Published OnlineFirst September 13, 2016; DOI: 10.1158/0008-5472.CAN-16-0873

Cancer Research

Molecular Evolution Patterns in Metastatic Lymph Nodes Reflect the Differential Treatment Response of Advanced Primary Lung Cancer

Sang-Won Um[1], Je-Gun Joung[2], Hyun Lee[1], Hojoong Kim[1], Kyu-Tae Kim[2], Jinha Park[1], D. Neil Hayes[3], and Woong-Yang Park[2,4]

Abstract

Tumor heterogeneity influences the clinical outcome of patients with cancer, and the diagnostic method to measure the tumor heterogeneity needs to be developed. We analyzed genomic features on pairs of primary and multiple metastatic lymph nodes from six patients with lung cancer using whole-exome sequencing and RNA sequencing. Although somatic single-nucleotide variants were shared in primary lung cancer and metastases, tumor evolution predicted by the pattern of genomic alterations was matched to anatomic location of the tumors. Four of six cases exhibited a branched clonal evolution pattern. Lymph nodes with acquired somatic variants demonstrated resistance to the cancer treatment. In this study, we demonstrated that multiple biopsies and sequencing strategies for different tumor regions are required for a comprehensive understanding of the landscape of genetic alteration and for guiding targeted therapy in advanced primary lung cancer. *Cancer Res; 76(22); 6568–76. ©2016 AACR.*

Introduction

Lung cancer is one of the most common malignancies in the world, accounting for about 18% of all cancer-related deaths (1). High lethality of advanced lung cancer may be closely involved with its heterogeneous nature, presenting mostly in advanced stages with a poor survival rate of less than 5% (2). Intratumoral heterogeneity has been one of the major difficulties in treating lung cancer successfully. Clonal diversity in lung cancer confers the possibility of variable treatment responses in patients (3). Comprehensive analysis on multiregional samples helps to understand the tumor heterogeneity, but the multiregional sample acquisition in patients with refractory cancer is clinically limited. Instead, we can utilize the multiple lymph node samples harboring subpopulation of the primary tumor. Lung cancers often spread first to lymph nodes near the tumor, known as hilar or interlobar lymph nodes (N1), but they may also spread to ipsilateral (N2) or contralateral (N3) mediastinal lymph nodes (4). Lymph node metastasis in lung cancer is regarded as a strong independent predictor of poor prognosis (4). Often primary tumor location can determine whether lung cancer will spread to the lymph nodes.

Understanding pathologic and molecular evolutionary processes of lung cancer may provide greater insight toward development of a suitable therapeutic guide (5). Although many mutations detected in single biopsies of primary tumors are common to several regions, other mutations are present only in specific regions, suggesting clonal evolution. Thus, intratumor heterogeneity (ITH) and intermetastatic heterogeneity are important to consider for tumor biopsy strategies and treatment plans (6). Recent works examining genomic tumor heterogeneity in lung cancer have mainly focused on how tumor heterogeneity and tumor evolution are linked to the anticancer response due to genomic alterations (7). However, the genomic heterogeneity between primary tumor and lymph node metastases still remains unclear.

In this study, we characterized genomic alterations and their associations with treatment response by applying whole-exome sequencing (WES) and RNA sequencing for multiple tumor regions from 6 patients who were suspected to have primary lung cancer with multiple lymph node metastases or distant metastasis (Table 1). To evaluate the tumor heterogeneity of somatic alterations, we classified mutations as trunk mutations presenting in all tumor regions or non-trunk mutations (private and shared) in specific region or subregions implying heterogeneous evolution. Our study demonstrates subclonal diversity of multiregional tumors and explores the connection between clonal patterns and treatment response.

Materials and Methods

Patients and sample preparation

This study was approved by the Institutional Review Board (IRB) of Samsung Medical Center (IRB approval no. SMC 2014-03-140). Written informed consent was obtained from each

[1]Division of Pulmonary and Critical Care Medicine, Department of Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea. [2]Samsung Genome Institute, Samsung Medical Center, Seoul, Korea. [3]Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina. [4]Department of Molecular Cell Biology, Sungkyunkwan University School of Medicine, Seoul, Korea.

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

S.-W. Um and J.-G. Joung contributed equally to this article.

Corresponding Authors: Woong-Yang Park, Samsung Genome Institute, Samsung Medical Center, 81 Irwon-ro, Gangnam-gu, Seoul 06351, Republic of Korea (South). Phone: 82-2-2148-9810; Fax: 82-2-2148-9819; E-mail: woongyang@skku.edu; and Sang-Won Um, sangwonum@skku.edu

doi: 10.1158/0008-5472.CAN-16-0873

Published OnlineFirst September 13, 2016; DOI: 10.1158/0008-5472.CAN-16-0873

Table 1. Summary of clinical features of six patients with lung cancer patients

Case #	Sex/age	Smoking	Histology	Analyzed lesions	Clinical stage
TH1	M/74	50 PY	SCLC	RML, 4R (N2), 7 (N2), 4L (N3)	T2bN3M0
TH2	M/80	6 PY	NSCLC, ADC	LLL, 4L (N2), liver (M1b)	T1bN3M1b
TH3	M/62	90 PY	NSCLC, SCC	RUL, 11Rs (N1), 4R (N2)	T2aN3M0
TH4	M/48	30 PY	NSCLC, SCC	RBI, 11Rs (N1), 7 (N2), 4L (N3)	T2aN3M0
TH5	F/66	Never	NSCLC, ADC	LUL, 11L (N1), 7 (N2), 4R (N3)	T1bN3M1b
TH6	M/64	30 PY	NSCLC, SCC	LUL, 4L (N2), 7 (N2), 2R (N3)	T4N3M1b

Abbreviations: ADC, adenocarcinoma; SCC, squamous cell carcinoma; RUL, right upper lobe; LLL, left lower lobe; LUL, left upper lobe; M, male; F, female; PY, pack-year.

patient. The study subjects were 6 Korean patients diagnosed with advanced lung cancer at Samsung Medical Center (Seoul, Korea). The included cases met the following criteria: patients were suspected to have primary lung cancer with multiple lymph node metastases and no previous treatment. Each case includes a primary tumor sample from bronchoscopic biopsy, hilar/interlobar or mediastinal lymph node (N1, N2, and N3) samples from endobronchial ultrasound guided-transbronchial needle aspiration (EBUS-TBNA), normal bronchial mucosal tissue from bronchoscopic biopsy, and peripheral blood (Supplementary Table S3). Biopsy samples were snap frozen and stored in liquid nitrogen until use. The study included 5 patients with non–small cell lung cancer (NSCLC) and 1 patient with small-cell lung cancer (SCLC; Table 1). Histology, clinical stage, and analyzed lesions are summarized in Table 1.

Isolation of genomic DNA and RNA

Genomic DNA and RNA in tissues were purified using ALLPrep DNA/RNA Mini Kit (Qiagen). Genomic DNA from peripheral blood was extracted by QIAamp DNA Blood Mini Kit (Qiagen). Genomic DNA concentration and purity were measured by a Nanodrop 8000 UV-Vis spectrometer (Thermo Scientific Inc.) and Qubit 2.0 Fluorometer (Life Technologies Inc.). To estimate DNA degradation, DNA median size and ΔC_t values were measured with a 2200 TapeStation Instrument (Agilent Technologies) and real-time PCR (Agilent Technologies), respectively. For RNA, the concentration and purity was measured by Nanodrop and Bioanalyzer (Agilent Technologies).

Whole-exome sequencing

Genomic DNA (1 μg) from each sample was sheared by Covaris S220 (Covaris) and used for the construction of a library using SureSelect XT Human All Exon v5 and a SureSelect XT Reagent Kit, HSQ (Agilent Technologies) according to the manufacturer's protocol. This kit is designed to enrich 335,756 exons of 21,058 genes, covering approximately 71 Mb of the human genome. After enriched exome libraries were multiplexed, the libraries were sequenced on a HiSeq 2500 sequencing platform (Illumina). Briefly, a paired-end DNA sequencing library was prepared through gDNA shearing, end-repair, A-tailing, paired-end adaptor ligation, and amplification. After hybridization of the library with bait sequences for 16 hours, the captured library was purified and amplified with an index barcode tag, and library quality and quantity were measured. Sequencing of the exome library was carried out using the 100-bp paired-end mode of the TruSeq Rapid PE Cluster kit and TruSeq Rapid SBS kit (Illumina).

Exome-seq data analysis

Sequencing reads were aligned to the UCSC hg19 reference genome (downloaded from http://genome.ucsc.edu) using Burrows-Wheeler Aligner (BWA; ref. 8), version 0.6.2 with default settings. PCR duplications were marked by Picard-tools-1.8 (http://picard.sourceforge.net/), data cleanup was followed by GATK and variants were identified with GATK-2.2.9 (https://www.broadinstitute.org/gatk/; Supplementary Table S4). Then, point mutations were identified by MuTect (https://github.com/broadinstitute/mutect) and VarScan 2 (http://varscan.sourceforge.net) with paired samples. Perl script and ANNOVAR (9) were used to annotate variants. Copy number variations were identified by EXCAVATOR (10). The phylogenetic analysis was performed on the basis of the binary sequence of somatic mutation profiles. The phylogenetic tree was obtained by the Wagner parsimony method.

RNA sequencing

Library construction for whole transcriptome sequencing was performed using a TruSeq RNA Sample Preparation v2 Kit (Illumina). Isolated total RNA was used in a reverse transcription reaction with poly (dT) primers using SuperScriptTM II reverse transcriptase (Invitrogen/Life Technologies) according to the manufacturer's protocol. Briefly, an RNA sequencing library was prepared through cDNA amplification, end-repair, 3′end adenylation, adapter ligation, and amplification. Quality and quantity of the library were measured by Bioanalyzer and Qubit. Sequencing of the transcriptome library was carried out using the 100-bp paired-end mode of the TruSeq Rapid PE Cluster kit and the TruSeq Rapid SBS Kit (Illumina).

RNA sequencing data analysis

Reads from the FASTQ files were mapped against the hg19 human reference genome by using TopHat version 2.0.6 (http://tophat.cbcb.umd.edu/; Supplementary Table S5). Raw read counts mapped into genes were measured with the BAM format file by HTSeq version 0.6.1 (11). Then, transcript abundance was measured for 18,161 coding genes, and then scarcely expressed genes were filtered out by the criteria of maximum read count >10 across all samples of each case. Read counts were normalized by the TMM (trimmed mean of M-values) normalization method (Supplementary Table S6). The differentially expressed genes were identified using the DESeq R package (www-huber.embl.de/users/anders/DESeq/). Fusion genes were detected by RNA-sequencing and Cancer Panel sequencing and they were visualized in genome regions using the IGV (Integrative Genome Viewer) tool (http://www.broadinstitute.org/igv/). The gene-set variance analysis (12) was performed with REACTOME gene sets to identify enriched pathways in each sample. Then, differential gene sets between sample groups were identified using limma R package.

Clonality analysis

Subclones were obtained by clustering cancer cell fractions by PyClone, which can deconvolute the tumor into subclones using a hierarchical Bayesian clustering model (13). The input data were generated with somatic single-nucleotide variations (SNV)

Published OnlineFirst September 13, 2016; DOI: 10.1158/0008-5472.CAN-16-0873

detected by MuTect and copy number variations corresponding to SNVs. Step filtering the SNVs with large credible intervals in the previous stage was applied to remove those with noninformative posterior distributions. The subclones between primary tumors and each lymph node were compared using the density plot of cancer cell fractions.

Results

Genomic alterations of metastatic lymph nodes correlated with anatomic region

Somatic SNVs including missense, nonsense, and splicing were identified by WES on samples from 6 patients with lung cancer (TH1–TH6; Fig. 1 and Supplementary Fig. S1). Patients were presented sorted by the rate of concordance in somatic mutations between tumors of each case. While the number of somatic SNVs in each patient varied from 65 to 493 (Supplementary Table S1), the rate of concordance in somatic mutations between distinct lymph nodes and primary lung cancer was 71.6% ± 19.1% for the six cases. Comparing concordance rate between the primary tumor and each node, patient TH1′s SCLC was the most homogeneous while patient TH6's squamous cell carcinoma (SCC) was the most heterogeneous (Fig. 1C). Patient TH3 with new distant metastasis to kidney during the follow-up had twice as many mutations (> 400) as the other cases. The large number of mutations may be associated with quite tobacco smoking of this patient [90 pack-years (PY); ref. 14]. The trees of dynamic tumor evolution show patterns of intermetastatic genetic heterogeneity in metastatic lymph nodes (Fig. 1B).

Anatomic location of the primary tumor could provide clues to advanced lung cancer tumor evolution. Three lymph node metastases in patient TH1 shared most mutations with the primary tumor in the right middle lobe (Fig. 1B). This suggests that multiple lymph node metastases might have occurred at the

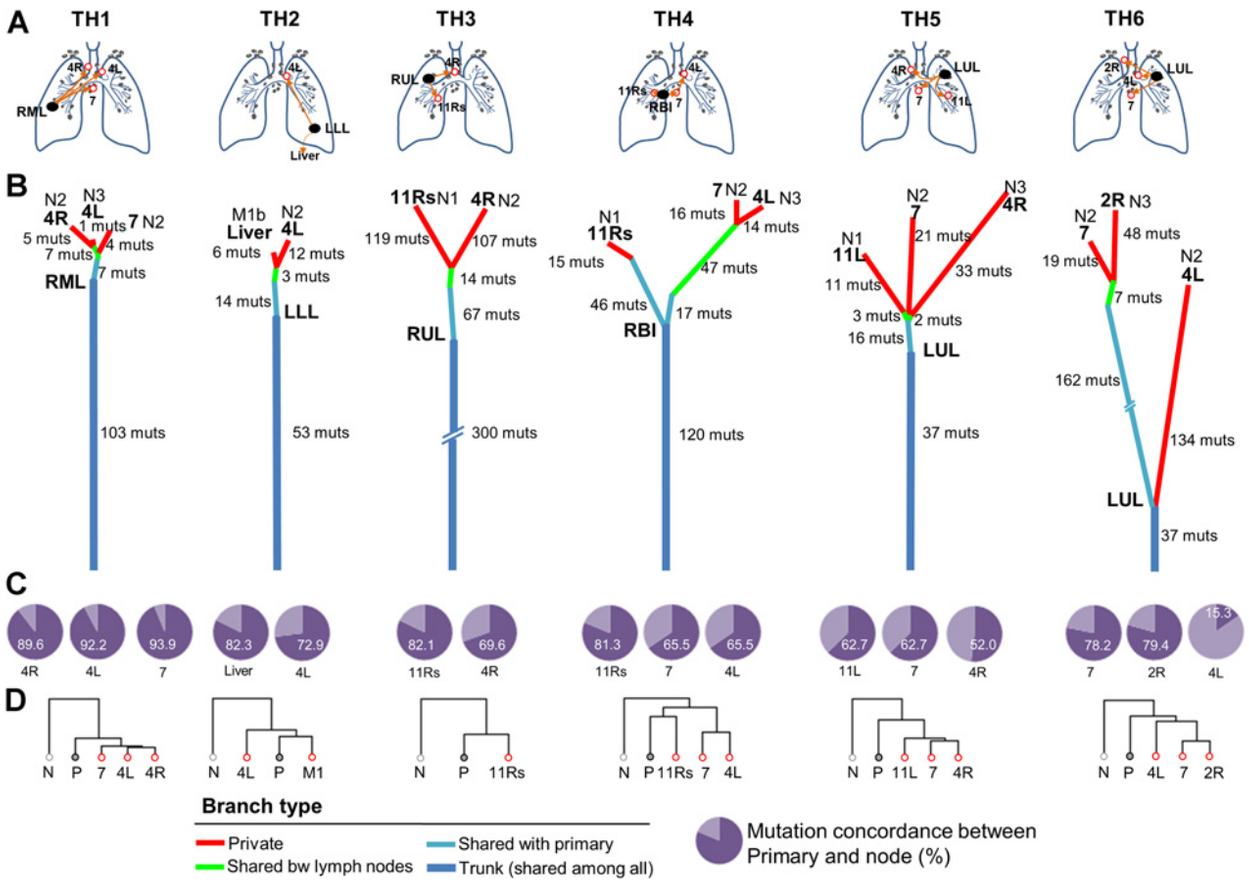

Figure 1.

The trees of dynamic tumor evolution representing patterns of intermetastatic mutational heterogeneity in lymph node metastases of advanced lung cancers. **A,** analyzed lesions for each case are displayed. The black and orange circles are primary lesions and metastasis sites, respectively. LUL, left upper lobe; RUL, right upper lobe; RML, right middle lobe; LLL, left lower lobe; RBI: right bronchus intermedius; 11L N1, left interlobar lymph node; 7 N2, subcarinal lymph node; 4L N2, left lower paratracheal lymph node; 4R N2, right lower paratracheal lymph node; 2R N3, right upper paratracheal lymph node; 4L N3, left lower paratracheal lymph node; 4R N3, right lower paratracheal lymph node; 11Rs N1, right superior interlobar lymph node; M1b, distant metastasis. **B,** branches in divergence trees include trunk mutations shared in all the tumors (thick blue line), mutations shared with primary site (thin blue line), shared mutations between lymph nodes (green line), and private mutations (red line). The number of somatic mutations (muts) in each branch is shown. **C,** tumor cases are listed according to the average concordance rate of mutations between primary site and lymph nodes. **D,** dendrogram representing the relative similarity between gene expression profiles. Cancer census genes (572) were used.

Published OnlineFirst September 13, 2016; DOI: 10.1158/0008-5472.CAN-16-0873

similar time points. The primary tumor site of patient TH4 was located in the right bronchus intermedius (RBI). Over half of the mutations in the primary tumor were shared with each lymph node. Among three lymph nodes, N1 shared the most mutations with the primary tumor and it is closely located near the primary site. Many mutations between N2 and N3 were concordant and N2 is closer to the primary tumor than N3. It is likely that metastasis of the N2 tumor occurred earlier than the N3 tumor. The dendrogram representing the relative similarity between gene expression profiles of cancer-related genes, generated by hierarchical clustering, show a pattern consistent with the branching pattern of mutations (Fig. 1D).

From the branching pattern of mutations, we analyzed mutation spectra analyzed for trunk and non-trunk mutations (Fig. 2A). A previous study reported that the proportion of C>A transversion was significantly decreased in late-occurring mutations from The Cancer Genome Atlas (TCGA) lung adenocarcinoma and lung squamous cell carcinoma smokers and former smokers (5). TH1 and TH3 patients showed typical mutation profile for heavy smokers, but their profiles are significantly decreased in non-trunk mutations ($P = 4.7 \times 10^{-5}$ and 5.1×10^{-3}, binomial test). The decreased proportion of C>A transversion is accompanied by an increased pattern of C>T and C>G mutations at TpCpW sites, which was reported to be induced by APOBEC cytidine deaminase (15).

Major mutations frequently observed in lung cancers were identified in six cases (Fig. 2B). TP53, ROS1, TSC2, SMARCA4, and ATM missense mutations, ARID1A nonsense mutations, and RB1 splicing mutations were commonly found in multiple regions of each case. However, many mutations including PIK3CA (H1047R) and HRAS (Q61R) were found in only specific regions. Both mutations were consistently found in the same samples in RNA sequencing (RNA-seq) data. Clinical sequencing on single biopsy of primary tumor might miss driver mutations present in low fractions of primary tumor. EML4-ALK fusion oncogenes were detected in TH2 (6 PY) and TH5 (never-smoking patient; Supplementary Fig. S2). This result supports previous studies on lung adenocarcinomas showing that EML4-ALK fusion was correlated with non- or light smoking (<10 PY; refs. 16, 17). Furthermore, EML4-ALK fusion is mutually exclusive for EGFR and KRAS mutations in patients with NSCLC (18–20) and this was also true for the patients in our study.

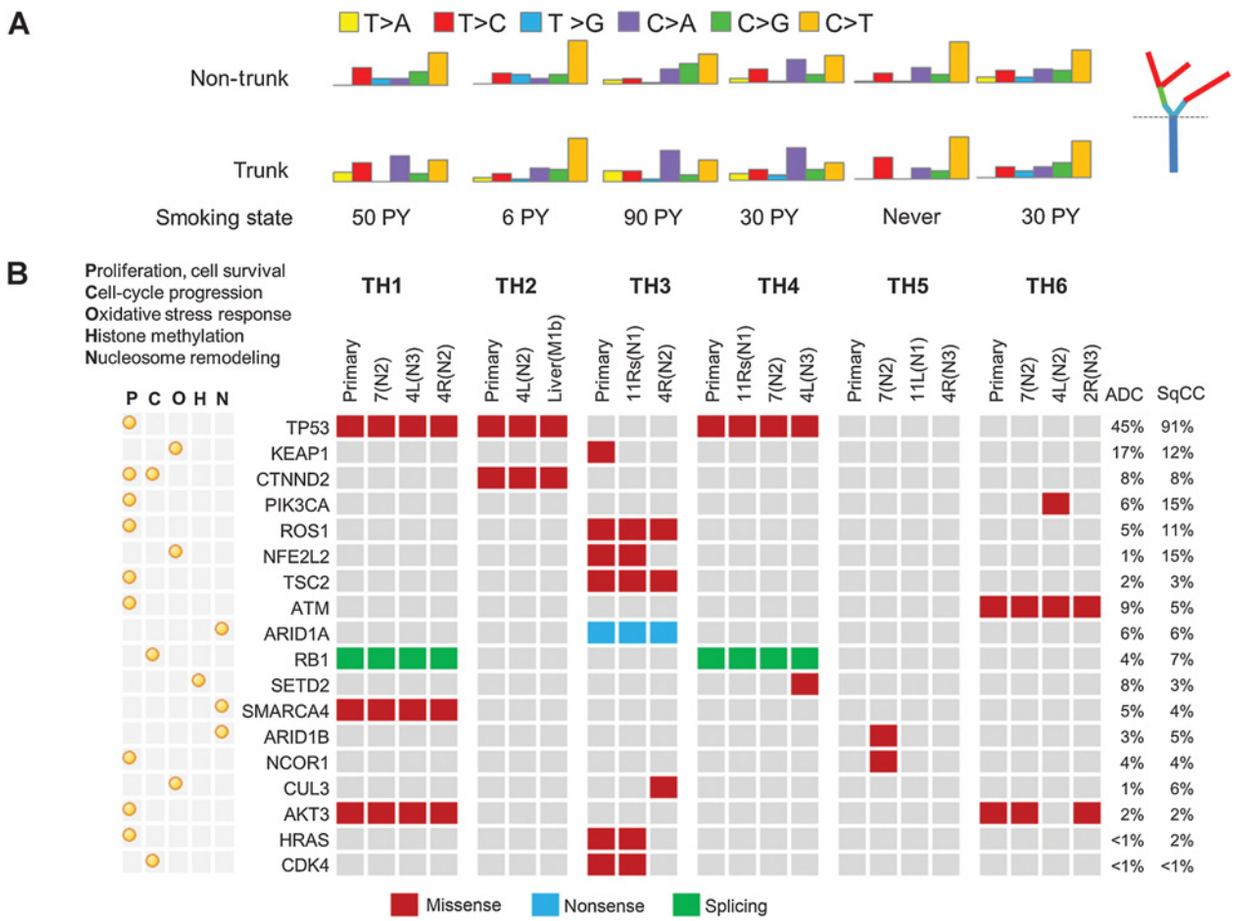

Figure 2.

The SNV spectrum and mutation profile. **A,** the spectrum of substitutions in trunk and non-trunk mutations. Each bar represents the proportion of each type of substitution. PY, pack-year. **B,** the heatmap of somatic mutations (including missense, nonsense, and splicing) detected in in lung cancer–associated genes. The listed genes were sorted by the mutation frequency observed in lung cancer cohorts of TCGA ADC and SCC.

Published OnlineFirst September 13, 2016; DOI: 10.1158/0008-5472.CAN-16-0873

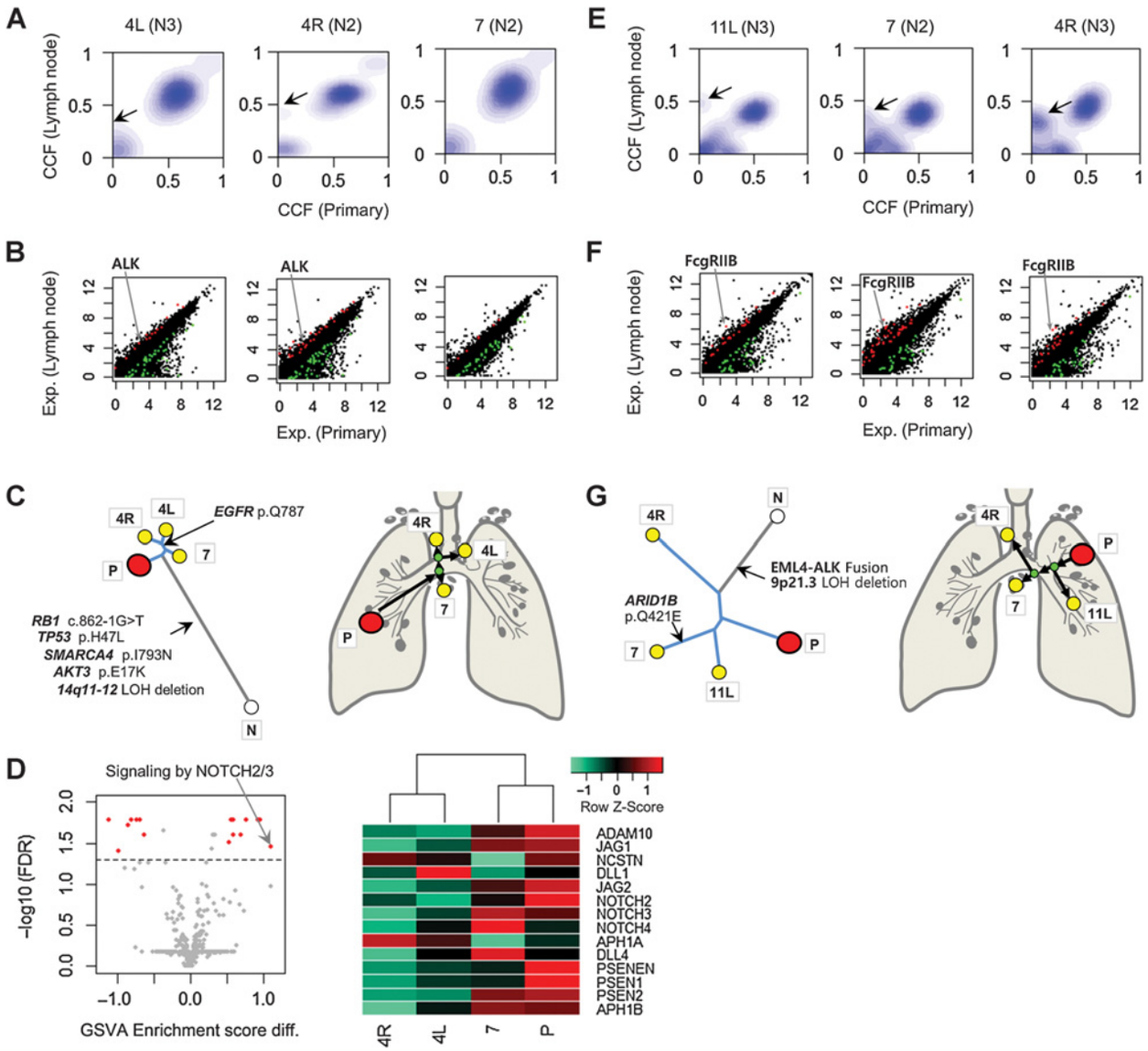

Figure 3.

The genomic and transcriptomic differences between multiple lymph nodes of patients TH1 (**A–D**) and TH5 (**E–G**). **A** and **E**, the density plots of cancer cell fraction (CCF) between the primary site and lymph node metastases. **B** and **F,** the scatter plots for similarity of gene expression profiles between the primary site and lymph node metastases. Differentially expressed cancer-related genes (>2-fold change) are highlighted in red and green. **C** and **G,** the phylogenetic trees were obtained on the basis of the binary sequence of somatic mutation profiles by the Wagner parsimony method. **D,** left, the volcano plots of gene-set variance analysis (GSVA) score difference. Significantly enriched gene sets are highlighted in red ($P_{adj} < 0.05$ and log ratio > ±0.5). Right, the heatmaps of gene expression. They include genes involved in "Signaling by NOTCH2/3" set for patient TH1.

Clonal evolution analysis from transcriptome profile of metastatic lymph nodes

Because lymph nodes obviously contain infiltrated nontumor cells, we need to select cancer-associated gene sets excluding genes related to stroma and lymphocytes for the correlation analysis. The primary tumor and three lymph nodes of patient TH1 had highly similar somatic SNV profiles (Fig. 1) and copy number alteration patterns (Fig. 4). However, minor differences were detected in clonality comparison (Fig. 3A). Unlike lymph node 7, 4L and 4R each have additional clones that might appear in late stage in tumorigenesis. Those clones contained common missense mutations of FAXDC2, HLCS, ADAM2, and ZNF585A in both nodes. While station 7 lymph node of patient TH1 did not exhibit the differential expression profile, gene expression in lymph nodes 4L and 4R was rather distinct to the primary tumor (Fig. 3B). ALK was highly expressed in 4L and 4R (>100-fold) compared with matched normal tissue, but not in lymph node 7. ALK was also highly expressed in 4L and 4R (>3-fold) and 7 (1.5-fold) comparing with primary cancer. Patient TH1 had RB1 (c.862-1G>T), TP53 (H47L), and AKT3 (E17K) mutations (Fig.

Published OnlineFirst September 13, 2016; DOI: 10.1158/0008-5472.CAN-16-0873

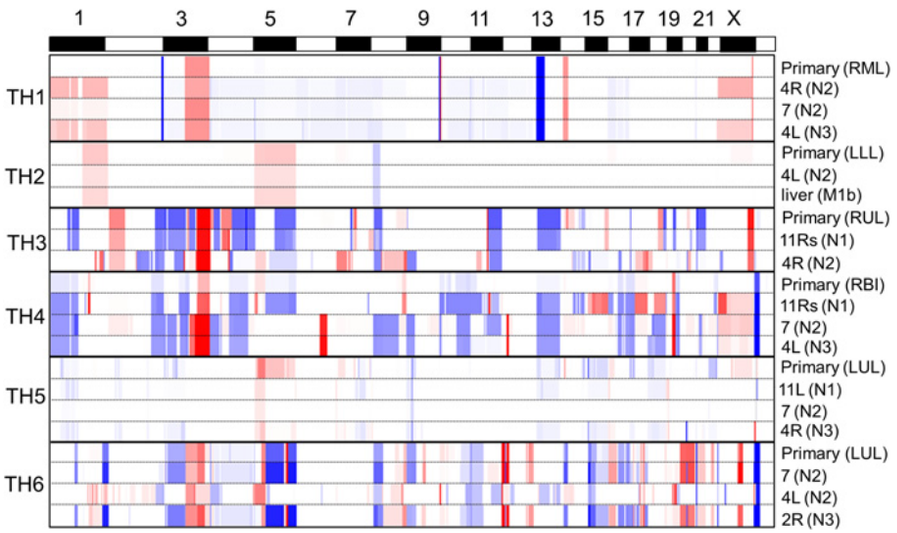

Figure 4.
The copy number alterations of primary and lymph nodes. Segmentation values ($\log_2$ ratios) are plotted. The amplification and deletion are presented by red and blue.

3C). It has been reported that RB1 is commonly mutated in SCLC (21). In addition, a silent EGFR mutation (Q787) was detected in 4L and 4R. On the basis of profile similarity, pathways showing differential expression patterns between group of 4L and 4R versus 7 were identified (Fig. 3D; Supplementary Table S2). Among them, genes belonging to "signaling by NOTCH2/3" pathway ($P_{adj} = 0.034$) showed lower levels of expression in nodes 4L and 4R than in 7. A previous study reported that activation of Notch signaling is associated with suppression of growth and arrest of the SCLC cells (22). Furthermore, Notch may be a key mediator of resistance to chemotherapy as well as radiotherapy (23).

TH5 had highly heterogeneous mutation profiles with additional subclones in three lymph nodes (Fig. 3E), and the gene expression profile of the primary tumor was different from each lymph node (Fig. 3F). Many cancer-related genes that exhibit high or low expression (> 2 fold change) were those expressed in human lymphocytes including Fc gamma receptor II-b (FcgRIIB/CD32B), IL21R, IL7R, and PTPRC. EML4-ALK fusion oncogenes were commonly detected in all tumors in TH5 patient and a novel ARID1B Q421E missense mutation was found in only lymph node 7 (Fig. 3G).

TH2 and TH3 had commonly additional subclones in metastatic tumors (Supplementary Fig. S3A and S3B). The p.L127V mutation in CUL3 known as a tumor suppressor in lung cancer (24) was detected in additional subclone of TH3 4R. Meanwhile TH4 and TH6 had additional subclones in specific metastatic tumors (Supplementary Fig. S3C and S3D). The expression patterns of lymph nodes were highly similar to corresponding clonal patterns. AKT2 was highly expressed in 4L of TH6 (>5.4-fold), but not in lymph node 7 and 2R (<1.0-fold). This gene plays an important role in lung cancer cell proliferation and colony growth (25). Significantly different activation status of pathways between lymph nodes and primary tumor was found (Supplementary Table S2).

Correlation of treatment responses with clonal evolution pattern of metastatic lymph nodes

Five of 6 patients in this study received treatment, and one patient, TH6, declined treatment. Clinical responses of 5 patients were quite personal to each treatment modality (Table 2). Patient TH1 was treated with etoposide/carboplatin #3 and radiation. Although mutation profiles of this patient were highly homogeneous for primary and multiple nodes, treatment responses showed a heterogeneous pattern. The primary tumor and lymph node 7 metastasis decreased in size while 4R and 4L lymph node metastases were not significantly responsive to treatment. It is worth noting that this result is consistent with the clonality pattern, as additional clones with ALK overexpression were present in 4R and 4L (Fig. 3E).

Patient TH2 had minor alterations in metastatic lymph nodes with trunk mutation like EML4-ALK fusion (Supplementary Fig. S2), which was target for crizotinib. Relatively homogeneous genomic alterations might explain the better response to the target therapy.

Patient TH3 was treated with concurrent chemoradiation therapy including weekly taxol/cisplatin. The patient did not respond to the treatment and had rapidly progressing systemic metastasis during follow-up. The patient's poor response to the treatment could be partly explained by higher load of mutation than other cases (Supplementary Fig. S1).

Patient TH4 was treated with concurrent chemoradiation therapy including weekly taxol/cisplatin. Complete resolution was seen at lymph node 4L and a decrease in primary tumor size was noticed. However, lymph node metastases 11Rs and 7 showed progression after an initial response. These heterogeneous progressions after treatment may be related to the heterogeneity of each lesion. Mutations in LPP and ABCB1 were found in additional subclones of nodes 11Rs and 7, respectively (Supplementary Table S1). Notably, P-glycoprotein ABCB1 is a known multidrug transporter and has been associated with drug resistance in human tumors (26). This may explain why lymph node 7 shows a unique treatment response, although its overall genomic and transcriptomic profiles are highly similar to those of lymph node 4L in this patient.

Patient TH5 was treated with gemcitabine/carboplatin #4, and the primary lesion responded to treatment while the three metastatic lymph nodes showed continuous tumor progression. Limited response of patient TH5 was associated with the high degree of tumor heterogeneity in metastatic lymph nodes, exhibiting additionally produced subclones (Fig. 1B).

Published OnlineFirst September 13, 2016; DOI: 10.1158/0008-5472.CAN-16-0873

Table 2. Summary of treatment modalities, clinical responses, and clonal evolution pattern in six patients

Case #	Treatment modality	Treatment response [clonal evolution pattern]
TH1	Etoposide/carboplatin #3 & radiotherapy	Heterogeneous response (primary tumor /7 lymph nodes: decrease in size but 4R /4L lymph nodes: no change in size) [overexpression of ALK in 4R and 4L]
TH2	Crizotinib for 9 months	Primary tumor, 4L lymph node, liver: decrease in size [homogeneous genomic alteration]
TH3	Concurrent chemoradiation therapy (weekly taxol/cisplatin #7)	Stable disease in primary tumor/lymph nodes but new metastasis in kidney [higher load of mutation]
TH4	Concurrent chemoradiation therapy (weekly taxol/cisplatin #7)	Heterogeneous progression (complete resolution of 4L lymph node and decrease in primary tumor but progression after initial response in 11Rs/7 lymph nodes) [additional subclones of 11Rs and 7 lymph node]
TH5	Gemcitabine/carboplatin #4	Heterogeneous progression (no progression in primary tumor but progression in 11L, 7, and 4R lymph nodes) [heterogeneous genomic alteration between multiple nodes]
TH6	The patient rejected any treatment	N.A.

Abbreviation: N.A., not applicable.

Evolution of distinct subclones in early metastasis to a distinct lymph node

The primary tumor of patient TH6 spread to lymph nodes 7, 4L, and 2R (Supplementary Fig. S4A). Two of these nodes (station 7 and 2R) have highly similar somatic mutation profiles, copy number alteration amplifications, and deletions, while 4L showed a distinctive alteration pattern (Fig. 4). Overall gene expression pattern of 4L is consistently different from those of other two nodes as measured by mutations, copy number variation (CNV), and clonality (Supplementary Fig. S4B and S4C). 4L is located relatively near the primary site (left upper lobe), yet the concordance of mutations with the primary tumor is very low (15.3% of primary mutations). Furthermore, this metastasis also showed low concordance rate with other lymph node metastases. This suggests that metastasis to 4L occurred in this patient at a relatively early stage of tumor progression compared with the other two metastases. The percentage of trunk mutations is only 8.2% of the total distinct mutations (Supplementary Fig. S5A). Among those mutations, a key mutation within a potential driver, ATM, was found. ATM plays a role as a regulator of the p53 signaling pathway, and it has been reported as a tumor suppressor gene in lung adenocarcinoma (27). ATM mutation may accelerate lung cancer formation by acting upstream of p53. Interestingly, only 4L lymph node has a PIK3CA driver mutation in this patient. PIK3CA H1047R mutation has been known to be one of most frequently observed PI3-kinase mutations in human cancers. Mutant PIK3CA proteins are known to activate the downstream signaling cascade and induce oncogenic transformation *in vitro* (28). NSCLCs harboring PIK3CA-activating mutations have shown sensitivity to the dual PIK3CA/mTOR inhibitor PI-103 (29). CNV pattern showed that 4L lymph node station had 10p15 and 19q13 chromosomal amplifications that appeared prominently, compared with other two lymph nodes. Amplification of Akt in human malignancies has been known to lead to hyperactivation of downstream signaling. 19q13 chromosomal amplifications, including AKT2, are associated with tumor aggressiveness (30). 12p13.33 amplification appeared in all the lymph nodes and in the primary tumor of this patient. A recent study reported that the recombination repair gene RAD52, within the 12p13.33 region, is associated with the risk of lung squamous cell carcinoma development (31). The expression level of genes belonging to the targetable PI3K/RTK/RAS oncogenic pathway were observed to be above background expression levels of the TCGA squamous cell carcinoma cohort ($n = 495$; Supplementary Fig. S4C). Most of these genes were highly expressed in the 4L lymph node. Among them, higher expression of *AKT2* (Supplementary Fig. S5B) and *RASA1* genes in 4L was significantly correlated with poor clinical outcome by predicting risk from the TCGA SCC cohort ($P = 0.012$ and $P = 0.015$, respectively). The previous study suggested that AKT activation may occur early in tumor progression of lung cancer (32). It is likely that PIK3CA mutation and ATK2 amplification evoke abnormal expression of downstream genes involved in oncogenic pathways.

Discussion

In this study, we characterized genomic alterations and their associations with treatment response by performing WES and RNA-seq on multiple biopsy specimens from different anatomic regions in advanced stage lung cancers. The patients with advanced stage lung cancer usually do not undergo surgery to receive a biopsy for the most representative lesion. Histologic diagnosis, genotyping, and treatment decision are usually based on a single biopsy specimen. However, tumor heterogeneity can be an important obstacle to successful targeted therapy in these patients as the current strategy of biopsy for the most representative lesion cannot reflect the whole landscape of tumor heterogeneity. Although previous studies have addressed the issue of tumor heterogeneity in patients with early-stage lung cancers (5, 6), they are not usually considered the candidate for the targeted therapy.

In this study, tumor heterogeneity and clonal diversity were associated with the heterogeneous treatment response in patients with primary lung cancer. The tumor heterogeneity could influence the clinical outcomes of patients with cancer, thus understanding the correlation between tumor heterogeneity and genetic alternations will lead to the establishment of a new strategy for targeted therapy in primary lung cancer. Targeted therapy showed an initial treatment response but disease progression during the treatment in patients with primary lung cancer (33, 34). This phenomenon may be attributable to the acquired resistance to the targeted therapy or clonal diversity with multiple therapeutic targets (35, 36). If we perceive multiple driver mutations in the tumor or metastatic lymph node, we may need to combine targeted therapies for the multiple therapeutic targets.

Mutation profiles obtained from WES or high-depth targeted panel sequencing will provide information for making treatment decisions for advanced lung cancers with metastasis. However, mutation profile alone is not sufficient for full understanding of

the treatment response and clinical phenotype. Additional information such as RNA expression and epigenetic features may be needed. TH1 has ALK expression as well as exon 20 mutations of EGFR in only 4R and 4L. Exon 20 mutations are associated with resistance to anti-EGFR therapies (37). In TH4, private mutation in ABCB1 known as a known multidrug transporter may explain partially why lymph node 7 shows poor treatment response. However, there may be yet unknown factors causing resistance to treatment in transcription level. Poor treatment response of three lymph nodes in TH5 may be due to the private mutation as well as high expression of the resistance gene. Lymph node 7 has ARID1B mutation that may be associated with treatment failure like in neuroblastoma (38). 11L and 4R showed higher expression level (> 14 fold change) of *ABCC3* gene than primary (< 8-fold). ABCC3 has been known to be a marker for multidrug resistance in NSCLC (39). Furthermore, pathway alterations according to aberrant gene expression could confer drug resistance. Patient TH6 has a PIK3CA H1047R mutation in only one region, as seen in another heterogeneity study (5). In this patient, an inhibitor targeting the PI3K/mTOR signaling axis as part of a combination therapy may be suggested. In patients TH1, TH4, and TH5, heterogeneous response or progression were associated with tumor heterogeneity in each lesion after cytotoxic chemotherapy or concurrent chemoradiation therapy. Thus, an integrative understanding of genomic and transcriptomic alterations from biopsies of multiple regions is critical for therapeutic decision making in advanced lung cancers with metastasis.

Interestingly, tumor heterogeneity of metastatic lymph nodes was associated with the proximity from the primary tumor and lymphatic drainage system. In patient TH4, the primary tumor (RBI) was in close proximity to 11Rs (N1; Fig. 1A). N1 shared more mutations with the primary tumor than N2 (station 7) or N3 (4L). Many mutations between N2 and N3 were concordant, but N2 was closer to the primary tumor than N3. Station 7 (subcarinal lymph node) is a midline lymph node considered as a crossroad where lymphatic vessels from all lobes of the lungs meet, and metastasis to the contralateral lymph nodes occur via station 7 (40). Likewise, in patient TH6, the primary tumor (left upper lobe) is likely to have metastasized first to 4L and then to the contralateral 2R via station 7 based on tumor heterogeneity and evolution pattern (Fig. 1A).

The determinants of stemness are able to be the contributors that affect therapy failure (41). The genetic and nongenetic mechanism as well as cancer stem cells (CSC) can be sources of tumor heterogeneity. The observation of transcriptional signatures associated with CSC from multiple tumors could be highly predictive in clinical outcome as well as therapeutic resistance in patients with advanced lung cancer. Analysis of branched clonal evolution pattern will lead us to understand mechanisms of tumor heterogeneity, the link with therapy resistance, and CSC biologies. Finally, our data can serve as new evidence for the old concept of the mediastinal drainage system based on intermetastatic tumor heterogeneity and clonal evolution.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: S.-W. Um, H. Kim, W.-Y. Park
Development of methodology: S.-W. Um, H. Kim, D.N. Hayes
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): S.-W. Um, H. Lee, J. Park
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): S.-W. Um, J.-G. Joung, H. Lee, K.-T. Kim, D.N. Hayes
Writing, review, and/or revision of the manuscript: S.-W. Um, J.-G. Joung, H. Lee, K.-T. Kim, D.N. Hayes, W.-Y. Park
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): H. Lee, J. Park
Study supervision: S.-W. Um, H. Kim, W.-Y. Park

Acknowledgments

We thank the research participants of Samsung Medical Center and patients in this research.

Grant Support

This work was supported by the Samsung Medical Center and Health Technology R&D Project, the Ministry of Health & Welfare (HI13C2096 to W.Y. Park).

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 March 29, 2016; revised August 3, 2016; accepted August 17, 2016; published OnlineFirst September 13, 2016.

References

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. CA Cancer J Clin 2011;61:69–90.
2. Teran MD, Brock MV. Staging lymph node metastases from lung cancer in the mediastinum. J Thorac Dis 2014;6:230–6.
3. Kris MG, Johnson BE, Berry LD, Kwiatkowski DJ, Iafrate AJ, Wistuba II, et al. Using multiplexed assays of oncogenic drivers in lung cancers to select targeted drugs. JAMA 2014;311:1998–2006.
4. Nosotti M, Palleschi A, Rosso L, Tosi D, Santambrogio L, Mendogni P, et al. Lymph node micrometastases detected by carcinoembryonic antigen mRNA affect long-term survival and disease-free interval in early-stage lung cancer patients. Oncol Lett 2012;4:1140–44.
5. de Bruin EC, McGranahan N, Mitter R, Salm M, Wedge DC, Yates L, et al. Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. Science 2014;346:251–6.
6. Zhang J, Fujimoto J, Wedge DC, Song X, Seth S, Chow CW, et al. Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing. Science 2014;346:256–9.
7. Hiley C, de Bruin EC, McGranahan N, Swanton C. Deciphering intratumor heterogeneity and temporal acquisition of driver events to refine precision medicine. Genome Biol 2014;15:453.
8. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;25:1754–60.
9. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 2010;38:e164.
10. Magi A, Tattini L, Cifola I, D'Aurizio R, Benelli M, Mangano E, et al. EXCAVATOR: detecting copy number variants from whole-exome sequencing data. Genome Biol 2013;14:R120.
11. Anders S, Pyl PT, Huber W. HTSeq–a Python framework to work with high-throughput sequencing data. Bioinformatics 2015;31:166–9.
12. Hanzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics 2013;14:7.
13. Roth A, Khattra J, Yap D, Wan A, Laks E, Biele J, et al. PyClone: statistical inference of clonal population structure in cancer. Nat Methods 2014;11:396–8.

Published OnlineFirst September 13, 2016; DOI: 10.1158/0008-5472.CAN-16-0873

14. Govindan R, Ding L, Griffith M, Subramanian J, Dees ND, Kanchi KL, et al. Genomic landscape of non-small cell lung cancer in smokers and never-smokers. Cell 2012;150:1121–34.
15. 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.
16. Inamura K, Takeuchi K, Togashi Y, Hatano S, Ninomiya H, Motoi N, et al. EML4-ALK lung cancers are characterized by rare other mutations, a TTF-1 cell lineage, an acinar histology, and young onset. Mod Pathol 2009;22: 508–15.
17. 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.
18. Soda M, Choi YL, Enomoto M, Takada S, Yamashita Y, Ishikawa S, et al. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. Nature 2007;448:561–6.
19. Koivunen JP, Mermel C, Zejnullahu K, Murphy C, Lifshits E, Holmes AJ, et al. EML4-ALK fusion gene and efficacy of an ALK kinase inhibitor in lung cancer. Clin Cancer Res 2008;14:4275–83.
20. Gainor JF, Varghese AM, Ou SH, Kabraji S, Awad MM, Katayama R, et al. ALK rearrangements are mutually exclusive with mutations in EGFR or KRAS: an analysis of 1,683 patients with non-small cell lung cancer. Clin Cancer Res 2013;19:4273–81.
21. The Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. Nature 2014;511:543–50.
22. Sriuranpong V, Borges MW, Ravi RK, Arnold DR, Nelkin BD, Baylin SB, et al. Notch signaling induces cell cycle arrest in small cell lung cancer cells. Cancer Res 2001;61:3200–5.
23. Capaccione KM, Pine SR. The Notch signaling pathway as a mediator of tumor survival. Carcinogenesis 2013;34:1420–30.
24. Dorr C, Janik C, Weg M, Been RA, Bader J, Kang R, et al. Transposon mutagenesis screen identifies potential lung cancer drivers and CUL3 as a tumor suppressor. Mol Cancer Res 2015;13:1238–47.
25. Attoub S, Arafat K, Hammadi NK, Mester J, Gaben AM. Akt2 knock-down reveals its contribution to human lung cancer cell proliferation, growth, motility, invasion and endothelial cell tube formation. Sci Rep 2015;5: 12759.
26. Wangari-Talbot J, Hopper-Borge E. Drug resistance mechanisms in non-small cell lung carcinoma. J Can Res Updates 2013;2:265–82.
27. Ding L, Getz G, Wheeler DA, Mardis ER, McLellan MD, Cibulskis K, et al. Somatic mutations affect key pathways in lung adenocarcinoma. Nature 2008;455:1069–75.
28. Kang S, Bader AG, Vogt PK. Phosphatidylinositol 3-kinase mutations identified in human cancer are oncogenic. Proc Natl Acad Sci U S A 2005; 102:802–7.
29. Zou ZQ, Zhang XH, Wang F, Shen QJ, Xu J, Zhang LN, et al. A novel dual PI3Kalpha/mTOR inhibitor PI-103 with high antitumor activity in non-small cell lung cancer cells. Int J Mol Med 2009; 24:97–101.
30. Altomare DA, Testa JR. Perturbations of the AKT signaling pathway in human cancer. Oncogene 2005;24:7455–64.
31. Lieberman R, Xiong D, James M, Han Y, Amos CI, Wang L, et al. Functional characterization of RAD52 as a lung cancer susceptibility gene in the 12p13.33 locus. Mol Carcinog 2016;55:953–63.
32. Balsara BR, Pei J, Mitsuuchi Y, Page R, Klein-Szanto A, Wang H, et al. Frequent activation of AKT in non-small cell lung carcinomas and preneoplastic bronchial lesions. Carcinogenesis 2004;25:2053–9.
33. Maemondo M, Inoue A, Kobayashi K, Sugawara S, Oizumi S, Isobe H, et al. Gefitinib or chemotherapy for non-small-cell lung cancer with mutated EGFR. N Engl J Med 2010;362:2380–8.
34. 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.
35. Yu HA, Arcila ME, Rekhtman N, Sima CS, Zakowski MF, Pao W, et al. Analysis of tumor specimens at the time of acquired resistance to EGFR-TKI therapy in 155 patients with EGFR-mutant lung cancers. Clin Cancer Res 2013;19:2240–7.
36. Su KY, Chen HY, Li KC, Kuo ML, Yang JC, Chan WK, et al. Pretreatment epidermal growth factor receptor (EGFR) T790M mutation predicts shorter EGFR tyrosine kinase inhibitor response duration in patients with non-small-cell lung cancer. J Clin Oncol 2012;30:433–40.
37. Hu C, Liu X, Chen Y, Sun X, Gong Y, Geng M, et al. Direct serum and tissue assay for EGFR mutation in non-small cell lung cancer by high-resolution melting analysis. Oncol Rep 2012;28:1815–21.
38. Sausen M, Leary RJ, Jones S, Wu J, Reynolds CP, Liu X, et al. Integrated genomic analyses identify ARID1A and ARID1B alterations in the childhood cancer neuroblastoma. Nat Genet 2013;45:12–7.
39. Zhao Y, Lu H, Yan A, Yang Y, Meng Q, Sun L, et al. ABCC3 as a marker for multidrug resistance in non-small cell lung cancer. Sci Rep 2013; 3:3120.
40. Libshitz HI, McKenna RJJr, Mountain CF. Patterns of mediastinal metastases in bronchogenic carcinoma. Chest 1986;90:229–32.
41. Kreso A, Dick JE. Evolution of the cancer stem cell model. Cell Stem Cell 2014;14:275–91.

Cancer Research
The Journal of Cancer Research (1916–1930) | The American Journal of Cancer (1931–1940)

AACR American Association for Cancer Research

Molecular Evolution Patterns in Metastatic Lymph Nodes Reflect the Differential Treatment Response of Advanced Primary Lung Cancer

Sang-Won Um, Je-Gun Joung, Hyun Lee, et al.

Cancer Res 2016;76:6568-6576. Published OnlineFirst September 13, 2016.