Genetic Deletions in Sputum as Diagnostic Markers for Early Detection of Stage I Non–Small Cell Lung Cancer

Ruiyun Li,[1] Nevins W. Todd,[2] Qi Qiu,[3] Tao Fan,[3] Richard Y. Zhao,[3] William H. Rodgers,[3] Hong-Bin Fang,[4] Ruth L. Katz,[5] Sanford A. Stass,[3] and Feng Jiang[3]

Abstract

Purpose: Analysis of molecular genetic markers in biological fluids has been proposed as a powerful tool for cancer diagnosis. We have characterized in detail the genetic signatures in primary non–small cell lung cancer, which provided potential diagnostic biomarkers for lung cancer. The aim of this study was to determine whether the genetic changes can be used as markers in sputum specimen for the early detection of lung cancer.

Experimental Design: Genetic aberrations in the genes *HYAL2, FHIT*, and *SFTPC* were evaluated in paired tumors and sputum samples from 38 patients with stage I non–small cell lung cancer and in sputum samples from 36 cancer-free smokers and 28 healthy nonsmokers by using fluorescence *in situ* hybridization.

Results: *HYAL2* and *FHIT* were deleted in 84% and 79% tumors and in 45% and 40% paired sputum, respectively. *SFTPC* was deleted exclusively in tumor tissues (71%). There was concordance of *HYAL2* or *FHIT* deletions in matched sputum and tumor tissues from lung cancer patients ($r = 0.82$, $P = 0.04$; $r = 0.84$, $P = 0.03$), suggesting that the genetic changes in sputum might indicate the presence of the same genetic aberrations in lung tumors. Furthermore, abnormal cells were found in 76% sputum by detecting combined *HYAL2* and *FHIT* deletions whereas in 47% sputum by cytology, of the cancer cases, implying that detecting the combination of *HYAL2* and *FHIT* deletions had higher sensitivity than that of sputum cytology for lung cancer diagnosis. In addition, *HYAL2* and *FHIT* deletions in sputum were associated with smoking history of cancer patients and smokers (both $P < 0.05$).

Conclusions: Tobacco-related *HYAL2* and *FHIT* deletions in sputum may constitute diagnostic markers for early-stage lung cancer.

Lung cancer is the most common cause of neoplasia-related deaths worldwide. The exceptionally high mortality rate associated with the disease demands effective early diagnostic strategies (1). Historically, chest X-ray and sputum cytology have been used for the early detection of lung cancer; however, they have not been proven effective in improving the overall survival as evaluated by the National Cancer Institute (2, 3). Although low-dose computed tomography seems promising as it can detect lung cancer at a smaller size compared with a chest X-ray, the technique is very expensive and the improved sensitivity is associated with high false-positive rate (3). The fluorescence bronchoscopy excels at detecting centrally occurring lung tumor (1). However, it is an invasive technique. There is thus an urgent need to develop reliable early diagnostic biomarkers for lung cancer that can be detected in noninvasively obtained samples.

We previously defined genetic signatures of primary lung cancers by microarray-based comparative genomic hybridization analyses, providing potential biomarkers for the early detection of lung cancer (4, 5). Furthermore, we developed a genetic test that could simultaneously evaluate diagnostic values of the signatures in clinical specimens (6). In addition, we validated the defined genetic changes in a large cohort of patients with lung cancer; specifically, we found that deletions of the *SP-A* gene in chromosome 10q22 might be a useful biomarker of poor prognosis in patients with early-stage lung cancer (7). Moreover, we analyzed tumor tissues and paired bronchial brushings from lung cancer patients using fluorescence *in situ* hybridization (FISH). Results of the study suggested that examination of the cellular components of bronchial brushings by FISH with tumor-specific DNA probes might aid lung cancer diagnosis, especially that of peripheral tumors (8). However, because bronchial brushing specimens are collected via an invasive bronchoscopic procedure, they are not suitable for use as surrogate materials in evaluating carcinogenic damage for large-scale population-based genetic studies in clinical setting.

Authors' Affiliations: Departments of [1]Surgery, [2]Medicine, and [3]Pathology, The University of Maryland School of Medicine; [4]Division of Biostatistics, The University of Maryland Greenebaum Cancer Center, Baltimore, Maryland; and [5]Department of Pathology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas

Received 6/29/06; revised 9/5/06; accepted 10/13/06.

Grant support: National Cancer Institute grant CA-113707 (F. Jiang).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Requests for reprints: Feng Jiang, Department of Pathology, The University of Maryland School of Medicine, 10 South Pine Street, MSTF 7th Floor, Baltimore, MD 21201-1192. Phone: 410-706-7056; Fax: 410-706-8414; E-mail: fjiang@som.umaryland.edu.

doi:10.1158/1078-0432.CCR-06-1593

The lung carcinogenesis is characterized by an accumulation of molecular genetic abnormalities resulting from repeated exposure of the respiratory tracts of chronic smokers to tobacco-related carcinogens (9). Therefore, the abnormalities, regardless of their location, may reflect the severity of the cancerization and may thus be targets for diagnosis (10). Because cigarette smokers have increased amounts of sputum that contain exfoliated cells from the bronchial tree and sputum is the most easily accessible biological fluid, molecular genetic analysis of sputum has been an active area of research on lung cancer biomarkers (11). Therefore, analysis of molecular genetic markers in sputum will provide a powerful diagnostic tool for the early detection of lung cancer.

We have previously reported that deletions in *HYAL2*, *FHIT*, and *SFTPC* were three of the most common genetic changes in primary lung cancer (4, 5). To determine whether the genetic changes can be used as markers in sputum specimen for the early detection of lung cancer, we studied the presence of these genetic changes in the sputum and tumors of 38 lung cancer patients and in the sputum of 36 smokers and 28 nonsmokers. We showed that there was high concordance between *HYAL2* or *FHIT* deletions in paired sputum and tumor tissues from lung cancer patients. Detecting combination of *HYAL2* and *FHIT* deletions in sputum had higher sensitivity than that of sputum cytology. Our findings, therefore, suggest that *HYAL2* or *FHIT* deletions may be new diagnostic markers and determination of the aberrations with FISH in sputum might be a complementary diagnostic tool for early-stage lung cancer.

Materials and Methods

Patients. Thirty-eight patients with stage I non – small cell lung cancer (NSCLC) were entered into the prospective study. All the patients were >18 years of age and had lung masses that were operable and had not received chemotherapy or radiation. Thirty-six current smokers who did not have lesions detectable by chest X-ray or computed tomography were studied as controls. To calculate pack-years, we multiplied the number of years smoked by the number of packs per day. Twenty-eight healthy nonsmokers were also recruited as "strict" control subjects. Written informed consent for participation was obtained through an institutional review board – approved protocol.

Collection and processing of tumor tissue and sputum. Tumor tissue was obtained by surgical resection from patients with lung cancer. After a sufficient number of sections were obtained to satisfy the requirements of a standard, comprehensive pathologic diagnosis, touch preparations were taken from actual residual tumors for subsequent FISH analysis.

Sputum was collected from all individuals using an ultrasonic nebulization technique (12, 13), which can maximize the yield of atypical cells from airway and decrease percentage of oral epithelial cells in the sputum. Furthermore, sputum collected by the same method from all the subjects would be equally informative and representative of deep bronchial cells. Twelve cytocentrifuge slides were made from each sputum in single preparation by using a cytospin machine (Shandon, Inc., Pittsburgh, PA) that can process 12 slides at one time. Two cytospin slides from each sample were prepared and stained with Papanicolaou stain for cytologic diagnosis using the classification of Saccomanno (12, 13). Positive cytology included both carcinoma *in situ* and invasive carcinoma. The main variables used to select slides for the FISH assay were high cellularity, good nuclear morphology, and lack of purulence, debris, and residual cytoplasm (14). The rest of the slides were fixed in fresh methanol and glacial acetic acid and stored at −20°C for FISH analysis. Because the 12 slides of each sputum were prepared from the same cell suspension in single preparation, all the remaining slides that were used for FISH test had the same cell quality as did the slides for cytologic evaluation.

Preparation of normal interphase and metaphase cells. Interphase and metaphase cells were prepared from peripheral lymphocytes from 10 control subjects as described previously (15). The cells were used as controls to test the efficiency of the FISH assay and to establish cutoff values to define abnormal specimens.

Fluorescence in situ hybridization. A specific probe for *FHIT* was prepared from a bacterial artificial chromosome clone, which was identified from the Celera database.[6] The clone was 200 kb in length and mapped to exons 5 to 8 of *FHIT*. The exons 5 to 8 have been proven to be the most frequently deleted regions of *FHIT* in lung cancer (16). To develop the unique probes that would specifically cover the genomic sequences of *HYAL2* and *SFTPC*, we used a strategy of long-distance PCR and degenerate oligonucleotide-primed PCR as described in our recent published work (6). Briefly, using the Celera database, we first identified clones that contained genomic sequences of several genes, including *HYAL2* and *SFTPC*, respectively. We then amplified considerably larger amplimers by long-distance PCR with specific primers for the target gene from the clone DNA. With this approach, several overlapping bidirectional PCRs with structurally complex genomic sequences were produced to span the whole genomic region of the target gene. The DNA of the long-distance PCR amplicons was then amplified by degenerate oligonucleotide-primed PCR (DOP PCR Master, Roche Diagnostics Corp., Indianapolis, IN) according to the manufacturer's instruction. The sizes of probes for *HYAL2* and *SFTPC* were 5 and 3 kb, respectively. Finally, the degenerate oligonucleotide-primed PCR – amplified DNA was labeled by nick translation reaction with fluorochrome-dUTP (Alexa Fluor, Molecular Probes, Eugene, OR). The exact size of each gene-specific probe developed by the protocol was confirmed by FISH on normal DNA fiber as described previously (6, 15). Furthermore, the chromosomal location of each probe was validated using centromeric probes for chromosomes 3 and 8 (CEP3 and CEP8, Vysis, Downers Grove, IL) in control metaphase and interphase cells, respectively.

Tricolor FISH was done using the *HYAL2* probe, which was labeled with green fluorescence; the *FHIT* probe, which was labeled with yellow fluorescence; and the CEP3 probe, which was labeled with red fluorescence and used as an internal control probe. Dual-color FISH was done using the *SFTPC* probe with green fluorescence and the CEP8 probe with red fluorescence. Each probe (100 ng) was mixed with a 30-fold excess of human Cot-1 DNA (Life Technologies, Rockville, MD) in 10 μL of LSI hybridization buffer (Vysis) and mounted on a slide. Hybridization and postwashing were done as described previously (7, 8, 15). The slides were examined under microscopes equipped with appropriate filter sets (Leica Microsystems, Buffalo, NY). Cells (200) were counted on each slide from all control and case specimens in the study. More or less signals from the *HYAL2* or *FHIT* probes than from the CEP3 probe indicated a gain or loss, respectively, of *HYAL2* or *FHIT* gene. Similarly, more or less signals from the *SFTPC* probe than from the CEP8 probe indicated a gain or loss, respectively, of *SFTPC* gene. The cutoff value was calculated from normal cells and defined as the mean number of cells ± 3 SD with an abnormal *HYAL2*, *FHIT*, or *SFTPC* signal pattern by FISH analysis.

Statistical analysis. The sample size was determined based on the three-group nested case-control study consisting of lung cancer cases, healthy nonsmokers who were considered as strict "control" population, and cancer-free smokers. The area under receiver operating characteristic curve (AUROC) of H0 (the null hypothesis) was set at 0.5. H1 represented the alternative hypothesis; accordingly, at least 28 patients were required in each category to show a minimum difference of interest between an AUROC of 0.75 versus an AUROC of 0.5 with 80% power at the 5% significance level (17). Therefore, the group assignment in the study with 38 cancer cases,

[6] http://myscience.appliedbiosystems.com.

Table 1. Characteristics of the subject population

	Primary NSCLC (n = 38)	Heavy smokers (n = 36)	Healthy nonsmokers (n = 28)
Age*	66 (45-79)	64 (43-83)	60 (34-78)
Sex			
Female	16	16	13
Male	22	20	15
Pack-years	36	48	
Smoking status			
Smoker	32	36	
Never smoker	6		28
Location of tumor			
Central	18		
Peripheral	20		
Stage	All are stage I		
Histology			
AC of lung	22		
SCC of lung	16		

Abbreviations: AC, adenocarcinoma; SCC, squamous cell carcinoma.
*Data are presented as median (range).

36 smokers, and 28 healthy nonsmokers should give enough statistical analysis power.

All statistical analyses were done using Statistical Analysis System software version 6.12 (SAS Institute, Cary, NC). Associations between deletions of the three genes and patients' clinical characteristics were analyzed using the Wilcoxon rank-sum test for continuous variables or Fisher's exact test for categorical variables. The correlations of genetic deletions between different specimens were analyzed using Spearman's rank correlation coefficient. All P values shown were two sided, and a P value of <0.05 was considered statistically significant.

Results

Patient characteristics. Characteristics of subjects enrolled in this study are compiled in Table 1. All the lung cancer patients had stage I NSCLCs, including 22 adenocarcinomas and 16 squamous cell carcinomas, as determined histologically according to WHO classification. The mean number of pack-years for cancer cases and smokers was 36 and 48, respectively.

Hybridization efficiency and cutoff values for FISH. Each gene-specific probe showed a single linear DNA molecule with juxtaposed color barcode signals as confirmed by FISH on normal DNA fiber. Dual-color FISH detected two green signals from the *SFTPC* probe and two red signals from the CEP8 probe in the control cells (Fig. 1A). Similarly, tricolor FISH detected two green signals from the *HYAL2* probe, two yellow signals from the *FHIT* probe, and two red signals from the CEP3 probe in the control metaphase and interphase cells (Fig. 1B). Loss of *HYAL2*, *FHIT*, and *SFTPC* signals was detected in 2% to 6%, 2% to 5%, and 3% to 5% of control cells, respectively. A specimen was therefore considered to have deletions in *HYAL2*, *FHIT*, or *SFTPC* if more than 10%, 9%, or 9% of its cells, respectively, had deletions in those genes.

Genetic changes in paired tumor and sputum samples from lung cancer patients. As shown in Table 2, the overall frequency of *HYAL2* deletions in tumors was higher than that in the paired sputum samples (84% versus 45%, respectively; P = 0.002). Seventeen of 32 (54%) cancer patients whose tumors had *HYAL2* deletions also had such deletions in their sputum. The patients whose tumors did not have *HYAL2* deletions had no such deletions in their sputum. Thus, 23 of 38 (60%) patients had concordant *HYAL2* deletion status in their tumors and sputum (r = 0.82, P = 0.04). Similarly, the frequency of *FHIT* deletions in tumors (79%) was significantly higher than in the paired sputum samples (40%; P = 0.001). Of the 30 tumors

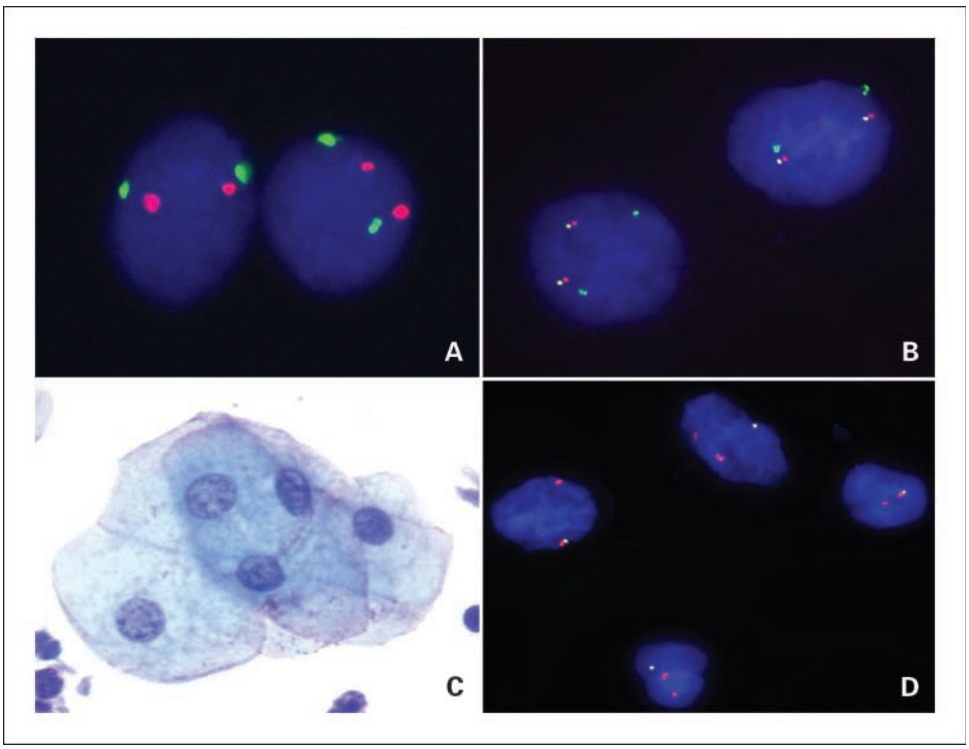

Fig. 1. FISH analyses of exfoliated cells from sputum specimens. *A,* dual-color FISH analysis of cells from a healthy nonsmoker shows two green signals from the *SFTPC* probe and two red signals from the CEP8 probe. Original magnification, ×1,000. *B,* tricolor FISH analysis of cells from healthy nonsmoker shows two green signals from the *HYAL2* probe, two yellow signals from the *FHIT* probe, and two red signals from the CEP3 probe. Original magnification, ×1,000. *C,* cytologic analysis of sputum sample from a lung cancer patient shows cells with atypical metaplastic changes. Original magnification, ×60. *D,* tricolor FISH analysis of bronchial epithelium prepared from the same lung cancer patient shows homozygous deletion of *HYAL2* (no green signals) and hemizygous deletion of *FHIT* (one yellow signals). There are two red signals from the CEP3 probe, indicating diploidy of chromosome 3. Original magnification, ×400.

Table 2. Prevalence of genetic deletions in paired tumor and sputum samples from NSCLC patients

Gene	No. subjects with deletions (%)	
	Tumor, no. positive (%)	Sputum, no. positive (%)
HYAL2	32/38 (84)	17/38 (45)
FHIT	30/38 (79)	15/38 (40)
SP-C	27/38 (71)	0/38 (0)

with a *FHIT* deletion, 15 (50%) had a deletion in the corresponding sputum sample, whereas in all of the 8 tumors without *FHIT* deletion, none had *FHIT* deletions in their sputum samples. Thus, there was concordance between *FHIT* deletions in paired sputum and tumor tissues from lung cancer patients ($r = 0.84$). The data suggested that deletion of *HYAL2* and *FHIT* in sputum might indicate the presence of the same genetic lesion in lung tumors. Although *SFTPC* deletions were found in 27 of 38 (71%) tumors, no such lesions were detected in the corresponding sputum samples, implying that the *SFTPC* deletion could be a genetic lesion that occurs exclusively in lung tumor tissues. Although there was concordance between HYAL2 and FHIT deletions in tumor tissues, there was not such concordance in the paired sputum ($r = 0.89$ and 0.22).

Genetic changes in sputum samples of lung cancer patients, cancer-free smokers, and nonsmokers. *HYAL2* deletions were found in 17 of 38 (45%) and 4 of 36 (11%) sputum samples from cancer patients and cancer-free smokers, respectively, but not in sputum from nonsmokers. The frequency of *HYAL2* deletion in sputum from cancer patients was statistically higher than that in sputum from smokers ($P = 0.001$). *FHIT* deletions were found in 15 of 38 (40%) cancer patients, 3 of 36 (8%) smokers, and none of 22 nonsmokers. The frequency of *FHIT* deletions in sputum from cancer patients was statistically higher than that in sputum from cancer-free smokers ($P = 0.003$).

The correlation between the genetic deletions in sputum and clinical characteristics of the subjects. The prevalence of *HYAL2* and *FHIT* deletions in the stage I lung cancer patients was associated with pack-years of smoking (all $P < 0.05$) but was independent of patient age, gender, histologic tumor type, and tumor size and location (all $P > 0.05$). The prevalence of *HYAL2* or *FHIT* deletions in the smokers was also related with pack-years of smoking (all $P < 0.05$) but was independent of patient age and gender. Therefore, tumor-related *HYAL2* and *FHIT* deletions in sputum might also be indicators of smoke-damaged epithelium.

Comparison of sputum cytology and FISH with HYAL2 or FHIT probes for the detection of lung cancer. As shown in Fig. 2 and Table 3, 18 of 38 (47%) sputum from lung cancer patients were diagnosed as positive for cancer by cytologic examination. Of the cytologically positive sputum samples, 8 (44%) had *HYAL2* deletions, 7 (39%) showed *FHIT* deletions, and 3 (16%) had both *HYAL2* and *FHIT* deletions. Of the 20 cytologically negative sputum samples from lung cancer patients, 5 (25%) had *HYAL2* deletions, 5 (25%) had *FHIT* deletions, and 1 (5%) had deletions for both genes (Fig. 1C and D). All the sputum from cancer-free smokers and healthy nonsmokers were negative by cytology. In the sputum from 36 cancer-free smokers, 3 (8%) had *HYAL2* deletions, 2 (6%) had *FHIT* deletions, and 1 (3%) had deletions for both genes, respectively. None of sputum samples from healthy nonsmokers was positive for any gene deletion. Overall, the sensitivity and specificity of cytologic examination were 47% and 100%, respectively, whereas the sensitivity and specificity of *HYAL2* and *FHIT* deletions for the detection of abnormal cells of sputum were 45% and 94%, 40%, and 95%, respectively. However, using ROC and their AUROCs, we found that combining both *HYAL2* and *FHIT* deletions had a sensitivity of 76% and a specificity of 92%, yielding higher sensitivity than that of sputum cytology ($P = 0.02$; Fig. 2; Table 3).

Discussion

Our current study clearly showed a high concordance of *HYAL2* or *FHIT* deletions in matched lung tumor and sputum specimens from primary lung cancer patients, implying that the genetic aberrations seen in sputum might reflect the presence of the same genetic events in lung tumors. Furthermore, combining both *HYAL2* and *FHIT* deletions yielded higher sensitivity than that of sputum cytology, suggesting that the detection of the aberrations might enhance diagnosis of lung cancer at an early stage.

HYAL2 is located in chromosome 3p21.3 that is the most frequent target of homozygous deletions in lung cancer (18). Recently, by studying specific nucleotide polymorphisms, Senchenko et al. (19) found that the smallest region of homozygous deletions in 3p21.3, including *HYAL2*, occurred in more than 90% of lung, renal, and breast tumors, suggesting that *HYAL2* might be a tumor suppressor gene. In the current study, deletion of *HYAL2* was the most frequently detected alteration of the three markers investigated in sputum and tumors in lung cancer patients. Furthermore, there was concordance between *HYAL2* deletions in primary tumors and paired sputum samples. In addition, *HYAL2* deletions were detected more frequently in sputum from cancer patients than in sputum from cancer-free smokers and were not detected in sputum from nonsmokers. Moreover, increased levels of *HYAL2* deletions in sputum were significantly associated with pack-years of smoking. Taken together, the findings suggested

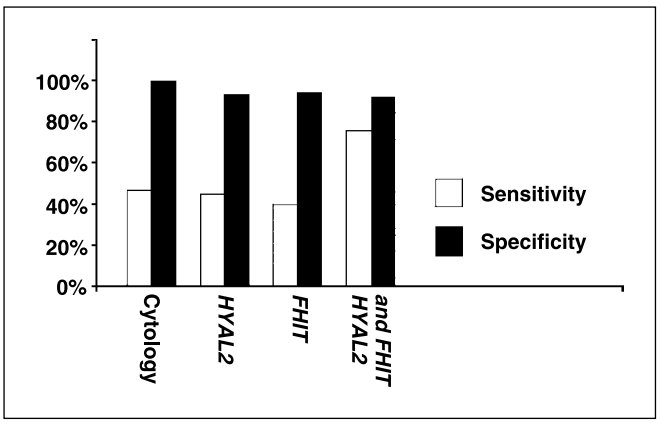

Fig. 2. Sensitivity and specificity of FISH and cytology of detecting abnormal cells in sputum specimens. Sensitivity of a test refers to the 38 lung cancer patients who have a positive result by the test. Specificity refers to the 64 people without lung cancer, including 36 cancer-free smokers and 28 healthy nonsmokers who have a negative result by the test. HYAL2, result of *HYAL2* probe; FHIT, result of *FHIT* probe; HYAL2 and FHIT, result from detecting the combination of *HYAL2* and *FHIT* deletions.

Table 3. Comparison of cytology and FISH with *HYAL2* or *FHIT* probes for the detection of lung cancer in sputum specimen

	38 sputum from lung cancer patients	
	18 cytology positive	20 cytology negative
No. sputum that are positive by *HYAL2* alone	8	5
No. sputum that are positive by *FHIT* alone	7	5
No. sputum that are positive by both *HYAL2* and *FHIT*	3	1

that tobacco exposure-related *HYAL2* deletion in sputum might be lung tumor–associated *HYAL2* deletion lesion and a biomarker for the diagnosis of lung cancer at its early stage.

The *FHIT* gene is located in chromosome 3p14.2, one of the most fragile human chromosome sites in human cancer (20). *FHIT* inactivation occurs in the majority of lung, gastric, cervical, and kidney carcinomas, and rescue of *FHIT* expression in such cancer cells suppresses their tumorigenicity, suggesting that *FHIT* is a tumor suppressor gene (20). In the current study, the presence of *FHIT* deletions was common in lung tumor tissues, and the prevalence of this genetic lesion in sputum was consistent with its prevalence in corresponding tumors from cancer patients. Although the findings from our study are consistent with previous reports (16, 20–23), we showed for the first time that detection of the smoking-related *FHIT* lesions in sputum might be a potential genetic test for the early detection of lung cancer.

In contrast to the distribution of *HYAL2* and *FHIT* deletions, the *SFTPC* deletion was found exclusively in the tumor specimens and did not appear in sputum samples from cancer patients, heavy smokers, and nonsmokers, suggesting that the *SFTPC* deletion might be more likely to be a late-stage lung tumor–associated marker. This agrees with previous reports in which *SFTPC* gene is located in chromosome 8p21.3, and deletion of the 8p21 occurs later in the process of lung tumorigenesis and is associated with the metastatic phenotype of the disease (24). Thus, the detection of *SFTPC* deletions might not be a biomarker in sputum for lung cancer diagnosis.

Although cytologic analysis of sputum failed to produce high accuracy in diagnosis of early-stage lung cancer, sputum has been considered as the best surrogate material for biomarker examinations for early detection of lung cancer (10, 11). Several molecular genetic approaches have been investigated in sputum to assist in the early diagnosis of lung cancer (25). Among them, interphase FISH has been considered as one of the most promising strategies because it can be directly applied to clinical samples and, thus, does not require microdissection or chromosomal and DNA and RNA preparation (26). Furthermore, the technique reveals cell-to-cell heterogeneity and allows the direct visualization of changes in chromosome or tumor-specific gene copy numbers. In addition, compared with other techniques, FISH is more quantitative because it detects genetic abnormalities by counting the number of signals that directly represent changes of gene. Moreover, if combined with cytologic analysis, FISH allows analysis of morphologic correlation with genetic aberrations. Recently, a commercial four-color FISH assay, LAVysion (Vysis), has been assessed for the diagnosis of lung cancer in sputum (14). However, the assay alone had low efficiency for the early detection of lung cancer because its sensitivity was not superior to that of cytology (14). The low efficiency of the test could be largely due to its probe selection, which was designed to detect the presence of chromosome aneusomy, whereas the development of aneuploidy may reflect late events in lung tumorigenesis (23). In the current study, the FISH test for the combined assessment of *HYAL2* and *FHIT* deletions detected abnormal cells not only in all the cytologically positive sputum but also in 55% cytologically negative specimens from lung cancer patients, suggesting that the FISH test can detect abnormal cells that escape cytologic examination. This observation was consistent with results of our previous study in which we directly compared efficiency of cancer cell detection between cytology, LAVysion, and dual-color FISH assays with probes designed for lung tumor–specific loci in bronchial brushings of early-stage NSCLCs (8). We showed that the frequency of allelic loss detected by dual-color FISH assay was much higher than that of aneuploidy detected by the LAVysion assay and of morphologic aberrations found by cytology. Together with the previous reports (8, 14), our current study, therefore, showed that the FISH test using lung cancer–associated genetic probes not only overcomes the limitation of the commercial LAVysion assay using chromosomal probes but also improves sputum cytology and, thus, shows great promise for increasing diagnostic value of sputum samples.

HYAL2 and *FHIT* deletions occur in the sputum of NSCLC patients with equal frequency among all histologic types and locations of lung tumor, suggesting that the genetic changes are not specific to histologic type and that detection of the abnormalities may be used to detect all histologic types of NSCLC, including peripheral tumors, such as adenocarcinomas. This is particularly important because adenocarcinomas, which arise from the smaller airways, are difficult to be detected by bronchoscopy or sputum cytology and have become more prevalent than squamous cell lung cancers worldwide.

Although the detection of combining *HYAL2* and *FHIT* deletions in sputum seems to be more sensitive than cytology in the detection of lung cancer, the sensitivity (76%) and specificity (92%) are still not yet efficient for routine clinical application. The result might indicate the need to develop a strategy for simultaneous assessment of a panel of tumor-specific biomarkers in sputum for highly sensitive and specific early detection of lung cancer. The fundamental mechanism supporting this premise is the "field defect" theory, which describes lung cancer as a disease of heterogeneity that develops from multiple foci and clonal aberrations. According to this theory, diagnostic efforts should be directed toward detection of multiple biomarkers rather than a single one to detect a clinically significant cancer phenotype with acceptable accuracy. To that end, we have developed a multiple FISH array that allows simultaneous analysis of 16 genetic probes (6). We will test sputum from lung cancer patients using the multiple FISH array with a panel of probes, including the ones for *HYAL2* and

FHIT, to determine the sensitivity and specificity of the assay and to establish diagnostic criteria for sputum specimens.

There was concordance between HYAL2 and FHIT deletions in tumor tissues; however, there was not such concordance in the paired sputum. One possible explanation for the observation might be the following. First, lung cancer is the end stage of multiple-step process characterized by the accumulation of numerous molecular genetic alterations (1). The genetic abnormalities increase with increasing severity of the tumorigenesis, and in tumor, there are accumulated multiple genetic changes, which usually exist together, including HYAL2 and FHIT deletions. Therefore, it is not surprising to find high concordance between the coexisting genetic changes in the tumor. Second, lung cancer develops from multiple, separate, clonally unrelated molecular genetic abnormalities that might occur in preneoplastic/preinvasive lesions of the airway; the clonally altered lesions of mildly abnormal bronchial epithelium are usually heterogeneous (22). Because HYAL2 and FHIT locate in two separate deleted sites of 3p, they may represent different clonal genetic aberrations and evolve more or less independently in preneoplastic/preinvasive lesions of lung cancer patients. Therefore, there was poor concordance for deletions of the genes detected in sputum specimens of the lung cancer patients, which contain exfoliated preneoplastic cells. The observation is also consistent with the finding by Park et al. (27) who showed that, although the tumors were homogeneous for some molecular changes, the clonally altered patches of mildly abnormal bronchial epithelium were usually heterogeneous. Furthermore, the observation of poor concordance between 3p deletions in sputa also provides strong evidence that a single biomarker unlikely provides sufficient sensitivity or specificity for the early detection of lung cancer in the sputum.

The *HYAL2* and *FHIT* deletions were found in four and three sputum from cancer-free smokers, respectively, thus resulting in a 9% false-positive rate. However, the deletions did not occur in the sputum obtained from healthy nonsmokers and there was an association of the deletions with number of smoking pack-years in cancer patients and heavy smokers, suggesting that the deletions might reflect chronic exposure to tobacco carcinogens. Therefore, long-term follow-up of the deletion-positive smokers with no malignant disease is needed to determine whether the result is real false positive or represents an early indicators of lung malignancy developed from tobacco-related epithelial damages before it is detectable by other means.

In summary, both *HYAL2* and *FHIT* deletions in sputum may constitute tumor-associated diagnostic biomarkers for early-stage lung cancer. Longitudinal studies of the deletion status in large populations and evaluation of additional genetic markers already identified in primary lung cancer are required to develop a panel of markers for reliable molecular genetic diagnostic tests that can be rapidly done on sputum for routine diagnosis of lung cancer at its early stage.

Acknowledgments

We thank Mildred Michalisko of the Department of Pathology for editorial review of this manuscript.

References

1. Hirsch FR, Brambilla E, Gray N, et al. Prevention and early detection of lung cancer—clinical aspects. Lung Cancer 1997;17:163–74.
2. Fontana RS, Sanderson DR, Woolner LB, Taylor WF, Miller WE, Muhm JR. Lung cancer screening: the Mayo program. J Occup Med 1986;28:746–50.
3. Frost JK, Ball WC, Jr., Levin ML, et al. Early lung cancer detection: results of the initial (prevalence) radiologic and cytologic screening in the Johns Hopkins study. Am Rev Respir Dis 1984;130:549–54.
4. Jiang F, Yin Z, Caraway NP, Li R, Katz RL. Genomic profiles in stage I primary non small cell lung cancer using comparative genomic hybridization analysis of cDNA microarrays. Neoplasia 2004;6:623–35.
5. Li R, Wang H, Bekele BN, Jiang F. Identification of putative oncogenes in lung adenocarcinoma by a comprehensive functional genomic approach. Oncogene 2005;18:2628–35.
6. Li R, Liu Z, Wang Z, Jiang F. Development of an *in situ* DNA chip and demonstration of its utility in the diagnosis of cancer. Lab Invest 2006;86:619–27.
7. Jiang F, Caraway NP, Nebiyou Bekele B, et al. Surfactant protein A gene deletion and prognostics for patients with stage I non-small cell lung cancer. Clin Cancer Res 2005;11:5417–24.
8. Barkan GA, Caraway NP, Jiang F, et al. The potential use of FISH to identify predictive markers in early-stage lung carcinomas. Cancer 2004;105:35–43.
9. Peters EJ, Morice R, Benner SE, et al. Squamous metaplasia of the bronchial mucosa and its relationship to smoking. Chest 1993;103:1429–32.
10. Thunnissen FB. Sputum examination for early detection of lung cancer. J Clin Pathol 2003;56:805–10.
11. Belinsky SA. Gene-promoter hypermethylation as a biomarker in lung cancer. Nat Rev Cancer 2004;4:707–17.
12. Saccomanno G, Archer VE, Auerbach O, Saunders RP, Brennan LM. Development of carcinoma of the lung as reflected in exfoliated cells. Cancer 1974;33:256–70.
13. Saccomanno G. Diagnostic Pulmonary Cytology. 2nd ed. Chicago (IL): American Society of Clinical Pathologists Press; 1986. p. 1–211.
14. Varella-Garcia M, Kittelson J, Schulte AP, et al. Multitarget interphase fluorescence *in situ* hybridization assay increases sensitivity of sputum cytology as a predictor of lung cancer. Cancer Detect Prev 2004;28:244–51.
15. Jiang F, Lin F, Price R, Katz RL. Rapid detection of IgH/BCL2 rearrangement in follicular lymphoma by interphase fluorescence *in situ* hybridization with bacterial artificial chromosome probes. J Mol Diagn 2002;4:144–9.
16. Burke L, Khan MA, Freedman AN, et al. Allelic deletion analysis of the FHIT gene predicts poor survival in non-small cell lung cancer. Cancer Res 1998;58:2533–6.
17. Dodd LE, Pepe MS. Partial AUC estimation and regression. Biometrics 2003;59:614–23.
18. Lerman MI, Minna JD. The 630-kb lung cancer homozygous deletion region on human chromosome 3p21.3: identification and evaluation of the resident candidate tumor suppressor genes. The International Lung Cancer Chromosome 3p21.3 Tumor Suppressor Gene Consortium. Cancer Res 2000;60:6116–33.
19. Senchenko VN, Liu J, Loginov W, et al. Discovery of frequent homozygous deletions in chromosome 3p21.3 LUCA and AP20 regions in renal, lung, and breast carcinomas. Oncogene 2004;29:5719–28.
20. Croce CM, Sozzi G, Huebner K. Role of FHIT in human cancer. J Clin Oncol 1999;17:1618–24.
21. Ho WL, Chang JW, Tseng RC, et al. Loss of heterozygosity at loci of candidate tumor suppressor genes in microdissected primary non-small cell lung cancer. Cancer Detect Prev 2002;26:343–9.
22. Wistuba II, Behrens C, Virmani AK, et al. High resolution chromosome 3p allelotyping of human lung cancer and preneoplastic/preinvasive bronchial epithelium reveals multiple, discontinuous sites of 3p allele loss and three regions of frequent breakpoints. Cancer Res 2000;60:1949–60.
23. Barrett MT, Sanchez CA, Galipeau PC, Neshat K, Emond M, Reid BJ. Allelic loss of 9p21 and mutation of the CDKN2/p16 gene develop as early lesions during neoplastic progression in Barrett's esophagus. Oncogene 1996;9:1867–73.
24. Kim TM, Yim SH, Lee JS, et al. Genome-wide screening of genomic alterations and their clinicopathologic implications in non-small cell lung cancers. Clin Cancer Res 2005;11:8235–42.
25. McWilliams A, Mayo J, MacDonald S, et al. Lung cancer screening: a different paradigm. Am J Respir Crit Care Med 2003;168:1167–73.
26. Jiang F, Katz RL. Use of FISH as a powerful diagnostic tool in cytology. Diagn Mol Pathol 2002;11:47–57.
27. Park IW, Wistuba II, Maitra A, et al. Multiple clonal abnormalities in the bronchial epithelium of patients with lung cancer. J Natl Cancer Inst 1999;91:1863–8.

Clinical Cancer Research

Genetic Deletions in Sputum as Diagnostic Markers for Early Detection of Stage I Non–Small Cell Lung Cancer

Ruiyun Li, Nevins W. Todd, Qi Qiu, et al.

Clin Cancer Res 2007;13:482-487.