Clinics in Oncology

Research Article
Published: 29 Jan, 2024

Targeted Therapy Directed Genes' Genomic Profiling in Circulating Tumor DNA of Lung Cancer Patients

Baia L[#], Songa J[#], Zhaia J, Lia J, Lua X, Zhoua Y and Zhoua J**

Department of Laboratory Medicine, West China Hospital, Sichuan University, China

[#]These authors contributed equally to this work

Abstract

Background: Circulating tumor DNA (ctDNA) is an easily accessible source to detect actionable gene alterations and discover drug-resistant mutations in lung cancer. This study investigated the ctDNA genomic profile for lung cancer patients to provide valuable information for clinical decision.

Methods: Six hundred and seventy-two lung cancer patients were enrolled in this retrospective study, and peripheral blood from each patient was collected for next generation sequencing based on Illumina NextSeq 550 platform, using a 23-gene targeted panel.

Results: About 73% of patients were detected at least one genetic alteration in plasma ctDNA, and 45.54% of patients obtained drug-sensitive/drug-resistance information. The most frequently mutated gene was *EGFR* (42.41%), *EGFR* mutation was associated with female, non-smokers and adenocarcinoma, while *KRAS* mutation tended to occur in males and smokers. Patients received targeted therapy before had higher frequency of *EGFR* mutation but lower mutation rate of *KRAS*, PTEN, KIT and FGFR2. Brain metastasis patients have a relatively lower positive rate (69.23%). Patients with *EGFR* mutations detected in ctDNA tended to have bone metastasis, while PIK3CA and *KRAS* mutations were enriched in liver metastasis, and FGFR1, HER2 for adrenal metastasis.

Conclusion: This study showed that mutation pattern and positive rate in plasma ctDNA were closely associated with smoking, gender, pathological type, treatment and metastasis. Some gene alterations might have organotropism preference about metastatic sites.

Keywords: Lung cancer; Circulating tumor DNA; Metastasis; Targeted therapy

OPEN ACCESS

Correspondence:

Yi Zhoua, Department of Laboratory Medicine, West China Hospital, Sichuan University, No. 37 Guoxue Alley, Chengdu, Sichuan 610041, China

Juan Zhoua, Department of Laboratory Medicine, West China Hospital, Sichuan University, No. 37, Guoxue Alley, Chengdu, Sichuan, 610041, China

Received Date: *26 Dec 2023*

Accepted Date: *22 Jan 2024*

Published Date: *29 Jan 2024*

Citation:

Baia L, Songa J, Zhaia J, Lia J, Lua X, Zhoua Y, et al. Targeted Therapy Directed Genes' Genomic Profiling in Circulating Tumor DNA of Lung Cancer Patients. Clin Oncol. 2024; 9: 2051.

ISSN: 2474-1663

Introduction

Lung cancer have made great progress on targeted therapy in recent years [1,2], although it is still the worldwide leading cause of cancer death [3-5], so dose in China [5]. The latest data from International Agency for Research on Cancer (IARC) shows that lung cancer accounts for 18.0% of the total cancer deaths in year 2020, almost twice as much as the second - colorectal cancer [6]. Hence, there is still a long way for lung cancer to go through even at the help of molecular-targeted agents. Since the first molecular targeted drug gefitinib was approved for lung cancer by US Food and Drug Administration (FDA) in 2002, there are already more than twenty targeted agents approved such as *EGFR* Tyrosine Kinase Inhibitor (TKI) erlotinib [7] and Osimertinib [8], ALK inhibitor Crizotinib [9], etc. Furthermore, many drugs against oncogenic proteins are now in clinical trial [1].

With the wide application of targeted therapy, specific genetic testing is necessary to guide decision-making of therapeutic regimen and enable patients receiving effective precise treatment [10]. Tumor tissue biopsy is still the gold standard source for molecular profiling, but it is unfeasible in about 30% of patients for different reasons [11], like wide spread tumor metastasis, inadequate tumor tissue, etc. By contrast, circulating tumor DNA (ctDNA) in peripheral blood is a valuable alternative to detect clinical actionable gene alterations owing to the minimal invasive operation. And ctDNA can better represent tumor heterogeneity than a scant tissue sample. A good consistency of oncogenic driver mutation between tissue and ctDNA was observed [12,13]. Furthermore, it's reported that ctDNA genetic analysis has a potential to enhance the positive detection rate remarkably [12,14,15]. With the advance of sensitivity and specificity, ctDNA genetic testing gradually becomes a routine clinical detection for patients with advanced stage lung cancer.

Currently, ctDNA genetic detection is relatively expensive; hence its cost and effect are particularly important. However, different positive rates were reported in several studies, varying from 70% to 90% [15-18], which might be caused by detection platform or population heterogeneity. In addition to overall positive rate, the actionable mutation rate of ctDNA detection, drug-resistant variant and performance difference under different metastasis condition also matter a lot when making clinical decisions. Moreover, the gene mutation profile of ctDNA, gene mutation's association with prognosis, and the pattern between uncommon alterations and drug resistance, still need further exploration for more real-world data and pharmaceutical development. Herein, in order to be assistant in clinical work, we investigated the ctDNA genomic profile in 672 lung cancer patients from southwest China with a 23-gene Next-Generation Sequencing (NGS) panel designed for guiding targeted therapy.

Material and Methods

Samples

From February 2017 to November 2019, 672 patients diagnosed with lung cancer at West China Hospital were enrolled in this retrospective study. Before the ctDNA genetic testing, we explained all aspects of the informed consent to patients and collected clinical characteristics at the same time. Each patient was drawn 20 ml of peripheral blood and stored in Streck tubes for testing. The cell-free DNA was extracted by magnetic bead method, and samples whose DNA content was no less than 40 ng and DNA concentration at least reached up to 0.42 ng/μl were qualified.

Next generation targeted sequencing

According to the manufacturer's protocol, cell-free DNA was enriched by hybrid-capture method and sequencing library was constructed with NovoPMTM Library Prep Kit for Solid Tumor Target Therapy (Novogene, Beijing, China). Then we performed Pair-End (PE150) sequencing on Illumina NextSeq 550 sequencer (Illumina, San Diego, CA) with a sequencing depth at an average level of 2,000 reads after removing duplicates, which covered all exons of the following 23 genes: *AKT1*, *BRAF*, *DDR2*, *EGFR*, *GNA11*, *GNAQ*, *HRAS*, *KIT*, *KRAS*, *MEK1*, *HER2*, *PDGFRA*, *PIK3CA*, *PTEN*, *SMO*, *TSC1*, *FGFR1*, *FGFR2*, *NRAS*, *MET*, *ALK*, *RET*, *ROS1*, as well as the intronic regions of *ALK*, *RET* and *ROS1* for fusion detection (Supplementary Table S1).

Bioinformatic data analysis

An in-house developed code was used for analyzing genomic alterations. Raw data was mapped to GRCh37 reference genome, and database 1000g2015aug_all and gnomAD were applied to filter germline mutation. The detection limit of genomic alteration was 0.5%, including Single Nucleotide Variant (SNV), Insertion and/or Deletion (InDel) and fusion, while Copy Number (CN) more than 2.7 was the cut-off value of copy number gain. But for *EGFR* 19DEL/L858R/T790M, when the three frequent and important *EGFR* mutations' variant frequency was less than 0.5% by NGS detection, a validation experiment was performed by Droplet Digital PCR out of clinical treatment consideration. To produce final reports for patients, we used NovoDB database to annotate medicine information for each variant. Cancer fraction was defined as the ratio of indicated Variant Allele Frequency (VAF) to the maximum VAF present in that sample. Cancer fraction was used for measuring mutation clonality, where mutations with cancer fraction no less than 0.5 were considered as dominant, otherwise subclonal.

Droplet digital PCR (ddPCR)

The extracted cell-free DNA was added to ddPCR super mix with primers and probes for pre-amplification according to *EGFR* T790M/L858R/19DEL assay kit manual from Shanghai Yuanqi Bio. The following amplification procedure was performed to prepare PCR-ready samples: 42°C, 5 min; 94°C, 5 min; (94°C, 15 s; 60°C, 25 s; 72°C, 40 s) 8 recycles; 72°C, 5 min. Then the PCR-ready samples were loaded to individual well of disposable droplet generator cartridge, and Bio-Rad QX200 Droplet Generator was used for generating droplets, which were transferred to a 96-well plate for PCR amplification in Bio-Rad thermal cycler. The PCR was performed following these steps: 95°C, 10 min; (94°C, 15 s; 58°C, 60 s) 40 recycles; 98°C, 10 min; 4°C, 5 min; set temperature change speed ≤ 2°C/s. Finally, droplet reading, and result analysis were conducted on Bio-Rad QX200 Droplet Reader with its bult-in software QuantaSoft.

Statistical analysis and visualization

All hypotheses testing about gene alterations' difference between different groups were used Chi-square test or Fisher exact test according to data requirements. Wilcoxon test was applied to variant allele frequency or cancer fraction comparison between two groups. Gene co-occurrence/mutual exclusivity analysis was performed by DISCOVER [19] method. An alteration was significantly associated with specific target organs if meeting the following two conditions: (a) the Chi-square test or Fisher exact test for categorical variables was statistically significantly; (b) the coefficient associated with gene alteration in multivariable logistical regression model adjusted for metastatic burden was statistically significant. P-value less than 0.05 were considered as statistically significantly through the whole work. Data cleaning, statistical analysis and figures were mainly generated in R with R packages tidyverse [20], readxl [21], ggplot2 [22], Complex Heatmap [23], Venn diagram [24], corrplot [25], R Color Brewer [26]. The protein mutation landscape of *EGFR* gene was created by ProteinPaint tool [27].

Results and Discussion

The characteristics of patients

Six hundred and seven two samples from lung cancer patients (319 males and 353 females) were analyzed in this study (Table 1, Table S2), with an average age at 63.5 (ranging from 28 to 90). The percentage of patients without smoking history in men and women is 54.4% and 46.1%, respectively. And the majority samples (80.51%, 541/672) were pathologically diagnosed as Non-Small Cell Lung Cancer (NSCLC), which consisted of adenocarcinoma (73.07%, 491/672), squamous cell carcinoma (4.61%, 31/672), adenosquamous carcinoma (0.89%, 6/672) and large cell carcinoma (0.15%, 1/672), except for subclass information loss of 12 NCSLC samples. Since most early-stage lung cancer patients can obtain tumor tissue and carried a relatively low level of ctDNA, ctDNA genomic test mostly serves patients with metastasis or unresectable tumor sites in order to see if qualified for targeted agents. In our study, 468 out of 672 patients (69.64%) were at stage IV, 32 patients (4.76%) at stage III and few at stage II/I (3 and 12 patients, respectively). Moreover, bone, brain, liver and adrenal, as the most common metastatic sites [28,29], were present in the 84.42% of our stage IV patients with metastatic sites information available. And it's quite common to see multiple metastatic sites occurred in the same patient, usually more than 6 different sites, which were classified as "widespread metastasis" in

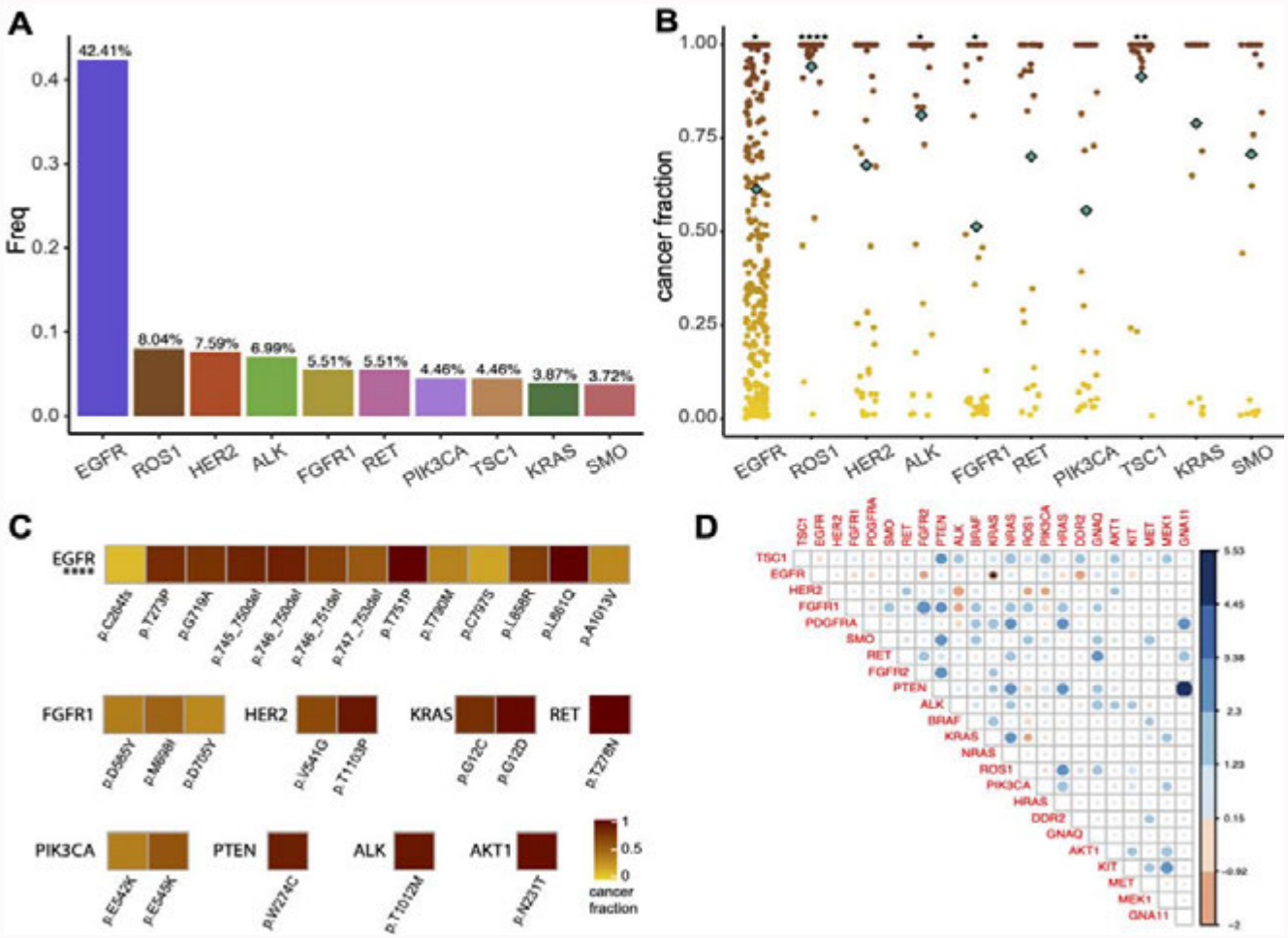

Figure 1: ctDNA gene alteration profile and its association with prognosis. (A) Genes on the x-axis are arranged in a descending order of frequency. The number above each bar is the exact mutation frequency of corresponding gene. (B) Cancer fraction of mutations in top 10 frequently mutated genes. Green diamond indicates mean value of each gene. Asterisks represents a significant difference between corresponding gene and the rest cases (*p<0.05, **p<0.01, ****p<0.0001). (C) Hotspot mutations' cancer fraction in indicated genes. Cancer fraction's significant variation across mutations in gene is marked with asterisks (*EGFR*, ****p<0.0001). (D) Co-occurrence or exclusivity analysis. The dot plot indicates the odds ratio of co-occurrence (blue) or exclusivity (red) using color-coding and circle size, and dot with asterisk is suggesting FDR-corrected statistical significance.

Table 1. Besides, 26.93% patients had underwent targeted therapy before ctDNA genetic detection.

ctDNA gene alterations' profile in lung cancer patients

In total, 493 patients (493/672, 73.36%) carried no less than one detectable ctDNA alteration in their peripheral blood, including SNV, InDel, gene fusion and amplification (Table S3). And 45.5% (306/672) patients carried actionable or drug-resistant mutations with a median variant allele frequency at 2.84%, hence obtained information about drug-sensitive or drug-resistance to guide clinical therapy. In terms of patients with detectable alterations, each patient had about 2 genomic alterations on average and the alteration number ranged from 1 to 7 in terms of individual (Supplementary Figure 1, 2). More than a half had at least two alterations at the same time and about 80.24% were missense mutations. Among all 672 patients, the top 10 altered genes in order were *EGFR* (42.41%), *ROS1* (8.04%), *HER2* (7.59%), *ALK* (6.99%), *FGFR1* (5.51%), *RET* (5.51%), *PIK3CA* (4.46%), *TSC1* (4.46%), *KRAS* (3.87%), *SMO* (3.72 %) (Figure 1A). Then followed by *PDGFRA* (3.57%), *PTEN* (3.57%), *AKT1* (2.83%), *BRAF* (2.83%), *MET* (2.53%), *DDR2* (2.38%), *KIT* (1.93%), *FGFR2* (1.79%), *GNAQ* (1.34%), *MEK1* (1.34%), *HRAS* (0.89%), *NRAS* (0.89%), *GNA11* (0.74%). The *EGFR* positive rate (42.41%) was close to the previous reported rate in Chinese patients [17]. Besides, cancer fractions were calculated to investigate clonal dominance of top 10 frequently mutated genes. Mutations observed in gene *EGFR*, *ROS1*, *ALK* and *TSC1* were more likely to be dominant, while *FGFR1* alterations seemed to be subclonal (Figure 1B). Most hotspot mutations (recurrent number >3) in this study were dominant clone, and the clonal dominance of *EGFR* hotspot mutations showed a significant variation (Figure 1C). And we found that *EGFR* and *KRAS* mutated exclusively in ctDNA genetic detection (p<0.01, Figure 1D). However, *KRAS* mutation rate was particularly low (3.87%) when compared with some other similar studies (12%~21%) [15-17], where *KRAS* was always the second or third most mutated gene. *KRAS* mutations were detected in 26 patients and many of them located in exon 2 (n=23), which included p.G12A (1), p.G12C (6), p.G12D (7), p.G12F (2), p.G12S (3), p.G12V (2), p.G13C (1) and p.G13D (1). The rest 3 mutations were all p.Q61H in exon 3. All the above were hot mutations and reported to constitute nearly 97% of *KRAS* mutation in NSCLC [30].

In addition to point mutation, there were 8 *RET* fusions and 7 *ALK* fusions. All *RET* rearrangements happened in exon 13 fused with exon 5 of LINC00486. As reported before [31], we also observed

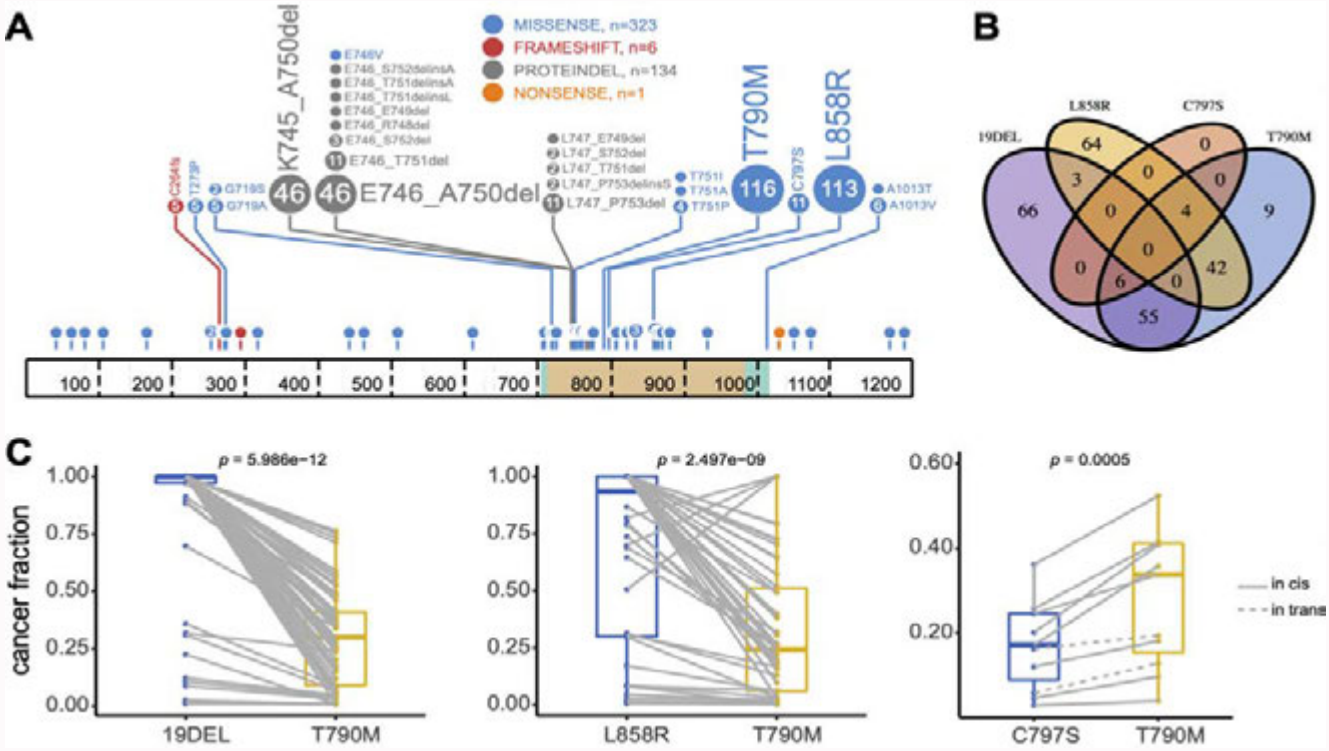

Figure 2: *EGFR* mutation spectrum and clonal dominance. (A) Protein paint of *EGFR* mutations. Different colored circles represent different mutation types, and numbers inside circles indicate the mutation numbers at specific positions. The colored region of *EGFR* protein schematic bar is catalytic domain of the protein tyrosine kinase. (B) Venn diagram of four main *EGFR* variants, *19DEL*, *L858R*, *T790M* and *C797S*. (C) Paired comparison of cancer fraction between variants *19DEL* and *T790M*, *L858R* and *T790M*, *C797S* and *T790M*. Dots with joint line are the matched, which comes from the same patient. Solid and dashed lines indicate cis-mutation and trans-mutation, respectively.

one RET fusion happened in a SCLC patient, accompanying with PTEN mutation and GNA11 mutation. *ALK* fusion detected in this study all broke in exon 20 of *ALK* and its partners included STRN (1), EML4 (6). The latter one was a well-known fusion partner in lung cancer and sensitive to *ALK* inhibitors [32]. As for gene amplification, they were observed only in two genes, *EGFR* (n=5) and *HER2* (n=3). One patient had both *EGFR* and *HER2* amplification.

Association between ctDNA alteration feature and clinical characteristics

Most genes' positive rates between female and male patients didn't have significant difference except for *EGFR* and *KRAS* (Supplementary Figure 3A). *EGFR* mutation is more frequent in female patients (47.88% *vs.* 36.36%, p<0.01) while *KRAS* positive detection rate is higher in male patients (6.58% *vs.* 1.42%, p<0.001). Besides, we found the positive detection rate of *HER2* was higher in patients without smoking history as well as *EGFR*, and the similar conclusion about *EGFR* was demonstrated before by a large Asian population study [33]. In contrast, higher positive rates in *BRAF*, *KRAS*, *DDR2* and *MEK1* were associated with patients with smoking history (Supplementary Figure 3B). The positive rate of ctDNA NGS test increased with disease stages (61.54%, 68.75% and 74.51% for stage I, III, IV, respectively), though no statistical significance probably due to large difference in sample size. There were only 3 patients of stage II, so it's not mentioned here. For similar reason, we only compared detection positive rate between adenocarcinoma and squamous cell carcinoma. Adenocarcinoma patients had a relative higher proportion than squamous cell carcinoma (74.13% *vs.* 58.06% p<0.05), so did *EGFR* positive rate (45.62% *vs.* 22.58%, p<0.01). Interestingly, *HER2* and *RET* alterations weren't detected in squamous cell carcinoma while they were the frequently mutated genes in adenocarcinoma (Supplementary Figure 3C).

ctDNA *EGFR* mutation features in lung cancer

EGFR mutation always plays an important role in targeted therapy of lung cancer, due to the successful application of *EGFR* tyrosine kinase inhibitors. In this cohort, we detected 464 mutations on *EGFR* in total, as well as 5 amplifications, accounting for more than half of the patients with positive results (58%, Supplementary Figure 2). The most frequent alteration in *EGFR* gene is exon 19 deletion (19DEL, 28%), and the deletion variants mainly occurred in four types including p.K745_A750del, p.E746_A750del, p.E746_T751del and p.L747_P753del (Figure 2A). Also, *EGFR* T790M and L858R mutation both made up around one quarter (Supplementary Figure 4). T790M always occurred in patients with drug-resistance to the first- or second-generation *EGFR-TKI* [34]. The 19DEL and L858R accompanying rate with T790M were 46.9% (61/130) and 40.7% (46/113) respectively (Figure 2B), but no statistical difference was achieved between them. Only 9 patients were detected T790M positive alone, not co-occurred with any one of the two above. However, eight of them either took first-generation *EGFR-TKI* before or had other uncommon *EGFR* mutations, or both, except for one whose clinical information was not available. Eleven C797S mutations were identified altogether with T790M in 10 patients (Figure 2B), while patient 0545 carried two different C797S cis mutations (c.T2389A and c.G2390C) at the same time. And most of the C797S mutations were in cis with T790M, which is a well-known cause of drug resistance to third-generation *EGFR-TKI* like Osimertinib [35-

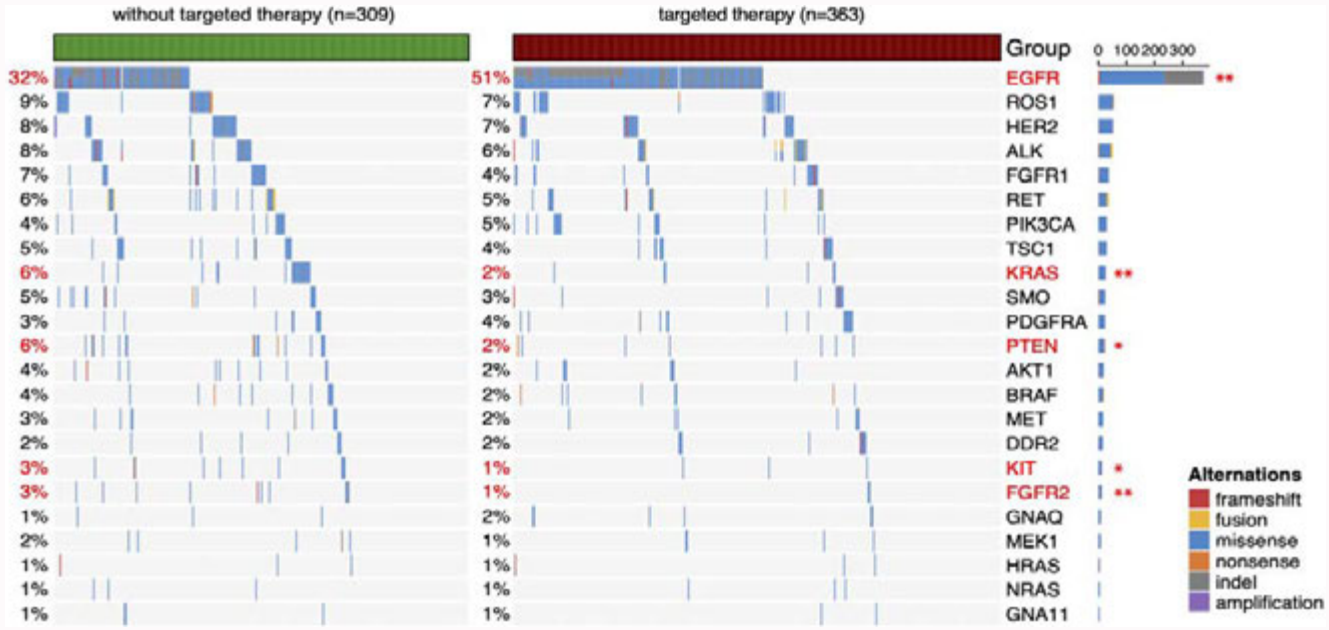

Figure 3: ctDNA mutation landscape grouped by targeted therapy status. The heatmap of 23 genes' alterations among 672 patients, and different colors correspond to 6 variation types: frameshift, fusion, missense, nonsense, indel, amplification. The side bars on the right indicate the total numbers of certain gene mutations. Genes in red are differently distributed between group "without targeted therapy" and "targeted therapy", and their significant levels are marked with * (p<0.05) or ** (p<0.01).

37]. Specifically, the mutation cancer fraction distribution of *EGFR* 19DEL, L858R, T790M and C797S, showed that 19DEL/L858R mutation was dominant to T790M and C797S' cancer fraction was significantly lower than that of T790M no matter whether they were in cis-structure or not (Figure 2C), which revealed the capability of ctDNA NGS test to reflect gene mutation clone evolution. Besides, A1013T/A1013V, T751A/T751I/T751P, G719A/G719S, T273P and C264fs were several variants with recurrent number no less than 5 (Figure 2A). Remarkably, 6 patients with mutation in the T751 residue (T751A/T751I/T751P) all had concurrent 19DEL alteration, and 4 out of the 6 patients among them still carried T790M mutation. In addition, some uncommon *EGFR* mutations [17,38], like I1050T, K806R, S768I and L861Q, always co-occurred with either *EGFR* hot mutations or other gene mutations.

Targeted therapy influenced ctDNA mutation spectrum

The spectrum of ctDNA alterations in patients treated with targeted drugs was distinguished from that in patients never taking targeted drugs (Figure 3). Patients received targeted therapy before testing included not only people who showed clinical drug-resistant signs and/or disease progression after a period of targeted treatment, but also those given empirical therapy such as the first line TKI without genetic testing. It was reasonable that targeted therapy group had much more *EGFR* mutations than patients without targeted therapy (51% *vs.* 32%, p<0.01). Notably, the co-existence *EGFR* mutations with other genes like *HER2*, *PIK3CA*, seemed more common in patients treated with targeted therapy, which suggested possible mechanism of resistance. But some gene mutations were less frequent in targeted therapy group, including *KRAS* (2% *vs.* 6%, p<0.01), PTEN (2% *vs.* 6%, p<0.05), KIT (1% *vs.* 3%, p<0.05) and FGFR2 (1% *vs.* 3%, p<0.01). Furthermore, mutation signatures were obviously different in patients treated by different administered therapy of Osimertinib (first-line n=11 *vs.* second-line n=50, Supplementary Figure 5). In patients treated with first-line Osimertinib, negative results were more common (54.55% *vs.* 14.00%, p<0.01), and Osimertinib targeted mutations (including 19DEL, L858R and T790M) were less frequently detected in plasma ctDNA (36.36% *vs.* 54.00%, p=0.29). Besides, patients receiving Osimertinib as second-line therapy had higher positive detection rate of C797S than first-line therapy (18.00% *vs.* 9.09%, p=0.67).

ctDNA mutation pattern in different metastatic sites

ctDNA mutation spectrum was found to be associated with different metastatic sites (Figure 4). The positive detection rate among adrenal (n=16), bone (n=156), brain (n=78) and liver (n=29) metastasis were 93.75%, 85.26%, 69.23%, 75.86%, respectively (Figure 4A). Apparently, brain metastasis had particularly high negative detection rate. Due to a small sample size, only a part of 23 genes' alteration was observed in adrenal metastasis, as well as in liver metastasis. As expected, brain metastasis' mutation spectrum had significantly less number of alterations per patient than others (Figure 4B). It might be explained by the blood-brain barrier which restrains tumor DNA from shedding into peripheral. High proportion of *EGFR* mutation was a generic feature for all groups (Figure 4C). Moreover, *EGFR* mutations were independently positively associated with bone metastasis, while *KRAS* and *PIK3CA* mutations were significantly associated with liver metastasis (Figure 4D and Table S4). Also, HER2 mutation was more likely to occur in patients with adrenal metastasis, as well as *FGFR1*.

Discussion

Our study analyzed 672 lung cancer patients' ctDNA profile and its correlation ship with clinical characteristics. Gene alterations were detected in 73.36% of patients, and nearly a half subjects (45.54%, 306/672) were benefit from this ctDNA NGS test. According to previous studies, positive detection rate varied with different gene number in the panel [15-18]. Our 23-gene panel's positive detection rate (73.36%) was close to that (73.9%) of a study with 37-gene panel

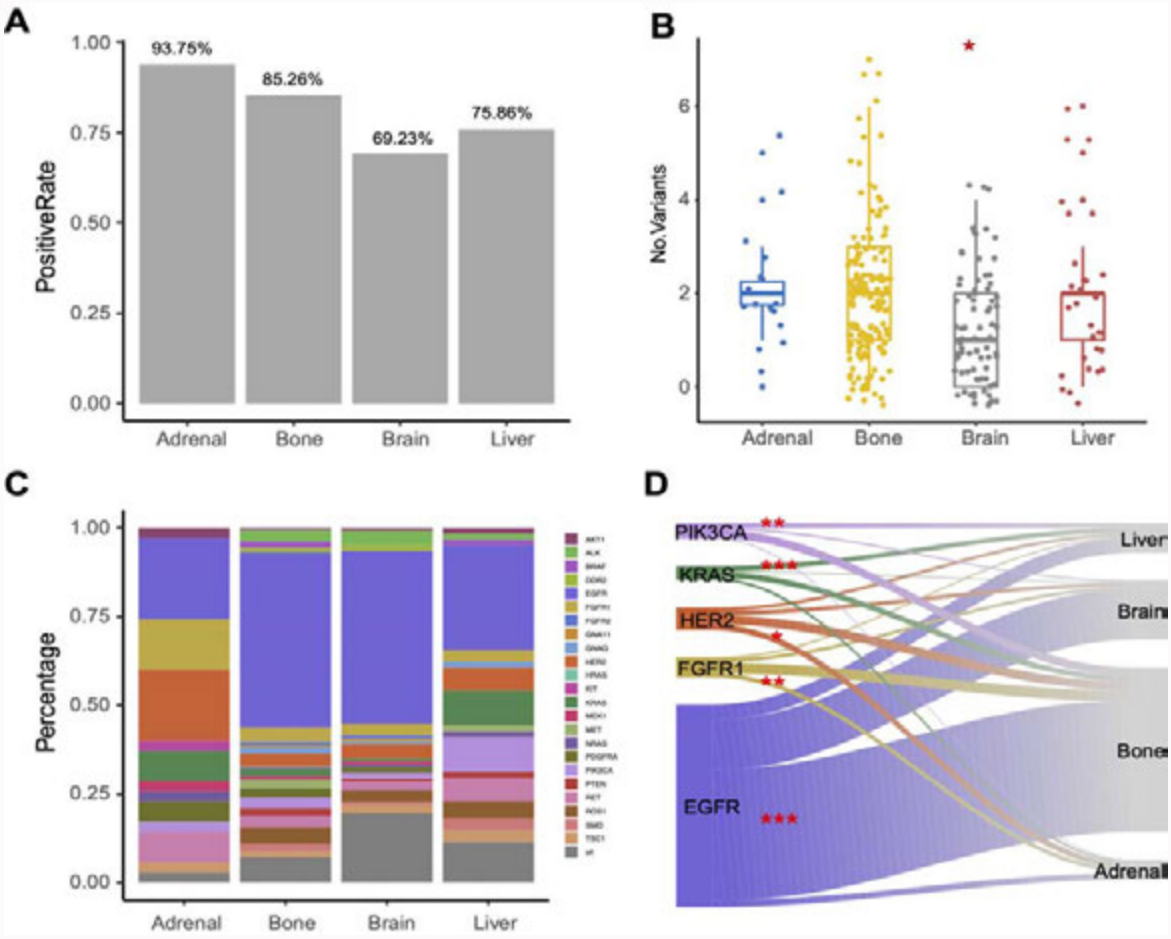

Figure 4: Mutation pattern among patients with different metastatic sites. (A) The positive rates of gene mutation in patients with adrenal, bone, brain, liver and multiple metastasis. (B) Number of variants per patient in different metastatic sites. (C) The composition of mutated genes in different metastasis. (D) Association of indicated genes with different metastasis. Positive association is marked with red asterisks while black asterisks represent negative association. (*p<0.05, **p<0.01, ***p<0.001).

[17], which demonstrated the multi-gene panel testing we used covered most frequently mutated genes in lung cancer. But one disadvantage was that our panel didn't contain *TP53* which is one of the most common genes in lung cancer [39], since it was designed for genes with targeted drugs on the market. Nevertheless, medical field develops so soon, more and more genes will have targeted molecular agents. Just like *TP53*, a recent study found a compound to recover p53 activity and it is on clinical trial now [40]. Maybe we should update our gene panel yearly, especially focus on genes which mutate frequently or have potential treatment, to provide a more comprehensive testing for patients. For instance, although *NTRK* gene fusion is rare in lung cancer with prevalence about 1%, it is recommended to be included in NGS testing panel since its targeted drugs (Entrectinib and Larotrectinib) were approved for solid tumors recently [41]. Although ctDNA genetic detection can always capture actionable mutations at a high concordance with tissue [12], 27% of patients still were detected with negative results in our cohort. It might be explained by the individual difference about tumor DNA shedding ability and ctDNA level. Besides, the positive rate of gene fusion or amplification was relative lower than that of mutation in ctDNA NGS analysis. There was some technology reason in addition to their own distribution feature in the population. ctDNA's fragment length is usually 160bp, and the short length makes gene fusion difficult to detect in some platform [42]. As for amplification, the low level of ctDNA in plasma challenges its detection [2]. So, if ctDNA genetic testing fails to detect targetable alteration, a subsequent tumor tissue genotyping is recommended [2,43].

The most frequently mutated gene *EGFR* was detected in 42.41% of all patients, far higher than others. As reported [16], *EGFR* mutation were associated with female, non-smoking history, adenocarcinoma. We also found *EGFR* and *KRAS* mutations were mutually exclusive in this study, and *KRAS* mutation tended to happen in males, which agrees with previous knowledge [44-47]. The ctDNA genetic test identified not only main type of *EGFR* variants like 19DEL, L858R and T790M, but also some uncommon mutations prompting possible drug resistance, such as I1050T, K806R, S768I and L861Q. Although not fully clarified, S768I was already demonstrated inferior tumor response rate [38] and occurred in the beginning of *EGFR-TKI* treatment [48]. All C797S were concomitant mutation of T790M, and we found one patient carried two different point mutations of C797S, both in cis with T790M. Similar situation was reported before [49], wherein cis and the other in trans. Besides, 6 *EGFR* mutations at T751 residue in this study all co-occurred with 19DEL. And T790M mutations existed in two-thirds of these 6 patients at the same time, it's hard to tell whether T751 residue mutation was associated with drug resistance or not. Mutation frequency in *EGFR* T751 residue seemed particular low in a recent large Chinese lung cancer patients'

study (1/1200) [17] compared with our cohort (6/672), hence there might be population heterogeneity. Whatever, its role needs further investigation in more rigorous research.

Targeted therapy and metastasis can influence mutation spectrum of ctDNA as we expected. The positive rate of *EGFR*, *KRAS*, *PTEN*, *KIT* and *FGFR2* mutation changed significantly between patients treated by targeted therapy or not. However, a sample bias may contribute to the higher rate of *EGFR* in targeted therapy group since patients in this group mostly were *EGFR* positive before. Except for *EGFR*, the other 4 genes' mutation rate was lower in targeted therapy group than that in patients without targeted therapy, suggesting that they may not tend to be secondary mutation acquired during *EGFR-TKI* treatment. In fact, *KRAS* mutations and defunction variants of PTEN are well-studied primary resistance mechanisms of gefitinib and erlotinib [50-53].

As for plasma ctDNA alteration in different metastatic sites, brain metastasis' positive rate is apparently the lowest (69.23%) while others can reach up to 80% like bone metastasis. Blood-brain barrier can affect ctDNA's shedding into blood [54]. Studies [55-57] show that cerebrospinal fluid could improve positive detection rate remarkably, as well as the diversity of variants. So cerebrospinal fluid could be a better source for ctDNA mutation detection than plasma in lung cancer patients with brain metastasis when it comes to sampling [58], especially patients received treatment before. Moreover, some gene mutations might have organotropism preference (*EGFR* for bone metastasis, *HER2* and *FGFR1* for adrenal metastasis, and *KRAS*, *PIK3CA* for liver metastasis), which might assist in predicting tumor metastatic sites. A study of large cancer patients' cohort has shown this kind of associations between somatic alterations and metastatic patterns [59]. It also indicates that tumor cell in lung adenocarcinoma patients with *EGFR* mutation tend to transfer to brain. However, our data didn't identify the same pattern, which might be caused by the plasma ctDNA we used for NGS, instead of tissue tumor DNA. A more comprehensive study about organ-specific metastasis in lung cancer should be carried out to elucidate this issue.

Conclusion

Altogether, ctDNA NGS analysis has showed great potential to identify therapeutic biomarker of lung cancer from real world data as well as possible drug-resistant alterations. Positive detection rate and mutation signature were closely associated with smoking history, gender, pathological type, treatment and metastatic sites. Plasma ctDNA alterations might play an important role in clinical outcome prediction, such as metastasis. Gene mutations' organotropism preference is still needed to be validated in large prospective cohort.

Acknowledgement

The authors would like to thank Qiongxia Hu, BiqinMou and Jing Wangfrom Precision Medicine Center, West China Hospital, for their technical support.

References

1. Hirsch FR, Scagliotti GV, Mulshine JL, Kwon R, Curran WJ Jr, Wu YL, et al. Lung cancer: Current therapies and new targeted treatments. Lancet. 2017;389(10066):299-311.
2. Aggarwal C, Rolfo CD, Oxnard GR, Gray JE, Sholl LM, Gandara DR. Strategies for the successful implementation of plasma-based NSCLC genotyping in clinical practice. Nat Rev Clin Oncol. 2021;18(1):56-62.
3. Bade BC, Dela Cruz CS. Lung cancer 2020: Epidemiology, etiology, and prevention. Clin Chest Med. 2020;41(1):1-24.
4. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. CA Cancer J Clin. 2015;65(2):87-108.
5. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394-424.
6. Latest global cancer data: Cancer burden rises to 19.3 million new cases and 10.0 million cancer deaths in 2020.
7. Cappuzzo F, Ciuleanu T, Stelmakh L, Cicenas S, Szczesna A, Juhasz E, et al. Erlotinib as maintenance treatment in advanced non-small-cell lung cancer: A multicenter, randomised, placebo-controlled phase 3 study. Lancet Oncol. 2010;11(6):521-9.
8. Janne PA, Yang JC, Kim DW, Planchard D, Ohe Y, Ramalingam SS, et al. AZD9291 in EGFR inhibitor-resistant non-small-cell lung cancer. N Engl J Med. 2015;372(18):1689-99.
9. Katayama R, Lovly CM, Shaw AT. Therapeutic targeting of anaplastic lymphoma kinase in lung cancer: A paradigm for precision cancer medicine. Clin Cancer Res. 2015;21(10):2227-35.
10. Lindeman NI, Cagle PT, Aisner DL, Arcila ME, Beasley MB, Bernicker EH, et al. Updated molecular testing guideline for the selection of lung cancer patients for treatment with targeted tyrosine kinase inhibitors: Guideline from the college of American pathologists, the international association for the study of lung cancer, and the association for molecular pathology. Arch Pathol Lab Med. 2018;142(3):321-46.
11. Esposito Abate R, Frezzetti D, Maiello MR, Gallo M, Camerlingo R, De Luca A, et al. Next generation sequencing-based profiling of cell free DNA in patients with advanced non-small cell lung cancer: Advantages and Pitfalls. Cancers (Basel). 2020;12(12):3804.
12. Leighl NB, Page RD, Raymond VM, Daniel DB, Divers SG, Reckamp KL, et al. Clinical utility of comprehensive cell-free DNA analysis to identify genomic biomarkers in patients with newly diagnosed metastatic non-small cell lung cancer. Clin Cancer Res. 2019;25(15):4691-700.
13. Odegaard JI, Vincent JJ, Mortimer S, Vowles JV, Ulrich BC, Banks KC, et al. Validation of a plasma-based comprehensive cancer genotyping assay utilizing orthogonal tissue- and plasma-based methodologies. Clin Cancer Res. 2018;24(15):3539-49.
14. Aggarwal C, Thompson JC, Black TA, Katz SI, Fan R, Yee SS, et al. Clinical implications of plasma-based genotyping with the delivery of personalized therapy in metastatic non-small cell lung cancer. JAMA Oncol. 2019;5(2):173-80.
15. Mack PC, Banks KC, Espenschied CR, Burich RA, Zill OA, Lee CE, et al. Spectrum of driver mutations and clinical impact of circulating tumor DNA analysis in non-small cell lung cancer: Analysis of over 8000 cases. Cancer. 2020;126(14):3219-28.
16. Chen H, Wang A, Wang J, He Z, Mao Y, Liu L. Target-based genomic profiling of ctDNA from Chinese non-small cell lung cancer patients: A result of real-world data. J Cancer Res Clin Oncol. 2020;146(7):1867-76.
17. Wen S, Dai L, Wang L, Wang W, Wu D, Wang K, et al. Genomic signature of driver genes identified by target next-generation sequencing in Chinese non-small cell lung cancer. Oncologist. 2019;24(11):e1070-81.
18. Thompson JC, Yee SS, Troxel AB, Savitch SL, Fan R, Balli D, et al. Detection of therapeutically targetable driver and resistance mutations in lung cancer patients by next-generation sequencing of cell-free circulating tumor DNA. Clin Cancer Res. 2016;22(23):5772-82.
19. Canisius S, Martens JW, Wessels LF. A novel independence test for somatic alterations in cancer shows that biology drives mutual exclusivity but chance explains most co-occurrence. Genome Biol. 2016;17(1):261.
20. tidyverse: Easily install and load the 'Tidyverse'.

21. Wickham H, Bryan J. Readxl: Read Excel Files. 2019.

22. Wickham H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. 2016.

23. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics. 2016;32(18):2847-9.

24. Chen H. VennDiagram: Generate High-Resolution Venn and Euler Plots. 2018.

25. Taiyun Wei, Simko V. R package "corrplot": Visualization of a Correlation Matrix. 2017.

26. Neuwirth E. RColorBrewer: Color Brewer Palettes. 2014.

27. Zhou X, Edmonson MN, Wilkinson MR, Patel A, Wu G, Liu Y, et al. Exploring genomic alteration in pediatric cancer using ProteinPaint. Nat Genet. 2016;48(1):4-6.

28. Milovanovic IS, Stjepanovic M, Mitrovic D. Distribution patterns of the metastases of the lung carcinoma in relation to histological type of the primary tumor: An autopsy study. Ann Thorac Med. 2017;12(3):191-8.

29. Riihimaki M, Hemminki A, Fallah M, Thomsen H, Sundquist K, Sundquist J, et al. Metastatic sites and survival in lung cancer. Lung Cancer. 2014;86(1):78-84.

30. Wood K, Hensing T, Malik R, Salgia R. Prognostic and predictive value in KRAS in non-small-cell lung cancer: A review. JAMA Oncol. 2016;2(6):805-12.

31. Devarakonda S, Sankararaman S, Herzog BH, Gold KA, Waqar SN, Ward JP, et al. Circulating tumor dna profiling in small-cell lung cancer identifies potentially targetable alterations. Clin Cancer Res. 2019;25(20):6119-26.

32. Lin JJ, Zhu VW, Yoda S, Yeap BY, Schrock AB, Dagogo-Jack I, et al. Impact of EML4-ALK variant on resistance mechanisms and clinical outcomes in ALK-positive lung cancer. J Clin Oncol. 2018;36(12):1199-206.

33. Shi Y, Au JS, Thongprasert S, Srinivasan S, Tsai CM, Khoa MT, 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(2):154-62.

34. Yu HA, Arcila ME, Rekhtman N, Sima CS, Zakowski MF, Pao W, et al. Analysis of tumor specimens at the time of acquired resistance to EGFR-TKI therapy in 155 patients with EGFR-mutant lung cancers. Clin Cancer Res. 2013;19(8):2240-7.

35. Thress KS, Paweletz CP, Felip E, Cho BC, Stetson D, Dougherty B, et al. Acquired EGFR C797S mutation mediates resistance to AZD9291 in non-small cell lung cancer harboring EGFR T790M. Nat Med. 2015;21(6):560-2.

36. Ercan D, Choi HG, Yun CH, Capelletti M, Xie T, Eck MJ, et al. EGFR mutations and resistance to irreversible pyrimidine-based EGFR inhibitors. Clin Cancer Res. 2015;21(17):3913-23.

37. Yu HA, Tian SK, Drilon AE, Borsu L, Riely GJ, Arcila ME, et al. Acquired resistance of EGFR-mutant lung cancer to a T790M-specific EGFR inhibitor: Emergence of a third mutation (C797S) in the EGFR tyrosine kinase domain. JAMA Oncol. 2015;1(7):982-4.

38. Chiu CH, Yang CT, Shih JY, Huang MS, Su WC, Lai RS, 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(5):793-9.

39. Testa U, Castelli G, Pelosi E. Lung cancers: Molecular characterization, clonal heterogeneity and evolution, and cancer stem cells. Cancers (Basel). 2018;10(8):248.

40. Chen S, Wu JL, Liang Y, Tang YG, Song HX, Wu LL, et al. Arsenic trioxide rescues structural p53 mutations through a cryptic allosteric site. Cancer Cell. 2021;39(2):225-39.e8.

41. Haratake N, Seto T. NTRK Fusion-positive non-small-cell lung cancer: The diagnosis and targeted therapy. Clin Lung Cancer. 2021;22(1):1-5.

42. Supplee JG, Milan MSD, Lim LP, Potts KT, Sholl LM, Oxnard GR, et al. Sensitivity of next-generation sequencing assays detecting oncogenic fusions in plasma cell-free DNA. Lung Cancer. 2019;134:96-9.

43. Bronkhorst AJ, Ungerer V, Holdenrieder S. The emerging role of cell-free DNA as a molecular marker for cancer management. Biomol Detect Quantif. 2019;17:100087.

44. Wang C, Yin R, Dai J, Gu Y, Cui S, Ma H, et al. Whole-genome sequencing reveals genomic signatures associated with the inflammatory microenvironments in Chinese NSCLC patients. Nat Commun. 2018;9(1):2054.

45. Zheng S, Wang X, Fu Y, Li B, Xu J, Wang H, et al. Targeted next-generation sequencing for cancer-associated gene mutation and copy number detection in 206 patients with non-small-cell lung cancer. Bioengineered. 2021;12(1):791-802.

46. Marchetti A, Martella C, Felicioni L, Barassi F, Salvatore S, Chella A, et al. 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. J Clin Oncol. 2005;23(4):857-65.

47. Linardou H, Dahabreh IJ, Bafaloukos D, Kosmidis P, Murray S. Somatic EGFR mutations and efficacy of tyrosine kinase inhibitors in NSCLC. Nat Rev Clin Oncol. 2009;6(6):352-66.

48. Leventakos K, Kipp BR, Rumilla KM, Winters JL, Yi ES, Mansfield AS. S768I mutation in EGFR in patients with lung cancer. J Thorac Oncol. 2016;11(10):1798-801.

49. Piotrowska Z, Nagy R, Fairclough S, Lanman R, Marcoux N, Gettinger S, et al. OA 09.01 characterizing the genomic landscape of EGFR C797S in lung cancer using ctDNA next-generation sequencing. J Thorac Oncol. 2017;12(11):S1767.

50. Pao W, Wang TY, Riely GJ, Miller VA, Pan Q, Ladanyi M, et al. KRAS mutations and primary resistance of lung adenocarcinomas to gefitinib or erlotinib. PLoS Med. 2005;2(1):e17.

51. Gkountakos A, Sartori G, Falcone I, Piro G, Ciuffreda L, Carbone C, et al. PTEN in lung cancer: Dealing with the problem, building on new knowledge and turning the game around. Cancers (Basel). 2019;11(8):1141.

52. Wang F, Diao XY, Zhang X, Shao Q, Feng YF, An X, et al. Identification of genetic alterations associated with primary resistance to EGFR-TKIs in advanced non-small-cell lung cancer patients with EGFR sensitive mutations. Cancer Commun (Lond). 2019;39(1):7.

53. Massarelli E, Varella-Garcia M, Tang X, Xavier AC, Ozburn NC, Liu DD, et al. KRAS mutation is an important predictor of resistance to therapy with epidermal growth factor receptor tyrosine kinase inhibitors in non-small-cell lung cancer. Clin Cancer Res. 2007;13(10):2890-6.

54. Cescon DW, Bratman SV, Chan SM, Siu LL. Circulating tumor DNA and liquid biopsy in oncology. Nature Cancer. 2020;1(3):276-90.

55. Zhang K, Dai Z, Liu S, Li D, Yang D, Cui S. [Clinical value of cerebrospinal fluid ctDNA in patients with non-small cell lung cancer meningeal metastasis]. Zhongguo Fei Ai Za Zhi. 2020;23(12):1039-48.

56. Ma C, Yang X, Xing W, Yu H, Si T, Guo Z. Detection of circulating tumor DNA from non-small cell lung cancer brain metastasis in cerebrospinal fluid samples. Thorac Cancer. 2020;11(3):588-93.

57. Li H, Xie Y, Lin Y, Yu T, Yin Z. [Different gene mutation spectrum of the paired CSF and plasma samples in lung adenocarcinoma with leptomeningeal metastases: the liquid biopsy based on circulating tumor DNA]. Zhongguo Fei Ai Za Zhi. 2020;23(8):646-54.

58. De Mattos-Arruda L, Mayor R, Ng CKY, Weigelt B, Martinez-Ricarte F, Torrejon D, et al. Cerebrospinal fluid-derived circulating tumor DNA

better represents the genomic alterations of brain tumors than plasma. Nat Commun. 2015;6:8839.

59. Nguyen B, Fong C, Luthra A, Smith SA, DiNatale RG, Nandakumar S, et al. Genomic characterization of metastatic patterns from prospective clinical sequencing of 25,000 patients. Cell. 2022;185(3):563-75.e511.