Int J Clin Exp Pathol 2016;9(12):12876-12882
www.ijcep.com /ISSN:1936-2625/IJCEP0038564

Original Article

Mutations in NSCLC identified by a next-generation sequencing targeted sequencing panel

Min Li[1], Lei Li[2], Mingzhu Wang[1], Hongjun Ma[2], Renya Zhang[2], Anna Zhu[3], Minying Sun[1], Zhenhua Chen[1], Yingsong Wu[1,3], Xuexi Yang[1,3], Ming Li[1,3]

[1]Institute of Antibody Engineering, School of Biotechnology, Southern Medical University, Guangzhou, China; [2]Department of Pathology, Affiliated Hospital of Jining Medical University, Shandong, China; [3]Guangzhou Darui Biotechnology Co. LTD, Guangzhou, China

Received August 23, 2016; Accepted October 10, 2016; Epub December 1, 2016; Published December 15, 2016

Abstract: Next-generation sequencing (NGS) has become a cost-effective approach to screening for a number of genes simultaneously in clinical use. The purpose of the present study is to screening for known mutations of cancer drug target genes in patients with non-small cell lung cancer (NSCLC) using an NGS targeted sequencing approach. Genomic DNA was extracted from sections of formalin-fixed paraffin-embedded (FFPE) tissue samples of 58 NSCLC patients. The Ion AmpliSeq Colon and Lung Cancer Research Panel v2 was used to screen. ARMS-PCR was used to validate the NGS results. NGS targeted sequencing revealed that 44 (75.9%) of the 58 carried mutations in 7 genes. Higher mutation frequencies were found in the EGFR (43.1%), TP53 (37.9%) and KRAS (12.1%) genes, and lower frequencies in the CTNNB1 (3.4%), NRAS (3.4%), SMAD4 (3.4%) and PIK3CA (1.7%) genes. A higher mutation rate was found in adenocarcinoma (ADK) samples compared to squamous cell carcinoma (SCC) samples, but the difference was not significant (P = 0.34). Female patients showed a significantly higher mutation rate in the EGFR gene than male patients (P = 0.03). A significantly higher number of EGFR mutations was observed in ADK samples compared to that in SCC samples (P = 0.02). The findings of the present study indicate that NGS targeted sequencing can effectively detect mutations in clinical tumor samples, thereby highlighting that NGS is a promising tool in personalized medicine.

Keywords: Next-generation sequencing, targeted sequencing, non-small cell lung cancer, personalized medicine

Introduction

Non-small cell lung cancer (NSCLC) is the leading cause of cancer-related death, accounting for 85%-90% of lung cancers, with a poor five-year survival rate of less than 20% [1]. The molecular basis of lung cancer is complex and heterogeneous. The underlying molecular mechanisms between the two major NSCLC histologic subtypes, namely, adenocarcinoma (ADK) and squamous cell carcinoma (SCC) are distinct [2]. For example, 25% of patients with ADK have mutations in the v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog (KRAS) gene, whereas these are extremely rare in SCCs [3]. In contrast, 35%-45% of patients with SCC carry mutations in the phosphoinositide-3-kinase (PI3K) or PTEN gene, but only 5%-10% of ADKs have mutations in these genes [4, 5]. Similar to other malignancies, tumorigenesis in lung cancer involves the activation of growth promoting genes such as EGFR, KRAS, BRAF, MEK1, HER2, MET, ALK, and RET as well as inactivation of tumor suppressor genes such as P53, PTEN and LKB1 [6].

Approximately 15%-20% NSCLC patients may benefit from personalized medicine based on the genetic background of the tumor [7]. Various drugs have been approved for the treatment of NSCLC patients who carry well-defined driver mutations such as EGFR mutations and ALK rearrangements [8-11]. Other driver mutations include alterations in the TP53, KRAS, LKB1, HER2, BRAF, MEK, ERK and AKT1 genes, as well as RET fusions [3, 11-13], and clinical trials have been designed to evaluate several compounds that are relevant to these mutations [14].

Table 1. Demographic and clinical features of NSCLC patients

Patient	Gender	Age	Histology	Stage	Family history	Smoking habits (years)
1	M	69	ADK	II	N	40
2	M	70	ADK	IV	N	50
3	F	56	ADK	IV	N	NS
4	F	55	ADK	II	N	NS
5	F	68	ADK	IV	N	NS
6	M	61	ADK	IV	N	30
7	M	78	ADK	II	N	40
8	M	71	ADK	III	N	40
9	M	61	ADK	IV	N	40
10	F	72	ADK	II	N	NS
11	F	77	ADK	IV	N	NS
12	M	56	ADK	II	N	NS
13	M	64	ADK	III	N	30
14	F	63	ADK	III	Y	NS
15	M	74	ADK	III	Y	55
16	M	88	ADK	II	N	60
17	M	45	ADK	IV	N	20
18	F	58	ADK	NA	N	20
19	F	58	ADK	IV	N	NS
20	F	25	ADK	I	N	NS
21	M	64	ADK	II	N	30
22	F	46	ADK	III	Y	NS
23	F	48	ADK	IV	N	NS
24	M	63	ADK	IV	N	30
25	F	46	ADK	IV	N	NS
26	F	68	ADK	IV	N	NS
27	M	64	SCC	III	N	NS
28	F	46	ADK	IV	Y	NS
29	M	67	SCC	IV	N	NS
30	M	50	ADK	III	N	20
31	M	68	SCC	IV	N	30
32	F	65	ADK	NA	N	NS
33	M	69	ADK	II	N	NS
34	M	74	ADK	II	N	40
35	M	74	ADK	IV	N	NS
35	M	65	SCC	IV	N	30
37	M	21	ADK	NA	N	5
38	M	46	ADK	IV	N	30
39	F	45	ADK	IV	N	NS
40	F	74	ADK	IV	N	NS
41	F	60	ADK	IV	N	NS
42	F	74	ADK	IV	N	NS
43	F	49	SCC	II	N	NS
44	M	61	ADK	IV	N	40
45	M	66	ADK	NA	N	40
46	F	52	ADK	III	N	NS

The advent of next-generation sequencing (NGS) has revolutionized the approach in detecting gene mutations. In particular, an optimized NGS workflow has been established for analyzing 22 lung cancer-related genes in critical samples such as DNA from formalin-fixed paraffin-embedded (FFPE) tissues and circulating free DNA (cfDNA) [15]. Detecting mutations in lung cancer-related genes from NSCLC patients based on the NGS targeted sequencing has been reported in various ethnic populations [11, 16-19]. However, investigations involving Chinese patients remain limited despite the high incidence of lung cancer and the leading cause of cancer-related death in China [20].

In the present study, we investigated 58 NSCLC patients using a NGS targeted sequencing panel that included 1,850 targeted sites in 22 cancer drug target genes. Our NGS analyses revealed that 75.9% of NSCLC patients carried mutations in 7 of these genes. These results suggest that NGS targeted sequencing can effectively detect mutations in clinical tumor samples, thereby highlighting this technology as a promising tool in personalized medicine.

Materials and methods

Patients

Fifty-eight NSCLC patients were recruited from the Affiliated Hospital of Jining Medical College in Jining, Shandong Province, China. The Ethical and Protocol Review Committee of Jining Medical College approved the study protocol. Written informed consent was obtained from all patients. The demographic and clinical features of the patients are shown in **Table 1**. There were 33 (56.9%) male patients and 25 (43.1%) female patients. The patients' average age was 60.7 ± 12.6 years, ranging from 21 to 88 years. Tissue samples were collected from these patients, including 41 lung biopsies, 11 metastases, 5 gross specimens, and 1 pleural effusion. There were 51 patients with ADK and 7

47	F	64	ADK	IV	N	NS
48	M	67	ADK	II	N	NS
49	M	46	ADK	IV	N	30
50	M	47	ADK	II	N	30
51	F	52	ADK	IV	N	NS
52	M	83	SCC	IV	Y	50
53	M	60	ADK	IV	N	30
54	F	73	ADK	III	N	NS
55	M	66	ADK	III	N	NS
56	M	53	SCC	II	N	30
57	M	57	ADK	NA	N	30
58	F	60	ADK	IV	Y	NS

Abbreviations: NSCLC: Non-small cell lung cancer; M: Male; F: Female; ADK: Adenocarcinoma; SCC: Squamous cell carcinoma; N: No; Y: Yes; NS: Never smoker; NA: Not available.

patients with SCC. One (1.7%) patient was diagnosed as stage I, 13 (22.4%) patients were diagnosed as stage II, 10 (17.2%) patients were diagnosed as stage III, and 29 (50.0%) patients were diagnosed as stage IV, whereas the stages of 5 (8.6%) patients were not determined. Six (10.3%) patients had a family history of lung cancer, whereas 52 (89.7%) had no family history. Twenty-seven (46.6%) patients had a smoking history (34.1 ± 11.6 years, ranging from 5 to 60 years), whereas 31 (53.4%) had none. Informed consent was obtained from all individual participants included in the study.

DNA preparation and Ion torrent PGM library preparation

Genomic DNA was extracted from sections of FFPE tissue samples using the QIAamp DNA Mini Kit (Qiagen, Valencia, CA, USA) according to the manufacturer's instructions. The Ion AmpliSeq Library Kit 2.0 (Life Technologies, Carlsbad, CA, USA) was used to construct an adapter-ligated library following the manufacturer's protocol and previous reports [21]. Then, quality control of libraries was performed with Qubit® 2.0 and a 2100 Bioanalyzer (Agilent Technologies, PaloAlto, CA, USA). There were a multiplexed of 16 libraries consisting of 50 pM prepared library per sample then amplified by emulsion PCR on Ion PI™ Ion Sphere™ Particles (ISPs) with the Ion One Touch™ 2 Instrument (Life Technologies, CA, USA). Template-positive ISPs were enriched (Ion One Touch™ ES Instrument), loaded onto an Ion 318™ Chip v2 (Ion PI™ Sequencing 200 Kit v3, Life Technologies, Carlsbad, CA, USA) and runned on PGM (Life Technologies, Carlsbad, CA, USA).

NGS and data analysis

Based on the Ion AmpliSeq Colon and Lung Cancer Research Panel v2 (Life Technologies), 22 tumor targeting drug-related genes were selected for the NGS targeted sequencing from the Oncomine Database, including ALK, AKT1, BRAF, CTNNB1, DDR2, EGFR, ERBB2, ERBB4, FBXW7, FGFR1, FGFR2, FGFR3, KRAS, MAP2K1, MET, NOTCH1, NRAS, PIK3CA, PTEN, SMAD4, STK11, and TP53. A total of 1,850 mutations in these genes were covered in the targeted sequencing. Targeted sequencing was performed on the Ion PGM platform (Life Technologies). The detailed methods of Ion Ampliseq Cancer Panel sequencing have been previously described [15, 19, 22].

NGS data analysis was performed with the Ion Reporter software (Life Technologies). Hotspot and targeted regions, together with the json parameters files associated with the Ion AmpliSeq Colon and Lung Cancer Panel v2 (Life Technologies), were loaded into the Variant Caller plugin. All identified variants were visually confirmed by using the Integrative Genomics viewer (IGV) [23]. To predict the effect of missense mutations on the corresponding protein and to calculate their conservation scores, variants were annotated using PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/), SIFT (http://sift.jcvi.org/), and Genomic Evolutionary Rate Profiling (GERP; http://www.broadinstitute.org/~mgarber/GERP/documentation.pdf) scores. Variants were named according to the Catalogue of Somatic Mutation in Cancer (COSMIC; http://cancer.sanger.ac.uk/cosmic) and Single Nucleotide Polymorphism Database (dbSNP; http://www.ncbi.nlm.nih.gov/snp/). Population frequencies of the variants were obtained from the NHLBI Exome Sequencing Project (http://evs.gs.washington.edu/EVS/).

Amplification refractory mutation system (ARMS)-PCR

To validate the mutations detected by NGS targeted sequencing, all patients were screened for 29 known EGFR mutations by using ARMS-PCR as previously described (Supplementary Tables 1, 2) [24].

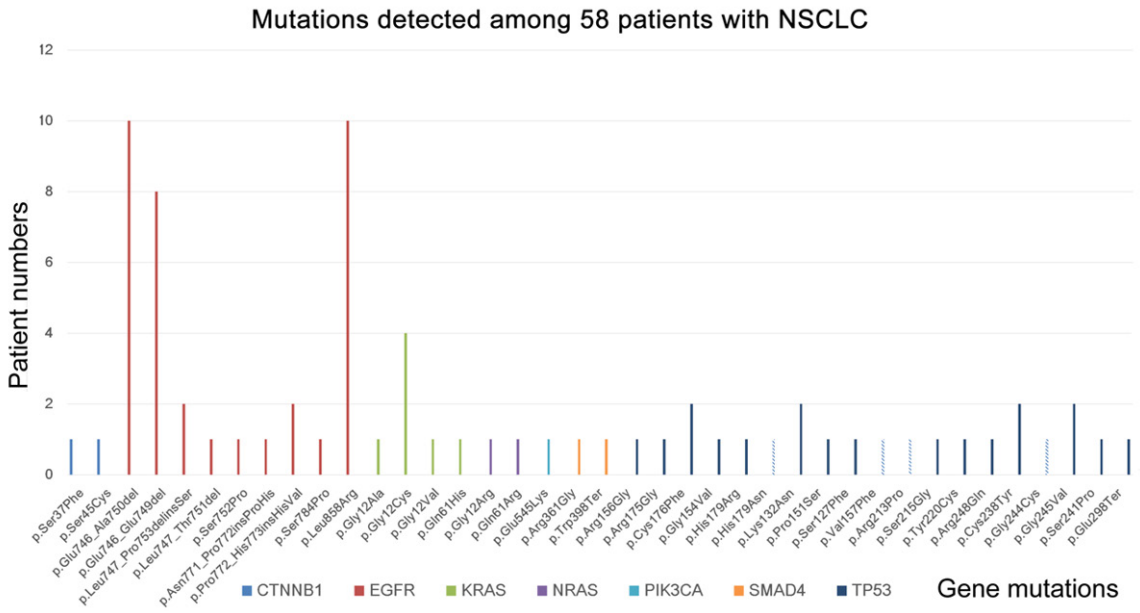

Figure 1. Mutations detected among 58 NSCLC patients. Distribution of the Mutations in 58 NSCLC patients.

Statistical analysis

Statistical analysis was performed using SAS (ver. 9.1; SAS Institute, Cary, NC, USA). Fisher's exact test was used to compare the mutation rates between male and female patients or between ADK and SCC samples. Female patients had a significantly higher mutation rate in the EGFR gene than male patients (60.0% and 30.3%, respectively; P = 0.03). A significantly higher number of EGFR mutations were detected in ADK samples compared to SCC samples (49.0% and 0.0%, respectively; P = 0.02). No significant difference in mutation rate of the other genes was observed between male and female patients or between ADK and SCC samples (all P > 0.05).

Results

Sequence coverage

For the sequencing data of 58 samples analyzed in the present study, the mean depth of coverage was 1,799 reads, ranging from 137 to 2,000 reads.

Mutations detected among 58 NSCLC patients

NGS targeted sequencing revealed that 44 (75.9%) of the 58 NSCLC patients carried mutations in 7 of the 22 genes screened in the present study (**Figure 1**). Among these patients, 24 patients (41.4%) had one mutation and 20 patients (34.5%) had two or more mutations, whereas 17 patients (29.3%) had mutations in two or more genes.

Distribution of the mutations in 58 NSCLC patients

Higher mutation frequencies were observed in the EGFR (43.1%), TP53 (37.9%), and KRAS (12.1%) genes, whereas lower frequencies were detected in the CTNNB1 (3.4%), NRAS (3.4%), SMAD4 (3.4%), and PIK3CA (1.7%) genes. No significant difference in mutation rate between males and females were observed (75.8% and 76.0%, respectively; P = 0.62). The ADK samples showed a higher mutation rate compared to the SCC samples (78.4% and 57.1%, respectively), although the difference was not significant (P = 0.34). Female patients had a significantly higher mutation rate in the EGFR gene than male patients (60.0% and 30.3%, respectively; P = 0.03). A significantly higher number of EGFR mutations were detected in ADK samples compared to SCC samples (49.0% and 0.0%, respectively; P = 0.02). No significant difference in mutation rate of the other genes was observed between male and female patients or between ADK and SCC samples (all P > 0.05; **Table 2**).

Table 2. Distributions of the mutations detected among 58 NSCLC patients

Gene	No. of patients (n = 58) (%)	Gender		Histology	
		Male (n = 33) (%)	Female (n = 25) (%)	ADK (n = 51) (%)	SCC (n = 7) (%)
Any gene	44 (75.9)	25 (75.8)	19 (76.0)	40 (78.4)	4 (57.1)
EGFR	25 (43.1)	10 (30.3)	15 (60.0)	25 (49.0)	0 (0.0)
TP53	22 (37.9)	12 (36.4)	10 (40.0)	18 (35.3)	4 (57.1)
KRAS	7 (12.1)	6 (18.2)	1 (4.0)	7 (13.7)	0 (0.0)
CTNNB1	2 (3.4)	1 (3.0)	1 (4.0)	2 (3.9)	0 (0.0)
NRAS	2 (3.4)	1 (3.0)	1 (4.0)	2 (3.9)	0 (0.0)
SMAD4	2 (3.4)	1 (3.0)	1 (4.0)	2 (3.9)	0 (0.0)
PIK3CA	1 (1.7)	1 (3.0)	0 (0.0)	1 (2.0)	0 (0.0)

Abbreviations: NSCLC: Non-small cell lung cancer; M: Male; F: Female; ADK: Adenocarcinoma; SCC: Squamous cell carcinoma.

Discussion

NGS targeted sequencing has become the preferred technique over the traditional mutation detection methods to simultaneously screen multiple mutations in various genes because it is relatively cheaper, faster, and requires less DNA [18]. In the present study, we investigated 58 NSCLC patients using a NGS targeted sequencing panel that included 1,850 targeted sites in 22 cancer drug target genes. Our NGS analyses revealed that 75.9% of NSCLC patients carried mutations in 7 of 22 interested genes (**Figure 1**). These results, together with -previous studies [11, 16-19, 25], suggest that NGS targeted sequencing can effectively detect mutations in clinical samples of NSCLC, thereby highlighting that NGS is a promising tool in NSCLC personalized medicine.

In the present study, the highest mutation frequency in Chinese NSCLC patients was observed in the EGFR gene (43.1%), followed by the TP53 (37.9%) and KRAS (12.1%) genes (**Table 2**). Our results were in agreement with previous studies [16-18, 26]. For instance, in an NGS analysis of 38 Belgian NSCLC patients [19], D'Haene and coworkers found that the most common mutations occurred in the KRAS (39.5%), TP53 (39.5%), and EGFR (10.5%) genes. In addition, Zhang et al. [26] screened 184 Chinese NSCLC patients and determined that EGFR mutations were the most prevalent (59.9%), followed by the KRAS (16.9%) and TP53 (12.7%) genes. It appears that the mutation frequencies in NSCLC patients vary among populations, particularly among different racial populations [16-18, 26]. Taken together, these findings indicated that the most common mutations in NSCLC patients regardless of population occur in the EGFR, TP53, and KRAS genes.

Female NSCLC patients showed significantly higher mutation rates in the EGFR gene than male patients (60% and 30%, respectively; P = 0.03; **Table 2**). Our results were in agreement with the findings of Vigneswaran et al. involving American NSCLC patients with broad ethnic diversity [17], which showed that EGFR mutations were more common in female patients than in male patients (59% vs. 41%). Moreover, Zhang et al. [26] also reported that Chinese female NSCLC patients had higher EGFR mutation frequencies than male patients (57.3% vs. 33.9%).

In the present study, a significantly higher number of EGFR mutations were observed in ADK samples compared to SCC samples (49% and 0%, respectively; P = 0.02; **Table 2**). Our findings are in line with those of Vigneswaran et al. [17], in which EGFR mutations were more common in ADK samples than in SCC samples (82% vs. 3%). Similarly, Zhang et al. also reported that EGFR mutations more commonly occur in ADK samples than in SCC samples (47.4% vs. 29.2%) [26].

The limitation of the present study was that we validated the NGS results by ARMS-PCR only for the EGFR mutations but not for the mutations in the six other genes, including CTNNB1, KRAS, NRAS, PIK3CA, SMAD4, and TP53. Although all the EGFR mutations identified by NGS targeted sequencing were validated by ARMS-PCR, further studies are required to rule out possible false-positives in the other genes screened in the present study.

In summary, we investigated 58 NSCLC patients using a NGS targeted sequencing panel that included 1,850 targeted sites in 22 cancer drug target genes. NGS targeted sequencing revealed that 75.9% of NSCLC patients carried mutations in 7 of these genes. Higher mutation frequencies were observed in the EGFR (43.1%), TP53 (37.9%), and KRAS (12.1%) genes, whereas lower frequencies were detected in the CTNNB1 (3.4%), NRAS (3.4%), SMAD4 (3.4%), and PIK3CA (1.7%) genes. Female patients showed a significantly higher mutation rate in the EGFR gene than male patients. The ADK samples showed a significantly higher number of EGFR mutations compared to SCC samples. These results suggest that NGS targeted sequencing can effectively detect mutations in clinical tumor samples, thereby highlighting NGS as a promising tool in personalized medicine.

Acknowledgements

This study was supported by Science and Technology Program of Guangdong (Grant No. 2015B020233009; Grant No. 2015A030401-040).

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Ming Li, Institute of Antibody Engineering, School of Biotechnology, Southern Medical University, NO. 1838 Guangzhou Road, Guangzhou, Guangdong, China. Tel: +86-20-61648550; Fax: +86-20-61648554; E-mail: lm20-062006@163.com

References

[1] Jemal A, Siegel R, Ward E, Hao Y, Xu J and Thun MJ. Cancer statistics, 2009. CA Cancer J Clin 2009; 59: 225-249.

[2] Giangreco A, Groot KR and Janes SM. Lung cancer and lung stem cells: strange bedfellows? Am J Respir Crit Care Med 2007; 175: 547-553.

[3] Heist RS and Engelman JA. SnapShot: non-small cell lung cancer. Cancer Cell 2012; 21: 448, e442.

[4] Yamamoto H, Shigematsu H, Nomura M, Lockwood WW, Sato M, Okumura N, Soh J, Suzuki M, Wistuba II, Fong KM, Lee H, Toyooka S, Date H, Lam WL, Minna JD and Gazdar AF. PIK3CA mutations and copy number gains in human lung cancers. Cancer Res 2008; 68: 6913-6921.

[5] Spoerke JM, O'Brien C, Huw L, Koeppen H, Fridlyand J, Brachmann RK, Haverty PM, Pandita A, Mohan S, Sampath D, Friedman LS, Ross L, Hampton GM, Amler LC, Shames DS and Lackner MR. Phosphoinositide 3-kinase (PI3K) pathway alterations are associated with histologic subtypes and are predictive of sensitivity to PI3K inhibitors in lung cancer preclinical models. Clin Cancer Res 2012; 18: 6771-6783.

[6] Larsen JE and Minna JD. Molecular biology of lung cancer: clinical implications. Clin Chest Med 2011; 32: 703-740.

[7] Cufer T, Ovcaricek T and O'Brien ME. Systemic therapy of advanced non-small cell lung cancer: major-developments of the last 5-years. Eur J Cancer 2013; 49: 1216-1225.

[8] Pao W, Miller V, Zakowski M, Doherty J, Politi K, Sarkaria I, Singh B, Heelan R, Rusch V, Fulton L, Mardis E, Kupfer D, Wilson R, Kris M and Varmus H. EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. Proc Natl Acad Sci U S A 2004; 101: 13306-13311.

[9] Kobayashi K and Hagiwara K. Epidermal growth factor receptor (EGFR) mutation and personalized therapy in advanced nonsmall cell lung cancer (NSCLC). Target Oncol 2013; 8: 27-33.

[10] Soda M, Choi YL, Enomoto M, Takada S, Yamashita Y, Ishikawa S, Fujiwara S, Watanabe H, Kurashina K, Hatanaka H, Bando M, Ohno S, Ishikawa Y, Aburatani H, Niki T, Sohara Y, Sugiyama Y and Mano H. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. Nature 2007; 448: 561-566.

[11] Vavala T, Monica V, Lo Iacono M, Mele T, Busso S, Righi L, Papotti M, Scagliotti GV and Novello S. Precision medicine in age-specific non-small-cell-lung-cancer patients: Integrating biomolecular results into clinical practice-A new approach to improve personalized translational research. Lung Cancer 2016; [Epub ahead of print].

[12] Olivier M, Hollstein M and Hainaut P. TP53 mutations in human cancers: origins, consequences, and clinical use. Cold Spring Harb Perspect Biol 2010; 2: a001008.

[13] Ye T, Pan Y, Wang R, Hu H, Zhang Y, Li H, Wang L, Sun Y and Chen H. Analysis of the molecular and clinicopathologic features of surgically resected lung adenocarcinoma in patients under 40 years old. J Thorac Dis 2014; 6: 1396-1402.

[14] Coco S, Truini A, Vanni I, Dal Bello MG, Alama A, Rijavec E, Genova C, Barletta G, Sini C, Burrafato G, Biello F, Boccardo F and Grossi F.

Next generation sequencing in non-small cell lung cancer: new avenues toward the personalized medicine. Curr Drug Targets 2015; 16: 47-59.

[15] Vanni I, Coco S, Truini A, Rusmini M, Dal Bello MG, Alama A, Banelli B, Mora M, Rijavec E, Barletta G, Genova C, Biello F, Maggioni C and Grossi F. Next-Generation sequencing workflow for NSCLC critical samples using a targeted sequencing approach by Ion torrent PGM platform. Int J Mol Sci 2015; 16: 28765-28782.

[16] Masago K, Fujita S, Muraki M, Hata A, Okuda C, Otsuka K, Kaji R, Takeshita J, Kato R, Katakami N and Hirata Y. Next-generation sequencing of tyrosine kinase inhibitor-resistant non-small-cell lung cancers in patients harboring epidermal growth factor-activating mutations. BMC Cancer 2015; 15: 908.

[17] Vigneswaran J, Tan YH, Murgu SD, Won BM, Patton KA, Villaflor VM, Hoffman PC, Hensing T, Hogarth DK, Malik R, MacMahon H, Mueller J, Simon CA, Vigneswaran WT, Wigfield CH, Ferguson MK, Husain AN, Vokes EE and Salgia R. Comprehensive genetic testing identifies targetable genomic alterations in most patients with non-small cell lung cancer, specifically adenocarcinoma, single institute investigation. Oncotarget 2016; 7: 18876-18886.

[18] D'Haene N, Le Mercier M, De Neve N, Blanchard O, Delaunoy M, El Housni H, Dessars B, Heimann P, Remmelink M, Demetter P, Tejpar S and Salmon I. Clinical validation of targeted next generation sequencing for colon and lung cancers. PLoS One 2015; 10: e0138245.

[19] Cai X, Sheng J, Tang C, Nandakumar V, Ye H, Ji H, Tang H, Qin Y, Guan H, Lou F, Zhang D, Sun H, Dong H, Zhang G, Liu Z, Dong Z, Guo B, Yan H, Yan C, Wang L, Su Z, Li Y, Jones L, Huang XF, Chen SY, Wu T and Lin H. Frequent mutations in EGFR, KRAS and TP53 genes in human lung cancer tumors detected by ion torrent DNA sequencing. PLoS One 2014; 9: e95228.

[20] Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, Jemal A, Yu XQ and He J. Cancer statistics in China, 2015. CA Cancer J Clin 2016; 66: 115-132.

[21] Zheng H, Wang Y, Tang C, Jones L, Ye H, Zhang G, Cao W, Li J, Liu L, Liu Z, Zhang C, Lou F, Li Y, Shi Z, Zhang J, Zhang D, Sun H, Dong H, Dong Z, Guo B, Yan HE, Lu Q, Huang X and Chen SY. TP53, PIK3CA, FBXW7 and KRAS mutations in esophageal cancer identified by targeted sequencing. Cancer Genomics Proteomics 2016; 13: 231-238.

[22] Bai X, Zhang E, Ye H, Nandakumar V, Wang Z, Chen L, Tang C, Li J, Li H, Zhang W, Han W, Lou F, Zhang D, Sun H, Dong H, Zhang G, Liu Z, Dong Z, Guo B, Yan H, Yan C, Wang L, Su Z, Li Y, Jones L, Huang XF, Chen SY and Gao J. PIK3CA and TP53 gene mutations in human breast cancer tumors frequently detected by ion torrent DNA sequencing. PLoS One 2014; 9: e99306.

[23] Thorvaldsdottir H, Robinson JT and Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform 2013; 14: 178-192.

[24] Cai W, Lin D, Wu C, Li X, Zhao C, Zheng L, Chuai S, Fei K, Zhou C and Hirsch FR. Intratumoral heterogeneity of ALK-Rearranged and ALK/EGFR coaltered lung adenocarcinoma. J Clin Oncol 2015; 33: 3701-3709.

[25] Xu X, Yang Y, Li H, Chen Z, Jiang G and Fei K. Assessment of the clinical application of detecting EGFR, KRAS, PIK3CA and BRAF mutations in patients with non-small cell lung cancer using next-generation sequencing. Scand J Clin Lab Invest 2016; 1-7.

[26] Zhang S, Xia B, Jiang H, Wang L, Xu R, Shi Y, Zhang J, Xu M, Cram DS and Ma S. Comprehensive profiling and quantitation of oncogenic mutations in non small-cell lung carcinoma using single molecule amplification and re-sequencing technology. Oncotarget 2016; [Epub ahead of print].

Supplementary Table 1. 29 *EGFR* mutations screened by ARMS-PCR

Exon	Cosmic ID	Mutation	Nucleotide change	Target sequence	Sequence following target sequence
18	6252	G719S	2155G>A	CAAAGTGCTG	AGC
	6253	G719C	2155G>T		TGC
	6239	G719A	2156G>C		GCC
19	6223	E746-A750del (1)	2235-2249 del 15	TATCAA	AACATCTC
	13551	E746-T751>I	2235-2252>AAT (complex)		AATATCTC
	6225	E746-A750del (2)	2236-2250 del 15		GACATCTC
	12728	E746-T751del	2236-2253 del 18		GTCTCCGA
	12384	E746-S752>V	2237-2250>T (complex)		GGTACATC
	12678	E746-T751>A	2237-2251 del 15		GGCATCTC
	12367	E746-S752>A	2237-2254 del 18		GGCTCCGA
	12422	L747-A750>P	2238-2248>GC (complex)		GGAGCCAACAT
	12419	L747-T751>Q	2238-2252>GCA (complex)		GGAGCAATCTC
	6220	E746-S752>D	2238-2255 del 18		GGATCCGAAAG
	6218	L747-E749del	2239-2247 del 9		GGAAGCAACAT
	12382	L747-A750>P	2239-2248 TTAAGAGAAG>C (complex)		GGAACCAACAT
	12383	L747-T751>P	2239-2251>C (complex)		GGAACCATCTC
	6254	L747-T751del	2239-2253 del 15		GGAATCTCCGA
	6255	L747-S752del	2239-2256 del 18		GGAACCGAAAG
	12387	L747-P753>Q	2239-2258>CA (complex)		GGAACAGAAAG
	6210	L747-T751>S	2240-2251 del 12		GGAATCATCTC
	12369	L747-T751del	2240-2254 del 15		GGAATCTCCGA
	12370	L747-P753>S	2240-2257 del 18		GGAATCGAAAG
20	6241	S768I	2303 G>T	GTGATGGCCA	T
	12376	V769-D770insASV	2307-2308 ins (GCCAGCGTG)		GCGTGGCCAGCGTGGAC
	12378	D770-N771insG	2310-2311 ins GGT		GCGTGGACGGTAAC
	12377	H773-V774insH	2319-2320 ins CAC		GCGTGGACAACCCCCACCACG
	6240	T790M	2369 C>T	CTCATCA	T
21	6224	L858R	2573 T>G	GATTTTGGGC	G
	6213	L861Q	2582 T>A		TGGCCAAACA

Supplementary Table 2. Mutations detected among 58 patients with NSCLC

Gene	Exon	Mutation	Patient	Gender	Age	Histology	*EGFR* (ARMS-PCR)
EGFR	19	p.Glu746_Ala750del	3	F	56	ADK	+
EGFR	19	p.Glu746_Ala750del	4	F	55	ADK	+
EGFR	19	p.Glu746_Ala750del	6	M	61	ADK	NA
EGFR	19	p.Glu746_Ala750del	16	M	88	ADK	NA
EGFR	19	p.Glu746_Ala750del	28	F	46	ADK	+
EGFR	19	p.Glu746_Ala750del	39	F	45	ADK	+
EGFR	19	p.Glu746_Ala750del	40	F	74	ADK	NA
EGFR	19	p.Glu746_Ala750del	45	M	66	ADK	NA
EGFR	19	p.Glu746_Ala750del	46	F	52	ADK	+
EGFR	19	p.Glu746_Ala750del	59	F	60	ADK	+
EGFR	19	p.Glu746_Glu749del	3	F	56	ADK	+
EGFR	19	p.Glu746_Glu749del	4	F	55	ADK	+
EGFR	19	p.Glu746_Glu749del	6	M	61	ADK	NA
EGFR	19	p.Glu746_Glu749del	39	F	45	ADK	+
EGFR	19	p.Glu746_Glu749del	40	F	74	ADK	NA
EGFR	19	p.Glu746_Glu749del	45	M	66	ADK	NA
EGFR	19	p.Glu746_Glu749del	46	F	52	ADK	+
EGFR	19	p.Glu746_Glu749del	59	F	60	ADK	+
EGFR	19	p.Leu747_Pro753delinsSer	11	F	77	ADK	NA
EGFR	19	p.Leu747_Pro753delinsSer	12	M	56	ADK	NA
EGFR	19	p.Leu747_Thr751del	44	M	61	ADK	NA
EGFR	19	p.Ser752Pro	11	F	77	ADK	NA
EGFR	20	p.Asn771_Pro772insProHis	41	F	60	ADK	NA

EGFR	20	p.Pro772_His773insHisVal	14	F	63	ADK	NA
EGFR	20	p.Pro772_His773insHisVal	18	F	58	ADK	NA
EGFR	20	p.Ser784Pro	9	M	61	ADK	NA
*EGF*R	21	p.Leu858Arg	12	M	56	ADK	+
EGFR	21	p.Leu858Arg	13	M	64	ADK	+
EGFR	21	p.Leu858Arg	26	F	68	ADK	+
EGFR	21	p.Leu858Arg	30	M	50	ADK	+
EGFR	21	p.Leu858Arg	32	F	65	ADK	+
EGFR	21	p.Leu858Arg	40	F	74	ADK	+
EGFR	21	p.Leu858Arg	47	F	64	ADK	+
EGFR	21	p.Leu858Arg	48	M	67	ADK	+
EGFR	21	p.Leu858Arg	51	F	52	ADK	+
EGFR	21	p.Leu858Arg	55	M	66	ADK	+

Abbreviations: NSCLC: Non-small cell lung cancer; M: Male; F: Female; ADK: Adenocarcinoma; SCC: Squamous cell carcinoma; NA: Not available. + refers to EGFR mutations detected by ARMS-PCR. - indicates no EGFR mutations detected by ARMS-PCR.