Polymorphisms in MicroRNAs Are Associated with Survival in Non–Small Cell Lung Cancer

Yang Zhao1, Qingyi Wei2, Lingming Hu1, Feng Chen1, Zhibin Hu1, Rebecca S. Heist3, Li Su4, Christopher I. Amos5, Hongbing Shen1, and David C. Christiani1,3,4

Abstract

Background: MicroRNAs (miRNA) play important roles in the regulation of eukaryotic gene expression and are involved in human carcinogenesis. Single-nucleotide polymorphisms (SNP) in miRNA sequence may alter miRNA functions in gene regulation, which, in turn, may affect cancer risk and disease progression.

Methods: We conducted an analysis of associations of 142 miRNA SNPs with non–small cell lung cancer (NSCLC) survival using data from a genome-wide association study (GWAS) in a Caucasian population from the Massachusetts General Hospital (Boston, MA) including 452 early-stage and 526 late-stage NSCLC cases. Replication analyses were further performed in two external populations, one Caucasian cohort from The University of Texas MD Anderson Cancer Center (Houston, TX) and one Han Chinese cohort from Nanjing, China.

Results: We identified seven significant SNPs in the discovery set. Results from the independent Caucasian cohort demonstrated that the C allele of rs2042253 (hsa-miRNA-5197) was significantly associated with decreased risk for death among the patients with late-stage NSCLC (discovery set: HR, 0.80; P = 0.007; validation set: HR, 0.86; P = 0.035; combined analysis: HR, 0.87; P = 0.007).

Conclusions: These findings provide evidence that some miRNA SNPs are associated with NSCLC survival and can be used as predictive biomarkers.

Impact: This study provided an estimate of outcome probability for survival experience of patients with NSCLC, which demonstrates that genetic factors, as well as classic nongenetic factors, may be used to predict individual outcome. Cancer Epidemiol Biomarkers Prev; 23(11); 2503–11. ©2014 AACR.

Introduction

Lung cancer is the leading cause of cancer-related deaths in the United States (1). More than 80% of lung cancers are non–small cell lung cancer (NSCLC; ref. 2). It is reported that the 5-year survival rate of patients with NSCLC ranges from 11% to 17% (3). Both environmental and genetic factors contribute to the mortality of patients with NSCLC. Although it is well known that the major risk factor for lung cancer–related deaths is smoking, several genes, including loci at 3p22.1, 5p14.1, 6q16.1, 7q31.31, 9p21.3, 11p15.1, and 14q24.3, have also been identified to have an influence on overall survival or prognosis through genetic association studies, as well as genome-wide association studies (GWAS; refs. 4, 5).

MicroRNAs (miRNA) are small, single-stranded, non–protein-coding RNAs that regulate gene expression. Although their biologic functions remain largely unknown, aberrant expression of miRNA seems to affect expression of various protein-coding oncogenes and tumor suppressors that are related to the etiology, diagnosis, and prognosis of many cancers (6–8). Recently, a growing body of evidence suggests that altered expression of some miRNAs can affect tumorigenesis and progression of NSCLC (9–20).

By altering the expression or maturation of a miRNA, single-nucleotide polymorphisms (SNP) within these 22-nucleotide sequences may lead to a dysregulation of gene expression that thereby contributes to cancer risk and survival (21–27). For example, Pu and colleagues reported that three miRNA SNPs (i.e., rs713065, rs6886834, and rs2234978) were associated with clinical outcomes in patients with early-stage NSCLC (28). The present study tested the hypothesis that SNPs in miRNAs contribute to survival of NSCLC. An association analysis was performed using genotype data from a GWAS in a Caucasian population with 452 patients with early-stage and 526 patients with late-stage NSCLC. Replication analyses were conducted in a Caucasian NSCLC cohort from The
University of Texas MD Anderson Cancer Center and a Han Chinese NSCLC cohort from Nanjing, China.

Materials and Methods

Study populations

The dataset from the discovery phase (The Harvard Lung Cancer Susceptibility Study GWAS) included 452 early-stage (I and II) and 526 late-stage (III and IV) patients with NSCLC recruited from Massachusetts General Hospital (Boston, MA). Details of participant recruitment for the study have been described previously (29). DNA was extracted from the whole blood and genotyped using the Illumina 610K Quad chip (Illumina).

Positive hits from the discovery set were validated in patients with NSCLC from two replication patient cohorts. One of the two cohorts comprised Caucasians in a lung cancer-control study at The University of Texas MD Anderson Cancer Center (Houston, TX). Sampling, genotyping, and quality control procedures have been described previously (5, 30, 31). The dataset included 788 newly diagnosed and histopathologically confirmed lung cancer cases. The second replication cohort was from a Han Chinese population in Nanjing, China (4, 32). This study included 609 patients with histopathologically or cytologically confirmed NSCLC.

A written informed consent was obtained from each subject at the time of recruitment, and the study was approved by the Institutional Review Boards of each participating institution.

Quality control in GWAS

We conducted systematic quality control (QC) on the raw genotyping data to filter both unqualified samples and SNPs (33). SNPs were excluded if they met any one of the following conditions: (i) SNPs that did not map on autosomal chromosomes; (ii) SNPs that had a call rate <95%; or (iii) SNPs that had minor allele frequency (MAF) <0.05. Samples with low call rates (<95%), ambiguous sex, familial relationships (PLHAT > 0.25), outliers in the principal component analysis, and extreme heterozygote rate (>6 SD from nearest neighbor) were removed. Finally, a total of 543,697 SNPs passed the general QC procedure.

Extracting of miRNA SNPs

A list of miRNAs was downloaded from an online miRNA database (miRBase: http://www.mirbase.org, release 18; refs. 34, 35). We used liftover (http://genome.ucsc.edu/util.html), an online tool to lift the version of assembly from Hg18 to Hg19. However, because of the short length of miRNA regions, only 16 SNPs in the Illumina 610K Quad chip could be matched to miRNAs.

To increase the opportunity to capture miRNA SNPs, we first performed a genotype imputation procedure. The reference CEU (Utah residents with ancestry from northern and western Europe) panel was downloaded from the 1000 Genomes Project (phase I, release 2010-6; http://www.1000genomes.org). MACH (http://www.sph.umich.edu/csg/abecasis/MACH/) was used to impute the ungenotyped SNPs (36). Among the 4,649,540 SNPs that passed QC, there were 142 miRNA SNPs. Similarly, the imputation procedure for the MD Anderson cohort was also performed by using MACH and the CEU reference panel from the 1000 Genomes Project. For the Nanjing cohort, the reference panel was based on HapMap phase II database (CHB+JPT, released July 17, 2006).

Statistical analysis

We performed a two-stage association analysis. In the first stage, survival analyses were performed on the basis of the discovery GWAS dataset. To reach satisfactory power, we used a significant level of 0.01. We also used the false discovery rate (FDR) to evaluate the proportion of false positives among our findings (37). The survival time was defined as the length of period (unit, month) from the time of diagnosis until death or the latest follow-up. Cox proportional hazards model analysis was performed on both early- and late-stage patients. For the 452 early-stage patients, covariates adjusted included age, sex, smoking status, cell type (adenocarcinoma, squamous, and the others), stage (I vs. II), and the top four principal components (PC). For the 526 late-stage patients, age, sex, smoking status, cell type (adenocarcinoma, squamous, and others), stage (III vs. IV), surgery (yes vs. no), and the top four PCs were adjusted in the multivariate Cox model. To remove the possible adverse influence of some long-term survivors and allow for easy comparisons with other similar studies, those late-stage patients with more than 5 years of overall survival were right-censored. The PCs included in the model were generated by the EIGENSTRAT analysis, which were used to control for the confounding effect of population stratification (38, 39). The significant SNPs observed in stage I were then evaluated in two independent cohorts in stage II with a significance level of 0.05. Meta-analysis was performed to synthesize the results from different study cohorts. To evaluate the association of the validated miRNA SNPs on NSCLC survival, we performed a time-dependent receiver operating characteristic (ROC) analysis to calculate the cumulative area-under-the-curve (AUC) of the miRNA SNPs, as proposed by Chambless and Diao (40).

We used PLINK 1.07 for GWAS data management and general statistical analysis (41). The “survival” package in R (PLINK plug-in; http://www.r-project.org/) was used to conduct the survival analysis. Meta-analysis was performed by using “metan” package in Stata (version 12). The time-dependent ROC analysis was performed using the “survAUC” package in R. The target miRNAs of the miRNAs, including miR-#-5p and miR-#-3p, were predicted by using TargetScan (42), and the predicted miRNAs were further queried for GO functional enrichments using CapitalBio Molecule Annotation System V4.0 (MAS, http://bioinfo.capitalbio.com/mas3/).

Results

Characteristics of the NSCLC cases in the discovery set and the two replication patient cohorts are described in Table 1. In the discovery set, mean ages of patients

Downloaded from cebp.aacrjournals.org on July 9, 2017. © 2014 American Association for Cancer Research.
MicroRNA SNPs and Survival of NSCLC

with early- and late-stage NSCLC were 67.86 and 63.47 years, respectively. The proportions of males with early and late-stage NSCLC were 47.35% and 49.24%, respectively. There were 174 (52.65%) current and 238 (38.50%) former smokers in the early-stage group. For the late-stage patients, the proportions of current and former smokers were 49.43% and 40.68%, respectively. For the histology type, more than half of the patients in both early and late stages were squamous (early stage, 29.87%; late stage, 15.02%) or adenocarcinoma (early stage, 44.03%; late stage, 45.25%).

We identified seven miRNA SNPs in the Harvard cohort; all were imputed. In the early-stage survival analysis, we identified three miRNA SNPs with P values < 0.01, assuming an additive genetic model. The G>A variation of rs11048315 (hsa-mir-4302) was associated with increased survival time (HR, 0.64; 95% CI, 0.55–0.79; P = 0.003; FDR q = 0.035). And only rs2042253 was significant at the level of 0.05 with the effect in the same direction as the discovery set (HR, 0.86; 95% CI, 0.74–0.99; P = 0.035). Thus, we used fixed effects model for data synthesis. Rs2042253 was significantly associated with the survival of patients with late-stage NSCLC (HR, 0.87; 95% CI, 0.80–0.96; P = 0.003). For rs7522956, there was a significant association between the genotype of rs7522956 and the survival of patients with early-stage NSCLC (HR, 1.22; 95% CI, 1.11–1.65; P = 0.003; FDR q = 0.115). Detailed results of the survival analysis are shown in Tables 2 and 3, as well as Supplementary Table S1 and Supplementary Figs. S1 and S2.

To validate these findings, we analyzed miRNA SNPs with P < 0.01 from the discovery set in the two replication populations. Of the seven SNPs, rs7522956 and rs2042253 were found in the MD Anderson cohort’s imputed dataset. Only rs2042253, was significant at the level of 0.05 with the effect in the same direction as the discovery set (HR, 0.86; 95% CI, 0.74–0.99; P = 0.035). And only rs2042253 was identified in the imputed dataset from the Nanjing cohort but was not significantly associated with survival (Supplementary Table S2). We then performed a meta-analysis on rs7522956 and rs2042253, which existed in the datasets of at least two of the three cohorts. The results are presented in Table 4. None of the effects of the two SNPs were significantly heterogeneous among the cohorts (rs7522956, P = 0.106; rs2042253, P = 0.219). Thus, we used fixed effects model for data synthesis. Rs2042253 was significantly associated with the survival of patients with late-stage NSCLC (HR, 0.87; 95% CI, 0.80–0.96; P = 0.003). For rs7522956, there was a significant association between the genotype of rs7522956 and the survival of patients with early-stage NSCLC (HR, 1.22; 95% CI, 1.05–1.42; P = 0.011). Several additional analyses were also performed. For the patients with early-stage NSCLC, rs7522956 was significantly associated with the progression-free survival.

| Table 1. Basic characteristics of the study populations from the three cohorts |
|---------------------------------|-----------------|-----------------|-----------------|
|                                | Harvard cohort (discovery) | MD Anderson cohort (replication) | Nanjing cohort (replication) |
|                                | Early stage | Late stage | Early stage | Late stage | Early stage | Late stage |
| Age Mean ± SD                  | 67.86 ± 9.66 | 63.47 ± 10.87 | 65.43 ± 9.85 | 60.23 ± 10.51 | 60.19 ± 10.38 |
| Sex Male                       | 214 (47.35%) | 259 (49.24%) | 122 (50.62%) | 331 (60.51%) | 439 (72.09%) |
| Female                         | 238 (52.65%) | 267 (50.76%) | 119 (49.38%) | 216 (39.49%) | 170 (27.91%) |
| Smoking status                 |              |              |              |              |              |
| Never                          | 20 (8.65%)   | 2 (0.89%)    | 0            | 0            | 232 (38.10)  |
| Former                         | 238 (52.65%) | 260 (49.43%) | 141 (58.51%) | 271 (54.77%) | 395 (64.86)  |
| Current                        | 174 (38.50%) | 214 (40.68%) | 100 (41.49%) | 276 (50.46%) | 304 (49.91)  |
| Histology                      |              |              |              |              |              |
| Adeno                          | 199 (44.03%) | 285 (54.18%) | 119 (49.38%) | 275 (50.27%) | 395 (64.86)  |
| SQC                            | 135 (29.87%) | 79 (15.02%)  | 74 (30.71%)  | 147 (26.87%) | 214 (35.14)  |
| Others                         | 118 (26.11%) | 162 (30.80%) | 48 (19.92%)  | 125 (22.85%) | 0            |
| Stage                          |              |              |              |              |              |
| I                              | 381 (84.29%) | —            | 176 (73.03%) | —            | —            |
| II                             | 71 (15.71%)  | —            | 65 (27.97%)  | —            | —            |
| III                            | —            | 238 (45.25%) | —            | 302 (55.21%) | 376 (61.74)  |
| IV                             | —            | 288 (54.75%) | 245 (44.79%) | 233 (38.26)  |              |

Abbreviations: adeno, adenocarcinoma; SQC, squamous carcinoma.
Table 2. Significant miRNA SNPs (P < 0.01) in the early-stage survival analysis (discovery set)

<table>
<thead>
<tr>
<th>miRNA</th>
<th>rs11048315</th>
<th>rs10416415</th>
<th>chr7:129197463</th>
<th>rs7522956</th>
</tr>
</thead>
<tbody>
<tr>
<td>Death</td>
<td>CC</td>
<td>CC</td>
<td>183 (48.67)</td>
<td>189 (55.26)</td>
</tr>
<tr>
<td>Censored N (%)</td>
<td>18 (48.67)</td>
<td>19 (50.00)</td>
<td>1.00 (reference)</td>
<td>1.00 (reference)</td>
</tr>
<tr>
<td>Median survival time (mo)</td>
<td>91.4</td>
<td>91.3</td>
<td>1.00 (reference)</td>
<td>1.00 (reference)</td>
</tr>
<tr>
<td>Unadjusted model</td>
<td>1.63</td>
<td>1.21</td>
<td>0.80</td>
<td>0.73</td>
</tr>
<tr>
<td>Adjusted model</td>
<td>1.73</td>
<td>1.21</td>
<td>0.80</td>
<td>0.73</td>
</tr>
<tr>
<td>P</td>
<td>0.26</td>
<td>0.005</td>
<td>0.004</td>
<td>0.94</td>
</tr>
<tr>
<td>HR (95% CI)</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>P</td>
<td>0.26</td>
<td>0.005</td>
<td>0.004</td>
<td>0.94</td>
</tr>
<tr>
<td>FDR</td>
<td>0.40</td>
<td>0.004</td>
<td>0.004</td>
<td>0.84</td>
</tr>
<tr>
<td>q</td>
<td>0.007</td>
<td>0.007</td>
<td>0.006</td>
<td>0.006</td>
</tr>
</tbody>
</table>

Discussion

There are an increasingly number of reports on the association between miRNA SNPs and survival of patients with lung cancer. Hu and colleagues reported that the C allele of rs11614913 (hsa-mir-196a2) was significantly associated with a decreased survival of patients with NSCLC (23). A meta-analysis published by Chen and colleagues demonstrated that hsa-mir-196a2 could also be a potential biomarker of lung cancer risk (43). Pu and colleagues reported that some miRNA-related SNPs (FZD4:rs713065, DROSHA:rs6886834, FAS:rs2234978) may be associated with NSCLC patients’ clinical outcomes through altered miRNA regulation of the target genes (28). In a Han Chinese population, Cheng and colleagues suggested that the functional SNP rs2240688A>C in CD133 could be a functional biomarker to predict risk and prognosis of lung cancer (44).
Table 3. Significant miRNA SNPs ($P < 0.01$) in the late-stage survival analysis (discovery set)

<table>
<thead>
<tr>
<th>SNP</th>
<th>Chr</th>
<th>Position (bp, Hg19)</th>
<th>Genotype</th>
<th>Death N (%)</th>
<th>Censored N (%)</th>
<th>Median survival time (mo)</th>
<th>Comparison</th>
<th>Unadjusted model HR (95% CI)</th>
<th>P</th>
<th>Adjusted model HR (95% CI)</th>
<th>P</th>
<th>FDR q</th>
<th>MiRNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs17111728</td>
<td>1</td>
<td>55691384</td>
<td>TT</td>
<td>372 (83.22%)</td>
<td>75 (16.78%)</td>
<td>13.9</td>
<td>1 vs. 0</td>
<td>0.74 (0.55–1.01)</td>
<td>5.71E–02</td>
<td>0.63 (0.46–0.86)</td>
<td>3.87E–03</td>
<td>5.71E–02</td>
<td>hsa-mir-4422</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TC</td>
<td>48 (76.19%)</td>
<td>15 (23.81%)</td>
<td>22.3</td>
<td>2 vs. 0</td>
<td>0.36 (0.09–1.43)</td>
<td>1.45E–01</td>
<td>0.25 (0.06–1.01)</td>
<td>5.20E–02</td>
<td>1.45E–01</td>
<td>hsa-mir-4422</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CC</td>
<td>2 (40%)</td>
<td>3 (60%)</td>
<td>—</td>
<td>Additive</td>
<td>0.71 (0.54–0.94)</td>
<td>1.72E–02</td>
<td>0.60 (0.45–0.8)</td>
<td>4.79E–04</td>
<td>1.72E–02</td>
<td>hsa-mir-4422</td>
</tr>
<tr>
<td>rs2042253</td>
<td>5</td>
<td>143059433</td>
<td>TT</td>
<td>223 (83.52%)</td>
<td>44 (16.48%)</td>
<td>13.5</td>
<td>1 vs. 0</td>
<td>0.88 (0.72–1.09)</td>
<td>2.47E–01</td>
<td>0.79 (0.64–0.97)</td>
<td>2.43E–02</td>
<td>2.47E–01</td>
<td>hsa-mir-5197</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TC</td>
<td>159 (84.13%)</td>
<td>30 (15.87%)</td>
<td>16.1</td>
<td>2 vs. 0</td>
<td>0.72 (0.46–1.11)</td>
<td>1.40E–01</td>
<td>0.65 (0.42–1.01)</td>
<td>5.44E–02</td>
<td>1.40E–01</td>
<td>hsa-mir-5197</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CC</td>
<td>22 (73.33%)</td>
<td>8 (26.67%)</td>
<td>23.4</td>
<td>Additive</td>
<td>0.87 (0.74–1.02)</td>
<td>8.70E–02</td>
<td>0.79 (0.67–0.94)</td>
<td>6.95E–03</td>
<td>8.70E–02</td>
<td>hsa-mir-5197</td>
</tr>
<tr>
<td>rs550894</td>
<td>11</td>
<td>6521940</td>
<td>CC</td>
<td>339 (83.91%)</td>
<td>65 (16.09%)</td>
<td>14.8</td>
<td>1 vs. 0</td>
<td>0.75 (0.58–0.98)</td>
<td>3.31E–02</td>
<td>0.74 (0.57–0.97)</td>
<td>2.89E–02</td>
<td>3.31E–02</td>
<td>hsa-mir-612</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AC</td>
<td>69 (74.19%)</td>
<td>24 (25.81%)</td>
<td>20.1</td>
<td>2 vs. 0</td>
<td>0.40 (0.13–1.25)</td>
<td>1.16E–01</td>
<td>0.30 (0.09–0.90)</td>
<td>4.03E–02</td>
<td>1.16E–01</td>
<td>hsa-mir-612</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AA</td>
<td>3 (50%)</td>
<td>3 (50%)</td>
<td>25.2</td>
<td>Additive</td>
<td>0.73 (0.57–0.92)</td>
<td>9.05E–03</td>
<td>0.70 (0.55–0.88)</td>
<td>2.98E–03</td>
<td>9.05E–03</td>
<td>hsa-mir-612</td>
</tr>
<tr>
<td>rs7227168</td>
<td>18</td>
<td>20513374</td>
<td>CC</td>
<td>320 (80.6%)</td>
<td>77 (19.4%)</td>
<td>16.4</td>
<td>2 vs. 0</td>
<td>1.38 (1.09–1.76)</td>
<td>8.49E–03</td>
<td>1.49 (1.17–1.91)</td>
<td>1.44E–03</td>
<td>1.38E–03</td>
<td>hsa-mir-4741</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TC</td>
<td>89 (85.58%)</td>
<td>15 (14.42%)</td>
<td>10.1</td>
<td>1 vs. 0</td>
<td>1.69 (0.87–3.28)</td>
<td>1.23E–01</td>
<td>1.31 (0.66–2.6)</td>
<td>4.43E–01</td>
<td>1.23E–01</td>
<td>hsa-mir-4741</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TT</td>
<td>9 (100%)</td>
<td>0 (0%)</td>
<td>11.9</td>
<td>Additive</td>
<td>1.35 (1.11–1.65)</td>
<td>3.16E–03</td>
<td>1.35 (1.11–1.65)</td>
<td>3.11E–03</td>
<td>1.35E–03</td>
<td>hsa-mir-4741</td>
</tr>
</tbody>
</table>
potential predictive value of hsa-mir-5197 (rs2042253) on the survival of patients with late-stage NSCLC. For the patients with early-stage NSCLC, although rs7522956 was not significant in the replication cohort, it can also improve the predictive ability of the model with clinical score only.

Hsa-mir-5197 (rs2042253), located on the long arm of chromosome 5, was significant in Caucasian populations from both Harvard and MD Anderson; the T>C variation of this SNP provided a protective effect on lung cancer survival. This novel miRNA is reported to have a high read frequency for pediatric acute lymphoblastic leukemia (ALL) through high-throughput sequencing (45). Although not significant in the Nanjing replication cohort, our meta-analysis showed that hsa-mir-4742 (rs7522956) was significantly associated with the survival of patients with early-stage NSCLCs, which was also associated with the PFS in early-stage patients in the Harvard cohort. Rs2042253 is located in the adjacent region of miR-5197-5p, also called a microRNA-offset RNA (moRNA; refs. 46–48), and rs7522956 is located in loop sequences of the mir-4742 gene (Supplementary Fig. S4_A). The nucleotide variations in pre-miRNAs may have an effect on the stability of the stem-loop structure, and even contribute to pre-miRNA processing via affecting recognition and cleavage of Drosha and Dicer. Although mature miRNA sequences can be generated by these pre-miRNAs with varied nucleotides, multiple isomiRs in miRNA processing and maturation processes may be regulated (49–53). These mature miRNAs, including miR-#-5p and miR-#-3p, have important biologic roles through contributing to basic multiple biologic processes (such as TGF-beta signaling pathway, cytokine–cytokine receptor interaction, and the insulin signaling pathway) and development of some human cancers (such as colorectal cancer, prostate cancer, and endometrial cancer; Supplementary Fig. S4_B). Further studies are needed to understand the roles of these two miRNA SNPs in the survival of patients with NSCLC.

The present study has several strengths. We used three relatively large datasets from three independent GWASs for the discovery and confirmation of the association between miRNA SNPs and overall survival. Thus, SNPs that were identified by the analysis should have a high probability of being true-positive findings. We also performed a time-dependent AUC to demonstrate that the identified miRNA SNP could be a biomarker for the survival of patients with NSCLC.

However, we acknowledge that there are several limitations of the present study. First, most microRNA-related variants may not be well covered by current GWAS platforms. Although we used imputed datasets for SNP extracting, it is possible that some miRNA SNPs identified in one cohort may not exist in the dataset of another cohort, because the SNPs may not in the reference panel or the imputed SNPs are of low quality. Second, although one miRNA SNP was validated in both Caucasian cohorts,
no SNPs were positive in the Nanjing Cohort. Possible reasons for this discrepancy include different ethnic background, demographic, and clinic characteristics of the cohorts, as well as potential gene–gene interactions between functional SNPs and gene–environmental interactions, which may result in the failure to replicate for some statistically significant SNPs in an independent dataset (54). Third, in vitro functional assays are needed for evidence of biologic plausibility of the identified miRNA SNPs.

In conclusion, the present study provides evidence that SNPs in some miRNAs are associated with NSCLC survival. Our result provides evidence for the application of miRNA SNPs as predictive biomarkers in future personalized medicine for patients with NSCLC. Further investigation is needed to illustrate the precise mechanism by which the miRNA SNPs affect NSCLC survival.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Authors' Contributions
Conception and design: Y. Zhao, Q. Wei, F. Chen, H. Shen, D.C. Christiani
Development of methodology: L. Su, C.I. Amos, D.C. Christiani
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): Q. Wei, L. Hu, R.S. Heist, L. Su, C.I. Amos, D.C. Christiani
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): Y. Zhao, R.S. Heist, C.I. Amos, D.C. Christiani
Writing, review, and/or revision of the manuscript: Y. Zhao, Q. Wei, R.S. Heist, C.I. Amos, H. Shen, D.C. Christiani
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): Z. Hu, L. Su, D.C. Christiani
Study supervision: Y. Zhao, Q. Wei, F. Chen, Z. Hu, H. Shen, D.C. Christiani

Acknowledgments
The authors thank Dr. Margaret Spitz for her suggestions on preparing this article. The authors also thank Dr. Li Guo and Dr. Guangfu Jin for their support of this work. The authors thank the participants and the physicians and staff of the Massachusetts General Hospital, the M.D. Anderson Cancer Center, and the Nanjing Medical University affiliated hospitals. The authors also greatly appreciate the constructive comments from the two anonymous reviewers.

Grant Support
This study was supported by NIH grants (R01CA092824, P50CA090578, and P30ES000002 to D.C. Christiani), the Natural Scientific Funding of China [NSFC, 30901232, 81373102, and NIH grants 5R01CA092824 (PI, D. C. Christiani) to Y. Zhao; NSFC 81072389 to F. Chen], Scientific Research Grants for High Education of Jiangsu Province (12KJB310003 to Y. Zhao), and The National Science Foundation for Distinguished Young Scholars of China (81225029 to Z. Hu).

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 April 15, 2014; revised July 17, 2014; accepted July 22, 2014; published OnlineFirst August 7, 2014.

References


Polymorphisms in MicroRNAs Are Associated with Survival in Non–Small Cell Lung Cancer

Yang Zhao, Qingyi Wei, Lingming Hu, et al.


Updated version  Access the most recent version of this article at: doi:10.1158/1055-9965.EPI-14-0389

Supplementary Material  Access the most recent supplemental material at: http://cebp.aacrjournals.org/content/suppl/2014/08/16/1055-9965.EPI-14-0389.DC1

Cited articles  This article cites 53 articles, 16 of which you can access for free at: http://cebp.aacrjournals.org/content/23/11/2503.full#ref-list-1

E-mail alerts  Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions  To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions  To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.