Abstract

Although most patients with small-cell lung cancer respond to chemotherapy, the survival time is highly diverse. We conducted a genome-wide analysis to examine whether germline genetic variations are prognostic factors in small-cell lung cancer patients treated with the same chemotherapy regimen. Genome-wide scan of single nucleotide polymorphisms (SNP) was performed using blood DNA to identify genotypes associated with overall survival in 245 patients treated with platinum-based chemotherapy, and the results were replicated in another independent set of 305 patients. Associations were estimated by Cox models and function of the variants was examined by biochemical assays. We found that rs1820453 T>G SNP within the promoter region of YAP1 on chromosome 11q22 and rs716274 A>G SNP in the region of downstream of DYNC2H1 on chromosome 11q22.3 are associated with small-cell lung cancer survival. In pooled analysis of 2 independent cohorts, the adjusted hazard ratio for patients with the rs1820453 TG or GG genotype was 1.49 (95% CI, 1.19–1.85; P = 0.0004) and 1.65 (95% CI, 1.36–2.01; P = 4.76 × 10^{-7}), respectively, compared with the TT genotype; and for patients with the rs716274 AG or GG genotype was 1.83 (95% CI, 1.47–2.29; P = 8.74 × 10^{-8}) and 2.96 (95% CI, 1.90–4.62; P = 1.59 × 10^{-5}), respectively, compared with the AA genotype. Functional analysis showed that the rs1820453 T>G change creates a transcriptional factor binding site and results in downregulation of YAP1 expression. These results suggest that YAP1 may play an important role in prognosis of small-cell lung cancer patients treated with platinum-based chemotherapy. Cancer Res 70(23); 9721–9. ©2010 AACR.

Introduction

Lung cancer, the leading cause of cancer-related deaths all over the world, is classified as small-cell lung cancer and non-small-cell lung cancer, with the latter occurring more frequently than the former. However, small-cell lung cancer, which accounts for ~20% of total lung cancer, is a very aggressive neuroendocrine malignancy characterized by high growth rate, rapid doubling time, propensity for the early development of widespread metastases, and poor prognosis (1–5). Although most patients with small-cell lung cancer respond to chemotherapy such as the standard first-line platinum-based drugs, the survival time among patients given the same regimen is highly diverse. Identification of prognostic factors is among the many steps toward personalized medicine, which stands to have a profound impact on the clinical care of cancer patients (6). Individual clinical characteristics including age, smoking history, and disease stage may affect prognosis of small-cell lung cancer patients (7), however, these clinical characteristics can only explain in small part the great heterogeneity of patient’s survival, indicating that there might be other factors determining disease prognosis.

In the recent years, evidence has been accumulated to show that genetic variations such as single nucleotide polymorphisms (SNP) in both tumor and host genomes are associated with patient’s survival. Screening for such variants at a genome-wide level is now possible due to advances in high-throughput genotyping methodology, and a few genome-wide association studies on complex diseases and survival of patients have been reported (8–10). However, to the best of our knowledge, no genome-wide association study on small-cell lung cancer survival has so far been reported.

Here we report a genome-wide study on the association between germline genetic variations and survival of small-cell lung cancer.

Subjects and Methods

Patients

The genome-wide discovery phase included 245 patients with small-cell lung cancer recruited at Cancer Hospital,
Chinese Academy of Medical Sciences (Beijing) between July 2000 and June 2008 (11). The replication phase included 305 patients with small-cell lung cancer who were recruited at Cancer Hospital of Jiangsu Province, the First Affiliated Hospital of Nanjing Medical University and Nanjing Thoracic Hospital (Nanjing), and 4 tertiary referral hospitals at Wuhan city, Hubei Province (Wuhan) between March 2002 and March 2008. All patients were ethnic Han Chinese. To be included in this study, all patients had to have cytologically confirmed small-cell lung cancer and received the first-line carboplatin (AUC 5–6, day 1) or cisplatin (60–80 mg/M², day 1) plus etoposide (100 mg/M², days 1–3) chemotherapy for at least 2 cycles. According to the Veterans’ Administration Lung Study Group, all patients were classified as having limited disease or extensive disease on the basis of the results of a physical examination; computed tomography scan of the chest, liver, and adrenal glands; a magnetic resonance imaging scan or computed tomography scan of the head; and a bone scan. All patients tolerated chemotherapy as defined by a Karnofsky performance score of at least 70, a life expectancy of at least 3 months, and having adequate organ function. Overall survival time of patients was measured from the date of treatment to the date of last follow-up or death. Whether and when a patient had died was obtained from inpatient and outpatient records, patients’ families, or local Public Security Census Register Office through follow-up telephone calls. The last date of follow-up was 30th June 2009 and no patients were lost to follow-up. Patients alive on the last follow-up date were considered censored. The median follow-up time was 19 months for patients in the discovery cohort and 23 months for patients in the replication cohort. Written informed consent was obtained from all patients and this study was approved by the Institutional Review Board of the Chinese Academy of Medical Sciences Cancer Institute, Nanjing Medical University, and Huazhong University of Science and Technology Tongji Medical College.

Genotyping and quality control
Genomic DNA was isolated from peripheral blood lymphocytes of each patient using a commercial Flexi Gene DNA extraction kit (Qiagen). A total of 440,093 SNPs were genotyped for 245 DNAs isolated from each patient recruited at Cancer Hospital, Chinese Academy of Medical Sciences (11) using Affymetrix GeneChip Human Mapping 5.0 set (Affymetrix) according to manufacturer’s instructions. We carried out systematic quality control on the raw genotyping data. No DNA sample failed to be genotyped because each individual had <10% of missing genotypes. No individuals were excluded for showing gender discrepancies based on their chromosomes X genotypes. SNPs were excluded by the following step-wise procedure: (1) 9,771 SNPs were not mapped on autosomal chromosomes, (2) 17,426 SNPs had a call rate of <0.95, (3) 146,843 SNPs had a minor allele frequency (MAF) <0.1 (the reason we chose this rigorous threshold was because this study sought to identify potential SNPs that can be used as prognostic biomarkers in clinics), and (4) 57 SNPs were excluded as the genotype distributions clearly deviated from that expected by Hardy–Weinberg equilibrium (chi-square >50, P < 5 × 10−13). After these quality control steps, the median sample call rate was 99.5% and a total of 265,996 SNPs were used in analysis of the discovery phase. In the replication phase, we used the MassArray system (Sequenom) to genotype the top 26 SNPs, which were selected from the discovery phase as tagSNPs when the r² is >0.8 based on the pair-wise linkage disequilibrium analysis using Haplovew software (Version 4.4; http://www.broad.mit.edu/mpg/haplovew).

Functional analysis of genetic variants
In the present study, we identified 3 significant SNPs, rs10895256, rs1820453, and rs716274. Because rs716274 is located in gene desert, functional studies were performed only for rs10895256 and rs1820453, which lie in the 5’ flanking region of YAP1. For electrophoretic mobility-shift assays, synthetic double-stranded and 3’ biotin-labeled oligonucleotides corresponding to the rs10895256 G/T or rs1820453 G/T sequences and H446 (human small-cell lung cancer cell line) nuclear extracts were incubated at room temperature for 20 minutes using the LightShift Chemiluminescent EMSA kit (Pierce). The reaction mixture was separated on 8% polyacrylamide gel electrophoresis, and the products were detected by Stabilized Streptavidin-Horseradish Peroxidase Conjugate (Pierce). For competition assays, unlabeled oligonucleotides at 10-fold or 100-fold molar excess were added to the reaction mixture before the addition of biotin-labeled probes.

For luciferase reporter gene assays, a 1,528-bp DNA fragment corresponding to the YAP1 5’-untranslated region and 5’-flanking region and containing the T allele at −875 position relative to the transcriptional start site (rs1820453T allele) was generated by polymerase chain reaction (primers are available upon request) and subcloned into the pGL3-Basic vector (Promega), and the resultant plasmid was designated as p-T. The p-T construct was then site-specifically mutated to create the p-G construct, which contains the G at −875 position (rs1820453G allele). These 2 constructs used in this study were restriction mapped and sequenced to confirm their authenticity. We seeded 5 × 10⁵ H446 cells per well in 48-well plates and transfected them with an empty pGL3-Basic vector (a promoterless control), p-T or p-G construct. pRL-SV40 plasmid (Promega) was cotransfected as an internal control. All transfections were carried out in triplicate. After 48 hours, cells were collected and analyzed for the luciferase activity with a Dual-Luciferase Reporter Assay System (Promega).

Analysis of YAP1 RNA
Total RNA was isolated from surgically removed normal lung tissues adjacent to tumors of 56 patients with small-cell lung cancer and then converted to cDNA using oligo (dT)₁₅ primer and Superscript II (Invitrogen). YAP1 RNA was measured by real-time quantitative reverse transcription-polymerase chain reaction in triplicate using the ABI 7900HT Real-Time PCR system based on the SYBR-Green method. The measurement of individual YAP1 RNA expression was calculated relative to that of GAPDH expression using a modification of the method described by Lehmann and Kreipe (12). The primers used for detecting YAP1
RNA are 5'-GGCGCTCTTCAACGCCGTCATGAAC-3' and 5'-CCTGTCGGGAGTGGGATTT-3'.

**Statistical analysis**

For each of the selected 265,996 SNPs, Cox regression under a log-additive genetic model was performed using 245 small-cell lung cancer patients recruited in Beijing with adjustment for covariates that might influence patients' survival, including age, sex, disease stage, and smoking status. The top 26 SNPs with \( P < 10^{-4} \) were advanced to replication in patients independently recruited from Nanjing and Wuhan. The same Cox models were used to analyze each of these 26 SNPs and Bonferroni correction was performed to correct the \( P \) value of single-locus results for the 26 testing. Hazard ratios (HR) and their 95% confidence intervals (CI) were calculated. Kaplan–Meier survival estimates were plotted and \( P \) values were assessed using log-rank tests. Student's \( t \) test was used to examine the differences in luciferase reporter gene expression and Mann–Whitney \( U \)-test was used to assess differences in \( YAP1 \) transcript abundance with different genotypes. All statistical tests were carried out in a 2-sided manner.

**Results**

**Patient characteristics**

The characteristics of small-cell lung cancer patients in the discovery and replication cohorts are shown in Table 1. In the discovery cohort, 186 (75.9%) patients died of the cancer but 59 (24.1%) patients had still survived until June 2009. Among the patients in the replication cohort, 191 (62.6%) died of the cancer but 114 (37.4%) had still survived. There were no significant differences in terms of age, sex, and smoking history in patients between the discovery and replication cohorts. However, the proportion of patients with limited disease was higher in the replication cohort (31.8%) than in the discovery cohort (16.7%). Among the select characteristics, only tumor stage showed significant association with patients' survival, with the median overall survival time of 31 months (95% CI, 24–39) for limited disease and 19 months (95% CI, 17–21) for extensive disease.

**Genetic variants associated with patient survival**

We found that only rs716274 SNP was significantly associated with overall survival using Cox models assuming the additive mode and adjusting for age, sex, disease stage, and smoking status after Bonferroni correction (\( P = 1.32 \times 10^{-7} \)). Because our sample size in the discovery cohort is relatively small and the statistical power might be limited, we therefore selected all top SNPs that had \( P < 10^{-4} \) for replication (Table 2).

Because some of these top SNPs are located in the same genomic regions and are in strong linkage disequilibrium, we were able to select tagSNPs to represent each region in tests for an independent association. According to the linkage disequilibrium analysis on the basis of \( r^2 \geq 0.8 \), 26 SNPs or tagSNPs were finally selected for validation in the replication cohort of 305 small-cell lung cancer patients. Among the 26 SNPs, 2 SNPs (rs10895256 G\( > \)T and rs716274 A\( > \)G) were confirmed to be significantly associated with overall survival of small-cell lung cancer patients in the replication cohort using Cox models adjusting for covariates (both \( P < 0.002 \) after Bonferroni correction; Supplementary Table 1).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Discovery Cohort (( n = 245 ))</th>
<th>Replication Cohort (( n = 305 ))</th>
<th>Pooled Cohort (( n = 550 ))</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Death</strong></td>
<td>186 (75.9)</td>
<td>191 (62.6)</td>
<td>377 (68.6)</td>
</tr>
<tr>
<td><strong>Survival</strong></td>
<td>59 (24.1)</td>
<td>114 (37.4)</td>
<td>173 (31.4)</td>
</tr>
<tr>
<td><strong>Median survival, months</strong></td>
<td>19</td>
<td>23</td>
<td>21</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>193 (78.8)</td>
<td>232 (76.1)</td>
<td>425 (77.3)</td>
</tr>
<tr>
<td>Female</td>
<td>52 (21.2)</td>
<td>73 (23.9)</td>
<td>125 (22.7)</td>
</tr>
<tr>
<td><strong>Age, y</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(&lt;50)</td>
<td>76 (31.0)</td>
<td>76 (24.9)</td>
<td>152 (27.6)</td>
</tr>
<tr>
<td>(&gt;50)</td>
<td>169 (69.0)</td>
<td>229 (75.1)</td>
<td>398 (72.4)</td>
</tr>
<tr>
<td><strong>Clinical stage</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Limited</td>
<td>41 (16.7)</td>
<td>97 (31.8)</td>
<td>138 (25.1)</td>
</tr>
<tr>
<td>Extensive</td>
<td>204 (83.3)</td>
<td>208 (68.2)</td>
<td>412 (74.9)</td>
</tr>
<tr>
<td><strong>Smoking status</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonsmoker</td>
<td>74 (30.2)</td>
<td>89 (29.2)</td>
<td>163 (29.6)</td>
</tr>
<tr>
<td>Smoker</td>
<td>171 (69.8)</td>
<td>216 (70.8)</td>
<td>387 (70.4)</td>
</tr>
</tbody>
</table>
rs10895256 SNP is located 2.8 kb from the upstream of the YAP1 gene on chromosome 11q22, whereas rs716274 SNP is located on chromosome 11q22.3 in the region of ~67 kb downstream of the DYNC2H1 gene.

In pooled analysis of the 2 independent cohorts, the median overall survival time for the rs10895256 GG, GT, and TT genotypes were 24 months (95% CI, 21–25), 18 months (95% CI, 16–21), and 11.5 months (95% CI, 9–14), respectively (Fig. IA). The adjusted HR for patients with the rs10895256 GT or TT genotype was 1.49 (95% CI, 1.19–1.85; \( P = 0.0004 \)) or 1.65 (95% CI, 1.36–2.01; \( P = 4.76 \times 10^{-5} \)) compared with the GG genotype (Table 3). The median overall survival for the rs716274 AA, AG, and GG genotypes were 24 months (95% CI, 22–26), 17 months (95% CI, 15–19) and 14 months (95% CI, 13–16), respectively (Fig. 1B). Patients with the rs716274 AG (HR = 1.83, 95% CI, 1.47–2.29; \( P = 8.74 \times 10^{-5} \)) or GG (HR = 2.96, 95% CI, 1.90–4.62; \( P = 1.59 \times 10^{-5} \)) genotype had an increased risk of death compared with the AA genotype (Table 3).
Identification of other variant in *YAP1* promoter region

Because rs10895256 SNP is located at 28 kb far from the transcriptional start site of *YAP1* and electrophoretic mobility shift assay showed that the binding pattern of nuclear proteins prepared from H446 cells was very similar for both rs10895256 G and rs10895256 T alleles (Fig. 2A, Lanes 2 and 4), this SNP might not affect transcriptional factor binding and thus is not functional relevant. We therefore examined a 2,806-bp potential promoter region of *YAP1* based on the genotyped SNPs derived from the International HapMap Project database (CHB population data) to find potentially functional variants. As a result, 1 additional SNP, rs1820453 T>G, was found in this region located at −875 bp from the transcriptional start site of *YAP1*. We thus genotyped this SNP in all of our study subjects and found that it was in complete linkage disequilibrium with rs10895256 SNP ($r^2 = 1.00$) and showed exactly the same association with overall survival time and HRs for death of patients as did the rs10895256 genotypes (Supplementary Table 2 and Supplementary Fig. 1).

Functional characterization of rs1820453 SNP

Regulatory sequences with discrete alleles might influence gene expression upon binding transcriptional activators or inhibitors that instruct their regulatory control. Therefore, we first examined whether the rs1820453 T>G SNP changes the binding pattern of nuclear proteins. EMSA showed that the binding pattern for rs1820453 G and rs1820453 T differed greatly; 1 additional DNA-protein complex (IV) was formed when the rs1820453 G probe was incubated with nuclear proteins from H446 cells (Fig. 2A, Lane 2) compared to the rs1820453 T probe under the same experimental conditions (Fig. 2A, Lane 7). Competition assays showed that the 100-fold excess of unlabeled rs1820453 G (Fig. 2B, Lane 4) but not rs1820453 T probe (Fig. 2B, Lane 5) adding to the reaction mixture significantly eliminated the DNA-protein complex formed by the interaction between the rs1820453 G and nuclear proteins, indicating that the binding are sequence specific.

We then examined whether the rs1820453 T>G change affected *YAP1* promoter activity. To do this, a set of luciferase

---

**Figure 1.** Kaplan–Meier estimates of overall survival of patients with small-cell lung cancer according to rs10895256 G>T (A) or rs716274 A>G (B) genotypes. The top, middle, and bottom represent discovery cohort from Beijing, replication cohort from Nanjing and Wuhan, and pooled cohort, respectively. All $P < 0.0001$ for the log-rank test.
Table 3. Hazard ratios of rs10895256 G>T and rs716274 A>G genotypes for death of patients with small-cell lung cancer in the discovery cohort from Beijing, replication cohort from Nanjing and Wuhan, and pooled cohort

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Discovery Cohort</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
<td>HR (95% CI)</td>
<td>P</td>
<td>No.</td>
<td>HR (95% CI)</td>
<td>P</td>
<td>No.</td>
<td>HR (95% CI)</td>
<td>P</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs10895256</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>142</td>
<td>1.00 (Reference)</td>
<td>0.0013</td>
<td>164</td>
<td>1.00 (Reference)</td>
<td>0.0091</td>
<td>306</td>
<td>1.00 (Reference)</td>
<td>0.0004</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GT</td>
<td>91</td>
<td>1.69 (1.23–2.31)</td>
<td>0.0003</td>
<td>113</td>
<td>1.37 (1.01–1.87)</td>
<td>0.0491</td>
<td>204</td>
<td>1.49 (1.19–1.85)</td>
<td>0.0004</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT</td>
<td>12</td>
<td>1.83 (1.32–2.54)</td>
<td>0.0003</td>
<td>28</td>
<td>1.61 (1.26–2.06)</td>
<td>0.0001</td>
<td>40</td>
<td>1.65 (1.36–2.01)</td>
<td>0.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs716274</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>147</td>
<td>1.00 (Reference)</td>
<td>0.0017</td>
<td>189</td>
<td>1.00 (Reference)</td>
<td>0.0001</td>
<td>336</td>
<td>1.00 (Reference)</td>
<td>0.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG</td>
<td>86</td>
<td>2.03 (1.49–2.78)</td>
<td>8.06×10⁻⁶</td>
<td>102</td>
<td>1.64 (1.19–2.25)</td>
<td>0.0023</td>
<td>188</td>
<td>1.83 (1.47–2.29)</td>
<td>8.74×10⁻⁸</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>12</td>
<td>2.84 (1.48–5.43)</td>
<td>0.0017</td>
<td>14</td>
<td>2.92 (1.58–5.40)</td>
<td>0.0006</td>
<td>26</td>
<td>2.96 (1.90–4.62)</td>
<td>1.59×10⁻⁶</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: HR, hazard ratio; CI, confidence interval.

*aNumber of patients.

*bCalculated with multivariate Cox models adjusting for age, sex, disease stage, and smoking status.
rs1820453 T-containing promoter has higher transcriptional activity.

When the TG and GG genotypes were pooled together in analysis, the difference between the TT genotype (0.020 vs. 0.012; P = 0.011) and nuclear extracts. The arrow indicates a DNA-protein complex. B, EMSA with biotin-labeled oligonucleotides containing the rs1820453G or rs1820453T allele and nuclear extracts from H446 cells. Lanes 1 and 6 show mobilities of the labeled oligonucleotides without nuclear extracts; lanes 2 and 7 show mobilities of the labeled oligonucleotides with nuclear extracts in the absence of competitor; lanes 3 and 8, 4 and 9, and 10 show mobilities of the labeled oligonucleotides with nuclear extracts in the presence of unlabeled rs1820453G or rs1820453T competitors. The arrow indicates an additional DNA-protein complex (IV) for the rs1820453G allele. C, reporter gene assays with constructs containing the rs1820453G (p-G) or rs1820453T (p-T) YAP1 promoter in H446 cells. All constructs were cotransfected with pRL-SV40 to standardize transfection efficiency. Luciferase levels of pGL3-Basic and pRL-SV40 were determined in triplicate. Fold increase was measured by defining the activity of the empty pGL3-Basic vector as 1. Data shown are the means ± SE from 3 independent transfection experiments, each performed in triplicate. The rs1820453G-containing YAP1 promoter drove significantly lower reporter gene expression than the rs1820453T promoter (P < 0.001). D, levels of YAPI RNA expression in lung tissues as a function of YAPI rs1820453 T-G genotype. Columns, mean; bars, ± SE normalized to GAPDH. Expression level among the GG (n = 3) or TG (n = 27) genotype was significantly (P = 0.003) lower than that among the TT genotype (n = 26).

Figure 2. Functional characterization of rs10895256 G→T and rs1820453 T→G SNPs in the YAPI gene. A, electrophoretic mobility-shift assays (EMSA) with biotin-labeled oligonucleotides containing the rs10895256G or rs10895256T allele and nuclear extracts from H446 cells. Lanes 1 and 3 show mobilities of the labeled oligonucleotides without nuclear extracts; lanes 2 and 4 show mobilities of the labeled oligonucleotides with nuclear extracts. The arrow indicates a DNA-protein complex. B, EMSA with biotin-labeled oligonucleotides containing the rs1820453G or rs1820453T allele and nuclear extracts from H446 cells. Lanes 1 and 6 show mobilities of the labeled oligonucleotides without nuclear extracts; lanes 2 and 7 show mobilities of the labeled oligonucleotides with nuclear extracts in the absence of competitor; lanes 3 and 8, 4 and 9, and 10 show mobilities of the labeled oligonucleotides with nuclear extracts in the presence of unlabeled rs1820453G or rs1820453T competitors. The arrow indicates an additional DNA-protein complex (IV) for the rs1820453G allele. C, reporter gene assays with constructs containing the rs1820453G (p-G) or rs1820453T (p-T) YAP1 promoter in H446 cells. All constructs were cotransfected with pRL-SV40 to standardize transfection efficiency. Luciferase levels of pGL3-Basic and pRL-SV40 were determined in triplicate. Fold increase was measured by defining the activity of the empty pGL3-Basic vector as 1. Data shown are the means ± SE from 3 independent transfection experiments, each performed in triplicate. The rs1820453G-containing YAP1 promoter drove significantly lower reporter gene expression than the rs1820453T-containing YAP1 promoter (P < 0.001). D, levels of YAPI RNA expression in lung tissues as a function of YAPI rs1820453 T-G genotype. Columns, mean; bars, ± SE normalized to GAPDH. Expression level among the GG (n = 3) or TG (n = 27) genotype was significantly (P = 0.003) lower than that among the TT genotype (n = 26).

Discussion

In this genome-wide study, we identified 2 genetic variants that are significantly associated with overall survival of small-cell lung cancer patients treated with platinum-based chemotherapy in 2 independent cohorts. To the best of our knowledge, this is the first report assessing the relationship between germline genetic variants and prognosis of small-cell lung cancer. In addition, although the functional significance of rs716274 A→G SNP remains to be elucidated, we showed that the rs1820453 T→G change in the YAPI promoter region results in decrease in the YAPI expression due to the creation of transcriptional factor binding site, suggesting this SNP to be
causative of variation in survival of patients with small-cell lung cancer.

The rs1820453 T>G SNP is located in the promoter region of YAP1 gene on chromosome 11q22, a region shown to be amplified in 5% to 10% of human cancers including lung, oral, esophageal, pancreatic, liver, and ovarian carcinomas; however, it has also been shown in a site of frequent loss of heterozygosity in breast pancreatic, liver, and ovarian carcinomas; however, it has also been shown in a site of frequent loss of heterozygosity in breast cancer (13–18). YAP1 encodes Yes-associated protein 1 (YAP1), a transcriptional coactivator that is important in P73-dependent apoptosis (19–22). It has been shown that cellular response to cytotoxic chemotherapy is critically dependent on the function of P73 and reduction of P73 levels in tumor cells by siRNA, or the expression of dominant-negative mutants, leads to a strong reduction of apoptosis induced by DNA-damaging agents such as cisplatin (20, 22). Upon DNA damage signaling, P73 accumulates to allow transcription of downstream apoptotic genes, in which YAP1 plays a crucial role in stabilizing and enhancing P73 activity. In the present study, we found that the rs1820453 T>G change forms a transcriptional factor (likely a transcriptional repressor) binding site in the promoter region of YAP1, resulting in considerably decreased expression of YAP1 in target tissues. This functional significance may be the underlying mechanism connecting the rs1820453 G allele (GG and GT genotypes) to poor survival of small-cell lung cancer patients, because downregulation of YAP1 would be expected to weaken the function of P73-dependent apoptosis pathway and suppresses chemotherapy-induced cancer cell death, which in turn results in faster cancer progression and shorter patients’ survival.

The rs716274 A>G SNP is located in ‘gene desert’, a site of ~67 kb downstream of DYNC2H1. Although DYNC2H1 mutations cause asphyxiating thoracic dystrophy and short rib-polydactyly syndrome type III (23), its role in cancer, especially in small-cell lung cancer, is currently unknown and further studies are required.

The present study has several strengths. Unlike other cancers such as non-small-cell lung cancer where surgical resection, chemotherapy with multiple anticancer drugs, and radiotherapy are often jointly used, small-cell lung cancer patients were mainly and simply treated with platinum-based chemotherapy. Studying such a homogeneously treated group of patients may have enhanced our ability to find associations if the main genetic factors that influence small-cell lung cancer survival do so primarily in the context of this particular regimen. Another major strength of our study is the 2-phase design in which small-cell lung cancer patients were recruited from different hospitals for SNP discovery and validation, which would largely reduce false positive findings from the genome-wide scan. Indeed, we examined 26 SNPs identified from the discovery cohort and only 2 were validated in the replication cohort. Third, we have characterized the function of rs1820453 SNP, making the association of this SNP with prognosis of small-cell lung cancer biologically plausible.

However, despite of aforementioned strengths, we also acknowledge several limitations of this study. First, the modest sample size of both discovery and replication cohorts might not have the optimal statistical power, so false negative and/or chance findings can not fully be excluded. It was noted that in the discovery phase only 1 SNP, rs716274, was statistically significantly associated with patients’ survival after Bonferroni correction \((P = 1.32 \times 10^{-7})\). Given this finding, we replicated 26 top SNPs and fortunately found that besides rs716274 SNP, rs1820453 SNP (in complete linkage disequilibrium with rs10895256) was also significant. Second, in vitro functional analysis and expression data can only indirectly support the association. While functional analysis showed rs1820453 T>G change creating a transcriptional inhibitor binding site, the identity of nuclear protein bound to this site remains to be clarified. Third, although we identified 2 SNPs in YAP1 associated with patients’ survival, we did not resequence whole YAP1 gene, which might miss other functional and causative YAP1 variants.

In conclusion, by using genome-wide interrogation and replication strategy, we identified 2 SNPs, rs1820453 T>G and rs716274 A>G, as genetic factors affecting overall survival of small-cell lung cancer patients treated with platinum-based chemotherapy. Although the function of rs716274 A>G SNP is unclear, rs1820453 T>G SNP within the promoter region of YAP1 causes downregulation of the gene, which might weaken P73-dependent apoptosis of cancer cells and suppress chemotherapy-induced cancer cell death, resulting in faster cancer progression and shorter patients’ survival. Larger studies are warranted to confirm the effects of the 2 SNPs in other ethnic patient cohorts, and more thorough functional analysis will verify the role of these SNPs in small-cell lung cancer development or progression.

Disclosure of Potential Conflicts of Interest

The authors declare no competing financial interests.

Grant Support

National High Technology Project (2006AA02A401) and National Basic Research Program (2004CB518701) to D. Lin.

Received 04/27/2010; revised 09/24/2010; accepted 09/24/2010; published OnlineFirst 11/30/2010.

References


