國立臺灣大學醫學院醫學檢驗暨生物技術學系
學士班學生論文

Department of Clinical Laboratory Sciences and Medical Biotechnology
College of Medicine
National Taiwan University
Bachelor's Thesis

以組織蛋白基因體學揭露SEL1L3在肺腺癌進程與免疫療法之角色

Tissue Proteogenomic Landscape Reveals the Role of Uncharacterized SEL1L3 in Progression and Immunotherapy Response in Lung Adenocarcinoma

盛琪雅
Chi-Ya Shen

指導教授：俞松良 博士
Advisor: Sung-Liang Yu, Ph.D.

中華民國 112 年 4 月
April, 2023

國立臺灣大學學士班學生論文
口試委員會審定書

以組織蛋白基因體學揭露 SEL1L3 在肺腺癌進程與免疫療法之角色

Tissue Proteogenomic Landscape Reveals the Role of Uncharacterized SEL1L3 in Progression and Immunotherapy Response in Lung Adenocarcinoma

本論文係盛琪雅君（學號 B08404002）在國立臺灣大學醫學檢驗暨生物技術學系完成之學士班學生論文，於民國 112 年 4 月 28 日承下列考試委員審查通過及口試及格，特此證明

口試委員(3 位)：

俞松度 （簽名）

（指導教授）

陳玉如　　楊雅清

系 主 任 ： 俞松度 （簽名）

中文摘要

肺癌是全球癌症相關死亡的首要原因，而肺腺癌是其中最常見的病理組織學亞型。為了完成人類蛋白質圖譜及補全肺腺癌未滿足的臨床需求，我們致力於研究未知蛋白的可能功能，並探索它們在腫瘤生物學中的應用。

在我們的研究中，我們利用台灣癌症登月計畫（Taiwan Cancer Moonshot）所建立的大型組織蛋白基因體學數據，篩選出一個位於四號染色體、功能未知的蛋白SEL1L3，其在肺腺癌研究族群中的表現量出現失調。在我們的研究族群中，相較於成對的癌旁正常組織，SEL1L3在90%肺腺癌患者的腫瘤部位有較高的表現。此外，存活分析結果顯示SEL1L3與較佳的臨床結果相關。有趣的是，抑制SEL1L3表現會降低細胞活性，並活化內質網壓力反應路徑，暗示了SEL1L3會參與調控細胞壓力。同時，SEL1L3高表現患者的免疫學特徵證實了SEL1L3對肺腺癌的免疫表型與較佳預後具有重要作用。

綜上所述，我們的研究闡明SEL1L3可能在調控細胞壓力、腫瘤免疫微環境中扮演關鍵角色。此研究結果為進一步深入探討SEL1L3訊息傳遞路徑奠定了基礎，也表明SEL1L3是潛在的肺腺癌免疫療法新興佐劑。

關鍵字: Sel-1 suppressor of Lin-12-like 3、非小細胞肺癌、染色體中心人類蛋白質體計畫、uPE1

Abstract

Lung cancer is the leading cause of cancer-related death worldwide, and lung adenocarcinoma is the most common histological subtype. The fundamental pursuit to complete the human proteome atlas and the unmet clinical needs in lung adenocarcinoma have prompted us to study the functional role of uncharacterized proteins and explore their implications in cancer biology.

In this study, we characterized SEL1L3, a previously uncharacterized protein encoded from chromosome 4 as a dysregulated protein in lung adenocarcinoma from the large-scale tissue proteogenomics data set established using the cohort of Taiwan Cancer Moonshot. SEL1L3 was expressed in abundance in the tumor parts compared with paired adjacent normal tissues in 90% of the lung adenocarcinoma patients in our cohorts. Moreover, survival analysis revealed the association of SEL1L3 with better clinical outcomes. Intriguingly, silencing of SEL1L3 imposed a reduction in cell viability and activation of ER stress response pathways, indicating a role of SEL1L3 in the regulation of cell stress. Furthermore, the immune profiles of patients with higher SEL1L3 expression were corroborated with its active role in immunophenotype and favorable clinical outcomes in lung adenocarcinoma.

Taken together, our study revealed that SEL1L3 might play a vital role in the regulation of cell stress, interaction with cancer cells and the immune microenvironment. Our research findings provide promising insights for further investigation of its molecular signaling network and also suggest SEL1L3 as a potential emerging adjuvant for immunotherapy in lung adenocarcinoma.

Keywords: Sel-1 suppressor of Lin-12-like 3, non-small cell lung cancer, Chromosome-centric Human Proteome Project, uPE1

Abbreviation

ANOVA	Analysis of variance
ATCC	American Type Culture Collection
ATF	Activating transcription factor
BCG	Bacillus Calmette-Guérin
Bip	Binding immunoglobulin protein
cDNA	Complementary DNA
C-HPP	Chromosome-centric Human Proteome Project
CT	Cycle threshold
CTLA4	Cytotoxic T lymphocyte-associated antigen-4
DC	Dendritic cells
DMEM	Dulbecco's Modified Eagle Medium
eIF2α	Eukaryotic translation initiation factor 2α
ER	Endoplasmic reticulum
ERAD	Endoplasmic reticulum-associated degradation
FBS	Fetal bovine serum
GO	Gene Ontology
GSEA	Gene set enrichment analysis
HPP	Human Proteome Project
HRD1	HMG-CoA reductase degradation protein 1
HRP	Horseradish peroxidase
ICIs	Immune checkpoint inhibitors
IgG	Immunoglobulin G
IRE1α	Inositol-requiring enzyme 1α
LC	Lethal concentration

MPs	Missing proteins
MTT	3-[4,5-dimethylthiazole-2-yl]-2,5-diphenyltetrazolium bromide
NSCLC	Non-small cell lung cancer
PD-1	Programmed cell death protein 1
PD-L1	Programmed cell death ligand 1
PERK	Protein kinase RNA-like endoplasmic reticulum kinase
RPMI	Roswell Park Memorial Institute Medium
RT-qPCR	Reverse transcription quantitative real-time polymerase chain reaction
SD	Standard deviation
SDS-PAGE	Sodium dodecyl-sulfate polyacrylamide gel electrophoresis
SEL1L	Sel-1 suppressor of Lin-12-like
shRNA	Short hairpin RNA
T/N	Tumor part compared with paired adjacent normal tissues
TBP	TATA-box binding protein
TCGA	The Cancer Genome Atlas
TERS	Transmissible ER stress
TIR	Tumor immune-relevant
uPE1	Uncharacterized protein existence level 1
XBP	X-box binding protein

Contents

List of Figures

List of Tables

List of Appendices

Chapter I: Introduction

1.1 The Human Proteome Project

A genome stores all genetic and hereditary information on an organism, and proteins execute physiological and biochemical reactions in a living cell [1, 2]. Both are of great importance to control the molecular mechanisms underlying human health and diseases [1, 2]. Since its initial release in 2000, the human reference genome covering the euchromatic portion has been continually improved by the Human Genome Project and Genome Reference Consortium [3, 4]. The remaining 8% of the genome has been completed in 2022, propelling the advances in molecular biology and human health [5]. To date, 93.2% of the gene-encoded predicted proteins have been characterized as protein existence level 1 (PE1) with the continued progress in human proteome [6]. The number of PE2, PE3, and PE4 missing proteins (MPs) without protein-level evidence has been reduced to 6.8% [6]. The remarkable progress on the completion of the human proteome may lay the foundation for the development of protein-driven precision medicine for disease management [2]. Moreover, numerous evidence supports the relatively low correlation between proteome and transcriptome, indicating the necessity to complete the human proteome atlas [7]. The Human Proteome Project (HPP) initiated by the Human Proteome Organization aims to map the entire human proteome in a systematic effort to uncover and study the potential roles of human proteome in diseases as well as biological systems [8, 9]. Specifically, the chromosome-centric HPP (C-HPP) aims to identify MPs that lack sufficient experimental data from mass spectrometry or other experimental methods and characterize the functions of annotated but uncharacterized proteins with protein existence level 1 (uPE1). Both steps are beneficial to the completion of the human proteome.

1.2 Lung cancer

Lung cancer remains the leading cause of cancer-related death despite the advances in early detection such as low-dose computed tomography and antitumor treatments, for example, targeted therapy and immunotherapy [10, 11]. The main obstacles to ameliorating the high mortality of lung cancer include metastatic dissemination and drug resistance, highlighting unmet clinical needs to reveal the cancer progression mechanisms by which aberrant gene expressions and pathways regulate oncogenic signalings and to discover potential biomarkers and druggable targets for prognosis and treatment prediction of lung cancers [10, 11]. Our research team has been actively involved in identification of proteins encoded in the chromosome 4 under the C-HPP leadership, including the 47 uPE1 proteins coded on chromosome 4 released in neXtProt 2022–02 [12, 13]. By employing lung cancer as a model, we have previously demonstrated that exploring the under-studied membrane proteome in lung cancer tissues and cell lines complements the coverage of human proteome [14, 15].

1.3 SEL1L3

In the quest to understand the remaining MPs and uPE1 proteins and bridge their implications to cancer biology, in this study, we identified the SEL1L family member 3 (SEL1L3) as a dysregulated uPE1 protein in chromosome 4 in lung adenocarcinoma based on our previous proteogenomic data from early stage nonsmoking lung cancer cohort [12, 16]. SEL1L3 protein belongs to the SEL1L (Sel-1 Suppressor of Lin-12-Like) family, and SEL1L is a well-characterized member in the family. SEL1L is located in the endoplasmic reticulum (ER) membrane and participates in the ER-associated degradation (ERAD) pathway [17]. The paralogous relationship between SEL1L and SEL1L3 postulates the potential functions of SEL1L3. Recently, the SEL1L3 transcript has been screened to be upregulated in pulmonary carcinoid patients

that developed postsurgical metastasis [18]. Another study showed that SEL1L3 is involved in the competitive endogenous RNA network with prognostic value in cholangiocarcinoma [19]. To the best of our knowledge, despite the clinical findings; the biological functions and the molecular mechanisms of SEL1L3 especially in lung adenocarcinoma have never been elucidated. Following its identification from the tissue proteomics data set, we uncovered the clinical relevance of SEL1L3 and the potential SEL1L3-mediated signaling pathways and immunotherapy response by utilizing the prospective cohort of Taiwan Cancer Moonshot [16]. Furthermore, *in vitro* functional assays and immune profiling were performed to characterize the functional role of SEL1L3. Our study provides new insights into the role of SEL1L3 in lung adenocarcinoma progression and as a prospective adjuvant agent for immunotherapy.

Chapter II: Materials and Methods

2.1 Reagents, plasmid constructs, and antibodies

RPMI-1640 medium, Dulbecco's modified Eagle's medium (DMEM), fetal bovine serum and penicillin-streptomycin were purchased from Life Technologies (Carlsbad, CA). TRC1 Scramble (pLKO.1-shScramble; ASN0000000004) and two SEL1L3-shRNA-containing lentiviral vectors (pLKO.1-shSEL1L3-1 and pLKO.1-shSEL1L3-2; TRCN0000159090 and TRCN0000160325) were obtained from the National RNAi Core Facility (Academia Sinica, Taipei, Taiwan). Goat antimouse IgG-HRP and goat antirabbit IgG-HRP secondary antibodies were purchased from Santa Cruz Biotechnology (Santa Cruz, CA). Anti-Bip (#3177), anti-PERK (#3192), anti-IRE1α (#3294), anti-eIF2α (#5324), and antiphospho-eIF2α (#3398) antibodies were purchased from Cell Signaling Technology (Danvers, MA). Anti-ATF4 (10835-1-AP), anti-XBP-1u (25997-1-AP), anti-XBP-1s (24868-1-AP), and anti-ATF6 (24169-1-AP) antibodies were purchased from Proteintech Group (Rosemont, IL). Antiphospho-PERK (MA5-15033) was purchased from Thermo Fisher Scientific (Waltham, MA). Antiphospho-IRE1α (ab243665) was purchased from Abcam (Cambridge, U.K.). Anti-β-actin (GT5512) antibody was purchased from GeneTex (Irvine, CA).

2.2 Cell culture and lentivirus transduction

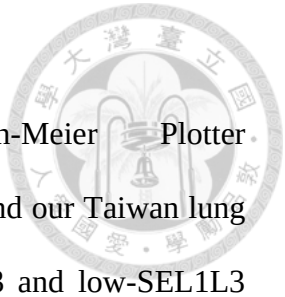
HOP-92, NCI-H23, NCI-H226, EKVX, NCI-H522, HOP-62, NCI-H322M, and NCI-H460 cells were purchased from the National Cancer Institute's (NCI) Developmental Therapeutics Program (Bethesda, MD). PC-9, H1299, H1568, BEAS-2B, H1650, H1437, A549, H3255, HCC827, and H838 cells were purchased from American Type Culture Collection (ATCC). CL1-0, CL1-5, and PE-089 were

established from Taiwanese lung cancer patients [20, 21]. Cells were cultured in the recommended medium at 37 °C in a humidified atmosphere of 5% CO_2. For the establishment of SEL1L3-knockdown stable cell lines, cells were infected with SEL1L3-shRNA-containing lentivirus in Polybrene (8 μg/mL) containing medium. After 24 h of infection, the cells were treated with 2 μg/mL puromycin and puromycin-resistant clones were selected.

2.3 Real-time quantitative RT-PCR

Total RNAs from the cultured cell lines were extracted by the TRIzol reagent (Life Technologies, Carlsbad, CA), and 1 μg of total RNAs was used in cDNA synthesis with random hexamer primers using Superscript III reverse transcriptase (Life Technologies, Carlsbad, CA). 10 ng cDNA was used as the template for mRNA expression detection in real-time quantitative RT-PCR on ABI prism 7900 sequence detection system (Life Technologies, Carlsbad, CA). The primers used in this study for SYBR Green real-time quantitative RT-PCR are described below.

SEL1L forward primer: 5'-GAG AAT ACG GCT GCC TGA TGA AG-3'; SEL1L reverse primer: 5'-CAG GTG CAG TTG TCC AAG ACC A-3'; SEL1L2 forward primer: 5'- TGC AGC CAA CAA ATA CCA CAA-3'; SEL1L2 reverse primer: 5'- AGC ACA GGT ATG TGG GCA TCT-3'; SEL1L3 forward primer: 5'-CCT GCT TGG CTT TCC TTA TGC-3'; SEL1L3 reverse primer: 5'-TGG CCT TGC AAT AAT GGA GTA A-3'; TBP forward primer: 5'-CAC GAA CCA CGG CAC TGA TT-3'; TBP reverse primer: 5'-TTT TCT TGC TGC CAG TCT GGA C-3'. The relative mRNA expression of target genes was determined as $-\Delta CT = -[CT_{target} - CT_{TBP}]$. The target/TATA-box binding protein (TBP) mRNA ratio was calculated as $2^{-\Delta CT} \times K$, in which K is a constant. All experiments were performed in triplicate.



2.4 Survival analysis

Survival analysis was performed in Kaplan-Meier Plotter (https://kmplot.com/analysis/) [22], the Human Protein Atlas [23], and our Taiwan lung cancer cohort data sets. Patients were separated into high-SEL1L3 and low-SEL1L3 groups, and overall survival analysis was carried out.

2.5 Western blot analysis

The preparations of whole-cell lysates and immunoblotting was performed as previously described [24, 25]. Cells were first lysed in lysis buffer (50 mM Tris-HCl (pH 7.4), 1% NP-40, 150 mM NaCl, 1 mM EDTA) containing protease and phosphatase inhibitor cocktails (Roche, Basel, Switzerland). Proteins were then electrophoretically separated on sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Immunoblotting was conducted with appropriate antibodies, visualized by a chemiluminescence assay kit (Merck Millipore, Burlington, MA), and detected by the FUJIFILM LAS-3000 ECL system.

2.6 3-[4,5-Dimethylthiazole-2-yl]-2,5-diphenyltetrazolium bromide (MTT) and colony formation assays

MTT and colony formation assays were used to quantitate viable cell numbers. For the MTT assay, 1000 cells were seeded into 96-well plates. Cell viability was evaluated by the MTT assay at indicated time points according to the manufacturer's protocol (Promega, Madison, WI). The absorbance at 570 nm was measured on a multiwell scanning spectrophotometer (Victor3; PerkinElmer, Boston, MA). For colony formation assay, 500 cells were seeded into each well of 6-well plate and treated with individual drugs with indicated concentration (10 μM carboplatin, 50 μM cisplatin, 1 μM etoposide, 10 nM gemcitabine, 5 nM paclitaxel, and 50 nM vinorelbine). After incubation at 37 °C for 7 days, the cell colonies were fixed with methanol, stained with

0.001% crystal violet, and then photographed. Colonies with a diameter greater than 1 mm were counted under an inverted microscope.

2.7 Transwell migration assays

Transwell chambers (8 μm pore size; BD Biosciences, Franklin Lake, NJ) were utilized to perform *in vitro* migration assays. First, 3000 cells suspended in serum free medium were seeded onto the upper chambers and 500 μL medium containing 10% FBS was placed in the lower chambers. After 20 h of incubation, upper chambers were swabbed with a cotton swab, fixed with methanol, and then stained with Giemsa solution (Sigma-Aldrich, St. Louis, MO). Subsequently, the cells attached to the lower surface of the chambers were counted under a light microscope. All experiments were assayed in triplicate.

2.8 Patient tumor specimens

The cohort of 98 patients with lung adenocarcinoma who underwent surgical resection was obtained from the Taichung Veterans General Hospital (Taichung, Taiwan) from March 2000 to January 2008. These investigations were approved by the Taichung Veterans General Hospital and written informed consent was obtained from all the patients.

2.9 Structure comparison between SEL1L and SEL1L3

BIOVIA Discovery Studio 2021 (Dassault Systems, San Diego, U.S.A.) software was used to perform all the calculations for protein structure comparison. The predicted full-length structures of the SE1L1 and SE1L3 proteins were obtained from the AlphaFold Protein Structure database [26, 27], and the protein superimposition was predicted using a molecular overlay method (50% electrostatic and 50% steric fields) and then the percentage of structural similarity was calculated.

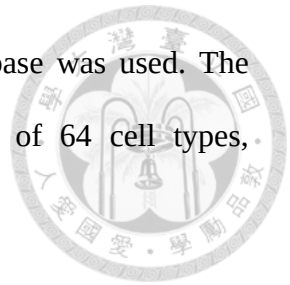
2.10 Pathway enrichment analysis

The proteomic data of the Taiwan Cancer Moonshot is available in the NCI Proteomics Data Commons (accession number: PDC000219 and PDC000220) [16]. The protein expression association between SEL1L3 and other 9447 proteins derived from the proteomic data of Taiwan Cancer Moonshot [16] was first analyzed by Pearson correlation. Next, the differentially expressed proteins between SEL1L3-high (n = 40) and SEL1L3-low (n = 49) groups were analyzed by Welch's t test, and a total of 445 SEL1L3-associated proteins were identified. Subsequently, the 445 proteins were subjected to the analysis of Gene Ontology (GO) pathway enrichment analysis by using MetaCore Analytical Suite (GeneGo, St. Joseph, MI). The pathways with a *p* value less than 0.05 were identified. The gene set enrichment analysis (GSEA) was performed using GSEA software (Broad Institute, Cambridge, Massachusetts) [28]. The proteomic profile of patients with high (T/N $\geq$ 1.3, n = 40) and low (T/N < 1.3, n = 49) SEL1L3 protein expression was uploaded into GSEA as unfiltered data, and the protein abundance from a ranked list in the gene set BIOCARTA_ERAD_PATHWAY was analyzed.

2.11 Estimation of an infiltrated immune population

CIBERSORTx (http://cibersortx.stanford.edu) [29] and xCell (https://xcell.ucsf.edu/) [30] are analytical web tools. CIBERSORTx is used to determine cell abundance and cell-specific gene expression profiles from bulk transcriptome while xCell is used to perform cell enrichment analysis from gene expression data sets [29, 30]. For the CIBERSORTx analysis, the proteomic data were extracted from the data set of Taiwan Cancer Moonshot [16]. The LM22 immune cell type signature matrix and the setting of 100 for "permutations for significance analysis" was applied for the CIBERSORTx analysis. For the xCell analysis, the analyzed

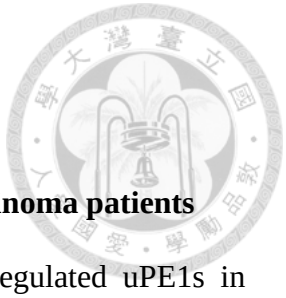
transcriptomic data from the Cancer Genome Atlas (TCGA) database was used. The xCell algorithm was used to estimate the comprehensive levels of 64 cell types, including 34 types of immune cells.



2.12 Statistical analysis

The statistical analyses used to compare the difference between the two groups are described in the corresponding figure legends. All analyses were performed using GraphPad Prism 9 (GraphPad Software, San Diego, CA) and R Programming (Johns Hopkins University, Baltimore, MD). Differentially expressed proteins were analyzed by Welch's t test and considered significant with a p value < 0.05. All tests were two-tailed, and p values < 0.05 were considered significant.

Chapter III: Results



3.1 Identification of the uPE1 protein SEL1L3 in lung adenocarcinoma patients

The mission of HPP prompted us to first identify the dysregulated uPE1s in chromosome 4 by utilizing the tissue proteomics data set established from the prospectively collected cohort of Taiwan Cancer Moonshot consisting of 89 patients with lung adenocarcinoma [16]. SEL1L3 was one of the dysregulated uPE1s, which was found to be more abundant in 90% (80/89) of the tumor parts compared with paired adjacent normal tissues (paired two-sample t test, $p = 1.13 \times 10^{-8}$, Figure 1, dark red and pale red). Notably, 45% (40/89, $p = 1.76 \times 10^{-15}$) of patients displayed statistically upregulated SEL1L3 expression in tumor compared to the adjacent normal tissues ($\log_2$T/N values > 0.38, Figure 1, dark red). However, the expression patterns were not associated with cancer stages (Figure 2A and Table 1). In our previous study, we identified a high-risk subtype of stage IA and IB tumors with late-like proteomic profiles which were enriched with clinical features of visceral pleural invasion, TP53 mutations, and higher mutational burden. Furthermore, the stage IB patients were enriched in lymph node metastasis and poor survival when compared among all IB and IA patients. Hence, a proteomics-based refined five-stage molecular classification was established to discriminate the diverse clinical trajectories of patients [16]. We adopted the refined five-stage classification to analyze the expression patterns of SEL1L3 again, but no significant differences between early stage, early late-like stage, and late stage were observed in relation to SEL1L3 expression (Figure 2B and Table 1). To determine whether SEL1L3 expression was correlated with clinical features, additional clinicopathological characteristics of patients were evaluated. Among all the 15 characteristics, we observed a significant correlation between SEL1L3 expression and

pathological differentiation (Table 1). The analysis revealed that 12.5% (5/40) of patients with high SEL1L3 expression (T/N ≥ 1.3) had poorly differentiated tumors whereas 39.6% (19/49) of patients with low SEL1L3 expression (T/N < 1.3) were identified as poorly differentiated, indicating that SEL1L3 could be a better prognosis marker. In the follow-up analysis of our cohort [16], 10% (4/40) of high SEL1L3 expression patients relapsed after surgical resection, and 27% (13/49) of patients with low SEL1L3 expression had disease recurrence (Figure 2C and Table 1, $p = 0.048$), consistent with their pathological features. Survival analysis by Kaplan–Meier Plotter suggested that high SEL1L3 expression was associated with prolonged overall survival ($p = 1.3 \times 10^{-6}$, Figure 3A). The SEL1L3 expression in tumor specimens from 98 Taiwan lung cancer patients was validated by RT-qPCR. Consistently, patients with high SEL1L3 expression showed prolonged overall survival with marginal significance due to the limited case number ($p = 0.12$, Figure 3B). In agreement with the above-mentioned results, a positive association between high SEL1L3 expression and prolonged overall survival was further validated by utilizing a larger cohort of the Human Protein Atlas [23] ($p = 0.40$, Figure 3C). Taken together, we propose that SEL1L3 expression might be positively associated with the favorable outcomes in lung adenocarcinoma.

3.2 Characterization of SEL1L3 with SEL1L protein family members

It is widely known that proteins in a family are derived from a common ancestral protein [31]. Typically, protein structures are highly conserved in families of homologous proteins and share high sequence similarities that frequently share common functional properties [31]. Due to the uncharacterized biological functionality of SEL1L3, we were interested in studying its potential domains and functional repeats. Motif analysis of SEL1L3 protein by MyHits

(https://myhits.sib.swiss/cgi-bin/motif_scan) showed presence of 5 Sel1 repeats (Figure 4), which are commonly shared among proteins in the SEL1L family [32, 33]. Surprisingly, protein sequence alignment revealed relatively low similarity between SEL1L family members (Figure 5). Similarity between SEL1L and SEL1L2 was 31%, and their similarities to SEL1L3 were 15% and 11%, respectively, which were contrary to the stereotypes. Structural superposition also indicated that SEL1L and SEL1L3 exhibited a low percentage in structural similarity (13.5%) with only 4 residue–residue pairs predicted to be structurally superimposed (Appendix 1). Subsequent analysis of the expression levels of SEL1L family members in 21 laboratory and primary lung cancer cell lines including A549, BEAS-2B, CL1-0, CL1-5, EKVX, HCC827, HOP-62, HOP-92, H838, H1299, H1437, H1568, H1650, H3255, NCI-H23, NCI-H226, NCI-H322M, NCI-H460, NCI-H522, PC-9, and PE-089 demonstrated that SEL1L levels positively correlated with SEL1L3 levels ($r = 0.57$, $p = 0.007$, Figures 6A to 6C, 7A), while no significant correlation was observed between SEL1L2 and SEL1L3 ($r = -0.11$, $p = 0.63$, Figures 6A to 6C, 7B), suggesting potential upstream coregulators of SEL1L and SEL1L3. Furthermore, we explored whether the redundancy and compensation of SEL1L family exist. First, two lentivirus-based SEL1L3-specific shRNAs were introduced into A549 cells, in which endogenous SEL1L3 was moderately expressed. The stable SEL1L3-silencing A549 cell clones (shSEL1L3-1, shSEL1L3-2) were generated to evaluate whether SEL1L3 deficiency influenced SEL1L and SEL1L2 expressions. Both SEL1L3-specific shRNAs significantly suppressed the expression of endogenous SEL1L3 even though the knockdown efficiency of shSEL1L3-2 was less than that of shSEL1L3-1. The mRNA levels of SEL1L and SEL1L2 were not induced by SEL1L3 knockdown (Figure 8). These results implied that there might be no compensation effect for SEL1L3 depletion, which is

likely due to the low similarity of SEL1L3 with the other two SEL1L family members.

3.3 Impact of SEL1L3 on cancerous phenotypes *in vitro*

To elucidate the functional role of SEL1L3 in lung adenocarcinoma, two SEL1L3-knockdown stable clones (shSEL1L3-1 and shSEL1L3-2) were assayed for cell proliferation, tumor growth, and cell migration. First, the suppression of SEL1L3 was determined by RT-qPCR and the mRNA expression of SEL1L3 was reduced by 95% and 89% in shSEL1L3-1 and shSEL1L3-2 clones, respectively (Figure 9A). The knockdown clones were used to perform MTT, colony formation, and transwell migration assays. We discovered that silencing of SEL1L3 significantly reduced cell proliferation (Figure 9B) and caused a decrease in the number of cell colonies (Figure 9C), suggesting that SEL1L3 was involved in the regulation of cell proliferation and tumorigenesis *in vitro*. However, the migratory capability of cancer cells did not show significant difference upon SEL1L3 knockdown (Figure 9D).

3.4 Enrichment analysis of SEL1L3-regulated pathways

The next intriguing question is that the low expression of SEL1L3 is associated with worse survival but the suppression of SEL1L3 expression inhibits *in vitro* tumor growth. These contradictory results implied that SEL1L3 might play different roles in tumor cells and the tumor microenvironment. To further elucidate the role of SEL1L3 in lung cancer progression and investigate the underlying molecular mechanisms of SEL1L3-mediated signaling pathways, proteins whose average expression levels significantly associated with SEL1L3 expression were first identified in 89 lung adenocarcinoma patients (Welch's t test, $p < 0.05$). Out of 445 proteins identified, 239 proteins (red dots, named high-SEL1L3 cluster) showed positive correlation while 206 proteins (blue dots, named low-SEL1L3 cluster) exhibited negative correlation with the protein expression of SEL1L3 (Appendix 2A). Next, the 445 proteins were subjected to

unsupervised hierarchical clustering analysis and grouped to two clusters (Appendix 2B). Proteins in each cluster were then subjected to Gene Ontology (GO) pathway enrichment analysis. The enriched pathways from the low-SEL1L3 cluster included mRNA processing, T cell receptor signaling pathway and negative regulation of cell growth (Appendix 3A, left panel). The negative correlation of SEL1L3 expression and the negative regulation of cell growth was consistent with our observation that SEL1L3 knockdown inhibited cell proliferation and tumor growth (Figures 9B and 9C). The enriched pathways from the high-SEL1L3 cluster were dominant in endoplasmic reticulum (ER) function and endoplasmic reticulum-associated protein degradation (ERAD) pathways (Appendix 3A, right panel). The ERAD pathway was also well corroborated with the gene set enrichment analysis (GSEA) and was enriched in the group of patients with high SEL1L3 protein expression with marginal significance (p = 0.084, Appendix 3B).

3.5 Immune cell profiling associated with SEL1L3 level

Due to the enrichment of immune-related pathways found in the low-SEL1L3 cluster (Appendix 3A, left panel), we utilized the proteomic data of Taiwan Cancer Moonshot to estimate the abundance of 22 immune cell types based on CIBERSORTx analysis [16, 29]. The patients with high SEL1L3 protein level (T/N < 1.3) presented higher infiltration levels of T cells CD8, T cells CD4 memory resting and activated, macrophages M0, and both dendritic cells (DCs) resting and activate in compared to low protein level of SEL1L3 (T/N ≥ 1.3). However, the differences in immune profile that are reported to be associated with better clinical outcomes in lung cancer did not reach the significance [34, 35] (Figure 10). For further validation in a larger cohort, the transcriptomic data of lung adenocarcinoma patients from the TCGA database was utilized for xCell analysis [30], and the abundance of 34 immune cell types was

measured. Consistent with our observation in CIBERSORTx analysis, patients with high SEL1L3 expression demonstrated significantly higher infiltration levels of CD4+ T-cells, CD8+ T-cells, and DCs (Figure 11). These findings supported the prolonged overall survival of patients with higher SEL1L3 expression and implied that SEL1L3 might be involved in immunomodulation upon tumorigenesis.

3.6 Role of SEL1L3 in ER stress and drug response

SEL1L family is defined by the presence of the conserved sequences known as Sel1 repeats, among which SEL1L serves as an indispensable adaptor for HRD1 as ERAD components [17]. Given that GSEA indicated ERAD pathway enrichment in the group of patients with high SEL1L3 protein expression (Appendix 3B), and that SEL1L participated in ERAD, we further characterized whether SEL1L3 plays a role in ER stress homeostasis. The expressions of proteins involved in ER stress response including Bip/GRP78, three ER stress transducers (PERK, IRE1α, ATF6) and the downstream effectors (eIF2α, ATF4, XBP-1u, XBP-1s) were analyzed by Western blot and the phosphorylation of PERK, eIF2α, and IRE1α were detected by specific antiphospho-serine/phospho-threonine antibodies [36]. SEL1L3-knockdown was accompanied by the down-regulation of Bip expression, and the up-regulation of PERK expression, PERK phosphorylation, and eIF2α phosphorylation in A549 cells (Figures 12A and 12B). Furthermore, we noticed that the alterations were more obvious in shSEL1L3-1 cells compared with shSEL1L3-2 cells in a dose dependent manner. The SEL1L3 expression was suppressed to 4% and 34% in shSEL1L3-1 and shSEL1L3-2 cells, respectively (Figure 12A). These results indicated that SEL1L3-knockdown might trigger ER stress through Bip/PERK/eIF2α axis in lung cancer cells.

Chemotherapy represents an extrinsic source of stress for cancer cells and is associated with ER stress regulation [37]. We thus explored whether the

shSEL1L3-mediated ER stress influences the response of chemotherapy using the colony formation assay. The lethal concentrations between 20% to 30% (LC20-LC30) of 6 most common chemotherapeutic drugs used in lung cancer such as carboplatin, cisplatin, etoposide, gemcitabine, paclitaxel, and vinorelbine were first determined and then the impact of SEL1L3 on drug response was evaluated. The selected LCs for carboplatin, cisplatin, etoposide, gemcitabine, paclitaxel, and vinorelbine were 10 μM, 50 μM, 1 μM, 10 nM, 5 nM, and 50 nM, respectively (Figure 13A). Intriguingly, SEL1L3 knockdown did not disturb the susceptibility of any chemotherapeutic drugs used in this study. (Figure 13B).

Chapter IV: Discussion

In our study, SEL1L3 derived from the SEL1L family is unveiled as one of the dysregulated uPE1s in lung adenocarcinoma by the proteogenomic study on the cohort of Taiwan Cancer Moonshot [16]. The clinicopathological analyses revealed that high expression levels of SEL1L3 were associated with prolonged overall survival and lower recurrence rate. We further performed *in vitro* functional assays to validate the potential SEL1L3-mediated signaling pathways. Our results indicated significant differences in cell proliferation and ER stress response upon SEL1L3 knockdown. Immune profiles in relation to SEL1L3 protein levels further supported the favorable clinical outcomes for high SEL1L3 expression patients (Figure 14).

Multiple studies support the link between ER stress and cancer [38-40]. Tumor cells are exposed to intrinsic and external factors such as hypoxia, acidification, nutrient deprivation, and accumulation of reactive oxygen species that alter protein homeostasis [38]. These conditions promote the accumulation of unfolded or misfolded proteins in the ER lumen, and thus fuel ER stress that produces endogenous or exogenous damage to cells [41]. Therefore, to maintain cell homeostasis, ER stress response pathways are activated that involves the engagement of Bip with accumulated misfolded proteins resulting in the release of three ER stress transducers (PERK, IRE1α, and ATF6) [37, 40]. Subsequently, these dissociated ER transducers then activate downstream pathways which increase the protein-folding capacity and decrease the protein-folding load in the ER thereby restoring cellular protein homeostasis [42, 43]. Even though ER stress response is considered cytoprotective, it is also reported to be involved in proapoptotic mechanisms [44]. Under unsolved ER stress conditions, the ER stress transducers initiate an alternative program which induces cell death [45]. Recent findings have indicated that ER stress response participates in the adaptive survival signaling, which

promotes resistance to chemotherapy in tumor cells [37, 46]. However, a thorough understanding of the relationship between ER stress response and chemotherapy resistance remains poorly understood [37, 46]. In our study, GO pathway enrichment analysis and GSEA analysis showed that the ERAD pathway is enriched in patients with high SEL1L3 expression (Appendix 3A and 3B). On the basis of our findings, we hypothesized that SEL1L3 might be involved in ER homeostasis. Our results revealed that expression levels of Bip, PERK-the ER stress transducer, and eIF2α-the downstream effector were responsive to SEL1L3 knockdown at protein level (Figures 12A and 12B). Interestingly, the colony forming ability of chemotherapeutic drug-treated clones showed no significant change upon SEL1L3 knockdown (Figure 13B). The reduction of cell proliferation and activation of ER stress response upon SEL1L3 knockdown implied that SEL1L3 engaged in the regulation of cell stress homeostasis. However, it might not be associated with resistance to chemotherapy at least in our assay conditions.

The SEL1L family is characterized by low sequence similarity between members which renders the determination of functional homologues difficult [32]. Although Sel1 repeats are commonly shared among proteins in the SEL1L family, the relatively low sequence and structural similarity between SEL1L and SEL1L3 implied that they might be responsible for distinct biological functions (Figure 5 and Appendix 1). Preliminary reports have implicated that SEL1L plays a critical role in facilitating ERAD, which is activated by accumulating ER stress to promote the degradation of unfolded proteins [17, 47]. SEL1L serves as the membrane-embedded ERAD adaptor for the E3 ubiquitin ligase HRD1, and is indispensable for HRD1 stability and substrate recruitment [48]. Our study revealed that SEL1L3 participated in the regulation of Bip/PERK/eIF2α axis, which is the initiation of ER stress response and ERAD that SEL1L engages is a

downstream pathway to maintain ER homeostasis [40, 49]. Further investigation in molecular and cellular biology is required to determine the involvement of SEL1L3 in ER stress response.

Recent studies have revealed the cell-extrinsic effects of ER stress [50]. Accumulating ER stress in cancer cells evokes the induction of ER stress response pathways and the release of ER stress-transmitting factors, resulting in a mixed proinflammatory/immunosuppressive phenotype of the recipient tumor-infiltrating immune cells [51, 52]. This intercellular signaling, termed transmissible ER stress (TERS), has been confirmed to defect T cell priming [50, 52]. In our study, while SEL1L3 knockdown inhibited cell proliferation (Figures 9B and 9C) and triggered ER stress response (Figures 12A and 12B), low expression of SEL1L3 was associated with unfavorable immune profiles (Figures 10 and 11) and worse overall survival (Figures 3A to 3C). Cancer cells with low SEL1L3 expression underwent chronic ER stress and might modulate the immunological landscape of tumor microenvironment via TERS, leading to unfavorable clinical outcomes. In the tumor microenvironment, there is a complex interaction between the immune system and tumor cells [34]. It has been revealed that immune escape is a fundamental trait of cancer [53]. Accumulating evidence has reported that cancer immunotherapies by immune checkpoint inhibitors (ICIs) targeting PD-L1, PD-1, and CTLA4 can enhance the antitumor activity and prolong survival by blocking immune-inhibitory signals [54]. SEL1L3 has been recently reported as a favorable prognostic marker in urothelial cancer [23], which is known to be immunogenic with a high tumor mutation burden [55] and responsive to immunotherapies including Bacillus Calmette-Guérin (BCG) therapy and ICIs [56]. According to previous reports, patients harboring high tumor mutation burden are associated with favorable response of immunotherapies [57], namely urothelial cancer,

melanoma, and lung adenocarcinoma [58]. However, 80% of patients among all tumor types are still unresponsive to ICI treatment [59], highlighting the urge to develop predictive biomarkers and regulators of immunotherapy response.

A recent study has suggested the potential role of SEL1L3 as a clinical biomarker of prognosis, risk assessment, and prediction of immunotherapy response in cutaneous melanoma patients [60]. In this study, melanoma patients in the TCGA cohort were classified into low, medium, and high immunity subtypes. A four-gene tumor immune-relevant (TIR) signature is further identified to be significantly associated with the overall survival of melanoma patients, and SEL1L3 is one of the genes in this signature. The mRNA expression levels of the four genes are positively correlated to the overall survival, activity of CD8+ T cells, and anti-CTLA4 response rate. This study implicated that SEL1L3 may be functionally involved in tumor progression and immune response. In lung adenocarcinoma, SEL1L3 was predicted to be a biomarker of prognosis and immune infiltration by a nine-gene risk model, in which SEL1L3 was found to be positively correlated with favorable outcomes [61]. Furthermore, the mRNA expressions of the nine genes are associated with CD8+ T cells, CD4+ T cells, macrophages, neutrophils, and dendritic cells, implying a high correlation with the level of immune cell infiltration. This study suggested the potential roles of SEL1L3 in tumor microenvironment and its prognostic value for lung adenocarcinoma. Consistently, our analysis revealed that high protein level of SEL1L3 was associated with lower recurrence rate (Figure 2C) and prolonged overall survival (Figures 3A to 3C). Additionally, by utilizing the proteomic data of Taiwan Cancer Moonshot [16], we found that patients with high protein level of SEL1L3 presented higher infiltration levels of T cells CD8, T cells CD4 memory resting, T cells CD4 memory activated, macrophages M0, dendritic cells resting, and dendritic cells activated (Figure 10). The

immune profiles have been reported to be associated with better clinical outcomes in lung cancer [34, 35]. Similar results have been obtained from the transcriptomic data of the TCGA melanoma and adenocarcinoma cohorts, in which positive correlations of SEL1L3 expression and the infiltration levels of CD8+ T cells, CD4+ T cells, macrophages, and dendritic cells are elucidated [60, 61]. Evidence of SEL1L3 as a potential tumor suppressor is provided at the transcriptome and proteome levels independently. The consistent conclusions corroborate the potential application of SEL1L3 in prediction of the survival outcome and immunotherapy response in various cancers.

Our findings provide insights for the further exploration and development of the functions and mechanisms of uPE1 protein, SEL1L3. In our report, a proteogenomic study on the cohort of Taiwan Cancer Moonshot was performed [16], and the dysregulated protein SEL1L3 was identified. We assessed the clinical relevance of SEL1L3 and the potential SEL1L3-mediated signaling pathways. Accordingly, *in vitro* functional assays and immune profiling in relation to SEL1L3 levels were carried out. Overall, this study revealed SEL1L3 might play a tumor suppression role via regulation of stress homeostasis and interaction of tumor and microenvironment (Figure 14).

Figures

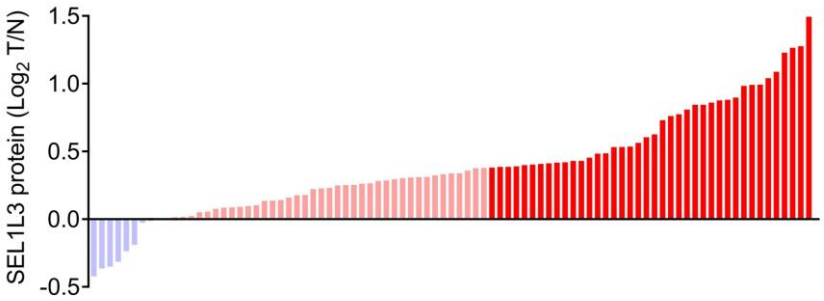


Figure 1. Upregulation of SEL1L3 protein in lung adenocarcinoma.

Relative SEL1L3 protein expressions of tumors (T) compared with adjacent normal tissues (N) of 89 lung adenocarcinoma patients in Log_2 scale. Dark red: Log_2 T/N value > 0.38. Pale red: 0 < Log_2T/N value ≤ 0.38. Blue: Log_2T/N value ≤ 0.

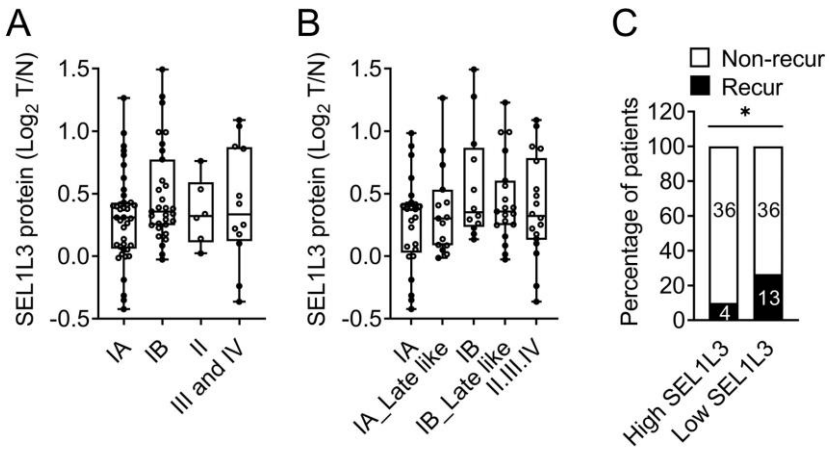


Figure 2. The correlation of SEL1L3 protein expression with tumor stages and recurrence rate.

(A and B) The correlation of SEL1L3 protein expression with tumor stages (A) and refined tumor stages (B). One-way ANOVA test was used to compare the difference between groups. (C) SEL1L3 expression was negatively associated with recurrence of surgery-treated lung adenocarcinoma patients. Patients with and without disease recurrence were represented by close bar and open bar respectively, and the patient numbers are shown in bars. High SEL1L3: Log_2T/N value > 0.38. Low SEL1L3: Log_2T/N value ≤ 0.38. *$p < 0.05$ by Chi-square test.

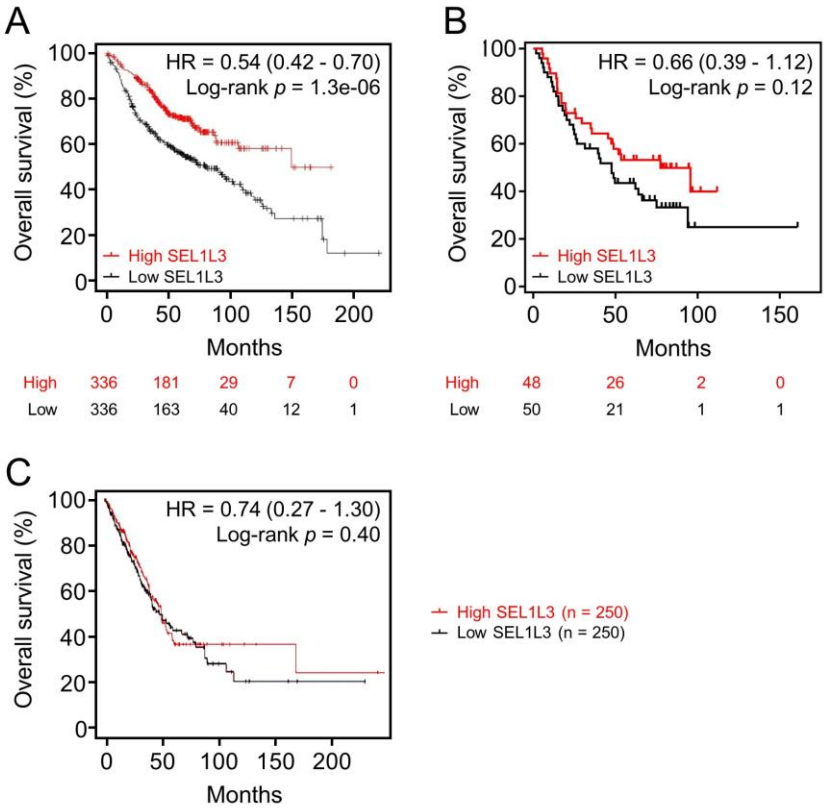


Figure 3. Positive correlation between SEL1L3 expression and prolonged overall survival.

(A to C) SEL1L3 expression was correlated with prolonged survival of lung adenocarcinoma patients. Kaplan–Meier analysis of overall survival for lung adenocarcinoma patients grouped by high-SEL1L3 or low-SEL1L3 mRNA expression using public database (A and C) and Taiwan lung adenocarcinoma patient cohorts (B). The mRNAs were extracted from 98 Taiwan lung adenocarcinoma patients, analyzed for SEL1L3 expression by RT-qPCR and grouped into high-SEL1L3 or low-SEL1L3 mRNA expression groups.

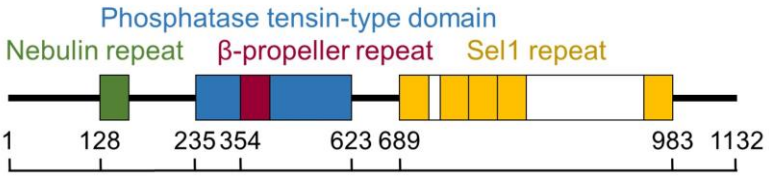


Figure 4. Schematic diagram of SEL1L3 protein domains and functional repeats.

Domains and functional repeats were identified by MyHits (https://myhits.sib.swiss/cgi-bin/motif_scan). The conserved Sel1 repeats are colored in yellow.

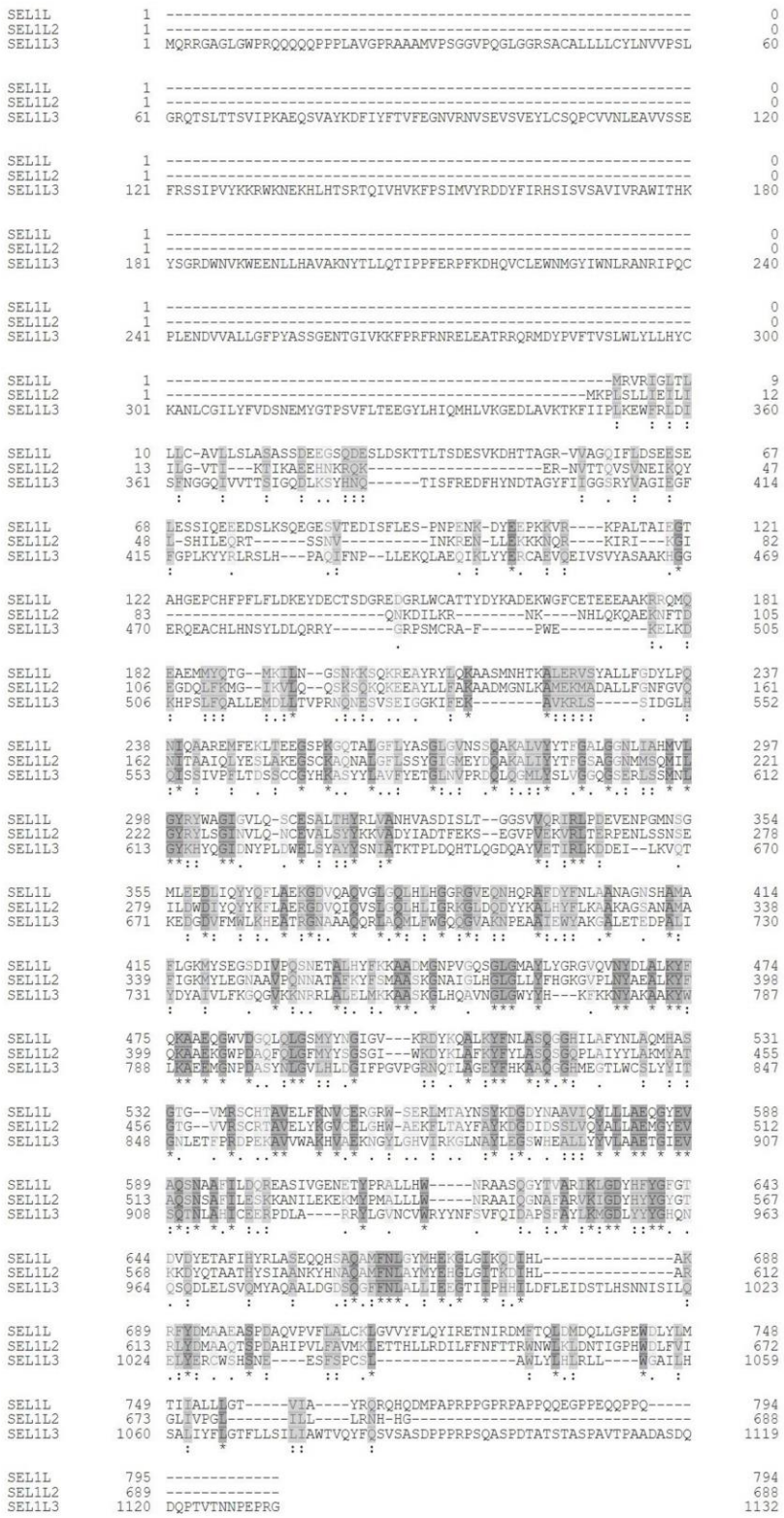


Figure 5. Comparison of the sequence similarity among SEL1L protein family.

Low similarity of SEL1L3 protein with members of SEL1L family, SEL1L and SEL1L2. Amino acid sequence of SEL1L3 (Q68CR1) was extracted from the Human

Peptide Atlas (http://www.peptideatlas.org/builds/human/). Analysis was performed using the Universal Protein Resource (UniProt, https://www.uniprot.org/align/). *: conserved residues; : conservative substitutions; • : semi-conservative substitutions.

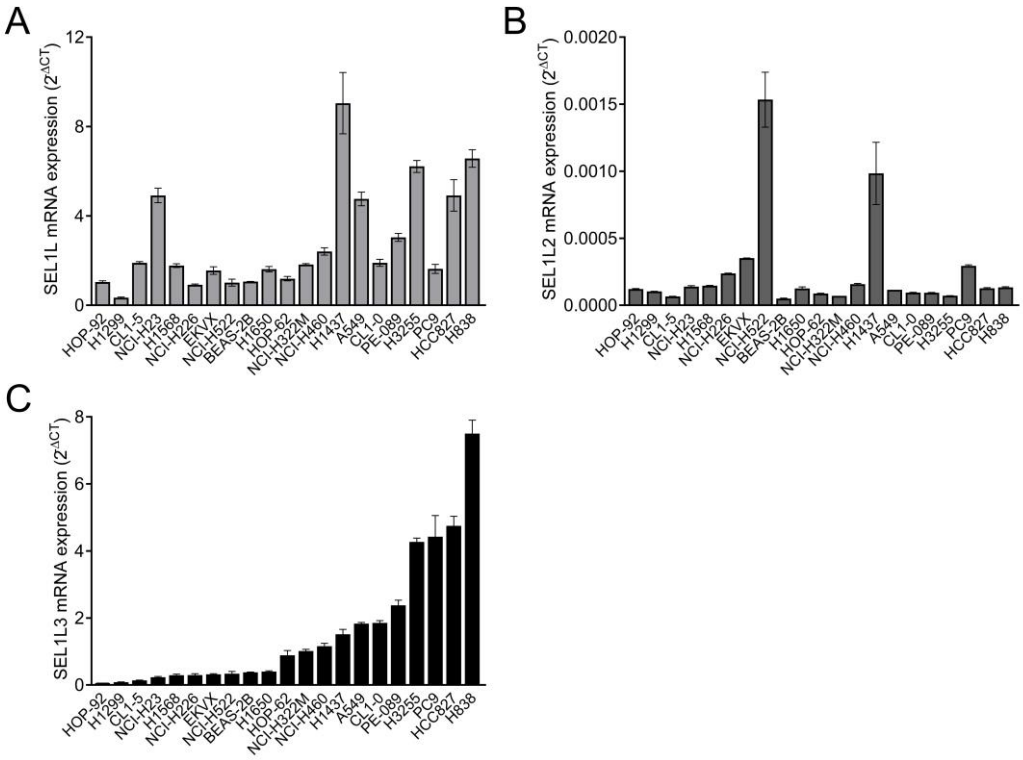


Figure 6. The mRNA expression levels of SEL1L family members.

The mRNA expression levels of SEL1L (A), SEL1L2 (B), and SEL1L3 (C) were detected in 21 lung cancer cell lines, A549, BEAS-2B, CL1-0, CL1-5, EKVX, HCC827, HOP-62, HOP-92, H838, H1299, H1437, H1568, H1650, H3255, NCI-H23, NCI-H226, NCI-H322M, NCI-H460, NCI-H522, PC-9, and PE-089 by RT-qPCR (n = 3).

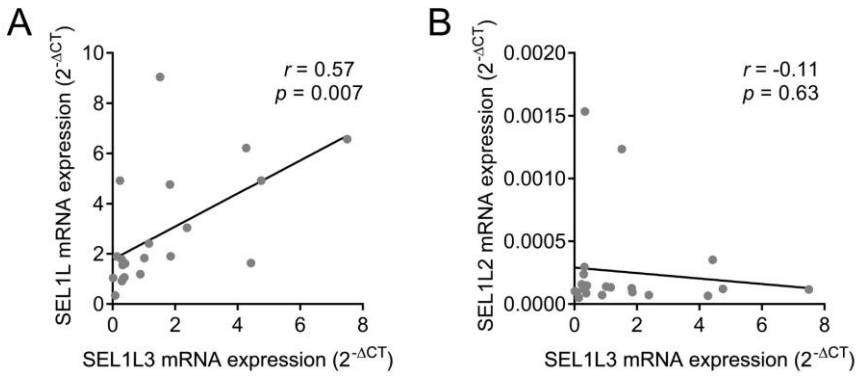


Figure 7. The transcriptional coregulation of SEL1L3 with SEL1L family members.

The mRNA expression levels of SEL1L, SEL1L2, and SEL1L3 were detected in 21 lung cancer cell lines, A549, BEAS-2B, CL1-0, CL1-5, EKVX, HCC827, HOP-62, HOP-92, H838, H1299, H1437, H1568, H1650, H3255, NCI-H23, NCI-H226, NCI-H322M, NCI-H460, NCI-H522, PC-9, and PE-089 by RT-qPCR (n = 3). The correlation of SEL1L3 and SEL1L is shown in left panel and that of SEL1L3 and SEL1L2 in right panel. *r*: Pearson's rank correlation coefficient.

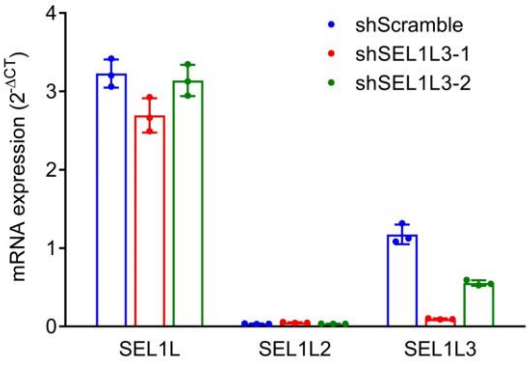


Figure 8. The effect of SEL1L3 knockdown on the expression of SEL1L and SEL1L2.

The mRNA expressions of SEL1L, SEL1L2, and SEL1L3 in shRNA scramble control (shScramble) and two SEL1L3-silencing clones (shSEL1L3-1, shSEL1L3-2) were assayed by RT-qPCR (mean ± SD). All the clones were established in A549 cells.

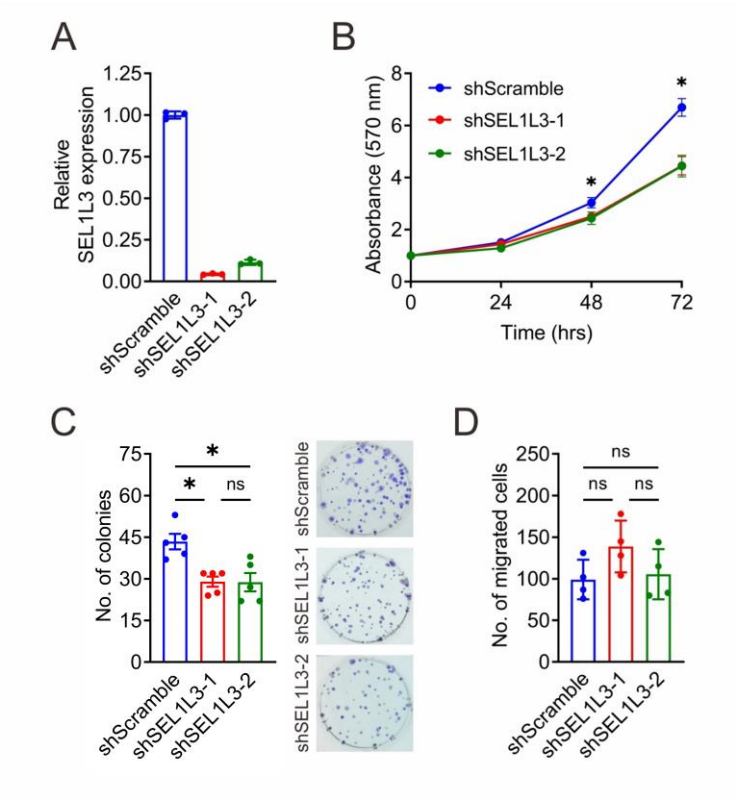


Figure 9. Cancerous phenotypic characterization of SEL1L3.

(A) The knockdown efficiency by SEL1L3-specific shRNAs in A549 cells was assayed by RT-qPCR (mean ± SD, n = 3). (B) SEL1L3 knockdown inhibited cell proliferation. The cell proliferation of A549 shRNA scramble control and SEL1L3-knockdown clones (shSEL1L3-1, shSEL1L3-2) was assessed by MTT assay (mean ± SD, n = 8, *$p < 0.05$ by one-way ANOVA and Tukey's multiple comparisons tests). (C) SEL1L3 knockdown inhibited tumor growth *in vitro*. Colonization ability of A549 shRNA scramble control and SEL1L3-knockdown clones was evaluated by anchorage-dependent colony formation assay. Left: quantification of colonies stained by crystal violet. Colonies with a diameter greater than 1 mm were counted by phase microscopy (mean ± SEM, n = 5, *$p < 0.05$ by one-way ANOVA and Tukey's multiple comparisons tests). Right: representative images of colonies at day 7. (D) Effect of SEL1L3 knockdown on cell migration. Migration ability of A549 shRNA scramble control and SEL1L3-knockdown

clones was measured by transwell migration assay (mean ± SD, n = 4).

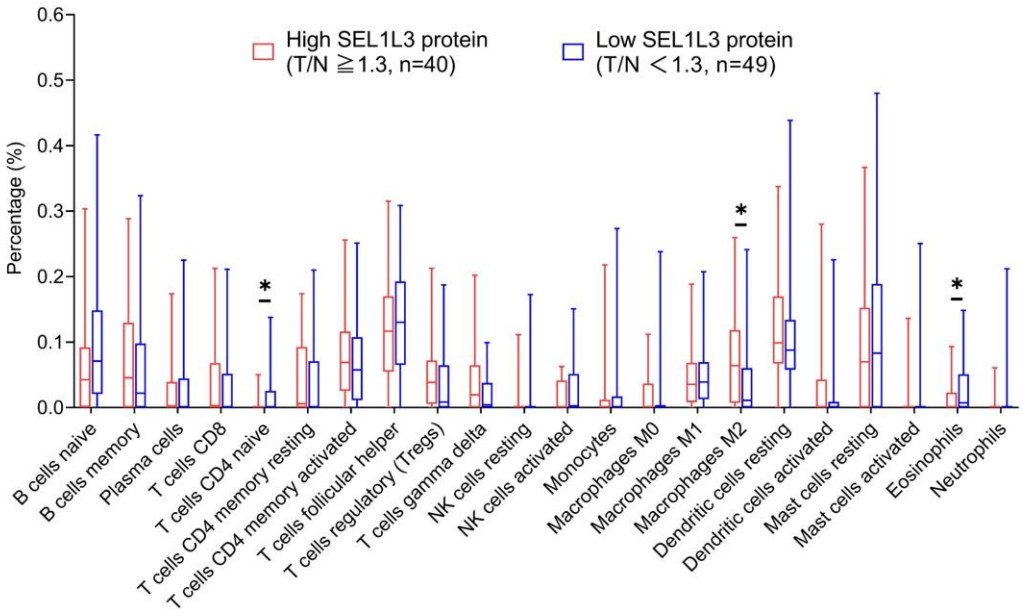


Figure 10. Estimation of immune profiles in relation to SEL1L3 levels by CIBERTSORTx.

Analysis was performed with proteomic data from Taiwan Cancer Moonshot. *$p < 0.05$ by Welch's t test.

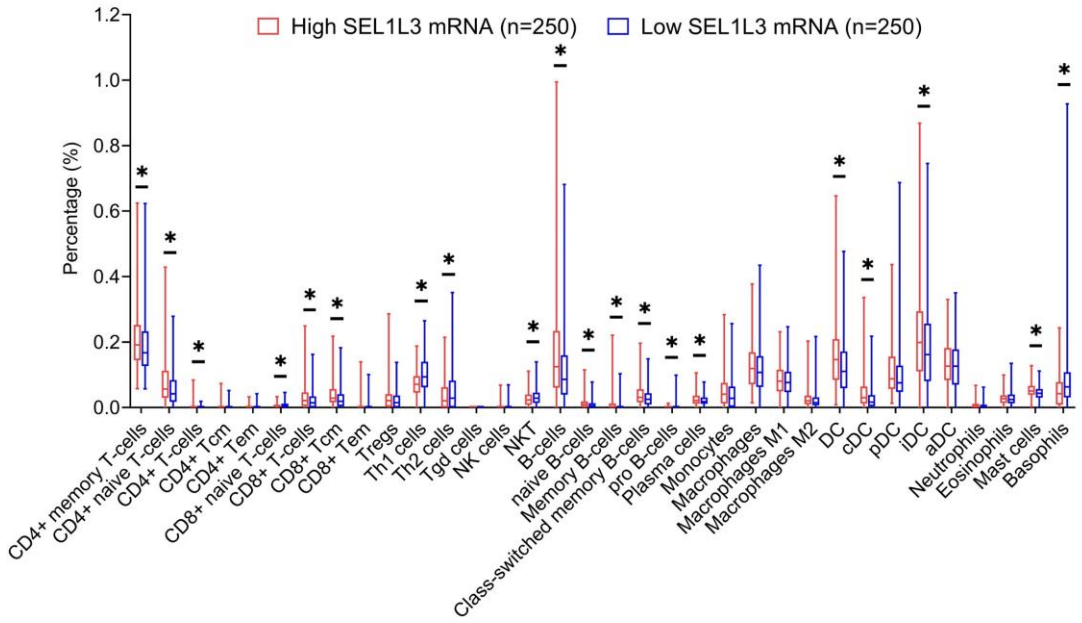


Figure 11. Estimation of immune profiles in relation to SEL1L3 levels by xCell.

Analysis was performed with transcriptomic data from the TCGA lung adenocarcinoma cohort. *$p < 0.05$ by Welch's t test.

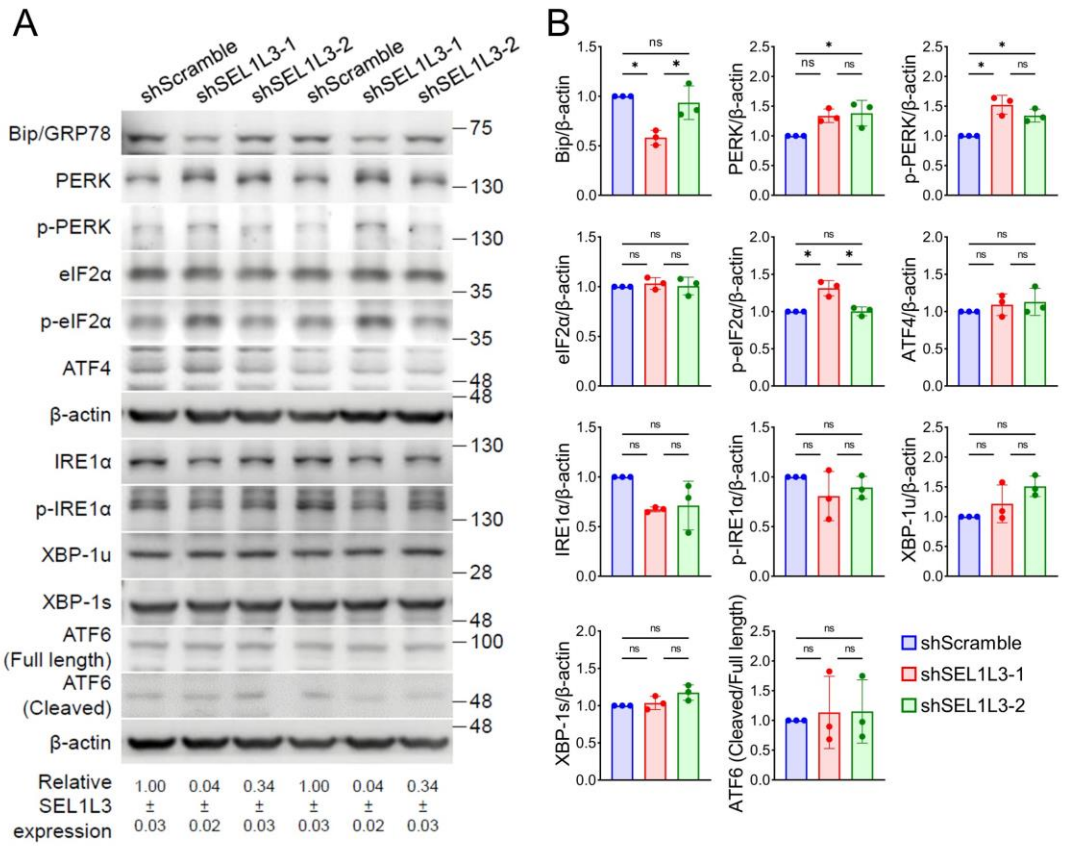


Figure 12. SEL1L3 is involved in ER stress response.

(A) Western blot analysis of the ER stress related mediators in A549 shRNA scramble control and SEL1L3-knockdown cells. The knockdown efficiency by SEL1L3-specific shRNAs in A549 cells was assayed by RT-qPCR (mean ± SD, n = 3). (B) Quantification of the signal intensities in Figure 12A by ImageJ software. (n = 3, *p < 0.05 by one-way ANOVA and Tukey's multiple comparisons tests).

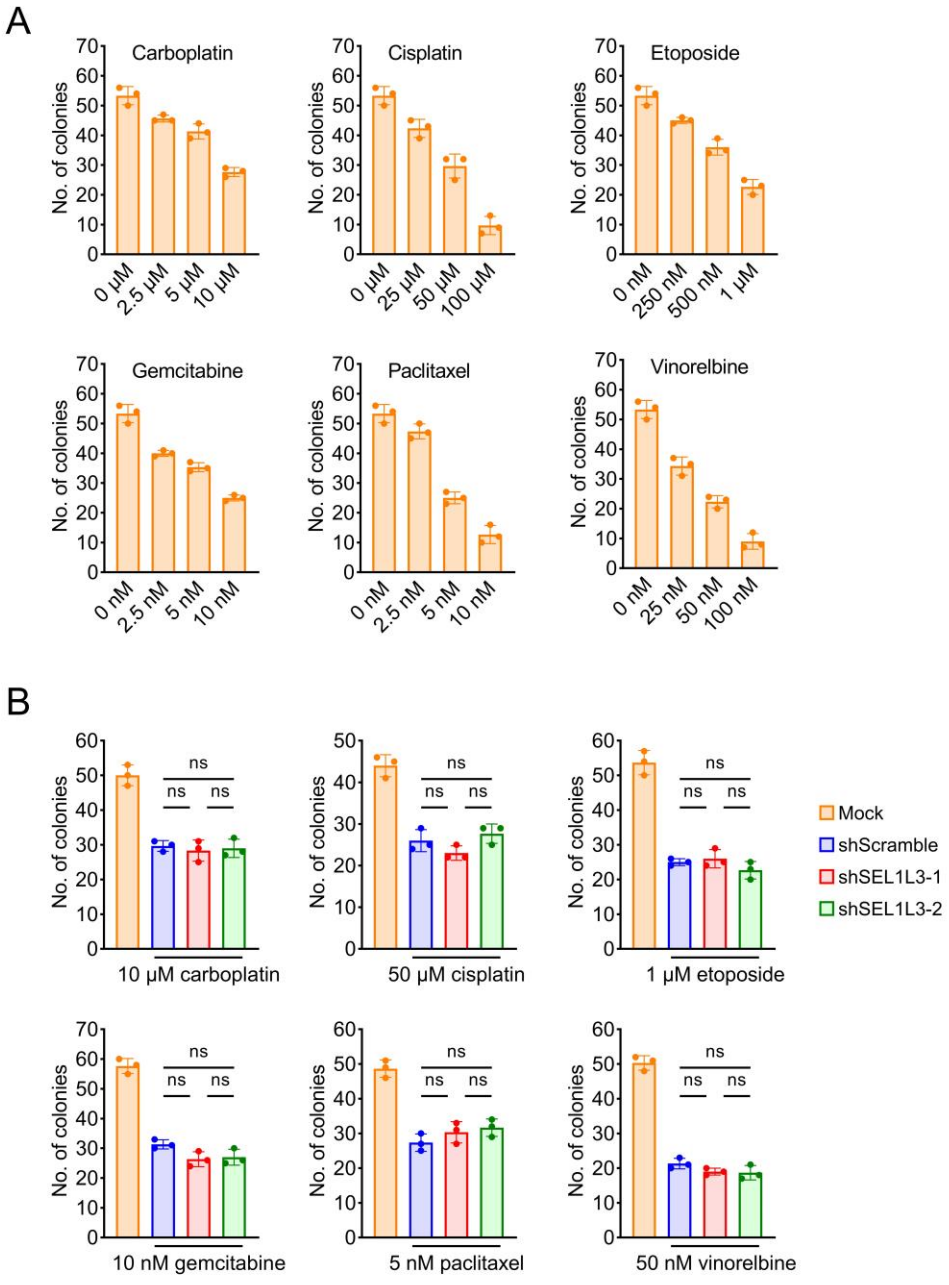


Figure 13. Drug responses upon SEL1L3 knockdown.

(A) Determination of lethal concentration of chemotherapeutic drugs. The cytotoxic effect of 6 chemotherapeutic drugs; carboplatin, cisplatin, etoposide, gemcitabine, paclitaxel, and vinorelbine in A549 shRNA scramble control cells assessed by anchorage-dependent colony formation assays. The treatment dosage of different therapeutic compounds is indicated below the respective figures (n = 3). (B) The

response of A549 shRNA scramble control and SEL1L3-knockdown clones to 6 chemotherapeutic drugs. Colony number of A549 shRNA scramble control and SEL1L3-knockdown cells in the indicated treatment conditions were evaluated by anchorage-dependent colony formation assay. (n = 3, $*p < 0.05$ by one-way ANOVA and Tukey's multiple comparisons tests).

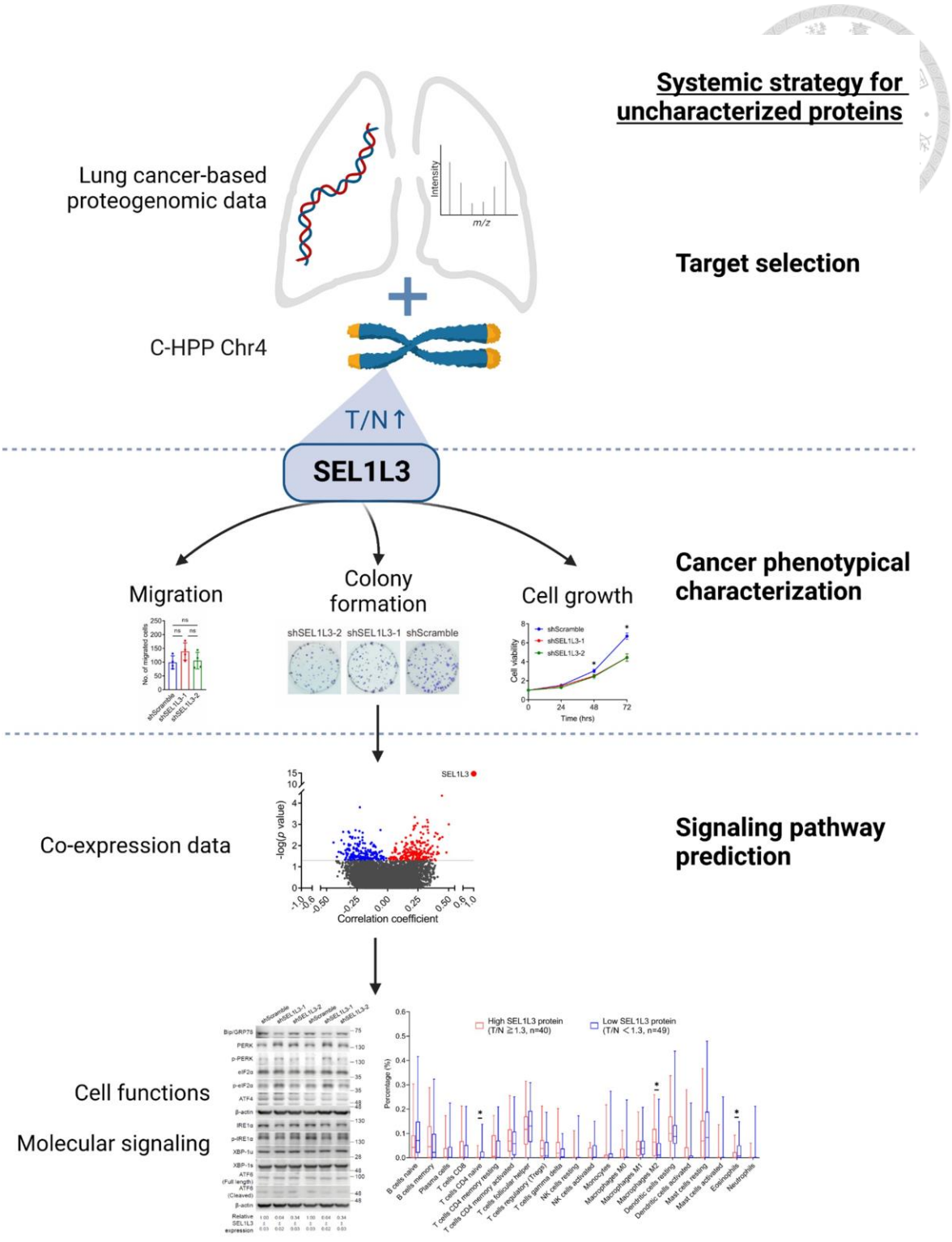


Figure 14. Schematic diagram of the systemic strategy for uncharacterized proteins.

Tables

Table 1. Clinicopathologic characteristics of 89 lung adenocarcinoma patients

Characteristic	**Total patients**	**T/N ≥ 1.3** Patient No. (%)	**T/N < 1.3** Patient No. (%)	***P* value**
Patient No.	n=89	n=40	n=49	
Age				0.116‡
< 60 years	30	10 (25.0%)	20 (40.8%)	
≥ 60 years	59	30 (75.0%)	29 (59.2%)	
Gender				0.233‡
Male	35	13 (32.5%)	22 (44.9%)	
Female	54	27 (67.5%)	27 (55.1%)	
Smoking status				0.239†
Non-smoker	77	35 (87.5%)	42 (85.7%)	
Ex-smoker	9	5 (12.5%)	4 (8.2%)	
Current smoker	3	0 (0.0%)	3 (6.1%)	
Pathology differentiation*				0.018†
Well	7	4 (10.0%)	3 (6.3%)	
Moderate	57	31 (77.5%)	26 (54.2%)	
Poor	24	5 (12.5%)	19 (39.6%)	
Stage				0.874†
IA	40	17 (42.5%)	23 (46.9%)	
IB	31	15 (37.5%)	16 (32.7%)	
II	6	2 (5.0%)	4 (8.2%)	

III-IV	12	6 (15.0%)	6 (12.2%)	
Refined stage				0.988‡
IA	25	11 (27.5%)	14 (28.6%)	
IB	12	6 (15.0%)	6 (12.2%)	
IA_Late like	15	6 (15.0%)	9 (18.4%)	
IB_Late like	19	9 (22.5%)	10 (20.4%)	
Late	18	8 (20.0%)	10 (20.4%)	
EGFR status				0.374†
Wild-type	13	6 (15.0%)	7 (14.3%)	
L858R	35	18 (45.0%)	17 (34.7%)	
Exon 19 deletion	31	14 (35.0%)	17 (34.7%)	
Others	10	2 (5.0%)	8 (16.3%)	
TP53 status				0.355‡
Wild-type	60	29 (72.5%)	31 (63.3%)	
Mutations	29	11 (27.5%)	18 (36.7%)	
Tumor size				0.515‡
> 2cm	66	31 (77.5%)	35 (71.4%)	
≤ 2cm	23	9 (22.5%)	14 (28.6%)	
Staging N				0.925†
N0	74	33 (82.5%)	41 (83.7%)	
N1	5	2 (5.0%)	3 (6.1%)	
N2	10	5 (12.5%)	5 (10.2%)	
Visceral pleural status				0.413†
P0	69	29 (72.5%)	40 (81.6%)	

P1	10	6 (15.0%)	4 (8.2%)	
P2	6	2 (5.0%)	4 (8.2%)	
P3	2	1 (2.5%)	1 (2.0%)	
NA	2	2 (5.0%)	0 (0.0%)	
Primary tumor location				0.297†
RUL	26	14 (35.0%)	12 (24.5%)	
LUL	22	11 (27.5%)	11 (22.4%)	
RML	6	3 (7.5%)	3 (6.1%)	
RLL	17	8 (20.0%)	9 (18.4%)	
LLL	18	4 (10.0%)	14 (28.6%)	
APOBEC signature				0.839‡
Yes	39	18 (45.0%)	21 (42.9%)	
No	50	22 (55.0%)	28 (57.1%)	
Proteomic subtype*				0.604‡
1	50	22 (61.1%)	28 (59.6%)	
2	11	6 (16.7%)	5 (10.6%)	
3	22	8 (22.2%)	14 (29.8%)	
Phosphorylation subtype*				0.524†
1	15	5 (16.7%)	10 (22.7%)	
2	27	10 (33.3%)	17 (38.6%)	
3	7	2 (6.7%)	5 (11.4%)	
4	25	13 (43.3%)	12 (27.3%)	
Recurrence				0.048†
Recurrence	17	4 (10.0%)	13 (26.5%)	

Non-recurrence	72	36 (90.0%)	36 (73.5%)

‡Chi-square test

†Fisher's exact test

*Some patients without information

P0: lack of pleural invasion beyond the elastic layer, P1: invasion beyond the elastic layer, P2: invasion to the surface of the visceral pleura, P3: invasion of the parietal pleura, RUL: right upper lobe, LUL: left upper lobe, RML: right middle lobe, RLL: right lower lobe; LLL: left lower lobe

References

1. *Finishing the euchromatic sequence of the human genome.* Nature, 2004. **431**(7011): p. 931-45.
2. Omenn, G.S., et al., *Progress Identifying and Analyzing the Human Proteome: 2021 Metrics from the HUPO Human Proteome Project.* Journal of Proteome Research, 2021. **20**(12): p. 5227-5240.
3. Lander, E.S., et al., *Initial sequencing and analysis of the human genome.* Nature, 2001. **409**(6822): p. 860-921.
4. Schneider, V.A., et al., *Evaluation of GRCh38 and de novo haploid genome assemblies demonstrates the enduring quality of the reference assembly.* Genome Res, 2017. **27**(5): p. 849-864.
5. Nurk, S., et al., *The complete sequence of a human genome.* Science, 2022. **376**(6588): p. 44-53.
6. Omenn, G.S., et al., *The 2022 Report on the Human Proteome from the HUPO Human Proteome Project.* Journal of Proteome Research, 2022.
7. Liu, Y., A. Beyer, and R. Aebersold, *On the Dependency of Cellular Protein Levels on mRNA Abundance.* Cell, 2016. **165**(3): p. 535-50.
8. *A gene-centric human proteome project: HUPO--the Human Proteome organization.* Mol Cell Proteomics, 2010. **9**(2): p. 427-9.
9. Adhikari, S., et al., *A high-stringency blueprint of the human proteome.* Nature Communications, 2020. **11**(1): p. 5301.
10. Herbst, R.S., D. Morgensztern, and C. Boshoff, *The biology and management of non-small cell lung cancer.* Nature, 2018. **553**(7689): p. 446-454.
11. Siegel, R.L., et al., *Cancer statistics, 2022.* CA Cancer J Clin, 2022. **72**(1): p. 7-33.

12. Zahn-Zabal, M., et al., *The neXtProt knowledgebase in 2020: data, tools and usability improvements.* Nucleic Acids Research, 2019. **48**(D1): p. D328-D334.

13. Paik, Y.K., et al., *The Chromosome-Centric Human Proteome Project for cataloging proteins encoded in the genome.* Nat Biotechnol, 2012. **30**(3): p. 221-3.

14. Kitata, R.B., et al., *Mining Missing Membrane Proteins by High-pH Reverse-Phase StageTip Fractionation and Multiple Reaction Monitoring Mass Spectrometry.* Journal of Proteome Research, 2015. **14**(9): p. 3658-3669.

15. Kitata, R.B., et al., *A data-independent acquisition-based global phosphoproteomics system enables deep profiling.* Nature Communications, 2021. **12**(1): p. 2539.

16. Chen, Y.J., et al., *Proteogenomics of Non-smoking Lung Cancer in East Asia Delineates Molecular Signatures of Pathogenesis and Progression.* Cell, 2020. **182**(1): p. 226-244 e17.

17. Bhattacharya, A. and L. Qi, *ER-associated degradation in health and disease - from substrate to organism.* J Cell Sci, 2019. **132**(23).

18. Miyanaga, A., et al., *Whole-exome and RNA sequencing of pulmonary carcinoid reveals chromosomal rearrangements associated with recurrence.* Lung Cancer, 2020. **145**: p. 85-94.

19. Zhang, C. and C. Ge, *A Simple Competing Endogenous RNA Network Identifies Novel mRNA, miRNA, and lncRNA Markers in Human Cholangiocarcinoma.* Biomed Res Int, 2019. **2019**: p. 3526407.

20. Chu, Y.W., et al., *Selection of invasive and metastatic subpopulations from a human lung adenocarcinoma cell line.* Am J Respir Cell Mol Biol, 1997. **17**(3): p. 353-60.

21. Huang, C.Y., et al., *Inhibition of Alternative Cancer Cell Metabolism of EGFR Mutated Non-Small Cell Lung Cancer Serves as a Potential Therapeutic Strategy.* Cancers (Basel), 2020. **12**(1).

22. Lánczky, A. and B. Győrffy, *Web-Based Survival Analysis Tool Tailored for Medical Research (KMplot): Development and Implementation.* J Med Internet Res, 2021. **23**(7): p. e27633.

23. Uhlén, M., et al., *Proteomics. Tissue-based map of the human proteome.* Science, 2015. **347**(6220): p. 1260419.

24. Chen, C.H., et al., *HLJ1 is an endogenous Src inhibitor suppressing cancer progression through dual mechanisms.* Oncogene, 2016. **35**(43): p. 5674-5685.

25. Xu, J., et al., *Targeting the insulin-like growth factor-1 receptor in MTAP-deficient renal cell carcinoma.* Signal Transduct Target Ther, 2019. **4**: p. 2.

26. Jumper, J., et al., *Highly accurate protein structure prediction with AlphaFold.* Nature, 2021. **596**(7873): p. 583-589.

27. Varadi, M., et al., *AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models.* Nucleic Acids Res, 2022. **50**(D1): p. D439-d444.

28. Subramanian, A., et al., *Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles.* Proceedings of the National Academy of Sciences, 2005. **102**(43): p. 15545-15550.

29. Newman, A.M., et al., *Determining cell type abundance and expression from bulk tissues with digital cytometry.* Nat Biotechnol, 2019. **37**(7): p. 773-782.

30. Aran, D., Z. Hu, and A.J. Butte, *xCell: digitally portraying the tissue cellular heterogeneity landscape.* Genome Biol, 2017. **18**(1): p. 220.

31. Orengo, C.A. and J.M. Thornton, *Protein families and their evolution-a structural perspective.* Annu Rev Biochem, 2005. **74**: p. 867-900.

32. Mittl, P.R. and W. Schneider-Brachert, *Sel1-like repeat proteins in signal transduction.* Cell Signal, 2007. **19**(1): p. 20-31.

33. Desiere, F., et al., *The PeptideAtlas project.* Nucleic Acids Research, 2006. **34**(suppl_1): p. D655-D658.

34. Fridman, W.H., et al., *The immune contexture in human tumours: impact on clinical outcome.* Nature Reviews Cancer, 2012. **12**(4): p. 298-306.

35. Bremnes, R.M., et al., *The Role of Tumor-Infiltrating Immune Cells and Chronic Inflammation at the Tumor Site on Cancer Development, Progression, and Prognosis: Emphasis on Non-small Cell Lung Cancer.* Journal of Thoracic Oncology, 2011. **6**(4): p. 824-833.

36. Walter, P. and D. Ron, *The Unfolded Protein Response: From Stress Pathway to Homeostatic Regulation.* Science, 2011. **334**(6059): p. 1081-1086.

37. Avril, T., E. Vauléon, and E. Chevet, *Endoplasmic reticulum stress signaling and chemotherapy resistance in solid cancers.* Oncogenesis, 2017. **6**(8): p. e373-e373.

38. Chen, X. and J.R. Cubillos-Ruiz, *Endoplasmic reticulum stress signals in the tumour and its microenvironment.* Nature Reviews Cancer, 2021. **21**(2): p. 71-88.

39. Clarke, Hanna J., et al., *Endoplasmic Reticulum Stress in Malignancy.* Cancer Cell, 2014. **25**(5): p. 563-573.

40. Giampietri, C., et al., *Cancer Microenvironment and Endoplasmic Reticulum Stress Response.* Mediators of inflammation, 2015. **2015**: p. 417281-417281.

41. Spencer, B.G. and J.W. Finnie, *The Role of Endoplasmic Reticulum Stress in*

Cell Survival and Death. Journal of Comparative Pathology, 2020. **181**: p. 86-91.

42. Yadav, R.K., et al., *Endoplasmic reticulum stress and cancer.* Journal of cancer prevention, 2014. **19**(2): p. 75-88.

43. Walter, P. and D. Ron, *The unfolded protein response: from stress pathway to homeostatic regulation.* Science, 2011. **334**(6059): p. 1081-6.

44. Oakes, S.A., *Endoplasmic Reticulum Stress Signaling in Cancer Cells.* The American journal of pathology, 2020. **190**(5): p. 934-946.

45. Mohamed, E., Y. Cao, and P.C. Rodriguez, *Endoplasmic reticulum stress regulates tumor growth and anti-tumor immunity: a promising opportunity for cancer immunotherapy.* Cancer immunology, immunotherapy : CII, 2017. **66**(8): p. 1069-1078.

46. Bahar, E., J.-Y. Kim, and H. Yoon, *Chemotherapy Resistance Explained through Endoplasmic Reticulum Stress-Dependent Signaling.* Cancers, 2019. **11**(3): p. 338.

47. Iida, Y., et al., *SEL1L protein critically determines the stability of the HRD1-SEL1L endoplasmic reticulum-associated degradation (ERAD) complex to optimize the degradation kinetics of ERAD substrates.* The Journal of biological chemistry, 2011. **286**(19): p. 16929-16939.

48. Hwang, J. and L. Qi, *Quality Control in the Endoplasmic Reticulum: Crosstalk between ERAD and UPR pathways.* Trends in biochemical sciences, 2018. **43**(8): p. 593-605.

49. Read, A. and M. Schröder, *The Unfolded Protein Response: An Overview.* Biology (Basel), 2021. **10**(5).

50. Rodvold, J.J., et al., *Intercellular transmission of the unfolded protein response*

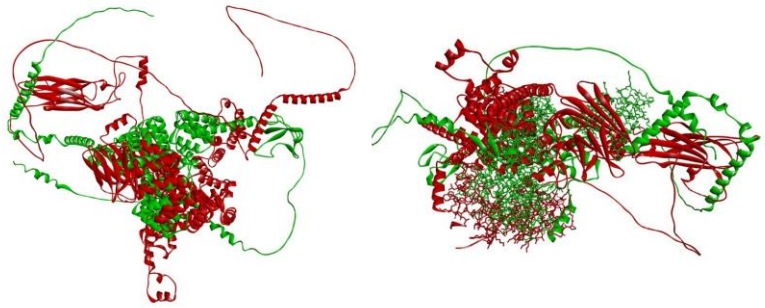
promotes survival and drug resistance in cancer cells. Sci Signal, 2017. **10**(482).

51. Mahadevan, N.R., et al., *Cell-extrinsic effects of tumor ER stress imprint myeloid dendritic cells and impair CD8⁺ T cell priming.* PLoS One, 2012. **7**(12): p. e51845.
52. Mahadevan, N.R. and M. Zanetti, *Tumor Stress Inside Out: Cell-Extrinsic Effects of the Unfolded Protein Response in Tumor Cells Modulate the Immunological Landscape of the Tumor Microenvironment.* The Journal of Immunology, 2011. **187**(9): p. 4403.
53. Rooney, M.S., et al., *Molecular and genetic properties of tumors associated with local immune cytolytic activity.* Cell, 2015. **160**(1-2): p. 48-61.
54. Gibney, G.T., L.M. Weiner, and M.B. Atkins, *Predictive biomarkers for checkpoint inhibitor-based immunotherapy.* The Lancet. Oncology, 2016. **17**(12): p. e542-e551.
55. Lawrence, M.S., et al., *Mutational heterogeneity in cancer and the search for new cancer-associated genes.* Nature, 2013. **499**(7457): p. 214-218.
56. Bilgin, B., et al., *An update on immunotherapy options for urothelial cancer.* Expert Opin Biol Ther, 2019. **19**(12): p. 1265-1274.
57. Liontos, M., et al., *DNA damage, tumor mutational load and their impact on immune responses against cancer.* Ann Transl Med, 2016. **4**(14): p. 264.
58. Yi, M., et al., *The role of neoantigen in immune checkpoint blockade therapy.* Exp Hematol Oncol, 2018. **7**: p. 28.
59. Chabanon, R.M., et al., *Mutational Landscape and Sensitivity to Immune Checkpoint Blockers.* Clinical Cancer Research, 2016. **22**(17): p. 4309.
60. Mei, Y., et al., *A four-gene signature predicts survival and anti-CTLA4*

immunotherapeutic responses based on immune classification of melanoma. Communications Biology, 2021. **4**(1): p. 383.

61. Zhong, H., et al., *Comprehensive Analysis of a Nine-Gene Signature Related to Tumor Microenvironment in Lung Adenocarcinoma.* Front Cell Dev Biol, 2021. **9**: p. 700607.

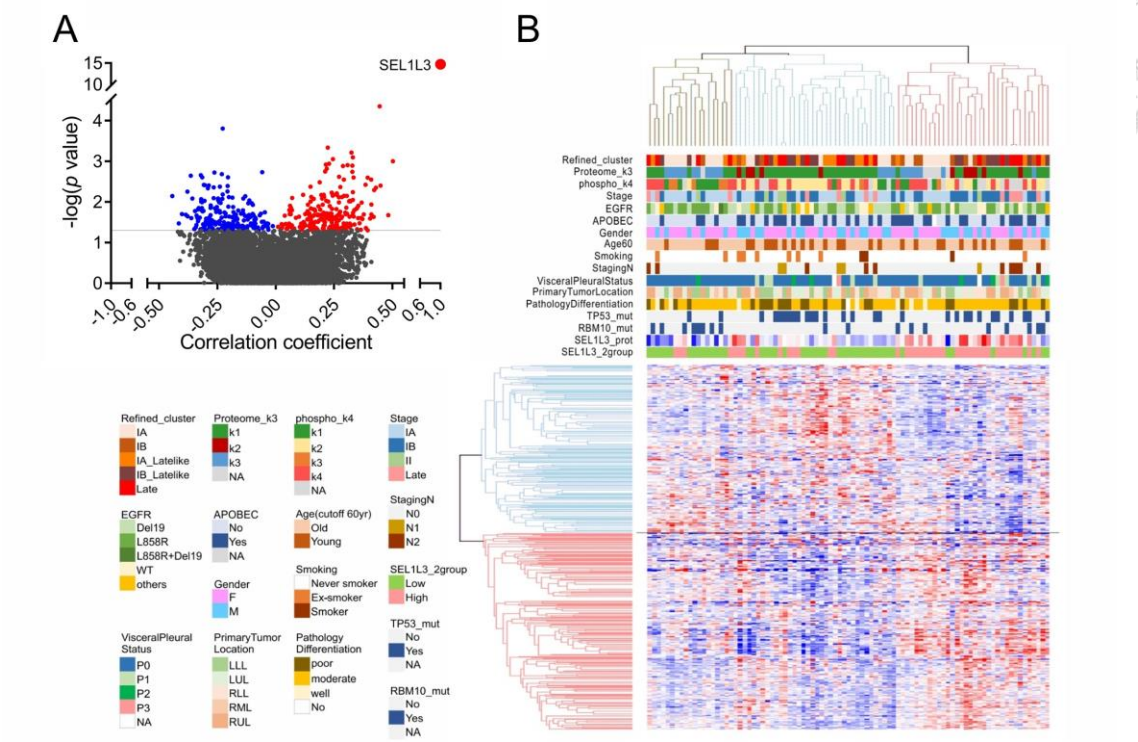
Appendices



Data by Chia-Wei Weng

Appendix 1. Structural similarity between SEL1L and SEL1L3.

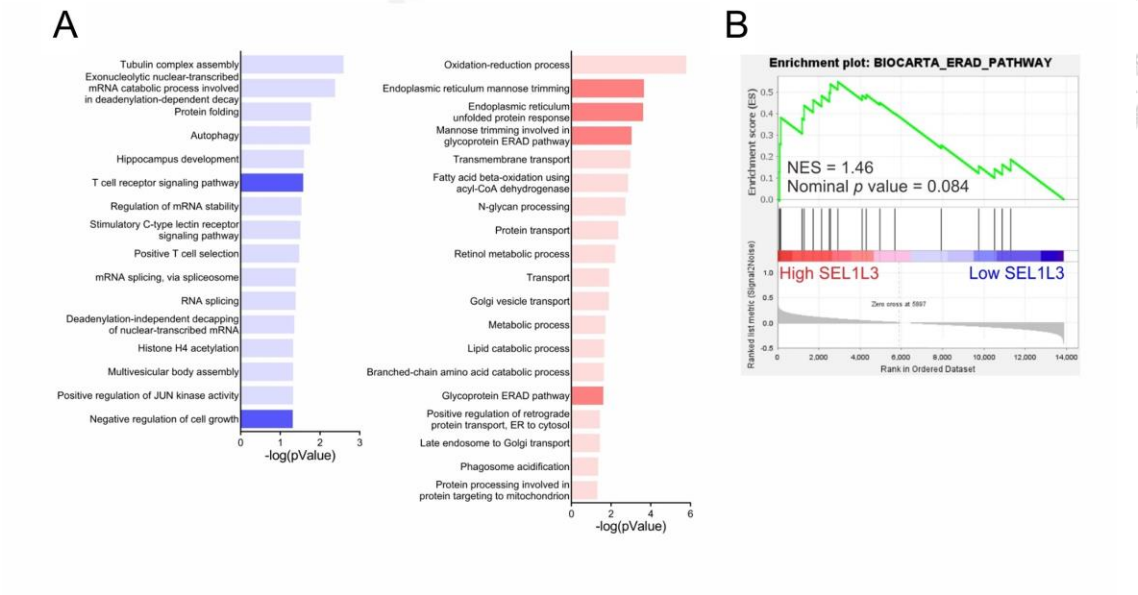
Protein structures of human SEL1L (green) and SEL1L3 (red) exhibited a low percentage in similarity. Structural similarity was determined by superimposing the three-dimensional structures of full-length SEL1L and SEL1L3 (left). 11 Sel1 repeats in SEL1L and 5 in SEL1L3 were shown as sticks (right).



Data by Wen-Hsin Chang and Yi-Ju Chen

Appendix 2. Identification of the differentially expressed proteins.

(A) Volcano plot of the proteins positively (red dots) or negatively (blue dots) correlated with SEL1L3 protein expression. The protein expression association was first analyzed by Pearson correlation. The differentially expressed proteins were then subjected to Welch's t test and considered as significance with p values < 0.05. Gray line: $p = 0.05$ by Welch's t test. (B) Unsupervised hierarchical clustering analysis of the 445 proteins significantly correlated with SEL1L3 protein expression level from Appendix 2A.



Data by Wen-Hsin Chang

Appendix 3. Analysis of SEL1L3-mediated signaling pathways.

(A) Pathway analyses of the upper blue protein cluster for the left panel or the lower red protein cluster for the right panel from Appendix 2B. (B) GSEA of the gene set BIOCARTA_ERAD_PATHWAY between patients with high SEL1L3 protein expression (High SEL1L3, T/N ≥ 1.3) and patients with low SEL1L3 protein expression (Low SEL1L3, T/N < 1.3). NES: normalized enrichment score.

Appendix 4. Publication

Journal of proteome research

pubs.acs.org/jpr Article

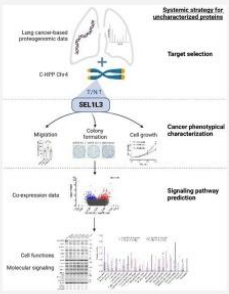
Tissue Proteogenomic Landscape Reveals the Role of Uncharacterized SEL1L3 in Progression and Immunotherapy Response in Lung Adenocarcinoma

Chi-Ya Shen, Wen-Hsin Chang, Yi-Ju Chen, Chia-Wei Weng, Prabha Regmi, Mickiela K. K. Kier, Kang-Yi Su, Gee-Chen Chang, Jin-Shing Chen, Yu-Ju Chen, and Sung-Liang Yu*

Cite This: https://doi.org/10.1021/acs.jproteome.2c00382 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The fundamental pursuit to complete the human proteome atlas and the unmet clinical needs in lung adenocarcinoma have prompted us to study the functional role of uncharacterized proteins and explore their implications in cancer biology. In this study, we characterized SEL1L3, a previously uncharacterized protein encoded from chromosome 4 as a dysregulated protein in lung adenocarcinoma from the large-scale tissue proteogenomics data set established using the cohort of Taiwan Cancer Moonshot. SEL1L3 was expressed in abundance in the tumor parts compared with paired adjacent normal tissues in 90% of the lung adenocarcinoma patients in our cohorts. Moreover, survival analysis revealed the association of SEL1L3 with better clinical outcomes. Intriguingly, silencing of SEL1L3 imposed a reduction in cell viability and activation of ER stress response pathways, indicating a role of SEL1L3 in the regulation of cell stress. Furthermore, the immune profiles of patients with higher SEL1L3 expression were corroborated with its active role in immunophenotype and favorable clinical outcomes in lung adenocarcinoma. Taken together, our study revealed that SEL1L3 might play a vital role in the regulation of cell stress, interaction with cancer cells and the immune microenvironment. Our research findings provide promising insights for further investigation of its molecular signaling network and also suggest SEL1L3 as a potential emerging adjuvant for immunotherapy in lung adenocarcinoma.



KEYWORDS: *Sel-1 suppressor of Lin-12-like 3, nonsmall cell lung cancer, Chromosome-centric Human Proteome Project, uncharacterized protein existence level 1*

■ INTRODUCTION

A genome stores all genetic and hereditary information on an organism, and proteins execute physiological and biochemical reactions in a living cell.[1,2] Both are of great importance to control the molecular mechanisms underlying human health and diseases.[1,2] Since its initial release in 2000, the human reference genome covering the euchromatic portion has been continually improved by the Human Genome Project and Genome Reference Consortium.[3,4] The remaining 8% of the genome has been completed in 2022, propelling the advances in molecular biology and human health.[5] To date, 93.2% of the gene-encoded predicted proteins have been characterized as protein existence level 1 (PE1) with the continued progress in human proteome (Omenn, G. S. et al., *J. Proteome Res.***2022**, in press). The number of PE2, PE3, and PE4 missing proteins (MPs) without protein-level evidence has been reduced to 6.8% (Omenn, G. S. et al., *J. Proteome Res.***2022**, in press). The remarkable progress on the completion of the human proteome may lay the foundation for the development of protein-driven precision medicine for disease management.[2] Moreover, numerous evidence supports the relatively low correlation between proteome and transcriptome, indicating the necessity to complete the human proteome atlas.[6] The Human Proteome Project (HPP) initiated by the Human Proteome Organization aims to map the entire human proteome in a systematic effort to uncover and study the potential roles of human proteome in diseases as well as biological systems.[7,8] Specifically, the chromosome-centric HPP (C-HPP) aims to identify MPs that lack sufficient experimental data from mass spectrometry or other experimental methods and characterize the functions of annotated but uncharacterized proteins with protein existence level 1 (uPE1). Both steps are beneficial to the completion of the human proteome.

Lung cancer remains the leading cause of cancer-related death despite the advances in early detection such as low-dose computed tomography and antitumor treatments, for example, targeted therapy and immunotherapy.[9,10] The main obstacles to ameliorating the high mortality of lung cancer include

Received: June 29, 2022
Published: November 9, 2022