**Table S1.** Comparison of the characteristics between the PLCO trial and the HLCS study

| Characteristics | PLCO | | HLCS | | $P^a$ |
|---|---|---|---|---|---|
| | Frequency | Deaths (%) | Frequency | Deaths (%) | |
| Total | 1185 | 798 (67.3) | 984 | 665 (67.5) | |
| Median overall survival (months) | 23.8 | | 39.9 | | |
| Age | | | | | |
| ≤ 71 | 636 | 400 (62.9) | 654 | 428 (65.4) | |
| > 71 | 549 | 398 (72.5) | 330 | 237 (71.8) | < 0.0001 |
| Sex | | | | | |
| Male | 698 | 507 (72.6) | 507 | 379 (74.7) | |
| Female | 487 | 291 (59.8) | 477 | 286 (59.9) | 0.0006 |
| Smoking status | | | | | |
| Never | 115 | 63 (54.8) | 92 | 52 (56.5) | |
| Current | 423 | 272 (64.3) | 390 | 266 (68.2) | |
| Former | 647 | 463 (71.6) | 502 | 347 (69.1) | 0.166 |
| Histology | | | | | |
| Adenocarcinoma | 577 | 348 (60.3) | 597 | 378 (63.3) | |
| Squamous cell carcinoma | 285 | 192 (67.4) | 216 | 156 (72.2) | |
| Others | 323 | 258 (79.9) | 171 | 131 (76.6) | < 0.0001 |
| Stage | | | | | |
| I-IIIA | 655 | 315 (48.1) | 606 | 352 (58.0) | |
| IIIB-IV | 528 | 482 (91.3) | 377 | 313 (83.0) | 0.003 |
| Missing | 2 | | – | | |

[a]Chi-square test for comparisons of the characteristics between the PLCO trail and the Harvard study. Abbreviations: PLCO, Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial; HLCS, Harvard lung cancer susceptibility study.

**Table S2.** List of 176 selected genes in the NLRP3 Inflammasome gene-set from MSigDB and Genecards

| Dataset | Name of pathway[a] | Selected genes | Number of genes |
|---|---|---|---|
| KEGG | NOD-like receptor signaling pathway | *BIRC2, BIRC3, CARD18, CARD6, CARD8, CARD9, CASP1, CASP5, CASP8, CCL11, CCL13, CCL2, CCL5, CCL7, CCL8, CHUK, CXCL1, CXCL2, ERBB2IP, HSP90AA1, HSP90AB1, HSP90B1, IKBKB, IKBKG, IL18, IL1B, IL6, IL8, MAP3K7, MAPK1, MAPK10, MAPK11, MAPK12, MAPK13, MAPK14, MAPK3, MAPK8, MAPK9, MEFV, NAIP, NFKB1, NFKBIA, NFKBIB, NLRC4, NLRP1, NLRP3, NOD1, NOD2, PSTPIP1, PYCARD, PYDC1, RELA, RIPK2, SUGT1, TAB1, TAB2, TAB3, TNF, TNFAIP3, TRAF6, TRIP6, XIAP* | 62 |
| REACTOME | Inflammasomes | *AIM2, APP, BCL2, BCL2L1, CASP1, HSP90AB1, LOC644816, MEFV, NLRC4, NLRP1, NLRP3, P2RX7, PANX1, PSTPIP1, PYCARD, TXN, TXNIP* | 17 |
| REACTOME | The NLRP3 inflammasome | *APP, CASP1, HSP90AB1, LOC644816, MEFV, NLRP3, P2RX7, PANX1, PSTPIP1, PYCARD, TXN, TXNIP* | 12 |
| GO | Regultation of NLRP3 inflammasome complex assembly | *ATAT1, CD36, EIF2AK2, GBP5, MEFV, NLRP2P, PYDC1, PYDC2, SIRT2, TLR4, TLR6* | 11 |
| Genecards (https://www.genecards.org/) | NLRP3_Inflammasome_Lung cancer | *TP53, EGFR, NLRP3, IL1B, TNF, IL6, CASP8, TGFB1, IFNG, IL10, TLR4, BRCA1, CCL2, NLRP1, PRKN, TP73, TLR2, BIRC3, MEFV, IL1RN, CASP1, SERPINA1, CXCL8, PDCD1, BCL2L1, CASP3, BCL2, TLR9, CCL3, NFKB1, CD36, MAPK1, JUN, FGF2, MAPK8, CASP9, MTOR, P2RX7, IL18, CTLA4, BCL10, RELA, CREB1, HSP90AA1, TLR5, NLRC4, NOD2, HMOX1, AGT, CTSB, CFLAR, FLT3, IL1A, ITGB1, CCL5, CLEC7A, CD274, FN1, NAIP, BIRC2, CARD8, CD209, STAT1, BTK, APAF1, HLA-G, CASR, PTPN22, SYK, MALT1, PSTPIP1, IFNA1, IL17A, C3, IL12B, HMGB1, APOA1, S100B, CASP7, CXCL1, SNCA, PIK3CG, LGALS3, NRG1, C5, ATF6, PML, IL1R1, IFIH1, S100A8, MYD88, S100A9, CD40, LEP, NPPA, TLR6, OLR1, IFNB1, DHX9, CAMP, TLR3, SAA1, JAK1, MAPK14, CCL4, NOD1, HSP90AB1, IL6ST, IRF3, CARD11, CYBB, MIF, IRF7, TREX1, TRIM21, RAC1, CARD9, DDX58, UBC, TLR7, H2AFX, UMOD, IRAK3, XDH, LY96, IL18BP, AGER, TNFAIP3, CD14, IL23A, HSPA1A, IL33, ORMDL3, CYBA, NCF1, IRAK4, IRAK1, LBP, IL27, MYO1C, IL37, ATP6V0A2* | 142 |
| Gene removed[a] | | *PRKN, BTK, CYBB, TLR7, IRAK1, LOC644816, IKBKG, IL8, TAB3, XIAP, NLRP2P* | 11 |
| Total genes | | | 176 |

[a]Gene in X chromosome was removed; Keyword for MSigDB and Genecards: NLRP3; Inflammasome. Organism: Homo sapiens. Abbreviation: MSigDB = Molecular signatures database.

**Table S3.** Associations of the first 10 principal components and OS of NSCLC in the PLCO trial

| PC* | Parameter Estimate | Standard Error | Chi-Square | *P* |
|---|---|---|---|---|
| PC1 | 4.821 | 1.353 | 12.697 | < 0.001 |
| PC2 | -0.681 | 1.228 | 0.308 | 0.579 |
| PC3 | -3.054 | 0.949 | 10.351 | 0.001 |
| PC4 | -2.837 | 1.246 | 5.184 | 0.023 |
| PC5 | -0.910 | 1.232 | 0.546 | 0.460 |
| PC6 | 1.355 | 1.252 | 1.172 | 0.279 |
| PC7 | -0.236 | 1.218 | 0.038 | 0.846 |
| PC8 | -1.684 | 1.322 | 1.622 | 0.203 |
| PC9 | -1.886 | 1.267 | 2.216 | 0.137 |
| PC10 | 0.347 | 1.240 | 0.078 | 0.180 |

*The top four PCs were used for the adjustment for potential population stratification in the multivariate analysis. Abbreviations: OS, overall survival; NSCLC, non-small cell lung cancer; PLCO, the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial; PC, principal component.

**Table S4.** Functional prediction analyses of the two validated SNPs

| SNP | Gene | Chr. | Allele | MAF | Regulome DB Score | HaploReg v4.1 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Promoter histone marks | Enhancer histone marks | DNase | Motifs changed | Selected eQTL hits | dbSNP function annotation |
| rs4733124 | *NRG1* | 8 | T/C | 0.19 | 4 | -- | BRST, SKIN | BRST[a], SKIN[b], BRST[c] | GZF1, Pax-4 | -- | intronic |
| rs11225211 | *BIRC3* | 11 | G/A | 0.16 | -- | -- | BLD | -- | 4 altered motifs | 2 hits | intronic |

[a]Breast variant Human Mammary Epithelial Cells; [b]Foreskin Fibroblast Primary Cells; [c]Breast Myoepithelial Primary Cells. Abbreviation: SNP = Single nucleotide polymorphisms; MAF = Minor allele frequency.

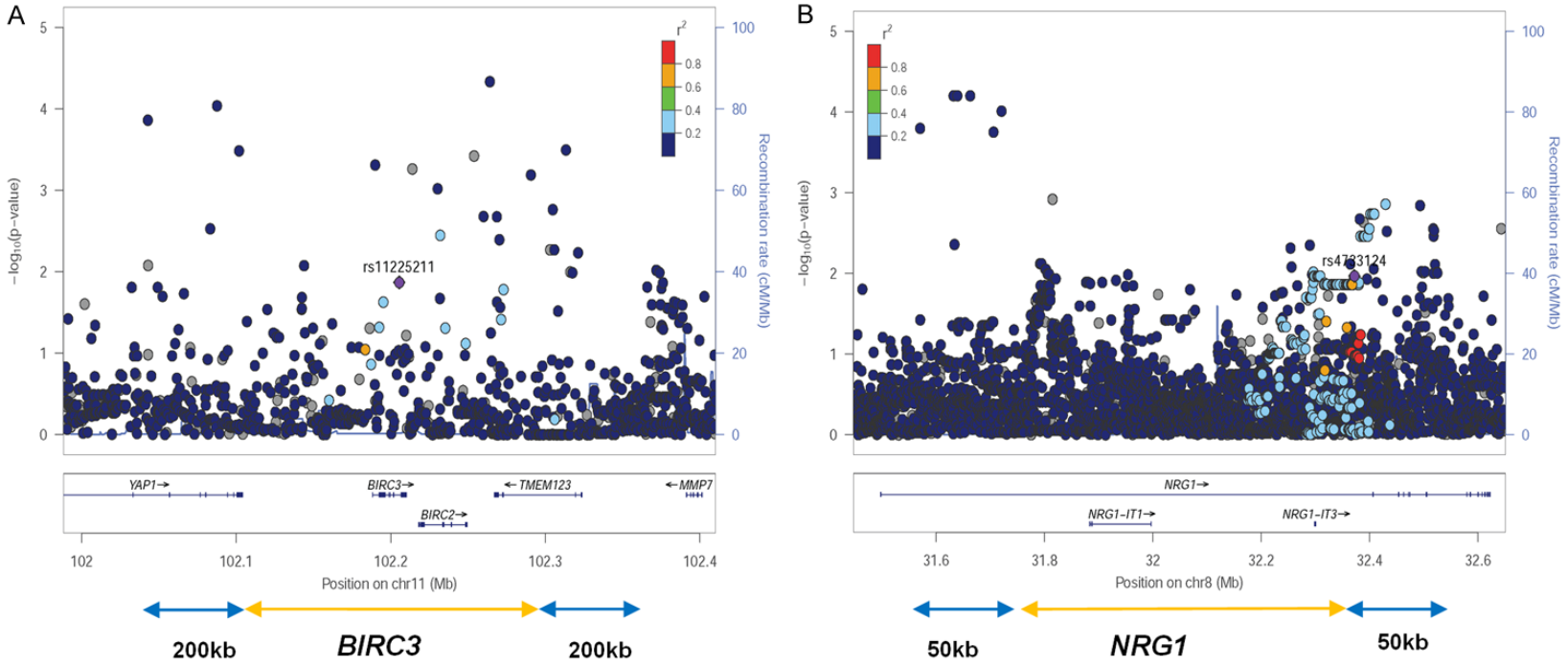


**Figure S1.** Regional association plots. The left-hand Y-axis shows the -log10 transformation of *P*-value of individual SNPs, which is plotted against the chromosomal base-pair position with an expansion of 500 KB in the flanks of the gene region. The right-hand Y-axis shows the recombination rate estimated for European populations from HapMap Data Rel 22/phase II. A. *XDH* rs141674738; B. *NRG1* rs4733124.

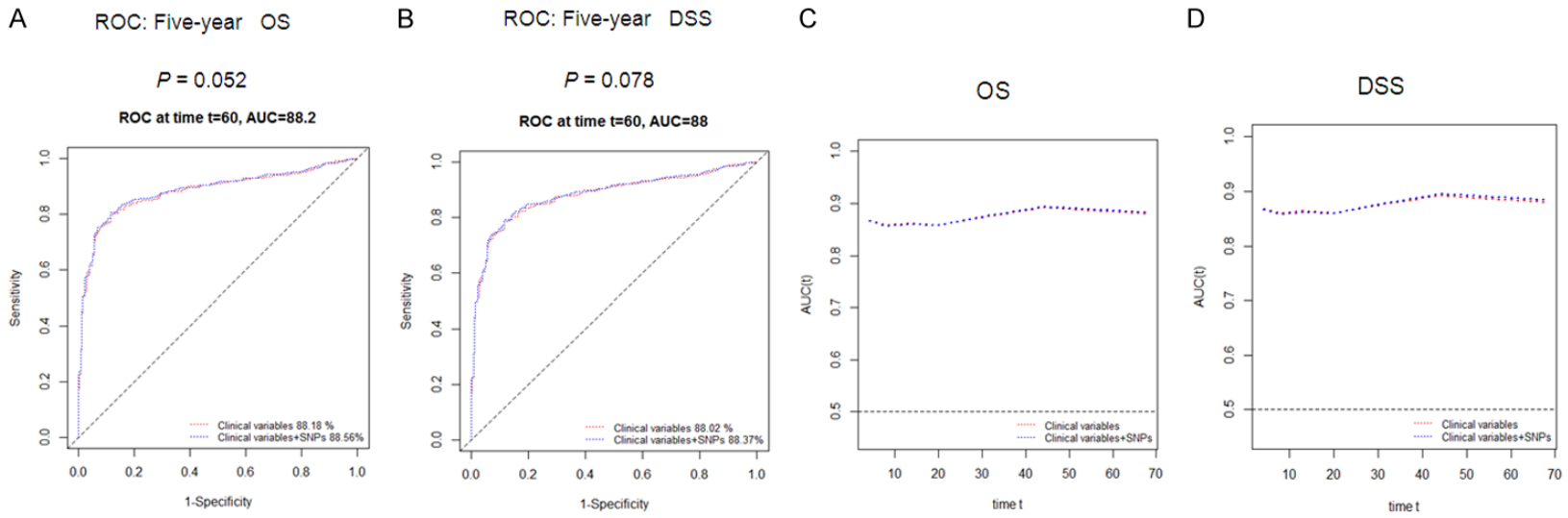


**Figure S2.** Receiver operating characteristic (ROC) curve and time-dependent area under the ROC curve (AUC) estimation for prediction of OS/DSS using the PLCO dataset. A. Five-year NSCLC overall survival rate, red line indicates clinical variables 88.18%; blue line indicates clinical variables + unfavorable genotypes 88.56%; B. Five-year NSCLC disease-specific survival rate, red line indicates clinical variables 88.02%; blue line indicates clinical variables + unfavorable genotypes 88.37%; C. Time-dependent AUC estimation (OS), based on age, sex, smoking status, histology, tumor stage, chemotherapy, radiotherapy, surgery, principal components and the protective genotypes of the two SNPs. time t = time (months); D. Time-dependent AUC estimation (DSS), based on age, sex, smoking status, histology, tumor stage, chemotherapy, radiotherapy, surgery, principal components and the protective genotypes of the two SNPs. time t = time (months).

**Table S5.** Stratified multivariate analyses for association between risk genotypes and OS/DSS in NSCLC patients in the PLCO trial

| Characteristics | No. of 1-2 risk genotypes | | No. of 0 risk genotype | | OS Multivariate analysis[a] | | $P_{inter}$ [b] | DSS Multivariate analysis[a] | | $P_{inter}$ [b] |
|---|---|---|---|---|---|---|---|---|---|---|
| | All | Death (%) | All | Death (%) | HR (95% CI) | *P* | | HR (95% CI) | *P* | |
| Age (years) | | | | | | | | | | |
| ≤ 71 | 503 | 334 (66.4) | 131 | 65 (49.6) | 0.80 (0.61-1.05) | 0.108 | | 0.92 (0.70-1.21) | 0.551 | |
| > 71 | 406 | 295 (72.7) | 133 | 94 (70.7) | 0.79 (0.62-1.00) | 0.051 | 0.088 | 0.77 (0.60-1.00) | 0.048 | 0.924 |
| Sex | | | | | | | | | | |
| Male | 533 | 399 (74.9) | 160 | 105 (65.6) | 0.80 (0.64-0.99) | 0.044 | | 0.86 (0.68-1.08) | 0.201 | |
| Female | 376 | 230 (61.2) | 104 | 54 (51.9) | 0.71 (0.53-0.97) | 0.030 | 0.787 | 0.73 (0.53-1.00) | 0.047 | 0.468 |
| Smoking status | | | | | | | | | | |
| Never | 81 | 43 (53.1) | 33 | 19 (57.6) | 1.04 (0.56-1.95) | 0.903 | | 1.01 (0.54-1.89) | 0.989 | |
| Current | 328 | 215 (65.6) | 87 | 50 (57.5) | 0.72 (0.53-1.00) | 0.046 | | 0.80 (0.58-1.12) | 0.198 | |
| Former | 500 | 371 (74.2) | 144 | 90 (62.50) | 0.81 (0.64-1.02) | 0.075 | 0.558 | 0.85 (0.66-1.08) | 0.178 | 0.613 |
| Histology | | | | | | | | | | |
| Adenocarcinoma | 429 | 271 (63.2) | 145 | 75 (51.7) | 0.62 (0.48-0.81) | 0.0004 | | 0.67 (0.51-0.88) | 0.004 | |
| Squamous | 220 | 147 (66.8) | 63 | 43 (68.3) | 1.08 (0.76-1.54) | 0.681 | | 1.29 (0.89-1.87) | 0.177 | |
| Others | 260 | 211 (81.2) | 56 | 41 (73.2) | 0.85 (0.60-1.20) | 0.342 | 0.032 | 0.80 (0.55-1.16) | 0.237 | 0.432 |
| Tumor stage | | | | | | | | | | |
| I-IIIA | 504 | 260 (51.6) | 149 | 54 (36.2) | 0.68 (0.51-0.92) | 0.012 | | 0.75 (0.54-1.04) | 0.089 | |
| IIIB-IV | 405 | 369 (91.1) | 115 | 105 (91.3) | 0.84 (0.67-1.05) | 0.119 | 0.097 | 0.85 (0.68-1.07) | 0.162 | 0.387 |
| Chemotherapy | | | | | | | | | | |
| No | 490 | 300 (61.2) | 147 | 66 (44.9) | 0.61 (0.46-0.81) | 0.0005 | | 0.64 (0.47-0.86) | 0.003 | |
| Yes | 419 | 329 (78.5) | 117 | 93 (79.5) | 1.01 (0.80-1.29) | 0.911 | 0.007 | 1.05 (0.82-1.33) | 0.719 | 0.010 |
| Radiotherapy | | | | | | | | | | |
| No | 582 | 363 (62.4) | 178 | 86 (48.3) | 0.75 (0.59-0.95) | 0.019 | | 0.80 (0.62-1.04) | 0.096 | |
| Yes | 327 | 266 (81.4) | 86 | 73 (84.9) | 0.88 (0.67-1.15) | 0.334 | 0.436 | 0.90 (0.69-1.19) | 0.457 | 0.564 |
| Surgery | | | | | | | | | | |
| No | 497 | 439 (88.3) | 137 | 125 (91.2) | 0.95 (0.77-1.17) | 0.614 | | 0.97 (0.78-1.20) | 0.763 | |
| Yes | 412 | 190 (46.1) | 127 | 34 (26.8) | 0.49 (0.34-0.72) | 0.0002 | 0.002 | 0.55 (0.36-0.83) | 0.004 | 0.012 |

[a]Adjusted for age, sex, stage, histology, smoking status, chemotherapy, radiotherapy, surgery, PC1, PC2, PC3 and PC4; One observations missing of *NRG1* rs4733124, One observations missing of *BIRC3* rs11225211; Two observations missing of tumor stage and eight observations missing of chemotherapy/radiotherapy/surgery in PLCO dataset. [b]$P_{inter}$: *P*-value for interaction analysis between characteristic and number of protective genotypes; Abbreviation: OS = Overall survival; NSCLC = Non-small cell lung cancer; HR = Hazards ratio; 95% CI = 95% Confidence interval.

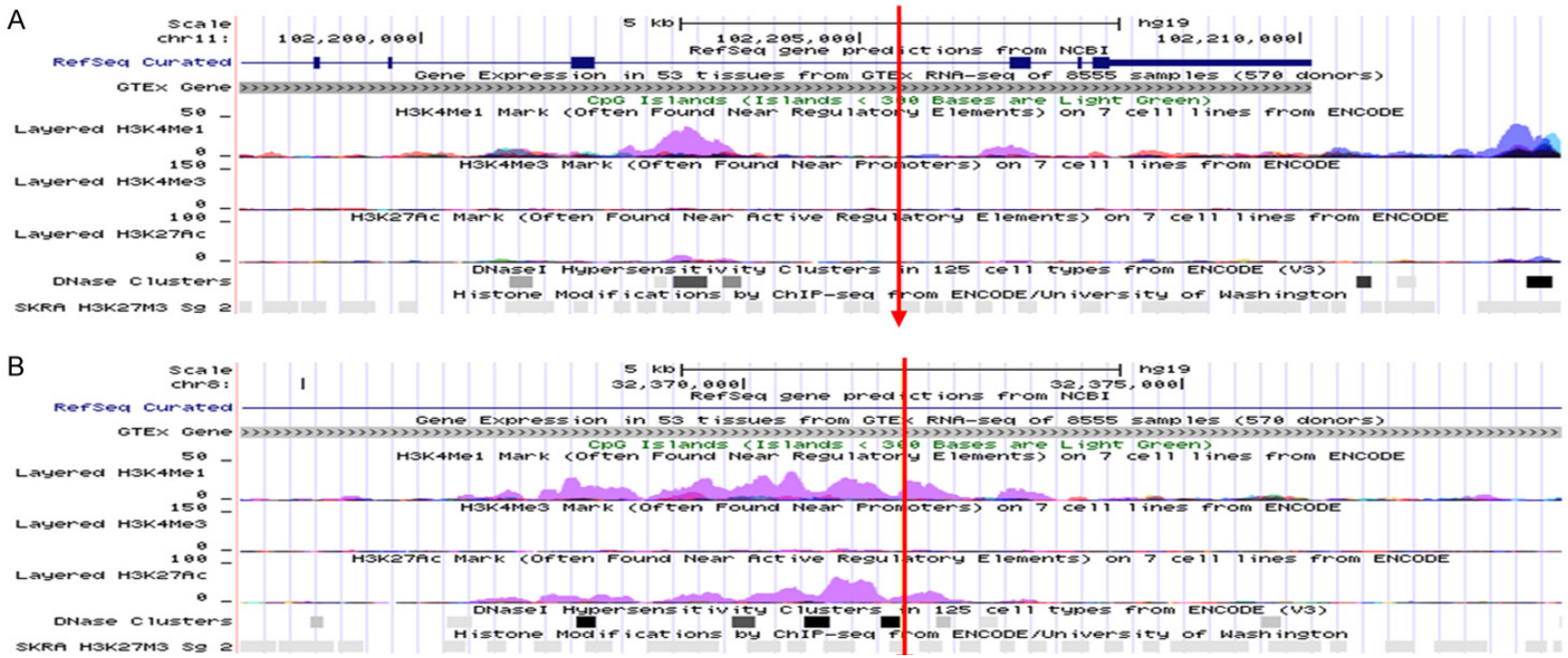


**Figure S3.** Functional prediction of SNPs in the ENCODE project. (A) Location and functional prediction of SNPs rs11225211, (B) Location and functional prediction of SNPs rs4733124.

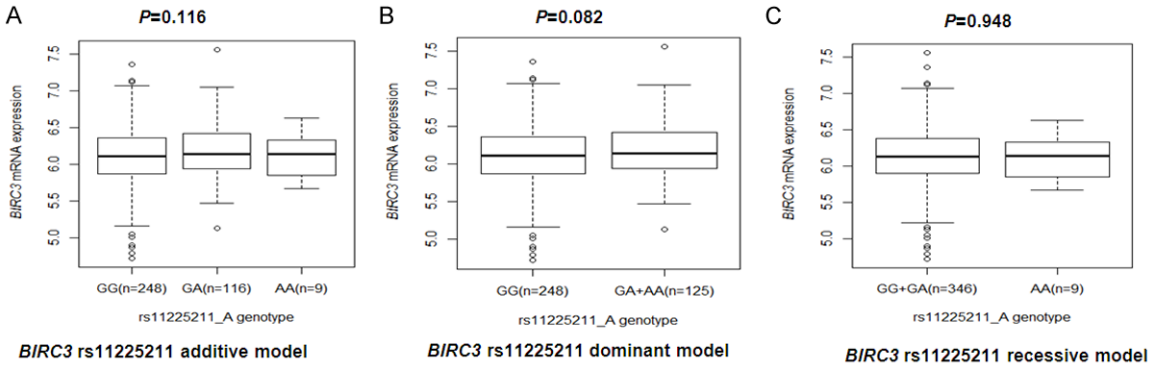


**Figure S4.** eQTL analysis of *BIRC3* rs11225211 genotypes and corresponding gene mRNA expression in 1000 Genomes project. All the data were from the 1000 Genome Project dataset. (A) rs11225211 additive model ($P$ = 0.116); (B) rs11225211 dominant model ($P$ = 0.082); and (C) rs11225211 recessive model ($P$ = 0.948). Abbreviations: eQTL, expression quantitative trait loci.

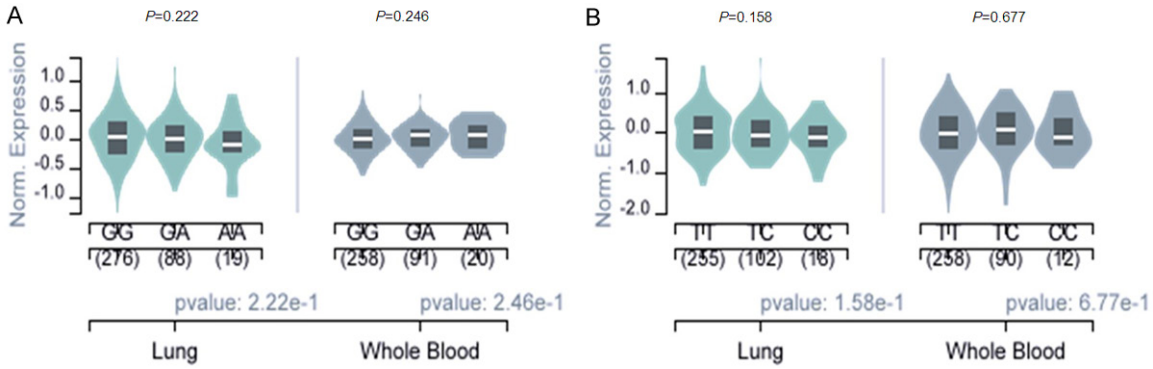


**Figure S5.** eQTL analysis of *BIRC3* rs11225211 and *NRG1* rs4733124 genotypes and corresponding gene mRNA expression in GTEx. All the data were from the GTEx dataset. A. rs11225211 in the lung tissues ($P$ = 0.222) and whole blood ($P$ = 0.246); B. rs4733124 in the lung tissues ($P$ = 0.158) and whole blood ($P$ = 0.677). Abbreviations: eQTL, expression quantitative trait loci.

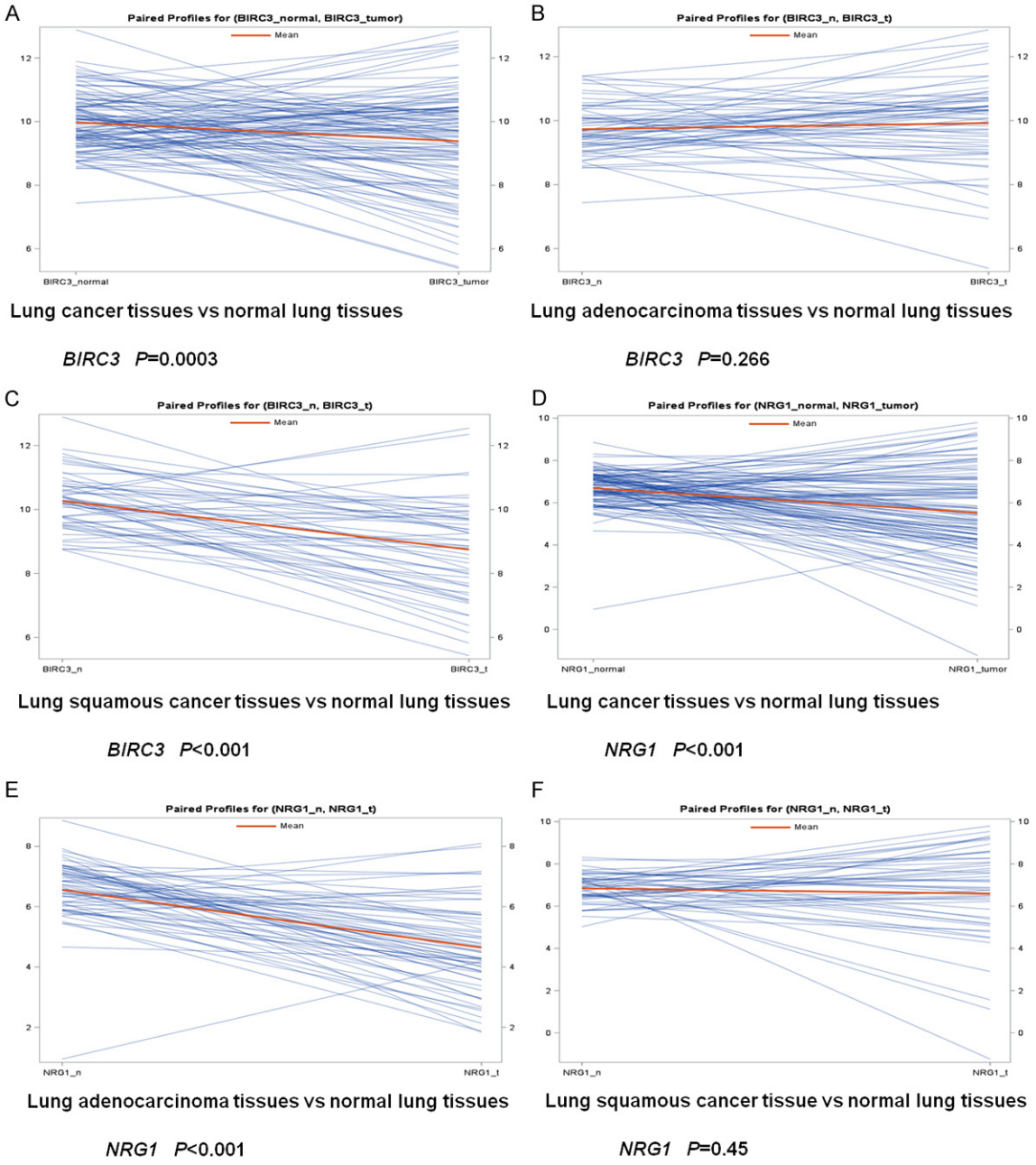


**Figure S6.** Comparison of mRNA expression levels of *BIRC3* and *NRG1* between lung cancer tissue and normal lung tissues in the TCGA dataset. A. The *BIRC3* mRNA expression levels in the lung cancer tissues were significantly lower than that in the normal lung tissues ($P < 0.001$); B. The *BIRC3* mRNA expression levels in the lung adenocarcinoma tissues were significantly lower than that in the normal lung tissues ($P < 0.001$); C. The *BIRC3* mRNA expression levels in the lung squamous tissues were significantly lower than that in the normal lung tissues ($P < 0.001$); D. The *NRG1* mRNA expression levels in the lung cancer tissues were significantly lower than that in the normal lung tissues ($P < 0.001$); E. The *NRG1* mRNA expression levels in the lung adenocarcinoma tissues were significantly lower than that in the normal lung tissues ($P < 0.001$); F. The *NRG1* mRNA expression levels in the lung squamous tissues were significantly lower than that in the normal lung tissues ($P = 0.45$); BIRC3_t = lung cancer tissues; BIRC3_n = adjacent normal lung tissues; NRG1_t = lung cancer tissues; NRG1_n = adjacent normal lung tissues.

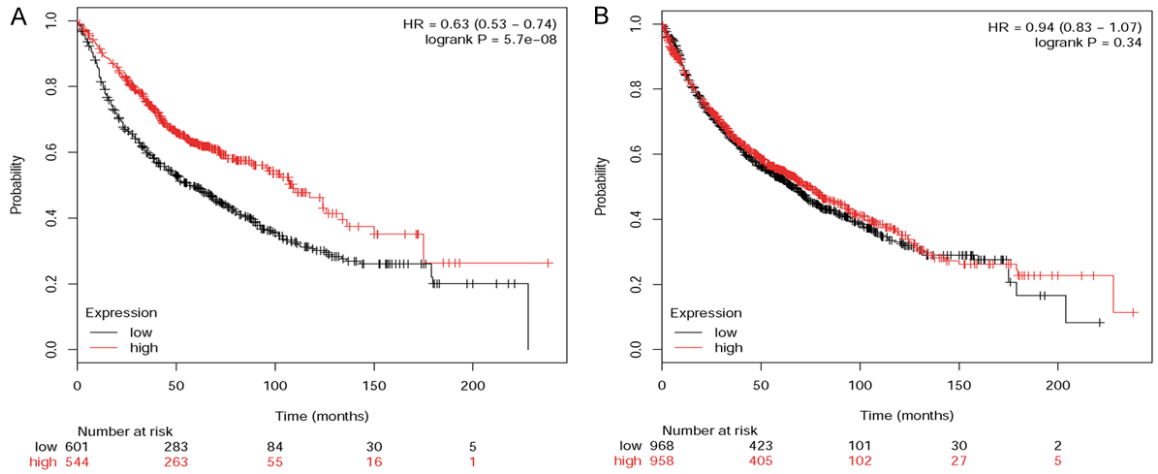


**Figure S7.** Kaplan-Meier analysis for patients with NSCLC by the two genes. Based on online survival analysis software (www.kmplot.com/analysis). A. High *BIRC3* expression were associated with poorer survival of NSCLC; B. *NRG1* expression were not associated with overall survival of NSCLC significantly.

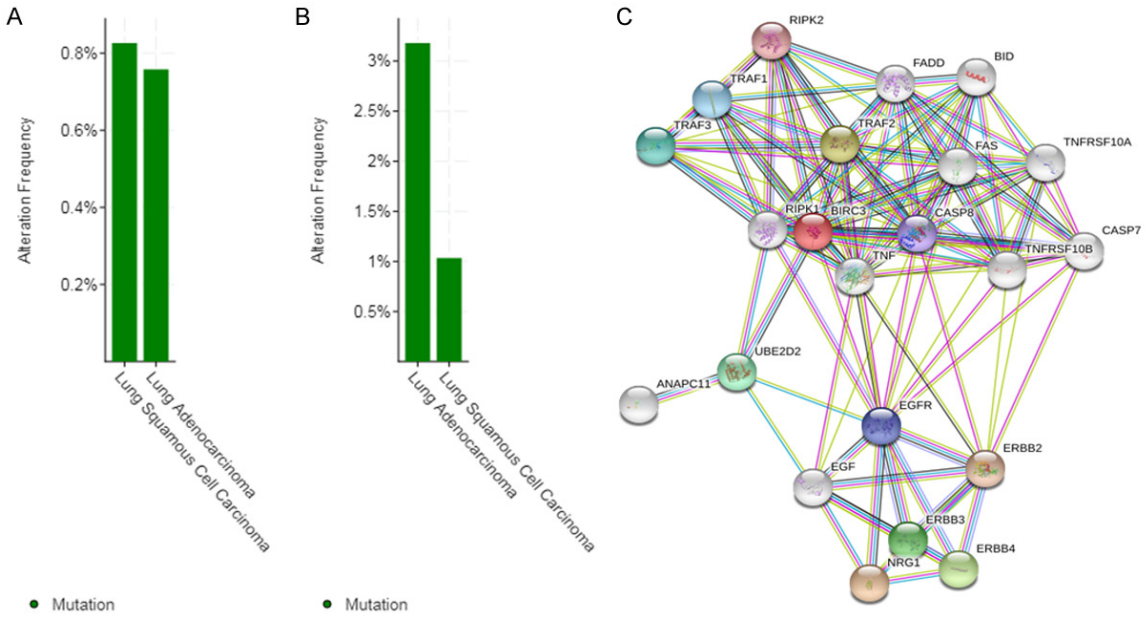


**Figure S8.** Mutation analysis of *BIRC3* and *NRG1* gene in non-small cell lung tumor tissues by using public available data in the database of the cBioportal for Cancer Genomics (http://www.cbioportal.org) and the network including the most frequently altered neighbor genes (https://string-db.org/cgi/network.pl?taskId = 36YKIt07ing0). A. *BIRC3* had low mutation frequency in LUSC and LUAD; B. *NRG1* had a relatively higher mutation frequency in LUSC and LUAD; C. *BIRC3* and *NRG1* were closely associated with altered neighbor genes.