

Association of polymorphisms in SLC gene family with risk of leprosy and malformation in Chinese familial and sporadic patients

Wang Xiaohua

1,JINAN University,2Dermatology hospital of southern medical university <https://orcid.org/0000-0001-7827-5491>

Junpu Mei

Dermatology hospital,southern medical university

Shouhong Que

Henan provincial xinyang central hospital

Ming Li

Dermatology hospital,southern medical university

Xueli wu

Dermatology hospital ,southern medical university

Junxin Huang

shenzhen center for chronic disease

Daocheng Zheng

Dermatology hospital, southern medical university

Bin Yang (✉ yangbin101@hotmail.com)

Research article

Keywords: Leprosy, solute carrier (SLC) super-family, Genes

Posted Date: June 5th, 2019

DOI: <https://doi.org/10.21203/rs.2.10132/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: Leprosy is a chronic infectious disease caused by *Mycobacterium leprae*, and usually affects the skin surface and peripheral nerves, and may lead to deformities if not be healed. Genes in solute carrier (SLC) super-family encode a series of the largest transport-proteins, and are reported to be associated with several connective tissue diseases. **Methods:** Here we performed a multiple-stage study to investigate the polymorphisms of SLC gene family in familial leprosy patients using whole exome sequencing data and sporadic leprosy patients using TaqMan custom genotyping assays, in China. **Results:** We found that, rs202071873 in SLC35E1 is associated with leprosy, rs3731865 in SLC11A1 and rs117820608 in SLC35E2B are associated with the degree of malformation in leprosy. **Conclusions:** This report of the polymorphisms in SLC gene family supports the view of a important role of human genetics in the natural development of the disease. **Key word** Leprosy; solute carrier (SLC) super-family; Genes

Background

Leprosy is a chronic infectious disease which mainly caused by directly contacting with *Mycobacterium leprae*, the common lesion regions are located on the skin and in the peripheral nerves. The clinical manifestations of leprosy are numbness skin damage, bulky nerve and even extremity handicapped. Due to the differences of immunity, patients with leprosy can exhibit various clinical symptoms. The leprosy on patients with stronger immune system are often developed to tuberculoid form (TT), consists of well-defined lesions, few bacilli, and vigorous cell-mediated immunity (CMI). In contrast to that, patients with weak immunity or immunity defect, tend to be developed to lepromatous leprosy (LL), which presents as many skin lesions with uncontrolled proliferation of leprosy bacilli and inefficient CMI. In addition, borderline leprosy within these two boundaries, can be classified as three subtypes: borderline lepromatous (BL), borderline tuberculoid (BT)^[1], and borderline (BB). The distinction of diverse symptoms are because of different degrees of defects in T cells, B cells and macrophages in leprosy patients, especially macrophages^[2-4]. Besides, there is strong evidence that different genetic basis of host is the primary cause of susceptibility subtypes^[5].

Genes in solute carrier (SLC) super-family encode a series of largest transport-proteins, defect of different SLC gene could cause different diseases. For instance, some Bartter-like syndromes result from defects of *SLC12A1*^[11], mutations in *SLC40A1*^[17] can lead to hemochromatosis type 4 (HFE4), and SNPs (rs26722 and rs16891982) in *SLC24A5*^[18] was found to be involved in reduced melanogenesis and thereby making the skin lighter. Specially, *SLC11A1*, a major candidate gene of leprosy susceptibility^[6], was found to highly express in mononuclear macrophages^[7] and restrictively express in myeloid cells^[7], neuronal cells^[8] and human exocrine tissues, including prepituitary gland, adrenal medulla and pancreatic island.

The 35th member of SLC super family (SLC35) has 6 sub-types (A-F)^[9, 10], and 31 SLC35 genes were found in human genome. SLC35 genes, which transport multiple types of monosaccharide nucleotide in the transporter cells, are essential genes in the process of glycosylation. In SLC35 defect research, SLC35 genes were found to participate in a variety of connective tissue diseases and genetic diseases^[19, 20, 21], and played an important role in the occurrence and metastasis of tumor in many pathological processes^[22, 23].

In this study, we applied a two-tiered procedure (**Figure 1**) that first sequences the whole exome of 29 individuals (19 leprosy patients and 10 unaffected control) from 8 families, and then selects 17 mutations of SLC gene family observed in the sequencing data and 33 candidate mutations based on susceptibility and resistance in Mendel-model, respectively. To determine whether the selected mutations were causal or susceptible to leprosy, we examined these mutations in 191 additional leprosy patients and 100 unaffected healthy controls to estimate the association between such mutations and leprosy patients and their multiple phenotype. Meanwhile, considering the low frequency of some polymorphic sites, gene-based burden test was also carried out to estimate the storage effect of such mutations.

Methods

Familial data

Eight families of Han Chinese were recruited from Guangdong province for exome sequencing (**Figure 2**). Parents and siblings of the proband in family I are unaffected, and the proband has 5 children, 1 of 2 sons and 1 of 3 daughters are suffered from leprosy, blood was obtained from the proband, his father and his affected son. Proband of family II has 4 sons and 3 daughters, one of his sons and one of his daughters are affected, in addition to the 3 affected patients, an unaffected son was also selected for sequencing. In family III, blood of proband, her affected daughter and, her unaffected son and granddaughter were enrolled. For family IV and V, there are only two patients and are all selected. When comes to family VI, both parents of the proband are affected, and the proband also has an affected brother and an unaffected brother, an unaffected sister. All members of this family were enrolled. In contrast, in family VII, only one of his unaffected parents and all his siblings were sequenced. And for family VIII, blood was obtained from the proband and her affected sister and daughter, unaffected mother and brother. Totally, 19 leprosy patients and 10 unaffected samples were enrolled. In the replication stage, a cohort of 196 patients and 100 healthy controls were collected, and clinical information was recorded in the process of diagnosis and follow-up telephone visit (**Supplementary Table 1**). All these subjects were recruited by Guangdong Dermatology Hospital from 2008 to 2014, and written informed consent was also collected from all subjects.

Exome sequencing and data analysis

Genomic DNA of each individuals was extracted from the peripheral blood with using the QIAamp DNA Blood Mini kit (QIAGEN GmbH, Hilden, Germany), under a strict quarantine measures to avoid the spread of leprosy. Pruned DNA was hybridized by the Agilent SureSelect Human All Exon V4 array to capture the coding sequencing of the DNA. All samples from 8 leprosy families were performed whole exome sequencing with 100bp pair-end reads using Illumina HiSeq

We captured the exonic regions of 29 samples from the 8 leprosy families by Agilent technology followed by sequencing on Illumina HiSeq 2000 platform (**Methods**). An average of 15.0 Gb paired-end short reads of 90bp was produced for each individual. After aligning to the human reference genome (hg19) and marking PCR duplication reads, we obtained ~143.9X fold coverage of the Exome sequencing array. The exome sequencing covered 99.67% of the target regions (genome segment of total exome sequencing array), furthermore, 98.70% of the target regions were sequenced by at least 10 times (**Supplementary Table 2**). We identified 20,924 coding SNVs and 589 coding InDels (Insertions and Deletions >=1bp) per sample on average (**Methods**), of these variants, 98.26% SNVs and 84.75% InDels (**Supplementary Table 3**) were found in dbSNP v142.

To validate their relationship, IBD was calculated of each family based on high confident SNPs (**Methods**) in autosomal chromosomes, and no family failed the test. Moreover, the gender for each individual was confirmed by examining the Y chromosome sequencing coverage.

To narrow down the variants set and include only potentially pathogenic variants (**Methods**), we focused only on nonsynonymous variants, splice acceptor and donor site mutations, and coding frameshift InDels. We also removed variants with >1% population frequency in east Asian in 1000 genomes project or ExAC, gnomAD. To reduce the effect of complicated genome sequences, variants located in segmental duplication region or simple repeat regions were also removed.

In summary, 39 putative events (occurred in 33 loci, **Table 1**) were extracted and 34 (87.2%) of them were validated in the corresponding family members, utilizing Sequenom MassARRAY System (**Supplementary Table 4a**). Specially, p.T1421A (CUL7) was found in both family 4 and 8, and p.V2169G (ARFGEF3) was presented in family 1, 7 and 8, simultaneously.

Variant validation and Replication

To deeply investigate the pathology of leprosy, in addition to 33 selected putative mutations above, 17 mutations (16 SNPs and 1 InDel, **Supplementary Table 4b**) that comes from SLC gene family were also genotyped by Sequenom MassARRAY System, in a larger Chinese leprosy cohort (including 15 exome sequencing samples that can be collected repeatedly, and 177 additional collected samples) and 100 healthy control. Among of these, 5 mutations that were not repeated in sequencing samples were discarded in following analysis.

We found a candidate susceptible SNP rs202071873 in SLC35E1, but no SNPs were found to satisfy the significant cutoff ($5.0E-08$) for single loci association analysis under various types of test, including chi-square (**Supplementary Table 5a**), fisher's test (**Supplementary Table 5b**) and logistic regression test (**Supplementary Table 5c**). To further our understanding, we used linear model (general allele model) to assess the effect of allele on the degree of malformation, nerve damage, response, and skin damage. All clinical information were transformed into quantity to simply the test (**Method**). And we found rs3731865 ($p=0.00822$, $\beta=-0.2923$) in SLC11A1, and rs117820608 ($p=0.033$, $\beta=0.3254$) in SLC35E2B to be suggestively associated with the degree of malformation. In addition, patients with genotype A/G at rs117820608 are tended to be more likely to cause disability ($p=0.0268$, $OR=3.939$), and no significant difference was found in the genotype of rs3731865. Meanwhile, in the gene-based burden test, SLC35E1 was showed to be associated with leprosy (**Table 2**).

Protein-protein interaction

For all 40 genes selected in our study, protein-protein interaction networks were predicted by DAPPLE (Disease Association Protein-Protein Link Evaluator). Direct protein-protein interaction was found between MAST2 and TEAD3, and SLC35E1-SLC35E2-SLC35E2B were found to be in an isolated network (**Supplementary Figure 1**).

Validation using Immunohistochemistry

To deeper investigate the role of SLC genes in the development of leprosy, we enrolled another 88 FFPE samples (including 28 healing samples, 12 healthy control, 16 multibacillary patients, 13 paucibacillary patients and 19 untreated patients) and examined the protein expression level of SLC35E1 and SLC35E2B in normal skin tissue of 12 healthy controls and skin lesions from different types of leprosy patients (**Figure 3**). And immunoreactive score (IRS, the product of dyeing strength and dyeing area) was used to estimate the protein expression level.

We also genotyped the polymorphic loci of SLC35E1 and SLC35E2B in these samples, but unfortunately only a few proportion of samples were succeed. Association analysis between the genotype and IRS was processed, but no significant result was found due to the limited genotype information (results not shown).

But for the IRS itself, we observed that, the protein expression of SLC35E2B varied prominently in different tissues of various samples (**Figure 4**). For SLC35E2B, all four types of tissues were observed to be obviously lower-expression in epidermis (**Table 3, Supplementary Table 8**). And in dermis tissues, there was no significant protein expression difference between patients and healthy samples, except for tissues from patients after treatment ($P=0.005$, wilcox rank sum test). Moreover, the protein expression level of SLC35E2B in dermis of un-treated patients were observed to be higher than that in healing patients ($P=0.0445$). Meanwhile, the difference of protein expression of SLC35E1 in diverse tissues in neither epidermis nor dermis were not obvious.

Discussion

Genetic factors have been recognized as an important contributor of leprosy after infected with leprosy bacillus. Genome wide association studies (GWAS) have identified several susceptibility common variants^[24, 25, 26] and lncRNA^[27] associated with leprosy. Meanwhile, immune response genes have been proved to be associated with pathogenesis of different forms of leprosy^[28]. However, these factors could only explain a small proportion of disease heritability.

In this study, we performed a two-stage study to investigate SLC gene polymorphisms in familial and sporadic leprosy patients. First, we performed exome sequencing in 8 leprosy families, and extracted rare inherited mutations based on the infection and immune response to the *Mycobacterium leprae*,

respectively. And meanwhile, we found a large amount of polymorphisms in SLC genes in patients of these 8 families, did not find any known causal rare mutations. These results suggested that, infection and fall ill could be caused by numbers of genes, and were not limited to rare mutations. Second, validated rare mutations and some specific common mutations of SLC gene that found in leprosy family, were replicated in larger cohort. Finally, 3 SLC genes (SLC11A1, SLC35E1, and SLC35E2B) were found to be suggestively associated with leprosy at different ways. Meanwhile, indirect protein-protein interactions were predicted between SLC35E1 and SLC35E2B.

SLC11A1 which translates a multi-pass membrane protein, is a member of the solute carrier family 11 family. **Mutations** in this gene have been reported to be associated with many disease^[32], including tuberculosis^[33, 34], leprosy^[35], **rheumatoid arthritis**^[34, 36] and **Crohn's disease**^[37, 38, 39]. Association analysis result of our data also found SLC11A1 to be significant associated with leprosy.

MAST2 and TEAD3 are protein coding genes of human genome, and implicate direct protein-protein interaction, but no additional mutations were found in expanded leprosy samples, There might be some interesting findings under further studies with larger sample size.

Conclusions

SLC35E1 and SLC35E2B are homologous genes that predicted to be located on membrane, and both belong to the solute carrier family 35 family, and thus far, no direct evidence showed these two genes are specifically linked to any disease. For SLC35E1 and SLC35E2B, we performed Immunohistochemistry in the skin lesions of multibacillary and paucibacillary, before and after treated, and skin tissues of healthy control, respectively. Our data showed that, protein expression level of SLC35E1 in diverse skin tissues were no-difference neither in epidermis nor dermis. In contrast, protein expression level of SLC35E2B in skin tissues selected from various types leprosy patients were significantly lower than that of the healthy controls, especially in epidermis. Meanwhile, in dermis, protein expression level of SLC35E2B of tissues collected from healing leprosy patients were lower than that of both healthy controls and un-treated leprosy patients. But there was no evidence in our results that can link the gene polymorphisms to the protein expression level. This may need more samples or should be examined in further study.

Moreover, our data show that, the pathology and genetics factor of leprosy varies along with diverse types of leprosy. And we found that the protein expression of SLC35E2B in epidermis was correlated with bacterial infection, which suggested that, SLC35E2B might play an important role in the occurrence of leprosy. Meanwhile, the protein expression difference between healing patients, un-treated and healthy controls suggested that, SLC35E2B might also play an important role in the treatment of leprosy, and can acted as an candidate marker to evaluate the curative effect of leprosy.

Totally, polymorphisms of SLC gene family present important potency in the development and treatment of leprosy. However, the association between polymorphisms the protein expression level is unclear, and deeper studies in larger cohort in different sub-types and functional studies are essential to verify our findings.

Abbreviations

SLC:solute carrier; CMI:cell-mediated immunity; TT:tuberculoid form; LL :epromatous leprosy ; BL: borderline lepromatous ; BB: borderline borderline; BT: borderline tuberculoid; GATK: Genome Analysis Toolkit; GWAS: Genome wide association studies

Declarations

Written informed consent was obtained from the patient for publication of this Case Report and any accompanying images. A copy of the written consent is available for review by the Editor of this journal.

Acknowledgements

We would like to thank the staffs of Dermatology Hospital, Southern Medical University and the participants in Guangdong Province who made contributions to this study.

Funding

Project supported by the Natural Science Foundation of Guangdong Province, (Grant No 2018A0303130318.)

Availability of data and materials

The datasets used and /or analysed during the current study available from the corresponding author on reasonable request.

Authors' contributions

XHW, JP and B Y Have designed this research . SHQ, ML, XLW and HJX have implemented this research and collect research data. XHW and JP have analyzed the research data and write this manuscript. B Y and DHZ have given technical support and administrated the research. All authors have read and approved the manuscript.

Ethical approval and consent to participate

Ethics approval was obtained from the Ethics Review Committee of Dermatology Hospital, Southern Medical University. All participants provided written informed consent in this research.

Consent for publication

Not Applicable.

Competing Interests

The authors declare no conflict of interest.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Declarations

Written informed consent was obtained from the patient for publication of this Case Report and any accompanying images. A copy of the written consent is available for review by the Editor of this journal.

References

- [1] D. S. Ridley and W. H. Jopling, "Classification of leprosy according to immunity. A five-group system," *International Journal of Leprosy and Other Mycobacterial Diseases*, vol. 34, no.3, pp. 255–273, 1966.
- [2] Zhang F, et al. Identification of two new loci at IL23R and RAB32 that influence susceptibility to leprosy. *Nat Genet*, 2011, 43:1247-51
- [3] Zhang FR, et al. Genomewide association study of leprosy. *N Engl J Med*, 2009,361: 2609-18
- [4] Geluk A, van Meijgaarden KE, Wilson L, et al. Longitudinal immune responses and gene expression profiles in type 1 leprosy reactions. *J Clin Immunol*, 2014,34(2):245-255.
- [5] G. A. Hansen, "Bacillus leprae," *Norsk Mag Laegevidenskaben*, vol. 9, pp. 1–21, 1874.
- [6] Casanova J-L, Abel L. Genetic dissection of immunity to mycobacteria: The human model. *Annu Rev Immunol*, 2002, 20:581-620.
- [7] Britton WJ, D.N.Lockwood. Leprosy. *Lancet*, 2004, 363:1209-1219.
- [8] Ali PM, Ramanujam K. Leprosy in Tw ins. *Int J Lepr*, 1966, 34: 404-407.
- [9] Fernando SL, Britton WJ. Genetic susceptibility to mycobacterial disease in humans. *Immunol Cell Biol*, 2006,84: 125-37.
- [10] Ananthanarayanan M, Ng OC, Boyer JL, et al. Characterization of cloned rat liver Nabile acid cotransporter using peptide and fusion protein antibodies. *Am J Physiol*, 1994, 267(4Pt 1): G637-G643.
- [11] Seyberth HW, Schlingmann KP. Bartter- and Gitelman-like syndromes: salt-losing tubulopathies with loop or DCT defects. *Pediatr Nephrol*. 2011;26(10):1789-802.
- [12] Chen N, Tranebjærg L, Rendtorff ND, Schrijver I. Mutation analysis of SLC26A4 for Pendred syndrome and nonsyndromic hearing loss by high-resolution melting. *J Mol Diagn*. 2011;13(4):416-26.
- [13] Sperandio MP, Andria G, Sebastio G. Lysinuric protein intolerance: update and extended mutation analysis of the SLC7A7 gene. *Hum Mutat*. 2008;29(1):14-21.
- [14] Qiu ZQ, Lu CX, Wang W, Qiu JJ, Wei M. [Mutation in the SLC37A4 gene of glycogen storage disease type Ib in 15 families of the mainland of China]. *Zhonghua Er Ke Za Zhi*. 2011;49(3):203-8.
- [15] Choi R, Park HD, Ko JM, et al. Novel SLC37A4 Mutations in Korean Patients With Glycogen Storage Disease Ib. *Ann Lab Med*. 2017;37(3):261-266.
- [16] Carlin MP, Scherrer DZ, De tommaso AM, Bertuzzo CS, Steiner CE. Determining mutations in G6PC and SLC37A4 genes in a sample of Brazilian patients with glycogen storage disease types Ia and Ib. *Genet Mol Biol*. 2013;36(4):502-6.
- [17] Chen SR, Yang LQ, Chong YT, et al. Novel gain of function mutation in the SLC40A1 gene associated with hereditary haemochromatosis type 4. *Intern Med J*. 2015;45(6):672-6.
- [18] Cook AL, Chen W, Thurber AE, Smit DJ, Smith AG, Bladen TG, et al. Analysis of cultured human melanocytes based on polymorphisms within SLC24A2/MATP, SLC24A2/NCKX5, and OCA2/P loci. *J Invest Dermatol*. 2008; 129(2):392-405
- [19] Song Z. Roles of the nucleotide sugar transporters (SLC35 family) in health and disease. *Mol Aspects Med*. 2013;34(2-3):590-600.
- [20] Sarangi A, Bupp K, Roth MJ. Identification of a retroviral receptor used by an envelope protein derived by peptide library screening. *Proc Natl Acad Sci USA*. 2007;104(26):11032-7.

- [21] Schlessinger A, Matsson P, Shima JE, et al. Comparison of human solute carriers. *Protein Sci.* 2010;19(3):412-28.
- [22] Orentas RJ, Yang JJ, Wen X, Wei JS, Mackall CL, Khan J. Identification of cell surface proteins as potential immunotherapy targets in 12 pediatric cancers. *Front Oncol.* 2012;2:194.
- [23] Ishida N, Kawakita M. Molecular physiology and pathology of the nucleotide sugar transporter family (SLC35). *Pflugers Arch.* 2004;447(5):768-75.
- [24] Mira MT, Alcaïs A, Nguyen VT, et al. Susceptibility to leprosy is associated with PARK2 and PACRG. *Nature.* 2004;427(6975):636-40.
- [25] Wang Z, Sun Y, Fu X, et al. A large-scale genome-wide association and meta-analysis identified four novel susceptibility loci for leprosy. *Nat Commun.* 2016;7:13760.
- [26] Liu H, Irwanto A, Tian H, et al. Identification of IL18RAP/IL18R1 and IL12B as leprosy risk genes demonstrates shared pathogenesis between inflammation and infectious diseases. *Am J Hum Genet.* 2012;91(5):935-41.
- [27] Fava VM, Manry J, Cobat A, et al. A genome wide association study identifies a lncRNA as risk factor for pathological inflammatory responses in leprosy. *PLoS Genet.* 2017;13(2):e1006637.
- [28] Jarduli LR, Sell AM, Reis PG, et al. Role of HLA, KIR, MICA, and cytokines genes in leprosy. *Biomed Res Int.* 2013;2013:989837.
- [29] Koriyama H, Nakagami H, Katsuya T, et al. Identification of evidence suggestive of an association with peripheral arterial disease at the OSBPL10 locus by genome-wide investigation in the Japanese population. *J Atheroscler Thromb.* 2010;17(10):1054-62.
- [30] Li H, Wetten S, Li L, et al. Candidate single-nucleotide polymorphisms from a genomewide association study of Alzheimer disease. *Arch Neurol.* 2008;65(1):45-53.
- [31] Sierra B, Triska P, Soares P, et al. OSBPL10, RXRA and lipid metabolism confer African-ancestry protection against dengue haemorrhagic fever in admixed Cubans. *PLoS Pathog.* 2017;13(2):e1006220.
- [32] Blackwell JM, Goswami T, Evans CA, et al. SLC11A1 (formerly NRAMP1) and disease resistance. *Cell Microbiol.* 2001;3(12):773-84.
- [33] Meilang Q, Zhang Y, Zhang J, et al. Polymorphisms in the SLC11A1 gene and tuberculosis risk: a meta-analysis update. *Int J Tuberc Lung Dis.* 2012;16(4):437-46.
- [34] Ates O, Dalyan L, Müsellim B, et al. NRAMP1 (SLC11A1) gene polymorphisms that correlate with autoimmune versus infectious disease susceptibility in tuberculosis and rheumatoid arthritis. *Int J Immunogenet.* 2009;36(1):15-9.
- [35] Brochado MJ, Gatti MF, Zago MA, Roselino AM. Association of the solute carrier family 11 member 1 gene polymorphisms with susceptibility to leprosy in a Brazilian sample. *Mem Inst Oswaldo Cruz.* 2016;111(2):101-5.
- [36] Niño-moreno P, Turrubiarres-martínez E, Ocegüera-maldonado B, et al. The Role of NRAMP1/SLC11A1 Gene Variant D543N (1730G/A) in the Genetic Susceptibility to Develop Rheumatoid Arthritis in the Mexican Mestizo population. *Rev Invest Clin.* 2017;69(1):5-10.
- [37] Chermesh I, Azriel A, Alter-koltunoff M, Eliakim R, Karban A, Levi BZ. Crohn's disease and SLC11A1 promoter polymorphism. *Dig Dis Sci.* 2007;52(7):1632-5.
- [38] Sechi LA, Gazouli M, Sieswerda LE, et al. Relationship between Crohn's disease, infection with *Mycobacterium avium* subspecies *paratuberculosis* and SLC11A1 gene polymorphisms in Sardinian patients. *World J Gastroenterol.* 2006;12(44):7161-4.
- [39] Zheng W, Rosenstiel P, Huse K, et al. Evaluation of AGR2 and AGR3 as candidate genes for inflammatory bowel disease. *Genes Immun.* 2006;7(1):11-8.

Tables

Chr	Pos	dbSNP id	refalt	Gene	frequency in gnomAD	carried families	function	phastCons (Y/N)	validated (Y/N)
chr1	17256835	.	C A	CROCC	1.66e-05	2	missense	Y	Y
chr1	17298098	rs368707716	C T	CROCC	0.0002	8	missense	Y	Y
chr1	46500794	.	C T	MAST2	2.46e-05	4	stopgain	Y	Y
chr1	52867106	rs182707743	G C	ORC1	0.0002	4	missense	Y	Y
<i>chr1</i>	<i>110169450</i>	<i>rs75736595</i>	<i>A C</i>	<i>AMPD2</i>	<i>0.0025</i>	<i>2</i>	<i>missense</i>	<i>Y</i>	<i>N</i>
						<i>7</i>			<i>N</i>
chr1	197070906	rs118010078	C T	ASPM	0.0009	4	missense	Y	Y
chr2	215843674	.	C T	ABCA12	4.065e-06	4	missense	Y	Y
chr3	31839284	rs11129464	A T	OSBPL10	0.0007	5	missense	N	Y
						<i>6</i>			<i>N</i>
chr4	67752	.	C G	ZNF595	3.403e-05	4	missense	N	Y
chr4	79350266	rs183594682	A T	FRAS1	0.0001	5	missense	Y	Y
chr4	114067107	rs118148584	C T	ANK2	0.0008	5	missense	N	Y
chr6	35443196	.	T A	TEAD3	1.218e-05	4	missense	Y	Y
chr6	35443218	.	G C	TEAD3	1.219e-05	4	missense	Y	Y
chr6	43007927	rs191578513	T C	CUL7	0.0009	4	missense	Y	Y
						<i>8</i>			<i>Y</i>
chr6	129475707	rs182958473	G T	LAMA2	0.0002	4	missense	Y	Y
chr6	138657595	rs372977543	T G	ARFGEF3	0.0076	1	missense	Y	Y
						<i>7</i>			<i>Y</i>
						<i>8</i>			<i>Y</i>
chr7	48312432	rs189969712	G T	ABCA13	0.0010	4	missense	N	Y
chr7	150920935	rs138611892	G A	ABCF2	6.952e-05	4	missense	Y	Y
chr9	14737576	.	G A	FREM1	NA	2	missense	N	Y
chr9	95526969	.	C T	BICD2	9.333e-05	5	missense	Y	Y
chr9	139916347	.	A T	ABCA2	3.751e-05	5	missense	Y	Y
chr11	74559452	rs78981281	C T	XRRA1	0.0014	4	missense	N	Y
chr14	74429785	rs187468747	T A	COQ6	0.0010	5	stopgain	N	Y
						<i>6</i>			<i>N</i>
chr15	55790053	rs61733359	T C	DYX1C1	0.0010	5	missense	Y	Y
chr16	682747	rs184792782	G A	WFIKKN1	1.666e-05	5	missense	Y	Y
chr16	2820693	.	C T	SRRM2	5.67e-05	5	missense	N	Y
chr16	15715636	rs35438466	C G	KIAA0430	0.0085	3	missense	Y	Y
<i>chr16</i>	<i>20998712</i>	.	<i>C T</i>	<i>DNAH3</i>	<i>1.218e-05</i>	<i>7</i>	<i>missense</i>	<i>Y</i>	<i>N</i>
chr16	50745233	rs1078327	C T	NOD2	0.0015	5	missense	N	Y
chr17	67013907	rs374862326	G A	ABCA9	0.0002	7	missense	N	Y
chr19	1053438	rs150012110	G A	ABCA7	0.0003	3	missense	Y	Y
chr19	10799893	rs145544698	G A	ILF3	0.0002	4	missense	Y	Y
chrX	118220887	.	G A	KIAA1210	0.0002	5	missense	N	Y

Table 1. Putative causal mutations extracted in multiple families, and italic events were failed in validation.

Association analysis												
loci	gene symbol	function	MAF		A1	A2	chi-square		fisher		logistic	
			case	control			P	OR	P	OR	P	OR
rs202071873	SLC35E1	missense	0.002551	0.02	T	C	0.0282	0.1253	0.04682	0.1253	0.06257	0.1231
QTL(based on malformation)												
loci	gene symbol	function	P	BETA								
rs3731865	SLC11A1	intronic	0.00822	-0.2923								
rs117820608	SLC35E2B	missense	0.03298	0.3254								
rs117820608 GA vs. GG												
degree of malformation	AA	GA	GG	fisher	0 vs. 1&2&3		1 vs. 2					
0	0	4	101	pvalue	0.02681	0.3345						
1	0	4	38	OR	3.939	0.4706						
2	0	7	31									
3	0	0	1									
rs3731865 CC vs. CG+GG												
degree of malformation	CC	CG	GG	fisher	0 vs. 1&2&3		1 vs. 2					
0	11	45	6	pvalue	0.148	0.0924						
1	3	12	5	OR	0.351	Inf						
2	0	18	5									
3	0	0	0									
Burden test												
gene symbol	all(P)	common(P)	rare(P)									
SLC35E1	0.02201362	1	0.02180415									

Table 2. Analysis result of SLC11A1, SLC35E1 and SLC35E2 on association analysis, QTL analysis based on malformation, and burden test based on different variation with different frequency.

SLC35E2B							
epidermis				dermis			
	vs. healthy control	vs. healing	vs. paucibacillary		vs. healthy control	vs. healing	vs. paucibacillary
un-treated	0.0006005	0.1437	-	un-treated	0.06033	0.0445	-
multibacillary	0.02172	-	0.2726	multibacillary	0.06931	-	0.4001
paucibacillary	0.04642	-	-	paucibacillary	0.3176	-	-
healing	0.001268	-	-	healing	0.005682	-	-

SLC35E1							
epidermis				dermis			
	vs. healthy control	vs. healing	vs. paucibacillary		vs. healthy control	vs. healing	vs. paucibacillary
un-treated	1	0.9286	-	un-treated	1	0.7469	-
multibacillary	0.6097	-	0.3175	multibacillary	1	-	0.1394
paucibacillary	0.1288	-	-	paucibacillary	0.4121	-	-
healing	0.986	-	-	healing	0.5698	-	-

Table 3. Wilcoxon rank sum test based on IRS of SLC35E2B and SLC35E1, in various type of skin tissues collected from paucibacillary and multibacillary, before treated and healing patients.

Figures

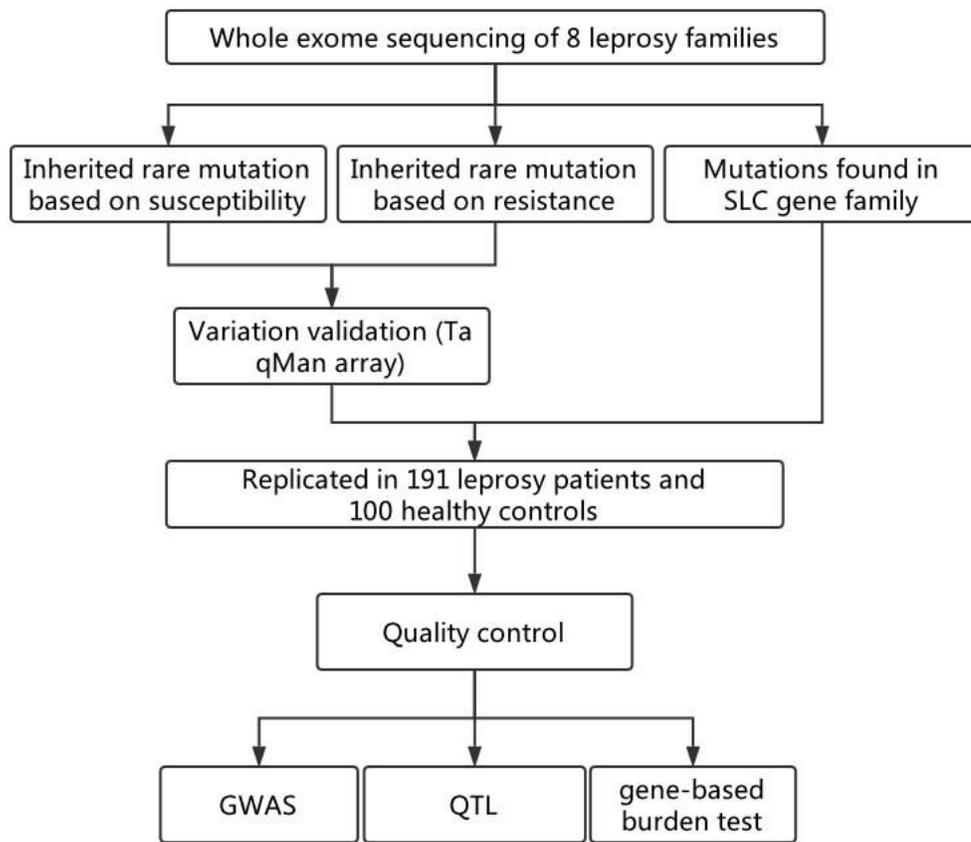


Figure 1

Data analysis workflow for estimating the association between SLC gene polymorphisms and leprosy.

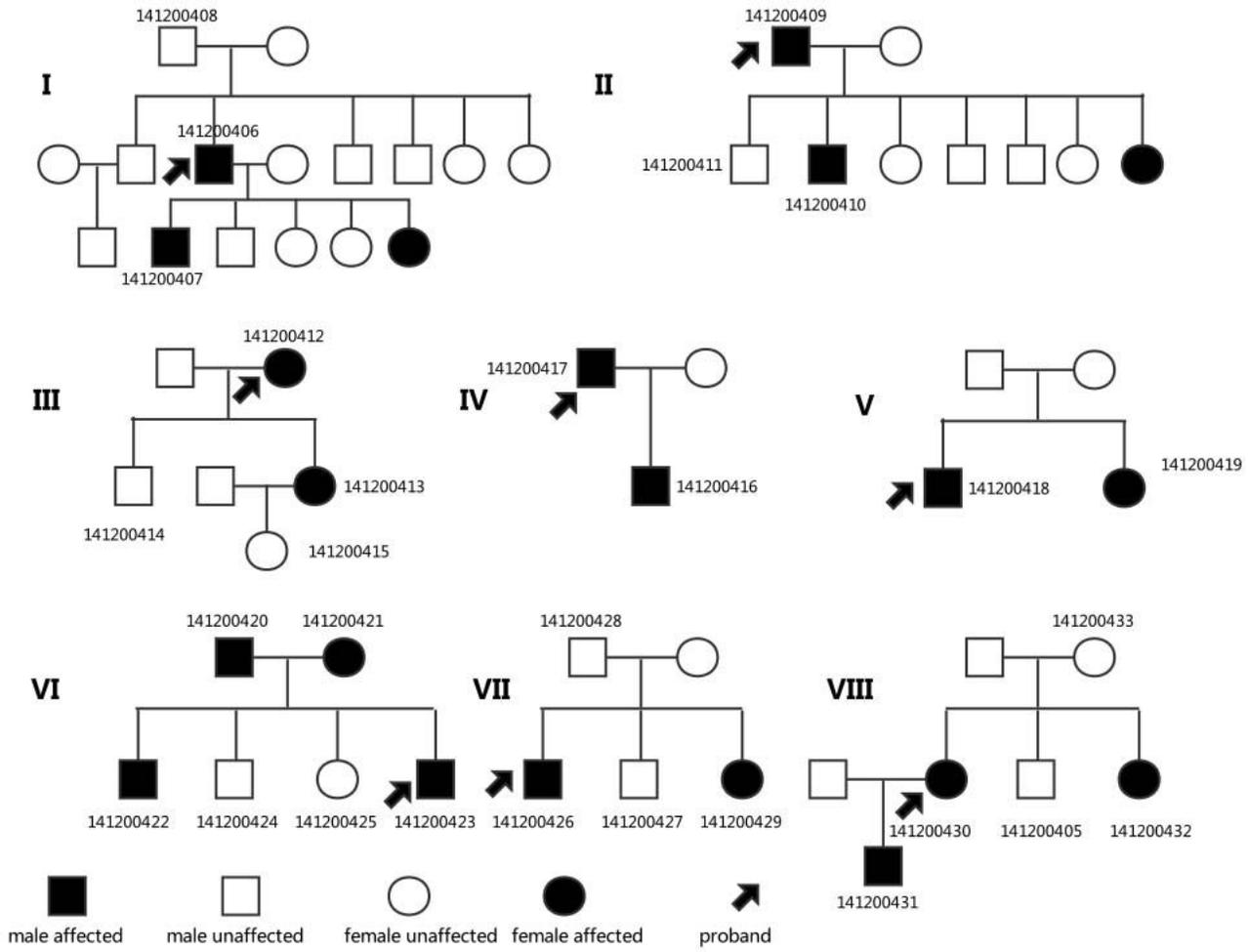


Figure 2

Pedigree of 8 selected leprosy families, samples with number id were performed whole exome sequencing. A square represents a male, and a circle represents a female. Samples filled with black are affected, and the patients that the black arrow points to are proband of the family.

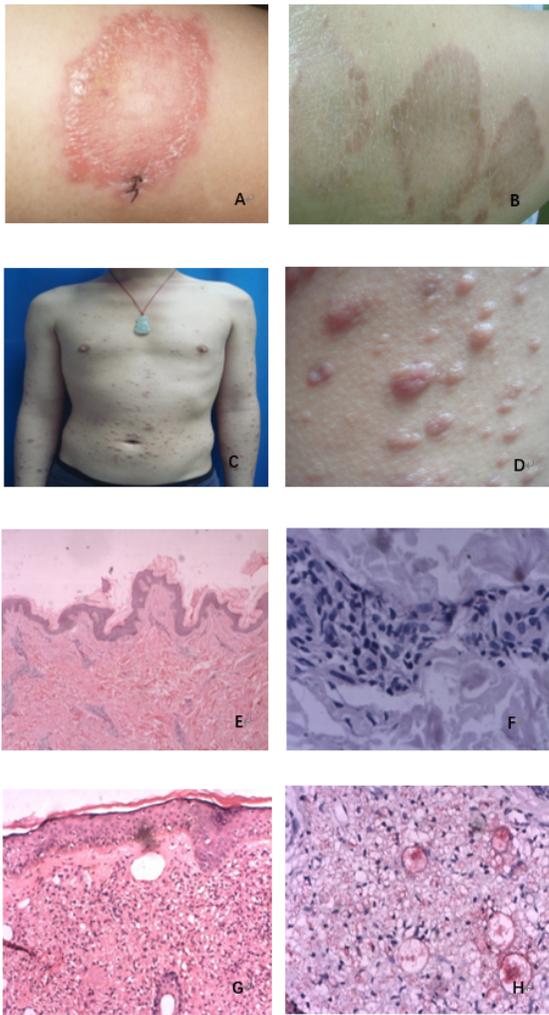


Figure 3

A-D: clinical pictures. E-H: Figures of pathological HE staining and acid fast staining. A) right arm of paucibacillary patients catches annular erythema, with unclear boundary, few scales, and itching. B) left upper limb of paucibacillary patients catches several erythemas with dry annular desquamation. C) trunk and limbs of multibacillary patients catch numbers of nodules. D) the surface of the nodules has a certain atrophy. E) paucibacillary patients(HE*20), epidermis atrophies, and many inflammatory cells gather in dermis. F) paucibacillary patients(HE*100), negative acid fast staining, Epithelioid cells infiltrated into the small cutaneous nerve cells. G) multibacillary patients(HE*20), thinning epidermis, non invasive zone found in the superficial layer of dermis, many inflammatory cells gather in dermis. H) multibacillary patients(HE*100), positive acid fast staining, agminated leprosy bacilli can be found in many foam cells.

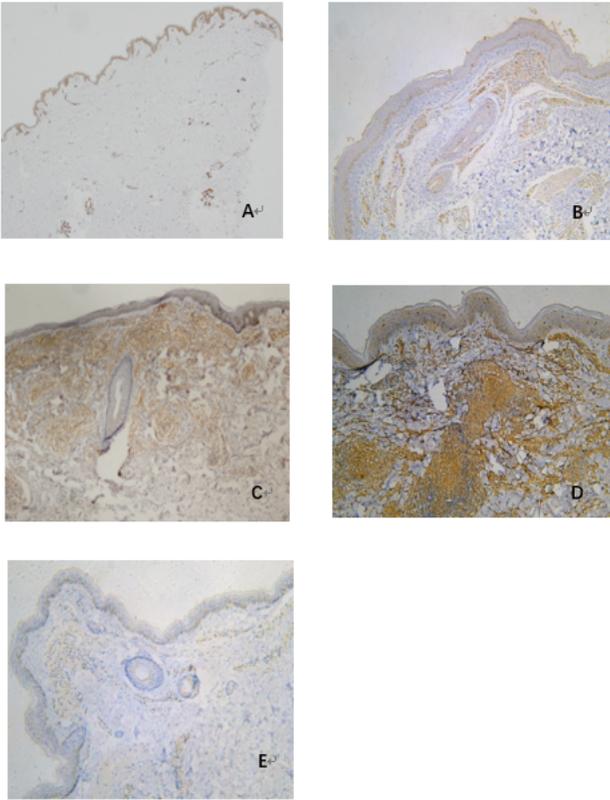


Figure 4

Results of staining immunohistochemistry with magnification 100 times. a) healthy control b) paucibacillary patients. c) multibacillary patients d) un-treated patients e) healing patient.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryTable5b.xlsx](#)
- [SupplementaryTable4a.xlsx](#)
- [SupplementaryTable6b.xlsx](#)
- [SupplementaryTable6a.xlsx](#)
- [SupplementaryTable6c.xlsx](#)
- [SupplementaryTable3.xlsx](#)
- [SupplementaryTable2.xlsx](#)
- [SupplementaryTable4b.xlsx](#)
- [SupplementaryTable1.xlsx](#)
- [SupplementaryTable5a.xlsx](#)
- [SupplementaryFigure1.pdf](#)
- [SupplementaryTable8.xlsx](#)
- [SupplementaryTable7.xlsx](#)
- [SupplementaryTable6d.xlsx](#)
- [SupplementaryTable5c.xlsx](#)