Original Article

Analysis of real-word mutations of lung cancer driver genes in five regions of China

Mengyao Sun[1#], Ye Guo[1#], Guoguang Shao[2], Xiumei Duan[3], Zhiguang Yang[2], Peng Zhang[2], Yunpeng Liu[2], Yutong Dong[4], Xu Wang[1], Yinghui Xu[1], Chao Sun[1], Kewei Ma[1]

[1]Department of Cancer Centre, [2]Department of Thoracic Surgery, [3]Department of Pathology, [4]Department of Gastroenteric Medicine, First Hospital of Jilin University, Changchun 130021, China

Contributions: (I) Conception and design: K Ma, M Sun, Y Guo; (II) Administrative support: G Shao, X Duan; (III) Provision of study materials or patients: X Duan, Z Yang, P Zhang; (IV) Collection and assembly of data: M Sun, Y Guo, Y Liu, Y Dong; (V) Data analysis and interpretation: M Sun, Y Guo; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Kewei Ma. Department of Cancer Centre, First Hospital of Jilin University, NO.71, Xinmin Street, Changchun 130021, China. Email: makw@jlu.edu.cn.

Background: The aim of this study was to analyse the epidemiological characteristics and clinical features of the three driver genes *EGFR*, *ALK* and *ROS1* in Chinese patients with non-small-cell lung cancer (NSCLC).

Methods: *EGFR* mutations, *ALK* fusions and *ROS1* rearrangements were detected simultaneously by quantitative real-time PCR. Subgroup analyses were performed for adenocarcinoma and squamous cancer. The Chi-square test and multivariate logistic regressive analysis were used to analyse the associations between gene alterations and clinical features.

Results: A total of 3,081 patients with pathologically confirmed NSCLC from five sites in China were enrolled, among whom 1,449 (47.03%) had *EGFR*, *ALK* and/or *ROS1* alterations. In adenocarcinoma, the alteration rates of *EGFR*, *ALK* and *ROS1* were 50.6% (1,193/2,360), 6.3% (148/2,360), and 1.6% (38/2,360), respectively. *EGFR* and *EML4-ALK* coexisted in 16 cases (0.5%), while *EGFR* and *ROS1* coexisted in 1 case (0.03%). Sex, smoking status, and tumour stage were significantly correlated with the *EGFR* mutation; age and smoking status were correlated with *EML4-ALK*; and age and tumour stage were correlated with *ROS1*. In squamous cancer, the alteration rates of *EGFR*, *ALK* and *ROS1* were 7% (34/488), 2.9% (14/488) and 0% (0/488), respectively. Sex and smoking history were associated with *EGFR*, and sex was the only independent predictor of *EGFR*. The *EGFR* gene mutation sites were mainly 19del (557/1,263; 44.1%) and 21 exon L858R (575/1,263; 45.5%). More uncommon *EGFR* mutation types were present in 10.4% (131/1,263) of patients. Patients with *EGFR*, *ALK*, and/or *ROS1* alterations had different epidemiological characteristics and clinical features.

Conclusions: This real-word study of alterations in driver genes in a large population in China revealed unique epidemiological characteristics and clinical features in Chinese patients with NSCLC.

Keywords: *EGFR*; *EML4-ALK*; *ROS1*; epidemiology; clinical features

Submitted May 30, 2019. Accepted for publication Sep 09, 2019.
doi: 10.21037/tcr.2019.10.28
View this article at: http://dx.doi.org/10.21037/tcr.2019.10.28

Introduction

Lung cancer is the most frequently diagnosed malignancy in the world, with the highest morbidity and mortality. Over the past decade, the rapid development of the identification of driver genes and the application of molecular targeted therapies has completely changed the use of traditional platinum-based combination therapy for non-small-cell lung cancer (NSCLC) (1-3). Several phase I, II, and III clinical trials have confirmed the effectiveness of drugs targeting different driver genes. Currently, *EGFR*, *ALK*, and *ROS1* are the targets for most clinically approved drugs with precise curative effects (4). Tyrosine kinase inhibitors (TKIs) specific against driver genes and exhibiting a high objective response rate (ORR) significantly prolong progression-free survival (PFS) and overall survival (OS).

Clinical experience with successful therapies directed against *EGFR*, *ALK* and *ROS1* has dramatically changed the treatment strategies for NSCLC. The status of the three genes affects and determines the choice of clinical treatment strategies. Therefore, accurate and extensive detection of driver gene mutations in suitable populations is of great significance for patients before they receive a specific targeted treatment. At present, the detection of *EGFR*, *ALK*, and *ROS1* alterations has reached a clinical consensus. Multiple clinical guidelines recommend the detection of the three genes in patients with non-squamous lung cancer and squamous cancer whose clinical features suggest that they are likely to have at least one alteration in one of the three genes (1,5). The more common strategy is to detect *EGFR* first. If the *EGFR* mutation is negative, then the detection of *EML4-ALK* is performed. This method of separately detecting the mutations has a long waiting period and incurs a high cost. For smaller specimens, there is often insufficient sample left for *ALK* and *ROS1* detection after *EGFR* detection. Therefore, detecting three genes separately cannot meet clinical and patient needs. The technology for detecting *EML4-ALK* and *ROS1* is complex and diverse and affects the consistency of results. In this study, *EGFR*, *ALK* and *ROS1* were simultaneously detected with the same qRT-PCR assay kit approved by the National Medical Products Administration (NMPA) for clinical use. Compared to other detection methods, qRT-PCR kits are easy to use, highly sensitive, and relatively inexpensive.

The key point is to encourage patients who are likely to benefit from targeted therapies to undergo gene testing. Driver gene mutations are common in China, which has the largest population of a single country worldwide. However, molecular epidemiological data from mainland China remain scarce. To obtain a more precise estimate of *EGFR*, *ALK*, and *ROS1* alterations in Chinese patients, we performed a multicentre, retrospective, large-scale population analysis to confirm the epidemiology and clinical features of the three driver genes and to provide important references for guiding clinical treatment.

Methods

Study design

This was a real-world, retrospective, multicentre, epidemiological study of the prevalence of *EGFR*, *ALK* and *ROS1* gene alterations, including uncommon mutations of *EGFR*, in Chinese patients. Subgroup analyses of adenocarcinoma and squamous cancer were performed. The primary objective of the study was to assess the frequency of *EGFR*, *ALK*, and *ROS1* gene alterations, whereas the secondary objective was to analyse the correlation between each of the three driver mutation statuses and clinical features.

Clinical samples

From January 1, 2017, to December 31, 2017, eligible patients from 5 centres, Shanxi, Beijing, Liaoning, Sichuan and Jilin, who had pathologically confirmed NSCLC with assessments of the *EGFR*, *ALK*, and *ROS1* genes were included in this study. Major demographic and clinical characteristics were collected, including age, sex, smoking status, pathology type and tumour stage. Pathological types and tumour stages were determined according to the 2015 World Health Organization (WHO) classification. The TNM stage was classified according to the 7th edition of the Union for International Cancer Control (UICC)/ American Joint Committee on Cancer (AJCC).

Mutation detection

Total RNA and DNA isolated from formalin-fixed paraffin-embedded (FFPE) tissues from each patient were used to qualitatively detect *EGFR* mutations, ALK fusion products and *ROS1* rearrangements simultaneously with the quantitative real-time (qRT)-PCR-based ADx-ARMS *EGFR/ALK/ROS1* Gene Joint Detection Kit (Amoy Diagnostics Co., Ltd., Xiamen, China). The recommended RNA and DNA concentration from FFPE samples is 10–500 ng/μL and 1.5–3 ng/μL, respectively. The qRT-PCR conditions for complementary DNA were as follows: one cycle of 95 ℃ for 5 min; 10 cycles of denaturation at

Table 1 Clinical features of 3,081 patients

Characteristics	Number	Ratio (%)
Age		
<60 years	1,331	43.2
≥60 years	1,750	56.8
Gender		
Male	1,780	42.4
Female	1,301	57.6
Smoking		
Yes	1,412	45.8
No	1,669	54.2
Histological		
Adenocarcinoma	2,360	76.6
Squamous	488	15.8
Others	233	7.6
Stage		
I	562	18.2
II	577	18.7
III	654	21.2
IV	819	26.6
NA	469	15.2

Smoking definitions: No, patients who have no history of smoking; Yes, patients who have history of smoking.

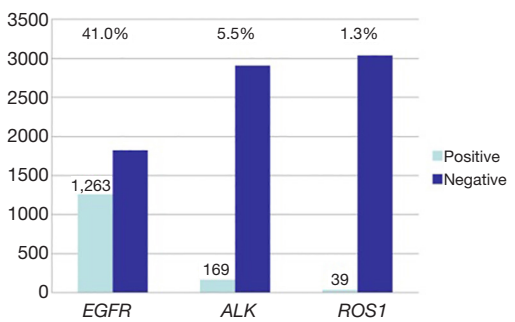

Figure 1 Positive rate of driver genes in NSCLC.

95 ℃ for 25 s, annealing at 64 ℃ for 20 s, and elongation at 72 ℃ for 20 s to ensure specificity; and up to 36 cycles of denaturation at 93 ℃ for 25 s, annealing at 60 ℃ for 35 s (data collection), and elongation at 72 ℃ for 20 s. An external control for each sample and an internal control for each tube were included to check the effects of DNA insufficiency or PCR inhibitors. For *EML4-ALK* fusion and *ROS1* rearrangement, the kit was used only for qualitative detection, not for genotyping. For *EGFR*, the following mutations could be detected: three in exon 18 (G719A, G719S, and G719C, which are referred to as G719X because the kit was unable to distinguish between these subtypes), 19 deletions in exon 19, two mutations in exon 20 (S768I, T790M), three insertions in exon 20, and two mutations in exon 21 (L858R, L861Q).

Statistical analysis

All statistical analyses were performed using SPSS software, version 22 (IBM SPSS Statistics. Inc. Chicago, IL, USA). Patients with tumours of undetermined mutation status were not included in these calculations. The relationship between gene mutations and clinical features was calculated using the Chi-square test. Statistical significance was defined as $P<0.05$. To best predict the relationship between gene mutations and clinical features, factors with $P<0.05$ in the Chi-square test were further analysed by multivariate logistic regression.

Results

Overall population characteristics and epidemiology (Table 1, Figure 1)

From January 1, 2017 to December 31, 2017, a total of 3,081 lung cancer tissue samples from patients from 5 regions were genetically tested. The overall population characteristics are shown in *Table 1*. A total of 42.2% (1,301/3,081) of patients were female, the median age at diagnosis was 61 years (ranging, 5 to 88 years), and 54.2% (1,669/3,081) of the patients had no history of smoking. Adenocarcinoma accounted for 76.6% (2,360/3,081), and squamous cancer accounted for 15.8% (488/3,081). The percentages of patients with stage I, II, III, and IV disease were 18.2%, 18.7%, 21.2%, and 26.6%, respectively (*Table 1*). Among the available 3,081 Chinese lung cancer patients, 1,449 were positive for driver mutations. The *EGFR* gene mutation rate was 41.0% (1,263/3,081), the *EML4-ALK* gene positive rate was 5.5% (169/3,081), and the *ROS1* gene positive rate was 1.3% (39/3,081). *EGFR* and *EML4-ALK* mutations coexisted in 16 cases (0.5%), and *EGFR* and *ROS1* mutations coexisted in 1 case (0.03%). Among patients with adenocarcinoma, the *EGFR* gene mutation rate was

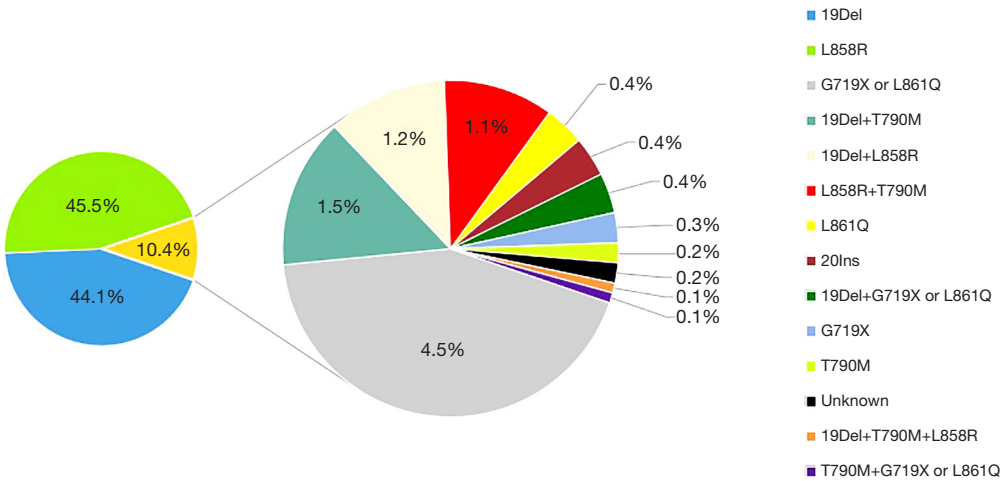

Figure 2 Summary of various type of *EGFR* mutations in NSCLC.

50.6% (1,193/2,360), which was significantly higher than the rate in patients with squamous carcinoma (7%, 34/488). The mutation rates of *ALK* and *ROS1* in adenocarcinoma were approximately 6.3% (148/2,360) and 1.6% (38/2,360), respectively. In squamous carcinoma, the *ALK* gene mutation rate was 2.9% (14/488). There were no *ROS1* gene mutations in squamous carcinoma.

The *EGFR* mutation types were mainly 19del and 21 exon L858R. The 19del mutation alone accounted for 44.1% (557/1,263); when combined with other mutations, it accounted for 47.1% (595/1,263). A total of 45.5% (575/1,263) of the samples had only the L858R mutation, and 47.7% (603/1,263) carried L858R alone or in combination with another mutation. Among the uncommon *EGFR* mutations, the G719X/L861Q mutation was the most frequently observed, accounting for 4.5% (n=57/1,263) of all patients with *EGFR* mutations. Other uncommon *EGFR* mutations are listed in *Figure 2*.

In all patients with EGFR mutations, the percentage of samples harbouring sensitive mutations alone was 96.6% (1,153/1,193), whereas 0.6% (7/1,193) had only a resistant mutation. The combination of sensitive and resistant mutations accounted for 2.7% (32/1,193).

Mutation analyses in adenocarcinoma samples subset (Tables 2,3)

EGFR gene

The frequency of *EGFR* mutations in adenocarcinoma was 50.6% (1,193/2,360).

Factors with a statistically significant association with *EGFR* mutation status (Chi-square test) were sex (P<0.001), smoking history (P<0.001), and tumour stage (P<0.001); age was not significantly correlated with *EGFR* mutation (P=0.330). Multivariate logistic regression further confirmed that female sex (P<0.001), no history of smoking (P=0.007), and stage III cancer (P=0.009) were independent predictive factors for *EGFR* mutation status. *EGFR* mutation was mutually exclusive with *EML4-ALK* fusion and *ROS1* rearrangement (P<0.001).

In conclusion, *EGFR* mutations were more likely to occur in females or patients with no history of smoking. The incidence of stage III patients was the lowest, and *EGFR* mutations were negatively correlated with *EML4-ALK* and *ROS1* mutations.

ALK gene

Of the 2,360 screened adenocarcinoma samples, 6.3% (148/2,360) harboured the *EML4-ALK* fusion.

Chi-square test analysis determined that age (P<0.001) and history of smoking (P=0.018) were significantly associated with *EML4-ALK* fusion. Multivariate logistic regression further confirmed that a younger age (P<0.001) and no history of smoking (P=0.047) are independent predictive factors for *EML4-ALK* mutation status. The median patient age was 56 years old, and the prevalence of this mutation was highest in stage IV patients (37%, 55/148; P<0.014). In conclusion, patients younger than 60 years old, who had no history of smoking and who had stage IV disease were prone to *EML4-ALK* gene fusion. There was no statistical association between *EML4-ALK* mutation and either sex (P=0.525) or *EGFR* mutation.

Table 2 Demographic and clinical characteristics of *EGFR*, *ALK*, and *ROS1* mutation in adenocarcinoma patients

Clinical feature	N	EGFR			ALK			ROS1		
		No. of positive (%)	95% CI	P	No. of positive (%)	95 CI	P	No. of positive (%)	95% CI	P
Gender				<0.001*			0.266			0.067
Male	1,164	451 (38.7)	25.7–30.1		66 (5.7)	4.1–6.6		13 (1.1)	0.5–1.7	
Female	1,196	742 (62.0)	36.1–40.5		82 (6.9)	5.1–7.8		25 (2.1)	1.3–2.8	
Age				0.330			<0.001*			0.004*
<60 years	1,068	530 (49.6)	30.8–35.4		108 (10.1)	7.5–10.8		26 (2.4)	1.5–3.3	
≥60 years	1,292	663 (51.3)	31.8–36.0		40 (3.1)	2.1–3.9		12 (0.9)	0.4–1.4	
Smoking				<0.001*			0.018*			0.067
No	1,456	853 (58.6)	38.4–38.8		105 (7.2)	5.5–8.0		29 (2.0)	1.2–2.7	
Yes	904	340 (37.6)	24.9–29.8		43 (4.8)	3.2–5.9		9 (1.0)	0.3–1.6	
Stage				<0.001*			0.111			0.182
I	490	291 (59.4)	33.9–40.7		23 (4.7)	2.7–6.3		4 (0.8)	0–1.6	
II	368	188 (51.1)	29.9–37.8		19 (5.2)	2.7–7.1		5 (1.4)	0.2–2.5	
III	461	201 (43.6)	26.9–33.9		33 (7.2)	4.5–8.9		11 (2.4)	1–3.7	
IV	710	355 (50.0)	30.5–36.2		55 (7.7)	5.4–9.0		15 (2.1)	1–3.1	
NA	331	158 (47.7)	28.2–36.5		18 (5.4)	2.8–7.5		3 (0.9)	0.1–1.9	
ALK				<0.001*			–			–
–	2,212	1,178 (53.3)	33.0–36.2		–	–		38 (1.7)	1.2–2.2	
+	148	15 (10.1)	4.7–13.7		–	–		0 (0)	0	
ROS1				<0.001*			–			–
–	2,322	1,171 (50.4)	32.3–35.4		148 (6.4)	5.1–6.9		–	–	
+	38	1 (2.6)	2.6–7.8		0 (0)	0		–	–	

*, P<0.05. P values are from χ^2 test, refer to overall comparisons across all subgroups, analyses as the results of an individual demographic and clinical factor. may be influenced by the others and may therefore not represent a true effect of that variable. Smoking definitions: No: patients who have no history of smoking; Yes: patients who have the history of smoking. CI, confidence interval.

The incidence of *EML4-ALK* fusion in *EGFR*-negative patients was 11.4% (133/1,167). Of note, 15 patients had double mutations in *EGFR* and *ALK*; for carcinoma, 11 patients were younger than 60 years old, and 4 patients were older than 60 years old. Eight patients had the 19del *EGFR* mutation, 5 had the L858R mutation, 1 had the L861Q or G719X mutation, and 1 had the 19del and T790M compound mutations.

ROS1 gene

A total of 1.7% (39/2,360) of the patients were positive for *ROS1* rearrangement, all of whom had adenocarcinoma. Approximately 3.3% (38/1,167) of patients without *EGFR* mutations harbour *ROS1* rearrangement.

The Chi-square test showed that *ROS1* rearrangement was correlated with age (P=0.004). Multivariate logistic regression confirmed that age (P=0.010) and tumour stage (P=0.024, 0.044) were independent predictive factors of *ROS1*. *ROS1* had no association with sex (P=0.443) or smoking status (P=0.334). The median age at diagnosis was 54 years. The prevalence of *ROS1* was higher in patients with stage III (29%, 11/38; P=0.024) and stage IV (39%, 15/38; P=0.044) than in those with other disease stages. In conclusion, *ROS1* gene mutations were more likely to occur in patients younger than 60 years old and in those with advanced lung cancer.

Table 3 Multivariate logistic regression analysis for *EGFR*, *ALK*, and *ROS1* status in adenocarcinoma patients

Variable	Regression coefficient (estimate)	SE	Odds ratio estimate (95% CI)	P
EGFR				
Intercept	−6.934	1.018		
Gender (female *vs.* male)	0.842	0.130	2.320 (1.798–2.994)	<0.001*
Age (<60 *vs.* ≥60 years)	−0.035	0.098	0.966 (0.797–1.171)	0.725
Smoking (no *vs.* yes)	0.356	0.133	1.428 (1.101–1.853)	0.007*
Stage				
II *vs.* I	−0.111	0.151	0.895 (0.666–1.203)	0.461
III *vs.* I	−0.369	0.142	0.691 (0.523–0.913)	0.009*
IV *vs.* I	−0.159	0.127	0.853 (0.665–1.095)	0.213
ALK (negative *vs.* positive)	2.688	0.313	14.705 (7.967–27.140)	<0.001*
ROS1 (negative *vs.* positive)	3.998	1.019	54.467 (7.385–401.716)	<0.001*
ALK				
Intercept	4.154	0.308		
Gender (female *vs.* male)	0.152	0.240	1.164 (0.728–1.862)	0.525
Age (<60 *vs.* ≥60 years)	−1.400	0.209	0.247 (0.164–0.372)	<0.001*
Smoking (no *vs.* yes)	−0.518	0.261	0.596 (0.357–0.994)	0.047*
Stage				
II *vs.* I	−0.184	0.322	0.832 (0.442–1.565)	0.568
III *vs.* I	−0.543	0.280	0.581 (0.336–1.005)	0.052
IV *vs.* I	−0.622	0.253	0.537 (0.327–0.881)	0.014*
ROS1				
Intercept	6.076	0.666		
Gender (female *vs.* male)	−0.364	0.475	0.695 (0.274–1.762)	0.443
Age (<60 *vs.* ≥60 years)	−0.953	0.368	0.386 (0.187–0.794)	0.010*
Smoking (no *vs.* yes)	−0.517	0.535	0.596 (0.209–1.701)	0.334
Stage				
II *vs.* I	−0.780	0.678	0.458 (0.121–1.730)	0.250
III *vs.* I	−1.336	0.592	0.263 (0.082–0.839)	0.024*
IV *vs.* I	−1.142	0.569	0.537 (0.105–0.972)	0.044*

CI, confidence interval; SE, standard error. *, P<0.05.

Mutation analyses in squamous samples subset (Tables 4,5)

In patients with squamous cancer, 7% (34/488) harboured *EGFR* mutations, and 2.8% (14/488) harboured *EML4-ALK* fusion. None of the samples had *ROS1* mutations. The same statistical method showed that *EGFR* mutation is more likely to occur in females (P<0.001) and in patients with no history of smoking (P=0.047). Of all the factors analysed in the squamous patient subgroup, female sex (P=0.004) was the only independent factor for *EGFR* mutation among. There was no statistical relationship between *ROS1* mutation and age (P=0.484), smoking history

Table 4 Demographic and clinical characteristics of *EGFR* mutation in squamous patients

Clinical feature	N	No. of positive (%)	95% CI	P
Gender				<0.001*
Male	447	25 (5.6)	3.3–7.3	
Female	41	9 (22.0)	7.0–29.0	
Age				0.584
<60 years	166	13 (7.8)	3.4–11.1	
≥60 years	322	21 (6.5)	3.6–8.7	
Smoking				0.047*
No	105	12 (11.4)	4.1–14.9	
Yes	383	22 (5.7)	3.2–7.6	
Stage				0.055
I	55	2 (3.6)		
II	171	14 (8.2)	3.7–11.4	
III	160	7 (4.4)	1.1–7.3	
IV	65	9 (13.8)	4.5–19.8	
NA	37	2 (5.4)		
ALK				0.981
–	474	33 (7.0)	4.4–8.7	
+	14	1 (7.1)	–7.6–21	
ROS1				a
–	488	34 (7.0)	4.4–8.6	
+	0			

P values are from χ^2 test, refer to overall comparisons across all subgroups, analyses as the results of an individual demographic and clinical factor may be influenced by the others and may therefore not represent a true effect of that variable. Smoking: patients have no history of smoking; patients have the history of smoking. a, unavailable P value because there was no positive for *EGFR* mutation in squamous patients with *ROS1* rearrangement. CI, confidence interval. *, P<0.05.

(P=0.756) or tumour stage (P=0.131). In addition, since only 2.8% of patients had *EML4-ALK* fusion and no *ROS1* rearrangements, we did not analyse the correlation between clinical features and gene alterations.

Discussion

EGFR gene

EGFR is considered the strongest biomarker in NSCLC. Studies across different regions indicate regional differences in the epidemiology of *EGFR*. Asia has the highest prevalence of *EGFR* mutations at 38–51% compared with 13% in Caucasians and 10–15% in North Americans and Europeans (6-9). We found that similar to Asia overall, China has a higher *EGFR* mutation rate than most other regions, and this rate is slightly lower than that in reports of cohorts containing specifically Asian populations. We confirmed that both adenocarcinoma and squamous cancer patients who are female and have no history of smoking are more likely to harbour *EGFR* mutations. Notably, adenocarcinoma patients at stage III comprise the smallest population of patients with an *EGFR* mutation and are negatively correlated with this mutation. We speculated that, except for stage IV patients, early stage (stage I–II) patients might account for the largest portion of those positive for an *EGFR* mutation. Unexpectedly, we found that more than half of early stage (stage I–II) lung adenocarcinomas were positive for *EGFR* mutations, demonstrating that early stage patients with *EGFR* mutations should not be ignored. In the CTong1104 study, TKIs were used as an adjuvant treatment and reduced tumour recurrence compared to chemotherapy alone (DFS: 28.7 *vs.* 18 months) for stage II–IIIA (N1–N2) patients (10). Updated guidelines in the 2018 Chinese Society of Clinical Oncology (CSCO) also strongly recommend molecular gene detection of tissues for stage I–IIIA patients after surgery (11). Our study, with the reported high incidence of *EGFR* mutations at an early stage, further confirmed the necessity of genetic testing among patients with early adenocarcinomas. When excluding the interaction of various clinical features not in line with a previous study that did not consider sex as an independent predictor (12), we found that sex, smoking status, and tumour stage were all independent predictors of EGFR in adenocarcinoma. Even in the subgroup of squamous cancer patients (which had a lower frequency of *EGFR* mutations), female was the only predictive factor for *EGFR* mutations.

Patients with certain features, including those with squamous cancer, those who were male and those with a history of smoking, were also found to have *EGFR* mutations, although with a lower prevalence. Therefore, these subsets of patients should not be exempt from genetic testing; otherwise, approximately 45% of the *EGFR* mutations reported in this study would have remained undetected. The prevalence of *EGFR* mutations was higher among Chinese patients than other ethnicities; the percentage of Chinese males with an *EGFR* mutation was 27.6%, which is in sharp contrast to the 8.2% reported

Table 5 Multivariate logistic regression analysis for EGFR in squamous patients

Variable	Regression coefficient (estimate)	SE	Odds ratio estimate (95% CI)	P
Intercept	–3.885	1.345		
Gender (female *vs.* male)	1.683	0.592	5.380 (1.687–17.155)	0.004*
Age (<60 *vs.* >60 years)	0.272	0.098	1.313 (0.612–2.816)	0.484
Smoking (no *vs.* yes)	–0.168	0.541	0.845 (0.293–2.441)	0.756
Stage				
II *vs.* I	0.827	0.783	2.287 (0.493–10.607)	0.291
III *vs.* I	0.095	0.872	1.100 (0.217–5.565)	0.908
IV *vs.* I	1.243	0.823	3.465 (0.690–17.392)	0.131
ALK (negative *vs.* positive)	0.391	1.109	1.479 (0.168–13.007)	0.742

CI, confidence interval; SE, standard error. *, P<0.05.

in the European male population, and the percentage of Chinese smokers with an *EGFR* mutation was approximately 26.3%, which is much higher than the 5.8% observed in European smokers (13). Therefore, *EGFR* mutation testing is warranted in males and in smokers, particularly for Chinese patients.

The squamous cancer subtype is rare in *EGFR*-positive patients, but this rate is 7% higher than the rates in Europe and America (14-16). We screened selected *EGFR* mutations in squamous cancer patients based on the analysed clinical features. The rate of *EGFR* mutations among females was approximately 22%. Patients who were non-smokers and harboured *EGFR* mutations accounted for 11.3%, and females who did not smoke accounted for 21.6% of the subgroup. Related studies reported that squamous cancer patients with *EGFR* mutations also significantly benefited from EGFR-TKIs, with a disease control rate of 50–70% (17-19). Although EGFR-TKIs were not as effective in squamous cancer patients with *EGFR* mutations as they were in adenocarcinoma patients, they can be used as an alternative treatment option. Therefore, *EGFR* gene testing in squamous cancer patients should be considered, especially for females or non-smokers.

The most common *EGFR* mutations, exon 19del and the L858R point mutation, are considered "classical mutations". The remaining *EGFR* mutations, including single mutations or those in combination with a classical *EGFR* mutation or other mutations, are termed "non-classical mutations" (20). Given the high incidence of *EGFR* mutations, non-classical mutations are not rare. Therefore, the epidemiology of non-classical mutations and the exploration of the efficiency of TKIs have certain clinical guidance. In this study involving Chinese patients, the rate of non-classical mutations was 10.4% lower than that in previous studies (12–15.5%) (20-22). The efficacy of EGFR-TKIs for treating patients harbouring non-classical *EGFR* mutations is varied (1). By reviewing related clinical trials and studies, we regarded G719X, L861Q, S768I and classical mutations alone or in combination with each other as sensitive type due to the sensitivity of these mutations to TKIs. We confirmed that the first two generations of TKI were ineffective against resistant mutation types, including 20ins and T790M either alone or combined with other mutations (21,23,24). However, the use of a third-generation EGFR-TKI was significantly effective against the T790M mutation, regardless of whether the tumour was primary or secondary. Our study results, along with the results of other clinical trials, show that even in the presence of resistant mutations, more than 97% of patients harbouring *EGFR* mutations could benefit from EGFR-TKIs. The premise is that undergoing molecular genetic testing can confirm whether patients could benefit from targeted therapies, which is of great significance for the development of individualized treatment options regarding EGFR-TKI in NSCLC patients.

ALK gene

Fusion of *ALK* with an upstream gene, *EML4*, was first found in lung adenocarcinoma patients in 2007 (25). Biomarker detection of *ALK* mutation for NSCLC is essential because of the sensitivity of this disease to the *ALK* fusion inhibitor, which significantly improves response rates and survival (3). The prevalence of the *EML4-ALK*

fusion product found in our study is in agreement with that in previous studies. Literature on *EML4-ALK* fusion that mostly covers East Asian patients indicates that the frequency of *EML4-ALK* fusion ranges from 3% to 7% in unselected patients with NSCLC (26-28). A larger population from Switzerland and America showed a prevalence of 3% (16/603) (29). In summary, unlike *EGFR* mutations, non-selected *EML4-ALK* fusion may not be influenced by ethnic differences.

EML4-ALK fusion defines a new molecular subset of NSCLC with distinct clinical features. We found that *EML4-ALK* fusion was more prevalent in adenocarcinoma patients with younger age, no smoking history, stage IV disease and no *EGFR* mutations. Various studies on the clinical features of NSCLC with *EML4-ALK* fusion yielded different results. It is widely recognized that light smokers or never smokers, younger age and lack of an *EGFR* mutation are related to the incidence of *EML4-ALK*, and adenocarcinomas are the major type that presents with this mutation, but *EML4-ALK* fusion can also be found in squamous cancer patients (26,27). An investigation conducted by Shaw *et al*. involving subpopulations demonstrated that in the group of never/light smokers, the frequency of *EML4-ALK* was 22% among those without EGFR mutations, and *EML4-ALK* fusion was found in 33% of patients (30), which is much higher than the frequency of EML4-ALK fusion in a non-selected population. These findings further suggested that smoking status and *EGFR* mutation influence the existence of *EML4-ALK* fusion. In our study, the median age of patients with *EML4-ALK* fusion is 56 years, which is younger than most NSCLC patients. Shaw *et al.* also showed that the majority of patients with *EML4-ALK* fusion had metastatic disease at the time of screening (30), which may reflect the aggressiveness of lung cancers with *EML4-ALK* fusion. This partially explains the result obtained from this study that stage IV was a predictive factor for *EML4-ALK* fusion.

EML4-ALK fusion was more prevalent in *EGFR*-negative patients, and the mutations were considered mutually exclusive. The ratio of *EML4-ALK* fusion in adenocarcinomas without *EGFR* mutation was significantly higher than that in all adenocarcinomas. Recent findings of the coexistence of *EGFR* and *ALK* mutations have been reported (31-33). We also found that approximately 1% of patients positive for *EML4-ALK* fusion had *EGFR* mutations. Studies have reported that ALK fusion genes in double mutations are all V1 type (31), which means that the V1 type *ALK* fusion gene is compatible with other types of *EGFR* mutations. Unlike the specificity of *EML4-ALK*, different types of *EGFR* mutations, including not only classical mutation types but also non-classical types, can be present in the double mutation, for example, the L861Q/G719X sensitive mutation and the T790M resistant mutation. Currently, there is no consensus on standard therapy for tumours with double mutations. According to case reports, different responses and resistance to EGFR-TKIs and *ALK* inhibitors have been described (34-36). The choice of first-line targeted drugs might be related to detection techniques, mutation abundance, and phosphorylation levels of *EGFR* and *ALK* (37). Further investigation is needed to determine the relative oncogenic role of the genomic changes of *EGFR* and *EML4-ALK*.

ROS1 gene

As a member of the tyrosine kinase receptor family, *ROS1* rearranges the tyrosine kinase region and promotes tumour cell growth and tumour formation (38). The prevalence of *ROS1* rearrangement is rare. In our study on Chinese patients, the overall prevalence of *ROS1* rearrangement was 1.3% in NSCLC, which is similar to previous findings (39-41), and up to 1.6% in lung adenocarcinoma, which is lower than the 2.4% prevalence reported in East Asian patients (42). We confirmed that age and tumour stage were independent predictive factors for both *ROS1* rearrangement and *EML4-ALK* fusion, which might be related to distant metastasis in most young patients suffering from lung cancer. Jin *et al.* reported that *ROS1* rearrangement is highly associated with the micropapillary component and aerogenous spread, the latter of which has been identified as a marker of aggressive tumour biology characteristics (43). This finding may result in patients at advanced stages (stages III, IV) tending to harbour *ROS1* rearrangements. Davies *et al.* identified *ROS1* fusions in squamous cell carcinoma histology for the first time (39). However, in our study, there were no *ROS1* rearrangements in patients with squamous cancer, which might be related to the low positive rate reported in the literature. Studies analysing the associations between *ROS1* and clinical features were varied, possibly due to a very low frequency of *ROS1*. The clinical features of *ROS1* fusion need further studies containing large samples to confirm our results.

Although *ROS1* is a distinct receptor tyrosine kinase, it is structurally similar to the *ALK* protein due to similarities in their kinase domains and ATP binding sites; furthermore, the *ALK* and *ROS1* kinase domains share 77% sequence

identity (39,44,45). This finding may explain why the clinical profile of patients with *ROS1* rearrangement is remarkably similar to that of patients with *ALK* fusion, including a young age of onset, no history of smoking, and an advanced tumour stage, and both are treatable with crizotinib. The similar characteristics between these two mutation types suggest that they may share a common pathogenesis, possibly relating to environmental or genetic risk factors.

Several clinical studies have shown that crizotinib has a high response rate of 72–80%. Although *ROS1* rearrangements represent a small fraction of the overall NSCLC population, the use of *ROS1* inhibitors significantly influences patient survival and prognosis. The newly updated guidelines in the CSCO indicated that *ROS1* testing ranges from optional to basic strategies (11). The accurate and extensive detection of *ROS1* rearrangement can help optimize individualized treatment.

Conclusions

In summary, this real-world analysis of *EGFR*, *EML4-ALK* and *ROS1* mutations in patients from five regions of China will guide clinical practice for lung cancer. The positive rate of *EGFR* mutation in a large population of China is far higher than that in most other regions. According to our clinical analysis, *EGFR* mutation testing should be considered not only for patients who are females, non-smokers and have adenocarcinoma but also for all adenocarcinoma patients who are male, are at an early stage of cancer, and have a history of smoking and even for squamous cancer patients with typical clinical features. The development of effective EGFR-TKIs for non-rare non-classical mutations requires additional clinical trials. Our results confirmed that patients who are younger, have stage IV disease and do not smoke are prone to *EML4-ALK* fusion and are highly recommended for gene testing. Although *ROS1* rearrangements are rare, the observed prevalence of *ROS1* fusion and clinical features of Chinese patients suggests that detection should be considered for all adenocarcinoma patients who are younger or have advanced stage. Additional studies on the epidemiological characteristics and clinical features of driver mutations are ongoing.

Objective evaluation

The investigation extensively enrolled patients with NSCLC at different sites in China. This study had widespread coverage and a large sample size. We used a detection method that can simultaneously detect three mutant genes, which reduces certain deviations. Moreover, the samples were from multiple centres, which makes this a representative study and indicates higher confidence in the results. The limitations of this study are related to missing information; some of the information was not available because the clinical data collection process was inconsistent, and there was no uniform quality control of the collected samples.

Acknowledgments

The authors thank the patients and data collectors from the five participating hospitals. The authors also wish to thank Amoy Diagnostics Co., Ltd., Xiamen, China, for providing the testing method.
Funding: None.

Footnote

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/tcr.2019.10.28). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. This study was approved by the Ethics Committee of the First Hospital of Jilin University. We have obtained written informed consent from all study participants. All of the procedures were performed in accordance with the Declaration of Helsinki (as revised in 2013) and relevant policies in China.

References

1. Ettinger DS, Aisner DL, Wood DE, et al. NCCN

Guidelines Insights: Non-Small Cell Lung Cancer, Version 5. 2018. J Natl Compr Canc Netw 2018;16:807-21.
2. Mitsudomi T, Morita S, Yatabe Y, et al. Gefitinib versus cisplatin plus docetaxel in patients with non-small-cell lung cancer harbouring mutations of the epidermal growth factor receptor (WJTOG3405): an open label, randomised phase 3 trial, Lancet Oncol 2010;11:121-8.
3. Solomon BJ, Mok T, Kim DW, et al. First-line crizotinib versus chemotherapy in ALK-positive lung cancer. N Engl J Med 2014;371:2167-77.
4. Gou LY, Wu YL. Prevalence of driver mutations in non-small-cell lung cancers in the People's Republic of China. Lung Cancer (Auckl) 2014;5:1-9.
5. Kalemkerian GP, Narula N, Kennedy EB, et al. Molecular Testing Guideline for the Selection of Patients With Lung Cancer for Treatment With Targeted Tyrosine Kinase Inhibitors: American Society of Clinical Oncology Endorsement of the College of American Pathologists/ International Association for the Study of Lung Cancer/ Association for Molecular Pathology Clinical Practice Guideline Update. J Clin Oncol 2018;36:911-9.
6. Zhang YL, Yuan JQ, Wang KF, et al. The prevalence of EGFR mutation in patients with non-small cell lung cancer: a systematic review and meta-analysis. Oncotarget 2016;7:78985-93.
7. Sekine I, Yamamoto N, Nishio K, et al. Emerging ethnic differences in lung cancer therapy. Br J Cancer 2008;99:1757-62.
8. Cortes-Funes H, Gomez C, Rosell R, et al. Epidermal growth factor receptor activating mutations in Spanish gefitinib-treated non-small-cell lung cancer patients. Ann Oncol 2005;16:1081-6.
9. Reinersman JM, Johnson ML, Riely GJ, et al. Frequency of EGFR and KRAS mutations in lung adenocarcinomas in African Americans. J Thorac Oncol 2011;6:28-31.
10. Zhong WZ, Wang Q, Mao WM, et al. Gefitinib versus vinorelbine plus cisplatin as adjuvant treatment for stage II-IIIA (N1-N2) EGFR-mutant NSCLC (ADJUVANT/ CTONG1104): a randomised, open-label, phase 3 study. Lancet Oncol 2018;19:139-48.
11. China Society of Clinical Oncology Guidelines Working Committee. China Society of Clinical Oncology (CSCO) Primary Lung Cancer Diagnosis and Treatment Guide. Beijing: People's Medical Publishing House, 2018:27-28.
12. Shi Y, Au JS, Thongprasert S, et al. A prospective, molecular epidemiology study of EGFR mutations in Asian patients with advanced non-small-cell lung cancer of adenocarcinoma histology (PIONEER). J Thorac Oncol 2014;9:154-62.
13. Rosell R, Moran T, Queralt C, et al. Spanish Lung Cancer, Screening for epidermal growth factor receptor mutations in lung cancer. N Engl J Med 2009;361:958-67.
14. Ohtsuka K, Ohnishi H, Fujiwara M, et al. Abnormalities of epidermal growth factor receptor in lung squamous-cell carcinomas, adenosquamous carcinomas, and large-cell carcinomas: tyrosine kinase domain mutations are not rare in tumors with an adenocarcinoma component. Cancer 2007;109:741-50.
15. Park SH, Ha SY, Lee JI, et al. Epidermal growth factor receptor mutations and the clinical outcome in male smokers with squamous cell carcinoma of lung. J Korean Med Sci 2009;24:448-52.
16. Miyamae Y, Shimizu K, Hirato J, et al. Significance of epidermal growth factor receptor gene mutations in squamous cell lung carcinoma. Oncol Rep 2011;25:921-8.
17. Hata A, Katakami N, Yoshioka H, et al. How sensitive are epidermal growth factor receptor-tyrosine kinase inhibitors for squamous cell carcinoma of the lung harboring EGFR gene-sensitive mutations? J Thorac Oncol 2013;8:89-95.
18. Fang W, Zhang J, Liang W, et al. Efficacy of epidermal growth factor receptor-tyrosine kinase inhibitors for Chinese patients with squamous cell carcinoma of lung harboring EGFR mutation. J Thorac Dis 2013;5:585-92.
19. Xu J, Chu T, Jin B, et al. Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitors in Advanced Squamous Cell Lung Cancer. Clin Lung Cancer 2016;17:309-14.
20. Lohinai Z, Hoda MA, Fabian K, et al. Distinct Epidemiology and Clinical Consequence of Classic Versus Rare EGFR Mutations in Lung Adenocarcinoma. J Thorac Oncol 2015;10:738-46.
21. Chiu CH, Yang CT, Shih JY, et al. Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitor Treatment Response in Advanced Lung Adenocarcinomas with G719X/L861Q/S768I Mutations. J Thorac Oncol 2015;10:793-9.
22. Shen YC, Tseng GC, Tu CY, et al. Comparing the effects of afatinib with gefitinib or Erlotinib in patients with advanced-stage lung adenocarcinoma harboring non-classical epidermal growth factor receptor mutations. Lung Cancer 2017;110:56-62.
23. O'Kane GM, Bradbury PA, Feld R, et al. Uncommon EGFR mutations in advanced non-small cell lung cancer, Lung Cancer 2017;109:137-44.
24. Tu HY, Ke EE, Yang JJ, et al. A comprehensive review of uncommon EGFR mutations in patients with non-small

cell lung cancer. Lung Cancer 2017;114:96-102.
25. Solomon B, Varella-Garcia M, Camidge DR, et al. ALK gene rearrangements: a new therapeutic target in a molecularly defined subset of non-small cell lung cancer. J Thorac Oncol 2009;4:1450-4.
26. Wong DW, Leung EL, So KK, et al. The EML4-ALK fusion gene is involved in various histologic types of lung cancers from nonsmokers with wild-type EGFR and KRAS. Cancer 2009;115:1723-33.
27. Inamura K, Takeuchi K, Togashi Y, et al. EML4-ALK fusion is linked to histological characteristics in a subset of lung cancers. J Thorac Oncol 2008;3:13-7.
28. Soda M, Choi YL, Enomoto M, et al. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. Nature 2007;448:561-6.
29. Perner S, Wagner PL, Demichelis F, et al. EML4-ALK fusion lung cancer: a rare acquired event. Neoplasia 2008;10:298-302.
30. Shaw AT, Yeap BY, Mino-Kenudson M, et al. Clinical features and outcome of patients with non-small-cell lung cancer who harbor EML4-ALK. J Clin Oncol 2009;27:4247-53.
31. Mizuno T. The biphasic role of microglia in Alzheimer's disease. Int J Alzheimers Dis 2012;2012:737846.
32. Masters CL, Beyreuther K. Alzheimer's centennial legacy: prospects for rational therapeutic intervention targeting the Abeta amyloid pathway. Brain 2006;129:2823-39.
33. Streit WJ. Microglia and neuroprotection: implications for Alzheimer's disease. Brain Res Brain Res Rev 2005;48:234-9.
34. Sasaki T, Koivunen J, Ogino A, et al. A novel ALK secondary mutation and EGFR signaling cause resistance to ALK kinase inhibitors. Cancer Res 2011;71:6051-60.
35. Popat S, Vieira de Araujo A, Min T, et al. Lung adenocarcinoma with concurrent exon 19 EGFR mutation and ALK rearrangement responding to erlotinib. J Thorac Oncol 2011;6:1962-3.
36. Kwak EL, Bang YJ, Camidge DR, et al. Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. N Engl J Med 2010;363:1693-703.
37. Lou NN, Zhang XC, Chen HJ, et al. Clinical outcomes of advanced non-small-cell lung cancer patients with EGFR mutation, ALK rearrangement and EGFR/ALK co-alterations. Oncotarget 2016;7:65185-95.
38. Takeuchi K, Soda M, Togashi Y, et al. RET, ROS1 and ALK fusions in lung cancer. Nat Med 2012;18:378-81.
39. Davies KD, Le AT, Theodoro MF, et al. Identifying and targeting ROS1 gene fusions in non-small cell lung cancer. Clin Cancer Res 2012;18:4570-9.
40. Bergethon K, Shaw AT, Ou SH, et al. ROS1 rearrangements define a unique molecular class of lung cancers. J Clin Oncol 2012;30:863-70.
41. Rimkunas VM, Crosby KE, Li D, et al. Analysis of receptor tyrosine kinase ROS1-positive tumors in non-small cell lung cancer: identification of a FIG-ROS1 fusion. Clin Cancer Res 2012;18:4449-57.
42. Chen YF, Hsieh MS, Wu SG, et al. Clinical and the prognostic characteristics of lung adenocarcinoma patients with ROS1 fusion in comparison with other driver mutations in East Asian populations. J Thorac Oncol 2014;9:1171-9.
43. Jin Y, Sun PL, Park SY, et al. Frequent aerogenous spread with decreased E-cadherin expression of ROS1-rearranged lung cancer predicts poor disease-free survival. Lung Cancer 2015;89:343-9.
44. Selinger CI, Li BT, Pavlakis N, et al. Screening for ROS1 gene rearrangements in non-small-cell lung cancers using immunohistochemistry with FISH confirmation is an effective method to identify this rare target. Histopathology 2017;70:402-11.
45. Solomon B. Validating ROS1 rearrangements as a therapeutic target in non-small-cell lung cancer. J Clin Oncol 2015;33:972-4.

Cite this article as: Sun M, Guo Y, Shao G, Duan X, Yang Z, Zhang P, Liu Y, Dong Y, Wang X, Xu Y, Sun C, Ma K. Analysis of real-word mutations of lung cancer driver genes in five regions of China. Transl Cancer Res 2019;8(7):2581-2592. doi: 10.21037/tcr.2019.10.28

