Discovery of predictive biomarker for recurrence in early stage *EGFR*-mutant lung adenocarcinoma using broad-panel Next-Generation Sequencing

Jae Seok Lee

Department of Medicine

The Graduate School, Yonsei University

Discovery of predictive biomarker for recurrence in early stage *EGFR*-mutant lung adenocarcinoma using broad-panel Next-Generation Sequencing

Jae Seok Lee

Department of Medicine

The Graduate School, Yonsei University

Discovery of predictive biomarker for recurrence in early stage *EGFR*-mutant lung adenocarcinoma using broad-panel Next-Generation Sequencing

Directed by Professor Hyo Sup Shim

The Doctoral Dissertation
submitted to the Department of Medicine,
the Graduate School of Yonsei University
in partial fulfillment of the requirements for the degree
of Doctor of Philosophy in Medical Science

Jae Seok Lee

December 2021

This certifies that the Doctoral Dissertation of Jae Seok Lee is approved.

Thesis Supervisor: Hyo Sup Shim

Thesis Committee Member#1: Se Hoon Kim

Thesis Committee Member#2: Jin Gu Lee

Thesis Committee Member#3: Jin Hur

Thesis Committee Member#4: Tae-Jung Kim

The Graduate School
Yonsei University

December 2021

ACKNOWLEDGEMENTS

I would like to express my heart-felt gratitude to my supervisor, Professor Hyo Sup Shim, for his great care and support for my doctoral dissertation and related studies. I would also like to express my sincere thanks to Professor Se Hoon Kim, Professor Tae-Jung Kim, Professor Jin Gu Lee, Professor Jin Hur for their valuable advice and warm encouragement, and to Kyung A Kim, Kisu Kim, and Nak Gyeong Ko for their generous support with data collection and statistical analyses.

<TABLE OF CONTENTS>

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Discovery of predictive biomarker for recurrence in early stage *EGFR*-mutant lung adenocarcinoma using broad-panel next-generation sequencing

Jae Seok Lee

Department of Medicine
The Graduate School, Yonsei University

(Directed by Professor Hyo Sup Shim)

Background. Despite complete surgical resection, the recurrence rate of *EGFR*-mutant early-stage lung adenocarcinoma remains high. No clear genomic predictive markers for survival or recurrence of *EGFR*-mutant early lung adenocarcinoma are currently available. Broad-panel next-generation sequencing (NGS) is increasingly used to identify driver mutations for targeted therapy and precision medicine strategies for patients with lung cancer. In combination with clinical parameters, the use of NGS provides avenues for biomarker discovery with relevance for risk assessment. Employing a pathway-centric approach using NGS to compare the genomic profiles of patients with and without recurrence, we aimed to identify genetic alterations to predict clinical outcomes in patients with early-stage lung adenocarcinoma.

Methods. Tissue samples from 105 patients with completely resected stage I *EGFR*-mutant lung adenocarcinoma were analyzed for 523 or 170 cancer-related genes by broad-panel NGS. Eleven canonical oncogenic pathways were analyzed and the number of pathway alterations (NPA) were determined. The correlations

between recurrence free survival (RFS) and co-occurring genetic aberrations, concurrent pathway alterations, and NPA were analyzed using the Kaplan-Meier method and log-rank test.

Results. The total number of pathogenic genetic alterations was higher in patients with recurrent disease (P<0.001). *EGFR/TP53* co-mutated patients showed shorter RFS than *EGFR* mutation only patients (P=0.045). The high NPA (≥3) group showed shorter OS (P=0.031) and RFS (P<0.0001) than the low NPA (1-2) group. Patients with certain concurrent pathway alterations, such as RTK/RAS-PI3K (P<0.001), RTK/RAS-p53 (P=0.0015), and RTK/RAS-cell cycle (P=0.0017) showed shorter RFS than those with RTK/RAS pathway alterations only. Amplification of *EGFR*, *RPS6KB1*, and *CHEK2* was correlated with worse OS (*RPS6KB1*, P<0.001) or RFS (P=0.023 in *EGFR*, P<0.0001 in *RPS6KB1*, P=0.0035 in *CHEK2*). High MSI score was correlated with worse RFS. *TP53* co-mutation (HR 2.17, P=0.0252), high NPA (HR 2.68, P<0.001), and higher pathologic stage (HR 3.27, P<0.001) were independent risk factors for recurrence on multivariate analysis.

Conclusions. Broad-panel NGS with pathway-centric analysis provided valuable information to facilitate risk stratification of resected early-stage *EGFR*-mutant lung adenocarcinoma. Increased number of genetic alterations, high NPA, co-occurring *TP53* mutation, concurrent PI3K, p53, cell cycle pathway alterations with RTK/RAS pathway alteration, and concurrent amplification of *EGFR*, *RPS6KB1*, *CHEK2* were correlated with worse clinical outcomes. The proposed approach may help identify patients at high risk for recurrence who may benefit from adjuvant EGFR-TKIs therapy or chemotherapy.

Key words: lung adenocarcinoma, early stage, *EGFR*, next-generation sequencing, recurrence, predictive biomarker

Discovery of predictive biomarker for recurrence in early stage *EGFR*-mutant lung adenocarcinoma using broad-panel next-generation sequencing

Jae Seok Lee

Department of Medicine
The Graduate School, Yonsei University

(Directed by Professor Hyo Sup Shim)

I. Introduction

Complete surgical resection with mediastinal lymph node dissection remains the gold standard for patients with early-stage non-small cell lung carcinoma (NSCLC)[1]. However, 20-50% of patients experience recurrence and eventually die of recurrent lung cancer (5 year overall survival, 58-73%)[2]. Adjuvant chemotherapy is recommended for patients with stage II and III lung cancer to prevent recurrence, but is not recommended for stage IA patients. The use of adjuvant chemotherapy for stage IB patients is controversial, and remains optional. It is recommended for stage IB patients with various clinicopathological factors such as lymphovascular invasion, visceral pleural invasion, and poor differentiation without considering genetic aberrations. However, clinicopathological factors alone are inadequate to stratify risk in patients with early-stage NSCLC to determining adjuvant therapy course. To fully understand the potential benefits of adjuvant chemotherapy or targeted therapy for patients with early-stage lung cancer, it is necessary to predict the risk of tumor recurrence

using additional genetic predictive factors.

Epidermal growth factor receptor (EGFR) mutations are the most common driver mutation in East Asian patients with lung adenocarcinoma (38.8-64.0%)[3], and several potent EGFR tyrosine-kinase inhibitors (TKI) have raised the median overall survival (OS) of patients with advanced stage disease to longer than 2.5 years[4]. Yet not all patients benefit equally from TKI treatment due to clinical and biological heterogeneity. Recently, the ADAURA clinical trial reported that adjuvant treatment of resected stage IB-IIIA *EGFR*-mutant NSCLC with the third generation EGFR-TKI, osimertinib, showed significantly better outcomes compared to a placebo[5]. However, considering the high cost, adverse effect, and long treatment period associated with EGFR-TKIs, it is controversial whether all stage I NSCLC should be treated with this approach. Thus, the selection of patient at high risk for recurrence is necessary when considering adjuvant TKI therapy. Studies on predictive genomic markers for early stage *EGFR*-mutant lung adenocarcinoma in patients who did not undergo adjuvant TKI therapy are limited. Recent studies were focused on risk stratification of TKI-treated, advanced stage *EGFR*-mutant NSCLC using next-generation sequencing[6, 7, 8]. Many studies have been conducted to discover novel predictive biomarkers for recurrence of resected early stage NSCLC using various methods, such as gene expression profiling, quantitative reverse transcriptase-polymerase chain reaction (PCR), microRNA assays, and mass spectroscopy. Despite this, genomic biomarkers and mechanism for recurrent NSCLC are not well established.

Broad-panel next-generation sequencing (NGS) allows the evaluation of multiple genes in parallel. This approach has recently been used to elucidate tumor biology, and to select targeted therapies to facilitate precision medicine strategies for patients with lung cancer[9]. Compared with traditional single-gene assays, NGS is a feasible and cost-effective method of diagnostic genomic profiling. However, it is unclear how the high-volume of genomic alterations identified by NGS interact and how they can be used alongside *EGFR* mutation

status and clinicopathological predictors to stratify risk for recurrence. In order to address this, we used broad-panel NGS to analyze stage I *EGFR*-mutant lung adenocarcinoma specimens to investigate the pathogenic genomic co-alterations and oncogenic pathway co-alterations, and their association with clinical outcomes.

II. MATERIALS AND METHODS

1. Patients and Sample collection

The study was approved by the Institutional Review Boards at Severance Hospital. A total of 410 patients with histologically confirmed stage I lung adenocarcinoma who underwent complete resection at Severance Hospital from 2004 to 2013 were screened for inclusion in this study. Of those patients, 256 (62.4%) individuals with *EGFR*-mutant stage I adenocarcinoma were selected by reviewing electronic pathologic reports and additional confirmation of *EGFR* mutation status by using the PNAClamp *EGFR* Mutation Detection Kit (Panagene Inc. Daejeon, Korea) or PANAMutyper *EGFR* Kit (Panagene Inc.) according to manufacturer's instructions. The electronic medical records were reviewed and age, sex, smoking history, pathologic stage, and the dates of recurrence and death/last follow up were collected. We excluded patients with non-adenocarcinoma histology, insufficient tumor specimen for complete molecular analysis, and those who received adjuvant chemoradiation therapy, TKI therapy, or any neoadjuvant treatment. Pathologic stages were reviewed and assigned according to the 8th American Joint Committee on Cancer (AJCC) criteria. Finally, 105 patients with resected stage I EGFR-mutant adenocarcinoma with full clinical information available in their medical records, including prognostic data of overall survival (OS) and recurrence free survival (RFS) by follow-up, who provided consent to participate in the study were enrolled.

2. Targeted Next-Generation Sequencing and interpretation

DNA was extracted from formalin-fixed, paraffin-embedded whole tissue sections using the QIAamp DNA kit (QIAGEN GmbH, Hilden, Germany), according to the manufacturer's protocol. Three versions of the NGS sequencing panel were used: the TrueSight Oncology 500 panel (Illumina, Inc., San Diego, CA, USA) was used to target 523 cancer-related genes in 99 (94.2%) cases; TrueSight Tumor 170 panel (Illumina, Inc., San Diego, CA, USA) was used to target 170 genes in five (4.8%) cases; and the NgeneBio panel (NgeneBio, Seoul, Korea) was used to target 46 genes for one case. The resulting libraries were PCR-amplified and purified with Agencourt AMPure XP beads (Beckman Coulter, Inc., CA, USA). Libraries were sequenced using the HiSeq 2500 platform (Illumina Inc., San Diego, CA, USA). Raw sequencing data were processed and variants were called using the Macrogen Inc. bioinformatics pipeline (Macrogen, Seoul, Korea). Annotations were performed by Clinical Genomics Workspace (PierianDx, Missouri, USA). Somatic mutations, including single nucleotide polymorphisms (SNPs), small insertions and deletions (Indels), copy number variations (CNVs), and gene rearrangements were identified. Pathogenic and likely pathogenic somatic mutations with variant allele frequencies (VAFs) greater than four percent were regarded as significant actionable mutations and were used for analysis. Fold changes over 1.5 or less than 0.5 were regarded as CNVs. Variant interpretation was based on recommendations from the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists[10].

The clonality index was derived by normalizing the variant allele frequency of each gene relative to that of the *EGFR* driver mutation. We defined each gene mutation with a clonality index of 0.6 and higher as clonal, and those with a clonality index less than 0.6 as subclonal[11]. We determined that this cutoff point was a reasonable discriminator between clonal and sub-clonal variants based on empiric data derived from whole genome sequencing studies in myeloid tumors[12].

Tumor mutational burden (TMB) was defined as the total number of nonsynonymous single-nucleotide or insertion/deletion mutations divided by the number of Mbs in the coding region captured by each panel (1.94 Mb and 0.53 Mb in the 523- and 170-gene panels, respectively). Microsatellite instability (MSI) was calculated using the percentile of unstable MSI sites in the samples from the 99 patient samples analyzed using the TrueSight Oncology 500 panel. Analysis of specific somatic alterations was performed using OncoKB to remove variants of unknown significance or likely passenger events. Therapeutic-actionability information was annotated using OncoKB knowledge base (http://www.OncoKB.org) which contains information about the oncogenic effects and treatment implication of variants in over 400 cancer genes. Then, each genomic alteration was stratified into one of four levels of clinical actionability[13].

3. Number of pathway alteration

We evaluated 11 recognized signaling pathways using the templates provided by the TCGA PanCancer Atlas project[14]. The pathways analyzed were receptor tyrosine kinase (RTK/RAS), Hippo, PI3K, TGFß, ß-catenin/Wnt, Myc, p53, oxidative stress response/Nrf2, Notch, cell cycle, and DNA damage. A tumor sample was considered “altered” in the specific pathway when one or more genes in the pathway contained a recurrent or known driver alteration. The status of specific pathways was determined to be either altered or wild type for each patient. The number of pathway alterations (NPA) for each sample was calculated as the total number of altered pathways[15]. We classified samples with one or two pathway alterations as ‘NPA low’, and those with three or more pathway alterations as ‘NPA high’, and grouped them accordingly.

4. Patient Follow-Up

Patients were followed with surveillance visits including a history interview,

physical examination, and chest radiography with chest CT scans. If symptoms or signs of tumor recurrence were detected, CT scans of chest and abdomen, PET-CT, and brain MRI evaluations were performed. Recurrences were distinguished from secondary primary lung cancers using the Martini and Melamed criteria[16]. The date of recurrence was defined as the date of histologic proof or, in patients whose diagnoses were based on clinicopathological findings, the date of identification by the physician. Patients were censored at the last follow-up appointment. OS was calculated from the date of diagnosis to the point of last follow-up or death.

5. Statistical analysis

Clinical and pathological parameters were evaluated by qui-square analysis or Fisher's exact test for categorical variables, and Spearman rank correlation test for continuous variables. Prognostic values were assessed by survival analysis. RFS and OS percentages were calculated using the Kaplan-Meier method, and the differences were tested by the log-rank test in the univariate analysis. Stepwise Cox regression analysis was conducted to identify the prognostic factor for recurrence in multivariate analysis. TMB and MSI were analyzed using Student's T-test. NPA was analyzed using Welch's test. Tests were two-sided for all calculations and $P < 0.05$ was considered statistically significant. Statistical analyses were conducted using R 3.5.1. and SAS 9.4.

III. RESULTS

1. Clinicopathological characteristics and *EGFR* mutation status of patients

The study included 105 patients with resected stage I *EGFR* mutant lung adenocarcinoma treated at our institution from 2004 to 2013. Recurrence was observed in 51 (19.9%) of the 256 patients with stage I *EGFR* mutant lung

adenocarcinoma patients (recurrent group). We performed broad-panel NGS analysis in all 51 patients in the recurrent group and 54 of the 205 patients without recurrence (non-recurrent group). The selected 54 patients in the non-recurrent group showed similar distribution of age, sex, follow-up time, and pathologic stage as the remaining 151 unselected non-recurrent patients. The median follow-up time was approximately 54 months (range 7-102). All patient characteristics are summarized in Table 1. Briefly, the median age of patients was 67 years (range 26-85 yrs.), 59 (56%) were female, and 74 (71%) were never-smokers. The median smoking history was 8 pack-years (range 0-110). The baseline pathological stage distribution was as follows: Stage IA1 8 (7.6%), Stage IA2 20 (19%), Stage IA3 35 (33.3%), and stage IB 42 (40%). Exon 21 p.L858R (L858R) was the most frequent *EGFR* alteration (54.3%), followed by exon 19 deletions (del19) (35.2%), and other rare mutations (10.5%). The recurrence rate was higher for stage IB patients than stage IA patients on univariate analysis (Table 2., HR 3.68, $P < 0.0001$).

2. Genetic landscape of resected *EGFR*-mutant adenocarcinoma according to recurrence status

Among the co-occurring mutations in patients with *EGFR*-mutant adenocarcinoma, *TP53* mutations were most frequent (16%), followed by *RBM10* mutations (8.5%). *RB1*, *BRCA1*, and *RAD51B* mutations were found in two patients, respectively. The number and type of concurrent CNVs were as follows: 25 (23.8%) *RPS6KB1* amplifications, 22 (21%) *EGFR* amplifications, 16 (15.2%) *CHEK2* amplifications, 12 (11.4%) *MDM2* amplifications, and 10 (9.5%) *MYC* amplifications. Individual genetic alterations in all 105 patients are shown in Figure 1.

Table 1. Patient characteristics (N=105)

Characteristics	**Total (n=105) No. (%)**	**Non-Recur (n=54)**	**Recur (n=51)**	**P-value**
Age at surgery, years, median (range)	67 (26-85)	67 (42-87)	63 (26-77)	0.447
Sex				0.893
Male	46 (44)	24 (44.6)	22 (43.1)	
Female	59 (56)	30 (55.6)	29 (56.9)	
Smoking status				0.980
Ever	31 (29)	16 (29.6)	15 (29.4)	
Never	74 (71)	38 (70.4)	36 (70.6)	
Pack-years (range)	8 (0-110)	7 (0-60)	9 (0-110)	0.687
Pathologic stage				<0.001
IA1	8 (7.6)	8 (14.8)	0 (0.0)	
IA2	20 (19)	18 (3.3)	2 (3.9)	
IA3	35 (33.3)	20 (37.0)	15 (29.4)	
IB	42 (40)	8 (14.8)	34 (66.7)	
EGFR mutation				0.001
L858R	57 (54.3)	38 (70.4)	19 (37.3)	
del19	37 (35.2)	15 (27.8)	22 (43.1)	
Other EGFR mutations	11 (10.5)	1 (1.9)	10 (19.6)	
Number of pathway alteration				<0.001
1	40 (38.1)	29 (53.70)	11 (21.57)	
2	32 (30.5)	18 (33.33)	14 (27.45)	
3	24 (22.9)	7 (12.96)	17 (33.33)	
4	9 (8.5)	0 (0.00)	9 (17.65)	
TMB, median (range)				
Synonymous	4.7 (0-49.7)	3.9 (0.7-49.8)	5.6 (0.8-40.1)	0.165
Nonsynonymous	3.1 (0-41.8)	2.4 (0-41.9)	3.9 (0.8-34.4)	0.169
MSI (% of unstable MSI sites), median (range)	3.2 (0-9.1)	2.5 (0-8.7)	3.3 (0.8-10.0)	0.028

Abbreviations: EGFR, epidermal growth factor receptor; TMB, tumor mutational burden; MSI, microsatellite instability

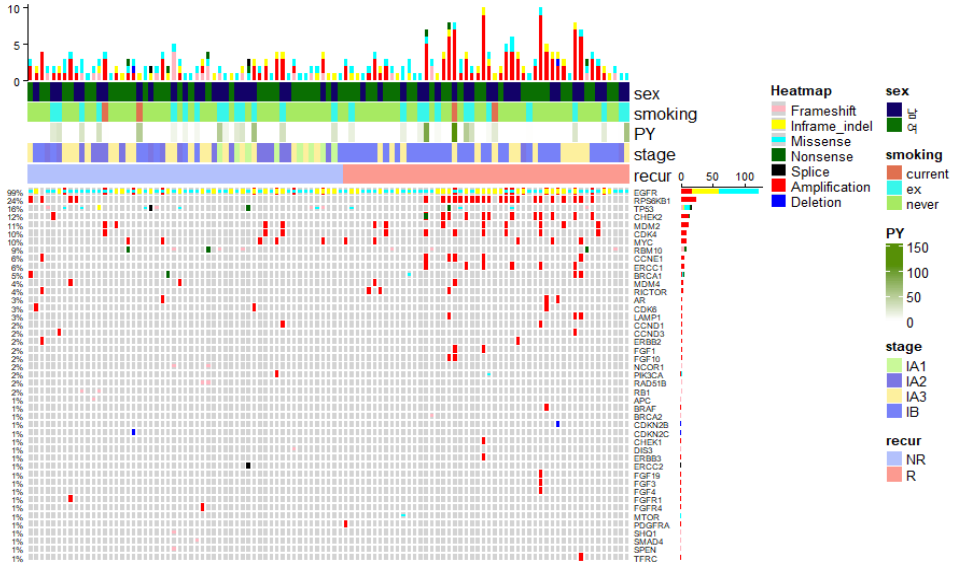


Figure 1. Genetic landscape of resected *EGFR*-mutant stage I lung adenocarcinoma according to disease recurrence status. The mutation data from all 105 patients are included according to disease recurrence status.

3. Number of genetic alterations and RFS

All samples had at least one pathogenic or likely pathogenic genetic alteration in SNP, INDEL, or CNV. The median number of genetic alterations in each sample was two (range, 1-10). We compared the number of genetic alterations between recurrent and non-recurrent groups. The median number of genetic alterations was three (range 1-10) in the recurrent group, and two (range 1-5) in the non-recurrent group. This shows a significant difference in the number of pathogenic genetic alterations between the groups (Fig. 2, $P < 0.001$).

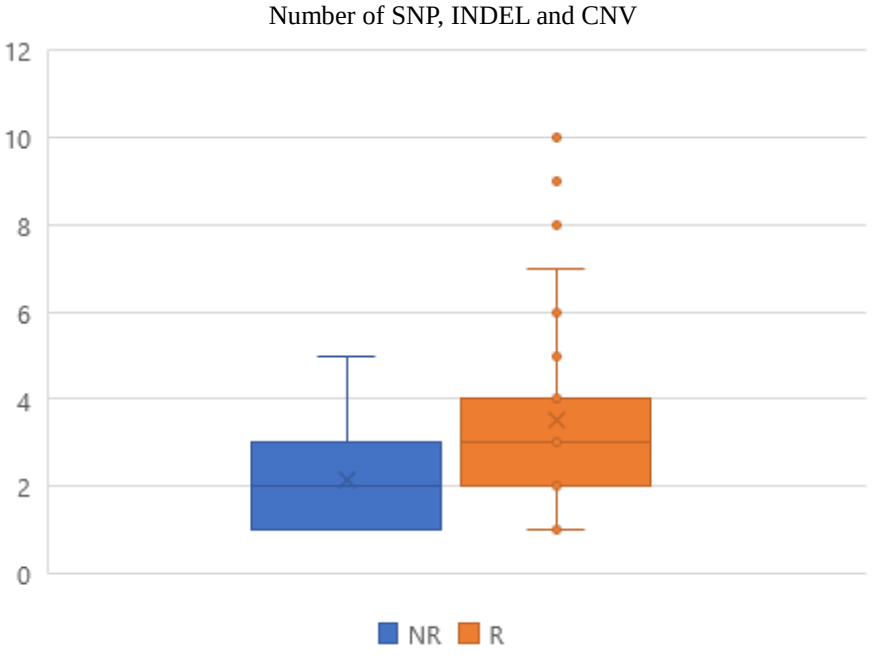
4. Co-occurring *TP53* mutations and RFS

TP53 co-mutation was observed in 17 (16%) samples. Pathogenic and likely pathogenic *TP53* mutations were spread over exons 4-10. *EGFR/TP53* co-mutated samples showed shorter RFS than *EGFR* mutation only samples (Fig. 3, $P = 0.045$). Univariate analysis showed the recurrence rate was higher for

EGFR/TP53 co-mutated samples than that of *EGFR* mutated samples (Table 2, HR 1.96, P = 0.0493).

5. Single gene CNVs

Univariate analysis showed amplification of three genes was associated with worse prognosis; *EGFR*, *RPS6KB1*, and *CHEK2*. *EGFR* amplification was in 22 (21%) patients. Average fold change was 2.25 (range 1.5-10.9) and average expected copy number of *EGFR* was 4.49. This group showed significant shorter RFS (P=0.023, Fig 4A).



Median (Total)		P-value (Total)
NR	R	P-value
2	3	0.000681

Figure 2. Comparison of number of SNPs, INDELs, and CNVs between recurrent (R) and non-recurrent (NR) groups. The number of SNP, INDEL, and CNV is significantly higher in the recurrent group.

Twenty-five (23.8%) samples had amplification of *RPS6KB1* with an average fold change of 1.72 (range 1.5-2.7) and average expected copy number of 3.44. This group showed significantly shorter OS (P = 0.0087, Fig. 4B), and shorter RFS (P < 0.0001, Fig. 4C). Sixteen (15.2%) samples had amplification of *CHEK2* with an average fold change of 1.65 (range 1.5-1.87) and an average expected copy number of *CHEK2* of 3.30. This group showed significant shorter RFS (P = 0.0035, Fig. 4D). Univariate Cox regression analysis indicated that the recurrence rate was higher than *EGFR* mutation only group for concurrent amplification of *EGFR* (Table 2, HR 2.11, P = 0.0265), *RPS6KB1* (HR 3.12, P < 0.0001), and *CHEK2* (HR 2.49, P = 0.0046).

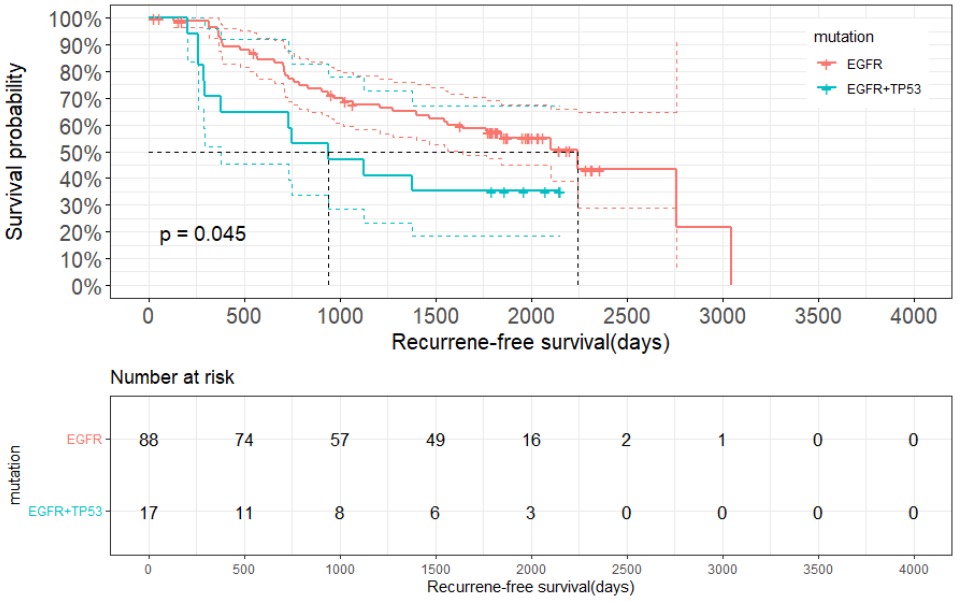


Figure 3. RFS in *EGFR-TP53* co-mutation group and *EGFR* mutation only group. The RFS was significantly shorter in the *EGFR+TP53* co-mutation group.

6. Pathway alterations and number of pathway alterations (NPA)

The p53 pathway was the most frequently altered pathway (n = 39, 37.1%), and TGFß and ß-catenin/Wnt were the least frequently altered pathways (both n=1,

0.9%). More than two genetic aberrations in RTK/RAS pathway occurred in 17 (33%) patients in the recurrent group, and eight (14.8%) patients in the non-recurrent group. A list of all 11 pathways alterations and their distribution is shown in Figure 5.

The NPA distribution was as follows: NPA1 40 (38.1%), NPA2 32 (30.5%), NPA3 24 (22.9%), and NPA4 9 (8.5%). Median NPA was 2 (range, 1-4) in all samples, 2.5 (range, 1-4) in the recurrent group, and 1.5 (range, 1-3) in the non-recurrent group. A heatmap of all altered genes in each signaling pathway is shown in Figure 6. OS and RFS showed significant differences between NPA low (1-2) and NPA high (≥ 3) groups (Fig. 7. P = 0.031 in OS, P < 0.0001 in RFS). The recurrence rate was higher for the NPA high (≥ 3) group than that of the NPA low (1-2) group (Table 2, HR 3.24, P < 0.0001).

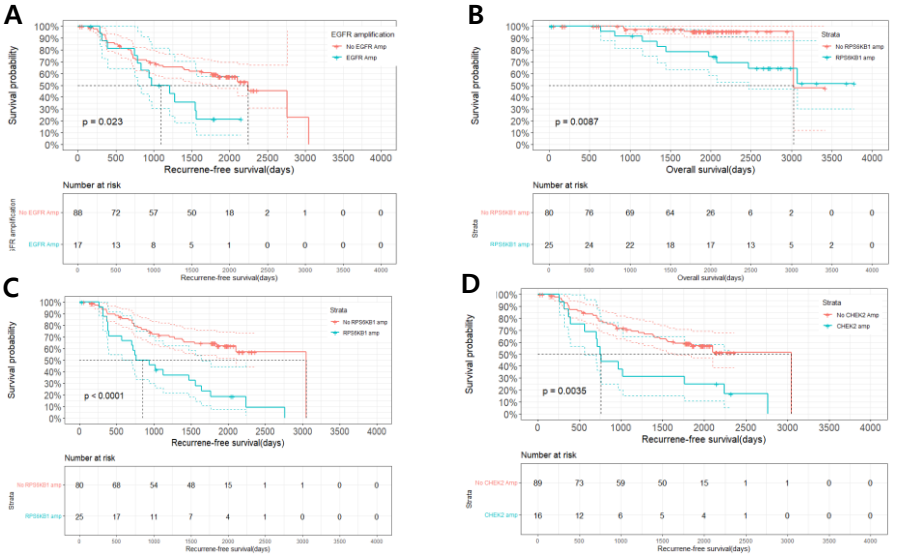


Figure 4. Amplification of *EGFR*, *RPS6KB1*, and *CHEK2*. A: *EGFR* amplification and RFS, B: *RPS6KB1* amplification and OS, C: *RPS6KB1* amplification and RFS, and D; *CHEK2* amplification and RFS.

Pathway	NAME	reference
1	RTK/RAS	
2	Hippo	
3	PI3K	
4	TGFß	
5	ß-catenin/Wnt	* 2018 Cell 173, 321-337(1~10)
6	Myc	
7	p53	
8	Nrf2	
9	Notch	
10	Cell cycle	
11	DNA_damage	http://www.pathwaycommons.org/

Pathway	NR	R
1	63	69
7	14	34
3	7	33
10	7	20
11	6	10
6	5	5
9	3	3
4	1	1
2	0	0
5	0	0
8	0	0

Figure 5. List and distributions of 11 pathway alterations. NR; non-recurrent group, R; recurrent group.

7. Co-occurring pathway alteration

The number and distribution of samples with RTK/RAS and co-occurring pathways was as follows: RTK/RAS-p53 pathway co-occurrence, 39 (37.1%) samples; RTK/RAS-PI3K, 29 (27.6%) samples; RTK/RAS-cell cycle, 22 (20.9%) samples; and RTK/RAS-DNA damage, 15 (14.2%) samples. Fourteen samples (13.3%) had two pairs of p53-cell cycle co-occurrence and seventeen (16.1%) had p53-PI3K co-occurrence.

Patients with RTK/RAS-PI3K pathway co-occurrence showed significantly shorter OS (P = 0.023) and RFS (P < 0.001) compared with RTK/RAS pathway only (Fig. 7A, B). Patients with RTK/RAS-p53 showed shorter RFS (P = 0.0015, Fig. 7C) than those with RTK/RAS pathway only, and those with RTK/RAS-cell cycle also showed shorter RFS than those with RTK/RAS pathway only (P = 0.0017, Fig. 7D). On univariate Cox regression analysis, the recurrence rate was higher for patients with RTK/RAS-PI3K (HR 1.60, P < 0.001), RTK/RAS-p53 (HR 1.69, P = 0.0057) and RTK/RAS-cell cycle (HR 1.74, P = 0.0072) pathway

co-occurrence than those with RTK/RAS pathway only (Table 2).

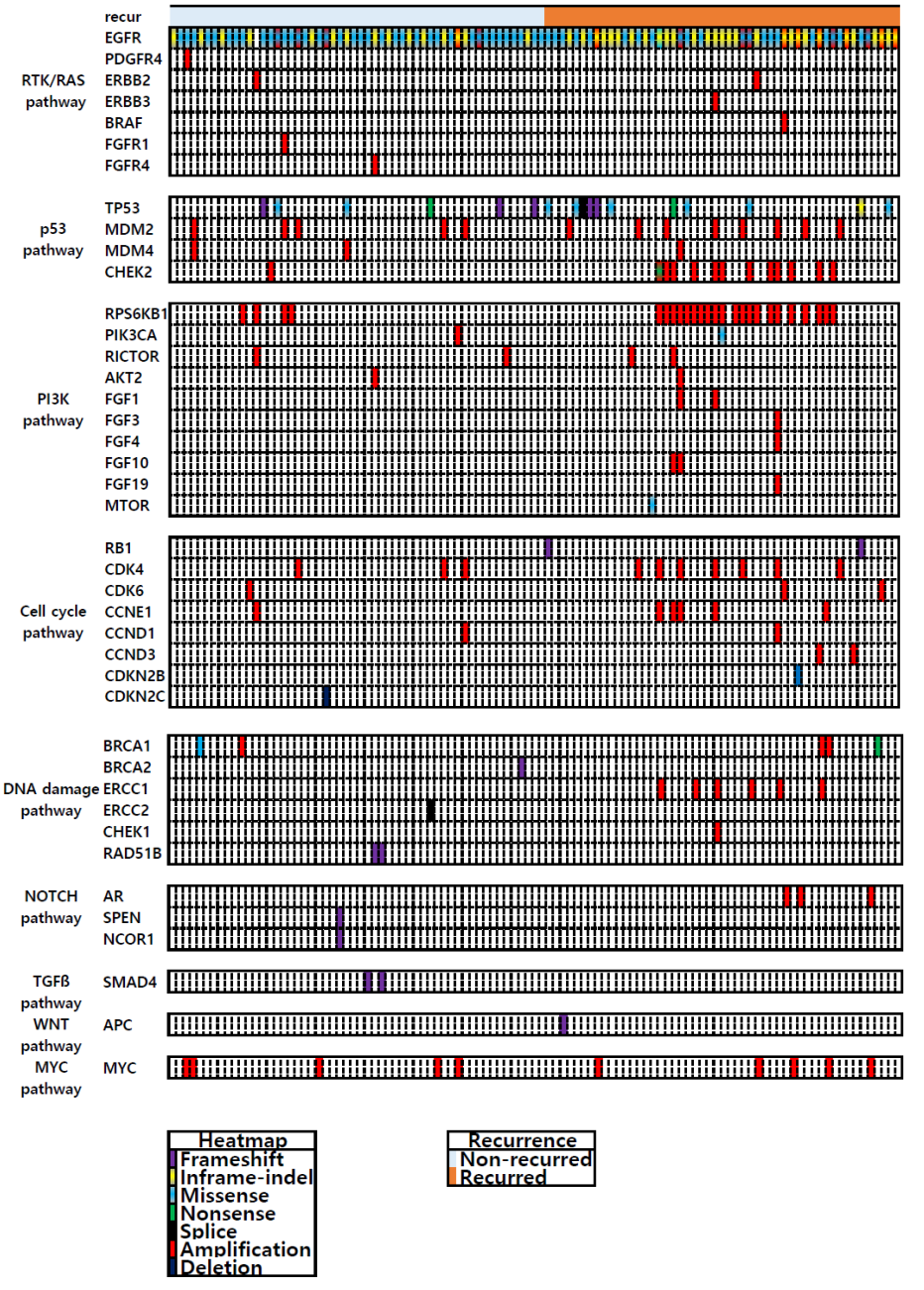


Figure 6. Heatmap of all altered genes in altered signaling pathways

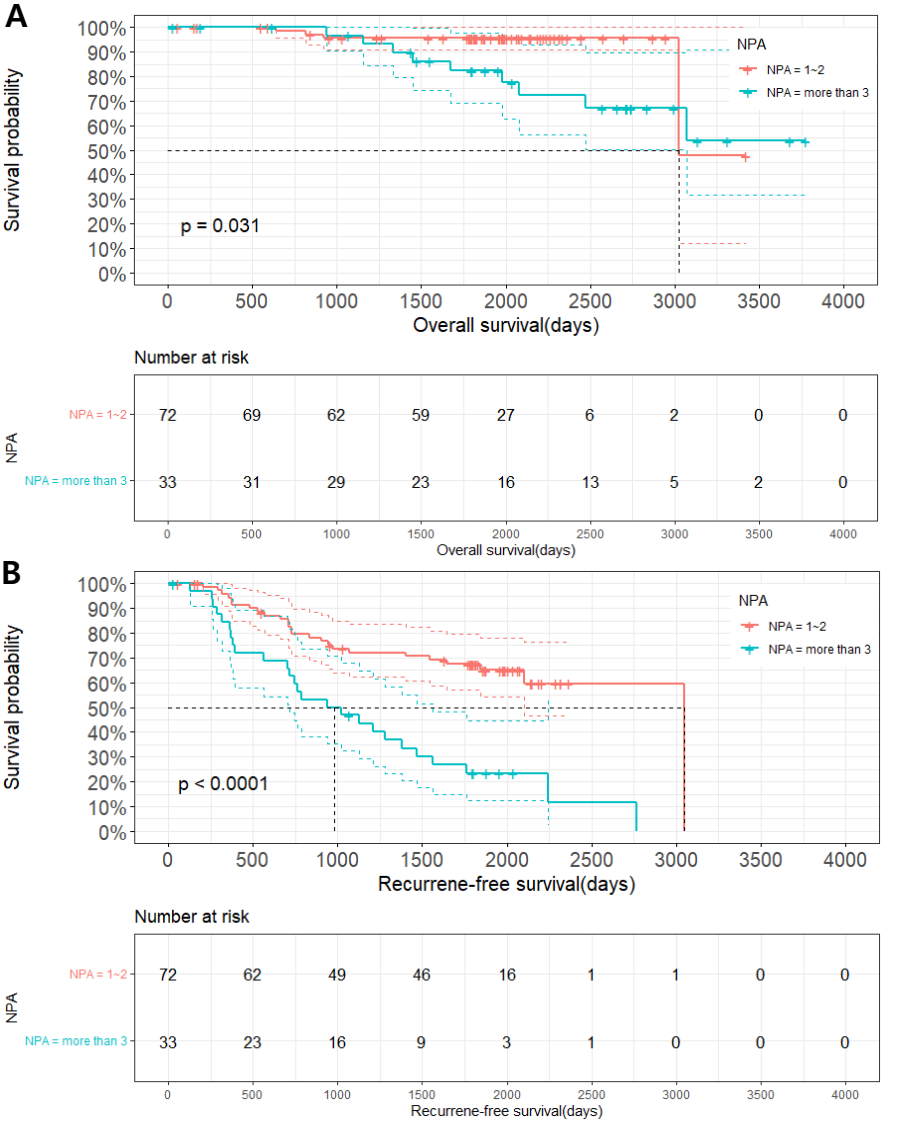


Figure 7. Associations of number of NPA and OS, RFS. OS (A) and RFS (B) was significantly shorter in the NPA high (NPA ≥ 3) group than NPA low (NPA 1-2) group.

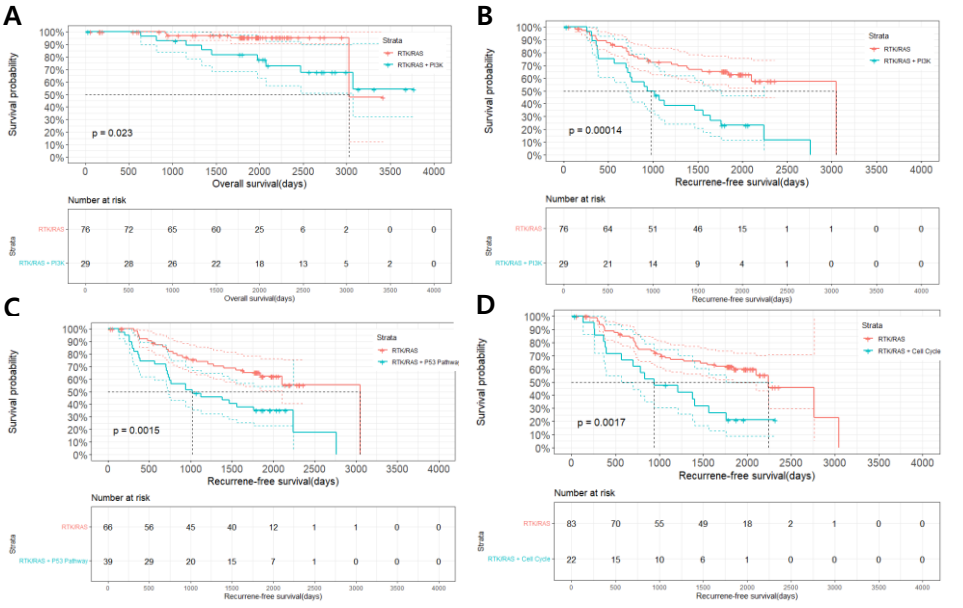


Figure 8. Co-occurring pathway alterations with RTK/RAS pathway alteration. A: RTK/RAS-PI3K and OS, B; RTK/RAS-PI3K and RFS, C: RTK/RAS-p53 and RFS, and D: RTK/RAS-cell cycle and RFS.

8. TMB and MSI

Average synonymous and nonsynonymous TMB were 6.2 mutations per megabase (mt/Mb) (range 0.78-49.77, median 4.7) and 4.3 mt/Mb (range 0-41.87, median 3.1), respectively. Median synonymous TMB was 3.92 mt/Mb (range 0.78-49.77) in the non-recurrent group, and 5.59 mt/Mb (range 0.78-40.07) in the recurrent group. Median nonsynonymous TMB was 2.36 mu/Mb (range 0-41.87) in the non-recurrent group, and 3.93 mt/Mb (range 0.78-34.35) in the recurrent group. Student's t-test showed median synonymous and nonsynonymous TMB were higher in the recurrent group, but there was no statistically significant difference between non-recurrent and recurrent groups (P = 0.1657 in synonymous, P = 0.1690 in nonsynonymous,).

The mean MSI score (percent of unstable MSI sites) was 3.3 (range 0-9.1,

median 3.2) in all 99 samples, 3.9 (range 0.8-9.1, median 3.28) in the recurrent group, and 2.8 (range 0-5.9, median 2.48) in the non-recurrent group, with Student's t-test indicating significant difference between groups (P = 0.0281).

9. Prognostic factors for recurrence based on multivariate analysis

Univariate analysis of RFS revealed the following negative prognostic factors for recurrence: pathologic stage (IB), *TP53* co-mutation, NPA, and concurrent PI3K, p53, cell cycle pathway alterations with RTK/RAS pathway, and amplifications of *EGFR*, *RPS6KB1*, *CHEK2*. Multivariate stepwise Cox regression analysis revealed that independent risk factors for recurrence included higher pathologic stage (HR 3.27 P < 0.001), *TP53* co-mutation (HR 2.17, P = 0.0252), and high NPA (HR 2.68, P < 0.001) (Table 2).

10. Nomogram for predicting recurrence

We constructed a nomogram model to predict the probability of recurrence in patients with resected stage I *EGFR*-mutant lung adenocarcinoma using the multivariate stepwise regression analysis results (Fig. 9). The three factors for the nomogram model were pathologic stage, *TP53* co-mutation, and NPA. The total score was determined based on the individual scores calculated using the nomogram. The two-, three-, and five-year recurrence probabilities in patients with resected stage I *EGFR*-mutant lung adenocarcinoma could be calculated using the nomogram.

Table 2. Prognostic factors for recurrence based on multivariate analysis of resected EGFR-mutant lung adenocarcinoma

		Univariate Analysis			Multivariate Analysis		
Category	**Variables**	**HR**	**95% CI**	**p-value**	**HR**	**95% CI**	**p-value**
Age	≥65 vs. <65*	0.91	0.52-1.59	0.74			
Sex	Male vs. female*	1.00	0.57-1.75	0.99			
Smoking history	Ever vs. non-smoker*	1.00	0.54-1.83	0.99			
Pathologic stage	IB vs. IA*	3.68	2.03-6.68	<0.0001	3.27	1.78-6.03	0.0001**
TP53 co-mutation	*TP53/EGFR* vs. *EGFR**	1.96	1.00-3.85	0.0493	2.17	1.10-4.29	0.0252**
NPA	High(3-4) vs. Low(1-2)*	3.24	1.84-5.68	<0.0001	2.68	1.50-4.76	0.0008**
PI3K pathway	RTK/RAS-PI3K vs. RTK/RAS*	1.60	1.25-2.05	0.0002			
p53 pathway	RTK/RAS-p53 vs. RTK/RAS*	1.69	1.16-2.46	0.0057			
Cell cycle pathway	RTK/RAS-cell cycle vs. RTK/RAS*	1.74	1.16-2.61	0.0072			
RPS6KB1 Amp.	*EGFR/RPS6KB1* amp. vs. *EGFR**	3.12	1.76-5.51	<0.0001			
EGFR Amp.	*EGFR/EGFR* amp. vs. *EGFR**	2.11	1.09-4.08	0.0265			
CHEK2 Amp.	*EGFR/CHEK2* amp. vs *EGFR**	2.49	1.32-4.70	0.0046			

Abbreviations: Amp, amplification; CI, confidence interval; HR, hazard ratio; EGFR, epidermal growth factor receptor; NPA, number of pathway alteration; RTK, receptor tyrosine kinase. *Reference variable. **p values were calculated using multivariate stepwise Cox regression analysis.

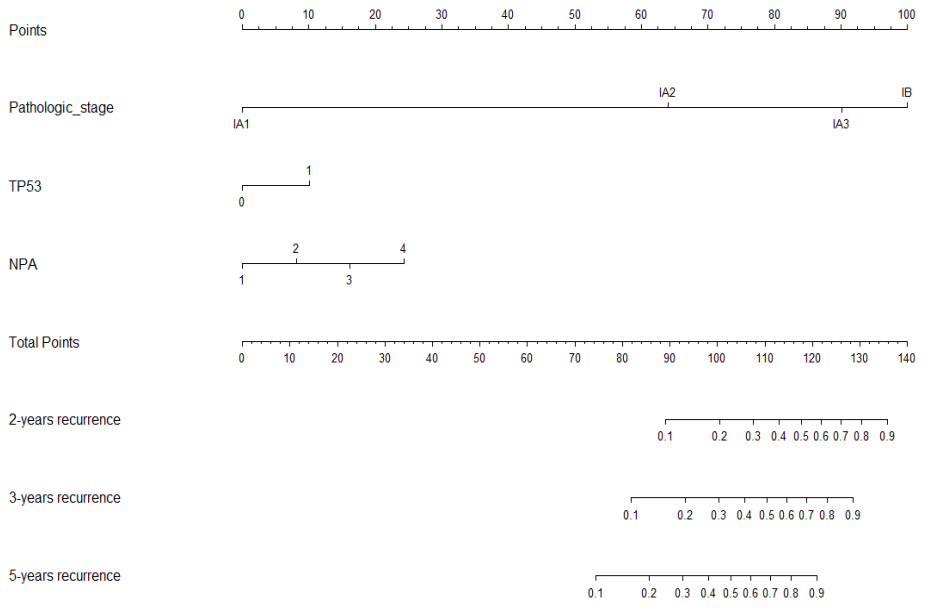


Figure 9. Nomogram for predicting recurrence in stage I, *EGFR*-mutant lung adenocarcinoma patients. Abbreviations: *TP53*, *TP53* co-mutation; NPA, Number of pathway alteration.

IV. DISCUSSION

We retrospectively evaluated the prognostic implication of genetic profiles in *EGFR* mutant, TKI untreated, resected stage I lung adenocarcinoma using broad-panel NGS. We analyzed early-stage tumor specimens with low tumor burden to identify the most potent genomic alterations associated with recurrence. Major findings were as follows: (1) total number of pathogenic genetic alterations was correlated with shorter RFS; (2) *EGFR-TP53* co-occurring mutation was a negative predictive marker for recurrence; (3) high NPA ($\geq$ 3) was correlated with shorter OS and RFS; (4) several co-occurring pathway alterations, such as RTK/RAS-PI3K, RTK/RAS-p53, and RTK/RAS-cell cycle, were associated with shorter RFS than the RTK/RAS pathway alteration only; (5) amplification of

EGFR, *RPS6KB1*, and *CHEK2* were correlated with worse OS (*RPS6KB1*) or RFS. (6) High MSI score was correlated with worse RFS; and (7) *TP53* co-mutation, high NPA, and higher pathologic stage were independent risk factors for recurrence on multivariate analysis.

Analysis of broad panel NGS usually depends on a gene-centric approach to lead clinical decision making for targeted therapy, but most genomic data could not be used. We presumed that the pathway-centric approach could provide additional genomic information to identify patients at high risk of recurrence following resection of lung adenocarcinoma. Therefore, we investigated 11 co-occurring oncogenic pathway alterations and NPA, and identified that certain co-occurring pathway alterations (PI3K, p53, and cell cycle) and high NPA (≥ 3) were associated with worse prognosis. We also found that high NPA was an independent prognostic marker for recurrence in the study population. Similar findings have been reported in a study of resected lung adenocarcinoma which identified NPA as an independent risk factor for poor DFS; and showed that cell cycle, Hippo, TGFß, and p53 pathway alterations were associated with poor DFS[15]. Another recent study reported that higher NPA, TMB, and copy number alterations are associated with increasing histologic subtype invasiveness and worse clinical outcomes in stage I-III lung adenocarcinoma[17].

The *RPS6KB1* gene encodes a member of the ribosomal S6 kinase family of serine/threonine kinases. The encoded protein responds to *mTOR* signaling (mammalian target of rapamycin) to promote protein synthesis, cell growth, and cell proliferation[18]. Amplification of *RPS6KB1* gene is reported in approximately in 10% of breast cancers, and is associated with worse clinical outcomes[19]. The prevalence of *RPS6KB1* amplification in lung cancer is little known. Hyperphosphorylation of RPS6KB1 (rather than overexpression) was recently correlated with worse clinical outcomes in NSCLC patients after surgical excision[20]. In our study, *RPS6KB1* gene amplification was frequently identified (23.8%) in patients with *EGFR* mutant lung adenocarcinoma, and was associated

with worse OS and RFS. Thus, co-occurring *EGFR* mutation/*RPS6KB1* amplification could potentially be used as a novel predictive marker for identifying patients at high risk for recurrence following resection of lung adenocarcinoma. However, further study of *RPS6KB1* amplification in a larger sample size with functional validation is needed to test this hypothesis.

EGFR-TP53 co-mutation is well-known and the most frequent co-mutation in all subtypes of lung cancer[21, 22]. The prognostic impact of *TP53* co-mutation is controversial; several studies suggested that *TP53* alterations may be associated with treatment resistance and shorter survival in patients with *EGFR* mutant lung cancer[23, 24, 25, 26]. On the other hand, a couple of studies focused on the prognostic implications of *TP53* co-mutation in patients with TKI-treated advanced-stage disease could not validate the prognostic impact of *TP53* co-mutation[27, 28]. It was recently suggested that early *TP53* co-mutation may influence clinical outcomes through facilitating genomic instability and the acquisition of additional co-occurring driver events in *EGFR*-mutant early-stage lung adenocarcinoma[29]. We found that *TP53* co-mutation was the most frequent SNP and it was an independent prognostic marker for recurrence in *EGFR*-mutant early-stage lung adenocarcinoma.

In our study, concurrent *EGFR* amplification and mutation was shown in 22 (21%) samples and was associated worse RFS. The prevalence of *EGFR* amplification in *EGFR* mutated NSCLC is varied (range 8-81%) and more frequently involved the mutant allele[30, 31, 32, 33, 34]. In one Asian cohort study and another Latino cohort study, patients with concurrent *EGFR* amplification and mutation had a better response to first/second-generation TKI compared with patients without *EGFR* amplifications[29, 30]. A recent study using amplicon-based targeted NGS reported high *EGFR* copy number to be associated with poor OS in *EGFR* mutant patients treated with EGFR-TKIs[35].

Checkpoint kinase 2 (CHEK2) encodes a serine/threonine kinase (Chk2) that is involved in the DNA damage-response (DDR) pathway. Being activated by

DNA damage-sensor proteins, ataxia-telangiectasia mutated (ATM) protein activates CHEK2 protein, after which CHEK2 phosphorylates effector proteins including BRCA1, p53 and cell division cycle 25C (CDC25C) that have important function in DNA repair, cell cycle regulation and apoptosis[36, 37]. A recent study demonstrated that high *CHEK2* gene expression is correlated with poor prognosis in patients with lung adenocarcinoma, and *CHEK2* pathway may be important for the proliferation of lung adenocarcinoma, especially in tumors with chromosomal instability (CIN)[38]. In our study, concurrent *EGFR* mutation/*CHEK2* amplification was found in 15% of patients and was associated with worse RFS.

The co-occurrence of genetic alterations in early-stage *EGFR* mutant lung adenocarcinoma may indicate functional synergies with *EGFR* mutation and may reflect resistance to therapy targeting *EGFR* mutations. Furthermore, understanding the detailed mechanisms, extent, and co-occurrence of the oncogenic alterations in the 11 selected pathways may facilitate the development of new therapeutic strategies to improve patient care, such as risk stratification of resected early-stage *EGFR*-mutant lung adenocarcinoma, and the development of novel candidates for combination targeted therapy.

Our study has several limitations. First, it is a retrospective study and germline mutations were not excluded because we did not collect blood or normal tissue samples. To prevent reporting false positive germline mutations, our data was filtered using a large human database and the house database of Macrogen Inc (https://www.macrogen.com). Second, we used DNA sequencing data alone, while previous pathway analyses were curated using RNA sequencing data[14]. Therefore, genes that were epigenetically silenced through hypermethylation would have been missed by our analysis. Third, our patient group consisted of ethnic Koreans only and data were not confirmed by using a large database such as TCGA, which includes other ethnicities.

V. CONCLUSION

Our results suggests that broad-panel NGS with a pathway-centric analysis strategy may provide valuable information for risk stratification of patients with resected early-stage *EGFR* mutant lung adenocarcinoma. Worse clinical outcome was correlated with increased number of genetic alterations, high number of pathway alterations, concurrent PI3K, p53, cell cycle pathway alterations, co-occurring *TP53* mutation, and concurrent amplification of *EGFR, RPS6KB1, CHEK2*. Independent risk factors for recurrence were higher pathologic stage, *TP53* co-mutation, and high NPA. These findings may help identify patients at high risk of recurrence who may benefit from adjuvant EGFR-TKIs therapy or chemotherapy.

REFERENCES

1. Chansky K, Detterbeck FC, Nicholson AG, et al. The IASLC lung cancer staging project: external validation of the revision of the TNM stage groupings in the eighth edition of the TNM classification of lung cancer. *J Thorac Oncol*. 2017;12(7):1109-21

2. Hung J-J, Jeng W-J, Hsu W-H, et al. Prognostic factors of postrecurrence survival in completely resected stage I non-small cell lung cancer with distant metastasis. *Thorax.* 2010;65(3):241-5

3. Midha A, Dearden S, McCormack R. EGFR mutation incidence in non-small-cell lung cancer of adenocarcinoma histology: a systematic review and global map by ethnicity (mutMapII). *Am J Cancer Res*. 2015;5(9):2892–911.

4. D. Planchard, S. Popat, K. Kerr, S. Novello, E.F. Smit, C. Faivre-Finn, et al. Metastatic non-small cell lung cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up, *Ann Oncol*. 2019;30(5):863-870

5. Wu YL, Tsuboi m, He J, John T, Grohe C, Majem M, et al. Osimertinib in resected EGFR-mutated Non small cell lung cancer. *N Engl J Med.* 2020;383:1711-23

6. P. Christopoulos, M. Kirchner, J. Roeper, F. Saalfeld, M. Janning, F. Bozorgmehr, et al. Risk stratification of EGFR+ lung cancer diagnosed with panel-based next-generation sequencing. *Lung Cancer.* 2020;148:115-112

7. Y. Zhao, S. Wang, Z. Yang, Y. Dong, Y. Wang, L. Zhang, et al. Co-Occurring potential actionable oncogenic drivers in non-small cell lung cancer. *Front Oncol*. 2021. doi; 10.3389/fonc.2021.665484/

8. Paolo B, Maria LR, Tessa S, Fabrizio T, Andrea C, Angela L, et al. Clinical and molecular features of EGFR mutation positive NSCLC patients treated with TKIs: Predictive and prognostic role of co-mutations. *Cancers.* 2021, 13, 2425. doi: 10.3390/cancers13102425

9. Korf BR, Rehm HL. New approaches to molecular diagnosis, *JAMA.* 2013;309:1511-21

10. Li MM, Datto M, Duncavage EJ, Kulkarni S, Lindeman NI, Roy S, et al. Standards and Guidelines for the Interpretation and Reporting of Sequence Variants in Cancer: A Joint Consensus Recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists. *J Mol Diagn*. 2017;19:4-23.

11. Yin P. Hung, Amanda Redig, Jason L. Hornick, Lynette M. ARID1A mutations and expression loss in non-small lung carcinomas: clinicopathologic and molecular analysis. *Modern Pathology.* 2020;33:2256-2268

12. Walter MJ, Shen D, Shao J, Ding L, White BS, Kandoth C, et al. Clonal diversity of recurrently mutated genes in myelodysplastic syndromes. *Leukemia.* 2013;27:1275-82

13. Chakravarty D, Gao J, Phillips SM, et al. OncoKB: a precision oncology knowledge base. *JCO Precis Oncol.* 2017;2017. doi.org/10.1200/PO.17.00011.

14. Sanchez-Vega F, Mina M, Armenia J, Chatila WK, Luna A, La KC, et al. Oncogenic signaling pathways in The Cancer Genomic Atlas. *Cell.* 2018;173:321-37

15. Jian Z, Francisco SV, Raul C, Kay ST, Whitney SB, Gregory DJ, et al.

Analysis of tumor genomic pathway alterations using broad-panel Next-Generation Sequencing in surgically resected lung adenocarcinoma. *Clin Cancer Res.* 2019;25:7475-84

16. Martini N, Bains MS, Burt ME, et al. Incidence of local recurrence and second primary tumors in resected stage I lung cancer. *J Thorac Cardiovsc Surg.* 1995;109(1):120-9

17. Raul C, Francisco SV, Kay ST, Brooke M, Jian Z, Gregory DJ, et al. The underlying tumor genomics of predominant histologic subtypes in lung adenocarcinoma. *J Thorac Oncol.* 2020;15(12);1844-56

18. Bahrami-b F, Ataie-Kachoie, Pourgholami MH, Morris DL. P70 Ribosomal protein S6 kinase(RPS6KB1): an update. *J Clin Pathol.* 2014;67(12):1019-25

19. Gizeh PT, Elin K, Marie AW, Birgit O, Birgitta H, Bo N, et al. Clinical potential of mTOR targets S6K1 and S6K2 in breast cancer. *Breast Cancer Res Treat.* 2011;128(3):713-23

20. Chen B, Yang L, Zhang R, Gan Y, Zhang W, Liu D, et al. Hyperphosphorylation of RPS6KB1, rather than overexpression, predicts worse prognosis in non-small cell lung cancer patients. *PLoS One.* 2017;12(8):e0182891

21. Cancer Genome Atlas Research N. Comprehensive molecular profiling of lung adenocarcinoma. *Nature.* 2014;511(7511):543-550

22. Cancer Genome Atlas Research N. Comprehensive genomic characterization of squamous cell lung cancers. *Nature.* 2012;489(7417):519-525

23. Canale M, Petracci E, Delmonte A, et al. Impact of TP53 mutations on outcome in EGFR-mutated patients treated with first-line tyrosine kinase

inhibitors. *Clin Cancer Res.* 2017;23(9):2195-2202

24. Bria E, Pilotto S, Amato E, et al. Molecular heterogeneity assessment by next-generation sequencing and response to gefitinib of EGFR mutant advanced lung adenocarcinoma. *Oncotarget.* 2015;6(14):12783-12795

25. Molina-Vila MA, Bertran-Alamillo J, Gasco A, et al. Nondisruptive p53 mutations are associated with shrter survival in patients with advanced non-small cell lung cancer. *Clin Cancer Res.* 2014;20(17):4647-4659

26. Clinical Lung Cancer Genome P, Network Genomic M. A genomics-based classification of human lung tumors. *Sci Transl Med.* 2013;5(209)ra153

27. Shepherd FA, Lacas B, Le Teuff G, et al. Pooled analysis of the prognostic and predictive effects of TP53 comutation status combined with KRAS or EGFR mutation in early-stage resected non-small-cell lung cancer in four trials of adjuvant chemotherapy. *J Cliln Oncol.* 2017;35(18):2018-2027

28. Labbe C, Cabanero M, Korpanty GJ, et al. Prognostic and predictive effects of TP53 co-mutation in patients with EGFR-mutated non-small cell lung cancer. *Lung Cancer.* 2017;111:23-29

29. Rahul N, Weiwei Z, Tong Z, Angela T, Alexis JK, YIN YL, et al. Elucidating the genomic architecture of Asian EGFR-mutant lung adenocarcinoma through multiregion exome sequencing. *Nat Commun.* 2018;9(1):216

30. Ruiz-Patino A, Castro CD, Ricaurte LM, Cardona AF, Rojas L, Zatarain-Barron ZL, et al. EGFR amplification and sensitizing mutations correlate with survival in lung adenocarcinoma patients treated with erlotinib (MutP-CLICaP). *Targ Oncol.* 2018;13(5):621-9

31. Shan L, Wang Z, Guo L, Sun H, Qiu T, Ling Y, et al. Concurrence of EGFR amplification and sensitizing mutations indicate a better survival benefit from EGFR-TKI therapy in lung adenocarcinoma patients. *Lung Cancer.* 2015;89(3):337-42

32. Lewandowska MA, Czubak K, Klonowska K, Jozwicki W, Kowalewski J, Kozlowski P. The use of a two-tiered testing strategy for the simultaneous detection of small EGFR mutations and EGFR amplification in lung cancer. *PLoS One.* 2015;10(2):e0117983

33. Varella-Garcia M, Mitsudomi T, Yatabe Y, Kosaka T, Nakajima E, Xavier AC, et al. EGFR and HER2 genomic gain in recurrent non-small cell lung cancer after surgery: impact on outcome to treatment with gefitinib and association with EGFR and KRAS mutations in a Japanese cohort. *J Thorac Oncol.* 2009;4(3):318-25

34. Miller VA, Riely GJ, Zakowski MF, Li AR, Patel JD, Heelan RT, et al. Molecular characteristics of bronchioloalveolar carcinoma and adenocarcinoma, bronchioloalveolar carcinoma subtype, predict response to erlotinib. *J Clin Oncol.* 2008;26(9):1472-8

35. Jiacong W, Pei M, Miente MT, Anke VR, Menno T, Frank S, et al. Clinical value of EGFR copy number gain determined by Amplicon-based targeted next negation sequencing in patients with EGFR-mutated NSCLC. *Target Oncol.* 2021;16(2):215-226

36. Antoni L, Sodha N, Collins I, Garrett MD. CHK2 kinase:cancer susceptibility and cancer therapy- Two sides of the same coin? *Nat Rev Cancer.* 2007;7(12):925-936

37. Perona R, Moncho-Amor V, Machado-Pinilla R, Belda-Iniesta C, Sanchez PI. Role of CHK2 in cancer development. *Clin Transl Oncol.* 2008;10(9):538-54

38. Tatsuro O, Mikihiro K, Kensaku I, Kazuki T, Masakazu K, Yosuke M, et al. Clinical significance of DNA damage response factors and chromosomal instability in primary lung adenocarcinoma. *Anticancer Research.* 2017;37:1729-1735

ABSTRACT(IN KOREAN)

EGFR **돌연변이 조기 폐 선암종에서 광범위 패널 차세대 염기 서열 분석을 이용한 재발 예측 바이오마커의 발견**

<지도교수 심 효 섭>

연세대학교 대학원 의학과

이 재 석

배경: 완전 수술절제에도 불구하고 *EGFR* 돌연변이가 있는 조기 폐 선암종의 재발율은 높은 편이며, *EGFR* 돌연변이가 있는 조기 폐 선암종 환자의 생존이나 재발을 예측할 수 있는 명확한 바이오마커는 없는 상황이다. 광범위 패널 차세대 염기 서열 분석은 폐암환자에서 표적치료를 위한 대상 유전자 변이를 찾기 위해 근래 널리 쓰이는 유전자 분석방법이며, 그 결과는 임상적 예후인자들과 함께 환자의 예후 예측에도 도움을 줄 수 있다. 본 연구에서는 차세대 염기 서열 분석의 결과를 경로(pathway)에 기반한 분석방법을 이용하여 접근하여 재발 환자군과 비재발 환자군에서의 차이를 비교하였고, 그 결과 재발 등 예후를 예측할 수 있는 유전적 변화를 알아내고자 하였다.

방법: 완전 절제수술을 받고 1기 (stage I), *EGFR* 돌연변이가 확인된 105명의 환자 조직을 이용하여 523개 또는 170개 유전자를 분석할 수 있는 광범위 차세대 염기 서열 분석을 시행하였다. 11개의 잘 알려진 종양형성 신호전달경로에 대한 분석을 하였으며 변경된 신호전달경로

개수를 세었다. *EGFR* 돌연변이와 동시 발생하는 유전적 변이, *EGFR* 돌연변이가 속해 있는 RTK/RAS 경로 변화와 동시 발생된 신호전달경로 변화, 그리고 변경된 신호전달경로 개수 등과 무재발 생존과의 관계를 분석하였다.

결과: 병원성 유전적 변화의 총 개수는 비재발 환자군에 비해 재발 환자군에서 더 높았다 (P<0.001). *EGFR/TP53* 동시 돌연변이가 있는 환자군은 *EGFR* 변이만 있는 환자군에 비해 짧은 무재발생존을 보였다. (P=0.045). 변경된 신호전달경로 개수가 높은(3개 이상) 환자군이 낮은(1-2개) 환자군에 비해 짧은 생존(P=0.031)과 무재발생존(P<0.0001)을 보였다. RTK/RAS-PI3K (P<0.001), RTK/RAS-p53 (P=0.0015), RTK/RAS-cell cycle (P=0.0017)과 같이 특정 신호전달경로의 변화가 RTK/RAS 신호전달경로 변화와 같이 일어났을 때 RTK/RAS 신호전달경로 변화만 있는 환자군에 비해 짧은 무재발생존을 보였다. *EGFR, RPS6KB1, CHEK2* 유전자 증폭은 그것이 없는 환자군에 비해 짧은 생존 (*RPS6KB1,* P<0.001) 또는 짧은 무재발생존 (P=0.023 in *EGFR,* P<0.0001 in *RPS6KB1,* P=0.0035 in *CHEK2*)을 보였다. 높은 미세위성불안정성 (microsatellite instability) 점수는 짧은 무재발생존과 연관을 보였다. *TP53* 동시 돌연변이 (위험비율 2.17, P=0.0252), 높은 변경된 신호전달경로 개수 (위험비율 2.68, P<0.001), 그리고 높은 병리적 병기 (위험비율 3.27, P<0.001)는 다변량 분석에서 재발을 예측할 수 있는 독립적인 위험 인자로 분석되었다.

결론: 경로 기반 분석을 통한 광범위 패널 차세대 염기 서열 분석 결과는 수술절제 받은 조기, *EGFR* 돌연변이 폐 선암종 환자에서 재발 등 예후 예측에 유용한 정보를 제공할 수 있다. 본 연구에서는 높은 유전적 변이의 총 개수, 높은 변경된 신호전달경로 개수, *TP53*

동시발생 돌연변이, RTK/RAS 신호전달경로 변화와 동시 발생하는 PI3K, p53, cell cycle 신호전달경로 변화, *EGFR, RPS6KB1, CHEK2* 유전자 증폭 등이 나쁜 예후와의 연관을 보였다. 본 연구결과를 활용하여 재발 고위험 환자군을 선별하여 추가적인 항암요법, EGFR-TKI 치료 등을 통해 더 나은 치료성적을 거둘 수 있을 것으로 기대한다.

핵심되는 말 : 폐 선암종, 조기 폐암, EGFR 돌연변이, 차세대 염기 서열 분석