Supplementary Appendix

Table of contents

MATERIALS AND METHODS

Study Subjects

The genome-wide discovery analysis of protein coding variants (Stage 1) included 1,670 individuals with leprosy and 2,321 control individuals from the northern region of China. The initial validation analysis (Stage 2) was done in additional 3,169 leprosy patients and 9,814 healthy controls from the northern region of China. And, the final replication analysis (Stage 3) was carried out in three independent samples from the southern region of China: (i) 906 individuals with leprosy and 878 control subjects from Sichuan province (Replication 1); (ii) 829 individuals with leprosy and 589 control subjects from Yunnan province (Replication 2) and (iii) 496 individuals with leprosy and 799 control subjects from Guizhou province (Replication 3).

Genotyping and quality control in the discovery stage

We carried out genome wide association study of protein coding variants using Illumina Infinium Human Exome Bead Chip (v1.0) array, which includes 242,102 putative functional coding variants discovered from > 12,000 exome and genome sequencing of multiple ethnicities and complex traits. The details of the variants and selection strategies are described on the exome array design webpage (http://genome.sph.umich.edu/wiki/Exome_Chip_Design). This array was augmented with additional 27,089 rare, recurrent (found in > 1 sample) non-synonymous variants detected by the whole-exome sequencing of 1,998 Chinese subjects (data not shown).

1,648 cases and 2,318 controls passed the sample quality control filters and were used in the discovery analysis. Of which, 802 cases and 980 controls were overlapped with previous genome-wide association studies (GWAS) (Wang et al., 2016), 846 cases and 1,338 controls were newly recruited.

Statistical analysis

3,984 samples after the familial relationship checking were assessed for population outliers and stratification by using a principal component analysis (PCA)-based approach. Firstly, 450 samples from 1000 Genome Project phase I (85 Utah residents

with Northern and Western European ancestry (CEU); 88 Yoruba in Ibadan, Nigeria (YRI); 89 Japanese in Tokyo, Japan (JPT); 97 Han Chinese in Beijing, China (CHB), and 91 Southern Han Chinese, China (CHS)) were analyzed together with our 3,984 samples. Variants in the 20 long-range linkage disequilibrium regions (Price et al., 2008) and variants with MAF < 1% were excluded, followed by LD pruning in PLINK v1.07 (Purcell et al., 2007), with the command option --indep-pairwise 1500 150 0.2. 20,588 variants remained for PCA analysis. 18 outliers were found (16 cases and two controls) and removed **(see Supplementary Figure S1).** Secondly, the PCA, using the same SNP filtering criteria, was applied to the 3,966 samples. No outliers or obvious population stratification were found at this round **(See Supplementary Figure S2).**

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). SNPs with MAF >= 1% were used to calculate the genetic relationship matrix (GRM). Association P values were assessed by the score test (glmm.score()) and odds ratio for the top 39 SNPs were estimated by the Wald test (glmm.wald()) in GMMAT.

In validation and replication stages, we included 5,400 cases and 12,080 controls (where 5,387 cases and 12,021 controls overlapped with previous GWAS studies (Wang et al., 2016)). We tested associations between phenotypes and single-variant genotypes using PLINK. Logistic regression model was used for those SNPs with MAF > 1%. Fisher's exact Test in PLINK was used in the association analysis of those SNPs with MAF < 1%.

The meta-analysis of the combined discovery + validation samples (4,817 cases and 12,132 controls) or discovery + validation + replication samples (7,048 cases and 14,398 controls) was performed using a fixed effect model (inverse variance method for SNPs with MAF > 1% and z-statistics combination method for SNPs with MAF < 1% in META_v.1.7.0 (Liu et al., 2010) software). Some SNPs that had discovery P-value < 1.0×10^{-3} (our validation selection threshold) located in the same LD block as the previously reported SNPs. To assess the independence of significant or

suggestive coding variants within the previously reported GWAS loci, we performed the conditional GMMAT analysis by setting the reported GWAS SNPs in the same region as co-variants when fitting the null model in GMMAT. The conditional GMMAT analysis was performed by using 802 cases and 980 controls that are overlapping between the current study and the previous GWAS (Wang et al., 2016).

We carried out gene-based test using SKAT-O (Lee et al., 2012) and burden test software. The gene-based tests were performed on all the QCed coding variants (including the rare variants whose frequency in all samples < 0.1%) after excluding all the variants within the MHC region as well as the coding variants with suggestive association ($P < 1.0 \times 10^{-3}$) in single variant-based association analysis. To accommodate the potential population stratification, we set the first principal component as co-variant in the SKAT-O and burden test. We performed two types of analyses: the first one was based on the coding variants with MAF < 5%, while the second was based on the coding variants with MAF < 1%. 10,643 and 10,054 genes with at least two variants were tested in the first and second types respectively. A gene-based association test was defined as exome-wide significant if the nominal P value was $< 4.7 \times 10^{-6}$, corresponding to a Bonferroni correction for 10,643 gene tests.

Variant selection for the validation analysis

For variants within previously reported GWAS loci, we first shortlisted those meeting the following criteria: 1) LD <= 0.2 with previously reported GWAS variants; 2) P-value $< 1 \times 10^{-3}$ in the discovery stage. Only one variant met the above criteria after checking the corresponding cluster plots. Secondly, for the above remaining SNPs, we performed unconditional and conditional analyses within the previously reported GWAS locus using 802 cases and 980 controls which overlapped with the previous GWAS data (see above). The difference between the adjusted and raw OR was smaller than 10% of the raw OR **(see Supplementary Table S1)**, showing that the current candidate SNP and the previously reported GWAS SNP were uncorrelated.

For variants outside the previously reported GWAS loci, the independent SNPs

with P-value or conditional P-value < 1×10^{-3} in a LD block were selected for validation. 38 variants met the above criteria after checking the corresponding cluster plots. Totally, 39 variants were followed in the validation stage.

Biological and Network Analysis

Protein structural analysis

The crystal structure templates of *IL23R, IL27* and *SLC29A3* were identified using BLAST against PDB with e-value of 1.0×10^{-10} in ANNOTATOR (Houten et al., 2003). Homology modeling based on the templates was performed using Modeller (Eswar et al., 2008) without and with loop refinement (Memoir (Ebejer et al., 2013) was used for *SLC29A3*). Homology models for the three genes, after checking their accuracy, and human crystal structure of *TYK2* were used to calculate the free energy change induced by the mutation of interest by FoldX plugin in YASARA (Van Durme et al., 2011). The average free energy change (ddG, kcal/mol) was obtained by running Foldx 5 times. Average ddG changes are viewed as significant if its absolute value is over 0.5 kcal/mol; positive free energy changes imply destabilizing the protein structure, while negative ones imply the stabilization.

GO-enrichment Analysis

GO-enrichment Analysis was performed with seven-gene set and 33-gene set, respectively. Seven-gene set include *IL23R, FLG, NCKIPSD, CARD9, SLC29A3, IL27* and *TYK2* from the current study. 33-gene set include seven genes from the current study, and 26 reported genes (*HLA-DRB1, LACC1, NOD2, PIPK2, TNFSF15, RAB32, IL12B, IL18R1, IL18RAP, BCL10, DEC1, BATF3, CCDC88B, CDH18, CIITA, EGR2, SOCS1, IL1RL1, GRM1, TNFSF8, HLA-DQB1, SYN2, BBS9, CTSB, MED30, PPARG*) from previous GWAS (Liu et al., 2013, Liu et al., 2015, Liu et al., 2012, Wang et al., 2016, Zhang et al., 2011, Zhang et al., 2009).

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 were listed in the supplement **(Supplementary Table S9).**

The enriched GO hierarchical subgraph (Alexa et al., 2006) induced from these top 30 significant GO terms was also shown in the supplement **(Supplementary Figure S11).**

Integrated Network Enrichment

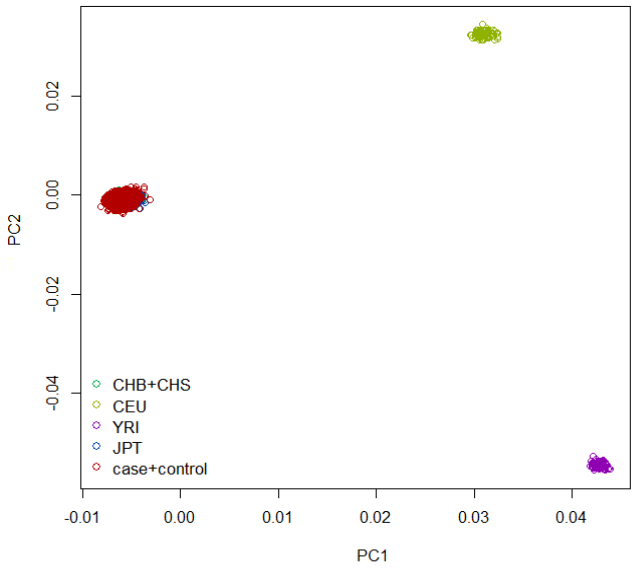
We analyzed the 33-gene set using GeneMANIA in Cytoscape with the default settings and identified a highly interactive gene network **(Supplementary Figure S12)** where all the significant sub-networks/functions are related to immunity and can be grouped into two clusters of innate and adaptive immunities. For each cluster, the top 15 significant biological processes/pathways were shown in **Supplementary Table S10,** and the top sub-network with the most significant (smallest false discovery rate (FDR)) functional annotation was highlighted in **Supplementary Figure S12.**

AUTHOR CONTRIBUTIONS

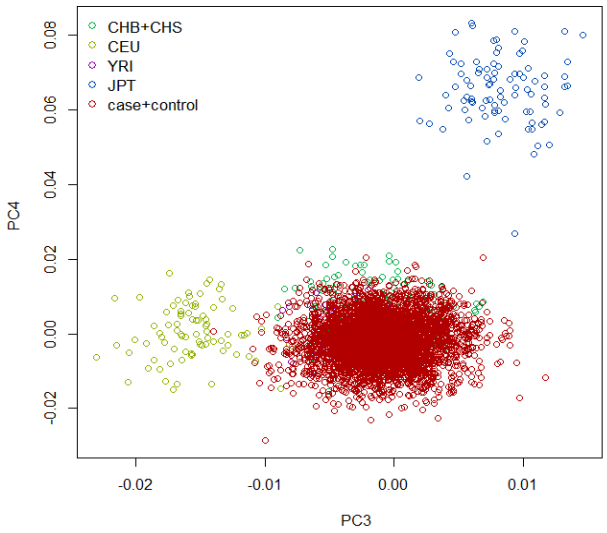
Z. F. R. conceived of this study and obtained the financial support. Z. F. R. and L. J. J. designed the study. C. S. M. Liu Jian, C. T. S., Y. M. W., C. C. K., K. S. S., Aung Tin, W. N. L., W. D. Y., Shi Li, Ning Yong, Z. Z. Y., Y. R. D., L. J. L., Yang Jun, Z. G. Z., Y. L. B. and S. J. P. undertook recruitment and collected phenotype data. Liu Hong, W. Z. Z. undertook related data handling and calculation, managed recruitment and obtained biological samples. Liu Hong, W. Z. Z., Y. G. Q., F. X. A., Wang Chuan, Y. Y. X., B. F. F., W. H. L., M. Z. H., S. Y. H., S. L. L., Y. J. B., L. J. H., N. G. Y., Y. Z. H., Zhao Qing, Wang Na, Y. W. J., C. X. J. conducted sample selection and performed the genotyping of all samples. L. J. J., W. Z. Z., L. Y., L. W. T., A. I., W. L., F. J. N., H. L., W. C. L., W. Y. M., S. M. S. and V. L. undertook data checking, statistical analysis and bioinformatics analyses. A. K. A. conducted out the eQTL analysis. Liu Hong was responsible for sample selection, genotyping and project management. Liu Hong and W. Z. Z. wrote the first draft. L. J. J., Z. F. R. and Li Yi revised the draft. All authors contributed to the final manuscript, with Z. F. R., L. J. J., Liu Hong, W. Z. Z. and Li Yi playing the key roles.

FIGURES

a) PC1 vs PC2 (3,966 Samples + 450 samples from 1000 Genomes Project)



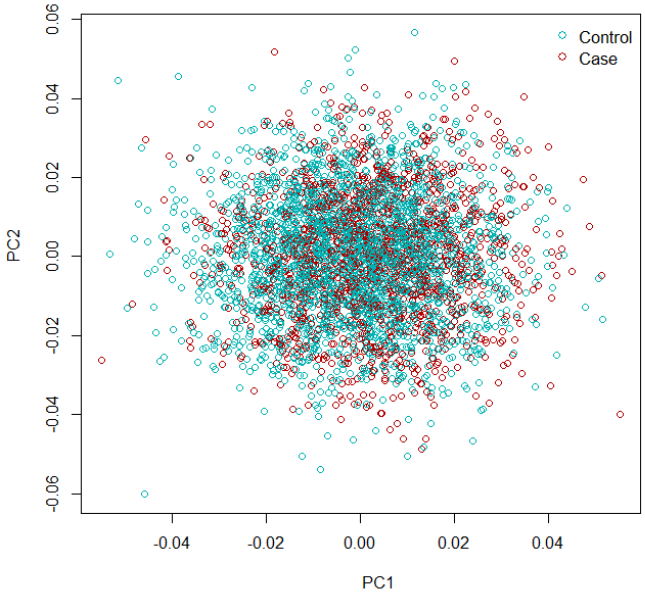
b) PC3 vs PC4 (3,966 Samples + 450 samples from 1000 Genomes Project)



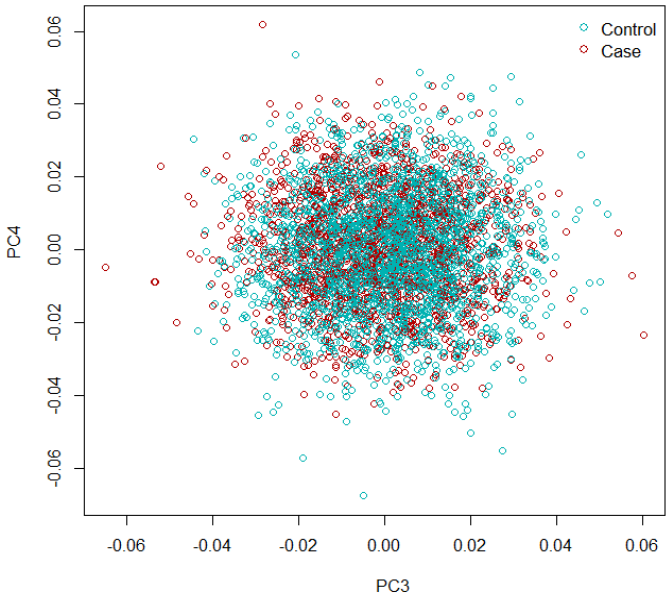
Supplementary Figure S1. PCA plot with 1000 Genomes Project samples.

Plots of first four principal components from the principal components analysis using 3,966 participant samples and 85 CEU samples, 88 YRI samples, 89 JPT samples, 97 CHB samples, 91 CHS samples from 1000 Genomes Project. a: plot of the first and second principal components; b: plot of the third and fourth principal components.

a) PC1 vs PC2 (1,648 Cases + 2,318 Controls)



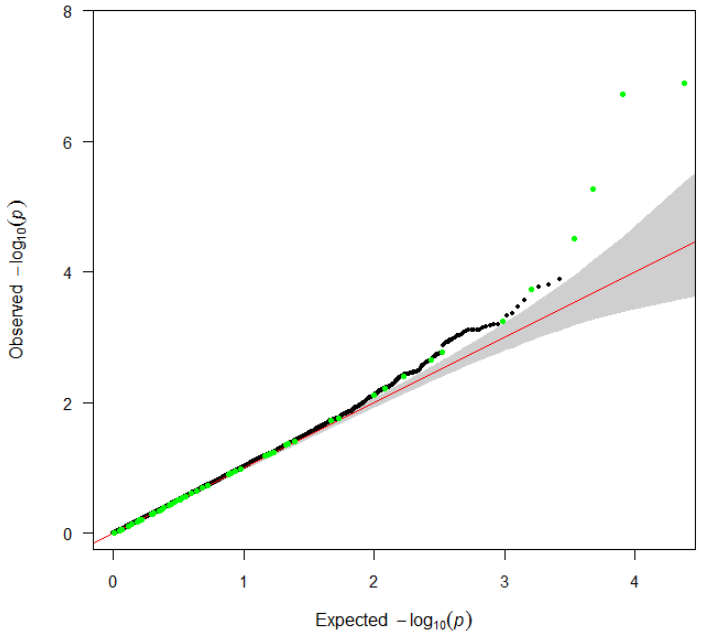
b) PC3 vs PC4 (1,648 Cases + 2,318 Controls)



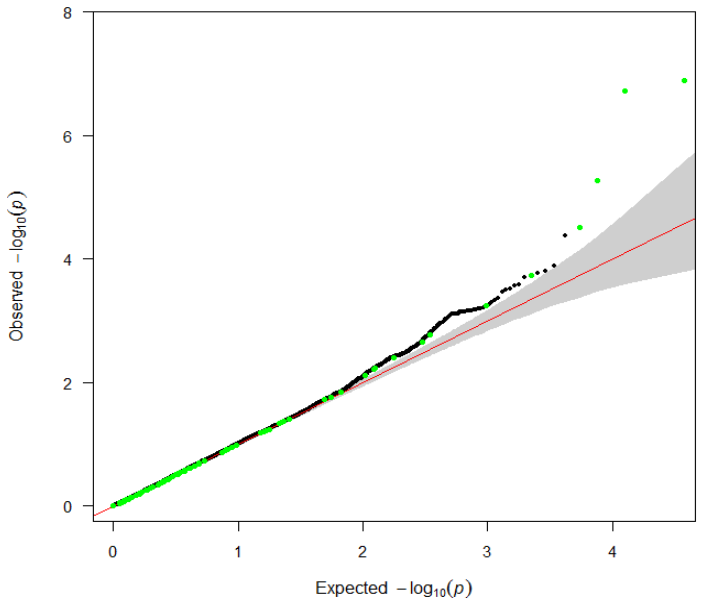
Supplementary Figure S2. PCA plot of 3,966 participant samples.

Plots of first four principal components from the principal components analysis using 1,648 cases, 2,318 controls. a: plot of the first and second principal components; b: plot of the third and fourth principal components. The red points are cases, the green points are controls.

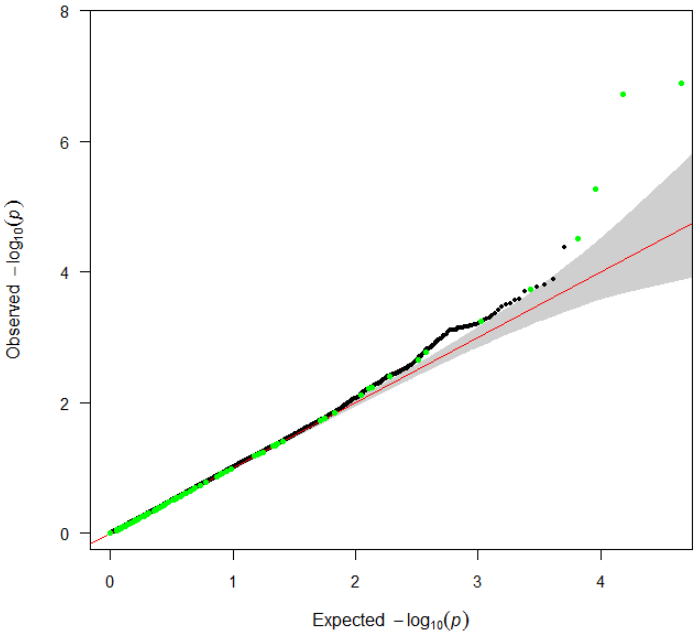
a) QQ plot for variants with MAF > 0.05



b) QQ plot for variants with MAF > 0.01



c) QQ plot for variants with MAF > 0.005

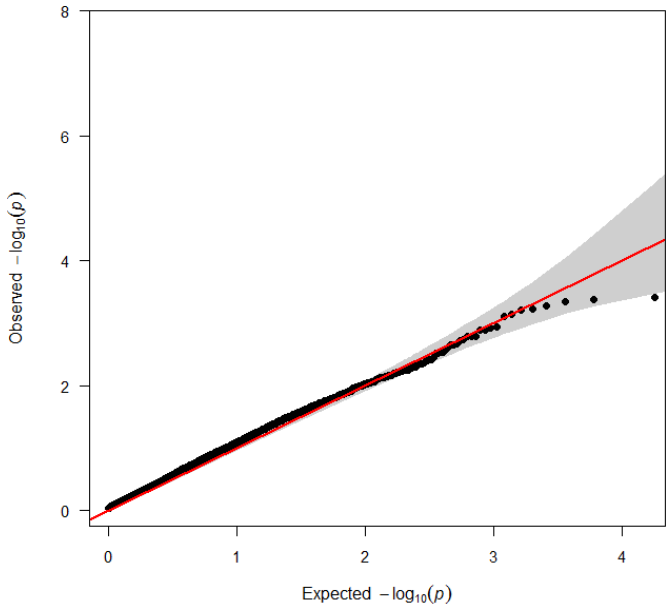


Supplementary Figure S3. Quantile-quantile plots of the P-values from GMMAT software.

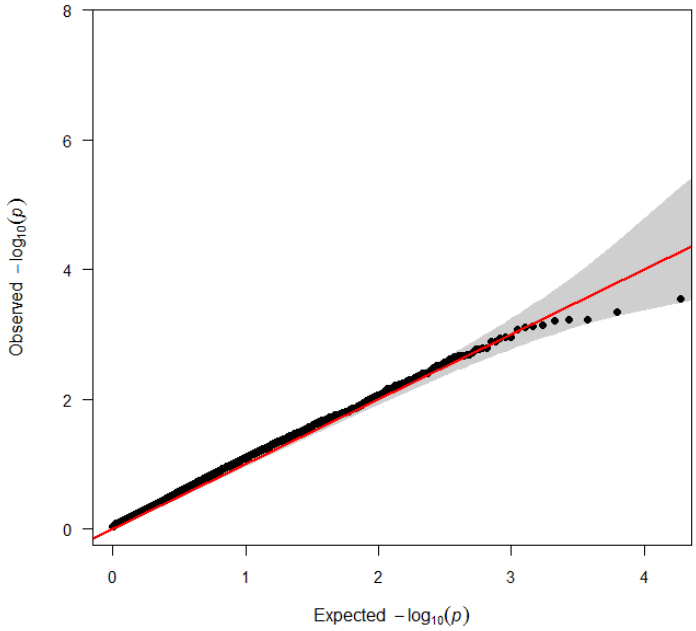
SNPs within the MHC region were removed, and SNPs in the previously reported GWAS loci were colored as green. PS: one SNP, rs3764147, which is a known GWAS SNP with p-value < 10^-20 in the current study, was removed for the ease of observation of the association signal.

a) QQ plot for 11,979 variants with MAF > 0. 05, lambdaGC = 0.99; b) QQ plot for 18,976 variants with MAF > 0. 01, lambdaGC = 0.97; c) QQ plot for 22,856 variants with MAF > 0.005, lambdaGC = 0.97.

a)



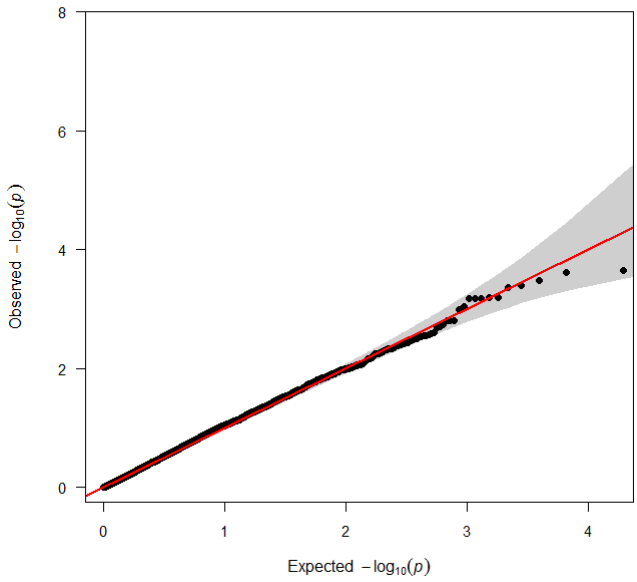
b)



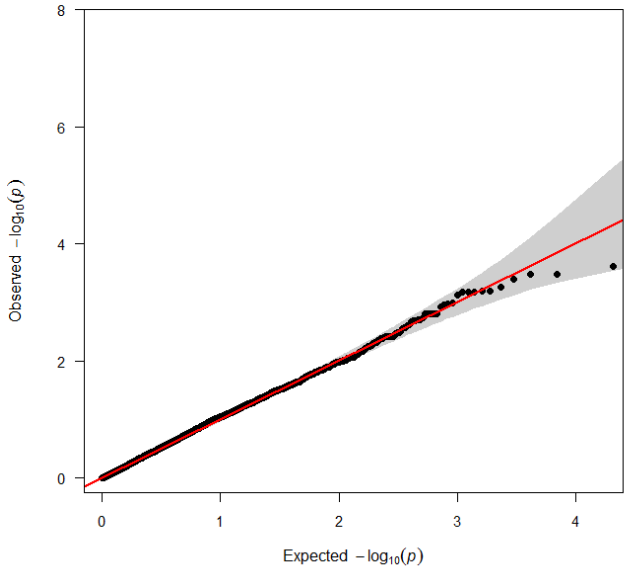
Supplementary Figure S4. Quantile-quantile plots of gene-based SKAT-O association P-values for the genes with at least two coding variants.

a) For variants with MAF < 1%, lambdaGC = 1.095; b) for variants with MAF < 5%, lambdaGC = 1.063.

a)



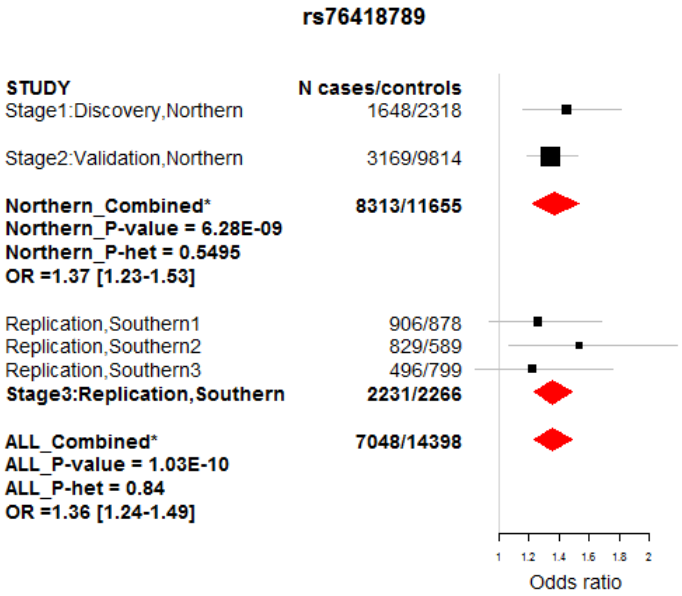
b)



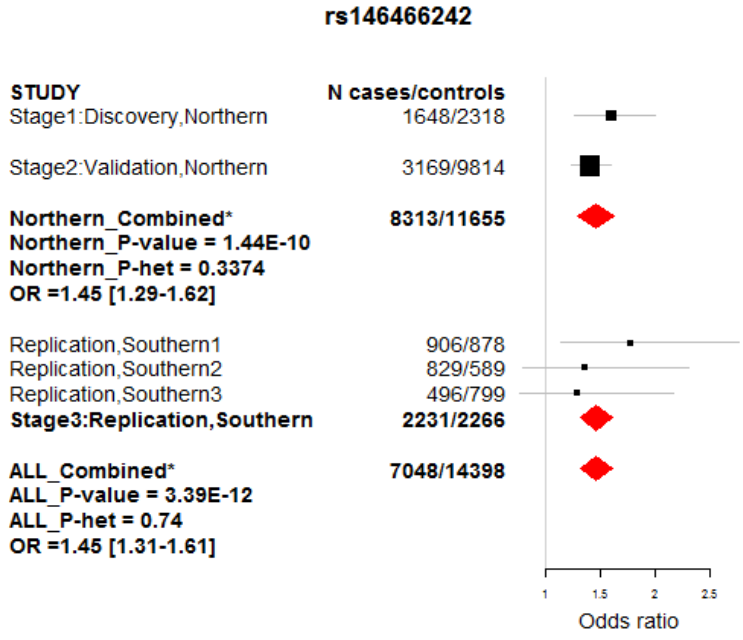
Supplementary Figure S5. Quantile-quantile plots of gene-based burden test association P-values for the genes with at least two coding variants.

a) For variants with MAF < 1%, lambdaGC = 1.073; b) For variants with MAF < 5%, lambdaGC = 1.10.

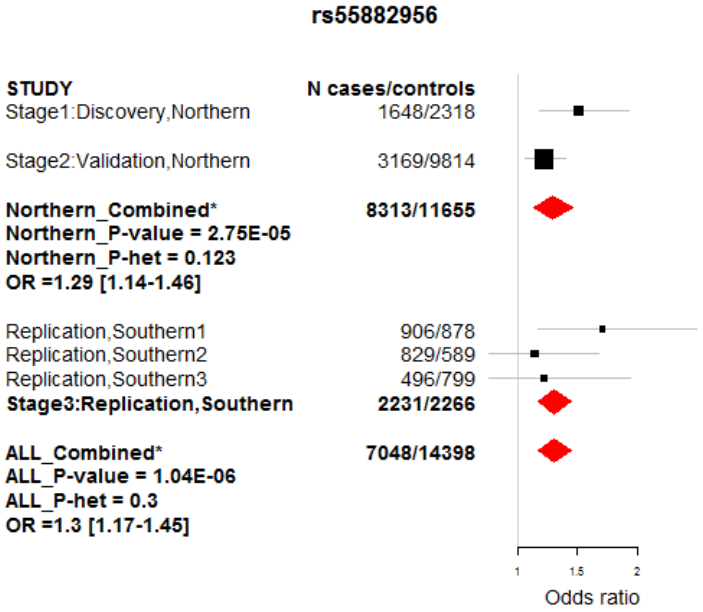
a: *IL23R* (low frequency)



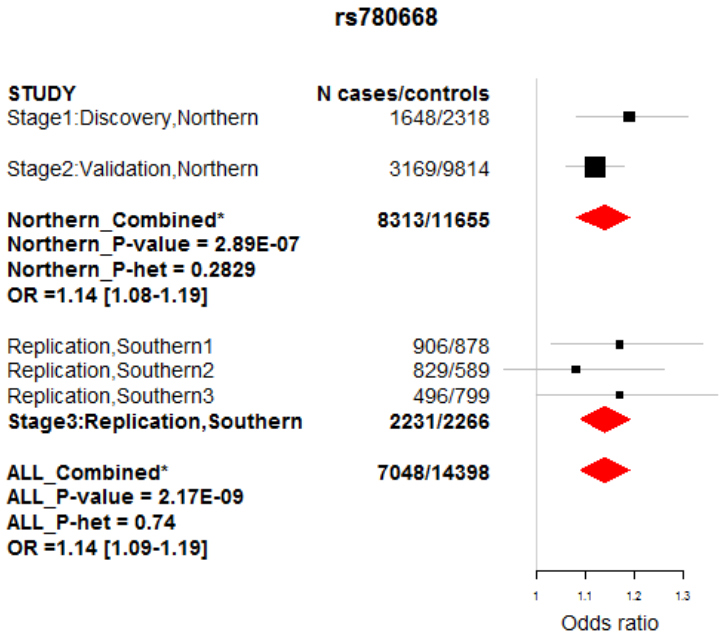
b: *FLG* (low frequency)



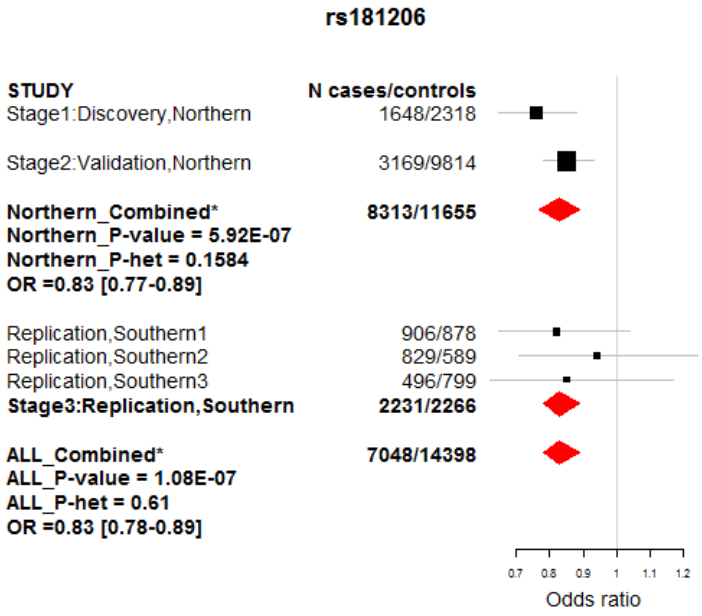
c: *TYK2* (low frequency)



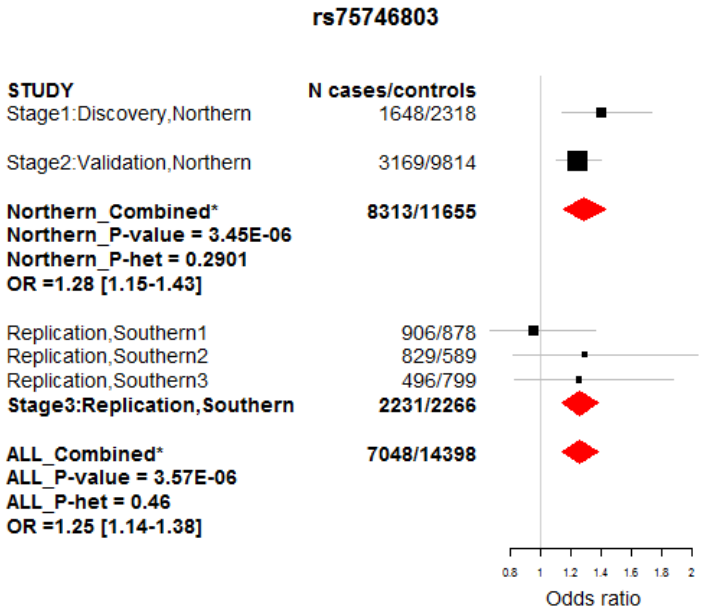
d: *SLC29A3* (common)



e: *IL27* (common)

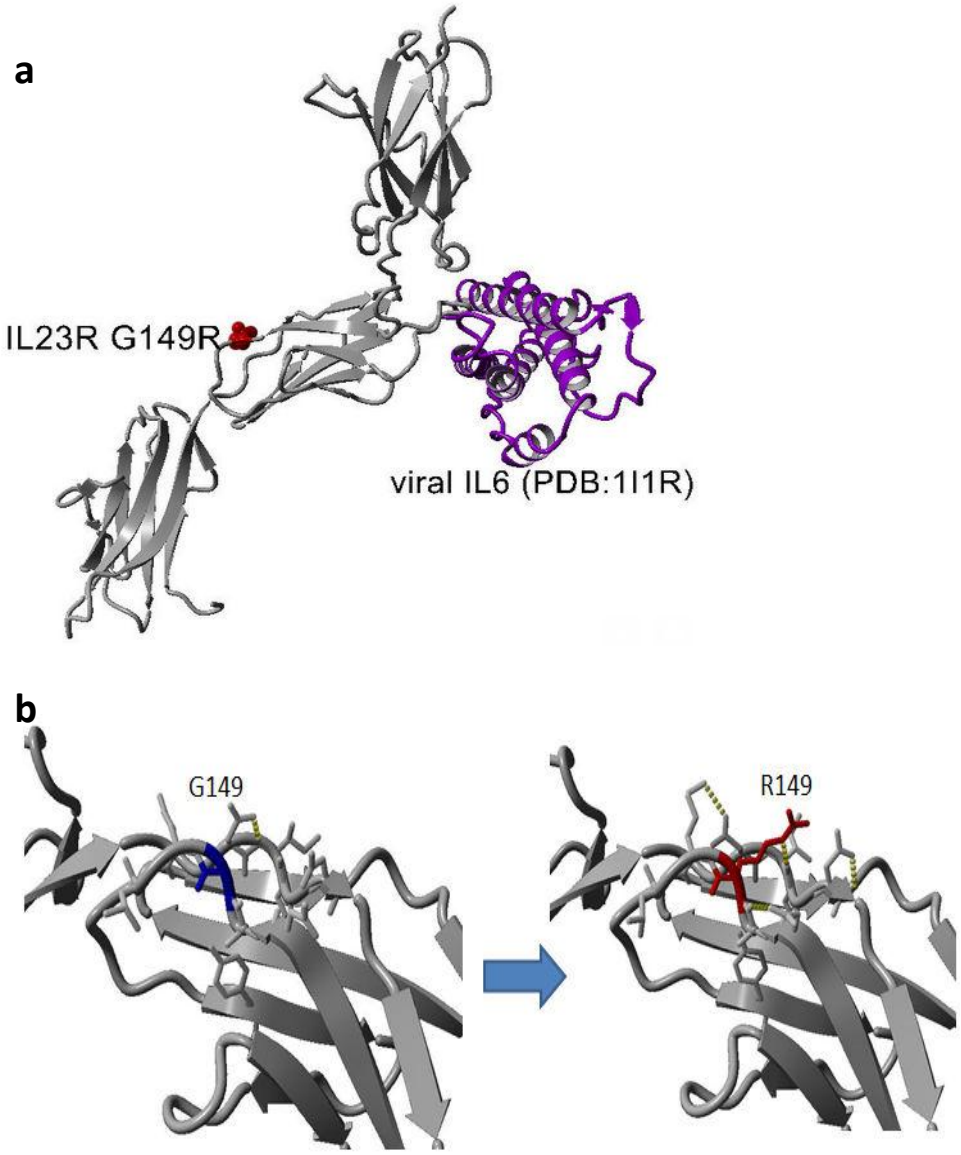


f: *USP49* (low frequency, suggestive loci)



Supplementary Figure S6. Forest plots of six loci that studied in the discovery, validation and replication samples.

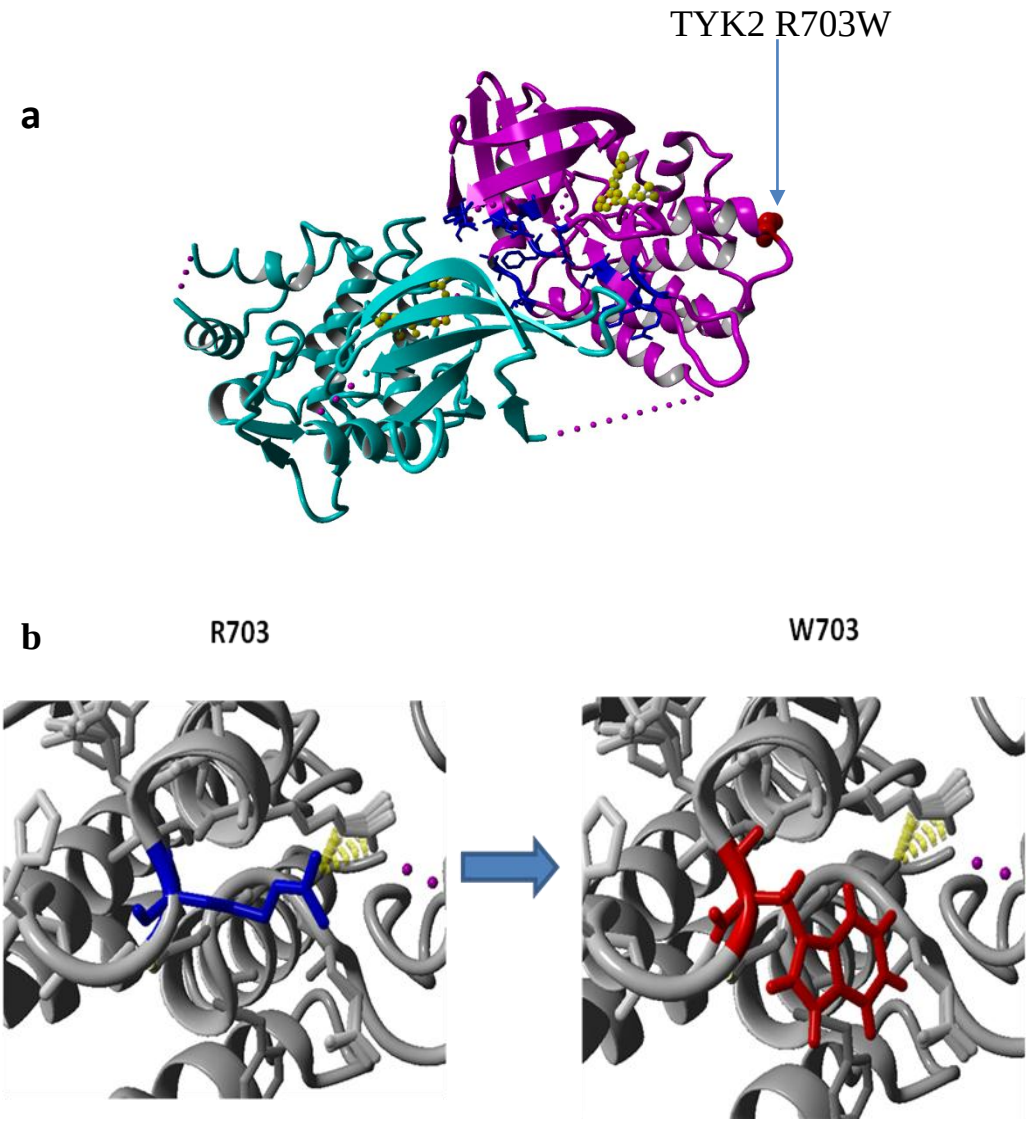
OR, odds ratio, presented with their 95% confidence intervals within the square bracket.



Supplementary Figure S7. Protein Structural impact of mutation p.G149R in IL23R.

(a) The IL23 model aligned with the template (PDB: 1I1R) with the ligand viral IL6. The IL23R_G149R is shown in red and the ligand is shown in purple. IL23R G149R is located far away from the interaction region (> 5 angstroms), hence the SNP may not directly affect the interaction with the ligand but may influence the protein stability.

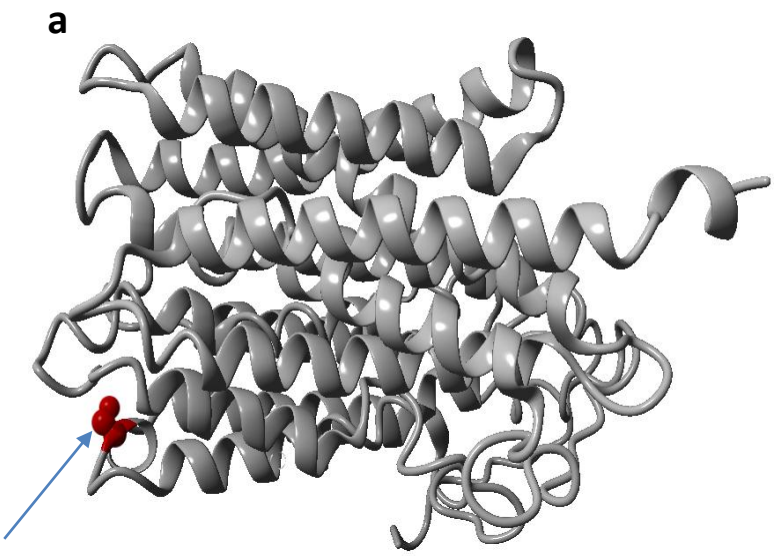
(b) The structure change induced by IL23R_G149R shown in YASARA with ddg = 5.69 ± 0.87 kcal/mol, indicating that G149R has a strong destabilizing effect on the protein structure.

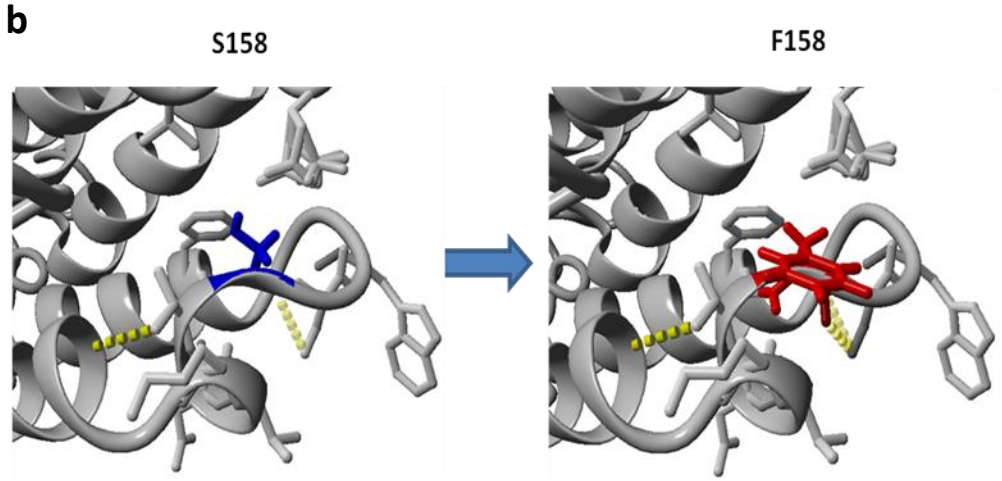


Supplementary Figure S8. Protein structural impact of mutation p. R703W in TYK2.

(a) The crystal structure of human TYK2 pseudokinase-kinase domains bound with ATP-competitive inhibitor (ligand 2TT) (PDB: 4OLI). The kinase domain is highlighted in cyan and pseudokinase domain is in magenta. The studied SNP is shown in red. 16 residues on the interaction interface between pseudokinase and kinase domains are shown in blue. Our SNP is located far from ligand 2TT as well as 16 residues of pseudokinase domain which are on the interface surface with kinase domain (i.e > 10 angstrom). This may indicate that our SNP may not directly affect the interaction with ligands and/or the kinase domain but influence the stability of TYK2 structure itself.

(b) The structure change induced by TYK2_R703W shown in YASARA with ddg = 1.66 ± 0.041 kcal/mol, indicating that R703W has a strong destabilizing effect on the protein structure.

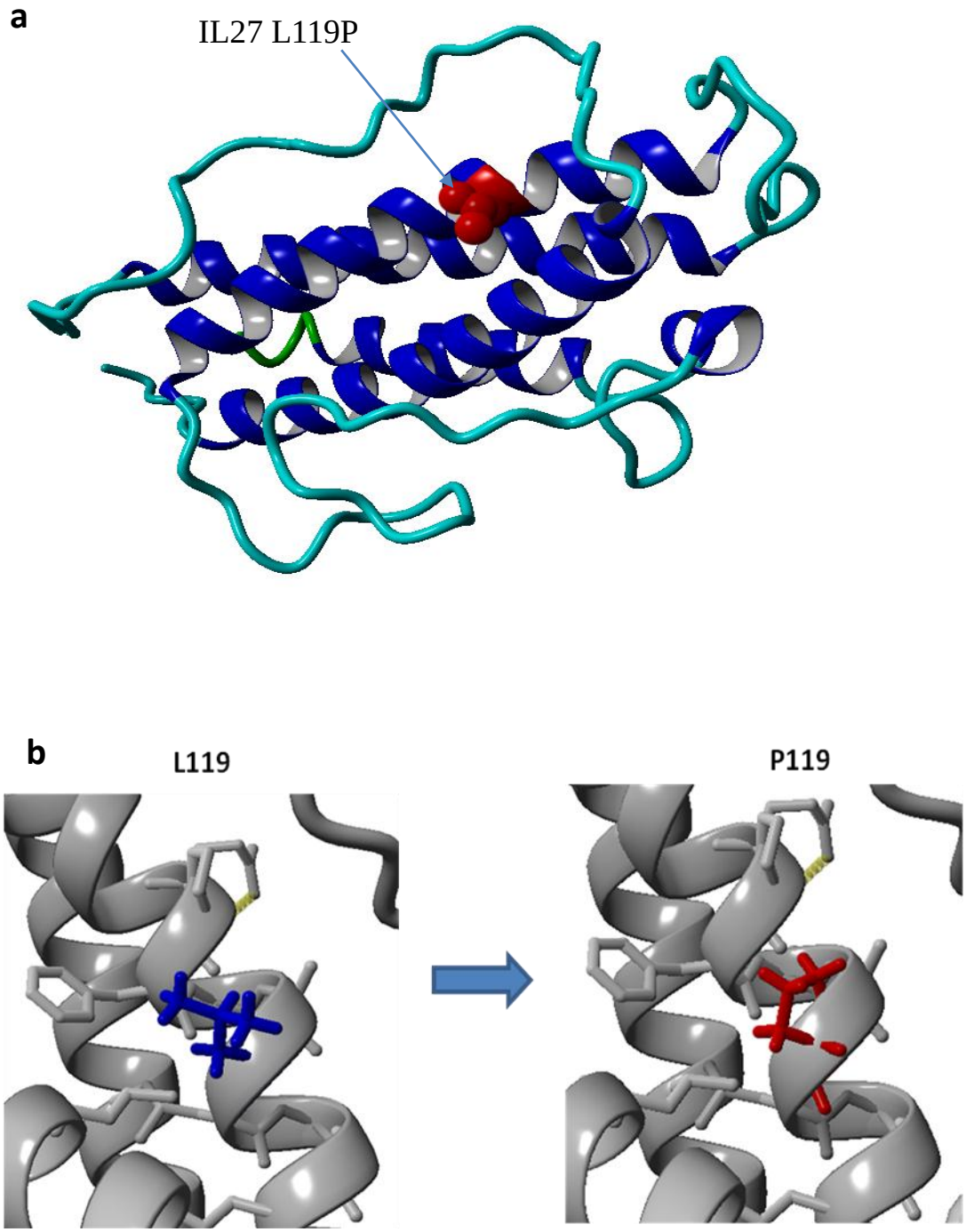




Supplementary Figure S9. Protein structural impact of mutation p. S158F in SLC29A3.

(a) The SLC29A3_S158F in crystal structure after homology modeling, shown in red.

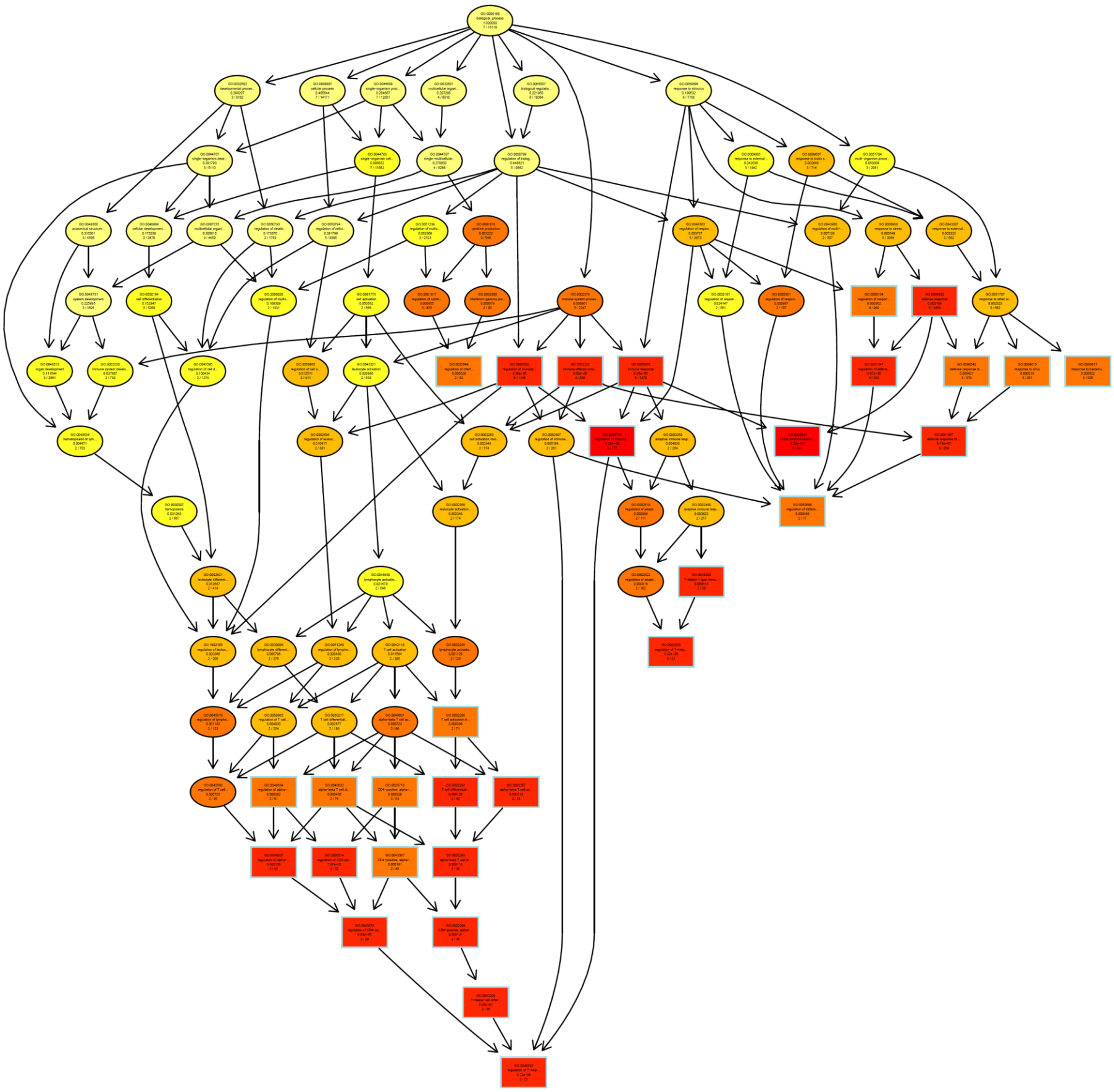
(b) The structure change induced by SLC29A3_S158F shown in YASARA, with ddg = 2.18 ± 0.21 kcal/mol, suggesting that S158F is significantly destabilizing the protein structure.



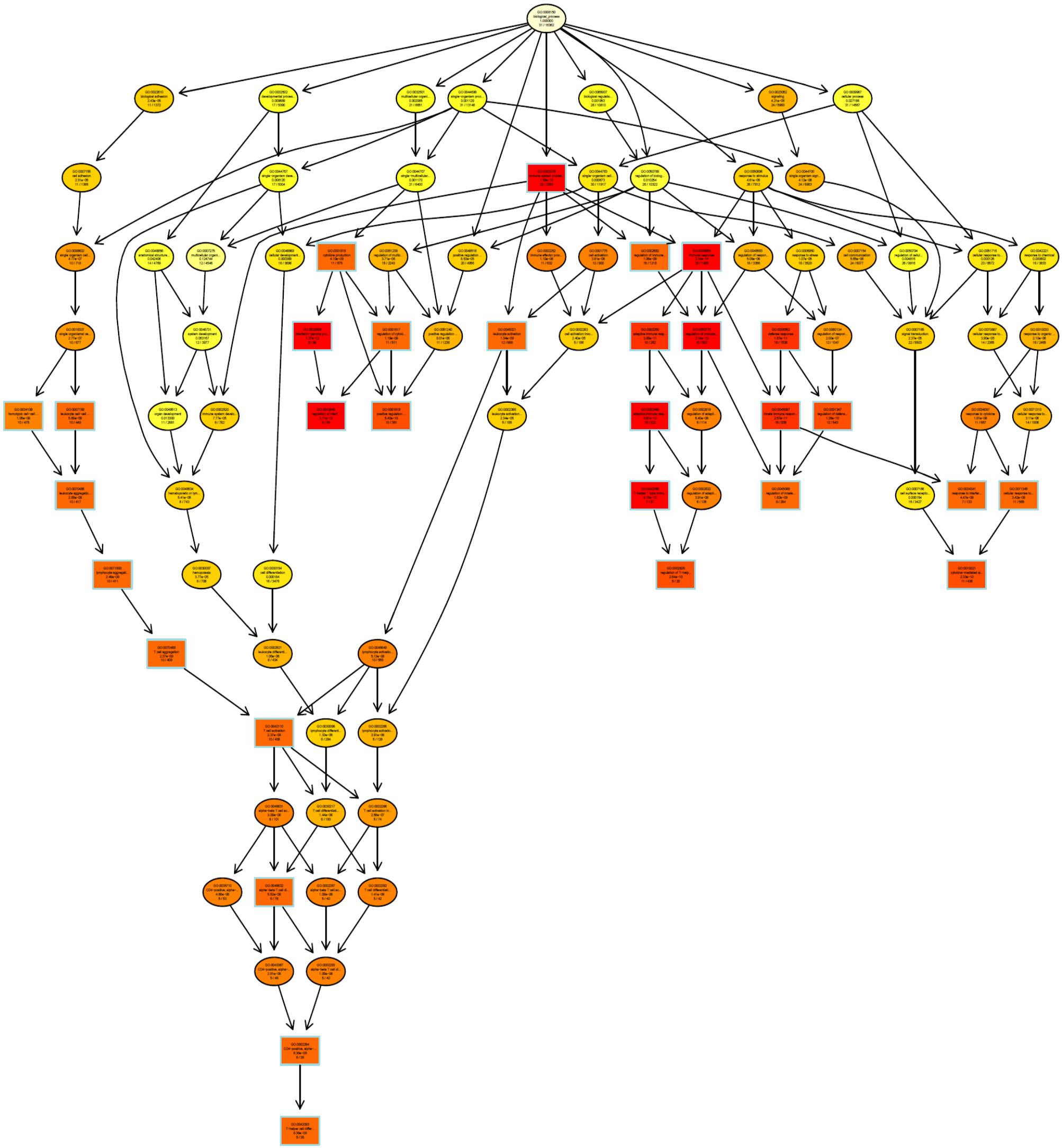
Supplementary Figure S10. Protein structural impact of mutation p. L119P in IL27.

(a) The IL27 L119P in crystal structure after homology modeling, shown in red. (b) The structure change induced by IL27_L119P shown in YASARA, with ddg = 4.37 ± 0.11 kcal/mol, suggesting that L119P has a significant destabilizing effect on the protein structure.

a) The enriched GO hierarchical subgraph induced from 7-gene set.



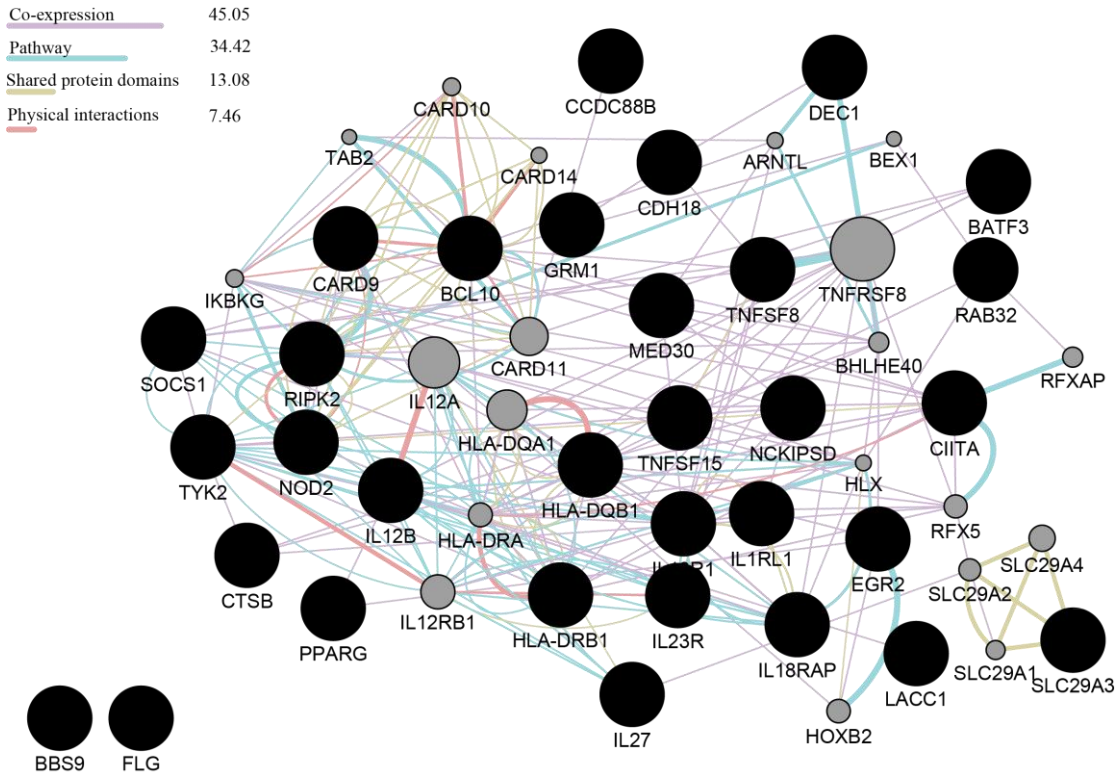
b) The enriched GO hierarchical subgraph induced from 33-gene set.



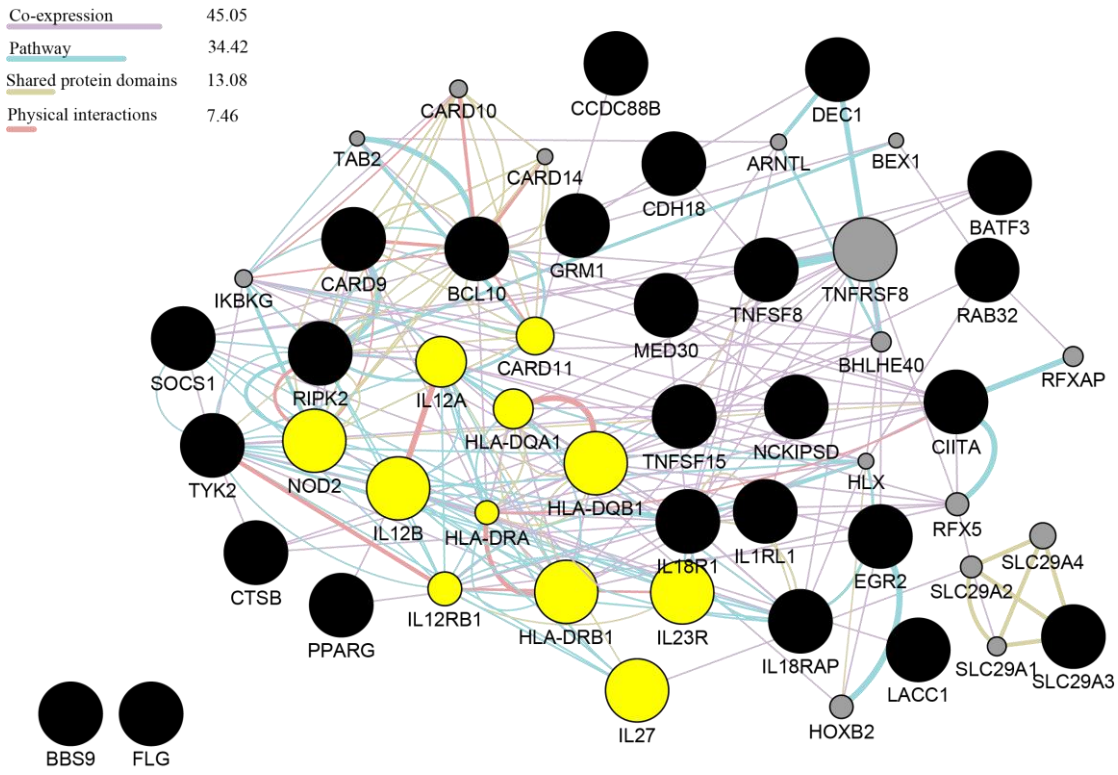
Supplementary Figure S11. The enriched GO hierarchical subgraph induced from top 30 significant GO terms.

a) The enriched GO hierarchical subgraph induced from 7-gene set; b) The enriched GO hierarchical subgraph induced from 33-gene set. In the figures, boxes indicate the 30 most significant GO terms, box colour represents the relative significance, ranging from dark-red (most significant) to light yellow (least significant). The arrows represent the hierarchical relationship, where "is-a" is in black arrow, and "part-of" is in red arrow.

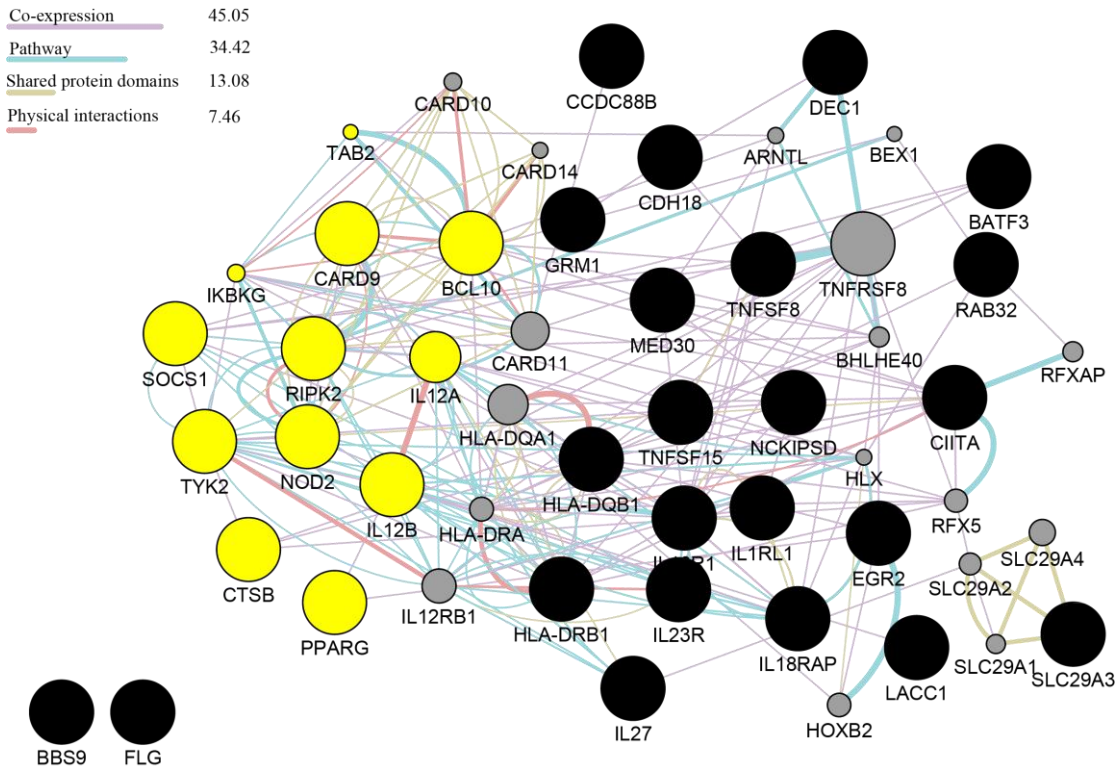
a) Composite network from GeneMANIA using 33 genes as query genes.



b) The sub-network that is related to adaptive immunity.



c) The sub-network that is related to innate immunity.



Supplementary Figure S12. The structure of enriched integrated network of 33-gene set from GeneMania.

a) Composite network from GeneMANIA using 33-gene set as query. In the figures, the black nodes represent initial query seed genes, the gray nodes were extended genes, the node size represents the strength of functional association with query genes. The edge weight represents the strength of functional association from the corresponding individual data source; b) The sub-network that is related to adaptive immune responses; c) The sub-network that is related to innate immune responses.

TABLES

Supplementary Table S1. Results of conditional analyses for variants ($P < 1 \times 10^{-3}$) within previously reported GWAS loci using overlapping samples of 802 cases and 980 controls.

Locus	Chr	Variant	Position	Alleles[1]	discovery stage[2]		Unadjusted[3]		Conditional analysis[3]			r^2	D'
					P value	OR	P value	OR	Conditioning Variant	P value	OR		
IL23R	1	**rs3762318**	67597119	G/A	NA	NA	1.02E-05	0.57	rs1884444	9.12E-04	0.64	0.20	1.00
	1	rs1884444	67633812	G/T	1.86E-04	0.83	5.00E-04	0.78	**rs3762318**	6.13E-02	0.86		
	1	**rs3762318**	67597119	G/A	NA	NA	1.02E-05	0.57	**rs76418789**	1.36E-05	0.58	0.00	0.93
	1	**rs76418789**	67648596	A/G	5.59E-04	1.45	3.07E-02	1.40	**rs3762318**	5.30E-02	1.35		
	1	rs1884444	67633812	G/T	1.86E-04	0.83	5.00E-04	0.78	**rs76418789**	1.36E-03	0.79	0.03	1.00
	1	**rs76418789**	67648596	A/G	5.59E-04	1.45	3.07E-02	1.40	rs1884444	9.51E-02	1.30		
IL18RAP/ IL18R1	2	**rs2058660**	103054449	T/C	NA	NA	2.95E-03	1.22	rs1420101	2.76E-02	1.23	0.36	0.83
	2	rs1420101	102957716	T/C	3.10E-05	1.24	5.10E-02	1.15	**rs2058660**	8.91E-01	0.99		
	2	**rs2058660**	103054449	T/C	NA	NA	2.95E-03	1.22	rs1014286	8.08E-01	1.03	0.78	0.96
	2	rs1014286	103149100	A/G	5.35E-06	1.25	6.70E-04	1.26	**rs2058660**	8.66E-02	1.23		
	2	rs1420101	102957716	T/C	3.10E-05	1.24	5.10E-02	1.15	rs1014286	9.97E-01	1.00	0.32	0.71
	2	rs1014286	103149100	A/G	5.35E-06	1.25	6.70E-04	1.26	rs1420101	5.24E-03	1.26		
HLA-DRB1	6	**rs9271100[4]**	32576478	T/C	NA	NA	3.15E-09	1.90	rs3200405	1.30E-03	1.92	0.74	0.97
	6	rs3200405[4]	32487309	T/C	3.46E-40	2.15	4.93E-07	1.78	**rs9271100**	9.45E-01	0.99		
	6	**rs9271100[4]**	32576478	T/C	NA	NA	3.15E-09	1.90	**HLA-DRB1 *15:01**	3.62E-03	1.58	0.57	0.99

	6	**HLA-DRB1 *15:01[4]**	32552064	P/A	NA	NA	5.29E-08	1.95	**rs9271100**	1.09E-01	1.33		
	6	**HLA-DRB1 *15:01[4]**	32552064	P/A	NA	NA	5.29E-08	1.95	rs3200405	2.31E-02	1.67	0.72	0.99
	6	rs3200405[4]	32487309	T/C	3.46E-40	2.15	4.93E-07	1.78	**HLA-DRB1 *15:01**	4.10E-01	1.19		
LACC1	13	**rs3764147**	44457925	G/A	NA	NA	1.69E-15	1.77	rs9567280	1.35E-13	1.83	0.24	0.95
	13	rs9567280	44411432	G/A	1.87E-07	1.47	1.61E-03	1.38	**rs3764147**	3.43E-01	0.90		

Note: Conditional analysis was based on the overlapping samples (802 cases and 980 controls) between the current study and the previous GWAS in the discovery stage;

[1]Minor allele/major allele;

[2]using all the samples in the discovery stage;

[3]using the overlapping samples (802 cases and 980 controls) in the discovery stage;

[4]These three variants using the overlapping samples (408 cases and 453 controls) in the discovery stage

OR, odds ratio is with respect to the minor allele;

NA, not applicable;

Bold indicates previously reported GWAS variants; Blue indicates independent variants which were selected for validation

Supplementary Table S2. The association results of all the 34 variants in discovery and validation samples.

Variant	Chr	Position	Function	Gene	AA	Type[1]	Alleles[2]	Discovery results				Validation				Combined		
								F_A	F_U	OR	P	F_A	F_U	OR	P	OR(f)	P(f)	Phet
rs13306061	1	7913445	missense	*UTS2*	R16Q	low freq	T/C	0.043	0.028	1.56	5.56E-04	0.037	0.035	1.06	4.62E-01	1.17	1.63E-02	0.01
rs76418789	1	67648596	missense	*IL23R*	G149R	low freq	A/G	0.061	0.043	1.45	5.59E-04	0.063	0.048	1.34	2.37E-06	1.37	6.28E-09	0.55
rs74518578	1	100155425	missense	*PALMD*	T537A	low freq	G/A	0.008	0.018	0.47	6.22E-04	0.010	0.012	0.85	2.53E-01	0.72	6.05E-03	0.02
rs3789604	1	114354942	coding-synonymous	*RSBN1*	R31R	common	G/T	0.179	0.214	0.80	1.68E-04	0.191	0.194	0.98	6.63E-01	0.92	1.68E-02	0.00
rs146466242	1	152275298	stop-gained	*FLG*	K4022*	low freq	A/T	0.056	0.035	1.59	4.17E-05	0.055	0.040	1.40	4.39E-07	1.45	1.44E-10	0.34
rs75906759	1	158724634	missense	*OR6K6*	H10R	low freq	G/A	0.030	0.018	1.69	7.31E-04	0.024	0.026	0.91	3.06E-01	1.07	3.84E-01	0.00
rs139768432	2	15644333	missense	*NBAS*	P297L	rare	A/G	0.003	0.000	NA	8.60E-04	0.000	0.000	1.56	5.66E-01	NA	7.51E-01	NA
rs3087403	2	100058870	missense	*REV1*	V138M	common	T/C	0.075	0.053	1.44	1.58E-04	0.065	0.065	1.00	9.76E-01	1.10	5.23E-02	0.00
rs11676273	2	105889349	coding-synonymous	*TGFBRAP1*	L100L	common	A/G	0.089	0.068	1.35	7.77E-04	0.082	0.073	1.13	2.15E-02	1.18	2.17E-04	0.09
rs145562243	3	48719549	missense	*NCKIPSD*	R176Q	rare	T/C	0.004	0.000	8.92	8.78E-04	0.006	0.001	3.93	1.31E-06	4.35	1.44E-08	NA
rs148995934	3	188426077	missense	*LPP*	G379E	low freq	A/G	0.016	0.028	0.56	7.00E-04	0.021	0.022	0.92	4.40E-01	0.81	1.76E-02	0.01
rs7717874	5	131007607	missense	*FNIP1*	I844V	rare	C/T	0.003	0.000	NA	4.29E-04	0.003	0.004	0.86	6.28E-01	NA	6.00E-01	NA
rs75746803	6	41773726	missense	*USP49*	W332C	low freq	G/C	0.066	0.048	1.40	8.34E-04	0.061	0.050	1.24	6.50E-04	1.28	3.45E-06	0.29
rs1887415	6	137519238	missense	*IFNGR1*	L467P	low freq	G/A	0.041	0.026	1.61	3.10E-04	0.034	0.029	1.21	1.84E-02	1.31	1.03E-04	0.07
rs2306093	9	100133973	missense	*C9orf174*	D1518N	common	A/G	0.085	0.110	0.76	7.26E-04	0.091	0.100	0.90	4.11E-02	0.86	4.43E-04	0.07
rs8176720	9	136132873	coding-synonymous	*ABO*	T98T	common	C/T	0.452	0.496	0.84	4.70E-04	0.462	0.473	0.95	9.95E-02	0.92	1.24E-03	0.03
rs149308743	9	139258965	missense	*CARD9*	R494H	rare	T/C	0.005	0.001	8.52	8.56E-05	0.005	0.001	4.25	1.77E-07	4.75	4.99E-10	NA
rs139905834	10	5248257	missense	*AKR1C4*	D156V	rare	T/A	0.000	0.004	0.00	3.23E-04	0.001	0.002	0.74	4.98E-01	NA	4.73E-01	NA
rs780668	10	73111408	missense	*SLC29A3*	S158F	common	T/C	0.465	0.421	1.19	3.42E-04	0.462	0.434	1.12	1.36E-04	1.14	2.89E-07	0.28
rs11591349	10	102744331	missense	*SEMA4G, MRPL43*	D597V	common	T/A	0.082	0.108	0.74	2.77E-04	0.088	0.096	0.91	6.58E-02	0.86	5.06E-04	0.04
rs73404785	11	1718833	missense	*KRTAP5-6*	C120G	low freq	G/T	0.008	0.019	0.44	2.60E-04	0.013	0.015	0.92	5.16E-01	0.78	2.11E-02	0.00
rs188675162	12	9875338	missense	*CLECL1*	Y130H	rare	G/A	0.005	0.001	5.00	7.55E-04	0.004	0.003	1.38	2.08E-01	NA	6.76E-01	NA
rs117605527	12	54894320	missense	*NCKAP1L*	H23Y	low freq	T/C	0.003	0.010	0.30	3.90E-04	0.006	0.007	0.91	7.14E-01	0.72	4.87E-02	0.01
rs191253798	12	96273421	splice-5	*CCDC38*	NA	rare	T/C	0.005	0.001	6.74	8.72E-04	0.003	0.001	2.26	4.59E-02	3.05	9.04E-04	NA

rs1001178	12	122691045	missense	*B3GNT4*	S58T	common	A/T	0.033	0.051	0.63	1.97E-04	0.042	0.042	1.01	9.42E-01	0.89	6.81E-02	0.00
rs41288291	13	42301395	missense	*KIAA0564*	R898K	low freq	T/C	0.016	0.008	2.16	6.75E-04	0.011	0.010	1.04	7.75E-01	1.27	4.49E-02	0.01
rs115121346	15	43548812	missense	*TGM5*	Q88R	rare	C/T	0.006	0.001	4.90	6.90E-04	0.003	0.003	1.28	3.28E-01	1.75	2.08E-02	NA
rs181206	16	28513403	missense	*IL27*	L119P	common	G/A	0.117	0.150	0.76	1.30E-04	0.118	0.135	0.85	4.68E-04	0.83	5.92E-07	0.16
rs2240154	19	1003172	missense	*GRIN3B*	T157M	common	T/C	0.478	0.439	1.17	9.56E-04	0.464	0.443	1.09	8.47E-03	1.12	5.47E-05	0.20
rs181899564	19	9091783	missense	*MUC16*	S11C	rare	C/G	0.009	0.003	3.08	5.28E-04	0.006	0.006	0.98	1.00E+00	1.32	1.25E-01	NA
rs55882956	19	10469919	missense	*TYK2*	R703W	low freq	A/G	0.053	0.036	1.51	4.96E-04	0.046	0.038	1.22	4.89E-03	1.29	2.75E-05	0.12
rs187900950	19	40097909	missense	*LGALS13*	P117L	rare	T/C	0.005	0.000	10.82	1.66E-04	0.003	0.002	1.97	4.38E-02	2.54	7.54E-04	NA
rs1111032	20	43378770	missense-near-splice	*KCNK15*	E95G	common	A/G	0.330	0.367	0.84	9.01E-04	0.334	0.348	0.94	4.63E-02	0.91	5.74E-04	0.08
rs437470	21	18937758	missense	*CXADR*	H195R	common	G/A	0.141	0.172	0.79	4.27E-04	0.159	0.156	1.02	5.49E-01	0.96	1.88E-01	0.00

[1]common: MAF ≥ 5% in controls, low freq: 1% ≤ MAF < 5% in controls, rare: MAF < 1% in controls;

[2]Minor allele/major allele;

AA, Amino acid change;

OR, odds ratio is based on the minor allele;

F_A, minor allele frequency in cases, F_U, minor allele frequency in controls;

(f) indicates results from fixed-effects meta analysis;

NA, not applicable

Supplementary Table S3. Baseline characteristics of cases and controls.

	CASES		CONTROLS	
	N	Male/ Female	N	Male/ Female
Discovery Study (North)	**1,648**	**1,330/318**	**2,318**	**1,722/596**
Validation (North)	**3,169**	**2,491/678**	**9,814**	**6,343/3,471**
Replication	**2,231**	**1,637/594**	**2,266**	**1,135/1,131**
Replication1 (South:Sichuan Province)	906	672/234	878	368/510
Replication2 (South:Yunnan Province)	829	605/224	589	317/272
Replication3 (South:Guizhou Province)	496	360/136	799	450/349
Total/Mean	**7,048**	**5,458/1,590**	**14,398**	**9,200/5,198**

Supplementary Table S4. The meta-analysis of to our knowledge previously unreported variants across three stages with and without gender adjustment.

Chr	Variant	Position	Alleles[1]	without gender adjustment			with gender adjustment		
				P(f)	OR(f)	Phet	P(f)	OR(f)	Phet
1	rs76418789	67648596	A/G	1.03E-10	1.36	0.84	1.00E-10	1.37	0.78
1	rs146466242	152275298	A/T	3.39E-12	1.45	0.74	2.18E-10	1.42	0.94
10	rs780668	73111408	T/C	2.17E-09	1.14	0.74	6.44E-09	1.14	0.81
16	rs181206	28513403	G/A	1.08E-07	0.83	0.61	1.52E-07	0.83	0.88
19	rs55882956	10469919	A/G	1.04E-06	1.30	0.30	3.47E-07	1.33	0.12

[1]Minor allele/major allele;

OR, odds ratio is with respect to the minor allele;

(f) indicates results from fixed-effect meta analysis.

Note: The association analysis of the two rare variants were done using Fisher's exact test, so the results with gender adjustment were not available.

Supplementary Table S5. Condition analysis between the leading GWAS variant rs3762318 and the coding variant rs76418789 in *IL23R* gene using 3,019 cases and 5,767 controls of northern Chinese.

Variant	Function	Alleles[1]	Association in overlapping samples[2]				Logistic conditional analysis			LD	
			F_A	F_U	P	OR	Conditioned on	P value	OR	D'	r^2
rs3762318[3]	Intergenic	G/A	0.0740	0.1089	1.05E-13	0.65	rs76418789	6.08E-13	0.66	1.000	0.006
rs76418789	missense	A/G	0.0633	0.0494	1.04E-04	1.30	rs3762318	9.06E-04	1.26		

[1]Minor allele/major allele

[2]Overlapping samples (3,019 cases and 5,767 controls) in the current study and our previous GWAS of leprosy, all samples are northern samples;

[3]Variant from previous GWAS.

F_A, minor allele frequency in cases; F_U, minor allele frequency in controls;

OR, odds ratio is based on the minor allele;

Supplementary Table S6. Haplotype analysis of the leading GWAS variant rs3762318 and the coding variant rs76418789 in *IL23R* gene.

Haplotype	rs3762318(GWAS)	rs76418789(coding)	Freq	F_A	F_U	P	OR
AG	A	G	0.8488	0.8626	0.8416	NA	NA
GG	G	G	0.0970	0.0741	0.1091	1.03E-12	0.66
AA	A	A	0.0542	0.0633	0.0494	1.03E-03	1.25

Note: yellow cells are minor allele, red characters are risk allele;

Freq, haplotype frequency in all samples;

F_A, haplotype frequency in cases; F_U, haplotype frequency in controls.

Supplementary Table S7. Variance explained by each variant.

a)To our knowledge previously unreported locus

No.	Variant	Loci	Variance explained (%)
1	rs76418789	*IL23R*	0.10
2	rs146466242	*FLG*	< 0.01
3	rs149308743	*CARD9*	< 0.01
4	rs780668	*SLC29A3*	0.30
5	rs181206	*IL27*	0.40
6	rs55882956	*TYK2*	0.10
7	rs145562243	*NCKIPSD*	0.30
	TOTAL		**1.20**

b)Known locus

No.	Variant	Loci	Variance explained (%)
1	rs7995004	*LACC1*	3.60
2	rs9302752	*NOD2*	2.10
3	rs6478109	*TNFSF15*	1.20
4	rs42490	*RIPK2*	0.80
5	rs3762318	*IL23R*	0.80
6	rs2275606	*RAB32*	0.60
7	rs2058660	*IL18RAP/IL1RL1*	0.70
8	rs6871626	*IL12B*	0.80
9	rs160451	*RIPK2*	0.50
10	rs8002861	*LACC1*	0.20
11	rs2221593	*BATF3*	0.40
12	rs58600253	*EGR2*	0.50
13	rs663743	*CCDC88B*	0.30
14	rs77061563	*LOC388210*	0.60
15	rs2735591	*BCL10*	0.10
16	rs9271100	*HLA-DR-DQ*	3.50
17	rs73058713	*CDH18*	0.10
18	rs10817758	*DEC1*	0.20
19	rs6807915	PPARG	<0.01
20	rs4720118	BBS9	<0.01
21	rs55894533	CTSB	<0.01
22	rs10100465	MED30	<0.01
	TOTAL		**17.01**

Supplementary Table S8. Protein function annotations of to our knowledge previously unreported locus.

Variant	Gene	Transcript	cDNA position	Protein	Amino acid change	Location	SIFT prediction (SIFT score)[1]	PolyPhen-2 (PolyPhen-2 Score)[2]
rs76418789	*IL23R*	ENST00000347310.5	c.445G>A	ENSP00000321345.5	p.Gly149Arg	Exon 4 of 11	deleterious(0)	probably damaging(0.991)
(chr1:67648596)		ENST00000371002.1	c.445G>A	ENSP00000360041.1	p.Gly149Arg	Exon 4 of 10	deleterious(0)	probably damaging(0.999)
		ENST00000371007.2	-	-	-	Intronic	-	-
		ENST00000448166.2	-	-	-	Intronic	-	-
rs146466242	*FLG*	ENST00000368799.1	c.12064A>T	ENSP00000357789.1	p.Lys4022Ter	Exon 3 of 3	-	-
(chr1:152275298)		ENST00000420707.1	-	-	-	Intronic	-	-
		ENST00000593011.1	-	-	-	Intronic	-	-
rs55882956	*TYK2*	ENST00000264818.6	c.2107C>T	ENSP00000264818.6	p.Arg703Trp	Exon 13 of 23	deleterious(0)	probably damaging(0.932)
(chr19:10469919)		ENST00000525621.1	c.2107C>T	ENSP00000431885.1	p.Arg703Trp	Exon 15 of 25	deleterious(0)	probably damaging(0.932)
		ENST00000529370.1	c.2107C>T	ENSP00000432728.1	p.Arg703Trp	Exon 15 of 17	deleterious(0)	probably damaging(0.979)
		ENST00000524462.1	c.1552C>T	ENSP00000433203.1	p.Arg518Trp	Exon 11 of 21	deleterious(0)	probably damaging(0.932)
		ENST00000533334.1	c.*149C>T	ENSP00000432320	-	3'UTR	-	-
		ENST00000525220	-	ENSP00000434931	-	Downstream	-	-
		ENST00000531620	-	-	-	Downstream	-	-
		ENST00000527481	-	ENSP00000466340	-	Upstream	-	-
		ENST00000530560	-	ENSP00000465291	-	Upstream	-	-
		ENST00000529412	-	-	-	Upstream	-	-
		ENST00000534228	-	-	-	Upstream	-	-
rs145562243	*NCKIPSD*	ENST00000294129.2	c.527G>A	ENSP00000294129.2	p.Arg176Gln	Exon 4 of 13	deleterious(0.03)	benign(0.006)
(chr3:48719549)		ENST00000341520.4	c.527G>A	ENSP00000342621.4	p.Arg176Gln	Exon 4 of 13	deleterious(0.02)	benign(0.016)
		ENST00000439518.1	c.527G>A	ENSP00000409675.1	p.Arg176Gln	Exon 4 of 5	deleterious(0)	benign(0.022)

		ENST00000416649.2	c.506G>A	ENSP00000389059.2	p.Arg169Gln	Exon 4 of 13	deleterious(0.02)	benign(0.014)
		ENST00000426678.1	c.179G>A	ENSP00000416904.1	p.Arg60Gln	Exon 4 of 5	deleterious(0)	benign(0.006)
		ENST00000453349.1	c.293G>A	ENSP00000408588.1	p.Arg98Gln	Exon 4 of 5	deleterious(0)	benign(0.006)
		ENST00000454134.1	c.*379G>A	ENSP00000416144	-	3'UTR	-	-
		ENST00000413374	-	ENSP00000396683	-	Upstream	-	-
		ENST00000415281	-	ENSP00000406442	-	Upstream	-	-
		ENST00000470006	-	-	-	Upstream	-	-
rs149308743	*CARD9*	ENST00000371732.5	c.1481G>A	ENSP00000360797.5	p.Arg494His	Exon 12 of 13	deleterious(0.01)	possibly damaging(0.855)
(chr9:139258965)		ENST00000481053.1	n.2339G>A	-	-	Non coding exon variant	-	-
		ENST00000485975.1	n.4099G>A	-	-	Non coding exon variant	-	-
		ENST00000460290.1	n.585G>A	-	-	Non coding exon variant	-	-
		ENST00000489932.2	c.*528G>A	ENSP00000451368	-	3'UTR	-	-
		ENST00000440944	-	ENSP00000392828	-	Downstream	-	-
		ENST00000563222	-	ENSP00000456621	-	Downstream	-	-
		ENST00000291775	-	ENSP00000291775	-	Downstream	-	-
		ENST00000429455	-	ENSP00000390705	-	Downstream	-	-
		ENST00000315908	-	ENSP00000323719	-	Downstream	-	-
		ENST00000563430	-	ENSP00000454556	-	Downstream	-	-
		ENST00000354753	-	ENSP00000346797	-	Downstream	-	-
		ENST00000371739	-	ENSP00000360804	-	Upstream	-	-
		ENST00000371738	-	ENSP00000360803	-	Upstream	-	-
		ENST00000371734.3	c.1441+40G>A	ENSP00000360799	-	Intronic	-	-
rs780668	*SLC29A3*	ENST00000373189.5	c.473C>T	ENSP00000362285.5	p.Ser158Phe	Exon 4 of 6	deleterious(0.01)	probably damaging(0.999)
(chr10:73111408)		ENST00000469204	-	-	-	Upstream	-	-
rs181206	*IL27*	ENST00000356897.1	c.356T>C	ENSP00000349365.1	p.Leu119Pro	Exon 4 of 5	tolerated(0.13)	possibly damaging(0.875)

(chr16:28513403)		ENST00000568075.1	c.-38T>C	ENSP00000455990	-	5'UTR	-	-
		ENST00000328423	-	ENSP00000327669	-	Downstream	-	-
		ENST00000431282	-	ENSP00000416094	-	Downstream	-	-
		ENST00000564831	-	ENSP00000457539	-	Downstream	-	-
rs75746803	*USP49*	ENST00000373009.3	c.996G>C	ENSP00000362100.3	p.Trp332Cys	Exon 1 of 5	tolerated(0.18)	possibly damaging(0.518)
(chr6:41773726)		ENST00000373010.1	c.996G>C	ENSP00000362101.1	p.Trp332Cys	Exon 6 of 10	tolerated(0.18)	possibly damaging(0.704)
		ENST00000394253.3	c.996G>C	ENSP00000377797.2	p.Trp332Cys	Exon 3 of 7	tolerated(0.18)	possibly damaging(0.518)
		ENST00000373006.1	c.996G>C	ENSP00000362097.1	p.Trp332Cys	Exon 4 of 7	tolerated(0.18)	benign(0.383)
		ENST00000297229.2	c.996G>C	ENSP00000297229.2	p.Trp332Cys	Exon 2 of 5	tolerated(0.18)	benign(0.383)
		ENST00000437061	-	ENSP00000410003	-	Downstream	-	-
		ENST00000423567	-	ENSP00000411603	-	Downstream	-	-
		ENST00000448078	-	ENSP00000389842	-	Upstream	-	-

Annotation. We annotated variants relative to Ensembl GRCh37 release 81 (http://grch37.ensembl.org/Homo_sapiens/Info/Index)

[1]SIFT scores range from 0 to 1. All scores ≤ 0.05 are predicted to be deleterious

[2]Polyphen-2 scores range from 0 (benign) to 1 (probably damaging).

Supplementary Table S9. The 30 most significant GO terms.

a) The 30 most significant GO terms by seven genes set.

GO.ID	Term	Annotated[1]	Significant[2]	Expected[3]	Classic[4]
GO:0050776	regulation of immune response	777	*CARD9, IL23R, IL27, NCKIPSD, TYK2*	0.34	5.00E-06
GO:0045087	innate immune response	861	*CARD9, IL23R, IL27, NCKIPSD, TYK2*	0.37	8.30E-06
GO:0002682	regulation of immune system process	1146	*CARD9, IL23R, IL27, NCKIPSD, TYK2*	0.5	3.40E-05
GO:0002825	regulation of T-helper 1 type immune response	21	*IL23R, IL27*	0.01	3.40E-05
GO:0031347	regulation of defense response	524	*CARD9, IL23R, IL27, TYK2*	0.23	3.60E-05
GO:0045622	regulation of T-helper cell differentiation	22	*IL23R, IL27*	0.01	3.70E-05
GO:0002252	immune effector process	598	*CARD9, IL23R, IL27, NCKIPSD*	0.26	6.00E-05
GO:0043370	regulation of CD4-positive, alpha-beta T cell differentiation	28	*IL23R, IL27*	0.01	6.10E-05
GO:0051607	defense response to virus	204	*CARD9, IL23R, IL27*	0.09	6.70E-05
GO:2000514	regulation of CD4-positive, alpha-beta T cell activation	32	*IL23R, IL27*	0.01	8.00E-05
GO:0006955	immune response	1379	*CARD9, IL23R, IL27, NCKIPSD, TYK2*	0.6	8.30E-05
GO:0002294	CD4-positive, alpha-beta T cell differentiation involved in immune response	36	*IL23R, IL27*	0.02	1.00E-04
GO:0042093	T-helper cell differentiation	36	*IL23R, IL27*	0.02	1.00E-04
GO:0006952	defense response	1459	*CARD9, IL23R, IL27, NCKIPSD, TYK2*	0.63	1.10E-04
GO:0002287	alpha-beta T cell activation involved in immune response	38	*IL23R, IL27*	0.02	1.10E-04
GO:0002293	alpha-beta T cell differentiation involved in immune response	38	*IL23R, IL27*	0.02	1.10E-04

GO:0042088	T-helper 1 type immune response	38	*IL23R, IL27*	0.02	1.10E-04
GO:0002292	T cell differentiation involved in immune response	40	*IL23R, IL27*	0.02	1.30E-04
GO:0046637	regulation of alpha-beta T cell differentiation	42	*IL23R, IL27*	0.02	1.40E-04
GO:0043367	CD4-positive, alpha-beta T cell differentiation	48	*IL23R, IL27*	0.02	1.80E-04
GO:0009615	response to virus	301	*CARD9, IL23R, IL27*	0.13	2.10E-04
GO:0035710	CD4-positive, alpha-beta T cell activation	53	*IL23R, IL27*	0.02	2.20E-04
GO:0046634	regulation of alpha-beta T cell activation	61	*IL23R, IL27*	0.03	2.90E-04
GO:0080134	regulation of response to stress	949	*CARD9, IL23R, IL27, TYK2*	0.41	3.60E-04
GO:0002286	T cell activation involved in immune response	71	*IL23R, IL27*	0.03	4.00E-04
GO:0098542	defense response to other organism	379	*CARD9, IL23R, IL27*	0.16	4.20E-04
GO:0046632	alpha-beta T cell differentiation	74	*IL23R, IL27*	0.03	4.30E-04
GO:0050688	regulation of defense response to virus	77	*IL23R, IL27*	0.03	4.70E-04
GO:0009617	response to bacterium	408	*CARD9, IL23R, IL27*	0.18	5.20E-04
GO:0032649	regulation of interferon-gamma production	82	*IL23R, IL27*	0.04	5.30E-04

b) The 30 most significant GO terms by 33 genes set.

GO.ID	Term	Annotated[1]	Significant[2]	Expected[3]	Classic[4]
GO:0006955	immune response	1465	20	2.78	3.10E-14
GO:0042088	T-helper 1 type immune response	37	7	0.07	4.20E-13
GO:0002376	immune system process	2362	22	4.48	1.60E-12
GO:0032649	regulation of interferon-gamma production	85	8	0.16	2.70E-12
GO:0050776	regulation of immune response	807	15	1.53	3.10E-12
GO:0002460	adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains	222	10	0.42	6.00E-12
GO:0032609	interferon-gamma production	96	8	0.18	7.40E-12

GO:0006952	defense response	1538	18	2.91	1.90E-11
GO:0045087	innate immune response	935	15	1.77	2.60E-11
GO:0002250	adaptive immune response	262	10	0.5	3.10E-11
GO:0031347	regulation of defense response	543	12	1.03	1.30E-10
GO:0019221	cytokine-mediated signaling pathway	438	11	0.83	2.30E-10
GO:0002825	regulation of T-helper 1 type immune response	20	5	0.04	2.60E-10
GO:0001819	positive regulation of cytokine production	351	10	0.67	5.40E-10
GO:0002682	regulation of immune system process	1218	15	2.31	1.10E-09
GO:0001817	regulation of cytokine production	511	11	0.97	1.20E-09
GO:0045321	leukocyte activation	668	12	1.27	1.30E-09
GO:0045088	regulation of innate immune response	284	9	0.54	1.80E-09
GO:0042110	T cell activation	409	10	0.77	2.40E-09
GO:0070489	T cell aggregation	409	10	0.77	2.40E-09
GO:0071593	lymphocyte aggregation	411	10	0.78	2.50E-09
GO:0070486	leukocyte aggregation	417	10	0.79	2.90E-09
GO:0071345	cellular response to cytokine stimulus	565	11	1.07	3.40E-09
GO:0001816	cytokine production	575	11	1.09	4.10E-09
GO:0034341	response to interferon-gamma	133	7	0.25	4.50E-09
GO:0046632	alpha-beta T cell differentiation	76	6	0.14	5.50E-09
GO:0007159	leukocyte cell-cell adhesion	449	10	0.85	5.80E-09
GO:0002294	CD4-positive, alpha-beta T cell differentiation involved in immune response	38	5	0.07	8.40E-09
GO:0042093	T-helper cell differentiation	38	5	0.07	8.40E-09
GO:0034109	homotypic cell-cell adhesion	478	10	0.91	1.10E-08

[1]The number of human genes annotated by the GO term

[2]The number of given gene set genes annotated by the GO term

[3]The expected probability of the gene set annotated by the GO term in random associations

[4]The P-value of the given gene set annotated by the GO term

Supplementary Table S10. Top 15 significant biological processes/pathways from GeneMANIA.

a) Top 15 significant biological processes/pathways that are related to adaptive immune responses

Feature	FDR	Genes in network	Genes in genome
regulation of T cell activation	7.09E-09	11	169
positive regulation of T cell activation	7.09E-09	10	131
regulation of lymphocyte activation	1.73E-08	11	208
positive regulation of lymphocyte activation	1.77E-08	10	155
positive regulation of leukocyte activation	2.76E-08	10	166
regulation of leukocyte activation	2.82E-08	11	232
positive regulation of cell activation	3.19E-08	10	172
T cell activation	3.39E-08	11	241
regulation of cell activation	5.35E-08	11	254
adaptive immune response	7.68E-08	9	135
T-helper 1 type immune response	1.80E-06	5	21
regulation of interferon-gamma production	2.01E-06	6	47
interferon-gamma production	2.92E-06	6	51
adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains	4.32E-06	7	98
positive regulation of T cell mediated immunity	1.64E-04	4	26

b) Top 15 significant biological processes/pathways that are related to innate immune responses

Feature	FDR	Genes in network	Genes in genome
regulation of innate immune response	7.09E-09	12	243
positive regulation of innate immune response	1.17E-06	9	188
positive regulation of defense response	4.46E-06	9	228
nucleotide-binding oligomerization domain containing signaling pathway	8.61E-06	5	30
positive regulation of NF-kappaB transcription factor activity	8.95E-06	7	111
cytoplasmic pattern recognition receptor signaling pathway	3.30E-05	5	40
nucleotide-binding domain, leucine rich repeat containing receptor signaling pathway	6.29E-05	5	47
pattern recognition receptor signaling pathway	6.36E-05	7	155
innate immune response-activating signal transduction	6.40E-05	7	157
activation of innate immune response	8.43E-05	7	165
regulation of I-kappaB kinase/NF-kappaB signaling	1.35E-04	7	183

I-kappaB kinase/NF-kappaB signaling	1.79E-04	7	196
positive regulation of I-kappaB kinase/NF-kappaB signaling	3.01E-04	6	137
toll-like receptor 5 signaling pathway	3.37E-03	4	65
toll-like receptor 10 signaling pathway	3.37E-03	4	65

Supplementary Table S11. Biological function annotations of to our knowledge previously unreported loci and associated variants.

Variants/locus	Other diseases	Known functions of genes
rs76418789 1p31.3 Variant location: Exon 4 of gene *IL23R*	Crohn's disease (Barrett et al., 2008, Franke et al., 2010, Huang et al., 2012, Kenny et al., 2012, Libioulle et al., 2007, McGovern et al., 2010a, McGovern et al., 2010b, Rioux et al., 2007, Wellcome Trust Case Control, 2007, Yang et al., 2014), psoriasis (Genetic Analysis of Psoriasis et al., 2010, Nair et al., 2009), ulcerative colitis (Anderson et al., 2011, Consortium et al., 2009, McGovern et al., 2010a, Silverberg et al., 2009), Vogt-Koyanagi-Harada syndrome (Hou et al., 2014), Behçet's disease (Mizuki et al., 2010, Remmers et al., 2010), ankylosing spondylitis (Australo-Anglo-American Spondyloarthritis et al., 2010, Evans et al., 2011), inflammatory bowel disease (Jostins et al., 2012, Kugathasan et al., 2008) The same Variant with Crohn's disease in Koreans (Yang et al., 2014). No LD with other disease associated variants.	The protein encoded by *IL23R* is a subunit of the receptor for IL23A/IL23, which pairs with the receptor molecule IL12RB1/IL12R beta1.The identification of the to our knowledge previously unreported independent association in the extracellular region (ECD) of *IL23R* gene, emphasized once again the essential role of IL23R as part of the IL-12–IL-23 and IFN-γ cascades in host defense against mycobacteria infections through IL23/Th17 pathway (Ottenhoff et al., 2005, Stockinger and Veldhoen, 2007)

rs146466242 1q21.3 Variant location: Exon 3 of gene *FLG*	Atopic dermatitis (Sun et al., 2011), psoriasis (Hu et al., 2012), ichthyosis vulgaris (Sandilands et al., 2006) The same Variant with psoriasis in Chinese (Hu et al., 2012). No LD with atopic dermatitis associated variant rs3126085 ($R^2 = 0.054$)[1]. No LD with ichthyosis vulgaris associated variant rs61816761 (R501X).	Profilaggrin and Filaggrin, encoded by *FLG* gene, play a pivotal role in skin barrier function by affecting formed stratum corneum and water binding (Sandilands et al., 2009, Scott and Harding, 1986). The stop-gain variant rs146466242 (p.Lys4022X), previously found in atopic dermatitis and psoriasis patients (Hu et al., 2012, Nemoto-Hasebe et al., 2009), generated a truncated protein with loss of C-terminal domain, leading to inhibited processing of profilaggrin to filaggrin peptides and impaired skin barrier function (Sandilands et al., 2007). We speculate the deficit of filaggrin caused by *FLG* mutation could enhance the entry of M. lepra that can otherwise trigger immune responses.
rs145562243 3p21.31 Variant location: Exon 4 of gene *NCKIPSD*	Not previously associated with any other trait.	The protein of *NCKIPSD* gene is localized exclusively in the cell nucleus and implicated in many functional processes, including signal transduction, maintenance of sarcomeres, assembly of myofibrils into sarcomeres, formation of stress fiber and so on (Lim et al., 2001, Ronty et al., 2007, Satoh and Tominaga, 2001, Teodorof et al., 2009). The rare variant rs145562243 (p. R176Q) was located in proline-rich region (PRD) of NCKIPSD N-terminus, whose over-expression was found to be related with abnormalities in vesicle formation and trafficking, leading to the defective endocytosis of Fcgamma receptor (FCGR)(Oh et al., 2013). The association of this variant with leprosy may give us a hint that this domain also plays a role in infectious diseases through influencing the endocytosis process and further studies are warranted to confirm it.

rs149308743 9q34.3 Variant location: Exon 12 of gene *CARD9*	Obesity (Voruganti et al., 2012), ulcerative colitis (Anderson et al., 2011, Barrett et al., 2009, Jostins et al., 2012, McGovern et al., 2010a), Crohn's disease (Anderson et al., 2011, Fearnhead et al., 2005, Jostins et al., 2012), Ankylosing spondylitis (Evans et al., 2011) No LD with Obesity, ulcerative colitis, Crohn's disease, Ankylosing spondylitis associated variants.	The identification of *NOD2*, *RIPK2* and *BCL10* as susceptibility genes of leprosy has highlighted the important role of CARD family in leprosy (Liu et al., 2013, Zhang et al., 2009). The caspase-recruitment domain (CARD)–containing adaptor protein CARD9, one member of CARD family, has been reported to be crucial for immune responses to various intracellular pathogens by integrating signals downstream of pattern recognition receptors (Bertin et al., 2000, Colonna, 2007, Glocker et al., 2009, Gross et al., 2006, Hruz and Eckmann, 2008, Hsu et al., 2007). CARD9 contains an N-terminal CARD domain (residues 7-98) and a C-terminal coiled-coil domain (residues 140-420) (Bertin et al., 2000). However, the functional rare variant rs149308743 (p.R494H) identified in our study, doesn't belong to any known domains and its function remains unclear. Our findings indicate that the unknown domain may also play a role in the pathogenesis of microbe infections.
rs780668 10q22.1 Variant location: Exon 4 of gene *SLC29A3*	Vitiligo (Tang et al., 2013), Systemic lupus erythematosus (Yang et al., 2013), dysosteosclerosis (Campeau et al., 2012) In LD with rs1417210 in Vitiligo in Asians ($R^2 = 0.207$)[2]; In LD with rs2252996 in Systemic lupus erythematosus in Asians($R^2 = 0.372$)[2]; No LD with dysosteosclerosis associated variants.	*SLC29A3* encodes the equilibrative nucleoside transporter 3 (ENT3), which mediates both influx and efflux of nucleosides across the membrane. Function loss of ENT3 has been shown to perturb lysosome function and macrophage homeostasis (Hsu et al., 2012). ENT3 deficiency mice appeared to have dysfunctional lysosomes incapable of normal cellular functions and overall defense, which showed susceptible to bacterial infection (Hsu et al., 2012), indicating that individuals carrying the variation in *SLC29A3* are more likely to develop leprosy after infection with *M. leprae*.

rs181206 16p11.2 Variant location: Exon 4 of gene *IL27*	Crohn's disease (Franke et al., 2010), Inflammatory bowel disease (Imielinski et al., 2009, Jostins et al., 2012), type 1 diabetes (Barrett et al., 2009, Plagnol et al., 2011) In LD with rs151181 in Crohn's disease in Asians (R^2 = 0.882)[2] In LD with rs8049439/rs26528 in Inflammatory bowel disease in Asians(R^2 = 0.468/0.294)[2] In LD with rs4788084 in type 1 diabetes in Asians(R^2 = 0.554)[2]	*IL27*, functions as a heterodimer containing p28 and EBI3, is one of the IFN-β downstream genes, leading to the suppression of lymphocyte response (Murugaiyan et al., 2009). It has been reported that *IL27* plays a critical role of the response to microbial pathogen in leprosy through the induction of immunosuppressive cytokine IL10 and IFN-βas well as the suppression of IFN-γ-induced antimicrobial activity against M.leprae (Teles et al., 2015). *TUFM*, encodes a protein which participates in protein translation in mitochondria. Reduction of TUFM resulted in enhanced IFN-I activation, thus demonstrating a partnership between NLRX1 and TUFM to control host antiviral responses (Lei et al., 2012). *SULT1A2* encodes one type of sulfotransferase, which is responsible for the sulfonation and activation of minoxidil and plays a key role in the phamacogenetics. *CCDC101* also known as *SGF29* or *STAF36*, is a subunit of two histone acetyltransferase complexes (Wang et al., 2008), which has been reported to promote cell survival (Schram et al., 2013). *SPNS1* is critically involved in necrotic or autophagic cell death by lysosomal acidification and trafficking during autophagy, and differentially acts in a pathway with Beclin 1 and p53 in the regulation of senescence (Sasaki et al., 2014).

rs55882956 19p13.2 Variant location: Exon 13 of gene *TYK2*	Crohn's disease (Franke et al., 2010), psoriasis (Genetic Analysis of Psoriasis et al., 2010), Multiple sclerosis (International Multiple Sclerosis Genetics et al., 2011), rheumatoid arthritis (Okada et al., 2014), inflammatory bowel disease (Jostins et al., 2012), type 1 diabetes (Wallace et al., 2010) No LD with Crohn's disease, psoriasis, Multiple sclerosis, rheumatoid arthritis, inflammatory bowel disease, type 1 diabetes associated variants.	*TYK2* gene encodes one member of the tyrosine kinase families and phosphorylates multiple cytokine receptors. TYK2 provides docking sites for STAT as part of the IL-12- and IL-23 signaling cascade and type I IFN signaling (Kilic et al., 2012). Mutation p.E775K within the *Tyk2* pseudokinase domain has been reported to alter susceptibility to infectious or autoimmune diseases through impaired cellular responses to IL-12, IL-23 and Type I IFNs (Shaw et al., 2003). Rs55882956 (p. R703W) which also located in the same domain, may play the same role in the pathogenesis of the infection for intracellular pathogens.

[1]R2 calculate used current study data

[2]R2 from SNAP(SNP Annotation and Proxy Search)Version 2.2

REFERENCES

Alexa A, Rahnenfuhrer J, Lengauer T. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. Bioinformatics 2006;22:1600-7.

Anderson CA, Boucher G, Lees CW, Franke A, D'Amato M, Taylor KD, et al. Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. Nat Genet 2011;43:246-52.

Australo-Anglo-American Spondyloarthritis C, Reveille JD, Sims AM, Danoy P, Evans DM, Leo P, et al. Genome-wide association study of ankylosing spondylitis identifies non-MHC susceptibility loci. Nat Genet 2010;42:123-7.

Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA, et al. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nat Genet 2009;41:703-7.

Barrett JC, Hansoul S, Nicolae DL, Cho JH, Duerr RH, Rioux JD, et al. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. Nat Genet 2008;40:955-62.

Bertin J, Guo Y, Wang L, Srinivasula SM, Jacobson MD, Poyet JL, et al. CARD9 is a novel caspase recruitment domain-containing protein that interacts with BCL10/CLAP and activates NF-kappa B. J Biol Chem 2000;275:41082-6.

Breslow NE, Clayton DG. Approximate Inference in Generalized Linear Mixed Models. J Am Stat Assoc 1993;88:9-25.

Campeau PM, Lu JT, Sule G, Jiang MM, Bae Y, Madan S, et al. Whole-exome sequencing identifies mutations in the nucleoside transporter gene SLC29A3 in dysosteosclerosis, a form of osteopetrosis. Hum Mol Genet 2012;21:4904-9.

Chen H, Wang C, Conomos MP, Stilp AM, Li Z, Sofer T, et al. Control for Population Structure and Relatedness for Binary Traits in Genetic Association Studies via Logistic Mixed Models. Am J Hum Genet 2016;98:653-66.

Colonna M. All roads lead to CARD9. Nat Immunol 2007;8:554-5.

Consortium UIG, Barrett JC, Lee JC, Lees CW, Prescott NJ, Anderson CA, et al. Genome-wide association study of ulcerative colitis identifies three new susceptibility loci, including the HNF4A region. Nat Genet 2009;41:1330-4.

Ebejer JP, Hill JR, Kelm S, Shi J, Deane CM. Memoir: template-based structure prediction for membrane proteins. Nucleic Acids Res 2013;41:W379-83.

Eswar N, Eramian D, Webb B, Shen MY, Sali A. Protein structure modeling with MODELLER. Methods Mol Biol 2008;426:145-59.

Evans DM, Spencer CC, Pointon JJ, Su Z, Harvey D, Kochan G, et al. Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. Nat Genet 2011;43:761-7.

Fearnhead NS, Winney B, Bodmer WF. Rare variant hypothesis for multifactorial inheritance: susceptibility to colorectal adenomas as a model. Cell Cycle 2005;4:521-5.

Franke A, McGovern DP, Barrett JC, Wang K, Radford-Smith GL, Ahmad T, et al. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. Nat

Genet 2010;42:1118-25.

Genetic Analysis of Psoriasis C, the Wellcome Trust Case Control C, Strange A, Capon F, Spencer CC, Knight J, et al. A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1. Nat Genet 2010;42:985-90.

Glocker EO, Hennigs A, Nabavi M, Schaffer AA, Woellner C, Salzer U, et al. A homozygous CARD9 mutation in a family with susceptibility to fungal infections. N Engl J Med 2009;361:1727-35.

Gross O, Gewies A, Finger K, Schafer M, Sparwasser T, Peschel C, et al. Card9 controls a non-TLR signalling pathway for innate anti-fungal immunity. Nature 2006;442:651-6.

Hou S, Du L, Lei B, Pang CP, Zhang M, Zhuang W, et al. Genome-wide association analysis of Vogt-Koyanagi-Harada syndrome identifies two new susceptibility loci at 1p31.2 and 10q21.3. Nat Genet 2014;46:1007-11.

Houten SM, van Woerden CS, Wijburg FA, Wanders RJ, Waterham HR. Carrier frequency of the V377I (1129G>A) MVK mutation, associated with Hyper-IgD and periodic fever syndrome, in the Netherlands. Eur J Hum Genet 2003;11:196-200.

Hruz P, Eckmann L. Caspase recruitment domain-containing sensors and adaptors in intestinal innate immunity. Curr Opin Gastroenterol 2008;24:108-14.

Hsu CL, Lin W, Seshasayee D, Chen YH, Ding X, Lin Z, et al. Equilibrative nucleoside transporter 3 deficiency perturbs lysosome function and macrophage homeostasis. Science 2012;335:89-92.

Hsu YM, Zhang Y, You Y, Wang D, Li H, Duramad O, et al. The adaptor protein CARD9 is required for innate immune responses to intracellular pathogens. Nat Immunol 2007;8:198-205.

Hu Z, Xiong Z, Xu X, Li F, Lu L, Li W, et al. Loss-of-function mutations in filaggrin gene associate with psoriasis vulgaris in Chinese population. Hum Genet 2012;131:1269-74.

Huang J, Ellinghaus D, Franke A, Howie B, Li Y. 1000 Genomes-based imputation identifies novel and refined associations for the Wellcome Trust Case Control Consortium phase 1 Data. Eur J Hum Genet 2012;20:801-5.

Imielinski M, Baldassano RN, Griffiths A, Russell RK, Annese V, Dubinsky M, et al. Common variants at five new loci associated with early-onset inflammatory bowel disease. Nat Genet 2009;41:1335-40.

International Multiple Sclerosis Genetics C, Wellcome Trust Case Control C, Sawcer S, Hellenthal G, Pirinen M, Spencer CC, et al. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature 2011;476:214-9.

Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP, Hui KY, et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature 2012;491:119-24.

Kenny EE, Pe'er I, Karban A, Ozelius L, Mitchell AA, Ng SM, et al. A genome-wide scan of Ashkenazi Jewish Crohn's disease suggests novel susceptibility loci. PLoS Genet 2012;8:e1002559.

Kilic SS, Hacimustafaoglu M, Boisson-Dupuis S, Kreins AY, Grant AV, Abel L, et al. A patient with tyrosine kinase 2 deficiency without hyper-IgE syndrome. J Pediatr 2012;160:1055-7.

Kugathasan S, Baldassano RN, Bradfield JP, Sleiman PM, Imielinski M, Guthery SL, et al. Loci on 20q13 and 21q22 are associated with pediatric-onset inflammatory bowel disease. Nat Genet 2008;40:1211-5.

Lee S, Wu MC, Lin X. Optimal tests for rare variant effects in sequencing association studies. Biostatistics 2012;13:762-75.

Lei Y, Wen H, Yu Y, Taxman DJ, Zhang L, Widman DG, et al. The mitochondrial proteins NLRX1 and TUFM form a complex that regulates type I interferon and autophagy. Immunity

2012;36:933-46.

Libioulle C, Louis E, Hansoul S, Sandor C, Farnir F, Franchimont D, et al. Novel Crohn disease locus identified by genome-wide association maps to a gene desert on 5p13.1 and modulates expression of PTGER4. PLoS Genet 2007;3:e58.

Lim CS, Park ES, Kim DJ, Song YH, Eom SH, Chun JS, et al. SPIN90 (SH3 protein interacting with Nck, 90 kDa), an adaptor protein that is developmentally regulated during cardiac myocyte differentiation. J Biol Chem 2001;276:12871-8.

Liu H, Bao F, Irwanto A, Fu X, Lu N, Yu G, et al. An association study of TOLL and CARD with leprosy susceptibility in Chinese population. Hum Mol Genet 2013;22:4430-7.

Liu H, Irwanto A, Fu X, Yu G, Yu Y, Sun Y, et al. Discovery of six new susceptibility loci and analysis of pleiotropic effects in leprosy. Nat Genet 2015;47:267-71.

Liu H, Irwanto A, Tian H, Fu X, Yu Y, Yu G, et al. Identification of IL18RAP/IL18R1 and IL12B as leprosy risk genes demonstrates shared pathogenesis between inflammation and infectious diseases. Am J Hum Genet 2012;91:935-41.

Liu JZ, Tozzi F, Waterworth DM, Pillai SG, Muglia P, Middleton L, et al. Meta-analysis and imputation refines the association of 15q25 with smoking quantity. Nat Genet 2010;42:436-40.

McGovern DP, Gardet A, Torkvist L, Goyette P, Essers J, Taylor KD, et al. Genome-wide association identifies multiple ulcerative colitis susceptibility loci. Nat Genet 2010a;42:332-7.

McGovern DP, Jones MR, Taylor KD, Marciante K, Yan X, Dubinsky M, et al. Fucosyltransferase 2 (FUT2) non-secretor status is associated with Crohn's disease. Hum Mol Genet 2010b;19:3468-76.

Mizuki N, Meguro A, Ota M, Ohno S, Shiota T, Kawagoe T, et al. Genome-wide association studies identify IL23R-IL12RB2 and IL10 as Behcet's disease susceptibility loci. Nat Genet 2010;42:703-6.

Murugaiyan G, Mittal A, Lopez-Diego R, Maier LM, Anderson DE, Weiner HL. IL-27 is a key regulator of IL-10 and IL-17 production by human CD4+ T cells. J Immunol 2009;183:2435-43.

Nair RP, Duffin KC, Helms C, Ding J, Stuart PE, Goldgar D, et al. Genome-wide scan reveals association of psoriasis with IL-23 and NF-kappaB pathways. Nat Genet 2009;41:199-204.

Nemoto-Hasebe I, Akiyama M, Nomura T, Sandilands A, McLean WH, Shimizu H. FLG mutation p.Lys4021X in the C-terminal imperfect filaggrin repeat in Japanese patients with atopic eczema. Br J Dermatol 2009;161:1387-90.

Oh H, Kim H, Chung KH, Hong NH, Shin B, Park WJ, et al. SPIN90 knockdown attenuates the formation and movement of endosomal vesicles in the early stages of epidermal growth factor receptor endocytosis. PloS one 2013;8:e82610.

Okada Y, Wu D, Trynka G, Raj T, Terao C, Ikari K, et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. Nature 2014;506:376-81.

Ottenhoff TH, Verreck FA, Hoeve MA, van de Vosse E. Control of human host immunity to mycobacteria. Tuberculosis (Edinb) 2005;85:53-64.

Plagnol V, Howson JM, Smyth DJ, Walker N, Hafler JP, Wallace C, et al. Genome-wide association analysis of autoantibody positivity in type 1 diabetes cases. PLoS Genet 2011;7:e1002216.

Price AL, Weale ME, Patterson N, Myers SR, Need AC, Shianna KV, et al. Long-range LD can confound genome scans in admixed populations. Am J Hum Genet 2008;83:132-5; author reply 5-9.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559-75.

Remmers EF, Cosan F, Kirino Y, Ombrello MJ, Abaci N, Satorius C, et al. Genome-wide association study identifies variants in the MHC class I, IL10, and IL23R-IL12RB2 regions associated with Behcet's disease. Nat Genet 2010;42:698-702.

Rioux JD, Xavier RJ, Taylor KD, Silverberg MS, Goyette P, Huett A, et al. Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. Nat Genet 2007;39:596-604.

Ronty M, Taivainen A, Heiska L, Otey C, Ehler E, Song WK, et al. Palladin interacts with SH3 domains of SPIN90 and Src and is required for Src-induced cytoskeletal remodeling. Exp Cell Res 2007;313:2575-85.

Sandilands A, O'Regan GM, Liao H, Zhao Y, Terron-Kwiatkowski A, Watson RM, et al. Prevalent and rare mutations in the gene encoding filaggrin cause ichthyosis vulgaris and predispose individuals to atopic dermatitis. J Invest Dermatol 2006;126:1770-5.

Sandilands A, Sutherland C, Irvine AD, McLean WH. Filaggrin in the frontline: role in skin barrier function and disease. J Cell Sci 2009;122:1285-94.

Sandilands A, Terron-Kwiatkowski A, Hull PR, O'Regan GM, Clayton TH, Watson RM, et al. Comprehensive analysis of the gene encoding filaggrin uncovers prevalent and rare mutations in ichthyosis vulgaris and atopic eczema. Nat Genet 2007;39:650-4.

Sasaki T, Lian S, Qi J, Bayliss PE, Carr CE, Johnson JL, et al. Aberrant autolysosomal regulation is linked to the induction of embryonic senescence: differential roles of Beclin 1 and p53 in vertebrate Spns1 deficiency. PLoS Genet 2014;10:e1004409.

Satoh S, Tominaga T. mDia-interacting protein acts downstream of Rho-mDia and modifies Src activation and stress fiber formation. J Biol Chem 2001;276:39290-4.

Schram AW, Baas R, Jansen PW, Riss A, Tora L, Vermeulen M, et al. A dual role for SAGA-associated factor 29 (SGF29) in ER stress survival by coordination of both histone H3 acetylation and histone H3 lysine-4 trimethylation. PloS one 2013;8:e70035.

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.

Shaw MH, Boyartchuk V, Wong S, Karaghiosoff M, Ragimbeau J, Pellegrini S, et al. A natural mutation in the Tyk2 pseudokinase domain underlies altered susceptibility of B10.Q/J mice to infection and autoimmunity. Proc Natl Acad Sci U S A 2003;100:11594-9.

Silverberg MS, Cho JH, Rioux JD, McGovern DP, Wu J, Annese V, et al. Ulcerative colitis-risk loci on chromosomes 1p36 and 12q15 found by genome-wide association study. Nat Genet 2009;41:216-20.

Stockinger B, Veldhoen M. Differentiation and function of Th17 T cells. Curr Opin Immunol 2007;19:281-6.

Sun LD, Xiao FL, Li Y, Zhou WM, Tang HY, Tang XF, et al. Genome-wide association study identifies two new susceptibility loci for atopic dermatitis in the Chinese Han population. Nat Genet 2011;43:690-4.

Tang XF, Zhang Z, Hu DY, Xu AE, Zhou HS, Sun LD, et al. Association analyses identify three susceptibility Loci for vitiligo in the Chinese Han population. J Invest Dermatol

2013;133:403-10.

Teles RM, Kelly-Scumpia KM, Sarno EN, Rea TH, Ochoa MT, Cheng G, et al. IL-27 Suppresses Antimicrobial Activity in Human Leprosy. J Invest Dermatol 2015;135:2410-7.

Teodorof C, Bae JI, Kim SM, Oh HJ, Kang YS, Choi J, et al. SPIN90-IRSp53 complex participates in Rac-induced membrane ruffling. Exp Cell Res 2009;315:2410-9.

Van Durme J, Delgado J, Stricher F, Serrano L, Schymkowitz J, Rousseau F. A graphical interface for the FoldX forcefield. Bioinformatics 2011;27:1711-2.

Voruganti VS, Laston S, Haack K, Mehta NR, Smith CW, Cole SA, et al. Genome-wide association replicates the association of Duffy antigen receptor for chemokines (DARC) polymorphisms with serum monocyte chemoattractant protein-1 (MCP-1) levels in Hispanic children. Cytokine 2012;60:634-8.

Wallace C, Smyth DJ, Maisuria-Armer M, Walker NM, Todd JA, Clayton DG. The imprinted DLK1-MEG3 gene region on chromosome 14q32.2 alters susceptibility to type 1 diabetes. Nat Genet 2010;42:68-71.

Wang YL, Faiola F, Xu M, Pan S, Martinez E. Human ATAC Is a GCN5/PCAF-containing acetylase complex with a novel NC2-like histone fold module that interacts with the TATA-binding protein. J Biol Chem 2008;283:33808-15.

Wang Z, Sun Y, Fu X, Yu G, Wang C, Bao F, et al. A large-scale genome-wide association and meta-analysis identified four novel susceptibility loci for leprosy. Nat Commun 2016;7:13760.

Wellcome Trust Case Control C. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447:661-78.

Yang SK, Hong M, Zhao W, Jung Y, Baek J, Tayebi N, et al. Genome-wide association study of Crohn's disease in Koreans revealed three new susceptibility loci and common attributes of genetic susceptibility across ethnic populations. Gut 2014;63:80-7.

Yang W, Tang H, Zhang Y, Tang X, Zhang J, Sun L, et al. Meta-analysis followed by replication identifies loci in or near CDKN1B, TET3, CD80, DRAM1, and ARID5B as associated with systemic lupus erythematosus in Asians. Am J Hum Genet 2013;92:41-51.

Zhang F, Liu H, Chen S, Low H, Sun L, Cui Y, et al. Identification of two new loci at IL23R and RAB32 that influence susceptibility to leprosy. Nat Genet 2011;43:1247-51.

Zhang FR, Huang W, Chen SM, Sun LD, Liu H, Li Y, et al. Genomewide association study of leprosy. N Engl J Med 2009;361:2609-18.