SOX2 expression is associated with FGFR fusion genes and predicts favorable outcome in lung squamous cell carcinomas

Objectives: SOX2 is a gene that encodes for a transcription factor, which functions as an activator or suppressor of gene transcription. SOX2 amplification and overexpression have been found in various types of tumors and play important roles in cancer cells. The aim of the study was to evaluate SOX2 expression and amplification in lung squamous cell carcinomas (SCCs) and to determine the relationship with main clinicopathologic features, patient prognosis, and common driver mutations.

Materials and methods: SOX2 protein levels were measured by immunohistochemistry, while SOX2 copy numbers were measured by fluorescence in situ hybridization in resected samples from 162 Chinese lung SCC patients. All patients were also analyzed for mutations in EGFR, HER2, BRAF, PIK3CA, NFE2L2, and FGFR fusion genes. Clinical characteristics, including age, sex, smoking status, stage, relapse-free survival (RFS), and overall survival (OS), were collected.

Results: SOX2 overexpression and amplification were observed in 58.6% and 45.9% of lung SCCs. Lung SCC patients with SOX2 overexpression were significantly associated with absence of malignant tumor family history (P=0.021), FGFR fusion gene (P=0.046), longer RFS (P=0.041), and OS (P=0.025). No correlation was found between SOX2 gene amplification and main clinicopathologic features, patient prognosis, or common driver mutations.

Conclusion: SOX2 overexpression and amplification are common in lung SCCs. SOX2 overexpression was associated with FGFR fusion genes and predicted favorable outcome in lung SCCs. The underlying relationship of SOX2 and FGFR still needs further investigation.

Keywords: lung squamous cell carcinoma, SOX2 amplification, protein expression, FGFR fusion gene, prognostic marker

Introduction

Lung cancer continues to be the leading cause of cancer-related deaths worldwide.1 Despite multidisciplinary cancer therapies having taken great strides during the past decade, the overall prognosis for lung cancer patients remains poor. In lung adenocarcinoma, small molecule inhibitors targeting activated EGFR and EML4-ALK have improved the response rates and progression-free survival;2,3 however, lung squamous cell carcinomas (SCCs) lack novel treatment strategies that target molecular abnormalities. Therefore, it is urgently needed to identify reliable prognostic biomarkers and develop targeted molecular therapies for lung SCCs.

SOX2 (SRY [sex-determining region Y] box 2) is a recently identified novel lineage-survival oncogene in SCCs, which is located on chromosome 3q26.33.4 The SOX2 protein is a transcription factor of 317 amino acids containing a high mobility group domain.5 It plays a crucial role in both the pluripotency regulation in embryonic stem cells and the morphogenesis and homeostasis of tracheobronchial epithelia.6 Moreover,
SOX2 amplification and protein expression have been found in various types of tumors,7–11 and recent studies have shown that SOX2 is responsible for cellular proliferation, tumor invasion and migration, self-renewal, and maintenance in cancer stem cell populations.12 Amplified and overexpressed SOX2 is frequently associated with higher tumor-node-metastasis (TNM) stage, metastasis, and poorer prognosis in SCCs.13–15 but with opposing consequences of better outcome in lung carcinomas.16,17 The conflicting results suggest that the role of SOX2 in the clinic is still not well defined. Furthermore, the relationship between SOX2 and common molecular abnormalities in lung SCCs is still largely unexplored.

In this study, we performed fluorescence in situ hybridization (FISH) and immunohistochemistry (IHC) in surgically resected lung SCC and correlated amplification and expression of SOX2 with the main clinicopathologic features, patient prognosis, and common driver mutations.

Materials and methods

Patients and samples

Lung SCCs in the form of formalin-fixed paraffin embedded tissue from patients who underwent surgical resection with curative intent at Fudan University Shanghai Cancer Center between January 2008 and December 2011 were obtained. Eligible patients were required to have sufficient tissue for immunohistochemical staining, FISH, and comprehensive mutational analyses. Patients who received neoadjuvant chemotherapy or had a history of malignant tumor were excluded.

Clinicopathologic variables collected for analyses included sex, age at diagnosis, smoking history, family history of malignant tumor, tumor differentiation, pathologic TNM stage in line with the seventh edition of the lung cancer staging system,18 and postoperative adjuvant chemotherapy/radiotherapy. Disease recurrence and survival were observed in the follow-up clinic or obtained via telephone. The average follow-up period was 40 months (median 42 months range [2, 88]). All patients provided written informed consent at the interview. This study was approved by the Committee for Ethical Review of Research (Fudan University Shanghai Cancer Center IRB#090977-1).

Immunohistochemistry

In brief, sections were deparaffinized by serial xylene washes and rehydrated in graded alcohols, then digested with protease K (0.5 mg/mL) at 37°C for 20 minutes. The slides were then dehydrated in ethanol. The probes were denatured for 5 minutes at 75°C before hybridization. Slides were hybridized at 37°C for 36 hours and washed in 2× SSC/0.3% NP-40 at 72°C for 2 minutes. Nuclei were counterstained with DAPI 1 counterstain (Vysis, Downers Grove, IL, USA). Analysis were performed using a fluorescence microscope (Olympus BX51TRF, Olympus Corporation, Tokyo, Japan) equipped with an Olympus BX-UCB filter set (Olympus) with single-band exciters for Texas red/rhodamine, fluorescein isothiocyanate, and DAPI (UV 360 nm). Copy number per cell for each gene was enumerated on at least 50 tumor cells for each case. In the absence of validated FISH scoring criteria for SOX2, a cutoff value of more than 4 gene copies/cell, or presence of gene clusters, was set to identify cases with increased gene copy number (FISH+). Examples of FISH patterns are shown in Figure 1.

Fluorescence in situ hybridization

FISH analysis of SOX2 amplification was performed using a spectrum orange-labeled probe CTD-2348H10 and SPOT- Light tissue pretreatment solution (Invitrogen, Carlsbad, CA, USA). Briefly, sections were deparaffinized by serial xylene washes and rehydrated in graded alcohols, then digested with protease K (0.5 mg/mL) at 37°C for 20 minutes. The slides were then dehydrated in ethanol. The probes were denatured for 5 minutes at 75°C before hybridization. Slides were hybridized at 37°C for 36 hours and washed in 2× SSC/0.3% NP-40 at 72°C for 2 minutes. Nuclei were counterstained with DAPI 1 counterstain (Vysis, Downers Grove, IL, USA).

Analyses were performed using a fluorescence microscope (Olympus BX51TRF, Olympus Corporation, Tokyo, Japan) equipped with an Olympus BX-UCB filter set (Olympus) with single-band exciters for Texas red/rhodamine, fluorescein isothiocyanate, and DAPI (UV 360 nm). Copy number per cell for each gene was enumerated on at least 50 tumor cells for each case. In the absence of validated FISH scoring criteria for SOX2, a cutoff value of more than 4 gene copies/cell, or presence of gene clusters, was set to identify cases with increased gene copy number (FISH+). Examples of FISH patterns are shown in Figure 1.

Mutational analyses

Comprehensive mutational analyses of EGFR, HER2, BRAF, FGFR, PIK3CA, and NFE2L2 were performed in lung SCCs.
In brief, frozen tissues were dissected into TRIzol (Life Technologies, Carlsbad, CA, USA) followed by total RNA extraction using standard protocol. Total RNA samples were reverse transcribed into cDNA using RevertAid First Strand cDNA Synthesis Kit (Fermentas, St Leon-Rot, Germany). \( EGFR \) (exons 18–22), \( HER2 \) (exons 18–21), \( BRAF \) (exons 11–15), \( FGFR (FGFR1, FGFR2, and FGFR3) \), \( PIK3CA \) (exons 9 and 20), and \( NFE2L2 \) (exon 2) were amplified by polymerase chain reaction using cDNA. Amplified products were analyzed by direct dideoxynucleotide sequencing.

### Statistical analyses

Difference in proportions was analyzed by Pearson’s chi-square test or Fisher’s exact test. Relapse-free survival (RFS) and overall survival (OS) of patients with positive or negative IHC and FISH were estimated using the Kaplan–Meier method. The log-rank test was used to determine survival differences between groups. Independent prognostic factors were identified through the Cox proportional hazards regression (forward likelihood ratio model). All tests were two-tailed. Statistical significance was set as \( P<0.05 \). All data were analyzed using the SPSS Version 19.0 Software (SPSS Inc., Chicago, IL, USA).

### Results

#### Patient characteristics

A total of 162 lung SCC samples were collected from 12 females and 150 males. The mean age of the patients was 61.6 years, ranging from 40 to 88 years. One hundred thirty-eight patients (85.2%) had a history of smoking, and
36 patients (22.2%) had a family history of malignant tumor. The mean diameter of tumor was 4.306 cm, ranging from 0.7 to 13.0 cm. About 68 (42.0%) patients presented with lymph node metastasis. The patient characteristics of this cohort are described in Table 1.

Correlation between SOX2 expression and clinicopathologic and molecular features

SOX2 expression status was successfully determined in 145 patients, and positive staining was observed in 85 tumors (Table 2). SOX2 overexpression was not associated with age, sex, smoking history, tumor differentiation, and pleural invasion. Positive SOX2 expression was significantly associated with absence of malignant tumor family history (P = 0.021). Although there was a trend that SOX2-positive patients had smaller tumor diameter (4.557 cm vs 4.021 cm, P = 0.115) and less advanced pathologic stage (P = 0.268), the differences did not reach a statistical significance. Interestingly, among the six detected common molecular abnormalities in non-small-cell lung cancer (NSCLC), a statistically significant association was observed between SOX2 expression and FGFR fusion gene. Of 162 samples, FGFR3–TACC3 was identified in 8 patients and BAG4–FGFR1 was found in 2 patients. We found that seven patients with FGFR3–TACC3 and two with BAG4–FGFR1 were positive for SOX2 staining and only one patient with FGFR3–TACC3 was negative for SOX2 expression (P = 0.046).

Correlation between SOX2 amplification and clinicopathologic and molecular features

We also examined SOX2 amplification using FISH in 111 cases. SOX2 amplification was detected in 50 sections, and SOX2 gene copy number ranged from 2.2 to 15.3 (mean, 4.46). No correlation was found between SOX2 gene amplification in SCC and the patient’s age, sex, smoking history, family history of malignant tumor, tumor diameter, tumor-infiltrating lymphocytes, pathologic stage, pleural invasion, and mutational status.

Our results revealed that SOX2 FISH results were correlated with SOX2 IHC results (Table 3, Spearman correlation coefficient = 0.453; P < 0.001). Altogether 94 samples were examined for both SOX2 expression and amplification. Of 40 SOX2 FISH positive cases, 33 (82.5%) showed positive SOX2 IHC and only 7 cases (17.5%) were SOX2 IHC negative. Of 54 SOX2 FISH negative cases, 34 cases (63.0%) were negative for SOX2 IHC, and 20 cases (37.0%) were SOX2 IHC positive.

Survival analysis

The log-rank test on the Kaplan–Meier survival analysis demonstrated that lung SCC patients with positive SOX2 expression had longer RFS (Figure 2, median survival 55 vs 36 months; log rank P = 0.041) and OS (Figure 2, median survival 65 vs 44 months; log rank P = 0.025) than subjects with negative SOX2 expression. In contrast, although the median RFS and OS time for patients with SOX2 FISH positive lung SCC were better than SOX2 FISH negative patients, the differences were not statistically significant (Figure 2, for RFS, median survival 43 vs 38 months, log rank P = 0.630; for OS, median survival 55 vs 43 months, log rank P = 0.141).

To determine the prognostic accuracy of SOX2 expression, we used multivariate Cox regression model adjusted for sex, age at diagnosis, smoking history, family history of malignant tumor, tumor differentiation, TNM stage, mutational status, and postoperative chemotherapy/radiotherapy. However, Cox

| Table 1 Patient characteristics |
|---------------------------------|
| Characteristics                         | n   | %  |
| ----------------------------- | ---- | --- |
| Total                         | 162  | 100 |
| Sex                           |      |    |
| Female                        | 12   | 7.4 |
| Male                          | 150  | 92.6|
| Smoking history               |      |    |
| <20 pack-years                | 35   | 20.1|
| 20–50 pack-years              | 81   | 50 |
| >50 pack-years                | 46   | 28.4|
| Family history of malignant tumor |      |    |
| Absent                        | 126  | 77.8|
| Present                       | 36   | 22.2|
| Tumor differentiation         |      |    |
| Well                          | 5    | 3.1 |
| Moderate                      | 70   | 43.2|
| Poor                          | 87   | 53.7|
| Pathologic stage              |      |    |
| I                             | 69   | 42.6|
| II                            | 46   | 28.4|
| III                           | 46   | 28.4|
| IV                            | 1    | 0.6 |
| Pleural invasion              |      |    |
| Absent                        | 125  | 77.2|
| Present                       | 37   | 22.8|
| Mutational status             |      |    |
| EGFR mutation                 | 5    | 3.1 |
| HER2 mutation                 | 1    | 0.6 |
| BRAF mutation                 | 1    | 0.6 |
| FGFR fusion                   | 10   | 6.2 |
| PIK3CA mutation               | 3    | 1.9 |
| NFE2L2 mutation               | 6    | 3.7 |
Table 2 Clinicopathologic and molecular characteristics according to SOX2 expression and amplification

| Variable                        | SOX2 protein expression | SOX2 amplification | P   |
|--------------------------------|-------------------------|--------------------|-----|
|                                | Negative (n=60) | Positive (n=85) |     | Negative (n=61) | Positive (n=50) |     |
| Median age, years              | 58.5163     | 61.5686   | 0.200 | 62.8521     | 60.8233   | 0.270 |
| Sex                            |             |           | 0.983 |             |           | 0.904 |
| Female                         | 5           | 7         |     | 4           | 3         |     |
| Male                           | 55          | 78        |     | 57          | 47        |     |
| Smoking history                |             |           | 0.005 |             |           | 0.861 |
| <20 pack-years                 | 12          | 20        |     | 10          | 9         |     |
| 20–50 pack-years               | 23          | 50        |     | 31          | 27        |     |
| >50 pack-years                 | 25          | 15        |     | 20          | 14        |     |
| Family history of malignant tumor |             |           | 0.021 |             |           | 0.370 |
| Absent                         | 44          | 75        |     | 43          | 39        |     |
| Present                        | 16          | 10        |     | 18          | 11        |     |
| Type of surgical resection     |             |           | 0.029 |             |           | 0.037 |
| Wedge-shape excision           | 0           | 2         |     | 1           | 1         |     |
| Segmentectomy                  | 0           | 0         |     | 0           | 0         |     |
| Pulmonary lobectomy            | 56          | 64        |     | 54          | 35        |     |
| Pneumonectomy                  | 4           | 17        |     | 5           | 13        |     |
| Tumor differentiation          |             |           | 0.997 |             |           | 0.586 |
| Well                           | 2           | 3         |     | 2           | 1         |     |
| Moderate                       | 26          | 37        |     | 27          | 18        |     |
| Poor                           | 32          | 45        |     | 32          | 31        |     |
| Mean tumor diameter, cm        | 4.557       | 4.021     | 0.115 | 4.193       | 4.598     | 0.341 |
| N status                       |             |           | 0.751 |             |           | 0.402 |
| N0                             | 33          | 49        |     | 33          | 31        |     |
| N1/2/3                         | 27          | 36        |     | 28          | 19        |     |
| Pathologic stage               |             |           | 0.268 |             |           | 0.289 |
| I                              | 22          | 39        |     | 22          | 23        |     |
| II–IV                          | 38          | 46        |     | 39          | 27        |     |
| Pleural invasion               |             |           | 0.667 |             |           | 0.139 |
| Absent                         | 44          | 65        |     | 50          | 35        |     |
| Present                        | 16          | 20        |     | 11          | 15        |     |
| Mutational status              |             |           |       |             |           |     |
| EGFR mutation                  | 2           | 2         |     | 2           | 2         | 1.000 |
| HER2 mutation                  | 1           | 0         |     | 1           | 0         | 1.000 |
| BRAF mutation                  | 0           | 1         |     | 1.000       | 0         | 0.000 |
| FGFR fusion                    | 1           | 9         |     | 0.046       | 2         | 0.406 |
| PIK3CA mutation                | 0           | 2         |     | 0.511       | 2         | 1.000 |
| NFE2L2 mutation                | 1           | 3         |     | 0.642       | 2         | 3     |

Proportional hazards regression analysis revealed that SOX2 protein expression was associated with age and TNM stage, and failed to demonstrate SOX2 protein expression to be an independent prognostic factor in lung SCC.

Discussion

This study evaluated the relationship of clinical pathologic features, patient prognosis, and common driver mutations with protein expression and copy number alterations of SOX2 in a cohort of patients with surgically resected lung SCC. For the first time, we report that high level SOX2 protein expression assessed by IHC is associated with FGFR fusion gene in lung SCC.

The correlation of SOX2 expression with clinicopathological characteristics has been investigated in several retrospective NSCLC series. High-level expression of SOX2 was reported to be associated with lower TNM grade, smaller tumor size, lower probability of invasion and metastasis, and former or current smoking history,\textsuperscript{19,20} but recently a meta-analysis showed no correlation between SOX2 expression and clinicopathological parameters such as age, sex, smoking, lymph node metastasis, and tumor stage.\textsuperscript{17} Here, our results showed that SOX2 positive expression was associated with

Table 3 Comparison of SOX2 amplification and expression

| SOX2 protein expression | SOX2 amplification | P   |
|-------------------------|-------------------|-----|
| Negative                | Negative          | 34  |
| Positive                | Positive          | 7   |
|                         |                   | 20  |
|                         |                   | 33  |
absence of malignant tumor family history. Our new results showed the probability that SOX2 overexpression has less relationship with genetic factors, but more association with environmental impacts, and the regulation of SOX2 protein remains to be further explored.

Recently, the prognostic role of SOX2 expression and amplification in NSCLC has been investigated in different studies, but the results were contradictory. Several studies associated SOX2 protein overexpression with prolonged survival in surgically resected lung SCC patients, but a poor outcome was shown in early stage lung adenocarcinomas in the study of Sholl et al. Meanwhile, Luca et al reported increased SOX2 gene copy number as an independent favorable prognostic factor in patients with stage I and II NSCLC, but Bricic et al showed a nonsignificant result. Here, we carried out a study in a Chinese cohort and proved that SOX2 protein overexpression predicts better RFS and OS in lung SCC, but is not an independent prognostic factor.

SOX2 has shown its potential not only to become a useful biomarker for prognosis in the clinic, but also to be a novel therapy option. Chen et al demonstrated that silencing of the SOX2 gene effectively induced apoptosis via the activation of death receptor and mitochondrial signaling pathways in human NSCLC cells. Dogan et al investigated that SOX2 knockdown using shRNA in lung adenocarcinoma cell lines decreased cell proliferation and increased cell sensitivity to erlotinib. Moreover, SOX2 expression decreased when

Figure 2 Relapse-free survival and overall survival in SOX2 amplification and gene expression positive and negative patients.

Notes: (A) Relapse-free survival according to SOX2 expression (P=0.041). (B) Overall survival according to SOX2 expression (P=0.025). (C) Relapse-free survival according to SOX2 amplification (P=0.630). (D) Overall survival according to SOX2 amplification (P=0.141).

Abbreviations: IHC, immunohistochemistry; FISH, fluorescence in situ hybridization.
treated with PI3K/AKT inhibitors. Therefore, utilizing SOX2 and its upstream or downstream proteins for cancer therapy could open a window to new therapeutic opportunities, and identification of the relationship between SOX2 protein expression or gene amplification and common driver genes may help to identify patients who might benefit from particular targeted therapies.

**FGFR** fusion gene was illustrated to be a new driver for a range of cancers. A previous study has reported that FGFR1/3 fusions occurred in 1.3% of patients with NSCLCs and in 3.5% of patients with lung SCC. FGFR fusions have been shown to sensitize cancer cells to FGFR kinase inhibitors PD173074 and pazopanib, suggesting that a new subset of cancers may be treatable with FGFR-targeted therapy. To our knowledge, this study was the first to represent the association of SOX2 gene expression with FGFR fusion in a clinical cohort. Although the underlying mechanism between SOX2 and FGFR still needs to be further developed, novel therapies targeting FGFR or even SOX2 may be promising.

**Conclusion**

In conclusion, our data demonstrated that positive SOX2 expression was significantly associated with absence of malignant tumor family history, FGFR fusion genes and had a favorable clinical outcome in lung SCC patients. Our finding emphasized the importance of SOX2 in lung SCC biology and encourages further exploration of novel therapeu tic combinations.

**Disclosure**

The authors report no conflicts of interest in this work.

**References**

1. Jemal A, Bray F, Center MM, et al. Global cancer statistics. *CA Cancer J Clin*. 2011;61(2):69–90.
2. Paez JG, Janne PA, Lee JC, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science*. 2004;304:1497–1500.
3. Kwak EL, Bang YJ, Camidge DR, et al. Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. *N Engl J Med*. 2010;363:1693–1703.
4. Bass AJ, Watanabe H, Mermel CH, et al. SOX2 is an amplified lineage-survival oncogene in lung and esophageal squamous cell carcinomas. *Carcinogenesis*. 2010;13:7663–7675.
5. Yuan P, Kadara H, Behrens C, et al. Sex determining region Y-Box 2 (SOX2) is a potential cell-lineage gene highly expressed in the pathogenesis of squamous cell carcinomas of the lung. *PLoS One*. 2010;5:e9112.
6. Li X, Wang J, Xu Z, et al. Expression of Sox2 and Oct4 and their clinical significance in human non-small-cell lung cancer. *Int J Mol Sci*. 2012;13:7663–7675.
7. Luca T, Giovanna F, Teresa TN, et al. Increased SOX2 gene copy number is associated with FGFR1 and PIK3CA gene gain in non-small-cell lung cancer and predicts improved survival in early stage disease. *PLoS One*. 2014;9(4):e95303.
8. Domman J, Kawabata S, Bergbower E, et al. SOX2 expression and amplification in lung SCC. *Genes Chromosomes Development*. 2014;21:657–664.
9. Han X, Fang X, Lou X, et al. Silencing SOX2 induced mesenchymal-epithelial transition and its expression predicts liver and lymph node metastasis of CRC patients. *PLoS One*. 2012;7:e41335.
10. Tian Y, Jia X, Wang S, et al. SOX2 oncogenes amplified and operate to activate AKT signaling in gastric cancer and predict immunotherapeutic responsiveness. *J Cancer Res Clin Oncol*. 2014;140:1117–1124.
11. Velcetti V, Schalper K, Yao X, et al. High SOX2 levels predict better outcome in non-small cell lung cancers. *PLoS One*. 2013;8:e61427.
12. Chen Y, Huang Y, Huang Y, et al. The prognostic value of SOX2 expression in non-small cell lung cancer: a meta-analysis. *PLoS One*. 2013;8:e71140.
13. Deterbeck FC, Boffa DJ, Tanoue LT. The new lung cancer staging system. *Chest*. 2009;136(1):260–271.
14. Wilbertz T, Wagner P, Petersen K, et al. SOX2 gene amplification and protein overexpression are associated with better outcome in squamous cell lung cancer. *Mod Pathol*. 2011;24:944–953.
15. Yuan P, Kadara H, Behrens C, et al. Sex determining region Y-Box 2 (SOX2) is a potential cell-lineage gene highly expressed in the pathogenesis of squamous cell carcinomas of the lung. *PLoS One*. 2010;5:e9112.
16. Sasaki H, Yokota K, Hikosaka Y, et al. Increased Sox2 copy number fusions have been shown to sensitize cancer cells to FGFR kinase inhibitors PD173074 and pazopanib, suggesting that a new subset of cancers may be treatable with FGFR-targeted therapy. To our knowledge, this study was the first to represent the association of SOX2 gene expression with FGFR fusion in a clinical cohort. Although the underlying mechanism between SOX2 and FGFR still needs to be further developed, novel therapies targeting FGFR or even SOX2 may be promising.

**Conclusion**

In conclusion, our data demonstrated that positive SOX2 expression was significantly associated with absence of malignant tumor family history, FGFR fusion genes and had a favorable clinical outcome in lung SCC patients. Our finding emphasized the importance of SOX2 in lung SCC biology and encourages further exploration of novel therapeutic combinations.

**Disclosure**

The authors report no conflicts of interest in this work.
