Table S1 List of 176 selected genes in the lymphocyte activation involved in immune response related gene-set used in the discovery analysis

Dataset	Name of pathway	Selected genes[a]	Number of genes
GO	LYMPHOCYTE_ACTIVATION_INVOLVED_IN_IMMN_IMMUNE_RESPONSE	*ABL1, ADA, AICDA, ANXA1, AP1G1, APBB1IP, APLF, ATAD5, ATP7A, BATF, BCL3, BCL6, C17orf99, CCL19, CCR6, CD180, CD19, CD1C, CD244, CD28, CD40, CD40LG, CD46, CD80, CD86, CDH17, CEACAM1, CLCF1, CLEC4D, CLEC4E, CORO1A, DLL1, DOCK10, DOCK11, EIF2AK4, EOMES, ERCC1, EXO1, EXOSC3, EXOSC6, F2RL1, FCER1G, FCGR2B, FOXP1, FOXP3, GAPT, GATA3, GPR183, HAVCR2, HLA-DMB, HLA-F, HLX, HMGB1, HSPD1, ICAM1, IFNA1, IFNA10, IFNA13, IFNA14, IFNA16, IFNA17, IFNA2, IFNA21, IFNA4, IFNA5, IFNA6, IFNA7, IFNA8, IFNB1, IFNE, IFNG, IFNK, IFNL1, IFNW1, IL10, IL12B, IL12RB1, IL18, IL18R1, IL2, IL23A, IL23R, IL27, IL27RA, IL4, IL4R, IL6, IRF4, ITFG2, ITGAL, ITM2A, JAK3, KLRF2, LAMP1, LCP1, LEF1, LFNG, LGALS1, LGALS3, LGALS9, LIG4, LILRB1, LOXL3, LY9, MAD2L2, MALT1, MFNG, MIR21, MLH1, MSH2, MSH6, MTOR, MYB, NBN, NDFIP1, NFKBID, NFKBIZ, NKX2-3, NLRP3, NOTCH2, NSD2, PARP3, PAXIP1, PGLYRP1, PGLYRP2, PGLYRP3, PGLYRP4, PLCG2, PLCL2, PRKCZ, PSEN1, PTGER4, PTK2B, PTPRC, RAB27A, RARA, RC3H1, RC3H2, RELB, RIF1, RIPK2, RNF168, RNF8, RORA, RORC, SEMA4A, SHLD1, SHLD2, SHLD3, SLAMF6, SLC11A1, SMAD7, SOCS5, SPN, STAT3, STAT6, SUPT6H, SWAP70, TBX21, TFRC, TGFB1, THOC1, TLR4, TNFSF13, TNFSF18, TNFSF4, TP53BP1, TSC1, UNC13D, UNG, VAMP7, XBP1, ZBTB7B, ZC3H12A, ZFPM1, ZNF683*	
Total			176

[a] Genes were selected based on online Hallmark datasets (http://software.broadinstitute.org/gsea/msigdb/search.jsp) and literatures. Keyword: lymphocyte AND immune. Organism: Homo sapiens.
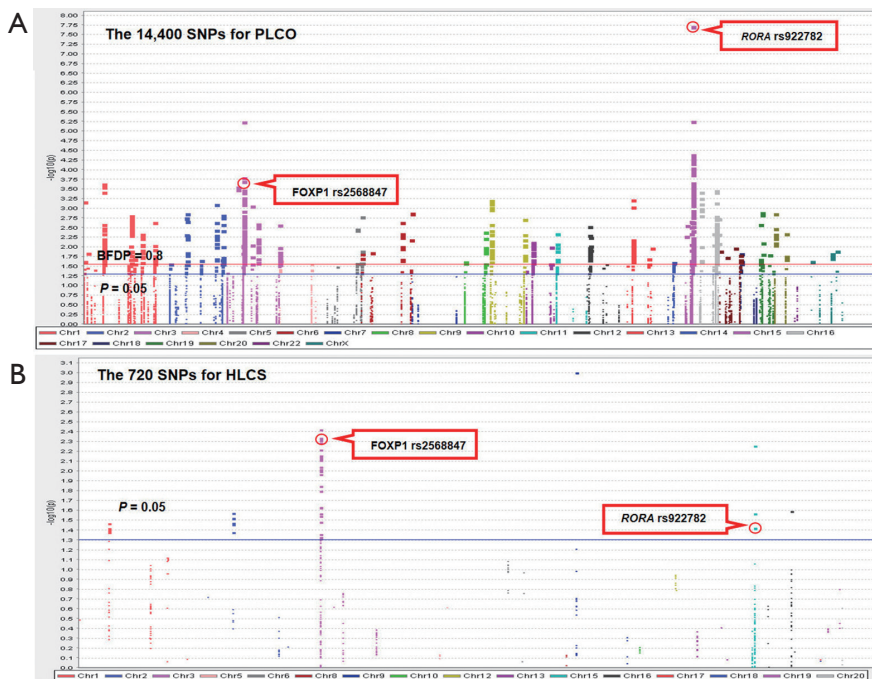

Figure S1 Manhattan plot for 14,400 SNPs of lymphocyte activation involved in immune response-related pathway genes in the PLCO trial (A). Manhattan plot for 720 SNPs in the HLCS dataset. (B). The red line represents BFDP=0.80 and the blue horizontal line indicates P=0.05. BFDP, Bayesian false-discovery probability; PLCO, Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial; HLCS, the Harvard Lung Cancer Susceptibility Study.
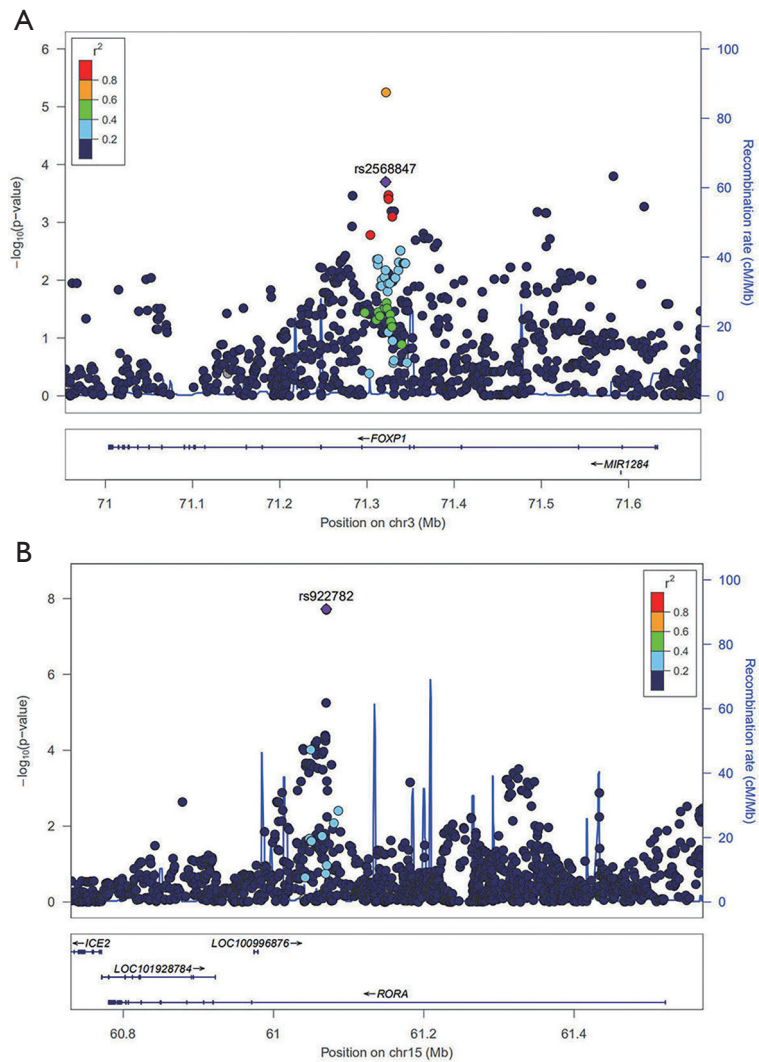
Figure S2 Regional association plots for two independent SNPs in the lymphocyte activation involved in immune response-related pathway genes. Regional association plots included 50 kb up or downstream of *FOXP1* (A) and *RORA* (B). Data points are labeled by color based on the level of LD of each pair of SNPs according to the hg19/1000 Genomes European population. The left y-axis represents the association P-value of each SNPs in the discovery cohort, potting as -log10 (*P*) against chromosomal base-pair position. The right y-axis represents the recombination rate calculated from HapMap Data Rel 22/phase II European population. The regional association plots were generated using LocusZoom (http://locuszoom.org/). SNPs, single-nucleotide polymorphisms; LD, linkage disequilibrium.

Table S2 Stratified analysis for associations between the unfavourable genotypes and survival of NSCLC in the PLCO

Characteristics	0 unfavourable genotype	1-2 unfavourable genotype	Multivariate analysis[b] for OS			Multivariate analysis[b] for DSS		
	Frequency[a]	Frequency[a]	HR (95% CI)	P	P_{inter}[c]	HR (95% CI)	P	P_{inter}[c]
Age (years)								
≤71	229	405	1.21 (0.98-1.50)	0.0715		1.24 (0.99-1.55)	0.0553	
>71	207	334	1.78 (1.43- 2.22)	<0.0001	0.0432	1.91 (1.51-2.41)	<0.0001	0.0196
Sex								
Male	251	444	1.51 (1.25- 1.83)	<0.0001		1.58 (1.28-1.94)	<0.0001	
Female	185	295	1.42 (1.11- 1.82)	0.0058	0.4293	1.46 (1.13-1.89)	0.0041	0.4891
Smoking status								
Never	28	86	1.49 (0.78- 2.85)	0.2293		1.51 (0.79- 2.90)	0.2133	
Current	152	265	1.62 (1.23- 2.13)	0.0006		1.76 (1.31- 2.36)	0.0002	
Former	256	388	1.38 (1.14- 1.67)	0.0012	0.4985	1.41 (1.14-1.73)	0.0012	0.4004
Histology								
Adeno	230	345	1.43 (1.13-1.80)	0.0025		1.40 (1.10-1.80)	0.0059	
Squamous	96	188	1.49 (1.08-2.05)	0.0147		1.70 (1.19-2.42)	0.0034	
Others	110	206	1.52 (1.15-1.99)	0.0028	0.5826	1.64 (1.22-2.19)	0.0009	0.3008
Tumor stage								
I-IIIA	254	400	1.38 (1.09-1.75)	0.0077		1.45 (1.11-1.90)	0.0060	
IIIB-IV	182	339	1.53 (1.25-1.86)	<0.0001	0.9359	1.62 (1.32-1.98)	<0.0001	0.5346
Chemotherapy								
No	253	385	1.69 (1.35-2.13)	<0.0001		1.81 (1.41- 2.32)	<0.0001	
Yes	183	354	1.22 (0.99- 1.50)	0.0564	0.0383	1.28 (1.03-1.57)	0.0240	0.0301
Radiotherapy								
No	273	488	1.40 (1.15- 1.72)	0.0009		1.51 (1.21-1.88)	0.0002	
Yes	163	251	1.50 (1.19- 1.89)	0.0005	0.7926	1.51 (1.19-1.92)	0.0007	0.7798
Surgery								
No	234	401	1.40 (1.17-1.68)	0.0002		1.45 (1.21-1.75)	<0.0001	
Yes	202	338	1.25 (0.94- 1.68)	0.1277	0.7185	1.33 (0.95-1.86)	0.0978	0.9648

[a], 10 missing date were excluded; [b], adjusted for age, sex, stage, histology, smoking status, chemotherapy, radiotherapy, surgery. PC1, PC2, PC3, and PC4; [c], P inter: P value for interaction analysis between characteristic and protective alleles. OS, overall survival; DSS, disease-specific survival; NSCLC, non-small cell lung cancer; PLCO, the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial; HR, hazards ratio; CI, confidence interval.
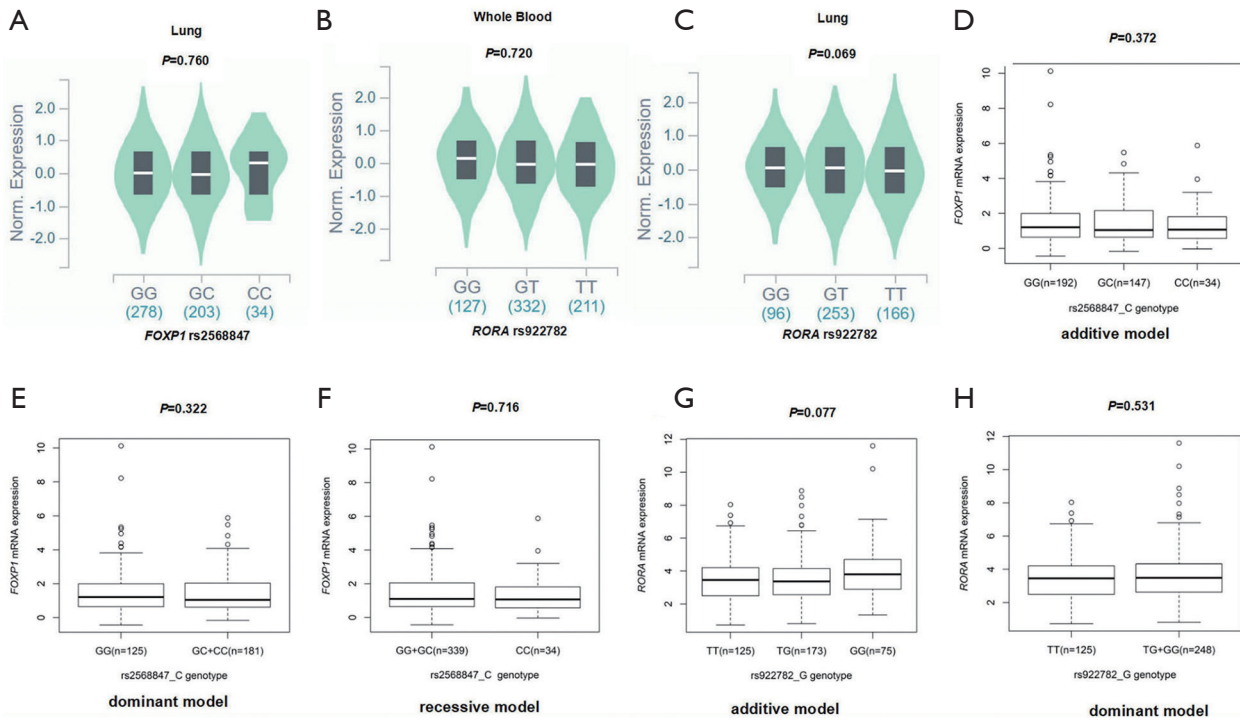
Figure S3 The eQTLs analysis for FOXP1 rs2568847 and RORA rs922782. The correlation of rs2568847 genotypes and FOXP1 mRNA expression in lung tissue samples (A), the correlation of rs922782 genotypes and RORA mRNA expression in whole blood samples (B), in lung tissue samples (C) from the GTEx database. The correlation of rs2568847 genotypes and FOXP1 mRNA expression in the additive (D), dominant (E) and recessive (F) model from the 1000 Genomes Project. The correlation of rs922782 genotypes and RORA mRNA expression in additive (G), dominant (H) model from the 1000 Genomes Project. eQTLs, expression quantitative trait loci; GTEx, Genotype-Tissue Expression project.

Table S3 Function prediction for FOXP1 rs2568847 and RORA rs922782

| SNP | Gene | Chr | RegDB[1] | Haploreg v4.1[2] | | | | | | |
				Promoter histone marks	Enhancer histone marks	DNAse	Proteins bound	Motifs changed	Selected eQTL hits	dbSNP func annot
rs2568847	*FOXP1*	3	7	--	BLD, BRN	--	--	BATF, Mef2, RXRA	--	Intronic
rs922782	*RORA*	15	5		4 tissues	--	--	Ets, Foxm1, Pou5f1	--	Intronic

1 RegDB: http://www.regulomedb.org; 2 Haploreg: http://archive.broadinstitute.org/mammals/haploreg/haploreg.php. SNP, single nucleotide polymorphism; Chr, chromosome; DNAse, deoxyribonuclease; eQTL, expression quantitative trait loci; dbSNP func annot, dbSNP function annotation.

Figure S4 Functional prediction of two independent SNPs in lymphocyte activation involved in immune response-related pathway genes in the ENCODE data. Location and functional prediction of *FOXP1* rs2568847 (A). Location and functional prediction of *RORA* rs922782. (B). The tracks of H3K4Me1, H3K4Me3, and H3K27Ac tracks demonstrated the genome-wide levels of enrichment of the mono-methylation of lysine 4, and tri-methylation of lysine 4 and acetylation of lysine 27, of the H3 histone protein. DNase clusters track illustrated DNase hypersensitivity sites. Tnx factor track showed DNA binding locations by transcription factor.

Figure S5 Mutation analysis of *FOXP1* and *RORA* in non-small cell lung tumor tissues by using publicly available data in the database of the cBioPortal for Cancer Genomics (http://www.cbioportal.org). (A) *FOXP1* had a relatively high mutation frequency in NSCLC. (B) *RORA* had a low mutation frequency in NSCLC.