INTRODUCTION

Leprosy is an ancient infectious disease caused by *Mycobacterium leprae* (M. leprae) that affects the skin and peripheral nerves. Although treatable, it remains as a major cause of disability and social stigma in many parts of the world, particularly in developing countries, with 201,758 new cases reported in 2015 (World Health Organization, 2016). To understand the genetic basis of leprosy susceptibility, candidate gene-based association analysis and genome-wide linkage and association studies (GWASs) have been performed, and 21 common risk variants were discovered, showing the involvement of the innate and adaptive immune responses (Liu et al., 2012, 2013, 2015; Misch et al., 2010; Wang et al., 2016; Wong et al., 2010; Zhang et al., 2009, 2011). However, these common risk variants are largely located within noncoding regions. Protein-coding variants, particularly low-frequency and rare ones, have not been investigated systematically, although these variants may
probably be the causal ones in most cases (Bodmer and Bonilla, 2008; Fearnhead et al., 2005; Frazer et al., 2009).

Here we performed a three-stage GWAS of protein-coding variants in leprosy. In the discovery stage, 40,491 coding variants were tested for single variant association in 1,648 leprosy patients and 2,318 healthy control subjects. We then validated and replicated our findings in four independent cohorts, totaling 5,400 leprosy patients and 12,080 healthy control subjects. In addition, we performed the biological and network analysis of seven genes identified in this study and 26 susceptibility genes implicated in the previous leprosy GWASs to provide more information on leprosy genetic pathogenesis.

RESULTS

Exome-wide discovery analysis

In the discovery stage (stage 1), we successfully genotyped 273,028 variants in 1,670 leprosy patients and 2,321 control individuals. Principal component analysis confirmed the Chinese ancestry of all participant samples, and a good genetic match between the patients and control subjects was observed (see Supplementary Figures S1, S2 online). After a series of sample and variant quality control filtering, 74,764 polymorphic coding variants were retained for association test. According to the result of the power analysis by Genetic Power Calculator (available at http://zzz.bwh.harvard.edu/gpc/cc2.html), the sample size of our discovery analysis was underpowered for detecting single variants with minor allele frequency (MAF) (in all samples) less than 0.1%, hence only 40,491 coding variants with MAF (in all samples) greater than 0.1% (38,068 nonsynonymous and 2,423 synonymous ones) were tested for single variant association in 1,648 leprosy patients and 2,318 healthy control subjects by GMMAT (Breslow and Clayton, 1993; Chen et al., 2016) software.

As expected, the exome-wide analysis showed strong and extensive associations within the major histocompatibility complex region (Figures 1, 2a). After removing all the variants within the major histocompatibility complex region, the quantile-quantile plot of the remaining single nucleotide polymorphisms (SNPs) showed a good fit with the expected null distribution (Figure 2b), indicating a minimal inflation of exome-wide association results due to population stratification ($\lambda_{SC} = 0.99$). Additional quantile-quantile plots of SNP sets with different MAF thresholds (>0.5%, >1%, and >5%) also showed consistent indication of minimal inflation (see Supplementary Figure S3 online).

The strongest association was obtained at rs3200405 ($P = 3.46 \times 10^{-10}$, odds ratio [OR] = 2.15) within HLA-DRB5, which is correlated to the previously reported GWAS SNP rs9271100 and HLA allele HLA-DRB1*15:01 (Liu et al., 2015; Rani et al., 1993; Wang et al., 2016) ($P_{conditional} = 0.95$, $r^2 = 0.74$, $D^\prime = 0.97$; $P_{conditional} = 0.41$, $r^2 = 0.72$, $D^\prime = 0.99$, respectively) (see Supplementary Table S1 online). Five coding variants (four common and one low frequency) with suggestive association ($P < 1.0 \times 10^{-3}$) were also found within the previously identified GWAS loci (Liu et al., 2012, 2013, 2015; Wang et al., 2016; Zhang et al., 2009, 2011). All four common coding variants were in high linkage disequilibrium with the previously reported SNPs, and only the low-frequency variant (rs76418789 in IL23R) showed independent association (see Supplementary Table S1). In addition, a moderate deviation of the quantile-quantile plot from the expected null distribution was observed at the tail of distribution (after removing all the SNPs within the major histocompatibility complex region), indicating the existence of associations (Figure 2, Supplementary Figure S3). There were 38 coding variants in, to our knowledge, previously unreported genes with suggestive association significance ($P < 1.0 \times 10^{-3}$).

To further investigate the role of low-frequency and rare coding variants, we carried out gene-based tests using SKAT-O (Lee et al., 2012b) and burden test methods (Lee et al., 2012a). However, both SKAT-O and burden test analyses failed to provide indication of true genetic association beyond expectation by chance (see Supplementary Figures S4, S5 online).

Validation analysis

A total of 39 independent coding variants were genotyped in an additional 3,169 leprosy patients and 9,814 healthy control subjects from the northern region of China (stage 2). Of the 34 successfully genotyped variants, 15 were found to show consistent associations between the discovery and validation samples ($P < 0.05$ in the validation samples and OR in the same direction as in the discovery stage), and six reached exome-wide significance ($P < 0.05/40,491 = 1.23 \times 10^{-6}$ based on Bonferroni correction, which is a conservative way to control for false positives as multiple tests are performed) (Duggal et al., 2008; Gibson, 2012) in the combined discovery and validation samples without evidence of heterogeneity: rs145562243 in NCKIPSD ($P = 1.44 \times 10^{-8}$, OR = 4.35), rs149308743 in CARD9 ($P = 4.99 \times 10^{-10}$, OR = 4.75), rs76418789 in IL23R ($P = 6.28 \times 10^{-13}$, OR = 1.37), rs146466242 in FLG ($P = 4.44 \times 10^{-10}$, OR = 1.45), rs780668 in SLC29A3 ($P = 2.89 \times 10^{-7}$, OR = 1.14), and rs181206 in IL27 ($P = 5.92 \times 10^{-7}$, OR = 0.83) (Table 1, Supplementary Table S2 online). In addition, two coding variants, rs55882956 in TYK2 ($P = 2.75 \times 10^{-5}$, OR = 1.29) and rs75746803 in USP49 ($P = 3.45 \times 10^{-6}$, OR = 1.28) showed consistent associations between the discovery and validation samples but were barely below the exome-wide significance in the combined samples (Table 1, Supplementary Table S2).

Replication analysis

Eight nonsynonymous coding variants were further selected for replication analysis using the following criteria: (i) showing consistent association in the validation stage at $P$-value less than 0.005 and (ii) had a $P$-value less than $5 \times 10^{-5}$ in the meta-analysis of the discovery and validation samples. The replication study (stage 3) was performed by using additional independent samples from three southern regions of China, totaling 2,231 leprosy patients and 2,266 healthy control subjects (see Supplementary Table S3 online). Five of the eight variants obtained significant associations in the replication samples alone (rs76418789 in IL23R, $P = 4.35 \times 10^{-3}$; rs146466242 in FLG, $P = 6.58 \times 10^{-3}$; rs55882956 in TYK2, $P = 1.15 \times 10^{-2}$; rs780668 in SLC29A3, $P = 2.05 \times 10^{-3}$; rs145562243 in NCKIPSD, $P = 3.16 \times 10^{-2}$). Out of the three nonsignificant variants, one (rs149308743 in CARD9, $P = 8.69 \times 10^{-1}$) is very rare, another (rs181206 in IL27, $P = 6.09 \times 10^{-2}$) is on the borderline, and the third variant (rs75746803 in USP49,
showed inconsistent genetic effect across the three replication samples.

We then performed the meta-analysis of all the samples from the three stages, consisting of a total of 7,048 patients and 14,398 control subjects, using the Fisher combination method for rare variants and a fixed-effect model for the others. Seven (two rare, three low-frequency, and two common) coding variants showed consistent associations without significant evidence of heterogeneity and reached the exome-wide significance in the combined samples (Table 1). rs75746803 in USP49 on chromosome locus 6p21.1 remained barely below exome-wide significance (MAF = 4.8%, \( P = 3.57 \times 10^{-6}; \text{OR} = 1.25 \) (Table 1, Figure 1, Supplementary Table S2, and Supplementary Figure S6 online).

We evaluated the impact of sex by performing the sex-adjusted meta association test across three stages. The unadjusted and sex-adjusted ORs of the two common and three low-frequency SNPs were similar (with <10% relative difference) (see Supplementary Table S4 online), indicating that the genetic effects of these variants are independent of the effects of sex (Aschengrau and Seage, 2006).

The two rare variants were extremely rare and almost absent in southern...
Table 1. Associations results of eight variants that were studied in all the discovery, validation and replication samples

<table>
<thead>
<tr>
<th>Variant/AA</th>
<th>Alleles</th>
<th>Gene</th>
<th>Function</th>
<th>Type</th>
<th>Stage</th>
<th>F_A</th>
<th>F_U</th>
<th>OR L95</th>
<th>U95</th>
<th>P</th>
<th>Phet</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs145562243 (chr3:48719549)</td>
<td>T/C</td>
<td>NCKIPSD</td>
<td>missense</td>
<td>rare</td>
<td>1. Discovery</td>
<td>0.0039</td>
<td>0.0004</td>
<td>8.92</td>
<td>2.26</td>
<td>35.13</td>
<td>8.78E-04 NA</td>
</tr>
<tr>
<td>AA: R176Q</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2. Validation</td>
<td>0.0057</td>
<td>0.0015</td>
<td>3.93</td>
<td>2.28</td>
<td>6.79</td>
<td>1.31E-06 NA</td>
</tr>
<tr>
<td>rs149308743 (chr9:139258965)</td>
<td>T/C</td>
<td>CARD9</td>
<td>missense</td>
<td>rare</td>
<td>1. Discovery</td>
<td>0.0055</td>
<td>0.0006</td>
<td>8.52</td>
<td>2.79</td>
<td>26.05</td>
<td>8.56E-05 NA</td>
</tr>
<tr>
<td>rs76418789 (chr1:67648596)</td>
<td>A/G</td>
<td>IL23R</td>
<td>missense</td>
<td>low freq</td>
<td>1. Discovery</td>
<td>0.0610</td>
<td>0.0432</td>
<td>1.45</td>
<td>1.16</td>
<td>1.81</td>
<td>5.59E-04 NA</td>
</tr>
<tr>
<td>AA: G149R</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2. Validation</td>
<td>0.0632</td>
<td>0.0479</td>
<td>1.34</td>
<td>1.19</td>
<td>1.52</td>
<td>2.37E-06 NA</td>
</tr>
<tr>
<td>rs146466242 (chr1:152275298)</td>
<td>A/T</td>
<td>FLG</td>
<td>stop-gained</td>
<td>low freq</td>
<td>1. Discovery</td>
<td>0.0558</td>
<td>0.0354</td>
<td>1.59</td>
<td>1.26</td>
<td>2.00</td>
<td>4.17E-05 NA</td>
</tr>
<tr>
<td>AA: K4022*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2. Validation</td>
<td>0.0552</td>
<td>0.0400</td>
<td>1.40</td>
<td>1.23</td>
<td>1.60</td>
<td>4.39E-07 NA</td>
</tr>
<tr>
<td>rs75746803 (chr6:41773726)</td>
<td>G/C</td>
<td>USP49</td>
<td>missense</td>
<td>low freq</td>
<td>1. Discovery</td>
<td>0.0664</td>
<td>0.0481</td>
<td>1.40</td>
<td>1.14</td>
<td>1.73</td>
<td>8.34E-04 NA</td>
</tr>
<tr>
<td>AA: W332C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2. Validation</td>
<td>0.0612</td>
<td>0.0500</td>
<td>1.24</td>
<td>1.10</td>
<td>1.40</td>
<td>6.50E-04 NA</td>
</tr>
<tr>
<td>rs55862956 (chr19:10469919)</td>
<td>A/G</td>
<td>TYK2</td>
<td>missense</td>
<td>low freq</td>
<td>1. Discovery</td>
<td>0.0529</td>
<td>0.0363</td>
<td>1.51</td>
<td>1.18</td>
<td>1.92</td>
<td>4.96E-04 NA</td>
</tr>
<tr>
<td>AA: R703W</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2. Validation</td>
<td>0.0464</td>
<td>0.0383</td>
<td>1.22</td>
<td>1.06</td>
<td>1.40</td>
<td>4.89E-03 NA</td>
</tr>
<tr>
<td>rs780668 (chr10:73111408)</td>
<td>T/C</td>
<td>SLC29A3</td>
<td>missense</td>
<td>common</td>
<td>1. Discovery</td>
<td>0.4648</td>
<td>0.4213</td>
<td>1.19</td>
<td>1.08</td>
<td>1.31</td>
<td>3.42E-04 NA</td>
</tr>
<tr>
<td>AA: S158F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2. Validation</td>
<td>0.4619</td>
<td>0.4338</td>
<td>1.12</td>
<td>1.06</td>
<td>1.18</td>
<td>1.36E-04 NA</td>
</tr>
</tbody>
</table>

(continued)
Chinese (stage 3), and as a result, their ORs could not be evaluated in the replication samples and are not included in Supplementary Table S4.

Conditional and haplotype analyses
To verify whether the low-frequency coding variant rs76418789 is independent of the previously reported GWAS SNP rs3762318 (Zhang et al., 2011) within IL23R, we performed conditional and haplotype analyses of the two variants using the overlapping samples between the current study and the datasets in the previously published GWASs (Liu et al., 2015; Zhang et al., 2009, 2011) (see Supplementary Table S9 and Supplementary Figures S7–S10 online). Consistently, the haplotype analysis showed that the AA haplotype (i.e., carrying the two risk alleles of the two variants) confers significantly higher risk than the AG haplotype (i.e., carrying only the risk allele of rs3762318), confirming further the independent risk effects of these two SNPs within IL23R (see Supplementary Table S6 online). We also probed the association evidence for the two SNPs within IL23R, all coding risk variants are within genes that have not been reported before to our knowledge (Table 1 and Figure 1), justifying the value of investigating exome-wide coding variants.

We performed gene ontology (GO) enrichment analysis of seven genes identified in this study and of those seven genes together with 26 susceptibility genes implicated in the previous leprosy GWASs. The analysis showed that the regulation of immune response (GO: 0007766, \( P = 5.00 \times 10^{-6} \)) and the immune response (GO: 0006955, \( P = 3.10 \times 10^{-14} \)) are the most enriched GO term for the two gene sets, respectively (Liu et al., 2012, 2013, 2015; Wang et al., 2016; Zhang et al., 2009, 2011) (see Supplementary Table S9 and Supplementary Figure S11 online). According to the GO hierarchy, the GO: 0007766 is annotated as the “Regulate” of the GO: 0006955, indicating that the leprosy genes discovered in this are closely connected (in terms of molecular function) with the ones implicated by the previous GWASs.

An integrated gene network analysis of all 33 susceptibility genes (identified by previous GWASs and these analyses) using GeneMANIA (Mostafavi et al., 2008) in Cytoscape tool.

### Table 1. Continued

<table>
<thead>
<tr>
<th>Variant/AA</th>
<th>Alleles</th>
<th>Gene</th>
<th>Function</th>
<th>Type</th>
<th>Stage</th>
<th>( F_{A}^{1} )</th>
<th>( F_{U}^{1} )</th>
<th>OR</th>
<th>L95</th>
<th>U95</th>
<th>( P )</th>
<th>Phet</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs181206</td>
<td>G/A</td>
<td>IL27</td>
<td>missense</td>
<td>common</td>
<td>1. Discovery</td>
<td>0.1171</td>
<td>0.1497</td>
<td>0.76</td>
<td>0.65</td>
<td>0.88</td>
<td>1.30E-04</td>
<td>NA</td>
</tr>
<tr>
<td>(chr16:28513403)</td>
<td>AA:L119P</td>
<td></td>
<td></td>
<td></td>
<td>2. Validation</td>
<td>0.1175</td>
<td>0.1346</td>
<td>0.85</td>
<td>0.78</td>
<td>0.93</td>
<td>4.68E-04</td>
<td>NA</td>
</tr>
<tr>
<td>3. Combined 1+2</td>
<td>NA</td>
<td>NA</td>
<td>0.83</td>
<td>0.77</td>
<td>0.89</td>
<td>5.92E-07</td>
<td>0.16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Replication1</td>
<td>0.0809</td>
<td>0.0969</td>
<td>0.82</td>
<td>0.65</td>
<td>1.04</td>
<td>9.42E-02</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Replication2</td>
<td>0.0720</td>
<td>0.0769</td>
<td>0.94</td>
<td>0.71</td>
<td>1.24</td>
<td>6.38E-01</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Replication3</td>
<td>0.0617</td>
<td>0.0721</td>
<td>0.85</td>
<td>0.62</td>
<td>1.17</td>
<td>3.13E-01</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Combined 4+5+6</td>
<td>NA</td>
<td>NA</td>
<td>0.86</td>
<td>0.71</td>
<td>1.04</td>
<td>6.09E-02</td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. All combined</td>
<td>NA</td>
<td>NA</td>
<td>0.83</td>
<td>0.78</td>
<td>0.89</td>
<td>1.08E-07</td>
<td>0.61</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: AA, Amino acid change; chr, chromosome; \( F_{A} \), minor allele frequency in patients; \( F_{U} \), minor allele frequency in control subjects; freq, frequency; L95, lower 95% confidence limit; NA, not applicable; OR, odds ratio with respect to the minor allele; Phet, \( P \)-value of heterogeneity; U95, upper 95% confidence limit.

1Minor allele/major allele.
2Common: minor allele frequency \( \geq 5\% \) in patients; low frequency: 1% \( \leq \) minor allele frequency < 5% in patients; rare: minor allele frequency < 1% in control subjects.
3Indicates results from fixed-effects meta-analysis.
4For rare variants in the three replications, \( F_{A} \) is count for A1A1/A1A2/A2A2 in patients; \( F_{U} \) is count for A1A1/A1A2/A2A2 in control subjects. Because of the large deviations of the estimated OR in the small samples, we reported only their association \( P \)-values, which were calculated by Fisher exact test.
5Combined \( P \)-values were calculated by fixed-effect inverse-variance method if ORs were estimated; otherwise by fixed-effect \( z \) statistics combination method.

### Biological and network analysis
All seven coding risk variants discovered in this study are nonsynonymous and were found to be deleterious by either SIFT tool (available from http://sift.bii.a-star.edu.sg), PolyPhen-2 tool (available from http://genetics.bwh.harvard.edu/pph2/bgi.shtml), or protein structural analysis (see Supplementary Table S8 and Supplementary Figures S7–S10 online). Except for rs76418789 within the known leprosy gene IL23R (Zhang et al., 2011), all coding risk variants are within genes that have not been reported before to our knowledge (Table 1 and Figure 1), justifying the value of investigating exome-wide coding variants.
skin barrier plays an important role in leprosy, and direct contact through skin might be an important route of infection for *M. leprae*.

Literature review also showed that NCKIPSD is not related to the immunity, although it has been included in the gene network. The protein of NCKIPSD belongs to the NCK family of adaptor proteins, major regulators of act in the cytoskeleton. It is a part of FCG receptor-dependent phagocytosis pathway and is implicated in many functional processes, such as assembly and maintenance of sarcomeres and stress fiber formation (Lim et al., 2001; Satoh and Tominaga, 2001). The rare variant rs145562243 (p.R176Q) was located in the proline-rich region (PRD) of the NCKIPSD N-terminus, whose overexpression was found to be related to abnormalities in vesicle formation and trafficking, leading to the defective endocytosis of FCG receptor (Oh et al., 2013). Together with the identification of RAB32 (Zhang et al., 2011), which has been shown to be involved in autophagy and phagocytic digestion of bacteria (Hirota and Tanaka, 2009; Seto et al., 2011; Spano and Galan, 2012), as well as LRRK2 (Zhang et al., 2009) as an interacting partner of RAB32 (Waschbusch et al., 2014) and related to cell death due to autophagic dysfunction and mitochondrial damage (Alegre-Abarrategui et al., 2009; Plowey et al., 2008), the discovery of NCKIPSD highlights the potential involvement of endocytosis/phagocytosis/autophagy in host defense against *M. leprae* infection.

We acknowledge that the results from the Gene Mania analysis or GO term enrichment analysis are broad in terms of functional inference, and future integrated analysis with other types of data, for example, gene expression profiles or epigenetic data, may provide more specific functional impacts of the reported variants.

**MATERIALS AND METHODS**

**Study subjects**

We performed a three-stage case-control analysis for this study. Stage 1 and stage 2 included 16,974 individuals from the northern region of China. Stage 3 included 2,231 leprosy patients and 2,266 healthy control subjects from southern China. All patients and healthy control subjects with written informed consent were recruited as previously described (Zhang et al., 2009). All individuals were of Chinese descent, and the clinical information of samples is summarized in Supplementary Table S3.

This study was approved by the institutional review board committee at the Shandong Provincial Institute of Dermatology and Venereology, Shandong Academy of Medical Sciences.

**Genotyping and quality control in the discovery stage**

We carried out this study using Illumina Infinium Human Exome Bead Chip (version 1.0) array (Illumina, San Diego, CA). SNPs went through the following quality control filters: call rate greater than 99%, MAF in all samples greater than 0.1%, and Hardy-Weinberg equilibrium *P*-value in control subjects greater than 1.0 × 10⁻⁸. We also excluded SNPs located in non-autosomal chromosomes and noncoding regions. A total of 40,491 genotyped SNPs were used for association analysis in the discovery stage.

The samples with call rate of less than 98%, one of related pairs (first-, second- or third-degree familial relationships) with lower call rate (seven), and population outliers based on the principal
component analysis method (16 patients and two control subjects) were excluded. Overall, 1,648 patients and 2,318 control subjects passed the sample quality control filters and were used in the discovery analysis.

Statistical analysis
In the discovery stage, we tested associations between phenotypes and single-variant genotypes using GMMAT_v0.7 (Breslow and Clayton, 1993; Chen et al., 2016). In the validation and replication stages, we used a logistic regression model for those SNPs with MAF greater than 1% and Fisher exact test for those SNPs with MAF less than 1%. The meta-analysis was performed using a fixed-effect model. We declared a single variant-trait association significant if the nominal p-value was less than 1.23 × 10^{-6} (0.05/40491). Assessment of OR heterogeneity across independent samples was carried out by evaluating the p-values from Cochran Q-statistics, and heterogeneity p-values less than 0.05 were considered significant.

Genotyping analysis and quality control in the validation and replication studies
SNP genotyping for the validation and replication stages were conducted at the Shandong Provincial Key Laboratory for Dermato-venereology using the Sequenom MassARRAY system, OpenArray custom genotyping assays on QuantStudio 12K Flex Real-Time PCR System (Applied Biosystems, Foster City, CA) and TaqMan custom genotyping assays on a 7900 HT Fast Real-Time PCR System (Applied Biosystems) according to the manufacturers’ instructions. The following quality control filters were adopted: variants with undetermined clusters or variants with call rate less than 90%, variants with significant deviation from Hardy-Weinberg equilibrium in control samples (Hardy-Weinberg equilibrium, P < 1 × 10^{-3}), and samples with call rate less than 95% were excluded.

Biological and network analysis
GO enrichment analysis. GO enrichment analysis was performed with seven identified genes set from the current study and 33-gene set, including seven genes from the current study and 26 reported genes from previous GWASs (Liu et al., 2012, 2013, 2015; Wang et al., 2016; Zhang et al., 2009, 2011), respectively.

GO enrichment was implemented by TopGO (Alexa et al., 2006), an R package, which calculates GO enrichment P-values for a given gene list. The 30 most significant GO terms are listed in the Supplementary Table S9.

Integrated network enrichment. GeneMANIA (Multiple Association Network Integration Algorithm) (Mostafavi et al., 2008) in Cytoscape was applied to the 33-gene set to carry out the integrated gene network analysis (see Supplementary Figure S12).

More detailed information on the study subjects, genotyping, quality control, variant selection, statistical analysis, and biological and network analyses is provided in the Supplementary Materials online.

CONFLICT OF INTEREST
The authors state no conflict of interest.

ACKNOWLEDGMENTS
We thank the individuals who participated in this project and Shandong Computer Science Center (National Supercomputer Center in Jinan), which provided us with the platform for statistical analysis. This work was funded by grants from the National Natural Science Foundation of China (81620108025, 81601387, 81602746, 31601040); the National Clinical Key Project of Dermatology and Venereology; the Shandong Provincial Advanced Taishan Scholar Construction Project; the Innovation Project of Shandong Academy of Medical Science; the Shandong Provincial Independent Innovation Project (ZR2015HZ001); the Shandong Province independent innovation and achievement transformation project (2014CGZH1307); the Natural Science Foundation of Shandong Province (BS2013Y010, 2014ZRK03128, ZR2014YL037, ZR2015Y1025, JQ201616, ZR2016HB47); the Key research and development program of Shandong Province (2015GSF118126, 2016ZDS07A06); the Science Development Funding Agency of Shandong Province, China (2010G0020258); a joint BII-GIS cat 3 platform grant from which VL has been partially paid (IAF311010); and the Agency of Science, Technology and Research (A*STAR) of Singapore.

SUPPLEMENTARY MATERIAL
Supplementary material is linked to the online version of the paper at www.jidonline.org, and at http://dx.doi.org/10.1016/j.jid.2017.08.004.

REFERENCES


Scott IR, Harding CR. Filaggrin breakdown to water binding compounds during development of the rat stratum corneum is controlled by the water activity of the environment. Dev Biol 1986;115:84–92.


This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/4.0/