Functional Role of S100A14 Genetic Variants and Their Association with Esophageal Squamous Cell Carcinoma

Hongyang Chen, Dianke Yu, Aiping Luo, Wen Tan, Chunpeng Zhang, Dan Zhao, Ming Yang, Zuniao Liu, Dongxin Lin, and Zhihua Liu

Abstract

S100 proteins have been implicated in various human diseases, including certain types of cancer. Among them, S100A14 is down-regulated in esophageal squamous cell carcinoma (ESCC). In this study, we sought to identify functional genetic variants in the S100A14 locus and assessed their associations with susceptibility to ESCC. Thirty individual DNA samples were sequenced to search for genetic variations in S100A14, and the function of the variants was investigated by a set of biochemical assays. A case-control analysis was performed in 1,021 patients with ESCC and 1,253 control subjects. Odds ratios and 95% confidence intervals (95% CI) were computed by logistic regression model. Four single nucleotide polymorphisms, 43A>G, 461G>A, 1493A>G, and 1545A>T, were identified in the S100A14 locus and they are in absolute linkage disequilibrium. Among them, the 461G>A change was shown to diminish a P53-binding site and is therefore associated with decreased expression of S100A14 in vitro and in vivo in the target tissues. Case-control analysis showed that the 461A allele was associated with susceptibility to ESCC among smokers, with the ORs being 2.01 (95% CI, 1.37–3.22) for the 461GA or 461AA genotype, respectively, compared with the 461GG genotype. The 461G>A change was shown to diminish a P53-binding site and its regulation, which may underlie phenotypic variation in susceptibility to ESCC. We further showed that the 461G>A SNP located in the 5'-untranslated region (5'-UTR) of S100A14 might function as a cancer suppressor working in the P53 pathway and play a role in esophageal carcinogenesis. [Cancer Res 2009;69(8):3451–7]

Introduction

Esophageal squamous cell carcinoma (ESCC), one of the most aggressive cancers, is spread worldwide, with a particularly high incidence rate in China (1). Epidemiologic studies have revealed that environmental factors, such as tobacco smoking, heavy alcohol drinking, and micronutrient deficiency, are linked to the etiology of this malignancy (2). In addition, chronic inflammation, which has been associated with carcinogenesis in various tissue sites, might also contribute to the risk of developing ESCC (3, 4). However, only a portion of individuals develop ESCC, suggesting that individual genetic makeup may play a role in esophageal carcinogenesis. Among genetic variations, single nucleotide polymorphisms (SNP) have been shown to contribute to individual susceptibility to ESCC (5–8).

S100A14, encoded by the S100A14 gene, is overexpressed in certain types of tumor, such as ovarian, breast, and uterine tumor, but down-regulated in some tumors, such as kidney, colon, and rectal tumor (17). We have previously shown that S100A14 is down-regulated in ESCC, suggesting that it may play a role in the development of ESCC (18). In view of the role that S100 proteins may play in cancer, we hypothesized that genetic variations in S100A14 might underlie phenotypic variation in susceptibility to ESCC. In this study, we sought to identify functional S100A14 variants and examined their association with susceptibility to ESCC in a Chinese population. By sequencing the 5'-untranslated region (UTR), coding region, and 3'-UTR of S100A14, we found four SNPs, among which the 461G>A SNP located in the 5'-UTR is associated with ESCC susceptibility. We further showed that the 461G>A change affects P53 binding and its regulation, which may be the mechanism underlying host susceptibility to the cancer.

Materials and Methods

SNP identification. Thirty DNA samples derived from WBCs of randomly selected healthy subjects (all were Han Chinese) were used to search for SNPs within the ~0.52 kb promoter region, 5'-UTR, coding region, and 3'-UTR of S100A14 (GenBank accession no. AF426828). In reference to the human S100A14 gene sequences, four sets of PCR primers were designed for SNP screening. SNPs were identified by directly sequencing the PCR products with ABI Prism Dye Terminator sequencing kits and ABI 3730 sequencer (Applied Biosystems).

Study subjects. The case-control analysis consisted of 1,021 patients with ESCC and 1,253 cancer-free controls. All subjects were unrelated Han Chinese and participants in previously published studies (7, 8, 19). Briefly, patients were consecutively recruited between July 1999 and July 2003 at the Chinese Academy of Medical Sciences Cancer Hospital (Beijing). All patients with histopathologically confirmed ESCC were enrolled, and there was no sex and age restriction. The exclusion criteria included other

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).
cancer(s) and previous chemotherapy or radiotherapy. Controls were cancer-free individuals randomly selected from a community cancer screening program for early detection of cancer based on a physical examination. Controls had no individual history of cancer and were frequency-matched to patients for sex and age. At recruitment, informed consent was obtained from each subject and personal data from each participant about demographic characteristics, such as sex and age, and related risk factors, including tobacco smoking, were collected. This study was approved by the Institutional Review Board of the Chinese Academy of Medical Sciences Cancer Institute.

Genotyping of S100A14 polymorphisms. Genotypes of −43A>G SNP were analyzed by the tetra-primer amplification refractory mutation system-PCR method (20), whereas genotypes of 461G>A, 1493A>G, and 1545A>T SNPs were determined by PCR-based RFLP. The primers for genotyping are shown in Supplementary Table S1. The genotypes distinguished by amplification refractory mutation system or RFLP were further confirmed by DNA sequencing. For quality control, a 10% masked, random sample of DNA from the patients and controls was tested twice by different persons and the results were concordant for all of the masked duplicate sets.

Construction of reporter gene plasmids. Five DNA fragments, corresponding to the −511 to +6, −203 to +6, −149 to +6, −90 to +6, and +1 to 466 regions, were generated by PCR and subcloned into pGL3-Basic vector (Promega). The resultant plasmids, designated P1, P2, P3, P4, or U-G, respectively, were sequenced to confirm containing exclusively wild-type alleles at −43 and 461 positions relative to transcriptional start site. The P-AG construct, containing −43A/461G, was then site-specifically mutated to create constructs P-GA, P-GG, and P-AA. All constructs were identical except for the different allele at the −43 or 461 polymorphic site. Two other pGL3-Control vector (Promega) constructs, p-UTR-AA and p-UTR-GT, contained the 1493A/1545A or 1493G/1545T allele of the S100A14 5′-UTR, respectively. All constructs used in this study were restriction mapped and sequenced to confirm their authenticity.

Cell culture. Human colon carcinoma cell lines HCT116/P53+/+ and HCT116/P53−/− were kindly provided by Dr. B. Vogelstein of Johns Hopkins University (Baltimore, MD). Human ESCC cell lines KYSE450 and KYSE150 were gifts from Dr. Y. Shimada of Kyoto University (Kyoto, Japan), and EC9706 was established in our own laboratory. HeLa, HCT116/P53+/+, and HCT116/P53−/− cells were maintained in DMEM, whereas KYSE450, KYSE150, and EC9706 cells were maintained in RPMI 1640 supplemented with 10% fetal bovine serum, 100 units/mL streptomycin, and 100 units/mL penicillin.

Transient transfection and luciferase assay. HeLa, KYSE450, HCT116/P53+/+, and HCT116/P53−/− cells were used for luciferase assays. Cells were plated in a multwell plate and grown to 60% to 70% confluence. Reporter plasmid (100 ng) was transfected to cells using Lipofectamine reagent (Life Technologies). Transfection efficiency was standardized by cotransfecting with 1.0 ng of pRL-SV40 (Promega). Luciferase levels of pGL3-Basic and pRL-SV40 were determined in triplicate. Fold increase was calculated by defining the activity of the empty pGL3-Basic vector as 1. Columns, mean of three independent transfection experiments, each performed in triplicate; bars, SE.

Electrophoretic mobility shift assays. Double-stranded oligonucleotides corresponding to the S100A14 −43A, −43G, 461G, 461A, or

Figure 1. Reporter gene assays with constructs containing the S100A14 promoter with different deletions or mutations. A, schematic representation of the S100A14 5′-flanking region and reporter gene constructs used in this study. B, luciferase expression of different constructs in HeLa or KYSE450 cells. C, luciferase expression in HCT116/P53+/+ or HCT116/P53−/− cells of different constructs with mutations at the −43 and 461 positions. 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. Columns, mean of three independent transfection experiments, each performed in triplicate; bars, SE. *, P < 0.001, compared with P-AG; **, P < 0.001, compared with P-GG.
P53-responsive element sequences (available upon request) were synthesized and 5’ end-labeled with biotin. Nuclear extracts were prepared from HCT116/P53+/+ and HCT116/P53+/− cells, which were treated with hydroxyacamptothecin (10 μM/L) for 15 h to induce P53 expression, using NE-PER extraction reagents (Pierce). The probes and nuclear extracts were incubated at 25°C for 20 min using the LightShift Chemiluminescent electrophoretic mobility shift assay (EMSA) kit (Pierce). The reaction mixture was separated on 6% PAGE and the products were detected by stabilized streptavidin/horseradish peroxidase conjugate (Pierce). For competition or supershift assays, unlabeled probes at 200-fold molar excess or antibodies against P53 (Santa Cruz Biotechnology) were added to the reaction before the addition of biotin-labeled probes.

Chromatin immunoprecipitation assays. HCT116/P53+/+ cells, grown in DMEM and treated with hydroxyacamptothecin (10 μM/L) for 15 h, were cross-linked in 1% formaldehyde for 10 min. DNA from the fixed-chromatin cells were then subjected to immunoprecipitation using a chromatin immunoprecipitation (ChIP) assay kit (Upstate) and antibodies against P53 or anti-mouse IgG. Purified DNA was analyzed by PCR with the primers 5’-gggacagtgaagaagttg-3’ and 5’-gggacagtggagacagaa-3’, which produced a 215-bp fragment of S100A14 5’-UTR containing the 461 polymorphic site.

Real-time analysis of S100A14 RNA. Total RNA was isolated from surgically removed normal esophageal tissues adjacent to the tumors of 32 individual patients and then converted to cDNA using oligo(dT)15 primer and SuperScript II (Invitrogen). S100A14 RNA was measured by real-time quantitative reverse transcription-PCR using the ABI Prism 7000 sequence detection system (Applied Biosystems) based on the SYBR Green method. Each assay was done in triplicate and the expression of individual S100A14 measurements was calculated relative to expression of β-actin using a modification of the method described by Lehmann and Kreipe (21). The primers used for S100A14 and β-actin are shown in Supplementary Table S1.

Hydroxyacamptothecin treatment and adenovirus infection. Logarithmically growing cells were treated with 10 μM/L of hydroxyacamptothecin for 3, 6, 9, 12, and 15 h. Exponentially growing cells were infected with adenovirus hP53 (SiBiono GeneTech, Co.) at a multiplicity of infection of 40.

Western blot analysis. Protein isolation and Western blot were performed as described (22). Antibodies against P53 or S100A14 (a gift from Dr. Iver Petersen, University Hospital Charité, Berlin, Germany) were used. β-Actin antibody (A5316; Sigma) was used to test for equal loading.

Statistical analysis. A χ² test was used to examine the differences in demographic variables, smoking, and genotype distribution of S100A14 polymorphisms between patients and controls. The associations between S100A14 genotypes and risk of developing ESCC were assessed by odds ratios (OR) and their 95% confidence intervals (95% CI), which were computed by logistic regression models using Statistical Analysis System software (version 9.2; SAS Institute). We tested the null hypothesis of multiplicative gene-smoking interaction by evaluating departure from multiplicative joint effect model by including the main effect variables and their product terms in the logistic regression model (23). A more than multiplicative joint effect was suggested when OR11 > OR10 × OR01. Departure from this multiplicative model was assessed by including the main effect variables and their product terms in the logistic regression model. A homogeneity test was done to compare the difference between smoking-related ORs among different genotypes or between the product of related ORs and joint effect OR. The degree of pairwise linkage disequilibrium between SNPs, as quantified by the disequilibrium coefficient D’ and r² values, was calculated by Haplovie software (24). All statistical tests were two-sided and P < 0.05 was considered significant.

Results

SNP identification and linkage disequilibrium analysis. Four SNPs, −43A>G, 461G>A, 1493A>G, and 1545A>T, were identified by sequencing S100A14 in DNA from 30 individuals (Supplementary Fig. S1). These SNPs are located in the promoter region (−43A>G, rs4072245), 5’-UTR (461G>A, rs11548103), and 3’-UTR (1493A>G, rs11548104 and 1545A>T, rs11548102) of S100A14, respectively, and they have been recorded in the National Center for Biotechnology Information SNP database. Sequencing and genotyping results showed that all these SNPs are in absolute linkage disequilibrium (D’ = 1.00; r² = 1.00), with the minor allelic frequencies for the −43G, 461A, 1493G, and 1545T allele being 0.298 in 1,253 controls.

Analysis of the S100A14 5’-flanking region. To investigate the preliminary role of the S100A14 5’-flanking region, a set of deletion constructs was generated (Fig. 1A) and transfected into HeLa and KYSE450 cells. Intriguingly, the P-AG construct, containing the full
length of the 5′-flanking region, displayed, respectively, a ~1.8-fold higher activity in KYSE450 or one-third activity in HeLa cells compared with P1, containing the same promoter region but lacking the 5′-UTR (Fig. 1B). These results, appearing contradictory, indicated that 5′-UTR plays a key role in the transcriptional regulation of S100A14 and exhibits differential efficacy in different cell types. Furthermore, almost no reporter activity driven by the U-G, which just retains the 5′-UTR, was detected, suggesting that the S100A14 5′-UTR is a regulator but not a transcriptional element. The transcriptional activity of P1 was comparable with that of P2 in KYSE450 cells but lower than P2 in HeLa cells, suggesting that the promoter region between −511 and −203 may play a minor negative role in the transcription. The reporter gene expression driven by P3, containing the fragment from −149 to the transcriptional start site, or P4, containing the fragment from −90 to the transcriptional start site, decreased dramatically compared with that driven by P2 (Fig. 1B), showing the region between −203 and −90 as a core S100A14 promoter region.

**Effects of genetic variants on S100A14 promoter activity and P53-binding activity.** We then investigated whether the −43A>G or 461G>A change influences S100A14 promoter activity. As shown in Fig. 1C, the reporter gene expression (mean ± SE) driven by P-AG (95.27 ± 1.69) or P-GG (105.79 ± 2.57) was significantly greater than that driven by P-AA (77.34 ± 2.79) or P-GA (81.81 ± 1.80) in HCT116/P53+/+ cells (all P < 0.001). However, the expression levels driven by the 461G-containing S100A14 promoters (71.25 ± 2.17 for P-AG and 106.05 ± 2.99 for P-GG) were comparable with those driven by the 461A-containing counterparts (77.08 ± 1.65 for P-AA and 102.99 ± 2.79 for P-GA) in HCT116/P53+/− cells (all P > 0.05). Transcriptional activities driven by the −43A-containing S100A14 promoters were comparable with those driven by the −43G-containing counterparts in HCT116/P53+/− cells but significantly lower than those driven by the −43G-containing counterparts in HCT116/P53+/− cells. These results indicated that the −43A>G and 461G>A variants are functionally significant and the effect is dependent on P53 status of the cells. The reporter gene expression driven by p-UTR-AA or p-UTR-GT did not differ significantly (Supplementary Fig. S2), suggesting that both 1493A>G and 1545A>T changes do not influence RNA stability.

EMSA was then performed to examine whether the −43A>G or 461G>A change affects the binding of transcriptional factors such as P53. Under our experimental conditions, two DNA-protein complexes (Fig. 2A, lane 2, bands I and II) were detected when the 461G probe was incubated with the HCT116/P53+/− nuclear extract, whereas only one complex was detected when this probe was incubated with the HCT116/P53+/− nuclear extract (Fig. 2A, lane 5, band I). For the 461A probe, however, only one DNA-protein complex corresponding to band I was detected with either HCT116/P53+/− or HCT116/P53−/− nuclear extracts (Fig. 2A, lanes 7 and 10). The DNA-protein complex (band II) formed by the 461G probe was also seen when the P53-responsive element probe was reacted with HCT116/P53+/− nuclear extract (Fig. 2B, lane 6).

In competition assays, the DNA-protein complexes formed by the 461G probe could be eliminated by unlabeled P53-responsive element or antibody against human P53 (Fig. 2B, suggesting that the protein bound to the 461G allele is P53. In addition, both the −43A and −43G probes were also able to bind a nuclear protein from both HCT116/P53+/− and HCT116/P53−/− cells (data not shown), the identity of which is unknown. Binding of P53 to the S100A14 461G allele was further shown by ChIP assays with antibody against human P53 or anti-mouse IgG using HCT116/P53+/− cells carrying the S100A14 461G/A genotype treated with hydroxycoptothecin.

**Effects of 461G>A change on S100A14 expression.** Because hydroxycoptothecin treatment enhances P53 to bind to the S100A14 5′-UTR as described above, we further examined whether the expression of S100A14 protein is correlated with P53 expression induction by hydroxycoptothecin. The results in Fig. 3A clearly indicated that S100A14 level was significantly increased in HCT116/P53+/− but not in HCT116/P53−/− cells treated with hydroxycoptothecin, showing that elevated S100A14 expression is dependent on the P53 level. We next examined the effect of S100A14 promoter genotypes on P53-induced S100A14 expression. To do this, cells with the S100A14 461GG (EC9706 and HaCaT) or 461AA (KYSE150) genotype were infected with adenovirus hP53. As shown in Fig. 3B, overexpression of P53 significantly elevated the S100A14 levels in EC9706 and HaCaT cells but not in KYSE150 cells, clearly indicating that the effect of P53 on S100A14 expression depends on the S100A14 allele. The quantitative data of P53 and S100A14
expression for the three independent experiments in different cells are shown in Supplementary Fig. S3.

We also measured the S100A14 RNA levels in surgically removed normal esophageal tissues from individuals with different S100A14 461G>A genotypes (Fig. 4), and the results showed that individuals with at least one 461A allele [20.17 ± 3.52 (n = 17) versus 9.97 ± 1.68 (GA, n = 13; AA, n = 2); P = 0.026].

**S100A14 variants and the risk of developing ESCC.** A case-control panel of 1,021 patients with ESCC and 1,253 controls were genotyped to assess the association between the S100A14 SNPs and risk of developing ESCC. As shown in Supplementary Table S2, the cases and controls were adequately matched in age and sex. However, more smokers were present in patients than in controls (61.8% versus 53.8%; OR, 1.37; 95% CI, 1.16–1.63; P < 0.001). The genotype results are shown in Table 1. Because the −43A>G, 461G>A, 1493A>G, and 1545A>T polymorphisms were in absolute linkage in our study population, only the 461G>A genotype data is shown. The allele frequencies for the 461A variant were 0.298 in controls and 0.331 in cases. The frequencies for the 461GG, GA, and AA genotypes in patients (44.7%, 44.7%, and 10.6%, respectively) differed significantly from those in controls (50.0%, 40.4%, and 9.6%, respectively; χ² = 6.76; P = 0.034; df = 2). Multivariate logistic regression analysis showed that individuals with at least one 461A allele (GA or AA genotype) had a 1.23-fold (95% CI, 1.04–1.46) increased risk for developing ESCC compared with those with the 461GG genotype (Table 1). Because smoking is an established etiologic factor for ESCC, the risk of ESCC associated with the 461G>A polymorphism was therefore assessed with stratification by smoking status and smoking levels (Table 2). The 461G>A polymorphism was associated with increased risk of ESCC only among smokers but not in nonsmokers. Furthermore, significant gene-smoking interaction was observed between 461G>A and smoking, with an OR of 2.01 (95% CI, 1.50–2.69) or 2.10 (95% CI, 1.37–3.22) for smokers with the 461GA or AA genotype versus OR of 1.37 (95% CI, 1.04–1.80) for smokers with the 461GG genotype or OR of 1.17 (95% CI, 0.89–1.55) or 1.13 (95% CI, 0.73–1.76) for nonsmokers with the 461GA or AA genotype, respectively (2.01 > 1.37 × 1.17 and 2.10 > 1.37 × 1.13; all P < 0.05, test for homogeneity). When smoking was stratified by pack-year value, more significant multiplicative joint effects were found between the 461GA or AA genotype and smoking in heavy smokers (>27 pack-years; Table 2). We also examined the association between 461G>A genotypes and ESCC invasion/metastasis (tumor stage) and differentiation (tumor grade) at the time of diagnosis. However, no significant association was detected (data not shown).

**Discussion**

In the present study, we investigated sequence variations in the S100A14 gene and their effects on susceptibility to ESCC. To summarize, we identified four common SNPs in the S100A14 locus, which are in complete linkage disequilibrium. Among them, the 461G>A SNP located in the 5′-UTR abolishes transcriptional factor P53 binding and thus associated with attenuated expression of S100A14. Because down-regulation of the S100A14 gene expression was seen in ESCC (18), a case-control analysis was conducted to examine whether this functional SNP is associated with susceptibility to this cancer. Indeed, we detected a 23% increased risk of developing ESCC for the 461AA genotype and this effect was confined in smokers, with the related risk increasing by 195% in heavy smokers compared with the 461GG genotype. These results are consistent with the common notion that genetic variants in genes in the P53 pathway may confer susceptibility to cancer.

Because the transcriptional regulation of S100A14 has not been elucidated prior to this study, we first analyzed the effects of different parts of 5′-UTR on this gene transcription. Our data clearly show that 5′-UTR functions as an important transcriptional regulatory element and that the 461G>A change located in this region has a significant effect on S100A14 expression. Reporter gene assays showed that the 461G-containing promoters yielded significantly higher transcriptional activity than 461A-containing promoters in HCT116/P53+/- but not HCT116/P53-/- cells, indicating that the differential transcriptional activity between the two alleles is dependent on P53. Although the oligonucleotides corresponding to the 461G>A SNP-containing sequence of the S100A14 5′-UTR did not seem to be within the range of reported typical P53 consensus sequences, EMSA did show that the 461G probe was able to bind P53, whereas the 461A probe was not. In accordance with these results, P53 expression introduced by exogenous adenovirus hP53 significantly elevated the S100A14 production levels in cells with the S100A14 461GG but not the S100A14 461AA genotype. Together, these data implicate that the

![Figure 4](https://www.aacrjournals.org/3455/fig4.jpg)

**Figure 4.** Levels of S100A14 RNA expression in human esophageal tissues as a function of S100A14 461G>A genotype. Columns, mean normalized to β-actin; bars, SE. Expression levels among the GA (n = 13) or AA (n = 2) genotypes were significantly lower than that among the GG genotype (n = 17). *P = 0.026.

---

**Table 1.** Allele and genotype frequencies of S100A14 461G>A polymorphism among patients and controls and their associations with the risk of ESCC

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Controls (n = 1,253)</th>
<th>Patients (n = 1,021)</th>
<th>OR* (95% CI)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. (%)</td>
<td>No. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G allele</td>
<td>1,759 (70.2)</td>
<td>1,367 (66.9)</td>
<td>0.82 (0.66–1.01)</td>
<td>0.019</td>
</tr>
<tr>
<td>A allele</td>
<td>747 (29.8)</td>
<td>675 (33.1)</td>
<td>1.00 (reference)</td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>627 (50.0)</td>
<td>455 (44.7)</td>
<td>1.00 (reference)</td>
<td></td>
</tr>
<tr>
<td>GA</td>
<td>505 (40.4)</td>
<td>457 (44.7)</td>
<td>1.24 (1.04–1.48)</td>
<td>0.012</td>
</tr>
<tr>
<td>AA</td>
<td>121 (9.6)</td>
<td>109 (10.6)</td>
<td>1.24 (0.94–1.66)</td>
<td>0.126</td>
</tr>
<tr>
<td>GA + AA</td>
<td>626 (50.0)</td>
<td>566 (55.3)</td>
<td>1.23 (1.04–1.46)</td>
<td>0.007</td>
</tr>
</tbody>
</table>

*Data were calculated by logistic regression, adjusted for sex, age, and smoking status.
DNA sequences around the 461G site in the S100A14 5'-UTR may correspond to a P53-binding motif. This notion is supported by observations that besides the high-affinity P53 sequence–specific DNA binding or P53-DNA structure–selective binding motifs, there may exist numerous low-affinity P53-binding motifs in genes that also play important roles in response to P53 regulation or action (25–27).

We found that under our experimental conditions, both 461G and 461A probes were able to interact with another unknown nuclear protein, with the interaction being much stronger for the 461G probe than the 461A probe. The same DNA-protein complex was also presented when the P21-P53BS1 probe was incubated in HCT116/P53+/+ nuclear extracts. These results suggest that this unknown nuclear protein may function as a partner in the regulation of expression of the S100A14 gene by P53. This phenomenon has been well documented (28, 29) and proposed to be an important mechanism for high-order control of cellular transcriptional regulation. Another interesting but unsolved finding in the present study is that the −43G allele showed a higher transcriptional activity than −43A in HCT116/P53+/− nuclear extracts. These results clearly suggest that the −43A>G polymorphic site is a nuclear protein–binding site; however, the A>G change apparently does not affect the binding ability (Supplementary Fig. S2). Taken together, these results suggest that the differential transcriptional activity between the S100A14 −43A and −43G variants may not depend on P53. However, further studies are needed to clarify this suggestion and the exact mechanism involved in the biological significance of −43A>G polymorphism.

If a gene plays a role in certain cancers, genetic polymorphisms of this gene that alter the level of protein expressed or function would be anticipated to have a substantial influence on disease activity. Indeed, our case-control analysis showed that individuals with at least one S100A14 461A allele (AA or GA genotype), which is associated with lower expression of S100A14, had a higher risk for developing ESCC compared to those with two 461G alleles. This result is in line with our previous finding that, among 12 other S100 proteins, the expression of S100A14 is significantly down-regulated in ESCC (18). Although the biological function of S100A14 has not been fully characterized, our data on ESCC suggests that this protein may act as a cancer suppressor. This notion is further supported by our functional examination of S100A14 variants indicating that S100A14 is a downstream gene of P53. Further studies are warranted to investigate S100A14’s functions within the P53 cancer suppressor pathway. Interestingly, we found that risk of ESCC related to the S100A14 461A allele was confined to smokers. Because smoking is an established risk factor for ESCC (33), interaction between smoking and genetic variants of S100A14 would be expected and biologically reasonable. Tobacco smoke contains numerous carcinogens and oxidative agents that can directly cause DNA damage and/or esophageal inflammation (2).

Therefore, to protect against carcinogenesis, smokers would need more efficient P53 pathways to repair damage or eliminate irreparable cells. If individuals smoke and carry the S100A14 461A allele that has reduced S100A14 expression, they would be expected to have the highest risk for developing ESCC. However, although the 461G>A SNP is of functional significance and most likely to be the susceptibility site, the observed association with ESCC might be by linkage disequilibrium with other SNPs rather than 461G>A alone.

Our study subjects in case-control analysis were from Beijing City and surrounding provinces and were believed to be a good presentation of all cases and population in this region of northern China. Having a large sample size, reproducible genotyping technique, and genotype frequencies that conform to the Hardy-Weinberg equilibrium, our results are unlikely to be due to selection bias or chance. However, lack of information on alcohol drink and hot beverage and/or food is a limitation of this study because these two risk factors might interact with S100A14 SNPs or act as confounding factors.

In summary, we have identified for the first time a genetic variant in the S100A14 gene (461G>A) that is associated with susceptibility to ESCC in a Chinese population in a manner of interaction with smoking. Functional analysis showed that the 461A allele contributes to significantly decreased expression of S100A14 in vitro and in vivo in the target tissues, which is most likely due to a diminished P53 regulation. These findings constitute strong evidence in support of the notion that S100A14 might function as a cancer suppressor working in the P53 pathway and play a role in esophageal carcinogenesis.

**Disclosure of Potential Conflicts of Interest**

No potential conflicts of interest were disclosed.

**Acknowledgments**

Received 11/5/08; revised 2/1/09; accepted 2/17/09; published OnlineFirst 4/7/09.