Molecular and Cellular Pathobiology

Cancer Research

Development of Lung Adenocarcinomas with Exclusive Dependence on Oncogene Fusions

Motonobu Saito[1,2], Yoko Shimada[1], Kouya Shiraishi[1], Hiromi Sakamoto[3], Koji Tsuta[4], Hirohiko Totsuka[5], Suenori Chiku[6], Hitoshi Ichikawa[3], Mamoru Kato[7], Shun-ichi Watanabe[8], Teruhiko Yoshida[3], Jun Yokota[1,9], and Takashi Kohno[1]

Abstract

This report delivers a comprehensive genetic alteration profile of lung adenocarcinomas (LADC) driven by *ALK*, *RET*, and *ROS1* oncogene fusions. These tumors are difficult to study because of their rarity. Each drives only a low percentage of LADCs. Whole-exome sequencing and copy-number variation analyses were performed on a Japanese LADC cohort ($n = 200$) enriched in patients with fusions ($n = 31$, 15.5%), followed by deep resequencing for validation. The driver fusion cases showed a distinct profile with smaller numbers of nonsynonymous mutations in cancer-related genes or truncating mutations in SWI/SNF chromatin remodeling complex genes than in other LADCs ($P < 0.0001$). This lower mutation rate was independent of age, gender, smoking status, pathologic stage, and tumor differentiation ($P < 0.0001$) and was validated in nine fusion-positive cases from a U.S. LADCs cohort ($n = 230$). In conclusion, our findings indicate that LADCs with *ALK*, *RET*, and *ROS1* fusions develop exclusively via their dependence on these oncogene fusions. The presence of such few alterations beyond the fusions supports the use of monotherapy with tyrosine kinase inhibitors targeting the fusion products in fusion-positive LADCs. *Cancer Res; 75(11); 2264–71. ©2015 AACR.*

Introduction

Lung adenocarcinoma (LADC) is the most frequent histologic type of lung cancer and its incidence is rising in Asian and Western countries. Oncogenic fusions of the protein tyrosine kinase genes *ALK*, *RET*, and *ROS1*, identified by us and others, are believed to drive the development of a subset (3%–4%, 1%–2%, and 1%–2%, respectively) of LADCs (1–4). These fusion-positive LADCs often, but not always, show mucinous–cribriform patterns (2, 5–7). In addition, fusion-positive LADCs tend to occur in young and non/light-smoking individuals (2, 4, 8–10), and show a high therapeutic response to tyrosine kinase inhibitors (TKI) that suppress the kinase activity of the fusion products (11–13). These results indicate that fusion-positive LADCs are a distinct LADC molecular entity. Despite recent large-scale genome sequencing studies in LADCs (14–16), the genetic profile of fusion-positive LADCs remains unknown due to the rarity of these tumors. Better genetic characterization of fusion-positive LADCs is required to improve therapeutic strategies. If other genetic abnormalities are detected, agents targeting these defects could be used in combination with TKIs to improve efficacy and outcome (11, 17, 18).

LADCs carrying *ALK*, *ROS1*, and *RET* fusions have already been shown to lack activating mutations in other oncogenes, such as *EGFR*, *KRAS*, *BRAF*, and *HER2/ERBB2* (2–4, 14–16); however, the mutational status of other genes frequently mutated in lung and other cancers, including those identified in the cancer gene census (CGC; ref. 19) or those identified as significantly mutated genes (SMG) in 12 common cancers (20), is unknown. These gene sets include tumor-suppressor genes, such as *TP53*, *CDKN2A*, *KEAP1*, and *STK11/LKB1*, and chromatin remodeling/modifying genes, such as *ARID1A* and *SMARCA4*, which are the targets of genetic loss-of-function aberrations in cancer cells (20). Notably, recent studies suggested that these deleterious aberrations are therapeutically targetable; drugs restoring the function of mutant p53 proteins are being developed (21, 22), and synthetic lethality-based therapies have been considered by us and others to treat cancers with *TP53*, *LKB1*, *ARID1A*, and *SMARCA4* deficiencies (23–28).

We performed the comparative genetic aberration profiling of oncogenic fusion-positive and -negative LADCs. Two hundred cases of snap-frozen surgical LADC tissues were subjected to whole-exome sequencing using a next-generation sequencer and to copy-number variation analysis using a DNA chip. The selected cases were enriched in oncogenic *ALK*, *RET*, or *ROS1* fusions ($n = 31$, 15.5%) and included 96 cases (48.0%) of activating mutations in *EGFR*, *KRAS*, *HER2*, *BRAF*, and *HRAS* oncogenes and 73

[1]Division of Genome Biology, National Cancer Center Research Institute, Tokyo, Japan. [2]Department of Organ Regulatory Surgery, Fukushima Medical University School of Medicine, Fukushima, Japan. [3]Division of Genetics, National Cancer Center Research Institute, Tokyo, Japan. [4]Division of Pathology and Clinical Laboratories, National Cancer Center Hospital, Tokyo, Japan. [5]Solution Development Department, New Project Development Division, Hitachi Government & Public Sector Systems, Ltd., Tokyo, Japan. [6]Science Solutions Division, Mizuho Information and Research Institute Inc., Tokyo, Japan. [7]Department of Bioinformatics, National Cancer Center Research Institute, Tokyo, Japan. [8]Division of Thoracic Surgery, National Cancer Center Hospital, Tokyo, Japan. [9]Cancer Genome Biology Group, Institute of Predictive and Personalized Medicine of Cancer, Barcelona, Spain.

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

Corresponding Author: Takashi Kohno, National Cancer Center Research Institute, 1-1 Tsukiji 5-chome, Chuo-ku, Tokyo 104-0045, Japan. Phone: 81-3-3547-5272; Fax: 81-3-3542-0807; E-mail: tkkohno@ncc.go.jp

doi: 10.1158/0008-5472.CAN-14-3282

AACR

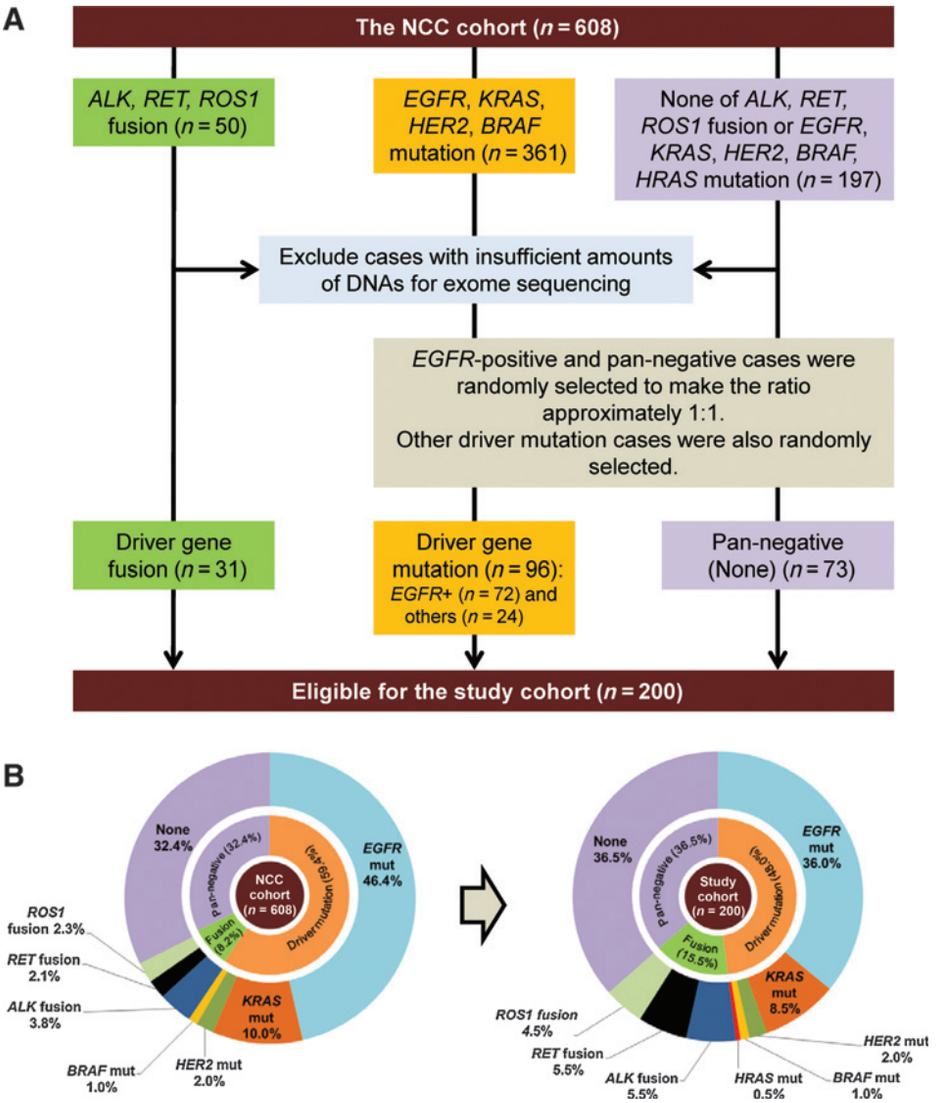

Figure 1.
Patient selection. A, two hundred surgically resected LADC cases (the study cohort) were selected from 608 consecutive cases (the NCC cohort). Patients were classified into three groups: "driver fusion," "driver mutation," and "pan-negative." The original cohort consists of 608 cases screened for driver mutations in *EGFR*, *KRAS*, *HER2*, and *BRAF* and for driver fusions involving *ALK*, *RET*, and *ROS1*. Other cases were classified as pan-negative cases. B, the study cohort, consisting of 200 cases enriched in *ALK*, *RET*, and *ROS1* fusion-positive cases ($n = 31$), was subjected to whole-exome sequencing. The study cohort includes 96 driver mutation cases, including activating mutations in *EGFR*, *KRAS*, *BRAF*, and *HER2*, and a case with an *HRAS* mutation detected in the present exome sequence.

cases (35.5%) without any such aberrations. The study revealed that fusion-positive LADCs have a unique genetic profile that includes fewer genetic aberrations than other LADCs.

Patients and Methods

Patients

A total of 200 LADC cases (the study cohort) were selected from 608 consecutive cases (the NCC cohort; NCC Biobank) who underwent surgical resection between 1997 and 2008 at the National Cancer Center Hospital, Tokyo, and for whom snap-frozen cancerous and noncancerous lung tissues were available (Fig. 1A and B). All of the 608 cases were screened for *EGFR*, *KRAS*, *BRAF*, and *HER2* hot spot mutations by the HRM method, and for *EML4*– and *KIF5B*–*ALK*, *KIF5B*– and *CCDC6*–*RET*, and *CD74*–, *EZR*– and *SLC34A2*–*ROS1* fusions by RT-PCR, as described previously (Supplementary Table S2; refs. 4, 7). In addition, a case with a novel type of *RET* fusion, *KIAA1468-RET*, which was detected by whole RNA sequencing, was included in this cohort.

Driver fusion study subjects ($n = 31$), that is, those with *ALK* ($n = 11$), *RET* ($n = 11$), or *ROS1* ($n = 9$) fusions were selected from all 50 fusion-positive cases in the NCC original cohort, that is, those with *ALK* fusions ($n = 23$), *RET* ($n = 13$), or *ROS1* ($n = 14$) fusions based on the criterion that sufficient amounts of genomic DNA for whole-exome sequencing were available (Fig. 1A). *EGFR*-positive ($n = 72$), other driver mutation ($n = 23$), and pan-negative cases ($n = 74$), with sufficient amounts of genomic DNA, were randomly selected from the NCC original cohort together with *EGFR*-positive ($n = 282$), other driver mutation ($n = 79$), and pan-negative cases ($n = 197$) to obtain a *EGFR*-positive case:pan-negative case ratio of approximately 1:1 and to make the total number of samples 200. The exome sequencing analysis revealed an activating *HRAS* mutation (Q61L), so this case was classified as driver mutation. Thus, the study cohort included 73 pan-negative cases and an *HRAS* mutation–positive case. The selection resulted in a cohort that was more enriched in driver fusion cases than the original cohort (Fig. 1B). The study subjects were diagnosed according to the seventh TNM classification of malignant tumors

(29, 30). The study was approved by the Institutional Review Boards of the NCC.

Genome copy analysis and tumor content estimation

Genome copy-number and allelic status were assessed in all 200 study cases by Illumina OMNI 2.5M array analysis using both cancerous and noncancerous lung DNA. Tumor cell content in each tumor sample and copy numbers for each gene were deduced using the Global Parameter Hidden Markov Model method (31).

Exome sequencing

Exome sequencing was conducted from 2.5 μg of cancerous or noncancerous DNA isolated from snap-frozen tissues. Exome capture was performed using the Agilent SureSelect Human All Exon 50-Mb, V4 or V5 according to the manufacturer's instructions. Exome sequencing was performed on the Illumina HiSeq 2000 platform using 75 bp paired-end reads (Illumina). Basic alignment and sequence quality control were conducted using the Picard and Firehose pipelines. The reads were aligned against the reference human genome from UCSC human genome 19 (Hg19) using the Burrows Wheeler Aligner Multi-Vision software package. Because duplicate reads were generated during the PCR amplification process, paired-end reads that aligned to the same genomic positions were removed using SAMtools.

Somatic SNVs were called by the MuTect program, which applies a Bayesian classifier to allow the detection of somatic mutations with a low allele frequency (32). Somatic InDel mutations were called by the GATK Somatic Indel Detector (33). SNV and InDel detection was supported by visual examination using the Integrative Genomics Viewer software (34).

Verification of somatic mutations by deep resequencing

Mutations in all coding exons of the following 28 genes were examined by targeted genome capture and massively parallel sequencing using an Illumina HiSeq 2000 system and the Haloplex Custom Enrichment Kit (Agilent Technologies): 10 representative cancer census genes (19), *AKT1*, *APC*, *CTNNB1*, *KEAP1*, *MAP2K1*, *MET*, *NRAS*, *PIK3CA*, *STK11* and *TP53*, and 18 SWI/SNF chromatin remodeling genes (35) whose mutations were detected in one or more tumors by exome sequencing, *ACTL6B*, *ARID1A*, *ARID1B*, *ARID2*, *BPTF*, *DPF1*, *EP400*, *HLTF*, *PBRM1*, *RAD54L2*, *SHPRH*, *SMARCA2*, *SMARCA4*, *SMARCAD1*, *SMARCAL1*, *SMARCB1*, *SMARCC1*, and *SRCAP*. Average read depths were approximately 1,000.

Cancer gene census

A list of somatic mutations from the CGC was downloaded from the most recently released COSMIC V70 (36).

Validation in a U.S. cohort

Validation analysis was performed using The Cancer Genome Atlas (TCGA) LADC study data (16). The TCGA cohort of 230 cases was selected from 678 patients with previously untreated LADC based on the tumor percentage, availability of clinical data, and availability of sufficient amounts of nucleic acid (37). Driver gene fusions and mutations were evaluated in all cases; thus, the 230 cases were subgrouped according to driver gene type: 9 driver fusion cases (3.9%) with *ALK* ($n = 3$), *RET* ($n = 2$), and *ROS1* ($n = 2$) fusions, 121 driver mutation cases (52.6%) with hot spot mutations in *EGFR* ($n = 26$), *KRAS* ($n = 74$), *HER2* ($n = 4$), *BRAF* ($n = 16$), and *HRAS* ($n = 1$), and 100 pan-negative cases (43.5%; Supplementary Fig. S4A). Among the cancer-related and SWI/SNF chromatin remodeling genes investigated in Fig. 2, information on *CDKN2A*, *RBM10*, *RB1*, *NF1*, *KEAP1*, *MET*, *MGA*, *U2AF1*, *PIK3CA*, *STK11*, *TP53*, *SMARCA4*, and *ARID1A* was available, and therefore was used in the analysis.

Statistical analysis

Statistical analyses of differences in genetic alterations, clinical and pathologic factors between the driver aberration groups, or smoking status were assessed by using the Kruskal–Wallis test, two-sided Mann–Whitney test, two-sided Fisher exact test or χ^2 test, and two-sided Spearman r test in GraphPad Prism 5 software (GraphPad Software). Multivariate regression analysis, including the number of nonsynonymous mutations per Mb and clinicopathologic factors (age, gender, smoking status, pathologic stage, and tumor differentiation), was conducted using JMP 10 software (SAS Institute). A P value of <0.05 was considered significant.

Results

Study cohort

The study cohort of 200 LADC cases included 31 cases (15.5%) with *ALK*, *RET*, or *ROS1* fusions, 96 cases (48.0%) with hot spot mutations in *EGFR*, *KRAS*, *HER2*, *BRAF*, or *HRAS*, and 73 cases (36.5%) without any of these driver gene aberrations (Table 1). The study cohort samples were selected from the original National Cancer Center (NCC) cohort samples ($n = 608$) to enrich fusion-positive cases (Fig. 1A and B; Supplementary Table S1). The driver gene aberrations detected were mutually exclusive, as predicted (14, 16, 38). There was no case of *NTRK1* fusion, a recently reported oncogene fusion (39), in this cohort. The 31, 96, and 73 cases were classified into three groups designated the "driver fusion," "driver mutation," and "pan-negative" groups, respectively (Fig. 1A and B).

Patient characteristics are summarized in Table 1. The driver fusion cases included a significantly higher frequency of young, female and never-smoker cases, and the pan-negative cases included a significantly higher frequency of older, male, and heavy smokers ($P = 0.024$ by Kruskal–Wallis test, 0.001 by χ^2 test, and <0.0001 by χ^2 test, respectively) than the other two groups, representing the characteristics of the original NCC cohort (Supplementary Table S2). In addition, the pan-negative cases showed a significant predominance of poor differentiation ($P = 0.0007$ by χ^2 test). These features are consistent with previous reports (2, 4–6, 8–10), indicating the authenticity and suitability of our study cohort to establish the genetic profile of LADCs driven by specific aberrations.

Genome-wide mutation profiling

Genomic DNA from cancerous and noncancerous lung tissues was subjected to whole-exome sequencing and SNP chip analyses. The average sequencing depth of the driver fusion, driver mutation, and pan-negative groups was similar; median: 106 (range, 84–218), 98 (82–216), and 104 (82–145), respectively ($P \geq 0.05$ by Kruskal–Wallis test; Supplementary Fig. S1A). The tumor contents deduced from genome-wide copy-number and allelic imbalance data obtained by SNP chip analysis were also similar ($P \geq 0.05$ by Kruskal–Wallis test; Supplementary Fig. S1B). Therefore, having ruled out

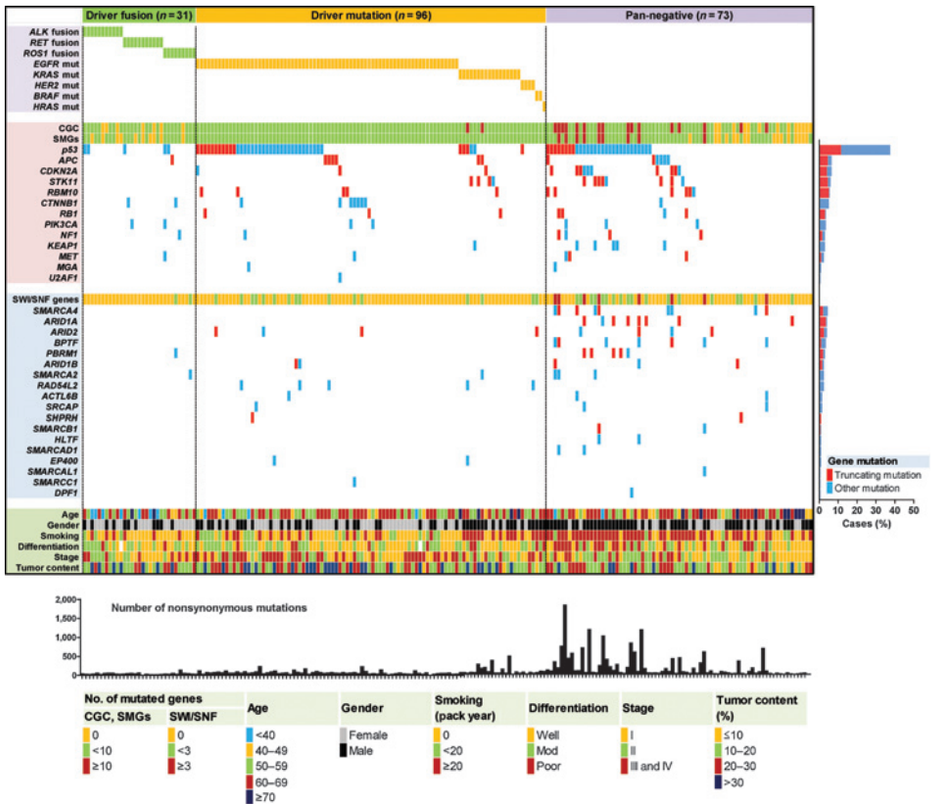

Figure 2.
Genetic aberration profile of lung adenocarcinoma. Genetic aberrations in LADCs according to driver aberration. Driver gene aberrations, numbers of aberrant CGC genes, SMGs and SWI/SNF chromatin remodeling genes, and aberrations in representative cancer-related and SWI/SNF chromatin remodeling genes are shown for each case with clinical characters. Bar chart (right) indicates fractions of cases with gene aberrations. The numbers of all nonsynonymous mutations are indicated with vertical bars at the bottom.

differential sensitivity between the groups, the sample set was found suitable to compare the genetic aberration profiles of the three groups. In addition, the SMGs deduced by the MutSigCV analysis were consistent with recent large-scale sequencing studies (14–16), supporting the authenticity of the present sample set (Supplementary Table S3).

Infrequent gene mutations in fusion-positive LADCs

Nonsynonymous mutations (missense, nonsense, indel, and splicing site mutations) based on driver genes are depicted in Fig. 2, focusing on CGC genes, SMGs, and SWI/SNF chromatin remodeling complex genes (20, 40). The driver fusion cases appeared to harbor fewer mutated genes. In fact, the median number of nonsynonymous mutations per Mb was lowest in driver fusion cases (0.37; range, 0–1.5) compared with driver mutation cases (0.65; 0–6.0) and pan-negative cases (0.87; 0–24.8; $P < 0.0001$ by Kruskal–Wallis test; Fig. 3A and Supplementary Fig. S2A). The median number of mutated CGC genes was smaller in driver fusion cases (1.0; range, 0–4) than in driver mutation cases (3.0; range, 1–12) and pan-negative cases (3.0; range, 0–46; $P < 0.0001$ by Kruskal–Wallis test). The median number of mutated SMGs was also smaller in driver fusion cases (1.0; range, 0–4) than in driver mutation cases (3.0; range, 1–7) and pan-negative cases (2.0; range, 0–25; $P < 0.0001$ by Kruskal–Wallis test). A positive correlation was observed between the number of nonsynonymous mutations in all genes and that in CGC genes or SMGs ($P < 0.0001$ by Spearman test; Supplementary Fig. S2B). Driver fusion was associated with fewer mutations regardless of age, gender, smoking status, pathologic stage, or tumor differentiation ($P < 0.0001$ by multivariate regression model analysis; Supplementary Table S4). Thus, driver fusion-positive LADCs seem to develop from the accumulation of a significantly smaller number of gene mutations than other types of LADCs.

Infrequent mutation of lung cancer–related genes in fusion-positive LADCs

Mutations in several genes known to contribute to lung carcinogenesis were examined. The *TP53* gene is a representative tumor-suppressor gene, included both in the CGC genes and SMGs, whose mutation occurs during the progression of LADCs following *EGFR* and *KRAS* mutations (41). *TP53* was the most frequently mutated gene in our study cohort (Table 2) and in the U.S. cases (20), with a mutation frequency of 41 of 96 (42.7%) in the driver mutation cases and 29 of 73 (39.7%) in the pan-negative cases. Notably, *TP53* mutations were significantly less frequent in the driver fusion group (5/31 or 16.1%, $P = 0.026$ by χ^2 test) than in the other two groups (Fig. 3B and Supplementary Fig. S3A). Truncating mutations in SWI/SNF chromatin remodeling complex genes, which are frequently observed in a variety of cancers (40), were not detected either (Figs. 2 and 3B and Supplementary Fig. S3B). The SMGs included 20 cellular process genes, most of which were less frequently mutated in fusion-positive cases than in mutation-positive and pan-negative cases (Supplementary Table S5). Two WNT signaling genes, *APC* and *CTNNB1*/*β-catenin*, and a PI(3)K signaling gene, *PIK3CA*, which activates PI(3)K signaling, are representative signaling genes known to be mutated in a small subset of LADCs (14, 16). Notably, these three genes were mutated irrespective of driver gene status (Fig. 3B, Table 2;

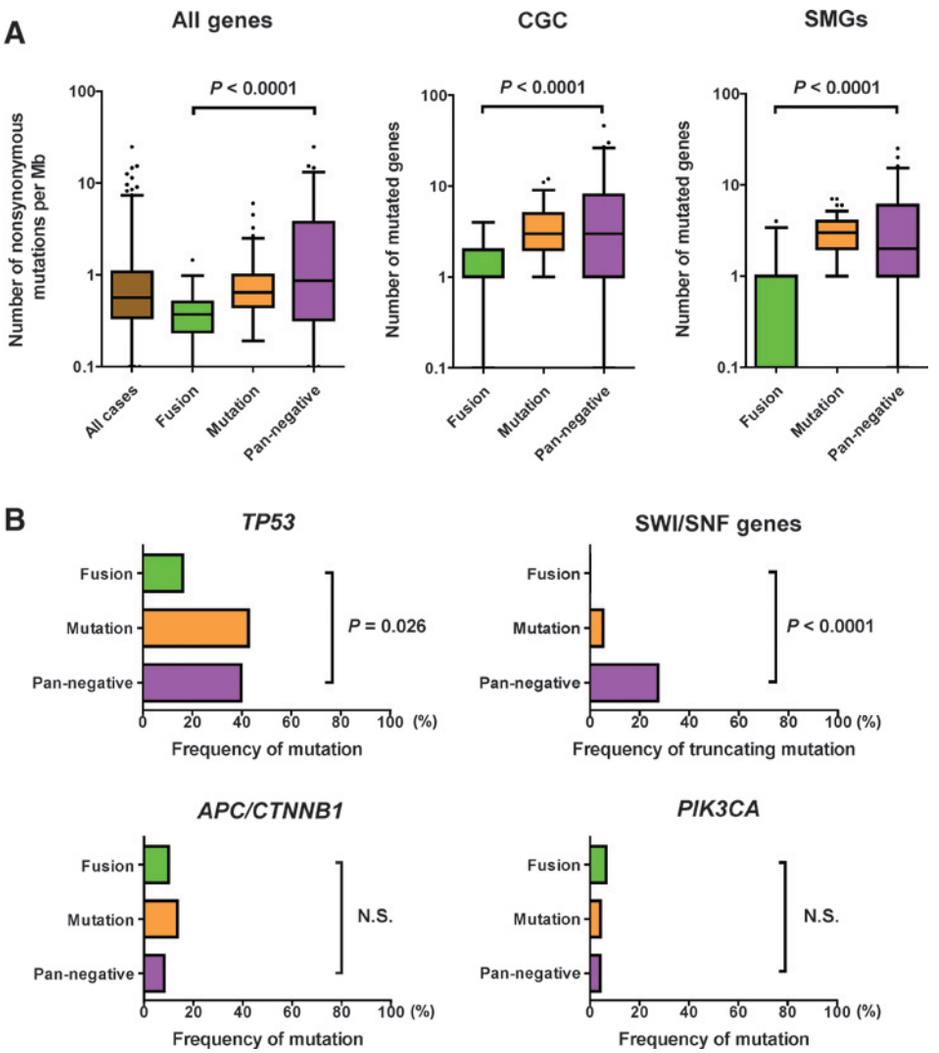

Figure 3.
Gene mutations by driver aberration. A, the numbers of nonsynonymous mutations per Mb in all genes and the numbers of CGC genes and SMGs with nonsynonymous mutations are shown. Whiskers represent the 5 to 95 percentiles and dots indicate the outliers. A *P* value was assessed by the Kruskal–Wallis test. B, frequency of mutation of the *TP53* gene, SWI/SNF chromatin remodeling genes, *APC* or *CTNNB1* genes, and *PIK3CA* gene. A *P* value was assessed by the χ^2 test; N.S., not significant.

Supplementary Table S5 and Supplementary Fig. S3C and S3D).

All 31 driver fusion cases and five driver mutation cases were subjected to deep resequencing of 28 genes with a mean depth of 1,000, using the same DNA samples. The 28 genes included 10 cancer-related genes and 18 SWI/SNF chromatin remodeling genes. All 19 nonsynonymous mutations detected by exome sequencing in these 36 cases were confirmed. In addition, an *ARID1A* missense mutation, which was not detected because of low depth by exome sequencing (i.e., no mutant reads among 27 reads), was detected in a driver fusion case. This resequencing verified that LADCs with oncogenic fusions develop through a pathway that involves only a small number of gene mutations.

Similar findings were also observed in a LADC cohort from the United States consisting of 230 cases (16). This cohort included 9 fusion-positive cases consisting of three *ALK* fusions, two *RET* fusions, and four *ROS1* fusions. Mutation distributions according to driver gene status are summarized in Supplementary Table S6. In this cohort, driver fusion cases carried fewer nonsynonymous mutations ($P = 0.046$ by Kruskal–Wallis test) and cancer-related gene mutations, including *TP53* ($P = 0.003$ by χ^2 test) mutations, than other cases (Supplementary Fig. S4A–A4C).

Copy-number gain in oncogenes

Genomic copy-number gains in nine oncogenes recently defined as amplified in histologic types of lung cancer (15) were examined to assess whether genomic aberrations other than mutations contribute to the development of fusion-positive LADCs (Supplementary Table S7). According to DNA chip analysis, these aberrations were infrequent regardless of the driver fusion.

Frequent gene mutations in pan-negative LADCs

The mutational characteristics of pan-negative cases were also investigated. These cases were associated with higher proportion of individuals of male gender, smokers, and individuals with poorly differentiated tumors (Table 1), and had the highest median number of nonsynonymous mutations per Mb among the three groups (Fig. 3A and Supplementary Fig. S2A). Furthermore, they had more frequent mutations in CGC genes ($P <$

Table 1. Clinical and pathologic characteristics of the 200 lung adenocarcinomas of the study cohort

		Driver fusion				Driver mutation							
Variable	**All**	**Total**	***ALK***	***RET***	***ROS1***	**Total**	***EGFR***	***KRAS***	***HER2***	***BRAF***	***HRAS***	**Pan-negative**	***P*[a]**
Cases, *n* (%)	200	31 (15.5)	11	11	9	96 (48.0)	72	17	4	2	1	73 (36.5)	
Age, y													0.024
Median	58.6	57	52	57	59	60	59	62	59	57	65	62	
Range	28–82	28–78	30–68	28–78	40–68	31–82	31–69	49–82	50–67	53–61		34–76	
Gender, *n* (%)													0.001
Male	108 (54.0)	11 (35.5)	4	6	1	46 (47.9)	28	12	3	2	1	51 (69.9)	
Female	92 (46.0)	20 (64.5)	7	5	8	50 (52.1)	44	5	1	0	0	22 (30.1)	
Smoking (pack year), *n* (%)													<0.0001
Never	88 (44.0)	22 (71.0)	5	9	8	46 (47.9)	38	6	2	0	0	20 (27.4)	
<20	26 (13.0)	3 (9.7)	3	0	0	18 (18.8)	18	0	0	0	0	5 (6.8)	
≥20	86 (43.0)	6 (19.3)	3	2	1	32 (33.3)	16	11	2	2	1	48 (65.8)	
TNM stage, *n* (%)													0.009
I	76 (38.0)	14 (45.2)	2	8	4	32 (33.3)	24	5	2	1	0	30 (41.1)	
II	59 (29.5)	9 (29.0)	6	2	1	21 (21.9)	12	8	0	0	1	29 (39.7)	
III	56 (28.0)	8 (25.8)	3	1	4	35 (36.5)	29	3	2	1	0	13 (17.8)	
IV	9 (4.5)	0	0	0	0	8 (8.3)	7	1	0	0	0	1 (1.4)	
Differentiation, *n* (%)													0.0007
Well	90 (45.0)	14 (45.2)	2	7	5	46 (47.9)	34	11	1	0	0	30 (41.1)	
Moderate	74 (37.0)	14 (45.2)	7	4	3	41 (42.7)	32	4	3	1	1	20 (27.4)	
Poor	34 (17.0)	2 (6.5)	1	0	1	8 (8.3)	5	2	0	1	0	23 (31.5)	
Unknown	2 (1.0)	1 (3.1)	1	0	0	1 (1.0)	1	0	0	0	0	0	

NOTE: Kruskal–Wallis test or two-sided χ^2 test, where appropriate.
[a]A *P* value was derived from the comparison among driver fusion, driver mutation, and pan-negative cases.

0.0001 by Kruskal–Wallis test; Fig. 3A) than the other cases, and most of the SMGs were more highly mutated in these cases than in the other groups (Supplementary Table S5). Among 127 SMGs, *TSHZ3*, *SETBP1*, *EPHA3*, and *NAV3* were preferentially mutated in pan-negative cases (Table 2). In addition, truncating mutations in SWI/SNF chromatin remodeling genes, such as *ARID1A* and *SMARCA4/BRG1*, which have been reported to be mutated in LADC (14, 16), and *PBRM*, another SWI/SNF gene frequently mutated in renal cell carcinoma (ccRCC; ref. 42), were significantly predominant in pan-negative cases ($P < 0.0001$ by Kruskal–Wallis test; Fig. 3B and Supplementary Fig. S3B; and Supplementary Table S5). On the other hand, *KEAP1*, *NF1*, and *RIT1* mutations, which were previously shown to occur frequently in pan-negative tumors in the TCGA cohort (16), were not frequent in our cases (Supplementary Table S6). Therefore, there might be a difference in the carcinogenic pathways of pan-negative cases between Asians and Europeans/Americans.

Table 2. Top 21 frequently mutated SMGs in the study cohort

	All cases (*n* = 200)	**Fusion (*n* = 31)**	**Mutation (*n* = 96)**	**Pan-negative (*n* = 73)**
Genes, *n* (%)				
TP53	75 (37.5)	5 (16.1)	41 (42.7)	29 (39.7)
EGFR	72 (36.0)	0	72 (75.0)	0
KRAS	17 (8.5)	0	17 (17.7)	0
NAV3	14 (7.0)	0	5 (5.2)	9 (12.3)
APC	13 (6.5)	2 (6.5)	8 (8.3)	1 (1.4)
CDKN2A	13 (6.5)	0	3 (3.1)	10 (13.7)
STK11	12 (6.0)	0	4 (4.2)	8 (11.0)
CTNNB1	11 (5.5)	2 (6.5)	8 (8.3)	1 (1.4)
EPHA3	11 (5.5)	0	3 (3.1)	8 (11.0)
TSHZ2	11 (5.5)	0	5 (5.2)	6 (8.2)
TSHZ3	11 (5.5)	0	2 (2.1)	9 (12.3)
MLL2	10 (5.0)	2 (6.5)	2 (2.1)	6 (8.2)
SETBP1	10 (5.0)	0	3 (3.1)	7 (9.6)
MLL3	9 (4.5)	0	4 (4.2)	5 (6.8)
SETD2	9 (4.5)	3 (9.7)	2 (2.1)	4 (5.5)
ARID1A	8 (4.0)	0	0	8 (11.0)
ATM	8 (4.0)	0	2 (2.1)	6 (8.2)
ATRX	8 (4.0)	0	3 (3.1)	5 (6.8)
EPPK1	8 (4.0)	1 (3.2)	0	7 (9.6)
PDGFRA	8 (4.0)	0	5 (5.2)	3 (4.1)
POLQ	8 (4.0)	0	3 (3.1)	5 (6.8)

Some pan-negative cases did not show mutations in cancer-related genes (Fig. 2). A small number of mutations were associated with cases with low tumor content, suggesting that the failure to detect mutations in some pan-negative cases was due to the low mutation detection power in tumors with low purity (Supplementary Fig. S5). To further address this point, we subjected 29 pan-negative cases, in which few or no cancer-related gene mutations were detected, to deep resequencing, in the same way as we did for fusion-positive cases. In addition to the nonsynonymous mutations already detected by exome sequencing, we additionally detected several mutations, including those with a lower mutation allele frequency than expected from tumor content (Supplementary Table S8). Therefore, some of the pan-negative tumors most likely had intratumor heterogeneity that hampered the detection of mutations by exome sequencing.

Discussion

This study compared gene aberrations based on driver gene status in a cohort enriched in LADCs with oncogenic *ALK*, *RET*, and *ROS1* fusions. The fusion-positive cases showed significantly fewer mutations than the other cases in all genes and in known cancer-related genes represented by the CGC genes and SMGs. The lower mutation rate was independent of age, gender, smoking status, pathologic stage, and tumor differentiation, and was validated in a LADC cohort consisting of 230 selected U.S. patients (Supplementary Fig. S4; ref. 16). The rate of nonsynonymous mutation in driver fusion LADCs was similar to that in ovarian, breast, brain, kidney, and hematopoietic tumors, for which mutation rates are low (20, 43, 44). Notably, fusion-positive LADCs had a lower frequency of C>A transversion, which is predominant

in tumors of ever-smokers and is a signature of cigarette smoke exposure, than others (Supplementary Fig S6A; refs. 20, 43, 44). Thus, LADCs with oncogene fusions develop through a distinct pathway that includes fewer gene aberrations than other LADCs. Interestingly, similar differential genomic profiles were observed in *PAX* fusion-positive and -negative rhabdomyosarcomas (45). This might be due to the strong ability of oncogene fusions to drive carcinogenesis; therefore, tumor cells with such fusions might not need many other genetic alterations to progress. The fact that LADC patients with oncogene fusions are younger than those without fusions in this and other populations (Table 1) might also reflect the oncogenic robustness of these fusions (2, 8–10, 38).

By contrast, pan-negative LADCs appear to develop from a large number of gene mutations, including mutations in *TP53* and other cancer-related genes. In addition, pan-negative cases carry truncating mutations in several SWI/SNF chromatin remodeling genes more frequently than other cases. Notably, truncating mutations in *PBRM1* were detected only in pan-negative cases. Truncating mutations in *PBRM1* and other SWI/SNF genes are frequent in ccRCC (42); therefore, some pan-negative LADC and ccRCC might develop through a common carcinogenic pathway. Smoking is linked to high numbers of mutations (Supplementary Fig. S6B and S6C), suggesting that the large numbers of gene mutations caused by exposure to tobacco carcinogens can lead to tumor development without any known driver oncogene alteration. In addition, a subset of pan-negative tumors were found to have high intratumor heterogeneity that hampered mutation detection by the exome sequencing method. Thus, the mutation profiles of pan-negative tumors require further investigation by more sensitive methods, such as deep whole-exome sequencing, to understand how these tumors develop.

The present study has implications for therapeutic approaches to fusion-positive LADCs. Personalized therapy targeting fusion products with TKIs has become the first-line therapeutic method in advanced and/or recurrent tumors (12, 46, 47). This study indicates that targeting fusion products is the best approach as only a small number of mutations have occurred in other genes. Notably, driver fusion cases (and also driver mutation cases) lack SWI/SNF chromatin remodeling gene aberrations. Recent studies demonstrated that deficiencies in *SMARCA4* and *ARID1A* make cancer cells treatable by inhibiting the activity of their paralogs, *SMARCA2* and *ARID1B*, based on synthetic lethality (23, 24, 28); however, LADC patients with driver fusions will not benefit from such therapies because their tumors have a low frequency of aberrations in *SMARCA4* and *ARID1A* genes. Notably, a small subset of fusion-positive LADCs carry gene mutations in genes, such as *PIK3CA*, *APC*, and *CTNNB1*, that could affect signal transduction, although how prevalent or specific these mutations are to fusion-positive LADCs remains unclear due to a small number of study subjects. This issue should be further investigated to unravel the carcinogenic pathways that lead to fusion-positive LADC and find efficient therapeutic targets for this type of cancer.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: M. Saito, J. Yokota, T. Kohno
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): Y. Shimada, K. Shiraishi, H. Sakamoto, K. Tsuta, H. Ichikawa, S.-i. Watanabe, T. Yoshida
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): M. Saito, K. Shiraishi, H. Totsuka, S. Chiku, M. Kato, T. Kohno
Writing, review, and/or revision of the manuscript: M. Saito, T. Kohno
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): Y. Shimada, H. Sakamoto, K. Tsuta, T. Yoshida
Study supervision: J. Yokota

Acknowledgments

The authors thank Sumiko Ohnami, Yoko Odaka, and Misuzu Okuyama for technical assistance.

Grant Support

This work was supported by the Advanced Research for Medical Products Mining Program of the National Institute of Biomedical Innovation (NIBIO); grants-in-aid from the Japan Agency for Medical Research and Development (AMED) for the Health, and Labor Sciences Research Expenses for Commission (the Practical Research for Innovative Cancer Control: H26-practical-general-007); a grant-in-aid from the Japanese Society for the Promotion of Science (JSPS) and Scientific Research (B: 26293200 and C: 15K10275), and the National Cancer Center Research and Development Fund (26A-1).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received November 10, 2014; revised March 2, 2015; accepted March 13, 2015; published OnlineFirst April 8, 2015.

References

1. Shaw AT, Hsu PP, Awad MM, Engelman JA. Tyrosine kinase gene rearrangements in epithelial malignancies. Nat Rev Cancer 2013;13:772–87.
2. Takeuchi K, Soda M, Togashi Y, Suzuki R, Sakata S, Hatano S, et al. RET, ROS1, and ALK fusions in lung cancer. Nat Med 2012;18:378–81.
3. Lipson D, Capelletti M, Yelensky R, Otto G, Parker A, Jarosz M, et al. Identification of new ALK and RET gene fusions from colorectal and lung cancer biopsies. Nat Med 2012;18:382–4.
4. Kohno T, Ichikawa H, Totoki Y, Yasuda K, Hiramoto M, Nammo T, et al. KIF5B-RET fusions in lung adenocarcinoma. Nat Med 2012;18:375–7.
5. Shaw AT, Yeap BY, Mino-Kenudson M, Digumarthy SR, Costa DB, Heist RS, et al. Clinical features and outcome of patients with non–small cell lung cancer who harbor EML4-ALK. J Clin Oncol 2009;27:4247–53.
6. Pan Y, Zhang Y, Li Y, Hu H, Wang L, Li H, et al. ALK, ROS1, and RET fusions in 1139 lung adenocarcinomas: a comprehensive study of common and fusion pattern-specific clinicopathologic, histologic, and cytologic features. Lung Cancer 2014;84:121–6.
7. Yoshida A, Kohno T, Tsuta K, Wakai S, Arai Y, Shimada Y, et al. ROS1-rearranged lung cancer: a clinicopathologic and molecular study of 15 surgical cases. Am J Surg Pathol 2013;37:554–62.
8. Tsuta K, Kohno T, Yoshida A, Shimada Y, Asamura H, Furuta K, et al. RET-rearranged non–small cell lung carcinoma: a clinicopathological and molecular analysis. Br J Cancer 2014;110:1571–8.
9. Bergethon K, Shaw AT, Ou SH, Katayama R, Lovly CM, McDonald NT, et al. ROS1 rearrangements define a unique molecular class of lung cancers. J Clin Oncol 2012;30:863–70.
10. Wang R, Hu H, Pan Y, Li Y, Ye T, Li C, et al. RET fusions define a unique molecular and clinicopathologic subtype of non–small cell lung cancer. J Clin Oncol 2012;30:4352–9.

11. Awad MM, Katayama R, McTigue M, Liu W, Deng YL, Brooun A, et al. Acquired resistance to crizotinib from a mutation in CD74-ROS1. N Engl J Med 2013;368:2395–401.
12. Drilon A, Wang L, Hasanovic A, Suehara Y, Lipson D, Stephens P, et al. Response to Cabozantinib in patients with RET fusion-positive lung adenocarcinomas. Cancer Discov 2013;3:630–5.
13. Shaw AT, Kim DW, Nakagawa K, Seto T, Crino L, Ahn MJ, et al. Crizotinib versus chemotherapy in advanced ALK-positive lung cancer. N Engl J Med 2013;368:2385–94.
14. Imielinski M, Berger AH, Hammerman PS, Hernandez B, Pugh TJ, Hodis E, et al. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. Cell 2012;150:1107–20.
15. Clinical Lung Cancer Genome P, Network Genomic M. A genomics-based classification of human lung tumors. Sci Translational Med 2013;5: 209ra153.
16. Cancer Genome Atlas Research N. Comprehensive molecular profiling of lung adenocarcinoma. Nature 2014;511:543–50.
17. Gainor JF, Shaw AT. Emerging paradigms in the development of resistance to tyrosine kinase inhibitors in lung cancer. J Clin Oncol 2013;31:3987–96.
18. Chong CR, Janne PA. The quest to overcome resistance to EGFR-targeted therapies in cancer. Nat Med 2013;19:1389–400.
19. Futreal PA, Coin L, Marshall M, Down T, Hubbard T, Wooster R, et al. A census of human cancer genes. Nat Rev Cancer 2004;4:177–83.
20. Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, et al. Mutational landscape and significance across 12 major cancer types. Nature 2013; 502:333–9.
21. Muller PA, Vousden KH. Mutant p53 in cancer: new functions and therapeutic opportunities. Cancer Cell 2014;25:304–17.
22. Khoo KH, Verma CS, Lane DP. Drugging the p53 pathway: understanding the route to clinical efficacy. Nat Rev Drug Discov. 2014;13:217–36.
23. Helming KC, Wang X, Wilson BG, Vazquez F, Haswell JR, Manchester HE, et al. ARID1B is a specific vulnerability in ARID1A-mutant cancers. Nat Med 2014;20:251–4.
24. Oike T, Ogiwara H, Tominaga Y, Ito K, Ando O, Tsuta K, et al. A synthetic lethality-based strategy to treat cancers harboring a genetic deficiency in the chromatin remodeling factor BRG1. Cancer Res 2013;73:5508–18.
25. Wilson BG, Helming KC, Wang X, Kim Y, Vazquez F, Jagani Z, et al. Residual complexes containing SMARCA2 (BRM) underlie the oncogenic drive of SMARCA4 (BRG1) mutation. Mol Cell Biol 2014;34:1136–44.
26. Liu Y, Marks K, Cowley GS, Carretero J, Liu Q, Nieland TJ, et al. Metabolic and functional genomic studies identify deoxythymidylate kinase as a target in LKB1-mutant lung cancer. Cancer Discov 2013;3:870–9.
27. Fang B. Development of synthetic lethality anticancer therapeutics. J Med Chem 2014;57:7859–73.
28. Hoffman GR, Rahal R, Buxton F, Xiang K, McAllister G, Frias E, et al. Functional epigenetics approach identifies BRM/SMARCA2 as a critical synthetic lethal target in BRG1-deficient cancers. Proc Natl Acad Sci U S A 2014;111:3128–33.
29. Goldstraw P, Crowley J, Chansky K, Giroux DJ, Groome PA, Rami-Porta R, et al. The IASLC lung cancer staging project: proposals for the revision of the TNM stage groupings in the forthcoming (seventh) edition of the TNM classification of malignant tumours. J Thorac Oncol 2007;2:706–14.
30. Groome PA, Bolejack V, Crowley JJ, Kennedy C, Krasnik M, Sobin LH, et al. The IASLC Lung Cancer Staging Project: validation of the proposals for revision of the T, N, and M descriptors and consequent stage groupings in the forthcoming (seventh) edition of the TNM classification of malignant tumours. J Thorac Oncol 2007;2:694–705.
31. Li A, Liu Z, Lezon-Geyda K, Sarkar S, Lannin D, Schulz V, et al. GPHMM: an integrated hidden Markov model for identification of copy number alteration and loss of heterozygosity in complex tumor samples using whole genome SNP arrays. Nucleic Acids Res 2011;39:4928–41.
32. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol 2013;31:213–9.
33. BROAD Institute. Available from: https://www.broadinstitute.org/gatk/
34. BROAD Institute. Available from: http://www.broadinstitute.org/igv/
35. HGNC (HUGO Gene Nomenclature Committee). Available from: http://www.genenames.org/
36. COSMIC (Catalogue of somatic mutations in cancer). Available from: https://cancer.sanger.ac.uk/cancergenome/projects/cosmic/
37. Cancer Genome Atlas Research N. Comprehensive genomic characterization of squamous cell lung cancers. Nature 2012;489:519–25.
38. Kohno T, Tsuta K, Tsuchihara K, Nakaoku T, Yoh K, Goto K. RET fusion gene: translation to personalized lung cancer therapy. Cancer Sci 2013; 104:1396–400.
39. Vaishnavi A, Capelletti M, Le AT, Kako S, Butaney M, Ercan D, et al. Oncogenic and drug-sensitive NTRK1 rearrangements in lung cancer. Nat Med 2013;19:1469–72.
40. Wilson BG, Roberts CW. SWI/SNF nucleosome remodellers and cancer. Nat Rev Cancer 2011;11:481–92.
41. Nakanishi H, Matsumoto S, Iwakawa R, Kohno T, Suzuki K, Tsuta K, et al. Whole genome comparison of allelic imbalance between noninvasive and invasive small-sized lung adenocarcinomas. Cancer Res 2009;69:1615–23.
42. Varela I, Tarpey P, Raine K, Huang D, Ong CK, Stephens P, et al. Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma. Nature 2011;469:539–42.
43. Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature 2013;499:214–8.
44. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, et al. Signatures of mutational processes in human cancer. Nature 2013; 500:415–21.
45. Shern JF, Chen L, Chmielecki J, Wei JS, Patidar R, Rosenberg M, et al. Comprehensive genomic analysis of rhabdomyosarcoma reveals a landscape of alterations affecting a common genetic axis in fusion-positive and fusion-negative tumors. Cancer Discov 2014;4:216–31.
46. Gerber DE, Minna JD. ALK inhibition for non–small cell lung cancer: from discovery to therapy in record time. Cancer Cell 2010;18:548–51.
47. Davies KD, Le AT, Theodoro MF, Skokan MC, Aisner DL, Berge EM, et al. Identifying and targeting ROS1 gene fusions in non–small cell lung cancer. Clin Cancer Res 2012;18:4570–9.