Polymorphisms in cancer-related pathway genes and lung cancer

Shin Yup Lee[1,2,10], Hyo-Gyoung Kang[3,4,10], Jin Eun Choi[3,4], Deuk Kju Jung[3,5], Won Kee Lee[6], Hyun Chul Lee[7], So Yeon Lee[2], Seung Soo Yoo[1,2], Jaehee Lee[2], Yangki Seok[8], Eung Bae Lee[8], Seung Ick Cha[2], Sukki Cho[9], Chang Ho Kim[2], Myung Hoon Lee[7] and Jae Yong Park[1,2,3]

Affiliations: [1]Lung Cancer Center, Kyungpook National University Medical Center, Daegu, Korea. [2]Dept of Internal Medicine, Kyungpook National University, Daegu, Korea. [3]Dept of Biochemistry and Cell Biology, Kyungpook National University, Daegu, Korea. [4]Dept of Cell and Matrix Research Institute, Kyungpook National University, Daegu, Korea. [5]BK21 Plus KNU Biomedical Convergence Program, Dept of Biomedical Science, Kyungpook National University, Daegu, Korea. [6]Dept of Biostatistics Center, Kyungpook National University, Daegu, Korea. [7]D&P Biotech, Inc., Daegu, Korea. [8]Dept of Thoracic Surgery, School of Medicine, Kyungpook National University, Daegu, Korea. [9]Dept of Thoracic and Cardiovascular Surgery, Seoul National University School of Medicine, Seoul, Korea. [10]Both authors contributed equally.

Correspondence: Jae Yong Park, Lung Cancer Center, Kyungpook National University Medical Center, 807, Hoguk-ro, Buk-gu, Daegu 702-210, Korea. E-mail: jaeyong@knu.ac.kr

ABSTRACT We evaluated the associations between potentially functional variants in a comprehensive list of cancer-related genes and lung cancer in a Korean population.

A total of 1969 potentially functional single nucleotide polymorphisms (SNPs) of 1151 genes involved in carcinogenesis were evaluated using an Affymetrix custom-made GeneChip in 610 nonsmall cell lung cancer patients and 610 healthy controls. A replication study was conducted in an independent set of 490 cases and 486 controls. 68 SNPs were significantly associated with lung cancer in the discovery set and tested for replication.

Among the 68 SNPs, three SNPs (corepressor interacting with RBPJ 1 (*CIR1*) rs13009079T>C, ribonucleotide reductase M1 (*RRM1*) rs1465952T>C and solute carrier family 38, member 4 (*SLC38A4*) rs2429467C>T) consistantly showed significant associations with lung cancer in the replication study. In combined analysis, adjusted odds ratio for *CIR1* rs13009079T>C, *RRM1* rs1465952T>C and *SLC38A4* rs2429467C>T were 0.69, 0.71 and 0.73, respectively ($p=4\times10^{-5}$, 0.01 and 0.001, respectively) under the dominant model. The relative mRNA expression level of *CIR1* was significantly associated with rs13009079T>C genotypes in normal lung tissues ($p_{trend}=0.03$).

These results suggest that the three SNPs, particularly *CIR1* rs13009079T>C, may play a role in the pathogenesis of lung cancer.

@ERSpublications
Single nucleotide polymorphisms in *CIR1*, *RRM1* and *SLC38A4* may play a role in the pathogenesis of lung cancer http://ow.ly/zwj43033L9j

This article has supplementary material available from erj.ersjournals.com

Received: Dec 04 2015 | Accepted after revision: July 01 2016 | First published online: Sept 01 2016

Support statement: This study was supported by the National R&D Program Programme for Cancer Control, Ministry of Health and Welfare (0720550-2).

Conflict of interest: None declared.

Introduction

Lung cancer, predominantly nonsmall cell lung cancer (NSCLC), is the leading cause of cancer deaths worldwide, with a 5-year survival rate of 16% [1]. Environmental exposure, primarily to cigarette smoke, is the major risk factor for lung cancer. It is well known that >90% of male and 75–85% of female lung cancer patients have a history of current or former smoking, suggesting that smoking is responsible for the majority of lung cancer [2]. However, among tobacco smokers only ~10–15% ultimately develop lung cancer, suggesting that host factors such as genetic susceptibility may play a role in the pathogenesis of lung cancer [3, 4]. Recently, genome-wide association studies (GWAS) have identified several chromosomal regions that contain genes associated with the risk of lung cancer [5–7], including 5p15.33, 6p21.33 and 15q25.1. However, there are major issues limiting the value of GWAS, such as low reproducibility, low predictability and unexplained heritability [8, 9]. Moreover, the biological mechanism underlying GWAS results remain unclear because it is likely that most of the variations identified in GWAS are not causal [8]. In fact, nearly 90% of the variants identified as phenotype-associated single nucleotide polymorphisms (SNPs) in GWAS have been located within intergenic or intronic regions, posing an obstacle to their interpretation [10, 11].

The evolution of normal cells to a neoplastic state is a characteristic multistep process with accumulating multiple genetic and epigenetic alterations, collectively resulting in the hallmarks of cancer: sustaining proliferative signalling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing angiogenesis, activating invasion and metastasis, and two other emerging hallmarks, *i.e.* reprogramming of energy metabolism and evading immune destruction [12]. Incipient cancer cells need to successively acquire these traits, enabling them to become tumorigenic and fully malignant, which could explain the multistep process of carcinogenesis. Although GWAS have identified several genetic susceptibility loci for lung cancer which were not previously anticipated to play a role in lung carcinogenesis, our expectation is that variants in genes known to participate in multiple steps of cancer development and progression could contribute to lung cancer susceptibility.

In our previous study [13], we reported that colony-stimulating factor 1 receptor (*CSF1R*) rs10079250A>G, tumor protein p63 (*TP63*) rs7631358G>A and corepressor interacting with RBPJ 1 (*CIR1*) rs13009079T>C may play roles in the susceptibility to lung cancer in never-smoking females, using a comprehensive list of potentially functional polymorphisms in cancer-related genes. In the present study, we performed a case–control study to identify genetic polymorphisms that are associated with lung cancer in a population comprising both males and females, and both smokers and never-smokers, using the same comprehensive collection of SNPs in cancer-related genes.

Materials and methods

Study populations

A discovery set comprised of a total of 610 NSCLC patients who were histologically confirmed at Kyungpook National University Hospital (KNUH), Daegu, Korea from January 2008 to June 2010. The control subjects (n=610) were randomly selected from a pool of healthy volunteers who visited the general health check-up centre at KNUH and matched with the cases for age and gender. This study was approved by the Institutional Review Board (IRB) of KNUH (KNUHBIO_09-1018) and written informed consent was obtained from all participants. Genomic DNA samples of cases and controls were provided by the National Biobank of Korea – KNUH (NBK), which is supported by the Ministry of Health, Welfare and Family Affairs of Korea. All materials derived from the NBK were obtained under IRB-approved protocols before any type of anticancer therapy, including systemic chemotherapy. Genomic DNA was extracted from peripheral blood lymphocytes using the QuickGene-810 system (Fujifilm, Tokyo, Japan). All case and control subjects were ethnic Koreans. A replication study was performed in an independent set of cases and control subjects. A total of 490 NSCLC patients who were diagnosed at KNUH from July 2010 to July 2012 were enrolled. Among the 486 controls for the replication study, genomic DNA samples from 135 subjects were provided by the NBK and 331 with matched data by the Korean Biobank Project (4851-307, KBP-2011-24) and the Korean Genome and Epidemiology study (4845-302) supported by the Korea Centers for Disease Control and Prevention.

Selection of polymorphisms and genotyping

We selected SNPs for the present study using public databases, as described previously [13, 14]. Briefly, we selected 1784 candidate genes involved in cancer-related pathways from the database of SABiosciences (www.sabiosciences.com/Cancer.php; online supplementary table S1). Using the dbSNP public database (www.ncbi.nlm.nih.gov/SNP) to select all the potentially functional SNPs, a total of 4215 SNPs with minor allele frequency ⩾5% in the HapMap-JPT (Japanese in Tokyo, Japan) data were collected for this study. Among those, 1969 SNPs of 1151 genes were genotyped using an Affymetrix custom-made GeneChip because other SNPs could not be applied to the platform. The captured and genotyped SNPs in the

discovery set are listed in online supplementary table S1. For validation, 68 SNPs with p<0.05 for genotype distribution in the discovery set were tested in the replication study (online supplementary table S2) using Sequenom's MassARRAY iPLEX assay (Sequenom, San Diego, CA, USA). For quality control, the genotyping analysis was performed "blind" with respect to the case/control status. Approximately 5% of samples in the validation set were randomly selected to be genotyped again. The results were 100% concordant.

RNA preparation and quantitative reverse transcription-PCR

CIR1, *RRM1* and Hairy and enhancer of split homolog-1 (*HES1*) mRNA expression levels were examined by quantitative reverse transcription PCR. Total RNA from tumour and paired nonmalignant lung tissues (n=107) was isolated using TRIzol (Invitrogen, Carlsbad, CA, USA). Real-time PCR with SYBR Green detection was performed using a LightCycler 480 (Roche Applied Science, Basel, Switzerland) with QuantiFast SYBR Green PCR Master Mix (Qiagen, Hilden, Germany). The following sets of primers were used to conduct the real-time PCR for *CIR1*, *RRM1* and the β-actin gene: *CIR1* forward, 5′-AGATCAGCCCTTTGGTATTCAG-3′; *CIR1* reverse, 5′-CCGAACTTGCATTGATTCCAG-3′; *RRM1* forward, 5′-ACCGCCCACAACTTTCTAG-3′; *RRM1* reverse, 5′-CCAGTAGCCCGAATACAACTC-3′; *HES1* forward, 5′-AACGCAGTGTCACCTTCC-3′; *HES1* reverse, 5′-TCAAGTTCCTGTTTAGAGTCCG-3′; β-actin gene forward, 5′-ATGCTATCACCTCCCCTGTGT-3′; β-actin gene reverse, 5′-TTGTTACAGGAAGTCCCTTGCC-3′. Each sample was run in duplicate. The relative *CIR1*, *RRM1* and *HES1* mRNA expression levels were normalised with β-actin expression and then calculated by the $2^{-\Delta\Delta C_t}$ method [15].

Statistical analysis

The cases and controls were compared using the t-test and the Chi-squared test, and for continuous variables and categorical variables, respectively. The Hardy–Weinberg equilibrium was tested by comparing observed and expected genotype frequencies using a Chi-squared test with 1 degree of freedom. Unconditional logistic regression analysis was used to calculate odds ratios (ORs) and 95% confidence intervals, with adjustment for possible confounding factors (gender as a nominal variable, and age and pack-years of smoking as continuous variables). Differences were considered statistically significant at p<0.05. All the analyses were performed using SAS for Windows version 9.2 (SAS Institute, Cary, NC, USA).

Results

Patient characteristics and clinical predictors

The clinical characteristics of case and control subjects in the discovery and validation sets are given in table 1. Age and gender were not significantly different between cases and controls in both sets. However, cases included a significantly higher fraction of smokers compared with controls in both sets, which was adjusted for in the multivariate logistic regression analysis. The discovery set included 33.4% squamous cell carcinoma and 47.9% adenocarcinoma, and the validation set 40.8% and 43.9%, respectively.

Associations between polymorphisms and the risk of lung cancer

Among the 1969 SNPs evaluated in the discovery set, 668 were excluded due to following reasons: 1) 86 with failed genotyping, 2) 283 with genotype call rate <95%, 3) 201 with minor allele frequency <5% or 4) 98 with deviation from Hardy–Weinberg equilibrium (p<0.05) in the controls. Accordingly, 1301 SNPs in 863 genes were analysed for the association study. Approximately 48% of all SNPs were located in the promoter region, 24% in exons (nonsynonymous SNPs), 16% in exon/intron boundaries, 6% in 5′-untranslated regions and 6% in 3′-untranslated regions.

Of the 1301 SNPs evaluated in the discovery set, 68 SNPs were significantly associated with lung cancer (p<0.05), as shown in online supplementary table S2, and tested for a replication study in an independent set of 490 cases and 486 controls. In the validation set, three SNPs (*CIR1* rs13009079T>C, ribonucleotide reductase M1 (*RRM1*) rs1465952T>C and solute carrier family 38, member 4 (*SLC38A4*) rs2429467C>T) consistantly showed significant associations with lung cancer in the same direction as the discovery set when adjusted for age, sex and smoking (table 2). In a combined analysis of discovery and replication cohorts, *CIR1* rs13009079T>C, *RRM1* rs1465952T>C and *SLC38A4* rs2429467C>T were associated with a significantly decreased risk of lung cancer (adjusted OR 0.69, 95% CI 0.58–0.83, $p=4\times10^{-5}$; adjusted OR 0.71, 95% CI 0.55–0.92, p=0.01 and adjusted OR 0.73, 95% CI 0.62–0.87, p=0.001, respectively, under the dominant model). In 114 adenocarcinoma patients whose epidermal growth factor (*EGFR*) mutation test results were available (positive in 24.6% cases tested by direct sequencing), *EGFR* mutation status was not significantly associated with the genotypes of any of the three SNPs (data not shown). We did not observe any significant heterogeneity in ORs between the discovery and validation sets.

 DOI: 10.1183/13993003.02040-2015

TABLE 1 Characteristics of the study population

	Discovery			Validation			Combined		
	Cases	Controls	p-value	Cases	Controls	p-value	Cases	Controls	p-value
Subjects	610	610		490	486		1100	1096	
Age years	61.4±8.7	60.9±7.9	0.35	61.4±8.6	60.8±6.7	0.20	61.4±8.7	60.8±7.4	0.12
Sex									
Male	381 (62.5)	381 (62.5)	1.00	357 (72.9)	355 (73.1)	0.95	738 (67.1)	736 (67.2)	0.98
Female	229 (37.5)	229 (37.5)		133 (27.1)	131 (26.9)		362 (32.9)	360 (32.8)	
Smoking status									
Current	289 (47.4)	187 (30.7)	4.4×10^{-9}	231 (47.1)	228 (46.9)	0.001	520 (47.3)	415 (37.9)	1.9×10^{-5}
Ex	116 (19.0)	178 (29.2)		151 (30.8)	109 (22.4)		267 (24.3)	287 (26.2)	
Never	205 (33.6)	245 (40.1)		108 (22.0)	149 (30.7)		313 (28.4)	394 (35.9)	
Pack-years[#]	38.6±19.3	30.9±16.1	2.7×10^{-9}	38.7±21.5	30.0±18.7	1.3×10^{-8}	38.7±20.4	30.5±17.4	2.1×10^{-16}
Histological type									
Squamous cell carcinoma	204 (33.4)			200 (40.8)			404 (36.7)		
Adenocarcinoma	292 (47.9)			215 (43.9)			507 (46.1)		
Large cell carcinoma	11 (1.8)			10 (2.0)			21 (1.9)		
Small cell carcinoma	94 (15.4)			54 (11.0)			148 (13.5)		
NSCLC/NOS	9 (1.5)			11 (2.2)			20 (1.8)		
Pathological stage									
I	129 (21.2)			107 (21.8)			236 (21.5)		
II	33 (5.4)			34 (6.9)			67 (6.1)		
III	213 (34.9)			148 (30.2)			361 (32.8)		
IV	235 (38.5)			201 (41.0)			436 (39.6)		

Data are presented as n, mean±SD or n (%), unless otherwise stated. NSCLC: nonsmall cell lung cancer; NOS: not otherwise specified. #: in current and former smokers.

Effect of SNPs on mRNA expression

To identify the functional effect of the SNPs, we evaluated the relationship between the genotypes of the *CIR1* rs13009079T>C and *RRM1* rs1465952T>C and mRNA expression in tumour and paired nonmalignant lung tissues. The *CIR1* expression level was significantly lower in tumour tissues than in nonmalignant tissues (p=0.02; figure 1a). In contrast, the *RRM1* expression level was significantly higher in tumour tissues than in nonmalignant tissues ($p=4\times10^{-10}$; figure 1b). The relative expression level of *CIR1* was significantly associated with rs13009079T>C genotypes in normal lung tissues ($p_{trend}=0.03$; figure 1c). However, the *RRM1* expression level among rs1465952T>C genotypes was not significantly different (data not shown). We further investigated whether *CIR1* mRNA expression was inversely correlated with the mRNA expression level of *HES1*, which is one of the most studied targets of the Notch signalling pathway, and plays important roles in the cell cycle, proliferation, differentiation, and survival and apoptosis in neuronal, endocrine, T-lymphocyte progenitors as well as diverse cancers [16]. In contrast to *CIR1* expression, the level of *HES1* expression was significantly higher in tumours than in normal lung tissues (p=0.0004; online supplementary figure S1). However, as shown in online supplementary figure S1, *CIR1* mRNA expression did not show significant inverse correlation with *HES1* mRNA expression (Pearson correlation coefficient=0.14, p=0.14). In addition, the *HES1* mRNA expression level was not significantly different according to *CIR1* rs13009079T>C genotype (data not shown).

Discussion

We evaluated 1301 SNPs in 863 candidate genes potentially involved in carcinogenesis to identify genetic variations associated with lung cancer using an Affymetrix custom-made GeneChip in a discovery set and replicated the result in a validation set. Three SNPs (*CIR1* rs13009079T>C, *RRM1* rs1465952T>C and *SLC38A4* rs2429467C>T) were consistently associated with lung cancer across both study sets. In addition, this study provides evidence that rs13009079T>C is a functional SNP that affects mRNA expression of the *CIR1* gene using clinical samples. These findings suggest that the three SNPs, particularly *CIR1* rs13009079T>C, may play a role in the pathogenesis of lung cancer.

Among the three SNPs (*CIR1* rs13009079T>C, *RRM1* rs1465952T>C and *SLC38A4* rs2429467C>T) significantly associated with lung cancer, *CIR1* rs13009079T>C showed the most significant association. Interestingly, we previously identified *CIR1* rs13009079T>C in lung cancer among never-smoking females [13]. The present study replicated the association in an extended set of subjects including both males and females, and both smokers and never-smokers. The mRNA expression using clinical samples further

TABLE 2 *CIR1*, *RRM1* and *SLC38A4* genotypes of cases and controls, and their associations with the risk of lung cancer

Genotype	Discovery					Validation						Combined				
	Cases	Controls	p-value[#]	Adjusted OR (95% CI)[¶]	p-value[¶]	Cases	Controls	p-value[#]	Adjusted OR (95% CI)[¶]	p-value[¶]	p_H-value[+]	Cases	Controls	p-value[#]	Adjusted OR (95% CI)[¶]	p-value[¶]
CIR1																
rs13009079																
TT	380 (62.3)	328 (53.8)	0.01	1.00		296 (63.2)	258 (54.9)	0.02	1.00			676 (62.7)	586 (54.3)	0.0004	1.00	
TC	200 (32.8)	238 (39.0)		0.72 (0.56–0.92)	0.01	142 (30.3)	184 (39.1)		0.68 (0.51–0.90)	0.01	0.76	342 (31.7)	422 (39.1)		0.70 (0.58–0.84)	0.0001
CC	30 (4.9)	44 (7.2)		0.53 (0.32–0.87)	0.01	30 (6.4)	28 (6.0)		0.89 (0.51–1.55)	0.67	0.17	60 (5.6)	72 (6.7)		0.67 (0.46–0.96)	0.03
Dominant			0.003	0.69 (0.54–0.87)	0.002			0.01	0.71 (0.54–0.92)	0.01	0.87			7×10^{-5}	0.69 (0.58–0.83)	4×10^{-5}
Recessive			0.09	0.60 (0.37–0.98)	0.04			0.40	1.02 (0.59–1.77)	0.93	0.16			0.29	0.77 (0.53–1.10)	0.14
Codominant			0.0004	0.72 (0.60–0.87)	0.001			0.05	0.80 (0.64–0.99)	0.04	0.47			0.0003	0.75 (0.65–0.87)	9×10^{-5}
RRM1																
rs1465952																
TT	534 (88.0)	510 (84.4)	0.04	1.00		430 (89.6)	409 (86.1)	0.25	1.00			964 (88.7)	919 (85.2)	0.01	1.00	
TC	67 (11.0)	92 (15.2)		0.66 (0.47–0.94)	0.02	49 (10.2)	65 (13.7)		0.66 (0.44–0.99)	0.04	1.00	116 (10.7)	157 (14.6)		0.68 (0.53–0.88)	0.004
CC	6 (1.0)	2 (0.3)		2.65 (0.52–13.41)	0.24	1 (0.2)	1 (0.2)		1.05 (0.07–16.92)	0.97	0.57	7 (0.6)	3 (0.3)		2.17 (0.56–8.48)	0.26
Dominant			0.07	0.71 (0.50–0.99)	0.04			0.10	0.67 (0.45–1.00)	0.05	0.83			0.02	0.71 (0.55–0.92)	0.01
Recessive			0.16	2.80 (0.55–14.15)	0.21			0.78	1.10 (0.07–17.65)	0.95	0.57			0.21	2.28 (0.58–8.89)	0.24
Codominant			0.01	0.77 (0.57–1.06)	0.11			0.11	0.68 (0.46–1.01)	0.06	0.63			0.04	0.76 (0.59–0.96)	0.02
SLC38A4																
rs2429467																
CC	377 (62.0)	339 (55.6)	0.03	1.00		282 (60.0)	244 (51.6)	0.02	1.00			659 (61.1)	583 (53.8)	0.003	1.00	
TC	198 (32.6)	221 (36.2)		0.78 (0.61–1.00)	0.05	157 (33.4)	200 (42.3)		0.69 (0.52–0.91)	0.01	0.52	355 (32.9)	421 (38.9)		0.74 (0.61–0.88)	0.001
TT	33 (5.4)	50 (8.2)		0.58 (0.36–0.94)	0.03	31 (6.6)	29 (6.1)		0.95 (0.55–1.64)	0.86	0.18	64 (5.9)	79 (7.3)		0.71 (0.50–1.02)	0.06
Dominant			0.02	0.75 (0.59–0.94)	0.01			0.01	0.72 (0.56–0.94)	0.01	0.82			0.001	0.73 (0.62–0.87)	0.001
Recessive			0.06	0.64 (0.40–1.02)	0.06			0.77	1.11 (0.65–1.88)	0.71	0.12			0.20	0.80 (0.57–1.13)	0.21
Codominant			0.003	0.77 (0.64–0.93)	0.01			0.05	0.82 (0.66–1.01)	0.07	0.66			0.001	0.79 (0.69–0.91)	0.001

Data are presented as n (%), unless otherwise stated. OR: odds ratio. [#]: two-sided Chi-squared test for genotype distribution between the cases and controls; [¶]: ORs (95% CIs) and p-values were calculated by unconditional logistic regression analysis, adjusted for age, sex and smoking status; [+]: Wald test for heterogeneity of adjusted ORs between the discovery and the validation sets.

DOI: 10.1183/13993003.02040-2015

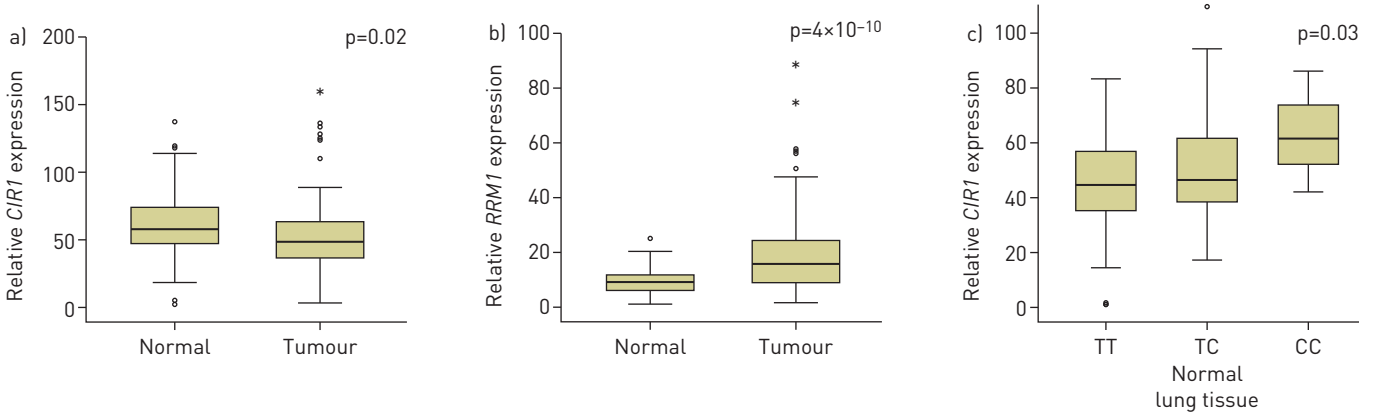

FIGURE 1 mRNA expression levels of a) *CIR1* and b) *RRM1* genes in normal lung and tumour tissues, and c) *CIR1* mRNA expression by rs13009079T>C genotype in normal lung tissues. *CIR1* mRNA expression levels and the associations with rs13009079T>C genotypes were examined in normal lung and paired tumour tissues (54 TT, 35 TC and five CC). mRNA expression level was normalised with that of the β-actin gene. The horizontal lines within the boxes represent the median values; the upper and lower boundaries of the boxes represent 75th and 25th percentiles, respectively; the upper and lower bars indicate the largest and smallest observed values, respectively, except outliers.

supported the functional property of *CIR1* rs13009079T>C. CIR1 is an evolutionarily conserved protein that is a component of the C promoter-binding factor 1 (CBF1, alias recombination signal binding protein for immunoglobulin kappa J region (RBPJ))-mediated transcription corepressor complex [17]. CIR1, with other components of the complex, binds to histone deacetylase (HDAC), participating in the recruitment of HDAC to DNA-bound CBF1 as a linker between CBF1 and the HDAC complex [17, 18]. The corepressor complex containing HDAC then leads to transcriptional repression of target genes. CBF1 is a central effector in Notch signalling [19, 20], which is dysregulated in many cancers. including lung cancer [21, 22]. Accumulating evidence has suggested that CBF1 represses Notch target gene expression in the absence of Notch signalling, but stimulation of the Notch receptor causes a switch in CBF1 function from a repressor to an activator [18, 19, 23]. Therefore, CIR1 participates in the transcriptional regulation of Notch target genes as a component of the CBF1-mediated corepressor complex. Based on the role of corepressors in the regulation of transcription, their altered expression or function may contribute to aberrant signalling, leading to dysregulated transcription of target genes, which has been associated with many cancers [24]. Although little is known about the role of CIR1 in the pathogenesis of cancer, altered CIR1 function may result in aberrant signalling pathways, including Notch signalling, leading to the development of cancer. In the present study, *CIR1* mRNA expression was significantly lower in tumour tissues compared with normal lung, which is consistent with mRNA sequencing data from The Cancer Genome Atlas database (cancergenome.nih.gov; data not shown), suggesting a potential tumour suppressor function of CIR1 in NSCLC. In agreement, the variant C allele, which was associated with a decreased risk of lung cancer, showed higher *CIR1* mRNA expression compared with the T allele in clinical samples. In contrast to *CIR1* expression, the level of *HES1* expression was significantly higher in tumours than in normal lung tissues. However, the inverse correlation between *CIR1* and *HES1* expression was not observed, probably because CIR1 does not directly inhibit *HES1* mRNA expression, but as a component of the CBF1 complex depending on Notch signalling. In addition, other pathways such as Hedgehog and Wnt signalling pathways also regulate *HES1* expression. It should be noted that a complicated cross-talk between the various signalling pathways and HES1, which is context- and cell-dependent, regulates *HES1* expression [16]. Further investigation is warranted to understand the role of CIR1 in the pathogenesis of lung cancer.

RRM1 plays an important role in DNA repair as well as its synthesis and also functions as a tumour suppressor gene, being a crucial determinant of the malignant tumour phenotype [25–28]. Therefore, altered expression of RRM1 may affect DNA repair capacity and tumour suppressor function, and thereby the risk of cancer. NSCLC patients with high RRM1 expression have been associated with significantly longer survival after surgery compared with those with low expression [28–30]. In contrast, high expression of RRM1 has been associated with reduced response and poor survival in patients with NSCLC treated with gemcitabine-containing chemotherapy [30–33], in line with the "double-edged sword" property of DNA repair genes. Genetic polymorphisms in *RRM1* have also been associated with the clinical outcome of lung cancer patients [34, 35]. In this study, *RRM1* rs1465952T>C was associated with a significantly decreased risk of lung cancer. It has been reported that *RRM1* rs1465952T>C is in the putative p53 response element and that the C allele is associated with stronger p53 binding than the T allele [36, 37], suggesting the C allele could lead to higher p53–DNA binding and transactivation capacity in response to DNA damage, which correlates with our finding of decreased risk of lung cancer. SLC38A4 transports both cationic and

neutral amino acids, and is found predominantly in liver. There is as yet no clear evidence that SLC38A4 has any role in carcinogenesis. Based on the PolyPhen algorithm [38], *SLC38A4* rs2429467C>T (Arg29Gly) might be a benign change. Further studies are needed to understand the biological function of *SLC38A4* in cancer, and to confirm the association between the SNP and the risk of lung cancer.

In this study, the association between *CIR1* rs13009079T>C, *RRM1* rs1465952T>C and *SLC38A4* rs2429467C>T and lung cancer was convincingly replicated in an independent set of patients, which would provide confidence in the results of genetic association studies [39, 40]. However, several limitations of our study should be considered. First, because all the subjects in the current study came from one country, our result needs to be verified in diverse populations with different ancestries. Second, although the three SNPs identified in this case–control study may contribute to genetic predisposition to lung carcinogenesis, they cannot be considered true genetic determinants of lung cancer as the analysis was performed on patients who had already been diagnosed with NSCLC. However, it would be impractical to perform a study for genetic determinants of lung cancer that involves sampling thousands of healthy subjects and then comparing the results between those who developed lung cancer in later years and those who remained healthy. Third, the three SNPs identified in our study may have not shown significant associations with lung cancer in other studies. The replication of such associations across studies often fails because of inadequate sample sizes with insufficient statistical power and/or imprecise estimation of the magnitude of the observed associations, different genetic structures (*i.e.* linkage disequilibrium) across populations or biased study designs [40]. A well-designed and properly powered study including diverse populations of different ancestries is warranted to validate our findings. Lastly, some associations were at marginal levels of statistical significance to avoid false-positive associations arising from multiple comparisons, probably because the modest sample size of both cohorts does not have optimal statistical power [39]. However, mRNA expression data supported the association between *CIR1* rs13009079T>C and lung cancer, although this study did not provide direct evidence of the role of CIR1 in the development of lung cancer.

In conclusion, this study showed that the three SNPs, particularly *CIR1* rs13009079T>C, were significantly associated with lung cancer, suggesting their potential role in the pathogenesis of lung cancer. Future studies are warranted regarding the biological role of those genes in the development and progression of lung cancer to understand the mechanism of association between the SNPs and lung cancer.

References

1 Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. *CA Cancer J Clin* 2013; 63: 11–30.
2 Subramanian J, Govindan R. Lung cancer in never smokers: a review. *J Clin Oncol* 2007; 25: 561–570.
3 Wu X, Zhao H, Suk R, *et al.* Genetic susceptibility to tobacco-related cancer. *Oncogene* 2004; 23: 6500–6523.
4 Kiyohara C, Otsu A, Shirakawa T, *et al.* Genetic polymorphisms and lung cancer susceptibility: a review. *Lung Cancer* 2002; 37: 241–256.
5 Hung RJ, McKay JD, Gaborieau V, *et al.* A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature* 2008; 452: 633–637.
6 McKay JD, Hung RJ, Gaborieau V, *et al.* Lung cancer susceptibility locus at 5p15.33. *Nat Genet* 2008; 40: 1404–1406.
7 Wang Y, Broderick P, Webb E, *et al.* Common 5p15.33 and 6p21.33 variants influence lung cancer risk. *Nat Genet* 2008; 40: 1407–1409.
8 Witte JS. Genome-wide association studies and beyond. *Annu Rev Public Health* 2010; 31: 9–20.
9 On beyond GWAS. *Nat Genet* 2010; 42: 551.
10 Boyle AP, Hong EL, Hariharan M, *et al.* Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* 2012; 22: 1790–1797.
11 Schaub MA, Boyle AP, Kundaje A, *et al.* Linking disease associations with regulatory information in the human genome. *Genome Res* 2012; 22: 1748–1759.
12 Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* 2011; 144: 646–674.
13 Kang HG, Lee SY, Jeon HS, *et al.* A functional polymorphism in *CSF1R* gene is a novel susceptibility marker for lung cancer among never-smoking females. *J Thorac Oncol* 2014; 9: 1647–1655.
14 Choi YY, Lee SY, Lee WK, *et al. RACK1* is a candidate gene associated with the prognosis of patients with early stage non-small cell lung cancer. *Oncotarget* 2015; 6: 4451–4466.
15 Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2 (-Delta Delta C(T)) Method. *Methods* 2001; 25: 402–408.
16 Rani A, Greenljaw R, Smith RA, *et al.* HES1 in immunity and cancer. *Cytokine Growth Factor Rev* 2016; 30: 113–117.
17 Hsieh JJ, Zhou S, Chen L, *et al.* CIR, a corepressor linking the DNA binding factor CBF1 to the histone deacetylase complex. *Proc Natl Acad Sci USA* 1999; 96: 23–28.
18 Lai EC. Keeping a good pathway down: transcriptional repression of Notch pathway target genes by CSL proteins. *EMBO Rep* 2002; 3: 840–845.
19 Hsieh JJ, Henkel T, Salmon P, *et al.* Truncated mammalian Notch1 activates CBF1/RBPJk-repressed genes by a mechanism resembling that of Epstein-Barr virus EBNA2. *Mol Cell Biol* 1996; 16: 952–959.
20 Kopan R, Ilagan MX. The canonical Notch signaling pathway: unfolding the activation mechanism. *Cell* 2009; 137: 216–233.
21 Ranganathan P, Weaver KL, Capobianco AJ. Notch signalling in solid tumours: a little bit of everything but not all the time. *Nat Rev Cancer* 2011; 11: 338–351.

DOI: 10.1183/13993003.02040-2015

22 Takebe N, Nguyen D, Yang SX. Targeting notch signaling pathway in cancer: clinical development advances and challenges. *Pharmacol Ther* 2014; 141: 140–149.
23 Lubman OY, Ilagan MX, Kopan R, *et al.* Quantitative dissection of the Notch:CSL interaction: insights into the Notch-mediated transcriptional switch. *J Mol Biol* 2007; 365: 577–589.
24 Vaiopoulos AG, Kostakis ID, Athanasoula K, *et al.* Targeting transcription factor corepressors in tumor cells. *Cell Mol Life Sci* 2012; 69: 1745–1753.
25 Fan H, Huang A, Villegas C, *et al.* The R1 component of mammalian ribonucleotide reductase has malignancy-suppressing activity as demonstrated by gene transfer experiments. *Proc Natl Acad Sci USA* 1997; 94: 13181–13186.
26 Gautam A, Li ZR, Bepler G. RRM1-induced metastasis suppression through PTEN-regulated pathways. *Oncogene* 2003; 22: 2135–2142.
27 Stubbe J. Ribonucleotide reductases in the twenty-first century. *Proc Natl Acad Sci USA* 1998; 95: 2723–2724.
28 Bepler G, Sharma S, Cantor A, *et al.* RRM1 and PTEN as prognostic parameters for overall and disease-free survival in patients with non-small-cell lung cancer. *J Clin Oncol* 2004; 22: 1878–1885.
29 Zheng Z, Chen T, Li X, *et al.* DNA synthesis and repair genes *RRM1* and *ERCC1* in lung cancer. *N Engl J Med* 2007; 356: 800–808.
30 Jordheim LP, Seve P, Tredan O, *et al.* The ribonucleotide reductase large subunit (RRM1) as a predictive factor in patients with cancer. *Lancet Oncol* 2011; 12: 693–702.
31 Bepler G, Sommers KE, Cantor A, *et al.* Clinical efficacy and predictive molecular markers of neoadjuvant gemcitabine and pemetrexed in resectable non-small cell lung cancer. *J Thorac Oncol* 2008; 3: 1112–1118.
32 Su C, Zhou S, Zhang L, *et al.* ERCC1, RRM1 and BRCA1 mRNA expression levels and clinical outcome of advanced non-small cell lung cancer. *Med Oncol* 2011; 28: 1411–1417.
33 Rosell R, Danenberg KD, Alberola V, *et al.* Ribonucleotide reductase messenger RNA expression and survival in gemcitabine/cisplatin-treated advanced non-small cell lung cancer patients. *Clin Cancer Res* 2004; 10: 1318–1325.
34 Bepler G, Zheng Z, Gautam A, *et al.* Ribonucleotide reductase M1 gene promoter activity, polymorphisms, population frequencies, and clinical relevance. *Lung Cancer* 2005; 47: 183–192.
35 Kim SO, Jeong JY, Kim MR, *et al.* Efficacy of gemcitabine in patients with non-small cell lung cancer according to promoter polymorphisms of the ribonucleotide reductase M1 gene. *Clin Cancer Res* 2008; 14: 3083–3088.
36 Bandele OJ, Wang X, Campbell MR, *et al.* Human single-nucleotide polymorphisms alter p53 sequence-specific binding at gene regulatory elements. *Nucleic Acids Res* 2011; 39: 178–189.
37 Noureddine MA, Menendez D, Campbell MR, *et al.* Probing the functional impact of sequence variation on p53-DNA interactions using a novel microsphere assay for protein-DNA binding with human cell extracts. *PLoS Genet* 2009; 5: e1000462.
38 Rebbeck TR, Spitz M, Wu X. Assessing the function of genetic variants in candidate gene association studies. *Nat Rev Genet* 2004; 5: 589–597.
39 Pharoah PD, Dunning AM, Ponder BA, *et al.* Association studies for finding cancer-susceptibility genetic variants. *Nat Rev Cancer* 2004; 4: 850–860.
40 Chanock SJ, Manolio T, Boehnke M, *et al.* Replicating genotype–phenotype associations. *Nature* 2007; 447: 655–660.