Int J Clin Exp Pathol 2020;13(3):525-535
www.ijcep.com /ISSN:1936-2625/IJCEP0107523

Original Article

Identifying a wide range of actionable variants using capture-based ultra-deep targeted sequencing in treatment-naive patients with primary lung adenocarcinoma

Lingfeng Chen[1,2], Minyan Chen[3,4], Jie Lin[1,2], Xiaoyan Chen[1,2], Xunbin Yu[1,2], Zhizhong Chen[1,2], Long Jin[1,2]

[1]Shengli Clinical Medical College, Fujian Medical University, Fuzhou, Fujian Province, China; [2]Department of Pathology, Fujian Provincial Hospital, Fuzhou, Fujian Province, China; Departments of [3]Breast Surgery, [4]General Surgery, Fujian Medical University Union Hospital, Fuzhou, Fujian Province, China

Received January 7, 2020; Accepted February 7, 2020; Epub March 1, 2020; Published March 15, 2020

Abstract: Precision medicine requires accurate multi-gene clinical diagnostics. In current clinical practice, the minimum confidence threshold for variant calling of targeted next-generation sequencing (NGS) on surgical specimens is set to 2%-5%. However, few studies have been conducted to identify a wide range of actionable variants using capture-based ultra-deep targeted sequencing, which has limit of detection (LOD) of 1%. The AmoyDx® Essential NGS panel for capture-based ultra-deep targeted sequencing (dual-indexed sequencing adapters with UMIs) was performed on 372 surgical specimens obtained from treatment-naive patients with primary lung adenocarcinoma, to detect actionable somatic driver mutations associated with each patient. Single-nucleotide variants, insertion/deletion events, and rearrangements were reported. Amplification-refractory mutation system (ARMS) assay and fluorescence in situ hybridization (FISH) were performed for the validation of hotspot mutations in EGFR and ALK, ROS1, and RET fusions. Potentially actionable variants were identified in 80.5% (352/437) of the nonsynonymous variants that were able to be sequenced, and were most commonly found in EGFR mutations (59.7%, 261/437), followed by KRAS mutations (5.5%, 24/437), PIK3CA mutations (3.7%, 16/437), ALK rearrangements (3.4%, 15/437), BRAF mutations (2.7%, 12/437), ERBB2 mutations (2.5%, 11/437), and RET rearrangements (2.3%, 10/437). A total of 7.2% (28/372) of the samples had multiple actionable mutations. Among the 93 triple-negative cases, which did not harbor mutations in EGFR, KRAS, or BRAF, gene fusions were detected in 26 cases (28%). Of the 328 samples, concordance of EGFR between the ARMS assay and NGS was observed in 318 samples (97.0%), and among 32 samples, concordance between ARMS/FISH test and NGS for ALK/ROS1/RET fusion genes was observed in 30 samples (93.8%). Here, we demonstrated that the capture-based ultra-deep targeted sequencing method, which has a LOD of 1% to profile a wide range of actionable variants in surgical specimens of treatment-naive lung adenocarcinoma patients, highlights the need for treatment-naive patients to undergo genomic profiling.

Keywords: Next-generation sequencing, unique molecular identifiers, amplification-refractory mutation system, lung adenocarcinoma, driver gene alterations

Introduction

Lung cancer is the most common cause of cancer-related deaths in China and many other countries. Targeted therapies directed at tumor cells harboring activating epidermal growth factor receptor (*EGFR*) mutations, anaplastic lymphoma kinase (*ALK*) fusions and v-ros UR2 sarcoma virus oncogene homolog 1 (*ROS1*) fusions have revolutionized the landscape of lung adenocarcinoma treatment [1-4]. In addition to *EGFR*, *BRAF*, *ALK*, and *ROS1*, there are several other abnormalities that could potentially be treated with drugs already approved for other malignancies or investigational agents, such as *ERBB2*, *MET*, and *KRAS* mutations and *RET* fusions [5-7].

In current clinical practice, single-gene assays, including, but not limited to, amplification-refractory mutation system (ARMS), Sanger sequencing, and fluorescence in situ hybridiza-

tion (FISH) are often applied in treatment-naive patients to outline driver mutations, such as *ALK* and *ROS1* fusions and *EGFR* mutations, to guide treatment decisions [8, 9]. These patients often undergo several single-gene tests to identify driver mutations. However, each of these conventional techniques is associated with its own disadvantages, including limitations in detecting certain types of aberrations, their low-throughput nature, and low sensitivity. Furthermore, serial testing takes time and depletes tumor tissue.

Single-gene assays have been challenged by more efficient next-generation sequencing (NGS) approaches. NGS allows for large-scale parallel sequencing and has been proven to be an accurate and effective tool for the parallel profiling of a large number of gene alterations including substitution mutations, insertion/deletion mutations, fusions, and amplifications [10]. It also allows the identification of novel mutations that cannot be identified by methods such as ARMS. Targeted NGS enables the generation of reliable data with sufficient sequencing depth in the targeted genes of interest [11]. But in addition to sufficient sequencing depth of coverage, it is also imperative to generate specific and sensitive data because NGS can produce erroneous results secondary to formalin-fixation artifacts [12], chemistry sequencing errors [13], or suboptimal coverage, and/or variant calling [14]. In current practice, the minimum confidence threshold for variant calling of targeted NGS on surgical specimens is set to 2%-5% [15-17].

Advances in bioinformatic field support have led to the development of NGS with high sensitivity, such as unique molecular identifiers (UMIs). Typically, PCR duplicates are identified as sequence reads that align to the same genomic coordinates using reference-based alignment. However, identical molecules can be independently generated during library preparation [18]. Incorporation of UMI adapters can improve accuracy and sensitivity by precisely remove bona fide PCR duplicates [18, 19]. By increasing the depth of sequencing and incorporation of UMI adapters, some less common and low frequency mutations can be discovered.

In this study, we performed capture-based ultra-deep targeted sequencing (dual-indexed sequencing adapters with UMIs) on 372 surgical specimens obtained from treatment-naive patients using the AmoyDx® Essential NGS panel (Amoy Diagnostics, Xiamen, China), which has limit of detection (LOD) of 1%, to identify actionable somatic driver mutations associated with each patient. We report here the sequence findings and validation using an ARMS assay (Amoy Diagnostics, Xiamen, China) and FISH in accordance with the manufacturer's protocol.

Materials and methods

Patient selection and ethics statement

375 surgical specimens of treatment-naive patients with primary lung adenocarcinomas were retrospectively registered in this study between January 2015 and January 2017. The present study was authorized by the Hospital Ethics committee and informed consent was obtained from all patients.

Tissue selection for mutation analysis

Clinical slides (paraffin sections) were reviewed to verify specimen adequacy, tumor content, and purity prior to DNA extraction and downstream testing. An estimate of tumor cell content was made by a diagnostic pathologist, with a requirement of ≥10% for the mutational analysis. Tumor morphology was determined by the diagnostic pathologist. If the DNA content and/or quality was too low for NGS-analysis, the case was excluded from the study.

Overview of test

We used the AmoyDx® Essential NGS panel (Amoy Diagnostics, Xiamen, China), which is approved by the Chinese National Medical Products Administration (NMPA) for qualitative detection of gene mutations in patients with non-small cell lung cancer (NSCLC) or colorectal cancer (CRC). The panel enables capture-based ultra-deep targeted sequencing (dual-indexed sequencing adapters with UMIs) for the following driver genes: EGFR, KRAS, BRAF, NRAS, ERBB2, PIK3CA, MET, ALK, ROS1, and RET (**Table 1** for gene lists and corresponding capture regions). Clinical target-capture sequencing on an Illumina NextSeq 500 was performed with DNA from formalin-fixed, paraffin-embedded (FFPE) tumor tissue. Single-

Table 1. Gene list and corresponding capture regions (with upstream and downstream sequence of each gene) of commercial kit of AmoyDx® Essential NGS panel (Amoy Diagnostics, Xiamen, China)

Gene	Number of coding exons	Detecting variant types
EGFR	17-24	SNVs and indel variants
ALK	20	Fusion, SNVs and indel variants
ROS1	32/34/35/36	Fusion, SNVs and indel variants
RET	8/11/12	Fusion, SNVs and indel variants
BRAF	15	SNVs and indel variants
MET	Intron14	SNVs and indel variants
HER2	20	SNVs and indel variants
KRAS	2/3/4	SNVs and indel variants
NRAS	2/3/4	SNVs and indel variants
PIK3CA	9/14/21	SNVs and indel variants

nucleotide variants (SNV), insertion/deletion events, and rearrangements were reported. This assay was performed before methods were developed to detect copy number variants (CNV) by NGS.

DNA and RNA extraction

DNA and RNA were extracted from scrolls of FFPE tissue comprising ten 10-μm sections using the AmoyDx FFPE DNA/RNA Extraction Kit (Amoy Diagnostics, Xiamen, China). The quality of extracted DNA was assessed using a Nanodrop (Thermo Fisher Scientific, USA) and a Quantus fluorometer (Promega, USA) readings; DNA with the minimum requirements of ≥100 ng total mass by Quantus fluorometry and A260/A280 ratios of 1.7-2.1 was used.

Library generation, enrichment, and sequencing

Extracted DNA (100-150 ng) was fragmented using a Covaris M220 Focused ultrasonicator (Thermo Fisher Scientific, USA) and quality control was performed using an Agilent Bioanalyzer 2100 (Agilent Technologies, USA), followed by fragment screening, end repair, and sequencing adapter ligation. The adapter-ligated DNA was amplified and fragments of 300-500 bp were selected by beads (Agencourt AMPure XP Kit, Beckman Coulter, USA) followed by hybridization with capture probe bait, hybrid selection with magnetic beads, and PCR amplification. Indexed samples were sequenced on a NextSeq 500 instrument (Illumina, USA) to obtain paired-end 150 bp reads.

Variant annotation and reporting

We defined sequencing quality control failure as any base in the specified targeted region sequenced at the unique depth of <500×. As the AmoyDx® Essential NGS panel has a LOD of 1%, variants detected in bases or regions with <500× coverage (unique depth) and/or with variant allele frequency (VAF) <1.0%, were excluded. Potentially actionable variants were classified (**Table 2**) according to the precision oncology knowledge base of OncoKB [20].

Mutational validation techniques

The ARMS assay (Amoy Diagnostics, Xiamen, China) has a LOD of 1% and was used to detect hotspot mutations in EGFR and ALK, ROS1, and RET fusions in the NGS assay validation cohort. ALK FISH analysis was conducted using the Vysis ALK break apart FISH probe (Vysis, Abbot molecular, USA), and RET FISH analysis was conducted using the RET break apart FISH probe (CytoTest, USA) in accordance with the manufacturers' instructions.

Statistical analysis

Descriptive statistics were calculated with Microsoft Office Excel 2010 (Microsoft Corporation, Redmond, WA, USA). The data were described as the number of mutations.

Results

Patient characteristics

Three cases were excluded from this study because the DNA content was too low for NGS-analysis. The characteristics of the 372 lung adenocarcinoma patients are presented in **Table 3**. The median age of patients was 62 years (range, 23-89 years). Overall, 51.1% of patients (190 cases) were male, and 20.7% patients (77/372) were diagnosed in advanced (tumor-node-metastasis III and IV) stages.

Overview of identified nonsynonymous variants

Among the 372 adenocarcinoma specimens, 437 nonsynonymous variants were called within the coding regions of the 10 sequenced

Table 2. Classification of identified variants according to clinical actionability

Variant Level	Significance
1	FDA-recognized biomarker predictive of response to a FDA approved drug in this indication
2	Standard care biomarker predictive of response to a FDA approved drug in this indication
3A	Compelling clinical evidence supports the biomarker as being predictive of response to a drug in this indication
3B	Compelling clinical evidence supports the biomarker as being predictive of response to a drug in another indication
4	Compelling biological evidence supports the biomarker as being predictive of response to a drug
5	Variants with unknown clinical significance/variants that are benign or likely benign

Table 3. Baseline characteristics of all patients (n=372) for which capture-based ultra-deep targeted sequencing was completed

Number of patients, n	372
Diagnosis, n (%)	
Adenocarcinoma	372 (100%)
Age at diagnosis, y (S.D.)	62 (11)
Sex	
Male, n (%)	190 (51.1%)
Female	182 (48.9%)
Stage at diagnosis	
I	235 (63.2%)
II	44 (11.8%)
III	60 (16.1%)
IV	17 (4.6%)
Unknown/undetermined	16 (4.3%)

genes by the standard vendor supplied data analysis pipeline. There was a mean of 1.17 nonsynonymous variants per specimen. A total of 90.3% (336/372) of cases had 1-3 called non-synonymous variants (**Figure 1A**). The nonsynonymous variants included 284 substitutions, 61 deletions, 23 insertions, 43 combined insertion/deletions, and 26 rearrangement events (**Figure 1B**).

Yield of actionable variants

The average coverage of the 372 samples analyzed was 31556× (range 6576×-53080×). The mean unique depth of coverage across the capture region was 4147× (range 525×-7411×). The mean percentage of on-target reads was 70.7% (range, 45.1%-79.3%).

A major objective of NGS in the clinic is to identify samples with potentially actionable variants (defined in **Table 2**). Because targeted therapies are often chosen based on molecular profiles rather than tumor histology, it may be most relevant to consider potentially actionable variants. By this measure, potentially actionable variants were identified in 80.5% (n=352) of the 437 nonsynonymous variants that were able to be sequenced. The most common genetic alterations were EGFR mutations (59.5%, 260/437), followed by KRAS mutations (5.5%, 24/437), PIK3CA mutations (3.7%, 16/437), ALK rearrangements (3.4%, 15/437), BRAF mutations (2.7%, 12/437), ERBB2 mutations (2.5%, 11/437), and RET rearrangements (2.3%, 10/437) (**Figure 1C**), revealing the different clinical characteristics of Chinese patients.

EGFR, KRAS, ERBB2, ALK, and RET oncogenes showed a striking enrichment of specific, well-established, activating variants. For EGFR, exon 19 deletions and p.L858R point mutations constituted 83.5% (218/261) of all detected variants (**Table 4**). Mutations besides exon 19 deletions and p.L858R variants, such as p.G719A (n=2), p.G719S (n=3), p.E709K (n=1), p.A750P (n=9), p.D761Y (n=1), exon 20 insertions (n=13), p.S768I (n=3), p.T790M (n=3), p.L861Q (n=7), and p.L833V (n=1) each represented ≤5% of the total EGFR mutation spectrum.

Within KRAS, 91.7% (22/24) of actionable mutations were at codon 12, and 72.7% (8/11) of actionable mutations were p.A775_G776insYVMA in ERBB2. Within ALK and RET rearrangements, 86.7% (13/15) and 60% (6/10) of actionable mutations were EML4-ALK and KIF5B-RET fusions, respectively (**Table 4**).

Cases of simultaneous p.T790M and p.L858R mutation occurrence in EGFR (n=3) showed a similar VAF in each case (17.9%-37.1% VAF versus 21%-41.8% VAF). EGFR p.R776H and p.R776C mutations showed a similar VAF in each case (8.4%-27.4% VAF versus 8.1%-27.3% VAF for activating mutations) [21].

One rationale for using NGS instead of single-gene testing is the possibility that actionable variants are present in more than one gene or

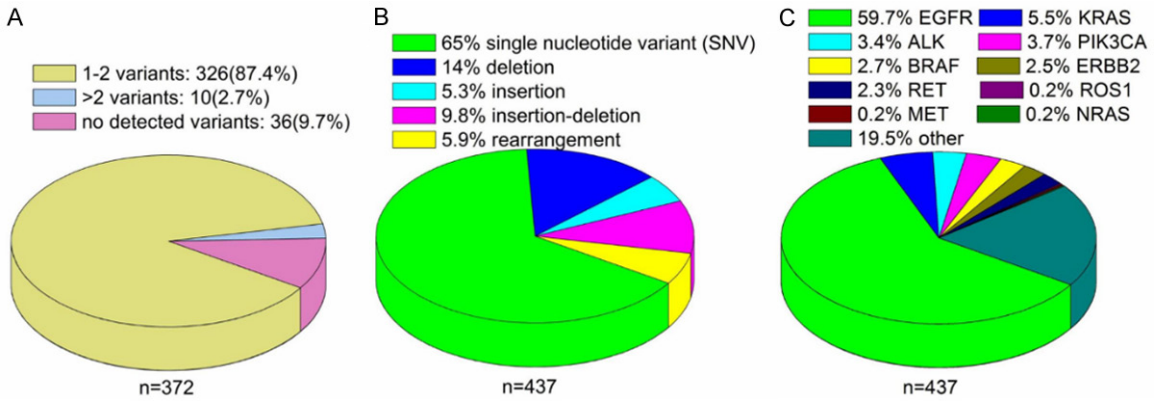

Figure 1. Overview of presumed somatic mutations detected by NGS. A. Number of called variants per sample for 372 lung adenocarcinoma samples. B. Nucleotide-level consequences of nonpolymorphic variants detected by NGS. C. Genes in which potentially actionable mutations were detected.

Table 4. Summary of all potentially actionable variants detected

Number of patients	Adenocarcinoma 372
EGFR	
Exon 18	
G719A	2
G719S	3
E709K	1
Exon 19	
del/delins	93
A750P	9
D761Y	1
Exon 20	
20-Ins	13
S768I	3
T790M	3
Exon 21	
L858R	125
L861Q	7
L833V	1
KRAS	
Codon 12	22
G13D	1
Q61H	1
PIK3CA	
E542K	4
E545K	7
H1047L	1
H1047R	4
BRAF	
V600E	7
G469A	1
D594G	1
D594N	1
L597Q	1
N581S	1
ALK fusion	
EML4-ALK_E6:A20	5
EML4-ALK_E13:A20	8
HIP1-ALK_H19:A20	1
HIP1-ALK_H28:A20	1
RET fusion	
KIF5B-RET_K15:R12	6
CCDC6-RET_C1:R12	3
CCDC6-RET_C8:R12	1
ROS1 fusion	
CD74-ROS1_C6:R34	1
ERBB2	
A775_G776insYVMA	8
G776delinsVC	1
P780_Y781insGSP	1
L755S	1
MET	
Exon 14 splice	1
NRAS	
G13R	1

The number of lung adenocarcinoma patients is stated in the first row. Subsequent rows indicate number of variants, not patients. Mutations are listed as amino acid changes in HGVS protein nomenclature. Non-actionable variants detected in any case were omitted.

pathway, rendering a multigene approach more efficient [22]. A total of 7.2% (28/372) of the samples had multiple actionable mutations, 27 samples had two actionable variants,

Table 5. Cases with more than one actionable variant of two genes

Case	*EGFR* (VAF)	*KRAS* (VAF)	*PIK3CA* (VAF)	*BRAF* (VAF)	Total
71	E746_A750del (45.2%)		E545K (1.2%)		2
115	E746_A750del (53.7%)		E545K (48.4%)		2
171	E746_A750del (19.1%)		H1047R (1.0%)		2
234	E746_A750del (46.7%)		E545K (26.3%)		2
64	L858R (15.3%)		E542K (16.2%)		2
81	L858R (22.4%)		E545K (1.3%)		2
76	L861Q (43.5%)		E545K (18.2%)		2
249	T751_I759delinsN (17.2%)		E542K (2.5%)		3
			H1047R (17.3%)		
256			E542K (1.7%)	N581S (33.9%)	2
340		G12V (2.5%)	E542K (2.5%)		2

The actionable variants are listed as amino acid changes in HGVS protein nomenclature (VAF, variant allele frequency).

Table 6. Cases with two actionable variants of *EGFR* gene

Case	*EGFR* mutation 1 (VAF)	*EGFR* mutation 2 (VAF)
16	L747_A750delinsP (43.4%)	A750P (51.8%)
32	L747_A750delinsP (22.1%)	A750P (25.6%)
50	L747_A750delinsP (4.9%)	A750P (15.6%)
189	L747_A750delinsP (14.1%)	A750P (15.8%)
191	L747_A750delinsP (33.8%)	A750P (35.5%)
245	L747_A750delinsP (10.6%)	A750P (11.3%)
288	L747_A750delinsP (15.8%)	A750P (15.7%)
346	L747_A750delinsP (6.1%)	A750P (5.8%)
17	L747_E749del (19.1%)	A750P (19.0%)
352	L858R (17.0%)	D761Y (17.1%)
39	L858R (42.1%)	L833V (41.8%)
371	L858R (21.5%)	S768I (21.9%)
308	L858R (17.7%)	S768I (16.7%)
132	L858R (31.3%)	T790M (32.3%)
149	L858R (37.1%)	T790M (41.8%)
272	L858R (17.9%)	T790M (21.0%)
179	L861R (8.3%)	G719A (8.2%)
253	E709K (20.2%)	G719S (19.7%)

The actionable variants are listed as amino acid changes in HGVS protein nomenclature (VAF, variant allele frequency).

and one sample had three actionable variants. In 10 samples, these actionable variants were present in more than one gene (**Table 5**). These are samples in which a single-gene approach might have missed therapeutically relevant information. Eighteen samples had two actionable variants of EGFR gene, and except for one sample, all of them had a similar VAF of two actionable variants (**Table 6**).

ALK, RET, and ROS1 gene fusion analysis

Gene fusion analysis was performed on DNA from lung adenocarcinomas in the 372 samples. Among the 93 triple-negative cases, which did not harbor mutations in EGFR, KRAS or BRAF, gene fusions were detected in 26 cases (28%), with 15 (16.1%) ALK gene fusions (eight EML4-ALK_E13:A20 fusions, five EML4-ALK_E6:A20 fusions, one HIP1-ALK_H19:A20, and one HIP1-ALK_H28:A20 fusion), 10 (10.8%) RET fusions (six KIF5B-RET_K15:R12 fusions, three CCDC6-RET_C1:R12 fusions, and one CCDC6-RET_C8:R12 fusion), and one (1.1%) CD74-ROS1_C6:R34 fusion (**Table 4**). A total of 28% cases of the 93 analyzed triple-negative lung adenocarcinomas harbored gene fusions, which was higher than the literature had reported [17, 23].

Mutational validation by ARMS and FISH

For the cases with ≥10% tumor cell content detected by pathological assessment, using capture-based ultra-deep targeted sequencing, we were able to detect and validate known activating driver mutations in EGFR, ALK, ROS1, and RET below 1% VAF by orthogonal methods. Of the 328 samples, concordance of EGFR between the ARMS assay and capture-based ultra-deep targeted sequencing was observed from 318 samples (97.0%) (**Table 7**). Among the 212 EGFR mutation-positive cases detected by the NGS assay, 202 cases (95.3%) yielded concordant results with the ARMS assay. All 10 discordant calls were due to a variant detected by NGS assay but not analyzed by ARMS assay because the ARMS assay

Table 7. Concordance between NGS and ARMS/FISH assay in the validation cohort

		NGS							
		EGFR mut	*EGFR* wt	*ALK* mut	*ALK* wt	*ROS1* mut	*ROS1* wt	*RET* mut	*RET* wt
ARMS or FISH assay	*EGFR* mut	202	0						
	EGFR wt	10*	116						
	ALK mut			14	0				
	ALK wt			1*	6**				
	ROS1 mut					1	0		
	ROS1 wt					0	6**		
	RET mut							9	0
	RET wt							1*	6**

*: detected variant not analyzed by ARMS or FISH assay. **: six same *ALK/ROS1/RET* fusion-negative cases. Wt: no mutation detected at investigated loci. Mut: mutant.

does not cover these mutation sites. Thus, exclusion of these variants implied a concordance of 100% between the NGS and ARMS assays.

Among the six ALK/ROS1/RET fusion-negative cases detected by the NGS assay, six cases yielded concordant results with the ARMS assay. Of the 32 samples, concordance of ALK/ROS1/RET fusion genes between the ARMS/FISH tests and the NGS assay was observed for 30 samples (93.8%) (**Table 7**). Among all the 26 ALK/ROS1/RET fusion-positive cases detected by the NGS assay, 24 cases (92.3%) yielded concordant results with ARMS (n=22), or FISH (n=2). Two discordant calls were due to a variant detected by the NGS assay but were not analyzed by the ARMS assay. Exclusion of these variants implied a concordance of 100%.

Discussion

In the assay described herein, only SNVs, deletions, insertions, and rearrangements were called by NGS, yet CNVs involving ERBB2 and MET, among others, are important in lung adenocarcinoma. This test focuses on potentially actionable variants because validated prognostic and predictive variants fall into this category at present. This work underlines the importance of reviewing submitted samples by pathologists at the time of intake.

In this study, we investigated a cohort of 372 treatment-naive lung adenocarcinoma patients using the AmoyDx® Essential NGS panel, which has limit of detection (LOD) of 1%, and performed a validation by ARMS and FISH. Using the panel and Amoy Diagnostics vendor-supplied bioinformatics pipeline, we were able to detect and validate known activating driver mutations in EGFR and ALK, ROS1, and RET fusion genes below 1% VAF. This study showed that when the LOD is 2%, approximately 1.9% (n=7) of patients (VAF range, 1.03%-1.68%) may miss their treatment opportunity, and when the LOD is 5%, about 7.5% (n=28) patients (VAF range, 1.03%-5.0%) may miss their treatment opportunity.

Potentially actionable variants were identified in 80.5% of the 437 nonsynonymous variants (**Figure 1**). In addition to detecting well-known actionable variants, assays based on complete gene sequencing have the potential to reveal noncanonical variants that may potentially be actionable. Our study demonstrated that targeted sequencing (dual-indexed sequencing adapters with UMIs) allows for large-scale parallel sequencing to accurately detect a wide range of actionable variants. We identified numerous well-established driver mutations that are not covered by commercially available single-gene testing kits, such as ERBB2 (exon 20 insertions, p.L755S), EGFR (c.2571_2573delinsTCG: p.L858R, c.2573_2574delinsGA: p.L858R, p.E709K, p.A750P, p.L833V, p.L861R, p.D761Y), MET c.3082+1G>T (exon-14 splice mutation), and BRAF (p.D594N, p.D594G, p.G469A, p.L597Q, p.N581S). Although the traditional ARMS method is commonly utilized to detect mutations in treatment-naive patients, all of the above mutations are not covered. If tested with commercially available single-gene testing kits, patients harboring such mutations may miss treatment opportunities. Therefore, when multiple genes need to be

Table 8. Placement of patients on matched therapy (16/372 patients placed on matched therapy)

Case	Diagnosis	Mutation	Matched therapy
19	lung adenocarcinoma	*EGFR* L747_T751del	Gefitinib
27	lung adenocarcinoma	*EGFR* L858R	Gefitinib
47	lung adenocarcinoma	*EGFR* L858R	Gefitinib
69	lung adenocarcinoma	*EGFR* E746_A750del	Gefitinib
109	lung adenocarcinoma	*EGFR* L747_P753delinsS	Gefitinib
141	lung adenocarcinoma	*EGFR* E746_A750del	Gefitinib
144	lung adenocarcinoma	*EGFR* L858R	Gefitinib
167	lung adenocarcinoma	*EGFR* L858R	Gefitinib
171	lung adenocarcinoma	*EGFR* E746_A750del *PIK3CA* H1047R	Gefitinib
185	lung adenocarcinoma	*EGFR* E746_A750del	Gefitinib
192	lung adenocarcinoma	*EGFR* H773_V774insH	Gefitinib
216	lung adenocarcinoma	*EGFR* L858R	Gefitinib
220	lung adenocarcinoma	*EGFR* L858R	Gefitinib
260	lung adenocarcinoma	*EML4-ALK*_E13:A20fusion	Crizotinib
285	lung adenocarcinoma	*EGFR* E746_A750del	Gefitinib
287	lung adenocarcinoma	*EGFR* L747_T751del	Gefitinib

The actionable variants are listed as amino acid changes in HGVS protein nomenclature.

tested, capture-based ultra-deep targeted sequencing should be recommended [17].

And race plays a role in the prevalence of certain genetic markers in lung adenocarcinoma [24]. In this research, the low number of detected variants per sample is consistent with similar reports [15, 25]. For the several investigated genes (e.g., EGFR, KRAS, ERBB2) the observed mutation sites and VAF in this Chinese cohort agree with previous reports [26, 27].

Together with a previous report [28], we observed that EGFR T790M mutation and original activating mutation L858R have a similar VAF in all three cases. For BRAF, we observed a general mutation frequency in adenocarcinoma (3.2%), with a 1.9% V600E mutation rate, which is similar with previous reports [29, 30] (1.9% versus ~2%). Despite the individually frequency of many potentially actionable variants defined in the current research is low (e.g., MET, ERBB2, BRAF, and PIK3CA), the high incidence of lung adenocarcinoma means that a large number of patients are affected [29].

A few notable discrepancies in our cohort are apparent. ROS1 gene fusions have been shown to be treatment-predictive for ALK inhibitor drugs [31, 32], and the National Comprehensive Cancer Network (NCCN) Guideline for NSCLC also recommends cabozantinib and vandetanib for patients with RET rearrangements. In the cohort of triple-negative lung adenocarcinoma patients (EGFR, KRAS, and BRAF mutation-negative), the number of cases with RET fusions were similar to that of ALK-positive cases, and specifically support the need for multiple fusion genes diagnostics in lung adenocarcinoma.

Together with several recent reports [15, 22, 25], the clinical validation of commercial capture-based ultra-deep targeted sequencing detected all known insertion/deletions, rearrangements, SNVs, and wild-type loci detected by orthogonal methods (n=360) (**Table 7**), indicating 100% concordance with known variants.

This test has not been available for long enough to allow for comprehensive clinical follow-up, but it will be important to confirm that treatment decisions have been made based on NGS results in the future. In this study, the patients were receiving routine clinical care (**Table 8**).

Although capture-based targeted sequencing has been widely regarded as a powerful tool to accurately detect a wide range of actionable variants, it still has limits [33, 34]. First, the accuracy of new platforms may be lower than conventional methods for the identification of specific mutation types. Second, the capture-based targeted sequencing protocol is commonly longer when compared with conventional methods such as FISH and ARMS. However, we should keep in mind that NGS will continue to be improved and optimized with respect to these disadvantages.

In conclusion, we demonstrate a preferable approach with a LOD of 1% to profile a wide range of actionable variants in surgical specimens of treatment-naive lung adenocarcinoma patients. This approach could further guide

more precise targeted therapy and help lung adenocarcinoma patients to achieve better clinical benefit.

Acknowledgements

This study was supported by The High-level Hospital Foster Grants from Fujian Provincial Hospital of China (grant number 2019HSJJ25), The Fujian Provincial Health Commission Young and Middle-aged Talents Training Project of China (grant number 2018-ZQN-4) and The Fujian Provincial Natural Science Foundation Project of China (grant number 2019J01182).

Disclosure of conflict of interest

None.

Abbreviations

NGS, next generation sequencing; UMI, unique molecular identifier; ARMS, amplification-refractory mutation system; FFPE, formalin-fixed, paraffin-embedded; FISH, fluorescence in situ hybridization; LOD, limit of detection; VAF, variant allele frequency.

Address correspondence to: Dr. Long Jin, Department of Pathology, Fujian Provincial Hospital, No. 134 Dong Street, Fuzhou 350001, Fujian Province, China. Tel: +86-15959090651; E-mail: jinlongdoctor@163.com

References

[1] Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW, Harris PL, Haserlat SM, Supko JG, Haluska FG, Louis DN, Christiani DC, Settleman J and Haber DA. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med 2004; 350: 2129-2139.

[2] Paez JG, Jänne PA, Lee JC, Tracy S, Greulich H, Gabriel S, Herman P, Kaye FJ, Lindeman N, Boggon TJ, Naoki K, Sasaki H, Fujii Y, Eck MJ, Sellers WR, Johnson BE and Meyerson M. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science 2004; 304: 1497-1500.

[3] Kwak EL, Bang YJ, Camidge DR, Shaw AT, Solomon B, Maki RG, Ou SH, Dezube BJ, Jänne PA, Costa DB, Varella-Garcia M, Kim WH, Lynch TJ, Fidias P, Stubbs H, Engelman JA, Sequist LV, Tan W, Gandhi L, Mino-Kenudson M, Wei GC, Shreeve SM, Ratain MJ, Settleman J, Christensen JG, Haber DA, Wilner K, Salgia R, Shapiro GI, Clark JW and Iafrate AJ. Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. N Engl J Med 2010; 363: 1693-1703.

[4] Bang YJ, Ou SHI, Camidge DR, Clark JW, Wilner KD, Tye L, Stephenson P, Varella-Garcia M, Iafrate AJ and Shaw AT. Clinical activity of crizotinib in advanced non-small cell lung cancer (NSCLC) harboring ROS1 gene rearrangement. Journal of Clinical Oncology, 30 (15_suppl), 7508. 10th Annual Meeting of the Japanese-society-of-medical-oncology. 2012.

[5] Farago AF and Azzoli CG. Beyond ALK and ROS1: RET, NTRK, EGFR and BRAF gene rearrangements in non-small cell lung cancer. Transl Lung Cancer Res 2017; 6: 550-559.

[6] Li BT, Shen R, Buonocore D, Olah ZT, Ni A, Ginsberg MS, Ulaner GA, Offin M, Feldman D, Hembrough T, Cecchi F, Schwartz S, Pavlakis N, Clarke S, Won HH, Brzostowski EB, Riely GJ, Solit DB, Hyman DM, Drilon A, Rudin CM, Berger MF, Baselga J, Scaltriti M, Arcila ME and Kris MG. Ado-trastuzumab emtansine for patients with HER2-mutant lung cancers: results from a phase II basket trial. J Clin Oncol 2018; 36: 2532-2537.

[7] Awad MM, Oxnard GR, Jackman DM, Savukoski DO, Hall D, Shivdasani P, Heng JC, Dahlberg SE, Jänne PA, Verma S, Christensen J, Hammerman PS and Sholl LM. MET exon 14 mutations in non-small-cell lung cancer are associated with advanced age and stage-dependent MET genomic amplification and c-Met overexpression. J Clin Oncol 2016; 34: 721-730.

[8] Thress KS, Brant R, Carr TH, Dearden S, Jenkins S, Brown H, Hammett T, Cantarini M and Barrett JC. EGFR mutation detection in ctDNA from NSCLC patient plasma: a cross-platform comparison of leading technologies to support the clinical development of AZD9291. Lung Cancer 2015; 90: 509-515.

[9] Wu YC, Chang IC, Wang CL, Chen TD, Chen YT, Liu HP, Chu Y, Chiu YT, Wu TH, Chou LH, Chen YR and Huang SF. Comparison of IHC, FISH and RT-PCR methods for detection of ALK rearrangements in 312 non-small cell lung cancer patients in Taiwan. PLoS One 2013; 8: e70839.

[10] Lipson D, Capelletti M, Yelensky R, Otto G, Parker A, Jarosz M, Curran JA, Balasubramanian S, Bloom T, Brennan KW, Donahue A, Downing SR, Frampton GM, Garcia L, Juhn F, Mitchell KC, White E, White J, Zwirko Z, Peretz T, Nechushtan H, Soussan-Gutman L, Kim J, Sasaki H, Kim HR, Park SI, Ercan D, Sheehan CE, Ross JS, Cronin MT, Jänne PA and Stephens PJ. Identification of new ALK and RET gene fusions from colorectal and lung cancer biopsies. Nat Med 2012; 18: 382-384.

[11] Han SW, Kim HP, Shin JY, Jeong EG, Lee WC, Lee KH, Won JK, Kim TY, Oh DY, Im SA, Bang YJ, Jeong SY, Park KJ, Park JG, Kang GH, Seo JS, Kim JI and Kim TY. Targeted sequencing of cancer-related genes in colorectal cancer using next-generation sequencing. PLoS One 2013; 8: e64271.

[12] Wong SQ, Li J, Tan AY, Vedururu R, Pang JM, Do H, Ellul J, Doig K, Bell A, MacArthur GA, Fox SB, Thomas DM, Fellowes A, Parisot JP and Dobrovic A; CANCER 2015 Cohort. Sequence artefacts in a prospective series of formalin-fixed tumours tested for mutations in hotspot regions by massively parallel sequencing. BMC Med Genomics 2014; 7: 23.

[13] Singh RR, Patel KP, Routbort MJ, Reddy NG, Barkoh BA, Handal B, Kanagal-Shamanna R, Greaves WO, Medeiros LJ, Aldape KD and Luthra R. Clinical validation of a next-generation sequencing screen for mutational hotspots in 46 cancer-related genes. J Mol Diagn 2013; 15: 607-622.

[14] Spencer DH, Abel HJ, Lockwood CM, Payton JE, Szankasi P, Kelley TW, Kulkarni S, Pfeifer JD and Duncavage EJ. Detection of FLT3 internal tandem duplication in targeted, short-read-length, next-generation sequencing data. J Mol Diagn 2013; 15: 81-93.

[15] Fisher KE, Zhang L, Wang J, Smith GH, Newman S, Schneider TM, Pillai RN, Kudchadkar RR, Owonikoko TK, Ramalingam SS, Lawson DH, Delman KA, El-Rayes BF, Wilson MM, Sullivan HC, Morrison AS, Balci S, Adsay NV, Gal AA, Sica GL, Saxe DF, Mann KP, Hill CE, Khuri FR and Rossi MR. Clinical validation and implementation of a targeted next-generation sequencing assay to detect somatic variants in non-small cell lung, melanoma, and gastrointestinal malignancies. J Mol Diagn 2016; 18: 299-315.

[16] Ma Y, Chen K, Yang Z and Guan M. Targeted sequencing reveals distinct pathogenic variants in Chinese patients with lung adenocarcinoma brain metastases. Oncol Lett 2018; 15: 4503-4510.

[17] Lindquist KE, Karlsson A, Levéen P, Brunnström H, Reuterswärd C, Holm K, Jönsson M, Annersten K, Rosengren F, Jirström K, Kosieradzki J, Ek L, Borg Å, Planck M, Jönsson G and Staaf J. Clinical framework for next generation sequencing based analysis of treatment predictive mutations and multiplexed gene fusion detection in non-small cell lung cancer. Oncotarget 2017; 8: 34796-34810.

[18] Hong J and Gresham D. Incorporation of unique molecular identifiers in TruSeq adapters improves the accuracy of quantitative sequencing. Biotechniques 2017; 63: 221-226.

[19] MacConaill LE, Burns RT, Nag A, Coleman HA, Slevin MK, Giorda K, Light M, Lai K, Jarosz M, McNeill MS, Ducar MD, Meyerson M and Thorner AR. Unique, dual-indexed sequencing adapters with UMIs effectively eliminate index cross-talk and significantly improve sensitivity of massively parallel sequencing. BMC Genomics 2018; 19: 30.

[20] Chakravarty D, Gao J, Phillips SM, Kundra R, Zhang H, Wang J, Rudolph JE, Yaeger R, Soumerai T, Nissan MH, Chang MT, Chandarlapaty S, Traina TA, Paik PK, Ho AL, Hantash FM, Grupe A, Baxi SS, Callahan MK, Snyder A, Chi P, Danila D, Gounder M, Harding JJ, Hellmann MD, Iyer G, Janjigian Y, Kaley T, Levine DA, Lowery M, Omuro A, Postow MA, Rathkopf D, Shoushtari AN, Shukla N, Voss M, Paraiso E, Zehir A, Berger MF, Taylor BS, Saltz LB, Riely GJ, Ladanyi M, Hyman DM, Baselga J, Sabbatini P, Solit DB and Schultz N. OncoKB: a precision oncology knowledge base. JCO Precis Oncol 2017; 2017.

[21] van Noesel J, van der Ven WH, van Os TA, Kunst PW, Weegenaar J, Reinten RJ, Kancha RK, Duyster J and van Noesel CJ. Activating germline R776H mutation in the epidermal growth factor receptor associated with lung cancer with squamous differentiation. J Clin Oncol 2013; 31: e161-164.

[22] Hagemann IS, Devarakonda S, Lockwood CM, Spencer DH, Guebert K, Bredemeyer AJ, Al-Kateb H, Nguyen TT, Duncavage EJ, Cottrell CE, Kulkarni S, Nagarajan R, Seibert K, Baggstrom M, Waqar SN, Pfeifer JD, Morgensztern D and Govindan R. Clinical next-generation sequencing in patients with non-small cell lung cancer. Cancer 2015; 121: 631-639.

[23] Takeuchi K, Soda M, Togashi Y, Suzuki R, Sakata S, Hatano S, Asaka R, Hamanaka W, Ninomiya H, Uehara H, Lim Choi Y, Satoh Y, Okumura S, Nakagawa K, Mano H and Ishikawa Y. RET, ROS1 and ALK fusions in lung cancer. Nat Med 2012; 18: 378-381.

[24] Midha A, Dearden S and McCormack R. 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 2015; 5: 2892-2911.

[25] Mäki-Nevala S, Sarhadi VK, Rönty M, Kettunen E, Husgafvel-Pursiainen K, Wolff H, Knuuttila A and Knuutila S. Hot spot mutations in Finnish non-small cell lung cancers. Lung Cancer 2016; 99: 102-110.

[26] Feng H, Wang X, Zhang Z, Tang C, Ye H, Jones L, Lou F, Zhang D, Jiang S, Sun H, Dong H, Zhang G, Liu Z, Dong Z, Guo B, Yan H, Yan C, Wang L, Su Z, Li Y, Nandakumar V, Huang XF, Chen SY and Liu D. Identification of genetic

mutations in human lung cancer by targeted sequencing. Cancer Inform 2015; 14: 83-93.
[27] Feng Y, Feng G, Lu X, Qian W, Ye J, Manrique CA, Ma C and Lu Y; written on behalf of the AME Lung Cancer Collaborative Group. Exploratory analysis of introducing next-generation sequencing-based method to treatment-naive lung cancer patients. J Thorac Dis 2018; 10: 5904-5912.
[28] Lou Y, Pecot CV, Tran HT, DeVito VJ, Tang XM, Heymach JV, Luthra R, Wistuba II, Zuo Z and Tsao AS. Germline mutation of T790M and dual/multiple EGFR mutations in patients with lung adenocarcinoma. Clin Lung Cancer 2016; 17: e5-11.
[29] Barlesi F, Mazieres J, Merlio JP, Debieuvre D, Mosser J, Lena H, Ouafik L, Besse B, Rouquette I, Westeel V, Escande F, Monnet I, Lemoine A, Veillon R, Blons H, Audigier-Valette C, Bringuier PP, Lamy R, Beau-Faller M, Pujol JL, Sabourin JC, Penault-Llorca F, Denis MG, Lantuejoul S, Morin F, Tran Q, Missy P, Langlais A, Milleron B, Cadranel J, Soria JC and Zalcman G; Biomarkers France contributors. Routine molecular profiling of patients with advanced non-small-cell lung cancer: results of a 1-year nationwide programme of the french cooperative thoracic intergroup (IFCT). Lancet 2016; 387: 1415-1426.
[30] Tissot C, Couraud S, Tanguy R, Bringuier PP, Girard N and Souquet PJ. Clinical characteristics and outcome of patients with lung cancer harboring BRAF mutations. Lung Cancer 2016; 91: 23-28.
[31] Shaw AT, Ou SH, Bang YJ, Camidge DR, Solomon BJ, Salgia R, Riely GJ, Varella-Garcia M, Shapiro GI, Costa DB, Doebele RC, Le LP, Zheng Z, Tan W, Stephenson P, Shreeve SM, Tye LM, Christensen JG, Wilner KD, Clark JW and Iafrate AJ. Crizotinib in ROS1-rearranged non-small-cell lung cancer. N Engl J Med 2014; 371: 1963-1971.
[32] Drilon A, Rekhtman N, Arcila M, Wang L, Ni A, Albano M, Van Voorthuysen M, Somwar R, Smith RS, Montecalvo J, Plodkowski A, Ginsberg MS, Riely GJ, Rudin CM, Ladanyi M and Kris MG. Cabozantinib in patients with advanced RET-rearranged non-small-cell lung cancer: an open-label, single-centre, phase 2, single-arm trial. Lancet Oncol 2016; 17: 1653-1660.
[33] Goodwin S, McPherson JD and McCombie WR. Coming of age: ten years of next-generation sequencing technologies. Nat Rev Genet 2016; 17: 333-351.
[34] Xuan J, Yu Y, Qing T, Guo L and Shi L. Next-generation sequencing in the clinic: promises and challenges. Cancer Lett 2013; 340: 284-295.