A 5-MicroRNA Signature for Lung Squamous Cell Carcinoma Diagnosis and hsa-miR-31 for Prognosis

Xiaogang Tan¹, Wenyan Qin³,⁵,⁶, Liang Zhang⁵,⁶, Jie Hang¹, Baozhong Li¹, Cuiyan Zhang¹, Junting Wan¹, Fang Zhou¹, Kang Shao¹, Yimin Sun¹,², Jiaping Wu³,⁶, Xin Zhang³,⁶, Bin Qiu³, Ning Li¹, Susheng Shi¹, Xiaoli Feng², Shouhua Zhao¹, Zong Wang¹, Xiaohong Zhao¹, Zhaoli Chen¹, Keith Mitchelson³,⁵,⁶, Jing Cheng³,⁴,⁵,⁶, Yong Guo³,⁵,⁶, and Jie He¹

Abstract

Purpose: Recent studies have suggested that microRNA biomarkers could be useful for stratifying lung cancer subtypes, but microRNA signatures varied between different populations. Squamous cell carcinoma (SCC) is one major subtype of lung cancer that urgently needs biomarkers to aid patient management. Here, we undertook the first comprehensive investigation on microRNA in Chinese SCC patients.

Experimental Design: MicroRNA expression was measured in cancerous and noncancerous tissue pairs strictly collected from Chinese SCC patients (stages I–III), who had not been treated with chemotherapy or radiotherapy prior to surgery. The molecular targets of proposed microRNA were further examined.

Results: We identified a 5-microRNA classifier (hsa-miR-210, hsa-miR-182, hsa-miR-486-5p, hsa-miR-30a, and hsa-miR-140-3p) that could distinguish SCC from normal lung tissues. The classifier had an accuracy of 94.1% in a training cohort (34 patients) and 96.2% in a test cohort (26 patients). We also showed that high expression of hsa-miR-31 was associated with poor survival in these 60 SCC patients by Kaplan–Meier analysis (P = 0.007), by univariate Cox analysis (P = 0.011), and by multivariate Cox analysis (P = 0.011). This association was independently validated in a separate cohort of 88 SCC patients (P = 0.008, 0.011, and 0.003 in Kaplan–Meier analysis, univariate Cox analysis, and multivariate Cox analysis, respectively). We then determined that the tumor suppressor Dicer1 is a target of hsa-miR-31. Expression of hsa-miR-31 in a human lung cancer cell line repressed Dicer1 activity but not Ppp2r2a or Lats2.


Introduction

Worldwide, there are more than 1,600,000 new cases of lung cancer and more than 1,370,000 attributable deaths each year, ranking it as the leading cause of cancer mortality. In China, lung cancer is the most frequent of the malignant tumors, initiating 520,000 new cases and causing 450,000 deaths per year; furthermore, the age-standardized incidence rate and mortality rate adjusted for the world standard population are both higher than the worldwide average levels (1). Non–small cell lung cancers (NSCLC) are the major types of lung cancer found worldwide (~80%). Among NSCLCs, squamous cell carcinoma (SCC, 40%) is one of the main pathologic subtypes in China. SCC patients with clinical stage IA disease have a 5-year survival rate of about 80%, whereas the 5-year survival rate for clinical stage II to IV disease ranges from 40% to less than 5% (2). The poor outcome of SCC patients could be explained, in part, by the lack of early diagnosis markers and the lack of prognostic indicators. Therefore, there is an urgent need to develop biomarkers for diagnosis and prognosis, to benefit Chinese SCC patients. Transbronchial lung biopsy (TBLB) has been explored as a method for early diagnosis of questionable lung infiltrations or densities; however, the diagnostic yield of TBLB is lower than that of open lung biopsy because the small amount of lung tissues

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).
MicroRNA Expression in Lung SCC

Translational Relevance

The poor outcome of patients with squamous cell carcinoma (SCC) is exacerbated by the lack of early diagnosis markers and the lack of prognostic indicators. The 5-microRNA classifier distinguishing SCC from normal lung tissues holds promise for early diagnosis of SCC among Chinese. The classifier may potentially improve the diagnostic value of transbronchial lung biopsy. hsa-miR-31 was identified conclusively as a prognostic factor and may be of use for the clinical management of SCC.

Obtained in biopsy. New biomarkers with high diagnostic potential to distinguish SCC carcinous from normal lung tissue should aid the acute diagnostic yield. Furthermore, novel prognostic biomarkers will ultimately allow SCC patients to receive personalized therapies.

MicroRNAs are a species of small (19–22 nucleotides) noncoding single-stranded RNA molecules that through partial sequence homology may interact with the 3’-untranslated region (3’-UTR) of target mRNA molecules (3, 4). Different microRNAs may function as tumor suppressors or oncogenes, and the deregulation of their expression is associated with the initiation and progression of cancers via the activation and/or repression of controlling pathways (5). Several recent studies have shown that selected microRNAs contributing to NSCLC progression can also be used to estimate prognosis (6–11). Although microRNA profiling of NSCLC has contributed to our understanding of the biology of these cancers, the microRNA signatures identified in NSCLC (including adenocarcinoma and SCC) were not consistent among different clinical studies. The variations were likely caused by factors such as the differences among the samples and methodologies used, including the ethnicity of patient groups, the pathologic subtypes collected, the sample collection methods, the technology platforms [microarray or quantitative reverse transcription PCR (qRT-PCR)], and the bioinformatics analysis used. The importance of carefully designed and appropriately controlled studies is illustrated by the well-established, although very different, biomarker profiles used currently to assess breast cancer prognosis with well-established, although very different, biomarker profiles used currently to assess breast cancer prognosis with the technology platforms [microarray or quantitative reverse transcription PCR (qRT-PCR)], and the bioinformatics analysis used. The importance of carefully designed and appropriately controlled studies is illustrated by the well-established, although very different, biomarker profiles used currently to assess breast cancer prognosis with Agenda’s MammaPrint and Genomic Health’s Oncotype DX genomic assays. The MammaPrint assay comprises 70 genes that were found to be related to distant recurrence of breast cancer, whereas Oncotype DX evaluates the activity of 21 genes that indicate the risk of breast cancer recurrence. Only one gene is common between the 2 tests, yet both have been used effectively by thousands of doctors to help guide treatment of more than 200,000 patients (http://www.oncotypedx.com/en-US/Breast/Healthcare-Professional/Overview.aspx, http://investor.genomic-health.com/release_detail.cfm?ReleaseID=386285, and http://www.mammaprint.co.uk/MammaPrint_Patient_Brochure_EU_lowres.pdf). To identify promising microRNA biomarkers for lung cancer subtypes, it is necessary that the study is critically designed and that an established and proven methodology is employed. We also ensured that only a specific pathologic cancer subtype (SSC) in one ethnic group (Han Chinese) was examined, and we carefully chose the best of collected patient samples to limit outlier contamination. The microarray platform and data analysis methods were also well established.

We conducted a genome-wide microRNA study (677 human microRNA probes) to ensure a broad profiling among cellular microRNAs. Principal component analysis (PCA) and support vector machine (SVM) analysis tools were then used to establish a powerful microRNA classifier that distinguished SCC from normal tissue: a minimal 5-element classifier (hsa-miR-210, hsa-miR-182, hsa-miR-486-5p, hsa-miR-30a, and hsa-miR-140-3p) was found. The patients chosen for our study had comprehensive follow-up data that allowed prognostic analysis, which then revealed that the high expression of hsa-miR-31 was associated with poor survival from SCC. This marker was confirmed by Kaplan–Meier analysis, univariate Cox analysis, and multivariate Cox analysis. Finally, we showed a molecular mechanism by which hsa-miR-31 might contribute to poor patient prognosis. hsa-miR-31 targets the tumor suppressor DICER1, resulting in the repression of its expression.

Materials and Methods

Patients and samples

The patients examined in this study underwent surgery at the Cancer Institute and Hospital, Chinese Academy of Medical Sciences, between 2000 and 2002. All patients (Supplementary Table S1) had not been previously treated by chemotherapy and radiotherapy when undergoing surgery and provided informed consent to participate in the study. The study was approved by the medical ethics committee of the Cancer Institute and Hospital, Chinese Academy of Medical Sciences.

Samples were snap frozen in liquid nitrogen immediately after resection and stored (minimum of 5 years) at −80°C until the extraction of RNA. Peripheral portions of the resected lung samples were paraffin embedded, sectioned, and hematoxylin and eosin stained by routine methods. The tumor cell concentrations were evaluated, and the tumor histology was independently confirmed by 2 pathologists (S. Shi and X. Feng). Follow-up information was extracted from the follow-up registry of the Cancer Institute and Hospital, Chinese Academy of Medical Sciences. For all the samples, clinicopathologic information [smoking, age, gender, pathologic subtype, tumor node metastasis (TNM) classification, tumor stage, lymph node stage, differentiation status, adjuvant therapy after surgery, and the duration of survival after surgery] was available.

We initially profiled the microRNA expression in a training cohort of 34 SCC patients for classifier establishment, using paired cancerous and adjacent normal samples, and then examined an additional 26 SCC patients as a
test cohort. The microRNA expression in this total of 60 SCC patients was used for prognostic analysis of their related survival data, and a further cohort of 88 SCC patients was used as an independent validation of the prognostic analysis.

**MicroRNA microarray analysis**

The microRNA analyses were conducted on a microarray using the CapitalBio platform (CapitalBio Corp.) that has been used in numerous previous studies (12, 13). Total RNA was extracted with TRIzol reagent (Invitrogen), and the low-molecular-weight RNA was isolated by a polyethylene glycol (PEG) solution precipitation method (14). The low-molecular-weight RNA was labeled by the T4 RNA ligase labeling method described by Thomson and colleagues (15). In brief, 4 µg of low-molecular-weight RNA isolated by a PEG solution precipitation method was labeled with 500 ng of 5'-phosphate-cytidyl-urydyl-cy3-3', using the T4 RNA ligase. The labeled RNA was hybridized with aldehyde-modified slide containing 924 mature mammalian microRNA probes (including 677 human microRNA sequences), with each probe present in triplicate spots. A 3-dimensional tilting agitator BioMixer II (CapitalBio) was used for overnight hybridization to increase probe signal intensity and improve signal uniformity. An automated SlideWarsher-8 (CapitalBio) was used to wash and dry the hybridized slides to reduce the dye blemishes that occur frequently during manual posthybridization slide washing. Arrays were scanned with a LuxScan 10K-A laser confocal scanner, and the images obtained were then analyzed with LuxScan 3.0 software (both from CapitalBio).

**Bioinformatics analysis**

For all samples, the average background values were subtracted from each of the replicate microRNA spots. Faint spots were filtered out if the expression signal was less than 1,500. A space- and intensity-dependent median-center normalization was done with Bioconductor. Differential expression was analyzed with the significance analysis of microarrays (SAM) (http://www-stat.stanford.edu/~tibs/SAM/index.html; ref. 16). The sufficient but minimal microRNA marker set to classify the cancerous and adjacent normal tissues was analyzed by the PCA and SVM methods (12). The 632 Bootstrap method was used to estimate the accuracy of each predicted model using the T4 RNA ligase. The labeled RNA was hybridized with aldehyde-modified slide containing 924 mature mammalian microRNA probes (including 677 human microRNA sequences), with each probe present in triplicate spots. A 3-dimensional tilting agitator BioMixer II (CapitalBio) was used for overnight hybridization to increase probe signal intensity and improve signal uniformity. An automated SlideWarsher-8 (CapitalBio) was used to wash and dry the hybridized slides to reduce the dye blemishes that occur frequently during manual posthybridization slide washing. Arrays were scanned with a LuxScan 10K-A laser confocal scanner, and the images obtained were then analyzed with LuxScan 3.0 software (both from CapitalBio).

**qRT-PCR analysis**

Total cellular RNAs were also subjected to qRT-PCR with microRNA specific primers for determination of microRNA expression. Total RNA extraction, reverse transcription, and quantitative PCR were done according to the procedures described previously (12, 22). The primers for hsa-miR-31 are as follows: RT primer, 5'-GCCTATCCACTG-CAGGGTCGCCGATTTGCACGGTATACGACAGGAT-3'; amplification primer, 5'-AGGCGAGATGCCTGGCATA-GCT-3' and 5'-CAAGCGAAGATGCTTGCCGATA-3'. For the quantification of Dicer1 mRNA, total cellular RNA was reverse transcribed with oligo(dT) primers. Amplification of Dicer1 was done with primers 5'-CCTCCCAT-CAACATACCAGTAGA-3' and 5'-ACGAGCATCAATGCAAAATC-3'. The highly conserved and universally expressed small nuclear RNA U6 and β-actin genes were used as endogenous qRT-PCR controls for hsa-miR-31 and Dicer1, respectively.

**Cell culture and transient transfection**

Human lung squamous cell carcinoma SK-MES-1 cells (2 × 10^5 per well) were plated into a 6-well plate 1 day before transfection. hsa-miR-31 or control microRNA (Dharmacon) at a concentration of 100 nmol/L was transfected into cells with Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. Cell lysates were prepared for Western blot assay at 48 hours posttransfection and luciferase reporter assay at 24 hours posttransfection.

**Western blot analysis**

SK-MES-1 cell lysates (20 µg) were separated by 10% SDS-PAGE and electroblotted onto polyvinylidene fluoride (PVDF) membrane (Pall). The PVDF membrane was blocked with TBS Tween-20 buffer containing 5% milk powder, and reaction with primary antibody was carried out at 4°C overnight with gentle shaking. After washing, secondary antibody conjugated with horseradish peroxidase (HRP) was added, and the reaction was carried out for...
1 hour at room temperature. SuperSignal West Pico Chemiluminescent Substrate (Pierce) was used for the detection according to the manufacturer’s instructions. Western blot results were scanned and quantified by NIH ImageJ software (http://rsbweb.nih.gov/ij/). Protein expression levels were normalized by the expression levels of β-tubulin.

Luciferase reporter assay
The 3′-UTRs of predicted microRNA target genes were amplified by PCR and cloned into the XbaI site of pRL-SV40 vector (Promega). Primers for the amplification of (i) DICER1 are 5′-GAGTGAATTTAGGACGAA-3′ and 5′-ATCAGCTGCTCAAGCTTAATG-3′, (ii) PPP2R2A 3′-UTR are 5′-TTTCAGAAGCAGTTAGGGTG-3′ and 5′-ACAGGAGGAGTATACCCG-3′, and (iii) LATS2 3′-UTR are 5′-GCTTCAATAGGCTTTTCAAGACC-3′ and 5′-CATACGAGACTACAGAAACGGAC-3′. The mutation of DICER1 3′-UTR sequence was undertaken with the QuickChange Site-Directed Mutagenesis Kit (Stratagene) with primers 5′-AGCTTCTGATCATAAAGAGGTTCTAGCTTAAAGTGAAAAGCTCATTGATGAGCT-3′ and 5′-ATCAGCTGACTGAGAAGCTTTTATGCAGACGT-3′. SK-MES-1 cells were transfected with a mixture of pGL3-control, pRL-SV40-3′-UTR, and miR-31 or control microRNA (Dharmacon). Cells were lysed at 24 hours posttransfection. Firefly and Renilla luciferase activities were assayed with the Dual-Glo Luciferase Assay System (Promega) according to the manufacturer’s instructions. Data were normalized to the control microRNA.

Immunohistochemical assay
Anti-DICER1 (catalog #ab82539) was purchased from Abcam. HRP-labeled goat anti-mouse IgG was purchased from Abgent. Immunohistochemistry was carried out on 4-μm, formalin-fixed, paraffin-embedded (FFPE) sections with anti-DICER1 as described (23). Briefly, the slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes. The slides were deparaffinized, rehydrated, and dripped in 3% hydrogen peroxide solution for 10 minutes.

Counterstaining was carried out with hematoxylin. Negative control was obtained by substituting the primary antibody with PBS. Immunohistochemical results were evaluated by 2 pathologists (S. Shi and X. Feng) without knowledge of patient information. The expression was assigned to one of the following categories according to the percentage of positive cells: 0 (0%–5%), + (6%–25%), ++ (26%–50%), and +++ (51%–100%). The expression was considered negative when the score was 0 and positive when the score was + or greater. An unpaired Student t test was used for analysis of data between groups. Values of P < 0.05 were considered statistically significant.

Results
A 5-microRNA signature can efficiently distinguish malignant SCC tissues from adjacent normal tissues in Chinese
We first investigated whether the microRNA expression could identify malignant SCC tissues in Chinese patients by examining primary lung cancers and their corresponding adjacent normal lung tissues collected a minimum distance of 5 cm from the tumor in 34 patients (training cohort). Twenty-two microRNAs (Supplementary Table S2) were found differentially expressed between cancer and normal tissues, which were further analyzed by the PCA–SVM strategy, and a minimum of 5-microRNA classifier set (hsa-miR-210, hsa-miR-182, hsa-miR-486-5p, hsa-miR-30a, and hsa-miR-140-3p), with the highest distinguishing values, was identified. The analysis of the original training group using this classifier had a predictive accuracy of 94.1%. The classifier was subsequently validated with an independent test cohort comprising another 26 SCC patients and displayed an accuracy of 96.2% with this group (Fig. 1). In the SCC tissues, 2 of the microRNAs (hsa-miR-210 and hsa-miR-182) were upregulated and 3 microRNAs (hsa-miR-486-5p, hsa-miR-30a, and hsa-miR-140-3p) were downregulated. These analyses indicated that a classifier with as few as 5-microRNA markers could efficiently distinguish malignant SCC tissues from normal tissues in Chinese patients.

High expression of hsa-miR-31 correlates strongly with low survival in Chinese SCC patients
We then examined the 22 differentially expressed microRNAs in 60 SCC patients (clinical characteristics listed in Table 1 and Supplementary Table S1) using Kaplan–Meier survival analysis for potential correlations between microRNA expression and survival prognosis. Kaplan–Meier analysis showed that a high expression of hsa-miR-31 was associated with poor survival in SCC patients (P = 0.007; log-rank test, 60 SCC patient training cohort; Fig. 2A). Subsequently, univariate Cox analysis with hsa-miR-31 and clinicopathologic factors (smoking, age, gender, differentiation, TNM classification, and adjuvant therapy after surgery) revealed that the hsa-miR-31 expression level had prognostic significance in the training cohort (P = 0.011; risk ratio: 2.827; 95% CI, 1.274–6.275; Table 2). Furthermore, multivariate Cox proportional hazard regression analysis of each of these parameters indicated that high hsa-miR-31 expression was a significantly unfavorable prognostic factor independent of all other clinicopathologic factors (P = 0.011; risk ratio 3.535; 95% CI, 1.274–6.275; Table 2). These results showed that high levels of hsa-miR-31 expression correlated strongly with poor survival of Chinese SCC patients.

Validation of the survival correlation of hsa-miR-31 with an independent cohort of Chinese SCC patients
The correlation of hsa-miR-31 with the prognosis of SCC was examined further with an independent set of 88 SCC
Tan et al.

patients. The hsa-miR-31 expression was determined by qRT-PCR. Kaplan–Meier survival analysis confirmed a significantly decreased survival of patients with high hsa-miR-31 expression \( (P = 0.008; \) log-rank test, 88 SCC patient test cohort; Fig. 2B). Both univariate \( (P = 0.011; \) risk ratio, 2.396; 95% CI, 1.218–4.715) and multivariate \( (P = 0.003; \) risk ratio, 3.071; 95% CI, 1.451–6.500) Cox proportional hazard regression analyses suggested that high hsa-miR-31 expression was an independent predictor of poor prognosis for SCC in Chinese (Table 3).

hsa-miR-31 targets DICER1 but not PPP2R2A and LATS2

We finally investigated the potential molecular targets of hsa-miR-31 by analysis of the target genes predicted by 4 algorithms (miRBase, MiRanda, TargetScan, and PicTar). We selected DICER1, PPP2R2A, and LATS2 for further investigation. The duplexes predicted to form between hsa-miR-31 and the 3 target genes are depicted in Figure 3A. Immunoblot analysis of DICER1, PPP2R2A, and LATS2 proteins showed that only DICER1 protein was reduced in SK-MES-1 cells transiently transfected with hsa-miR-31 (Fig. 3B). We carried out a dual luciferase reporter assay to confirm whether hsa-miR-31 target DICER1 mRNA. A significant decrease in luciferase activity was observed when we cotransfected hsa-miR-31 and luciferase reporter gene linked to the wild-type 3'-UTR of DICER1, whereas specific mutation of the hsa-miR-31 binding sequence in DICER1 3'-UTR resulted in only slightly downregulation of the reporter gene (Fig. 3C). We also constructed luciferase reporters linked to the 3'-UTR of PPP2R2A and of LATS2, but no significant effects on these reporter genes were observed in the presence of cotransfected hsa-miR-31 (data not shown). The immunoblot and reporter assays independently confirmed that DICER1 is a target of hsa-miR-31 in human cells. Interestingly, quantification of DICER1 mRNA indicated a certain downregulation of the level of DICER1 mRNA expression in SK-MES-1 cells transfected with hsa-miR-31 as compared with control microRNA, even though the \( P \) value is not significant, suggesting that hsa-miR-31 may act mainly by inhibiting the translation of DICER1 (Fig. 3D).

Expression of DICER1 is inversely correlated with hsa-miR-31 level in SCC tumor tissues

To further determine the relationship between DICER1 and hsa-miR-31 in SCC tumor tissues, we carried out the immunohistochemical assay in 46 SCC samples for which the corresponding FFPE tumor tissues were available. We found that the level of DICER1 was markedly lower in SCC tumor tissues with high miRNA expression. The DICER1 expression was scored negative in 83% (20 of 24) tumor tissues with high hsa-miR-31 levels, whereas only 58% (12 of 22) tumor tissues with low hsa-miR-31 level were negative for DICER1 \( (P = 0.036; \) Fig. 3E and F). These results indicated that the expression of DICER1 is inversely correlated with hsa-miR-31 level in SCC tumor tissues.

Discussion

The identification and validation of novel biomarkers for different cancers comprise a significant area of practical cancer research. Multiplex molecular diagnostic and prognostic tests for cancers have been intensively investigated during the past 10 years, and now including genomic DNA, mRNAs, microRNAs, DNA methylation, and proteins. Two demonstrable analytic successes are Agendia’s Mammaprint and Genomic Health’s Oncotype DX genomic assays, which are both used by thousands of doctors in clinical practice. Several recent clinical studies have proposed sets of differentially expressed microRNAs as promising markers for monitoring the development of lung cancer and its prognosis, in particular for NSCLC \( (6–11) \). An important consideration is how to achieve a reliable and universal

Figure 1. A 5-microRNA classifier could distinguish malignant SCC lesions from normal tissues. The PCA–SVM strategy was used to construct a classifier, which was used to classify the cancerous tissue and adjacent normal tissue. Red dots represent the cancerous tissue and blue crosses represent the adjacent normal tissue. The classifier had a predictive accuracy of 94.1% in the original training cohort (34 SCC patients; A) and 96.2% in the independent test cohort (another 26 SCC patients; B).
signature, considering that results from different studies that identified biomarker species often varied. The reason for this occurrence is complex but likely arises from the differences in patient ethnicities, pathologic sample subtypes, sample collection methods, technology platforms, bioinformatics approaches, as well as the different experience and expertise of the investigators with each of the methodologic variables.

We, therefore, took particular care in the design of this study to ensure the identification of reliable biomarkers and used only established methods that reveal valuable information. We assayed only Chinese SCC patients to minimize ethnicity and pathologic subtype effects and to amplify the molecular homogeneity of tumor specimens. The patients were strictly selected from individuals who had not been previously treated with chemotherapy or radiotherapy when undergoing surgery to avoid therapy influences. A microarray containing 677 mature human microRNA probes was used to determine expression levels. Although this microarray does not represent the latest newly discovered microRNAs (The miRBase database version 16 contains 1,212 mature human microRNA; http://www.mirbase.org/), the study still revealed new and important information about microRNA useful for the management of SCC. The MicroArray Quality Control Consortium (MAQC I) study previously reported the high reproducibility of expression analysis results produced by the microarray platform our group used here (24). More recently, the MAQC II study elucidated that the execution of a gene expression classifier depended largely on the proficiency of the researchers and on the different methods used (25). The PCA–SVM strategies used here for lung cancer microRNA profiling analysis are well established in our group and were used previously in a study of esophageal cancer (12).

The present study suggests that a minimal 5-microRNA classifier, consisting of hsa-miR-210, hsa-miR-182, hsa-miR-486-5p, hsa-miR-30a, and hsa-miR-140-3p, can distinguish malignant SCC tissues from normal tissues. This classifier is entirely new and has important molecular pathologic applications, because changes in the levels of the classifier species in tumor tissues may be detected as biomarkers. One particular powerful application of the 5-microRNA molecular classifier could be with TBLB, which has been explored as a method for early diagnosis of questionable lung infiltrations or densities (26). The diagnostic yield of TBLB is lower than that of open lung biopsy because the amount of lung tissue obtained is smaller.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Training cohort (n = 60)</th>
<th>Test cohort (n = 88)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>17–77</td>
<td>42–81</td>
</tr>
<tr>
<td>Median</td>
<td>59.0</td>
<td>62.5</td>
</tr>
<tr>
<td>Sex, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>50 (83)</td>
<td>79 (90)</td>
</tr>
<tr>
<td>Female</td>
<td>10 (17)</td>
<td>9 (10)</td>
</tr>
<tr>
<td>Smoking, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoking</td>
<td>49 (82)</td>
<td>60 (68)</td>
</tr>
<tr>
<td>Nonsmoking</td>
<td>11 (18)</td>
<td>28 (32)</td>
</tr>
<tr>
<td>Stage (TNM) classification, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>21 (35)</td>
<td>27 (31)</td>
</tr>
<tr>
<td>II</td>
<td>17 (28)</td>
<td>31 (35)</td>
</tr>
<tr>
<td>III</td>
<td>22 (37)</td>
<td>30 (34)</td>
</tr>
<tr>
<td>Differentiation, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well</td>
<td>5 (8)</td>
<td>4 (5)</td>
</tr>
<tr>
<td>Moderate</td>
<td>33 (55)</td>
<td>44 (50)</td>
</tr>
<tr>
<td>Poor</td>
<td>22 (37)</td>
<td>40 (45)</td>
</tr>
</tbody>
</table>

Table 1. Clinical characteristics of the training and test cohorts for prognostic analysis in SCC patients

![Figure 2. Kaplan–Meier survival curves for SCC patients. High levels of hsa-miR-31 expression correlated with a lower survival rate in the training cohort (60 SCC patients; A), whereas low expression correlated with a high survival rate in an independent test cohort (88 SCC patients; B). The log-rank test P value was 0.007 for the training cohort and 0.008 for the test cohort.](#)
However, molecular detection methods require less amounts of tissue than traditional pathology and could improve the diagnostic accuracy of endoscopic procedures. The use of the microRNA classifier to improve the diagnostic yield after TBLB and avoid unnecessary surgery would be a significant advance. Although the TBLB tissue is scant, the RNA from the sample is sufficient for microRNA amplification and analysis, which could use U6 RNA from the same sample as an internal control for normalization. An alternative technical approach is to compare the expression of microRNAs from patients with common lung tissue control, which could be supplied in a kit form. It is not necessary to take the normal tissue from the same patient for analysis. A custom-based array chip could be designed to determine the 5-microRNA expression levels, using experiences learned from our previous array-based molecular diagnostic studies (27–29).

Furthermore, our study identified that the elevated expression of hsa-miR-31 correlates inversely with the survival of Chinese SCC patients. This single marker could be conveniently measured by qRT-PCR. It is notable that hsa-miR-31 is not a member of the 5-microRNA classifier identified in the same patient samples. The reason might be that the molecules involved in the initial occurrence of SCC and the later malignant effects are different. We also observed that TNM staging (I/II/III) was not consistently correlated with the SCC survival in our cohorts (Tables 2 and 3). Resection surgery cannot be successfully done on patients with TNM stage IV and part IIIb lung cancer; hence, samples of these stages were not available to us. The absence of these stage samples raises the possibility of unknown statistical bias in our Cox regression analysis. Yet, our findings suggest it may be beneficial to incorporate molecular markers into the lung cancer staging system to aid personalized treatment of patients, particularly at early to intermediate stages.

Recently, Yang and colleagues reported a profiling study on microRNA expression in SCC among 23 Chinese patients (30). Their conclusions were not concordant with our findings. Our study was the first comprehensive investigation of a statistically significant cohort of 148 Chinese patients with extensive clinicopathologic data. Interestingly, the 5 microRNAs comprising our SCC classifier have individually been reported to have roles in 3 other studies. Mascaux and colleagues reported a microRNA expression during bronchial squamous carcinogenesis (31). They used a total of 60 Caucasian biopsies, with 6 samples from each category: normal bronchial epithelium of nonsmokers, normal normofluorescent bronchial epithelial tissue of smokers, histologically normal but hypofluorescent bronchial epithelium of smokers, hyperplasia, metaplasia, mild, moderate, and severe dysplasia, carcinoma in situ, and SCC. The hsa-miR-486 was found dysregulated during lung carcinogenesis. Raponi and colleagues studied the prognostics of SCC using a cohort of 61 Caucasian SCC samples and 10 matched normal lung samples (32). hsa-miR-210 and hsa-miR-182 were among the microRNAs differentially expressed between SCC and normal lung samples in their cohort. They also reported a different microRNA from ours, hsa-miR-146b alone having prognostic value for SCC. In addition, hsa-miR-210 was one of the microRNA markers found in sputum by Xing and colleagues for early detection of SCC in Caucasians (33). The fact that some of our 5 microRNAs are common to 3 other studies with Caucasian patients is persuasive and it brings some confidence to the universality of our observations, but this assertion must be tested further.

MicroRNA-31 has been shown to be an oncogene in a murine cell line (ED-1 cell) through its repression of 2 tumor suppressors, PPP2R2A and LATS2 (34). Here, we examined the molecular role of hsa-miR-31 in a human SCC lung cancer cell line (SK-MES-1 cell) and found different results, which may be due to using cell lines from different species. Interestingly, our findings indicated that hsa-miR-31 suppressed the expression of DICER1 in the human SCC cell line, but no effect was found on PPP2R2A or LATS2. The expression of DICER1 has been shown to be downregulated in human lung cancer and its reduced expression correlates with shortened postoperative survival, which is consistent with our data (35). This novel

### Table 2. Postoperative survival of SCC patients in relation to clinicopathologic characteristics and microRNA expression analyzed by the Cox proportional hazard regression model in a training cohort of 60 cases

<table>
<thead>
<tr>
<th>Univariate analysis</th>
<th></th>
<th>Multivariate analysis</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HR (95% CI)</strong></td>
<td><strong>P</strong></td>
<td><strong>HR (95% CI)</strong></td>
<td><strong>P</strong></td>
</tr>
<tr>
<td>Smoking Yes/no</td>
<td>1.014 (0.385–2.668)</td>
<td>0.978</td>
<td>1.859 (0.581–5.951)</td>
</tr>
<tr>
<td>Age ≥59</td>
<td>1.111 (0.528–2.338)</td>
<td>0.781</td>
<td>1.149 (0.505–2.617)</td>
</tr>
<tr>
<td>Sex M/F</td>
<td>1.265 (0.481–3.311)</td>
<td>0.634</td>
<td>1.997 (0.637–6.256)</td>
</tr>
<tr>
<td>TNM I/II/III</td>
<td>1.794 (1.131–2.845)</td>
<td>0.013</td>
<td>1.325 (0.775–2.267)</td>
</tr>
<tr>
<td>Differentiation</td>
<td>1.283 (0.656–2.508)</td>
<td>0.466</td>
<td>2.053 (0.882–5.130)</td>
</tr>
<tr>
<td>Adjuvant therapy</td>
<td>0.833 (0.374–1.855)</td>
<td>0.655</td>
<td>1.163 (0.459–2.950)</td>
</tr>
<tr>
<td>hsa-miRNA-31 High/low</td>
<td>2.827 (1.274–6.275)</td>
<td>0.011</td>
<td>3.535 (1.340–9.328)</td>
</tr>
</tbody>
</table>
finding that hsa-miR-31 can downregulate DICER1 is significant, considering DICER1 plays an important role in general microRNA expression and processing (3). Our results suggest that hsa-miR-31 may act as a negative regulator in the feedback loop controlling general microRNA expression and may have important implications for the systems biology of the cellular regulation of microRNA expression. The 3’ UTR homology to hsa-miR-31 was identical across all 3 genes (DICER1, PPP2R2A, and LATS2; Fig. 3A). It is particularly interesting why only DICER1 and
Table 3. Postoperative survival of SCC patients in relation to clinicopathologic characteristics and microRNA expression analyzed by the Cox proportional hazard regression model in a test cohort of 88 cases

<table>
<thead>
<tr>
<th></th>
<th>Univariate analysis</th>
<th>Multivariate analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR (95% CI)</td>
<td>P</td>
</tr>
<tr>
<td>Smoking</td>
<td>Yes/no</td>
<td>1.989 (0.908–4.355)</td>
</tr>
<tr>
<td>Age</td>
<td>≥60</td>
<td>1.169 (0.606–2.254)</td>
</tr>
<tr>
<td>Sex</td>
<td>M/F</td>
<td>0.969 (0.343–2.737)</td>
</tr>
<tr>
<td>TNM</td>
<td>I/II/III</td>
<td>1.263 (0.839–1.899)</td>
</tr>
<tr>
<td>Differentiation</td>
<td>Well/moderate/poor</td>
<td>2.543 (1.354–4.779)</td>
</tr>
<tr>
<td>Adjuvant therapy</td>
<td>Yes/no</td>
<td>3.060 (1.342–6.976)</td>
</tr>
<tr>
<td>hsa-miRNA-31</td>
<td>High/low</td>
<td>2.396 (1.218–4.715)</td>
</tr>
</tbody>
</table>

Table 3. Postoperative survival of SCC patients in relation to clinicopathologic characteristics and microRNA expression analyzed by the Cox proportional hazard regression model in a test cohort of 88 cases

not other genes (PPP2R2A and LATS2) were regulated by hsa-miR-31 in our study. One possible explanation is that microRNAs bind through partial sequence homology to the 3’-UTR of target genes, resulting at the same time in some degree of mRNA degradation and translation inhibition (4). Furthermore, the target gene sequences flanking the hsa-miR-31 binding site may affect the functions of hsa-miR-31. Grimson and colleagues also reported that the seed matches are not always sufficient for microRNA repression (36).

Significantly, the upregulation of hsa-miR-31 in human colorectal (37), liver (38), and head and neck tumors (39), as well as squamous cell carcinomas of the tongue (40), has been previously reported. Paradoxically, the downregulation of hsa-miR-31 is associated with malignant progression in human carcinomas of the breast (41), prostate (42), ovary (43), and stomach (44). These seemingly conflicting roles of hsa-miR-31 indicate that it may affect crucial events related to the progression of these tumors through complicated regulatory mechanisms (45). Further elucidation of the biological actions of hsa-miR-31 may prove significant to improve diagnosis and treatment of several types of human cancers. Our observation of its regulation of DICER1 activity may potentially explore wider biological functions of hsa-miR-31 in the development of other cancers.

In summary, we have carefully designed and investigated the microRNA expression profile in a large cohort of Chinese SCC patients, using cryopreserved archival tissues stored for at least 5 years for which extensive clinicopathologic data were available. We revealed a 5-microRNA classifier (hsa-miR-210, hsa-miR-182, hsa-miR-486-5p, hsa-miR-30a, and hsa-miR-140-3p) that can distinguish cancerous SCC lesions from adjacent normal tissues. We showed that a high level of hsa-miR-31 expression strongly correlated with a low survival for SCC patients. We showed that hsa-miR-31 directly targets the DICER1 3’-UTR and repressed the expression of DICER1 but not PPP2R2A or LATS2. The diagnostic 5-microRNA classifier and prognostic hsa-miR-31 should prove to be useful for the management of SCC among Chinese patients.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

The authors thank Meihua Xiong, Shujie Wang, and Xiaoyu Zhang for their excellent technical assistance.

Grant Support

This work was supported by grants from the National High Technology Research and Development Program of China (861 Program; No. 2006AA02A401) and National Key Technology R&D Program (No. 2006BAI1B01). 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.

Received February 27, 2011; revised August 19, 2011; accepted August 19, 2011; published OnlineFirst September 2, 2011.

References


A 5-MicroRNA Signature for Lung Squamous Cell Carcinoma Diagnosis and hsa-miR-31 for Prognosis

Xiaogang Tan, Wenyan Qin, Liang Zhang, et al.

Clin Cancer Res 2011;17:6802-6811. Published OnlineFirst September 2, 2011.

Updated version
Access the most recent version of this article at:
doi:10.1158/1078-0432.CCR-11-0419

Supplementary Material
Access the most recent supplemental material at:
http://clincancerres.aacrjournals.org/content/suppl/2011/09/02/1078-0432.CCR-11-0419.DC1

Cited articles
This article cites 43 articles, 12 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/17/21/6802.full#ref-list-1

Citing articles
This article has been cited by 5 HighWire-hosted articles. Access the articles at:
http://clincancerres.aacrjournals.org/content/17/21/6802.full#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.