

## Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Zhang F-R, Huang W, Chen S-M, et al. Genomewide association study of leprosy. *N Engl J Med* 2009;361:2609-18. DOI: 10.1056/NEJMoa0903753.

## SUPPLEMENTARY INFORMATION

The authors' affiliations are as follows: Shandong Provincial Institute of Dermatology and Venereology, Shandong Academy of Medical Science, Jinan, Shandong, China (F.R.Z., S.M.C., H.L., X.X.Y., T.S.C., L.Z., G.Q.Y., Y.X.Y., B.Q.S., C.Y.W., Y.L., X.A.F., Y.P.S., H.Q.T., H.X.L., W.L.D., J.J.L.); Institute of Dermatology and Department of Dermatology at No.1 Hospital, Anhui Medical University, Hefei, Anhui, China; The Key Laboratory of Gene Resource Utilization for Severe Diseases, Ministry of Education and Anhui Province, Hefei, Anhui, China (L.D.S., Y.C., S.Y., C.Z., J.W.H., C.Q., Z.Z., H.C., H.F.Z., X.B.Z., Y.L.C., F.S.Z., W.S.L., S.M.H., X.J.Z., J.J.L.); Shanghai-MOST Key Laboratory of Health and Disease Genomics, Chinese National Human Genome Center at Shanghai, Shanghai, China (H.W., M.S., Q.Y.J., Y.W.); Genome Institute of Singapore, Singapore (J.J.L., Y.L., H.Q.L., T.E., N.Y., M.L.H.); Jiangsu Provincial Center for Disease Control and Prevention, Nanjing, Jiangsu, China (Y.H.T., L.H.Z.); Wenshan Institute of Dermatology, Wenshan, Yunnan, China (R.D.Y., H.L.); Anhui Provincial Institute of Dermatology, Hefei, Anhui, China (Q.W.); Weifang Institute of Dermatology, Weifang, Shandong, China (J.Y.L.); Linyi Institute of Dermatology, Linyi, Shandong, China (X.Z.); Yantai Institute of Dermatology, Yantai, Shandong, China (Y.L.J.) Jining Institute of Dermatology, Jining, Shandong, China (L.G.M.); Zaozhuang Institute of Dermatology, Zaozhuang, Shandong, China (G.Y.); Qingdao Center for Disease Control and Prevention, Qingdao, Shandong, China (Z.X.J.); Jinan Institute of Dermatology, Jinan, Shandong, China (X.D.W.); Weihai Center for Disease Control and Prevention (J.P.Y.); Rizhao

Institute of Dermatology (Z.H.H); Zibo Center for Disease Control and Prevention (C.H.G.); Qixia County Skin Disease Control Station (Y.Q.L.); Linqu County Skin Disease Control Station (R.Y.L.); Zhucheng County Skin Disease Control Station (D.M.W.); Zoucheng County Skin Disease Control Station (D.W.); Linshu County Skin Disease Control Station (J.X.L.); Feixian County Skin Disease Control Station (W.K.C.); Sishui County Skin Disease Control Station (H.C.C.); Tengzhou County Skin Disease Control Station (Y.P.L.); Shouguang Skin Disease Control Station (W.G.Y). Laixi County Skin Disease Control Station (S.Y.W.); Pingdu County Skin Disease Control Station (K.J.W).

## **METHODS**

### ***SUBJECTS***

The GWAS series consisted of 738 persons affected by leprosy and 1276 controls of Chinese Han collected in Anhui, Shandong and Jiangsu provinces. The diagnosis of leprosy was based on medical records stored in local leprosy control institutions and clinical assessments at the time of blood taken (looking for evidence of leprosy such as claw hand, lagothermalomas or foot drop, etc). Demographic characteristics, clinical subtypes and age at onset of the disease were also collected from medical records. The classification of the patients was based on clinical and histological criteria (Ridley & Jopling, 1966). All the patients were classified into two clinical subtypes: multibacillary (MB) form including patients with lepromatous (LL), BL and borderline (BB) leprosy and paucibacillary (PB) form including patients with borderline tuberculoid (BT) and tuberculoid (TT) leprosy. The patients were grouped into four age-at-diagnosis classes (0-15 years, 16-25 years, 26-35 years, >35 years) for heterogeneity analysis. All the controls were healthy individuals without leprosy, autoimmune and systemic disorders and family history of leprosy (including first-, second- and third-degree relatives).

Three independent samples were used in the replication study. The first one included 2,164 cases and 4,373 controls of Chinese Han recruited from Anhui, Shandong and Jiangsu provinces in the north/middle of China. The second sample included 304 cases and 709 controls of Chinese Han collected from Yunnan province, in the south of China. The third sample of 786 cases and 873 controls was recruited

from the Chinese Minority groups in Yunnan province, including the subjects of Chinese Chuang (355 cases and 426 controls); Miaos (197 cases and 220 controls); Yizu (169 cases and 200 controls), and several other smaller Chinese ethnic groups (65 cases and 27 controls, in total). All the cases and controls were recruited using the uniform criteria and matched regarding to ethnic origin and geographic area. The clinical and demographic information were collected using the same questionnaire as the samples used in the GWAS. Ancestry was determined by self-report through questionnaire where a multiple-choice question (Han, Chuang, Miaos, Yizu and others) was provided. All the cases and the controls were recruited in 2006 to 2008 and with written informed consent. The study was approved by the institutional IRB committees at the Shandong Provincial Institute of Dermatology and Venereology, Shandong Academy of Medical Science and the Anhui Medical University.

#### ***DNA EXTRACTION AND NORMALIZATION***

EDTA anticoagulated venous blood samples were collected from all participants. Genomic DNA was extracted from peripheral blood lymphocytes by standard procedures (Miller et al, 1988) using Flexi Gene DNA kits (QIAGEN, Germany). Genomic DNAs were diluted to working concentrations of 50 ng/μl for genome-wide genotyping and 15-20 ng/μl for the replication study (diluted in 10 mM Tris/1 mM EDTA). DNA samples were surveyed for quality both by a Nanodrop Spectrophotometer (ND-1000) and the electrophoresis assay.

### ***GENOTYPING OF GWAS***

The whole genome genotyping was conducted at Chinese National Human Genome Center at Shanghai, China. Approximately 200ng of genomic DNA was used for genotyping analysis. Briefly, each sample was whole-genome amplified, fragmented, precipitated and resuspended in appropriate hybridization buffer. Denatured samples were hybridized on prepared Illumina Human 610-Quad BeadChip. After hybridization, the BeadChip oligonucleotides were extended by a single labeled base, which was detected by fluorescence imaging with an Illumina Bead Array Reader. Normalized bead intensity data obtained for each sample were loaded into the Illumina BeadStudio 3.2 software, which converted fluorescence intensities into SNP genotypes.

### ***GENOTYPING OF REPLICATION STUDY***

Genotypes for top 93 SNPs selected for follow-up analysis were conducted at the Key Laboratory of Gene Resource Utilization for Severe Diseases, Anhui Medical University, Hefei, China. 88 SNPs from non-MHC region were genotyped using the Sequenom MassArray system (San Diego, USA) according to manufacture's instructions. None of the nine SNPs selected from the MHC region was successful in designing the Sequenom MassARRAY Assay (San Diego, USA) and were subjected to the analysis by TaqMan Assay by Design in a 7900 HT Fast Real-Time PCR System (Applied Biosystem Foster City, CA, USA), but only 5 SNPs were successfully genotyped.

Approximately 15ng of genomic DNA was used to genotype each sample using the Sequenom MassArray system (San Diego, USA). The sample DNA was amplified by multiplex PCR reaction and the PCR products were then used for locus-specific single-base extension reaction. The resulting products were desalted and transferred to a 384-element SpectroCHIP array. Allele detection was performed using MALDI-TOF MS. The mass spectrograms were analysed by the Sequenom MassARRAY TYPER software (San Diego, USA).

The primers and probes of TaqMan Assay were ordered from Shanghai GeneCore BioTechnologies Co., Ltd in China. Reactions were set up for PCR using the TaqMan® Universal PCR Master Mix (2x), No AmpErase® UNG, and appropriate primers and probes. The experiment data and results were generated using a 7900HT Fast System by performing an amplification run and an allelic discrimination run. An amplification run was performed using a Standard Curve (AQ) plate document to generate real-time PCR data. Then the real-time PCR data could be used to analyze and troubleshoot the PCR data by the allelic discrimination assay, if needed. An allelic discrimination run was conducted using an Allelic Discrimination (AD) plate document. The SDS software analyzed the data, and then we assigned allele calls automatically.

### ***QUALITY CONTROL***

For the genome-wide analysis, the clustering of the genotypes was carried out with the Gen-Call software version 6.2.0.4, which assigns a quality score to each locus and

an individual genotype confidence score that is based on the distance of a genotype from the center of the nearest cluster. All the intensity-only SNPs and the SNPs on the X, Y and mitochondria chromosomes (43902 SNPs) and the SNPs with call-rate lower than 90%, or MAF < 1% in either cases or controls (83616 SNPs), or showing significant deviation from Hardy-Weinberg Equilibrium in the controls (p-value  $\leq 10^{-7}$ ) (1440 SNPs), or having bad clusters (59) were removed.

We genotyped a total of 738 cases and 1276 controls. 14 cases and 16 controls were removed due to their low genotype call rates (< 98%). We examined potential genetic relatedness of the samples using pair-wise IBS-based method implemented in PLINK 1.04 software.<sup>1</sup> A total of 32 controls and 18 cases were removed due to potential genetic relatedness.

Finally, 491,883 SNPs in 1934 samples (706 cases and 1228 controls) remained for further statistical analysis.

### ***THE PRINCIPAL COMPONENTS ANALYSIS***

The remaining 1934 samples (after removal of samples with low call rates and familial relationships) were subsequently assessed for population stratification using a principal component analysis-based method implemented in the software package EIGENSTRAT. The original script from EIGENSTRAT was modified to extract the principal components for plotting in R. PCA analysis was performed twice in our data. Firstly, it was used to identify genetic outliers. We combined the 1934 samples with 206 reference HapMap samples and performed PCA analysis. The HapMap samples

were drawn from: Yoruba in Ibadan, Nigeria (YRI) (57), Japanese in Tokyo, Japan (JPT) (44), Chinese Han in Beijing (45), China (CHB) and CEPH (Utah residents with ancestry from northern and western Europe) (CEU) (45). We employed a very stringent criterion for removing potential genetic outliers. 3 samples (controls) with  $PC3 > 0.07$  were removed as population outliers from GWA analysis. After removing the 3 outliers, PCA was utilized again to assess population stratification in our samples. This was done by performing PCA on the 1931 samples of GWA study (706 cases, 1225 controls). The second PCA analysis suggested some genetic mismatch between the 706 cases and 1225 controls) (Supplement Figure 2).

The genotype data in replication study was subjected to the same quality control analysis. All the 93 SNPs analyzed in the replication study, the cluster patterns of the genotyping data from the Illumina, Sequenom and TaqMan analyses were checked to confirm their good quality.

### ***GENOME-WIDE ASSOCIATION ANALYSIS***

The genotypes of 491,883 SNPs in 706 cases and 1225 controls were used for the genome-wide association analysis. Cochran-Armitage trend test was performed to evaluate the association between genotype and disease phenotypes. Two methods were adopted to minimize the inflationary effect of population stratification on the GWA results. Firstly, we performed a conditional trend test by using the EIGENTSTRAT program where the top 10 principal components from the PCA analysis of the 1931 samples were used as ancestry index to correct for implicit

population sub-structure within our samples. This correction has effectively minimized the inflationary effect of population stratification on the genome-wide association results ( $\lambda_{gc}=1.03$ ). Secondly, we created a perfectly matched sample of 706 cases and 514 controls by removing 711 control samples that did not match well with the cases and performed the second genome-wide analysis without correction for population stratification. The analysis yielded a small  $\lambda_{gc}$  value of 1.03. For both the genome-wide association analyses, the distribution of the logarithms of the observed p values largely fitted its null distribution, except at the tail of distribution ( $p<10^{-3}$ ) where the observed p values were much smaller than what are expected under the null hypothesis of no association (Supplement Figure 4). The Q-Q plot was used to evaluate the overall significance of the GWA results.

#### ***ASSOCIATION ANALYSIS OF THE TOP 1000 SNPS WITH CORRECTION FOR BOTH STRATIFICATION AND GENDER IMBALANCE***

Given that the cases and the controls were not matched on gender ratio and the sex is a risk effect (there were about 4 male patients to one female patient), we performed the association analysis of the top 1000 SNPs from both the genome-wide analyses by adjusting for both population stratification and gender by adding the top ten principle components (generated in EIGENSTRAT) and gender as co-variables in a logistic regression model. The association p-values were calculated using score test, similar with the Cochran-Armitage trend test. A direct comparison showed that the association results of the top 1000 SNPs from the analysis with gender adjustment and

the one without gender adjustment were very similar, indicating that the gender imbalance between the cases and the controls had a minimal impact on the association results.

### ***REPLICATION ANALYSIS***

To select SNPs for the replication analysis, we first identified top 60 non-MHC genomic regions of association whose p value of top SNP was less than  $5.0 \times 10^{-4}$  in either GWAS\_1931 or GWAS\_1220 analysis and then selected top 1 or 2 SNPs within each identified region, with the exception of LRRK2 and NOD2 regions where more than two top SNPs were selected. In addition, 8 non-MHC SNPs were selected whose p value of association was greatly improved by incorporating additional 2269 Chinese subjects of other GWAS studies as controls into the first-stage GWA analysis. With 4 SNPs failed in the assay design and additional 3 SNPs failed in the genotyping analysis, 88 non-MHC SNPs were successfully genotyped in the replication samples. Within the MHC region, 9 SNPs were selected to cover the two independent associations identified within the region as well as the previously suggested susceptibility locus LTA. The 9 MHC SNPs were designed for TaqMan SNP assay (all failed in Sequenom iPLEX assay design), and 5 SNPs were successful genotyped. Therefore, a total of 93 SNPs were successfully genotyped and analyzed in the replication (Supplement table 1). In the replication study, the association analysis of the 93 SNPs was performed in the three replication samples individually as well as the combined sample consisting of the GWA sample of 1220 subjects and all the

three replication samples. For rs1873613 within LRRK2, the joint analysis was also performed by combined the GWAS sample of 1931 subjects and the two replication samples of Chinese Han. For the replication analysis, the analysis was performed using additive logistic regression model, with gender and age as the covariate. For the joint analysis based on the 1931 GWAS subjects, the analysis was performed using additive logistic regression model, with gender, age and the ten principle components as the covariates. The 10 principle components for the replication samples were set to zero as there was no way to assess population stratification of the replication samples given the small number of selected SNPs typed. All p values reported are without correction for multiple testing in the replication study.

#### ***TEST FOR INDEPENDENCE***

Independence test of associated SNPs within the MHC region as well identified non-MHC loci was performed by step-wise conditional analysis, i.e., form a co-variate set, then the most significant SNP given the current co-variate set was selected and added to the current co-variate set until the adjusted smallest p-values is bigger than 0.01.

#### ***HETEROGENEITY TEST***

Heterogeneity test between the odds ratios in different patient subgroups was performed by using the logistic regression analysis of the cases (case-only analysis) where subclinical phenotypes were used as the outcome variable.<sup>2</sup>

### ***LD PATTERN ANALYSIS***

To identify susceptibility genes underlying various associations, we analyzed the linkage disequilibrium (LD) pattern around the risk-associated SNPs and determined the LD block where the risk-associated SNPs located. We then investigated the gene or genes that were covered by the LD blocks (Supplement Figure 6). Only one gene was identified within each LD block at each of the identified associations, with the exception of the association locus within MHC where multiple genes are implicated. For the association within the MHC region, the susceptibility gene, HLA-DRB1, was referred according to the well-established association at the gene. LD structures and haplotype block plots were generated using Haploview software v3.2.<sup>3</sup>

### ***NETWORK ANALYSIS***

The 8 Identified genes were investigated for network analysis by Ingenuity Pathway Analysis (IPA) software (Ingenuity Systems, [www.ingenuity.com](http://www.ingenuity.com)) using an unsupervised analysis. In order to form networks, IPA queried the Ingenuity Knowledge Base for interactions between the identified genes and all other gene objects stored in the knowledge base, to generate a set of networks. Within each network, genes/gene products were represented by nodes, with the biological relationship between the nodes represented by edges. All edges were supported by at least one reference from the literature, or from canonical information stored in the IPA knowledge base. A significance score was assigned to each network formed, based on the hypergeometric distribution and calculated with the right tailed Fisher's

Exact Test. The score indicated the likelihood of the identified genes forming a network together, based on the IPA knowledge base, due to random chance.

## REFERENCE

1. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81(3):559-75.
2. Garcia-Closas M, Hall P, Nevanlinna H, et al. Heterogeneity of breast cancer associations with five susceptibility loci by clinical and pathological characteristics. *PLoS Genet* 2008;4(4):e1000054.
3. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;21(2):263-5.

## SUPPLEMENTARY TABLES AND FIGURES

Supplementary Table 1 **The association results of all the 93 SNPs from the two genome-wide analyses and the replication analysis**

SNP	chr	pos	allele		MAF	GWAS_1931		GWAS_1220		Replication1		Replication2		Replication3		Combined_all	
			Major/minor	gene	control	P	OR	P	OR	P	OR	P	OR	P	OR		
rs10493579	1	76275783	G/A	LOC729766	0.08	3.59E-05	0.61	2.34E-03	0.47	7.71E-01	1.02	4.70E-01	0.81	8.45E-01	1.04	3.57E-01	0.94
rs17633857	1	194170533	A/G	KCNT2	0.03	1.00E-03	1.52	3.87E-03	2.59	7.94E-01	1.02	4.66E-02	0.58	4.79E-01	0.89	8.57E-01	1.01
rs4388693	1	229286987	A/G	FAM89A	0.24	NA	NA	7.34E-04	0.63	7.87E-01	0.99	9.59E-01	1.01	6.33E-01	0.95	3.91E-02	0.91
rs17744934	3	9284199	G/A	SRGAP3	0.15	3.61E-04	0.72	2.03E-02	0.69	2.11E-01	0.92	5.68E-02	1.34	1.64E-03	1.32	5.83E-01	1.02
rs13063276	3	9287536	G/A	SRGAP3	0.19	1.60E-05	0.7	3.65E-02	0.77	3.38E-01	0.94	1.30E-01	1.23	1.74E-02	1.21	9.80E-01	1.00
rs7633016	3	46703663	G/A	ALS2CL	0.27	5.61E-07	1.43	1.16E-02	1.40	4.12E-02	1.10	5.86E-01	0.92	8.64E-01	1.01	9.94E-03	1.10
rs4076927	3	46705753	A/G	ALS2CL	0.27	4.65E-07	1.44	1.28E-02	1.38	4.15E-02	1.09	6.09E-01	0.92	9.02E-01	1.01	1.10E-02	1.09
rs9844990	3	95860101	G/A	LOC389137	0.40	1.16E-04	1.37	8.74E-03	1.31	6.18E-01	0.98	1.93E-01	1.14	9.31E-05	1.37	1.10E-02	1.09
rs1447862	3	95895941	G/A	LOC389137	0.40	3.94E-04	1.36	3.16E-02	1.23	5.51E-01	0.97	1.37E-01	1.18	3.80E-06	1.45	1.09E-02	1.08
rs4235455	4	101448476	G/A	EMCN	0.29	2.20E-05	1.19	1.65E-02	1.43	9.93E-01	1.00	2.42E-01	0.88	6.37E-01	1.04	6.22E-01	1.02
rs10491238	5	12217978	G/A	CTNND2	0.24	1.54E-06	0.81	3.15E-02	0.67	8.50E-01	1.01	6.27E-01	1.08	7.81E-01	1.03	8.47E-01	0.99
rs9312793	5	12235708	A/G	CTNND2	0.26	8.61E-07	0.84	1.08E-02	0.63	9.16E-01	1.01	7.09E-01	1.06	6.69E-01	0.96	5.21E-01	0.98
rs1390566	5	18660238	G/A	LOC646273	0.38	6.60E-06	1.46	1.72E-05	1.75	1.91E-02	0.91	1.73E-01	1.19	1.89E-01	0.90	3.70E-01	0.97
rs1845830	5	18665166	G/A	LOC646273	0.38	7.05E-06	1.46	1.74E-05	1.75	1.76E-02	0.91	1.98E-01	1.18	1.69E-01	0.90	3.50E-01	0.97
rs3749642	5	73232396	G/A	RGNEF	0.30	2.16E-05	1.25	5.80E-03	1.47	5.06E-01	0.97	1.48E-01	0.87	3.57E-01	0.93	8.46E-01	0.99
rs283626	5	73237008	A/G	RGNEF	0.33	1.33E-06	1.27	1.82E-03	1.57	3.05E-01	0.96	1.26E-01	0.85	8.86E-01	0.99	8.58E-01	0.99
rs2250107	5	123324046	A/G	KRT18P16	0.10	4.96E-05	0.63	2.23E-03	0.49	3.91E-01	0.93	3.04E-01	1.38	2.04E-01	1.28	1.79E-01	0.91
rs2546839	5	123343048	A/G	KRT18P16	0.11	1.43E-03	0.76	4.59E-03	0.54	4.07E-01	0.93	5.04E-01	1.23	6.61E-01	1.08	1.53E-01	0.91
rs2517467	6	30997239	A/G	HLA-B	0.27	2.08E-04	1.42	4.36E-04	1.71	9.85E-01	1.00	2.49E-01	1.22	5.97E-02	0.73	1.81E-01	1.06

rs753725	6	30998850	G/A	VAR52	0.27	2.29E-04	1.42	4.45E-04	1.71	9.81E-01	1.00	2.82E-01	1.21	8.77E-02	0.75	1.64E-01	1.06
rs9264868	6	31379580	A/G	HLA-B	0.10	NA	NA	1.96E-04	2.12	2.77E-02	1.13	4.94E-01	1.12	5.95E-01	0.95	2.33E-03	1.14
rs1800683	6	31648050	G/A	LTA	0.42	3.96E-01	0.88	7.57E-01	1.02	5.09E-02	1.08	2.18E-01	1.13	1.01E-01	0.88	3.92E-01	1.03
rs602875	6	32681607	A/G	HLA-DRB1	0.32	3.27E-09	0.55	3.47E-04	0.58	8.81E-22	0.64	3.29E-01	0.85	6.32E-03	0.77	5.35E-27	0.67
rs4730520	7	111703220	G/A	ZNF277	0.13	1.31E-03	1.44	1.20E-02	1.80	2.35E-01	1.08	2.57E-01	1.24	2.61E-01	0.89	1.08E-01	1.08
rs42490	8	90847650	G/A	RIPK2	0.42	6.00E-04	0.66	1.23E-03	0.66	2.45E-13	0.74	2.40E-01	0.88	1.21E-02	0.83	1.38E-16	0.76
rs40457	8	90892832	A/G	RIPK2	0.28	2.96E-04	0.71	1.43E-02	0.73	2.45E-10	0.74	6.79E-01	0.97	2.03E-02	0.84	1.34E-12	0.77
rs7014909	8	98919103	G/A	LAPTM4B	0.19	4.59E-05	1.31	3.96E-02	1.51	4.77E-01	0.97	4.53E-01	1.12	5.09E-01	0.94	7.33E-01	1.02
rs13270024	8	142270659	A/G	DENND3	0.14	3.43E-05	1.45	2.56E-02	1.53	5.67E-02	1.11	7.67E-01	1.03	1.70E-01	1.14	4.74E-04	1.16
rs13261781	8	142270794	G/A	DENND3	0.14	4.10E-05	1.46	2.83E-02	1.51	6.36E-02	1.11	7.37E-01	1.04	8.07E-02	1.19	2.85E-04	1.17
rs1348644	9	9168137	A/G	PTPRD	0.39	1.37E-03	0.8	8.02E-03	0.70	2.76E-01	1.05	8.23E-03	1.40	6.16E-01	1.04	3.79E-01	1.04
rs999648	9	13054793	G/A	MPDZ	0.14	8.10E-05	1.44	4.96E-03	1.54	6.33E-02	0.91	6.83E-01	0.93	5.61E-01	1.06	7.63E-01	0.99
rs4325000	9	17921364	G/A	SH3GL2	0.38	6.33E-05	0.77	4.70E-02	0.77	6.89E-01	0.98	4.87E-01	1.06	1.84E-01	1.11	8.10E-01	0.99
rs1111782	9	27202966	G/A	TEK	0.42	1.56E-02	1.2	2.00E-01	1.08	9.32E-01	1.00	3.22E-01	0.88	2.22E-03	1.27	1.64E-01	1.04
rs3737187	9	27203772	A/G	TEK	0.43	3.43E-02	1.17	2.54E-01	1.05	9.86E-01	1.00	2.23E-01	0.86	2.50E-03	1.27	1.85E-01	1.04
rs489383	9	29562017	G/A	LOC401497	0.32	6.57E-03	0.67	1.00E-07	0.47	4.59E-01	0.97	6.09E-01	0.93	5.03E-01	0.94	1.36E-02	0.92
rs303549	9	105079214	A/G	CYLC2	0.06	5.51E-04	0.53	3.28E-02	0.52	8.64E-01	0.98	6.84E-01	0.91	2.68E-02	1.64	7.17E-01	0.97
rs10982385	9	116532838	A/C	TNFSF15	0.44	1.25E-03	1.18	6.09E-02	1.28	4.94E-05	1.18	1.54E-01	1.16	5.66E-02	1.16	8.73E-08	1.19
rs4574921	9	116578155	A/G	TNFSF15	0.32	1.23E-05	1.24	2.38E-03	1.46	9.23E-09	1.27	2.29E-02	1.29	1.79E-05	1.39	2.17E-16	1.31
rs10114470	9	116587593	A/G	TNFSF15	0.47	3.36E-07	0.76	1.47E-04	1.60	6.63E-08	1.25	4.95E-01	1.08	9.87E-05	1.35	5.42E-14	1.28
rs6478108	9	116598524	G/A	TNFSF15	0.46	8.60E-09	0.74	4.55E-04	1.54	1.80E-11	1.32	3.95E-02	1.30	8.20E-07	1.47	3.39E-21	1.37
rs10904914	10	6844599	A/G	LOC439949	0.01	4.18E-04	2.46	3.56E-03	4.55	8.81E-02	0.74	3.78E-01	0.49	4.58E-01	0.56	3.67E-01	0.87
rs2462692	10	7184261	C/A	SFMBT2	0.06	NA	NA	1.03E-09	0.17	4.35E-01	0.71	1.00E-01	0.56	6.20E-01	1.62	5.15E-06	0.39
rs7896133	10	96454720	G/A	CYP2C18	0.09	3.49E-03	1.5	1.97E-03	1.91	6.27E-01	1.03	4.41E-01	1.21	1.50E-01	1.20	4.61E-02	1.12
rs4086116	10	96697192	G/A	CYP2C9	0.08	4.26E-03	1.48	1.49E-03	2.02	5.12E-01	1.05	5.66E-01	1.17	1.41E-01	1.21	3.00E-02	1.13
rs6583967	10	96804465	A/G	CYP2C8	0.05	5.59E-03	1.61	7.19E-03	1.86	4.05E-01	1.08	7.50E-01	1.12	2.10E-01	1.19	3.13E-02	1.16

rs10786189	10	96866686	G/A	CYP2C8	0.06	3.66E-02	1.41	2.01E-01	1.26	4.91E-01	1.06	3.29E-01	1.34	3.47E-01	1.14	7.96E-02	1.12
rs11193639	10	109461372	A/G	SORCS1	0.20	7.95E-05	1.45	1.10E-01	1.22	7.83E-01	1.02	9.91E-01	0.98	7.61E-01	1.04	2.24E-01	1.05
rs853600	10	119889137	C/A	CASC2	0.17	5.85E-05	0.76	2.46E-03	0.62	8.00E-02	0.91	3.63E-03	1.59	1.37E-01	1.17	3.02E-01	0.96
rs509360	11	61305135	A/G	C11orf9	0.36	4.65E-02	0.88	1.72E-01	0.84	7.42E-03	1.12	2.22E-01	0.88	9.20E-01	0.99	1.99E-01	1.04
rs174576	11	61360086	C/A	FADS2	0.34	4.10E-02	0.87	1.18E-01	0.83	7.06E-03	1.12	3.20E-01	0.90	8.69E-01	0.99	1.95E-01	1.04
rs670358	11	64348255	G/A	FHAD1	0.44	8.68E-04	0.82	2.35E-02	0.74	6.80E-01	1.02	7.76E-01	1.04	6.28E-01	1.03	5.86E-01	0.98
rs916606	11	64372813	G/A	EHD1	0.49	4.61E-04	0.81	3.91E-02	0.79	8.29E-01	0.99	6.70E-01	1.05	7.53E-01	1.02	3.29E-01	0.97
rs495942	11	85490056	G/A	PICALM	0.25	1.70E-05	0.8	7.00E-02	0.71	4.08E-02	0.91	7.61E-01	0.97	1.11E-02	0.78	1.60E-04	0.86
rs4963910	12	25736132	G/A	LOC645233	0.05	9.08E-06	0.35	2.00E-03	0.36	5.14E-02	1.61	4.56E-01	0.62	2.69E-03	0.33	3.88E-02	0.71
rs11174235	12	38587878	G/A	SLC2A13	0.18	1.77E-02	0.74	1.61E-02	0.55	2.08E-01	0.93	5.15E-03	0.67	6.57E-02	1.16	1.01E-01	0.93
rs11174238	12	38589361	A/C	SLC2A13	0.19	1.63E-02	0.74	1.86E-02	0.55	3.76E-01	0.95	6.51E-03	0.68	4.35E-02	1.18	2.18E-01	0.94
rs1873613	12	38838684	A/G	LRRK2	0.25	6.25E-04	0.67	9.37E-03	0.67	3.62E-03	0.87	1.39E-03	0.65	8.92E-01	1.01	5.10E-05	0.86
rs10878220	12	38895147	G/A	LRRK2	0.25	5.40E-05	0.64	2.80E-04	0.57	2.73E-02	0.90	9.26E-04	0.63	9.79E-01	1.00	1.29E-04	0.86
rs1491938	12	38931897	G/A	LRRK2	0.29	6.74E-05	0.65	3.56E-03	0.65	3.44E-02	0.91	1.21E-02	0.72	8.66E-01	1.01	8.72E-04	0.89
rs12820920	12	38974348	A/G	LRRK2	0.28	5.42E-05	0.65	8.27E-03	0.66	2.95E-02	0.90	2.73E-02	0.76	7.31E-01	0.97	4.87E-04	0.88
rs2384679	12	114716811	C/A	MED13L	0.47	5.57E-04	0.9	9.39E-02	0.85	9.55E-01	1.00	6.29E-01	0.92	7.57E-01	0.98	3.05E-01	0.96
rs4767374	12	114720946	A/G	MED13L	0.47	5.48E-04	0.9	9.31E-02	0.85	8.04E-01	0.99	7.19E-01	0.93	8.32E-01	0.98	2.67E-01	0.96
rs3913471	12	116457223	G/A	KSR2	0.45	3.26E-05	0.77	2.48E-02	0.73	9.64E-01	1.00	2.07E-02	0.78	2.10E-01	1.10	2.74E-01	0.97
rs2347291	12	125590863	A/G	LOC728173	0.07	2.69E-05	0.6	1.51E-01	0.67	1.86E-01	1.14	6.85E-01	0.89	8.02E-01	1.06	9.85E-01	1.01
rs9533634	13	43295815	A/G	CCDC122	0.24	3.15E-03	0.76	1.43E-01	0.85	9.48E-08	0.77	2.45E-01	0.85	7.31E-05	0.69	4.77E-12	0.76
rs3088362	13	43331630	C/A	CCDC122	0.26	1.51E-09	1.72	2.00E-06	1.87	6.64E-23	1.53	4.69E-04	1.60	1.11E-02	1.26	1.36E-31	1.52
rs3764147	13	43355925	A/G	C13orf31	0.31	1.69E-10	1.81	4.06E-07	1.97	1.46E-37	1.70	7.29E-06	1.74	4.54E-08	1.55	3.72E-54	1.68
rs10507522	13	43377000	A/G	C13orf31	0.31	2.22E-07	0.6	4.17E-05	0.55	3.32E-18	0.66	3.07E-02	0.75	1.05E-02	0.81	4.64E-24	0.68
rs9568798	13	52512555	G/A	OLFM4	0.23	3.86E-05	1.31	7.15E-03	1.57	1.12E-01	1.08	3.79E-01	1.09	2.57E-01	1.11	1.76E-03	1.13
rs12880888	14	39966526	A/G	FBXO33	0.08	2.25E-05	0.64	8.40E-02	0.65	2.82E-01	0.91	5.29E-01	1.19	2.29E-01	1.24	2.90E-01	0.93
rs1952915	14	40015091	G/A	FBXO33	0.08	9.80E-05	0.64	2.89E-01	0.76	2.06E-01	0.90	1.47E-01	1.49	1.67E-01	1.28	4.17E-01	0.95

rs2748141	14	51384867	G/A	GNG2	0.03	5.11E-06	2.6	5.46E-02	1.82	5.80E-01	1.06	5.95E-01	0.84	4.44E-01	0.82	1.48E-01	1.14
rs12878420	14	51406432	C/A	GNG2	0.45	1.69E-04	0.8	1.12E-02	0.67	4.25E-01	0.97	3.91E-01	0.91	1.49E-01	1.12	3.82E-01	0.97
rs9920704	15	82211371	A/G	ADAMTSL3	0.25	1.02E-04	0.73	5.28E-02	0.73	7.08E-01	1.02	7.81E-01	0.98	6.74E-02	1.18	7.63E-01	1.01
rs1426165	15	82243056	A/G	ADAMTSL3	0.46	9.33E-05	0.74	1.44E-02	0.74	6.82E-02	1.08	4.26E-01	1.12	6.25E-02	1.15	8.01E-02	1.06
rs9302752	16	49276604	A/G	NOD2	0.29	1.05E-10	1.78	1.42E-09	2.28	3.83E-28	1.59	7.11E-02	1.26	9.21E-05	1.44	3.77E-40	1.59
rs7194886	16	49282694	G/A	NOD2	0.14	4.12E-06	1.56	4.43E-07	2.25	5.26E-18	1.56	1.31E-02	1.51	1.86E-07	1.77	1.77E-30	1.63
rs8057341	16	49295481	A/G	NOD2	0.22	4.22E-03	1.4	5.22E-02	1.33	2.23E-05	1.22	4.51E-01	1.10	1.49E-01	0.86	5.53E-05	1.17
rs3135499	16	49323628	A/C	NOD2	0.21	2.54E-03	1.41	9.21E-02	1.26	8.16E-05	1.20	5.59E-01	1.07	1.13E-01	0.84	2.52E-04	1.16
rs12443665	16	78166002	C/A	MAF	0.20	6.35E-01	1.1	8.42E-01	1.05	7.78E-01	1.01	3.33E-01	0.84	5.31E-02	0.85	4.29E-01	0.97
rs1559341	16	78168467	G/A	MAF	0.11	6.93E-05	1.38	6.76E-02	1.67	1.20E-04	1.26	4.21E-01	0.90	8.98E-01	1.01	3.49E-04	1.18
rs16950734	16	78173101	G/A	MAF	0.46	3.99E-03	0.75	3.77E-02	0.71	9.39E-01	1.00	3.11E-01	1.15	8.87E-01	1.01	7.04E-01	0.99
rs8048274	16	78173882	A/C	MAF	0.45	8.83E-04	1.22	1.27E-01	1.32	6.31E-01	1.02	8.93E-01	0.98	8.22E-01	1.01	2.09E-01	1.04
rs2671666	17	44891520	A/G	NGFR	0.33	1.28E-05	1.42	7.89E-04	1.66	7.55E-01	0.99	3.78E-01	0.87	1.83E-01	0.89	9.78E-01	1.00
rs2412101	17	44896220	G/A	NGFR	0.10	1.52E-03	1.36	2.41E-02	1.80	4.40E-01	0.95	5.27E-01	0.88	2.82E-01	0.87	4.53E-01	0.97
rs410852	19	59492183	G/A	LILRA3	0.08	5.83E-03	0.6	1.98E-01	0.69	4.04E-01	1.08	9.49E-01	1.02	7.24E-02	1.14	2.73E-01	1.06
rs6057007	20	9837990	G/A	PAK7	0.20	8.54E-04	0.81	3.57E-01	0.83	2.87E-01	0.95	4.27E-01	0.87	1.90E-01	0.88	4.62E-02	0.92
rs656111	20	9978931	C/A	ANKRD5	0.33	1.52E-04	1.5	5.94E-02	1.26	4.99E-01	0.97	3.44E-01	1.12	8.03E-01	0.98	6.29E-01	1.01
rs1119133	20	11251203	A/G	LOC441940	0.34	6.41E-04	0.74	6.46E-02	0.78	7.77E-01	0.99	7.40E-01	1.04	3.45E-01	0.93	2.54E-01	0.96
rs1109400	20	60439800	G/A	C20orf151	0.11	9.46E-04	1.43	5.93E-01	1.16	9.43E-01	1.00	4.90E-01	0.87	4.03E-01	1.12	4.05E-01	1.04
rs9974808	21	32307639	A/G	HUNK	0.33	1.89E-05	0.75	6.45E-03	0.67	1.41E-02	0.89	3.35E-01	0.87	4.74E-01	0.94	3.66E-04	0.87
rs12166331	22	41451370	A/G	A4GALT	0.35	1.32E-02	1.24	5.84E-02	1.34	1.60E-01	0.94	9.80E-01	0.98	9.94E-01	1.00	7.82E-01	0.99
rs3788622	22	43566454	G/A	ARHGAP8	0.43	3.27E-05	1.2	6.50E-02	1.31	3.48E-01	1.04	1.04E-01	0.82	2.18E-01	1.10	1.11E-01	1.06

NA: no GWAS data

**Supplemental Table 2** Results of the conditional association of the confirmed loci

Region	SNP	Chr	Position	Allele	Nominal P value	Conditional Pvalue (Top SNP)
MHC	rs7759127	6	31348967	C/A	1.50E-07	2.74E-06
	rs2524070	6	31352499	A/G	1.97E-06	2.43E-05
	rs2853930	6	31363403	C/A	4.74E-07	9.64E-06
	rs2524132	6	31372891	A/G	1.95E-07	3.73E-06
	rs9264868	6	31379580	G/A	1.18E-07	3.51E-06
	rs28490179	6	32626983	G/A	1.42E-13	2.46E-01
	rs9270856	6	32678817	A/G	3.56E-17	9.55E-02
	rs9271348	6	32691720	G/A	1.71E-15	1.94E-02
	rs9271366	6	32694832	G/A	1.46E-17	*
	rs35367950	6	32732638	A/G	2.48E-16	3.89E-03
Chr 8	rs42490	8	90847650	A/G	1.12E-19	*
	rs40457	8	90892832	G/A	2.36E-15	1.19E-06
Chr 9	rs10982385	9	116532838	C/A	1.03E-08	1.96E-03
	rs4574921	9	116578155	G/A	5.20E-18	4.34E-03
	rs10114470	9	116587593	G/A	5.92E-17	4.76E-01
	rs6478108	9	116598524	A/G	1.44E-24	*
Chr 13	rs9533634	13	43295815	G/A	3.91E-15	7.48E-02
	rs3088362	13	43331630	A/C	1.05E-34	7.28E-04
	rs3764147	13	43355925	G/A	6.56E-60	*
	rs10507522	13	43377000	G/A	2.84E-28	5.77E-05
Chr 16	rs9302752	16	49276604	G/A	3.73E-44	*
	rs7194886	16	49282694	A/G	9.28E-33	9.52E-04
	rs8057341	16	49295481	G/A	1.26E-06	1.24E-04
	rs3135499	16	49323628	C/A	5.88E-06	3.01E-04

<sup>a</sup> Minor allele given first, major allele second.

\* Top SNP/2 SNPs used in the conditional analysis of other SNPs in the same region.

**Supplement Table 3.** The association results of the SNPs around PARK2/PACRG and LTA from the GWAS analyses

snp.idx	chr	pos	allele	pval_1931	OR	pval_1220	OR
rs10806762	6	163014913	A/C	2.19E-01	1.11	2.19E-01	1.11
rs2803086	6	163020585	A/G	5.35E-01	0.94	5.35E-01	0.94
rs4510651	6	163027087	A/G	4.39E-01	0.93	4.39E-01	0.93
rs1893537	6	163042210	A/C	2.09E-01	1.12	2.09E-01	1.12
rs2849525	6	163046642	A/C	9.32E-01	1.03	9.32E-01	1.03
rs10806763	6	163052519	A/G	2.19E-01	1.11	2.19E-01	1.11
rs2849518	6	163054238	A/G	9.27E-01	1.03	9.27E-01	1.03
rs2156259	6	163062262	A/G	3.82E-01	0.91	3.82E-01	0.91
rs2276201	6	163069487	A/G	3.45E-01	0.90	3.45E-01	0.90
rs9347684	6	163071814	A/G	2.07E-01	1.12	2.07E-01	1.12
rs13195186	6	163079177	A/G	2.74E-01	0.89	2.74E-01	0.89
rs6900416	6	163082050	A/G	2.87E-01	0.89	2.87E-01	0.89
rs7764309	6	163082340	A/C	7.68E-01	1.10	7.68E-01	1.10
rs4568424	6	163091786	A/C	6.97E-01	0.96	6.97E-01	0.96
rs9347686	6	163091960	A/G	2.91E-01	0.91	2.91E-01	0.91
rs7744306	6	163096141	A/G	3.58E-01	1.08	3.58E-01	1.08
rs9458643	6	163101139	A/G	6.56E-01	0.96	6.56E-01	0.96
rs10806767	6	163132289	A/G	6.92E-01	0.96	6.92E-01	0.96
rs1333955	6	163133444	A/G	7.52E-01	0.97	7.52E-01	0.97
rs1041632	6	163139881	A/G	5.33E-01	1.06	5.33E-01	1.06
rs4458655	6	163141782	A/G	7.63E-01	1.04	7.63E-01	1.04
rs1333958	6	163153251	A/C	9.28E-01	1.01	9.28E-01	1.01
rs7758664	6	163161043	A/G	9.12E-01	1.01	9.12E-01	1.01
rs6917499	6	163168579	A/G	7.60E-01	1.03	7.60E-01	1.03
rs9347691	6	163169635	A/G	8.02E-01	0.98	8.02E-01	0.98
rs949902	6	163170768	A/C	1.83E-01	0.80	1.83E-01	0.80
rs1333962	6	163188894	A/G	7.23E-01	1.03	7.23E-01	1.03
rs9365499	6	163191823	A/G	2.19E-01	0.89	2.19E-01	0.89
rs9456814	6	163193183	A/G	7.17E-01	1.03	7.17E-01	1.03
rs9456815	6	163193316	A/G	3.81E-01	1.10	3.81E-01	1.10
rs12665063	6	163194296	A/G	1.50E-01	1.19	1.50E-01	1.19
rs9346937	6	163199969	A/G	2.32E-01	0.90	2.32E-01	0.90
rs4709655	6	163200194	A/G	1.90E-01	0.81	1.90E-01	0.81
rs9346938	6	163206342	A/G	5.40E-01	0.95	5.40E-01	0.95
rs13215569	6	163207269	A/G	1.78E-01	1.68	1.78E-01	1.68
rs7760425	6	163209057	A/G	7.42E-01	0.97	7.42E-01	0.97

※This association locus was observed in the regulatory region shared by PARK2 and PACRG in Vietnamese and Brazilian samples. 13 SNPs listed in the original paper located at 6q25-26(about from 163,000,000 to 163,150,000).

**All the SNPs located at LTA gene in 610 chips**

snp.idx	chro	pos	allele	gene	pval_1931	OR	pval_1220	OR
rs1041981	6	31648763	A/C	LTA coding	4.10E-02	0.835929	4.10E-02	0.835929
rs1800683	6	31648050	A/G	LTA 5'	4.07E-02	0.835643	4.07E-02	0.835643
rs2009658	6	31646223	C/G	LTA 5'	3.50E-01	1.118316	3.50E-01	1.118316
rs2229092	6	31648736	A/C	LTA coding	2.57E-02	2.062517	2.57E-02	2.062517
rs2229094	6	31648535	A/G	LTA coding	5.38E-01	1.068578	5.38E-01	1.068578
rs2844482	6	31647746	A/G	LTA 5'	3.90E-01	1.108345	3.90E-01	1.108345

rs2844484	6	31644203	A/G	LTA 5'	1.44E-01	1.138388	1.44E-01	1.138388
rs2857708	6	31641585	A/G	LTA 5'	8.18E-02	1.264562	8.18E-02	1.264562

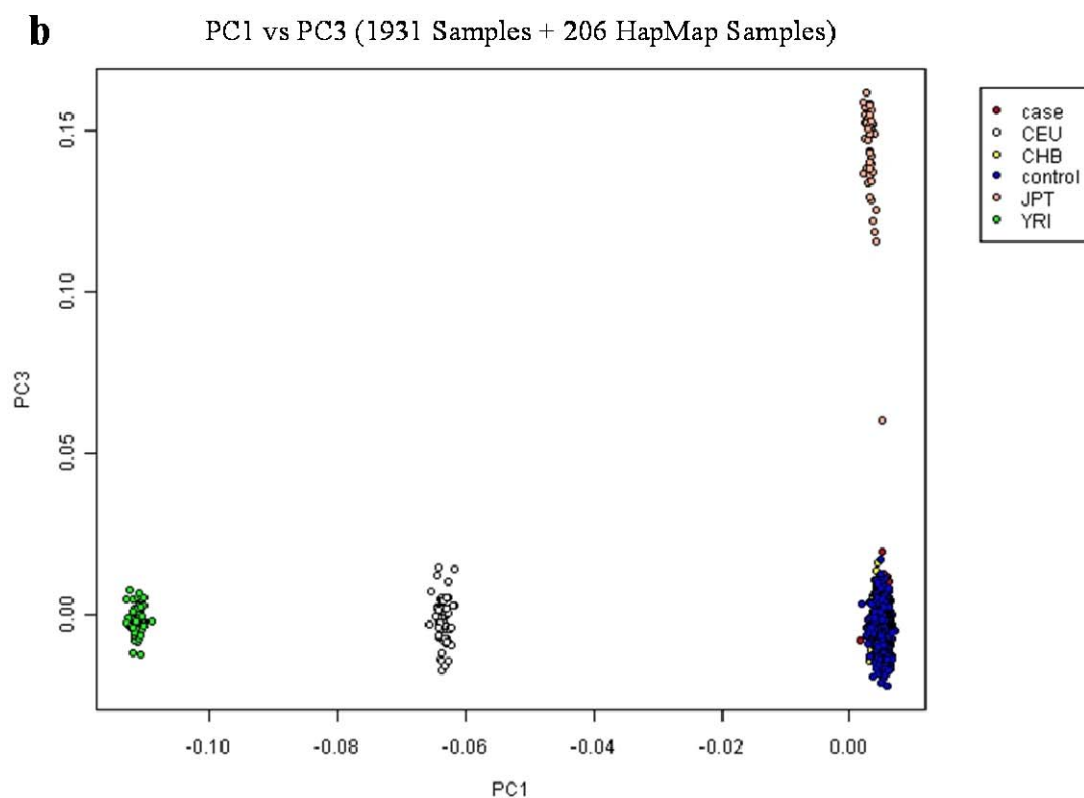
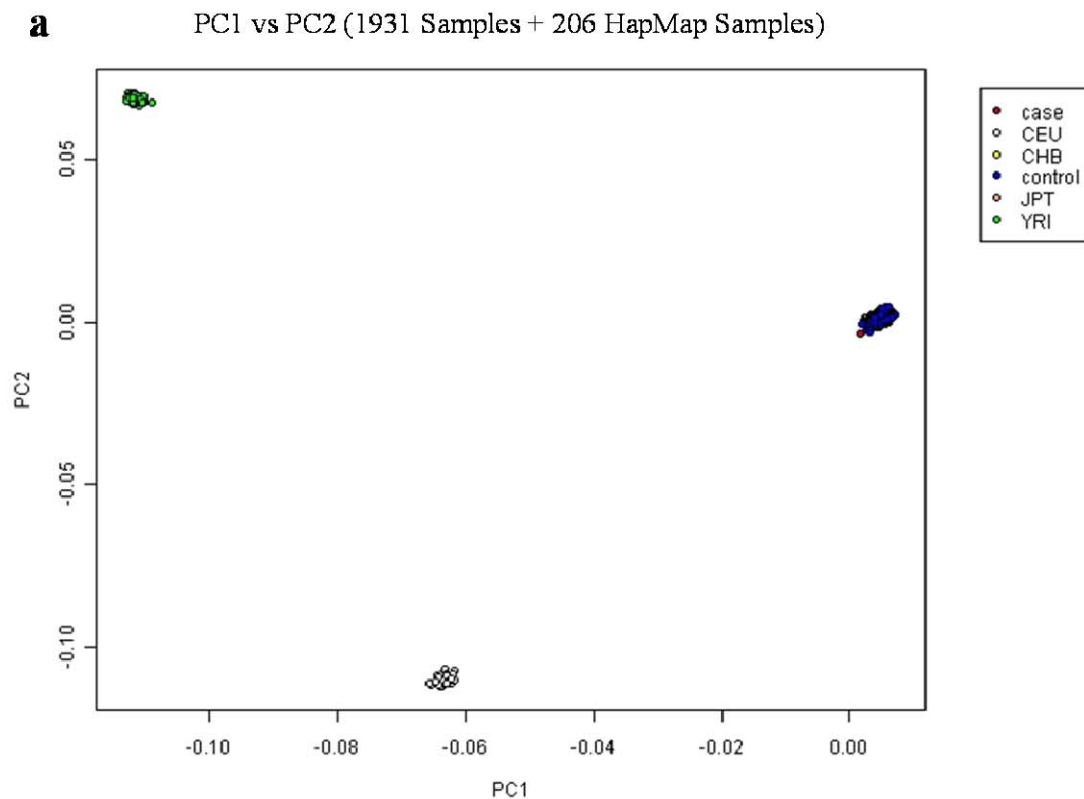
※In the original paper, only LTA+80(rs2239704) was indentified to be associated with leprosy.

**Supplement Table 4. The association results of top 10 SNPs within 10p13 from the GWAS analysis**

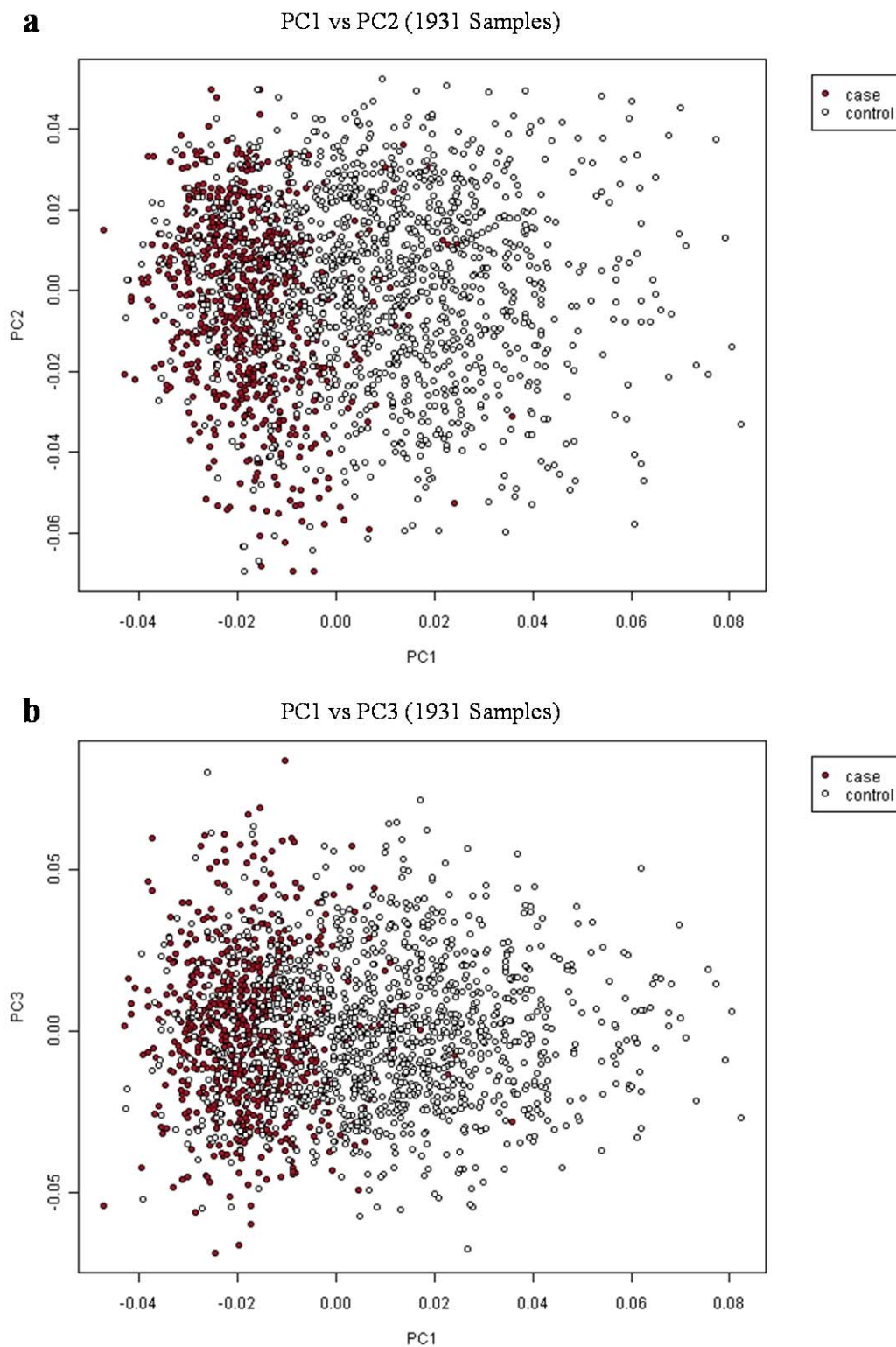
CHR	SNP	Poc	P_GWAS	OR_all	P_pauci	OR_pauci
10	rs1571212	12548962	2.32E-03	0.71	6.88E-02	0.81
10	rs10906156	12569201	7.34E-03	0.65	8.41E-02	0.75
10	rs7085304	13347283	6.73E-03	1.25	1.71E-03	1.29
10	rs11258729	14018379	6.53E-03	1.92	1.76E-01	1.33
10	rs11258736	14021197	9.97E-03	1.84	1.96E-01	1.31
10	rs10906548	14023013	6.28E-03	1.91	1.16E-01	1.38
10	rs10906549	14028972	2.47E-03	2.05	5.84E-02	1.47
10	rs12260246	15556408	2.16E-03	0.69	6.13E-02	0.79
10	rs12251305	15557286	3.61E-03	0.7	6.94E-02	0.8
10	rs2167451	15559928	9.66E-03	0.79	2.25E-02	0.8
10	rs753538	16478388	2.24E-03	1.37	1.59E-01	1.16
10	rs1334592	16487010	1.06E-03	1.58	6.62E-02	1.27
10	rs10904861	17070781	2.36E-03	0.7	5.05E-03	0.71
10	rs2291521	17072287	8.18E-04	0.68	3.36E-03	0.7
10	rs12414709	17081089	7.70E-03	0.76	4.46E-02	0.81

※The result of top SNPs in GWAS\_pauci located at 10p13

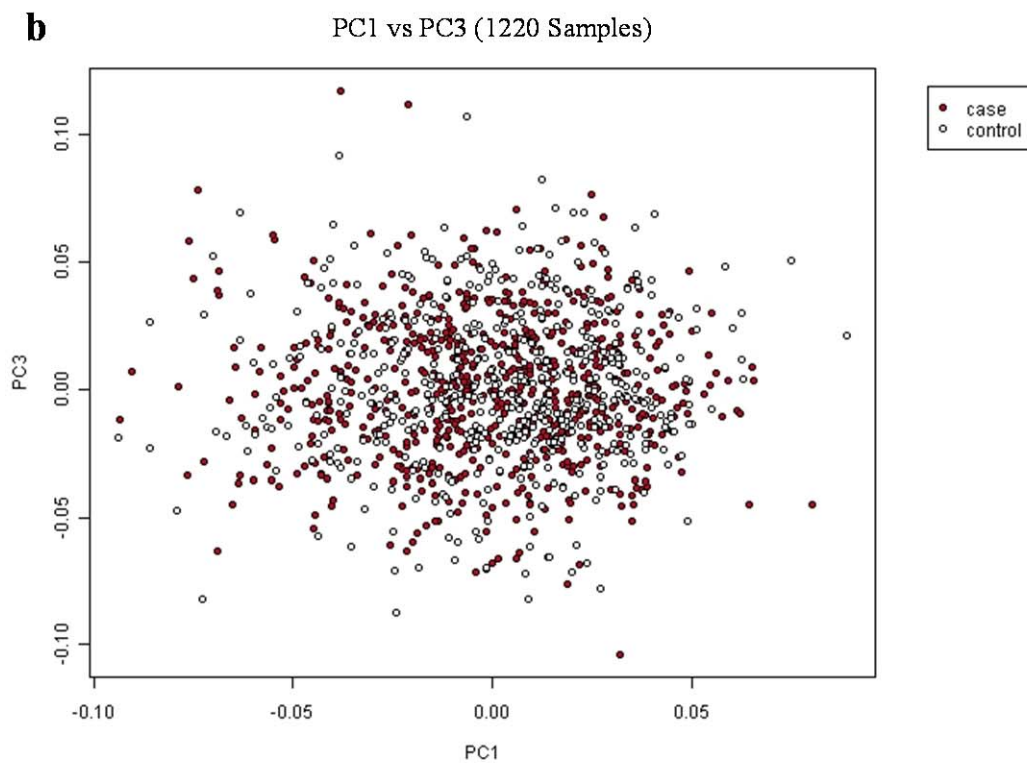
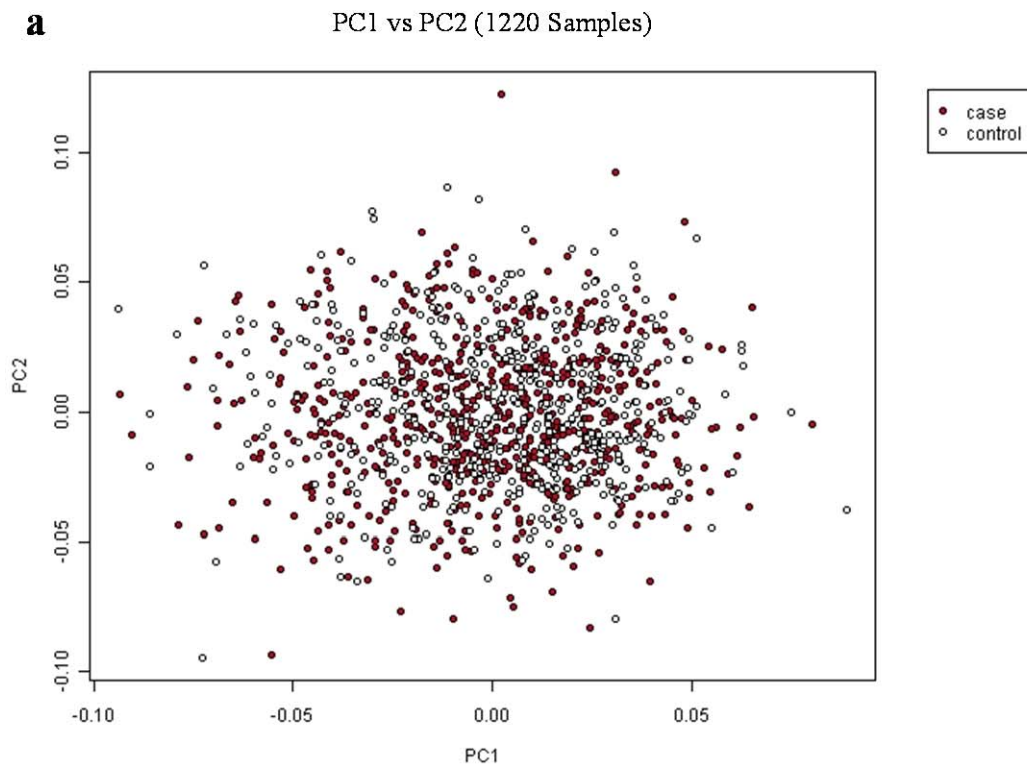
**Supplement Figure 1.** Plots of first three principal components from the principal components analysis using 1931 participant samples and 206 HapMap samples. Cases are in orange, controls are in light blue, CHB samples are in blue, JPT samples are in white, CEU samples are in yellow and YRI samples are in green. a: plot of the first and second principal components; b: plot of the first and third principal components.



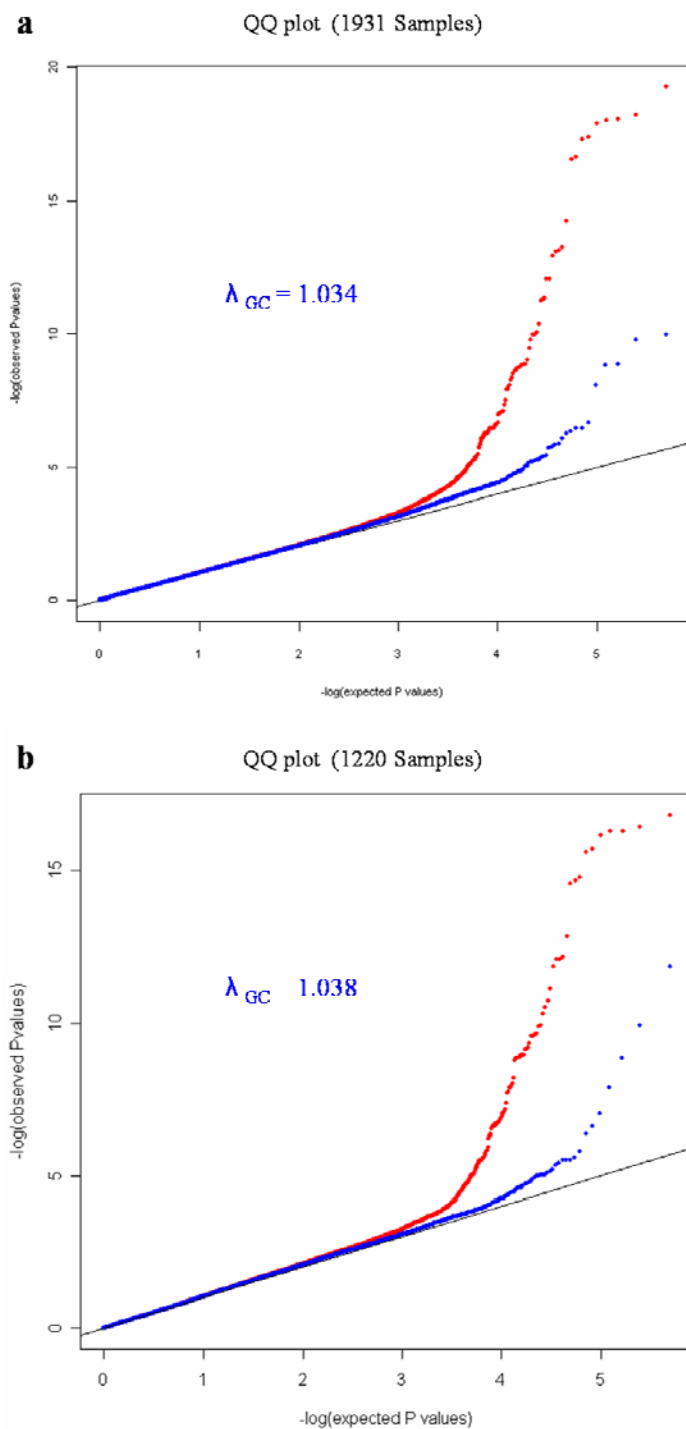
**Supplement Figure 2.** Plots of first three principal components from the principal components analysis using 1931 participant samples - 706 cases, 1225 controls. a: plot of the first and second principal components; b: plot of the first and third principal components. The red points are cases and the white points are controls.



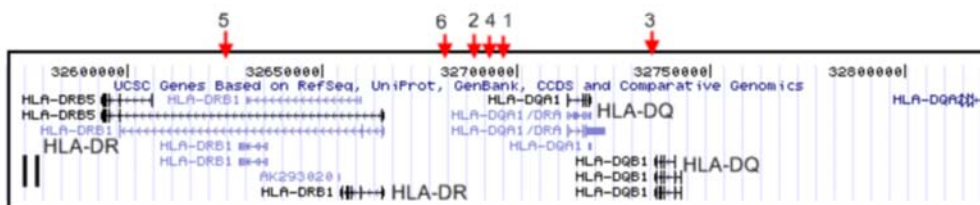
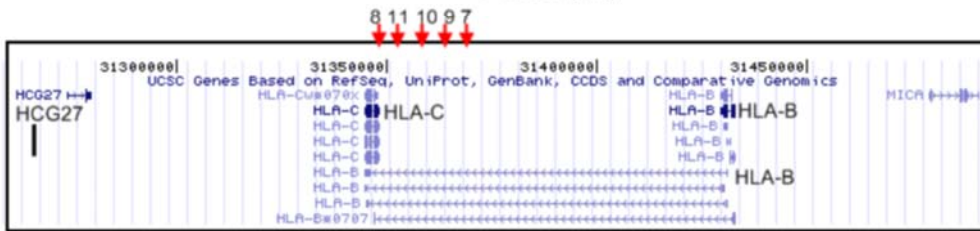
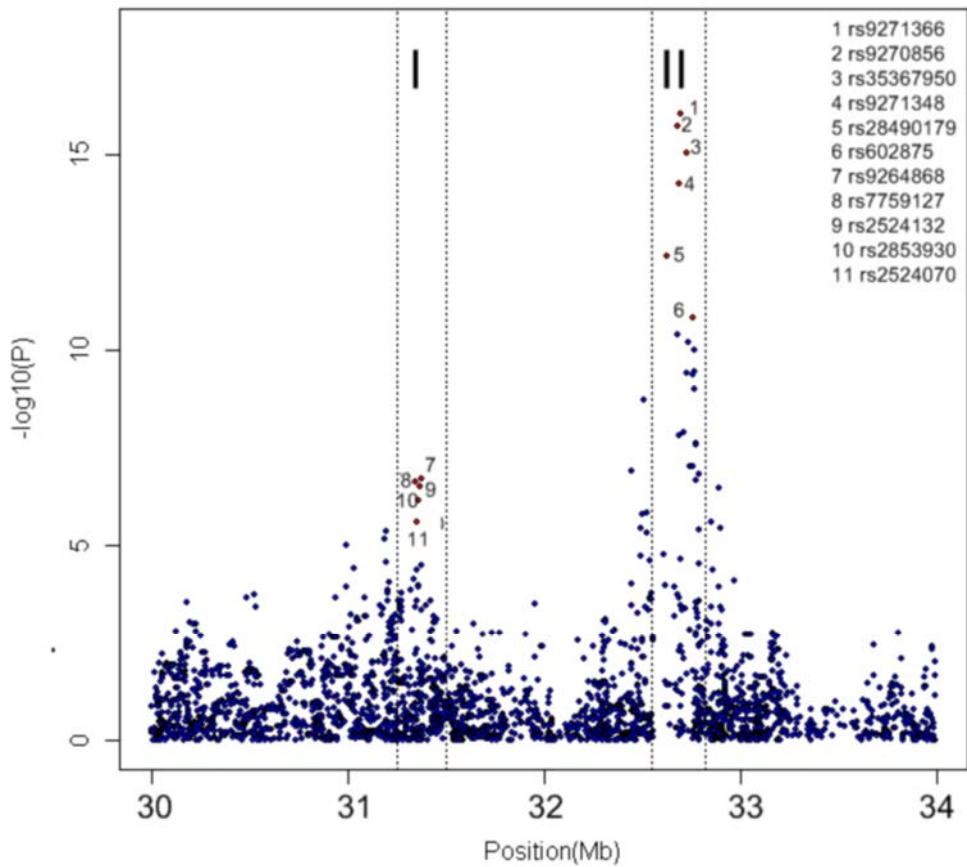
**Supplement Figure 3.** Plots of first three principal components from the principal components analysis using 1220 matched samples - 706 cases, 514 controls. a: plot of the first and second principal components; b: plot of the first and third principal components. The red points are cases and the white points are controls.



**Supplement Figure 4:** a. The quantile-quantile (Q-Q) plots of the eigenstrat-corrected P values for association in 1931 samples. The plot in red is for the P values from all the 491,833 SNPs, whereas the plot in blue is for P values for SNPs excluding the 4,637 SNPs within the MHC region (chr. 6: 25–34 Mb). b. The quantile-quantile (Q-Q) plots of the observed P values for association in the matched 1220 samples. The plot in red is for the P values from all the 492,109 SNPs, whereas the plot in blue is for P values for SNPs excluding the 4,679 SNPs within the MHC region (chr. 6: 25–34 Mb).  $\lambda_{GC}$  shown in both plots was calculated by removing SNPs within the MHC region.



**Supplement Figure 5.** Association results within the MHC region from the GWAS\_1220 analysis. The P values (trend test) of all the MHC SNPs within chr 6:25-34 Mb from the GWAS analysis in 1220 samples are presented. The upper panel shows two association signals within the MHC region. The top six SNPs within HLA class II region and the top five SNPs within HLA class I region showing the smallest p value are listed. All the eleven SNPs are colored by red. The bottom panels show the two genomic regions by listing the nine known genes and the eleven SNPs. The positions of eleven SNPs are indicated by red arrows.



**Supplement Figure 6:** The association results ( $-\log_{10}(P)$ ) and LD patterns of the 7 identified non-MHC loci in Chinese population. The LD patterns ( $D'$ ) were created in the Haploview by using the genotyping data (only SNPs with  $MAF > 0.01$ ) from the Hapmap project. For each locus, the upper panel shows the association results of all the SNPs within each region with the validated SNPs to be indicated in red. The bottom panel shows the LD ( $D'$ ) pattern in Chinese population, and the positions of all the validated SNPs are indicated by red arrows. (6a): RIPK2 at 8q21 (6b): TNFSF15 at 9q32 (6c): LRRK2 at 12q12. (6d): CCDC122 and C13orf31 at 13q14 (6e): NOD2 at 16q12

