Lung Cancer 134 (2019) 108–116

Contents lists available at ScienceDirect

Lung Cancer

journal homepage: www.elsevier.com/locate/lungcan

Circulating tumor DNA detection is correlated to histologic types in patients with early-stage non-small-cell lung cancer

Bin Zhang[a,1], Xueliang Niu[b,c,1], Qiang Zhang[a,1], Chunli Wang[b,c], Bo Liu[b,c], Dongsheng Yue[a], Chenguang Li[a], Giuseppe Giaccone[a,d], Shiyong Li[c,e], Liuwei Gao[a], Hua Zhang[a], Jian Wang[f,g], Huanming Yang[f,g], Renhua Wu[b,c], Peixiang Ni[b,c], Changli Wang[a,***], Mingzhi Ye[c,e,h,**], Weiran Liu[i,*]

[a] Department of Lung Cancer, Tianjin Lung Cancer Center, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Key Laboratory of Cancer Prevention and Therapy, Tianjin's Clinical Research Center for Cancer, Tianjin, China
[b] Tianjin Medical Laboratory, BGI-Tianjin, BGI-Shenzhen, Tianjin, China
[c] BGI Genomics, BGI-Shenzhen, Shenzhen, China
[d] Georgetown University, Washington, District of Columbia, USA
[e] BGI-Guangzhou Medical Laboratory, BGI-Shenzhen, Guangzhou, China
[f] BGI-Shenzhen, Shenzhen, China
[g] James D. Watson Institute of Genome Sciences, Hangzhou, China
[h] BGI-Guangzhou, Guangzhou Key Laboratory of Cancer Trans-Omics Research, Guangzhou, China
[i] Department of Anesthesiology, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Key Laboratory of Cancer Prevention and Therapy, Tianjin's Clinical Research Center for Cancer, Tianjin, China

ARTICLE INFO

Keywords:
Circulating tumor DNA
Next-generation sequencing
Non-small-cell lung cancer
Gene mutations
Classification

ABSTRACT

Objectives: Circulating tumor DNA (ctDNA) testing in plasma in patients with non-small-cell lung cancer (NSCLC) has the potential to be a supplemental or surrogate tool for tissue biopsy. Detection of genomic abnormalities in ctDNA and their association with clinical characteristics in early-stage NSCLC need to be clarified.
Materials and methods: Here, we comprehensively analyzed gene variations of 48 tumor tissues and 48 matched preoperative (pre-op) plasma and 25 postoperative (post-op) plasma from early-stage NSCLC patients using a targeted 546 genes capture-based next generation sequencing (NGS) assay.
Results: In early-stage NSCLC, the average mutation allele frequency (MAF) in pre-op plasma ctDNA was lower than that in tissue DNA (tDNA). The concordant gene variations between pre-op ctDNA and tDNA were difficult to detect. However, we found the tissue- pre-op plasma concordant ctDNA mutation detection ratio in lung squamous cell carcinoma (LUSC) was much higher than that in lung adenocarcinoma (LUAD). We also established a LUSC-LUAD classification model by a least absolute shrinkage and selection operator (LASSO) based approach to help separate LUAD from LUSC based on ctDNA profiling. This model included 14 gene mutations and extracted an accuracy of 89.2% in the training set and 91.5% in the testing set. Correlation analysis showed tDNA-ctDNA concordant ratio was related to histologic subtype, gene mutations and tumor size in early-stage NSCLC.
Conclusion: This study suggests histology subtype and gene mutations could affect ctDNA detection in early-stage NSCLC. NGS-based ctDNA profile has the potential utility in LUSC-LUAD classification.

* Corresponding author at: Department of Anesthesiology, Tianjin Medical University Cancer Institute and Hospital, Huan-Hu-Xi Road, Ti-Yuan-Bei, He Xi District, Tianjin, 300060, China.
** Corresponding author at: BGI-Shenzhen, Beishan Industrial Zone, Yantian District, Shenzhen, 518083, China.
*** Corresponding author at: Department of Lung Cancer, Tianjin Lung Cancer Center, Tianjin Medical University Cancer Institute and Hospital, Huan-Hu-Xi Road, Ti-Yuan-Bei, He Xi District, Tianjin 300060, China.
E-mail addresses: wangchangli@tjmuch.com (C. Wang), yemingzhi@genomics.cn (M. Ye), wrliu@tmu.edu.cn (W. Liu).
[1] These authors contributed equally to this study.

https://doi.org/10.1016/j.lungcan.2019.05.034
Received 19 December 2018; Received in revised form 28 May 2019; Accepted 30 May 2019
0169-5002/ © 2019 Elsevier B.V. All rights reserved.

1. Introduction

Lung cancer is the leading cause of cancer-related deaths worldwide [1]. LUAD and LUSC are the most common histologic subtypes of NSCLC. Several genetic alterations drive tumor growth in adenocarcinoma, such as EGFR mutations, ALK and ROS1 translocations are found in adenocarcinoma histology and these can be successfully targeted with small molecule tyrosine kinase inhibitors [2–5]. In East Asians close to 50% of patients with adenocarcinoma harbor EGFR mutations, and most of these mutations are found in never-smokers. In contrast, squamous cell lung cancer still represents a tumor subtype that is most common in heavy smokers and no targetable driver mutations have so far been identified in this tumor type. Because of the association with smoking, squamous cell histology has a higher number of mutations than adenocarcinomas [6,7]. Genomic characterization of lung cancer based on tumor tissue of patients with resected disease has been comprehensively studied, but limited analysis of blood ctDNA in early NSCLC is available.

Circulating tumor DNA (ctDNA) is composed of small fragments of nucleic acid that are released from apoptotic or necrotic tumor cells and circulate in blood [8,9]. Detection of ctDNA has several advantages over conventional tumor biopsies. Sampling ctDNA from blood is noninvasive and can be repeated over time, which allows potentially for early diagnosis of cancer, identifying minimal residual disease or recurrence, real-time monitoring of responses to treatment, and predicting the prognosis [10–17]. Considering tumor heterogeneity, analysis of ctDNA could theoretically also provide a more comprehensive and representative information of multiple tumor deposits. ctDNA–-based liquid biopsy is a promising surrogate and potential supplement to tumor biopsies.

Most studies of ctDNA in NSCLC focused on late-stage patients, where the tumor burden is high. Limited data is available on the evaluation of ctDNA detection in early-stage NSCLC. In this study, we investigated the ctDNA mutation profile using targeted NGS in matched tumor and plasma samples from 48 NSCLC patients who underwent surgery for stage I-IIIA at our institution. We investigated the concordance between tissue and plasma mutations in LUSC and LUAD and we analyzed the correlation between ctDNA gene mutations and clinical features. Furthermore, a LUSC-LUAD classification model was built based on gene mutations detected in ctDNA.

2. Materials and methods

2.1. Patient information

Forty-eight patients with NSCLC who underwent radical surgery at Tianjin Medical University Cancer Institute and Hospital between October 2014 and December 2015 were included in this study (Table 1). Patients who received neoadjuvant chemotherapy before surgery or had a prior cancer diagnosis were excluded. In the cohort, 31 patients with stage IB-IIIA received platinum-based double chemotherapy after surgery, the other 17 patients with stage IA-IB did not receive adjuvant chemotherapy. The median follow-up time was 46.5 (range 6–50) months. Tumor samples were collected immediately after surgical resection and snap-frozen in liquid nitrogen. Blood samples were collected in EDTA tubes and plasma was isolated within 3 h of blood draw 1 day before surgery and 2–5 days (average 3.8 days) after surgery. DNA from peripheral blood lymphocytes (PBMC) or the adjacent non-tumorous tissues from each patient was used as corresponding normal control to exclude germline mutations and ensure that only somatic mutations were identified from tumor tissue or ctDNA samples.

This study protocol was reviewed and approved by the Institutional Review Board of Tianjin Medical University Cancer Institute and Hospital. Written informed consent was obtained from all patients.

Table 1
Clinical and molecular features.

Characteristics (48 total samples)	No.
Age at surgery	
≤60	23
> 60	25
Gender	
Male	30
Female	18
Smoking history	
Never	16
Ever	32
Tumor size	
≤4	30
>4	18
pN	
N0	30
N1 + N2	18
Stage	
I	22
II	11
III	15
Histology type	
Adenocarcinoma	26
Squamous cell carcinoma	22
Tumor location	
Central type	20
Peripheral type	28
Recurrence and metastasis	
Yes	16
No	28
Unknown	4
Adjuvant chemotherapy	
Yes	31
No	17
Driver mutations(tumor tissue)	
EGFR L858R	3
EGFR 19del	12
EGFR G719A	1
EGFR T790M + EGFR 19del	1
EGFR L858R + EGFR 19del	2
KRAS	7
ERBB2	3
BRAF	1
PIK3CA	4

2.2. Sample collection and processing

The patient blood samples collected in EDTA tubes were centrifuged at 1600 *g*/min for 10 min, and peripheral blood lymphocytes were separated and stored at -80 °C until use. The plasma was collected by centrifuging the supernatant from the blood samples again at 18,000 *g*/min for 5 min and was stored at -80 °C until use.

DNA from fresh tumor tissues that contained at least 60% tumor cells, as assessed by H&E staining was extracted using QIAamp® DNA mini Kit (Qiagen, Germany) following the manufacturer's protocol. The plasma cell free-DNA (cfDNA) was extracted from at least 2 ml plasma using the QIAamp® Circulating Nucleic Acid Kit (Qiagen, Germany) according to the vendor's instructions and the concentration of genomic DNA and cfDNA was measured by Qubit 2.0 BR Kit and Qubit 2.0 HS Kit (Invitrogen, USA) according to the respective protocols.

2.3. Library preparation and next generation sequencing

Next generation sequencing (NGS) was performed using Illumina Hiseq 2500/4000 platforms (Illumina, San Diego, USA), following the manufaturer's instructions. For the library construction, genomic DNA from fresh tumor tissue or peripheral blood cells was disrupted into fragments (about 250 bps) by ultrasound (30 s followed by 30 s rest (repeated 3 times) and short centrifugation, repeated 3 times). The DNA fragments were end-repaired, terminated with a cohesive adenine-tail

and ligated to the Illumina-indexed adaptors, according to the library construction protocol [18]. Agilent 2100 Bioanalyzer (Agilent Technologies) was used to determine the size and concentration of the library. The library for ctDNA was carried out using KAPA LTP Library Preparation Kit for Illumina Platform (Kapa Biosystems) as per the manufacturer's instructions without modification [19].

Low quality reads were removed from sequencing data using SOAPnuke v1.2.0. After that, the clean sequencing data were mapped to the human genome (hg19) using BWA aligner v0.6.2 [20] and Samtools v0.1.19. PCR duplicate reads were marked using Picard v1.98. The alignment refinement was performed using GATK v2.3–9 [21]. Next, the candidate SNVs were called using Bayesian Model. In order to filter the SNVs with strand bias and read location bias, we used the Fisher Exact Test and Kolmogorov-Simrnov Test [22]. Then SNVs in local control set which constructed by 4725 blood cell samples were filtered. In addition, we scored SNVs according to GC%, adjacent SNVs and InDels, multiple mapping locations, base quality, mapping quality and duplication. Finally, SNVs with low score were removed. Candidate InDels were extracted from the CIGAR information in the bam files. Then Candidate InDels were called using local assembly by de Bruijn approach [23]. Then, we removed the InDels in local control set (blood-cell samples). Additionally, we checked the InDels in simple repeat regions of human genome again, on account of more sequencing errors that may have occurred in these regions.The somatic SNV/InDels calling in the ctDNA samples sequencing data was processed with the same method as fresh tumor tissue sequencing data variant calling approach. Finally, we used ANNOVAR for the Somatic SNVs and InDels annotation.

We used the Oseq-T panel (a high coverage of all cancers, all exons of 546 genes, panel size 1.76 Mb, Nimblegen, USA) for the cancer-related genes capture. TMB (tumor mutation burden) was calculated as follow: (No. of total mutations- No. of driver mutations)/panel size.

2.4. Droplet digital PCR

ddPCR was performed to verify the NGS sequencing data in this study. According to the quick reference, QuantStudio™ 3D Digital PCR Master Mix v2 and TaqMan® Assay, samples were thawed to room temperature and inverted 10 times. DNA samples were diluted and the concentration of target sequence in the final reaction was 200 - 2000 copies/μL. Thereafter, the ddPCR reaction mix was prepared according to the protocol and immediately loaded onto QuantStudio TM 3D Digital PCR Chip. After that, PCR was performed using the ProFlex™ 2x Flat PCR System, with the anneal/extend temperature of 56.0 °C in stage 2. Finally, we analyzed the data using the QuantStudio™ 3D AnalysisSuite™ Software.

2.5. LASSO modeling

For modeling of 48 LUSC and LUAD patients, 169 candidate MGs (Mutation Genes) were initially included in the training feature sets and transformed into a 0, 1 mutation-distance matrix (M), which $M\_ij = 1$ indicates the i_th sample with mutation in j_th MG and 0 indicates non-mutations. For comprehensively evaluating the prognostic ability of multi-platform molecules to respective cancer types, we utilized a Least Absolute Shrinkage and Selection Operator (LASSO) based fast modeling algorithm to establish a performance pipeline. Here, the hosmerlemeshow test was proposed to select the optimal and suboptimal numbers of potential mutation factors along with LASSO algorithm for the predictive model. The LASSO was first proposed in 1996 by Tibshirani, it uses the linear formula to shrink the coefficients and select the most significant variables by tuning parameters that control the degree of penalty, which can be determined by cross-validation. According to the nature of the constraint, the LASSO algorithm tends to produce most coefficients as zero, and it improves the overall prediction accuracy by allowing a small amount of bias to reduce the variance of the predicted values. The selection of the optimal number of mutation factors was performed by cross-validation in this study. To account for the over-fitting problem and to select the best predictive features, the model was built based on a 5-fold cross-validation method with 10,000 times resampling. The best predictive model was further fitted to the training set itself and tested with a validation data set.

2.6. Statistical analysis

The heatmaps were plotted using R software (pheatmap package). Pearson correlation between mutation genes and clinical phenotypes were performed using R (v3.2) with the cor () function. We observed a minimal significance P value under 0.05 with Pearson correlation value 0.4 as cut off value. We used student's t test or Chi-square test to compare the difference of clinical factors (tumor size, age, tumor stage, gender, smoking history) between the two lung cancer subtypes.

The Wilcoxon signed rank test was adopted to test the statistical significance of tissue-blood-overlap mutation distributions of the two lung cancer subtypes. As formula,

$$\chi^2 = (S_1 + S_2)^2/(V_1 + V_2)$$

Where S_1 and S_2 is the rate of tissue-blood-overlap mutations in two cancer subtype groups, V_1 and V_2 is the variance of S_1 and S_2.

3. Results

3.1. Landscape of somatic mutations and driver mutations and validation by ddPCR

All of the samples (48 tumor tissues, 48 pre-operation (pre-op) and 25 post-operation (post-op) plasma samples) were sequenced using target capture gene panel. The panel which included 546 cancer-related genes (Supplementary Table 1) was able to detect SNVs and InDels, but not amplifications or translocations. The NGS data showed that 45 tissue samples contained one or more mutations; 34 pre-op and 20 post-op plasma samples had one or more variants (Supplementary Table 2–4). The median read depth of PBMC, tumor samples, pre-op samples and post-op plasma samples were 426 (range 62–794), 482 (range 278–958), 866 (range 392–1602) and 1254 (range 824–1916), respectively (Supplementary Fig. 1).

There were less gene mutations detected in plasma than in tumor tissues, despite the higher depth of reads utilized in plasma samples. The median mutation allele frequency (MAF) in tumor DNA (tDNA) was 19.5% (range: 0.4%–79.6%). Most of the mutations (360/446; 80.7%) in tDNA were SNVs, with the remaining mutations being InDels (86/446, 19.3%). We found that some mutations in ctDNA were from clonal haematopoietic mutations [24,25]. So, we excluded these mutations in ctDNA in the following analysis. The median MAF in pre-op plasma ctDNA was 2.1% (range: 0.5%–33.4%). The median MAF for discordant mutations with tissues was lower than that of concordant mutations (Fig. 1A). We also found patients with stage III NSCLC had higher MAF in pre-op plasma ctDNA, compared to patients with stage I and II NSCLC (Fig. 1B). The median MAF in post-op plasma ctDNA was 1.2% (range: 0.6%–14.0%). Common driver gene mutations [EGFR: 39.6% (19/48), KRAS: 14.6% (7/48), BRAF: 2.1% (1/48); ERBB2: 6.3% (3/48), PIK3CA: 8.3% (4/48)] were found in 62.5% (30/48) individuals (tumor tissue samples)(Table 1). The distribution of driver gene mutations in different types of samples are shown in Fig. 1C. From these data, it appears that the detection ratio of common driver gene mutations is much lower in ctDNA [14.6% (7/48) in pre-op plasma and 4.0% (1/25) in post-op plasma] than in tDNA (62.5%, 30/48).

The sensitivity of ddPCR assay is higher than NGS assay. We performed ddPCR testing (QuantStudio™ 3D Digital PCR System, Thermo Fisher, USA) for 23 samples with EGFR, PIK3CA and TP53 hotspot mutations to confirm the NGS results (Supplementary Table 5). The results showed that there was a good detection concordance using NGS

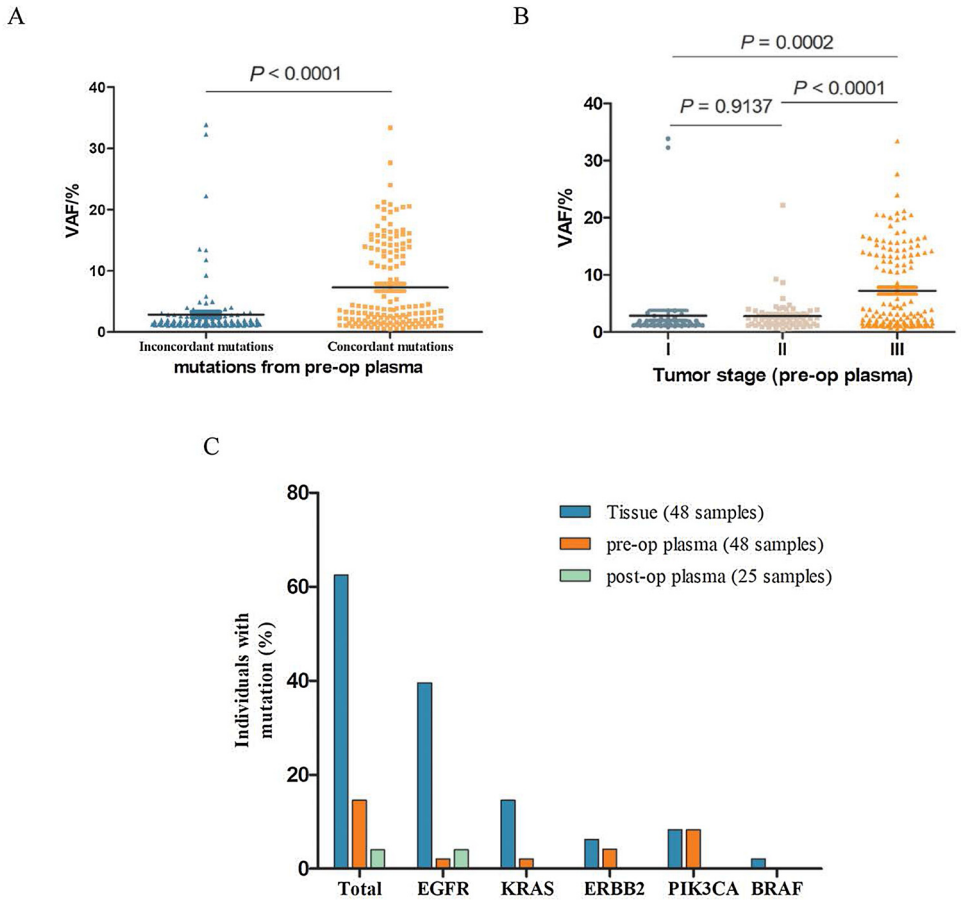

Fig. 1. The common driver genes detection ratio in different types of samples.
The figure showed the detecting rates of common driver gene mutations in different type of samples (blue: tumor tissue, orange: pre-op plasma, green: post-op plasma). The common driver gene mutations ratio was much lower in ctDNA than in tDNA. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

and ddPCR (tissue, 100%; plasma, 92.3%). The MAFs detected by the ddPCR approach were comparable to those detected by NGS. Most of the gene mutations in tissue samples listed in Supplementary Table 5 not detected in matched plasma by NGS were also not tested by ddPCR, suggesting it was not as relative low sensitivity of NGS assay. Thus, we believe the NGS data in this study is reliable.

3.2. Mutation concordance in matched tDNA and ctDNA samples

Of 48 matched tDNA and pre-op plasma ctDNA sample pairs, 15 contained concordant mutations (at least one mutation identified in both tDNA and ctDNA) and 3 had no variants in any of the 546 genes in either tDNA or ctDNA. We analyzed the concordance rates for all the gene mutations between tissue and pre-op plasma ctDNA. From the data showed in Table 2, the mutation concordance rate was 37.50% (95% CI 23.95%–52.65%), with a sensitivity of 37.50% (95% CI 22.73%–54.20%), specificity of 37.50% (95% CI 8.52%–75.51%), and positive predictive value (PPV) of 75.00% (95% CI 60.57%–85.42%).

3.3. Comparison of tissue-plasma concordant ctDNA mutations in LUSC and LUAD

For further study of the mutational concordance, we first compared the pre-op plasma ctDNA mutations in lung squamous cell carcinoma (LUSC) and lung adenocarcinoma (LUAD). We evaluated the distribution of the variants in tDNA and pre-op ctDNA, respectively. There were less mutations in LUAD than in LUSC which were sampled from either tumor tissue or plasma (Supplementary Fig. 2). Next, the tissue-plasma concordant ctDNA mutations in LUSC and LUAD were analyzed. We observed that the concordant mutations in tDNA and pre-op plasma ctDNA in LUAD were much less than those in LUSC (13 versus 129) (Fig. 2A). The tissue-plasma concordant ctDNA mutation detection ratio (the concordant mutation number in ctDNA was divided by the total mutation number in matched tDNA) in LUAD was much lower than that in LUSC ($P = 0.005$ Fig. 2B). The concordant ctDNA mutations were detected in 15% (4/26) of LUAD and 50% (11/22) of LUSC (Supplementary Table 6). The sensitivity was 18.18% (95% CI 5.19%–40.28%) in LUAD, and 61.11% (95% CI 35.75%–82.70%) in LUSC. In the post-op plasma, the tissue-plasma concordant ctDNA mutations were found only in one sample (Supplementary Table 7). Next, we compared the baseline clinical characteristics such as tumor size, age at surgery, tumor stage, gender and smoking status of LUSC patients and LUAD patients. There was no significant difference in tumor size, tumor stage, age at surgery between LUSC patients and LUAD patients ($P = 0.073$, $P = 0.961$, $P = 0.053$), except for gender and smoking status (more female and never-smokers in adenocarcinoma, $P = 0.002$, $P = 0.008$) (Fig. 2C-G). The data showed that the difference of the tissue-plasma concordant ctDNA mutation detection ratio in pre-op plasma was mainly correlated to histological subtype.

3.4. Pre-op plasma ctDNA displays high tissue specificity

To estimate if the pre-op plasma ctDNA has the potential to be a predictive marker for histological classification of NSCLC, we applied a least absolute shrinkage and selection operator (LASSO) based approach to build LUSC-LUAD classification models to distinguish the two types of lung cancer and extract the most predictive mutation genes feature based on ctDNA. We trained the LASSO model by proposing a mutation matrix from all pre-op ctDNA mutations with 10,000 times re-sampling over these mutation genes.

We randomly divided the study cohort into two sets, the training set and the test set. In total, a classifier with the highest bootstrap accuracy was identified as the optimal model and the predictors were then tested in the test cohort. The model performance is highlighted in Fig. 3A-B.

Table 2
Gene mutation concordance in tissue DNA and pre-op plasma ctDNA.

Sample ID	Histology type	Concordant mutations	Tissue only	Plasma only	Total	True negatives	True positives	Percentage of Plasma only	Percentage of Tissue only
194C	LUAD	0	5	0	5	0.00	0.00	0.00	1.00
196C	LUAD	0	4	3	7	0.00	0.00	0.43	0.57
198C	LUAD	3	2	2	7	0.00	1.00	0.00	0.00
199C	LUAD	0	5	3	8	0.00	0.00	0.38	0.63
211C	LUSC	15	3	7	25	0.00	1.00	0.00	0.00
219C	LUSC	0	0	0	0	1.00	0.00	0.00	0.00
222C	LUSC	2	1	3	6	0.00	1.00	0.00	0.00
229C	LUSC	0	4	0	4	0.00	0.00	0.00	1.00
241C	LUSC	6	6	4	16	0.00	1.00	0.00	0.00
251C	LUAD	0	6	3	9	0.00	0.00	0.33	0.67
253C	LUAD	0	4	3	7	0.00	0.00	0.43	0.57
254C	LUAD	0	14	0	14	0.00	0.00	0.00	1.00
256C	LUAD	0	3	0	3	0.00	0.00	0.00	1.00
264C	LUAD	0	9	1	10	0.00	0.00	0.10	0.90
265C	LUSC	0	15	1	16	0.00	0.00	0.06	0.94
271C	LUAD	0	15	0	15	0.00	0.00	0.00	1.00
278C	LUSC	57	11	11	79	0.00	1.00	0.00	0.00
288C	LUAD	0	2	0	2	0.00	0.00	0.00	1.00
289C	LUSC	0	0	0	0	1.00	0.00	0.00	0.00
290C	LUAD	0	6	0	6	0.00	0.00	0.00	1.00
292C	LUAD	5	5	7	17	0.00	1.00	0.00	0.00
294C	LUAD	4	2	0	6	0.00	1.00	0.00	0.00
295C	LUSC	11	14	3	28	0.00	1.00	0.00	0.00
299C	LUSC	0	8	4	12	0.00	0.00	0.33	0.67
301C	LUAD	0	5	3	8	0.00	0.00	0.37	0.63
302C	LUSC	2	15	2	19	0.00	1.00	0.00	0.00
308C	LUSC	19	8	15	42	0.00	1.00	0.00	0.00
313C	LUAD	0	2	0	2	0.00	0.00	0.00	1.00
317C	LUSC	3	8	0	11	0.00	1.00	0.00	0.00
323C	LUAD	0	12	6	18	0.00	0.00	0.33	0.67
324C	LUAD	0	6	3	9	0.00	0.00	0.33	0.67
330C	LUAD	0	5	2	7	0.00	0.00	0.29	0.71
338C	LUAD	0	10	1	11	0.00	0.00	0.09	0.91
360C	LUSC	0	4	0	4	0.00	0.00	0.00	1.00
361C	LUAD	0	7	0	7	0.00	0.00	0.00	1.00
408C	LUAD	0	2	1	3	0.00	0.00	0.33	0.67
410C	LUAD	1	5	7	13	0.00	1.00	0.00	0.00
457C	LUSC	5	1	5	11	0.00	1.00	0.00	0.00
594C	LUSC	0	11	1	12	0.00	0.00	0.08	0.92
619C	LUSC	4	3	1	8	0.00	1.00	0.00	0.00
623C	LUSC	0	9	4	13	0.00	0.00	0.31	0.69
624C	LUSC	0	0	0	0	1.00	0.00	0.00	0.00
659C	LUSC	5	2	3	10	0.00	1.00	0.00	0.00
667C	LUSC	0	19	1	20	0.00	0.00	0.05	0.95
670C	LUAD	0	3	4	7	0.00	0.00	0.57	0.43
674C	LUAD	0	3	0	3	0.00	0.00	0.00	1.00
697C	LUSC	0	16	0	16	0.00	0.00	0.00	1.00
715C	LUAD	0	2	1	3	0.00	0.00	0.33	0.67
Total		142	302	115	559	3	15	5	25

Notably, the best model extracts an accuracy of 89.2% in the training set and 91.5% in the test set, which ensures that the LASSO approach is robust to any random resampling errors. A total of 14 best predictive features are listed in the heatmap (Fig. 3C), including: TP53、SLIT2、NOTCH3、MTOR、LIFR、MRE11A、ARID2、ERCC3、KCNH2、CDC25C、RB1、ALK、NFE2L2 and FBXW7. Based on ctDNA mutation features, we established a LUSC-LUAD classification model which could aid in the histological classification of NSCLC especially in cases where pathological diagnosis does not provide a clear-cut definition.

3.5. ctDNA gene mutations and clinical features

In order to further understand the relationship between ctDNA gene mutations, the tissue-plasma concordant ctDNA mutation detection ratio and clinical features, the Pearson correlation analysis was used to evaluate the correlation coefficients. The relationship between pre-op ctDNA gene mutations and clinical features is shown in Fig. 4. The high tissue-pre-op plasma concordant ctDNA mutation detection ratio had a relatively high correlation with large tumor size (Pearson's $r = 0.71$) and LUSC histology (Pearson's $r = -0.45$). Moreover, the concordant tDNA-ctDNA mutation detection ratio was related to *TP53*, *ROS1*, *PIK3CA*, *KMT2D*, *EPHB1*, *CDKN2A* gene mutations (Pearson's $r = 0.42$-0.96). *TP53* gene mutations in pre-op plasma ctDNA was also found to be associated with tumor size (Pearson's $r = 0.67$) and LUSC histology (Pearson's $r = -0.40$). Furthermore, we analyzed the TMB (tumor mutation burden) of tDNA. LUSC had a higher TMB than LUAD ($P = 0.024$, t test) (Supplementary Fig. 3).

4. Discussion

In this study, we investigated the ctDNA mutation profile in 48 tDNA, 48 pre-op ctDNA and 25 post-op ctDNA samples from operable (stage I-IIIA) NSCLC patients using NGS. Two or more gene mutations were detected in 45(93.7%)fresh tumor tissue samples. At least two gene mutations were detected in 28 (59.6%) and a single gene mutation was detected in 6 pre-operative plasma. 31.2% (15/48) matched tDNA and pre-op plasma ctDNA sample pairs contained concordant mutations. These data suggest that it is feasible to detect ctDNA mutations from pre-op plasma by target sequencing in patients with early-stage

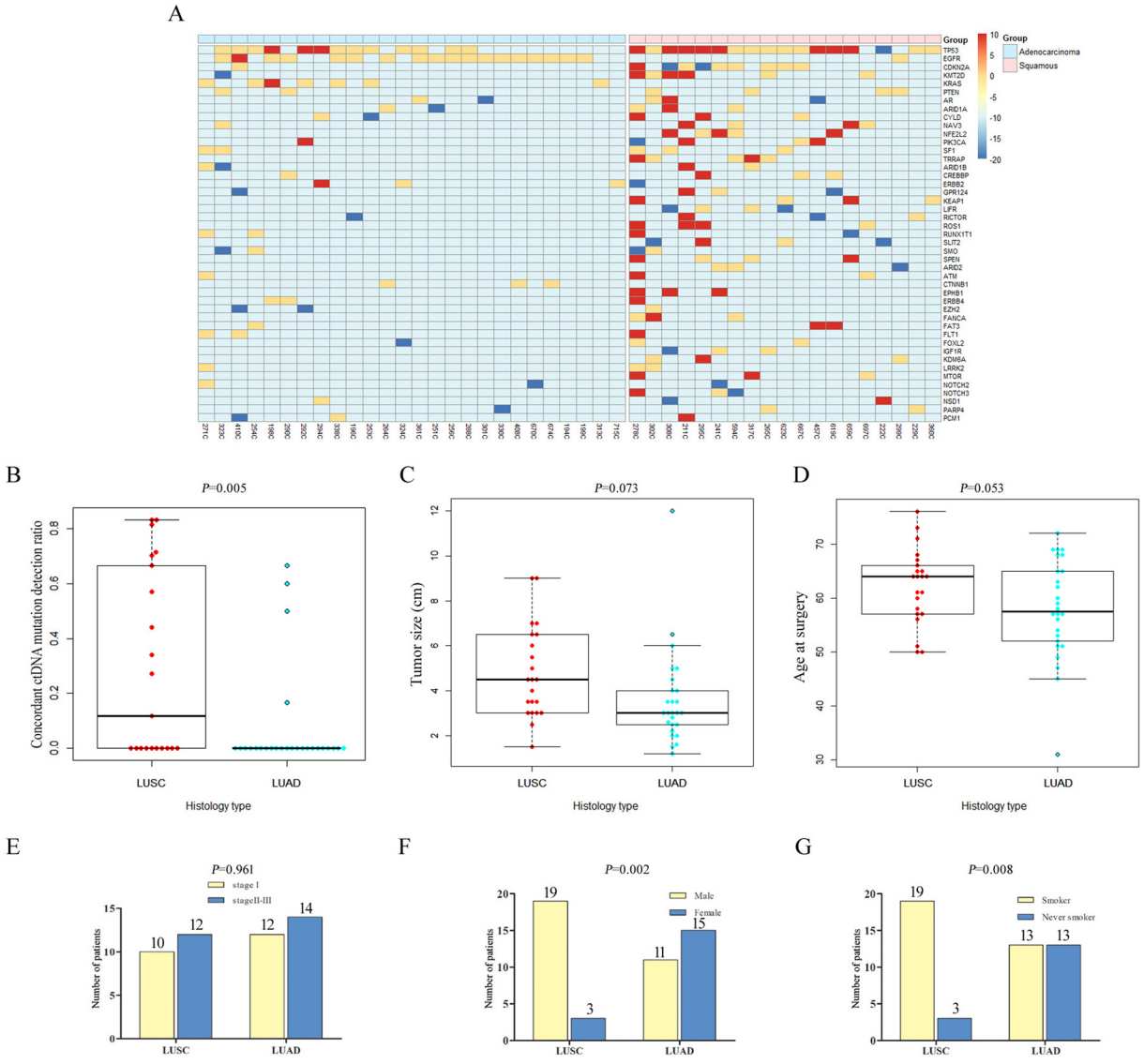

Fig. 2. Summary of gene mutations in matched tDNA and pre-op plasma ctDNA sample pairs.
(A) For 45 patients (each column), altered genes (rows) with mutations from tDNA and ctDNA (pre-op plasma) are shown. The red barcode indicates tissue-plasma concordant gene mutations. The yellow barcode indicates tissue particular gene mutations. The blue barcode indicates pre-op plasma particular gene mutations. (B) Tissue-plasma concordant ctDNA mutation detection ratio was calculated and a significant difference between LUSC and LUAD was found ($P = 0.005$). (C–G) There was no significant difference of tumor size, ages at surgery and tumor stage between LUSC and LUAD patients ($P = 0.073$, $P = 0.053$, $P = 0.961$). There were significant differences of gender and smoking status between LUSC and LUAD patients ($P = 0.002$, $P = 0.008$). The P values were calculated by using the Student's *t*-test or the Chi-square test. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

NSCLC.

Notably, the concentration of cfDNA from post-op plasma was significantly higher than pre-op plasma ($P = 1.436 \times 10^{-6}$, Supplementary Fig. 4, Supplementary Table 8). This effect can be caused by surgery trauma, as described in some studies [26,27]. Some studies have shown post-op ctDNA can identify minimal residual disease and monitor recurrence [28,29]. The time period of posttreatment plasma samples collected was different between in our study and in Aadel A. Chaudhuri's study. In that study, plasma samples were taken every 2–6 months after curative-intent treatment, at which time ctDNA detection would not be affected by surgery. If patients had detectable ctDNA at the time points, it suggested possible minimal residual disease or recurrence. But our plasma samples were collected 2–5 days after complete resection, surgery could influence ctDNA release and ctDNA could not be completely cleared at the time points. Considering the short follow-up time (less than 5 years), small sample size (n = 25), and single sampling time, the survival data was immature in our cohort. Notably, concordant post-op ctDNA mutations were detected in one patient in our cohort, and sternal metastasis was found only one month after surgery in the patient. Therefore, early monitoring post-op ctDNA, multi-point sampling and specific assay panels could be needed for using post-op plasma ctDNA profiling to detect minimal residual disease and relapse or monitor for treatment response.

The average mutation frequency in pre-op ctDNA was lower than that in tDNA, and there were less gene mutations in pre-op ctDNA than in tDNA. Additionally, there were discordant tDNA -pre-op ctDNA gene mutations in 33 sample pairs. This may be caused by lung cancer heterogeneity and the sensitivity of the technology [30–32]. It was reported the mutant allele fractions of ctDNA were higher in metastatic cancer than in earlier stage cancer [33]. Consisting with the observation, we found the mean MAF of ctDNA was higher in patients with stage III than in patients with stage I and II in NSCLC.

It was reported EGFR mutation frequency range was 20–76% in LUAD patients in the Asia-Pacific region [34]. EGFR mutation

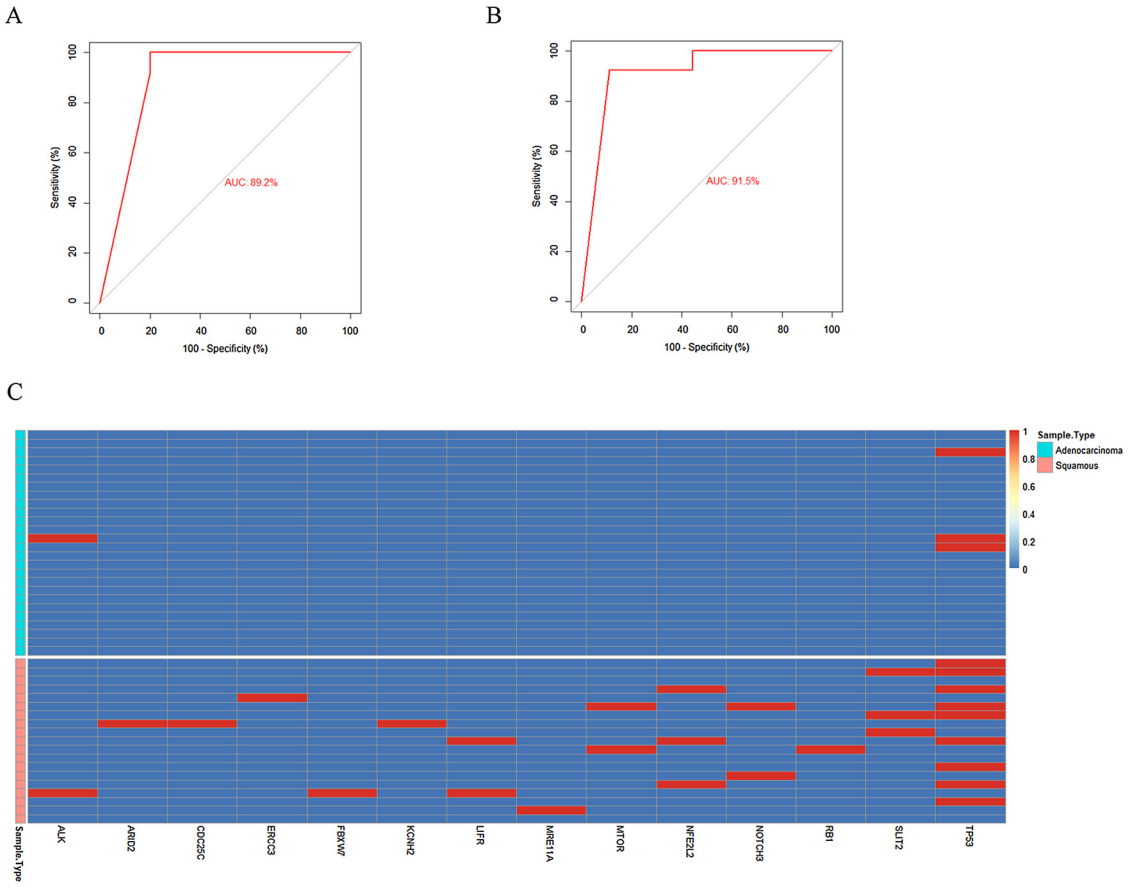

Fig. 3. The distribution of best predictive mutation genes among all pre-op plasma samples.
(A, B) The predictive performance of mutation gene features in training and testing sets with L1 feature selection procedure. The x-axis indicates penalty function Log (lambda) and y-axis indicates mean-squared prediction error PE for each feature groups which calculated numerically. (C) The red barcode indicates a match and a blue indicates a mismatch. The samples above the white line are LUAD samples ($N = 26$), the samples down the white line are LUSC samples ($N = 19$). It could observe that 12 features were observed in only LUSC cohort. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

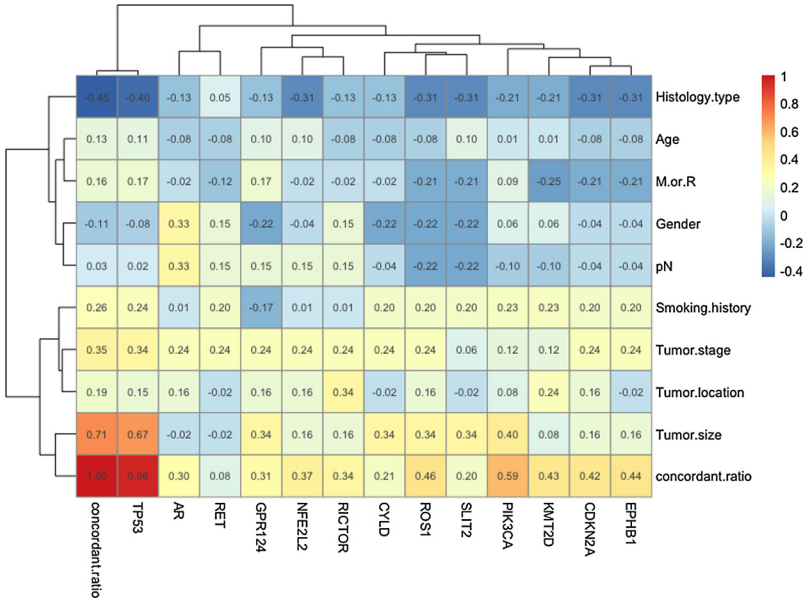

Fig. 4. Pearson correlation analysis between gene mutations and clinical features.
The cut-off values for Pearson correlation coefficient was set at 0.40, Pearson's $r \geq 0.40$ indicates positive correlation, and Pearson's $r \leq -0.40$ indicates negative correlation. M or R: metastasis or recurrence.

frequency was 69.2% (18/26) in our LUAD tissue samples, higher than the reported 47%, but was within the range. And our sample size (n = 26, LUAD) was so small that could not represent the EGFR mutation rate in whole population. However, the detection of EGFR mutations in pre-op plasma ctDNA was quite low in our cohort. We verified our NGS data by testing EGFR, PIK3CA and TP53 hotspot mutations using ddPCR method in 23 samples. The ddPCR showed good concordant results with NGS. Most of the MAFs detected by the ddPCR approach were comparable to those detected by NGS. The EGFR mutations not detected in ctDNA by NGS were also not detected by ddPCR, suggesting it was not as relative low sensitivity of NGS assay. It could be attributed to low ctDNA input because total cfDNA were extracted from 2 ml plasma. ctDNA often accounts for a small fraction of (< 1%) total cfDNA in early stage lung cancer [11]. And, we find most researches use small panel including dozes of genes to detect plasma EGFR mutations. While we used 546-genes panel in the study. We think it may not appropriate to use big panel to detect driver gene mutations from plasma.

Notably, we found the tissue-pre-op plasma concordant ctDNA mutation detection ratio in LUSC was much higher than that in LUAD. Extensive intratumor heterogeneity, the same tumor containing many different subpopulations of cells, is presented in both LUSC and LUAD. Single tumor region analysis cannot cover all gene mutations. But LUSC was reported to carry more clonal mutations which were thought to be present in all cancer cells than did LUAD [35]. So, single-region test from LUSC could have less effect on the tDNA-ctDNA concordant mutation detection ratio than that from LUAD. Additionally, tumor shed DNA fragments containing tumor-specific mutations into the circulation. The release of ctDNA from apoptotic or necrotic tumor cells is related to tumor location, size and vascularity. As we know, tumor location of LUSC is different from LUAD. The majority of LUSC is central type and most LUAD is peripheral type. And, abundant necrotic cells are often observed in LUSC not in LUAD [36]. We indeed observed a lot of necrotic cells in LUSC with high tDNA-pre-op ctDNA concordant mutation detection ratio in our cohort (Supplementary Figure 5). This result is consistent with the most recent report that ctDNA detection is associated with histological subtype [29]. These findings suggest not only the difference of ctDNA release between LUSC and LUAD, but also a potential superior application of ctDNA as a feasible biomarker in LUSC.

The difference of tissue-plasma concordant ctDNA mutation pools between LUSC and LUAD implicated that a prediction model could be built to distinguish LUSC and LUAD by ctDNA sequencing. We used a LASSO based approach to build the LUSC-LUAD classification models. The study implies that mutated genes in ctDNA are largely distinct in LUSC and LUAD. The model included 14 gene mutations. If not less than 3 of the 14 gene mutations, except TP53 and ALK mutations, are detected in one sample, it is highly possible this subtype is LUSC. Interestingly, mutations in LUSC were more similar to other squamous carcinoma such as head and neck squamous cell carcinoma than to mutations in LUAD. These results supported the potential use of ctDNA genotyping. Moreover, this analysis might aid in the histological classification of NSCLC especially in cases where immunohistochemistry does not provide a clear-cut definition.

Pearson correlation analysis also demonstrated that tissue-pre-op plasma concordant ctDNA mutation detection ratio had a relatively high correlation with tumor histology type. More concordant ctDNA mutations were found in LUSC, confirming the report that non-adenocarcinoma histology was an independent predictor of ctDNA detection in early-stage NSCLC [29]. Consistent with most studies [37,38], the concordant ctDNA mutation detection ratio correlated with tumor size. Moreover, we found that the tDNA-ctDNA concordant ratio was related to TP53, ROS1, PIK3CA, KMT2D, EPHB1, and CDKN2A gene mutations, suggesting the gene mutations could represent clonal somatic mutations of tumor evolution [29]. Also, it is possible that the passive release of ctDNA into the bloodstream more often involves these gene mutations.

There are a few limitations in the study. It is a single center and small sample size analysis. Large sample size studies are needed to validate the findings. Additionally, the median follow-up time is 46.5 (range 6–50) months, relatively short, and 4 patients lost follow-up. Therefore, survival analysis related data are not mature.

5. Conclusions

In general, we found that ctDNA detection is correlated to histology subtype and gene mutations in early-stage NSCLC in the cohort, increasing our knowledge of the characteristics of ctDNA in early-stage NSCLC. In addition, we built the LUSC-LUAD classification models based on ctDNA profile. Our findings indicate the potential application of ctDNA detection by NGS for clinical diagnosis and histologic classification in early-stage NSCLC. It should be noted that these findings must be confirmed in large sample sizes and prospective clinical trials.

Conflicts of interest

The authors have declared no conflicts of interest.

Acknowledgements

This work was supported by National Key Research and Development Program of China (2016YFC0905501,2016YFC0905500); and National Natural Science Foundation of China (81772484, 81472182); and Tianjin Cancer Hospital Clinical Trial Project (C1705); and Guangzhou Science and Technology Project (201400000004-5); and Pearl River Nova Program of Guangzhou (201506010065); and Tianjin Municipal Science and Technology Special Funds for Enterprise Development (14ZXLJSY00320).

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.lungcan.2019.05.034.

References

[1] R.L. Siegel, K.D. Miller, A. Jemal, Cancer Statistics, CA. Cancer J. Clin. 67 (1) (2017) 7–30 2017.
[2] J.G. Paez, P.A. Janne, J.C. Lee, S. Tracy, H. Greulich, S. Gabriel, P. Herman, F.J. Kaye, N. Lindeman, T.J. Boggon, K. Naoki, H. Sasaki, Y. Fujii, M.J. Eck, W.R. Sellers, B.E. Johnson, M. Meyerson, EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy, Science 304 (5676) (2004) 1497–1500.
[3] E.L. Kwak, Y.J. Bang, D.R. Camidge, A.T. Shaw, B. Solomon, R.G. Maki, S.H. Ou, B.J. Dezube, P.A. Janne, D.B. Costa, M. Varella-Garcia, W.H. Kim, T.J. Lynch, P. Fidias, H. Stubbs, J.A. Engelman, L.V. Sequist, W. Tan, L. Gandhi, M. Mino-Kenudson, G.C. Wei, S.M. Shreeve, M.J. Ratain, J. Settleman, J.G. Christensen, D.A. Haber, K. Wilner, R. Salgia, G.I. Shapiro, J.W. Clark, A.J. Iafrate, Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer, N. Engl. J. Med. 363 (18) (2010) 1693–1703.
[4] A.T. Shaw, D.W. Kim, K. Nakagawa, T. Seto, L. Crino, M.J. Ahn, T. De Pas, B. Besse, B.J. Solomon, F. Blackhall, Y.L. Wu, M. Thomas, K.J. O'Byrne, D. Moro-Sibilot, D.R. Camidge, T. Mok, V. Hirsh, G.J. Riely, S. Iyer, V. Tassell, A. Polli, K.D. Wilner, P.A. Janne, Crizotinib versus chemotherapy in advanced ALK-positive lung cancer, N. Engl. J. Med. 368 (25) (2013) 2385–2394.
[5] A.T. Shaw, S.H. Ou, Y.J. Bang, D.R. Camidge, B.J. Solomon, R. Salgia, G.J. Riely, M. Varella-Garcia, G.I. Shapiro, D.B. Costa, R.C. Doebele, L.P. Le, Z. Zheng, W. Tan, P. Stephenson, S.M. Shreeve, L.M. Tye, J.G. Christensen, K.D. Wilner, J.W. Clark, A.J. Iafrate, Crizotinib in ROS1-rearranged non-small-cell lung cancer, N. Engl. J. Med. 371 (21) (2014) 1963–1971.
[6] N. Cancer genome atlas research, comprehensive genomic characterization of squamous cell lung cancers, Nature 489 (7417) (2012) 519–525.
[7] N. Cancer genome atlas research, comprehensive molecular profiling of lung adenocarcinoma, Nature 511 (7511) (2014) 543–550.
[8] S. Jahr, H. Hentze, S. Englisch, D. Hardt, F.O. Fackelmayer, R.D. Hesch, R. Knippers, DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells, Cancer Res. 61 (4) (2001) 1659–1665.
[9] M. Stroun, J. Lyautey, C. Lederrey, A. Olson-Sand, P. Anker, About the possible origin and mechanism of circulating DNA apoptosis and active DNA release, Clin. Chim. Acta 313 (1-2) (2001) 139–142.
[10] C. Bettegowda, M. Sausen, R.J. Leary, I. Kinde, Y. Wang, N. Agrawal, B.R. Bartlett,

H. Wang, B. Luber, R.M. Alani, E.S. Antonarakis, N.S. Azad, A. Bardelli, H. Brem, J.L. Cameron, C.C. Lee, L.A. Fecher, G.L. Gallia, P. Gibbs, D. Le, R.L. Giuntoli, M. Goggins, M.D. Hogarty, M. Holdhoff, S.M. Hong, Y. Jiao, H.H. Juhl, J.J. Kim, G. Siravegna, D.A. Laheru, C. Lauricella, M. Lim, E.J. Lipson, S.K. Marie, G.J. Netto, K.S. Oliner, A. Olivi, L. Olsson, G.J. Riggins, A. Sartore-Bianchi, K. Schmidt, M. Shih l, S.M. Oba-Shinjo, S. Siena, D. Theodorescu, J. Tie, T.T. Harkins, S. Veronese, T.L. Wang, J.D. Weingart, C.L. Wolfgang, L.D. Wood, D. Xing, R.H. Hruban, J. Wu, P.J. Allen, C.M. Schmidt, M.A. Choti, V.E. Velculescu, K.W. Kinzler, B. Vogelstein, N. Papadopoulos, L.A. Diaz Jr., Detection of circulating tumor DNA in early- and late-stage human malignancies, Sci. Transl. Med. 6 (224) (2014) 224ra24.

[11] L.A. Diaz Jr., A. Bardelli, Liquid biopsies: genotyping circulating tumor DNA, J. Clin. Oncol. 32 (6) (2014) 579–586.

[12] E. Crowley, F. Di Nicolantonio, F. Loupakis, A. Bardelli, Liquid biopsy: monitoring cancer-genetics in the blood, Nat. Rev. Clin. Oncol. 10 (8) (2013) 472–484.

[13] M. Murtaza, S.J. Dawson, D.W. Tsui, D. Gale, T. Forshew, A.M. Piskorz, C. Parkinson, S.F. Chin, Z. Kingsbury, A.S. Wong, F. Marass, S. Humphray, J. Hadfield, D. Bentley, T.M. Chin, J.D. Brenton, C. Caldas, N. Rosenfeld, Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA, Nature 497 (7447) (2013) 108–112.

[14] I. Garcia-Murillas, G. Schiavon, B. Weigelt, C. Ng, S. Hrebien, R.J. Cutts, M. Cheang, P. Osin, A. Nerurkar, I. Kozarewa, J.A. Garrido, M. Dowsett, J.S. Reis-Filho, I.E. Smith, N.C. Turner, Mutation tracking in circulating tumor DNA predicts relapse in early breast cancer, Sci. Transl. Med. 7 (302) (2015) 302ra133.

[15] N. Karachaliou, C. Mayo-de-Las-Casas, M.A. Molina-Vila, R. Rosell, Real-time liquid biopsies become a reality in cancer treatment, Ann. Transl. Med. 3 (3) (2015) 36.

[16] C. Alix-Panabieres, K. Pantel, Clinical applications of circulating tumor cells and circulating tumor DNA as liquid biopsy, Cancer Discov. 6 (5) (2016) 479–491.

[17] A.M. Aravanis, M. Lee, R.D. Klausner, Next-generation sequencing of circulating tumor DNA for early cancer detection, Cell 168 (4) (2017) 571–574.

[18] S.R. Head, H.K. Komori, S.A. LaMere, T. Whisenant, F. Van Nieuwerburgh, D.R. Salomon, P. Ordoukhanian, Library construction for next-generation sequencing: overviews and challenges, Biotechniques 56 (2) (2014) 61–64 66, 68, passim.

[19] S. Xu, F. Lou, Y. Wu, D.Q. Sun, J.B. Zhang, W. Chen, H. Ye, J.H. Liu, S. Wei, M.Y. Zhao, W.J. Wu, X.X. Su, R. Shi, L. Jones, X.F. Huang, S.Y. Chen, J. Chen, Circulating tumor DNA identified by targeted sequencing in advanced-stage non-small cell lung cancer patients, Cancer Lett. 370 (2) (2016) 324–331.

[20] H. Li, R. Durbin, Fast and accurate long-read alignment with burrows-wheeler transform, Bioinformatics 26 (5) (2010) 589–595.

[21] A. McKenna, M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, M.A. DePristo, The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data, Genome Res. 20 (9) (2010) 1297–1303.

[22] Z. Kan, H. Zheng, X. Liu, S. Li, T.D. Barber, Z. Gong, H. Gao, K. Hao, M.D. Willard, J. Xu, R. Hauptschein, P.A. Rejto, J. Fernandez, G. Wang, Q. Zhang, B. Wang, R. Chen, J. Wang, N.P. Lee, W. Zhou, Z. Lin, Z. Peng, K. Yi, S. Chen, L. Li, X. Fan, J. Yang, R. Ye, J. Ju, K. Wang, H. Estrella, S. Deng, P. Wei, M. Qiu, I.H. Wulur, J. Liu, M.E. Ehsani, C. Zhang, A. Loboda, W.K. Sung, A. Aggarwal, R.T. Poon, S.T. Fan, J. Wang, J. Hardwick, C. Reinhard, H. Dai, Y. Li, J.M. Luk, M. Mao, Whole-genome sequencing identifies recurrent mutations in hepatocellular carcinoma, Genome Res. 23 (9) (2013) 1422–1433.

[23] G. Narzisi, J.A. O'Rawe, I. Iossifov, H. Fang, Y.H. Lee, Z.H. Wang, Y.Y. Wu, G.J. Lyon, M. Wigler, M.C. Schatz, Accurate de novo and transmitted indel detection in exome-capture data using microassembly, Nat. Methods 11 (10) (2014) 1033–1036.

[24] M. Xie, C. Lu, J. Wang, M.D. McLellan, K.J. Johnson, M.C. Wendl, J.F. McMichael, H.K. Schmidt, V. Yellapantula, C.A. Miller, B.A. Ozenberger, J.S. Welch, D.C. Link, M.J. Walter, E.R. Mardis, J.F. Dipersio, F. Chen, R.K. Wilson, T.J. Ley, L. Ding, Age-related mutations associated with clonal hematopoietic expansion and malignancies, Nat. Med. 20 (12) (2014) 1472–1478.

[25] G. Genovese, A.K. Kahler, R.E. Handsaker, J. Lindberg, S.A. Rose, S.F. Bakhoum, K. Chambert, E. Mick, B.M. Neale, M. Fromer, S.M. Purcell, O. Svantesson, M. Landen, M. Hoglund, S. Lehmann, S.B. Gabriel, J.L. Moran, E.S. Lander, P.F. Sullivan, P. Sklar, H. Gronberg, C.M. Hultman, S.A. McCarroll, Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence, N. Engl. J. Med. 371 (26) (2014) 2477–2487.

[26] N. Guo, F. Lou, Y. Ma, J. Li, B. Yang, W. Chen, H. Ye, J.B. Zhang, M.Y. Zhao, W.J. Wu, R. Shi, L. Jones, K.S. Chen, X.F. Huang, S.Y. Chen, Y. Liu, Circulating tumor DNA detection in lung cancer patients before and after surgery, Sci. Rep. 6 (2016) 33519.

[27] F. Diehl, K. Schmidt, M.A. Choti, K. Romans, S. Goodman, M. Li, K. Thornton, N. Agrawal, L. Sokoll, S.A. Szabo, K.W. Kinzler, B. Vogelstein, L.A. Diaz Jr., Circulating mutant DNA to assess tumor dynamics, Nat. Med. 14 (9) (2008) 985–990.

[28] A.A. Chaudhuri, J.J. Chabon, A.F. Lovejoy, A.M. Newman, H. Stehr, T.D. Azad, M.S. Khodadoust, M.S. Esfahani, C.L. Liu, L. Zhou, F. Scherer, D.M. Kurtz, C. Say, J.N. Carter, D.J. Merriott, J.C. Dudley, M.S. Binkley, L. Modlin, S.K. Padda, M.F. Gensheimer, R.B. West, J.B. Shrager, J.W. Neal, H.A. Wakelee, B.W. Loo Jr., A.A. Alizadeh, M. Diehn, Early detection of molecular residual disease in localized lung cancer by circulating tumor DNA profiling, Cancer Discov. 7 (12) (2017) 1394–1403.

[29] C. Abbosh, N.J. Birkbak, G.A. Wilson, M. Jamal-Hanjani, T. Constantin, R. Salari, J. Le Quesne, D.A. Moore, S. Veeriah, R. Rosenthal, T. Marafioti, E. Kirkizlar, T.B.K. Watkins, N. McGranahan, S. Ward, L. Martinson, J. Riley, F. Fraioli, M. Al Bakir, E. Gronroos, F. Zambrana, R. Endozo, W.L. Bi, F.M. Fennessy, N. Sponer, D. Johnson, J. Laycock, S. Shafi, J. Czyzewska-Khan, A. Rowan, T. Chambers, N. Matthews, S. Turajlic, C. Hiley, S.M. Lee, M.D. Forster, T. Ahmad, M. Falzon, E. Borg, D. Lawrence, M. Hayward, S. Kolvekar, N. Panagiotopoulos, S.M. Janes, R. Thakrar, A. Ahmed, F. Blackhall, Y. Summers, D. Hafez, A. Naik, A. Ganguly, S. Kareht, R. Shah, L. Joseph, A. Marie Quinn, P.A. Crosbie, B. Naidu, G. Middleton, G. Langman, S. Trotter, M. Nicolson, H. Remmen, K. Kerr, M. Chetty, L. Gomersall, D.A. Fennell, A. Nakas, S. Rathinam, G. Anand, S. Khan, P. Russell, V. Ezhil, B. Ismail, M. Irvin-Sellers, V. Prakash, J.F. Lester, M. Kornaszewska, R. Attanoos, H. Adams, H. Davies, D. Oukrif, A.U. Akarca, J.A. Hartley, H.L. Lowe, S. Lock, N. Iles, H. Bell, Y. Ngai, G. Elgar, Z. Szallasi, R.F. Schwarz, J. Herrero, A. Stewart, S.A. Quezada, K.S. Peggs, P. Van Loo, C. Dive, C.J. Lin, M. Rabinowitz, H. Aerts, A. Hackshaw, J.A. Shaw, B.G. Zimmermann, T.R. consortium, P. consortium, C. Swanton, Phylogenetic ctDNA analysis depicts early-stage lung cancer evolution, Nature 545 (7655) (2017) 446–451.

[30] M. Gerlinger, A.J. Rowan, S. Horswell, M. Math, J. Larkin, D. Endesfelder, E. Gronroos, P. Martinez, N. Matthews, A. Stewart, P. Tarpey, I. Varela, B. Phillimore, S. Begum, N.Q. McDonald, A. Butler, D. Jones, K. Raine, C. Latimer, C.R. Santos, M. Nohadani, A.C. Eklund, B. Spencer-Dene, G. Clark, L. Pickering, G. Stamp, M. Gore, Z. Szallasi, J. Downward, P.A. Futreal, C. Swanton, Intratumor heterogeneity and branched evolution revealed by multiregion sequencing, N. Engl. J. Med. 366 (10) (2012) 883–892.

[31] M. Jamal-Hanjani, G.A. Wilson, S. Horswell, R. Mitter, O. Sakarya, T. Constantin, R. Salari, E. Kirkizlar, S. Sigurjonsson, R. Pelham, S. Kareht, B. Zimmermann, C. Swanton, Detection of ubiquitous and heterogeneous mutations in cell-free DNA from patients with early-stage non-small-cell lung cancer, Ann. Oncol. 27 (5) (2016) 862–867.

[32] K. Pantel, Blood-based analysis of circulating cell-Free DNA and tumor cells for early cancer detection, PLoS Med. 13 (12) (2016) e1002205.

[33] J. Phallen, M. Sausen, V. Adleff, A. Leal, C. Hruban, J. White, V. Anagnostou, J. Fiksel, S. Cristiano, E. Papp, S. Speir, T. Reinert, M.B.W. Orntoft, B.D. Woodward, D. Murphy, S. Parpart-Li, D. Riley, M. Nesselbush, N. Sengamalay, A. Georgiadis, Q.K. Li, M.R. Madsen, F.V. Mortensen, J. Huiskens, C. Punt, N. van Grieken, R. Fijneman, G. Meijer, H. Husain, R.B. Scharpf, L.A. Diaz, S. Jones, S. Angiuoli, T. Orntoft, H.J. Nielsen, C.L. Andersen, V.E. Velculescu, Direct detection of early-stage cancers using circulating tumor DNA, Sci. Transl. Med. 9 (403) (2017).

[34] A. Midha, S. Dearden, R. McCormack, EGFR mutation incidence in non-small-cell lung cancer of adenocarcinoma histology: a systematic review and global map by ethnicity (mutMapII), Am. J. Cancer Res. 5 (9) (2015) 2892–2911.

[35] M. Jamal-Hanjani, G.A. Wilson, N. McGranahan, N.J. Birkbak, T.B.K. Watkins, S. Veeriah, S. Shafi, D.H. Johnson, R. Mitter, R. Rosenthal, M. Salm, S. Horswell, M. Escudero, N. Matthews, A. Rowan, T. Chambers, D.A. Moore, S. Turajlic, H. Xu, S.M. Lee, M.D. Forster, T. Ahmad, C.T. Hiley, C. Abbosh, M. Falzon, E. Borg, T. Marafioti, D. Lawrence, M. Hayward, S. Kolvekar, N. Panagiotopoulos, S.M. Janes, R. Thakrar, A. Ahmed, F. Blackhall, Y. Summers, R. Shah, L. Joseph, A.M. Quinn, P.A. Crosbie, B. Naidu, G. Middleton, G. Langman, S. Trotter, M. Nicolson, H. Remmen, K. Kerr, M. Chetty, L. Gomersall, D.A. Fennell, A. Nakas, S. Rathinam, G. Anand, S. Khan, P. Russell, V. Ezhil, B. Ismail, M. Irvin-Sellers, V. Prakash, J.F. Lester, M. Kornaszewska, R. Attanoos, H. Adams, H. Davies, S. Dentro, P. Taniere, B. O'Sullivan, H.L. Lowe, J.A. Hartley, N. Iles, H. Bell, Y. Ngai, J.A. Shaw, J. Herrero, Z. Szallasi, R.F. Schwarz, A. Stewart, S.A. Quezada, J. Le Quesne, P. Van Loo, C. Dive, A. Hackshaw, C. Swanton, T.R. Consortium, Tracking the evolution of Non-small-cell lung cancer, N. Engl. J. Med. 376 (22) (2017) 2109–2121.

[36] R. Caruso, A. Parisi, A. Bonanno, D. Paparo, E. Quattrocchi, G. Branca, M. Scardigno, F. Fedele, Histologic coagulative tumour necrosis as a prognostic indicator of aggressiveness in renal, lung, thyroid and colorectal carcinomas: a brief review, Oncol. Lett. 3 (1) (2012) 16–18.

[37] A.M. Newman, S.V. Bratman, J. To, J.F. Wynne, N.C.W. Eclov, L.A. Modlin, C.L. Liu, J.W. Neal, H.A. Wakelee, R.E. Merritt, J.B. Shrager, B.W. Loo, A.A. Alizadeh, M. Diehn, An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage, Nat. Med. 20 (5) (2014) 552–558.

[38] D.A. Haber, V.E. Velculescu, Blood-based analyses of cancer: circulating tumor cells and circulating tumor DNA, Cancer Discov. 4 (6) (2014) 650–661.