A Functional Genetic Variant in microRNA-196a2 Is Associated with Increased Susceptibility of Lung Cancer in Chinese

Tian Tian,[1,2] Yongqian Shu,[2,3] Jiaping Chen,[1,2] Zhibin Hu,[1,2] Lin Xu,[2,4] Guangfu Jin,[1] Jie Liang,[1] Ping Liu,[2,3] Xiaoyi Zhou,[1] Ruifen Miao,[1] Hongxia Ma,[1] Yijiang Chen,[2,3] and Hongbing Shen[1,2]

[1]Department of Epidemiology and Biostatistics; [2]Cancer Center; [3]Department of Oncology, The First Affiliated Hospital, Nanjing Medical University; [4]Department of Thoracic Surgery, Jiangsu Cancer Hospital, Nanjing, People's Republic of China

Abstract

microRNAs (miRNA) are a new class of non-protein-coding, small RNAs that function as tumor suppressors or oncogenes. They participate in diverse biological pathways and function as gene regulators. Recently, we conducted a survey of common single nucleotide polymorphisms (SNP) in miRNA sequences and reported that, among four SNPs (rs2910164, rs2292832, rs11614913, and rs3746444) in pre-miRNAs, rs11614913 in *miR-196a2* might affect mature *miR-196a* expression and target mRNA-binding activity and was significantly associated with non-small cell lung cancer survival. However, it remains largely unknown whether miRNA SNPs may alter lung cancer susceptibility. In the current study, we evaluated associations between the above four SNPs in pre-miRNAs and lung cancer susceptibility in a case-control study of 1,058 incident lung cancer patients and 1,035 cancer-free controls in a Chinese population. We found that *miR-196a2* rs11614913 variant homozygote CC was associated with ~25% significantly increased risk of lung cancer compared with their wild-type homozygote TT and heterozygote TC (odds ratio, 1.25; 95% confidence interval, 1.01-1.54). However, no significant effects were observed on the association between the other three SNPs and lung cancer risk. These findings suggest that functional SNP rs11614913 in *miR-196a2* could also contribute to lung cancer susceptibility. (Cancer Epidemiol Biomarkers Prev 2009;18(4):1183–7)

Introduction

Lung cancer is the leading cause of cancer-related deaths in the world and the incidence rate has been increasing significantly in the last two decades in China (1). Cigarette smoking is the most important risk factor for lung cancer and the genetic and epigenetic damage caused by tobacco smoke is primarily considered as one of the mechanisms of lung cancer development (2). Molecular epidemiologic studies showed that there were hundreds of genes involved in lung carcinogenesis (3, 4), such as *p53, Rb*, and *Ras* (2, 5). Although focusing on known genes might yield further understanding in lung cancer development, newly developed markers such as noncoding small RNAs may lead novel insight into the biological mechanism of lung cancer (6).

microRNAs (miRNA) are 21- to 24-nucleotide-long small noncoding RNA gene products that regulate gene expression by base pairing with target mRNAs at the 3′-untranslated region, leading to mRNA cleavage or translational repression (7-9). It has been suggested that miRNAs are involved in various biological processes, including cell proliferation, cell death, stress resistance, and fat metabolism (10). Moreover, several recent reports show that miRNAs participate in human tumorigenesis as tumor suppressors or oncogenes (11-13). For example, miRNA *let-7*, targeting the oncogene *Ras*, is down-regulated in lung cancer (14), whereas *miR-17-92* cluster at 13q31.3 was reported to be overexpressed in lung cancer (15). Single nucleotide polymorphisms (SNP) or mutations in miRNA sequence may alter miRNA expression and/or maturation. Recently, we performed a screening for common SNPs in miRNA sequence and identified four SNPs (rs2910164, rs2292832, rs11614913, and rs3746444) located at the pre-miRNA regions of *miR-146a, miR-149, miR-196a2*, and *miR-499*, respectively (16). We found that, among the above four SNPs, the rs11614913 SNP in *miR-196a2* was associated with shortened survival time of non-small cell lung cancer through altering the expression of mature *miR-196a* and binding activity of target mRNA (16). In the present study, we hypothesized that this functional SNP, rs11614913 T/C in *miR-196a2*, was also associated with lung cancer susceptibility. To test this hypothesis, we performed genotyping analyses for rs11614913 T/C and the other three common SNPs (rs2910164 C/G, rs2292832 C/T, and rs3746444 A/G) located at pre-miRNA regions and evaluated their associations with the susceptibility of lung cancer in a case-control study of 1,092 lung cancer

Received 9/5/08; revised 1/20/09; accepted 2/10/09; published OnlineFirst 3/17/09.

Grant support: Key Grant of National Natural Science Foundation of China 30730080, National Outstanding Youth Science Foundation of China grant 30425001, China National Key Basic Research Program grant 2002CB512902, and Natural Science Foundation of Jiangsu Province grant BS2006005.

Note: T. Tian and Y. Shu contributed equally to this work.

Requests for reprints: Hongbing Shen, Department of Epidemiology and Biostatistics, Nanjing Medical University, 140 Hanzhong Road, Nanjing 210029, People's Republic of China. Phone: 86-25-868-62756; Fax: 86-25-868-62756. E-mail: hbshen@njmu.edu.cn

doi:10.1158/1055-9965.EPI-08-0814

cases and 1,064 cancer-free controls in a Chinese population.

Materials and Methods

Subjects. This is an ongoing molecular epidemiologic study of lung cancer conducted in Nanjing, China and the subjects recruitment was approved by the Institutional Review Board of Nanjing Medical University. Briefly, all subjects were genetically unrelated ethnic Han Chinese and were from Nanjing City and the surrounding regions in Jiangsu Province of eastern China. The incident lung cancer patients were histopathologically diagnosed and recruited between July 2002 and April 2008 at the Cancer Hospital of Jiangsu Province (Nanjing), The First Affiliated Hospital of Nanjing Medical University, and the Nanjing Thoracic Hospital, without the restrictions of age, sex, and histology. The exclusion criteria included previous cancer, metastasized cancer, and previous radiotherapy or chemotherapy. Cancer-free controls were randomly selected from a pool of 30,000 individuals who participated in a community-based screening program for noninfectious diseases conducted in Jiangsu Province during the same period as the cases were recruited. The control subjects had no history of cancer and were frequency matched to the cases on age ($\pm$5 years), sex, and residential area (urban or countryside). Each participant was scheduled for an interview after written informed consent was obtained, and a structured questionnaire was administered by interviewers to collect information on demographic data and environmental exposure history. Those who had smoked <1 cigarette per day and <1 year in their lifetime were defined as nonsmokers; otherwise, they were considered as smokers. Those smokers who quit for >1 year were considered former smokers. Pack-years smoked [(cigarettes per day / 20) $\times$ years smoked] were calculated to indicate the cumulative smoking dose. After interview, ~5 mL venous blood sample was collected from each participant.

SNP Identification. We searched in silicon by blasting 400 miRNAs and their surrounding sequences with the dbSNP database and identified 273 common (minor allele frequency > 0.05) genetic variants located in pre-miRNAs and their surrounding regions in all the ethnicity groups (Asian, European, and African) and found that only five SNPs located at the pre-miRNA regions denoted as class A/B (class A: located in 5p or 3p mature miRNA regions and class B: located in other regions of pre-miRNAs) were unique for Chinese populations. The other SNPs classified as class C (located in 100 bp flanking regions of the pre-miRNAs), class D (located in ~100-200 bp flanking regions), and class E (located in ~200-450 bp flanking regions). It was reported that SNPs in pre-miRNAs could alter miRNA processing, expression, and/or binding to target mRNA (17). One class B SNP (rs6505162) that did not affect hydrogen band and predicted that secondary structure free energy was not included in this study (18). Therefore, we finally selected four pre-miRNA SNPs for genotyping in our studies.

Genotyping. The identification of four SNPs in the pre-miRNA regions were described previously (16), that is, common (minor allele frequency > 0.05) genetic variants for Chinese populations located in pre-miRNAs of 400 known human miRNAs, which may alter miRNA processing, expression, and/or binding to target mRNAs. We used PCR-RFLP assay to achieve the genotypes of the four SNPs (rs2910164 C/G, rs2292832 C/T, rs11614913 T/C, and rs3746444 A/G) (16). Two research assistants independently read the gel pictures and performed the repeated assays if they did not reach a consensus on the tested genotype. In addition, 10% of the samples were randomly selected to perform the repeated assays for each locus, and the results were 100% concordant. DNA quality or quantity was insufficient and failed for genotyping in 34 lung cancer cases and 29 controls; thus, the final analyses included 1,058 cases and 1,035 controls.

Statistical Analysis. Differences in demographic variables, selected variables, and frequencies of the genotypes between cases and controls were evaluated by using the Student's t test (for continuous variables) and χ^2 test (for categorical variables). The associations between pre-miRNA SNPs and lung cancer risk were estimated by computing the odds ratios and 95% confidence intervals from logistic regression analyses. The potential gene-environment interaction was also evaluated by logistic regression analysis and tested by comparing the changes in deviance (-2 log likelihood) between the models of main effects with or without the interaction term. Hardy-Weinberg equilibrium was tested by a goodness-of-fit χ^2 test to compare the observed genotype frequencies with the expected ones among the control subjects. All the statistical analyses were done with Statistical Analysis System software (version 9.1.3; SAS Institute).

Results

The characteristics of the 1,058 lung cancer patients and 1,035 controls included in the analysis are summarized in Table 1. There were no statistically significant differences between cases and controls in terms of the frequency distribution of sex and age. As expected, smoking was a significant risk factor for lung cancer. About 49.1% of the cases were current smokers, which were significantly higher than that of the controls (33.7%; $P < 0.0001$), and lung cancer cases are more frequent to be heavy smokers (>30 pack-years) than the controls (34.3% versus 17.0%; $P < 0.0001$). Of the 1,058 lung cancer cases, 486 (45.9%) were adenocarcinoma, 342 (32.3%) squamous cell carcinoma, 102 (9.6%) small cell carcinoma, and 128 (12.1%) large cell, mixed cell, or undifferentiated carcinomas.

The genotype distributions of the four SNPs in the cases and controls are shown in Table 2. The observed genotype frequencies for these four polymorphisms were all in agreement with that expected under the Hardy-Weinberg equilibrium in the controls (P = 0.853 for *miR-146a* rs2910164, 0.855 for *miR-149* rs2292832, 0.700 for *miR-196a2* rs11614913, and 0.404 for *miR-499* rs3746444). In the recessive genetic model, we found that the rs11614913 variant homozygote CC of *miR-196a2* was associated with a significantly increased risk of lung cancer compared with its wild-type homozygote TT and heterozygote CT (odds ratio, 1.25; 95% confidence

Table 1. Distribution of selected variables in lung cancer cases and controls

Variable	Cases (n = 1,058), n (%)	Controls (n = 1,035), n (%)	P
Age, y (mean ± SD)	59.78 ± 10.04	59.66 ± 9.83	0.777
Age, y			
≤60	550 (52.0)	530 (51.2)	0.722
>60	508 (48.0)	505 (48.8)	
Sex			
Male	790 (74.7)	780 (75.4)	0.714
Female	268 (25.3)	255 (24.6)	
Smoking status			
Nonsmokers	373 (35.3)	575 (55.6)	<0.0001
Former smokers	165 (15.6)	111 (10.7)	
Current smokers	520 (49.1)	349 (33.7)	
Pack-years of smoking			
0	373 (35.3)	575 (55.6)	<0.0001
1-30	322 (30.4)	284 (27.4)	
>30	363 (34.3)	176 (17.0)	
Stage			
I	230 (21.7)		
II	130 (12.3)		
III	372 (35.2)		
IV	212 (20.0)		
Unclassified	114 (10.8)		
Histologic types			
Adenocarcinoma	486 (45.9)		
Squamous cell	342 (32.3)		
Small cell	102 (9.6)		
Other carcinoma	128 (12.1)		

interval, 1.01-1.54). However, we did not find any main effects between the other three SNPs and lung cancer risk. We also calculated false-positive report probability (19) value for the SNP rs11614913. When the assumption of prior probability was 0.25 or 0.1, the value of false-positive report probability is 0.17 and 0.39, respectively. Considering the functional relevance of rs11614913, the value of false-positive report probability could be as high as 0.17, which suggested that our finding would have a 83% probability to represent a true association. To evaluate the gene-smoking interactions, we performed stratification analyses by smoking status and found that the risk of lung cancer associated with rs11614913 CC genotype were slightly more evident among smokers (odds ratio, 1.31; 95% confidence interval, 0.98-1.74) than nonsmokers (odds ratio, 1.13; 95% confidence interval, 0.82-1.56). However, there were no significant interactions between this genotype and smoking status by logistic regression analysis. In addition, we performed stratification analyses according to subjects' age, sex, histologic types, and disease stages to examine the association between rs11614913 genotype and lung cancer risk and found no heterogeneity of the risk in each stratum (data not shown).

Table 2. Main effects of pre-miRNA SNPs on lung cancer risk

Genotypes	Lung cancer cases (n = 1,058), n (%)	Controls (n = 1,035), n (%)	Odds ratio (95% confidence interval)*	P
miR-146a rs2910164				
GG	360 (34.0)	364 (35.2)	1.00 (reference)	
CG	510 (48.2)	502 (48.5)	1.03 (0.85-1.24)	0.783
CC	188 (17.8)	169 (16.3)	1.13 (0.87-1.45)	0.364
miR-149 rs2292832				
TT	463 (43.8)	470 (45.4)	1.00 (reference)	
CT	472 (44.6)	453 (43.8)	1.06 (0.88-1.27)	0.546
CC	123 (11.6)	112 (10.8)	1.12 (0.84-1.48)	0.457
miR-196a2 rs11614913				
TT	293 (27.7)	307 (29.7)	1.00 (reference)	
CT	512 (48.4)	519 (50.1)	1.03 (0.85-1.26)	0.747
CC	253 (23.9)	209 (20.2)	1.27 (0.99-1.62)	0.056
TT/CT	805 (76.1)	826 (79.8)	1.00 (reference)	
CC	253 (23.9)	209 (20.2)	1.25 (1.01-1.54)	0.038
miR-499 rs3746444				
AA	781 (73.8)	755 (73.0)	1.00 (reference)	
AG	253 (23.9)	254 (24.5)	0.96 (0.79-1.18)	0.712
GG	24 (2.3)	26 (2.5)	0.89 (0.51-1.57)	0.692

*Adjusted for age, sex, and pack-years of smoking.

Discussion

In this case-control study of lung cancer in Chinese, we found, for the first time, that variant genotype CC of *miR-196a2* rs11614913 was associated with significantly increased risk of lung cancer. Together with our previous observations that rs11614913 CC predicts poorer survival of non-small cell lung cancer, elevated mature *miR-196a* expression, and increased target mRNA binding (16), we conclude that this functional variant rs11614913 was a candidate biomarker of both lung cancer occurrence and survival. In stratified analysis in the present study, we did not find any significant associations among different subgroups. Because of the small sample size in the subgroups, these findings were primary and need to be validated in further studies.

Our understanding of miRNA expression patterns and function in normal or neoplastic human cells is just starting to emerge. Researchers first identified a link between miRNA expression and human malignancy with the observation that there was a down-regulation or deletion of miRNAs *miR-15a* and *miR-16-1* in patients with B-cell chronic lymphocytic leukemia (20). In 2005, Johnson et al. observed an inverse relationship between miRNA *let-7* and RAS protein expression in human lung cancer cell lines (21). Microarray analysis of human cancerous tissues also showed decreased *let-7* expression in lung cancer but not in adjacent normal lung tissue (22). Other links between lung cancer and miRNAs have been reported, including high expression of *mir-17-92* cluster in lung cancer (15). It has been also shown that miRNA expression patterns are associated with the biological and clinical behavior of human solid tumors, including lung cancer (6). However, the precise mechanisms regulating miRNA expression are ambiguous and the low penetrance genetic effect of miRNA SNPs to cancer diagnosis and prognosis is largely unknown.

During the preparation for this study, Jazdzewski et al. reported that the rs2910164 SNP in *pre-mir-146a* could reduce mature *mir146a* expression and affect target mRNA binding (23), which is similar to our recent findings on the rs11614913 SNP (16). rs2910164 was also found to be associated with susceptibility to papillary thyroid carcinoma (23). However, we did not find such an association in our lung cancer patients. A newly published epidemiologic study also reported a positive association between miRNA-binding SNPs and carcinogenesis, in which Landi et al. performed a case-control study that examined the relationship of eight polymorphisms within miRNA-binding sites and the risk of sporadic colorectal cancer and found that the variant alleles in *CD86* and *INSR* 3′-untranslated region were significantly associated with the risk of CRC (24). To date, no tangible evidence showed associations between miRNA SNPs and lung carcinogenesis.

We set an example by using the *LSP1* 3′-untranslated region as the target of 3p mature *miR-196a2* and showed that rs11614913 influence target mRNA binding (16). We also showed rs11614913 can influence mature *miR-196a* expression, which partly directs the cleavage of *HOX* gene cluster (25). Several studies reported that some *HOX* genes were related to lung development and lung cancer metastasis (26, 27). For example, Hamada et al. revealed that abnormal expression of *HOXD3* gene in human lung cancer A549 cells enhanced invasion and metastasis through coordinate expression of metastasis-associated molecules (26). Therefore, the miRNA SNP may play an important role in lung cancer development and survival through influencing the expression/maturation of miRNAs and the binding of their mRNA targets. Further functional evaluation of the rs11614913 SNP, *miR-196a2* miRNA, and also its target mRNAs in lung cancer development, invasion, and metastasis are warranted. A later case-control set from the same study would suffer the same bias as the initial set, and validations from different well-designed studies for the findings of our present study will be more convincing.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We thank our collaborators in the Department of Thoracic & Cardiac Surgery, The First Affiliated Hospital of Nanjing Medical University (Yijiang Chen) and the Department of Respiratory Diseases, Nanjing Thoracic Hospital (Yi Zeng).

References

1. Yang L, Parkin DM, Li L, Chen Y. Time trends in cancer mortality in China: 1987-1999. Int J Cancer 2003;106:771–83.
2. Sekido Y, Fong KM, Minna JD. Molecular genetics of lung cancer. Annu Rev Med 2003;54:73–87.
3. Meyerson M, Carbone D. Genomic and proteomic profiling of lung cancers: lung cancer classification in the age of targeted therapy. J Clin Oncol 2005;23:3219–26.
4. Granville CA, Dennis PA. An overview of lung cancer genomics and proteomics. Am J Respir Cell Mol Biol 2005;32:169–76.
5. Minna JD, Roth JA, Gazdar AF. Focus on lung cancer. Cancer Cell 2002;1:49–52.
6. Yanaihara N, Caplen N, Bowman E, et al. Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. Cancer Cell 2006;9:189–98.
7. Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T. Identification of novel genes coding for small expressed RNAs. Science 2001;294:853–8.
8. Lau NC, Lim LP, Weinstein EG, Bartel DP. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. Science 2001;294:858–62.
9. Lee RC, Ambros V. An extensive class of small RNAs in *Caenorhabditis elegans*. Science 2001;294:862–4.
10. Ambros V. MicroRNA pathways in flies and worms: growth, death, fat, stress, and timing. Cell 2003;113:673–6.
11. Lu J, Getz G, Miska EA, et al. MicroRNA expression profiles classify human cancers. Nature 2005;435:834–8.
12. Esquela-Kerscher A, Slack FJ. Oncomirs—microRNAs with a role in cancer. Nat Rev Cancer 2006;6:259–69.
13. Kumar MS, Lu J, Mercer KL, Golub TR, Jacks T. Impaired microRNA processing enhances cellular transformation and tumorigenesis. Nat Genet 2007;39:673–7.
14. Takamizawa J, Konishi H, Yanagisawa K, et al. Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. Cancer Res 2004;64:3753–6.
15. Hayashita Y, Osada H, Tatematsu Y, et al. A polycistronic microRNA cluster, miR-17-92, is overexpressed in human lung cancers and enhances cell proliferation. Cancer Res 2005;65:9628–32.
16. Hu Z, Chen J, Tian T, et al. Genetic variants of miRNA sequences and non small cell lung cancer survival. J Clin Invest 2008;118:2600–8.
17. Duan R, Pak C, Jin P. Single nucleotide polymorphism associated with mature miR-125a alters the processing of pri-miRNA. Hum Mol Genet 2007;16:1124–31.
18. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 2003;31:3406–15.
19. Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. J Natl Cancer Inst 2004;96:434–42.

20. Calin GA, Dumitru CD, Shimizu M, et al. Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc Natl Acad Sci U S A 2002;99:15524–9.
21. Johnson SM, Grosshans H, Shingara J, et al. RAS is regulated by the let-7 microRNA family. Cell 2005;120:635–47.
22. Eder M, Scherr M. MicroRNA and lung cancer. N Engl J Med 2005;352:2446–8.
23. Jazdzewski K, Murray EL, Franssila K, Jarzab B, Schoenberg DR, de la Chapelle A. Common SNP in pre-miR-146a decreases mature miR expression and predisposes to papillary thyroid carcinoma. Proc Natl Acad Sci U S A 2008;105:7269–74.
24. Landi D, Gemignani F, Naccarati A, et al. Polymorphisms within micro-RNA-binding sites and risk of sporadic colorectal cancer. Carcinogenesis 2008;29:579–84.
25. Yekta S, Shih IH, Bartel DP. MicroRNA-directed cleavage of HOXB8 mRNA. Science 2004;304:594–6.
26. Hamada Ji, Omatsu T, Okada F, et al. Overexpression of homeobox gene HOXD3 induces coordinate expression of metastasis-related genes in human lung cancer cells. Int J Cancer 2001;93:516–25.
27. Volpe MV, Pham L, Lessin M, et al. Expression of Hoxb-5 during human lung development and in congenital lung malformations. Birth Defects Res Part A Clin Mol Teratol 2003;67:550–6.

A Functional Genetic Variant in microRNA-196a2 Is Associated with Increased Susceptibility of Lung Cancer in Chinese

Tian Tian, Yongqian Shu, Jiaping Chen, et al.

Cancer Epidemiol Biomarkers Prev 2009;18:1183-1187.