Predictors of Survival in Never-Smokers with Non–Small Cell Lung Cancer: A Large-Scale, Two-Phase Genetic Study

Xia Pu, Yuanqing Ye, Margaret R. Spitz, Liang Wang, Jian Gu, Scott M. Lippman, Margaret R. Spitz, North American Lung Cancer Survival Study Group, Lippman, and Xifeng Wu

Abstract

Purpose: Lung cancer in never-smokers (LCINS) is increasingly recognized as a distinct disease from that in ever-smokers owing to substantial differences in etiology, clinical characteristics, and prognosis. Therefore, we aimed to identify prognostic markers specific for LCINS.

Experimental Design: First, 11,930 single-nucleotide polymorphisms (SNP) in 904 inflammation-related genes were genotyped, and their associations with overall survival in 411 patients with LCINS at MD Anderson Cancer Center were analyzed. Next, validation of the top 27 SNPs in 311 patients with LCINS at Mayo Clinic was conducted.

Results: Three SNPs (IL17RA:rs879576, BMPRA:rs698141, and STY:rs290229) were validated (P < 0.05), and two SNPs (CD74:rs1056400 and CD38:rs10805347) reached borderline significance (P = 0.08) in the Mayo Clinic population. We validated a survival-tree created in the MD Anderson population exploring gene–gene interactions in the Mayo Clinic population. This survival-tree stratified patients into subsets with significantly different risks of death: patients with the rs1056400_GG/rs698141_GA+AA genotype had significantly higher risk of death in both MD Anderson (HR:2.32, 95%CI: 1.58–3.41) and Mayo (HR:1.97, 95%CI: 1.11–3.50) populations compared with those with the rs1056400_GG/rs698141_GG or rs1056400_GA + AA genotype. We evaluated these five SNPs in 996 ever-smokers from MD Anderson and found no significant associations.

Conclusions: Our study provides strong evidence that inflammation-related genetic variations can affect clinical outcomes in LCINS, which may lead to significant biologic insight into these outcomes. Clin Cancer Res; 18(21): 5983–91. ©2012 AACR.
Translational Relevance

Lung cancer in never-smokers (LCINS) is increasingly recognized as a distinct disease from that in ever-smokers owing to substantial differences in etiology, clinical characteristics, and prognosis. Identification of specific prognostic and predictive markers for LCINS beyond the general markers for lung cancer is warranted. Inflammation plays an important role in cancer initiation and progression, as well as influence patients’ clinical outcomes. In this study, we conducted a large-scale two-phase study to identify and validate inflammation-related genetic variations as prognostic markers specific for LCINS using two independent lung cancer patient cohorts. Our study provides strong evidence that inflammation-related genetic variations can affect clinical outcomes in LCINS, which may lead to significant biologic insight into these outcomes and guide personalized treatment and follow-up care regimens for LCINS.

resulting in carcinogenesis and formation of invasive and/or metastatic phenotypes (9–14). Inflammatory cells and related signaling molecules could also be used by tumor to facilitate its progression and metastasis by generating a favorable microenvironment as well as promoting genetic instability and angiogenesis (10). Moreover, poorer survival was found in patients with cancer with elevated inflammatory markers (15).

The lung is a frequent site of infection and occasional site of chronic inflammation owing to environmental exposures. Furthermore, accumulating evidence shows that inflammation is associated with clinical outcomes of various cancers, including lung cancer (16–19). Previous studies have explored associations between selected inflammation gene polymorphisms and lung cancer prognosis, with inconsistent results potentially due to small sample sizes (20, 21). In the present study, we genotyped a comprehensive panel of inflammation-related single-nucleotide polymorphisms (SNP) in a large number of never-smokers with lung cancer at The University of Texas MD Anderson Cancer Center with comprehensive epidemiological, clinical, and follow-up information to determine the effect of inherited genetic variations on clinical outcome of this cancer. To validate our findings, we studied these SNPs in an independent patient population from Mayo Clinic. To our knowledge, this is the first study to systematically evaluate the associations of genetic variations in major inflammation-related genes with overall survival of LCINS that uses a 2-stage study design.

Materials and Methods

Study population and data collection

Written informed consent to participate in the study was obtained from all participants. The study protocol was approved by the Institutional Review Boards at MD Anderson (Houston, TX) and Mayo Clinic (Rochester, MN). Patients at MD Anderson had newly diagnosed, histologically confirmed non–small cell lung cancer (NSCLC) enrolled in an ongoing lung cancer study initiated in 1995. Demographic information, smoking status and environmental exposures were collected from the patients by trained staff using a structured questionnaire in-person interviews. Patients’ clinical and follow-up data were abstracted from medical records. The primary analysis focused on never-smokers; a group of ever-smokers with lung cancer was also studied to compare the effect of SNPs by smoking status. Patients at Mayo Clinic had newly diagnosed, histopathologically confirmed primary NSCLC who are never smokers. A structured questionnaire was used to collect detailed epidemiologic data on the patients. These patients participated in a long-term follow-up study from 1997 to 2008 described in detail previously (22, 23). All analyses were restricted to Caucasian patients to minimize effects of population structure.

SNP selection

Compilation of the genes involved in the inflammatory response was conducted on the basis of a published panel of inflammation-associated genes (24) and a database of diabetes and inflammation genes [T1DBase (http://www.t1dbase.org); University of Cambridge, Cambridge, UK]. Tagging SNPs for candidate genes based on data from an European population were identified using data from the International HapMap Project, based on National Center for Biotechnology Information B36 assembly and dbSNP b126. For each gene, sequences 10 kb before the transcription start site and 10 kb after the transcription end site were included in the tag SNP selection using the Tagger pairwise method (Broad Institute, Cambridge, MA) with an $r^2$ threshold of 0.8 and minor allele frequency of at least 0.05 (25). The compiled SNP list was sent to Illumina (San Diego, CA) for designability analysis using their array design tool, which uses a proprietary algorithm to predict successful design of the genotyping assay for each SNP. Each SNP is scored from 0 to 1 with higher scores having a higher likelihood of assay success. A score of 0.4 was considered as the cut-off value. Only SNPs that exceeded the threshold score (>0.4) were considered designable. In total, 11,930 SNPs (Supplementary Table S1) were included for construction of an Infinium II iSelect Custom Genotyping BeadChip (Illumina).

Genotyping and quality control

Genomic DNA was extracted from peripheral blood samples obtained from the study patients using a QIAamp DNA Mini Kit (QIAGEN, Valencia, CA). iSelect Custom Genotyping BeadChips were processed according to the Infinium II assay protocol (Illumina). Standard quality control procedures were conducted including only SNPs with genotyping data in more than 95% of all samples and samples with genotyping data for more than 95% of all SNPs. SNPs selected for validation were genotyped at Mayo Clinic’s Genotyping Core Facility using a Fluidigm Dynamic Array (Fluidigm, South San Francisco, CA) and a HumanHap317 BeadChip (Illumina) according to a standard protocol and using quality control measures.
**Statistical analysis**

Overall survival, defined as the duration from diagnosis to death of any cause or the last follow-up, was the primary endpoint of our analysis. Never-smokers were defined as patients who had smoked fewer than 100 cigarettes over their lifetimes, ever-smokers are those who had smoked more than 100 cigarettes over their lifetimes. The effect of each SNP on survival was assessed using multivariate Cox proportional hazards regression analysis adjusted for age, sex, clinical stage, and treatment regimen. Kaplan–Meier survival curves and log-rank tests were used to evaluate the effect of variants on the time to death. Multiple testing adjustment was conducted to calculate $P$-value after adjusting for 11,930 tests in the discovery phase setting a false discovery rate (FDR) of 10% using the "q-value" package in R (26). Between-study heterogeneity for each SNP was examined using the $\chi^2$-based Q test (27). A fixed effects model was used to estimate the summarized effect on survival with absence of heterogeneity ($P > 0.05$). Survival tree analysis was conducted to identify higher-order gene–gene interactions affecting survival using the STREE software (http://c2s2.yale.edu/software/stree/). STREE uses a log-rank statistical method to select optimal and subsequent splits of datasets.

**Results**

**Patient characteristics**

In the MD Anderson study, we identified 411 never-smokers with NSCLC (Table 1). Sixty-seven percent of them were women, and adenocarcinoma was the most common histology (77%). The mean age at diagnosis was 61.5 years. The median survival time (MST) was 23.2 months, and the median follow-up time (MFT) was 54.2 months. Most of the cases (77%) were diagnosed at a late stage (stage III/IV). Fifty-three percent of the patients received only chemotherapy, 33% underwent surgery, and 24% received radiotherapy. At the time of the current study, 276 (67%) of the patients had died. In the Mayo Clinic study, 311 never-smokers with NSCLC were identified and included as the validation population (Table 1). The mean age was 61.7 years, with the majority being female (73%). Sixty-one percent of the patients had late-stage disease at diagnosis. Fifty-nine percent of the patients received chemotherapy, 53% underwent surgery, and 25% received radiotherapy. At the time of this study, 182 (59%) of the patients had died. Because of the greater proportions of patients with early-stage disease and who had undergone surgery in the Mayo Clinic population than in the MD Anderson population, the MST (44.6 months) and MFT (73.6 months) were longer.

---

**Table 1. Characteristics of the never-smokers with lung cancer**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>MD Anderson (discovery)</th>
<th>Mayo Clinic (validation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MST, mo</td>
<td>23.2</td>
<td>44.6</td>
</tr>
<tr>
<td>MFT, mo</td>
<td>54.2</td>
<td>73.6</td>
</tr>
<tr>
<td>Mean age, y (SD)</td>
<td>61.5 (13.0)</td>
<td>61.7 (13.1)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>135 (33)</td>
<td>84 (27)</td>
</tr>
<tr>
<td>Female</td>
<td>276 (67)</td>
<td>227 (73)</td>
</tr>
<tr>
<td>Stage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>93 (23)</td>
<td>105 (34)</td>
</tr>
<tr>
<td>II</td>
<td>15 (4)</td>
<td>15 (5)</td>
</tr>
<tr>
<td>III</td>
<td>91 (22)</td>
<td>90 (29)</td>
</tr>
<tr>
<td>IV</td>
<td>212 (52)</td>
<td>101 (32)</td>
</tr>
<tr>
<td>Histology</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>316 (77)</td>
<td>213 (68)</td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>28 (7)</td>
<td>14 (5)</td>
</tr>
<tr>
<td>Non–small cell carcinoma</td>
<td>41 (10)</td>
<td>18 (6)</td>
</tr>
<tr>
<td>Bronchoalveolar carcinoma</td>
<td>20 (5)</td>
<td>11 (4)</td>
</tr>
<tr>
<td>Other</td>
<td>6 (1)</td>
<td>55 (18)</td>
</tr>
<tr>
<td>Treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surgery</td>
<td>135 (33)</td>
<td>165 (53)</td>
</tr>
<tr>
<td>Radiotherapy</td>
<td>100 (24)</td>
<td>77 (25)</td>
</tr>
<tr>
<td>Chemotherapy</td>
<td>218 (53)</td>
<td>182 (59)</td>
</tr>
<tr>
<td>Concurrent chemoradiation</td>
<td>38 (9)</td>
<td>36 (12)</td>
</tr>
<tr>
<td>Vital status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dead</td>
<td>276 (67)</td>
<td>182 (59)</td>
</tr>
<tr>
<td>Alive</td>
<td>135 (33)</td>
<td>129 (41)</td>
</tr>
<tr>
<td>Total</td>
<td>411</td>
<td>311</td>
</tr>
</tbody>
</table>
months) were longer in the former population than in the latter.

Main effect of individual SNP on survival in the discovery, replication, and combined analysis

In the discovery phase, after carrying out quality control measures, 11,689 SNPs were included in our analysis. Of these SNPs, 1,538 were significantly associated with overall survival \((P < 0.05\), with 14 of these variants being significant at the \(P < 10^{-4}\) level. These associations remained significant after removal of patients who reported passive smoking exposures from the analysis. We selected 27 top SNPs for validation in the Mayo Clinic population. The entire 27 SNPs remained significant after multiple comparisons adjustment in the MD Anderson discovery population \((q < 0.1\). Eighteen SNPs had a consistent direction of the effect \(HR \) same direction) in both populations (Table 2). Of these 18, 3 SNPs [interleukin 17 receptor A \((IL17RA)\);rs879576, bone morphogenetic protein 8A \((BMP8A)\);rs698141, and spleen tyrosine kinase \((SYK)\);rs290229] in the Mayo population were significant \((P < 0.05\) with an additional 2 \((CD74;rs1056400 and CD38;rs10805347)\) reaching borderline significance \((P < 0.1\)

The most significant SNP was rs879576, a synonymous variant in the last exon of the proinflammatory cytokine \(IL17RA\). Rs879576 was associated with a significantly decreased risk of death in the discovery phase \(HR\): 0.57; 95% confidence interval \((CI)\) 0.41–0.78; \(P = 5.49 \times 10^{-4}\), validation phase \(HR\): 0.65; 95%CI: 0.44–0.94; \(P = 0.023\) and combined population \(HR\): 0.60; 95%CI: 0.47–0.77; \(P = 4.13 \times 10^{-5}\); Table 2). This decreased risk of dying resulted in enhanced survival duration. Compared with patients with variant genotypes, a prolonged MST was observed in patients with the common homozygous genotype in both discovery \((31 \pm 20\) months, \(P = 0.066\), log-rank test) and validation \((46 \pm 34\) months, \(P = 0.069\), log-rank test) phases. (Fig. 1A and B)

Rs290229 is an intronic SNP in \(SYK\), a gene that encodes for a nonreceptor type \(Tyr\) protein kinase. This SNP was associated with a significantly increased risk of death in both the MD Anderson \(HR\): 1.58; 95%CI: 1.23–2.03; \(P = 3.03 \times 10^{-4}\), Mayo Clinic \(HR\): 1.43; 95%CI: 1.01–2.02; \(P = 0.046\)\], and combined \(HR\): 1.53; 95%CI: 1.25–1.87; \(P = 4.15 \times 10^{-5}\)\] populations. Although not significant, both study populations had the same trend of decreased MST \(HR\) \(1\)C and D)\]

Rs698141 is located in intron of \(BMP8A\), a gene involved in cytokine signaling transduction. Patients who had at least one variant allele had a nearly 2-fold increase in risk of death in both the MD Anderson \(HR\): 1.89; 95%CI: 1.33–2.68; \(P = 4.04 \times 10^{-4}\) and Mayo Clinic \(HR\): 1.73; 95%CI: 1.03–2.91; \(P = 0.038\)\] and combined \(HR\): 1.84; 95%CI: 1.37–2.46; \(P = 4.29 \times 10^{-5}\)\] populations (Table 2). The MST was 23 months in patients with the common homozygous genotype and 16 months in patients with the heterozygous or homozygous variant genotypes in the MD Anderson population \(P = 9.1 \times 10^{-4}\).

Table 2. SNPs with the same trend in both the MD Anderson and Mayo Clinic populations

<table>
<thead>
<tr>
<th>Position</th>
<th>Gene</th>
<th>SNP</th>
<th>Model</th>
<th>MD Anderson (discovery)</th>
<th>Mayo Clinic (validation)</th>
<th>Combined analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chr22:15969246</td>
<td>IL17RA</td>
<td>rs879576DME</td>
<td>OME</td>
<td>0.57 (0.41–0.76)</td>
<td>0.65 (0.44–0.94)</td>
<td>0.60 (0.47–0.77)</td>
</tr>
<tr>
<td>Chr9:26274293</td>
<td>SYK</td>
<td>rs290229DME</td>
<td>OME</td>
<td>1.58 (1.23–2.03)</td>
<td>1.43 (1.01–2.02)</td>
<td>1.53 (1.25–1.87)</td>
</tr>
<tr>
<td>Chr1:39373384</td>
<td>BMP8A</td>
<td>rs698141DME</td>
<td>OME</td>
<td>1.89 (1.33–2.68)</td>
<td>1.73 (1.03–2.91)</td>
<td>1.84 (1.37–2.46)</td>
</tr>
<tr>
<td>Chr21:9237964</td>
<td>IFNA8</td>
<td>rs4978115REC</td>
<td>OME</td>
<td>1.77 (1.35–2.36)</td>
<td>1.84 (1.37–2.46)</td>
<td>1.84 (1.37–2.46)</td>
</tr>
</tbody>
</table>

Abbreviations: ADD, additive model; DOM, dominant model; p-het, \(P\) for heterogeneity test; REC, recessive model.

aAdjusted according to age, clinical stage, and treatment regimen.

bCombined (meta) analysis based on fixed effects model. Boldface. \(P < 0.1\).
We also observed a similar longer MST (24 months) in the Mayo population ($P = 0.044$, log-rank test; Fig. 1E). CD74:rs1056400 (3’-untranslated region) and CD38: rs10805347 (intronic) were significantly associated with an increased risk of death in the MD Anderson population but were borderline significant in the Mayo Clinic population (Table 2). Although not statistically significant, the trend of differing survival times by genotype was observed (Figs. 1G–J).

Stratified analysis of individual SNPs on survival

The majority of never-smokers with lung cancer are female with adenocarcinoma. We conducted a subgroup analysis of survival in female patients with adenocarcinoma. The results were similar to those of the overall analysis of all study patients (Supplementary Table S2). We further conducted a stratified analysis of the 5 top SNPs according to disease stage. Specifically, we combined the MD Anderson and Mayo Clinic patients and stratified them according to early-stage (I and II) and late-stage (III and IV) lung cancer. The results showed that all 5 SNPs were significantly associated with survival in the patients with late-stage, an association that was comparable with or even stronger than that in the overall population. Because of the limited sample size and number of deaths in the patients with early-stage, this association was not as robust. However, the same trend of effect for all 5 SNPs was observed in the patients with early-stage (Supplementary Table S3). In addition, stratified analysis was also conducted in treatment subgroups (chemotherapy only, surgery only, surgery plus chemotherapy, radiation ± chemotherapy) and similar effects were observed (data not shown). To identify potential differences by gender, we conducted a stratified analysis. The results in the female population were similar to the combined population. In male cases, 3 SNPs were not significant in Mayo Clinic population, which might be due to the limited number of male cases (Supplementary Table S4).

Main effects of individual SNPs on survival in ever-smokers

We next analyzed overall survival in the 996 ever-smokers at MD Anderson to assess the effects of the 5 SNPs described earlier on survival according to smoking status. The ever-smokers were slightly older than the never-smokers (mean age, 64.8 years vs. 61.5 years) and had a smaller proportion of women (42% vs. 67%) and adenocarcinoma cases (52% vs. 77%). The treatment regimens in the 2 groups were similar. None of the SNPs validated in the never-smokers were significantly associated with survival in the ever-smokers (Table 3). We further stratified the ever-smoker patients into former and current smokers and did not observe any significant associations within these subgroups (data not shown).

Survival tree analysis

Survival tree analysis was used to identify higher order gene–gene interactions among these 5 SNPs in modulating risk of death. Using the MD Anderson never-smoker population as a training set, we identified 2 SNPs (CD74: rs1056400 and BMP8A:rs698141) potentially having gene–gene interactions. Patients with the rs1056400_GG/rs698141_GA+AA genotype (node 3) had a 2.32-fold greater risk of death (HR: 2.32; 95%CI: 1.58–3.41; \( P = 1.72 \times 10^{-3} \)) and significantly shorter MST (14 months vs. 23 months; \( P = 4.5 \times 10^{-4}, \) log-rank test) than did patients with the rs1056400_GG/rs698141_GG or rs1056400_GA+AA genotype (nodes 1 and 2). This tree model was validated in the Mayo Clinic population: patients with the rs1056400_GG/rs698141_GA+AA genotype (node 3) had a nearly 2-fold greater risk of death (HR: 1.97; 95%CI: 1.11–3.50; \( P = 0.02 \)) and a strikingly shorter MST (by 26 months) than did patients with the rs1056400_GG/rs698141_GG or rs1056400_GA+AA genotype (nodes 1 and 2; \( P = 0.029, \) log-rank test; Fig. 2).

Discussion

NSCLC in never-smokers is unique from that in ever-smokers due to distinct clinical, histological, and genetic characteristics. These attributes warrant specific investigation of never-smokers. Although we are in the era of the genome-wide association study (GWAS), the coverage of certain genetic region on commercial available GWAS chips is not sufficient for detailed genetic analysis; this limits the power of GWAS to identify all genetic determinants. Thus study design based on prior knowledge focusing on known cancer relations is indispensable. In this context, we conducted a 2-stage, discovery-validation study to identify genetic predictors of overall survival in never-smokers with lung cancer using a pathway-based approach. By systematically evaluating SNPs in major inflammatory pathways, we found 5 SNPs in CD74, CD38, SYK, BMP8A, and IL17RA that were significantly associated with overall survival in never-smokers.
these patients. Furthermore, we analyzed and validated a survival tree model in predicting survival that takes gene–gene interactions into consideration. In comparing the associations of SNPs with survival in ever- and never-smokers, we provided evidence of distinct roles for inflammatory genetic determinants of prognosis in never-smokers with lung cancer. Moreover, we conducted a stratified analysis by clinical stage and treatments. The similar results observed in the subgroup analysis indicated the global prognostic role of these top markers regardless of treatments received.

Two SNPs—IL17RA:rs879576 and BMP8A:rs698141—are related to cytokine signaling. IL17RA is an isoform of the interleukin (IL)-17 receptors. In the presence of IL-17 ligands, these receptors can activate various downstream signaling pathways to induce macrophage recruitment, angiogenesis, and inflammatory lung diseases (28, 29). In our study, IL17RA:rs879576 was associated with a consistent protective effect against death and corresponding prolonged MSTs in both the MD Anderson and Mayo Clinic populations. This is a synonymous SNP located in the last exon of IL17RA that may influence the structure and/or regulation of its host gene. BMP8A is a member of the transforming growth factor β superfamily (30). BMP proteins play important roles in cell differentiation, proliferation, survival, and apoptosis and are implicated in tumor cell migration, metastasis, and angiogenesis in various cancers (31–34). Rs698141 is located in the first intron of BMP8A, and not in any obvious functional elements. Therefore, it is most likely linked with other functional SNPs that result in BMP8A altered function. Authors have reported that tobacco smoking can lead to immunosuppression and downregulation of proinflammatory cytokines specifically in the lung tissues, suggesting important roles for cytokines in lung pathology (35). Cytokine signaling pathway variants were predominant in our validated SNPs highlighted the potential roles of cytokines in determining prognosis for LCINS.

SYK belongs to the Syk family of tyrosine kinases and plays an oncogenic role in different cancers (36). In lung cancer cells, SYK is silenced owing to hypermethylation in its promoter region (37). SYK:rs290229 was associated with an increased risk of death and reduced survival in our populations. This SNP is located in an intron; it is possible that this SNP tagged another causal variant that affects the function of SYK. Further deep sequencing would be

![Survival Analysis](image.png)

Figure 2. Potential gene–gene interactions among SNPs validated in the survival tree analysis. A, survival tree analysis results and Kaplan–Meier estimates in the MD Anderson population (discovery phase). B, survival tree analysis results and Kaplan–Meier estimates in the Mayo Clinic population (validation phase). MST, median survival time in months. N = $A/B$, A: number of patients with event, B: total number of patients.
warranted to identify the potential casual locus responsible for this finding.

Two other SNPs—CD38:rs10805347 and CD74:rs1056400—were borderline significant in our validation. CD74 is a member of a class of polypeptides involved in antigen presentation that is a potential therapeutic target and prognostic factor for cancer (38–41) with involvement in lung adenocarcinoma. Our results suggested the potential prognostic role of CD74:rs1056400 about overall survival in patients with lung cancer. CD38 is a multifunctional single-chain type II transmembrane glycoprotein, related to the development of viral infections, diabetes, and cancer (42). Studies have shown a prognostic role for CD38 in patients with leukemia (43). We observed a consistent protective effect for CD38:rs10805347 against death, which indicated a potential role for this gene in solid cancers in addition to leukemia.

In the current study, we aimed at identifying specific prognostic markers for never smokers. Although incidence is increasing, LCINS represents only approximately 10% of all lung cancer cases. Thus, to identify a homogeneous never-smoking patient cohort with adequate demographic/clinical variables is a challenge. In this study, we were able to identify relatively large and well-characterized study populations from 2 study sites with complete collection of clinical and epidemiologic data that enabled us to recruit a sufficient study population. This provided an important resource contributing to the understanding of this disease, which has emerged as a major public health problem tracking smoking and smoking cessation rate. Interestingly, none of the 5 SNPs were significantly associated with overall survival in ever-smokers, providing additional evidence of LCINS as a distinct disease and requires identifying specific prognostic markers.

Moreover, to control for potential false discoveries, we have adopted both a FDR-based multiple testing adjustment in the discovery phase and an independent external validation, which largely reduced the likelihood of false-positive results. Another significant finding in our study was the identification and validation of a survival tree, which has proven to be a powerful analytic tool about survival in patients with cancer based on higher-order gene–gene interactions (37–39). The survival tree analysis stratified the Mayo Clinic patients into significantly different risk subgroups in a manner similar to that in the MD Anderson patients. Beyond the effect of a single SNP on survival, the survival tree takes into account the complicated interactions of genes, which are yet not discovered and has high predictive power about patients’ prognosis that may be clinically applicable.

In conclusion, this is the first large-scale study to examine the association of SNPs in 800 inflammation-related genes with survival in never-smokers with lung cancer. The identified individual SNPs and the survival tree may be applicable to future modeling of clinical outcome for prediction of survival following validation in other independent populations of never-smokers with lung cancer.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors’ Contributions

Conception and design: X. Pu, W.K. Hong, J.D. Minna, P. Yang, X. Wu

Development of methodology: X. Pu, J. Gu, W.K. Hong, P. Yang, X. Wu

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): J. Pu, M.R. Spitz, L. Wang, J.A. Roth, P. Yang, X. Wu

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): X. Pu, Y. Ye, M.R. Spitz, L. Wang, M.A.T. Hildebrandt, P. Yang, X. Wu

Writing, review, and/or revision of the manuscript: X. Pu, Y. Ye, M.R. Spitz, L. Wang, J. Gu, S.M. Lippman, M.A.T. Hildebrandt, W.K. Hong, J.D. Minna, J.A. Roth, P. Yang, X. Wu

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): M.R. Spitz, S.M. Lippman, J.A. Roth, P. Yang, X. Wu

Study supervision: P. Yang, X. Wu

Principal investigator of large program project NCI Special Program of Research Excellence grant that provided part of the support for this research: J.D. Minna

Provided funding: P. Yang, X. Wu

Grant Support

This research is supported by National Cancer Institute grants [R01 CA111646, P50 CA070907, and R01 CA055769] and in part by the NIH through MD Anderson’s Cancer Center Support Grant [CA016672].

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received March 5, 2012; revised July 31, 2012; accepted August 25, 2012; published OnlineFirst September 13, 2012.

References


29. Zhou Y, Toh ML, Ziloial S, Miossec P. IL-17A versus IL-17F induced intracellular signal transduction pathways and modulation by IL-17RA and IL-17RC RNA interference in AGS gastric adenocarcinoma cells. Cytokine 2007;38:157–64.