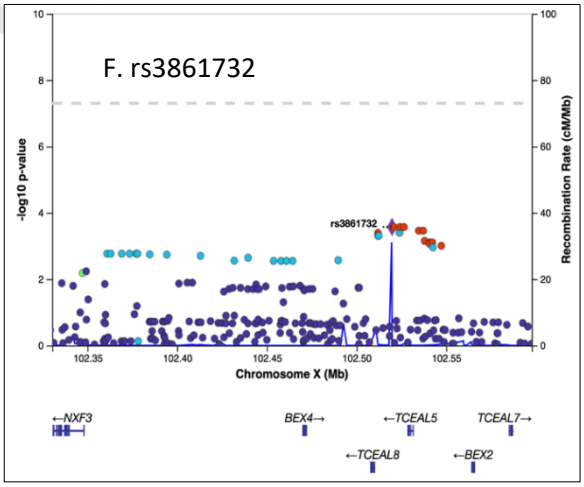












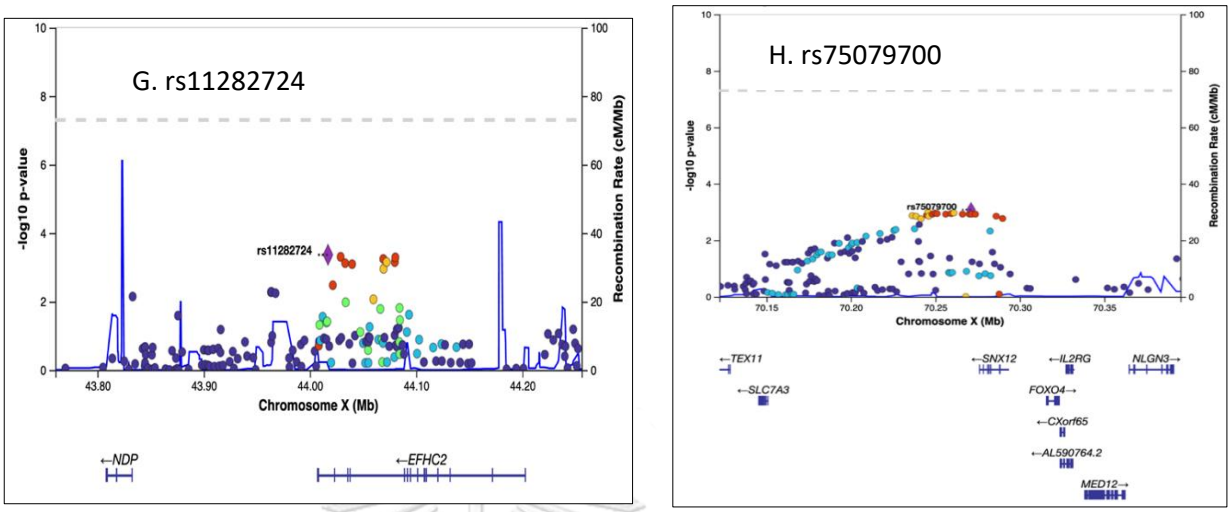


Figure 19 The Locus zoom plots of 8 novel SNPs (A-H) show the location and SNPs surrounding LD (r2).

Next, we tested the gene expression correlation with significant associated SNPs from our data. Interestingly, our novel identified SNPs including rs11094246 on *NHSL2*, rs138858396 on *WAS*, rs5961374 on *NLGN4X*, and rs75079700 on *SNX12*. are significantly correlated with gene expression in specific tissues. This result confirms that SNPs might be associated with the pathogenesis of SLE patients. However, experimental validation is needed to confirm this finding.

Function annotation GTEx result

Table 8 Results of eQTL analyses for novel loci in multiple human tissues.

Gene	**SNP**	***P***	**NES**	**Tissue**
NHSL2	rs11094246	2.40E-48	0.53	Esophagus - Mucosa
NHSL2	rs11094246	6.50E-34	0.36	Skin - Sun Exposed (Lower leg)
NHSL2	rs11094246	1.00E-27	0.36	Skin - Not Sun Exposed (Suprapubic)
NHSL2	rs11094246	1.90E-23	0.23	Cells - Cultured fibroblasts
NHSL2	rs11094246	5.50E-22	0.38	Colon - Transverse
NHSL2	rs11094246	2.50E-19	0.48	Liver

NHSL2	rs11094246	2.20E-18	0.29	Adipose - Visceral (Omentum)
NHSL2	rs11094246	5.50E-18	0.23	Lung
NHSL2	rs11094246	5.00E-17	0.18	Muscle - Skeletal
NHSL2	rs11094246	8.10E-17	0.33	Artery - Aorta
NHSL2	rs11094246	3.70E-16	0.31	Esophagus - Muscularis
NHSL2	rs11094246	5.20E-16	0.23	Thyroid
NHSL2	rs11094246	4.20E-15	0.22	Whole Blood
NHSL2	rs11094246	4.80E-13	0.27	Heart - Atrial Appendage
NHSL2	rs11094246	9.10E-13	-0.2	Whole Blood
NHSL2	rs11094246	2.50E-12	-0.18	Thyroid
NHSL2	rs11094246	5.20E-12	0.32	Stomach
NHSL2	rs11094246	1.80E-11	0.3	Colon - Sigmoid
NHSL2	rs11094246	2.40E-11	-0.19	Nerve - Tibial
NHSL2	rs11094246	9.70E-11	0.18	Artery - Tibial
NHSL2	rs11094246	1.10E-10	0.22	Heart - Left Ventricle
NHSL2	rs11094246	1.10E-09	-0.3	Spleen
NHSL2	rs11094246	1.60E-09	0.31	Spleen
NHSL2	rs11094246	1.50E-08	0.15	Testis
NHSL2	rs11094246	1.90E-08	0.3	Pituitary
NHSL2	rs11094246	2.40E-08	0.15	Adipose - Subcutaneous
NHSL2	rs11094246	5.00E-08	0.41	Cells - EBV-transformed lymphocytes
NHSL2	rs11094246	6.20E-08	0.28	Small Intestine - Terminal Ileum
NHSL2	rs11094246	2.40E-07	0.26	Pancreas
NHSL2	rs11094246	7.90E-07	0.29	Artery - Coronary
NHSL2	rs11094246	0.0000012	0.26	Adrenal Gland
NHSL2	rs11094246	0.0000012	-0.23	Nerve - Tibial

NHSL2	rs11094246	0.0000026	0.18	Breast - Mammary Tissue
NHSL2	rs11094246	0.0000032	-0.19	Thyroid
NHSL2	rs11094246	0.000004	-0.31	Prostate
NHSL2	rs11094246	0.0000049	0.23	Prostate
NHSL2	rs11094246	0.0000063	0.23	Esophagus - Gastroesophageal Junction
NHSL2	rs11094246	0.000013	0.2	Brain - Hippocampus
NHSL2	rs11094246	0.000018	-0.36	Cells - EBV-transformed lymphocytes
NHSL2	rs11094246	0.000033	-0.18	Pituitary
NHSL2	rs11094246	0.000069	-0.058	Esophagus - Muscularis
NHSL2	rs11094246	0.000078	-0.094	Muscle - Skeletal
NHSL2	rs11094246	0.0001	-0.12	Esophagus - Gastroesophageal Junction
NHSL2	rs11094246	0.00011	0.13	Thyroid
NHSL2	rs11094246	0.00016	-0.26	Esophagus - Gastroesophageal Junction
NHSL2	rs11094246	0.00017	0.11	Nerve - Tibial
NHSL2	rs11094246	0.00021	-0.21	Esophagus - Muscularis
NHSL2	rs11094246	0.00024	-0.081	Esophagus - Muscularis
NHSL2	rs11094246	0.00044	-0.16	Adipose - Subcutaneous
WAS	rs138858396	1.30E-21	0.4	Skin - Not Sun Exposed (Suprapubic)
WAS	rs138858396	7.60E-18	0.36	Adipose - Subcutaneous
WAS	rs138858396	1.30E-17	0.38	Lung
WAS	rs138858396	1.10E-16	0.39	Cells - Cultured fibroblasts
WAS	rs138858396	3.90E-16	0.38	Adipose - Visceral (Omentum)
WAS	rs138858396	5.60E-15	0.32	Thyroid
WAS	rs138858396	1.10E-14	0.32	Skin - Sun Exposed (Lower leg)
WAS	rs138858396	1.20E-14	0.39	Breast - Mammary Tissue

WAS	rs138858396	6.60E-14	0.34	Nerve - Tibial
WAS	rs138858396	1.40E-11	0.34	Esophagus - Muscularis
WAS	rs138858396	1.90E-11	0.32	Esophagus - Mucosa
WAS	rs138858396	2.50E-11	0.41	Stomach
WAS	rs138858396	9.00E-11	0.29	Artery - Tibial
WAS	rs138858396	2.00E-10	0.12	Muscle - Skeletal
WAS	rs138858396	2.30E-10	0.25	Muscle - Skeletal
WAS	rs138858396	3.10E-10	0.39	Spleen
WAS	rs138858396	4.40E-10	0.34	Colon - Transverse
WAS	rs138858396	3.90E-09	0.23	Adipose - Subcutaneous
WAS	rs138858396	4.00E-09	-0.34	Brain - Cerebellum
WAS	rs138858396	4.20E-09	0.35	Colon - Sigmoid
WAS	rs138858396	1.00E-08	0.39	Pancreas
WAS	rs138858396	3.50E-08	0.21	Adipose - Subcutaneous
WAS	rs138858396	3.50E-08	0.19	Artery - Tibial
WAS	rs138858396	8.60E-08	0.15	Esophagus - Mucosa
WAS	rs138858396	1.50E-07	0.41	Small Intestine - Terminal Ileum
WAS	rs138858396	2.10E-07	0.19	Skin - Sun Exposed (Lower leg)
WAS	rs138858396	2.30E-07	0.13	Esophagus - Mucosa
WAS	rs138858396	2.60E-07	0.2	Cells - Cultured fibroblasts
WAS	rs138858396	2.70E-07	0.073	Whole Blood
WAS	rs138858396	6.20E-07	-0.25	Brain - Cerebellar Hemisphere
WAS	rs138858396	0.0000012	0.16	Whole Blood
WAS	rs138858396	0.0000014	0.57	Vagina
WAS	rs138858396	0.0000015	0.26	Artery - Aorta
WAS	rs138858396	0.0000016	0.36	Artery - Coronary

WAS	rs138858396	0.0000021	-0.11	Heart - Left Ventricle
WAS	rs138858396	0.0000025	0.28	Esophagus - Gastroesophageal Junction
WAS	rs138858396	0.0000027	-0.12	Heart - Atrial Appendage
WAS	rs138858396	0.0000036	0.093	Artery - Tibial
WAS	rs138858396	0.0000037	0.27	Pituitary
WAS	rs138858396	0.0000064	0.16	Breast - Mammary Tissue
WAS	rs138858396	0.0000072	0.1	Esophagus - Muscularis
WAS	rs138858396	0.0000087	0.19	Breast - Mammary Tissue
WAS	rs138858396	0.000011	0.45	Cells - EBV-transformed lymphocytes
WAS	rs138858396	0.000015	0.15	Artery - Tibial
WAS	rs138858396	0.000016	0.16	Skin - Not Sun Exposed (Suprapubic)
WAS	rs138858396	0.000022	0.23	Heart - Atrial Appendage
WAS	rs138858396	0.000026	-0.15	Brain - Cerebellum
WAS	rs138858396	0.00003	0.16	Skin - Sun Exposed (Lower leg)
WAS	rs138858396	0.000036	0.36	Brain - Cerebellar Hemisphere
WAS	rs138858396	0.000045	0.076	Whole Blood
WAS	rs138858396	0.000052	0.1	Colon - Transverse
WAS	rs138858396	0.000055	-0.18	Brain - Putamen (basal ganglia)
WAS	rs138858396	0.000074	0.16	Nerve - Tibial
WAS	rs138858396	0.000079	0.18	Breast - Mammary Tissue
WAS	rs138858396	0.00017	0.17	Testis
WAS	rs138858396	0.00018	0.093	Lung
WAS	rs138858396	0.00019	0.048	Testis
WAS	rs138858396	0.00021	-0.17	Skin - Sun Exposed (Lower leg)
WAS	rs138858396	0.0003	0.14	Skin - Not Sun Exposed (Suprapubic)
WAS	rs138858396	0.00038	0.1	Nerve - Tibial

NLGN4X	rs5961374	0.0000075	-0.14	Testis
NLGN4X	rs5961374	0.000014	-0.15	Heart - Atrial Appendage
NLGN4X	rs5961374	0.000019	-0.18	Heart - Left Ventricle
TCEAL5	rs3861732	0.00002	0.3	Nerve - Tibial
TCEAL5	rs3861732	0.000081	-0.32	Esophagus - Gastroesophageal Junction
TCEAL5	rs3861732	0.00036	-0.15	Artery - Tibial
TCEAL5	rs3861732	0.00056	-0.15	Artery - Tibial
SNX12	rs75079700	3.20E-13	-0.32	Skin - Sun Exposed (Lower leg)
SNX12	rs75079700	1.00E-12	0.21	Testis
SNX12	rs75079700	5.60E-08	0.09	Lung
SNX12	rs75079700	1.10E-07	-0.25	Brain - Nucleus accumbent (basal ganglia)
SNX12	rs75079700	2.80E-07	-0.23	Skin - Not Sun Exposed (Suprapubic)
SNX12	rs75079700	0.0000057	-0.29	Brain - Hippocampus
SNX12	rs75079700	0.000016	0.077	Thyroid
SNX12	rs75079700	0.000022	-0.23	Brain - Putamen (basal ganglia)
SNX12	rs75079700	0.000055	-0.28	Brain - Cerebellar Hemisphere
SNX12	rs75079700	0.0001	-0.19	Esophagus - Mucosa
SNX12	rs75079700	0.00012	-0.18	Whole Blood
SNX12	rs75079700	0.00018	-0.18	Adipose - Subcutaneous
SNX12	rs75079700	0.00024	-0.075	Nerve - Tibial

NES: normalized effect size. The threshold of *P* values is 0.05/47≈0.001.

Discussion

Our study focuses on the SNPs on X-chromosomes in Thai SLE patients. The significant p-value is the first critical to determining susceptibility loci. The p-value can classify true loci affected to trials by random chance and testing alternative hypotheses [56]. After which the OR value can represent risk allele who carry will have a chance to disease (OR > 1) and the protective allele to less a chance to developing to disease (OR < 1) [57].

According to our results, we identified several known SLE susceptible loci in Thai SLE patients such as *IRAK1-MECP2-TMEM187* region, *TLR-7*, *TLR-8* [58, 59]. These showed the reliability of our analysis processing. The *IRAK1* is a signaling protein affecting the innate and adaptive immune system, especially the interleukin-1 receptor, transcription factor NF-κB [60]. While *MECP2* (rs1734791, p = 4.63 x 10-7, OR = 0.68) are key role for supporting suppression genes on X chromosome [7], the decrease of *MECP2* can stimulate risk to lupus [60]. The *IRAK1-MECP2-TMEM187* region is neighbor genes located on Xq28. The SNPs rs1059702 on *IRAK1* is the highest significance of SLE in the East Asian population, and the SNPs on *ARHGAP4, NAA10, RENBP, HCFC1, TMEM187, IRAK1*, and *MECP2* showed strong LD (r2 > 0.2) inheritable with rs1059702 [60]. Consistent with the previous study, the SNPs on *ARHGAP4, NAA10, RENBP, HCFC1, TMEM187, IRAK1*, and MECP2 were associated in Thai SLE patients.

Another important significant locus is rs3853839 C>G on TLR7. These SNPs have been found to increases *TLR7* expression and IFN (interferon type I) release. The IFN I can induce a new-form transition (TR) B-cell. This type of B-cells produces auto-antibody to self-antigen, progressing to SLE pathogenesis [61].

In additional SNPs on *FAM9B* were report in the Asian SLE such as rs1876415 (*p*-value = 4.6 x 10^{-4}, OR = 1.18) and SNPs nearly *FAM9B* (rs5934505, p *p*-value = 5.6×10^{-16}) have been reported an affected to decrease testosterone level in males [62, 63]. Low testosterone concentration increases the risk of autoimmune disease in males. Testosterone's role in the immune system was identified that could suppression BAFF (B-cell activating factor) cytokine or TNFSF13B (Tumor necrosis factors Superfamily Member 13b) [64]. The BAFF can increase splenic B cell survival and differentia [65]. Therefore, overexpress of BAFF can induce massive B cells and lead

to autoantibody-producing [66]. This finding related to our result shows loci at X:9165034 on *FAM9B* in SLE patients. The *CXorf61* is highly expressed in tumor or cancer cells but it is still unknown in SLE [67, 68]. This mechanism can explain why men have a protective from the autoimmune condition more than women.

Novel discovered SLE susceptible alleles are also identified on *GPR101* at Xq26.3. These alleles are highly repeated loci and are shown specifically associated with the male X chromosome in Thai SLE patients. Previously, GWAS study reported a novel SNP, rs13440883 (p-value = 7.53 × 10− 9, OR = 1.16) within *GPR173* (G protein-coupled receptor 173) and upstream of *GPR19* are susceptibility to SLE [13, 69]. Correlated with our finding, the *GPR101* encodes G protein-coupled receptor 101 and receptors of leukocytes, including neutrophils, monocytes, and macrophages. A recent study covered *GPR101* as an immunoregulator, when combined with N-3 docosapentaenoic acid-derived resolvin D5 (RvD5n-3 DPA). Their ability can inhibit leukocyte transmigration to the inflammatory site, promote macrophage eliminate cell death and antigen. Lack of *Gpr101* led to neutrophil migration to increase inflammation in arthritis mice [70]. Leukocytes play an essential role in SLE. They were stimulating tissue inflammation by releasing proinflammatory cytokines such as Type I interferons (IFNs), tumor necrosis factor-α (TNF-α), interleukin-6 (IL-6), and IL-1β there is influence on the severity in SLE [71, 72]. Leukocyte apoptosis is reported high level in SLE and without adequate clearance [73]. This abnormality inducible nuclear autoantigen exposure led to auto-antibody production, which is associated with severity level in SLE [74]. The RvD5n-3 DPA amino acid interaction with the *GPR101* receptor can promote macrophage clearance apoptotic cell; this mechanism might be elevated auto-antibody production [70, 75]. Moreover, the *GPR101* co-action with RvD5n-3 DPA inhibits the invasion of neutrophils to the inflammatory site, which reduces inflammation levels [70]. These covered functions of *GPR101* are the essential key to understanding the SLE in males. The *GPR101* is immunoregulation and specific in male patients. We might assume the *GPR101* is crucial for males to developing SLE. However, there is no reported mutation of *GPR101* in SLE patients, and our SNPs rs6528443 closets *GPR101* not provided in GTEx. Despite this, these findings are necessary for future investigations. According

to the above, these results represent informative data that could be useful for further Genomics Thailand Project.

The novel susceptible loci identified in our study express *NHSL2* or NHS like 2 is an association with Nance-Horan syndrome cause by Cataract 40, X-linked disorder. NHS protein is necessary for cell morphology and inhibits actin cytoskeletal that can cause cell mobilization and adhesion disorder. Moreover, loss of NHS protein decreased the WAVE complex in immune cells led to disturbing immune synapse (IS) formation due to immune dysfunction [76]. Consistent with Riccardo Papa, et al.,2021 reported actin formation problem is the cause of immune dysfunction and autoinflammation.

The *WAS* genes had been reported to affect T-cell structure impairment and promote T-cell migration. Our study identified two SNPs associated with actin formation, including rs11094246 at *NHSL2* and rs138858396 at *WAS*. These findings may have influenced SLE development.

NLGN4X is a cause of autism spectrum disorder (ASD) and was identified escape from XCI [9, 77, 78]. *TCEAL7* and *TCEAL6* are high expressions in ovarian cancer cells motivated by genes escape from X chromosome inactivation (XCI) [79]. This event may include association with SLE. These genes need to identify in the immune system in the future. In black SLE patients are rich in anti-nuclear antibodies, anti-Sm, and anti-RNP antibodies [80]. Those complex with proteins (70 Kd, A, C) translation from the U1 gene coincided with our study, which found SNPs rs77418624 at U1 or 3' SNX12 locus in the Thai SLE population[81].

Interestingly, the miR-891A which has been reported in exosome extracted from nasopharyngeal carcinoma [82]. It inhibits Th1 and Th17 cell differentiation while promoting Treg cell differentiation by decreasing the activity of extracellular signal-regulated kinases (ERK), signal transducer, and activator of transcription (STAT) 1 and STAT3 and increasing the activity of STAT5 in exosome-treated T cells [83]. The polymorphisms on miR-891A might be associated with low-level T-reg cells found among SLE patients [84].

Furthermore, XCI has determined significant with immune action in females according to introduction part, the SNPs on X chromosome reported escaped from XCI are significant with SLE. Previous studies determined SNPs escaped XCI such as

rs887369 on *CXorf21 [35]* Epigenetic features control the XCI, including the transcription of the long noncoding RNAs (lncRNA), X-inactive specific transcript (XIST) encoded by an X-linked gene [85, 86]. The XIST RNA spreads covered the X chromosome and recruiting the polycomb repressive complex 1 (PRC1) and 2 (PRC2) to action for monoubiquitylation of lysin 119 on the histone H2A (H2AK119ub1) and trimethylation of the lysine 27 on histone H3 (H3K27me3) on the Xi [85, 86]. The extinction of histone acetylation, H3K27ac, is the beginning of the XCI event [85]. To confirm the XCI is a factor of SLE we can compare the XCI status of healthy and SLE by epigenetic features or DNA methylation analysis. The X active (Xa) and X inactive (Xi) can be detected by CpG islands methylation to observe increase methylation on the Xi and to mapping, location genes escape from XCI [87]. Next, determine SNPs escape XCI by probes [78]. This experiment is needed to be future performed.

Sex bias immune responses have been researched for several years. The sex hormone is the earliest evidence cause of this bias [88, 89]. Oestrogen is the main character autoimmune induce in females, and testosterone is a protective immune overaction in males. These two sex hormones are reported as factors of bias sex in autoimmune disease [90]. Currently, the sex hormone has competition; the X chromosome becomes a potential role immune function [91]. There is an independent contribution immune response in autoimmune disease [91]. However, the X chromosome factor has a piece of support evidence which is Klinefelter syndrome. The patients who carry extra X chromosomes XX, Y risk developing lupus and Sjogren's syndrome equal to females [92]. This evidence could increase the credibility of X chromosome induce autoimmune disease.

Although these results are promising, replications in large cohorts are still required. In addition, it is also a challenge to link the associated SNPs to target genes. More detailed studies are needed to perform to better understand the underlying genetic regulation in those disease-associated loci.

CHAPTER V
CONCLUSION

Conclusions

In the present, there are many reports of SLE susceptibility loci that can increase understanding about SLE and other autoimmune diseases. Nevertheless, the genetic variation in ethnicity influences the incident rate and mortality rate differently from each region. Our study explores specific 12 loci in the Thai SLE population, distinct and duplicate in the Asian SLE population reported. The risk loci are known as *IRAK1-MECP2-TMEM187* region also susceptibility in Thai SLE patients. Moreover, we identified novel susceptibility genes in Thai SLE patients confirm from the GTEx project, including *NLGN4X*, *WAS*, *SNX12*, and *NHSL2*

Besides, other loci are control cells - EBV-transformed lymphocytes that involved B-cell proliferation and T-cell activity, including rs138858396 and rs11094246, which is specific in Thai SLE.

We also identified novel SNPs rs6528443, located on *GPR101* gene are susceptible to males SLE. *GPR101* was recently found significant with leukocytes regulator. We expect this finding will help improve SLE therapy in Thai populations.

Future Perspective

In the future, this finding should be replicated in the Genomic Thailand database to confirms the SNPs specific in the Thai population and compare GWAS results with other populations to identify specific SNPs in populations that could be explained the difference severity pathogenesis. In conjunction with SNPs, function tests such as Expression quantitative trait loci (eQTLs) determine candidate variants that act on genes and explain how is affect traits [93, 94].

Nonetheless, there is another enjoyable method to confirm the effect of rs6528443 loci nearby GPR101, which is CRISPR/cas9 gene-editing method [93]. To test SNPs, affect endogenous gene expression, we can knock in SNPs to cell lines and delete target SNPs by Cas9-sgRNAs. Then compare gene expression in a cell line with and without deletion SNPs target [95].

Since SLE has varied phenotypes and severity levels, we can determine the SNPs associated with different phenotypes by P-value and OR analysis. We might classify sub phenotype of SLE patients according to ACR 2019 criteria into six classes of SLE, including constitutional, hematologic, neuropsychiatric, serosal, musculoskeletal, and renal [96]. Next, select SNPs with a significant p-value in each sub phenotype and determine risk allele with OR value less than one and protective allele with OR more than 1 to predict and prevent severity in SLE patients.

In addition, future investigations should increase more sample sizes. This is allowing the discovery of more significant numbers of SNPs and empowers the study.

APPENDIX

Table 9 List of SNPs association significant at $p < 1 \times 10^{-3}$ from Meta-analysis of the X chromosome from primary data and secondary data.

dbSNP	Position	Ref/Alt	Locus	Annotation	MAF	OR	P	Q	I^2
rs1482816	5746390	G/A	NLGN4X	none	0.61	0.79	0.0008216	0.758	0
rs34500934	5746650	T/TA	NLGN4X	none	0.61	0.79	0.0008068	0.809	0
rs6639513	5750225	C/G	CXorf61	none	0.14	0.77	0.0001743	0.807	0
rs5961374	5751679	C/T	NLGN4X	none	0.61	0.78	0.0002499	0.748	0
rs62583770	5755065	A/G	NLGN4X	none	0.61	0.79	0.0005083	0.687	0
rs77345530	5759179	G/GC	NLGN4X	none	0	0.79	0.000614	0.689	0
rs6639515	5761897	T/C	NLGN4X	none	0.61	0.80	0.0008424	0.723	0
rs2128516	5762654	G/A	NLGN4X	none	0.61	0.79	0.0003652	0.721	0
rs66523284	5762967	TAA/T	NLGN4X	none	0	0.79	0.000729	0.687	0
rs6638568	5763035	C/T	NLGN4X	none	0.61	0.79	0.0004611	0.756	0
rs6638569	5763146	A/T	NLGN4X	none	0.61	0.79	0.0004611	0.756	0
rs6639516	5763908	G/A	NLGN4X	none	0.6	0.80	0.0007879	0.717	0
rs10521577	5764199	G/A	NLGN4X	none	0.6	0.79	0.0004611	0.756	0
rs35807547	5764379	T/C	NLGN4X	none	0.61	0.80	0.0007879	0.717	0
rs1384520	5767410	G/T	NLGN4X	none	0.61	0.79	0.0004611	0.756	0
rs6639517	5768475	G/A	NLGN4X	none	0.61	0.79	0.0004611	0.756	0
rs1482814	5770197	T/C	NLGN4X	none	0.61	0.80	0.0007879	0.717	0

rs6638570	5771479	G/A	NLGN4X	none	0.61	0.79	0.0005468	0.736	0
rs6639518	5771535	C/T	NLGN4X	none	0.61	0.79	0.0005984	0.788	0
rs5961375	5771896	C/T	NLGN4X	none	0.61	0.79	0.0004511	0.765	0
rs5961874	5772030	A/C	NLGN4X	none	0.61	0.79	0.0005754	0.764	0
rs544805987	5772108	A/AG	NLGN4X	none	0.3379	0.79	0.0006458	0.778	0
X	5772110	C/TTATTTA	NLGN4X	none	0	0.79	0.0006458	0.778	0
rs1967021	5776196	T/C	NLGN4X	none	0.61	0.79	0.0006529	0.765	0
rs1384519	5783856	A/T	NLGN4X	none	0.62	0.80	0.0009174	0.742	0
rs6639526	5787915	C/T	NLGN4X	none	0.62	0.80	0.0009169	0.668	0
rs6638573	5788258	C/A	NLGN4X	none	0.61	0.79	0.0005798	0.719	0
rs6639527	5788335	A/G	NLGN4X	none	0.61	0.79	0.0005798	0.719	0
rs6639528	5788371	C/T	NLGN4X	none	0.61	0.79	0.0006566	0.779	0
rs5961875	5788576	T/C	NLGN4X	none	0.62	0.80	0.000831	0.655	0
rs11094840	5789556	T/C	NLGN4X	none	0.62	0.80	0.0008766	0.663	0
rs6638574	5790104	G/T	NLGN4X	none	0.62	0.80	0.0009755	0.619	0
rs6638901	9164601	G/A	FAM9B	none	0.75	1.25	0.000399	0.463	0
rs539040862	9165034	A/AAAAAT	none	none	0.353	1.33	1.14E-05	0.494	0
rs4830405	9165158	C/T	FAM9B	none	0.23	1.28	9.24E-05	0.502	0
rs5979044	9165588	A/G	FAM9B	none	0.23	1.25	0.0003522	0.536	0
rs5933713	9165712	G/A	FAM9B	none	0.21	1.30	5.23E-05	0.414	0
rs3853839	12907658	C/G	TLR7	none	0.76	0.75	0.0002031	0.933	0
rs5935442	12923109	C/T	TLR8-AS1	none	0.79	0.78	0.000909	0.622	0
rs5935443	12923197	G/T	TLR8-AS1	none	0.76	0.78	0.0008967	0.610	0
rs3764879	12924697	C/G	TLR8-AS1	none	0.79	0.77	0.0007644	0.644	0
rs3764880	12924826	A/G	TLR8	none	0.79	0.77	0.0008428	0.595	0

rs371216363	22492192	TA/T	none	none	0	3.64	0.0001114	0.391	0
rs35069992	22494052	AT/ATT	ZNF645	none	0	2.94	0.0004309	0.638	0
rs5963154	39918362	A/C	BCOR	none	0.71	0.79	0.000612	0.410	0
rs11282724	44016996	C/CCCGCCA	EFHC2	none	0	0.80	0.0004312	0.877	0
rs6417888	44028922	A/G	EFHC2	none	0.46	0.80	0.0004937	0.883	0
rs2050399	44033173	T/C	EFHC2	none	0.46	0.81	0.0007564	0.858	0
rs6609283	44039681	T/C	EFHC2	none	0.46	0.81	0.0008071	0.852	0
rs5952559	44069078	A/G	EFHC2	none	0.46	0.80	0.0005709	0.865	0
rs61419118	44072240	T/C	EFHC2	none	0.46	0.81	0.0007007	0.871	0
rs1335101	44079894	C/A	EFHC2	none	0.46	0.81	0.0007082	0.843	0
rs4824814	44080389	G/C	EFHC2	none	0.46	0.80	0.0005126	0.857	0
rs138858396	48518415	G/A	WAS	none	0.07	3.10	0.0002422	0.441	0
rs77418624	70245729	A/G	U1	none	0.16	0.78	0.0009871	0.854	0
rs75079700	70271108	C/CA	SNX12	none	0.16	0.78	0.000821	0.833	0
rs11094246	71361418	C/T	NHSL2	none	0.04	0.60	0.0001871	0.345	0
rs7472405	71362904	G/T	NHSL2	none	0.04	0.62	0.0003965	0.322	0
rs7471188	71365595	T/C	NHSL2	none	0.04	0.62	0.0002299	0.705	0
rs138487857	71372268	T/TTTAGG	BX119917.1	none	0	0.64	0.0003941	0.492	0
rs6525581	71373407	G/A	NHSL2	none	0.04	0.64	0.0003554	0.469	0
rs34444248	71374155	TG/T	NHSL2	none	0.04	0.66	0.0005319	0.548	0
rs7877671	71374461	T/C	NHSL2	none	0.04	0.66	0.0005319	0.548	0
rs67861709	71376567	C/T	NHSL2	none	0.04	0.66	0.000593	0.535	0
rs7472697	71377253	A/G	NHSL2	none	0.04	0.65	0.0005494	0.628	0
rs7886775	71377489	G/T	NHSL2	none	0.04	0.65	0.0004878	0.576	0
rs7062862	71378104	C/T	NHSL2	none	0.04	0.66	0.0007081	0.593	0

rs7884806	71379702	/	NHSL2	none	0.04	0.66	0.0005319	0.548	0
rs7880917	71379853	C/T	NHSL2	none	0.04	0.66	0.0005319	0.548	0
rs7881332	71380152	C/T	NHSL2	none	0.04	0.66	0.0005319	0.548	0
rs78869251	71382485	C/T	NHSL2	none	0.04	0.66	0.0007081	0.593	0
rs35776454	71383577	T/TG	NHSL2	none	0.04	0.65	0.0007112	0.641	0
rs7884010	71384147	A/G	NHSL2	none	0.95	0.67	0.0007545	0.607	0
rs112461116	71384522	AAG/A	FLJ44635	none	0	0.66	0.0009581	0.657	0
rs6621212	101076547	C/T	NXF5	none	0.03	1.87	0.0006937	0.516	0
rs182367921	101322284	G/A	TCEAL2	none	0.04	1.79	0.0007568	0.463	0
rs6616350	101370038	G/T	TCEAL2	none	0.04	1.82	0.0006441	0.471	0
rs6621358	101399659	G/A	TCEAL6	none	0.04	1.86	0.0004028	0.413	0
rs6621359	101401138	T/C	TCEAL6	none	0.04	1.86	0.0004028	0.413	0
rs184359519	101402214	CTT/C	TCEAL6	none	0	1.85	0.0004118	0.415	0
rs6621364	101414340	T/C	BEX5	none	0.03	1.86	0.0003939	0.419	0
rs6621368	101434691	G/A	BEX5	none	0.04	1.86	0.0003884	0.430	0
rs145342903	101445481	G/A	NXF2B	none	0.03	1.86	0.0003884	0.430	0
rs201958906	101541243	C/T	NXF2B	none	0	1.82	0.000626	0.435	0
X	101572655	T/C	NXF2	none	0	1.85	0.0004116	0.419	0
rs5987713	102511915	G/T	TCEAL5	none	0.11	1.32	0.0003983	0.425	0
rs6621636	102512040	T/C	TCEAL8	none	0.11	1.32	0.0004959	0.412	0
rs6621637	102512398	A/C	TCEAL8	none	0.11	1.32	0.0004948	0.411	0
rs3861732	102519802	G/A	TCEAL5	none	0.11	1.33	0.0002678	0.427	0
rs7886956	102520782	G/T	TCEAL5	none	0.11	1.33	0.0002678	0.427	0
rs6621638	102523960	T/G	TCEAL5	none	0.11	1.32	0.0003921	0.398	0
rs6616454	102524043	G/A	TCEAL5	none	0.11	1.32	0.0003921	0.398	0

rs5987715	102524207	C/T	TCEAL5	none	0.11	1.33	0.0002678	0.427	0
rs6621639	102526480	G/C	TCEAL5	none	0.11	1.33	0.0002678	0.427	0
rs6621642	102534681	T/C	TCEAL5	none	0.11	1.32	0.0003471	0.422	0
rs6621643	102537217	C/T	TCEAL5	none	0.11	1.32	0.0003471	0.422	0
rs12556222	102537946	C/T	TCEAL5	none	0.11	1.30	0.0007	0.437	0
rs374500312	102540678	T/C	none	none	0	1.30	0.0007818	0.444	0
rs201589816	102540682	T/TAAA	TCEAL5	none	0	1.30	0.0008615	0.437	0
rs150873941	102542290	C/T	TCEAL5	none	0.11	1.30	0.0007818	0.444	0
rs5987719	102547309	C/T	TCEAL5	none	0.11	1.29	0.0009867	0.411	0
rs139840812	115661589	C/G	CXorf61	none	0.81	1.38	0.0002095	0.569	0
rs5905339	115663664	G/A	CXorf61	none	0.85	1.36	0.0003627	0.562	0
rs6608697	115668252	A/C	CXorf61	none	0.85	1.36	0.0003627	0.562	0
rs1015041	115670998	C/T	CXorf61	none	0.85	1.36	0.0003627	0.562	0
rs12399468	115681326	G/A	CXorf61	none	0.15	1.38	0.0002705	0.559	0
rs12384658	115740289	T/G	CXorf61	none	0.75	1.29	0.0004364	0.473	0
rs9724300	115740847	G/C	CXorf61	none	0.75	1.27	0.0009	0.432	0
rs12398129	115778625	C/A	CXorf61	none	0.14	1.39	0.0001723	0.392	0
rs11260309	115780226	C/T	CXorf61	none	0.14	1.39	0.0001741	0.386	1.3689
rs6648751	121447671	T/C	GRIA3	none	0.11	1.45	0.0001435	0.460	0
rs1383661	121455087	C/G	GRIA3	none	0.11	1.37	0.0007469	0.436	0
rs1413645	136348411	T/A	GPR101	none	0.65	3.50	1.13E-05	0.566	0
rs5975891	136351546	G/C	GPR101	none	0.65	3.33	2.95E-05	0.479	0
rs5975892	136351600	G/T	GPR101	none	0.65	3.44	1.47E-05	0.566	0
rs5974666	136351855	G/C	GPR101	none	0.65	3.44	1.47E-05	0.566	0
rs1334504	136353300	C/T	GPR101	none	0.65	3.46	1.35E-05	0.561	0

rs6528443	136353993	T/C	GPR101	none	0.65	3.55	8.71E-06	0.590	0
rs6528444	136354195	A/G	GPR101	none	0.65	3.47	1.29E-05	0.551	0
rs4829610	136355613	C/T	GPR101	none	0.65	3.44	1.49E-05	0.564	0
rs4829611	136355922	T/C	GPR101	none	0.65	3.53	9.60E-06	0.541	0
rs1413644	136355984	G/A	GPR101	none	0.65	3.55	8.85E-06	0.614	0
rs5929811	136356727	T/G	GPR101	none	0.65	3.53	9.76E-06	0.599	0
rs1334503	136359531	C/G	GPR101	none	0.65	3.32	3.33E-05	0.823	0
rs4829612	136361353	G/A	GPR101	none	0.65	3.51	1.08E-05	0.597	0
rs1334502	136362917	A/G	GPR101	none	0.65	3.38	1.86E-05	0.762	0
rs1334501	136363022	T/C	GPR101	none	0.65	3.33	2.34E-05	0.725	0
rs5929812	136365198	G/T	GPR101	none	0.66	3.29	3.87E-05	0.765	0
rs6635415	136365743	G/A	GPR101	none	0.66	3.18	7.55E-05	0.869	0
rs6528445	136365978	T/C	GPR101	none	0.65	3.35	3.56E-05	0.896	0
rs5931114	136367940	A/G	GPR101	none	0.66	3.23	6.06E-05	0.851	0
rs5975897	136368384	A/G	GPR101	none	0.66	3.21	8.25E-05	0.773	0
rs144196112	136470484	G/A	ZIC3	none	0.17	1.29	0.000667	0.696	0
rs5975924	136475200	C/A	ZIC3	none	0.17	1.28	0.0009843	0.829	0
rs12388481	145287904	A/T	MIR891A	none	0.3	0.79	0.0008615	0.693	0
rs72608998	145288065	G/T	MIR891A	none	0.3	0.79	0.0008615	0.693	0
rs5965701	145295410	A/G	MIR891A	none	0.34	0.79	0.0005232	0.745	0
rs10562670	145297000	TAC/T	MIR891A	none	0.31	0.79	0.0009213	0.725	0
rs7052503	145297247	C/T	MIR891A	none	0.33	0.78	0.0002456	0.739	0
rs2285037	152816206	C/T	ATP2B3	none	0.24	2.65	0.0008811	0.973	0
rs5987186	153194459	A/G	ARHGAP4	none	0.69	0.77	0.0005277	0.610	0
rs2071128	153195393	G/A	NAA10	none	0.67	0.77	0.0004045	0.395	0

rs2071129	153195921	T/G	NAA10	none	0.68	0.79	0.0008206	0.459	0
rs2071131	153196345	G/A	NAA10	none	0.67	0.77	0.0004064	0.395	0
rs2269370	153196429	C/A	NAA10	none	0.68	0.77	0.0003547	0.453	0
rs17422	153227426	A/G	HCFC1	none	0.77	0.76	0.0007227	0.548	0
rs2266890	153247722	C/T	TMEM187	missense	0.73	0.70	3.37E-06	0.565	0
rs7350355	153247745	A/G	TMEM187	missense	0.73	0.71	5.02E-06	0.611	0
rs6571303	153247954	C/T	TMEM187	synonymous	0.73	0.71	3.96E-06	0.583	0
rs13397	153248248	G/A	TMEM187	synonymous	0.69	0.71	2.27E-06	0.416	0
rs5945173	153250172	G/A	TMEM187	none	0.72	0.71	3.75E-06	0.543	0
rs6643808	153252147	T/C	TMEM187	none	0.74	0.71	1.16E-05	0.420	0
rs6643809	153252908	T/C	TMEM187	none	0.74	0.71	1.01E-05	0.423	0
rs6643656	153254605	C/G	TMEM187	none	0.82	0.68	2.64E-06	0.659	0
rs6655269	153256435	G/A	TMEM187	none	0.71	0.71	4.26E-06	0.442	0
rs5986947	153256505	G/C	TMEM187	none	0.71	0.73	4.38E-05	0.629	0
rs12353692	153260032	G/T	TMEM187	none	0.23	0.73	9.46E-05	0.544	0
rs35059571	153264624	A/AT	IRAK1	none	0	0.73	9.34E-05	0.906	0
rs11795678	153265728	G/A	IRAK1	none	0.2	0.73	9.94E-05	0.883	0
rs5986948	153266172	T/C	IRAK1	none	0.2	0.72	1.52E-05	0.683	0
rs5945386	153269755	G/T	IRAK1	none	0.18	0.73	0.0001151	0.843	0
rs4898375	153273226	A/G	IRAK1	none	0.2	0.71	1.74E-05	0.879	0
rs633	153274228	C/T	IRAK1	none	0.18	0.72	4.36E-05	0.925	0
rs12400188	153275075	G/A	GPR101	none	0.66	0.72	5.88E-05	0.934	0
rs3027898	153275890	C/A	IRAK1	none	0.18	0.72	5.88E-05	0.934	0
rs731642	153277507	A/G	IRAK1	none	0.19	0.72	2.53E-05	0.663	0
rs2239673	153277889	C/T	IRAK1	none	0.18	0.73	6.84E-05	0.937	0

rs763737	153278307	G/A	IRAK1	none	0.18	0.72	5.81E-05	0.912	0
rs1059703	153278829	G/A	IRAK1	none	0.2	0.71	1.41E-05	0.926	0
rs146868205	153279822	TAAAA/T	IRAK1	none	0	0.70	1.25E-05	0.947	0
rs5945174	153279858	G/A	IRAK1	none	0.18	0.70	1.36E-05	0.963	0
rs7061789	153280475	G/A	IRAK1	none	0.18	0.71	3.77E-05	0.972	0
rs1059702	153284192	A/G	IRAK1	missense	0.2	0.68	1.82E-07	0.580	0
rs1059701	153284483	G/A	IRAK1	none	0.18	0.72	1.07E-05	0.652	0
rs2734647	153292180	T/C	MECP2	3'-UTR	0.2	0.70	1.32E-06	0.559	0
rs1624766	153317154	C/T	MECP2	none	0.18	0.76	0.0006222	0.920	0
rs1734787	153325446	A/C	MECP2	none	0.81	0.73	6.48E-05	0.910	0
rs1734791	153330920	A/T	MECP2	intronic	0.81	0.69	4.63E-07	0.669	0
rs4898376	153343006	C/T	MECP2	none	0.18	0.76	0.0005202	0.917	0
rs3831674	153348218	CT/C	MECP2	none	0	0.74	8.36E-05	0.641	0
rs2239464	153348431	A/G	MECP2	none	0.18	0.72	1.61E-05	0.734	0
rs5945393	153349428	G/A	MECP2	none	0.18	0.76	0.0003976	0.592	0
rs12841797	153370114	T/G	MECP2	none	0	0.76	0.0003194	0.563	0
rs5945233	153939325	T/A	GAB3	none	0.19	0.78	0.0002153	0.563	0
rs2664169	153939663	T/C	GAB3	none	0.81	0.79	0.0006657	0.479	0
rs2664170	153945602	G/A	GAB3	none	0.79	0.78	0.000328	0.524	0
rs5987015	153947981	/	GAB3	none	0.78	0.78	0.0001852	0.546	0
rs5987016	153948160	C/T	GAB3	none	0.78	0.78	0.0001852	0.546	0
rs2728723	153948687	G/A	GAB3	none	0.77	0.78	0.0001852	0.546	0
rs2728526	153949217	C/T	GAB3	none	0.78	0.78	0.0001852	0.546	0
rs2664172	153949614	C/G	GAB3	none	0.78	0.78	0.000208	0.531	0
rs1605895	153949793	G/T	GAB3	none	0.78	0.78	0.000208	0.531	0

rs1848763	153950635	G/T	GAB3	none	0.78	0.78	0.000208	0.531	0
rs2728528	153952147	C/T	GAB3	none	0.78	0.78	0.000208	0.531	0
rs35659282	153963293	CTAAT/C	GAB3	none	0.78	0.78	0.000208	0.531	0
rs2728725	153963756	T/C	GAB3	none	0.78	0.78	0.0001852	0.546	0
rs142295494	153972357	C/T	GAB3	none	0.08	0.74	4.28E-05	0.500	0
rs2664160	153998497	A/G	DKC1	none	0.84	0.79	0.0007592	0.710	0
rs1800533	154005148	G/A	DKC1	none	0.08	0.77	0.0003221	0.560	0
rs145403890	154009154	GC/G	MPP1	none	0.08	0.76	0.0001724	0.378	8.7616
rs2221730	154019083	T/C	MPP1	none	0.79	0.79	0.0002664	0.893	0
rs1126762	154020114	C/A	MPP1	none	0.79	0.80	0.0003173	0.888	0
rs2728536	154020918	T/C	MPP1	none	0.79	0.79	0.0002635	0.897	0
rs2048294	154022646	A/T	MPP1	none	0.79	0.78	0.0001286	0.912	0
rs4898396	154022818	C/A	MPP1	none	0.21	0.78	0.0001286	0.912	0
rs73641113	154022877	T/C	MPP1	none	0.13	0.79	0.0002443	0.664	0
rs1848762	154022952	G/T	MPP1	none	0.79	0.78	0.0001335	0.912	0
rs5945115	154023890	A/T	MPP1	none	0.78	0.78	9.75E-05	0.914	0
rs2664167	154024359	A/G	MPP1	none	0.79	0.78	0.0001554	0.916	0
rs2664168	154024573	T/G	MPP1	none	0.77	0.78	0.0001206	0.914	0
rs2728538	154025165	A/T	MPP1	none	0.8	0.78	0.0001916	0.757	0
rs6643707	154042428	C/T	MPP1	none	0.77	0.80	0.0007241	0.539	0
rs5945247	154048269	C/G	MPP1	none	0.13	0.80	0.000969	0.564	0
rs5987037	154048289	T/C	MPP1	none	0.81	0.80	0.0008479	0.670	0

Ref = Allele Reference
Alt = Allele Alternate
MAF = Minor Allele Frequency
OR = Odd ratio
P = P-value
Q = p-value for Cochrane's Q statistic
I^2 = heterogeneity index (0-100)

REFERENCES

1. Vollenhoven, L.A.R.v., *Advanced Handbook of Systemic Lupus Erythematosus*. 2018: Springer International Publishing Switzerland. 1.
2. **Frances Rees, M.D., Matthew J Grainge, Peter Lanyon, Weiya Zhang**, *The worldwide incidence and prevalence of systemic lupus erythematosus: a systematic review of epidemiological studies* Rheumatology, 2017. **56**(11): p. 1945–1961.
3. Megan R W Barber , C.D., Titilola Falasinnu , Alberta Hoi , Anselm Mak, et al. , ***Global epidemiology of systemic lupus erythematosus***. Nature Reviews Rheumatology, 2021.
4. Satirapoj, B., et al., *Predictors of renal involvement in patients with systemic lupus erythematosus.* Asian Pac J Allergy Immunol, 2007. **25**(1): p. 17-25.
5. Myles J. Lewis, A.S.J., *The effect of ethnicity and genetic ancestry on the epidemiology, clinical features and outcome of systemic lupus erythematosus* Rheumatology, 2016. **56**(suppl_1).
6. Mousavi, M.J., M. Mahmoudi, and S. Ghotloo, *Escape from X chromosome inactivation and female bias of autoimmune diseases.* Molecular Medicine, 2020. **26**(1): p. 127.
7. Renaudineau, W.H.B.a.Y., *Epigenetics and autoimmune diseases: the X chromosome-nucleolus nexus.* Frontier in Genetics, 2015. **6**(22).
8. Scofield, R.H., et al., *Klinefelter's syndrome (47,XXY) in male systemic lupus erythematosus patients: support for the notion of a gene-dose effect from the X chromosome.* Arthritis Rheum, 2008. **58**(8): p. 2511-7.
9. Taru Tukiainen, A.-C.V., Angela Yen, Manuel A. Rivas, et al., *Landscape of X chromosome inactivation across human tissues.* 12 October 2017.
10. David G. Wang, J.-B.F., Chia-Jen Siao, Anthony Berno, Peter Young and Ron Sapolsky, et al., *Large-Scale Identification, Mapping, and Genotyping of Single-Nucleotide Polymorphisms in the Human Genome.* Science, 1998. **280**(5366): p. 1077-1082.
11. Shastry, B.S., *Single Nucleotide Polymorphisms*. Methods in Molecular Biology™ Vol. 578. 2009: Humana Press, Totowa, NJ.
12. Andries T. Marees, H.d.K., Sven Stringer, Florence Vorspan, et al., *A tutorial on conducting genome-wide association studies:*

Quality control and statistical analysis. Psychiatric Research Published June 2018. **27**(2).

13. Huoru Zhang, Y.Z., Yong-Fei Wang1, David Morris, *Meta-analysis of GWAS on both Chinese and European populations identifies GPR173 as a novel X chromosome susceptibility gene for SLE.* Arthritis Research & Therapy, 2018.
14. Xianyong Yin, K.K., Hiroyuki Suetsugu, So-Young Bang,Leilei Wen, et al. , *Meta-analysis of 208370 East Asians identifies 113 susceptibility loci for systemic lupus erythematosus.* Annual of Rheumatic disease, 2020. **80**(5).
15. James Bentham, D.L.M., Deborah S Cunninghame Graham, Christopher L Pinder, et al., *Genetic association analyses implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of systemic lupus erythematosus.* Nature Genetics, 2015. **47**: p. 1457–1464

16. Tsao, M.L.L.R.a.B.P., *Chapter 11 - Genes and genetics in human SLE*, in *Systemic Lupus Erythematosus (Second Edition)*. 2021.
17. Zhengwei Zhu, Z.L., Xuejun Zhang, et al. , *Discovery of a novel genetic susceptibility locus on X chromosome for systemic lupus erythematosus.* Arthritis Research & Therapy volume, 2015. **17**: p. 349.
18. Hye-Ri Shin, W.K.C., In-Cheol Baek, Na Yeong Lee, et al., *Polymorphisms of IRAK1 Gene on X Chromosome Is Associated with Hashimoto Thyroiditis in Korean Children.* Endocrinology,, 2020. **161**(8).
19. Gang Liu, Y.T., Zhiqian Gao, Young-Jun Park and Edward Abraham, *Variant IL-1 Receptor-Associated Kinase-1 Mediates Increased NF-κB Activity.* The Journal of Immunology, 2007. **179**(6): p. 4125-4134.
20. Wang, D., *The essential role of G protein-coupled receptor (GPCR) signaling in regulating T cell immunity*. Immunopharmacology and Immunotoxicology, 2018. **40**(3): p. 187-192.
21. Christou, E.A.A., et al., *Sexual dimorphism in SLE: above and beyond sex hormones.* Lupus, 2018. **28**(1): p. 3-10.
22. Erin E Carter, S.G.B., Ann E Clarke, *The global burden of SLE: prevalence, health disparities and socioeconomic impact.* Nature Reviews Rheumatology, 2016. **12**: p. 605–620.
23. Julie S. Nusbaum, M.I.M.J.S., MD; et al., *Systemic Lupus Erythematosus*. Genes and genetics in human SLE. February 2020, Carolina, Charleston, SC, United States: Department of Medicine, Division of Rheumatology & Immunology, Medical University of South
24. Diane L. Kamen MD, M., *Environmental Influences on Systemic Lupus Erythematosus Expression.* Rheumatic Disease Clinics of North America, 2014. **40**(3): p. 401-412.
25. He, Y. and A.H. Sawalha, *Drug-induced lupus erythematosus: an update on drugs and mechanisms*. Current Opinion in Rheumatology, 2018. **30**(5).
26. Ginzler, N.K.a.E.M., *Is ethnicity linked to the severity of SLE manifestations?* Nature Reviews Rheumatology, 2019. **15**: p. 515–516
27. Bancha Satirapoj, J.W., Nantana Youngprang, Boonraksa Laonapaporn, Chitrada Thongdee, Sumonkarn Lapkitticharеonchai and Jayanton Patumanond, *Predictors of Renal Involvement in Patients with Systemic Lupus Erythematosus* Asian Pac J Allergy Immunol, 2007. **25**: p. 17-25.
28. Mu, W., Zhang, Wei, *Chapter 8 - Molecular Approaches, Models, and Techniques in Pharmacogenomic Research and Development*, in *Pharmacogenomics*, Y.-W.F. Lam, Cavallari, Larisa H., Editor. 2013, Academic Press: San Diego. p. 273-294.
29. *Allele*. Available from: https://www.genome.gov/search?terms=allele.
30. *Infinium Asian Screening Array-24 v1.0 BeadChip*. 2020; Available from: https://www.illumina.com/products/by-type/microarray-kits/infinium-asian-screening.html.
31. Shaun Purcell, B.N., Kathe Todd-Brown, Lori Thomas, Manuel A. R. Ferreira, et al., *PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses.* American Journal of Human Genetics, 2007. **87**(3): p. 559-575.
32. Consortium, I.H., *The International HapMap Project.* Nature volume 2003. **426**:

p. 789–796
33. Consortium*, T.G.P., *A map of human genome variation from population-scale sequencing.* Nature, 2010. **467**.
34. Laura Carrel, A.A.C., Karrie C. Goglin, and Huntington F. Willard, *A first-generation X-inactivation profile of the human X chromosome.* PNAS, 1999. **96**(25): p. 14440-14444.
35. Odhams, C.A., et al., *Interferon inducible X-linked gene CXorf21 may contribute to sexual dimorphism in Systemic Lupus Erythematosus.* Nature Communications, 2019. **10**(1): p. 2164.
36. Souyris, M., et al., *Female predisposition to TLR7-driven autoimmunity: gene dosage and the escape from X chromosome inactivation.* Seminars in Immunopathology, 2019. **41**(2): p. 153-164.
37. Bentham, J., et al., *Genetic association analyses implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of systemic lupus erythematosus.* Nature Genetics, 2015. **47**(12): p. 1457-1464.
38. *Chapter 12 - Single Nucleotide Polymorphisms and Applications.* . Advanced Topics in Forensic DNA Typing: Methodology. 2012.
39. Shastry, B.S., *SNPs: Impact on Gene Function and Phenotype*, B.S. Shastry, Editor. 2009, Humana Press, Totowa, NJ. p. 3-22.
40. Maggie Haitian Wanga, H.J.C., Kristel Van Steen, *Statistical methods for genome-wide association studies.* Statistical methods for genome-wide association studies, 2019: p. 53-60.
41. Jonathan Marchini, B.H., *Genotype imputation for genome-wide association studies.* Nature Reviews Genetics, 02 June 2010. **11**: p. 499–511.
42. Knight, C.M.L.a.J., *Introduction to Genetic Association Studies.* Cold Spring Harbor Protoc, November 8, 2020.
43. *Single Nucleotide Polymorphism (SNP).* [cited 2021 8 Oct]; Available from: https://www.sciencephoto.com/contributor/xx9/.
44. Sonthalia, S., et al., *Collagen Vascular Diseases*. 2013. p. 62.
45. Tiao, J., et al., *Using the American College of Rheumatology (ACR) and Systemic Lupus International Collaborating Clinics (SLICC) criteria to determine the diagnosis of systemic lupus erythematosus (SLE) in patients with subacute cutaneous lupus erythematosus (SCLE).* J Am Acad Dermatol, 2016. **74**(5): p. 862-9.
46. Lachance, J., *Hardy–Weinberg Equilibrium and Random Mating*, in *Encyclopedia of Evolutionary Biology*, R.M. Kliman, Editor. 2016, Academic Press: Oxford. p. 208-211.
47. Olivier Delaneau, J.M. and e.a. Jean-François Zagury, *A linear complexity phasing method for thousands of genomes.* Nature Methods, 2012. **9**(2).
48. O. Delaneau, J.M., JF. Zagury. *SHAPIT*. 17 January 2018; Available from: http://mathgen.stats.ox.ac.uk/genetics_software/shapeit/shapeit.html#reference.
49. Bryan Howie, C.F., Matthew Stephens, Jonathan Marchini, Gonçalo R Abecasis, *Fast and accurate genotype imputation in genome-wide association studies through pre-phasing.* Nature Genetics, 2012. **44**: p. 955–959.
50. Andrew D Skol, L.J.S., Gonçalo R Abecasis, Michael Boehnke, *Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies.* Nature Genetics, 2006. **38**: p. 209–213.

51. Delaneau, O., et al., *Integrating sequence and array data to create an improved 1000 Genomes Project haplotype reference panel.* Nature Communications, 2014. **5**(1): p. 3934.
52. James R Staley, E.J., Stephen Kaptoge, Adam S Butterworth, Michael J Sweeting, Angela M Wood and Joanna M M Howson, *A comparison of Cox and logistic regression for use in genome-wide association studies of cohort and case-cohort design.* European Journal of Human Genetics, 2017. **25**: p. 854–862.
53. Ehret, G.B., *Genome-Wide Association Studies: Contribution of Genomics to Understanding Blood Pressure and Essential Hypertension.* Current Hypertension Reports, 2010. **12**(1): p. 17-25.
54. *Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans.* Science, 2015. **348**(6235): p. 648-60.
55. Yang, J., et al., *Genomic inflation factors under polygenic inheritance.* European journal of human genetics : EJHG, 2011. **19**(7): p. 807-812.
56. Su, Y.-J., et al., *The association among leukocyte apoptosis, autoantibodies and disease severity in systemic lupus erythematosus.* Journal of translational medicine, 2013. **11**: p. 261-261.
57. Waltoft, B.L., et al., *The importance of distinguishing between the odds ratio and the incidence rate ratio in GWAS.* BMC Medical Genetics, 2015. **16**(1): p. 71.
58. Chaim O Jacob 1, J.Z., Don L Armstrong, Mei Yan, Jie Han, Xin J Zhou, James A Thomas, et al., *Identification of IRAK1 as a risk gene with critical role in the pathogenesis of systemic lupus erythematosus.* PNAS 2009. **106**: p. 6256-6261.
59. Amr H. Sawalha , R.W., Shizhong Han,Jennifer A. Kelly, et al., *Common Variants within MECP2 Confer Risk of Systemic Lupus Erythematosus.* PLOS ONE, 2008. **3**(3): p. e1727.
60. Kenneth M Kaufman, J.Z., Jennifer A Kelly, Travis Hughes, Adam Adler, Elena Sanchez, Joshua O Ojwang., et al, *Fine mapping of Xq28: both MECP2 and IRAK1 contribute to risk for systemic lupus erythematosus in multiple ancestral groups.* Annals of the Rheumatic Diseases 2012. **72**(3).
61. al., T.Z.e., *CD8 + T-cell senescence and skewed lymphocyte subsets in young Dyskeratosis Congenita patients with PARN and DKC1 mutations.* JCLA, 2020. **34**(9).
62. Claes Ohlsson, H.W., Kathryn L. Lunetta ,Lisette Stolk, et al., *Genetic Determinants of Serum Testosterone Concentrations in Men.* PLoS Genetics, 2011. **7**(10).
63. Zhang, Y., et al., *Meta-analysis of GWAS on two Chinese populations followed by replication identifies novel genetic variants on the X chromosome associated with systemic lupus erythematosus.* Hum Mol Genet, 2015. **24**(1): p. 274-84.
64. Anna S. Wilhelmson, A.S.W., Marta Lantero Rodriguez, Alexandra Stubelius, Per Fogelstrand, Inger Johansson, Matthew B. Buechler, Steve Lianoglou, Varun N. Kapoor, et al., *Testosterone is an endogenous regulator of BAFF and splenic B cell number.* Nature Communications, 2018.
65. Marcel Batten, J.G., Teresa G. Cachero, Fang Qian, Pascal Schneider, Jurg Tschopp, Jeffrey L. Browning and Fabienne Mackay, *BAFF mediates survival of peripheral immature B lymphocytes.* The Journal of experimental medicine, 2000(10): p. 1453–1466.

66. Fabien B. Vincent, E.F.M., Pascal Schneider & Fabienne Mackay, *The BAFF/APRIL system in SLE pathogenesis.* Nature Reviews Rheumatology, 2014. **10**: p. 365–373
67. Claudia Paret, P.S., Kirsten Vormbrock, et al., *CXorf61 is a target for T cell based immunotherapy of triple-negative breast cancer.* Oncotarget, 2015. **6**: p. 25356-25367.
68. Takashi Fukuyama, T.H., Mitsuhiro Takenoyama, et al., *Identification of a New Cancer/Germline Gene, KK-LC-1, Encoding an Antigen Recognized by Autologous CTL Induced on Human Lung Adenocarcinoma.* Cancer Research, 2006. **66**(9).
69. Wanling Yang, N.S., Dong-Qing Ye,Qiji Liu,Yan Zhang,Xiao-Xia Qian, Nattiya Hirankarn, et al., *Genome-Wide Association Study in Asian Populations Identifies Variants in ETS1 and WDFY4 Associated with Systemic Lupus Erythematosus.* PLoS Genetics, 2010. **6**(2): p. e1000841.
70. Magdalena B. Flak, D.S.K., Agua Sobrino, James Smith, Kimberly Pistorius, Francesco Palmas and Jesmond Dalli, *GPR101 mediates the pro-resolving actions of RvD5n-3 DPA in arthritis and infections.* J Clin Invest, 2019. **130**(1): p. 359-373.
71. Kaplan, M.J., *Neutrophils in the pathogenesis and manifestations of SLE.* Nature Reviews Rheumatology, 2011. **7**(12): p. 691-699.
72. Jiang, S.-J., et al., *A potential peptide derived from cytokine receptors can bind proinflammatory cytokines as a therapeutic strategy for anti-inflammation.* Scientific Reports, 2019. **9**(1): p. 2317.
73. van Bavel, C.C., et al., *Apoptosis-induced histone H3 methylation is targeted by autoantibodies in systemic lupus erythematosus.* Annals of the Rheumatic Diseases, 2011. **70**(1): p. 201.
74. Huggins, M.L., et al., *Antibodies from systemic lupus erythematosus (SLE) sera define differential release of autoantigens from cell lines undergoing apoptosis.* Clin Exp Immunol, 1999. **118**(2): p. 322-8.
75. Dalli, J., R.A. Colas, and C.N. Serhan, *Novel n-3 Immunoresolvents: Structures and Actions.* Scientific Reports, 2013. **3**(1): p. 1940.
76. Simon P. Brooks, M.C., Hao R. Tang, et al., *The Nance–Horan syndrome protein encodes a functional WAVE homology domain (WHD) and is important for co-ordinating actin remodelling and maintaining cell morphology* Human Molecular Genetics, 2010. **9**(12): p. 2421–2432.
77. Thien A Nguyen 1, K.W., Saurabh Pandey 2, Alexander W Lehr 3, Yan Li , et al., *A Cluster of Autism-Associated Variants on X-Linked NLGN4X Functionally Resemble NLGN4Y.* Neuron, 2020. **106**(5): p. 759-768.e7.
78. Cotton, A.M., et al., *Analysis of expressed SNPs identifies variable extents of expression from the human inactive X chromosome.* Genome Biology, 2013. **14**(11): p. R122.
79. Jeremy Chien, J.S., Rajeswari Avula, Heyu Zhang, Wanguo Liu, Lynn C Hartmann, et al., *Epigenetic silencing of TCEAL7 (Bex4) in ovarian cancer.* Oncogene, 2005. **24**: p. 5089–5100.
80. F C Arnett , R.G.H., M G Roebber, J B Harley, M Reichlin, *Increased frequencies of Sm and nRNP autoantibodies in American blacks compared to whites with systemic lupus erythematosus.* The Journal of Rheumatology, 1988.

15(12): p. 1773-1776.

81. P. Migliorini, C.B., V. Rocchi & S. Bombardieri, *Anti-Sm and anti-RNP antibodies.* Autoimmunity, 2005. **38**.
82. Francesco Ingenito, G.R., Alessandra Affinito,Silvia Nuzzo, OrcID,Iolanda Scognamiglio,Cristina Quintavalle and Gerolama Condorelli *The Role of Exo-miRNAs in Cancer: A Focus on Therapeutic and Diagnostic Applications.* Molecules Science, 2019. **20**(19).
83. Baldari, C.U.a.C.T., *Regulation of T Cell Activation and Differentiation by Extracellular Vesicles and Their Pathogenic Role in Systemic Lupus Erythematosus and Multiple Sclerosis.* molecules, 2017. **22**(225).
84. Wenli Li, C.D., Hanbo Yang, and Guochun Wang, *The Regulatory T Cell in Active Systemic Lupus Erythematosus Patients: A Systemic Review and Meta-Analysis.* Fronteir in Immunology, 2019.
85. Fang, H., C.M. Disteche, and J.B. Berletch, *X Inactivation and Escape: Epigenetic and Structural Features.* Frontiers in Cell and Developmental Biology, 2019. **7**(219).
86. Youness, A., C.-H. Miquel, and J.-C. Guéry, *Escape from X Chromosome Inactivation and the Female Predominance in Autoimmune Diseases.* International journal of molecular sciences, 2021. **22**(3): p. 1114.
87. Sharp, A.J., et al., *DNA methylation profiles of human active and inactive X chromosomes.* Genome Res, 2011. **21**(10): p. 1592-600.
88. Jaillon, S., K. Berthenet, and C. Garlanda, *Sexual Dimorphism in Innate Immunity.* Clinical Reviews in Allergy & Immunology, 2019. **56**(3): p. 308-321.
89. Jaillon, S., K. Berthenet, and C. Garlanda, *Sexual Dimorphism in Innate Immunity.* Clin Rev Allergy Immunol, 2019. **56**(3): p. 308-321.
90. Klein, S.L. and K.L. Flanagan, *Sex differences in immune responses.* Nature Reviews Immunology, 2016. **16**(10): p. 626-638.
91. Libert, C., L. Dejager, and I. Pinheiro, *The X chromosome in immune functions: when a chromosome makes the difference.* Nature Reviews Immunology, 2010. **10**(8): p. 594-604.
92. R. Hal Scofield, G.R.B., Bahram Namjou,Robert P. Kimberly, et al., *Klinefelter's syndrome (47,XXY) in male systemic lupus erythematosus patients: Support for the notion of a gene-dose effect from the X chromosome.* ARTHRITIS & RHEUMATISM, 2008. **58**.
93. Cano-Gamez, E. and G. Trynka, *From GWAS to Function: Using Functional Genomics to Identify the Mechanisms Underlying Complex Diseases.* Frontiers in Genetics, 2020. **11**(424).
94. van der Wijst, M.G.P., et al., *Single-cell RNA sequencing identifies celltype-specific cis-eQTLs and co-expression QTLs.* Nature Genetics, 2018. **50**(4): p. 493-497.
95. Wu, S., et al., *Genome-wide association studies and CRISPR/Cas9-mediated gene editing identify regulatory variants influencing eyebrow thickness in humans.* PLoS genetics, 2018. **14**(9): p. e1007640-e1007640.
96. Martin Aringer, K.C., David Daikh, Ralph Brinks, Marta Mosca, et al., *2019 European League Against Rheumatism/American College of Rheumatology Classification Criteria for Systemic Lupus Erythematosus.* Arthritis & Rheumatology, 2019. **71**(9): p. 1400-1412.

VITA

NAME	Krisana Jaiwan
DATE OF BIRTH	3 March 1994
PLACE OF BIRTH	Lampang, Thailand.
INSTITUTIONS ATTENDED	School of Health Science, Mae Fah Luang University
HOME ADDRESS	97/5 village No.4 Sub-district Namcho. District Mae Tha. Lampang Province 52150