Journal of Dermatological Science 84 (2016) 322–329

Contents lists available at ScienceDirect

Journal of Dermatological Science

journal homepage: www.jdsjournal.com

Fine mapping of the GWAS loci identifies *SLC35D1* and *IL23R* as potential risk genes for leprosy

Guo-Dong Li[a,d,1], Dong Wang[a,1], Deng-Feng Zhang[a], Qun Xiang[a,d], Jia-Qi Feng[a,b], Xiao-An Li[c], Yu-Ye Li[b], Yong-Gang Yao[a,d,*]

[a] *Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences & Yunnan Province, Kunming Institute of Zoology, Kunming, Yunnan 650223, China*
[b] *Department of Dermatology, The First Affiliated Hospital of Kunming Medical University, Kunming, Yunnan 650032, China*
[c] *Yuxi City Center for Disease Control and Prevention, Yuxi, Yunnan 653100, China*
[d] *Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming, Yunnan 650204, China*

ARTICLE INFO

Article history:
Received 26 July 2016
Received in revised form 23 September 2016
Accepted 28 September 2016

Keywords:
Leprosy
SLC35D1
IL23R
GWAS
Susceptibility
eQTL

ABSTRACT

Background: Previous genome-wide association study (GWAS) identified two new leprosy associated loci (1p31.3 [rs3762318] and 6q24.3 [rs2275606]). However, there were insufficient validations in independent populations.
Objective: To validate the association and to map the potentially causal variants/genes underlying the association between the confirmed GWAS hit and leprosy.
Methods: We genotyped 10 variants in the regions encompassing the two loci in 1110 Han Chinese subjects with and without leprosy, followed by expression quantitative trait loci (eQTL), mRNA expression profiling, and network analysis. We further sequenced the exon region of four genes that were located in the confirmed GWAS hit region in 80 leprosy patients and 99 individuals without leprosy.
Results: We validated the positive association of rs3762318 with multibacillary leprosy ($P = 7.5 \times 10^{-4}$), whereas the association of rs2275606 could not be validated. eQTL analysis showed that both the GWAS locus rs3762318 and one surrounding positively associated SNP rs2144658 ($P = 1.8 \times 10^{-3}$) significantly affected the mRNA expression of a nearby gene *SLC35D1*, which might be involved in metabolism. Moreover, *SLC35D1* was differentially expressed in skin tissues of leprosy patients, and the differential expression pattern was consistent among leprosy subtypes. Rare damaging missense variants in *IL23R* were significantly enriched in leprosy patients.
Conclusion: Our results supported the positive association between the GWAS reported rs3762318 and leprosy, and *SLC35D1* and *IL23R* might be the causal genes.

1. Introduction

Leprosy (also called Hansen's disease) is a chronic infectious disease caused by *Mycobacterium leprae* (*M. leprae*). This pathogenic bacterium mainly affects the skin, peripheral nerves and other areas including eyes, nose and testicles [1]. Although the epidemic of leprosy is now under effective control, new cases of leprosy are detected every year, with more than 20,000 cases worldwide in the past 8 years [2]. Evidently, leprosy remains an important public health problem.

Population-based studies had suggested a positive association between leprosy and host genetic background [3–5]. Many susceptible genes of leprosy have been identified in the past years, such as *TLR1*, *TLR2* [6], *IFNG* [7], *CFH* [8], and *NOD2* [9]. In a recent genome-wide association study (GWAS), Zhang et al. [10] reported two genome-wide significant leprosy susceptible loci, with index single nucleotide polymorphisms (SNPs) of rs3762318 (chromosome 1p31.3) and rs2275606 (chromosome 6q24.3). However, there was no validation of this result in other Han Chinese populations.

In this study, we aimed to validate the association and to fine-map the potential causal variants/genes. We genotyped 10 SNPs covering the two GWAS loci in 1110 subjects, followed by expression quantitative trait loci (eQTL) and differential expression analyses [11–13]. We validated the association of rs3762318 with

* Corresponding author at: Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan, 650223, China.
E-mail addresses: yaoyg@mail.kiz.ac.cn, ygyaozh@gmail.com (Y.-G. Yao).
[1] These authors contributed equally to this study.

http://dx.doi.org/10.1016/j.jdermsci.2016.09.018
0923-1811/

leprosy, and identified *SLC35D1* as the potentially causal gene. Exome sequencing of four genes that were located in the validated GWAS hit region showed that *IL23R* had an enrichment of missense variants with potential pathogenicity in leprosy patients.

2. Materials and methods

2.1. Subjects

A total of 527 leprosy patients (279 multibacillary leprosy [MB] and 248 paucibacillary leprosy [PB]) and 583 healthy controls were collected from the Yuxi Prefecture, Yunnan Province, Southwest China. These subjects had been described in our previous studies [8,14–18]. Leprosy was diagnosed on the basis of clinical and histopathological features and/or bacteriological index (if available). The healthy controls had no history of leprosy, HIV and tuberculosis. Eighty leprosy patients were collected from the Wenshan Prefecture, Yunnan Province, and were sequenced for the exons of four genes that were located in the GWAS hit region, together with 99 individuals without leprosy. Informed consents conforming to the tenets of the Declaration of Helsinki were obtained from all participants. This study was approved by the institutional review board of the Kunming Institute of Zoology.

2.2. SNP selection, genotyping and sequencing

Genomic DNA samples from the patients and controls were extracted from whole blood by using the AxyPrep™ Blood Genomic DNA Miniprep Kit (Axygen, USA). The two SNPs (rs3762318 and rs2275606) identified by the previous GWAS [10] were selected. To fine-map the potential causal variants/genes, another eight SNPs (rs2144658, rs17129664, rs6670134, rs10789224, rs1321157, rs7539625, rs11465817, rs10889677) were selected according to the SNP information in public database (NCBI dbSNP, http://www.ncbi.nlm.nih.gov/projects/SNP/; HapMap, http://hapmap.ncbi.nlm.nih.gov/, phase 3, CHB dataset [19]) under a rational of minor allele frequency (MAF) >5%. Among these selected SNPs, seven (rs2144658, rs6670134, rs17129664, rs1321157, rs7539625, rs11465817 and rs10889677) were tag SNPs, one (rs10789224) was previously reported to be associated with Crohn's disease [20]. The detailed information of each SNP was included in Table S1. All these 10 SNPs were genotyped by using the SNaPshot assay (SNaPshot® Multiplex System, Life Technologies, California, USA) following the same procedure described in our previous studies [14–16,18]. Primers for genotyping were listed in Table S2. The genotyping data were read by two researchers (G.-D.L. and Q.X.) independently by using the GeneMarker software [21].

Table 1
The allele and genotype frequencies of 10 SNPs in 527 leprosy patients and 583 healthy controls.

SNP	Allele/Genotype		No. of controls	Leprosy vs. CN		PB vs. CN		MB vs. CN		
				No.	*P**	No.	*P**	No.	*P**	OR (95%CI)
rs2144658	Allelic	A/G	253/913	194/852	0.07	110/386	0.83	84/466	**1.8E−03**	**0.65 (0.50–0.85)**
	Genotype	AA/AG/GG	23/207/353	21/152/350	0.07	12/86/150	0.83	9/66/200	**2.1E−03**	–
	Dominant	AA + AG/GG	230/353	173/350	0.03	98/150	0.99	75/200	**5.1E−04**	**0.58 (0.42–0.79)**
	Recessive	AA/AG + GG	23/560	21/502	0.95	12/236	0.56	9/266	0.628	*NS*
rs17129664	Allelic	T/C	136/1030	117/937	0.68	64/432	0.48	53/505	0.178	*NS*
	Genotype	TT/TC/CC	3/130/450	8/101/418	0.12	5/54/189	0.13	3/47/229	0.125	–
	Dominant	TT + TC/CC	133/450	109/418	0.39	59/189	0.76	50/229	0.100	*NS*
	Recessive	TT/TC + CC	3/580	8/519	0.09	5/243	0.04	3/276	0.354	*NS*
rs6670134	Allelic	T/C	376/790	316/738	0.25	144/352	0.20	172/386	0.553	*NS*
	Genotype	TT/TC/CC	53/270/260	57/202/268	0.03	26/92/130	0.05	31/110/138	0.151	–
	Dominant	TT + TC/CC	323/260	259/268	0.04	118/130	0.04	141/138	0.180	*NS*
	Recessive	TT/TC + CC	53/530	57/470	0.34	26/222	0.53	31/248	0.349	*NS*
rs3762318	Allelic	C/T	104/1062	65/983	0.02	41/455	0.67	24/528	**7.5E−04**	**0.46 (0.29–0.73)**
	Genotype	CC/CT/TT	4/96/483	4/57/463	0.03	2/37/209	0.84	2/20/254	**1.1E−03**	–
	Dominant	CC + CT/TT	100/483	61/463	0.01	39/209	0.61	22/254	**3.2E−04**	**0.42 (0.26–0.68)**
	Recessive	CC/CT + TT	4/579	4/520	0.88	2/246	0.85	2/274	0.950	*NS*
rs10789224	Allelic	C/T	385/781	319/735	0.16	148/348	0.20	171/387	0.324	*NS*
	Genotype	CC/CT/TT	59/267/257	48/223/256	0.32	23/102/123	0.34	25/121/133	0.593	–
	Dominant	CC + CT/TT	326/257	271/256	0.13	125/123	0.14	146/133	0.322	*NS*
	Recessive	CC/CT + TT	59/524	48/479	0.57	23/225	0.71	25/254	0.591	*NS*
rs1321157	Allelic	A/G	438/728	383/671	0.55	187/309	0.96	196/362	0.326	*NS*
	Genotype	AA/AG/GG	84/270/229	69/245/213	0.80	39/109/100	0.79	30/136/113	0.330	–
	Dominant	AA + AG/GG	354/229	314/213	0.70	148/100	0.78	166/113	0.732	*NS*
	Recessive	AA/AG + GG	84/499	69/458	0.53	39/209	0.62	30/249	0.138	*NS*
rs7539625	Allelic	G/A	534/632	487/567	0.85	245/251	0.18	242/316	0.343	*NS*
	Genotype	GG/GA/AA	120/294/169	115/257/155	0.83	60/125/63	0.40	55/132/92	0.489	–
	Dominant	GG + GA/AA	414/169	372/155	0.88	185/63	0.29	187/92	0.233	*NS*
	Recessive	GG/GA + AA	120/463	115/412	0.61	60/188	0.25	55/224	0.766	*NS*
rs11465817	Allelic	C/A	492/674	472/582	0.22	231/265	0.10	241/317	0.696	*NS*
	Genotype	CC/CA/AA	103/286/194	105/262/160	0.47	52/127/69	0.24	53/135/91	0.893	–
	Dominant	CC + CA/AA	389/194	367/160	0.30	179/69	0.12	188/91	0.847	*NS*
	Recessive	CC/CA + AA	103/480	105/422	0.34	52/196	0.26	53/226	0.635	*NS*
rs10889677	Allelic	C/A	321/845	315/739	0.22	142/354	0.65	173/385	0.136	*NS*
	Genotype	CC/CA/AA	45/231/307	46/223/258	0.45	19/104/125	0.82	27/119/133	0.331	–
	Dominant	CC + CA/AA	276/307	269/258	0.22	123/125	0.55	146/133	0.170	*NS*
	Recessive	CC/CA + AA	45/538	46/481	0.54	19/229	0.98	27/252	0.331	*NS*
rs2275606	Allelic	A/G	289/877	283/771	0.27	135/361	0.30	148/410	0.438	*NS*
	Genotype	AA/AG/GG	35/219/329	37/209/281	0.54	17/101/130	0.56	20/108/151	0.723	–
	Dominant	AA + AG/GG	254/329	246/281	0.30	118/130	0.29	128/151	0.523	*NS*
	Recessive	AA/AG + GG	35/548	37/490	0.49	17/231	0.64	20/259	0.513	*NS*

MB, multibacillary leprosy; PB, paucibacillary leprosy; OR, odds ratios; CI, confidence interval; CN, healthy controls; NS, not significant.
* $P < 0.005$ (0.05/10) was set as statistically significant after Bonferroni correction; P value < 0.005 was marked in bold.

The exons and flanking regions of four genes (*SLC35D1*, *C1orf141*, *IL23R* and *IL12RB2*) that were located in the validated GWAS hit region were sequenced in an independent cohort including 80 leprosy patients and 99 individuals without leprosy, following the same approach as described in our recent study [18]. Briefly, the targeted regions were captured by NimbleGene SeqCap Kit and were sequenced by the paired-end (2×150 bp) reads on the HiSeq 2500 sequencer (Illumina). Clean sequence reads were aligned to the reference genome hg19 by Burrows-Wheeler Aligner (http://bio-bwa.sourceforge.net/) [22] and Samtools (http://samtools.sourceforge.net/) [23]. SNP calling was performed by GATK (Genome Analysis Toolkit, https://www.broadinstitute.org/gatk/) [24]. To improve the statistic power, all available variants within the four genes were retrieved from the Exome Aggregation Consortium (ExAC, http://exac.broadinstitute.org/), which contains 4327 East Asians, as another control for comparison. The potential pathogenicity of variants in these genes was predicted by using an *in silico* program affiliated prediction (SIFT [25,26], PolyPhen2 HumDiv, PolyPhen2 HumVar [27], LRT [28] and MutationTaster [29]).

2.3. Statistical analyses

Statistical power was calculated by using the Quanto software [30]. The Hardy-Weinberg equilibrium (HWE) was estimated by the PLINK software [31]. A P value < 0.05 was regarded as a deviation from the HWE. The PLINK [31] was also used to calculate the frequencies of the alleles, genotypes and haplotypes of the 10 analyzed SNPs. Bonferroni correction for multiple testing ($P < 0.005$) was applied to avoid false positive result.

2.4. Expression quantitative trait loci (eQTL) and differential expression analyses

As the majority of reported GWAS loci were located in non-coding region, biological relevance behind the statistical association of certain locus with leprosy might be attributed to its effect on gene expression. eQTL analyses were thus performed by using HaploReg V4.1 [32] to evaluate the effect of leprosy-associated SNPs on gene expression. The eQTL effect of SNPs was based on the previous report [12] and the GTEx dataset [13]. To characterize the mRNA expression levels of studied genes in leprosy skin lesions, we retrieved the reported Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) dataset GSE74881 [11], which contains 66 leprosy patients and 9 healthy controls, and compared the gene expression profiling between patients and controls following the procedure in our previous study [33]. Additionally, an interaction network was created to further explore whether the differentially expressed gene was involved in leprosy by using the GeneMANIA webserver (http://genemania.org/) [34]. Detailed information about the methods of the interaction network was described in the original studies [34,35]. In brief, two algorithms were used in the GeneMANIA [34]: the algorithm based on linear regression was used to calculate a single, composite functional association network from multiple networks derived from different genomic or proteomic data sources; the label propagation algorithm was used to predict gene function given this composite network [34,35].

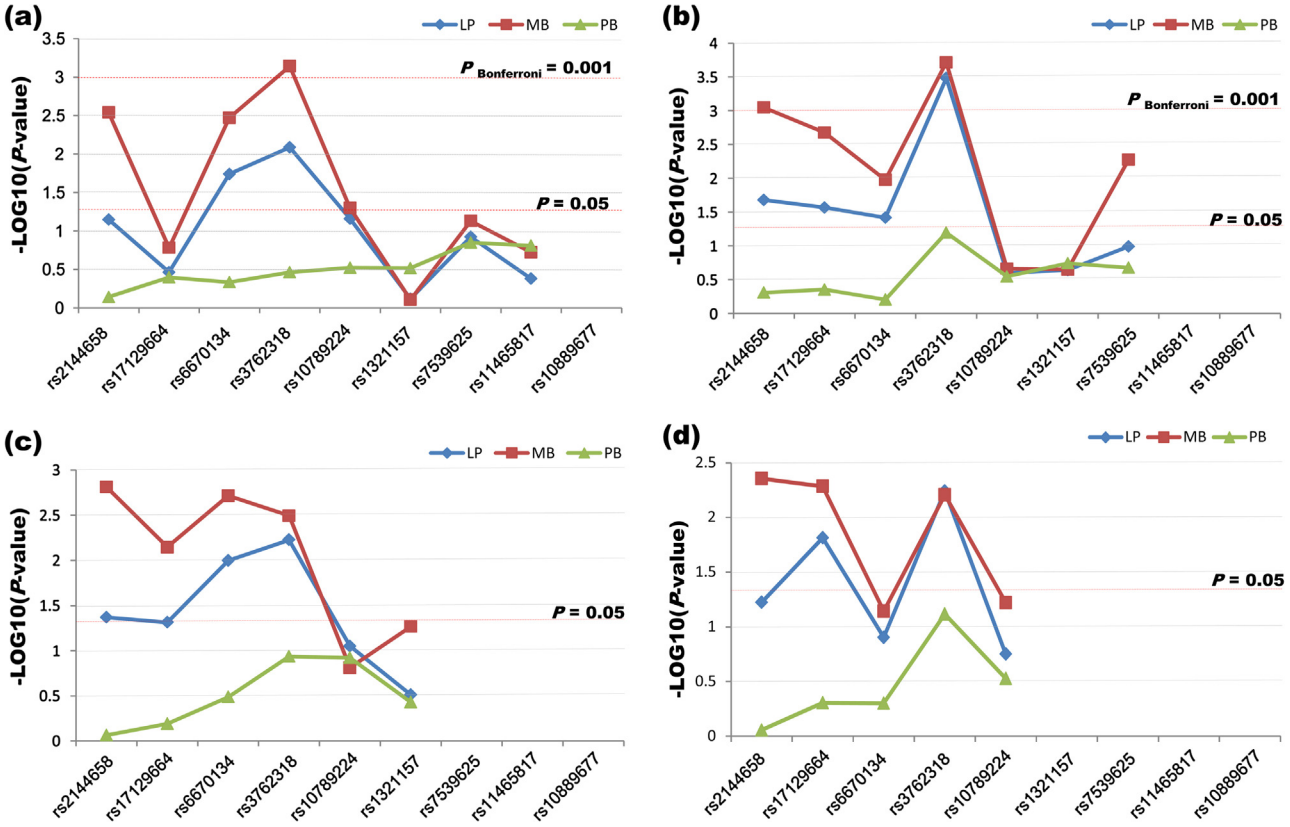

Fig. 1. Sliding window haplotype analysis. A total of 36 windows were found for the 9 SNPs with 2–9 SNPs width (sliding by 1 SNP). Results of 2-SNPs (a), 3-SNPs (b), 4-SNPs (c), and 5-SNPs (d) windows were shown. A P value < 0.0014 [0.05/36 windows] was regarded as significant after Bonferroni correction. SNP rs2275606 was not included as it was located in another chromosome. LP, leprosy *per se*; MB, multibacillary leprosy; PB, paucibacillary leprosy.

3. Result

3.1. Association of rs3762318 and rs2144658 with leprosy

The genotyping call rate of each SNP was above 99.6% in our samples. The minor allele frequency (MAF) of all SNPs in our control samples ranged from 0.09 to 0.46 (Table S3). The power to detect the odds ratio (OR) value as 1.5 for risk allele was expected to be above 82% (Fig. S1). There was no deviation from HWE in controls for all 10 SNPs (Table S3).

The frequencies of the alleles and genotypes of the 10 analyzed SNPs were summarized in Table 1. SNPs rs3762318 and rs2144658 showed significant associations with MB at the allelic ($P = 1.8 \times 10^{-3}$, OR = 0.65, 95% CI: [0.50–0.85] for rs2144658; $P = 7.5 \times 10^{-4}$, OR = 0.46, 95% CI: [0.29–0.73] for rs3762318) and genotypic levels ($P = 2.0 \times 10^{-3}$ for rs2144658; $P = 1.0 \times 10^{-3}$ for rs3762318), especially when the dominant model was considered (AA + AG vs. GG, $P = 5.1 \times 10^{-4}$, OR = 0.58, 95% CI: [0.42–0.79] for rs2144658; CC + CT vs. TT, $P = 3.2 \times 10^{-4}$, OR = 0.42, 95% CI: [0.26–0.68] for rs3762318). The association test of these two SNPs and leprosy survived the Bonferroni correction (Table 1). No significant association was observed between the other SNPs and leprosy *per se* (including its subtypes) after the Bonferroni correction.

(a)

Index	SNP ID	LD (r^2)	Tissues	*P*-value
rs3762318	rs12564219	0.94	Adrenalgland	7.89E-06
rs2144658	rs2144658	1	Lymphoblastoid	2.97E-06
	rs2144657	1	Lymphoblastoid	2.79E-06
	rs7520634	1	Lymphoblastoid	2.90E-06
	rs10789223	1	Lymphoblastoid	2.55E-06
	rs11208998	1	Lymphoblastoid	2.79E-06
	rs4655674	1	Lymphoblastoid	3.11E-06
	rs10789221	1	Lymphoblastoid	2.50E-06
	rs6703469	0.98	Lymphoblastoid	3.10E-06
	rs2144655	0.98	Lymphoblastoid	4.60E-06
	rs6691094	0.86	Lymphoblastoid	1.78E-06
	rs2145812	0.77	Lymphoblastoid	2.39E-06
	rs6679223	0.77	Lymphoblastoid	2.70E-06
	rs2815345	0.73	Lymphoblastoid	6.08E-06
	rs10736406	1	Skin	7.13E-06
	rs10732837	1	Skin	7.13E-06
	rs7544189	1	Skin	7.30E-06
	rs4655515	1	Skin	7.13E-06
	rs7551197	0.98	Skin	3.32E-06
	rs12140736	0.98	Skin	3.87E-06
	rs35028503	0.95	Skin	3.89E-06
	rs10749770	0.92	Skin	9.99E-06
	rs6672126	0.71	Skin	9.64E-06

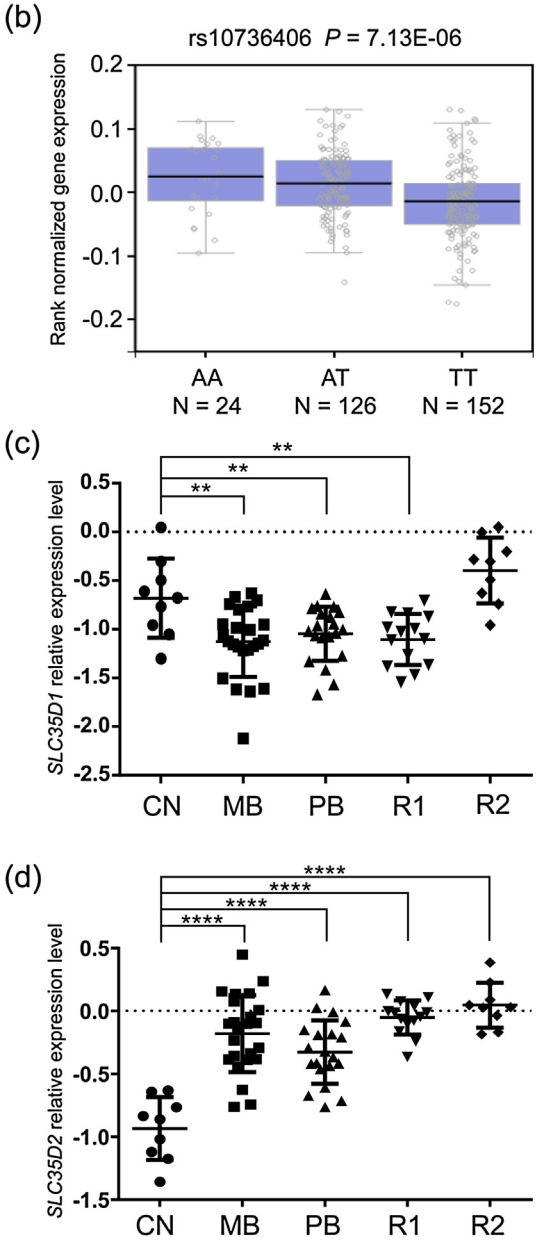

Fig. 2. Expression quantitative trait loci (eQTL) and gene expression analyses. (a) SNPs in high LD with rs3762318 and rs2144658 showed a significant eQTL effect on *SLC35D1* mRNA level. All these SNPs are located in *C1orf141*, except for rs2815345 in *SLC35D1*. The calculation of r^2 value was based on the Asian (ASN) dataset (including Han Chinese in Beijing [CHB] and Japanese in Tokyo [JPT]) of 1000 Genomes Project Phase 1 population [56] from HaploReg V4.1 (http://www.broadinstitute.org/mammals/haploreg/haploreg.php) [32]. (b) The minor allele of the proxy (rs10736406) for rs2144658 ($r^2 = 1$) is associated with a higher *SLC35D1* mRNA level in skin tissue. Significant mRNA expression changes of *SLC35D1* (c) and *SLC35D2* (d) were found in multibacillary leprosy (MB), paucibacillary leprosy (PB), type I reaction (R1), and type II reaction (R2) patients compared to healthy controls (CN). The GSE74481 [11] from GEO database (http://www.ncbi.nlm.nih.gov/geo/) was used for data-mining analysis. ** $P < 0.01$, **** $P < 0.0001$, two-tailed Student's *t*-test.

In the sliding haplotype analysis, several significant haplotypic associations in sliding windows of 2-SNPs and 3-SNPs were observed after the Bonferroni correction ($P < 0.0014$ [0.05/36 windows]) (Fig. 1).

3.2. Expression quantitative trait loci (eQTL) analyses

The eQTL analyses were performed to characterize the biological effect of the positively associated SNPs on gene expression. We observed positive eQTL signals of the two leprosy risk SNPs on *SLC35D1* only. Although we found no eQTL effect of rs3762318 on expression level of nearby genes, a highly linked SNP of rs3762318 (rs12564219, $r^2 = 0.94$) showed a significant ($P = 7.9 \times 10^{-6}$) eQTL effect on gene *SLC35D1* (1p31.3) in the adrenal gland (Fig. 2a). The other leprosy associated SNP rs2144658 also showed a significant ($P = 2.97 \times 10^{-6}$) eQTL effect on *SLC35D1* mRNA level in lymphoblastoid. The minor allele of rs10736406, which is the proxy for rs2144658 ($r^2 = 1$), was associated with a higher *SLC35D1* mRNA level in skin tissue (Fig. 2b). Furthermore, other SNPs in high linkage disequilibrium (LD) with rs2144658, including SNP rs2815345 that was located within *SLC35D1*, also had a significant eQTL effect on *SLC35D1* mRNA level in lymphoblastoid and skin tissues (Fig. 2a).

3.3. Gene expression and interaction network analyses

The validated GWAS hit region (1p31.3) [10] contained four adjacent genes (*IL12RB2*, *IL23R*, *C1orf141* and *SLC35D1*) (Fig. S2). To identify which gene might be the causal gene, dataset GSE74481 [11] from the GEO was retrieved and differential gene expression pattern between leprosy patients and healthy controls was analyzed. A significantly lower mRNA expression level of *SLC35D1*was found in MB, PB and type I reaction (R1) patients relative to healthy controls (Fig. 2c). In addition, *SLC35D2*, an important paralog of *SLC35D1*, showed a higher mRNA expression level in MB, PB, R1 and type II reaction (R2) patients than in healthy controls (Fig. 2d), suggesting a possible involvement of the SLC35D family in leprosy. The other three genes, *IL12RB2*, *IL23R* and *C1orf141*, were unfortunately not included in this GEO dataset [11], and we could not compare their expression patterns.

To further evaluate whether *SLC35D1* could interact with other leprosy risk genes, we performed an interaction network by using the GeneMANIA webserver [34]. Genetic interactions with these leprosy susceptibility genes *CCDC122* [9], *CCDC88B* [36] and *MRC1* [14], and co-expression with *MRC1* [14] and *IL18R1* [37] were found for *SLC35D1* (Fig. 3). All these lines of evidence suggested that *SLC35D1* should be categorized as a leprosy associated gene.

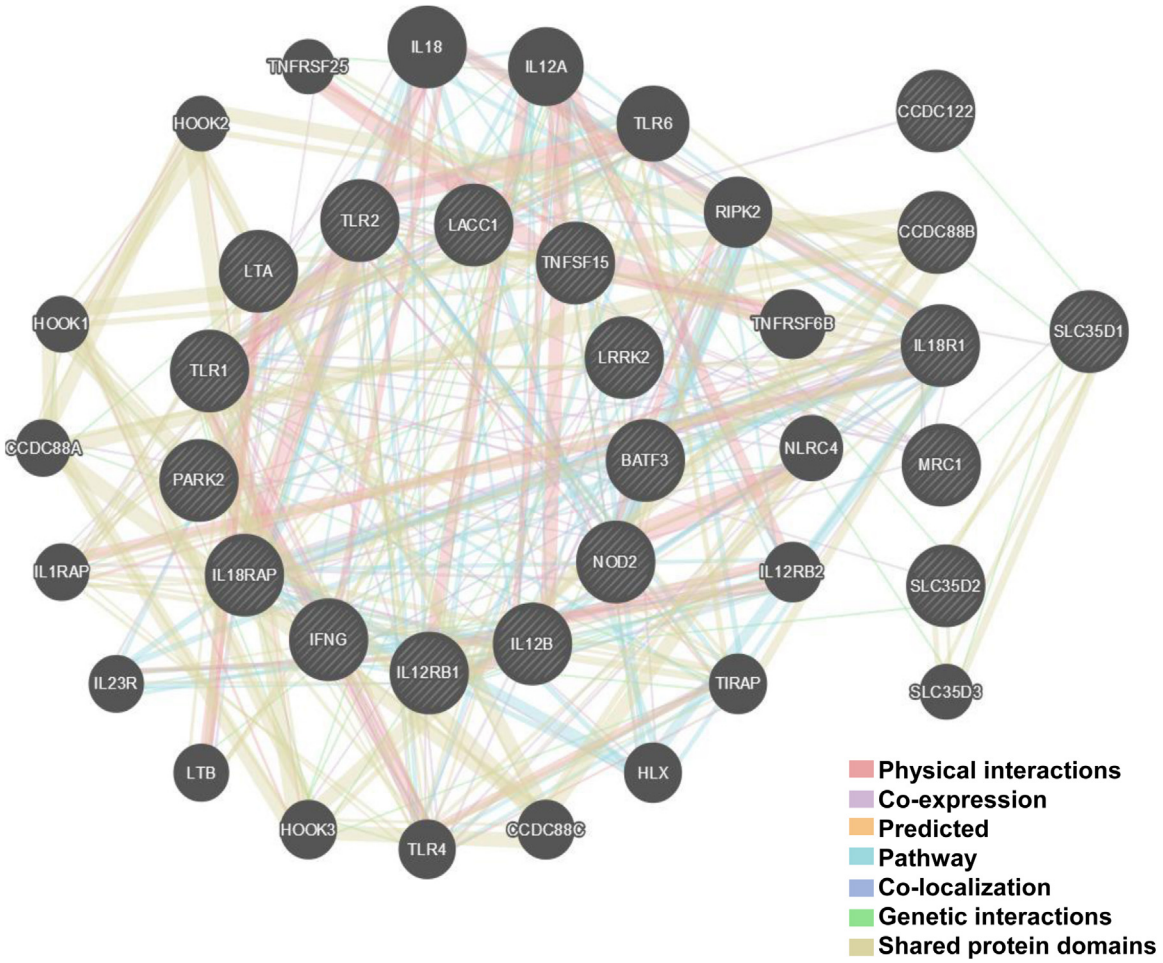

Fig. 3. Interaction network between *SLC35D1* and the GWAS reported genes [9,36,37]. The interaction network was performed by GeneMANIA webserver (http://genemania.org/) [34]. The black circles with slash refer to the query proteins, and the other proteins refer to predicted ones.

3.4. *Rare* IL23R *missense variants were enriched in leprosy patients*

Considering the important role of *IL23R* in immune regulatory pathway albeit we found no evidence at the expression level [38], and to refine the involvement of the confirmed GWAS hit in leprosy, we sequenced the exons and flanking regions of four genes (*SLC35D1*, *C1orf141*, *IL23R* and *IL12RB2*) that were located in the GWAS hit region for genetic variants in 80 leprosy patients and 99 controls (Tables 2 and S4). Among the identified 8 rare missense variants, 4 variants (rs199561187 [p.Y236C] in *SLC35D1*; rs76418789 [p.G149R], rs199542433 [p.L372F] and rs201752419 [p.I373F] in *IL23R*) were predicted to be damaging (Table 2). Although our sample size was relatively small, we were able to find that three of the four damaging variants (rs199561187 in *SLC35D1*, $P = 5.03 \times 10^{-3}$; rs199542433 [$P = 0.02$] and rs201752419 [$P = 0.02$] in *IL23R*) showed a significant association with leprosy. Note that none of the associations survived the Bonferroni correction for multiple tests. The association between the two rare *IL23R* damaging variants and leprosy could be confirmed (rs199542433, $P = 9.08 \times 10^{-3}$; rs201752419, $P = 9.08 \times 10^{-3}$) when the ExAC East Asians (http://exac.broadinstitute.org; accessed on November 2, 2015) were used as a control (Table 2). This result suggested that rare missense variants, rather than regulatory variants in *IL23R* might affect leprosy.

4. Discussion

The development and onset of leprosy was affected by host genetic background [3–5]. Based on the available knowledge, we tentatively concluded that leprosy was caused by many genes/loci with trivial effects [33], and there might be no major effect gene affecting onset and clinical presentation of leprosy. The recent GWASs of leprosy in Han Chinese have provided a valuable list of risk loci [9,10,36], however, limited number of genes were well replicated [16,39,40]. Moreover, despite the success of GWAS in identifying leprosy risk loci, to understand the underlying pathological role and to interpret the biological function of these risk loci/genes in leprosy are difficult, especially when the risk loci reside in the non-coding regions with limited annotation and unknown function.

In this study, we attempted to validate the most recently reported GWAS loci (rs3762318 and rs2275606) [10] in independent cohorts from Southwest China. We increased the SNP density to achieve a higher coverage of the regions harboring the two leprosy associated loci. We confirmed the significant genetic association between rs3762318 and MB, and identified an additional SNP rs2144658 that was associated with MB. Unfortunately, we failed to replicate the association between rs2275606 and leprosy. This result was consistent with the recent family-based association study by using the FBATdosage method that rs3762318, but not rs2275606, was associated with leprosy in Vietnamese [41]. To further characterize the effect of the genotypes and to fine-map the potential causal gene/variants, we performed eQTL, gene expression analyses, as well as sequencing analysis of the exon and flanking regions of four adjacent genes in the validated GWAS hit region. We found more evidence to suggest that *SLC35D1* and *IL23R* might be the causal genes.

The SNP rs3762318 is located in chromosome 1p31.3, where resides four adjacent genes (*IL12RB2*, *IL23R*, *C1orf141* and *SLC35D1*). *C1orf141* has been reported to be associated with Vogt-Koyanagi-Harada syndrome [42] and psoriasis [43], but the function of the corresponding product of *C1orf141* has not been annotated. SNPs rs3762318 and rs2144658 in *C1orf141*, which were positively associated with MB in this study, significantly affected the gene expression level of a nearby gene *SLC35D1*. In addition, other SNPs in high LD with rs3762318 and rs2144658, such as rs12564219, which is located in *SLC35D1*, also showed a significant eQTL effect on *SLC35D1* expression level in adrenal gland, lymphoblastoid and skin tissues (Fig. 2a). Further gene expression analyses showed a significant expression difference of *SLC35D1* and its paralog *SLC35D2* between leprosy patients and healthy controls. These results indicated that *SLC35D1* might be a potential causal gene for leprosy.

SLC35D1 is a member of the solute carrier family, which is expressed in endoplasmic reticulum (ER) membrane [44]. The proteins encoded by *SLC35D1* and its paralog *SLC35D2* play an important role in the translocation of nucleotide sugars [45], which are high-energy donor substrates. Genetic association between member of solute carrier family (*SLC11A1*) and leprosy had been previously reported [46]. In this study, a lower expression level of *SLC35D1* mRNA was found in leprosy patients relative to healthy controls, suggesting an association between *SLC35D1* and leprosy. In contrast, a higher expression level of *SLC35D2* mRNA was found in leprosy patients, this might be caused by a negative feedback autoregulation under the pressure of energy requirement. Considering the dependence of *M. leprae* on host energy production and nutritional products, there might be a potential metabolic association between *SLC35D1* and leprosy. Of note, *SLC35D1* represented an edge node in the network, and the majority of the central nodes were well-recognized leprosy susceptibility genes, which are mainly immune-related genes. Previous studies have shown that those edge nodes, such as *SLC35D1* [44], *SLC35D2* [47], *CCDC88B* [48] and *MRC1* [49], play a major function in metabolism. We suggested that the metabolic

Table 2
Rare missense variants in the *SLC35D1*, *C1orf141*, *IL23R* and *IL12RB2* genes.

Gene	Chr	ID in dbSNP	Function	Ref/Alt	MAF[a]	Leprosy[b]	Control[b]	*P*-value[c]	EA[d]	*P*-value[c]	OR
SLC35D1	chr1:67507941	rs199561187	missense*	T/C	0.0008	2/160	0/198	**5.03E-03**	39/8652	0.17	2.80
C1orf141	chr1:67581127	rs2273682	missense	T/C	0.0371	10/160	26/198	0.14	1152/8390	**4.83E-03**	0.42
IL23R	chr1:67648538	rs539497366	missense	T/A	0.0002	1/160	1/198	**0.01**	11/8654	0.20	4.94
	chr1:67648596	rs76418789	missense*	G/A	0.0112	11/160	15/198	0.08	443/8654	0.28	1.37
	chr1:67705932	rs199542433	missense*	G/C	0.0002	3/160	2/198	**0.02**	21/8620	**9.08E-03**	7.82
	chr1:67705933	rs201752419	missense*	A/T	0.0002	3/160	2/198	**0.02**	21/8618	**9.08E-03**	7.82
IL12RB2	chr1:67795372	rs185002143	missense	G/A	0.0002	1/160	0/198	**5.03E-03**	0/8654	**1.82E-02**	NA
	chr1:67796346	rs78198420	missense	A/T	0.0012	1/160	1/198	**0.01**	55/8626	1.00	0.98

Chr, Chromosome; Ref, Reference allele; Alt, Alternate allele; MAF, Minor allele frequency; OR, odds ratio; NA, not applicable.
*Missense variants were rated as damaging when at least two of five prediction algorithms (SIFT [25,26], PolyPhen2 HumDiv, PolyPhen2 HumVar [27], LRT [28] and MutationTaster [29]) showed a potential pathogenicity.
[a] Minor allele frequency (MAF) from the 1000 Genomes Project [55], variants with MAF <0.05 were regarded as rare.
[b] Allele counts in 80 leprosy patients and 99 controls.
[c] *P*-values were calculated by using the Fisher's exact test and *P*-values < 0.05 were marked in bold.
[d] Allele counts in East Asians from ExAC (http://exac.broadinstitute.org; accessed on November 2, 2015).

pathways, with interactions with the immune system, might contribute to leprosy development, as hypothesized in our previous studies [15,17].

IL12RB2 is a receptor for *IL12*, which plays a key role in immune response and has been reported to be associated with different autoimmune diseases [50,51]. However, there is no robust evidence to support the association between *IL12RB2* and leprosy [52,53]. Note that we found two rare missense variants in *IL12RB2* showed a significant association with leprosy in our sequenced leprosy patients (Table 2), but the associations did not survive the Bonferroni correction for multiple tests. Moreover, these two missense variants were predicted to be not damaging. Therefore, the potential role of *IL12RB2* in leprosy awaits further study.

The *IL23R* gene is located nearby *IL12RB2* and encodes the IL23 receptor. IL23R and its ligand IL23 play important roles in immune regulatory pathway [38]. Copy number variation (CNV) in *IL23R* has been reported to be positively associated with PB [54]. In this study, three rare damaging missense variants were found in *IL23R* and two of them showed a nominally significant association with leprosy, suggesting a potential involvement of *IL23R* protein alteration in leprosy.

The current study has several limitations. First, although our sample size was modest, there is still a possibility of insufficient statistical power to detect the association between leprosy and risk SNPs, especially when the odds ratio was small. Second, we did not validate our result in independent samples. Moreover, there is a low density of SNP coverage in the GWAS hit region containing rs2275606. Finally, the expression profiling did not cover the other three genes in the validated 1p31.3 region, and we could not exclude the potential effect of these genes on leprosy.

In summary, we analyzed 10 SNPs in the previous GWAS hit regions in Han Chinese with and without leprosy from Southwest China, and sequenced a small portion of samples for rare missense variants. The positive association between the GWAS loci rs3762318 (chromosome 1p31.3) and MB was validated, and we found a surrounding SNP rs2144658 was significantly associated with MB. Rare damaging missense variants in *SLC35D1* and *IL23R* might be associated with leprosy. Our results suggested that the *SLC35D1* and *IL23R* genes might be the causal genes underlying the association between the GWAS hit and leprosy. Further studies are essential to elucidate the potential role of *SLC35D1* and *IL23R* in leprosy.

Conflict of interest

The authors declare no conflicts of interest.

Author contributions

D.W. and Y.-G.Y. designed the study; X.-A.L. and Y.-Y.L. collected the samples and clinical information; G.-D.L. and J.-Q.F performed the experiments; G.-D.L., D.W., D.-F.Z. and Q.X. analyzed data; G.-D. L., D.W. and Y.-G.Y. drafted the manuscript. All authors approved the final version of the manuscript.

Acknowledgements

We thank the participants in this study. This study was supported by the National Natural Science Foundation of China (31271346 and 81573034) and Yunnan Province (2014FB177). The funders had no role in study design, data collection and analyses, decision to publish, or preparation of the manuscript.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jdermsci.2016.09.018.

References

[1] R.R. Jacobson, J.L. Krahenbuhl, Leprosy, Lancet 353 (9153) (1999) 655–660.
[2] WHO, Global leprosy update, 2014: need for early case detection, Wkly. Epidemiol. Rec. 90 (36) (2015) 461–474.
[3] M.O. Moraes, C.C. Cardoso, P.R. Vanderborght, A.G. Pacheco, Genetics of host response in leprosy, Lepr. Rev. 77 (3) (2006) 189–202.
[4] E.A. Misch, W.R. Berrington, J.C. Vary, T.R. Hawn Jr., Leprosy and the human genome, Microbiol. Mol. Biol. Rev. 74 (4) (2010) 589–620.
[5] A. Alcais, M. Mira, J.L. Casanova, E. Schurr, L. Abel, Genetic dissection of immunity in leprosy, Curr. Opin. Immunol. 17 (1) (2005) 44–48.
[6] S.R. Krutzik, M.T. Ochoa, P.A. Sieling, S. Uematsu, Y.W. Ng, A. Legaspi, et al., Activation and regulation of Toll-like receptors 2 and 1 in human leprosy, Nat. Med. 9 (5) (2003) 525–532.
[7] C.C. Cardoso, A.C. Pereira, V.N. Brito-de-Souza, I.M. Dias-Baptista, V.C. Maniero, J. Venturini, et al., IFNG +874 T > A single nucleotide polymorphism is associated with leprosy among Brazilians, Hum. Genet. 128 (5) (2010) 481–490.
[8] D.F. Zhang, D. Wang, Y.Y. Li, Y.G. Yao, Mapping genetic variants in the CFH gene for association with leprosy in Han Chinese, Genes Immun. 15 (7) (2014) 506–510.
[9] F.R. Zhang, W. Huang, S.M. Chen, L.D. Sun, H. Liu, Y. Li, et al., Genomewide association study of leprosy, N. Engl. J. Med. 361 (27) (2009) 2609–2618.
[10] F. Zhang, H. Liu, S. Chen, H. Low, L. Sun, Y. Cui, et al., Identification of two new loci at IL23R and RAB32 that influence susceptibility to leprosy, Nat. Genet. 43 (12) (2011) 1247–1251.
[11] F. Belone Ade, P.S. Rosa, A.P. Trombone, L.R. Fachin, C.C. Guidella, S. Ura, et al., Genome-wide screening of mRNA expression in leprosy patients, Front. Genet. 6 (2015) 334.
[12] T. Lappalainen, M. Sammeth, M.R. Friedländer, P.A. 't Hoen, J. Monlong, M.A. Rivas, et al., Transcriptome and genome sequencing uncovers functional variation in humans, Nature 501 (7468) (2013) 506–511.
[13] GTEx Consortium, The genotype-tissue expression (GTEx) pilot analysis: multitissue gene regulation in humans, Science 348 (6235) (2015) 648–660.
[14] D. Wang, J.Q. Feng, Y.Y. Li, D.F. Zhang, X.A. Li, Q.W. Li, et al., Genetic variants of the MRC1 gene and the IFNG gene are associated with leprosy in Han Chinese from Southwest China, Hum. Genet. 131 (7) (2012) 1251–1260.
[15] Y.L. Xiang, D.F. Zhang, D. Wang, Y.Y. Li, Y.G. Yao, Common variants of OPA1 conferring genetic susceptibility to leprosy in Han Chinese from Southwest China, J. Dermatol. Sci. 80 (2) (2015) 133–141.
[16] D. Wang, L. Xu, L. Lv, L.Y. Su, Y. Fan, D.F. Zhang, et al., Association of the LRRK2 genetic polymorphisms with leprosy in Han Chinese from Southwest China, Genes Immun. 16 (2) (2015) 112–119.
[17] D. Wang, L.Y. Su, A.M. Zhang, Y.Y. Li, X.A. Li, L.L. Chen, et al., Mitochondrial DNA copy number, but not haplogroup, confers a genetic susceptibility to leprosy in Han Chinese from Southwest China, PLoS One 7 (6) (2012) e38848.
[18] D. Wang, G.D. Li, D.F. Zhang, L. Xu, X.A. Li, X.F. Yu, et al., Genetic variants of the MAVS, MITA and MFN2 genes are not associated with leprosy in Han Chinese from Southwest China, Infect. Genet. Evol. 45 (2016) 105–110.
[19] International HapMap Consortium, The international HapMap project, Nature 426 (6968) (2003) 789–796.
[20] C. Lippert, J. Listgarten, R.I. Davidson, S. Baxter, H. Poon, C.M. Kadie, et al., An exhaustive epistatic SNP association analysis on expanded Wellcome Trust data, Sci. Rep. 3 (2013) 1099.
[21] M.M. Holland, W. Parson, GeneMarker(R) HID: a reliable software tool for the analysis of forensic STR data, J. Forensic Sci. 56 (1) (2011) 29–35.
[22] H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform, Bioinformatics 25 (14) (2009) 1754–1760.
[23] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, et al., The sequence alignment/map format and SAMtools, Bioinformatics 25 (16) (2009) 2078–2079.
[24] A. McKenna, M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, et al., The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data, Genome Res. 20 (9) (2010) 1297–1303.
[25] P. Kumar, S. Henikoff, P.C. Ng, Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm, Nat. Protoc. 4 (7) (2009) 1073–1081.
[26] P.C. Ng, S. Henikoff, SIFT: predicting amino acid changes that affect protein function, Nucleic Acids Res. 31 (13) (2003) 3812–3814.
[27] I.A. Adzhubei, S. Schmidt, L. Peshkin, V.E. Ramensky, A. Gerasimova, P. Bork, et al., A method and server for predicting damaging missense mutations, Nat. Methods 7 (4) (2010) 248–249.
[28] S. Chun, J.C. Fay, Identification of deleterious mutations within three human genomes, Genome Res. 19 (9) (2009) 1553–1561.
[29] J.M. Schwarz, D.N. Cooper, M. Schuelke, D. Seelow, MutationTaster2: mutation prediction for the deep-sequencing age, Nat. Methods 11 (4) (2014) 361–362.

[30] W.J. Gauderman, Sample size requirements for matched case-control studies of gene-environment interaction, Stat. Med. 21 (1) (2002) 35–50.
[31] S. Purcell, B. Neale, K. Todd-Brown, L. Thomas, M.A. Ferreira, D. Bender, et al., PLINK: a tool set for whole-genome association and population-based linkage analyses, Am. J. Hum. Genet. 81 (3) (2007) 559–575.
[32] L.D. Ward, M. Kellis, HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease, Nucleic Acids Res. 44 (D1) (2016) D877–D881.
[33] D.F. Zhang, D. Wang, Y.Y. Li, Y.G. Yao, Integrative analyses of leprosy susceptibility genes indicate a common autoimmune profile, J. Dermatol. Sci. 82 (1) (2016) 18–27.
[34] D. Warde-Farley, S.L. Donaldson, O. Comes, K. Zuberi, R. Badrawi, P. Chao, et al., The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function, Nucleic Acids Res. 38 (Web Server issue) (2010) W214–220.
[35] S. Mostafavi, D. Ray, D. Warde-Farley, C. Grouios, Q. Morris, GeneMANIA: a real-time multiple association network integration algorithm for predicting gene function, Genome Biol. 9 (2008) S4.
[36] H. Liu, A. Irwanto, X. Fu, G. Yu, Y. Yu, Y. Sun, et al., Discovery of six new susceptibility loci and analysis of pleiotropic effects in leprosy, Nat. Genet. 47 (3) (2015) 267–271.
[37] H. Liu, A. Irwanto, H. Tian, X. Fu, Y. Yu, G. Yu, et al., Identification of IL18RAP/IL18R1 and IL12B as leprosy risk genes demonstrates shared pathogenesis between inflammation and infectious diseases, Am. J. Hum. Genet. 91 (5) (2012) 935–941.
[38] C. Parham, M. Chirica, J. Timans, E. Vaisberg, M. Travis, J. Cheung, et al., A receptor for the heterodimeric cytokine IL-23 is composed of IL-12Rbeta1 and a novel cytokine receptor subunit, IL-23R, J. Immunol. 168 (11) (2002) 5699–5708.
[39] C. Sales-Marques, H. Salomao, V.M. Fava, L.E. Alvarado-Arnez, E.P. Amaral, C.C. Cardoso, et al., NOD2 and CCDC122-LACC1 genes are associated with leprosy susceptibility in Brazilians, Hum. Genet. 133 (12) (2014) 1525–1532.
[40] J.H. Xiong, C. Mao, X.W. Sha, Z. Jin, H. Wang, Y.Y. Liu, et al., Association between genetic variants in NOD2, C13orf31, and CCDC122 genes and leprosy among the Chinese Yi population, Int. J. Dermatol. 55 (1) (2016) 65–69.
[41] A. Cobat, L. Abel, A. Alcais, E. Schurr, A general efficient and flexible approach for genome-wide association analyses of imputed genotypes in family-based designs, Genet. Epidemiol. 38 (6) (2014) 560–571.
[42] S. Cao, S.P. Chee, H.G. Yu, S. Sukavatcharin, L. Wu, A. Kijlstra, et al., Investigation of the association of Vogt-Koyanagi-Harada syndrome with IL23R-C1orf141 in Han Chinese Singaporean and ADO-ZNF365-EGR2 in Thai, Br. J. Ophthalmol. 100 (3) (2016) 436–442.
[43] X. Zuo, L. Sun, X. Yin, J. Gao, Y. Sheng, J. Xu, et al., Whole-exome SNP array identifies 15 new susceptibility loci for psoriasis, Nat. Commun. 6 (2015) 6793.
[44] M. Muraoka, M. Kawakita, N. Ishida, Molecular characterization of human UDP-glucuronic acid/UDP-N-acetylgalactosamine transporter, a novel nucleotide sugar transporter with dual substrate specificity, FEBS Lett. 495 (1–2) (2001) 87–93.
[45] N. Ishida, T. Kuba, K. Aoki, S. Miyatake, M. Kawakita, Y. Sanai, Identification and characterization of human Golgi nucleotide sugar transporter SLC35D2, a novel member of the SLC35 nucleotide sugar transporter family, Genomics 85 (1) (2005) 106–116.
[46] M.J. Brochado, M.F. Gatti, M.A. Zago, A.M. Roselino, Association of the solute carrier family 11 member 1 gene polymorphisms with susceptibility to leprosy in a Brazilian sample, Mem. Inst. Oswaldo Cruz 111 (2) (2016) 101–105.
[47] J.I. Sesma, C.R. Esther Jr., S.M. Kreda, L. Jones, W. O'Neal, S. Nishihara, et al., Endoplasmic reticulum/golgi nucleotide sugar transporters contribute to the cellular release of UDP-sugar signaling molecules, J. Biol. Chem. 284 (18) (2009) 12572–12583.
[48] E. Matsushita, N. Asai, A. Enomoto, Y. Kawamoto, T. Kato, S. Mii, et al., Protective role of Gipie, a Girdin family protein, in endoplasmic reticulum stress responses in endothelial cells, Mol. Biol. Cell 22 (6) (2011) 736–747.
[49] A.P. Moseman, E.A. Moseman, S. Schworer, I. Smirnova, T. Volkova, U. von Andrian, et al., Mannose receptor 1 mediates cellular uptake and endosomal delivery of CpG-motif containing oligodeoxynucleotides, J. Immunol. 191 (11) (2013) 5615–5624.
[50] J.H. Namkung, J.E. Lee, E. Kim, S. Kim, S. Kim, E.S. Shin, et al., Association of single nucleotide polymorphisms in the IL-12 (IL-12A and B) and IL-12 receptor (IL-12Rbeta1 and beta2) genes and gene-gene interactions with atopic dermatitis in Koreans, J. Dermatol. Sci. 57 (3) (2010) 199–206.
[51] G.M. Hirschfield, X. Liu, C. Xu, Y. Lu, G. Xie, Y. Lu, et al., Primary biliary cirrhosis associated with HLA, IL12A, and IL12RB2 variants, N. Engl. J. Med. 360 (24) (2009) 2544–2555.
[52] G.A. Silva, M.P. Santos, I. Motta-Passos, A.L. Boechat, A. Malheiro, R. Ramasawmy, et al., Polymorphisms assessment in the promoter region of IL12RB2 in Amazon leprosy patients, Hum. Immunol. 75 (6) (2014) 592–596.
[53] H. Ohyama, K. Ogata, K. Takeuchi, M. Namisato, Y. Fukutomi, F. Nishimura, et al., Polymorphism of the 5' flanking region of the IL-12 receptor beta2 gene partially determines the clinical types of leprosy through impaired transcriptional activity, J. Clin. Pathol. 58 (7) (2005) 740–743.
[54] S. Ali, A.K. Srivastava, R. Chopra, S. Aggarwal, V.K. Garg, S.N. Bhattacharya, et al., IL12B SNPs and copy number variation in IL23R gene associated with susceptibility to leprosy, J. Med. Genet. 50 (1) (2013) 34–42.
[55] 1000 Genomes Project Consortium, G.R. Abecasis, D. Altshuler, A. Auton, L.D. Brooks, R.M. Durbin, et al., A map of human genome variation from population-scale sequencing, Nature 467 (7319) (2010) 1061–1073.
[56] 1000 Genomes Project Consortium, G.R. Abecasis, A. Auton, L.D. Brooks, M.A. DePristo, R.M. Durbin, et al., An integrated map of genetic variation from 1092 human genomes, Nature 491 (7422) (2012) 56–65.

Supplementary Materials

Table S1. The detailed information of the 10 SNPs analyzed in this study.

Gene	SNP	Allele	MAF[a]	Chr.	Position	Location
C1orf141	rs2144658	A/G	0.208	1	67106212	intron
	rs17129664	T/C	0.113	1	67117615	intron
	rs6670134	T/C	0.345	1	67126431	intron
	rs3762318	C/T	0.207	1	67131436	intron
IL23R-C1orf141	rs10789224	C/T	0.292	1	67139451	intron
IL23R	rs1321157	A/G	0.369	1	67188427	intron
	rs7539625	G/A	0.488	1	67207082	intron
	rs11465817	C/A	0.44	1	67255414	intron
	rs10889677	C/A	0.268	1	67259437	3'-UTR
RAB32-C6orf103	rs2275606	A/G	0.256	6	146597814	intron

[a] Minor allele frequency (MAF) in CHB (Han Chinese in Beijing, China) was obtained from HapMap (http://hapmap.ncbi.nlm.nih.gov/).
Allele - minor / major; Chr. - chromosome

Table S2. Primer information for genotyping the 10 SNPs by using the SNaPshot assay.

Gene	SNP	Primer(5'-3')[a]
C1orf141	rs2144658	F: ATAATTGACTGCTAGAACAAAAACATC
		R: TTATCTTTTATGTTAACACTAGGTCTCACT
		E: $(gact)_5$GCAGAGTCTTCACAATACAATAGTC
	rs17129664	F: TATCTGTTCCTTTCAAAAAATGTATCT
		R: AGAACAATGCCCATTCTCAC
		E: ct$(gact)_9$TTCTTATTAATATCATTGGGGAGGA
	rs6670134	F: AAATGGATTGAAGAAGAAAAAGG
		R: TTTATTCTTTTGAGCTATGCTCTG
		E: $(gact)_2$CTAGCTAAAAGGTGAGTAAAGGAGC
	rs3762318	F: TATTCTTCAGGTTATGAGAGCAAG
		R: GTGACAGAGTGAGACTCCTTCTC
		E: act$(gact)_{13}$TGTTTAAGAAAGACAGCTCTGATGGCATTG
C1orf141-IL23R	rs10789224	F: CACTCTCCCTGTCACCACA
		R: TTTTTGTGAATATGAGTTTTCTTAGAAC
		E: ct$(gact)_{13}$AGTCCACACAACATAGCTCAAAAAC
IL23R	rs1321157	F: TCACATGTCTACTCTGTCTGGATG
		R: TAAATAACTCTTCTAAGATTACACAGCTAGC
		E: ct$(gact)_8$GAGTTGGATACCGTCATCCAACAGT
	rs7539625	F: TACAACAAACCAAACTTGGAAT
		R: ACTGTAGCCACTGATGGTTG
		E: ct$(gact)_{10}$TAAGCTCAACTTTCATTATGCTTTA
	rs11465817	F: AAAAGAAGCAGAGCAATAGAGATG
		R: AGACCCTGACTCAAAAAAAAAAT
		E: ATTAAGTAAGAGATGAAAACTTTAG
	rs10889677	F: ATCACCATGTAAGAATTCCCG
		R: TATGTGTGACCATGAAGCATG
		E: ct$(gact)_{12}$TTTAATTTTAGCCATTCTTCTGCCT
RAB32-C6orf103	rs2275606	F: ATTCATTTGTCAGAGAAGTTGAAAG
		R: TTTAAATAAAAGAAATCCCCTCC
		E: t$(gact)_{12}$CTCTGCAGGAAGCTGCCTCTTGCCTGCACA

[a] In the "$(agct)_n$", n means repeats of "agct". F: forward primer; R: reverse primer; E: extension primer.

Table S3. Allele frequencies of the 10 SNPs in in 527 leprosy patients and 583 healthy controls

Gene	SNP	Allele	MAF [a]		*P*-value [b]
			Control	Leprosy	
C1orf141	rs2144658	A/G	0.22	0.19	0.33
	rs17129664	T/C	0.12	0.11	0.07
	rs6670134	T/C	0.32	0.30	0.18
	rs3762318	C/T	0.09	0.06	1.00
C1orf141-IL23R	rs10789224	C/T	0.33	0.30	0.45
IL23R	rs1321157	A/G	0.38	0.36	0.79
	rs7539625	G/A	0.46	0.46	0.74
	rs11465817	C/A	0.42	0.45	0.93
	rs10889677	C/A	0.28	0.30	0.84
RAB32-C6orf103	rs2275606	A/G	0.25	0.27	0.91

[a] MAF, minor allele frequency

[b] *P*-value of the Hardy-Weinberg equilibrium test

Table S4. Entire list of genetic variants in the *SLC35D1*, *C1orf141*, *IL23R* and *IL12RB2* genes in 80 leprosy patients and 99 individuals without leprosy (Control)

Gene	Chr	ID in dbSNP	Function	Ref/Alt	MAF [a]	Leprosy [b]	Control [b]	*P*-value [c]	EA [d]	*P*-value [c]	OR
SLC35D1	chr1:67470044	-	synonymous	A/G	NA	1/160	0/198	5.03E-03	NA/NA	NA	NA
	chr1:67474880	rs12061148	intron	T/C	0.1981	4/160	7/178	0.04	41/528	0.02	0.30
	chr1:67507879	rs150492296	intron	C/T	0.0054	3/160	3/198	0.02	212/8646	1.00	0.76
	chr1:67507941	rs199561187	missense*	T/C	0.0008	2/160	0/198	5.03E-03	39/8652	0.17	2.80
	chr1:67512903	rs1024229	intron	G/T	0.7095	115/160	118/198	1.00	5188/8204	0.03	1.49
	chr1:67512920	rs1024230	intron	A/C	0.7095	114/160	118/198	1.00	5379/8464	0.05	1.42
	chr1:67515563	rs3048488	intron	CAT/C	NA	23/160	52/198	0.27	NA/NA	NA	NA
C1orf141	chr1:67558739	rs6588237	synonymous	C/T	0.9207	150/160	193/198	1.00	8412/8650	0.01	0.42
	chr1:67558756	rs41313272	missense	C/T	0.1162	7/160	10/198	0.06	NA/NA	NA	NA
	chr1:67558763	rs41299551	synonymous	A/G	0.1070	7/160	10/198	0.06	563/8642	0.33	0.66
	chr1:67558821	-	missense	G/A	NA	2/160	0/198	5.03E-03	NA/NA	NA	NA
	chr1:67560897	rs10889653	intron	A/G	0.9323	150/160	193/198	1.00	8366/8608	0.03	0.43
	chr1:67560956	rs11208997	missense	C/T	0.9385	150/160	193/198	1.00	8404/8646	0.03	0.43
	chr1:67561090	rs72933970	missense	T/C	0.1162	7/160	10/198	0.06	428/5022	0.06	0.49
	chr1:67561909	rs11801593	intron	G/A	0.1162	7/160	10/198	0.06	NA/NA	NA	NA
	chr1:67569090	rs143579777	intron	T/C	0.0118	4/160	5/182	0.03	18/438	0.47	0.60
	chr1:67577367	rs6588240	intron	A/T	0.8053	143/160	167/180	1.00	424/478	0.89	1.07
	chr1:67581017	rs200518928	intron	T/A	NA	1/160	2/198	0.02	NA/NA	NA	NA
	chr1:67581127	rs2273682	missense	T/C	0.0371	10/160	26/198	0.14	1152/8390	4.83E-03	0.42
	chr1:67591596	-	splice site	C/T	NA	2/158	NA/NA	NA	NA/NA	NA	NA

	chr1:67591611	rs11336462	intron	GA/G	NA	59/124	NA/NA	NA	NA/NA	NA	NA
	chr1:67592840	rs7525652	intron	G/A	0.0314	2/160	0/198	5.03E-03	67/8654	0.36	1.62
IL23R	chr1:67633812	rs1884444	missense	G/T	0.5302	116/160	131/198	1.00	5613/8638	0.05	1.42
	chr1:67648460	rs6687620	intron	T/C	0.9521	160/160	198/198	1.00	NA/NA	NA	NA
	chr1:67648538	rs539497366	missense	T/A	0.0002	1/160	1/198	0.01	11/8654	0.20	4.94
	chr1:67648596	rs76418789	missense*	G/A	0.0112	11/160	15/198	0.08	443/8654	0.28	1.37
	chr1:67666483	-	synonymous	A/G	NA	1/160	0/198	5.03E-03	1/8654	0.04	54.42
	chr1:67672765	rs7539625	intron	G/A	0.3816	94/160	116/198	1.00	4574/8612	0.17	1.26
	chr1:67685387	rs7530511	missense	T/C	0.8760	160/160	196/198	1.00	8506/8652	0.12	NA
	chr1:67685443	rs7518660	intron	G/A	0.4111	40/160	45/198	0.23	2238/8612	0.86	0.95
	chr1:67705932	rs199542433	missense*	G/C	0.0002	3/160	2/198	0.02	21/8620	9.08E-03	7.82
	chr1:67705933	rs201752419	missense*	A/T	0.0002	3/160	2/198	0.02	21/8618	9.08E-03	7.82
IL12RB2	chr1:67794033	-	synonymous	T/C	NA	1/160	0/198	5.03E-03	NA/NA	NA	NA
	chr1:67795319	rs1495963	synonymous	T/C	0.8502	123/160	165/198	1.00	6820/8650	0.56	0.89
	chr1:67795372	rs185002143	missense	G/A	0.0002	1/160	0/198	5.03E-03	0/8654	0.02	NA
	chr1:67795456	-	intron	A/C	NA	1/160	0/198	5.03E-03	NA/NA	NA	NA
	chr1:67796346	rs78198420	missense	A/T	0.0012	1/160	1/198	0.01	55/8626	1.00	0.98
	chr1:67833501	rs2252596	intron	G/A	0.1262	38/160	32/198	0.17	NA/NA	NA	NA
	chr1:67845648	rs146207069	intron	CTTCCT/C	NA	9/160	5/198	0.03	378/8640	0.43	1.30
	chr1:67845727	rs78878673	synonymous	G/C	0.0014	2/160	0/198	5.03E-03	91/8652	0.69	1.19
	chr1:67852335	rs2228420	synonymous	G/A	0.4135	45/160	73/198	0.37	2931/8644	0.13	0.76
	chr1:67852381	rs189685637	intron	A/T	0.0014	2/160	0/198	5.03E-03	91/8652	0.69	1.19
	chr1:67861520	rs2229546	synonymous	C/A	0.5551	52/160	78/198	0.40	NA/NA	NA	NA

Chr, Chromosome; Ref, Reference allele; Alt, Alternate allele; MAF, Minor allele frequency; OR, odds ratio; NA, no data available

[a] Minor allele frequency (MAF) from 1000 Genomes Project [1], variants with MAF < 0.05 were regarded as rare.
[b] Allele counts in 80 leprosy patients and 99 controls
[c] *P*-values were calculated by using the Fisher's exact test
[d] Allele counts in East Asians from the ExAC (http://exac.broadinstitute.org; accessed on November 2, 2015)
[*] Missense variants were rated as damaging when at least two of five prediction algorithms (SIFT [2, 3], PolyPhen2 HumDiv, PolyPhen2 HumVar [4], LRT [5] and MutationTaster [6]) showed a potential pathogenicity

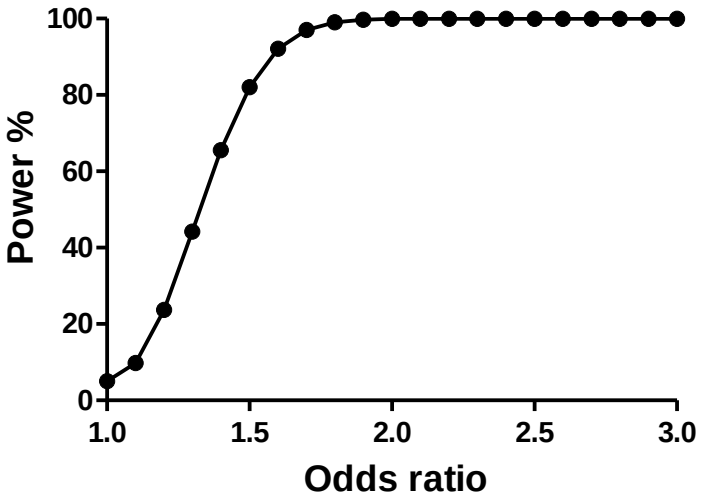

Figure S1. Statistical power estimates for the case-control association analysis. Statistical power was computed using the Qunto software [7] under the gene only hypothesis and log additive model, with the following parameters: risk allele frequency = 0.1; sample size = 527 cases vs. 583 controls; range of OR from 1.0 to 3.0 in increments of 0.1; two-sided type I error rate = 0.05.

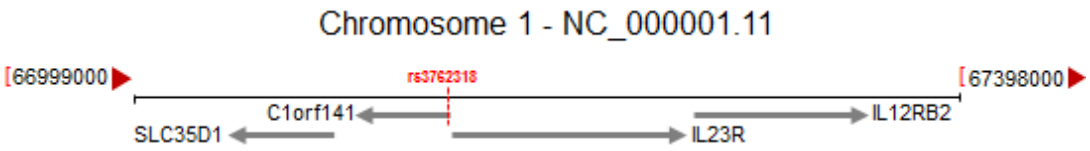

Figure S2. Schematic diagram of the GWAS [8] hit (rs3762318) and its adjacent genes. The schematic diagram was based on the genomic sequence (NC_000001.11 Chromosome 1 Reference GRCh38.p7 Primary Assembly) from NCBI (http://www.ncbi.nlm.nih.gov/gene/).

Reference

[1] C. Genomes Project, G.R. Abecasis, D. Altshuler, A. Auton, L.D. Brooks, R.M. Durbin, et al., A map of human genome variation from population-scale sequencing, Nature 467(7319) (2010) 1061-1073.

[2] P. Kumar, S. Henikoff, P.C. Ng, Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm, Nat. Protoc. 4(7) (2009) 1073-1081.

[3] P.C. Ng, S. Henikoff, SIFT: predicting amino acid changes that affect protein function, Nucleic Acids Res. 31(13) (2003) 3812-3814.

[4] I.A. Adzhubei, S. Schmidt, L. Peshkin, V.E. Ramensky, A. Gerasimova, P. Bork, et al., A method and server for predicting damaging missense mutations, Nat Methods 7(4) (2010) 248-249.

[5] S. Chun, J.C. Fay, Identification of deleterious mutations within three human genomes, Genome Res. 19(9) (2009) 1553-1561.

[6] J.M. Schwarz, D.N. Cooper, M. Schuelke, D. Seelow, MutationTaster2: mutation prediction for the deep-sequencing age, Nat Methods 11(4) (2014) 361-362.

[7] W.J. Gauderman, Sample size requirements for matched case-control studies of gene-environment interaction, Stat. Med. 21(1) (2002) 35-50.

[8] F. Zhang, H. Liu, S. Chen, H. Low, L. Sun, Y. Cui, et al., Identification of two new loci at IL23R and RAB32 that influence susceptibility to leprosy, Nat. Genet. 43(12) (2011) 1247-1251.