Prevalence of driver mutations in non-small-cell lung cancers in the People’s Republic of China

Lan-Ying Gou1,2
Yi-Long Wu1
1Guangdong Lung Cancer Institute, Guangdong General Hospital and Guangdong Academy of Medical Sciences, 2Southern Medical University, Guangzhou, People’s Republic of China

Abstract: Lung cancer is a leading cause of cancer-related mortality worldwide and in the People’s Republic of China. Recently, the pathological proportions of the various forms of lung cancer have changed. A shift to a preponderance of adenocarcinoma at the expense of squamous cell carcinoma is observable. Treatment decisions have historically been based on tumor histology, and evolution of our molecular understanding of cancer has led to development of targeted therapeutic agents. It is essential to further understand mutations that drive cancer development (driver mutations) in relevant genes and their effects on cancer cell proliferation and survival. The epidemiology of lung cancer in the People’s Republic of China has been extensively reviewed elsewhere. However, molecular epidemiological data from mainland China are scarce. Consequently, we herein review the prevalence of driver mutations in Chinese patients.

Keywords: lung cancer, driver mutation, prevalence, EGFR, EML4-ALK, KRAS, ROS1, PIK3CA, BRAF, RET, HER2

Introduction

As reported in the Chinese Cancer Registry Annual Report in 2012,1 lung cancer is the leading cause of cancer-related death in the People’s Republic of China. Mortality from this cancer was 45.57 per 100,000 of the national population, being 61.00 for 100,000 males and 29.77 for 100,000 females. Male mortality was 2.05-fold that of female mortality. A shift in the relative frequency of lung cancer type from squamous cell cancer to adenocarcinoma is evident both domestically and overseas in countries where numbers of male smokers remain high.2 Most Chinese women do not smoke. However, in most areas, the frequency of smoking is increasing among women. Wang et al3 studied 32,845 patients with newly diagnosed lung cancer between 1998 and 2007 (Table 1). In terms of cancer subtype, the proportion of squamous cell cancers decreased annually from 30.41% (333/1,095) in 1998 to 24.16% (638/2,641) in 2007, while the proportion of adenocarcinomas increased from 42.83% (469/1,095) to 46.80% (1,236/2,641) over the same period. In women, the decline was more marked; the proportions of squamous cell cancer to adenocarcinoma were changed to 14.77% (925/6,262) and 60.83% (3,809/6,262), respectively, between 1998 and 2007. Alamoudi et al4 and Chang et al5 also found that the principal subtype of lung cancer that was histologically categorized apparently changed from squamous cell cancer to adenocarcinoma. However, the reason for this change requires further assessment.

Over the past few decades, translational research has clarified many of the molecular mechanisms for development, growth, and metastasis of lung cancer. Driver mutations occur in genes encoding signaling proteins critical in terms of cellular proliferation...
and survival. In lung adenocarcinomas, such mutations occur in the EGFR, EML4-ALK, KRAS, ROS1, PIK3CA, BRAF, RET, and HER2 genes. Tyrosine kinase inhibitors (TKIs) targeting these genes have been developed. Such drugs include gefitinib, erlotinib, and crizotinib, which were used in the IPASS, BR.21, and Profile 1007 clinical trials, respectively.6-9 Many other new compounds, including the MET inhibitor tivantinib10 and the phosphatidylinositol 3-kinase inhibitor XL147,11 which directly or indirectly target these driver mutations, have been developed and are undergoing clinical trials. Non-smoking-related adenocarcinoma of the lung has its peculiar epidemiologic, clinical, and biological characteristics. Many experts, including Sun et al12 and Yano et al,13 have named it as a distinct entity. In recent years, many studies have shown that EGFR, EML4-ALK, HER2, KRAS, and BRAF mutations are mainly found in non-smoking-related adenocarcinoma.

The People’s Republic of China has the highest population of any country. However, molecular epidemiological data from mainland China remain scarce. A comprehensive review of the prevalence of driver mutations is essential for development of personalized therapy targeting non-small-cell lung cancer (NSCLC). Consequently, we herein review the prevalence of driver mutations and the epidemiology of resistance mechanisms in Chinese patients with NSCLC.

### Prevalence of driver mutations in lung cancers

#### EGFR mutations

Currently, a promising treatment strategy for advanced NSCLC features small molecules targeting epidermal growth factor receptor (EGFR) mutations.14 The prevalence of EGFR mutations in pulmonary adenocarcinoma among female non-smoking Asian patients was first revealed by a subgroup analysis of the Iressa Survival Evaluation in Lung Cancer trial.15 Further, several Phase III studies7,16,17 have compared first-line EGFR TKIs such as gefitinib and erlotinib with double platinum-based chemotherapy, used to treat NSCLC patients harboring activating EGFR mutations. It was found that EGFR TKIs were associated with significantly higher response rates and afforded longer progression-free survival than did traditional chemotherapy in selected populations.

Currently, adenocarcinoma of the lung in non-smokers is recognized as a distinct disease entity because of its peculiar epidemiologic, clinical, and biological characteristics; many studies have comprehensively analyzed the major known driver mutations in female non-smoking Asian patients with pulmonary adenocarcinoma. EGFR mutations are associated with smoking status and particular histological subtypes of disease. Wu et al18 performed a meta-analysis of up-to-date individual patient data from six medical centers of mainland China. Of 506 patients with NSCLC, the EGFR mutation rate in smokers was 15.1%, and lower than in non-smokers (45.5%; P<0.00001). Male patients showed a lower mutation rate than female patients (23.1% versus 42.9%; P<0.0001). An et al19 found that EGFR mutation rates varied by smoking status and the histological lung cancer subtype; these authors analyzed differences in EGFR mutations among patients varying in these parameters. In non-smokers with adenocarcinoma, EGFR was the most frequently altered gene (in 49.8%, 114/229). The mutation rates of EGFR in patients with adenocarcinoma and squamous cell cancer were 40.3% (140/347) and 4.4% (6/144), respectively.

In summary, EGFR mutations occur at high frequencies in female patients, non-smokers, those of Asian ethnicity, and patients with adenocarcinoma. Earlier Chinese studies19-27 recorded the frequencies of EGFR mutations in patients with NSCLC, adenocarcinoma, and squamous cell cancer (Table 2). Such mutations include exon 18 G719A, G719V, and G719D; exon 19 E746–A750 del, E746–S752 del, and L747–A750 (751, 753) del; exon 20 T790M and R776H; and exon 21 L858R, L858M, and L861R. As reported earlier, most EGFR mutations occur in exons 19 and 21. We have summarized all previous Chinese data in Figure 1D.

#### KRAS mutations

The KRAS oncogene is often mutated during carcinogenesis. Many reports have linked KRAS mutations to lung cancer.28,29

<table>
<thead>
<tr>
<th>Table 1 Distribution of pathological types of lung cancer according to gender in Beijing 1998–2007</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pathological type</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>AD</td>
</tr>
<tr>
<td>SCC</td>
</tr>
<tr>
<td>SCLC</td>
</tr>
</tbody>
</table>

Abbreviations: SCC, squamous cell cancer; AD, adenocarcinoma; SCLC, small cell lung cancer.
### Table 2: Frequency of EGFR and KRAS mutation and ALK rearrangements in different subgroups

<table>
<thead>
<tr>
<th>Study</th>
<th>Techniques</th>
<th>EGFR mutation</th>
<th>KRAS mutation</th>
<th>ALK rearrangements</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>AD</td>
<td>SCC</td>
<td>NSCLC</td>
</tr>
<tr>
<td>Ren et al(^{11})</td>
<td>Sequencing</td>
<td>70.2%</td>
<td>(73/104)</td>
<td>1.9%</td>
</tr>
<tr>
<td>Sun et al(^{14})</td>
<td>Sequencing</td>
<td>78.8%</td>
<td>(41/52)</td>
<td>5.4%</td>
</tr>
<tr>
<td>An et al(^{19})</td>
<td>Sequencing</td>
<td>40.3%</td>
<td>(140/347)</td>
<td>7.4%</td>
</tr>
<tr>
<td>Li et al(^{20})</td>
<td>Sequencing</td>
<td>43.5%</td>
<td>(100/230)</td>
<td>16.5%</td>
</tr>
<tr>
<td>Xu et al(^{21})</td>
<td>ARMS</td>
<td>48.1%</td>
<td>(149/310)</td>
<td>3.0%</td>
</tr>
<tr>
<td>Li et al(^{21})</td>
<td>Sequencing</td>
<td>44.2%</td>
<td>(42/95)</td>
<td>1.5%</td>
</tr>
<tr>
<td>Shaozhang et al(^{22})</td>
<td>Sequencing</td>
<td>45.8%</td>
<td>(33/72)</td>
<td>7.7%</td>
</tr>
<tr>
<td>Sun et al(^{23})</td>
<td>Sequencing</td>
<td>32.5%</td>
<td>(26/80)</td>
<td>7.4%</td>
</tr>
<tr>
<td>Wu et al(^{24})</td>
<td>Sequencing</td>
<td>57.3%</td>
<td>(90/157)</td>
<td>4.6%</td>
</tr>
<tr>
<td>Xu et al(^{25})</td>
<td>Sequencing</td>
<td>47.5%</td>
<td>(317/676)</td>
<td>5.2%</td>
</tr>
<tr>
<td>Gao et al(^{26})</td>
<td>ARMS</td>
<td>40.9%</td>
<td>(38/93)</td>
<td>10.4%</td>
</tr>
<tr>
<td>Li et al(^{27})</td>
<td>HRM</td>
<td>4.3%</td>
<td>(3/69)</td>
<td>3.1%</td>
</tr>
<tr>
<td>Guang et al(^{28})</td>
<td>HRM</td>
<td>5.6%</td>
<td>(98/1735)</td>
<td>5.6%</td>
</tr>
<tr>
<td>Wang et al(^{29})</td>
<td>ARMS</td>
<td>5.8%</td>
<td>(16/288)</td>
<td>4.7%</td>
</tr>
<tr>
<td>Sun et al(^{30})</td>
<td>Sequencing</td>
<td>48.4%</td>
<td>(675/1,395)</td>
<td>4.3%</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>48.4%</td>
<td>(675/1,395)</td>
<td>4.3%</td>
</tr>
</tbody>
</table>

**Abbreviations:** NSCLC, non-small-cell lung cancer; AD, adenocarcinoma; SCC, squamous cell carcinoma; ARMS, amplification refractory mutation system; HRM, high resolution melting analysis; RT-PCR, reverse transcriptase polymerase chain reaction; qPCR, quantitative polymerase chain reaction; IHC, immunohistochemistry; FISH, fluorescence in situ hybridization; RACE, rapid cloning of cDNA ends.
However, the frequency of \textit{KRAS} mutations is lower in Eastern Asia than in the West. The \textit{KRAS} mutation frequency is 30% or over in NSCLC patients from the USA and some other Western countries, but less than 10% in patients from Japan, Korea, Taiwan, and Hong Kong (Table 2).\textsuperscript{20,24,25,30–34} Other research\textsuperscript{19,26,35} showed that \textit{KRAS} mutations occurred more often in males than females, in patients with adenocarcinoma compared with other forms of lung cancer, and in smokers compared with non-smokers.

Several \textit{KRAS} mutations are known. Codon 12 mutations include G12A, G12C, G12D, G12R, G12S, G12V, G12V/C, and G12V/D. Codon 13 mutations include G13C and G13D. Guan et al\textsuperscript{35} found \textit{KRAS} mutations in codons 12 (81.32%, 74/1,935), 13 (5.49%, 51/1,935), 59 (1.10%, 11/1,935), 60 (3.30%, 61/1,935), and both 12 and 13 (5.49%, 31/1,935). The only codon 59 mutation was G59G. The codon 60 mutations were Q60H and Q60L. Mutations in both codons 12 and 13 consisted of G12D/G13D (n=3), G12C/G13C (n=1), and G12V/G13C (n=1). An et al\textsuperscript{19} enrolled 524 patients and found that 27 had \textit{KRAS} mutations, all in codon 12, including G12C (44.4%, 12/27), G12V (18.5%, 5/27), G12D (18.5%, 5/27) and others (18.5%, 5/27).

In summary, most studies support the idea that \textit{KRAS} mutations occur in Chinese smokers with adenocarcinoma. We analyzed studies\textsuperscript{19,22–27,34,35} performed in Chinese patients with lung cancer to identify the subtypes and frequencies of \textit{KRAS} mutations, and found that most occurred in codons 12, 13, 59, and 60, especially codon 12 (88%; Figure 1F). Further, \textit{EGFR} mutations were never found in tumors with \textit{KRAS} mutations, suggesting that these mutations are mutually exclusive.

\textbf{ALK fusion genes}

\textit{ALK} is a receptor tyrosine kinase that is not normally expressed in the lung.\textsuperscript{36} Fusions of \textit{ALK} with an upstream gene, \textit{EML4}, were found in NSCLC patients in 2007.\textsuperscript{37} The frequency of \textit{EML4-ALK} translocations ranges from 3% to 7% in unselected patients with NSCLC.\textsuperscript{38–40} Detection methods include reverse-transcriptase polymerase chain reaction (RT-PCR), fluorescence in situ hybridization, and immunohistochemistry. As with \textit{EGFR} mutations, the frequency of \textit{EML4-ALK} fusions is elevated in patients with adenocarcinoma, in younger adult patients, and in those who have never smoked (<100 cigarettes in a lifetime) or who are light smokers (≤15 pack-years). \textit{EML4-ALK} translocations are generally found in tumors wild-type for \textit{EGFR} and \textit{KRAS}.\textsuperscript{38} The \textit{EML4-ALK} translocation occurs in a low proportion of lung cancers, ie, in 4%–11.6% of unselected Asian patients with early stage NSCLC.\textsuperscript{23,41–43} However, only small numbers of \textit{EML4-ALK}-positive patients have been studied, because patient enrolment in most cited studies was non-selective. Thus, the relationship between...
gene fusion and clinical and pathological profiles requires further study.

A careful review of the literature in the Chinese language revealed that ALK fusion proteins are frequently present in patients with lung cancer (Table 2). This table lists the frequencies of ALK fusions in patients with NSCLC, adenocarcinoma, and squamous cell cancer, respectively. In summary, the results show that risk factors for development of the EML4-ALK fusion are female gender, adenocarcinoma, non-smoking status, and relative youth. Patients with EML4-ALK fusions are always wild-type for EGFR and KRAS.

EML4-ALK fusions are caused by various small inversions within the short arm of chromosome 2. At least nine variants have been identified. Exons 20–29 of ALK may be fused to EML4 exon 13 (variant 1, V1), exon 20 (V2), exon 6 (V3a), exon 6 with an 11 amino acid insertion (V3b), exon 14 with an additional 11 nucleotide insertion of unknown origin at nucleotide 50 of exon 20 of ALK (V4), exon 2 (V5), exon 13 (V6), exon 14 with the fusion beginning at nucleotide 13 of exon 20 of ALK (V7), exon 15 (also termed “V4” [V8]), and exon 18 (also termed “V5” [V9]). The details can be found in a review by Horn and Pao. For example, Wong et al used RT-PCR to show that the frequencies of the fusion variants 1, 2, 3, and 9 were 2/13, 2/13, 8/13, and 1/13, respectively, in Chinese patients with lung cancer. Zhang et al found that variants 1, 2, and 3 were each present alone in individual Chinese patients with NSCLC using RT-PCR. Li et al found that the frequencies of EML4-ALK variants 1, 2, 3, and 4 were 2/7, 1/7, 3/7, and 1/7, in Chinese patients with lung adenocarcinoma using RT-PCR. Similarly, Li et al found that the frequencies of fusion variants 1, 2, 3, 5, and 9 were 14/22, 3/22, 3/22, 1/22, and 1/22, respectively. No Chinese patient has been found to be positive for variant 6, 7, or 8 (Figure 1E).

**ROS1 mutation**

ROS1 rearrangements were initially identified in a human glioblastoma cell line. Such rearrangements have been identified in approximately 1%–2% of NSCLC patients using different genotyping techniques (Table 3). Several recent studies used whole-transcriptome and whole-genome sequencing to detect ROS1 rearrangements, and several novel fusion partners were identified. Such methods are not yet readily applicable in the clinic. Therefore, ROS1 screening strategies have been largely informed by experience with ALK testing, for which the three most commonly used methods of detection are fluorescence in situ hybridization, RT-PCR, and immunohistochemistry. ROS1 fusions appear to be more common in patients with adenocarcinoma. Cai et al found that the ratios of females to males and never-smokers to smokers in a ROS1 fusion-positive group were 5:3. No significant differences in age (P=0.866), gender (P=0.479), smoking history (P=1.0), histological tumor type (P=0.148), or pathological stage (P=0.475) were evident between ROS1 fusion-positive and fusion-negative patients.

**General spectra of mutations in PIK3CA, BRAF, RET, and HER2**

Recently, many studies have analyzed the spectra of well known driver mutations in non-smokers with adenocarcinoma of the lung. Table 3 shows the oncogenic driver mutations in the ROSI, BRAF, PIK3CA, RET, and HER2 genes of such highly selected Chinese patients. Targeted therapies are developing rapidly, and molecular data in conjunction with clinical and pathological features suggest that prospective genotyping of lung adenocarcinomas in smokers, in terms of changes in the genes listed above, may facilitate targeted therapy in almost all cases.

As with EGFR, the HER2 protein is a member of the HER family of receptor tyrosine kinases. The protein forms homodimers or heterodimers with other members of the same family. HER2 is overexpressed in about 20% of patients with NSCLC, but gene amplification occurs in only 2%. Early clinical trials with trastuzumab, a humanized monoclonal antibody against HER2 (which is effective to treat breast cancers in which HER2 is amplified), exerted only slight effects in unselected patients with NSCLC. Xu et al and others have shown that the frequency of the PIK3CA mutation in NSCLC patients is 2%–4%, and the mutation occurs less often in those with adenocarcinomas compared with other forms of lung cancer.

We have summarized the data from the Chinese studies, and describe the general spectrum of mutations in EGFR, ALK, KRAS, PIK3CA, BRAF, RET, and HER2 in patients with NSCLC, adenocarcinoma, and squamous cell cancer, respectively (Figure 1A–C).
<table>
<thead>
<tr>
<th>Study</th>
<th>Techniques</th>
<th>PREVALENCE</th>
<th>PIK3CA</th>
<th>BRAF</th>
<th>RET</th>
<th>HER2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xu et al²⁶</td>
<td>Sequencing</td>
<td></td>
<td>2.7%</td>
<td>7.2%</td>
<td>3.7%</td>
<td>0.7%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(18/667)</td>
<td>(14/194)</td>
<td>(32/861)</td>
<td>(6/861)</td>
</tr>
<tr>
<td>An et al²⁹</td>
<td>HRM</td>
<td></td>
<td>4.2%</td>
<td>5.8%</td>
<td>4.4%</td>
<td>2.3%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(13/307)</td>
<td>(7/121)</td>
<td>(20/452)</td>
<td>(7/307)</td>
</tr>
<tr>
<td>Ren et al³¹</td>
<td>Sequencing</td>
<td></td>
<td>1.9%</td>
<td>1.9%</td>
<td>0%</td>
<td>2.3%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(2/104)</td>
<td>(2/104)</td>
<td></td>
<td>(2/104)</td>
</tr>
<tr>
<td>Sun et al³⁴</td>
<td>Sequencing</td>
<td></td>
<td>7.6%</td>
<td>0</td>
<td>0%</td>
<td>1.5%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(4/52)</td>
<td></td>
<td></td>
<td>(8/320)</td>
</tr>
<tr>
<td>Li et al³⁰</td>
<td>Sequencing</td>
<td></td>
<td>3.5%</td>
<td>3.0%</td>
<td>0%</td>
<td>1.5%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(8/230)</td>
<td>(7/230)</td>
<td></td>
<td>(0/230)</td>
</tr>
<tr>
<td>Zheng et al³⁶</td>
<td>Sequencing</td>
<td></td>
<td>2.2%</td>
<td>2.2%</td>
<td>1.5%</td>
<td>1.6%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(9/409)</td>
<td>(9/409)</td>
<td>(6/409)</td>
<td>(1/609)</td>
</tr>
<tr>
<td>Xu et al³⁷</td>
<td>Sequencing</td>
<td></td>
<td>0%</td>
<td>0%</td>
<td>1.6%</td>
<td>1.6%</td>
</tr>
<tr>
<td>Cai et al³⁸</td>
<td>RT-PCR</td>
<td></td>
<td>3.0%</td>
<td>0</td>
<td>2.0%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(7/231)</td>
<td>(0/19)</td>
<td>(8/392)</td>
<td>(0/392)</td>
</tr>
<tr>
<td>Rinkunas et al³⁹</td>
<td>FISH, IHC</td>
<td></td>
<td>3.3%</td>
<td>0</td>
<td>1.6%</td>
<td>1.6%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(8/246)</td>
<td>(0/226)</td>
<td>(9/556)</td>
<td>(9/556)</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>3.1%</td>
<td>1.9%</td>
<td>3.3%</td>
<td>6.7%</td>
</tr>
</tbody>
</table>

**Abbreviations:** NSCLC, non-small-cell lung cancer; AD, adenocarcinoma; SCC, squamous cell carcinoma; HRM, high resolution melting analysis; RT-PCR, reverse transcriptase polymerase chain reaction; IHC, immunohistochemistry; FISH, fluorescence in situ hybridization.
Ongoing clinical research on EGFR-mutated NSCLCs will improve the survival of patients with somatic mutations such as T790M, MET, or IGF-1R.

**Conclusion**

Currently, non-smoker lung adenocarcinoma is the principal subtype of lung cancer, and is recognized as a distinct entity because of its peculiar biological characteristics. The frequency of EGFR mutations ranges from 28.0% in unselected Chinese patients with NSCLC to 48.5% in those with lung adenocarcinoma. The KRAS mutation frequency is lower in the People’s Republic of China than in Western countries. The frequency of the EML4-ALK fusion is 6.4% and this mutation is often found in females, patients with adenocarcinoma, non-smokers, and younger patients. ROS1 fusions appear to be more common in patients with adenocarcinoma. These observations have changed the treatment strategies for lung cancer. Genetic testing prior to treatment is now considered essential, to allow the best treatment option to be selected. First, EGFR mutation status should be tested; almost 70% of adenocarcinoma tumors have such a mutation. If the result is negative, further molecular testing is required. Treatment options include EGFR TKIs such as gefitinib or erlotinib or irreversible TKIs such as afatinib (if the T790M mutation is detected). If the tumor is wild-type in terms of EGFR, ALK translocations, found in about 30% of EGFR-negative tumors, should be sought. If a translocation is present, treatment would include an ALK inhibitor. If the tumor is negative for an ALK translocation, more rare mutations (in KRAS, ROSI, PIK3CA, BRAF, RET, or HER2) should be sought. Also, new driver mutations remain to be discovered.

**Disclosure**

The authors report no conflicts of interest in this work.
References

51. Hirsch FR, Langer CJ. The role of HER2/neu expression and trastu
47. Rimkunas VM, Crosby KE, Li D, et al. Analysis of receptor tyrosine
46. Birchmeier C, Sharma S, Wigler M. Expression and rearrangement of
42. Wong DW, Leung EL, So KK, et al. The EML4-ALK fusion gene is
driver genes in Chinese female lung adenocarcinoma[abstract].