

## Research Paper

# Mutational profiling of non-small-cell lung cancer patients resistant to first-generation EGFR tyrosine kinase inhibitors using next generation sequencing

Ying Jin<sup>1,2</sup>, Yang Shao<sup>3</sup>, Xun Shi<sup>1</sup>, Guangyuan Lou<sup>1</sup>, Yiping Zhang<sup>1</sup>, Xue Wu<sup>3</sup>, Xiaoling Tong<sup>3</sup>, Xinmin Yu<sup>1,4</sup>

<sup>1</sup>Department of Medical Oncology, Zhejiang Cancer Hospital, Hangzhou, China

<sup>2</sup>Zhejiang Key Laboratory of Radiation Oncology, Hangzhou, China

<sup>3</sup>Geneseeq Technology Inc., Toronto, Ontario, Canada

<sup>4</sup>Zhejiang Key Laboratory of Diagnosis and Treatment Technology of Thoracic Oncology, Hangzhou, China

**Correspondence to:** Xinmin Yu, **email:** yu\_xinm@sina.cn

**Keywords:** non-small-cell lung cancer, epithelial growth factor receptor, tyrosine kinase inhibitor, drug resistance, next generation sequencing

**Received:** April 29, 2016

**Accepted:** July 28, 2016

**Published:** August 12, 2016

## ABSTRACT

**Patients with advanced non-small-cell lung cancer (NSCLC) harboring sensitive epithelial growth factor receptor (EGFR) mutations invariably develop acquired resistance to EGFR tyrosine kinase inhibitors (TKIs). Identification of actionable genetic alterations conferring drug-resistance can be helpful for guiding the subsequent treatment decision. One of the major resistant mechanisms is secondary *EGFR-T790M* mutation. Other mechanisms, such as *HER2* and *MET* amplifications, and *PIK3CA* mutations, were also reported. However, the mechanisms in the remaining patients are still unknown. In this study, we performed mutational profiling in a cohort of 83 NSCLC patients with TKI-sensitizing *EGFR* mutations at diagnosis and acquired resistance to three different first-generation EGFR TKIs using targeted next generation sequencing (NGS) of 416 cancer-related genes. In total, we identified 322 genetic alterations with a median of 3 mutations per patient. 61% of patients still exhibit TKI-sensitizing *EGFR* mutations, and 36% of patients acquired *EGFR-T790M*. Besides other known resistance mechanisms, we identified *TET2* mutations in 12% of patients. Interestingly, we also observed *SOX2* amplification in *EGFR-T790M* negative patients, which are restricted to Icotinib treatment resistance, a drug widely used in Chinese NSCLC patients. Our study uncovered mutational profiles of NSCLC patients with first-generation EGFR TKIs resistance with potential therapeutic implications.**

## INTRODUCTION

Lung cancer is the leading cause of cancer death in China as well as worldwide [1]. Approximately 80% of lung cancers are non-small-cell lung carcinoma (NSCLC), and the overall 5-year relative survival rate for this cohort is less than 20% [2]. Patients with advanced NSCLC have an extremely high mortality rate. During the past decades, genomic research has increased our understanding of the molecular characterization of cancer [3-5]. The treatment strategy for advanced NSCLC has changed dramatically from the traditional chemotherapy depending on

pathologic histology to personalized precision medicine based on the oncogenic drivers [6].

The epidermal growth factor receptor (*EGFR*) gene mutations are found in ~10% of lung adenocarcinomas in Caucasian population [3], but in 30% ~ 50% of Asian population [7, 8], which define a substantial population that can benefit from the use of EGFR tyrosine kinase inhibitors (TKIs). Several randomized phase III clinical trials have shown the superiority in the overall response rate (ORR) and median progression-free survival (PFS) of EGFR TKI treatment over chemotherapy as first-line therapy for patients with TKI-sensitizing *EGFR* mutations

[9, 10]. However, the vast majority of patients inevitably experienced acquired resistance in less than one year, limiting the overall survival advantage of EGFR TKI treatment over chemotherapy [11, 12].

Currently, the known mechanisms of acquired resistance are as follows [13-17]: 1) the secondary gatekeeper *EGFR* T790M mutation which increases ATP affinity and subsequently prevents drug binding to the kinase domain; 2) activation of members of downstream signaling pathways such as RAS-RAF-ERK MAPK pathway and PI3K/AKT/mTOR pathway; 3) activation of bypass signaling through receptor tyrosine kinase such as MET; 4) changes in tumor histology with tumor cells displaying features of small-cell lung cancer or epithelial-mesenchymal transition (EMT). The above mechanisms account for about 70% of acquired resistance, with 30% of remaining patients demonstrating unknown resistant mechanisms.

The introduction of next generation sequencing (NGS) into cancer genetic interrogation achieved tremendous successes in acquiring cancer genomic information comprehensively and efficiently [18]. It demonstrates great potentials in identifying genetic aberrations that can be used to match targeted drugs and monitoring acquired genetic changes during the treatment with limit amount of tumor materials. To take advantage of this technology, we performed targeted NGS with a gene panel covering 416 cancer-related genes to profile genetic characteristics of 83 non-small cell lung cancer (NSCLC) patients after they developed systematically progress to the first generation EGFR TKI treatments, including erlotinib, gefitinib and icotinib. Besides *EGFR* T790M mutations, a variety of other previously known and novel genetic alterations were identified that might be potentially related to their primary and acquired resistance to treatments.

## RESULTS

### An overall characterization of cancer-related mutations identified in all patients

We analyzed either genomic DNAs from formalin-fixed paraffin embedded (FFPE) samples or pleural effusions, or circulating tumor DNAs (ctDNA) from plasma samples from 83 Chinese NSCLC patients with stage IV diseases at the time of developing drug resistance to the first generation of EGFR TKIs, erlotinib, gefitinib or icotinib. These patients were identified with TKI-sensitizing *EGFR* mutations prior to treatments and their characteristics were summarized in Table 1. The choice of collecting different tumor materials depends on clinical risks that would impose on the patients by the operation. 45 patients (54.2%) patients were undertaken blood withdrawing for testing ctDNA, while in others tumor tissues or pleural effusions were obtained through biopsies. Prior-treatment histology analysis confirmed that

68 patients (81.9%) were adenocarcinoma and 4 (4.8%) were squamous cell carcinoma. The rest 11 patients cannot be clearly distinguished based on histology appearance.. Half of patients were subjected to icotinib treatment upon diagnosis largely because of its lower cost compared to the other two options [19].

A total of 322 cancer-related genetic mutations were detected in these patients with a median of 3 mutations per patient and a range of 1-10 mutations per patient (Figure 1A). 87 genes within the 416-gene panel were involved. Majority of mutations (47%) were missense mutations and other types of mutations were also uncovered (Figure 1B).

### *EGFR* mutational status in all patients

30 of 83 patients (36.1%) were detected with *EGFR* T790M mutation and all of them except one were found harboring *EGFR* activating mutation either exon 19 deletion (19del) or L858R (Figure 2). 6 of them were accompanied with the copy number gain of *EGFR* and one of them harbors C797S mutation, which will exert resistance to the third generation EGFR TKI, AZD9291 [20]. Uncommon mutations including S752F and N826S were also identified in one case each, which might be related to the resistance to gefitinib and erlotinib according to previous reports [21, 22].

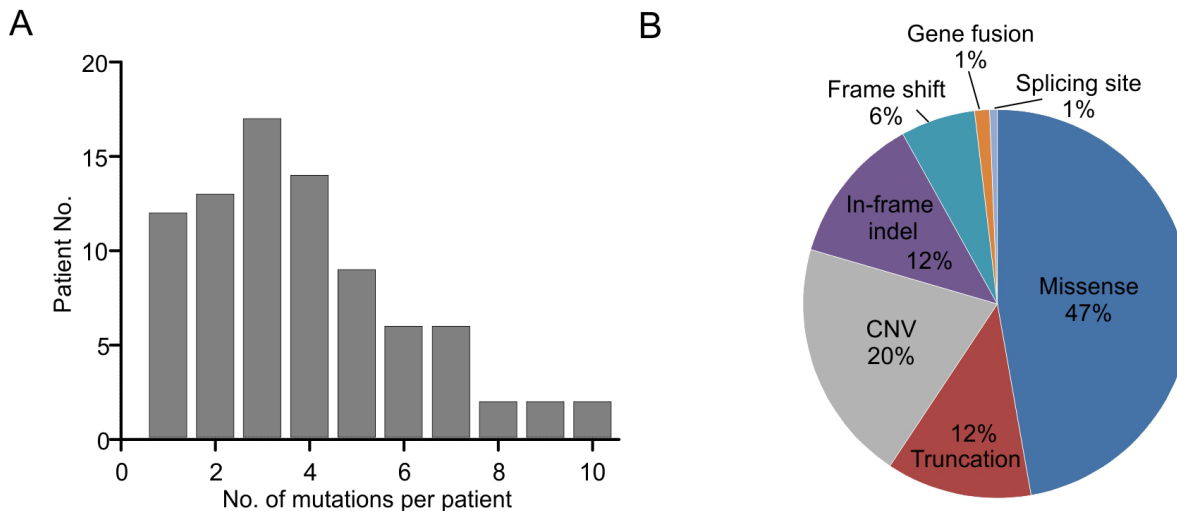
As to the other *EGFR*-T790M negative (T790M-) patients, in addition to the presence of 19del (23%) and L858R (17%), a variety of other infrequent *EGFR* mutations that were suggested less sensitive to the first generation TKIs were identified, including M766delinsMASV, D770delinsDNPH, L861Q and G719A [23, 24], as well as R776C mutation that was previously reported to be more sensitive to erlotinib than gefitinib [21]. One patient carries doublet mutations E709K/G719A, which were reported as oncogenic drivers but maintaining the same responses to the first generation TKIs [25, 26]. 27 patients were not identified with any *EGFR* mutations. Statistical analysis indicates that the presence of T790M did not show clear association with gender or targeted drugs administered (Table 2).

### Top mutated genes in T790M- patients

In T790M absent group, other genetic mutations are potentially responsible for drug resistance. To clarify this, we generated a co-mutation plot by dividing the patients into two groups according to their T790M statuses for 11 most frequently mutated genes in T790M- patients. T790M positive (T790M+) group showed limit number of mutations in these genes other than *EGFR* and *TP53* mutations compared to T790M- group, indicating that T790M alone is the dominant resistant mechanism in this group (Figure 3). In T790M- group, *EGFR* (26 out of 53 patients, 49%) and *TP53*

**Table 1: Patients' characteristics**

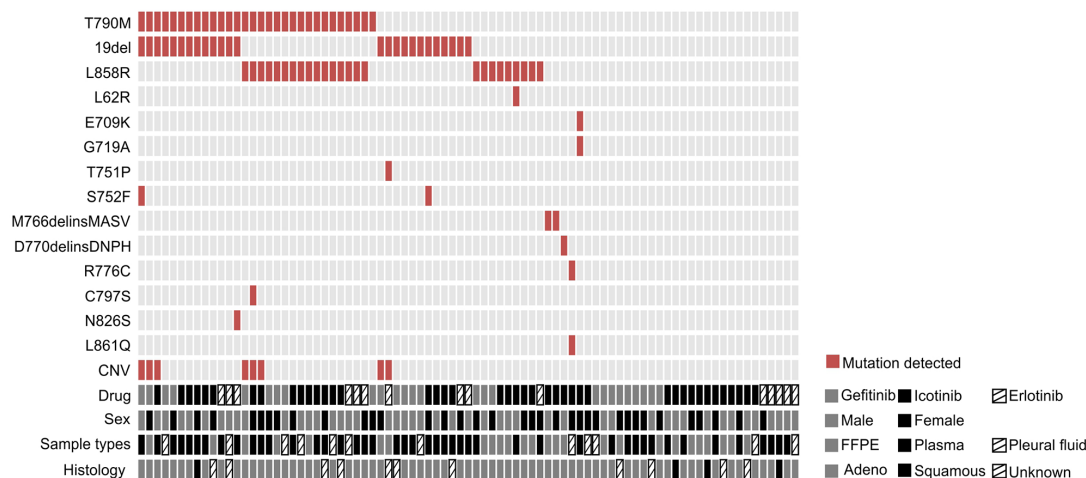
characteristic	
Sex, No. (%)	
Female	36 (43.3)
Male	47 (56.6)
Age, years	
Median	61
Range	29~85
Histology, No. (%)	
Adenocarcinoma	68 (81.9)
Squamous cell carcinoma	4 (4.8)
unknown	11 (13.3)
Sample type, No. (%)	
FFPE	26 (31.3)
Plasma	45 (54.2)
Pleural effusions	12 (14.5)
EGFR-TKI history, No. (%)	
Gefitinib	26 (31.3)
Icotinib	42 (50.6)
Erlotinib	15 (18.1)



**Figure 1: Mutation types and mutation number identified in 83 patients.** **A.** The number of mutations identified in each patient was plotted to a histogram. **B.** Total mutations detected in 83 patients were classified according to the mutation types.

(27, 50%) are still the most frequently mutated genes (Figure 3 and Figure 4). Other top mutated genes can be classified into following categories: 1) the activation and amplification of receptor tyrosine kinases (RTKs), including *ERBB2*, *MET* and *FGFR1*, and the rearrangement of *ALK* gene with *EML4*, *HERC1* and *HIP1*, respectively; 2) the activation of members of downstream RAS-RAF-ERK MAPK pathway, mainly

including *KRAS*, *NRAS*, *MAP2K1/2*; 3) the activation of PI3K-AKT/mTOR pathway, correlated to the activated mutations of *PIK3CA*, *AKT*, *TSC1/2* and amplification of *SOX2* the downstream transcriptional factor; 4) disruption of epigenetic regulators, most noticeable *TET2* and *DNMT3A*; 5) the inactivation or copy number loss of tumor suppressor genes, including *APC*, *RBI* and *PTEN* (Figure 4 and Supplementary Table S1).



**Figure 2: Comutation plot of EGFR mutations in 83 patients.** Each vertical line of blocks represents a patient. Patient features, including the drug they used, their sexes, tumor sample types that collected and histology types, were aligned below the mutation plot.

**Table 2: Comparison of the most frequently mutated genes among patients with different characteristics**

Mutated Genes	<i>EGFR</i> T790M No. (%)	<i>TET2</i> No. (%)	<i>SOX2</i> No. (%)	<i>MET</i> No. (%)	<i>KRAS</i> No. (%)	<i>ALK</i> No. (%)
<b>Sex</b>						
Female	12 (33.3)	4 (11.1)	1 (2.8)	2 (5.6)	1 (2.8)	0 (0)
Male	18 (38.3)	6 (12.8)	4 (8.5)	2 (4.3)	3 (6.4)	3 (6.4)
P value	0.818	1.000	0.382	1.000	0.629	0.254
<b>EGFR-TKI History</b>						
Gefitinib	9 (34.6)	4 (15.4)	0 (0)	4 (15.4)	1 (3.8)	0 (0)
Icotinib	15 (35.7)	6 (14.3)	5 (11.9)	0 (0)	2 (4.8)	1 (2.4)
Erlotinib	6 (40.0)	0 (0)	0 (0)	0 (0)	1 (6.7)	2 (13.3)
P value	0.910	0.358	0.109	0.015*	1.000	0.111

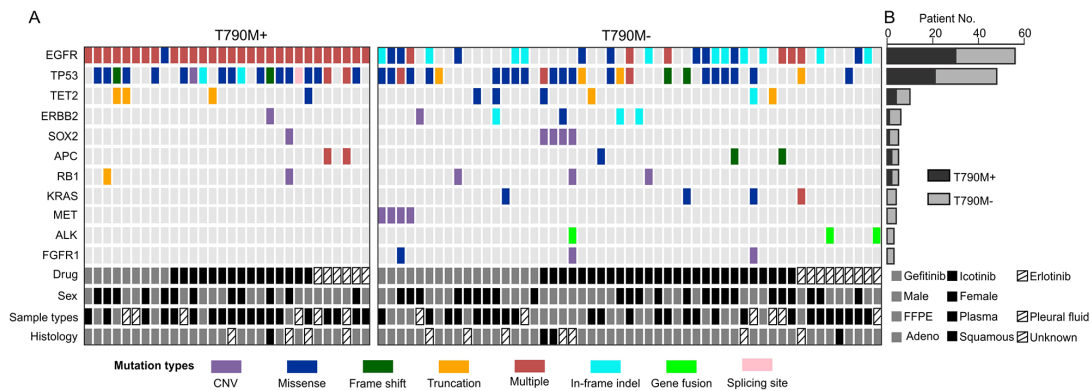
“%” indicates the percentage of each specific mutation detected in the population of defined category. The statistical differences of mutation frequency in different groups were tested by Fisher’s exact test.

It is noticeable that *MET* amplification only occurred in gefitinib treated group and statistical analysis found a significantly difference of its occurrence in three treatment groups ( $p < 0.05$ ) (Table 2). Meanwhile, *SOX2* amplification specifically presented in icotinib group and more frequently in male, but is not statistically different probably due to the small sample size (Table 2). The occurrence of *KRAS* activating point mutations is reported being mutually exclusive to *EGFR* mutations in treatment naïve patients possible due to their overlapping pathways in the onset of cancer [27, 28]. However, we observed one case that carries not only *KRAS* G13C and V14I, but also *EGFR* 19del, T751P and amplification (Figure 3A, Supplementary Table S1). It is also showed that the rearrangement of *ALK* only happens in cases without *EGFR* mutations, indicating its possible role in EGFR TKIs resistance.

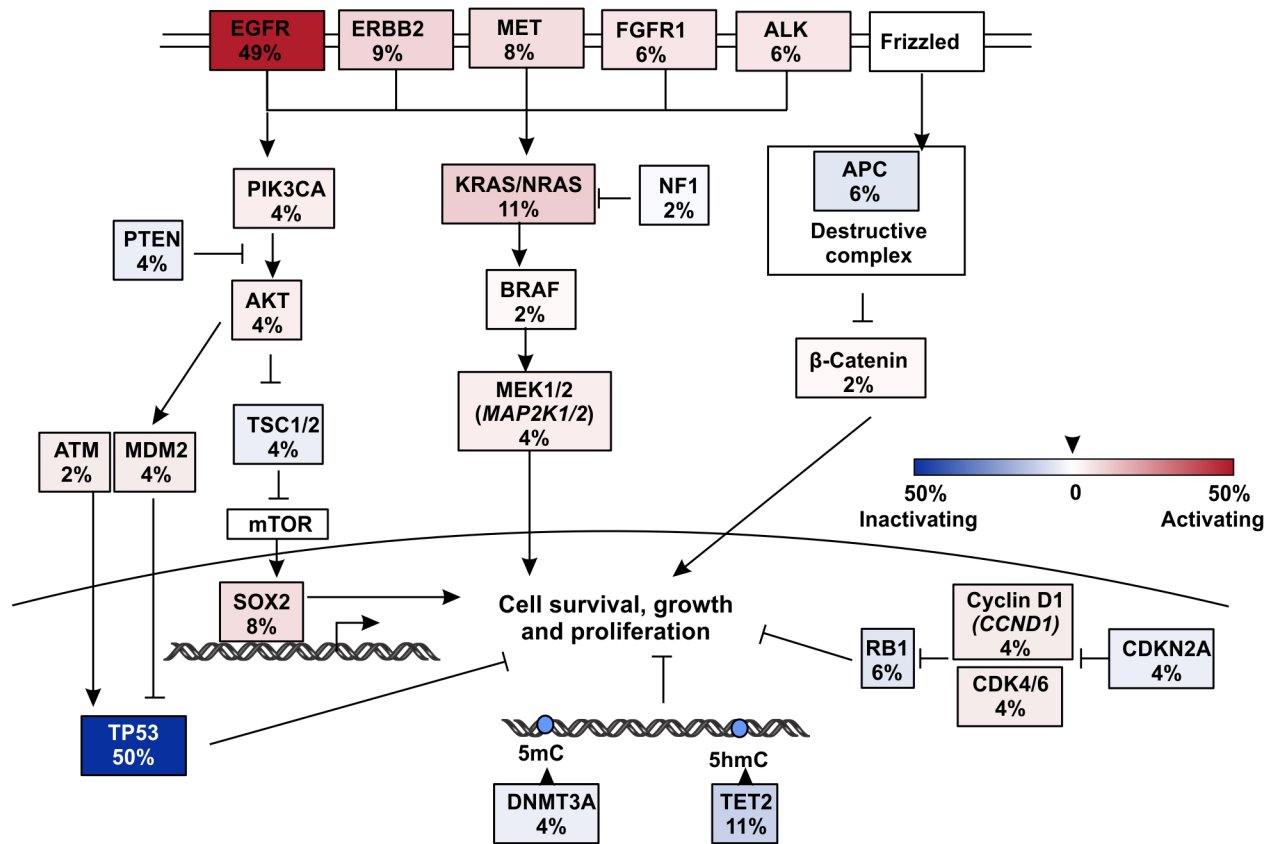
We next determined the association of most commonly mutated genes with ages of patients at the time of NGS testing. Patients with *TP53*, *TET2* and *SOX2* mutations are all significantly older than those without mutations on these genes (Supplementary Figure S1). In contrast, the presence of *EGFR* T790M is not correlated with ages.

## DISCUSSION

This study profiled genetic backgrounds of 83 patients with advanced NSCLC using a targeted pan-cancer NGS panel after they developed resistance to three different first generation EGFR TKIs. To our knowledge, it is one of the largest panels in the clinical practice that covers most up-to-date cancer-related genes. Similar to previous reports, we identified *EGFR* T790M as the most abundant secondary



**Figure 3: T790M+ and T790M- groups demonstrated different mutation spectrums.** A. Top 11 mutated genes (with at least 3 mutations identified in 83 patients) in T790M- group were selected and plotted against the T790+ group in order to compare the occurrence of different mutations between these two groups. Each vertical line of blocks represents a patient with patient features list at the bottom. Mutation types were differentiated by block colors. Multiple mutation types (red blocks) indicate that the patient have more than one mutations on the same gene. B. The number of mutated patients in T790M+ and T790M- groups was stacked for each gene. Each bar represents the gene on the left.



**Figure 4: Pathways that were influenced by mutations in EGFR TKI resistant but T790M- patients.** Somatic mutations in all 53 T790M- patients were summarized and only key mutated genes were listed.

mutation in developing drug resistance to the treatment of erlotinib, gefitinib and icotinib, and it is observed irrespective of sex, age and treatment options. The study depicts other diverse mechanisms that potentially are responsible for EGFR TKI resistance. The most frequent mechanisms in addition to EGFR T790M are the activation of other RTKs, including ERBB2, MET, FGFR1 and ALK (totally 29% in T790M-

patients), RAS/MEK/ERK pathway (totally 17%), as well as PIK3CA/AKT/mTOR pathway (totally 20%). RTKs are currently the most druggable targets. The representative drugs in market or in clinical trials are osimertinib (AZD9291) for EGFR T790M mutation [29], afatinib for ERBB2 amplification, crizotinib for ALK gene fusions and MET amplification, and AZD4547 for FGFR1 mutations.



Due to the lack of comprehensive genetic information before treatment, it is hard to conclude newly acquired resistant mechanisms solely from post-treatment results. But a few putative candidates that correlate to the resistance of the first generation EGFR TKI come to surface by comparing the genetic spectrums between *EGFR* T790M+ and T790- groups. *EGFR* T790M+ group did not encompass any mutations on genes *KRAS*, *MET*, *ALK* and *FGFR1* (Figure 3A), while all these genes have been suggested resulting in poor responses to EGFR TKIs in NSCLC patients [30-32]. Because both *MET* and *EGFR* reside on chromosome 7, a previous study suggested that polysomy of entire chromosome would result in the *in cis* amplification of *MET* and *EGFR* genes and it is hard to distinguish them by FISH [33]. Here in our study, 3 out of 4 cases with *MET* amplification were not escorted with *EGFR* amplification, and 1 case shows significantly higher amplification of *MET* compared to *EGFR*. Therefore, *MET* is more likely acting as a bypass signaling pathway that exerts resistance to EGFR TKIs. Moreover, it shows statistically higher occurrence in gefitinib-treated group rather than erlotinib and icotinib groups. Further studies with larger sample sizes might be needed in order to elucidate this finding.

ICotinib is the second line first generation EGFR TKI that was approved by China Food and Drug Administration (CFDA) in 2011, and because of its compatible efficacy and side effects as gefitinib [34], as well as its lower price, it is widely used in treating NSCLC in China. As far as we know, our study was the first to examine the genetic spectrums of icotinib-resistant NSCLC patients compared to other first generation EGFR TKIs, and found that *SOX2* amplification were private to icotinib treatment resistance. *SOX2* is frequently amplified at a percentage of around 20% in small cell lung cancer (SCLC) and NSCLC [35, 36]. It functions as a transcription factor that inhibits cell differentiation [37] and promotes cell cycling [38]. In addition, it is recognized as a downstream molecule of AKT/mTOR pathway that exerts controlling over cell survival [39].

There are several limitations in this study: firstly, since this was a retrospective analysis and most of patients were outpatients, we were unable to collect all the patients' subsequent treatment strategies and outcomes; Secondly, since there was no comparison with the pre-treatment genetic background, it was not certain whether these alterations have emerged as a result of acquired resistance, or they have been there since the beginning of the treatment; Thirdly, the relatively limited sample size of this study could decrease the statistical power in analysis. Thus, we are ready to conduct a prospective research using NGS technology to compare the pre-treatment and acquired resistance genetic profiles to further uncover the mechanism of resistance to EGFR-TKIs.

In summary, the study depicted the genetic landscapes comprehensively in Chinese NSCLC

population resistant to first generation TKI treatments including icotinib. Our analysis demonstrates new perspectives for further study of resistance and putting forward corresponding relevant tactics against the challenge of disease progression.

## MATERIALS AND METHODS

### Patient enrollment and sample preparation

Between Jan 2015 to Dec 2015, a total of 83 patients with stage IV NSCLC in Zhejiang Cancer Hospital, China, were undergoing tumor biopsies or blood withdrawing by the time of acquiring resistance to the first generation EGFR TKIs, including 26 formalin-fixed paraffin-embedded (FFPE) samples, 45 serum samples and 12 serous effusions. Acquired resistance to EGFR TKIs was evaluated by "Jackman criteria" in each patient [40]. For FFPE samples, only samples harbored tumor cell content above 20% were considered qualified and included. Written consents from all patients were collected according to the ethic regulations of Zhejiang Cancer Hospital. Collected samples were sent to the core facility of Nanjing Shihe Jiyin Biotechnology Inc. (Nanjing, China) for targeted NGS analysis.

### DNA extraction

5-8 of 10µm tissue sections from tumor FFPE samples were used for genomic DNA extraction with QIAamp DNA FFPE Tissue Kit (QIAGEN) following the manufacturer's instructions. Plasma was extracted from 5-10 ml peripheral blood in EDTA-coated tubes within 2 hours of blood withdrawing, and circulating cell free DNA (cfDNA) was extracted using the QIAamp Circulating Nucleic Acid Kit (QIAGEN). Genomic DNA of cellular sediments of pleural effusions and whole blood samples were prepared with DNeasy Blood & Tissue kit (QIAGEN). Whole blood DNA was sequenced together with tumor or cfDNA samples for the purpose of identifying germline mutations. The DNA quality was assessed by Nanodrop2000 (Thermo Fisher Scientific) and the quantity was measured by dsDNA HS Assay Kit (Life Technologies) on Qubit 2.0.

### Library preparation and sequencing

Extracted tumor genomic DNA was fragmented into 300~350bp using Covaris M220 instrument (Covaris). Sequencing libraries were prepared with KAPA Hyper Prep kit (KAPA Biosystems) with optimized protocols. In brief, cfDNA or sheared tissue DNA were experienced with end-repairing, A-tailing, adapter ligation and size selection using Agencourt AMPure XP beads (Beckman Coulter). Libraries were then subjected to PCR amplification and purification before targeted enrichment.

DNA libraries from different samples were marked with unique indices during library preparation and up to 2 µg of different libraries were pooled together for targeted enrichment. Human cot-1 DNA (Life Technologies) and xGen Universal blocking oligos (Integrated DNA Technologies) were added to block nonspecific binding of library DNA to targeted probes. Customized xGen lockdown probes panel (Integrated DNA Technologies) were used to targeted enrich for 416 predefined genes. The hybridization reaction was performed by using NimbleGen SeqCap EZ Hybridization and Wash Kit (Roche). Dynabeads M-270 (Life Technologies) was used to capture probe-bind fragments, followed by library amplification with Illumina p5 (5' AAT GAT ACG GCG ACC ACC GA 3') and p7 primers (5' CAA GCA GAA GAC GGC ATA CGA GAT 3') in KAPA HiFi HotStart ReadyMix (KAPA Biosystems), and purification by Agencourt AMPure XP beads. Library quantification was analyzed by KAPA Library Quantification kit (KAPA Biosystems). The size distribution of libraries was measured by Agilent Technologies 2100 Bioanalyzer (Agilent Technologies). The enriched libraries were sequenced on Hiseq 4000 NGS platforms (Illumina) to coverage depths of at least 100x, 300x, 3000x after removing PCR duplicates for blood, FFPE/pleural effusion, and ctDNA, respectively.

### Annotation and interpretation of sequencing results

Trimmomatic [41] was used for sequencing data quality control. The reads with a quality below the threshold of 15, as well as N bases were removed before mapping to reference sequence hg19 (Human Genome version 19) using Burrows-Wheeler Aligner (BWA) [42] with optimized parameters. Genome Analysis Toolkit (GATK) [43] was used for indels local realignment and base quality score recalibration. SNPs/indels were identified using VarScan2 (MAF<10%) (<http://dkoboldt.github.io/varscan/>) and HaplotypeCaller/UnifiedGenotyper in GATK (MAF>10%). SNPs were filtered out with dbSNP and 1000 Genome data sets. Germline mutations in tumor tissues or ctDNA were identified by comparing to the matched whole blood DNA. A mutation was called when at least 3 mutated reads were found in the sample on different strands with good quality scores and manually inspected in Integrative Genomics Viewer (IGV, Broad Institute). Genomic fusions were identified by FACTERA [44] with default parameters. Copy number variations (CNVs) were detected using ADTEx (<http://adtex.sourceforge.net>) with default parameters. Proposed discrete wavelet transform (DWT) was used to reduce intrinsic noise.

### CONFLICTS OF INTEREST

All authors declare no conflict of interest.

### REFERENCES

1. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, Jemal A, Yu XQ and He J. Cancer statistics in China, 2015. *CA Cancer J Clin.* 2016; 66:115-132.
2. Strand TE, Rostad H, Moller B and Norstein J. Survival after resection for primary lung cancer: a population based study of 3211 resected patients. *Thorax.* 2006; 61:710-715.
3. Cancer Genome Atlas Research N. Comprehensive molecular profiling of lung adenocarcinoma. *Nature.* 2014; 511:543-550.
4. Wagle N, Berger MF, Davis MJ, Blumenstiel B, Defelice M, Pochanard P, Ducar M, Van Hummelen P, Macconail LE, Hahn WC, Meyerson M, Gabriel SB and Garraway LA. High-throughput detection of actionable genomic alterations in clinical tumor samples by targeted, massively parallel sequencing. *Cancer discovery.* 2012; 2:82-93.
5. Lawrence MS, Stojanov P, Mermel CH, Robinson JT, Garraway LA, Golub TR, Meyerson M, Gabriel SB, Lander ES and Getz G. Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature.* 2014; 505:495-501.
6. Boolell V, Alamgeer M, Watkins DN and Ganju V. The Evolution of Therapies in Non-Small Cell Lung Cancer. *Cancers.* 2015; 7:1815-1846.
7. Li S, Li L, Zhu Y, Huang C, Qin Y, Liu H, Ren-Heidenreich L, Shi B, Ren H, Chu X, Kang J, Wang W, Xu J, Tang K, Yang H, Zheng Y, et al. Coexistence of EGFR with KRAS, or BRAF, or PIK3CA somatic mutations in lung cancer: a comprehensive mutation profiling from 5125 Chinese cohorts. *British journal of cancer.* 2014; 110:2812-2820.
8. Shi Y, Au JS, Thongprasert S, Srinivasan S, Tsai CM, Khoa MT, Heeroma K, Itoh Y, Cornelio G and Yang PC. A prospective, molecular epidemiology study of EGFR mutations in Asian patients with advanced non-small-cell lung cancer of adenocarcinoma histology (PIONEER). *Journal of thoracic oncology.* 2014; 9:154-162.
9. Mok TS, Wu YL, Thongprasert S, Yang CH, Chu DT, Saijo N, Sunpaweravong P, Han B, Margono B, Ichinose Y, Nishiwaki Y, Ohe Y, Yang JJ, Chewaskulyong B, Jiang H, Duffield EL, et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *The New England journal of medicine.* 2009; 361:947-957.
10. Rosell R, Carcereny E, Gervais R, Vergnenegre A, Massuti B, Felip E, Palmero R, Garcia-Gomez R, Pallares C, Sanchez JM, Porta R, Cobo M, Garrido P, Longo F, Moran T, Insa A, et al. Erlotinib versus standard chemotherapy as first-line treatment for European patients with advanced EGFR mutation-positive non-small-cell lung cancer (EURTAC): a multicentre, open-label, randomised phase 3 trial. *The Lancet Oncology.* 2012; 13:239-246.
11. Ge L and Shi R. Progress of EGFR-TKI and ALK/ROS1 inhibitors in advanced non-small cell lung cancer. *International journal of clinical and experimental medicine.* 2015; 8:10330-10339.

12. Russo A, Franchina T, Ricciardi GR, Picone A, Ferraro G, Zanghi M, Toscano G, Giordano A and Adamo V. A decade of EGFR inhibition in EGFR-mutated non small cell lung cancer (NSCLC): Old successes and future perspectives. *Oncotarget*. 2015; 6:26814-26825. doi: 10.18632/oncotarget.4254.
13. Balak MN, Gong Y, Riely GJ, Somwar R, Li AR, Zakowski MF, Chiang A, Yang G, Ouerfelli O, Kris MG, Ladanyi M, Miller VA and Pao W. Novel D761Y and common secondary T790M mutations in epidermal growth factor receptor-mutant lung adenocarcinomas with acquired resistance to kinase inhibitors. *Clinical cancer research*. 2006; 12:6494-6501.
14. Correction: PTEN Loss Contributes to Erlotinib Resistance in EGFR-Mutant Lung Cancer by Activation of Akt and EGFR. *Cancer research*. 2015; 75:1922.
15. Becker K and Xu Y. Management of tyrosine kinase inhibitor resistance in lung cancer with EGFR mutation. *World journal of clinical oncology*. 2014; 5:560-567.
16. Remon J, Moran T, Majem M, Reguart N, Dalmau E, Marquez-Medina D and Lianes P. Acquired resistance to epidermal growth factor receptor tyrosine kinase inhibitors in EGFR-mutant non-small cell lung cancer: a new era begins. *Cancer treatment reviews*. 2014; 40:93-101.
17. Yun CH, Mengwasser KE, Toms AV, Woo MS, Greulich H, Wong KK, Meyerson M and Eck MJ. The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. *Proceedings of the National Academy of Sciences of the United States of America*. 2008; 105:2070-2075.
18. Buermans HP and den Dunnen JT. Next generation sequencing technology: Advances and applications. *Biochim Biophys Acta*. 2014; 1842:1932-1941.
19. Zhang C, Zhang H, Shi J, Wang D, Zhang X, Yang J, Zhai Q and Ma A. Trial-Based Cost-Utility Analysis of Icotinib versus Gefitinib as Second-Line Therapy for Advanced Non-Small Cell Lung Cancer in China. *PloS one*. 2016; 11:e0151846.
20. Thress KS, Paweletz CP, Felip E, Cho BC, Stetson D, Dougherty B, Lai Z, Markovets A, Vivancos A, Kuang Y, Ercan D, Matthews SE, Cantarini M, Barrett JC, Janne PA and Oxnard GR. Acquired EGFR C797S mutation mediates resistance to AZD9291 in non-small cell lung cancer harboring EGFR T790M. *Nat Med*. 2015; 21:560-562.
21. Kancha RK, von Bubnoff N, Peschel C and Duyster J. Functional analysis of epidermal growth factor receptor (EGFR) mutations and potential implications for EGFR targeted therapy. *Clinical cancer research*. 2009; 15:460-467.
22. Chou TY, Chiu CH, Li LH, Hsiao CY, Tzen CY, Chang KT, Chen YM, Perng RP, Tsai SF and Tsai CM. Mutation in the tyrosine kinase domain of epidermal growth factor receptor is a predictive and prognostic factor for gefitinib treatment in patients with non-small cell lung cancer. *Clinical cancer research*. 2005; 11:3750-3757.
23. Watanabe S, Minegishi Y, Yoshizawa H, Maemondo M, Inoue A, Sugawara S, Isobe H, Harada M, Ishii Y, Gemma A, Hagiwara K and Kobayashi K. Effectiveness of gefitinib against non-small-cell lung cancer with the uncommon EGFR mutations G719X and L861Q. *Journal of thoracic oncology*. 2014; 9:189-194.
24. Yasuda H, Park E, Yun CH, Sng NJ, Lucena-Araujo AR, Yeo WL, Huberman MS, Cohen DW, Nakayama S, Ishioka K, Yamaguchi N, Hanna M, Oxnard GR, Lathan CS, Moran T, Sequist LV, et al. Structural, biochemical, and clinical characterization of epidermal growth factor receptor (EGFR) exon 20 insertion mutations in lung cancer. *Science translational medicine*. 2013; 5:216ra177.
25. Kobayashi Y, Togashi Y, Yatabe Y, Mizuuchi H, Jangchul P, Kondo C, Shimoji M, Sato K, Suda K, Tomizawa K, Takemoto T, Hida T, Nishio K and Mitsudomi T. EGFR Exon 18 Mutations in Lung Cancer: Molecular Predictors of Augmented Sensitivity to Afatinib or Neratinib as Compared with First- or Third-Generation TKIs. *Clinical cancer research*. 2015; 21:5305-5313.
26. Chen Z, Feng J, Saldivar JS, Gu D, Bockholt A and Sommer SS. EGFR somatic doublets in lung cancer are frequent and generally arise from a pair of driver mutations uncommonly seen as singlet mutations: one-third of doublets occur at five pairs of amino acids. *Oncogene*. 2008; 27:4336-4343.
27. Shigematsu H, Lin L, Takahashi T, Nomura M, Suzuki M, Wistuba II, Fong KM, Lee H, Toyooka S, Shimizu N, Fujisawa T, Feng Z, Roth JA, Herz J, Minna JD and Gazdar AF. Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. *Journal of the National Cancer Institute*. 2005; 97:339-346.
28. Marchetti A, Martella C, Felicioni L, Barassi F, Salvatore S, Chella A, Camplese PP, Iarussi T, Mucilli F, Mezzetti A, Cuccurullo F, Sacco R and Buttitta F. EGFR mutations in non-small-cell lung cancer: analysis of a large series of cases and development of a rapid and sensitive method for diagnostic screening with potential implications on pharmacologic treatment. *Journal of clinical oncology*. 2005; 23:857-865.
29. Wang S, Cang S and Liu D. Third-generation inhibitors targeting EGFR T790M mutation in advanced non-small cell lung cancer. *Journal of hematology & oncology*. 2016; 9:34.
30. Pao W, Wang TY, Riely GJ, Miller VA, Pan Q, Ladanyi M, Zakowski MF, Heelan RT, Kris MG and Varmus HE. KRAS mutations and primary resistance of lung adenocarcinomas to gefitinib or erlotinib. *PLoS medicine*. 2005; 2:e17.
31. Ware KE, Hinz TK, Kleczko E, Singleton KR, Marek LA, Helfrich BA, Cummings CT, Graham DK, Astling D, Tan AC and Heasley LE. A mechanism of resistance to gefitinib mediated by cellular reprogramming and the acquisition of an FGF2-FGFR1 autocrine growth loop. *Oncogenesis*. 2013; 2:e39.



32. Sequist LV, Waltman BA, Dias-Santagata D, Digumarthy S, Turke AB, Fidias P, Bergethon K, Shaw AT, Gettinger S, Cospoer AK, Akhavanfard S, Heist RS, Temel J, Christensen JG, Wain JC, Lynch TJ, et al. Genotypic and histological evolution of lung cancers acquiring resistance to EGFR inhibitors. *Science translational medicine*. 2011; 3:75ra26.
33. Cappuzzo F, Hirsch FR, Rossi E, Bartolini S, Ceresoli GL, Bemis L, Haney J, Witta S, Danenberg K, Domenichini I, Ludovini V, Magrini E, Gregorc V, Doglioni C, Sidoni A, Tonato M, et al. Epidermal growth factor receptor gene and protein and gefitinib sensitivity in non-small-cell lung cancer. *J Natl Cancer Inst*. 2005; 97:643-655.
34. Shi Y, Zhang L, Liu X, Zhou C, Zhang L, Zhang S, Wang D, Li Q, Qin S, Hu C, Zhang Y, Chen J, Cheng Y, Feng J, Zhang H, Song Y, et al. Icotinib versus gefitinib in previously treated advanced non-small-cell lung cancer (ICOGEN): a randomised, double-blind phase 3 non-inferiority trial. *The Lancet Oncology*. 2013; 14:953-961.
35. Karachaliou N, Rosell R and Viteri S. The role of SOX2 in small cell lung cancer, lung adenocarcinoma and squamous cell carcinoma of the lung. *Translational lung cancer research*. 2013; 2:172-179.
36. Rudin CM, Durinck S, Stawiski EW, Poirier JT, Modrusan Z, Shames DS, Bergbower EA, Guan Y, Shin J, Guillory J, Rivers CS, Foo CK, Bhatt D, Stinson J, Gnad F, Haverty PM, et al. Comprehensive genomic analysis identifies SOX2 as a frequently amplified gene in small-cell lung cancer. *Nature genetics*. 2012; 44:1111-1116.
37. Okumura-Nakanishi S, Saito M, Niwa H and Ishikawa F. Oct-3/4 and Sox2 regulate Oct-3/4 gene in embryonic stem cells. *The Journal of biological chemistry*. 2005; 280:5307-5317.
38. Chen Y, Shi L, Zhang L, Li R, Liang J, Yu W, Sun L, Yang X, Wang Y, Zhang Y and Shang Y. The molecular mechanism governing the oncogenic potential of SOX2 in breast cancer. *The Journal of biological chemistry*. 2008; 283:17969-17978.
39. Singh S, Trevino J, Bora-Singhal N, Coppola D, Haura E, Altiook S and Chellappan SP. EGFR/Src/Akt signaling modulates Sox2 expression and self-renewal of stem-like side-population cells in non-small cell lung cancer. *Molecular cancer*. 2012; 11:73.
40. Jackman D, Pao W, Riely GJ, Engelman JA, Kris MG, Janne PA, Lynch T, Johnson BE and Miller VA. Clinical definition of acquired resistance to epidermal growth factor receptor tyrosine kinase inhibitors in non-small-cell lung cancer. *Journal of clinical oncology*. 2010; 28:357-360.
41. Bolger AM, Lohse M and Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014; 30:2114-2120.
42. Li H and Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009; 25:1754-1760.
43. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernysky AM, Sivachenko AY, Cibulskis K, Gabriel SB, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature genetics*. 2011; 43:491-498.
44. Newman AM, Bratman SV, Stehr H, Lee LJ, Liu CL, Diehn M and Alizadeh AA. FACTERA: a practical method for the discovery of genomic rearrangements at breakpoint resolution. *Bioinformatics*. 2014; 30:3390-3393.