

บทความวิจัย (Research Article)**An exploratory genome-wide association study of leprosy susceptibility genes in a Thai cohort**

Sukanya Wattanapokayakit^{1*}, Ekawat Pasomsub², Punna Kunhapan¹, Suphannee Phatsadon³, Punnarai Veeraseatakul⁴, Krisada Mahotarn⁵, Taisei Mushiroda⁶ and Surakameth Mahasirimongkol¹

¹ Medical Life Sciences Institute, Department of Medical Sciences, Ministry of Public Health, Nonthaburi, 11000 Thailand

² Faculty of Medicine Ramathibodi Hospital, Mahidol University, Bangkok, 10400 Thailand

³ Sirinathorn Hospital, Ministry of Public Health, Nonthaburi, 11000 Thailand

⁴ Regional Medical Sciences Center 1 Chiang Mai, Department of Medical Sciences, Ministry of Public Health, Chiang Mai, 50180 Thailand

⁵ Rajprachasamasai Institute, Department of Disease Control, Ministry of Public Health, Samut Prakan, 10130 Thailand

⁶ Center for Genomic Medicine, RIKEN Yokohama Campus, Yokohama City, Kanagawa, 230-0045 Japan

* Correspondence author, E-mail: sukanya.w@dmsc.mail.go.th

Health Science, Science and Technology Reviews. 2026;19(2):68-79.

Received: 29 January 2026; Revised: 22 July 2026; Accepted: 3 August 2026

Abstract

Leprosy is a chronic granulomatous infectious disease caused by *Mycobacterium leprae*. Although candidate gene studies have explored genetic susceptibility to leprosy, only a limited number of genomic loci have been independently replicated. In contrast, genome-wide association studies (GWAS) have provided clear evidence that host genetic factors—varying across different ethnic populations—play a critical role in determining human susceptibility to infection and in influencing the progression of infectious diseases. We performed genome-wide association study involving 344 individuals with leprosy - 103 with paucibacillary form and 241 multibacillary form - and 200 control individuals, using the Illumina Omni express platform. This study highlights a potential genome-wide association signal for leprosy on chromosome 18 (18q21.33–18q22.1) supported by suggestive evidence (P -values $< 1 \times 10^{-5}$). Within this region, the lead variant rs9949621 demonstrated the strongest association (P -values $< 4.75 \times 10^{-6}$, Odds Ratio [OR] = 0.53, 95% Confidence Interval [CI] = 0.40-0.70) and is located near the *HMSD* gene. The association peak extended across a broader region encompassing several *SERPIN* family genes (*SERPINB2*, *SERPINB10*, *HMSD*, and *SERPINB8*), where multiple single nucleotide polymorphisms (SNPs) consistently showed P -values below 5.0×10^{-5} . Collectively, this suggests that variants within this gene cluster may influence genetic susceptibility to leprosy. Importantly, all associated SNPs appeared to confer a protective effect, as the minor alleles were more frequent in controls than in cases, indicating reduced disease risk in carriers. Our genome-wide association study (GWAS) suggests a potential novel link between *HMSD* and several genes within the *SERPIN* family and susceptibility to leprosy in the Thai population. While these results highlight a promising genetic association, they will need to be confirmed through independent replication studies and supported by functional investigations. Additional high-

resolution genotyping or sequencing analyses will also be necessary to refine these findings and clarify their biological significance.

Keywords: Leprosy, Genome-wide association studies, GWAS

Introduction

Leprosy is a chronic skin infection caused by *Mycobacterium leprae* (*M. Leprae*), with infected individuals serving as the primary reservoir of transmission. The organism primarily resides in the skin, peripheral nerves, and nasal mucosa of infectious patients, and is believed to be transmitted mainly through respiratory droplets from the upper respiratory tract during prolonged close contact with untreated individuals. The incubation period is typically 3–5 years. Following exposure, most individuals mount an effective immune response against *M. leprae* and do not develop clinical disease; however, those with an impaired or inadequate immune response may develop leprosy. Notably, only a small proportion of exposed individuals, estimated at approximately 3–5%, are susceptible to infection and subsequently develop clinical manifestations of the disease, suggesting an important role for host genetic factors in disease susceptibility [1,2]. While the World Health Organization (WHO) reported more than 172,717 new cases in 2024, the global burden of disease has been steadily declining [3]. A similar trend is observed in Southeast Asia, where leprosy control has improved significantly. In Thailand, 80 new cases were reported in 2024, with 155 patients registered for treatment [4]. The close alignment between incidence and prevalence in Thailand indicates effective disease management.

Transmission occurs primarily through the airborne route; however, most individuals are naturally resistant, and the bacteria are eliminated by innate and adaptive immune responses. It is estimated that only about 5% of infected individuals develop clinical disease. To guide treatment, the WHO classifies leprosy into two forms based on bacterial burden: paucibacillary (PB) and multibacillary (MB). Crucially, paucibacillary (PB) patients typically mount a robust Th1-mediated cellular immune response, resulting in tight pathogen containment and localized, sparse skin lesions. Conversely, the multibacillary (MB) phenotype is characterized by a predominant Th2 humoral response that is permissive to extensive bacterial replication—a distinct immunopathological divergence driven primarily by host genetic determinants.

Despite major advances in the control of leprosy, our understanding of its immunogenetic basis initially centered on the role of the major histocompatibility complex (MHC) genes, which were recognized as key determinants of host susceptibility following infection with *M. leprae*. However, progress in genetic epidemiology has expanded this perspective, demonstrating that leprosy risk is influenced by a broader range of genes and biological pathways beyond the MHC region [5-9]. Evidence for host genetic involvement first emerged from family and population studies, which revealed clustering of cases among relatives. In South Indian populations, linkage analysis identified a susceptibility locus on chromosome 10p13 [5], while subsequent research in Tamil Nadu reported another locus on chromosome 20p12 [6]. Further linkage analysis highlighted a region on chromosome 6q25 in Vietnamese population [7].

Expanding beyond India, genome-wide association studies (GWAS) in Chinese cohorts identified significant associations between single SNPs in several immune-related genes—including *CCDC122*, *C13orf31*, *NOD2*, *TNFSF15*, *RIPK2*, and the classical MHC loci on chromosome 6 [8]. Complementary evidence from Vietnamese populations pointed to a cluster of interleukin receptor genes, including *IL18R1*, *IL18RAP*, and

IL1RL1, on chromosome 2q12.1, which play a role in inflammatory signaling [9]. More recently, GWAS conducted in African cohorts from Malawi and Mali identified a novel susceptibility locus at chromosome 10q24, with *ACTR1A* proposed as a candidate gene influencing host response to *M. leprae* [10].

Taken together, these findings underscore that genetic susceptibility to leprosy is highly heterogeneous and shaped by population-specific risk factors. Rather than being governed by a single pathway, leprosy risk arises from a complex genetic architecture involving both classical immune genes and novel, previously unrecognized loci, reflecting the diverse evolutionary pressures acting on human populations. Consequently, relying solely on genetic discoveries from other well-studied populations, such as Chinese or African cohorts, risks overlooking suggestive protective or causative loci unique to the Southeast Asian and Thai genetic backgrounds.

Sample Collection

A total of 544 Thai participants were enrolled, including 344 leprosy patients (103 PB and 241 MB cases) and 200 healthy controls from endemic areas of northeastern Thailand. Detailed baseline demographics are summarized in Table 1. Notably, confirmation of *Mycobacterium leprae* infection relied on the microscopic detection of AFB via slit-skin smears rather than bacterial sequencing.

Table 1. Demographic Characteristics of the Study Cohort.

Characteristic	Paucibacillary (PB) (n=103)	Multibacillary (MB) (n=241)	Healthy Controls (n=200)
Sex, n (%)			
Male	39 (38%)	145 (60%)	102 (51%)
Female	64 (62%)	96 (40%)	98 (49%)
Age (Years)			
Mean	50	52	49
Range	18-78	18-89	18–89

Eligibility was strictly determined according to the following inclusion and exclusion criteria:

Inclusion Criteria:

Case Group (Index cases / Leprosy Patients)

Patients were diagnosed in accordance with the World Health Organization (WHO) diagnostic criteria. To be included, individuals had to meet at least one of the two following conditions:

- Clinical Presentation: Manifestation of at least two of the three primary clinical signs: characteristic leprosy skin lesions, anesthesia (loss of sensation either directly within the skin lesions or along peripheral nerve pathways damaged by leprosy), or enlarged peripheral nerves.

- Laboratory Confirmation: Positive detection of acid-fast bacilli (AFB) via a slit-skin smear examination alone was sufficient for a definitive diagnosis.

Patients were further categorized and selected based on the following specific criteria:

- Paucibacillary (PB) Leprosy Cases: Presenting with 1 to 5 skin lesions.

- Multibacillary (MB) Leprosy Cases: Presenting with more than 5 skin lesions.
- General Criteria for PB and MB Cases: Male or female participants aged 15 to 90 years who were either treatment-naïve or had received chemotherapy for a cumulative duration not exceeding 6 months.

Control Group (Community Contacts)

Healthy controls comprised individuals residing within high-risk, leprosy-endemic regions of northeastern Thailand. Eligible participants must have fit into one of the following community contact categories relative to an index case:

- Household Contacts: Individuals co-habiting in the same residence as an index case.
- Neighboring Contacts: Individuals living in households immediately adjacent to or surrounding an index case.
- Social Contacts: Individuals maintaining regular social networks or close community relationships with an index case.

Exclusion criteria:

- Children under 2 years of age
- Individuals with active tuberculosis (TB), a history of TB, or suspected TB infection
- Pregnant women
- Individuals who declined to participate in the study (or individuals unwilling to provide informed consent)

Endemic areas for sample collection:

- Sirindhorn Hospital, Khon Kaen Province
- Non Somboon Colony, Khon Kaen Province
- Office of Disease Prevention and Control Region 5, Nakhon Ratchasima Province
- Office of Disease Prevention and Control Region 6, Khon Kaen Province
- Office of Disease Prevention and Control Region 7, Ubon Ratchathani Province
- Office of Disease Prevention and Control Region 12, Department of Disease Control
- High-risk areas: Districts in the Northeastern region (based on the areas and definitions of the Chemoprophylaxis project)

This research was approved by the Ethical Review Committee for Research in Human Subjects, Ministry of Public Health (Document No. 54/2009, Ref. No. 94/2551) on 15 July 2009. Data collection was initiated after obtaining official ethical approval.

DNA extraction and microarray-based SNP genotyping

Genomic DNA was isolated from peripheral blood samples (5ml in EDTA) obtained from all participants using the QIAamp® DNA Blood Mini Kit (QIAGEN, Hilden, Germany) according to the manufacturer's protocol. Microarray-based SNP genotyping was subsequently performed using the Infinium HD Ultra Illumina OmniExpress Assay platform (Illumina, San Diego, CA, USA) to interrogate 731,442 genetic markers. Following hybridization and imaging on an Illumina iScan system, genotype calling and data quality assessment were executed using the Genotyping Module in Illumina GenomeStudio Software, with samples achieving a call rate below 95% excluded from downstream analyses.

Data Quality Control and Statistical Analysis

Quality control of the genotyping data, encompassing 544 Thai participants (103 PB and 241 MB cases and 200 controls) and 731,442 genetic markers, was executed in multiple sequential stages. During the pre-SNP and sample quality control phase using PLINK (v1.90b5.4), samples with a call rate below 0.98, gender mismatches, or high relatedness (identity-by-descent > 0.9) were excluded. SNPs located on sex chromosomes were also removed. Principal component analysis (PCA) was then conducted using IPCAPS v1.1.8 with an iterative pruning threshold of 0.03 to identify and remove outlier samples [11]. Following this step, 512 samples (98 PB, 227 MB, and 187 controls) and 712,726 SNPs remained for downstream analysis.

Post-SNP and sample quality control was subsequently applied to the 512 remaining samples. This included filtering based on missing data (--mind 0.1 and --geno 0.05), Hardy–Weinberg equilibrium (--hwe 1e-06), and minimum allele frequency (--maf 0.01). In addition, identity-by-descent estimation (--ibd 0.185) was used to remove related samples up to the second- and third-degree level. After these procedures, 466 samples (91 PB, 214 MB, and 161 controls) and 605,763 SNPs remained.

In the final phase of our study, we performed a genome-wide association study (GWAS) to identify genetic variants that may contribute to susceptibility to leprosy. The analysis was carried out using FaST-LMM (Factored Spectrally Transformed Linear Mixed Models), a linear mixed model framework that efficiently accounts for population structure, hidden relatedness, and sample heterogeneity, thereby minimizing the likelihood of false-positive associations [12]. For each single-nucleotide polymorphism (SNP), we calculated the odds ratio (OR) and the corresponding 95% confidence interval (CI) using PLINK (v1.90b5.4) to evaluate the strength and direction of association with disease status. To illustrate the genome-wide distribution and significance of association signals, we generated Manhattan plots, which display the $-\log_{10}(P\text{-value})$ for each SNP, facilitating the identification of genomic loci with the strongest associations. In addition, quantile–quantile (Q–Q) plots were produced to compare observed versus expected P -value distributions, allowing assessment of potential inflation or deflation of test statistics. All visualizations were created in R (version 4.4.3), providing a comprehensive graphical representation of the GWAS results.

Results

We analyzed a total of 344 individuals diagnosed with leprosy, consisting of 103 paucibacillary (PB) cases and 241 multibacillary (MB) cases, together with 200 unaffected controls. Quality control (QC) for both samples and SNPs was performed as detailed in the Data Quality Control and Statistical Analysis section. To account for potential population stratification, principal component analysis (PCA) was conducted using IPCAPS (v1.1.8), enabling the identification and exclusion of outlier samples that deviated from the main study population (Figure 1). Following final SNP- and sample-level filtering, a clean dataset comprising 466 samples (91 PB, 214 MB, and 161 controls) and 605,763 SNPs was retained for genome-wide association testing. The association results comparing the combined leprosy cohort (both PB and MB subtypes) against healthy controls were visualized using Manhattan and Q–Q plots (Figures 2A and 2B). The genomic inflation factor ($\lambda = 1.007$) confirmed the absence of systematic inflation due to population stratification or confounding artifacts. Subgroup-stratified GWAS analyses comparing MB vs. controls, PB vs. controls, and MB vs. PB did not reveal any genome-wide significant or notably associated genes (Supplementary Figures 1–3, respectively).

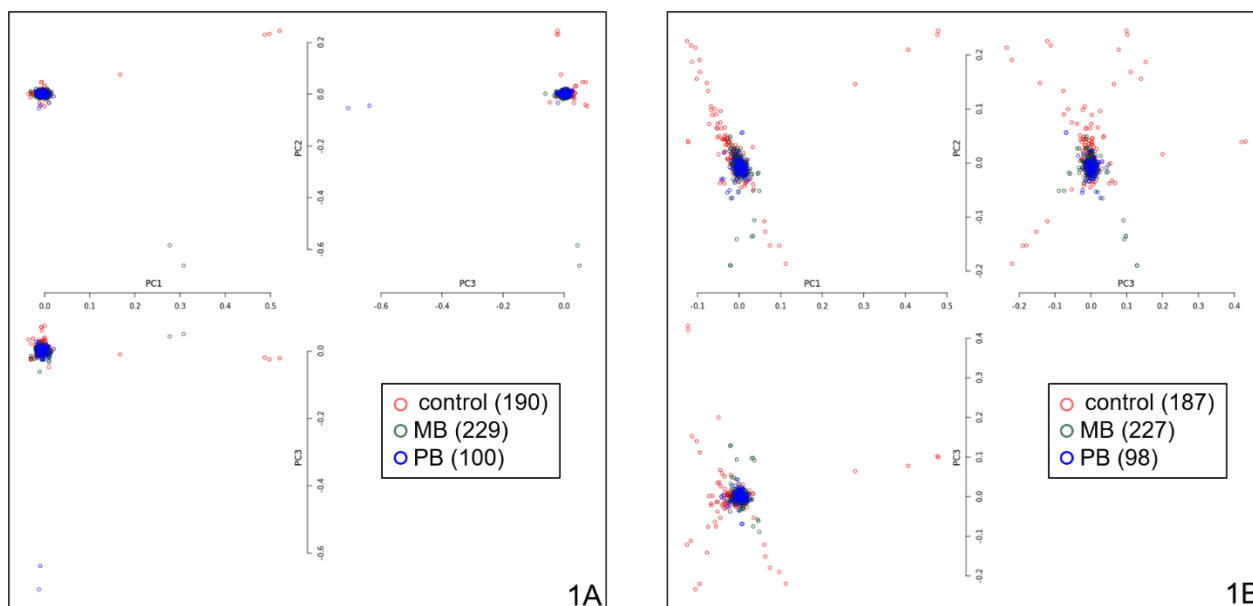


Figure 1. Principal component analysis (PCA) of the study population using IPCAPS v1.1.8. **(A)** All samples prior to outlier removal. **(B)** Final dataset after removing outliers, including 98 PB cases, 227 MB cases, and 187 controls.

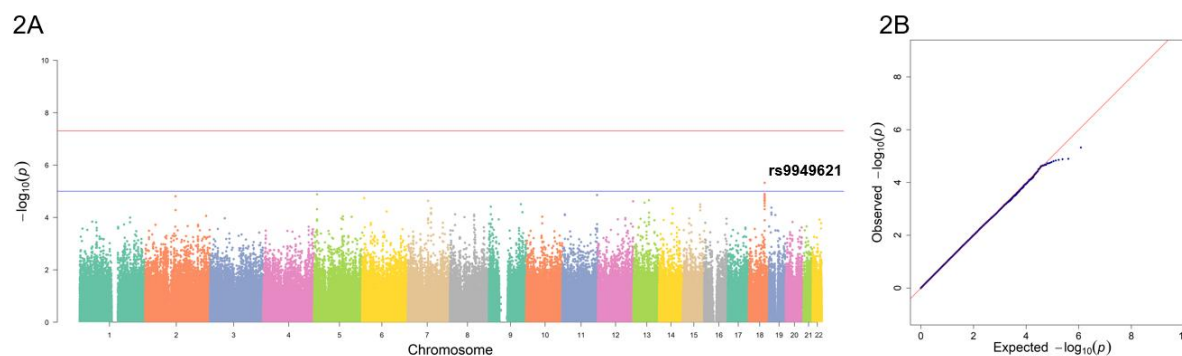


Figure 2. Genome-wide association results for leprosy. **(A)** Manhattan plot showing the $-\log_{10}(P\text{-values})$ of SNPs across the genome. The strongest signal mapped to the lead variant rs9949621 on chromosome6:63962163, located adjacent to the *HMSD* gene ($P\text{-values} < 4.75 \times 10^{-6}$, OR = 0.53, 95% CI = 0.40-0.70). **(B)** Quantile-quantile (Q-Q) plot comparing observed versus expected P -values. The genomic inflation factor (λ) was estimated at 1.007363

Table 2. Genotype and allele frequencies in 305 leprosy cases (PB+MB) and 161 controls, highlighting a strong association on chromosome 18 (rs9320032–rs17072304; chr18:63,893,977–63,966,335).

SNP	Position (hg38)	Gene (Closest Gene)	Allele	Case			MAF ^a	Control			MAF ^a	Odds ratio	95% CI	P
rs9320032	63893977	SERPINB2	T/C	TT 152	TC 120	CC 33	0.30	TT 51	TC 75	CC 33	0.44	0.56	0.43-0.74	2.39E-05
rs6102	63903236	SERPINB2	T/G	TT 152	TG 120	GG 33	0.30	TT 53	TG 73	GG 33	0.44	0.58	0.44-0.76	4.83E-05
rs9966202	63912143	SERPINB10	G/A	GG 152	GA 120	AA 33	0.30	GG 51	GA 74	AA 34	0.45	0.56	0.42-0.73	2.14E-05
rs17072097	63915517	SERPINB10	T/G	TT 151	TG 119	GG 33	0.31	TT 51	TG 74	GG 34	0.45	0.56	0.42-0.74	2.27E-05
rs11152409	63937045	(SERPINB10)	T/C	TT 152	TC 120	CC 33	0.30	TT 51	TC 74	CC 34	0.45	0.56	0.42-0.73	2.14E-05
rs9955763	63947201	(HMSD)	A/G	AA 151	AG 121	GG 33	0.31	AA 50	AG 76	GG 33	0.45	0.56	0.42-0.74	2.90E-05
rs8090046	63954852	HMSD	G/A	GG 148	GA 125	AA 32	0.31	GG 49	GA 78	AA 32	0.45	0.56	0.42-0.74	3.65E-05
rs9953837	63958124	HMSD	T/C	TT 152	TC 122	CC 31	0.30	TT 49	TC 79	CC 31	0.44	0.54	0.41-0.72	1.87E-05
rs9962574	63958884	HMSD	G/T	GG 151	GT 123	TT 31	0.30	GG 49	GT 78	TT 32	0.45	0.54	0.41-0.72	1.48E-05
rs9949621	63962163	(HMSD)	T/G	TT 151	TG 123	GG 31	0.30	TT 49	TG 75	GG 35	0.46	0.53	0.40-0.70	4.75E-06
rs1006754	63964213	(HMSD, SERPINB8)	C/T	CC 152	CT 122	TT 31	0.30	CC 49	CT 79	TT 31	0.44	0.54	0.41-0.72	1.87E-05
rs2009989	63966010	(HMSD, SERPINB8)	T/C	TT 152	TC 122	CC 31	0.30	TT 49	TC 78	CC 32	0.45	0.54	0.41-0.71	1.27E-05
rs17072304	63966335	(HMSD, SERPINB8)	T/C	TT 147	TC 124	CC 34	0.31	TT 47	TC 78	CC 34	0.46	0.55	0.41-0.72	1.69E-05

Note: ^a MAF is minor allele frequency.

A suggestive association signal was detected on chromosome 18, where a cluster of SNPs spanning from rs9320032 to rs17072304 (genomic region chr18:63,893,977–63,966,335) exceeded the suggestive threshold for statistical significance (P -value $< 5.0 \times 10^{-5}$) (Table 2). Within this region, the strongest signal mapped to the lead SNP rs9949621 (P -value of 4.75×10^{-6}). To evaluate the local linkage disequilibrium (LD) structure, we generated regional association plots using LocusZoom [13]. As shown in Figure 3, the lead SNP is strongly supported by a cluster of neighboring variants in high LD, reinforcing confidence in the observed association. Comprehensive GWAS results are available in Supplementary Table 1.

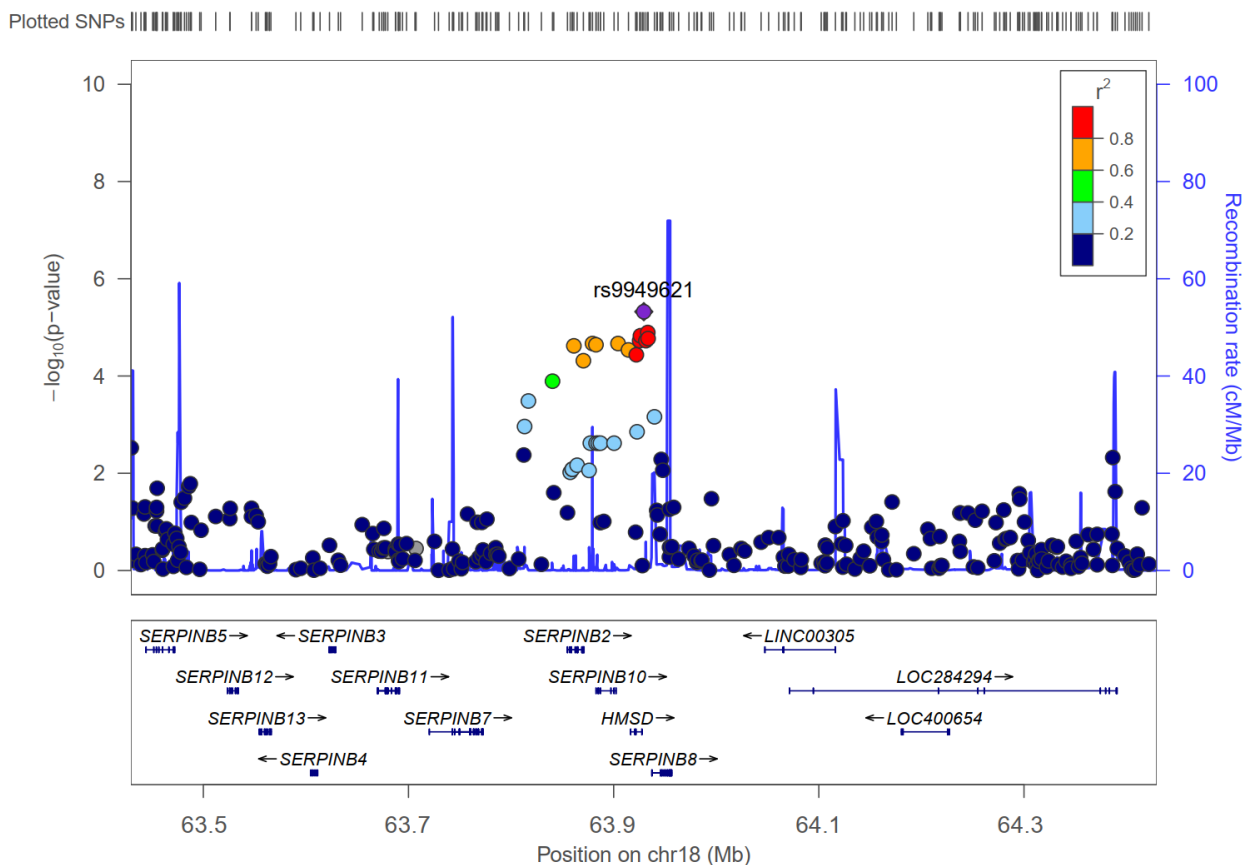


Figure 3. Regional association plot of the chromosome 18 locus. LocusZoom plot [13] illustrating the local linkage disequilibrium (LD) structure surrounding the lead SNP, rs9949621 (P -value of 4.75×10^{-6}). The associated genomic region (chr18:63,893,977–63,966,335) features a cluster of variants spanning from rs9320032 to rs17072304 that exceed the suggestive statistical significance threshold (P -value $< 5.0 \times 10^{-5}$)

This cluster of SNPs is located in close proximity to the HMSD gene and the SERPIN gene family, implicating this genomic locus as a strong candidate for genetic susceptibility to leprosy (Table 2). Interestingly, the SNPs identified within this susceptibility region all appeared to exert a protective effect rather than increasing disease risk. Specifically, the alternative alleles (minor alleles) were observed more frequently in the control group than in the leprosy cases, and the corresponding odds ratios and 95% confidence intervals were consistently below 1. This indicates that individuals carrying these minor alleles were less likely to develop leprosy, supporting a protective role of these variants.

The leprosy-associated SNPs identified in our Thai cohort are suggestive protective variants (enriched in controls relative to cases) that have not been reported in previous studies—where instead entirely different causative regions were highlighted—matching summary statistics (such as odds ratios and confidence intervals) for these specific alleles are unavailable in the literature for other populations. To further contextualize these findings, we compared the minor allele frequencies (MAF) of the significant chromosome 18 variants in our Thai dataset with the MAF values reported in global reference populations (Table 3). Remarkably, the control group in our study displayed a much higher frequency of the alternative alleles compared not only with the leprosy cases from this cohort, but also with global population datasets. This consistent pattern reinforces the hypothesis that these variants may confer population-specific protection against leprosy in Thais.

Table 3. Comparison of minor allele frequencies (MAF) for significant chromosome 18 variants in the Thai dataset with those reported in global reference populations.

SNP	Position (hg38)	Gene (Closest Gene)	Allele	Minor Allele Frequency			
				This Study		GnomAD (v4.1.0) ^a	
				Case	Control	All Population	East Asian
rs9320032	63893977	<i>SERPINB2</i>	T/C	0.30	0.44	0.18	0.22
rs6102	63903236	<i>SERPINB2</i>	T/G	0.30	0.44	0.21	0.18
rs9966202	63912143	<i>SERPINB10</i>	G/A	0.30	0.45	0.26	0.22
rs17072097	63915517	<i>SERPINB10</i>	T/G	0.31	0.45	0.21	0.17
rs11152409	63937045	(<i>SERPINB10</i>)	T/C	0.30	0.45	0.28	0.22
rs9955763	63947201	(<i>HMSD</i>)	A/G	0.31	0.45	0.25	0.22
rs8090046	63954852	<i>HMSD</i>	G/A	0.31	0.45	0.27	0.23
rs9953837	63958124	<i>HMSD</i>	T/C	0.30	0.44	0.30	0.23
rs9962574	63958884	<i>HMSD</i>	G/T	0.30	0.45	0.34	0.23
rs9949621	63962163	(<i>HMSD</i>)	T/G	0.30	0.46	0.35	0.24
rs1006754	63964213	(<i>HMSD</i> , <i>SERPINB8</i>)	C/T	0.30	0.44	0.29	0.23
rs2009989	63966010	(<i>HMSD</i> , <i>SERPINB8</i>)	T/C	0.30	0.45	0.29	0.23
rs17072304	63966335	(<i>HMSD</i> , <i>SERPINB8</i>)	T/C	0.31	0.46	0.20	0.23

Note: ^a <https://gnomad.broadinstitute.org/>

Discussion

In this study, we identified a suggestive protective signal on chromosome 18 by expanding our analysis to a 0.1 Mb genomic window (chr18:63,880,000–63,980,000). Given that this locus did not clear the conservative genome-wide significance threshold ($P < 5.0 \times 10^{-8}$) and our cohort lacks an independent replication sample, the association between this region and leprosy susceptibility must be treated as preliminary. Because the incidence of leprosy has dropped markedly in Thailand, recruiting a massive cohort poses unique logistical restrictions. Consequently, our study is constrained by this relatively modest sample size, which limits our statistical power to detect rare or small-effect genetic variants. This investigation should therefore be viewed as an exploratory baseline designed primarily to capture localized, high-impact signals within the Thai population. Furthermore, we emphasize that our classification of these alleles as "protective" is inferred strictly

from statistical parameters—specifically, differences in minor allele frequencies (MAF) and calculated odds ratios (OR) within this case-control sample—and must not be interpreted as definitive biological protection without direct functional validation. Nevertheless, the evaluated region encompasses a cluster of four biologically compelling candidate genes: *SERPINB2*, *SERPINB10*, *HMSD*, and *SERPINB8*. These genes belong to the serpin (serine protease inhibitor) superfamily, known to participate in inflammation, apoptosis, and host–pathogen interactions, all processes central to leprosy pathogenesis.

SERPINB8 encodes PI8, a protease inhibitor that regulates furin-like proteases involved in cell signaling, immune responses, and pathogen infectivity. Mutations in *SERPINB8* have been associated with exfoliative ichthyosis, a skin disorder marked by impaired keratinocyte adhesion and mechanical fragility, highlighting its role in maintaining epidermal barrier integrity. Considering that *M. leprae* invades skin and peripheral nerves, *SERPINB8* dysfunction may modulate host susceptibility by affecting epidermal structure and immune defense [14,15].

SERPINB10 (bomapin) is primarily expressed in myeloid cells and functions as a redox-sensitive regulator of cell fate, promoting proliferation under growth-supportive conditions while enhancing apoptosis under limited growth factor availability. This suggests a critical role in balancing myeloid cell expansion and survival, with dysregulation potentially influencing innate immune responses against *M. leprae*, particularly in macrophages and other myeloid-derived cells [16].

HMSD encodes a serpin-related minor histocompatibility antigen. Although its protease inhibitory function is not fully understood, HMSD plays a well-documented immunological role: its peptide products can be presented by HLA molecules and recognized by cytotoxic T cells. Expression in immune-related cells suggests relevance for antigen presentation and host defense against intracellular pathogens, including *M. leprae* [17].

SERPINB2 (plasminogen activator inhibitor type 2, PAI-2) classically regulates fibrinolysis but is also implicated in immune processes such as macrophage differentiation and inflammatory responses. Its expression is elevated during infection and inflammation, suggesting a role in modulating host immune responses to *M. leprae* [18–20].

Early genetic studies primarily focused on the major histocompatibility complex (MHC), but subsequent research demonstrated that host susceptibility involves a much broader array of genes and biological pathways. For instance, linkage analyses in South India identified susceptibility loci on chromosomes 10p13 and 20p12 [6,7], while studies in Vietnamese populations highlighted chromosome 6q25 [8]. Genome-wide association studies (GWAS) have since expanded these insights. In Chinese cohorts, immune-related loci including *CCDC122*, *C13orf31*, *NOD2*, *TNFSF15*, *RIPK2*, and the MHC region were heavily implicated [9]. Meanwhile, studies in Vietnam associated variants within the interleukin receptor cluster (*IL18R1*, *IL18RAP*, *IL1RL1*) on 2q12.1 with leprosy risk [10], and an evaluation of African cohorts revealed a novel locus at 10q24 with *ACTR1A* as a strong candidate gene [11].

The divergence between our findings and those of previously reported leprosy GWAS in other populations can be attributed primarily to differences in statistical power and ancestral genetic backgrounds. For example, the landmark GWAS in China utilized a discovery cohort of 670 cases and 555 controls, which expanded into a massive multi-stage replication cohort totaling 4,473 cases and 2,707 controls [4]. This

substantial sample size provided vastly superior statistical power compared to our relatively small Thai cohort. Furthermore, the minor allele frequencies (MAFs) observed in our Thai control dataset—derived from the Northeastern region—distinctly deviate from both East Asian and global reference populations, as detailed in Table 3. This underscores the profound impact of population-specific genetic architectures on downstream association signals. Collectively, these data highlight the substantial genetic heterogeneity of leprosy susceptibility across populations. In the Thai cohort, we mapped a suggestive susceptibility locus to chromosome 18, identifying *SERPINB2*, *SERPINB10*, *HMSD*, and *SERPINB8* as candidate genes for further functional validation. Given their roles in immune regulation, apoptosis, and epithelial barrier maintenance, variants in this region may influence host defense against *M. leprae*. The higher frequency of potentially protective alleles among controls further emphasizes their relevance, supporting the need for functional studies to elucidate the underlying biological mechanisms.

In conclusion, our study identifies a potential candidate region on chromosome 18 associated with leprosy susceptibility in the Thai population, underscoring that host susceptibility is governed by an intricate, population-specific genetic architecture. Given our modest sample size, the absence of an independent replication cohort, and the lack of direct functional evidence validating the roles of *HMSD* or *SERPIN* families, this baseline findings are exploratory and should be interpreted with caution. Nevertheless, this work provides a crucial starting point that opens new avenues for leprosy research in Thailand. Future investigations utilizing larger, independent cohorts, alongside robust functional molecular analyses, are critically needed to validate these preliminary signals and determine their precise biological significance in host immunity.

Reference

- [1] Hastings RC, editor. Leprosy. Edinburgh: Churchill Livingstone; 1985. p. 31-36, 189-194, 322-323.
- [2] McDougall AC, Jopling WH. Handbook of leprosy. 5th ed. Oxford: Heinemann Professional Publishing; 1988.
- [3] World Health Organization. Global leprosy (Hansen disease) update, 2024: Beyond zero cases – what elimination of leprosy really means. Wkly Epidemiol Rec. 2025;100(37):365-384.
- [4] Department of Disease Control, Ministry of Public Health. สถานการณ์โรคเรื้อนในประเทศไทย ปี พ.ศ. 2567 [Internet]. Nonthaburi: Department of Disease Control; 2024 [cited 2025 Aug 28]. Available from: https://rajpracha.ddc.moph.go.th/Rpsi_pages/Nonthaburi/satus_lep.html.
- [5] Siddiqui MR, Meisner S, Tosh K, Balakrishnan K, Ghei S, Fisher SE, et al. A major susceptibility locus for leprosy in India maps to chromosome 10p13. Nat Genet. 2001;27(4):439-41.
- [6] Tosh K, Meisner S, Siddiqui MR, Balakrishnan K, Ghei S, Golding M, et al. A region of chromosome 20 is linked to leprosy susceptibility in a South Indian population. J Infect Dis. 2002;186(8):1190-3.
- [7] Mira MT, Alcaïs A, Van Thuc N, Thai VH, Huong NT, Ba NN, et al. Chromosome 6q25 is linked to susceptibility to leprosy in a Vietnamese population. Nat Genet. 2003;33(3):412-5.
- [8] Zhang FR, Huang W, Chen SM, Sun LD, Liu H, Li Y, et al. Genomewide association study of leprosy. N Engl J Med. 2009;361(27):2609-18.
- [9] Fava VM, Dallmann-Sauer M, Orlova M, Correa-Macedo W, Van Thuc N, Thai VH, et al. Deep resequencing identifies candidate functional genes in leprosy GWAS loci. PLoS Negl Trop Dis. 2021;15(12):e0010029.

- [10] Gilchrist JJ, Auckland K, Parks T, Mentzer AJ, Goldblatt L, Naranbhai V, et al. Genome-wide association study of leprosy in Malawi and Mali. *PLoS Pathog.* 2022;18(9):e1010312.
- [11] Chaichoompu K, Abegaz F, Tongsimma S, Shaw PJ, Sakuntabhai A, Pereira L, et al. IPCAPS: an R package for iterative pruning to capture population structure. *Source Code Biol Med.* 2019;14:2.
- [12] Lippert C, Listgarten J, Liu Y, Kadie CM, Davidson RI, Heckerman D. FaST linear mixed models for genome-wide association studies. *Nature Methods.* 2011;8(10):833-5.
- [13] Boughton AP, Welch RP, Flickinger M, VandeHaar P, Taliun D, Abecasis GR, et al. LocusZoom.js: interactive and embeddable visualization of genetic association study results. *Bioinformatics.* 2021 Sep;37(18):3017-3018.
- [14] Leblond J, Laprise MH, Gaudreau S, Grondin F, Kisiel W, Dubois CM. The serpin proteinase inhibitor 8: an endogenous furin inhibitor released from human platelets. *Thromb Haemost.* 2006;95(2):243-52.
- [15] Pigors M, Sarig O, Heinz L, Plagnol V, Fischer J, Mohamad J, et al. Loss-of-Function Mutations in SERPINB8 Linked to Exfoliative Ichthyosis with Impaired Mechanical Stability of Intercellular Adhesions. *Am J Hum Genet.* 2016;99(2):430-6.
- [16] Przygodzka P, Ramstedt B, Tengeli T, Larsson G, Wilczynska M. Bomapin is a redox-sensitive nuclear serpin that affects responsiveness of myeloid progenitor cells to growth environment. *BMC Cell Biol.* 2010;11:30.
- [17] Kawase T, Akatsuka Y, Torikai H, Morishima S, Oka A, Tsujimura A, et al. Alternative splicing due to an intronic SNP in HMSD generates a novel minor histocompatibility antigen. *Blood.* 2007;110(3):1055-63.
- [18] Medcalf RL, Stasinopoulos SJ. The undecided serpin. The ins and outs of plasminogen activator inhibitor type 2. *Febs j.* 2005;272(19):4858-67.
- [19] Kruithof EK, Baker MS, Bunn CL. Biological and clinical aspects of plasminogen activator inhibitor type 2. *Blood.* 1995;86(11):4007-24.
- [20] Perazzoli S, Fiuza MF, De Moraes PC, Heck R, Vianna FS, Bonamigo RR. Prevalence of genetic variants in SERPINB2 and PKNOX1 genes in erythema nodosum leprosum patients from southern Brazil. *J Infect Dev Ctries.* 2025;19(7):1083-8.