Carcinogenesis vol.28 no.9 pp.1971–1977, 2007
doi:10.1093/carcin/bgm156
Advance Access publication August 11, 2007

Genetic polymorphisms in the Rb-binding zinc finger gene *RIZ* and the risk of lung cancer

Kyong-Ah Yoon[1], Sohee Park[2], Bin Hwangbo[1], Hyoung Doo Shin[3], Hyun Sub Cheong[3], Hai-Rim Shin[2] and Jin Soo Lee[1,4,*]

[1]Research Institute and Hospital, National Cancer Center, Goyang, Gyeonggi, Korea, [2]National Cancer Control Research Institute, Goyang, Gyeonggi 410-769, Korea, [3]Department of Genetic Epidemiology, SNP Genetics, Inc., Seoul, Korea and [4]Department of Epidemiology, The Graduate School of Public Health, Seoul National University, Seoul, Korea

*To whom correspondence should be addressed. Tel: +82 31 920 1601; Fax: +82 31 920 1520;
Email: jinslee@ncc.re.kr.

Histone methyltransferase (HMT) enzymes that methylate the lysine of histones are involved in chromatin-mediated gene expression. Previously, we reported that a novel polymorphism of *SUV39H2*, the HMT that is required for the methylation of H3-K9, was associated with an increased risk of lung cancer in Koreans. The retinoblastoma protein-interacting zinc finger gene *RIZ* (*PRDM2*) is also a member of a histone/protein-methyltransferase superfamily, and the inactivation of *RIZ* in many cancers was detected as frameshift mutations, hypermethylation and missense mutations. In this study, we show the association of *RIZ* polymorphisms with the risk of lung cancer. In a hospital-based study of 335 lung cancer patients and 335 age- and gender-matched healthy controls, 120 polymorphisms of *RIZ* were screened. Of the 120 genotyped single nucleotide polymorphisms (SNPs), 42 SNPs were selected for the statistical analysis based on their frequency (>5%) and linkage disequilibrium [LD; only a representative SNP was analyzed if there were absolute LDs (r^2 = 1)]; this resulted in three LD blocks. The *+92337G>A* and *+95701C>A* polymorphisms showed a statistically significant association with the reduced risk of lung adenocarcinomas after correcting the *P* values for multiple testing [for carrying one variant allele versus none, adjusted odds ratio (aOR) = 0.55 (95% CI = 0.38–0.78), corrected *P* = 0.04; aOR = 0.54 (95% CI = 0.38–0.77), corrected *P* = 0.02, respectively]. One haplotype (Ht5) in LD block 3 of *RIZ* was significantly associated with the reduced risk of lung adenocarcinomas (aOR = 0.28, 95% CI = 0.13–0.58) as well as overall lung cancer (aOR = 0.50, 95% CI = 0.30–0.82). This study suggested that *RIZ* polymorphisms may be important predictive markers for lung cancer susceptibility.

Introduction

Lung cancer is the leading cause of cancer mortality worldwide and also in Korea. Although cigarette smoking is an established risk factor for lung cancer, genetic diversity plays an important role in determining the ultimate outcome following exposure to tobacco carcinogens. Certain genetic polymorphisms of several genes have been associated with individual susceptibility to lung cancer due to their ability to modify the effect of tobacco smoke carcinogens (1,2). Molecular epidemiologic studies have reported the relationships of lung cancer with polymorphisms of genetic susceptibility genes, including metabolizing enzymes (cytochrome P450s, glutathione *S*-transferases) and DNA repair enzymes (hOGG1, XRCC1), to elucidate the correlation with lung cancer susceptibility (3–7). Furthermore, the possible association of cancer susceptibility and genetic variations in the genes involved in the structure of the chromatin and histone methylation has been investigated in cancer studies (8–10). Recently, the association between a polymorphism of SMYD3, a histone H4 lysine 4-specific methyltransferase, and cancer was reported. The tandem repeat polymorphism of SMYD3 increased the risk of colorectal cancers, hepatocellular carcinomas and breast cancers (8). Cebrian *et al.* (9) reported a preliminary observation regarding the association with breast cancer for variants in DNA methyltransferase and histone methyltransferases (HMTs). We also reported the increased lung cancer risk associated with a polymorphism of *SUV39H2*, one of the HMTs (10).

HMT class enzymes contain a conserved catalytic core domain termed the SET (Suvar3-9, Enhancer of zeste, Trithorax) domain, which shares sequence homology with an independently described sequence motif, the PR (PRDI-BFI and RIZ) domain (11). Isolated as a retinoblastoma-binding (Rb) protein, RIZ contains the Rb-binding motif, the nuclear hormone receptor-binding motif and the PR domain (12,13). The *RIZ* gene that is located on human chromosome 1p36 produces two mRNAs, RIZ1 that contains the PR domain and RIZ2 that lacks this domain (14–16). RIZ1 but not RIZ2 has tumor suppressive properties and is frequently silenced in many human cancers, including breast, liver, colon and lung cancers (17–19). It has been shown that promoter hypermethylation of *RIZ1* is a common mechanism involved in the inactivation of the *RIZ1* gene (20–22). Frameshift mutations of the *RIZ* gene frequently occur in microsatellite instability-positive tumors of the colon, stomach, endometirum and pancreas (23–25). Missense mutations of RIZ1 that target the PR domain were also reported in human diffuse large B-cell lymphoma (DLBL) (26). Moreover, RIZ1 mutant mice showed a high incidence of DLBL and a broad spectrum of unusual tumors. RIZ1 deficiency also accelerated tumorigenesis in p53 heterozygous mutant mice (26). These findings suggest that genetic variations in the *RIZ* gene could be associated with tumor formation.

To test the hypothesis that genetic polymorphisms of the *RIZ* gene are associated with the risk of lung cancer, we analyzed the polymorphisms and haplotypes of *RIZ* in a Korean population.

Abbreviations: aOR, adjusted odds ratio; LD, linkage disequilibrium; SNP, single nucleotide polymorphism.

Materials and methods

Study population

This is a hospital-based matched case–control study. Three hundred and thirty-five cases were recruited from the patients with histologically confirmed lung cancer who visited the National Cancer Center in Korea and voluntarily participated in a health questionnaire survey conducted from May 2002 to July 2003. They donated blood for genetic tests after signing the informed consent form that was approved by the institutional review board. There were no recruitment restrictions with regard to gender or cancer stage; however, only subjects who were not older than 70 years were recruited. None of these cancer patients had received previous chemotherapy or radiotherapy prior to their recruitment. For comparison, a total of 335 control subjects were individually matched with lung cancer patients for age (±3 years) and gender. These control subjects without a prior history of cancer were recruited from the visitors of our institution for a cancer-screening program. Information on demographic characteristics including gender, age, smoking habits and family history of cancers was obtained from self-administered questionnaires (for controls) or a personal interview (for cases) administered by a trained personnel after written informed consent was obtained.

Genotyping RIZ polymorphisms

A total of 120 single nucleotide polymorphisms (SNPs) of the *RIZ* gene (NM_012231) were selected from the International HapMap Project data, (www.hapmap.org) for this study. Genomic DNA was extracted from the peripheral blood by using a QIAamp DNA Blood Mini Kit (Qiagen, Valencia, CA) in accordance with the manufacturer's instructions. For a highly multiplexed SNP genotyping assay, the GoldenGate genotyping assay (Illumina Inc., San Diego, CA) that combined an oligonucleotide ligation and allele-specific extension reaction was performed (27).

Statistical analysis

To test for the differences in demographic characteristics between lung cancer cases and controls, Pearson's χ^2 test categorical variables and Wilcoxon rank-sum test were performed for continuous variables. With regard to smoking habits, the former and present smoking status, the number of cigarettes smoked per day and the time of starting and quitting were investigated. Individuals who had either formerly or currently smoked >100 cigarettes during their lifetime were defined as ever smokers. In order to distinguish them from current smokers, former smokers were defined as those who had ceased smoking for ≥1 year at the time of sample collection. Never smokers were defined as subjects who had smoked <100 cigarettes during their lifetime. As a measure of cumulative smoking exposure, pack-years was defined as the average number of packs (20 cigarettes/pack) of cigarettes smoked per day multiplied by the total number of years of smoking.

SNPs with duplicate error or a minimum call rate under 98% were excluded for statistical analysis. Hardy–Weinberg equilibrium was tested for the genotyping results of controls. A level of $P < 0.01$ was accepted as statistically significant for the Hardy–Weinberg equilibrium test. If a polymorphism was not in Hardy–Weinberg equilibrium ($P < 0.01$), it was also excluded for further analysis.

We employed a widely used measure of linkage disequilibrium (LD) between all pairs of biallelic loci, Lewontin's D' ($|D'|$) (28) and r^2. LD blocks were identified using Haploview software (29). The $|D'|$ values for all pairs of SNPs were calculated and the haplotype blocks were estimated using the confidence interval method (30). Haplotypes of each block and individual were inferred using the algorithm developed by Stephens *et al.* (31), which (PHASE) uses a Bayesian approach incorporating a priori expectations for haplotype reconstruction. Phase probabilities of each site were calculated for each individual by this software and the haplotype with the highest probability for each sample was used for the further analysis. Genetic effects of inferred haplotypes were analyzed in same way as SNPs. The relationship between *RIZ* polymorphisms and the lung cancer risk was analyzed using both unconditional and conditional multiple logistic regression models while controlling for family history, pack-years (continuous value) and smoking status (current smoker, former smoker, never smoker) as covariates. We also analyzed the association between *RIZ* polymorphisms and the risk of adenocarcinoma and squamous cell carcinoma, which were the two most common histologic types in our study subjects. For subset analyses by cancer histologic type or smoking status, only unconditional multiple logistic regression was used because the sample size becomes small when matched pairs were maintained in the analyses. Analyses for the association between haplotype and the lung cancer risk were performed using unconditional logistic regression at the individual level, where the covariate was defined by the number of copies (0, 1 or 2) of each haplotype that a subject carried, and the model also included family history of cancer, smoking status and pack-years as covariates. To achieve the optimal correction for multiple testing of 42 SNPs, the effective number of independent marker loci (36.593) in *RIZ* was calculated with the SNPSpD software on the basis of the spectral decomposition (SpD) of matrices of pairwise LD between SNPs. The resulting number of independent marker loci was applied to correct for multiple testing (32). All reported *P* values are two sided. Statistical software SAS version 9.2 was used for statistical analyses (SAS Institute Inc., Cary, NC).

Results

The demographic distribution of lung cancer patients and healthy controls is shown in Table I. Lung cancer patients were more probably ever smokers than the controls ($P < 0.001$). The distribution of the family history of cancer was significantly different between the cases and controls ($P < 0.05$). Of the 335 lung cancer patients, 189 (56%) had adenocarcinomas, 77 (23%) had squamous cell carcinomas and 29 (9%) had small cell lung cancer. Forty patients (12%) have other histologic type that included unspecified NSCLC, large cell carcinoma, mixed large adenocarcinoma, adenosquamous carcinoma and sarcomatoid carcinoma.

Among a total of 120 SNPs of the *RIZ* gene (NM_012231), selected from the International HapMap Project data, 21 (18%) showed a monomorphic pattern in the Korean population. After excluding additional five SNPs due to Hardy–Weinberg disequilibrium ($P < 0.01$) and a low minimum call rate (<98%), 94 SNPs were analyzed to measure the LD (Figure 1A). Only a representative SNP was selected if there were absolute LDs ($r^2 = 1$) and 42 SNPs were finally selected for the statistical analysis. The genotype data of the 42 SNPs were used to estimate the LD block and haplotype structure of the *RIZ* gene. The common haplotypes (>5%) in LD blocks are indicated in Figure 1B. Three LD blocks were determined from the results of LD analysis using the Haploview software (Figure 1C).

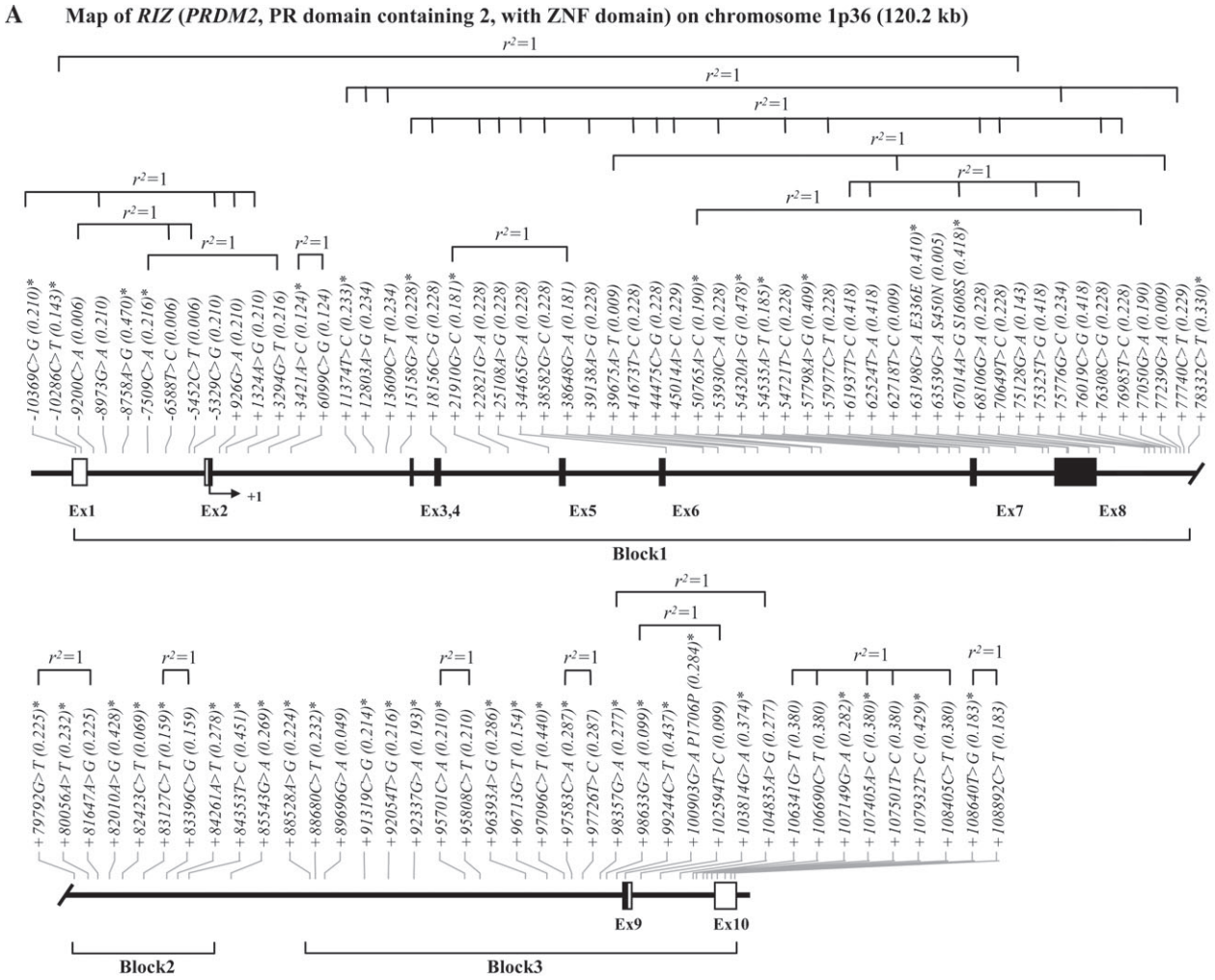
Table II presents the minor allele frequencies of the 42 SNPs among lung cancer patients and normal controls, and estimated odds ratios of lung cancer risk between subjects carrying one variant allele and subjects carrying no variant allele of each SNP, based on a multiple logistic regression model (log additive model) fit by controlling for smoking status, pack-years and family history of cancer as covariates. There were signals found in LD block 3 genotypes differentiating lung cancer patients from controls, specifically for the comparison of adenocarcinoma cases and controls. Ten of 19 SNPs in LD block 3 showed a significant association ($P < 0.05$) with the risk of lung adenocarcinoma, although eight of these 10 significant SNPs became no longer significant after the correction for multiple testing (data not shown).

Specifically, the variant genotypes of *+92337G>A* and *+95701C>A* were highly associated with a decreased risk of lung adenocarcinoma when compared with the normal controls [for carrying one variant allele versus none, adjusted odds ratio (aOR) = 0.55, 95% CI = 0.38–0.78, $P = 0.001$ and aOR = 0.54, 95% CI = 0.38–0.77, $P = 0.0006$, respectively]. To further evaluate the protective effects of *+92337G>A* and *+95701C>A*, we performed subset analyses based on the histologic cell type using four alternative models

Table I. Demographic characteristics among lung cancer patients and normal controls

Phenotype	Normal controls ($n = 335$)	Lung cancer cases		
		All ($n = 335$)	Adenocarcinoma ($n = 189$)	Squamous cell carcinoma ($n = 77$)
Age (median, range)	58 (28–73)	58 (25–70)	57 (25–70)	59 (29–68)
Gender (*n*, %)				
Male	226 (67.5%)	226 (67.5%)	96 (50.8%)	75 (97.4%)
Female	109 (32.5%)	109 (32.5%)	93 (49.2%)	2 (2.6%)
Smoking status (*n*, %)*				
Never smoker	139 (41.5%)	116 (34.6%)	96 (50.8%)	4 (5.2%)
Ever smoker	189 (56.4%)	218 (65.1%)	93 (49.2%)	73 (94.8%)
Current	86 (25.7%)	143 (42.7%)	55 (29.1%)	47 (61.0%)
Former	103 (30.7%)	75 (22.4%)	38 (20.1%)	26 (33.8%)
Unknown	7	1	0	0
Pack-years (median, range)[a]	21 (0.3–126)	34.5 (0.08–135)	26.25 (0.08–87.5)	40 (3.8–102.5)
Family history (*n*, %)*				
Yes	166 (49.6%)	128 (38.3%)	77 (40.7%)	25 (32.5%)
No	167 (49.9%)	205 (61.4%)	111 (58.7%)	51 (66.2%)
Unknown	2	2	0	1

*$P < 0.05$ from Pearson's χ^2 test for the difference between lung cancer patients and controls.
[a]Pack-years of smoking were for ever smokers only. $P < 0.05$ from Wilcoxon's rank-sum test for the difference between lung cancer patients and controls.



B Haplotypes in *RIZ*

Block1

Haplotype	-10369C>G	-10286C>T	-8758A>G	-7509C>A	+3421A>C	+11374T>C	+15158G>A	+21910G>C	+50765A>C	+54320A>G	+54535A>T	+57798A>G	+63198G>A	+67014A>G	+78332C>T	Frequency
Block1-ht1	C	C	G	C	A	T	G	G	A	G	A	A	G	A	C	0.323
Block1-ht2	G	C	A	A	A	C	A	G	A	A	A	G	A	G	T	0.195
Block1-ht3	C	T	G	C	A	T	G	G	A	G	A	A	G	A	C	0.143
Block1-ht4	C	C	A	C	C	T	G	G	A	A	A	A	G	A	C	0.104
Block1-ht5	C	C	A	C	A	T	G	C	C	A	T	G	A	G	T	0.092
Block1-ht6	C	C	A	C	A	T	G	C	C	A	T	G	A	G	C	0.081
Others	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.063

Block2

Haplotype	+79792G>T	+80056A>T	+82010A>G	+82423C>T	+83127C>T	+84261A>T	+84353T>C	+85543G>A	Frequency
Block2-ht1	G	A	G	C	C	A	T	A	0.265
Block2-ht2	G	T	A	C	C	T	T	G	0.229
Block2-ht3	T	A	A	C	C	A	C	G	0.219
Block2-ht4	G	A	G	C	T	A	C	G	0.159
Block2-ht5	G	A	A	T	C	A	C	G	0.069
Others	-	-	-	-	-	-	-	-	0.058

Block3

Haplotype	+88528A>G	+88680C>T	+91319C>G	+92054T>G	+92337G>A	+95701C>A	+96393A>G	+96713G>T	+97096C>T	+97583C>A	+98357G>A	+98633G>A	+99244C>T	+100903G>A	+103814G>A	+107149G>A	+107405A>C	+107932T>C	+108640T>G	Frequency
Block3-ht1	A	C	C	T	G	C	A	G	C	C	G	G	C	G	G	G	A	T	T	0.344
Block3-ht2	G	C	C	T	G	C	A	G	C	C	G	G	C	G	G	G	A	T	T	0.135
Block3-ht3	A	T	G	G	A	A	G	T	T	A	A	G	T	A	A	A	C	C	T	0.078
Block3-ht4	A	C	C	T	G	C	A	G	T	C	G	A	T	G	A	G	C	C	T	0.075
Block3-ht5	A	T	G	G	A	A	G	G	T	A	A	G	T	A	A	A	C	C	G	0.062
Others	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.305

C. LDs among *RIZ* SNPs

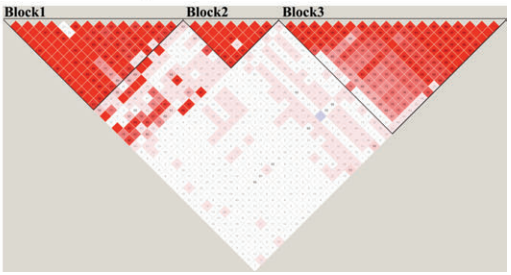


Fig. 1. Gene maps, haplotypes and LDs among *RIZ* SNPs. Coding exons are marked by black blocks, and 5′ and 3′ UTRs by white blocks. The first base of the translational start site is denoted as nucleotide +1. Asterisks indicate the SNPs that were used for the statistical analysis. (**A**) SNPs genotyped in the study subjects. The frequencies of the SNPs were based on findings in 670 subjects (335 lung cancer patients and 335 normal controls). (**B**) Haplotypes of *RIZ*. Haplotypes with a frequency >0.05 are presented. (**C**) Three LD blocks in *RIZ* using Haploview.

Table II. Subgroup analysis of *RIZ* polymorphisms associated with the risk of lung cancer

Loci	Position	rs#	Minor allele frequency			Lung cancer versus controls		Adenocarcinoma versus controls	
			Controls (n = 322)	Lung cancer (n = 331)	Adenocarcinoma (n = 187)	OR (95% CI)	P	OR (95% CI)	P
Block 1									
−10369C>G	Exon 1	rs2495061	0.207	0.215	0.195	0.96 (0.73–1.27)	0.76	0.92 (0.66–1.29)	0.63
−10286C>T	Exon 1	rs17393663	0.149	0.140	0.142	0.94 (0.69–1.28)	0.68	0.93 (0.65–1.33)	0.69
−8758A>G	Intron 1	rs6690270	0.478	0.459	0.489	0.97 (0.77–1.21)	0.78	1.05 (0.81–1.36)	0.72
−7509C>A	Intron 1	rs6693939	0.214	0.221	0.195	0.95 (0.72–1.25)	0.70	0.88 (0.63–1.22)	0.44
+3421A>C	Intron 2	rs12755924	0.110	0.134	0.134	1.27 (0.90−1.79)	0.18	1.24 (0.83−1.85)	0.29
+11374T>C	Intron 2	rs16852866	0.227	0.240	0.214	1.00 (0.76–1.30)	0.97	0.92 (0.67–1.27)	0.61
+15158G>A	Intron 2	rs2277	0.222	0.236	0.214	1.00 (0.76−1.30)	0.98	0.95 (0.69−1.31)	0.76
+21910G>C	Intron 4	rs2294484	0.188	0.177	0.168	0.94 (0.71–1.24)	0.65	0.88 (0.63–1.23)	0.47
+50765A>C	Intron 6	rs1406417	0.193	0.190	0.184	0.99 (0.75−1.31)	0.95	0.95 (0.69−1.32)	0.76
+54320A>G	Intron 6	rs2359756	0.489	0.465	0.495	0.96 (0.77−1.20)	0.70	1.03 (0.79−1.33)	0.85
+54535A>T	Intron 6	rs2884788	0.186	0.186	0.182	1.00 (0.75–1.32)	0.98	0.98 (0.71–1.36)	0.89
+57798A>G	Intron 7	rs1203677	0.405	0.415	0.384	0.98 (0.78−1.23)	0.87	0.92 (0.70−1.19)	0.51
+63198G>A	Exon 8	rs1203678	0.410	0.414	0.390	0.97 (0.77–1.22)	0.79	0.93 (0.71–1.21)	0.57
+67014A>G	Exon 8	rs1203651	0.416	0.423	0.393	0.97 (0.78−1.22)	0.81	0.91 (0.70−1.18)	0.48
+78332C>T	Intron 8	rs1203639	0.322	0.339	0.307	1.03 (0.80−1.32)	0.82	0.94 (0.70−1.25)	0.65
Block 2									
+79792G>T	Intron 8	rs1203638	0.236	0.219	0.211	0.85 (0.65−1.12)	0.25	0.85 (0.62–1.16)	0.30
+80056A>T	Intron 8	rs11807320	0.228	0.236	0.249	1.03 (0.79–1.35)	0.83	1.13 (0.83–1.54)	0.45
+82010A>G	Intron 8	rs2281169	0.430	0.423	0.439	1.01 (0.80–1.27)	0.94	1.05 (0.81–1.36)	0.71
+82423C>T	Intron 8	rs16852988	0.056	0.082	0.070	1.64 (1.03−2.60)	0.04	1.30 (0.76−2.22)	0.35
+83127C>T	Intron 8	rs2744692	0.165	0.156	0.163	0.94 (0.70–1.28)	0.71	0.96 (0.68–1.37)	0.83
+84261A>T	Intron 8	rs742355	0.281	0.273	0.278	0.96 (0.75−1.23)	0.73	0.98 (0.74−1.31)	0.90
+84353T>C	Intron 8	rs2014788	0.452	0.456	0.447	1.00 (0.80–1.25)	0.98	0.95 (0.74–1.24)	0.72
+85543G>A	Intron 8	rs2281168	0.262	0.272	0.278	1.10 (0.85−1.42)	0.48	1.13 (0.84−1.52)	0.41
Block 3									
+88528A>G	Intron 8	rs2281164	0.230	0.220	0.223	1.03 (0.78–1.35)	0.86	0.96 (0.70–1.31)	0.79
+88680C>T	Intron 8	rs1203673	0.247	0.215	0.190	0.82 (0.63–1.07)	0.14	0.69 (0.50–0.96)	0.03
+91319C>G	Intron 8	rs2697985	0.230	0.196	0.166	0.80 (0.61−1.05)	0.11	0.65 (0.46−0.91)	0.01
+92054T>G	Intron 8	rs2244634	0.234	0.196	0.166	0.78 (0.59–1.03)	0.07	0.63 (0.45–0.88)	0.007
+92337G>A	Intron 8	rs2281161	0.213	0.172	0.134	0.74 (0.55–0.98)	0.04	**0.55 (0.38–0.78)**	**0.001**
+95701C>A	Intron 8	rs2744690	0.234	0.184	0.147	0.72 (0.54–0.95)	0.02	**0.54 (0.38–0.77)**	**0.0006**
+96393A>G	Intron 8	rs2744688	0.307	0.260	0.227	0.81 (0.63−1.04)	0.10	0.63 (0.47−0.86)	0.003
+96713G>T	Intron 8	rs2744687	0.161	0.145	0.136	0.91 (0.66–1.25)	0.55	0.78 (0.54–1.14)	0.20
+97096C>T	Intron 8	rs2245197	0.446	0.427	0.401	0.96 (0.77–1.20)	0.72	0.83 (0.64–1.08)	0.17
+97583C>A	Intron 8	rs2245213	0.304	0.264	0.249	0.85 (0.67–1.09)	0.19	0.74 (0.55–0.99)	0.04
+98357G>A	Intron 8	rs2235516	0.292	0.258	0.235	0.86 (0.67−1.11)	0.24	0.73 (0.54–0.99)	0.04
+98633G>A	Intron 8	rs7551586	0.095	0.100	0.099	1.07 (0.72–1.59)	0.73	1.09 (0.69–1.71)	0.71
+99244C>T	Intron 8	rs2235515	0.450	0.418	0.401	0.88 (0.70−1.11)	0.28	0.82 (0.63−1.06)	0.13
+100903G>A	Exon 9	rs1046331	0.300	0.264	0.235	0.84 (0.66–1.08)	0.17	0.71 (0.53–0.95)	0.02
+103814G>A	Intron 9	rs1810474	0.388	0.353	0.329	0.87 (0.68−1.10)	0.23	0.76 (0.58−1.00)	0.05
+107149G>A	Intron 9	rs2744682	0.295	0.266	0.243	0.88 (0.68–1.13)	0.30	0.76 (0.57–1.01)	0.06
+107405A>C	Intron 9	rs2697970	0.391	0.363	0.340	0.88 (0.70–1.12)	0.31	0.79 (0.60–1.04)	0.09
+107932T>C	Exon 10	rs2697967	0.443	0.409	0.388	0.87 (0.69−1.10)	0.25	0.79 (0.61–1.03)	0.08
+108640T>G	Exon 10	rs2697963	0.194	0.168	0.144	0.84 (0.62–1.13)	0.24	0.69 (0.48–0.99)	0.04

The ORs (95% CI) and corresponding P values shown were estimated from a log additive model using unconditional multiple logistic regression, controlling for family history, smoking status and pack-years as covariates. P values under 0.001 are indicated in bold strokes.

(the codominant, log additive, dominant and recessive model) (Table III). From the log additive model, subjects carrying a variant allele of *+92337G>A* and *+95701C>A* showed significantly reduced risks for lung adenocarcinoma after correcting P values for multiple testing of 42 SNPs (aOR = 0.55, corrected P = 0.04; aOR = 0.54, corrected P = 0.02, respectively).

From 19 SNPs of the *RIZ* gene form the LD block 3, five common haplotypes (>5% frequency) were identified with an accumulated frequency of 69.5% in the controls (Figure 1B). The common haplotypes were labeled and categorized as Ht1–Ht5 based on their computer-estimated frequency, and other rare haplotypes were grouped together in the analyses as shown in Table IV. Subjects carrying a copy of Ht5 haplotype showed a 50% decreased risk of lung cancer (aOR = 0.50, 95% CI = 0.30–0.82). Although the haplotype frequencies were low, Ht5 haplotype of block 3 also showed a significantly decreased risk of lung adenocarcinoma when compared with the normal controls (aOR = 0.28, 95% CI = 0.13–0.58) (Table IV). Whereas Ht1 represents that all 19 SNPs were the wild-type allele, Ht5 represents that only three SNPs were wild-type allele. The haplotypes of LD block 1 or block 2 did not yield any statistically significant association.

Discussion

Our case–control study revealed that the *RIZ* polymorphisms showed a significant association with lung cancer, particularly with a lower risk of lung adenocarcinomas. The retinoblastoma protein-interacting zinc finger gene *RIZ* is a member of a nuclear protein-methyltransferase superfamily and is also known to play an important role in a variety of

Table III. Association of two polymorphisms of *RIZ* and lung cancer risk according to histologic cell type in lung cancer patients

SNP	Genotype	Frequency		Codominant model			Log additive model			Dominant model			Recessive model		
		Controls, *N* (%)	LC, *N* (%)	OR (95% CI)	*P*	Corrected *P*	OR (95% CI)	*P*	Corrected *P*	OR (95% CI)	*P*	Corrected *P*	OR (95% CI)	*P*	Corrected *P*
Overall															
+92337G>A	*GG*	199 (61.8)	228 (68.9)	1			1			1			1		
	AG	109 (33.9)	92 (27.8)	0.75 (0.53–1.06)	0.1	ns	0.74 (0.55–0.98)	0.04	ns	0.72 (0.51–1.01)	0.05	ns			
	AA	14 (4.3)	11 (3.3)	0.51 (0.21–1.23)	0.13	ns							0.56 (0.24–1.34)	0.19	ns
+95701C>A	*CC*	188 (58.4)	221 (66.8)	1			1			1			1		
	AC	117 (36.3)	98 (29.6)	0.73 (0.52–1.03)	0.08	ns	0.72 (0.54–0.95)	0.02	ns	0.70 (0.50–0.97)	0.03	ns			
	AA	17 (5.3)	12 (3.6)	0.49 (0.22–1.10)	0.08	ns							0.54 (0.24–1.21)	0.14	ns
Adenocarcinoma															
+92337G>A	*GG*	199 (61.8)	140 (74.8)	1			1			1			1		
	AG	109 (33.9)	44 (23.5)	0.55 (0.36–0.84)	0.006	ns	**0.55 (0.38–0.78)**	**0.001**	**0.04**	0.52 (0.35–0.78)	0.002	0.06			
	AA	14 (4.3)	3 (1.6)	0.28 (0.08–0.98)	0.05	ns							0.33 (0.09–1.16)	0.08	ns
+95701C>A	*CC*	188 (58.4)	136 (72.7)	1			1			1			1		
	AC	117 (36.3)	47 (25.1)	0.54 (0.36–0.82)	0.003	ns	**0.54 (0.38–0.77)**	**0.0006**	**0.02**	**0.51 (0.34–0.76)**	**0.0008**	**0.03**			
	AA	17 (5.3)	4 (2.1)	0.30 (0.10–0.90)	0.03	ns							0.36 (0.12–1.09)	0.07	ns
Squamous															
+92337G>A	*GG*	199 (61.8)	44 (57.9)	1			1			1			1		
	AG	109 (33.9)	29 (38.2)	1.33 (0.73–2.44)	0.35	ns	1.09 (0.67–1.79)	0.73	ns	1.24 (0.69–2.22)	0.48	ns			
	AA	14 (4.3)	3 (3.9)	0.66 (0.14–2.93)	0.58	ns							0.59 (0.14–2.59)	0.48	ns
+95701C>A	*CC*	188 (58.4)	43 (56.6)	1			1			1			1		
	AC	117 (36.3)	30 (39.5)	1.33 (0.73–2.43)	0.35	ns	1.08 (0.66–1.76)	0.77	ns	1.23 (0.69–2.20)	0.49	ns			
	AA	17 (5.3)	3 (3.9)	0.63 (0.14–2.75)	0.54	ns							0.56 (0.13–2.41)	0.44	ns

LC: lung cancer cases; ns: *P* value is not significant at 5% significance level; corrected *P* value is obtained by the correction for multiple testing; The ORs (95% CI) and corresponding *P* values were derived from a logistic analysis controlling for smoking status, pack-years and family history as covariates. Odds ratios that are statistically significant (corrected $P < 0.05$) are indicated in bold strokes.

Table IV. Association analysis among haplotypes of *RIZ* LD block 3 and lung cancer risk according to histologic cell type in lung cancer patients

Haplotype		Frequency *N*, (%)				Lung cancer versus controls		Adenocarcinoma versus controls		Squamous versus controls	
		Controls ($n = 322$)	Lung cancer ($n = 331$)	Adenocarcinoma ($n = 187$)	Squamous ($n = 76$)	OR (95% CI)	*P*	OR (95% CI)	*P*	OR (95% CI)	*P*
Ht1	0 copy	149 (46.3)	139 (42.0)	73 (39.0)	37 (48.7)	1		1		1	
	1 copy	136 (42.2)	143 (43.2)	82 (43.9)	30 (39.5)	1.16 (0.82–1.63)	0.40	1.24 (0.83–1.84)	0.29	0.96 (0.52–1.78)	0.90
	2 copies	37 (11.5)	49 (14.8)	32 (17.1)	9 (11.8)	1.36 (0.82–2.25)	0.23	1.83 (1.05–3.20)	0.03	0.78 (0.31–1.96)	0.59
Ht2	0 copy	239 (74.2)	251 (75.8)	143 (76.5)	59 (77.6)	1		1		1	
	1 copy	75 (23.3)	73 (22.1)	42 (22.5)	16 (21.1)	0.90 (0.62–1.32)	0.60	0.93 (0.60–1.44)	0.74	0.79 (0.40–1.57)	0.50
	2 copies	8 (2.5)	7 (2.1)	2 (1.1)	1 (1.3)	0.84 (0.29–2.48)	0.75	0.42 (0.09–2.02)	0.28	0.76 (0.08–7.54)	0.81
Ht3	0 copy	271 (84.2)	285 (86.1)	162 (86.6)	67 (88.2)	1		1		1	
	1 copy	47 (14.6)	45 (13.6)	25 (13.4)	9 (11.8)	0.95 (0.60−1.51)	0.84	0.86 (0.51−1.47)	0.59	0.93 (0.39−2.21)	0.87
	2 copies	4 (1.2)	1 (0.3)	0	0	0.24 (0.03–2.25)	0.21	—	0.98	—	0.99
Ht4	0 copy	277 (86.0)	279 (84.3)	159 (85.0)	60 (78.9)	1		1		1	
	1 copy	45 (14.0)	52 (15.7)	28 (15.0)	16 (21.1)	1.14 (0.73–1.79)	0.57	1.15 (0.68–1.93)	0.61	1.41 (0.68–2.89)	0.36
	2 copies	0	0	0	0	—	—	—	—	—	—
Ht5	0 copy	273 (84.8)	301 (90.9)	178 (95.2)	65 (85.5)	1		1		1	
	1 copy	49 (15.2)	29 (8.8)	9 (4.8)	11 (14.5)	**0.50 (0.30–0.82)**	**0.007**	**0.28 (0.13-0.58)**	**0.0007**	0.78 (0.34−1.78)	0.56
	2 copies	0	1 (0.3)	0	0	—	—	—	—	—	—
Others	0 copy	164 (50.9)	163 (49.2)	89 (47.6)	35 (46.1)	1	—	1	—	1	—
	1 copy	122 (37.9)	132 (39.9)	76 (40.6)	32 (42.1)	1.18 (0.84–1.66)	0.34	1.12 (0.76–1.66)	0.56	1.72 (0.92–3.22)	0.09
	2 copies	36 (11.2)	36 (10.9)	22 (11.8)	9 (11.8)	1.03 (0.61–1.75)	0.90	1.12 (0.62–2.04)	0.71	1.16 (0.45–2.96)	0.68

The ORs (95% CI) and corresponding *P* values were derived from logistic analysis controlling for smoking status, pack-year and family history as covariates.

cancers including lung cancer (12,15). Frequent epigenetic inactivation of by promoter hypermethylation and frameshift mutations that result in a truncated PR-interacting domain generated the inactivation of the *RIZ* gene in many cancers (21–24). The functional importance and genetic alteration of *RIZ* in human cancers make it feasible to investigate the association between *RIZ* polymorphisms and lung cancer risk.

We studied 120 SNPs of the *RIZ* gene that were successfully genotyped in 335 Korean lung cancer patients and 335 healthy controls. The controls recruited among the cancer screenees at our hospital showed a higher prevalence of a family history of cancer than the lung cancer cases. Therefore, the association between the SNPs and the risk of lung cancer was analyzed by controlling for the family history of cancer, as well as other potential confounding variables such as smoking status and pack-years. Furthermore, to correct for the inflated false-positive (type I error) rate in multiple testing of SNPs, we applied a statistical correction method based on the SpD of matrices of the pairwise LD between SNPs. After the correction, *+92337G>A* and *+95701C>A* still showed a statistically significant association with a lower risk of lung adenocarcinomas. Furthermore, the analysis of frequent haplotypes of *RIZ* revealed that Ht5 of LD block 3 that included these two polymorphisms showed a significant protective effect on lung cancer risk. Although we found additional interesting polymorphisms associated with the lung cancer risk in LD block 3, they were no longer significant when the correction for multiple testing was applied.

As *+92337G>A* and *+95701C>A* are located on intron of the longest transcript variant, *RIZ1* but not *RIZ2*, the variant allele of these polymorphisms could be associated with tumor suppressive function of RIZ1. Polymorphisms in introns may also impact the gene function by affecting the splice donor–acceptor site, or regions nearby as well as regulatory motifs within the introns. Further biological and/or functional evidence is needed to confirm the genetic effects of *RIZ* polymorphisms on lung cancer.

The reduced lung cancer risk associated with these two polymorphisms is more apparent in adenocarcinoma group than in squamous cell carcinoma group in this case–control study. Although we have no clear answer to explain the association of RIZ polymorphisms and adenocarcinoma, we suspect that it may be related to different demographic features of adenocarcinoma and squamous cell carcinoma as shown in Table I. Adenocarcinoma group has higher prevalence to female and never smoker than squamous cell carcinoma group. More than 90% of squamous cell carcinoma patients are ever smokers. While there are differences in demographics and smoking histories between the two groups, the possible association between squamous cell carcinoma and RIZ polymorphisms cannot be ruled out. As the number of patients with squamous cell carcinoma ($n = 77$) is much smaller than that of adenocarcinoma patients ($n = 189$), the effect of RIZ polymorphisms on squamous cell carcinoma can be examined by additional studies with larger sample sizes.

As *RIZ* was known as a tumor suppressor gene that silenced in many cancers, we hypothesize that *RIZ* polymorphisms may show a strong association with cancer susceptibility not only in lung cancer but also in other cancers. Further epidemiologic studies in larger populations are required to test our hypothesis.

Despite a protective effect of the *RIZ* polymorphisms, our study has the following limitations: first, this is a hospital-based case–control study, where the controls were recruited from among those that visited the hospital for a cancer-screening program. Therefore, it is not surprising that the prevalence of a family history of cancer is higher in the control group than in the case group. Second, this study only considers a Korean population that may limit the application of these findings to other ethnic populations.

In conclusion, this is the first study to show a significant association between polymorphisms of the *RIZ* gene and lung cancer risk, particularly the risk of adenocarcinomas. These results suggest that the presence of the variant allele in *RIZ* may be a protective factor for the development of lung cancer and could be an important marker of genetic susceptibility to lung cancer.

Funding

The National Cancer Center Research Grant (0410110-3 to J.S.L.; 0110250 to H.-R.S)

Acknowledgements

We thank to Jin Hee Kim and Hyea-Jin Gil (National Cancer Center) for their excellent technical assistance.

Conflict of Interest Statement: None declared.

References

1. Miller,M.C.,3rd, *et al*. (2001) Genetic variability in susceptibility and response to toxicants. *Toxicol. Lett.*, **120**, 269–280.
2. Ingelman-Sundberg,M. (2001) Genetic variability in susceptibility and response to toxicants. *Toxicol. Lett.*, **120**, 259–268.
3. Larsen,J.E., *et al*. (2006) CYP1A1 Ile462Val and MPO G-463A interact to increase risk of adenocarcinoma but not squamous cell carcinoma of the lung. *Carcinogenesis*, **27**, 525–532.
4. Wenzlaff,A.S. *et al*. (2005) CYP1A1 and CYP1B1 polymorphisms and risk of lung cancer among never smokers: a population-based study. *Carcinogenesis*, **26**, 2207–2212.
5. Wenzlaff,A.S. *et al*. (2005) GSTM1, GSTT1 and GSTP1 polymorphisms, environmental tobacco smoke exposure and risk of lung cancer among never smokers: a population-based study. *Carcinogenesis*, **26**, 395–401.
6. Le Marchand,L. *et al*. (2002) Association of the hOGG1 Ser326Cys polymorphism with lung cancer risk. *Cancer Epidemiol. Biomarkers Prev.*, **11**, 409–412.
7. Ito,H. *et al*. (2004) Gene-environment interactions between the smoking habit and polymorphisms in the DNA repair genes, APE1 Asp148Glu and XRCC1 Arg399Gln, in Japanese lung cancer risk. *Carcinogenesis*, **25**, 1395–1401.
8. Tsuge,M. *et al*. (2005) A variable number of tandem repeats polymorphism in an E2F-1 binding element in the 5′ flanking region of SMYD3 is a risk factor for human cancers. *Nat. Genet.*, **37**, 1104–1107.
9. Cebrian,A. *et al*. (2006) Genetic variants in epigenetic genes and breast cancer risk. *Carcinogenesis*, **27**, 1661–1669.
10. Yoon,K.A. *et al*. (2006) Novel polymorphisms in the SUV39H2 histone methyltransferase and the risk of lung cancer. *Carcinogenesis*, **27**, 2217–2222.
11. Rea,S. *et al*. (2000) Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature*, **406**, 593–599.
12. Huang,S. *et al*. (1998) The PR domain of the Rb-binding zinc finger protein RIZ1 is a protein binding interface and is related to the SET domain functioning in chromatin-mediated gene expression. *J. Biol. Chem.*, **273**, 15933–15939.
13. Buyse,I.M. *et al*. (1995) The retinoblastoma protein binds to RIZ, a zinc-finger protein that shares an epitope with the adenovirus E1A protein. *Proc. Natl Acad. Sci. USA*, **92**, 4467–4471.
14. Buyse,I.M. *et al*. (1996) Physical mapping of the retinoblastoma interacting zinc finger gene RIZ to D1S228 on chromosome 1p36. *Genomics*, **34**, 119–121.
15. Huang,S. (1999) The retinoblastoma protein-interacting zinc finger gene RIZ in 1p36-linked cancers. *Front. Biosci.*, **4**, D528–D532.
16. Liu,L. *et al*. (1997) The retinoblastoma interacting zinc finger gene RIZ produces a PR domain-lacking product through an internal promoter. *J. Biol. Chem.*, **272**, 2984–2991.
17. He,L. *et al*. (1998) RIZ1, but not the alternative RIZ2 product of the same gene, is underexpressed in breast cancer, and forced RIZ1 expression causes G2-M cell cycle arrest and/or apoptosis. *Cancer Res.*, **58**, 4238–4244.
18. Jiang,G. *et al*. (1999) Decreased RIZ1 expression but not RIZ2 in hepatoma and suppression of hepatoma tumorigenicity by RIZ1. *Int. J. Cancer*, **83**, 541–546.
19. Chadwick,R.B. *et al*. (2000) Candidate tumor suppressor RIZ is frequently involved in colorectal carcinogenesis. *Proc. Natl Acad. Sci. USA*, **97**, 2662–2667.
20. Tokumaru,Y. *et al*. (2003) Biallelic inactivation of the RIZ1 gene in human gastric cancer. *Oncogene*, **22**, 6954–6958.
21. Du,Y. *et al*. (2001) Hypermethylation in human cancers of the RIZ1 tumor suppressor gene, a member of a histone/protein methyltransferase superfamily. *Cancer Res.*, **61**, 8094–8099.

22. Oshimo,Y. *et al*. (2004) Frequent epigenetic inactivation of RIZ1 by promoter hypermethylation in human gastric carcinoma. *Int. J. Cancer*, **110**, 212–218.
23. Sakurada,K. *et al*. (2001) RIZ, the retinoblastoma protein interacting zinc finger gene, is mutated in genetically unstable cancers of the pancreas, stomach, and colorectum. *Genes Chromosomes Cancer*, **30**, 207–211.
24. Piao,Z. *et al*. (2000) Frequent frameshift mutations of RIZ in sporadic gastrointestinal and endometrial carcinomas with microsatellite instability. *Cancer Res.*, **60**, 4701–4704.
25. Pan,K.F. *et al*. (2004) Detection of frameshift mutations of RIZ in gastric cancers with microsatellite instability. *World J. Gastroenterol.*, **10**, 2719–2722.
26. Steele-Perkins,G. *et al*. (2001) Tumor formation and inactivation of RIZ1, an Rb-binding member of a nuclear protein-methyltransferase superfamily. *Genes Dev.*, **15**, 2250–2262.
27. Oliphant,A. *et al*. (2002) BeadArray technology: enabling an accurate, cost-effective approach to high-throughput genotyping. *Biotechniques*, Suppl: 56–8 60–61.
28. Hedrick,P.W. (1987) Gametic disequilibrium measures: proceed with caution. *Genetics*, **117**, 331–341.
29. Barrett,J.C. *et al*. (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*, **21**, 263–265.
30. Gabriel,S.B. *et al*. (2002) The structure of haplotype blocks in the human genome. *Science*, **296**, 2225–2229.
31. Stephens,M. *et al*. (2001) A new statistical method for haplotype reconstruction from population data. *Am. J. Hum. Genet.*, **68**, 978–989.
32. Nyholt,D.R. (2004) A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am. J. Hum. Genet.*, **74**, 765–769.

Received February 5, 2007; revised June 27, 2007; accepted June 28, 2007