**JBR**

*Research Paper*

# Prognostic assessment of apoptotic gene polymorphisms in non–small cell lung cancer in Chinese

Songyu Cao[a,△], Cheng Wang[a,△], Xinen Huang[b], Juncheng Dai[a], Lingmin Hu[a], Yao Liu[a], Jiaping Chen[a], Hongxia Ma[a], Guangfu Jin[a,d], Zhibin Hu[a,d], Lin Xu[c], Hongbing Shen[a,d, ✉]

[a]*Department of Epidemiology and Biostatistics, Modern Toxicology Laboratory of Ministry of Education, School of Public Health, Nanjing Medical University, Nanjing, Jiangsu 210009, China;*
[b]*Department of Oncology, Nanjing Medical University Affiliated Cancer Hospital of Jiangsu Province; Nanjing, Jiangsu 210009, China;*
[c]*Department of Thoracic Surgery, Nanjing Medical University Affiliated Cancer Hospital of Jiangsu Province, Nanjing, Jiangsu 210009, China;*
[d]*Section of Clinical Epidemiology, Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Cancer Center, Nanjing, Jiangsu 210009, China.*

## Abstract

Apoptosis plays a key role in inhibiting tumor growth, progression and resistance to anti-tumor therapy. We hypothesized that genetic variants in apoptotic genes may affect the prognosis of lung cancer. To test this hypothesis, we selected 38 potentially functional single nucleotide polymorphisms (SNPs) from 12 genes (*BAX, BCL2, BID, CASP3, CASP6, CASP7, CASP8, CASP9, CASP10, FAS, FASLG* and *MCL1*) involved in apoptosis to assess their prognostic significance in lung cancer in a Chinese case cohort with 568 non-small cell lung cancer (NSCLC) patients. Thirty-five SNPs passing quality control underwent association analyses, 11 of which were shown to be significantly associated with NSCLC survival ($P < 0.05$). After Cox stepwise regression analyses, 3 SNPs were independently associated with the outcome of NSCLC (*BID* rs8190315: $P = 0.003$; *CASP9* rs4645981: $P = 0.007$ and *FAS* rs1800682: $P = 0.016$). A favorable survival of NSCLC was significantly associated with the genotypes of BID rs8190315 AG/GG (adjusted HR = 0.65, 95% CI: 0.49-0.88), CASP9 rs4645981 AA (HR = 0.22, 95% CI: 0.07-0.69) and FAS rs1800682 GG (adjusted HR = 0.67, 95% CI: 0.46-0.97). Time-dependent receptor operation curve (ROC) analysis revealed that the area under curve (AUC) at year 5 was significantly increased from 0.762 to 0.819 after adding the risk score of these 3 SNPs to the clinical risk score. The remaining 32 SNPs were not significantly associated with NSCLC prognosis after adjustment for these 3 SNPs. These findings indicate that *BID* rs8190315, *CASP9* rs4645981 and *FAS* rs1800682 polymorphisms in the apoptotic pathway may be involved in the prognosis of NSCLC in the Chinese population.

**Keywords:** apoptosis, polymorphisms, non-small cell lung cancer (NSCLC), prognosis

# INTRODUCTION

Over the past decades, lung cancer has been the most common malignancy and the leading cause of cancer-related deaths around the world. In 2008, there were estimated 1.61 million new cases and 1.38 million deaths from lung cancer globally[1]. In China, lung cancer is the most common cancer and the leading cause of cancer death, the majority (about 80%) of which are non-small cell lung cancer (NSCLC)[2]. The prognosis of lung cancer remains poor, with a 5-year overall survival rate less than 15%. The tumor node metastasis (TNM) staging system is currently used as a guide for prognosis prediction of lung cancer; however, the wide range of survival time has been almost always observed in clinical practice for patients with the same clinical stage, indicating heterogeneity of prognoses among lung cancer patients[3]. Thus, it is critically important to identify biomarkers that may facilitate personalized treatment of lung cancer and ultimately improve the prognoses of patients.

It is well established that apoptosis plays an integral part in tumor growth, progression and resistance to therapy[4]. Defects in apoptosis are implicated in tumor progression and metastasis through maintaining survival of tumor cells, leading to clonal expansion within tumor and further invading surrounding tissues[5]. In addition, radiation and chemotherapy can induce apoptotic cell death in tumors, including lung carcinoma[5]. For example, the induction of FasL and the upregulation of Fas are frequently observed following cisplatin treatment to different tumor cell lines, indicating that anti-cancer therapy kills target cells by the induction of apoptosis[6]. Therefore, multiple components in the apoptotic pathway are evaluated as potential markers to predict lung cancer prognosis[7]. For example, several well designed studies revealed that high bcl-2 expression may be an independent favorable marker for lung cancer survival.

Because of the pivotal role of the apoptotic program in tumor progression and anti-tumor therapy, it is biologically plausible to hypothesize that genetic variants in apoptosis-related genes may be involved in tumor prognosis. There are several pathways of apoptosis like the Bcl-2 protein family, TP53 dependent pathway and TNF-regulated pathway. Some studies investigated the associations between the polymorphisms of these pathway genes and tumor progression, and yielded intriguing findings. For example, *BCL2* -938 C>A variant (rs2279115) is a regulatory polymorphism in the *BCL2* inhibitory promoter, which modulates the promoter activity of *BCL2* by destroying the binding affinity of nuclear protein SP-1 and result in increased expression of Bcl-2[8]. *BCL2* -938 C>A is one of the most extensively investigated polymorphisms in the apoptotic pathway that is associated with multiple tumor prognosis, including chronic lymphocytic leukemia[8], breast cancer[9], oropharyngeal squamous cell carcinoma[10] and acute myeloid leukemia[11]. Sreeja et al. found that *p53* Pro72Pro genotype is an independent risk factor favoring the development of lung carcinoma and that the Arg72Pro genotype is independently associated with a poorer prognosis of lung cancer[12]. Besides, TNFRSF1B rs1061622 (T>G) GG genotype was an independent prognosis predictor of better overall NSCLC survival[13]. Especially, in respect to lung cancer, a series of studies were performed to explore the prognostic significance of the polymorphisms of the apoptotic genes for early stage NSCLC in Korean patients[14-16], which interestingly showed that multiple variants (*FAS* rs1800682, *CASP7* rs2227310, *CASP9* rs4645981, *TNFRSF10B* rs1047266, *TNFRSF1A* rs4149570 and *PPP1R13L* rs1005165) were significantly related to the prognosis of NSCLC in their case cohort. However, evidence is still limited to the demonstration of the effects of apoptotic gene-related polymorphisms on the prognosis of lung cancer. In this study, we systematically selected 38 potentially functional SNPs from 12 genes in the apoptotic pathway, including *BAX, BCL2, BID, CASP3, CASP6, CASP7, CASP8, CASP9, CASP10, FAS, FASLG* and *MCL1* to assess their prognostic significance for lung cancer in a Chinese case cohort of 568 NSCLC patients.

# SUBJECTS AND METHODS

## Study subjects

Patients with histologically confirmed NSCLC were recruited from the Cancer Hospital of Jiangsu Province and the First Affiliated Hospital of Nanjing Medical University, Nanjing, China, from July 2003 to April 2008. All patients recruited in this study were Han Chinese with no prior history of other cancers or previous chemotherapy or radiotherapy. Those patients, who had complete demographic and exposure information such as age, sex and cigarette smoking, were selected in the case cohort. We identified patients who smoked < 1 cigarette per day and < 1 year in their lifetime as nonsmokers; otherwise, they were considered as smokers. All the subjects were prospectively followed up every 3 months by contacting patients or their family members until death or the last follow-up (July 2009). The maximum follow up duration was 72 months and the median follow-up duration was 18.8 months. Each patient was required to

donate 5-mL venous blood for DNA extraction after signing informed consent. This study protocol was approved by the local institutional review board at authors´affiliated institutions.

## SNP selection and genotyping

In this study, we systematically searched the SNPs from 12 genes in the apoptotic pathway, including *BAX, BCL2, BID, CASP3, CASP6, CASP7, CASP8, CASP9, CASP10, FAS, FASLG* and *MCL1*, according to the NCBI dbSNPs (Build 36). The criteria for inclusion of the SNPs were: i) minor allele frequency (MAF) $\geqslant 0.05$ in Chinese population; ii) mapping to 5´ flanking regions ( $\leqslant$ 2,000 bp), 5´ untranslated regions (UTRs), coding regions, or 3´ UTRs of selected genes, as studies have suggested that SNPs of these regions have potential function[17]. Two SNPs (rs4645980 and rs4645981) in the *CASP9* gene without the frequencies of Chinese population in dbSNPs were also included due to polymorphic status reported in a previous study[18]. As a result, 38 SNPs from 12 genes were selected to be genotyped in this study (**Supplementary Table 1** available online).

Genomic DNA was extracted from a leukocyte pellet by proteinase K digestion and was followed by phenol-chloroform extraction and ethanol precipitation. DNA samples were regularly stored at -40℃ . Genotyping was performed by using Illumina Golden Gate platform (Illumina, San Diego, CA, USA) at Berkeley Biotech (Taizhou, Jiangsu, China). All selected SNPs were firstly evaluated for chip design and SNPs with score < 0.50 were excluded. Before genotyping, DNA quantity and quality were assessed by using both fluorometer and agarose gel. Quality control was performed according to the standard operation criteria.

## Statistical analysis

Hardy-Weinberg equilibrium was examined by $\chi^2$ goodness-of-fit test comparing observed genotype frequency to expected one. We calculated the survival time from the date of diagnosis to the date of patient death or the last follow-up. We compared the survival time by using Kaplan-Meier method and log-rank test in different subgroups by demographical variables, clinical features and genotypes. Crude and adjusted hazard ratio (HR) and their 95% confidence intervals (CIs) were further assessed by using univariate and multivariate Cox regression analyses, respectively. Cox regression analyses were also used to calculate trend HRs and 95%CIs and their respective trend *P* values. To define predictors of NSCLC prognosis, Cox stepwise regression model was performed with a significance level of 0.050 for entering and 0.051 for

removing the respective variables. The heterogeneity between subgroups was assessed with the Chi-square-based *Q* test and the heterogeneity was considered significant when *P* < 0.10. Bonferroni correction for α (0.05/number of test) was used to test multiple comparisons in both single variable analysis or multiple cox regression. Risk score analysis was performed to assess the combined effect of SNPs in a manner of dichotomizing the combined genotypes. We analyzed the association of risk score with the outcome of patients by using Cox model and time-dependent receiver-operator characteristic (ROC) curves for censored data and calculated the value of area under curves (AUCs)[19]. The time-dependent performances for different risk scores were evaluated by plotting (t, AUC (t)) for different cutoffs of follow up time. Linkage disequilibrium (LD) analysis was performed by Haploview 4.2 software to measure degree to which alleles at two loci are associated. All statistical analyses were carried out by Statistical Analysis System software (version 9.1.3; SAS Institute, Cary, NC, USA) and R software (version 2.10.1; The R Foundation for Statistical Computing, Vienna, Austria).

# RESULTS

## Patient characteristics and clinical features

As described elsewhere[20], totally 568 patients, who had enough quantity and high quality DNA, were included. During the follow-up period, 311 patients died of NSCLC and 3 died from other causes which were considered as censored data in the final analyses. The overall median survival time (MST) was 24.8 months. In this case cohort, the median age was 60 (range, 25-83) years and 76.4% patients (*n* = 434) were males and 64.6% (*n* = 367) were smokers. For histological types, 353 (62.2%) cases had adenocarcinoma, 184 (32.4%) squamous cell carcinomas and 31 (5.4%) others, including large cell, undifferentiated and mixed-cell carcinomas. The survival time was not statistically different among different strata by age, sex, smoking and histology (*P* > 0.05). As expected, the survival time was significantly shorter for patients with advanced stage and surgical tumor resection could significantly improve the prognosis of NSCLC patients (*P* < 0.001).

## Genotyping results and association with NSCLC survival

The details of genotyping results for 38 SNPs are summarized in **Supplementary Table 1** available online. Those SNPs with genotyping call rate < 90% (*n* = 2) or deviated from Hardy-Weinberg equilibrium (*P* < 0.05:

3 SNPs) were ruled out from further analyses. As a result, 35 SNPs from 12 genes in the apoptotic pathway were finally determined to test their association with NSCLC survival in additive, dominant and recessive models in 568 patients. Furthermore, 11 SNPs located at 5 genes were identified to be significantly associated with NSCLC survival in every genetic model ($P < 0.05$) (*Supplementary Table 1* available online).

### Identification of SNPs independently associated with NSCLC prognosis

The LD analyses between SNPs for 10 genes including two or more SNPs each are shown in *Supplementary Fig. 1* available online. SNPs at the same gene, e.g., *CASP 6, CASP7* and *FAS*, were significantly associated with NSCLC prognosis possible due to LD between SNPs. At the same time, clinical features might confound association results. Therefore, stepwise Cox regression analyses were performed to further determine SNPs that were independently associated with the prognosis of NSCLC together with clinical features. Three SNPs were finally included in multivariate Cox model (BID rs8190315: $P = 0.003$; *CASP9* rs4645981: $P = 0.007$; *FAS* rs1800682: $P = 0.016$) accompanied with the variables of stage, surgical operation and chemo- or radio-therapy (*Supplementary Table 2* available online). The remaining 32 SNPs were not significantly associated with NSCLC prognosis after adjustment for these 3 SNPs ($P >$ 0.05).

Associations between the above 3 independent loci and NSCLC survival are further described in *Table 1* and *Fig. 1A–C*. A favorable survival of NSCLC was significantly associated genotypes of BID rs8190315 AG/GG (adjusted HR = 0.65, 95% CI: 0.49-0.88), *CASP9* rs4645981 AA (HR = 0.22, 95% CI: 0.07-0.69) and FAS rs1800682 GG (adjusted HR = 0.67, 95% CI: 0.46-0.97). When age and sex were forced into the predictive model as fixed factors, the multivariate Cox regression model revealed that these 3 SNPs together with stage, surgical operation and chemo- or radio-therapy could independently predict the prognosis of NSCLC with HRs of 0.64 (95% CI: 0.48-0.86) for rs8190315, 0.21 (95% CI: 0.07-0.65) for rs4645981 and 0.63 (95% CI: 0.43-0.92) for rs1800682 (*Table 2*).

### Combined effects and stratified analyses

We combined the 3 SNPs (BID rs8190315, CASP9 rs4645981 and FAS rs1800682) to assess overall joint effect on NSCLC survival. As shown in Table 3, the more favorable genotypes the patients carried, the longer they survived, suggesting that a locus-dosage effect between combined genotypes and NSCLC survival ($P$ for trend = 1.58 $\times$ $10^{-5}$). After categorization of the combined genotypes into favorable or unfavorable genotypes, patients with favorable genotypes had a MST of 36.1 months, which was significantly longer than that of 22.4 months in those with unfavorable

**Table 1** Polymorphisms in apoptotic genes independently associated with non–small cell lung cancer (NSCLC) survival

| SNP[*] | Patients | Deaths | MST (Months) | Crude HR (95% CI) | Adjusted HR (95% CI)[*] |
|---|---|---|---|---|---|
| *BID* rs8190315 | $n = 568$ | $n = 311$ | | | |
| AA | 455 | 256 | 23.1 | 1.00 | 1.00 |
| AG | 108 | 52 | 32.9 | 0.75 (0.56-1.01) | 0.68 (0.50-0.91) |
| GG | 5 | 3 | 41.5 | 0.63 (0.20-1.95) | 0.42 (0.13-1.32) |
| Per allele[†] | | | | 0.76 (0.58-0.99) | 0.67 (0.51-0.88) |
| AA | 455 | 256 | 23.1 | 1.00 | 1.00 |
| AG + GG | 113 | 55 | 32.9 | 0.74 (0.56-0.99) | 0.65 (0.49-0.88) |
| *CASP9* rs4645981 | $n = 568$ | $n = 311$ | | | |
| GG | 414 | 239 | 23.5 | 1.00 | 1.00 |
| GA | 141 | 69 | 26.4 | 0.88 (0.67-1.15) | 0.84 (0.64-1.10) |
| AA | 13 | 3 | NA | 0.23 (0.07-0.72) | 0.21 (0.07-0.66) |
| Per allele[‡] | | | | 0.75 (0.59-0.95) | 0.72 (0.57-0.90) |
| GG + GA | 555 | 308 | 24.3 | 1.00 | 1.00 |
| AA | 13 | 3 | NA | 0.24 (0.08-0.74) | 0.22 (0.07-0.69) |
| FAS rs1800682[†] | $n = 564$ | $n = 309$ | | | |
| AA | 205 | 113 | 24.0 | 1.00 | 1.00 |
| AG | 282 | 165 | 23.1 | 0.99 (0.78-1.26) | 0.95 (0.74-1.21) |
| GG | 77 | 31 | 41.5 | 0.66 (0.44-0.98) | 0.65 (0.43-0.97) |
| Per allele[†] | | | | 0.87 (0.73-1.03) | 0.85 (0.71-1.01) |
| AA + AG | 487 | 278 | 23.7 | 1.00 | 1.00 |
| GG | 77 | 31 | 41.5 | 0.66 (0.45-0.95) | 0.67 (0.46-0.97) |

SNP: single nucleotide polymorphism; MST: median survival time; HR: hazard ratio; CI: confidence interval. [†]Derived from trend test ($d.f. = 1$) using logistic regression analyses; [‡]Four samples were failed to be genotyped for *FAS* rs1800682.
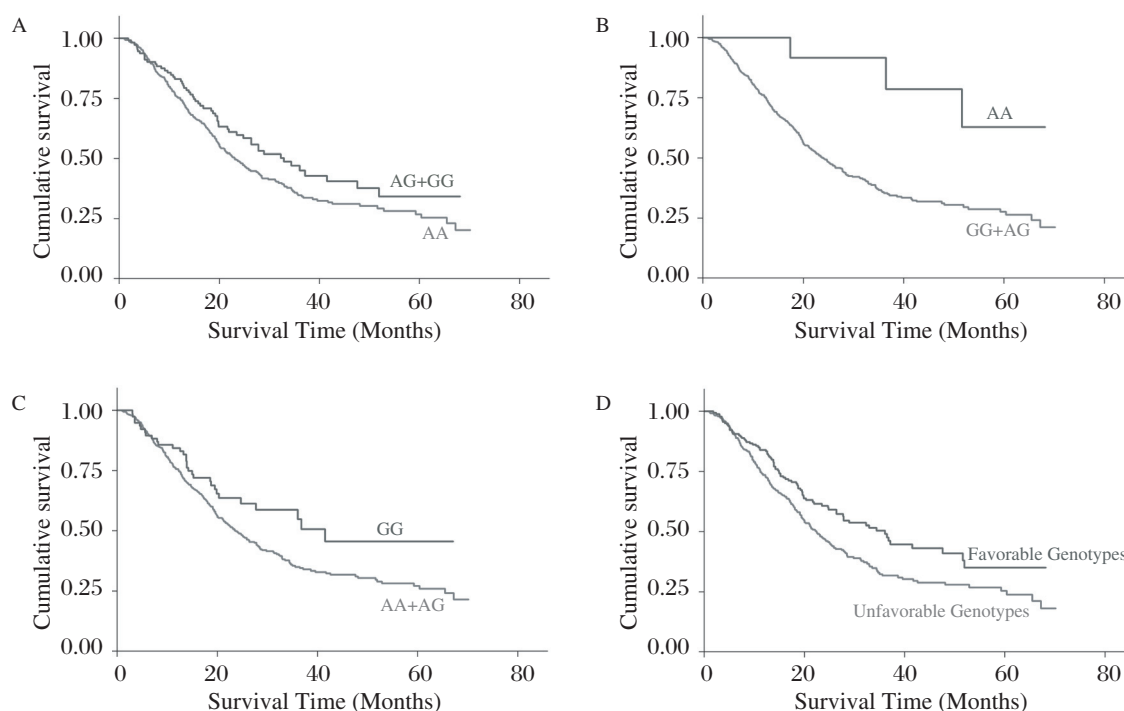
**Fig. 1 Kaplan-Meier plots of survival for non-small cell lung cancer (NSCLC) of more than 3 independent loci.** A: *BID* rs8190315 A>G (log-rank $P$ = 0.045); B: CASP9 rs4645981 G>A (log-rank $P$ = 0.007); C: FAS rs1800682 (log-rank $P$ = 0.026); D: Combined genotypes (log-rank $P$ = 0.002). For combined genotypes, patients were dichotomized by carrying favorable genotypes (*BID* rs8190315AG/GG, *CASP9* rs4645981AA or *FAS* rs1800682GG) or not.

genotypes (log-rank $P$ = 2.39 $\times$ 10$^{-3}$) (**Fig. 1D**). After adjusted for age, sex, smoking, histology, stage, surgical operation and chemo- or radio-therapy, the risk of deaths was significantly decreased by 40% (HR = 0.60, 95% CI: 0.46-0.77) for patients with favorable genotypes, compared with those without favorable genotypes (**Table 3**). Furthermore, stratification analyses revealed that there were no significant differences among the different strata of age, sex, smoking, histology, stage, surgical operation and radio- or chemo-therapy status ($P$ for heterogeneity > 0.10) (**Supplementary Table 3** available online).

The predictive ability of the 3 SNPs was further evaluated by using time-dependent ROC analysis, which was performed by estimating the value of area

under the curve (AUC) according to time-dependent sensitivity and specificity. The risk score of combined genotypes was significantly associated with survival in a regression model ($P$ = 0.003), indicating the capacity to predict the outcome of NSCLC patients. As shown in **Fig. 2**, time-dependent AUCs revealed that the combination of convention factors and combined genotypes performed consistently better than either one alone. The AUC at the survival time of 60 months was 0.762 for the risk score of clinical factors (stage, surgical operation and chemo- or radio-therapy), but 0.819 for the combination of clinical factors and genotype risk score. It seems that AUC decreased with survival time after 54 months when the model included only combined genotypes, and there were maybe

**Table 2 Multivariate Cox regression model for non-small cell lung cancer (NSCLC) outcome prediction**

| Variable | $\beta$ | SE | HR | 95% CI | $P$ |
|---|---|---|---|---|---|
| Age | 0.00 | 0.006 | 1.00 | 0.99-1.02 | 0.467 |
| Sex (female vs male) | -0.09 | 0.139 | 0.91 | 0.70-1.20 | 0.517 |
| Stage ( IV vs III vs II vs I ) | 0.52 | 0.075 | 1.68 | 1.45-1.94 | 3.98 $\times$ 10$^{-12}$ |
| Surgical operation (yes vs no) | -0.61 | 0.143 | 0.55 | 0.41-0.72 | 2.06 $\times$ 10$^{-5}$ |
| Chemo- or radio-therapy (yes vs no) | -0.46 | 0.176 | 0.63 | 0.45-0.89 | 9.30 $\times$ 10$^{-3}$ |
| rs8190315 (AG + GG vs AA) | -0.45 | 0.151 | 0.64 | 0.48-0.86 | 3.01 $\times$ 10$^{-3}$ |
| rs4645981 (AA vs GA + GG) | -1.58 | 0.584 | 0.21 | 0.07-0.65 | 6.93 $\times$ 10$^{-3}$ |
| rs1800682 (GG vs AG + AA) | -0.46 | 0.191 | 0.63 | 0.43-0.92 | 1.57 $\times$ 10$^{-2}$ |

$\beta$: regression coefficient; SE: standard error; HR: hazard ratio; CI: confidence interval.

***Table 3*** **Combined effects of polymorphisms in apoptotic genes on non-small cell lung cancer (NSCLC) survival**

| Genetype | Patients[†] | Deaths | MST (months) | Crude HR (95% CI) | Adjusted HR (95% CI)[‡] |
|---|---|---|---|---|---|
| Combined genotypes[*] | | | | | |
| 0 | 382 | 224 | 22.4 | 1.00 | 1.00 |
| 1 | 162 | 81 | 28.9 | 0.76 (0.59-0.97) | 0.65 (0.50-0.84) |
| 2 | 19 | 4 | NA[§] | 0.23 (0.09-0.63) | 0.24 (0.09-0.65) |
| 3 | 1 | 0 | NA[§] | -- | -- |
| Locus trend | | | | 0.67 (0.54-0.83) | 0.60 (0.48-0.76) |
| Dichotomized group[*] | | | | | |
| 0 | 382 | 224 | 22.4 | 1.00 | 1.00 |
| 1-3 | 182 | 85 | 36.1 | 0.68 (0.53-0.87) | 0.60 (0.46-0.77) |

MST: median survival time; HR: hazard ratio; CI: confidence interval. [*]*BID* rs8190315AG/GG, *CASP9* rs4645981AA and *FAS* rs1800682GG were assumed as favorable genotypes, and "0" - "3" represent the number of favorable genotype; [†]Total 564 patients were successfully genotyped for all three SNPs; [‡]Adjusted for age, gender, smoking, histology, stage and chemo- or radio-therapy; [§]NA (not available) means that MST could not be calculated.

two reasons. First, there were only 33 of 568 patients alive after a 54-month follow up, and this sample size was relatively small and may affect the stability of the AUC model. Second, the prognosis value of combined genotypes would be impaired with the time, while the clinical factors played more important roles (***Fig. 2***).

## DISCUSSION

In this case cohort study, we assessed the significance of genetic variants in apoptotic genes on the prognosis of NSCLC by analyzing 38 potentially functional SNPs from 12 genes in 568 Chinese patients. Three SNPs, including *BID* rs8190315, *CASP9* rs4645981 and *FAS* rs1800682, and their combinations were identified as one of the independent prognostic factors for NSCLC in multivariate Cox regression model. They could significantly improve the ability to predict the prognosis of NSCLC by increasing the AUC from 0.762 to 0.819 after adding the risk score of these 3 SNPs to clinical factors (stage, operation

and chemo- or radio-therapy).

Once cleaved by activated caspase-8, Bid, a member of the BH3-domain-only subgroup of the BCL-2 family, is activated to release cytochrome C from the mitochondria to induce apoptosis[21]. As Bid processing can link the extrinsic and intrinsic cell death pathways through caspase-8 activation and amplify death receptor signaling[22], Bid overexpression has been proposed as a potential therapy signal for the management of lung cancer[22]. In this study, we found that the AG or GG genotype of rs8190315, a nonsynonymous polymorphism located at the codon 56 of *BID* gene resulting in the substitution of Ser to Gly, might predict a favorable prognosis for Chinese NSCLC patients. Interestingly, Lee et al.[16] recently reported a borderline significant association between *BID* rs8190315 and NSCLC survival in early stage patients in Korea. However, the potential function of the variant is not clear and needs to be clarified in further studies.
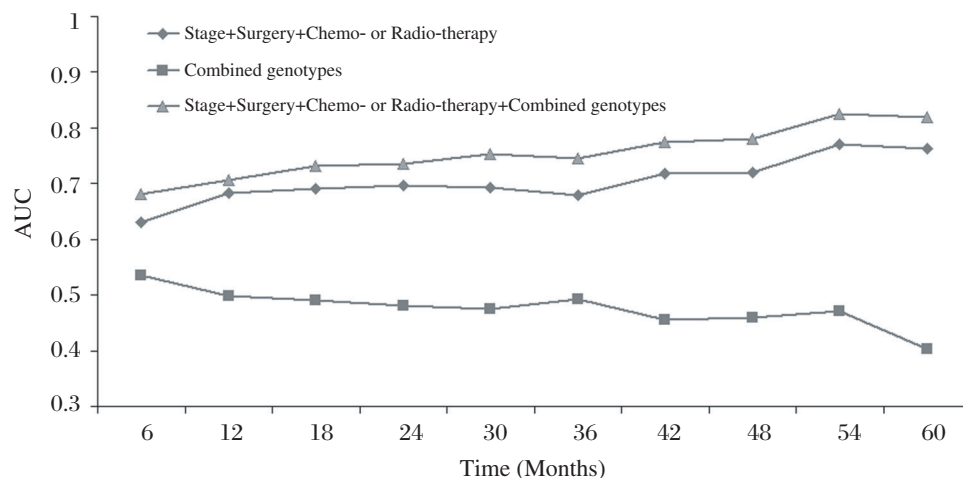


***Fig. 2*** **Time-dependent receiver-operator characteristic (ROC) curve analysis.** The figure shows the time-dependent area under curves (AUCs) of predictive capacity to non-small cell lung cancer (NSCLC) outcome for clinical factors and combined genotypes separately or in aggregate.

Capase-9, the initiator caspase of the intrinsic pathway of apoptosis (i.e., the mitochondrial pathway) and being activated by binding with Apaf-1 and cytochrome C within the apoptosome, cleaves and activates the effectors caspases-3 and -7 to transmit apoptotic signal to downstream components[11]. Luciferase assay showed that the polymorphism rs4645981 (G>A), located at the position of 712 bps upstream of the transcriptional start point of the *CASP9* gene, could influence the promoter activity of *CASP9* in lung cancer cell lines[16], which might result in apoptosis signal change and be involved in tumor initiation and progression. In our NSCLC case cohort, patients with rs4645981 AA genotype were shown to have a better outcome than those with GG or GA genotypes, although the conflicting findings were reported in early stage NSCLC patients in Korea[21]. The low frequency of variant homozygotes may allow for discrepant associations and therapy, and further study with large sample size are warranted to confirm these findings.

The death inducing signaling complex (DISC) is a receptor platform formed by Fas combining with FADD (Fas-associated death domain protein) and caspase-8. DISC is a pivotal trigger of apoptosis, which, once assembled, initiates the induction of programmed cell death[23]. The expression level of Fas has been considered to be an important prognostic factor for the clinical outcome of NSCLC patients[24,25]. The SNP rs1800682 (-670A>G) in the promoter region of the FAS gene was found to be related to the different expression of FAS and to affect the survival of early-stage NSCLC patients in Korean population[14]. In contrast, we found a significant association between rs1800682 variant genotypes and favorable prognosis in this study.

TNM stage is the most consistent prognostic factor in NSCLC patients[26]. However, as patients within the same stage may have very different survivals, better prognostic information is needed[3]. In this study, we evaluated the significance of genetic variants in apoptotic genes on NSCLC prognosis. The effect of individual variant is relatively modest in predicting the outcome of tumor patients. The cumulative effect of multiple variants based on a pathway approach was evaluated in aggregate. The combined genotypes were proven to be more significant on the prognosis of NSCLC patients in our case cohort. Furthermore, we applied time-dependent ROC analysis to reveal the performance of genotype information on the prediction of NSCLC outcome. The combined genotypes of apoptotic genes could significantly improve the predictive ability of clinical parameters (clinical stage and treatment information) for NSCLC outcome. There-

fore, if confirmed in future studies, it may be helpful to optimize treatment programs and predict prognosis for NSCLC patients by applying the genetic information of the apoptotic pathway in clinical practice.

The limitations of this study need to be addressed. First, this study was sought to discover potential effective variants responsible for the prognosis of NSCLC in a Chinese case cohort, by using a pathway-based candidate SNP approach. It was biologically driven, but might result in false positive findings due to multiple comparisons. After adjustment for at least 35 tests in a Bonferroni manner, none of the single SNPs remained significant in association with NSCLC survival in this study ($P > 1.42 \times 10^{-3}$). To better evaluate our findings, we performed an analysis of false positive report probability (FPRP) according to Wacholder et al.[27]. As shown in ***Supplementary Table 4*** available online, the FPRPs were low (< 0.50) for the positive findings of 3 SNPs when prior probability was assigned to be 0.1 or higher, but the association might be less convincing when prior probability was set to 0.01 or lower. Second, potential selection bias might influence the effect estimation of SNPs regarding that some of the patients without clinical and follow-up information or without high quality DNA were excluded from genotyping analysis. Third, we had a moderate sample size in this study, which might afford enough statistical power for general comparisons, but it might be not large enough for the cases that the genotype frequency was low or the patients were divided into subgroups. Therefore, attention should be paid to interpreting these results before they are replicated by other well-designed studies in future.

In summary, the findings in this study indicate that *BID* rs8190315, *CASP9* rs4645981 and *FAS* rs1800682 polymorphisms in the apoptotic pathway may be involved in the prognosis of NSCLC in Chinese population, which provides additional candidate biomarkers for individualized clinical treatment. Our findings remain preliminary and further efforts in population or functional studies are required to confirm these findings and elucidate the mechanism of the association.

## References

[1] Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer* 2010; 127: 2893-917.

[2] Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA Cancer J Clin* 2011; 61: 69-90.

[3] Molina JR, Yang P, Cassivi SD, Schild SE, Adjei AA. Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivorship. *Mayo Clin Proc* 2008; 83:

584-94.

[4] Cotter TG. Apoptosis and cancer: the genesis of a research field. *Nat Rev Cancer* 2009; 9: 501-7.

[5] Vermeulen K, Van Bockstaele DR, Berneman ZN. Apoptosis: mechanisms and relevance in cancer. *Ann Hematol* 2005; 84: 627-39.

[6] Lowe SW, Lin AW. Apoptosis in cancer. *Carcinogenesis* 2000; 21: 485-95.

[7] Singhal S, Vachani A, Antin-Ozerkis D, Kaiser LR, Albelda SM. Prognostic implications of cell cycle, apoptosis, and angiogenesis biomarkers in non-small cell lung cancer: a review. *Clin Cancer Res* 2005; 11: 3974-86.

[8] Nuckel H, Frey UH, Bau M, Sellmann L, Stanelle J, Durig J, et al. Association of a novel regulatory polymorphism (-938C>A) in the BCL2 gene promoter with disease progression and survival in chronic lymphocytic leukemia. *Blood* 2007; 109: 290-7.

[9] Bachmann HS, Otterbach F, Callies R, Nuckel H, Bau M, Schmid KW, et al. The AA genotype of the regulatory BCL2 promoter polymorphism ( 938C>A) is associated with a favorable outcome in lymph node negative invasive breast cancer patients. *Clin Cancer Res* 2007; 13: 5790-7.

[10] Lehnerdt GF, Franz P, Bankfalvi A, Grehl S, Kelava A, Nuckel H, et al. The regulatory BCL2 promoter polymorphism (-938C>A) is associated with relapse and survival of patients with oropharyngeal squamous cell carcinoma. *Ann Oncol* 2009; 20: 1094-9.

[11] Moon JH, Sohn SK, Lee MH, Jang JH, Kim K, Jung CW, et al. BCL2 gene polymorphism could predict the treatment outcomes in acute myeloid leukemia patients. *Leuk Res* 2010; 34: 166-72.

[12] Sreeja L, Syamala V, Raveendran PB, Santhi S, Madhavan J, Ankathil R. p53 Arg72Pro polymorphism predicts survival outcome in lung cancer patients in Indian population. *Cancer Invest* 2008; 26: 41-6.

[13] Guan X, Liao Z, Ma H, Qian J, Liu Z, Yuan X, et al. TNFRSF1B +676 T>G polymorphism predicts survival of non-small cell lung cancer patients treated with chemoradiotherapy. *BMC Cancer* 2011; 11: 447.

[14] Park JY, Lee WK, Jung DK, Choi JE, Park TI, Lee EB, et al. Polymorphisms in the FAS and FASL genes and survival of early stage non-small cell lung cancer. *Clin Cancer Res* 2009; 15: 1794-800.

[15] Yoo SS, Choi JE, Lee WK, Choi YY, Kam S, Kim MJ, et al. Polymorphisms in the CASPASE genes and survival in patients with early-stage non-small-cell lung cancer. *J Clin Oncol* 2009; 27: 5823-9.

[16] Lee EB, Jeon HS, Yoo SS, Choi YY, Kang HG, Cho S, et al. Polymorphisms in apoptosis-related genes and survival of patients with early-stage non-small-cell lung cancer. *Ann Surg Oncol* 2010; 17: 2608-18.

[17] Sauna ZE, Kimchi-Sarfaty C. Understanding the contribution of synonymous mutations to human disease. *Nat Rev Genet* 2011; 12: 683-91.

[18] Park JY, Park JM, Jang JS, Choi JE, Kim KM, Cha SI, et al. Caspase 9 promoter polymorphisms and risk of primary lung cancer. *Hum Mol Genet* 2006; 15: 1963-71.

[19] Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics* 2000; 56: 337-44.

[20] Zhang M, Hu Z, Huang J, Shu Y, Dai J, Jin G, et al. A 3´-untranslated region polymorphism in IGF1 predicts survival of non-small cell lung cancer in a Chinese population. *Clin Cancer Res* 2010; 16: 1236-44.

[21] Li H, Zhu H, Xu CJ, Yuan J. Cleavage of BID by caspase 8 mediates the mitochondrial damage in the Fas pathway of apoptosis. *Cell* 1998; 94: 491-501.

[22] Fukazawa T, Walter B, Owen-Schaub LB. Adenoviral Bid overexpression induces caspase-dependent cleavage of truncated Bid and p53-independent apoptosis in human non-small cell lung cancers. *J Biol Chem* 2003; 278: 25428-34.

[23] Scott FL, Stec B, Pop C, Dobaczewska MK, Lee JJ, Monosov E, et al. The Fas-FADD death domain complex structure unravels signalling by receptor clustering. *Nature* 2009; 457: 1019-22.

[24] Koomagi R, Volm M. Expression of Fas (CD95/APO-1) and Fas ligand in lung cancer, its prognostic and predictive relevance. *Int J Cancer* 1999; 84: 239-43.

[25] Volm M, Koomagi R. Relevance of proliferative and pro-apoptotic factors in non-small-cell lung cancer for patient survival. *Br J Cancer* 2000; 82: 1747-54.

[26] Brundage MD, Davies D, Mackillop WJ. Prognostic factors in non-small cell lung cancer: a decade of progress. *Chest* 2002; 122: 1037-57.

[27] Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. *J Natl Cancer Inst* 2004; 96: 434-42.

**JBR**

# Prognostic assessment of apoptotic gene polymorphisms in non–small cell lung cancer in Chinese

Songyu Cao[a,△], Cheng Wang[a,△], Xinen Huang[b], Juncheng Dai[a], Lingmin Hu[a], Yao Liu[a], Jiaping Chen[a], Hongxia Ma[a], Guangfu Jin[a,d], Zhibin Hu[a,d], Lin Xu[c], Hongbing Shen[a,d, ✉]

[a]*Department of Epidemiology and Biostatistics, Modern Toxicology Laboratory of Ministry of Education, School of Public Health, Nanjing Medical University, Nanjing, Jiangsu 210009, China;*
[b]*Department of Oncology, Nanjing Medical University Affiliated Cancer Hospital of Jiangsu Province; Nanjing, Jiangsu 210009, China;*
[c]*Department of Thoracic Surgery, Nanjing Medical University Affiliated Cancer Hospital of Jiangsu Province, Nanjing, Jiangsu 210009, China;*
[d]*Section of Clinical Epidemiology, Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Cancer Center, Nanjing, Jiangsu 210009, China.*

***Supplementary Table 1*** **The results of single nucleotide polymorphisms (SNPs) selection, genotyping and analyses with non-small cell lung cancer (NSCLC) survival**

| Gene | Location | SNP | Chr | Position | Alleles | Reference MAF[a] | Call rate | MAF in patients | P for HWE[b] | Analysis | Log rank P[a] | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | Additive | Dominant | Recessive |
| *BAX* | 5′ flanking | rs11667351 | 19 | 54147966 | A:C | 0.056 | 100.0 | 0.079 | 1.000 | Include | 0.662 | 0.475 | 0.682 |
| *BAX* | 5′ flanking | rs4645878 | 19 | 54149750 | G:A | 0.058 | 100.0 | 0.077 | 1.000 | Include | 0.725 | 0.553 | 0.682 |
| *BCL2* | 3′ UTR | rs1564483 | 18 | 58945634 | G:A | 0.378 | 99.6 | 0.321 | 0.640 | Include | 0.740 | 0.687 | 0.450 |
| *BCL2* | 5′ flanking | rs2279115 | 18 | 59137817 | C:A | 0.433 | 99.6 | 0.400 | 1.000 | Include | 0.706 | 0.458 | 0.546 |
| *BID* | 3′ UTR | rs2305001 | 22 | 16598210 | A:G | 0.122 | 99.6 | 0.102 | 0.712 | Include | 0.301 | 0.206 | 0.228 |
| *BID* | S56G | rs8190315 | 22 | 16606764 | A:G | 0.125 | 100.0 | 0.104 | 0.828 | Include | 0.128 | 0.045 | 0.471 |
| *CASP10* | S59S | rs3900115 | 2 | 201758922 | A:G | 0.189 | 93.1 | 0.215 | 0.941 | Include | 0.533 | 0.318 | 0.447 |
| *CASP10* | L479I | rs13006529 | 2 | 201790704 | A:T | 0.200 | 99.8 | 0.227 | 0.066 | Include | 0.678 | 0.523 | 0.460 |
| *CASP3* | 3′ UTR | rs1049216 | 4 | 185787083 | G:A | 0.200 | 100.0 | 0.193 | 0.690 | Include | 0.591 | 0.387 | 0.761 |
| *CASP6* | 3′ UTR | rs3182325 | 4 | 110829240 | A:G | 0.525 | 99.6 | 0.486 | 0.313 | Include | 0.038 | 0.764 | 0.019 |
| *CASP6* | 3′ UTR | rs1042891 | 4 | 110829535 | G:A | 0.489 | 99.6 | 0.434 | 0.981 | Include | 0.060 | 0.795 | 0.020 |
| *CASP7* | 5′ flanking | rs12415607 | 10 | 115428194 | C:A | 0.398 | 100.0 | 0.402 | 0.902 | Include | 0.062 | 0.115 | 0.032 |
| *CASP7* | 5′ flanking | rs11196418 | 10 | 115428456 | G:A | 0.111 | 100.0 | 0.085 | 1.000 | Include | 0.579 | 0.298 | 0.803 |
| *CASP7* | 5′ UTR | rs28411397 | 10 | 115429520 | G:C | 0.500 | 100.0 | 0.403 | 0.821 | Include | 0.043 | 0.112 | 0.020 |
| *CASP7* | D4E | rs11593766 | 10 | 115447254 | A:C | 0.067 | 100.0 | 0.139 | 0.107 | Include | 0.037 | 0.011 | 0.455 |
| *CASP7* | D255E | rs2227310 | 10 | 115479142 | C:G | 0.420 | 86.1 | 0.373 | <0.001 | Exclude | -- | -- | -- |
| *CASP7* | 3′ UTR | rs4353229 | 10 | 115479579 | A:G | 0.411 | 99.6 | 0.407 | 1.000 | Include | 0.012 | 0.010 | 0.028 |
| *CASP7* | 3′ UTR | rs10787498 | 10 | 115479640 | A:C | 0.244 | 99.8 | 0.220 | 0.822 | Include | 0.465 | 0.303 | 0.695 |
| *CASP7* | 3′ UTR | rs12247479 | 10 | 115480050 | G:A | 0.089 | 98.4 | 0.120 | 1.000 | Include | 0.442 | 0.202 | 0.732 |
| *CASP7* | 3′ UTR | rs1127687 | 10 | 115480099 | G:A | 0.200 | 99.8 | 0.188 | 0.218 | Include | 0.064 | 0.019 | 0.529 |
| *CASP8* | 5′ flanking | rs6747918 | 2 | 201805820 | G:A | 0.211 | 99.3 | 0.234 | 0.081 | Include | 0.844 | 0.691 | 0.613 |
| *CASP8* | 3′ UTR | rs1045494 | 2 | 201860026 | A:G | 0.211 | 100.0 | 0.227 | 0.874 | Include | 0.966 | 0.818 | 0.849 |
| *CASP8* | 3′ UTR | rs13113 | 2 | 201860407 | T:A | 0.455 | 99.8 | 0.466 | 0.641 | Include | 0.702 | 0.888 | 0.463 |

*Supplementary Table 1* **The results of single nucleotide polymorphisms (SNPs) selection, genotyping and analy–ses with non-small cell lung cancer (NSCLC) survival (continued)**

| Gene | Location | SNP | Chr | Position | Alleles | Reference MAF[a] | Call rate | MAF in patients | P for HWE[b] | Analysis | Log rank P[a] | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | Additive | Dominant | Recessive |
| *CASP8* | 3′ flanking | rs1035140 | 2 | 201860736 | A:T | 0.467 | 99.8 | 0.471 | 0.723 | Include | 0.784 | 0.927 | 0.535 |
| *CASP8* | 3′ flanking | rs1035142 | 2 | 201861323 | C:A | 0.267 | 99.8 | 0.305 | 0.819 | Include | 0.967 | 0.944 | 0.794 |
| *CASP9* | R221Q | rs1052576 | 1 | 15705130 | G:A | 0.378 | 97.5 | 0.362 | 0.589 | Include | 0.239 | 0.240 | 0.126 |
| *CASP9* | V28A | rs1052571 | 1 | 15723200 | G:A | 0.378 | 97.5 | 0.228 | 0.001 | Exclude | -- | -- | -- |
| *CASP9* | 5′ flanking | rs4645981 | 1 | 15724070 | G:A | 0.168 | 100.0 | 0.147 | 0.901 | Include | 0.016 | 0.071 | 0.007 |
| *CASP9* | 5′ flanking | rs4645980 | 1 | 15724263 | C:A | 0.410 | 86.8 | 0.440 | <0.001 | Exclude | -- | -- | -- |
| *CASP9* | 5′ flanking | rs4645978 | 1 | 15724621 | A:G | 0.375 | 99.6 | 0.367 | 0.532 | Include | 0.241 | 0.172 | 0.170 |
| *FAS* | 5′ flanking | rs2234767 | 10 | 90739236 | G:A | 0.341 | 99.5 | 0.288 | 0.522 | Include | 0.037 | 0.069 | 0.024 |
| *FAS* | 5′ flanking | rs1800682 | 10 | 90739943 | A:G | 0.378 | 99.3 | 0.387 | 0.238 | Include | 0.083 | 0.472 | 0.026 |
| *FAS* | 3′ UTR | rs1468063 | 10 | 90765271 | G:A | 0.367 | 100.0 | 0.374 | 0.076 | Include | 0.237 | 0.961 | 0.102 |
| *FASLG* | 5′ flanking | rs763110 | 1 | 170894121 | G:A | 0.300 | 100.0 | 0.235 | 0.952 | Include | 0.378 | 0.816 | 0.169 |
| *MCL1* | 3′ UTR | rs878471 | 1 | 148814371 | G:A | 0.333 | 99.5 | 0.405 | 0.836 | Include | 0.523 | 0.357 | 0.352 |
| *MCL1* | 5′ flanking | rs3738484 | 1 | 148818954 | C:A | 0.250 | 99.3 | 0.402 | 1.000 | Include | 0.524 | 0.285 | 0.473 |
| *MCL1* | 5′ flanking | rs3738485 | 1 | 148819016 | C:G | 0.307 | 98.4 | 0.400 | 0.963 | Include | 0.602 | 0.364 | 0.477 |
| *MCL1* | 5′ flanking | rs9803935 | 1 | 148819246 | A:C | 0.322 | 99.5 | 0.400 | 0.892 | Include | 0.579 | 0.300 | 0.637 |

*MAF* minor allele frequency

[a]Minor allele frequencies in Chinese population according to NCBI dbSNPs;

[b]Hardy-Weinberg equilibrium (HWE) was tested by goodness-of-fit $\chi^2$ test;

[c]Log rank tests were performed in additive, dominant and recessive models, respectively.

*Supplementary Table 2* **Results of stepwise Cox regression analysis identifying predictive factors for non-small cell lung cancer (NSCLC) prognosis**

| Step | Variables | β | SE | HR | 95% CI | P |
|---|---|---|---|---|---|---|
| 1 | Stage ( Ⅳ *vs* Ⅲ *vs* Ⅱ *vs* Ⅰ ) | 0.51 | 0.074 | 1.67 | 1.44-1.93 | $6.80\times10^{-12}$ |
| 2 | Surgical Operation (yes *vs* no) | -0.61 | 0.143 | 0.54 | 0.41-0.72 | $1.92\times10^{-5}$ |
| 3 | rs8190315 (AG+GG *vs* AA) | -0.45 | 0.150 | 0.64 | 0.48-0.86 | $2.85\times10^{-3}$ |
| 4 | rs4645981 (AA *vs* AG+GG) | -1.59 | 0.583 | 0.21 | 0.07-0.64 | $6.52\times10^{-3}$ |
| 5 | Chemo- or Radio-therapy (yes *vs* no) | -0.46 | 0.176 | 0.63 | 0.45-0.89 | $8.40\times10^{-3}$ |
| 6 | rs1800682 (GG *vs* AG+AA) | -0.46 | 0.190 | 0.63 | 0.44-0.92 | $1.58\times10^{-2}$ |

*β* regression coefficient, *SE* standard error, *HR* hazard ratio, *CI* confidence interval.

*Supplementary Table 3* **Stratified analyses for the associations of combined genotypes with non-small cell lung cancer (NSCLC) survivals**

| Variable | Favorable genotypes* | | Adjusted HR (95% CI)[†] | P for heterogeneity |
|---|---|---|---|---|
| | No (Deaths/Patients) | Yes (Deaths/Patients) | | |
| Age (years) | | | | 0.308 |
| ≤ 60 | 114/191 | 41/93 | 0.53 (0.37-0.76) | |
| > 60 | 110/191 | 44/89 | 0.69 (0.48-0.98) | |
| Sex | | | | 0.412 |
| Male | 176/294 | 64/138 | 0.63 (0.47-0.84) | |
| Female | 48/88 | 21/44 | 0.49 (0.290.83) | |
| Smoking | | | | 0.903 |
| Never | 72/129 | 37/69 | 0.62 (0.41-0.94) | |
| Ever | 152/253 | 48/113 | 0.60 (0.44-0.84) | |
| Histology | | | | 0.209 |
| Adenocarcinoma | 138/229 | 53/122 | 0.85 (0.55-1.33) | |
| Squamous cell | 70/129 | 28/53 | 0.52 (0.38-0.72) | |
| Others[‡] | 16/24 | 4/7 | 0.66 (0.20-2.17) | |
| Stage | | | | 0.823 |
| Ⅰ - Ⅱ | 58/146 | 17/67 | 0.53 (0.30-0.93) | |
| Ⅲ - Ⅳ | 166/236 | 68/115 | 0.57 (0.42-0.76) | |

***Supplementary Table 3*** **Stratified analyses for the associations of combined genotypes with non-small cell lung cancer (NSCLC) survivals (continued)**

| Variable | Favorable genotypes[*] | | Adjusted HR (95% CI)[†] | P for heterogeneity |
|---|---|---|---|---|
| | No (Deaths/Patients) | Yes (Deaths/Patients) | | |
| Surgical Operation | | | | 0.476 |
|   Never | 104/130 | 50/69 | 0.64 (0.45-0.91) | |
|   Ever | 120/252 | 35/113 | 0.53 (0.36-0.77) | |
| Chemo- or Radio-therapy | | | | 0.907 |
|   Never | 35/73 | 11/32 | 0.64 (0.30-1.37) | |
|   Ever | 189/309 | 74/150 | 0.61 (0.46-0.80) | |

*HR* hazard ratio, *CI* confidence interval

[*]*BID* rs8190315AG/GG, *CASP9* rs4645981AA and *FAS* rs1800682GG are presumed as favorable genotypes and total 564 patients were successfully genotyped for all 3 SNPs;

[†]Adjusted for age, gender, smoking, histology, stage, surgical operation and chemo- or radio-therapy;

[‡]Other carcinomas include large cell, undifferentiated and mixed-cell carcinomas.

***Supplementary Table 4*** **The false positive report probability (FPRP) for significant associations of apoptotic gene variants and non-small cell lung cancer (NSCLC) survival**

| Comparison | HR(95%CI)[*] | P[†] | HR[‡] | Prior probability | | | |
|---|---|---|---|---|---|---|---|
| | | | | 0.25 | 0.10 | 0.01 | 0.001 |
| *BID* rs8190315 | 0.65 (0.49-0.88) | 0.0046 | 0.65 | 0.031 | 0.087 | 0.513 | 0.914 |
| *CASP9* rs4645981 | 0.22 (0.07-0.69) | 0.0092 | 0.22 | 0.054 | 0.145 | 0.651 | 0.950 |
| *FAS* rs1800682 | 0.67 (0.46-0.97) | 0.0358 | 0.67 | 0.169 | 0.0379 | 0.870 | 0.985 |

[*]Observed adjusted hazard ratios (HRs) and 95% confidence intervals (CIs);

[†]P values corresponding to adjusted HRs and 95% CIs;

[‡]HRs used for FPRP estimation were derived from the observed HRs.
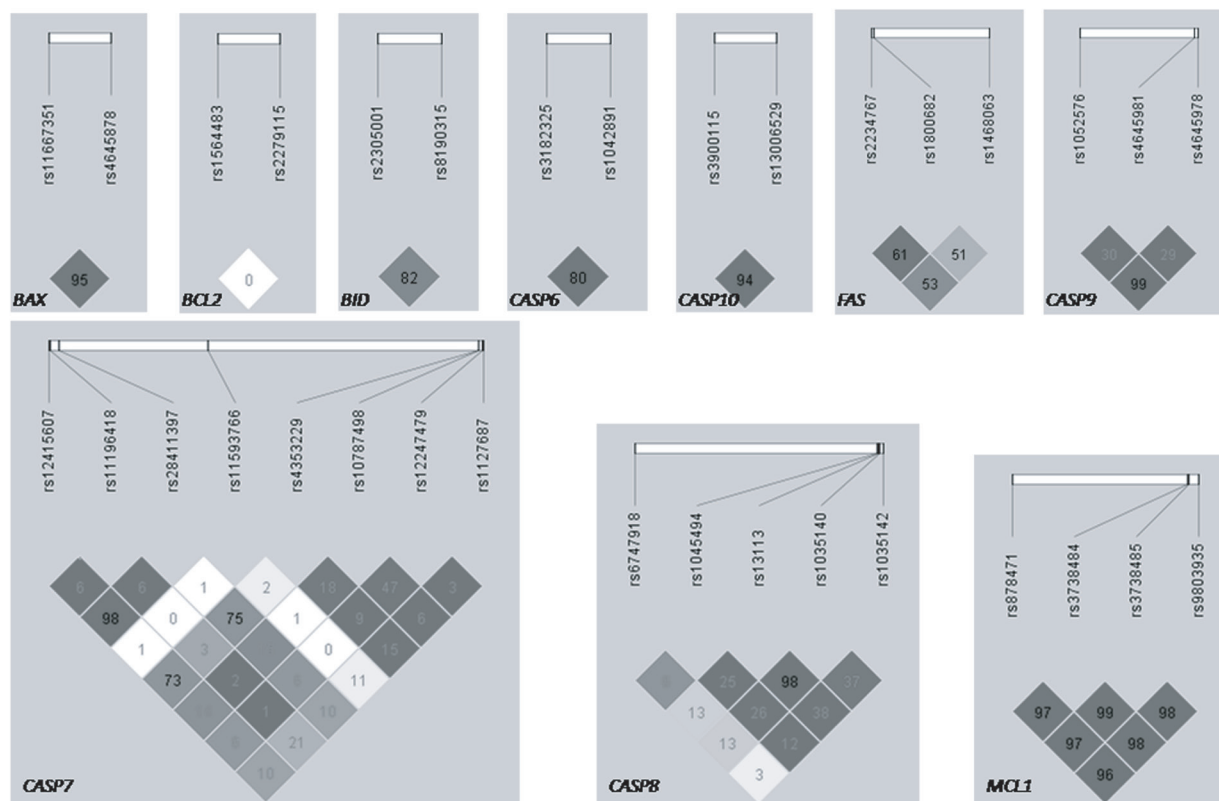


***Supplementary Fig. 1*** Linkage disequilibrium (LD) plots for genes with two or more SNPs included in this study based on geno–type data in 568 patients. The LD value (r2) between two SNPs was shown in the diamond.