이학 박사 학위논문

Establishment and characterization of 29 cell lines derived from lung cancer patient:

Multi-omics analysis based upon immune biomarker PD-L1

폐암환자 유래 29개 세포주의 수립과 특성 분석:

면역 생물 지표 PD-L1에 근거한 통합 유전체 분석

2023년 8월

서울대학교 대학원
의학과 협동과정 종양생물학 전공
박 지 민

Establishment and characterization of 29 cell lines derived from lung cancer patient:

Multi-omics analysis based upon immune biomarker PD-L1

A thesis submitted to the Department of Cancer Biology for the Degree of Doctor of Philosophy in Graduate School of Medicine at Seoul National University

April 2023

Ji min Park

Confirming the Ph.D. Dissertation written by
Ji min Park

July 2023

Chair _____ (Seal)
Vice Chair _____ (Seal)
Examiner _____ (Seal)
Examiner _____ (Seal)
Examiner _____ (Seal)

Abstract

박 지 민 (Ji min Park)

의학과 협동과정 종양생물학과 전공 (Cancer Biology Major)

The Graduate School of Medicine

Seoul National University

Despite advances in cancer diagnosis and treatment techniques, lung cancer is a disease with high incidence and mortality both domestically and globally. Most of the recently introduced immuno-anticancer drugs show limited effects, side effects, and drug resistance. So, it is crucial to find advanced oncogenic and malignant mechanisms of lung cancer development through analysis of genetic mutations and signal transduction in immune oncology. After establishing 29 cell lines derived from the pleural fluid of lung cancer patients, we suggest clinical treatment strategy by analyzing the characteristics of the patient characteristics. As personalized treatment for cancer patients has progressed, further research on the molecular carcinogenic pathway of tumors and biomarkers and signal transduction pathways of tumor genes was needed. A short tandem repeat (STR) test was confirmed to be the same cell line as the patient, and then characterization study was performed. With the established cell lines, the characteristics of gene mutations and genetic information for each cell line were identified to reveal the

correlation with genetic mutations. As expression of PD-L1 is examined as a screening method for the progress of immunotherapy, the established cell lines were analyzed by dividing the cell lines based on PD-L1 expression through western analysis as a major analysis decision classification. The classified cell lines according to the level of expression of PD-L1 were identified to confirm the genetic mutations and the difference in expression level through whole exome sequencing (WES), and RNA sequencing on the transcriptome. This study reveals the correlation between the expression of PD-L1 and the signal transduction pathway. Several previous studies showed the same trend in terms of the correlation between PD-L1 expression and EGFR mutation. This study further supports the correlation between PD-L1 at protein level and mRNA level in established lung cancer cell lines. In addition, it identifies the correlation between PD-L1 expression and genetic mutation, drug sensitivity. Individual anticancer treatment responses are predicted by identifying the correlation between cell biological characteristics and the expression of PD-L1, which is a key biomarker for immunotherapy. The classification according to the biomarker PD-L1 identifies the relations of genetic variation with drug sensitivity thereby contributing to expanding clinical treatment access and through this, significant results are achieved for

immuno-chemotherapy, also contributing to the future cancer therapy. Based upon precision oncology, the generation of anticancer treatment is becoming personalized. This study shows the necessity and potential applications of anticancer treatment in the immunotherapy.

Keyword: Lung cancer cell lines, characterization, Genomic alteration, PD-L1 expression, Immune related gene, Drug sensitivity

Student Number : 2015-22057

Table of Contents

List of Tables

List of Figures

Introduction

Lung cancer remains the most common cause of cancer related mortality worldwide. Approximately 85% of patients have a group of histological subtypes collectively known as NSCLC, of which Lung adenocarcinoma(LUAD) and Lung Squamous cell carcinoma(LUSC) are the most common subtype[1].

Immunotherapy have demonstrated durable responses and represent a new treatment approach for patients with lung cancer, with promising benefit also seen in patients with advanced lung cancer. Due to the distinct feature of lung cancer related immune microenvironment, clinical guidance is needed on chemotherapy concerning agent, appropriate patient selection, sequencing of therapies, and genetic alteration.

Epithelial growth factor receptor (EGFR) mutation status was reported to be associated with programmed death-ligand (PD-L1) expression[2]. And, the molecular function, mechanism of PD-L1 regulation by EGFR activation and by the potential other subtype for customized treatment created by the combination of target therapy and immune related mutant is considerably related to treatment in lung cancer.

The approach of combination study on immune checkpoint protein

and immune related mutation analysis of lung cancer patients can be a good direction therapy. The high level of PD-L1 contributes to immune tolerance, consistently several alterations in cancer cells are known to impair immune surveillance including lung cancer[3]. This feature could be targeted to assess to understand the contribution of lung cancer genetics to those patients who are treated with immunotherapy. Recently, some studies have proposed that genetic alteration in lung cancer, such as EGFR, ALK, and ROS1 including fusion gene EML4-ALK, is associated with a high level of immune checkpoint protein[4-6]. Also, TP53/K-ras co-mutated lung cancer has been proposed to be associated with high level expression of PD-L1 [7].

We wonder how PD-L1 expression and tumor immune related gene function was affected by different fusion gene and high level PD-L1 expression in EGFR positive NSCLC. Also, these studies were largely unknown. The relationship between the immune related ligand such as PD-L1 expression (immune check point protein) and tumor genomic alteration in 29 lung cancer patient derived cell lines was validated using the immune related gene signature[8-11]. The prognostic role of immune related gene signature was also evaluated [12, 13].

The final goal of this study is to confirm the accessibility of lung

cancer treatment in immunotherapy by discovering the relationship between PD-L1 expression gene mutation, and transcriptome analyses in 29 cell lines from patient with lung cancer.

In the study, in order to explore the mechanism of immune related signaling pathway in NSCLC, we identified the immune related signaling pathway specific differential expression genes (DEGs) in PD-L1 positive cell lines and the proportion of immune related gene in mRNA level between absence of PD-L1 expression NSCLC and PD-L1 positive cell line in patient derived lung cancer.

In addition, mRNA level in immune related genes and PD-L1 protein expression level in NSCLC were constructed respectively to predict the potential relationship between underlying immune gene signatures and possible therapeutic targets to patient with NSCLC in the future.

Materials and Methods

Table 1. Overall methodological schematics

Experimental approaches	Method
• Establishment and maintenance of patient-derived cell lines derived from lung cancer pleural fluid (Growth properties, morphology)	• Growth pattern, doubling time, cell morphology in vitro (microscopy) • Mycoplasma test (check a negative cross-contamination)
• Genomic DNA extraction and DNA fingerprinting analysis	• Genomic DNA extraction and DNA fingerprinting analysis using 16 short tandem repeat markers
• Confirmation of PD-L1 expression in the established cell line	• Protein extraction and western blotting
• Genome analysis through WES (Whole Exome Sequencing)	•Oncoplot, mutation trait table
• Analysis of correlation between transcript gene PD-L1 (CD274) and protein PD-L1 through RNA sequencing	• Pearson correlation coefficient analysis using TPM
• TOP 30 genes transcriptome analysis through RNA sequencing	• p heatmap using readcount : Analysis of difference in expression level of genes
• Analysis of changes in immune-related genes through RNA sequencing	• Expression level analysis of immune related genes: Bubble plot, Enhanced volcano plot using TPM
• SPIA analysis of genes expressed in transcriptome analysis (analysis of activated/inhibited cell signaling pathways)	• SPIA Two evidence plot (pGDFR<0.05) using TPM
• Signal transduction pathway analysis of gene expression differences between PD-L1 expressing cell lines and non-expressing cell lines	• GSEA (Gene set enrichment analysis)
• Anticancer drug sensitivity analysis through 16 anticancer drug	• Drug sensitivity test

tests • Correlation study between expression of EMT marker proteins and PD-L1	• Protein extraction and western blotting
expressing cell line	• Protein extraction and western blotting
• Statistical analysis	• R program version 3.3.1 with various packages

2.1. Establishment and maintenance of human NSCLC cell lines

The research protocol was reviewed and approved by the institutional review board of the Seoul National University Hospital (IRB No. 1102-098-357). The study was performed in accordance with the Declaration of Helsinki. Written informed consent was obtained from all patients enrolled in this study. Cell lines were established from malignant pleural effusion (MPE) derived from pathologically proven lung cancer. A total of 29 MPE specimens of human lung cancer from 29 different patients who underwent a pleural intervention for the palliation of dyspnea were obtained from Seoul National University Hospital (Seoul, Korea). MPE samples were directly transferred from the operating room to the laboratory for cell culture. Tumor cells were spun down by centrifugation the MPE sample at 300 rpm for 5 min, and re-suspended with Opti-MEM I (GIBCO, CA, USA) supplemented with 5% fetal bovine serum (GE Healthcare Life Sciences, Buckinghamshire, UK) and 1% Penicillin/Streptomycin. Gathered

cells were then seeded into T-25cm^2 flasks (Corning, NY, USA). Confined-area trypsinization or scraping method was used to attain a pure tumor cell population when stromal cells like mesothelial cells or fibroblasts grew in the initial culture. 29 lung cancer cell line was maintained in RPMI 1640 medium containing 10% FBS and 1% (v/v) penicillin streptomycin (10,000U/ml). All established cells were incubated in 5% CO2 air at 37° C. The initial passage was sub-cultured when a significant amount tumor cell growth was showed, and successive passages were given at sub-confluence after trypsinization. When one culture population contains both floating and adherent cells, floating cells were gathered by centrifuging the medium and dispersed by pipetting. All established cell lines introduced in this study will be deposited in the Korean Cell Line Bank (http://cellbank.snu.ac.kr) at the initial passage in order to be distributed to researchers worldwide.

2.2. Growth properties and morphology in vitro

To calculate doubling time, the density of 5×10^4/ml to 2×10^5/ml viable cells were seeded into 96 well cell culture plate with a volume of 200 μl, and cell viability was calculated daily for 7-14 days. Once 10 μl of EZ-Cytox solution (Daeil Lab, Seoul, Korea) was added, the plate was incubated at 37 ° C for 2 h, and the proper density of each well was calculated by Multiskan™ GO

Microplate Spectrophotometer (Thermo Fisher Scientific, MA, USA) at 450 nm wavelength. Acquired growth rate values were calibrated with GraphPad Prism 5 (GraphPad Software, CA, USA). The morphology of each cell lines was obtained using phase-contrast microscopy. Mycoplasma contamination was checked by the 16S-rRNA-genebased polymerase chain reaction (PCR) amplification method using e-Myco Mycoplasma PCR Detection Kit (Intron Biotechnology, Gyeonggi, Korea).

2.3. Genomic DNA extraction and DNA fingerprinting analysis

Genomic DNA extraction was performed using QIAamp DNA Mini kit (Qiagen). Genomic DNA extracted from each lung cancer cell line was amplified using an AmpFISTR identifiler Polymerase Chain Reaction (PCR) Amplification Kit (Applied Biosystems, CA, USA). A single cycle of PCR amplified 15 short tandem repeat markers (CSF1PO, D2S1338, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D19S433, D21S11, FGA, TH01, TPOX and vWA) and an Amelogenin gender-determining marker containing highly polymorphic microsatellite markers. Amplified PCR products were analyzed by an ABI 3500XL Genetic analyzer (Applied Biosystems).

2.4. RNA extraction and RNA sequencing

Total cellular RNA was extracted from 29 lung cancer cell lines

and cell pellets of cell lines using TRIzol reagent(ambion by Invitrogen, CA, USA)according to the protocol provided by the manufacturer(Qiagen RNeasy Kit).

Sequencing libraries were prepared using the Illumina TruSeq Stranded Total RNA Library Prep Kit. Fifty-one million reads were obtained from the cell lysates. Following base-calling and alignment with the Tuxedo Suite, rejected reads were analyzed using FusionMap, ChimeraScan and Defuse with default parameters for RNA and alignment to GRCh37.72. The output was filtered to include in-frame fusions, with at least one rescued read and two unique seed reads, and exclude known, recurrent artifacts.

2.5. Drug sensitivity test

At density of 2×10^5 to 4×10^5 viable cells from each cell line were seeded into well of 96 well plate in triplicate to measure drug sensitivity of 18 compounds. After 72 h-incubation at 37 ° C, 10ul EZ-Cytox solution was added to well of each seeded lung cancer cells. After 2 h-incubation at 37 ° C, optical density of EZ-Cytox-treated cells was calculated by Multiskan™ GO Microplate Spectrophotometer (Thermo Fisher Scientific). These steps were repeated in triplicate. The half maximal effective concentration (EC50) were measured by GraphPad Prism 5 (GraphPad Software). The maximum concentration and solvent are as follows.

2.6. Protein extraction and Western blotting analysis

Cells were lysed in EzRIPA buffer (ATTO Co., Tokyo, Japan) and centrifuged at 16,000 × g for 30 min to remove debris. Protein concentration of each cell line was determined by SMART micro BCA protein assay kit (Intron biotechnology). Cell lysates (10 μg) were boiled in sodium dodecyl sulfate (SDS) sample buffer (Nacalai Tesque, Inc.) and proteins were then electrophoresed on a 4–12% Bis–Tris gel (Invitrogen) at 70 V for 3 h and were transferred onto PVDF membranes (Invitrogen) by electro–blotting in condition under a constant current of 80 mA at 4 ° C overnight. The membranes were blocked with 1.5% to 2.0% skim milk at room temperature for 1 h and then incubated over night at 4 ◦C in a shaker with primary antibodies. The Membranes were washed several times with PBS containing 0.05% Tween 20. The membranes were incubated for 1h at room temperature with horseradish peroxidase conjugated secondary antibodies (Jackson Immunoresearch, PA, USA, 1:5000). Detection was performed using a chemiluminescent working solution, WESTZOLTM (Intron biotechnology), the membrane were finally developed using Fuji RX film (Fujiflm, Tokyo, Japan) for 1–5 min. The primary antibody dilution factor and secondary antibody information are provided in the table below.

Table 2 . List of antibodies used western blotting

Target	2nd species	Company	Cat no.	Dilution
EGFR	Rabbit	Cell signaling	#4267	1:1000
HER2	Rabbit	Cell signaling	#2165	1:1000
PD-L1	Rabbit	Cell signaling	#12984	1:1000
STAT3	Mouse	Cell signaling	#9139	1:1000
YAP1	Rabbit	Abcam	ab76252	1:5000
Vimentin	Rabbit	Cell signaling	#5741	1:2000
E-cadherin	Mouse	Abcam	#1416	1:500
Slug	Rabbit	Cell signaling	#9585	1:1000
Snail	Rabbit	Cell signaling	#3879	1:1000
ZO-1	Rabbit	Abcam	ab59720	1:250
m-TOR	Rabbit	Cell signaling	#2938	1:1000
B-actin	Mouse	Invitrogen	MA5-15739	1:5000

2.7. Whole Exome Sequencing

SureSelect sequencing libraries were prepared according to manufacturer' s instructions (Agilent sureselect all Exon kit 50 Mb) using the Bravo automated liquid handler. Tree micrograms of genomic DNA were fragmented to a median size of 150 bp using the Covaris-S2 instrument (Covaris, Woburn, MA). The adapter ligated DNA was amplified by PCR, and the PCR product quality was assessed by capillary electrophoresis (Bioanalyzer, Agilent). The hybridization buffer and DNA blocker mix were incubated for 5 min at 95 ° C and then for 10 minutes at 65 ° C in a thermal cycler. The hybridization mixture was added to the bead suspension and incubated for 30 min at RT while mixing. The

beads were washed, and DNA was eluted with 50 ml SureSelect elution buffer (Agilent). The flow cell loaded on HISEQ 2500 sequencing system (Illumina).

2.8. Statistical analysis

Statistical analysis was performed using R program version 3.3.1 (R Foundation for Statistical Computing, Vienna, Austria) with various packages including maftools, Performance Analytics, survminer, survival, iplot, gplot, and lattice. Data analysis was performed using R statistical software (http://www.R-project.org/) and Differential gene expression analyses were performed using the DESeq2 R package. The correlations between PD-L1 expression and other variables were determined by Pearson or Spearman correlation coefficients. Differences were considered to be statistically significant at a two-side $p < 0.05$. Fisher's exact test was used to analyze GO analysis of various genes. A multivariate analysis of variance (MANOVA) model was applied to the drug response data matrix with various factors such as the mutational status and the three different transcriptional subtypes. Approximate F value, p-value and Pillai's trace score were obtained for each of the factors/drug pairs. A value of $p<0.05$ was considered statistically significant. Clustering methods include: Ward's minimum variance method, complete linkage method, k-

means method, and single linkage method.

Results

3.1 Patient derived lung cancer cell lines show various morphological and molecular patterns.

The established 29 cell lines were all derived from lung cancer patients and the morphologies of the cell lines were categorized into three subtypes: polygonal, oval, and fibroblast-like. (Figure 2, Table 3A). The doubling time of each cell lines varied with a range of 32 to 106 h (Table 3A). Most of the 29 cell lines had adherent growth patterns, and among them, there were cell lines with two growth patterns: adherent and floating pattern (SNU-3652, SNU-3652S). DNA fingerprinting analysis identified a heterogeneous distribution of 15 tetranucleotide repeat loci and an Amelogenin gender determining marker in each cell line, and confirmed 29 cell lines without cross contamination (Table 4).

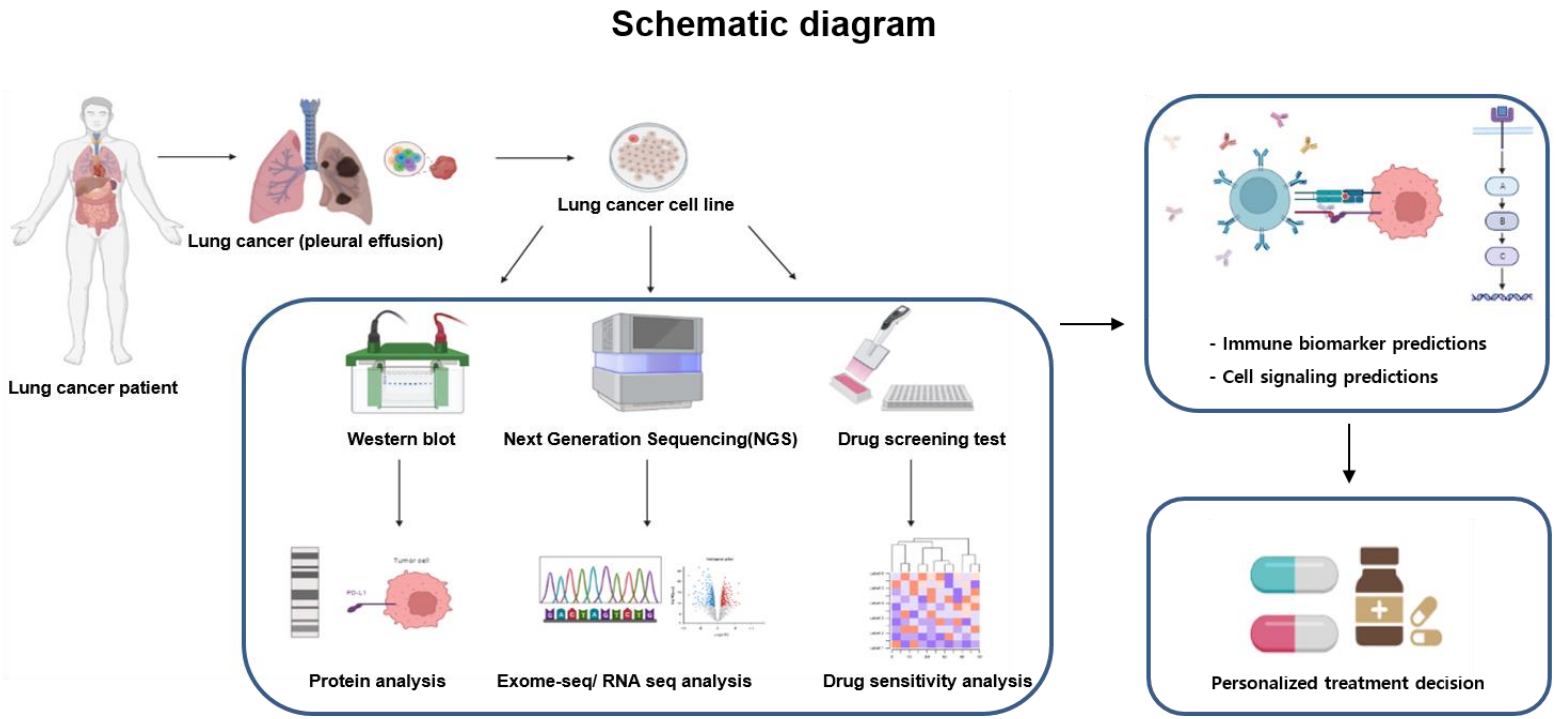

Figure 1. The schematic diagram of experiment design

A Schematic Diagram that provides an overview of research on 29 patient derived lung cancer cell line

Figure 2. The morphologies of patient derived lung cancer cell lines exhibited heterogeneous growth patterns. Scale bar=250 μM

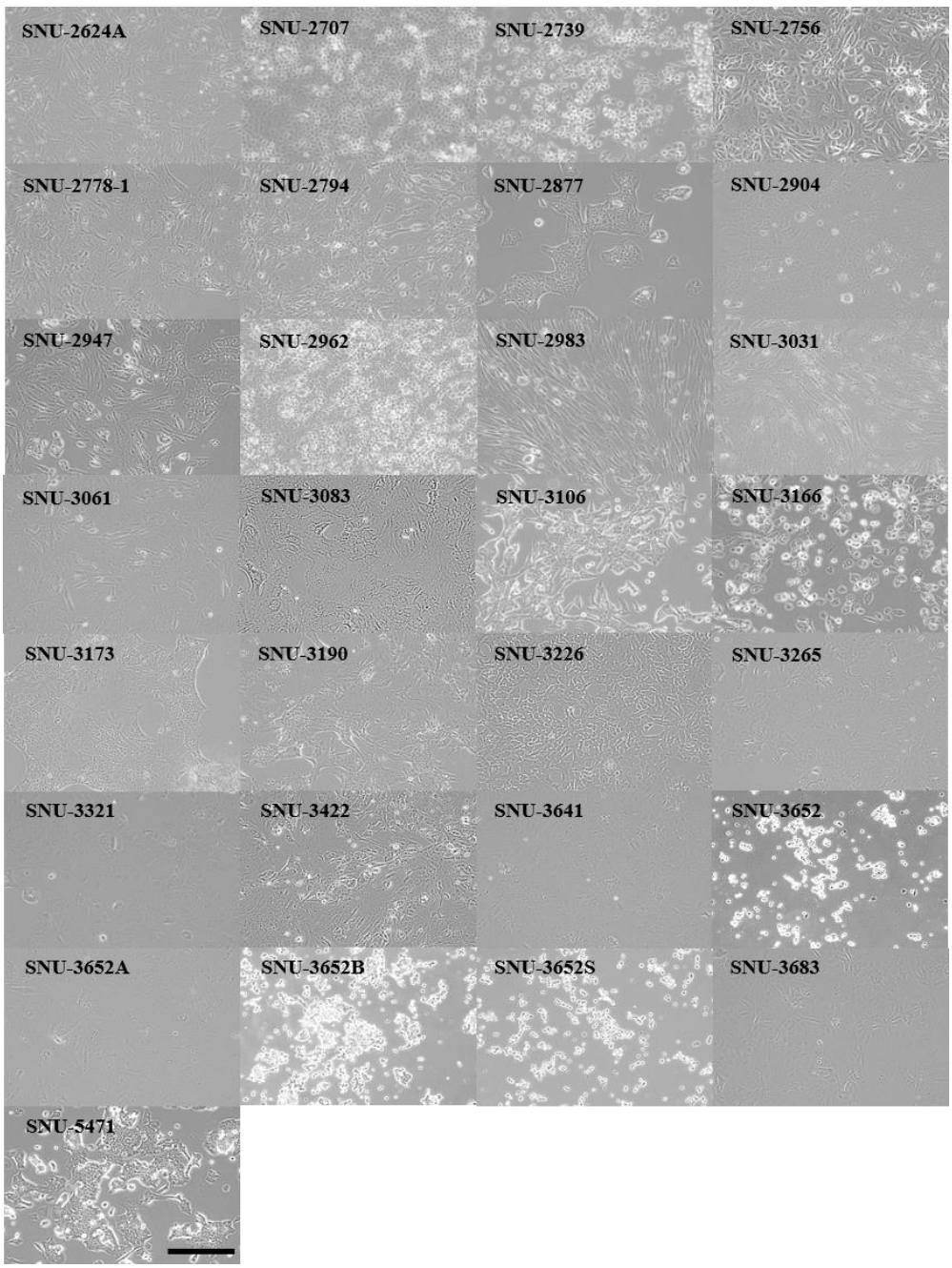

Table 3A. Clinicopathologic characteristics of the established patient derived lung cancer cell lines.

Cell Line	Culture site	Sex	Age	Pathology	Growth pattern	Doubling Time(hr)	Cell Morphology
SNU-2624A	pleural effusion	M	33	NSCLC(ADC)	Adherent	55.45	Oval
SNU-2707	pleural effusion	M	50	NSCLC(ADC)	Adherent	84.55	Oval
SNU-2739	pleural effusion	F	57	NSCLC(ADC)	Adherent	24.52	Oval
SNU-2756	pleural effusion	M	69	NSCLC(ADC)	Adherent	75.19	Fibroblast-like
SNU-2778-1	pleural effusion	M	55	NSCLC(ADC)	Adherent	86.63	Polygonal
SNU-2794	pleural effusion	F	50	NSCLC(ADC)	Adherent	106.2	Fibroblast-like
SNU-2877	pleural effusion	F	58	NSCLC(ADC)	Adherent	93.71	Oval
SNU-2904	pleural effusion	F	44	NSCLC(ADC)	Adherent	50.71	Oval
SNU-2947	pleural effusion	F	65	NSCLC(ADC)	Adherent	67.6	Oval
SNU-2962	pleural effusion	F	58	NSCLC(ADC)	Adherent	73.01	Polygonal
SNU-2983	pleural effusion	M	66	NSCLC(ADC)	Adherent	59.11	Oval
SNU-3031	pleural effusion	F	51	NSCLC(ADC)	Adherent	47.14	Oval
SNU-3061	pleural effusion	M	47	NSCLC(ADC)	Adherent	-	Oval
SNU-3083	pleural effusion	M	69	NSCLC(ADC)	Adherent	60.94	Oval
SNU-3106	pleural effusion	F	52	NSCLC(ADC)	Adherent	52.27	Oval
SNU-3166	pleural effusion	F	33	NSCLC(ADC)	Adherent	46.04	Polygonal
SNU-3173	pleural effusion	M	46	NSCLC(ADC)	Adherent	61.62	Polygonal

SNU-3190	pleural effusion	F	66	NSCLC(ADC)	Adherent	66.56	Oval
SNU-3226	pleural effusion	M	67	NSCLC(ADC)	Adherent	32.53	Oval
SNU-3265	pleural effusion	F	70	NSCLC(ADC)	Adherent	11.73	Polygonal
SNU-3321	pleural effusion	M	50	NSCLC(ADC)	Adherent	-	Oval
SNU-3422	pleural effusion	F	58	NSCLC(ADC)	Adherent	73.7	Oval
SNU-3641	pleural effusion	M	77	NSCLC(ADC)	Adherent	85.57	Oval
SNU-3652	pleural effusion	M	57	SCLC	Adherent/Floating	62.76	Oval
SNU-3652A	pleural effusion	M	57	SCLC	Adherent	64.4	Oval
SNU-3652B	pleural effusion	M	57	SCLC	Adherent	53.93	Polygonal
SNU-3652S	pleural effusion	M	57	SCLC	Adherent/Floating	66.01	Polygonal
SNU-3683	pleural effusion	M	57	NSCLC(ADC)	Adherent	62.11	Oval
SNU-5471	pleural effusion	F	45	NSCLC(ADC)	Adherent	79.01	Oval

* NSCLC(ADC) :Non-small cell lung cancer(Adenocarcinoma), *SCLC : Small cell lung cancer

Table 3B. Clinicopathologic characteristics of the established patient derived lung cancer cell lines.

Cell Line	PD-L1 Status	EGFR Status	TP53 Status	HER2 Status
SNU-2624A	Absence	WT	WT	WT
SNU-2707	Positive	WT	WT	WT
SNU-2739	Absence	WT	WT	ALT
SNU-2756	Positive	WT	WT	ALT
SNU-2778-1	Positive	WT	WT	ALT
SNU-2794	Absence	WT	ALT	WT
SNU-2877	Positive	WT	WT	ALT
SNU-2904	Positive	ALT	ALT	WT
SNU-2947	Positive	ALT	ALT	WT
SNU-2962	Positive	ALT	WT	WT
SNU-2983	Positive	WT	ALT	ALT
SNU-3031	Positive	WT	ALT	ALT
SNU-3061	Absence	ALT	WT	WT
SNU-3083	Positive	WT	ALT	WT
SNU-3106	Positive	WT	WT	ALT
SNU-3166	Absence	WT	ALT	ALT
SNU-3173	Positive	WT	WT	ALT
SNU-3190	Positive	WT	ALT	WT
SNU-3226	Positive	WT	WT	ALT
SNU-3265	Positive	ALT	WT	WT
SNU-3321	Absence	ALT	WT	WT
SNU-3422	Absence	WT	WT	ALT
SNU-3641	Positive	ALT	WT	ALT
SNU-3652	Absence	WT	WT	WT
SNU-3652A	Positive	WT	WT	WT
SNU-3652B	Absence	WT	WT	WT
SNU-3652S	Absence	WT	WT	WT

SNU-3683	Positive	ALT	WT	ALT
SNU-5471	Positive	WT	WT	ALT

*ALT: Alteration, *WT : Wild type

Table 3C. Clinicopathologic characteristics of the established patient derived lung cancer cell lines.

Clinicopathologic Characteristics of 29 Lung Cancer Cell Lines

Cell Line	Major Mutation Event(s)	EGFR TKI Naïve vs Resistance
SNU-2624A	BRAF(V600G), KIF5B-RET	
SNU-2707		
SNU-2739	ERBB2(P1170A)	
SNU-2756	ERBB2(P1170A)	
SNU-2778-1	ERBB2(P1170A), KIF5B-RET	
SNU-2794	TP53(A281A)	
SNU-2877	ERBB2(P1170A), BRCA1(G1038G), EML4-ALK	
SNU-2904	EGFR(T790M),TP53(T234H)	EGFR TKI (iressa) Resistance
SNU-2947	TP53(S20F), BRCA1(G1038G), EGFR(T790M)	EGFR TKIs(iressa/osimertinib) Resistance
SNU-2962	EGFR(G746_Ala750d), BRCA1(G1038G), PIK3CA(G545L), BRCA1(G1038G)	EGFR TKI (iressa) Resistance transformed SCLC
SNU-2983	TP53(C242P), ERBB2(P1170A), BRCA1(G1038G)	
SNU-3031	TP53(A239S), ERBB2(P1170A), BRCA1(G742G), BRCA1(G1038G)	
SNU-3061	EGFR(L858A), BRCA2(L467*), BRCA1(G1038G), APC(V1125A), APC(V1107A)	EGFR TKI (Tarceva) Resistance
SNU-3083	RET(G691S), TP53(A213L), BRCA1(G1038G)	

Cell Line	Major Mutation Event(s)	EGFR TKI Naïve vs Resistance
SNU-3106	ERBB2(P1170A), BRCA1(G1038G), EML4-ALK	
SNU-3166	TP53(H178G), ERBB2(P1170A), BRCA1(G1038G), EML4-ALK	
SNU-3173	RET(G691S), ERBB2(P1170A), RET(G691S)	EGFR TKI Naïve
SNU-3190	BRAF(V600G), TP53(A280L)	
SNU-3226	ERBB2(P1170A), BRCA1(G1038G)	EGFR TKI (iressa) Resistance
SNU-3265	EGFR(L858A)	EGFR TKI Naïve
SNU-3321	ALK(L1404V), EGFR(T790M), EGFR(L858A), BRCA1(G1038G), ALK(L1404I)	EGFR TKI (iressa) Resistance
SNU-3422	BRAF(V600G), ERBB2(P1170A)	
SNU-3641	EGFR(Leu858A), ERBB2(Pro1170A), BRCA1(Glu1038G)	EGFR TKI (iressa) Resistance
SNU-3652	RB1(C706P)	
SNU-3652A	RB1(C706P)	
SNU-3652B	RB1(C706P)	
SNU-3652S	RB1(C706P)	
SNU-3683	EGFR(G746_A750), EGFR(T790M), ERBB2(P1170A), BRCA1(G1038G), EGFR(G701_A705), EGFR(G693_A697)	EGFR TKI (poziotinib) Resistance
SNU-5471	ERBB2(A775_Gly776), BRCA1(G1038G)	

Table 4. DNA Fingerprinting Analysis using 15 STR loci and Amelogenin for newly established 29 SNU Lung cancer cell lines

Cell Name	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	TH01	D13S317	D16S539
SNU-2624A	13,14	30	8,10	10	15,16	9	8,12	9
SNU-2707	11	30	9,11	12	17	9	13	10,11
SNU-2739	10,16	32	12	10,11	15	9	8	9,11
SNU-2756	15	28,30	8,9	12	15	6,9	12	12
SNU-2778-1	13,16	32.2	12	10,11	15,17	8	8	9,10
SNU-2794	13,15	26,29	10	9,11	15,16	8	8,11	9
SNU-2877	12,14	29,31.2	8,11	12	17	9	12	11,13
SNU-2904	12	30,32.2	12	9,13	16	9,9.3	11	12
SNU-2947	13,15	29	8,10	10,12	16,17	6,10	8	12
SNU-2962	11	30	10	14,15	16	7,9	11	13
SNU-2983	14	29	11	12.13	17	9	8	11
SNU-3031	15,16	29,32.2	11	10,11	16	6,9	8	11
SNU-3061	10,12	30,32.2	8,11	12	15	7	8,9	12
SNU-3083	11	30	10	14,15	16	7,9	11	13
SNU-3106	12,13	33	10,12	10	16	7,9	8	10,12
SNU-3166	13	31,32.2	8	12	15	9	8	10
SNU-3173	11,14	30	8	11,12	17	9	10	10
SNU-3190	12,14	29	10,12	10,11	17,18	7,9	8,9	11,12
SNU-3226	10,15	30,31	10	11,12	15	6	12	9
SNU-3265	14	29,32.2	10,11	11,12	18	7,9	8	9,11
SNU-3321	12	29,31	11	10,12	15	7,9.3	9	10,11
SNU-3422	10,11	29,30	10	11,12	15	7,9	9	10
SNU-3641	12,14	30	8,11	12	15	6,9	8	11
SNU-3652	13,15	29,33.2	8,11	12,13	15	9	8	9
SNU-3652A	13,15	29,33.2	8,11	12	15	9	8	9
SNU-3652B	13,15	29,33.2	8,11	12	15	9	8	9
SNU-3652S	13,15	29,33.2	8,11	12	15	9	8	9
SNU-3683	11,13	30	10	13	15	7	10	10
SNU-5471	12,16	32.2	9,11	11,12	15,18	9	8,11	12

continued

Cell Name	D2S1338	D19S433	Vwa	TPOX	D18S51	Amelogenin	D5S818	FGA
SNU-2624A	19,21	13,14	18	8,10	13,16	X,Y	9	18,21
SNU-2707	17,26	13,14.2	19	8,10	13,14	X	11,13	24,25
SNU-2739	17,22	13,14.2	16,18	8,11	20	X	11,12	20,21
SNU-2756	19	16	17	11	15	X	12	21
SNU-2778-1	23,26	14,15.2	14	8	20	X	10,12	22,23
SNU-2794	17,18	14,15.2	18	8,11	12,16	X	10,12	20,23
SNU-2877	23,25	15.2	14,17	8	12	X	9,11	22,23
SNU-2904	18,22	13,13.2	14	8	14	X	11	22
SNU-2947	17,24	13	17	11	14	X	10,11	23,24
SNU-2962	17,23	13	16	8,11	17	X	11	23
SNU-2983	26	14	17	8	16,21	X,Y	11,13	25
SNU-3031	17,19	12,14	14,17	8	14	X	11	18,22
SNU-3061	23,25	13,14	17	8	14,16	X,Y	11	22,24
SNU-3083	17,23	13	16	8,11	17	X	11	23
SNU-3106	19,24	13,14	16	9,12	13	X	10,12	24
SNU-3166	19	14.2	14	9	18	X	11,12	18
SNU-3173	18,19	14,14.2	16	9,11	16,17	X,Y	10,12	23,24
SNU-3190	20,23	14	16,18	8	13,19	X	11,13	18
SNU-3226	19,20	13,14	16	8,11	16	X,Y	9,10	17,24
SNU-3265	18,23	13,15.2	18	8	14,17	X	9,11	21
SNU-3321	20,23	14.2	18	9,10	13,14	X,Y	10,12	21,24
SNU-3422	18,25	14	15,17	8,11	16	X	11	21,22,23
SNU-3641	20,25	14.2,15.2	17	8,11	14	X	13	22
SNU-3652	24	13,14	17	8	14,15	X,Y	10	23
SNU-3652A	24	13,14	17	8	14,15	X,Y	10	23
SNU-3652B	24	13,14	17	8	14,15	X,Y	10	23
SNU-3652S	24	13,14	17	8	14,15	X,Y	10	23
SNU-3683	17	13,14	14,18	11	18	X,Y	13	20,24
SNU-5471	19,21	13	16	8	13,15	X	10,11	22,25

Table 5. Mycoplasma test by Cycleave PCR Testing

Cell Name	25000 기준
NC	Undetermined
PC	24.64016533
SNU-2624A	Undetermined
SNU-2707	Undetermined
SNU-2739	Undetermined
SNU-2756	Undetermined
SNU-2778-1	Undetermined
SNU-2794	Undetermined
SNU-2877	Undetermined
SNU-2904	Undetermined
SNU-2947	Undetermined
SNU-2962	Undetermined
SNU-2983	Undetermined
SNU-3031	Undetermined
SNU-3061	Undetermined
SNU-3083	Undetermined
SNU-3106	Undetermined
SNU-3166	Undetermined
SNU-3173	Undetermined
SNU-3190	Undetermined
SNU-3226	Undetermined
SNU-3265	Undetermined
SNU-3321	Undetermined
SNU-3422	Undetermined
SNU-3641	Undetermined
SNU-3652	Undetermined
SNU-3652A	Undetermined
SNU-3652B	Undetermined
SNU-3652S	Undetermined
SNU-3683	Undetermined
SNU-5471	Undetermined

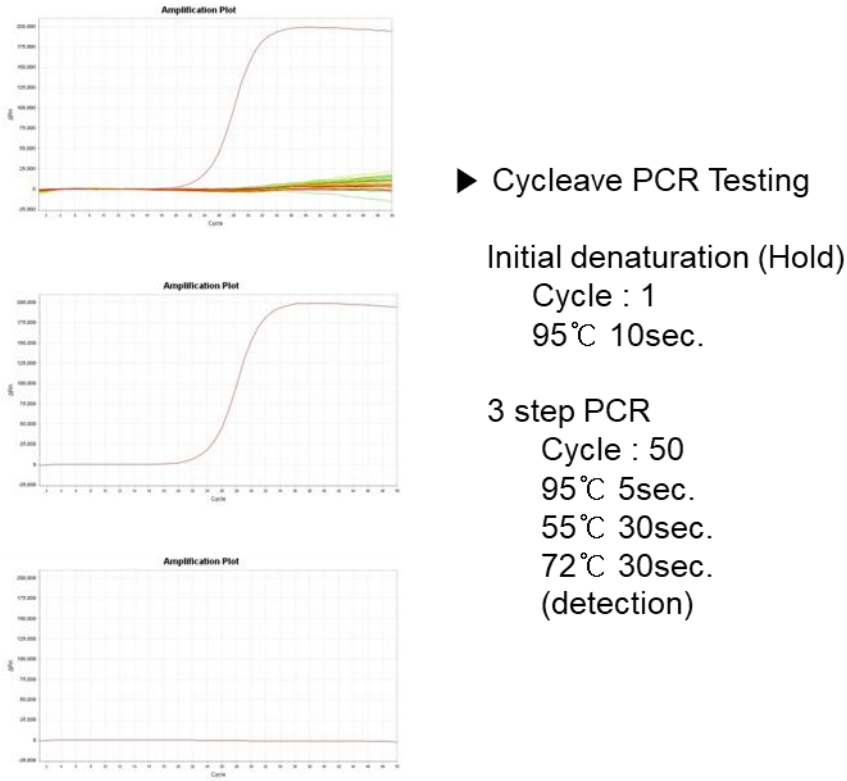

Figure 3. Established 29 patients derived lung cancer cell lines mycoplasma test by the 16S-rRNA-gene-based polymerase chain reaction (PCR) amplification method

The experiment was performed through Cycleave PCR Testing, and it was confirmed that there was no cross contamination of each cell line. ⊿Rn(Delta Rn) : The magnitude of the signal generated by a given set of PCR conditions.

Although each cell lines were confirmed to be not cross-contaminated, the STR authentication between the cells obtained from lung cancer patients and matched cell lines has not been performed, which limited the association of clinical information on section. All 29 cell lines were confirmed to be free of contamination from mycoplasma (Table 5, Figure 3). The clinic pathological information are also reported in Table 3A. For instance, Table 3A are included gender, age, mutation characteristics, and pathology information.

Blood samples from 15 lung cancer patients and matched 15 established lung cancer cell lines were analyzed for various mutations on genomic DNA (Fig 6). No EGFR mutation was detected in the cohort of blood sample gDNA, the 15 established lung cancer cell lines was detected in 5 patients (33%). Among 15 patients, ALK mutations were detected in 5 patients (33%). As a result, 15 established lung cancer cell line has more various genetic mutations, fusion gene than normal lung cell line.

3.2 PD-L1 positive cell line has more various genetic mutations, fusion gene than absence of PD-L1 expression cell line. (Correlation between PD-L1 expression and clinicopathologic features)

Several studies have shown that K-ras and Met mutations have high PD-L1 expression [2]. And, some studies show that EGFR and TP53 mutations occur more frequently in patients with high PD-L1 expression [3]. The relationship between the expression level of PD-L1 and key mutation genes in the 29 lung cancer cell lines we established was performed. we divided the expression of PD-L1 into three categories at the protein level, which is shown in Fig 4. A schematic structure of this study is shown in Fig 1 with detailed description of the experiment component specified.

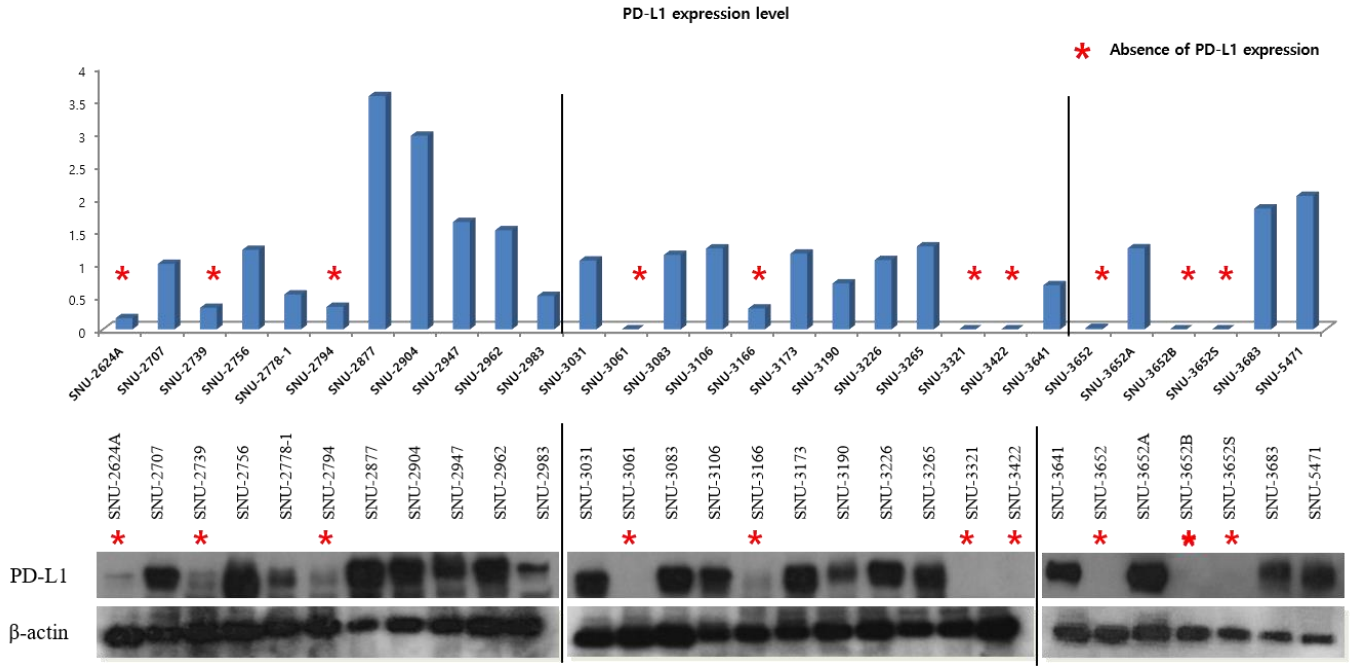

Figure 4. Normalized PD-L1 expression level status of Western blotting analysis using loading control as β-actin.

Patient derived 29 lung cancer cell lines were classified PD-L1 status in two types by western blotting analysis, dividing PD-L1 positive cell lines (n=19) and absence of PD-L1 (n=10).

To investigate PD-L1 expression level, we assessed PD-L1 protein level to 29 cell lines. And then, we divide protein level into three levels via western blot.

We are categorized in the PD-L1 expression level into PD-L1 high expression, PD-L1 low expression and absence of PD-L1 expression. As a result of subtyping by protein level, 19 cell lines had PD-L1 positive expression, 4 cell lines had PD-L1 low expression, and 6 cell lines without expression (Table 3B, Fig 4).

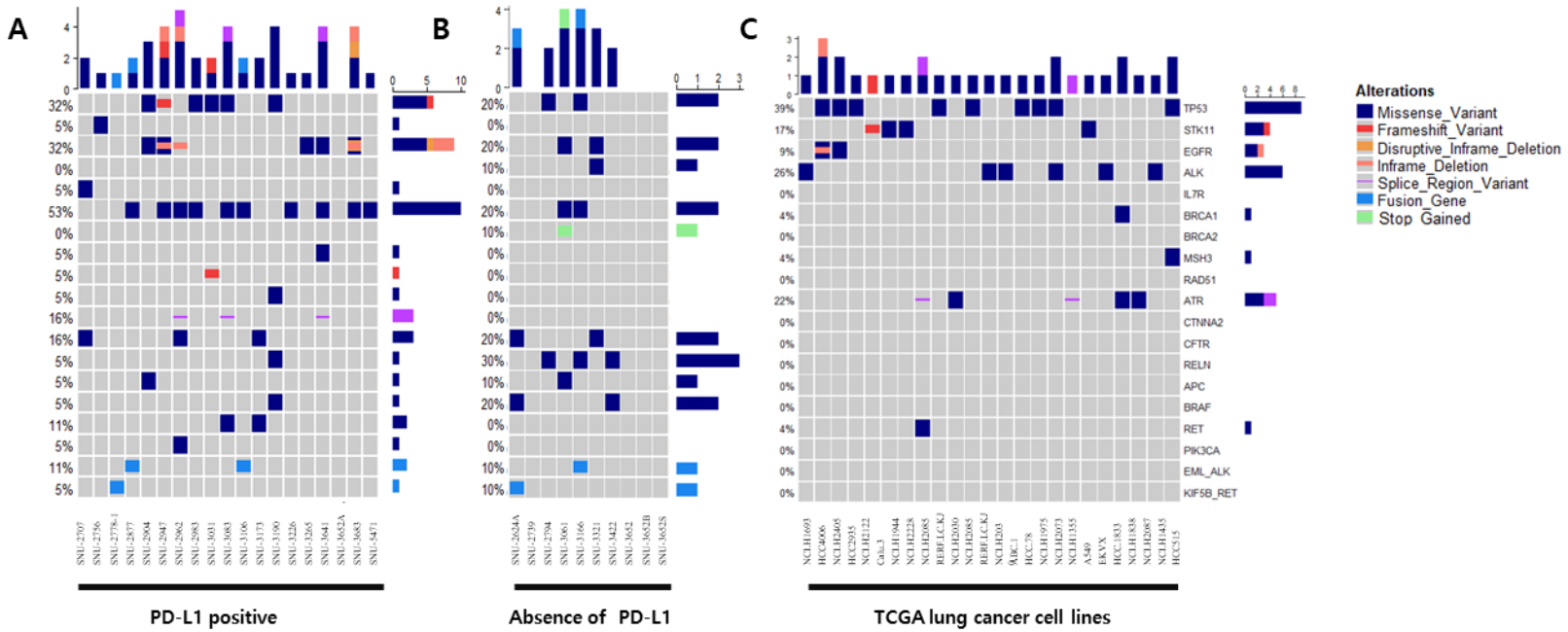

Figure 5. Mutational landscape of the established lung cancer cell lines.

A. Mutational landscape of the established lung cancer cell lines harboring PD-L1 expression (n=19)

B. Mutational landscape of the established lung cancer cell lines with Absence of PD-L1 expression (n=10)

C. Mutational landscape of TCGA generated data (n=29)

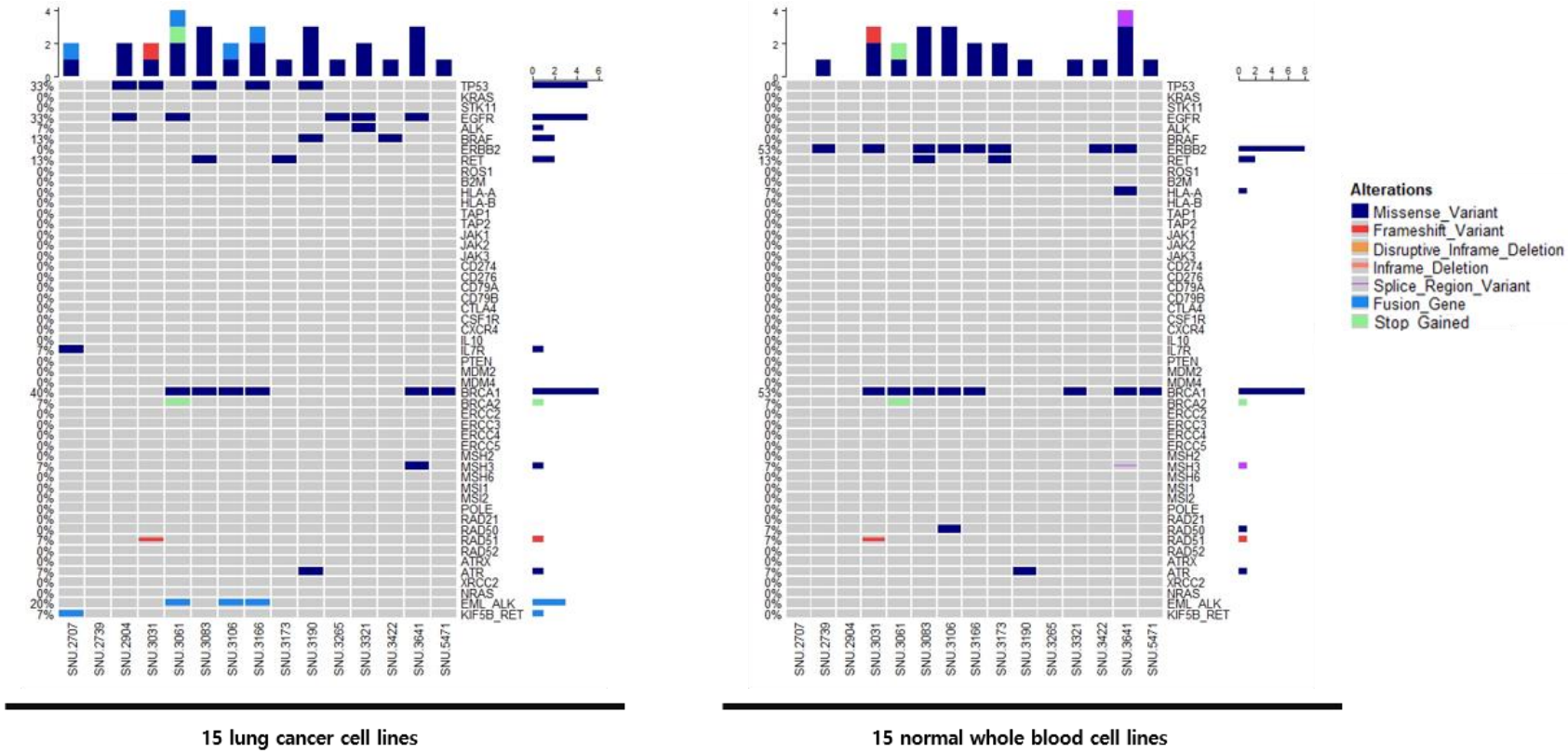

Figure 6. Mutational landscape of the 15 established lung cancer cell lines paired whole blood cell lines

Taken together, PD-L1 status was shown on the oncoplot shown through WES (whole exome sequencing). The precise exclusion of the germinant mutations was not feasible since the DNA from matched blood or normal tissue was unavailable when WES of established cell lines was performed. Alternatively, we have referred to the Clinvar database (https://www.ncbi.nlm.nih.gov/clinvar) in order to avoid misestimating the mutation frequency due to ascended germinant/ benign mutations. The pathogenicity of the representative driver mutations was manually examined, and all marked mutations for calculating the frequency of mutated tumor genes in Fig. were formerly reported as “Pathogenic”, “Likely Pathogenic” or “Drug response” in the Clinvar database.

Of the 29 cell lines, 9 cell lines had EGFR mutations, and 8 cell lines out of the 9 cell lines had EGFR mutations (Table 3C). We confirmed that it was an EGFR mutant type. In addition, in the cell line of the mutant type of TP53, there were 10 cell lines out of 29 cell lines, except for 3 cell lines. The 7 cell lines of mutant type of TP53 were PD-L1 positive cell lines (Table 3C). In lung cancer, the fusion gene accounts for a significant portion of study that is a gene mutation that occurs frequently[14]. In some studies fusion gene occurs frequently in lung cancer, which also results in a

higher level of PD-L1 expression than absence of PD-L1 expression[4-7, 11]. The mutation profile showed that the cell lines with a fusion gene including ALK mutation had PD-L1 expression (Table 3C). In other words, all cell lines with a fusion gene among 29 cell lines had mainly PD-L1 expression, and there were two types of EML-ALK(echinoderm microtubule-associated protein-like 4-anaplastic lymphoma kinase) and KIF5B-RET(kinesin family member 5B-ret proto-oncogene) fusion genes. Detection was performed by PCR to confirm the presence or absence of the fusion gene in the 29 established cell lines. The cell lines with KIF5B (exon16)-RET were SNU-2624A and SNU-2778-1, and the cell lines with EML4 (exon6)-ALK were SNU-2877, SNU-3106 and SNU-3166 (Figure 7). Among the cell lines we established, two cell lines with EML3-ALK and one cell line with KIF5B-RET could be identified in the cell line with PD-L1 expression (Figure 5A). In the absence of PD-L1 cell line, one cell line with EML3-ALK and one cell line with KIF5B-RET could be identified (Figure 5B). Also, other studies show that about half of lung. Another research show that genetic analysis of the EGFR, KRAS, and ERBB2 genes demonstrated that the EML4-ALK fusion gene is mutually exclusive with mutations in the previous represented genes [14]. We did not find that 9 of the 29 cell lines

contain the EML4-ALK[15] fusion gene in cell lines with EGFR mutations. The presence of EML4-ALK appears to coexist with TP53 mutations at a low frequency in lung cancer [8]. We were able to find that all of the BRCA1 mutations were common in the SNU-2877, SNU-3106, and SNU-3166 cell lines with the fusion gene EML4-ALK, and among them, SNU-3166 had TP53 mutation. And, when the key mutation of the established cell line and the trait of TCGA lung cancer were compared, a similar degree could be seen (Figure 5C).

3.3 Transcriptomic analysis correlated with value of protein level in PD-L1 expression.

We first sought to identify correlations of cell surface PD-L1 expression to protein from mRNA level. Consistent with previous report[9], PD-L1(CD274) mRNA levels are correlated to cell surface PD-L1 protein levels in human lung cancer cell lines(Pearson correlation coefficient, R = 0.36, p=0.058)(Fig 8). Correlation analysis of mRNA expression was performed to see the relationship between PD-L1 expression in the 29 established cell lines and each EGFR/TP53/HER2 mutations. As a result, the expression of EGFR and PD-L1 had a significant mRNA expression, and the rest mutations were not shown significant result (Fig 9).

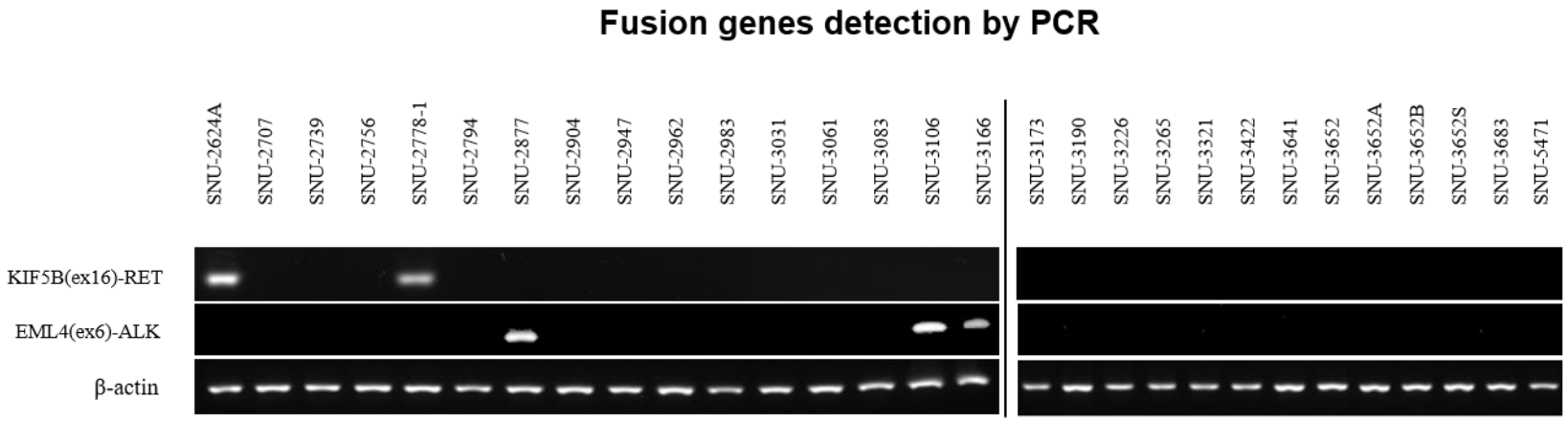

Figure 7. Fusion genes detection by PCR

Representative DNA gel images revealing the expression of the KIF5B-RET fusion transcript and EML4-ALK fusion transcript as determined by reverse transcription polymerase chain reaction.

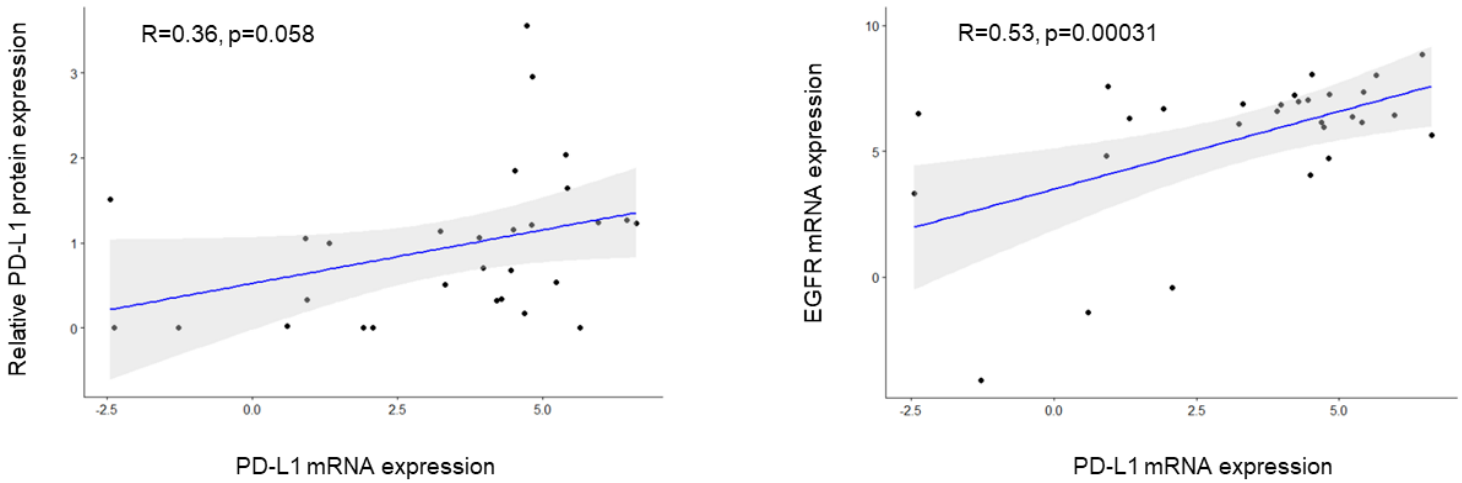

Figure 8. mRNA level of transcriptome analysis correlated with value of protein level in PD-L1 expression.

The Pearson correlation coefficient (R) with p values between the transcriptomic PD-L1 mRNA expression and relative PD-L1 protein expression are indicated on the top of the correlation graph. Each dot represents each cell lines. Also, correlation between EGFR mutation and PD-L1 expression, the results of Pearson correlation coefficient were able to confirm significant results. The confidence interval is designated by shading along with the correlation graph.

As previously mentioned, few studies focus on the association between PD-L1 expression and EGFR mutational status[10]. With regard to EGFR mutation status, 6 cell lines harbored a missense mutation and 4 cell lines had an in frame deletion mutation. 16 cell lines harboring PD-L1 high expression were shown to have EGFR gene mutation. Given the positive association between EGFR mutation and high PD-L1 expression in lung cancer cell lines, we examined the expression of PD-L1 in NSCLC cell lines whose EGFR mutation status had been previously determined. Cell surface expression of PD-L1 was also significantly higher in NSCLC cell lines positive for activating EGFR mutations than in those with wild-type EGFR. We were able to obtain significant results of PD-L1 TPM and EGFR TPM values through Pearson correlation analysis(R=0.53, p=0.00031)(Fig 8).

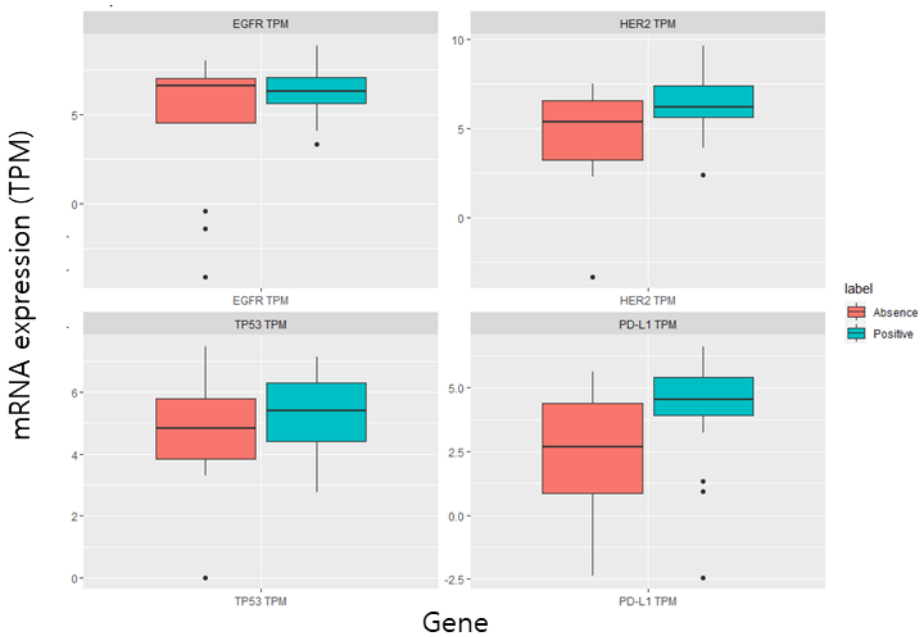

Figure 9. Comparison of mRNA expression levels of specific mutation genes according to PD-L1 groups.

Within the PD-L1 group divided by protein expression, the PD-L1 TPM value was significantly higher in the PD-L1 expression positive group compared to the absence of PD-L1 group.

TP53 mRNA expression (TPM) and HER2 mRNA expression (TPM) also showed high mRNA expression in the PD-L1 positive group except for EGFR mRNA expression (TPM).

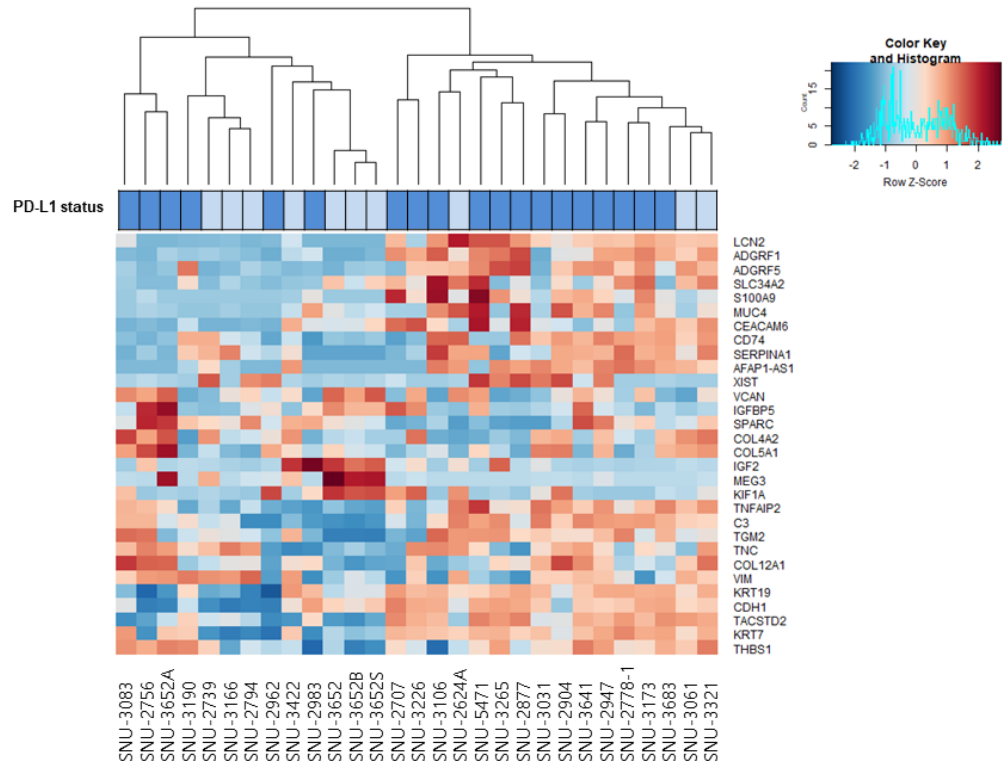

Figure 10. P-heatmap showing the TOP30 gene with a large difference expression level in the group divided into classifications according to PD-L1 status.

Among the genes with a large difference in expression level between the PD-L1 groups, genes related to lung cancer and PD-L1 are shown along with immune related genes (KRT19, KRT7, XIST, TACSTD2, VIM). In addition, genes specifically overexpressed in small cell lung cancer cell lines are also shown (IGF2, MEG3).

Table 6. P-heatmap TOP 30 Gene description

Gene Name	Function/ Clinical implication
MEG3	MEG3 is considered as a tumor suppressive gene and downregulated in NSCLC cells with DDP resistance. MEG3 can activate p53 to inhibit the tumor development and enhance the DDP-mediated apoptosis via Wnt pathway[12, 13]
UCHL1	UCHL1 promotes the expression of PD-L1 in NSCLC cell lines[15]
IGFBP5	Our research indicates that IGF1R could be a potential biomarker for early prediction of drug response and clinical evolution of NSCLC patients. IGFBPs were described to promote or suppress tumor growth in various tissues and contexts /IGFBP5 was suggested to function as a tumor suppressor[16]
KRT17	proliferation gene in lung cancer[17]
FBN2	We successfully clustered and constructed predictive models for TME-associated genes and helped guide immunotherapy strategies[18].
COL4A2	These known fibrosis-associated markers were used to identify populations that might be critical for fibrosis formation following RT[19].
SPARC	Secreted protein acidic and rich in cysteine (SPARC)[20]
XIST	XIST and TSIX were differentially expressed in subtypes of BC patients, and their levels were correlated to PD-L1 expression level[21].
LCN2	Lipocalin 2 (LCN2), an innate immune protein, plays a pivotal role in promoting sterile inflammation by regulating immune responses[22].
S100A9	tumor oncogene/ S100A9, an immunosuppressive molecule was much higher in NKTCL patients both in serum and tumor stroma. Elevated level of S100A9 was associated with advanced stage, poor overall response and early recurrence[23].
MUC4	MUC4 was highly glycosylated, and the expression of EGFR was modulated by MUC4 /MUC4 was reported to modulate the expression of EGFR [24].

ADGRF1	DGRF1 (formerly named GPR110) is a cell surface protein that is overexpressed in lung cancer/ ADGRF1 mRNA expression was positively correlated with E-cadherin (CDH1) and negatively correlated with vimentin and N-cadherin (CDH2), suggesting a potential role of ADGRF1 in the epithelial–mesenchymal transition process[25].
AFAP1-AS1	AFAP1-AS1 was highly expressed in NPC cells and was positively correlated with the expression of PD-1, a key molecular marker of tumor immune evasion[26].
CD74	Expression of CD74, the Receptor for Macrophage Migration Inhibitory Factor, in Non-Small Cell Lung Cancer[27].
TNFAIP2	TNFAIP2 may be a potential tumorigenesis gene in ESCC[28].
TGM2	aberrant expression of TGM2 has been demonstrated to be linked with a series of aggressive phenotypes of tumor cells, such as tumor growth, metastasis, epithelial-mesenchymal transition, and cancer stem cell property[29].
TGFBI	The high-TGFBI-expression group had poorer clinical responses than did the low-TGFBI-expression group ($p < 0.0001$) [30].
KRT7	In lung cancer CTCs, KRT7 and TTF-1 levels have been found to be elevated[31].
LAMC2	LAMC2 promotes the proliferation of cancer cells and induce infiltration of macrophages in non-small cell lung cancer[32].
TACSTD2	TROP2 (also known as TACSTD2) is a transmembrane glycoprotein that is highly expressed in many cancers, and is a promising molecular target for the treatment of various malignancies[33].
FN1	Along these lines we found that PD-L1 transfer rates positively correlated with fibronectin (FN1) mRNA expression levels[34].
VIM	The current study revealed that TUB, GLU, VIM and PD-L1 were overexpressed in CTCs from NSCLC patients[35].

In another research, TP53 mutant was correlated with positive PD-L1 expression in lung cancer patient [2, 36]. However, TP53 TPM values and PD-L1TPM values did not have significant result established 29 lung cancer cell lines. As a result, No statistically significant correlations between the PD-L1 protein level and TP53 mRNA status, as based on Pearson correlation analysis, could be demonstrated when the complete cohort was analyzed. We inferred that the downstream signals of EGFR might regulate PD-L1 expression. Therefore, PD-L1 status was a significant prognostic factor for patient with EGFR mutant type. Because, EGFR mutant type cell line in genomic alteration level or read count value in mRNA might contribute to the PD-L1 expression level in lung cancer. Subsequently, we attempted to report the relationship between the top 30 genes expressed at the established 29 cell line mRNA level and the PD-L1 protein level. Thus, the p-heat map showing the PD-L1 protein status was presented as an analysis of the difference in expression level of the p-heat map. Analysis of 29 established lung cancer cell lines identified expression correlation between PD-L1 expression and immune related gene.

Moreover, we identified expression correlation between 30 genes and PD-L1 protein level, immune related genes and protein level, and total previously mentioned genes for 29 established lung cell

lines.

LAMC2 gene is highly expressed in lung cancer with PD-L1 expression among immune specific genes [32]. In the 29 lung cancer cell lines, it was found that the LAMC2 gene was mainly expressed in the cell line where PD-L1 was expressed, which is also related to the expression of PD-L1. The KRT17 gene is also known to be highly expressed in lung cancer [17]. Overexpression of KRT17 promotes proliferation and invasion of non-small cell lung cancer and indicates poor prognosis. Also, we found that the KRT17 gene was highly expressed in the cell line with PD-L1 expression. And, the UCHL1 gene is known to promote PD-L1 expression in NSCLC [15]. However, in our cell line, it did not have an important meaning regardless of the presence or absence of PD-L1 expression. Another notable gene, XIST gene level is known to be related to PD-L1 expression level [21]. By examining the expression of the XIST gene, it was confirmed that it was expressed in a cell line with PD-L1 expression. Genes such as FN1 and VIM are also highly expressed in lung cancer, and some studies show that they are related to PD-L1[35, 37].

We detected that the genes expression was only in the cell line with PD-L1 expression. Transmembrane proteins such as TACSTD2 are also expressed in many carcinomas, which is also

known to be highly expressed in lung cancer [38]. We were confirmed to be expressed in cell lines with PD-L1 expression in 29 lung cancer cell lines. In conclusion, our findings showed that LAMC2, KRT17, XIST, FN1, VIM, TACSTD2 genes are expected to be considered as markers of PD-L1 expression. Conversely, in the case of IGF1 and MEG3 genes, it could be confirmed that they were expressed only in small lung cancer cell lines without PD-L1 expression[12, 39, 40]. Several recent studies have determined the expression of PD-L1 protein in SCLC with a range of 0.0–82.8% PD-L1-positive detection rates. Among the genes that were expressed in the cell lines without PD-L1 expression in the 29 cancer cell line, IGF1 and MEG3 were confirmed that the most notable gene is the tumor suppressor gene [41]. Among the 29 established cell lines, small cell lung cancer (SNU-3652 series) was found to not express PD-L1 in 3 cell lines (SNU-3652, SNU-3652B, SNU-3652S) except for one(SNU-3652A) of the 4 cell lines. In conclusion, we were able to identify genes important for connectivity at the protein level and mRNA level of lung cancer cell lines according to the presence or absence of PD-L1 expression. It was confirmed that the expression level of each gene may vary according to PD-L1 status. By confirming the level of the genes involved in PD-L1 expression, it will be possible to

get closer to the accessibility of treatment for patients receiving immunotherapy as a PD-L1 up/down predictive marker. We described the top 30 genes in the p-heat map in detail in Table 6.

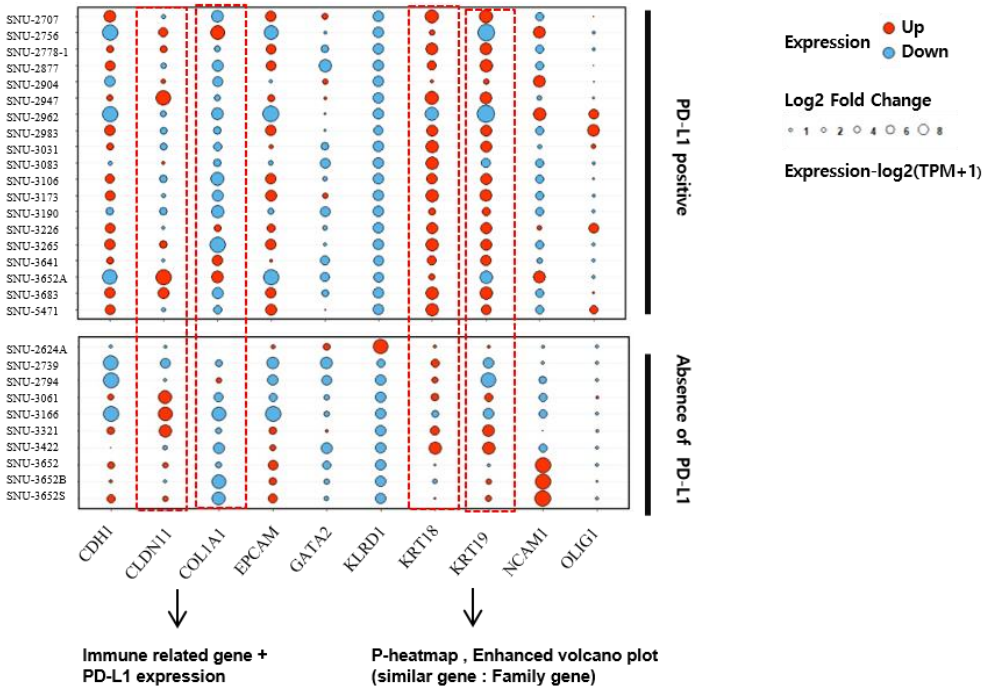

Figure 11. Bubble plot according to the PD-L1 status group for the selected genes that showed differences in expression levels of immune related genes.

Among the immune related genes related to lung cancer and PD-L1 expression, genes that showed specific discrimination are shown according to the cell lines divided into groups.

To characterize the immune related gene profile of 29 lung cancer cell lines with PD-L1 expression, we analyzed the expression of immune related genes using TPM and found all immune related genes to be significantly lower in absence of PD-L1 expression cell lines relative to PD-L1 positive expression cell lines, suggesting PD-L1 status annotation in 29 lung cancer cell lines. In immune related genes, CLDN11 and COL1A1 were specifically up-regulated in PD-L1 positive cell line, showing important results. According to one study, CLDN11 is highly expressed in lung cancer, which has resistance to anti-cancer drugs [39]. In 29 lung cancer cell line, we confirmed that all cell lines in CLDN11 up-regulated were PD-L1 positive cell lines, but the gene down-regulated were in cell lines without PD-L1 expression (Fig 11). In conclusion, COL1A1 can be a potential immunity-related biomarker and therapeutic target in lung cancer[42]. Through these results, we found that among the expression of PD-L1 and immune related genes, CLDN11 and COL1A1 have significant bio-markers, which are immune related genes expressed in PD-L1 positive patients.

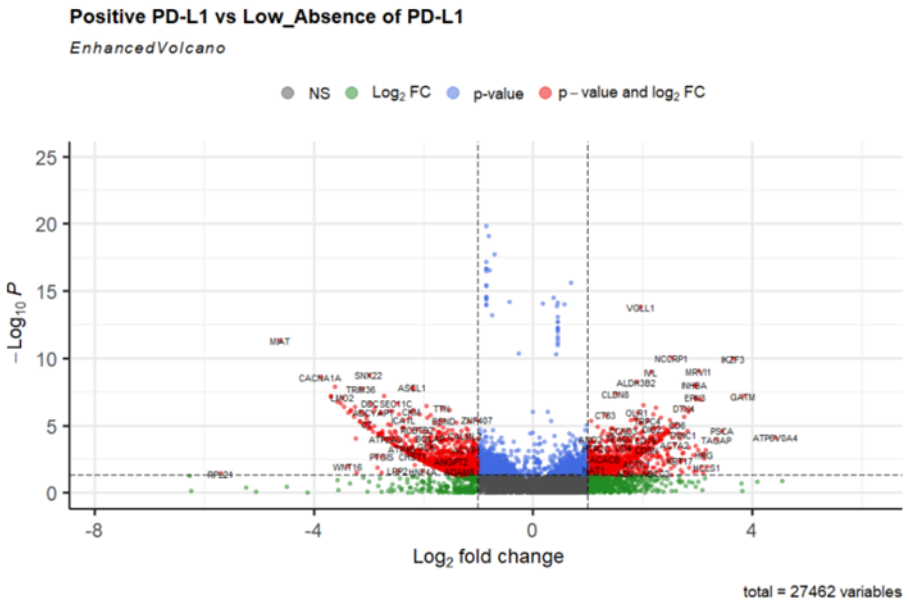

Table . Volcano plot of the all differentially Up-regulated expression Genes

Gene	pvalue
VGLL1	1.49E-14
KRT16P2	9.54E-05
CLDN8	4.13E-08
KRT75	7.64E-05
KRT6A	0.000117744
S100A8	4.34E-05
OLIG1	5.15E-05

Gene identified Bubble plot

Gene identified P-Heatmap

Figure 12. Transcriptional analysis of enhanced volcano plot.

Comparative analysis of gene expression profiles between PD-L1 positive cell lines and absence of PD-L1 cell lines in patients derived lung cancer. Volcano plot presented for illustrating differentially expressed genes (DEGs) in paired PD-L1 positive cell lines/ absence of PD-L1 cell lines. Log2 (fold change) in the non-ILs versus absence

of PD-L1 cell lines is represented on the horizontal axis. Right and left red dots represent the significantly upregulated and downregulated genes, respectively. An FDR-adjusted p value of 0.05 is indicated by using the dark gray dotted line.

Table 7. Enhanced Volcano plot expressed Gene Description

Gene	Function/ Clinical implication
KRT17	The expression level of KRT17 in NSCLCs was significantly higher than normal lung tissues. /Overexpression of KRT17 up-regulated β-catenin activity and levels of Wnt target genes, such as cyclin D1, c-Myc, and MMP7. KRT17 promoted EMT by up-regulating Vimentin, MMP-9, and Snail expression and down-regulating E-cadherin expression[17].
PHYHD1	Phytanoyl-CoA dioxygenase domain containing 1. Oxidation-reduction process; metabolic process/ Dioxygenase activity; oxidoreductase activity; metal ion binding[43].
RAG1	The recombination activating gene (RAG)1. DNA damage caused by RAG1/2 activity in pre–B cells was able to downmodulate RAG1/2 expression and activity, confirming the existence of a negative feedback regulatory mechanism[44].
KRT81	KRT81- a miR-SNP associated with recurrence in non-small-cell lung cancer. KRT81 has emerged as a promising immunohistochemical marker for the identification of squamous cell lung carcinoma[45-47].
KCNJ15	KCNJ15 was significantly downregulated in RCC, and this low expression was an independent prognostic factor for clear cell RCC (ccRCC). KCNJ15 overexpression significantly inhibited RCC cell proliferation, migration, and colony formation, arrested the cell cycle and induced apoptosis of RCC cells in vitro[48].
CLDN16	Via regulating the expression of CLDN16, which may be involved in the chemical carcinogenesis pathway. Co-expressed genes of CLDN16 were enriched in chemical carcinogenesis pathway and metabolism pathways related to cytochromeP450 (CYP). CLDN16 was remarkably overexpressed in LUSC whereas not in LUAD, which was consistent with our knowledge that LUSC was much more associated with smoking. Moreover, co-expression genes with CLDN16 mostly enriched in chemical carcinogenesis pathways[49].
ABCA4	ABCA4, ABCC1, and ABCC show significant associations with lung cancer susceptibility[50].
EMILIN1	Elastin microfibrilinterface located protein 1. A connective tissue glycoprotein associated with elastic fibres. Increased expression levels of EMILIN1 and FBN1 were associated with low proliferation. Suppressive role of emilin 1 is related to the grade of growing breast tumors, and associated with increased hypoxia in the tumor microenvironment

	followed by elevated unfolding and degradation of tissue proteins[51].
PCOLCE	PCOLCE significantly upregulated in gastric cancer patients compared to normal gastric samples. And the increased expression of PCOLCE mRNA was closely linked to shorter overall survival (OS), progress-free survival (PFS) in all gastric cancers. Besides, PCOLCE expression displayed a tight correlation with infiltrating levels of macrophages and dendritic cells (DCs) in gastric cancer. Moreover, PCOLCE expression was positively correlated with diverse immune marker sets in gastric cancer[52].
DIO2	NSCLC patients DIO2 was upregulated in tumor tissue and its expression was higher in smoking patients[53].
HPCAL4	HPCAL4 expression was not present in the normal lung cell lines, but expression was seen in the SCLC cells. This may be a result of the neuroendocrine nature of SCLC[54].
RAP1B	Rap1b overexpression rescued the aggressive characteristics of CRC cells[55].
ACACB	ACACB (also known as ACC2) is responsible for fatty acid metabolism[56].
GPD1	The expression of GPD1 was downregulated in ccRCC tissues, and overexpression of GPD1 inhibited the progression of ccRCC both in vivo and in vitro[57].
AVPR1A	Arginine vasopressin receptor 1A (AVPR1A) acts as receptor for arginine vasopressin, which belongs to the subfamily of G-protein coupled receptors. vasopressin receptor 1a, which is normally associated with maintenance of blood pressure and fluid balance, can help activate signaling pathways involved in the development of castration resistance in prostate cancer[58].
MPZ	Mast cell activated MPZ might play an important role[59].

Table 8A. Volcano plot of the all differentially up-regulated expression genes

Gene	p-value
VGLL1	1.49E-14
KRT16P2	9.54E-05
CLDN8	4.13E-08
KRT75	7.64E-05
KRT6A	0.000117744
S100A8	4.34E-05
OLIG1	5.15E-05

Table 8B. Volcano plot of the all differentially down-regulated expression genes

Gene	p-value
DAZ1	1.42E-20
NRSN1	8.59E-20
AMER3	1.76E-18
GPC5-AS1	6.53E-18
GHRH	1.97E-17
IGSF21	2.18E-17
CACNG5	2.84E-17
KCNK17	3.28E-17
CHRNB3	3.39E-16
TUSC8	4.21E-16
FAM181A	2.54E-15
C7orf33	2.75E-15

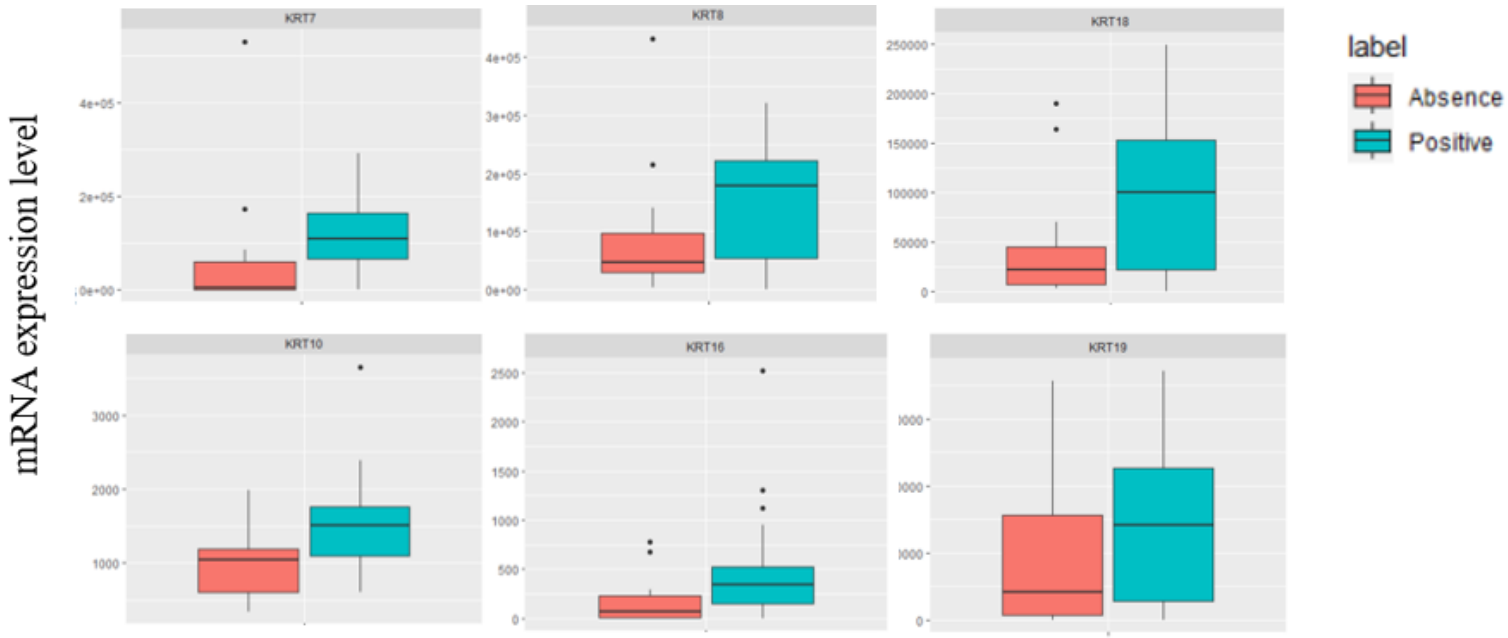

Figure 13. Comparison of KRT family gene mRNA expression in two type cell lines

(PD-L1 Positive cell line versus Absence of PD-L1 cell line)

KRT family genes are associated with PD-L1 expression in lung cancer. PD-L1 expression positive cell lines have different KRT family gene expression than absence of PD-L1 cell lines.

We showed a comparative enhanced volcano plot according to the presence or absence of PD-L1 expression in 29 cell lines (Fig 12). When comparing the genes between the PD-L1 positive cell line and the absence of PD-L1 expression cell line, the gene expressed in PD-L1 expression positive cell line through p-heat map and bubble plot were commonly represented. For instance, KRT17 had a high expression level in the PD-L1 positive cell line as shown in the p-heat map, and CLDN11, CLDN16 was also expressed in the PD-L1 positive cell line in the bubble plot. In the case of KRT81, a previous study showed that it was related to recurrence in lung cancer, and emerged as a promising immunohistochemical marker for the identification of squamous cell lung carcinoma[60]. According to the study, the up-regulation of KRT81 is related to PD-L1 expression in lung cancer cell lines. In order to investigate the relationship between the expression of KRT family genes and PD-L1, the mRNA of expressed KRT family genes was compared. The result shows KRT7, KRT8, KRT10, KRT16, KRT18, and KRT19 among the KRT family were more expressed in the PD-L1 positive cell lines than in the absence of PD-L1 cell lines (Fig 13). And, among other notable up-regulated genes, KCNJ15 is known to be associated with EMT as a tumor suppressor gene [61]. Through western blotting, the connectivity

of the PD-L1 positive cell line was confirmed as an EMT marker. This result is further supported by the mRNA level. In addition, ABCA4 gene, including ABCC1, and ABCC, showed significant associations with lung cancer susceptibility[50]. The ABCA4 gene also showed a correlation with PD-L1 expression. We described expressed genes in the Enhanced Volcano Plot in detail (Table 7). We represented Tables 8A and 8B genes that summarize only the major genes of the enhanced volcano plot.

3.4 PD-L1 positive cell lines are associated with immune related signaling pathway and immune related gene.

We conducted a SPIA pathway analysis, which calculates the connections of pathways by considering their topology and the expression levels of their gene (Fig 14). For comparison in the two groups of PD-L1 positive cell line and absence of PD-L1 expression, the normal value of TPM was calculated using THE HUMAN PROTEIN ATLAS Downloadable data.
(https://www.proteinatlas.org/about/download)

A log2 fold change was used using lung normal tissue TPM, and it was calculated by adding +1 to the TPM value of each cell line. PD-L1 expression can be modulated by various signals in cancer cells, which exert a critical role in tumorigenesis. It is known that signaling pathways related to PD-L1 expression include PI3K-

AKT pathway, JAK-STAT pathway, MAPK pathway, Hedge-hog pathway, WNT pathway, and NF kappa B pathway[62-65].

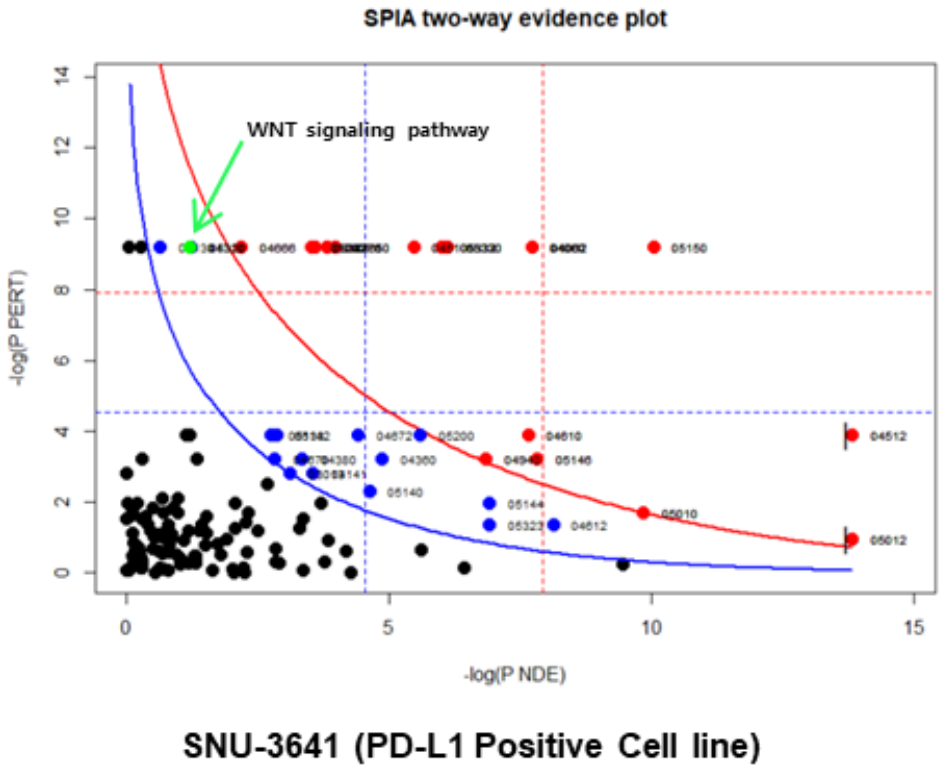

Figure 14. SPIA two evidence plot showing activated signaling pathway of PD-L1 /PD-1 related various signaling pathways in PD-L1 positive cell lines. Within the SPIA evidence plot, a single point represents each pathway and the coordinates are the log of overrepresentation p-value (pNDE) and the p-value from perturbations (pPERT). The oblique lines in the plot show the significance regions based on the combined evidence. The pathways at the right of the red curve are significant after Bonferroni correction of the global p-values, obtained by combining the pPERT and pNDE using the normal inversion method. The pathways at the right of the blue curve line are significant

after a FDR correction of the global p-values.

Table 9A. Significant pathway of PD-L1 positive cell lines (pGFWER<0.05)

Cell Name	Name	Pathway ID	pGFWER	Status
SNU-3641	Wnt signaling pathway	4310	0.047231	Activated

Table 9B. Significant pathway of absence of PD-L1 cell lines (pGFWER<0.05)

Cell Name	Pathway	Pathway ID	pGFWER	Status
SNU-3422	NF-kappa B signaling pathway	4064	6.07E-05	Inhibited
SNU-3422	MAPK signaling pathway	4010	0.012863	Inhibited
SNU-3652	NF-kappa B signaling pathway	4064	2.04E-04	Inhibited
SNU-3652	MAPK signaling pathway	4010	0.005562	Inhibited
SNU-3652B	NF-kappa B signaling pathway	4064	1.28E-05	Inhibited
SNU-3652B	MAPK signaling pathway	4010	0.031167	Inhibited
SNU-3652S	NF-kappa B signaling pathway	4064	1.97E-05	Inhibited
SNU-3652S	MAPK signaling pathway	4010	0.022021	Inhibited

Cell lines containing PD-L1 expression showed expressions in WNT signaling pathway. In particular, we were able to confirm activation in SNU-3641(Table 9A). In parallel with previous findings, all WNT pathways were inhibited in the absence of PD-L1 expression cell line. In the future immunotherapy in lung cancer patient with PD-L1 expression positive, it is expected that WNT signaling pathway will be effective to select as the target drug of PD-L1 expression. It is reported that NF-Kappa signaling pathway is activated in the PD-L1 expression cell line[66, 67]. In the 29 lung cancer cell line established, SPIA analysis represent that not only NF-Kappa signaling pathway (SNU-3422, SNU-3652, SNU-3652B, SNU-3652S) but also MAPK signaling pathway (SNU-3422, SNU-3652, SNU-3652B, SNU-3652S) also was inhibited in the absence of PD-L1 cell line (Table 9B). Although, the NF-Kappa signaling pathway and the MAPK signaling pathway were not activated in the PD-L1 positive cell line, they were inhibited in the absence of PD-L1 cell line in the opposite meaning. These patterns of pathway activity were also observed in GSEA analysis. We performed GSEA analysis and SPIA two way evidence plot to find out the signaling pathways that are related to expression level of PD-L1 at the mRNA level.

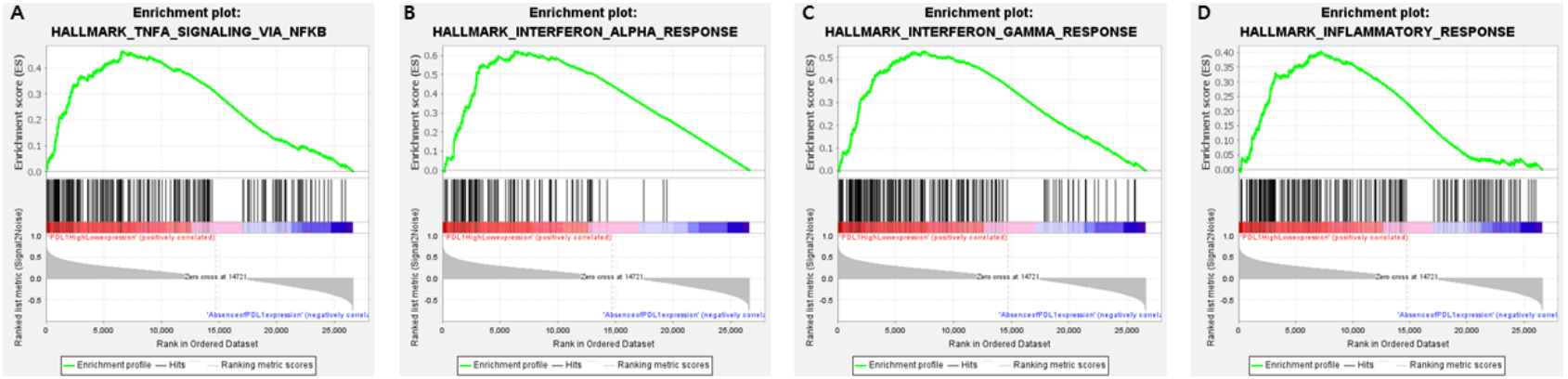

Figure 15A. GSEA analysis in PD-L1 positive cell lines versus Absence of PD-L1 cell lines

Enrichment plots present (A) TNF pathway, (B) Interferon alpha pathway, (C) Interferon gamma pathway, and (D) Inflammatory pathway. GSEA analysis revealed that immune related pathway is upregulated in PD-L1 positive groups compared to the Absence of PD-L1 cell lines. The statistical settings for GSEA analysis is as follows (Number of permutations=1000, Permutation type=phenotype, Chip platform=MSigDB.v.7.4.chip, Enrichment statistic=weighted, Max size: exclude larger sets=500, Min size: exclude smaller sets=15).

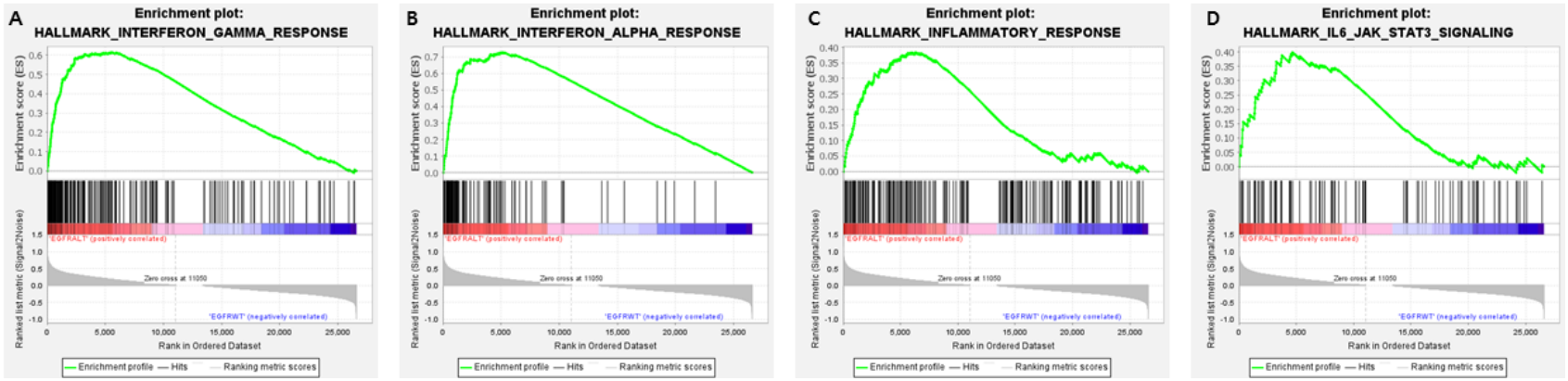

Figure 15 B. GSEA analysis in EGFR-mutant type cell lines versus EGFR- wild type cell lines

Enriched signaling pathways in PD-L1 positive cell lines showed similarly enriched immune related signaling pathways in EGFR mutant cell lines.

Enrichment plots present (A) TNF pathway, (B) Interferon alpha pathway, (C) Interferon gamma pathway, and (D) Inflammatory pathway

Through GSEA analysis was performed in all 29 cell lines, we discover that 19 PD-L1 positive expression cell lines was activated in immune related signaling pathway such as inflammatory response, TNF alpha signaling via NFKB, interferon gamma and Interferon alpha pathway signaling pathway (Fig 15 A). Thus, as shown in Fig 14A, PD-L1 expression positive cell lines were significantly correlated with immune related signaling pathway [68]. The absence of PD-L1 expression cell lines didn't have significant signaling pathway. The GSEA analysis revealed that immune related pathway such as interferon alpha, interferon gamma pathway signaling pathway is up-regulated in PD-L1 positive cell lines compared to absence of PD-L1 expression cell lines. Then, when GSEA analysis was performed according to the EGFR mutation type, the signaling pathways enriched in the PD-L1 positive cell line came out the same (Fig 15 B). As mentioned above, it shows the relationship between PD-L1 expressed cell lines and EGFR mutation.

Supervised cluster analysis and gene set enrichment analysis between the PD-L1 positive and absence of PD-L1 in lung cancer cell lines revealed a correlation between PD-L1 expression and signaling pathways/genes related to cell activation/inhibition. These results indicate that immune related signaling is the

dominant downstream signal. The dominant downstream signals responsible for PD-L1 positive lung cancer cell lines. Furthermore, immune related signaling pathway may control the expression of PD-L1 and several genes related to 29 established lung cancer cell lines. Our findings suggest that immune related signaling pathway is crucial for determining PD-L1 expression, which contributes to the targeted therapies against lung cancers. To address the potential effect of tumor mesenchymalization on the expression of PD-L1, 29 lung cancer cell lines were assessed for the expression of E-cadherin, Vimentin, snail, slug and ZO-1. EMT has been identified in many different tumor type, including lung cancer[69, 70]. PD-L1 has been shown to induce epithelial to mesenchymal transition (EMT) [71]. Cell surface expression of PD-L1 in 29 cell lines was quantified by western blot. We demonstrated that PD-L1 expression classification of lung cancer is categorized as it presents 2 expression levels. We also assessed the expression of EMT in PD-L1 high expression cell lines, PD-L1 low expression cell lines and absence of PD-L1 expression cell lines. Among them, 19 cell line harboring PD-L1 expression of 29 cell lines was not significantly related to EMT.

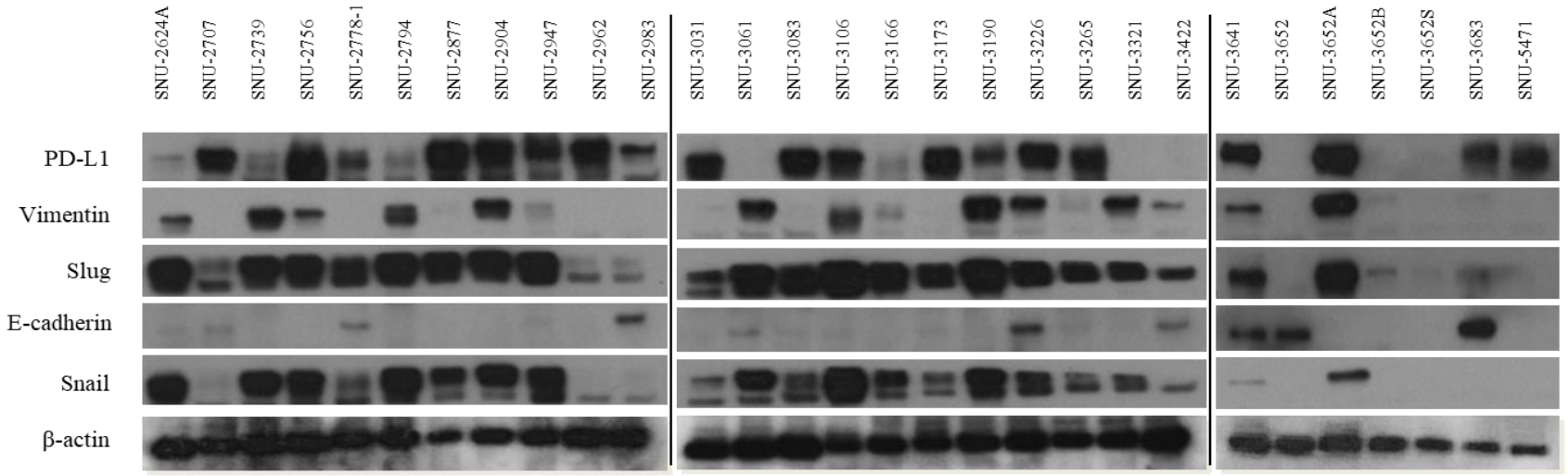

Figure 16. Western blotting analysis in 29 patient derived lung cancer cell lines.

The protein expression of PD-L1, Vimentin, E-cadherin, Slug and Snail in 29 cell lines was determined by western blotting using the antibodies. The protein level of EMT marker was augmented as well on PD-L1 positive cell lines. The increased PD-L1 expression barely affected the expression of Vimentin, E-cadherin, Slug and Snail.

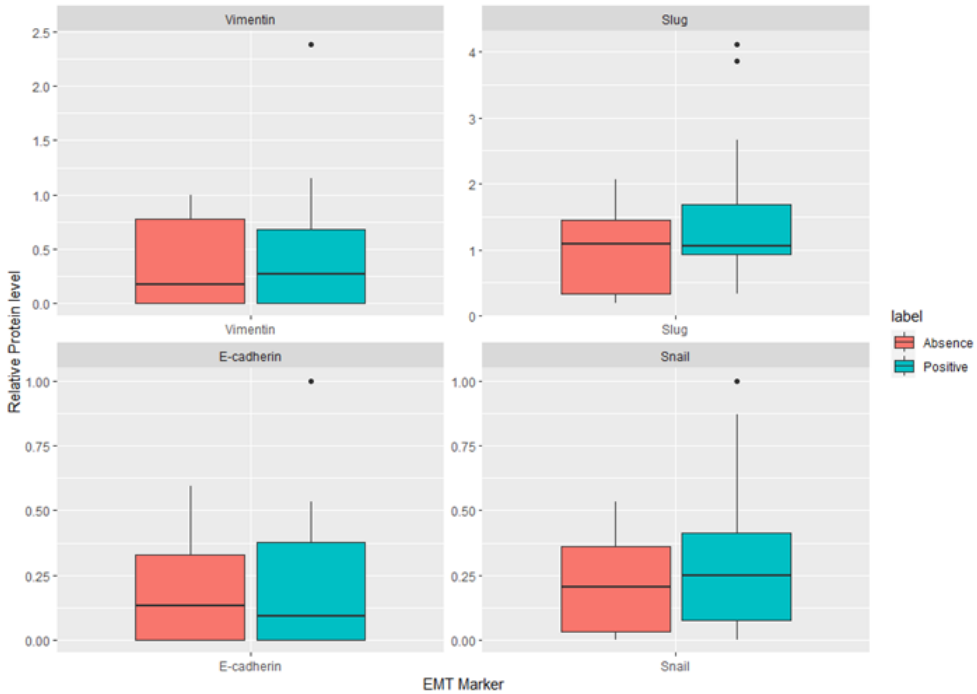

Figure 17. Comparison of the average expression of EMT markers in the cell lines of the two groups)

The protein level of EMT marker was augmented as well on both lines. The increased PD-L1 expression barely affected the expression of Vimentin, E-cadherin, Slug and Snail. Especially, Slug, an EMT marker, increased in expression in the PD-L1 positive cell line, and decrease in E-cadherin. Other EMT markers, such as Vimentin or Snail, showed no significant differences between groups.

As shown in Fig.16, the expression levels of PD-L1 and EMT expression level is not correlated to be up-regulated as the expression of PD-L1 increase. For instance, the high expression of PD-L1 in 29 cell lines were markedly increased to EMT maker such as Vimentin, Slug in especially SNU-2756, SNU-3641 and SNU-3652A in common. Conversely, in the cell lines with negative PD-L1 expression, the expression of the EMT marker, such as E-cadherin and Snail, was down-regulated [72]. Although a maker that showed that PD-L1 expression was not related to EMT, in principle EMT marker were found to be correlated in cells with PD-L1 expression. These results revealed that the signaling pathways of related EMT marker served an important role in inducing the expression of PD-L1, but not absence of PD-L1 expression cell lines. In addition, we explored a correlation the influence of PD-L1 expression with immune related signaling pathway. As mentioned, high expression of PD-L1 was associated with the presence of EGFR mutations in NSCLC and was an independent negative prognostic factor for this disease[73]. Also, it is reported that EGFR is involved in the regulation of PD-L1 expression and cell proliferation via the IL-6/JAK/STAT3 signaling pathway in NSCLC[64]. Some studies showed that the JAK/STAT3 pathway participates in cancer cell survival,

proliferation and progression by regulating multiple processes, such as epithelial-mesenchymal transition(EMT), Which is required for tumor metastasis, and molecular signals that control other cancer hall markers [74]. According to previously analyzed GSEA analysis, we revealed the JAK-STAT3 pathway as well as the immune related signaling pathway. EMT increases the capacities of migration and invasion in lung cancer, which resulted in up-regulation of PD-L1 expression via a mechanism that is dependent on NF-κB activation [75]. Through RNA seq analysis (GSEA, SPIA two way evidence analysis), it was found that NF-κB among the signaling pathways related to expression of PD-L1 was related only to cell lines expressing PD-L1. Western blotting results showed that EMT induction led to the up-regulation of PD-L1 but not Vimentin, Snail in this PD-L1 expression positive cell lines. These results indicated that cytokine-mediated EMT promoted the expression of PD-L1. As a result of confirming various EMT markers and comparing their values quantitatively, slug and E-cadherin showed some significant results, but no significant results were shown between the other marker groups (Fig 17). The characteristics of cell lines with or without PD-L1 expression could be explained in relation to EMT, and as a result, cell lines with or without EGFR mutation could show similar results.

Interestingly, significant results of EMT according to EGFR status could also give an important meaning of the role of established cell lines. In summary, all our data contribute to a novel finding that PD-L1 expression is at least partly regulated by NF-κB signaling when induced EMT. EMT has become a focus of research due to its vital role in tumor progress throughout the body and complexity in a variety of immune processes. In this study, we explored the influence of EMT in immune-related process through the study of PD-L1 expression regulation in patient derived lung cancer cell line.

3.5 Lung cancer cell line harboring PD-L1 expression has sensitivity to EGFR-TKI but PD-L1 expression cell lines with EGFR mutation shows resistance to EGFR-TKI. (Relationship between anti-cancer drug response and PD-L1 expression)

In order to examine the reactivity of 29 lung cancer cell lines for 18 anti-drugs, we investigated 15 different drugs responses with action of various mechanisms.

The 29 lung cancer cell lines derived from patient were successfully screened in experimental triplicate, and the concentration of each 18 drug was measured with an appropriate AUC value. The class of the agent and mechanism of drug along with the maximum concentration are listed in the Table 10.

Table 10. Chemotherapeutic agents and targeted agents for lung cancer treated in this study

Drug	Class of the anti-cancer agent	Mechanism of action	Max conc.(uM)
Erlotinib Hydrochloride	Epidermal growth factor receptor- tyrosine kinase inhibitor	EGFR ex19 deletion, ex21(L858R) substitution mutation inhibition	100
Crizotinib	Anaplastic lymphoma kinase(ALK) inhibitor	ALK(anaplastic lymphoma kinase) and ROS1 inhibition	10
Everolimus	protein kinase inhibitor	rapamycin(mTOR) inhibition	100
Paclitaxel	Anti-microtubule Agents(cytoskeletal drug)	Anti-microtubule agent	50
Trametinib	MEK inhibitor	MEK1 and MEK2 inhibition	50
ICG-001	beta catenin inhibitor	a co-activator of Wnt/beta-catenin-mediated transcription inhibition	100
Alectinib	Anaplastic lymphoma kinase(ALK) inhibitor	ALK(anaplastic lymphoma kinase) inhibition	10
Ceritinib	Anaplastic lymphoma kinase(ALK) inhibitor	ALK(anaplastic lymphoma kinase) inhibition	10
Dacomitinib	Epidermal growth factor receptor- tyrosine kinase inhibitor	EGFR mutation inhibition	10
Gefitinib	Epidermal growth factor receptor- tyrosine kinase inhibitor	EGFR mutation inhibition	100
Afatinib	Epidermal growth factor receptor- tyrosine kinase inhibitor	Tyrosine kinase inhibiton	50
Regorafenib	multi-kinase inhibitor	EGFR, HER2 inhibition	100
Buparlisib(BKM120)	PI3K inhibitor	PI3K inhibition	100
Apitolisib	PI3K/ mTOR inhibition	PI3K/ mTOR inhibition	50
Cyclopamine	hedgehog signaling inhibitor	hedgehog signaling inhibition	50
Vistusertib(AZD2014)	mTOR inhibitor	mTOR inhibition	5
MK-5108	Aurora A inhibitor	Aurora kinase inhibition	100
Olaparib	PARP inhibitor	PARP inhibition	50

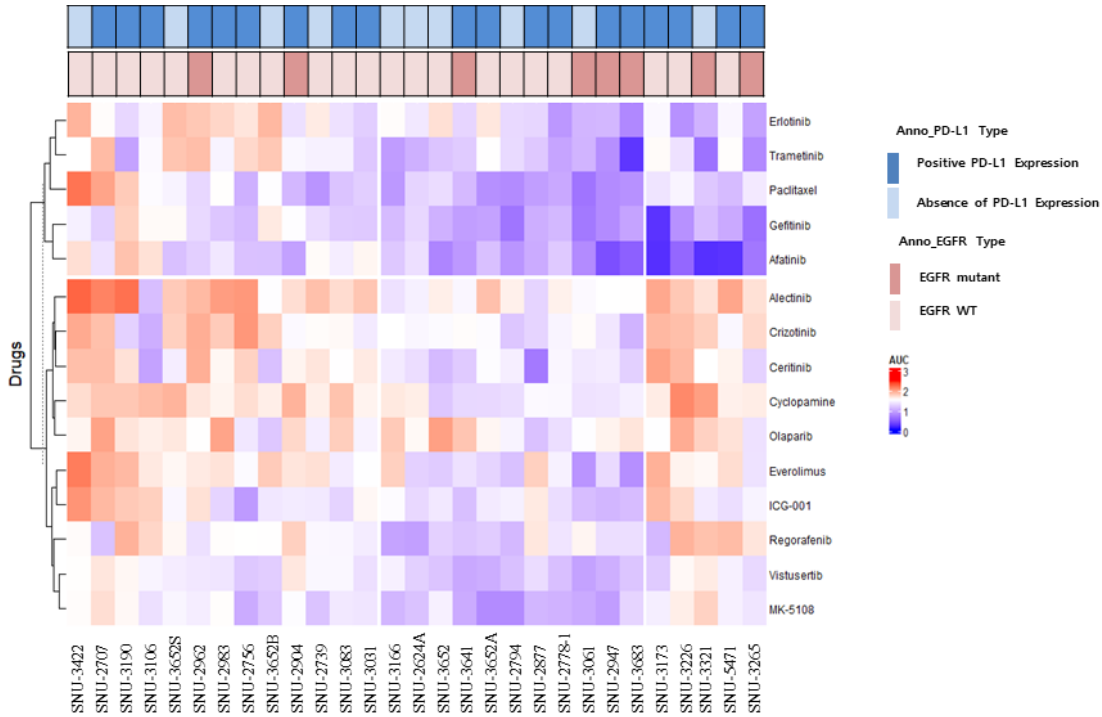

Figure 18. Drug response heatmap caused by molecular target diversity

Patient derived 29 lung cancer cell lines. See also Table 11.

Heatmap of patient-derived lung cancer cell lines exhibited heterogeneous distribution of 15 compounds according to their molecular characteristics. The names of the compounds are provided on the right. The cell lines and the drugs were k-means clustered based on the AUC values across the drug panel. The mutational status of PD-L1, EGFR are specified above the heatmap.

Table 11-A. AUC of 29 established lung cancer cell lines

Cell Name	Erlotinib	Crizotinib	Everolimus	Paclitaxel	Trametinib	ICG-001	Alectinib
SNU-2624A	1.43	1.51	1.28	1.30	1.07	1.33	1.47
SNU-2707	1.58	1.87	1.95	2.01	1.90	1.91	2.16
SNU-2739	1.67	1.59	1.72	0.89	1.49	1.43	1.87
SNU-2756	1.70	2.06	1.54	1.07	1.76	0.93	2.06
SNU-2778-1	0.91	1.50	1.48	1.02	1.00	1.36	1.64
SNU-2794	1.33	1.21	1.18	0.84	1.34	1.48	1.64
SNU-2877	1.31	1.30	1.80	0.96	1.23	1.67	1.30
SNU-2904	1.38	1.54	1.70	1.11	1.52	1.44	1.73
SNU-2947	1.12	1.42	1.34	0.84	0.86	1.11	1.57
SNU-2962	1.84	1.96	1.67	1.32	1.88	1.72	1.91
SNU-2983	1.77	1.82	1.71	1.58	1.49	1.30	2.04
SNU-3031	1.26	1.44	1.56	1.26	1.22	1.61	1.85
SNU-3061	1.10	1.59	0.89	0.69	1.10	1.16	1.53
SNU-3083	1.38	1.60	1.45	1.19	1.48	1.29	1.74
SNU-3106	1.50	1.06	1.68	1.54	1.53	1.80	1.16
SNU-3166	1.58	1.57	1.80	0.91	0.94	1.54	1.38
SNU-3173	1.53	1.91	1.95	1.40	1.59	1.90	1.99
SNU-3190	1.32	1.28	1.91	1.82	0.98	1.83	2.23
SNU-3226	0.87	1.88	1.63	1.50	1.39	1.76	1.84
SNU-3265	0.98	1.75	1.39	1.44	0.83	1.49	1.72
SNU-3321	1.08	1.80	1.60	1.22	0.67	1.44	1.71
SNU-3422	1.93	1.98	2.18	2.22	1.57	2.08	2.28
SNU-3641	1.31	1.58	1.38	1.13	1.24	1.17	1.51
SNU-3652	1.72	1.54	1.24	1.35	1.20	1.43	1.65
SNU-3652A	1.69	1.55	1.30	0.86	1.57	1.44	1.87
SNU-3652B	1.91	1.80	1.82	1.56	1.84	1.42	1.55
SNU-3652S	1.89	1.80	1.61	1.49	1.85	1.51	1.83
SNU-3683	0.81	1.09	0.86	0.87	0.30	1.15	1.57
SNU-5471	1.51	1.52	1.74	1.14	1.58	1.36	2.00

Continued

Table 11-B. AUC of 29 established lung cancer cell lines

Cell Name	Ceritinib	Dacomitinib	Gefitinib	Afatinib	Regorafenib	Cyclopamine	Vistusertib
SNU-2624A	1.36	1.87	1.32	1.37	0.96	1.66	1.27
SNU-2707	1.88	1.98	1.27	1.38	1.19	1.84	1.69
SNU-2739	1.70	2.10	1.35	1.59	1.52	1.65	1.52
SNU-2756	1.79	2.07	1.14	1.18	1.57	1.73	1.23
SNU-2778-1	1.51	1.69	1.12	1.23	1.40	1.53	1.17
SNU-2794	1.47	1.70	0.69	0.90	1.12	1.36	1.17
SNU-2877	0.72	1.41	1.06	1.04	1.69	1.53	1.35
SNU-2904	1.62	1.51	1.56	0.97	1.80	1.94	1.68
SNU-2947	1.45	1.72	0.84	0.45	1.37	1.40	1.07
SNU-2962	1.95	1.94	1.33	1.27	1.38	1.69	1.42
SNU-2983	1.61	2.02	1.21	1.42	1.58	1.63	1.40
SNU-3031	1.67	1.97	1.23	1.62	1.45	1.63	1.37
SNU-3061	1.42	1.76	0.72	0.83	1.62	1.39	0.98
SNU-3083	1.57	1.74	1.24	1.46	1.52	1.86	1.52
SNU-3106	0.98	1.79	1.59	1.72	1.77	1.89	1.50
SNU-3166	1.44	2.04	1.14	1.22	0.99	1.68	1.48
SNU-3173	2.01	1.27	0.29	0.27	1.12	1.66	1.24
SNU-3190	1.71	2.04	1.80	1.86	1.94	1.84	1.61
SNU-3226	1.91	0.92	0.87	0.60	1.94	2.13	1.60
SNU-3265	1.29	1.47	0.66	0.71	1.70	1.66	1.33
SNU-3321	1.57	0.68	1.16	0.27	1.87	2.03	1.67
SNU-3422	1.88	2.30	1.47	1.72	1.59	1.74	1.58
SNU-3641	1.24	1.64	0.95	0.92	1.21	1.32	1.02
SNU-3652	1.14	2.08	1.08	0.79	1.28	1.22	1.20
SNU-3652A	1.55	1.96	0.98	1.18	1.25	1.33	1.05
SNU-3652B	1.17	1.46	1.67	1.18	1.57	1.65	1.26
SNU-3652S	1.46	1.85	1.59	1.17	1.61	1.93	1.45
SNU-3683	1.27	1.28	0.99	0.56	1.36	1.47	1.21
SNU-5471	1.63	2.02	1.03	0.29	1.90	1.64	1.48

Continued

Table 11-C. AUC of 29 established lung cancer cell lines

Cell Name	MK-5108	Olaparib
SNU-2624A	1.19	1.60
SNU-2707	1.73	2.01
SNU-2739	1.19	1.46
SNU-2756	1.04	1.42
SNU-2778-1	1.09	1.37
SNU-2794	0.84	1.49
SNU-2877	1.11	1.18
SNU-2904	1.55	1.76
SNU-2947	0.95	1.63
SNU-2962	1.46	1.57
SNU-2983	1.58	2.01
SNU-3031	1.41	1.49
SNU-3061	1.03	1.55
SNU-3083	1.43	1.81
SNU-3106	1.38	1.65
SNU-3166	1.11	1.83
SNU-3173	1.46	1.56
SNU-3190	1.60	1.70
SNU-3226	1.65	1.97
SNU-3265	1.41	1.38
SNU-3321	1.79	1.79
SNU-3422	1.58	1.62
SNU-3641	1.04	1.85
SNU-3652	1.48	2.02
SNU-3652A	0.83	1.61
SNU-3652B	1.22	1.22
SNU-3652S	1.51	1.68
SNU-3683	1.30	1.70
SNU-5471	1.48	1.71

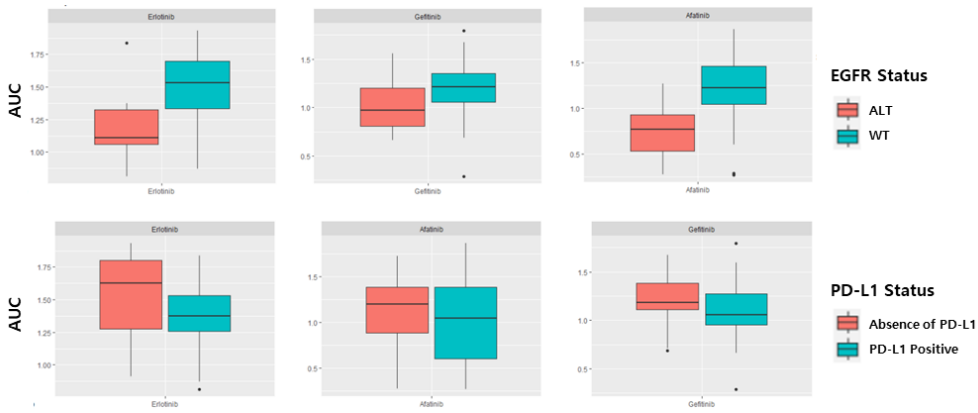

Figure 19. Box plot of four EGFR-targeting drugs

Cell lines between the PD-L1 type groups show EGFR targeting drug response.

EGFR ALT and WT are clearly more sensitive in cell lines with EGFR mutations than cell lines divided to PD-L1 expression type, suggesting more resistance to three drugs (Erlotinib, Afatinib, Gefitinib).

As a first validation, drug reactivity was compared with the annotation of PD-L1 status through k-means clustering of AUC values. The cell line with PD-L1 expression showed a tendency toward resistance to drug reaction as a whole compared to the cell line without the expression. When looking at the overall drug heat map, the cell lines that had a high response to the EGFR target drug were total EGFR mutant type cell lines, Cell lines of the EGFR mutant type except for one cell line showed a high status of PD-L1 expression (Fig 18). However, one of the drugs targeting EGFR TKI (Dacomitinib) showed resistance reactivity in most cell lines regardless of EGFR type. The only cell lines with good responsiveness to Dacomitinib were SNU-3226 and SNU-3321 patients, they showed PD-L1 expression, and consisted of EGFR wild type. With this result, the cell lines with PD-L1 high expression have EGFR mutations, and the cell lines with PD-L1 high expression and EGFR mutations, except for Dacomitinib, have high sensitivity to EGFR target drug. Finally, the results of studies on the association between TP53 mutation and PD-L1 expression showed a trend in our results, but did not show a significant association with drugs. In general, there was no significant correlation between TP53 mutation and PD-L1 expression cell

lines with ALK mutation (SNU-3321) and EML4-ALK fusion gene (SNU-2877, SNU-3106, SNU-3166), except for SNU-3321, all three cell lines with fusion gene has good reactivity to ALK inhibitors (Alectinib, Crizotinib, Ceritinib). They were all confirmed to have PD-L1 expression. There are studies that do not have both fusion gene and EGFR mutations in lung cancer [14]. Among the 29 cell lines we established, the cell line containing the fusion gene did not have both the fusion gene and the EGFR mutation, and it was additionally confirmed that all of the cell lines had PD-L1 expression.

We also detected that cell lines between the PD-L1 type groups show EGFR targeting drug response. EGFR ALT and WT are clearly more sensitive in cell lines with EGFR mutations than cell lines divided to PD-L1 expression type, suggesting more resistance to three drugs (Erlotinib, Afatinib, Gefitinib) (Fig 19). In addition, we compared to the drug correlation of each cell line in the PD-L1 expression status through performance analysis for visualizing a correlation matrix. The drug correlation of each cell line in the PD-L1 expression positive group is shown in Fig20. The Western blotting analysis shows that the AUC value of the drug matches the PD-L1 expression level.

As a result, it was confirmed that the higher the expression of PD-

L1, the higher the significance of the drug response. No significant result was confirmed in the absence of PD-L1 cell line group.

3.6 Multiomics in either PD-L1 positive or absence cell lines provide treatment direction of lung cancer instead of tumor proportion score or tumor mutation burden.

We observed to association EGFR mutant type and the TPM value of EGFR at the mRNA level with PD-L1 TPM value for predicting drug resistance. Among the drugs targeting TKI, Erlotinib and Afatinib were confirmed to have a negative relationship with PD-L1 (CD274) at the mRNA level in EGFR mutant type cell lines. When increasing EGFR TPM value, the PD-L1 TPM value increases, and in addition, the drug resistance also increases with reference to 1(AUC value) (Fig 20).

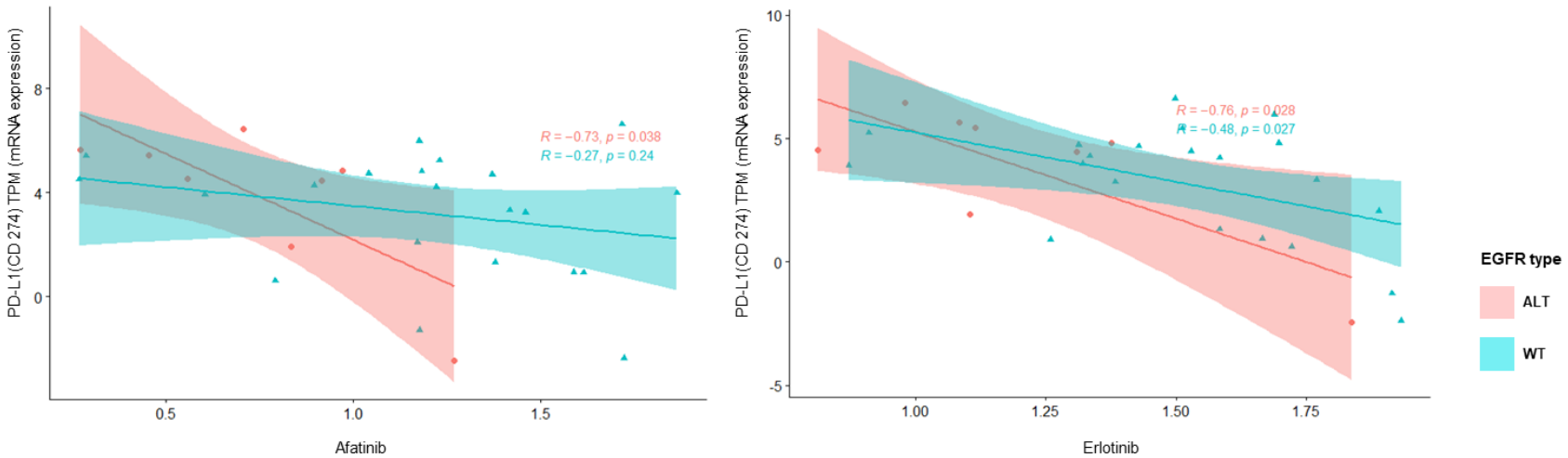

Figure 20. The direct correlation between AUC of drug (Afatinib, Erlotinib) and PD-L1 mRNA expression (TPM) is inversed by the mutational status of EGFR.

Cell lines with EGFR mutant type are separated from the correlation graph and colored in red.

The Pearson correlation coefficient (R) with p values is re-calculated according to the mutational status of EGFR.

The two results show that Erlotinib and Afatinib were the ones that had increasingly the higher the TPM value of EGFR in PD-L1 expression positive cell lines and that reconfirmed drug resistant characteristics. In the future, for lung cancer patients with PD-L1 expression and EGFR mutation, the use of a TKI drug other than Erlotinib or Afatinib would be expected to obtain better results in the treatment of lung cancer patients with PD-L1 expression. To indicate a resumption of these results, we identified EGFR-mutant type cell lines were exceedingly resistant to EGFR targeting drugs (Afatinib, Erlotinib). These EGFR-mutant type cell lines showed even better correlate to PD-L1 expression at the mRNA level than EGFR-wild type cell lines. In conclusion, our work emphasizes the significance of integrating the mutational status with transcriptomic subtypes using statistical models such as Pearson correlation coefficient and MANOVA for the accurate assessment of drug responses in lung cancer patients with MPE.

3.7 CNV patterns of EGFR TKI naive/ resistance cell line, PD-L1 positive and negative cell lines.

Copy number variation analysis were used to investigate chromosomal aberrations within lung cancer patient derived cell lines. In order to assess whether genetic changes underlie the different phenotypes of PD-L1 positive and negative cell lines, we employed low-coverage whole exome sequencing to establish copy number alterations. Out of 15 cell lines, 10 cell lines were normalized to each cell line-matched whole blood reference gDNA. Copy number variation(CNV) profiles demonstrated that PD-L1 positive cell lines showed more extensive chromosomal abnormalities and intense CNV changes at specific loci (Fig 21).

Our copy number variation (CNV) analysis revealed that gain at chromosomes 5, 7 and X and loss at chromosome 4, 8.

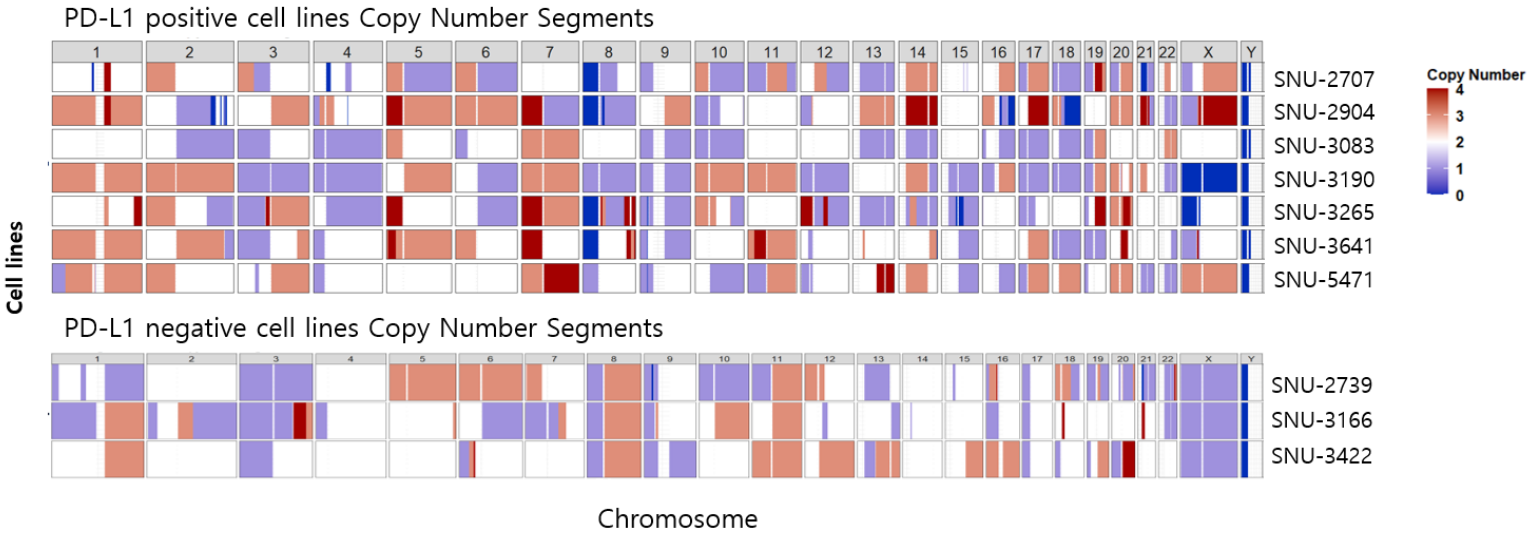

Figure 21 A. Copy number variation analysis performed by whole exome sequencing.

Heat map of copy number variation segments mean value in PD-L1 positive/negative cell lines.

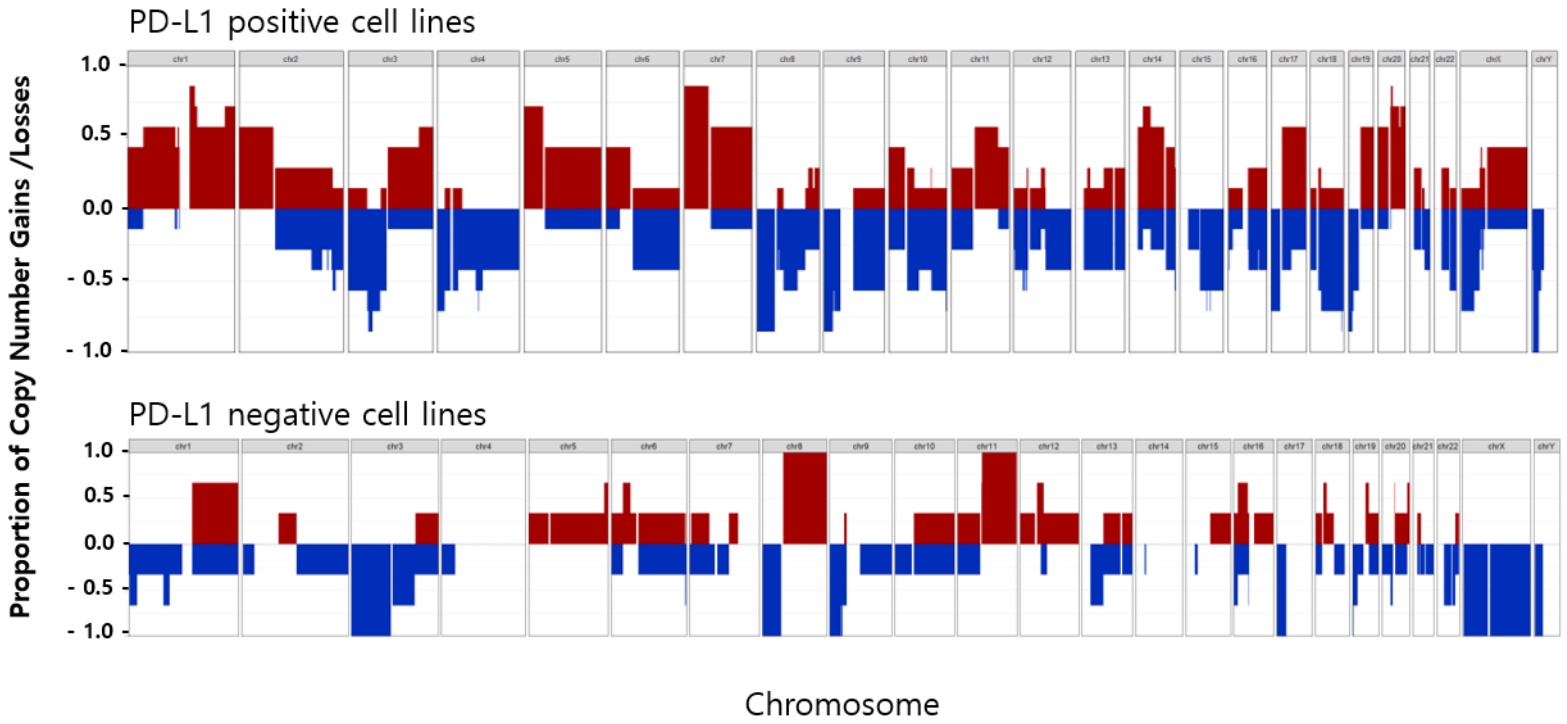

Figure 21 B. Copy number variation analysis performed by whole exome sequencing.

Proportion of copy number gains/ losses in PD-L1 positive/negative cell lines.

3.8 immune cell types signature analysis to 29 established lung cancer cell lines.

To characterize intratumoral immune cell subtype states, we scored immune cell types to TPM expression signatures and used normalized lung tissues TPM to identify of immune signature sets. We selected 14 immune cell types. Analysis of lung cancer derived cell lines revealed the gene sets related to T cells, CD8 T cells, exhausted CD8 T cells, Treg cells, B cells, macrophages, CD45, mast cells, DCs, NK cells, CD56dim NK cells, cytotoxic cells, Th1 cells, and neutrophils between the PD-L1 positive group and PD-L1 negative group (Fig. 22). Compared with the PD-L1 negative group, a higher level of immune related genes was detected in PD-L1 positive group (Fig 22).

In PD-L1 positive cell lines corresponding to SNU-3226, SNU-2707, and SNU-2756, the expression of genes related to immune cell markers was higher than that of PD-L1 negative cell lines. In particular, TBX21 related to Th1 cells was high in SNU-3226, and the expression of S100A12 gene related to neutrophils was high in SNU-2707. In addition, in SNU-2756, the expression of SH2D1A, a gene related to the T cell type, was high, and the expression of the IL21R gene related to NK 56dim cells has high expression.

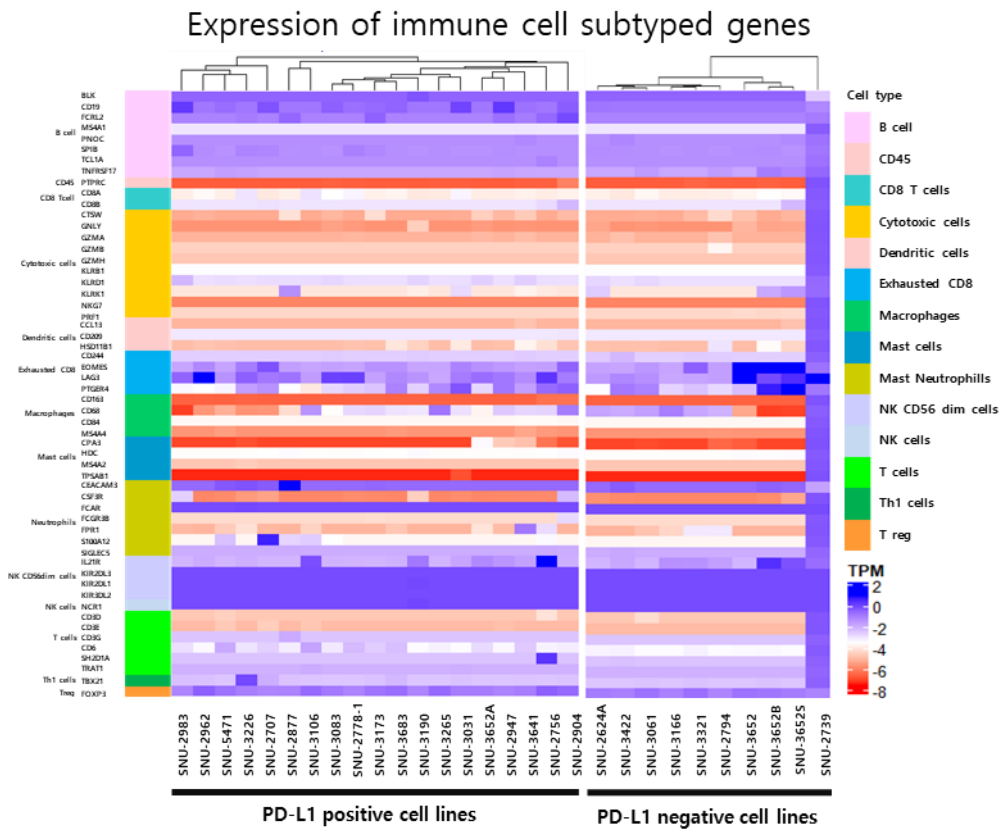

Figure 22. Expression of immune cell subtyped genes in 29 lung cancer patient derived cell lines

Association of Immune cell subtyped genes and clinicopathological characteristics with PD-L1 status in our 29 established cell lines.

Discussion

Tumor cells can evade the immune response through up-regulation of PD-L1, with diverse human malignant tumors showing elevated levels of PD-L1 protein, including NSCLC. Blockade of PD-1 and PD-L1 interaction using monoclonal antibodies has effective clinical responses in patient with various tumor types. Although there are some differences between the expression of PD-L1 in the cell line and the expression in the patient, the presence or absence of PD-L1 expression in the cell line in the in vitro model may be a prediction through the integrated analysis of genomic analysis and transcriptomic analysis to determine the accessibility of clinical treatment. Thus, 29 cell lines were characterized, and interesting results could be presented as a result. Characterization of our 29 lung cancer cell lines began with an examination of cell surface expression of PD-L1 through protein level. We found that all of the cell lines exhibited different level of protein, which verified correlation with both mRNA analysis and value of protein level. In the present study, the first specifically conducted in a cohort of molecularly selected NSCLC, we showed that expression of PD-L1 differs according to protein level and mutation trait. According to genomic

analysis, the types and amounts of mutations in cell lines with PD-L1 expression were higher than those in cell lines without PD-L1 expression, and so was the occurrence of fusion gene. Among them, in the case of EGFR mutation and TP53 mutation, we know that a lot of genomic alteration occurs in tumors with PD-L1 expression, and our cell lines also showed the same pattern. To strengthen the connectivity of genomic analysis and transcriptomic analysis, when looking at the correlation at the DNA level and RNA level, its connectivity could be sufficiently shown, and it could also be shown in the immune related signaling pathway at the protein level and transcriptomic level. After showing the correlation on the central dogma significantly, we were curious about the context of drug reactivity, and a study related to drug response was also conducted. Cell lines without PD-L1 expression did not have significantly sensitive drug reactivity compared to cell lines with high PD-L1 expression. As the previous study results show, if grafted to clinical immunotherapy, patients with PD-L1 expression may be more responsive to the drug, in the case of patients without PD-L1 expression knowing about drug-resistant drugs for approaching treatment would be one of the results. And, the enhancement of immune related genes and immune related signaling pathways shown in transcriptome analysis is likely to

bring boosting findings of immunotherapy as an approach to clinical treatment. Among the many advantages of combined immunotherapy during chemotherapy, it has a long survival period and good therapeutic efficacy compared to other chemotherapy. By scoring PD-L1, not only patients with PD-L1 expression receive immunotherapy, but also in patients with PD-L1 expression but low TPS or without PD-L1 expression, an in vitro model was used to provide appropriate treatment access. The analysis of transcriptomic analysis in human lung cancer cell lines will also be an important outcome among treatment trials to approach the results of signaling pathway analysis involving immune related genes, which are also helpful for them. Cell lines harboring EGFR mutations had higher levels of PD-L1 expression when compared to the EGFR wild-type population and presence of TP53 mutations or other kinds of mutations were associated with increased PD-L1 protein levels. Another interesting finding was the evidence that levels of PD-L1 were higher in cell lines with TP53 mutation. Although the association was not statistically significant and although we were not able to evaluate whether a different immune check point expression impacted on efficacy of agent, PD-L1 expression levels were higher in positive TP53 mutation cell lines when compared with cell lines harboring wild-type TP53.

Importantly, EGFR/TP53 mutation positive cell lines affected transcriptional level, which were distinctive to NSCLC cell lines with other driver mutations such as BRACA1 or fusion genes. [76] The association between PD-L1 expression and EGFR/TP53 mutations may have clinical relevance to predict the efficacy of PD-L1 immune checkpoints inhibitors. In conclusion, our study showed that immune check point protein were differently expressed in oncogene addicted NSCLC potentially modulating sensitivity to targeted agents. Our findings represent the rationale to choose the integration of highly prevalent genetic and transcriptional characteristics of patient derived lung cancer cell lines to various drug responses. In summary, the present study demonstrated that it can be predicted in advance from a patient-derived cell line instead of what is needed to diagnose an actual patient. In other words, it can be predicted in advance from a patient-derived cell line instead of what is needed to diagnose an actual patient. EGFR mutation can induce the expression of PD-L1 via the IL6/JAK-STAT3 signaling pathway in EGFR-mutant NSCLC cell lines, which can be inhibited by EGFR-TKIS. In addition, down-regulation of PD-L1 was associated with inhibited proliferation and enhanced apoptosis of EGFR-mutant NSCLC cells. Furthermore, our and other data suggest that expression of PD-L1

with EGFR TKIs may be a promising therapeutic strategy to extend the overall survive and to delay development of drug resistance. Considering the retrospective nature of our investigation, additional prospective studies are warranted.

Patient derived cell lines help researchers better understand cancer biology and genetics, allowing identification of biomarkers for clinical diagnosis and prediction of drug response.

Some characteristics study to patient-derived cell lines have been used extensively in lung cancer research, while others have recently emerged, such as organoids, and their exploitation remains in infancy. This study has the advantages and limitations specific to each preclinical model and provides an overall summary of these observations. The characteristics study using cell line are easy and fast to culture, making them the first choice for approaching large scale studies, but the lack of heterogeneity and cell to cell interactions within a 2D environment limits the translational potential of findings from cell line studies. However, the study of patient-derived cell lines can be accomplished in as little as in a few weeks and may therefore have the potential to inform clinical decisions during the patient's lifetime. Compared to pre-clinical cell lines study, other study retains tumor heterogeneity and the in-situ tumor microenvironment, but are

more time-consuming and expensive, therefore hampering large-scale experiments.

Nevertheless, among all the patient-derived cancer models, cell lines may best recapitulate tumor biology and microenvironment, especially when using a better option for clinical decision making in precision medicine. While there is no ideal model for lung cancer research, study design must consider all the advantages and disadvantages of different models, as well as the compatibility with different experiments, before selecting one to answer relevant biological questions.

Despite the limitations of the characteristic study, we tried to find noble findings. As a result, we were able to find that KRT family genes were always high on the transcriptome level in PD-L1 positive cell lines. And there were several references[77-81].

Although we could not find clinical results that directly linked KRT family gene, PD-L1 expression, and immunotherapy, the characteristics study the established cell line is considered to be the first initiation report.

Reference

1. Bai, Y., et al., *PD-L1 expression and its effect on clinical outcomes of EGFR-mutant NSCLC patients treated with EGFR-TKIs.* Cancer Biology & Medicine, 2018. **15**(4): p. 434.
2. Albitar, M., et al., *Correlation of MET gene amplification and TP53 mutation with PD-L1 expression in non-small cell lung cancer.* Oncotarget, 2018. **9**(17): p. 13682.
3. Mogi, A. and H. Kuwano, *TP53 mutations in nonsmall cell lung cancer.* Journal of Biomedicine and Biotechnology, 2011. **2011**.
4. Chang, Y.-S., et al., *The Fusion Gene Landscape in Taiwanese Patients with Non-Small Cell Lung Cancer.* Cancers, 2021. **13**(6): p. 1343.
5. Hong, S., et al., *Upregulation of PD-L1 by EML4-ALK fusion protein mediates the immune escape in ALK positive NSCLC: Implication for optional anti-PD-1/PD-L1 immune therapy for ALK-TKIs sensitive and resistant NSCLC patients.* Oncoimmunology, 2016. **5**(3): p. e1094598.
6. Takahashi, T., et al., *Clinicopathologic features of non-small-cell lung cancer with EML4-ALK fusion gene.* Annals of surgical oncology, 2010. **17**: p. 889-897.
7. Koivunen, J.P., et al., *EML4-ALK fusion gene and efficacy of an ALK kinase inhibitor in lung cancer.* Clinical cancer research, 2008. **14**(13): p. 4275-4283.
8. Inamura, K., et al., *EML4-ALK lung cancers are characterized by rare other mutations, a TTF-1 cell lineage, an acinar histology, and young onset.* Modern Pathology, 2009. **22**(4): p. 508-515.
9. Sumimoto, H., et al., *RAS–mitogen-activated protein kinase signal is required for enhanced PD-L1 expression in human lung cancers.* PloS one, 2016. **11**(11): p. e0166626.
10. Tang, Y., et al., *The association between PD-L1 and EGFR status and the prognostic value of PD-L1 in advanced non-small cell lung cancer patients treated with EGFR-TKIs.* Oncotarget, 2015. **6**(16): p. 14209.
11. Yokota, K., et al., *KIF5B/RET fusion gene in surgically-treated adenocarcinoma of the lung.* Oncology reports, 2012. **28**(4): p. 1187-1192.
12. Sun, R., et al., *Long non-coding RNA in drug resistance of non-small cell lung cancer: a mini review.* Frontiers in Pharmacology, 2019. **10**: p. 1457.
13. Yu, P., et al., *Research progress regarding long-chain non-coding RNA in lung cancer: A narrative review.* Journal of Thoracic Disease, 2022. **14**(8): p. 3016.
14. Takahashi, T., et al., *Clinicopathologic features of non-small-cell lung cancer with EML4-ALK fusion gene.* Annals of surgical oncology, 2010. **17**(3): p. 889-897.
15. Mao, R., et al., *Ubiquitin C-terminal hydrolase L1 promotes*

expression of programmed cell death-ligand 1 in non-small-cell lung cancer cells. Cancer science, 2020. **111**(9): p. 3174-3183.
16. Alfaro-Arnedo, E., et al., *IGF1R acts as a cancer-promoting factor in the tumor microenvironment facilitating lung metastasis implantation and progression.* Oncogene, 2022. **41**(28): p. 3625-3639.
17. Wang, Z., et al., *Overexpression of KRT17 promotes proliferation and invasion of non-small cell lung cancer and indicates poor prognosis.* Cancer management and research, 2019. **11**: p. 7485.
18. Xu, C., et al., *Identification of a novel tumor microenvironment prognostic signature for bladder urothelial carcinoma.* Frontiers in Oncology, 2022. **12**.
19. Li, L., et al., *Mesenchymal stem cell transplantation attenuates cardiac fibrosis associated with isoproterenol-induced global heart failure.* Transplant International, 2008. **21**(12): p. 1181-1189.
20. Aoi, W., et al., *A novel myokine, secreted protein acidic and rich in cysteine (SPARC), suppresses colon tumorigenesis via regular exercise.* Gut, 2013. **62**(6): p. 882-889.
21. Salama, E.A., R.E. Adbeltawab, and H.M. El Tayebi, *XIST and TSIX: novel cancer immune biomarkers in PD-L1-overexpressing breast cancer patients.* Frontiers in oncology, 2020. **9**: p. 1459.
22. Rodvold, J.J., N.R. Mahadevan, and M. Zanetti, *Lipocalin 2 in cancer: when good immunity goes bad.* Cancer letters, 2012. **316**(2): p. 132-138.
23. Major, A., P. Porcu, and B.M. Haverkos, *Rational Targets of Therapy in Extranodal NK/T-Cell Lymphoma.* Cancers, 2023. **15**(5): p. 1366.
24. Li, W., et al., *MUC4 modulates human glioblastoma cell proliferation and invasion by upregulating EGFR expression.* Neuroscience letters, 2014. **566**: p. 82-87.
25. Xie, K., et al., *Polymorphisms in genes related to epithelial-mesenchymal transition and risk of non-small cell lung cancer.* Carcinogenesis, 2017. **38**(10): p. 1029-1035.
26. Tang, Y., et al., *Co-expression of AFAP1-AS1 and PD-1 predicts poor prognosis in nasopharyngeal carcinoma.* Oncotarget, 2017. **8**(24): p. 39001.
27. McClelland, M., et al., *Expression of CD74, the receptor for macrophage migration inhibitory factor, in non-small cell lung cancer.* The american journal of pathology, 2009. **174**(2): p. 638-646.
28. Xie, Y. and B. Wang, *Downregulation of TNFAIP2 suppresses proliferation and metastasis in esophageal squamous cell carcinoma through activation of the Wnt/β-catenin signaling pathway.* Oncology reports, 2017. **37**(5): p. 2920-2928.
29. Agnihotri, N., S. Kumar, and K. Mehta, *Tissue transglutaminase as a central mediator in inflammation-induced progression of breast cancer.* Breast Cancer Research, 2013. **15**(1): p. 1-9.
30. Nakazawa, N., et al., *High stromal TGFBI in lung cancer and intratumoral CD8-positive T cells were associated with poor prognosis and therapeutic resistance to immune checkpoint inhibitors.* Annals of surgical oncology, 2020. **27**: p. 933-942.

31. Maly, V., et al., *Circulating tumor cells in diagnosis and treatment of lung cancer.* in vivo, 2019. **33**(4): p. 1027-1037.
32. Liu, M., et al., *LAMC2 promotes the proliferation of cancer cells and induce infiltration of macrophages in non-small cell lung cancer.* Annals of Translational Medicine, 2021. **9**(17).
33. Inamura, K., et al., *Association of tumor TROP2 expression with prognosis varies among lung cancer subtypes.* Oncotarget, 2017. **8**(17): p. 28725.
34. Hinterleitner, C., et al., *Platelet PD-L1 reflects collective intratumoral PD-L1 expression and predicts immunotherapy response in non-small cell lung cancer.* Nature Communications, 2021. **12**(1): p. 7005.
35. Katsarou, S.D., et al., *Detyrosinated α-Tubulin, Vimentin and PD-L1 in Circulating Tumor Cells (CTCs) Isolated from Non-Small Cell Lung Cancer (NSCLC) Patients.* Journal of Personalized Medicine, 2022. **12**(2): p. 154.
36. Zhang, C., et al., *Non-small-cell lung cancer patients harboring TP53/KRAS co-mutation could benefit from a PD-L1 inhibitor.* Future Oncology, 2022. **18**(27): p. 3031-3041.
37. Hinterleitner, C., et al., *Platelet PD-L1 reflects collective intratumoral PD-L1 expression and predicts immunotherapy response in non-small cell lung cancer.* Nature communications, 2021. **12**(1): p. 1-16.
38. Inamura, K., *Diagnostic and therapeutic potential of microRNAs in lung cancer.* Cancers, 2017. **9**(5): p. 49.
39. Maruhashi, R., et al., *Chrysin enhances anticancer drug-induced toxicity mediated by the reduction of claudin-1 and 11 expression in a spheroid culture model of lung squamous cell carcinoma cells.* Scientific reports, 2019. **9**(1): p. 1-14.
40. Nallasamy, P., et al. *PD-L1, inflammation, non-coding RNAs, and neuroblastoma: Immuno-oncology perspective*. in *Seminars in cancer biology*. 2018. Elsevier.
41. Alfaro-Arnedo, E., et al., *IGF1R acts as a cancer-promoting factor in the tumor microenvironment facilitating lung metastasis implantation and progression.* Oncogene, 2022: p. 1-15.
42. Geng, Q., et al., *COL1A1 is a prognostic biomarker and correlated with immune infiltrates in lung cancer.* PeerJ, 2021. **9**: p. e11145.
43. He, P., et al., *The spirotetramat inhibits growth and reproduction of silkworm by interfering with the fatty acid metabolism.* Pesticide Biochemistry and Physiology, 2022. **188**: p. 105282.
44. Ochodnicka-Mackovicova, K., et al., *The DNA damage response regulates RAG1/2 expression in pre-B cells through ATM-FOXO1 signaling.* The Journal of Immunology, 2016. **197**(7): p. 2918-2929.
45. Robles, A. and B. Ryan, *KRT81 miR-SNP rs3660 is associated with risk and survival of NSCLC.* Annals of Oncology, 2016. **27**(2): p. 360-361.
46. Lee, S., et al., *A genetic variation in microRNA target site of KRT81 gene is associated with survival in early-stage non-small-cell lung*

cancer. Annals of Oncology, 2015. **26**(6): p. 1142-1148.
47. Campayo Guillaumes, M., et al., *A dual role for KRT81: a miR-SNP associated with recurrence in non-small-cell lung cancer and a novel marker of squamous cell lung carcinoma.* PLoS One, 2011, vol. 6, num. 7, p. e0022509, 2011.
48. Liu, Y., et al., *Loss of KCNJ15 expression promotes malignant phenotypes and correlates with poor prognosis in renal carcinoma.* Cancer Management and Research, 2019: p. 1211-1220.
49. Fan, J., et al., *Genome-wide analysis of expression quantitative trait loci identified potential lung cancer susceptibility variants among Asian populations.* Carcinogenesis, 2019. **40**(2): p. 263-268.
50. Lee, D., et al., *Pathway-based analysis using genome-wide association data from a Korean non-small cell lung cancer study.* PloS one, 2013. **8**(6): p. e65396.
51. Rabajdova, M., et al., *The crucial role of emilin 1 gene expression during progression of tumor growth.* Journal of cancer research and clinical oncology, 2016. **142**: p. 2397-2402.
52. Xiang, A., et al., *PCOLCE is potent prognostic biomarker and associates with immune infiltration in gastric cancer.* Frontiers in Molecular Biosciences, 2020. **7**: p. 544895.
53. Canals, J., et al., *1167P DIO2 is implicated in the antitumor effect of the lung embryonic stem cell conditioned medium and impacts prognosis in non-small cell lung cancer.* Annals of Oncology, 2021. **32**: p. S937.
54. Henderson, L., et al., *Genomic and gene expression profiling of minute alterations of chromosome arm 1p in small-cell lung carcinoma cells.* British journal of cancer, 2005. **92**(8): p. 1553-1560.
55. Fan, M., et al., *MicroRNA-30b-5p functions as a metastasis suppressor in colorectal cancer by targeting Rap1b.* Cancer Letters, 2020. **477**: p. 144-156.
56. Hoeks, J., et al., *Peroxisome proliferator-activated receptor-γ coactivator-1 and insulin resistance: acute effect of fatty acids.* Diabetologia, 2006. **49**: p. 2419-2426.
57. Liu, R., et al., *A HIF1a-GPD1 feedforward loop inhibits the progression of renal clear cell carcinoma via mitochondrial function and lipid metabolism.* Journal of Experimental & Clinical Cancer Research, 2021. **40**(1): p. 188.
58. Zhao, N., et al., *Arginine vasopressin receptor 1a is a therapeutic target for castration-resistant prostate cancer.* Science translational medicine, 2019. **11**(498): p. eaaw4636.
59. Huang, R., et al., *Co-expression Analysis of Genes and Tumor-Infiltrating Immune Cells in Metastatic Uterine Carcinosarcoma.* Reproductive Sciences, 2021. **28**(9): p. 2685-2698.
60. Campayo, M., et al., *A dual role for KRT81: a miR-SNP associated with recurrence in non-small-cell lung cancer and a novel marker of squamous cell lung carcinoma.* PloS one, 2011. **6**(7): p. e22509.
61. Liu, Y., et al., *Loss of KCNJ15 expression promotes malignant phenotypes and correlates with poor prognosis in renal carcinoma.*

Cancer Management and Research, 2019. **11**: p. 1211.
62. Stutvoet, T.S., et al., *MAPK pathway activity plays a key role in PD-L1 expression of lung adenocarcinoma cells.* The Journal of pathology, 2019. **249**(1): p. 52-64.
63. Gattinoni, L., Y. Ji, and N.P. Restifo, *Wnt/β-catenin signaling in T-cell immunity and cancer immunotherapy.* Clinical Cancer Research, 2010. **16**(19): p. 4695-4701.
64. Zhang, N., et al., *The EGFR pathway is involved in the regulation of PD-L1 expression via the IL-6/JAK/STAT3 signaling pathway in EGFR-mutated non-small cell lung cancer.* International journal of oncology, 2016. **49**(4): p. 1360-1368.
65. Saigi, M., et al., *MET-oncogenic and JAK2-inactivating alterations are independent factors that affect regulation of PD-L1 expression in lung cancer.* Clinical Cancer Research, 2018. **24**(18): p. 4579-4587.
66. He, J., et al., *Development of PD-1/PD-L1 pathway in tumor immune microenvironment and treatment for non-small cell lung cancer.* Scientific reports, 2015. **5**(1): p. 13110.
67. Gianchecchi, E., D.V. Delfino, and A. Fierabracci, *Recent insights into the role of the PD-1/PD-L1 pathway in immunological tolerance and autoimmunity.* Autoimmunity reviews, 2013. **12**(11): p. 1091-1100.
68. Ritprajak, P. and M. Azuma, *Intrinsic and extrinsic control of expression of the immunoregulatory molecule PD-L1 in epithelial cells and squamous cell carcinoma.* Oral oncology, 2015. **51**(3): p. 221-228.
69. Sato, M., D.S. Shames, and Y. Hasegawa, *Emerging evidence of epithelial-to-mesenchymal transition in lung carcinogenesis.* Respirology, 2012. **17**(7): p. 1048-1059.
70. Mani, V., et al., *Epithelial-to-Mesenchymal Transition (EMT) and Drug Response in Dynamic Bioengineered Lung Cancer Microenvironment.* Advanced Biosystems, 2019. **3**(1): p. 1800223.
71. Funaki, S., et al., *Chemotherapy enhances programmed cell death 1/ligand 1 expression via TGF-β induced epithelial mesenchymal transition in non-small cell lung cancer.* Oncology Reports, 2017. **38**(4): p. 2277-2284.
72. Shah, P.P. and S.S. Kakar, *Pituitary tumor transforming gene induces epithelial to mesenchymal transition by regulation of Twist, Snail, Slug, and E-cadherin.* Cancer letters, 2011. **311**(1): p. 66-76.
73. Azuma, K., et al., *Association of PD-L1 overexpression with activating EGFR mutations in surgically resected nonsmall-cell lung cancer.* Annals of oncology, 2014. **25**(10): p. 1935-1940.
74. Jin, W., *Role of JAK/STAT3 signaling in the regulation of metastasis, the transition of cancer stem cells, and chemoresistance of cancer by epithelial-mesenchymal transition.* Cells, 2020. **9**(1): p. 217.
75. Xu, D., et al., *PD-L1 Expression Is Regulated By NF-kappaB During EMT Signaling In Gastric Carcinoma.* Onco Targets Ther, 2019. **12**: p. 10099-10105.

76. Gao, Z.Z., et al., *Optimization strategies for expression of a novel bifunctional anti-PD-L1/TGFBR2-ECD fusion protein.* Protein Expression and Purification, 2022. **189**.
77. Gallo, M., et al., *Clinical utility of circulating tumor cells in patients with non-small-cell lung cancer.* Translational lung cancer research, 2017. **6**(4): p. 486.
78. Girard, L., et al., *An expression signature as an aid to the histologic classification of non-small cell lung cancer.* Clinical Cancer Research, 2016. **22**(19): p. 4880-4889.
79. Gómez-Morales, M., et al., *Differential immunohistochemical localization of desmosomal plaque-related proteins in non-small-cell lung cancer.* Histopathology, 2013. **63**(1): p. 103-113.
80. Liu, L., et al., *Detection of circulating cancer cells in lung cancer patients with a panel of marker genes.* Biochemical and biophysical research communications, 2008. **372**(4): p. 756-760.
81. Tian, Y., et al., *Single-cell transcriptomic profiling reveals the tumor heterogeneity of small-cell lung cancer.* Signal Transduction and Targeted Therapy, 2022. **7**(1): p. 346.

국문 초록

암 진단 및 치료 기법이 발달했음에도 불구하고 폐암은 국내 및 전 세계적으로 높은 발생률과 사망률을 나타내는 질환이다. 최근 도입된 면역항암제는 제한적인 효과와 부작용, 약제 내성을 나타내고 있다. 폐암 발생의 암 화와 악성 기작의 유전적 변이와 신호전달의 분석을 통해 면역항암에 필요한 폐암의 종양 유전자 표적과 그에 알맞은 신호전달경로의 규명이 필요하다. 폐암 환자 흉수로부터 유래한 세포주 수립한 후에 특성을 분석함으로서 조금 더 임상 치료에 가까운 치료 전략을 제시할 수 있다. 면역 항암 치료의 PD-L1은 종양 표지자로서 그를 기준으로 유전자 바이오마커와 신호전달 경로를 분석, 연구해 보았다. 수립한 세포주들은 환자의 특성을 잘 반영하기에 향후 면역 항암 치료의 표적이 될 수 있는 바이오마커와 신호전달경로를 임상에서 채택하여 적용하게 될 가능성을 제시하고자 한다.

암 환자들의 개인맞춤화치료가 진행되면서 종양의 분자적 발암 경로 및 종양 유전자의 바이오마커와 신호전달경로의 더 깊은 연구가 필요하여 분자적 발암 신호전달 경로, 유전체 분석을 수행하였다. 폐암 환자에서 유래한 세포주 29개를 수립한 후 수립한 세포주의 STR(Short Tandem Repeat) 검사를 통해 환자와 동일한 세포주인 것을 확인한 후 특성분석을 수행하였다. 수립한 세포주를 가지고 유전자의 돌연변이와 각 세포주에 대한 유전정보의 따른 특징을 규명하여 유전변이와의 상관관계를 밝히고자 하였다. 면역 항암 치료의 진행 여부의 검진방법으로 PD-L1의 발현 정도를 보듯이 우리가 수립한 세포주에서도 주요 분석 결정 분

류로 western analysis를 통해 PD-L1 발현 기준으로 세포주를 나누어 분석 시행 하였다. PD-L1의 발현 정도에 따라 나누어진 세포주들은 protein coding 영역인 exome 영역만을 보는 whole exome sequencing (WES), Transcriptome상의 RNA sequencing을 통해 유전 변이와 발현된 유전자들의 발현량의 차이를 확인한 후 PD-L1의 발현 상관관계와 신호전달경로의 상관관계를 밝히고자 하였다. 여러 연구에서 밝혀진 PD-L1의 발현과 EGFR의 돌연변이의 상관관계가 우리가 수립한 세포주에서 같은 경향성을 보여주었고, Protein level과 mRNA level에서의 PD-L1의 상관관계가 있음을 규명할 수 있었다. 또 한, 폐암 환자 흉수 유래로 수립한 29개의 폐암 세포주의 PD-L1의 발현의 차이와 돌연변이의 상관관계가 항암제 내성이나 민감성에서 어떠한 반응성을 보이는지를 연구 분석하여 그와 관련한 기전과 target 하는 메카니즘의 약물 반응성 및 신호전달 경로, 유전자들을 규명할 수 있었다. 항암 약제의 민감성과 유전정보의 분석을 통해 약물의 validation을 할 수 있었으며, 유의미한 결과를 통해 폐암 세포주의 특성을 분석함으로 면역-화학 병용 치료의 면역 항암 요법의 전략 중 하나로 사용될 것으로 사료 된다.

폐암 환자 유래 세포주의 유전체 정보를 가지고 세포생물학적 특성과 면역 치료의 주요 생물 지표인 PD-L1의 발현 상관관계를 파악함으로서 개인의 항암 치료 반응을 예측할 수 있도록 도울 수 있다. 생물 지표인 PD-L1에 따른 분류를 통해 유전자 변이와 약제 감수성의 상관관계를 규명하여 임상적 치료 접근성을 가까이 갈 수 있으며, 이를 통해 유전자

변이와 신호전달 경로가 면역-화학 치료 요법의 유의미한 결과를 도출 가능하다. 항암 치료의 세대가 정밀 종양학에 기인한 개인 맞춤화 되어 가고 있듯이 이 연구는 항암 치료의 활용 가능성과 필요성을 보여주는 결과물이 될 것이다.