





Article

A Multi-Omics Network of a Seven-Gene Prognostic Signature for Non-Small Cell Lung Cancer

Qing Ye ^{1,2}, Brianne Falatovich ¹, Salvi Singh ^{1,2}, Alexey V. Ivanov ^{1,3}, Timothy D. Eubank ^{1,4} 
and Nancy Lan Guo ^{1,5,*} 

¹ West Virginia University Cancer Institute, Morgantown, WV 26506, USA; qiy@mix.wvu.edu (Q.Y.); bf00014@mix.wvu.edu (B.F.); ss0083@mix.wvu.edu (S.S.); aivanov@hsc.wvu.edu (A.V.I.); tdeubank@hsc.wvu.edu (T.D.E.)

² Lane Department of Computer Science and Electrical Engineering, West Virginia University, Morgantown, WV 26506, USA

³ Department of Biochemistry, School of Medicine, West Virginia University, Morgantown, WV 26506, USA

⁴ Department of Microbiology, Immunology and Cell Biology, School of Medicine, West Virginia University, Morgantown, WV 26506, USA

⁵ Department of Occupational and Environmental Health Sciences, School of Public Health, West Virginia University, Morgantown, WV 26506, USA

* Correspondence: lguo@hsc.wvu.edu; Tel.: +1-304-293-6455

Abstract: There is an unmet clinical need to identify patients with early-stage non-small cell lung cancer (NSCLC) who are likely to develop recurrence and to predict their therapeutic responses. Our previous study developed a qRT-PCR-based seven-gene microfluidic assay to predict the recurrence risk and the clinical benefits of chemotherapy. This study showed it was feasible to apply this seven-gene panel in RNA sequencing profiles of The Cancer Genome Atlas (TCGA) NSCLC patients ($n = 923$) in randomly partitioned feasibility-training and validation sets ($p < 0.05$, Kaplan–Meier analysis). Using Boolean implication networks, DNA copy number variation-mediated transcriptional regulatory network of the seven-gene signature was identified in multiple NSCLC cohorts ($n = 371$). The multi-omics network genes, including *PD-L1*, were significantly correlated with immune infiltration and drug response to 10 commonly used drugs for treating NSCLC. ZNF71 protein expression was positively correlated with epithelial markers and was negatively correlated with mesenchymal markers in NSCLC cell lines in Western blots. PI3K was identified as a relevant pathway of proliferation networks involving *ZNF71* and its isoforms formulated with CRISPR-Cas9 and RNA interference (RNAi) profiles. Based on the gene expression of the multi-omics network, repositioning drugs were identified for NSCLC treatment.

Keywords: non-small cell lung cancer; prognostic gene signature; multi-omics; CRISPR-Cas9; RNA interference; Boolean implication networks; chemotherapy; immune checkpoint inhibitor; radiotherapy; drug reposition



Citation: Ye, Q.; Falatovich, B.; Singh, S.; Ivanov, A.V.; Eubank, T.D.; Guo, N.L. A Multi-Omics Network of a Seven-Gene Prognostic Signature for Non-Small Cell Lung Cancer. *Int. J. Mol. Sci.* **2022**, *23*, 219. <https://doi.org/10.3390/ijms23010219>

Academic Editors: Amelia Casamassimi, Monica Rienzo, Ciro Abbondanza and Erika Di Zazzo

Received: 9 November 2021

Accepted: 23 December 2021

Published: 25 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Lung cancer has the second highest cancer incidence rate and the top cancer-related mortality worldwide. An estimate from the American Cancer Society shows that, in 2021, there were 235,760 lung cancer cases (119,100 men and 116,660 women) in the US [1]. In the US, 60–65% of newly diagnosed lung cancer cases are never-smokers or former smokers [2]. About 84% of all known lung cancers cases are non-small-cell lung cancer (NSCLC) [3]. Major histological subtypes of NSCLC include lung adenocarcinoma (LUAD, 50% of NSCLC cases), squamous cell carcinoma (LUSC, 30%), and large cell carcinoma (5–10%) [4]. The overall 5-year survival rate after initial diagnosis for NSCLC is less than 15% due to extensive invasion, limited therapeutic response, recurrence, and metastasis [5]. Patients with NSCLCs that are in stage II and above are treated with chemotherapy. Patients in stages

III and IV are treated with additional radiotherapy [6]. Although adjuvant chemotherapy for stage II and stage III disease improves overall survival rate by 10–15% [7], early-stage NSCLC has a dismal prognosis, with a 5-year mortality rate of 40% in stage I, 66% in stage II, and 75% in stage IIIA, as a result of recurrence [8]. These data suggest that many early-stage tumors harbor occult recurrence despite chemotherapy. The FDA recently approved atezolizumab for stage II and IIIA NSCLC after platinum-based chemotherapy in patients with PD-L1 > 1% [9]. There is an unmet clinical need to identify early-stage NSCLC patients who are likely to develop recurrence and to predict the clinical benefits of chemotherapy, with or without subsequent immunotherapy depending on their PD-L1 expression levels.

Our previous study developed a qRT-PCR-based seven-gene microfluidic assay for NSCLC prognosis and prediction of chemotherapeutic benefits, including *ABCC4*, *CCL19*, *SLC39A8*, *CD27*, *FUT7*, *DAG1*, and *ZNF71* [10]. This gene assay can inform selection of specific chemotherapy for individual patients to improve cancer care. Zinc finger protein 71 (*ZNF71*) protein expression quantified with AQUA was identified as a good prognostic marker of NSCLC [10]. As a transcriptional repression domain, the Krüppel-associated box (KRAB) is commonly found in human zinc finger protein-based transcription factors, specifically KRAB zinc finger proteins (KRAB-ZFPs) [11–13]. KRAB-ZFPs-mediated transcriptional repression is involved in cell proliferation, differentiation, apoptosis, and cancer [14]. Emerging evidence links epithelial to mesenchymal transition (EMT) to tumor progression, metastasis, and resistance to cancer therapy [15–18]. We reported that *ZNF71* KRAB isoform was associated with EMT in NSCLC tumors and cell lines [19].

Revelation of genomic and transcriptomic landscape has shed lights to NSCLC subtypes and treatment [20]. An examination of TCGA multi-omics data showed that these seven prognostic genes have more DNA copy number variations (CNV) than mutations in NSCLC tumors (Figures A1–A7). Here, we sought to investigate 1) the feasibility of applying the seven-gene signature for NSCLC prognosis using next-generation sequencing (NGS) data of TCGA; and 2) the potential functional mechanisms involving the molecular networks of these seven genes. In order to understand how the molecular networks enfold these seven prognostic genes, we developed a novel Boolean implication network methodology for modeling whole-genome copy number variation (CNV) and gene expression profiles ($n = 371$) in NSCLC tumors. CRISPR-Cas9 ($n = 78$) and RNAi ($n = 92$) screening data were used to identify proliferation genes in the multi-omics network associated with the seven-gene panel. The identified network genes were further examined for their association with response to radiotherapy in TCGA patient tumors ($n = 966$) and 10 commonly used NSCLC therapeutic regimens in the Cancer Cell Line Encyclopedia (CCLE) NSCLC panel ($n = 117$). In myeloid dendritic cells, macrophages, neutrophils, CD4+ T cells, CD8+ T cells, and B cells, immunological infiltration linked to the discovered genes was studied. The association between mRNA expression and protein expression of these network genes in NSCLC CCLE panel and TCGA tumors was analyzed.

EMT is highly dynamic and reversible. EMT is still unknown and controversial as to its role in cancer outcome [21,22]. There was a reported link between KRAB-ZFPs and EMT [23]. Tripartite motif containing 28 (TRIM28) protein is a universal co-factor for KRAB-ZFP transcription factors [24]. TRIM28 contributes to EMT and might be involved in lung cancer metastasis [25]. We previously reported that *ZNF71* KRAB was associated with EMT and patient outcomes [19]. Nevertheless, the function of *ZNF71* KRAB in EMT is not known. To better understand the molecular mechanisms of *ZNF71* and its isoforms, this study investigated the correlation between *ZNF71* protein expression and EMT markers in NSCLC cell lines. Furthermore, molecular networks formulated with proliferation genes co-expressed with *ZNF71*, *ZNF71* KRAB, and *ZNF71* KRAB-less isoforms in NSCLC tumors were constructed with the Boolean implication network algorithm, respectively. The gene expression of these proliferation networks was further analyzed with L1000 Connectivity Map (CMap) [26,27] to identify matched functional pathways in shRNA knockdown, CRISPR-Cas9 knockout, or overexpression assays in NSCLC cell lines. Finally, the 14-

gene EMT classifier constructed in our previous study for characterizing NSCLC tumor EMT states [19] was combined with the identified multi-omics network of the seven-gene signature to generate mechanistic hypotheses and discover repositioning drugs to reverse EMT, to enhance drug sensitivity, and to inhibit ICIs and NSCLC proliferation. Such a revelation will provide insight for designing therapeutic strategies for NSCLC intervention in future research.

2. Results

2.1. A Multi-Omics Network of the Seven-Gene NSCLC Prognostic and Predictive Signature

Our previous retrospective longitudinal clinical studies developed a seven-gene qRT-PCR microfluidic assay for NSCLC prognostic and prediction of chemotherapeutic benefits [10], including *ABCC4*, *CCL19*, *SLC39A8*, *CD27*, *FUT7*, *DAG1*, and *ZNF71*. Here, we sought to investigate whether it is feasible to apply this seven-gene signature in NGS data for NSCLC prognosis. TCGA-LUAD ($n = 515$) and TCGA-LUSC ($n = 501$) datasets were combined as TCGA NSCLC data ($n = 923$), keeping patient samples with sufficient survival information. The combined TCGA NSCLC dataset was randomly partitioned into a training ($n = 462$) and a testing set ($n = 461$) for building and validating the prognostic model, respectively. A multivariate Cox model was built using the seven prognostic genes in the training set to calculate coefficients for the gene covariates and the risk-score for each patient. This model was applied to the testing set using the same gene coefficients and cutoff for patient stratification. In Kaplan–Meier analysis, the patients who had a risk-score less than -0.74 survived significantly longer than the patients who had a risk-score equal or greater than -0.74 in both training and testing sets ($p = 0.016$, HR: 1.454 [1.072, 1.971] in training set, $p = 0.049$, HR: 1.366 [1.001, 1.864] in the testing set; Figure 1A,B) in Kaplan–Meier analysis. More details were included in Supplementary File S1.

The Boolean implication network approach was used to create a multi-omics association network of *ABCC4*, *CCL19*, *SLC39A8*, *CD27*, *FUT7*, *DAG1*, and *ZNF71*. The following genes were selected to build the network. First, the genome-scale analysis identified genes that had significant ($p < 0.05$, z -tests) co-expression with at least two of the seven original genes in both the GSE31800 and GSE28582 NSCLC patient cohorts. Second, genes with a significant ($p < 0.05$, z -tests) CNV-mediated gene expression (GE) associated with at least two of the seven original genes in both GSE31800 and GSE28582 cohorts were identified. Third, we screened for genes in both GSE31800 and GSE28582 that had a significant ($p < 0.05$, z -tests) co-occurrence of DNA copy number aberrations with at least two of the seven original genes, but no genes were discovered. In addition, proliferation genes were defined as genes with a significant dependency score in 50% of the tested NSCLC cell lines in CRISPR-Cas9 ($n = 78$) or RNAi ($n = 92$). *ERH*, *IFITM3*, *NFS1*, and *KRT17* were identified as NSCLC proliferation genes. The final multi-omics network (Figure 1C) contained genes having an association with at least two of the seven prognostic genes, with each edge indicating an observed co-expression or CNV-mediated gene expression dysregulation in NSCLC tumors in multiple patient cohorts. *PD-L1* (*CD274*) was co-expressed with multiple signature genes. Down-regulation of *CD274* was associated with up-regulation of *DAG1*, and down-regulation of *DAG1* was associated with up-regulation of *CD274*. Up-regulation of *CD274* was associated with up-regulation of *CCL19* and down-regulation of *CD27* (Supplementary File S2 datasheet 2). It was reported that X-ray irradiation induced miR-574-3p expression in NSCLC A549 cells, which, in turn, suppressed enhancer of rudimentary homolog (ERH) protein production and delayed cell growth [28]. Interferon-induced transmembrane protein 3 (IFITM3) protein expression was expressed in the cytoplasm of metastatic lung adenocarcinoma tumor tissues and was correlated with higher tumor grade [29]. Single-cell sequencing profiles of *IFITM3* and *ISG15* characterize an interferon signaling cluster of macrophages in metastasis-bearing lungs [30]. *NFS1* cysteine desulfurase is under positive selection in lung adenocarcinomas and protects cancer cells from ferroptosis, a non-apoptotic form of cell death, in response to oxidative damage [31]. Keratin 17 (*KRT17*), reported to promote lung adenocarcinoma via enhancing cell proliferation

and invasion [32], showed a paradoxical anti-metastatic role of *KRT17* both in vitro and in vivo in cancer xenografts in a single-cell lineage study [33]. These reports in the literature support the relevance of the identified multi-omics network of the seven-gene signature in NSCLC.

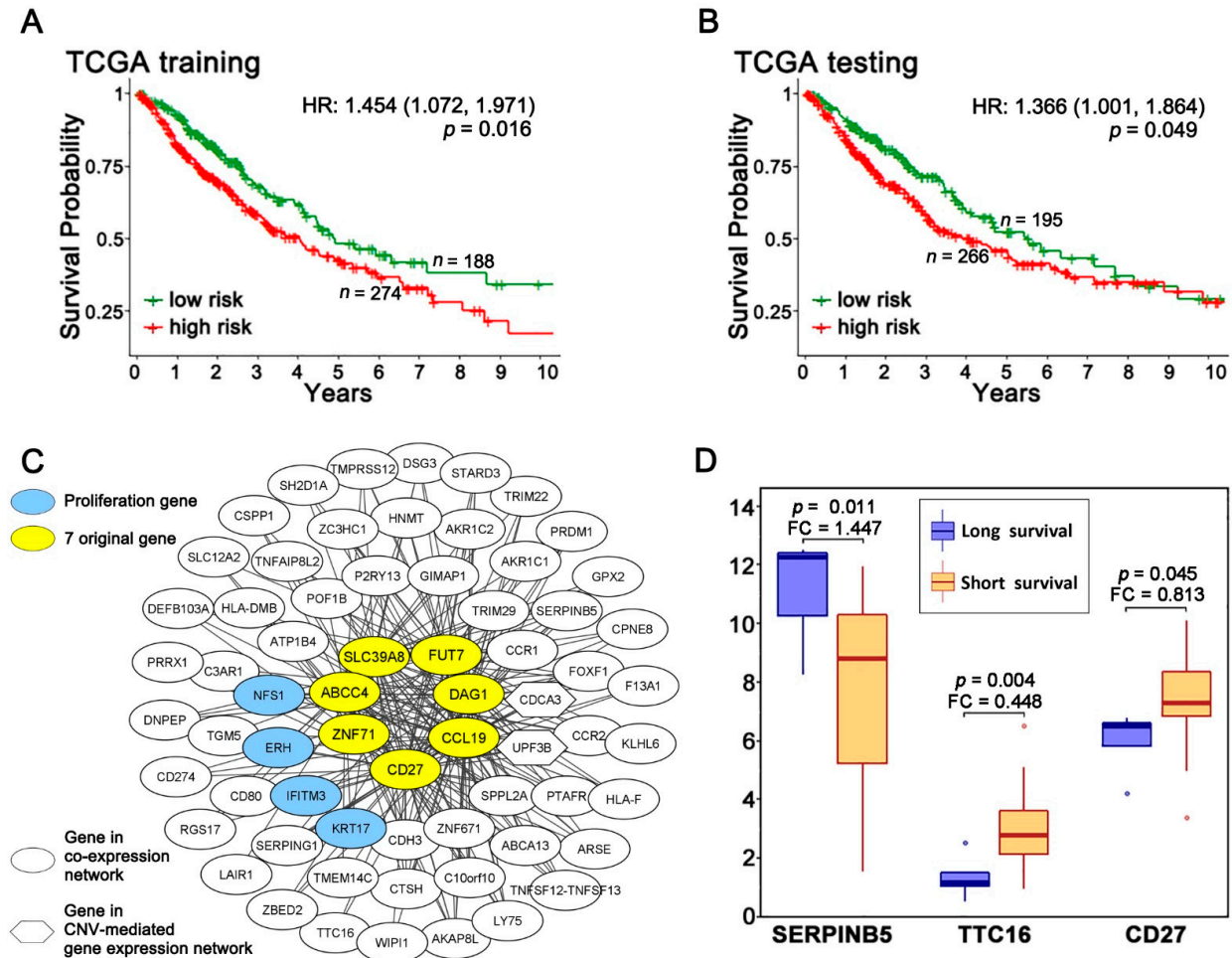


Figure 1. The 7-gene signature and associated multi-omics network. Kaplan–Meier analysis of 7-gene signature in training (A) and testing (B) sets randomly partitioned from the combined TCGA-LUSC and TCGA-LUAD data. Patients were grouped by risk-scores with the cutoff value of -0.74 . The plots show the survival outcome of the first 10 years after surgery. (C) The multi-omics network of the 7-gene signature in NSCLC tumors. All proliferation genes in the networks were identified from CRISPR-Cas9/RNAi screening data in CCLE NSCLC cell lines, with more details provided in Supplementary File S2. (D) Genes associated with radiotherapy response. In TCGA-LUSC and TCGA-LUAD, these genes showed significantly different expression levels ($p < 0.05$, two-sample t -tests) between the long survival group (>58 months; $n = 144$) and the short survival group (<20 months; $n = 186$). Analysis was conducted on stage III or IV patients who had received radiotherapy.

Stage III and IV NSCLC patients receive radiation therapy [6]. To identify genes associated with the response to radiotherapy, TCGA-LUAD and TCGA-LUSC stage III and IV patients who had received radiotherapy were included in the analysis. There was significant differential expression ($p < 0.05$, two sample t -tests) in genes identified in the multi-omics network (Figure 1C), including *SERPINB5*, *TTC16*, and *CD27*. These differences were observed in the short survival (<20 months) and long survival (>58 months) patient groups (Figure 1D). There was a higher level of expression of *CD27* and *TTC16* among the short-survival patients, and higher levels of *SERPINB5* expression among the long-survival patients.

The correlation of the expression of 65 genes (*DEFB103A* was not available) from the final network with immune infiltration in TCGA-LUAD ($n = 515$) and TCGA-LUSC ($n = 501$) patients was assessed with TIMER 2.0 (Figure 2A,B) [34–36]. The immune infiltration analysis included myeloid dendritic cells, macrophages, neutrophils, CD4+ T cells, CD8+ T cells, and B cells. All the available genes in the 7-gene network had a significant correlation with immune infiltration in at least one examined cell types.

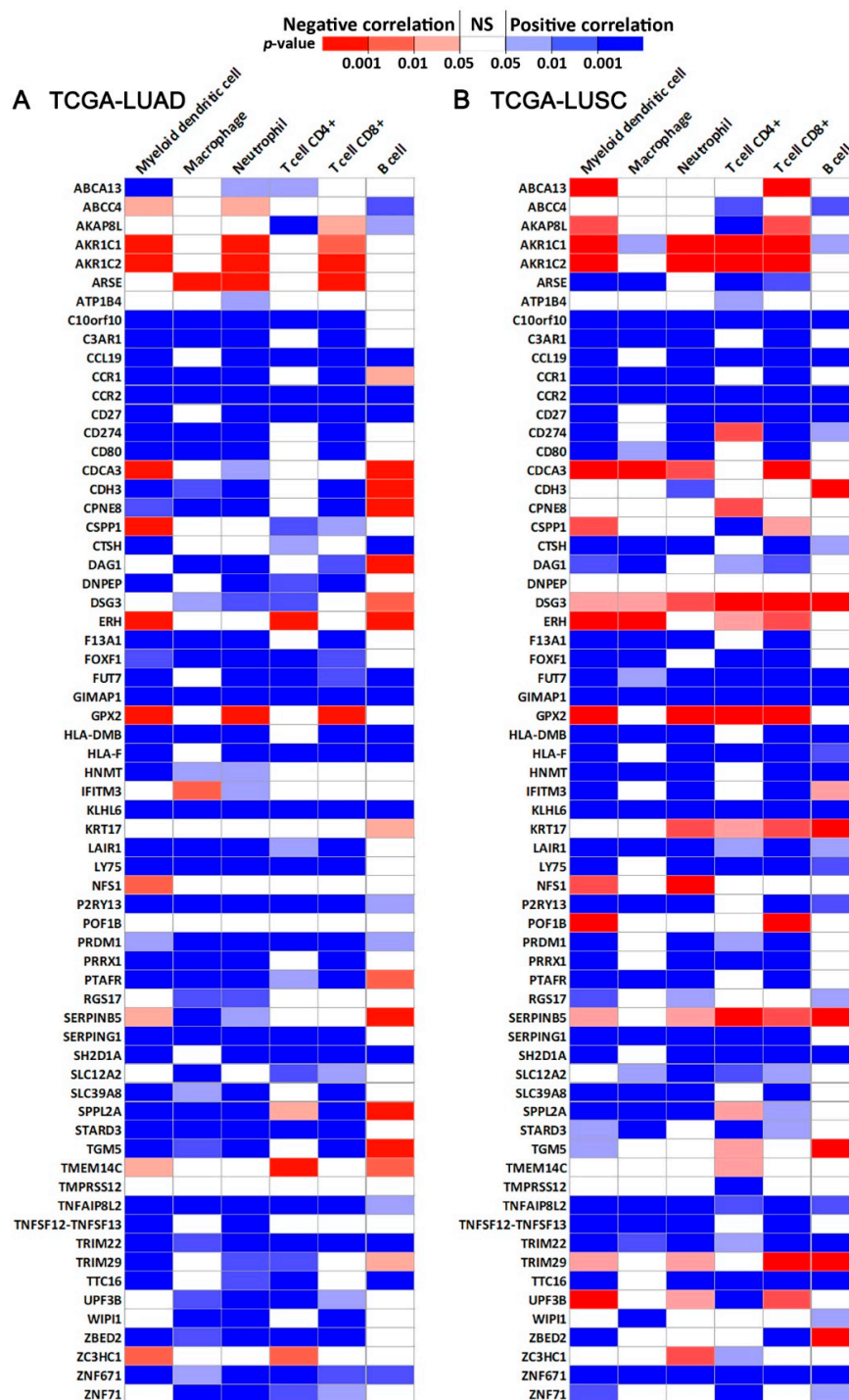


Figure 2. Association between immune infiltration and expression of the multi-omics network genes (Figure 2C). (A) The correlation of gene expression with immune infiltration level in TCGA-LUAD patients ($n = 515$) assessed with TIMER 2.0. (B) The correlation of gene expression with immune infiltration level in TCGA-LUSC patients ($n = 501$) assessed with TIMER 2.0. NS: not statistically significant.

This study also examined the association with drug activities of 10 commonly used therapeutic regimens for treating NSCLC, including carboplatin, cisplatin, paclitaxel (Taxol), pemetrexed, docetaxel, gemcitabine, vinorelbine, etoposide, gefitinib, and erlotinib. There were three histological subtypes for the CCLE NSCLC cell lines: adenocarcinoma, squamous cell carcinoma, large cell carcinoma, and adenosquamous carcinoma. Our study selected genes associated with drug response based on their differential expression of mRNA and protein in sensitive and resistant NSCLC cell lines. In the identified seven-gene network (Figure 1C), numerous genes were associated with drug sensitivity or resistance to the studied drugs in the CCLE NSCLC panel (Tables 1 and 2).

Table 1. Genes (shown in Figure 1C) with a significant differential mRNA expression ($p < 0.05$; two sample t -tests) and a fold change < 0.5 or > 2 in sensitive versus resistant CCLE NSCLC cell lines ($n = 117$) to the studied drugs. Blue font indicates higher mRNA expression in sensitive cell lines, and red font indicates high mRNA expression in resistant cell lines.

Drug	PRISM	GDSC1	GDSC2
Carboplatin	<i>IFITM3, ATP1B4, PRRX1, PRRX1</i>		
Cisplatin	<i>FUT7</i>		<i>POF1B, ZNF671, ZBED2, C10orf10</i>
Docetaxel	<i>ERH, KLHL6</i>	<i>KLHL6, IFITM3, C10orf10</i>	<i>SERPING1</i>
Erlotinib	<i>KLHL6, TRIM29, IFITM3, CCR2, TNFSF12-TNFSF13, ATP1B4</i>	<i>IFITM3, SERPINB5, PRDM1, RGS17, CD80, AKR1C1, TRIM29</i>	<i>DEFB103A, CSPP1, TRIM29, SERPINB5, AKR1C2, KLHL6, ATP1B4, SERPING1, DSG3</i>
Etoposide	<i>PRDM1, TRIM29</i>	<i>ATP1B4, SERPING1, ABCA13, ARSE</i>	
Gefitinib	<i>DEFB103A, PTAFR, ZEB1, ERH, F13A1</i>	<i>DEFB103A, KLHL6, TRIM29, RGS17, ATP1B4, PRRX1</i>	<i>CSPP1, KLHL6, TRIM29, ATP1B4, SERPING1, ARSE</i>
Gemcitabine	<i>ARSE, GPX2, RGS17, LAIR1, TRIM22</i>	<i>SERPING1</i>	<i>ARSE, C10orf10</i>
Paclitaxel	<i>PTAFR, TRIM29, RAB25, ESRP1, SERPINB5, KLHL6, CD80, SH2D1A</i>	<i>CD27</i>	<i>C10orf10</i>
Pemetrexed		<i>SH2D1A</i>	
Vinorelbine	<i>FUT7, TNFSF12-TNFSF13, DSG3</i>		

Table 2. Genes (shown in Figure 1C) with a significant differential protein expression ($p < 0.05$; two sample t -tests) and a fold change < 0.5 or > 2 in sensitive versus resistant CCLE NSCLC cell lines ($n = 63$) to the studied drugs. Blue font indicates higher protein expression in sensitive cell lines, and red font indicates higher protein expression in resistant cell lines.

Drug	PRISM	GDSC1	GDSC2
Carboplatin	<i>TGM5</i>		
Cisplatin	<i>TRIM29</i>	<i>CD274, HLA-F</i>	<i>LY75, PRRX1, F13A1, CTSH</i>
Docetaxel	<i>SERPINB5</i>	<i>AKR1C2, SERPINB5</i>	<i>AKR1C2, AKR1C1, CPNE8</i>
Erlotinib	<i>ARSE, HLA-DMB, TRIM22, LY75, TGM5</i>	<i>AKR1C1</i>	<i>POF1B, SERPINB5, F13A1</i>
Etoposide	<i>HLA-F, ABCA13, TGM5</i>	<i>SERPINB5</i>	
Gefitinib		<i>SLC39A8</i>	
Gemcitabine	<i>TGM5, HNMT, CTSH, HLA-F, SERPINB5, CDH3, HNMT, ABCA13, AKR1C2, GPX2, AKR1C1</i>	<i>SERPINB5, AKR1C1</i>	<i>AKR1C1</i>
Paclitaxel	<i>SERPINB5, CDH3, HLA-F, SLC39A8</i>		<i>PRRX1, LY75, CTSH, TRIM29</i>
Pemetrexed			
Vinorelbine		<i>SERPINB5, AKR1C2, AKR1C1, F13A1, HLA-F</i>	<i>CTSH</i>

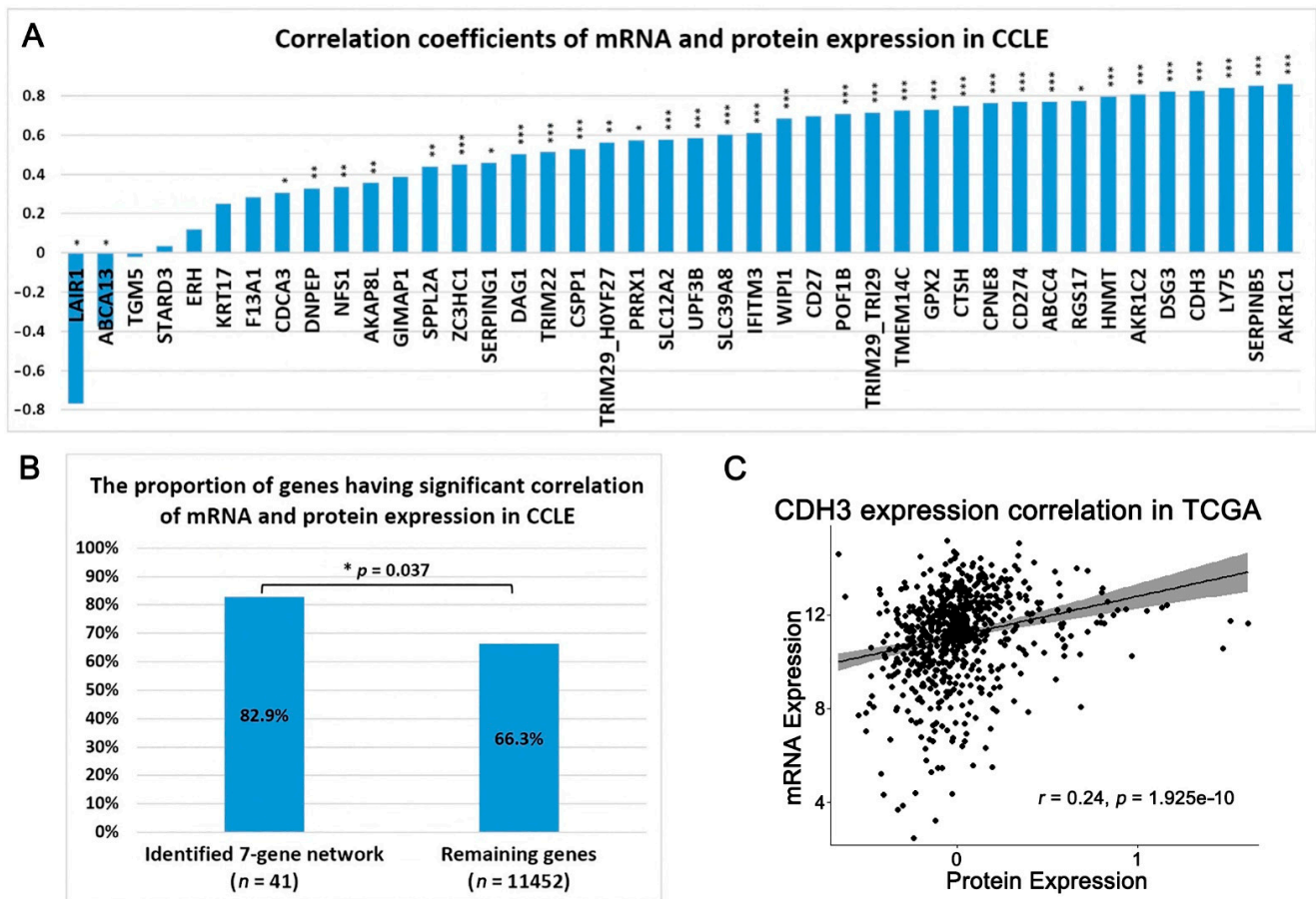


Figure 3. The correlation between mRNA and protein expression of genes in the multi-omics network (shown in Figure 1C). (A) The bar chart shows the correlation coefficients of mRNA and protein expression from the same gene in NSCLC cell lines. The asterisk (*) indicates the significance level (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). (B) The proportion of genes having a significant correlation of mRNA and protein expression in NSCLC cell lines ($p < 0.05$, Pearson's correlation). The proportion of genes having a significant correlation between mRNA and protein expression is significantly higher in identified 7-gene network than the remaining genes in the genome (two-tailed proportion z-test, $p < 0.05$). (C) CDH3 had a significant positive correlation between mRNA and protein expression in combined TCGA-LUSC and TCGA-LUAD patient tumors ($n = 685$).

Next, we examined the correlation between mRNA expression and protein expression of the seven-gene network in NSCLC cell lines and patient tumors. In CCLE NSCLC cell lines ($n = 63$), 41 genes in the seven-gene network had both mRNA and protein expression measurements. A total of 34 out of 41 genes had significant correlation between mRNA expression and protein expression in NSCLC cell lines ($p < 0.05$, Pearson's correlation; Figure 3A). TRIM29 had two corresponding protein expression quantifications, both significantly correlated with its mRNA expression. The proportion of genes with a significant mRNA and protein expression correlation is 82.9% in the identified seven-gene network, which is significantly higher than 66.3% in the remaining 11,452 genes in the genome (two-tailed proportion z-test $p < 0.05$, Figure 3B). In TCGA-LUAD and TCGA-LUSC tumors, only CDH3 and CD274 had both mRNA expression and protein expression. The protein expression of CDH3 was positively correlated with its mRNA expression in TCGA NSCLC patient tumors ($p < 0.05$, Pearson's correlation; Figure 3C), consistent with the observation in CCLE NSCLC cell lines (Figure 3A).

2.2. Gene Co-Expression Networks of ZNF71 and Isoforms

In the study cohorts, *ZNF71* was associated with a more favorable survival outcome among patients who received cisplatin, carboplatin, and Taxol, indicating its linkage with chemosensitivity [10]. While *ZNF71* mRNA expression was not associated with survival in NSCLC patient tumors analyzed in qRT-PCR, a higher *ZNF71* protein expression measured with AQUA linked to a longer overall survival in two separate NSCLC cohorts ($n = 291$) using tissue microarrays [10]. The *ZNF71* antibody (Abcam; ab87250) used in our previous study [10] was discontinued. Protein level of overexpressed *ZNF71* (GeneTex, Cat. No. GTX116553) was detected in our previous study [19].

Gene *ZNF71* gene is composed four exons, in which exons 1 and 2 encode the 180nt 5'UTR and the first 11 amino acids of the protein. The next 43 amino acids (aa) in exon 3 are the majority of the KRAB repression domain. As the longest exon, exon 4 codes for the remainder of the protein, such as the putative 13-zinc-finger DNA binding domain and the predicted ~4 kb 3'UTR. It is possible to splice out the third KRAB-containing exon to produce a shorter in-frame protein encoded by exon 4, i.e., an alternative KRAB-less isoform. Nearly half of the 800 C2H2-type *ZNF* genes in humans contain a single, evolutionarily conserved KRAB region, which could be spliced in alternative fashion [11–13]. Our study demonstrated that *ZNF71* KRAB expression was associated with EMT in cell lines and tumors from NSCLC patients [19].

ZNF71 protein expression was not available in the CCLE database [37]. Here, *ZNF71* protein expression (GeneTex, Cat. No. GTX116553) was detected in A549 and H441 cell lines in whole cell lysate and nuclear fractions isolated from the same cell lines (Figure 4), consistent with its function as a transcription factor. *ZNF71* protein expression was positively correlated with epithelial markers, including E-cadherin and Cytokeratin, and was negatively correlated with mesenchymal markers, including ZEB1 and Vimentin in Western blots. These results are consistent with the observed positive association between elevated *ZNF71* protein expression and a good prognosis of NSCLC [10].

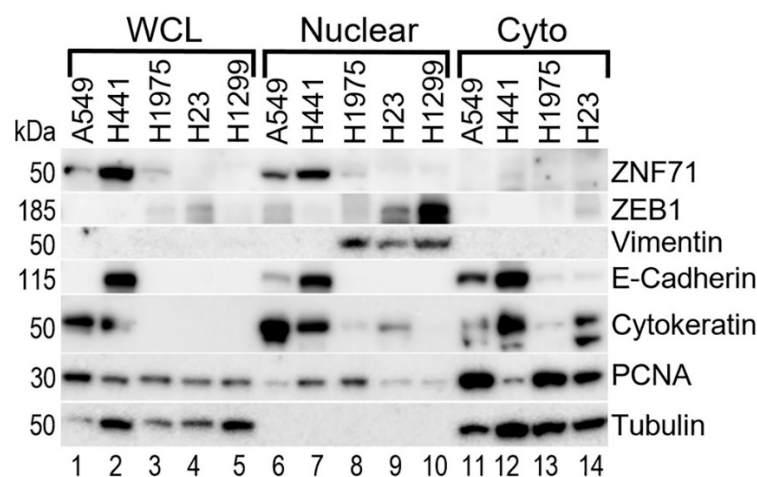


Figure 4. Correlation of expression of *ZNF71* with EMT markers in five NSCLC cell lines. Western blotting analysis of *ZNF71*, mesenchymal and epithelial markers in the indicated cell lines. PCNA and Tubulin—loading controls. WCL—whole cell lysate, Nuclear and Cyto—nuclear and cytoplasmic fractions isolated from the same cell lines.

To illustrate the molecular networks of *ZNF71* and its isoforms, the expression of *ZNF71* overall and isoforms was analyzed in the NGS dataset GSE81089 of NSCLC tumor samples ($n = 197$) [38]. Proliferation networks of *ZNF71*, *ZNF71* KRAB, and *ZNF71* KRAB-less were constructed, respectively, with the Boolean implication network algorithm. First, genes that had a significant co-expression ($p < 0.05$, z-tests) with *ZNF71*, *ZNF71* KRAB, or *ZNF71* KRAB-less in GSE81089 were identified in the genome-scale analysis, respectively. Next, the proliferation genes which meet the criteria were selected based

on CRISPR-Cas9/RNAi screening data in CCLE NSCLC cell lines (more details were provided in Supplementary File S3). These proliferation genes were pinpointed in the gene co-expression networks to form the proliferation networks. There were six categories of the constructed proliferation networks: (1) when *ZNF71* was up-regulated or not down-regulated (Figure 5A); (2) when *ZNF71* was down-regulated or not up-regulated (Figure 5B); (3) when *ZNF71* KRAB was up-regulated or not down-regulated (Figure 5C); (4) when *ZNF71* KRAB was down-regulated or not up-regulated (Figure 5D); (5) when *ZNF71* KRAB-less was up-regulated or not down-regulated (Figure 5E); and (6) when *ZNF71* KRAB-less was down-regulated or not up-regulated (Figure 5F). The gene expression signatures containing up- and down-regulated NSCLC proliferation genes from these six networks were input into CMap, respectively, to identify significant ($p < 0.05$) perturbagen signatures and functional pathways. The common significant (connectivity score > 0.9) functional pathways were summarized for *ZNF71* (Table 3), *ZNF71* KRAB (Table 4), and *ZNF71* KRAB-less (Table 5). More details were provided in Supplementary File S3.

The functional pathways associated with the identified proliferation networks of *ZNF71*, *KRAB*, and *KRAB*-less isoforms were largely different, although with some overlaps, i.e., CP_FGFR_INHIBITOR between *ZNF71* and *KRAB*, and PID_CIRCADIAN_PATHWAY for *ZNF71* and KEGG_CIRCADIAN_RHYTHM_MAMMAL for *KRAB*-less. When the up-regulated and down-regulated proliferation genes in each network (Figure 5) were uploaded to CMap separately, significant ($p < 0.05$, connectivity score > 0.9) perturbagen signatures and functional pathways were identified (Supplementary File S3 datasheet 14), showing that KD_AHSP_PATHWAY, KD_PHOSPHOLIPASES, and KD_PI3K_SIGNALING were commonly associated with *ZNF71*, *KRAB*, and *KRAB*-less proliferation networks. These results generated hypotheses for future mechanistic studies of *ZNF71* and its isoforms.

Table 3. The common significant ($p < 0.05$, connectivity score > 0.9) functional pathways from CMap using *ZNF71* up-regulated network genes (Figure 5A) and *ZNF71* down-regulated network genes (Figure 5B) as input.

src_set_id	ZNF71 Upregulated Network Genes		ZNF71 Downregulated Network Genes	
	Cell Line	Type	Cell Line	Type
BIOCARTA_CTL_PATHWAY	A549	TRT_SH.CGS	A549	TRT_SH.CGS
	HCC515	TRT_SH.CGS	HCC515	TRT_SH.CGS
CP_FGFR_INHIBITOR (PD-173074, dovitinib, brivanib, orantinib)	A549	TRT_CP	A549	TRT_CP
KD_RNA_POLYMERASE_ENZYMES (POLR2A, POLR2C, POLR2D, POLR2E, POLR2F, POLR2I, POLR2K)	A549	TRT_SH.CGS	HCC515	TRT_SH.CGS
KEGG_GALACTOSE_METABOLISM	HCC515	TRT_SH.CGS	HCC515	TRT_SH.CGS
PID_CIRCADIAN_PATHWAY	A549	TRT_SH.CGS	A549	TRT_SH.CGS
PID_INTEGRIN2_PATHWAY	HCC515	TRT_SH.CGS	A549	TRT_SH.CGS
PID_S1P_S1P2_PATHWAY	HCC515	TRT_SH.CGS	HCC515	TRT_SH.CGS
REACTOME_CHYLOMICRON_MEDIATED_LIPID_TRANSPORT	HCC515	TRT_SH.CGS	HCC515	TRT_SH.CGS
REACTOME_JNK_C_JUN_KINASES_PHOSPHORYLATION_AND_ACTIVATION_MEDIATED_BY_ACTIVATED_HUMAN_TAK1	HCC515	TRT_SH.CGS	HCC515	TRT_SH.CGS
REACTOME_PURINE_RIBONUCLEOSIDE_MONOPHOSPHATE_BIOSYNTHESIS	A549	TRT_SH.CGS	A549	TRT_SH.CGS
			HCC515	TRT_SH.CGS
SULFONYLUREA (glipizide, glibenclamide, gliquidone)	HCC515	TRT_CP	HCC515	TRT_CP

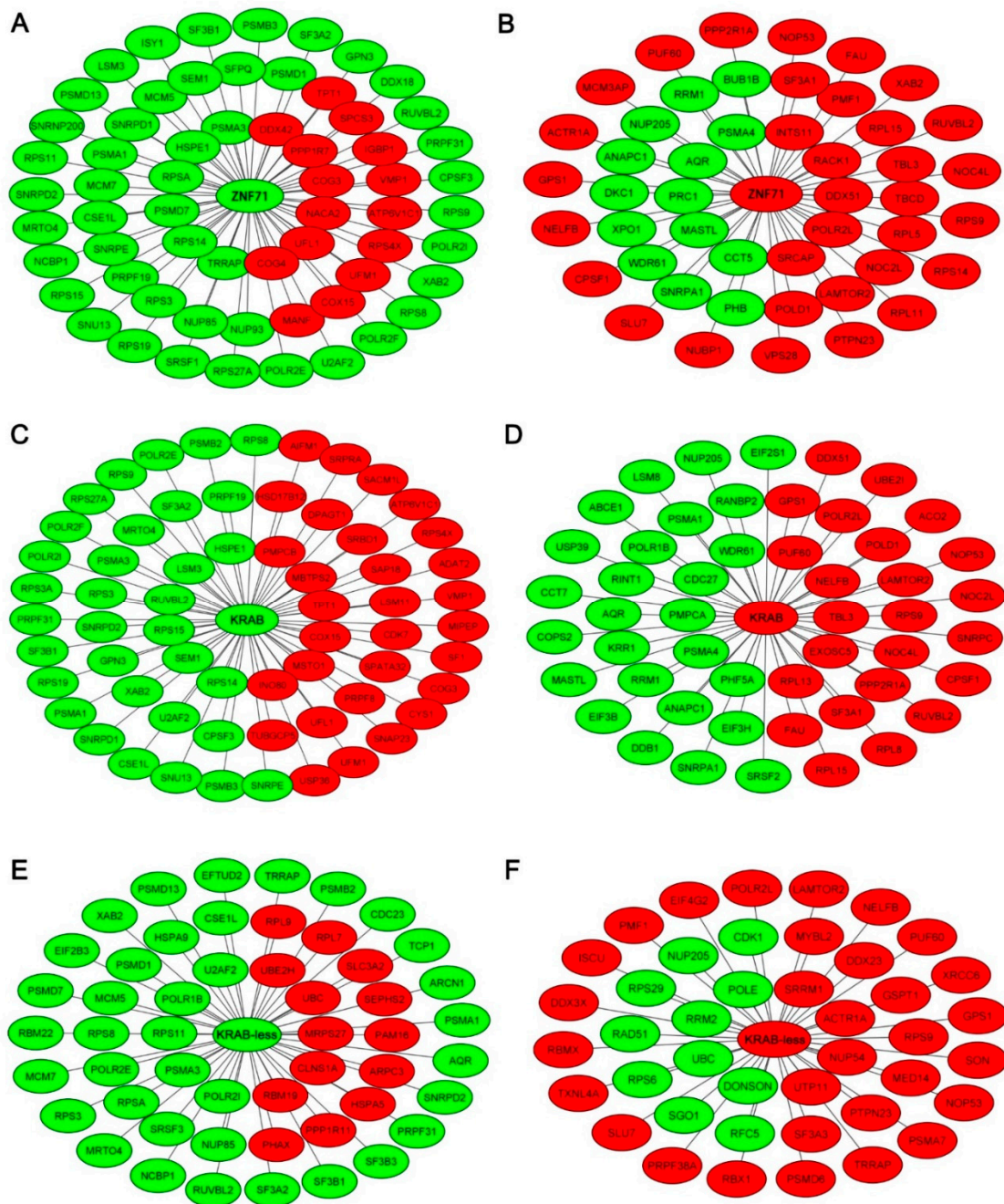


Figure 5. *ZNF71* gene level, *ZNF71* KRAB, and *ZNF71* KRAB-less co-expression networks in patient cohort GSE81089. The green color indicates up-regulation or non-downregulation, and the red color indicates downregulation or non-upregulation. All proliferation genes in the networks were identified from CRISPR-Cas9/RNAi screening data in CCLE NSCLC cell lines, with more details provided in Supplementary File S3. (A) Proliferation genes that had a significant ($p < 0.05$, z-tests) co-expression with up-regulated *ZNF71*. (B) Proliferation genes that had a significant ($p < 0.05$, z-tests) co-expression with down-regulated *ZNF71*. (C) Proliferation genes that had a significant ($p < 0.05$, z-tests) co-expression with up-regulated *ZNF71* KRAB isoform. (D) Proliferation genes that had a significant ($p < 0.05$, z-tests) co-expression with down-regulated *ZNF71* KRAB isoform. (E) Proliferation genes that had a significant ($p < 0.05$, z-tests) co-expression with up-regulated *ZNF71* KRAB-less isoform. (F) Proliferation genes that had a significant ($p < 0.05$, z-tests) co-expression with down-regulated *ZNF71* KRAB-less isoform. Details provided in Supplementary File S3.

Table 4. The common significant ($p < 0.05$, connectivity score > 0.9) functional pathways from CMap using *ZNF71* KRAB upregulated network genes (Figure 5C) and *ZNF71* KRAB downregulated network genes (Figure 5D) as input.

src_set_id	KRAB Upregulated Network Genes		KRAB Downregulated Network Genes	
	Cell Line	Type	Cell Line	Type
BIOCARTA_CTL_PATHWAY	HCC515	TRT_SH.CGS	HCC515	TRT_SH.CGS
			A549	TRT_SH.CGS
BIOCARTA_GLYCOLYSIS_PATHWAY	HCC515	TRT_SH.CGS	A549	TRT_SH.CGS
BIOCARTA_SET_PATHWAY	A549	TRT_OE	A549	TRT_OE
CP_CCK_RECEPTOR_ANTAGONIST (devazepide, LY-225910, SR-27897)	HCC515	TRT_CP	HCC515	TRT_CP
CP_FGFR_INHIBITOR (PD-173074, dovitinib, brivanib, orantinib)	A549	TRT_CP	HCC515	TRT_CP
KD_INTEGRIN_SUBUNITS_BETA (ITGB1, ITGB4, ITGB5)	A549	TRT_XPR	A549	TRT_XPR
PID_DNA_PK_PATHWAY	A549	TRT_SH.CGS	A549	TRT_SH.CGS
REACTOME_NFKB_IS_ACTIVATED_ AND_SIGNALS_SURVIVAL	A549	TRT_SH.CGS	HCC515	TRT_SH.CGS
REACTOME_REGULATION_ OF_COMPLEMENT_CASCADE	HCC515	TRT_SH.CGS	HCC515	TRT_SH.CGS
SODIUM/GLUCOSE_COTRANS PORTER_INHIBITOR (maackiain, phloretin)	A549	TRT_CP	A549	TRT_CP
SULFONYLUREA (glipizide, glibenclamide, gliquidone)	HCC515	TRT_CP	HCC515	TRT_CP
TGF_BETA_RECEPTOR_INHIBITOR	HCC515	TRT_CP	HCC515	TRT_CP

Table 5. The common significant ($p < 0.05$, connectivity score > 0.9) functional pathways from CMap using *ZNF71* KRAB-less upregulated network genes (Figure 5E) and *ZNF71* KRAB-less downregulated network genes (Figure 5F) as input.

src_set_id	KRAB-Less Upregulated Network Genes		KRAB-Less Downregulated Network Genes	
	Cell Line	Type	Cell Line	Type
ABL_KINASE_INHIBITOR	A549	TRT_CP	HCC515	TRT_CP
BIOCARTA_BLYMPHOC YTE_PATHWAY	A549	TRT_SH.CGS	A549	TRT_SH.CGS
KD_CYCLINS (CCNL1, CCND1, CCNA1, CCNH)	HCC515	TRT_SH.CGS	A549	TRT_OE
KEGG_CIRCADIAN_RHY THM_MAMMAL	A549	TRT_SH.CGS	A549	TRT_SH.CGS
REACTOME_TRAF3_DEPEN DENT_IRF_ACTIVATION_PATHWAY	A549	TRT_SH.CGS	A549	TRT_SH.CGS

2.3. Functional Pathways Associated with the Seven-Gene Multi-Omics Network and Discovery of Repositioning Drugs

ZNF71 KRAB was reported to be associated with EMT in our previous study [19]. Protein expression of *ZNF71* was positively correlated with epithelial markers and negatively correlated with mesenchymal markers in the nuclear fraction of NSCLC cells (Figure 4). The identified seven-gene multi-omics network (Figure 1C) and 14 EMT markers were used to

discover connected perturbagen signatures and repositioning drugs with potential implications in improving NSCLC treatment. The following mechanisms of action were considered to reverse EMT, to enhance drug sensitivity, to induce immune response, and to inhibit NSCLC proliferation through (1) up-regulation of epithelial markers; (2) up-regulation of drug sensitive genes; (3) down-regulation of mesenchymal markers; (4) down-regulation of proliferation genes; and (5) down-regulation of *CD27* and *PD-L1* (*CD274*). The 14-gene EMT classifier was used in our previous study [19] to define the EMT states of NSCLC tumors, including three mesenchymal markers (*ZEB1*, *VIM*, and *FN1*) and 11 epithelial markers (*CDH1*, *EPCAM*, *ESRP1*, *ESRP2*, *DDR1*, *CTNNB1*, *CD24*, *CLDN7*, *KRT8*, *KRT19*, and *RAB25*). The drug sensitive genes were selected if a gene's mRNA expression was sensitive to one or more of the studied 10 NSCLC therapeutic regimens and was not resistant to any studied drugs in PRISM or GDSC1/2 datasets. The drug sensitive genes and the epithelial markers were excluded from the list if they were also proliferation genes. Here, the proliferation genes referred to those with a significant dependency score (< -0.5) in CRISPR-Cas9 or RNAi in more than 10 NSCLC cell lines. The proliferation gene list also excluded the epithelial markers (Figure 6A).

With the above defined up- and down-regulated gene lists, significant ($p < 0.05$, connectivity score > 0.9) functional pathways were identified with CMap (Table 6). The details of results were provided in Supplementary File S2, sheet 4. Specifically, shRNA knock-down of the following pathways matched the seven-gene multi-omics network, including KD_APOLIPOPROTEINS, KD_CYCLINS, KD_LYSINE_ACETYLTRANSFERASES, KD_NON_HOMOLOGOUS_END_JOINING, KD_V_TYPE_ATPASES, and OE_NFKB_ACTIVATION (Table 6). After checking the gene co-expression networks in NSCLC patient tumors in GSE31800 and GSE28582, XRCC5 from KD_NON_HOMOLOGOUS_END_JOINING pathway was found to have the same gene co-expression patterns in patient tumors as in shRNA knock-down in A549 cell line (Figure 6B; Table 6). A co-downregulation between XRCC5 and *CDCA3*, *ERH*, and *KRT17*, respectively, was found in both NSCLC tumors and LINCS L1000 CMap NSCLC cell lines (Figure 6B).

Table 6. The identified significant ($p < 0.05$, connectivity score > 0.9) functional pathways from shRNA knock-down experiments in CMap using the up-regulated and down-regulated gene lists in Figure 6A as input.

src_set_id	Cell_Iname	Pert_Type	Genes
KD_APOLIPOPROTEINS	A549	TRT_SH.CGS	<i>APOB</i> , <i>APOC2</i> , <i>APOE</i>
KD_CYCLINS	A549	TRT_SH.CGS	<i>CCNL1</i> , <i>CCND1</i> , <i>CCNA1</i> , <i>CCNH</i>
	HCC515	TRT_SH.CGS	
KD_LYSINE_ACETYLTRANSFERASES	A549	TRT_SH.CGS	<i>KAT6B</i> , <i>KAT6A</i> , <i>NCOA3</i>
KD_NON_HOMOLOGOUS_END_JOINING	A549	TRT_SH.CGS	<i>RAD50</i> , <i>FEN1</i> , <i>XRCC4</i> , <i>XRCC5</i>
KD_V_TYPE_ATPASES	HCC515	TRT_SH.CGS	<i>ATP6V1A</i> , <i>ATP6V0C</i> , <i>ATP6V0B</i> , <i>ATP6V1F</i>
OE_NFKB_ACTIVATION	HCC515	TRT_SH.CGS	<i>CD40</i> , <i>FADD</i> , <i>LTBR</i> , <i>TNFRSF10A</i> , <i>TNFRSF10B</i> , <i>TNFRSF1A</i> , <i>BCL10</i>

Based on the designed mechanisms of action, 39 repositioning drugs were identified with CMap ($p < 0.05$, Supplementary File S4, datasheet 7). From this list, drugs that have lower average IC_{50} and EC_{50} in CCLE NSCLC cell lines were selected for potential inhibition of the growth of lung cancer cells with a safer dose. In the original PRISM drug screen, eight doses ranging from 0.0006 μ M to 10 μ M were tested on most of the compounds [39]. Five small molecules (dasatinib, FK-888, homosalate, lestaurtinib, and

penfluridol) had a small average IC_{50} or EC_{50} value (smaller than or close to 1 μM) in the PRISM drug screen and were selected for drug repurposing (Figure 6C, more details provided in Supplementary File S4). Dasatinib is a targeted therapy used to treat chronic myelogenous leukemia and acute lymphoblastic leukemia. FK 888 is a selective, high-affinity tachykinin NK1 receptor antagonist [40]. Homosalate is an organic compound commonly used in sunscreen that absorbs shortwave UVB rays, which helps prevent DNA damage and decreasing the risk of skin cancer. Lestaurtinib is a tyrosine kinase inhibitor and is structurally related to staurosporine. It has undergone clinical trials for the treatment of various cancers, including leukemia [41] and refractory pediatric neuroblastoma [42]. Penfluridol is a first-generation antipsychotic drug used to treat schizophrenia [43]. These results provided evidence for future drug repurposing study to improve NSCLC treatment in combination with clinical therapies.

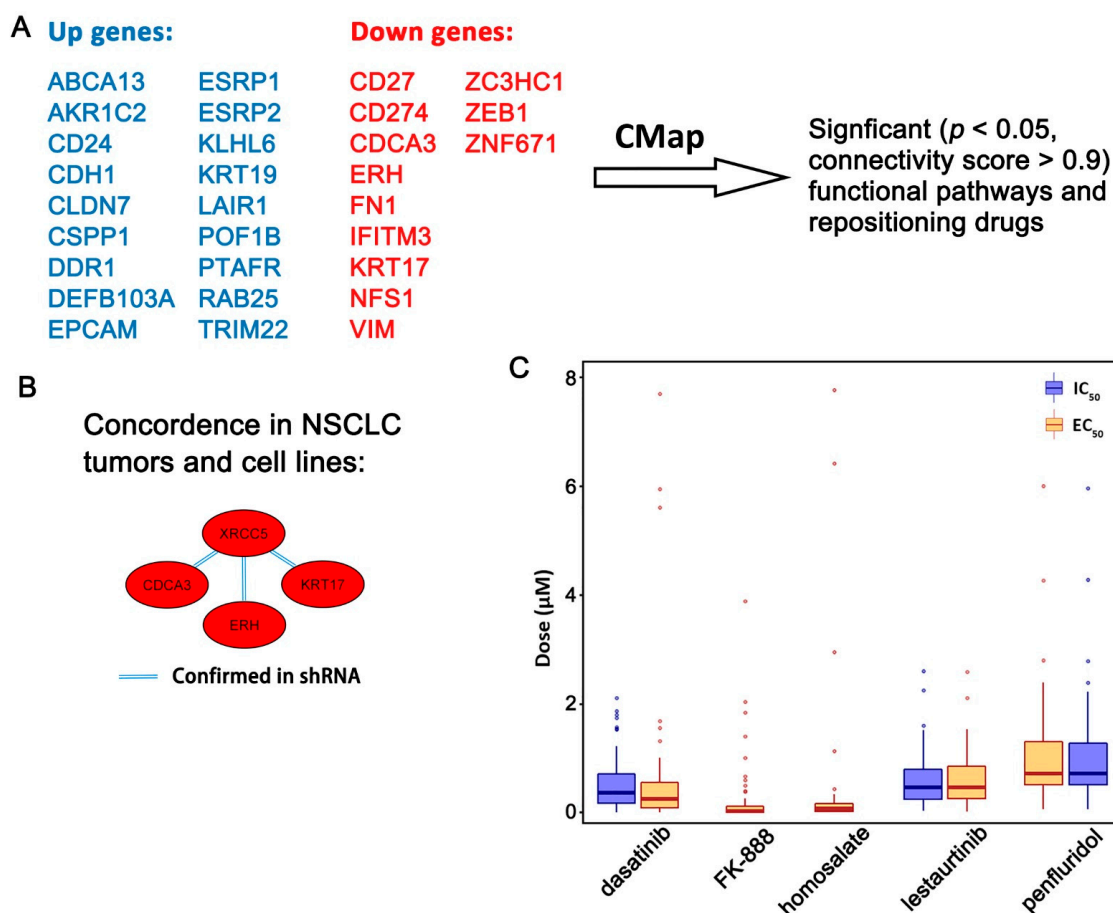


Figure 6. Identification of functional pathways associated with the 7-gene multi-omics network and discovery of repositioning drugs. (A) Selection of significant functional pathways and repositioning drugs based on the identified 7-gene multi-omics proliferation network with CMap. (B) Confirmed co-expression network of *XRCC5* in NSCLC patient tumors in GSE31800 or GSE28582 showing concordant co-expression patterns after shRNA knock-down of *XRCC5* in A549 cell line. The red color indicates down-regulation or non-upregulation of genes. The double line indicates that the co-expression relations ($p < 0.05$, z -tests) observed in patient tumors were confirmed in shRNA knockdown experiments in LINCS L1000 CMap NSCLC cell lines. (C) Small molecules that had a low average concentration of drug response in the CCLE NSCLC cell lines. Measurements with a drug activity value larger than 10 μM were considered outliers and were removed from the plot. Dasatinib EC_{50} measurements had 2.38% of outliers. FK-888 EC_{50} measurements had 14.49% of outliers. Homosalate EC_{50} measurements had 7.25% of outliers. Penfluridol had 1.25% of outliers in both IC_{50} and EC_{50} measurements. Detailed information was provided in Supplementary File S4.

3. Discussion

In both males and females, lung cancer remains the leading cause of cancer-related death. Lung cancer is difficult to manage clinically due to its complex etiology and a lack of biomarkers for the selection of therapy in individual patients. A National Cancer Institute report shows that 56% of lung cancer cases are diagnosed at the distant (metastatic) stage [44]. Novel biomarkers can aid the selection of appropriate chemotherapy and immunotherapy for treatment of advanced NSCLC. For early-stage NSCLC, it remains a challenge to improve patient survival due to the varying performance of lymphadenectomy and insufficient understanding of molecular features harboring recurrence [45]. The development of clinically applicable prognostic biomarkers for early-stage NSCLC could offer a solution by identifying tumors prone to recurrence. Thus, a significant unmet need exists for a reliable diagnostic method to (1) identify stage I NSCLC who might benefit from adjuvant chemotherapy; and (2) select specific therapeutic regimens that a NSCLC patient might respond to.

We sought to meet these critical needs with the development of a qRT-PCR based seven-gene microfluidic assay for early-stage NSCLC prognosis and prediction of benefits of chemotherapy [10]. The seven-gene assay was constructed from the feasibility-training cohort with 83 patients based on accurate patient stratification ($p = 0.0043$), and it has been validated in independent patient cohorts including a clinical trial JBR.10 [46] ($n = 248$, $p < 0.0001$) according to Kaplan–Meier analysis. In both training ($p = 0.035$) and validation ($p = 0.0049$) sets, patients who received adjuvant chemotherapy had significantly better disease-specific survival than patients who did not receive chemotherapy in the predicted chemotherapy-benefit group. Patients receiving chemotherapy in either set did not show a survival benefit in the predicted non-benefit group [10]. The ability of this gene assay to predict the recurrence risk and recommend early implementation of chemotherapy in stage I NSCLC patients can potentially improve patient care. In this study, we showed that it is feasible to apply this seven-gene panel for NSCLC prognosis using NGS profiles of the TCGA patient tumors randomly partitioned into training and testing sets. In future studies, the development of an NGS-based seven-gene assay will need to be defined with specific NGS data generation protocols, data processing methods, and model parameters from the training set that are validated in external patient cohorts for potential clinical applications.

Within this seven-gene assay, we recently found that *ZNF71 KRAB* isoform is associated with EMT in NSCLC tumors and cell lines, i.e., expressed higher in top 50% of *ZEB1* expressing cells [19]. *ZNF71* protein expression was positively correlated with epithelial markers, including E-cadherin and Cytokeratin, and was negatively correlated with mesenchymal markers, including *ZEB1* and Vimentin in NSCLC cell lines in Western blots. These results were consistent with the observed positive association between *ZNF71* protein expression and prolonged NSCLC patient survival [10]. Together, these results supported the transcriptional repression role of the KRAB domain.

Approximately 3% of the human genome is composed of zinc finger proteins (ZFPs), which are the largest transcription factor family in human cells and most eukaryotes [47,48]. ZFPs contain a repeating structure of two histidine and two cysteine amino acid residues (i.e., C2H2) that coordinate a zinc ion, which is capable of binding to DNA, RNA, or cellular proteins [49]. A number of biological functions have been predicted for ZFPs as a result of their molecular structure, including DNA repair, protein degradation, signal transductions, cell migration, apoptosis regulation, lipid binding, and transcription regulation [48,49]. The KRAB-ZFPs family of transcriptional repressors plays a variety of roles, most importantly the silencing of transposable elements [50]. KRAB-ZFPs are composed of an N-terminal KRAB domain and a C-terminal C2H2-type zinc finger array. The recruitment of KRAB-associated protein 1, also known as tripartite motif protein 28, KAP1/TRIM28, plays an essential role in the repressor function of KRAB-ZEPs. KAP1/TRIM28 acts as a scaffold complex comprised of histone methyl transferase (SETDB1), heterochromatin protein-1 (HP-1), nucleosome remodeling and deacetylation (NuRD), and DNA methyl transferase [13]. The constructed repressor complex triggers heterochromatin production

when KRAB-ZFPs bind KAP1/TRIM28. KAP1/TRIM28 is a universal co-factor for KRAB-ZFP transcription factors [24] and is involved in a variety of cellular functions, ranging from promoting cell proliferation [51] to mediating anti-proliferative activities [52]. The KAP1/TRIM28 protein is involved in tumorigenesis by formatting heterochromatin, mediating DNA damage response, inhibiting *p53* activity, regulating EMT, and maintaining stem cell pluripotency and genome stability [53]. TGF- β treatment induces KAP1/TRIM28 mRNA and protein expression. KAP1/TRIM28 deficiency blocks TGF- β induced EMT and reduces cell migration and invasion. KAP1/TRIM28 regulates E-cadherin and N-cadherin promoters, implying that KAP1/TRIM28 contributes to EMT and is involved in lung cancer metastasis [25]. Upregulation of RB-associated KRAB zinc finger (BRAK) and its association with poor prognosis in NSCLC was reported [51]. Zinc finger protein 668 (ZNF668) down-regulates Snail and upregulates E-cadherin and zonula occludens-1, which, in turn, suppresses NSCLC invasion and migration [54]. ZNF668 protein expression was downregulated in tumors compared with normal lung tissues and was negatively associated with lymph node metastasis in NSCLC [54].

The role of *ZNF71* and its isoforms in EMT is unclear. EMT may be linked to significant splicing changes, and as a consequence, it could potentially skew splicing for inclusion of the KRAB domain [55,56]. To better understand the functions of *ZNF71* and its isoforms, genome-scale co-expression networks containing all the genes with a statistically significant co-expression association relation with *ZNF71* overall, *KRAB*, and *KRAB-less* were constructed, respectively. Proliferation networks were dissected based on CRISPR-Cas9/RNAi screening data in CCLE NSCLC panel. Significant functional pathways matching the up- and down-regulation signatures of these proliferation networks were identified with L1000 CMap. The results showed that multiple different functional pathways were associated with *ZNF71*, *KRAB*, and *KRAB-less* isoforms in shRNA knock-down or overexpression assays in NSCLC cell lines, respectively. Three pathways, including KD_AHSP_PATHWAY, KD_PHOSPHOLIPASES, and KD_PI3K_SIGNALING in shRNA knock-down assays of NSCLC cell lines, were commonly associated with *ZNF71*, *KRAB*, and *KRAB-less* isoforms. These results provided hypotheses of future molecular and cellular studies of *ZNF71* and its isoforms in NSCLC.

It has been found that the seven-gene panel interacted with major inflammatory and cancer signaling hallmarks including *TNF*, *PI3K*, *NF- κ B*, and *TGF- β* [10]. The protein expression of CD27 quantified with ELISA had a strong correlation with its mRNA expression in NSCLC tumors [10]. As a new generation of immunotherapy target [57,58], CD27 is currently being tested in phase I/II clinical trials for multiple tumor types with promising results [59,60]. The treatment effectiveness of cancer vaccines and immunotherapy can be improved by CD27 agonist antibodies, either alone or in combination with anti-PD1 antibodies [61,62]. There was reported synergy between PD-1 blockade and CD27 stimulation for CD8+ T-cell driven anti-tumor immunity [62], indicating the therapeutic potential of CD27 in neoadjuvant PD-1 blockade in early-stage NSCLC [63].

Novel biomarker discovery is complicated by combinations of chemotherapy and immunotherapy, and multiple intricate interactions relevant to immune responses should be considered in this process [64]. The NSCLC tumor immune microenvironment can be represented as a multi-dimensional component network of immune and stromal cells as well as blood vessels, interacting with epithelial tumor cells in a convoluted manner. An integrated analysis of genome-scale CNV and gene expression profiles in NSCLC tumors was conducted using the Boolean implication network methodology to model this complicated molecular machinery [65]. The presented Boolean implication network algorithm is conceptually innovative and can model molecular circuits, such as feedback loops, which are important in biological switches and signaling motifs that regulate large-scale cellular processes [66]. The multi-omics network involving the seven-gene signature was dissected. All the network genes, including *PD-L1 (CD274)*, had a significant correlation with immune infiltration in at least one immune cell type of TCGA NSCLC tumors, including myeloid dendritic cells, macrophages, neutrophils, CD4+ T cells, CD8+ T cells, and B cells. Among

the network genes, down-regulation of *SERPINB5* by small interfering RNA increased the resistance of NSCLC H460 cells to ionizing radiation as reported in the literature [67], consistent with the observed association between *SERPINB5* up-regulation and prolonged survival in TCGA NSCLC patients receiving radiotherapy. Low XRCC5 protein expression was more prevalent among squamous cell carcinoma compared with lung adenocarcinoma compared with normal lung tissues in Lee et al. [68]. The cell division cycle-associated gene (*CDCA3*), mRNA, and protein expressions were increased in NSCLC compared with normal tissue, and high levels of *CDCA3* was associated with poor prognosis in Adams et al. [69]. The mRNA and protein expressions of numerous network genes were associated with drug sensitivity or resistance to 10 commonly used therapeutic regimens for treating NSCLC. These results further substantiate the potential clinical implications of the multi-omics network of the seven-gene signature.

The multi-omics network genes had a higher percentage (82.9%) of correlated mRNA and protein expression in CCLE NSCLC cell lines compared with the rest of the genome (66.3%). The correlation between mRNA and protein expression levels in complex biological systems is in general poor [70]. Many studies reported dysregulated gene expression as evidence of the consequent dysregulation of protein expression. Nevertheless, this is not necessarily true in many circumstances. Current RNA sequencing and reverse-phase protein array technologies have revealed genome-scale mRNA and protein expression and their correlation to better understand the transcriptional and translational processes. When reverse-phase protein arrays or commercial antibodies are not available for a research project, the mRNA expression can be used as a surrogate of the protein expression for genes with highly correlated mRNA and protein expression levels.

Molecular pathway and network approaches are promising for the discovery of novel therapeutic targets and repositioning drug candidates [71]. Drug repositioning is essential for pharmaceutical research and development (R & D) due to the proven toxicity profiles [71]. To design therapeutic strategies to reverse EMT, enhance drug sensitivity, and inhibit NSCLC proliferation, significant functional pathways and repositioning drugs were identified with L1000 CMap. The 14-gene EMT classifier to characterize NSCLC tumors [10], the multi-omics network of the seven-gene signature, and CRISPR-Cas9/RNAi screening data were used to generate up- and down-regulated gene lists as input to CMap. ShRNA knock-down of the following pathways were linked to these mechanisms, including KD_APOLIPOPROTEINS, KD_CYCLINS, KD_LYSINE_ACETYLTRANSFERASES, KD_NON_HOMOLOGOUS_END_JOINING, KD_V_TYPE_ATPASES, and OE_NFKB_ACTIVATION. Nuclear factor kappa B (NF- κ B) is a master regulator in NSCLC pathogenesis and progression [72,73]. Activated NF- κ B signaling may influence the progression of a lung tumor either positively via triggering immune surveillance [74] or negatively via promoting inflammation and tumorigenesis [75]. Activation of NF- κ B stimulates transcription of cyclin D1, a key regulator of G1 checkpoint control [76]. NNK, a tobacco-specific nitrosamine, can stimulate proliferation through activation of the NF- κ B and subsequent upregulation of cyclin D1 in a normal human bronchial cell [77]. Inhibition of NF- κ B results in loss of induction of EMT master-switch transcription factors in NSCLC [78]. NF- κ B activation has been associated with resistance to chemotherapeutics and tyrosine kinase inhibitors in EGFR mutant lung cancer [72]. Protein expression of V-ATPase is associated with drug resistance and higher tumor grade in NSCLC [79]. Dasatinib and lestaurtinib used for leukemia treatment were identified as repurposing drugs for NSCLC. Homosalate, FK 888, and penfluridol were discovered as potential repositioning drugs that can kill NSCLC cells, maintaining the expression of drug sensitivity genes and epithelial markers, and inhibit mesenchymal markers, proliferation genes, and ICIs for NSCLC treatment. We developed a novel pipeline to identify repositioning drugs based on 1) network-based identification of genes with prognostic and chemo-predictive implications, 2) pre-selection of repositioning drug candidates using CMap, and 3) further selection of repositioning drugs that can effectively kill NSCLC cells with the PRISM and GDSC1/2 drug activity databases. The pipeline presented in this study can aid pharmaceutical R & D in determining disease

relevance and designing of clinical trials of repositioning drugs to minimize the risk of failures, to expedite drug development, and to reduce costs.

4. Materials and Methods

4.1. Patient Cohorts

4.1.1. TCGA NSCLC Patient Cohorts

The TCGA-LUAD ($n = 515$) and TCGA-LUSC ($n = 501$) data were obtained from LinkedOmics [80] (<http://linkedomics.org/>, accessed on 28 April 2021). Normalized gene-level reverse phase protein array data, gene expression, and patient clinical information were available from the source.

4.1.2. NSCLC Patient Cohort GSE31800

A previous study [81] measured DNA copy number profiles in 271 samples from NSCLC tumors (GSE31800). A total of 49 samples had matched custom microarray gene expression profiles [81].

4.1.3. NSCLC Patient Cohort GSE28582

This study included a publicly available NSCLC patient cohort ($n = 100$) [82,83] (GSE28582). All 100 samples had both SNP array profiled DNA copy number variation and microarray gene expression data.

4.1.4. NSCLC Patient Cohort GSE81089

A total of 197 patients with sufficient survival information out of 199 NSCLC patient tumor samples from GSE81089 [38] were included in this study. The transcripts per million (TPM) reads of *ZNF71* at the isoform level were generated from the raw RNA-seq data with the Salmon method [84].

4.2. Data Pre-Processing

4.2.1. CNV Data Pre-Processing

The CNV data from GSE31800 [81] and GSE28582 [82,83] cohorts were matched with the converted Hg38 chromosome locations. Details of the CNV data pre-processing were provided in our previous study [65]. The pre-processed CNV data were categorized into 3 levels: 1 (amplification), -1 (deletion), and 0 (normal).

4.2.2. Gene Expression Data Pre-Processing

In order to categorize the level of gene expression, we used 27 housekeeping genes (*ACTB*, *B2M*, *CDKN1B*, *ESD*, *FLOT2*, *GAPDH*, *GRB2*, *GUSB*, *HMBS*, *HPRT1*, *HSP90AB1*, *IPO8*, *LDHA*, *NONO*, *PGK1*, *POLR2A*, *PPIA*, *PPIH*, *PPP1CA*, *RHOA*, *RPL13A*, *SDCBP*, *TBP*, *TFRC*, *UBC*, *YAP1*, and *YWHAZ*) [10,85–88] to define the thresholds of gene expression level. The total percentage of over-expression and under-expression samples for all the housekeeping genes was fixed to be 30%, and the number n that meets the following Equation (1) was computed for each gene [88,89]:

$$\begin{aligned} & (\text{number of samples have expression} > \text{mean} + n \times \text{std}) \\ & + (\text{number of samples have expression} < \text{mean} - n \times \text{std}) \\ & = 30\% \text{ of total number of samples} \end{aligned} \quad (1)$$

The averaged n value of all 27 housekeeping genes was 0.874 in the gene expression data of GSE31800 and 0.977 in the gene expression data of GSE28582, respectively. These n values were used to categorize each gene's expression in each dataset. Details of gene expression normalization using these housekeeping genes were provided in our previous publication [65].

4.3. Boolean Implication Networks

In the present study, an implication network algorithm based on prediction logic [90] was used. This algorithm was proposed by Guo et al. [91,92]. Details of the implication network algorithm and its application in CNV and gene expression analysis were provided in our previous study [65]. For an implication rule to be considered significant, it must have a scope and precision values that are greater than the threshold values, as well as being the largest among all the rules for the specific pair of variables. Based on the significance level of statistical tests, the threshold for scope and precision will vary. Based on the sample size, the threshold value was calculated using a one-tailed z-test using the ideal value of z. The z value used for this study was 1.64 (95% confidence interval, $\alpha = 0.05$, one-tailed z-tests) [65].

4.4. Cancer Cell Line Encyclopedia (CCLE)

CCLE cell line RNA-seq data were obtained from DepMap 20Q2 (https://figshare.com/articles/dataset/DepMap_20Q2_Public/12280541, accessed on 1 April 2021) [93]. Gene expression data quantified with the GTEx pipelines [94] were obtained from the CCLE data portal (https://data.broadinstitute.org/ccle/CCLE_RNAseq_081117.rpkm.gct, accessed on 1 April 2021). A total of 117 NSCLC cell lines were included in this analysis.

Proteomics data for CCLE were downloaded from the Gygi lab (<https://gygi.hms.harvard.edu/publications/ccle.html>, accessed on 1 October 2021) [37]. A panel of 63 NSCLC cell lines were included in this analysis.

4.5. CRISPR-Cas9 Knockout Assays

Project Achilles has quantified CRISPR-Cas9 gene knockout effects in CCLE [95,96] available in version DepMap 20Q2 (https://figshare.com/articles/dataset/DepMap_20Q2_Public/12280541, accessed on 1 April 2021) [93]. CERES [95] was used to process the screening data for CRISPR-Cas9. Genes are defined as essential if their knockout significantly impact the cellular growth in a cell line; otherwise, they are defined as non-essential. In each cell line, gene knockout effects were normalized so that the median non-essential gene knockout effect was 0, and the median essential gene knockout effect was -1 . Genome-scale CRISPR-Cas9 knockout results in 78 NSCLC cell lines were included in this study. If a gene has a normalized dependency score less than -0.5 in a sample, it is considered to have a significant effect in CRISPR-Cas9 knockout.

4.6. RNAi Knockdown Assays

Project Achilles also contains genome-scale RNAi screening data of CCLE (<https://depmap.org/R2-D2/>, accessed on 1 April 2021) [97]. The average gene dependency scores in each cell line for short hairpin RNA (shRNA) libraries were estimated using the DEMETER2 method [97]. The gene dependency scores were standardized such that the median of the across-cell-line average dependency scores of the positive control gene set was -1 , and the median of the averaged scores of the negative control gene set was 0. Same as CRISPR-Cas9, if a sample has normalized dependence score less than -0.5 , that indicates it has a significant effect in RNAi knockdown. A total of 92 NSCLC cell lines were included in the RNAi analysis.

4.7. Immune Infiltration Estimation

TIMER 2.0 [34–36] was used to compute the association of gene expression and immune infiltration in multiple cell types using a variety of immune deconvolution methods.

4.8. PRISM Drug Response in CCLE

The PRISM molecular barcoding and multiplexed screening method quantified the growth inhibitory activity of 4518 drugs in 578 human cancer cell lines [39]. The PRISM dataset was obtained from the Cancer Dependency Map portal (<https://depmap.org/portal/download/>, accessed on 1 April 2021). In this study, we focused on the activities

of nine drugs for treating NSCLC, including carboplatin, cisplatin, paclitaxel, docetaxel, gemcitabine, vinorelbine, etoposide, gefitinib, and erlotinib. In the experiments, different levels of doses were designed for the treatments by each drug. A cell line was directly defined as resistant if its IC₅₀ or EC₅₀ value is higher than the maximum dose; a cell line was defined as sensitive if its IC₅₀ or EC₅₀ value is lower than the minimum dose. The remaining cell lines with IC₅₀ or EC₅₀ value in-between the minimum and maximum doses were divided into groups of resistant, sensitive, or partial response. Details of drug response categorization were provided in our previous studies [19,65].

4.9. Genomics of Drug Sensitivity in Cancer (GDSC1/2)

Drug screening data were available from the Genomics of Drug Sensitivity in Cancer (GDSC) Project (https://www.cancerrxgene.org/downloads/bulk_download, accessed on 15 April 2021) [98]. In more than 1000 human cancer cell lines, a broad spectrum of anti-cancer therapeutic compounds was screened in the GDSC Project. Details of GDSC1/2 and drug response categorization were described in our previous publications [19,65].

4.10. Drug Repurposing Using Connectivity Map (CMap)

CMap uses transcriptional expression signatures to connect disease, cell physiology, and therapeutics. CMap contains a library of over 1.5M mRNA expression profiles from ~5000 small-molecule compounds and ~3000 genetic reagents, screened in multiple cell types. Using these screening data, CMap generates hypotheses of functional pathways and repositioning drugs through the matching of gene expression signatures containing up-regulated and/or down-regulated gene lists as database queries. CMap (<https://clue.io/>, accessed on 1 October 2021) [26,27] was used to find perturbagen signatures and functional pathways that connected with the input gene expression signatures.

4.11. Western Blots

NSCLC cell lines H441, H23, and H1299 (kind gift of Dr. Scott Weed, WVU), A549 and H1975 (kind gift of Dr. Ivan Martinez, WVU) were grown in DMEM (Corning, cat. # 15-018-CV) supplemented with 10% FBS (HyClone, UT, USA), 2mM L-glutamine (Corning, Steuben County, NY, USA; cat. # 25-005-CI), and 1 × Antibiotic Antimycotic Solution (Corning, Steuben County, NY, USA; cat. # 30-004-CI). All cells were maintained at 37 °C in a 5% CO₂ incubator. Whole/total cell lysates were prepared in nonreducing Laemmli buffer as described in [99]. Nuclear-cytoplasmic fractionation was performed using ProteoJET Cytoplasmic and Nuclear Protein Extraction Kit (Fermentas, cat. # K0311) as described in [51].

Protein concentration was quantified by Pierce BCA Protein Assay (ThermoFisher, Waltham, MA, USA; cat # 23225). Lysates with equal amount of total protein were separated on 4% to 12% Bis-Tris NuPAGE Novex gels (Invitrogen, Waltham, MA, USA) and transferred to a polyvinylidene difluoride (PVDF) membrane. Protein bands were detected using standard chemiluminescence techniques using GE Healthcare Amersham Imager 680.

ZNF71 antibody was purchased from GeneTex (Cat. No. GTX116553). The following antibodies were included in Western blots: ZEB1 (Sigma, St. Louis, MO, USA; HPA027524), Vimentin (V-9; Santa Cruz Biotechnology, Dallas, TX, USA; sc-6260), E-cadherin (BD Biosciences, Franklin Lakes, NJ, USA; cat# 610181), pan-Cytokeratin (C11, Santa Cruz Biotechnology, Dallas, TX, USA; sc-8018), PCNA (eBioSciences, 14-6748-81), and Tubulin α (Sigma, St. Louis, MO, USA; T9026). Standard chemiluminescence was used to detect the protein bands.

4.12. Statistical Methods

Rstudio version 1.4.1106 [100] was used as the main software in the statistical analysis of this study. Student's *t*-tests were used to assess differences in mRNA and protein expression between two groups; statistical significance was defined as having a two-sided *p*-value < 0.05. The multivariate Cox regression analysis was used to compute risk-scores. Kaplan–Meier analysis was applied with the survival package in R to perform the survival

analysis. In Kaplan–Meier analyses, log-rank tests were used to determine the difference in survival probability between different groups. The relationship between two sample groups were evaluated with Pearson’s correlation test. One-tailed two proportion z-test was used to compare if two proportions are the same.

5. Conclusions

This study substantiated the feasibility of applying the seven-gene signature developed with qRT-PCR microfluidic assay to TCGA NGS data for NSCLC prognosis. CNV mediated transcriptional regulatory network of the seven-gene signature was identified in multiple NSCLC cohorts using Boolean implication networks. The multi-omics network genes, including *PD-L1 (CD274)*, were significantly correlated with immune infiltration and drug sensitivity or resistance to 10 commonly used therapeutic regimens for treating NSCLC. Dasatinib, lestaurtinib, homosalate, FK 888, and penfluridol were discovered as potential repositioning drugs that can kill NSCLC cells, maintain the expression of drug sensitivity genes and epithelial markers, and inhibit mesenchymal markers, proliferation genes, and ICIs for NSCLC treatment. ShRNA knock-down of NF- κ B and cyclin D1 pathways were identified as associated with the multi-omics network for designing therapeutic strategies for NSCLC intervention. PI3K was identified as a potential pathway related to proliferation networks involving *ZNF71* and its isoforms for future mechanistic study. We acknowledge that the identified functional pathways and repositioning are based on the matching of the gene expression signatures with screening data in the CMap databases. Although shRNA knockdown and overexpression experimental results were included in the similarity matching in the CMap queries, these results from bioinformatics analysis will need to be further validated in molecular and cellular experiments.

6. Patents

The seven-gene microfluidic assay was filed under US Non-Provisional Patent Pub. No. US 2021-0254173 A1. The results in this study were filed under US provisional patent application 63/265,649.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/ijms23010219/s1>.

Author Contributions: Conceptualization and methodology, N.L.G.; software, N.L.G., Q.Y., and S.S.; formal analysis, Q.Y. and S.S.; investigation, B.F.; resources, A.V.I. and T.D.E.; writing—original draft preparation, N.L.G., Q.Y., and A.V.I.; visualization, Q.Y.; supervision, N.L.G. and A.V.I.; funding acquisition, N.L.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the NIH R01/R56LM009500, P20RR16440 and ARRA Supplement, as well as NIGMS/XLerateHealth LLC (4UT2GM130174) I2P Award to N.L.G.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data included in this analysis are publicly available with data access provided in the manuscript.

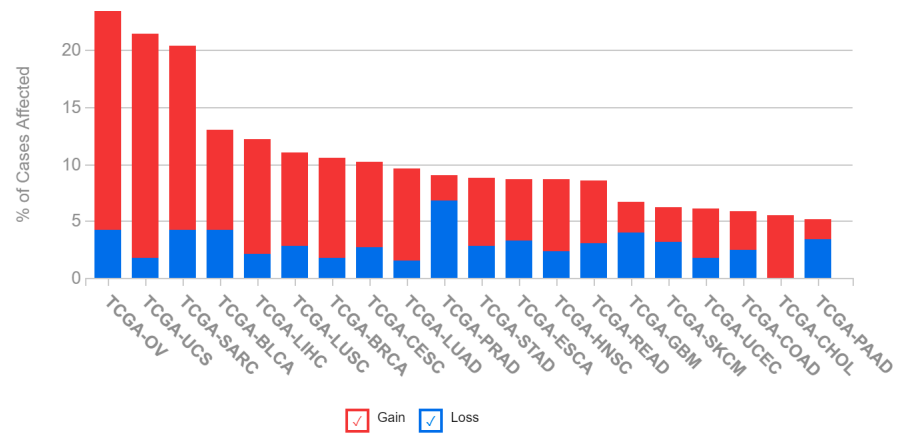
Acknowledgments: We thank Scott Weed and Ivan Martinez at West Virginia University for providing cell lines. The Office of Research and Graduate Education at West Virginia University Health Sciences Center provided institutional support.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

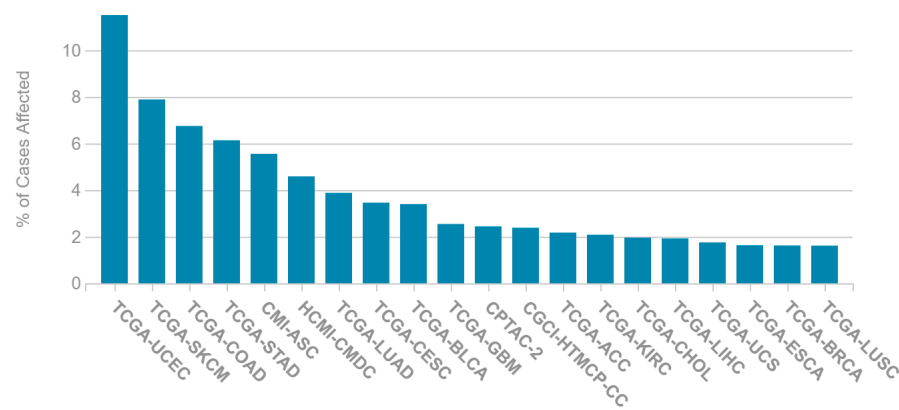
Appendix A

The Cancer Genome Atlas (TCGA) datasets were used in the following results. The Genomic Data Commons Data Portal was used to retrieve the percentage of DNA copy

number variation (CNV) and mutation of the studied genes in multiple cancer types (<https://portal.gdc.cancer.gov/>, accessed on 15 October 2021). TCGA-LUAD symbolizes lung adenocarcinoma in TCGA. TCGA-LUSC symbolizes lung squamous cell carcinoma in TCGA.

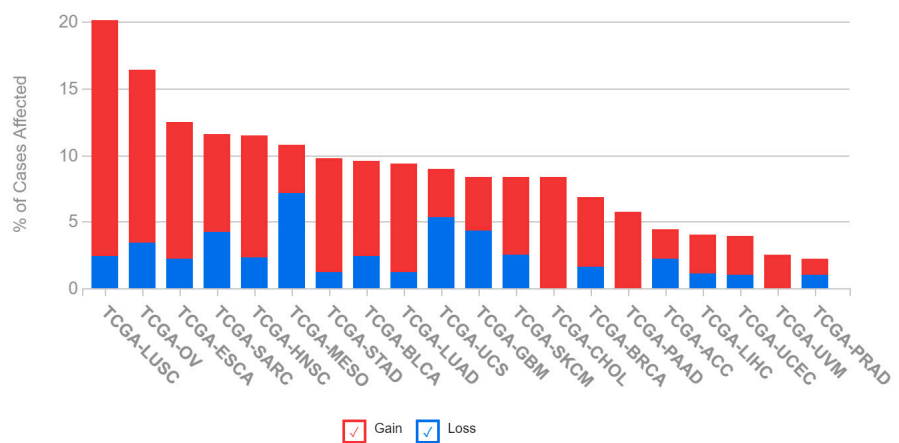


(A)



(B)

Figure A1. CNV and mutation distribution of *ABCC4* in different cancer types. (A) CNV distribution. (B) Mutation distribution.



(A)

Figure A2. Cont.

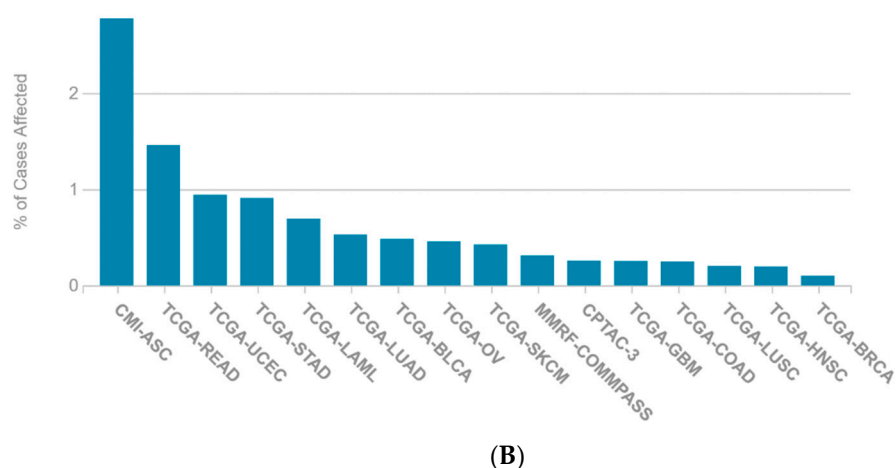


Figure A2. CNV and mutation distribution of *CCL19* in different cancer types. (A) CNV distribution. (B) Mutation distribution.

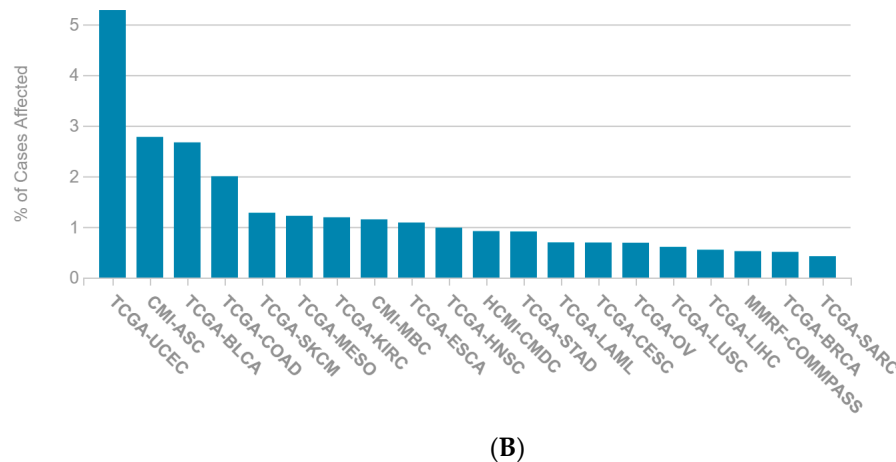
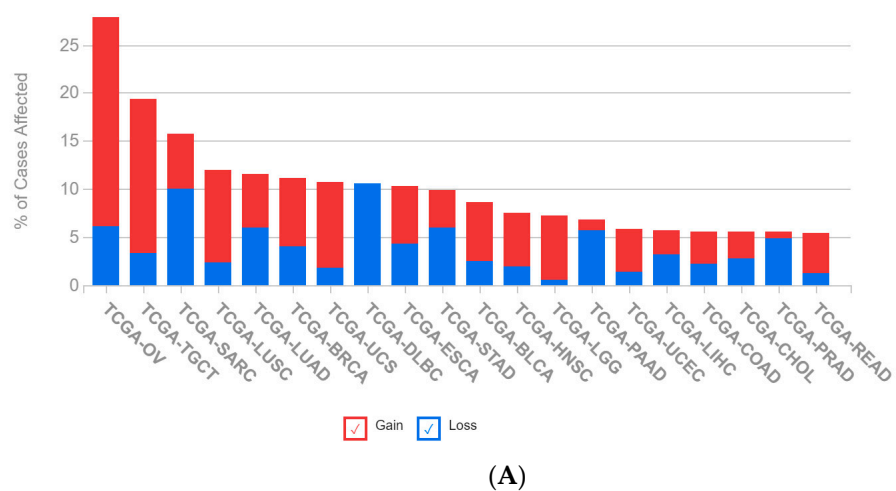
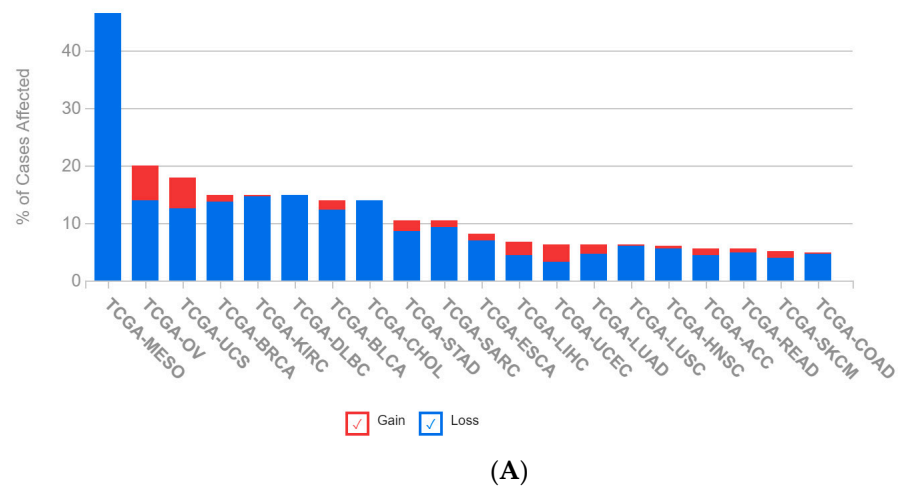
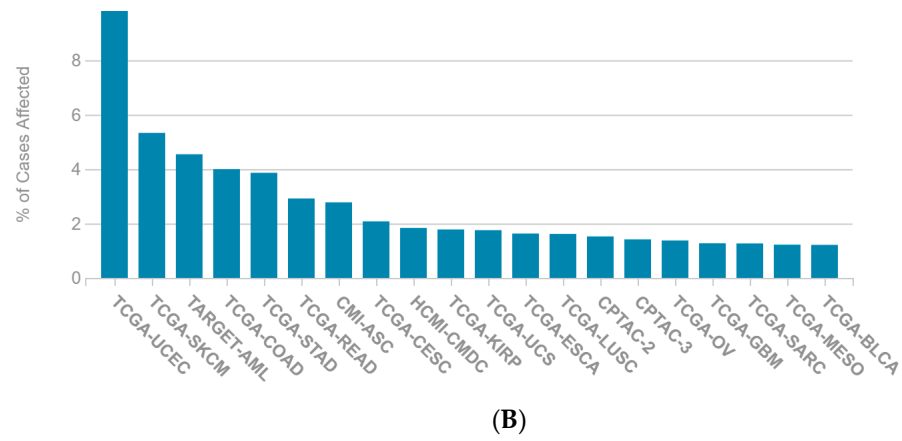


Figure A3. CNV and mutation distribution of *CD27* in different cancer types. (A) CNV distribution. (B) Mutation distribution.

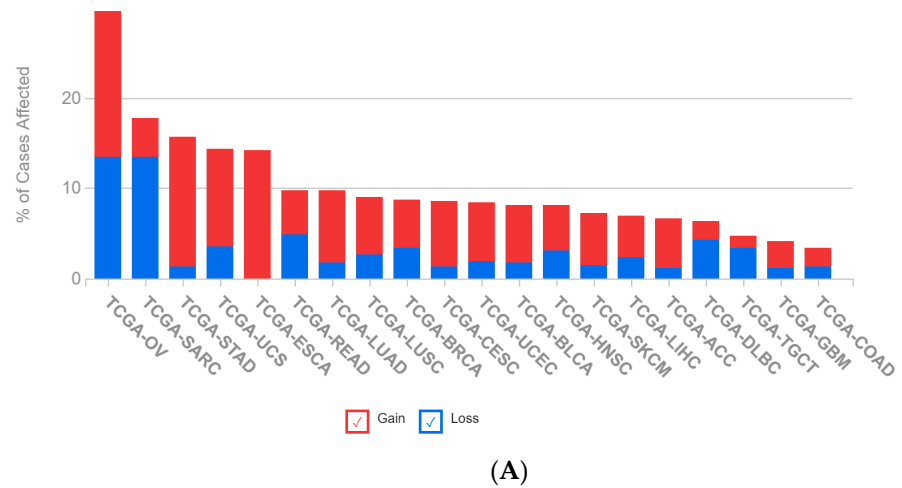


(A)



(B)

Figure A4. CNV and mutation distribution of *DAG1* in different cancer types. (A) CNV distribution. (B) Mutation distribution.



(A)

Figure A5. Cont.

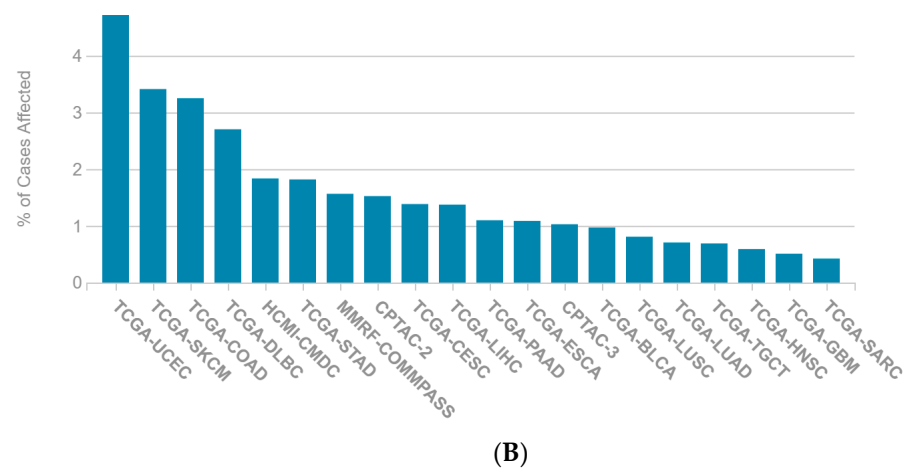


Figure A5. CNV and mutation distribution of *FUT7* in different cancer types. (A) CNV distribution. (B) Mutation distribution.

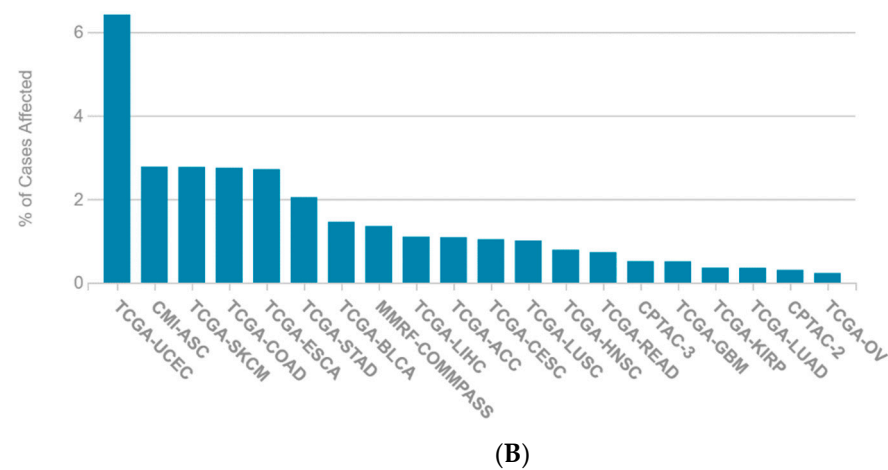
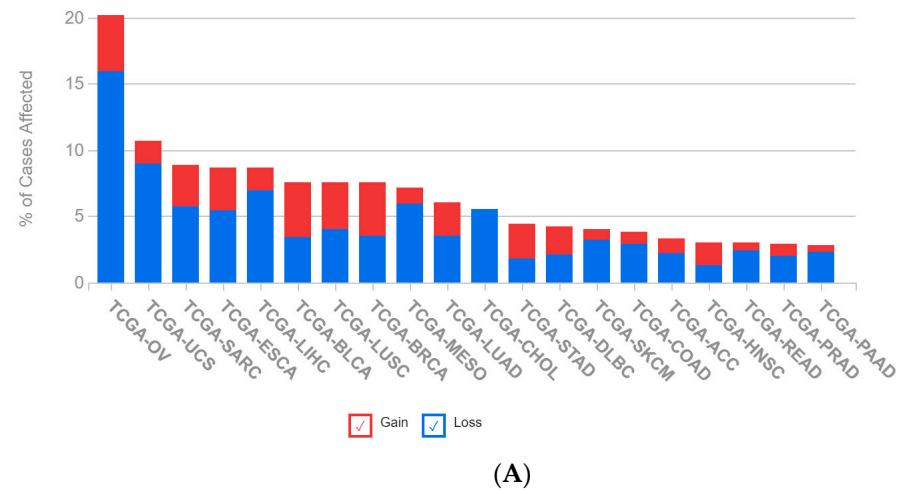


Figure A6. CNV and mutation distribution of *SLC39A8* in different cancer types. (A) CNV distribution. (B) Mutation distribution.

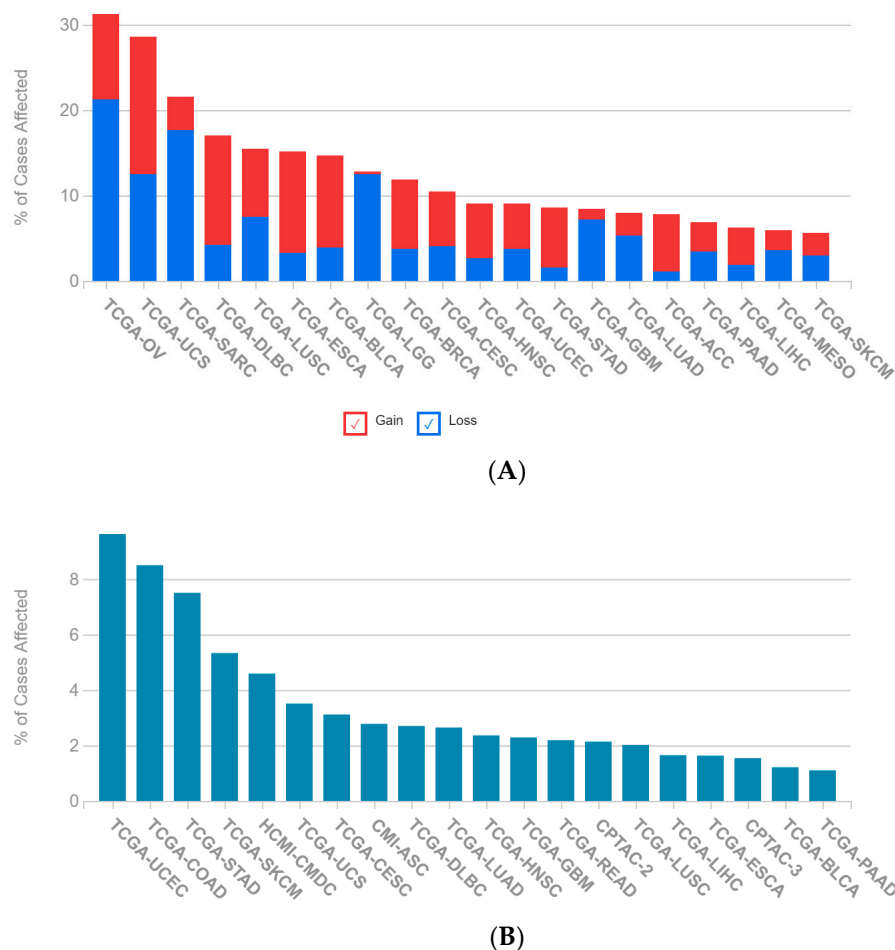


Figure A7. CNV and mutation distribution of *ZNF71* in different cancer types. **(A)** CNV distribution. **(B)** Mutation distribution.

References

- Zhang, W.; Laborde, P.M.; Coombes, K.R.; Berry, D.A.; Hamilton, S.R. Cancer genomics: Promises and complexities. *Clin. Cancer Res.* **2001**, *7*, 2159–2167.
- Mahmoud, S. Non-Smokers Are at Increasing Risk for Lung Cancer. Available online: <https://www.mdlinx.com/article/non-smokers-are-at-increasing-risk-for-lung-cancer/lfc-3757> (accessed on 19 September 2021).
- Spira, A.; Wood, A.J.J.; Ettinger, D.S. Multidisciplinary Management of Lung Cancer. *N. Eng. J. Med.* **2004**, *350*, 379–392. [[CrossRef](#)]
- Kim, H.S.; Mitsudomi, T.; Soo, R.A.; Cho, B.C. Personalized therapy on the horizon for squamous cell carcinoma of the lung. *Lung Cancer* **2013**, *80*, 249–255. [[CrossRef](#)] [[PubMed](#)]
- Wood, S.L.; Pernemalm, M.; Crosbie, P.A.; Whetton, A.D. Molecular histology of lung cancer: From targets to treatments. *Cancer Treat. Rev.* **2015**, *41*, 361–375. [[CrossRef](#)] [[PubMed](#)]
- Zheng, Y.; Jaklitsch, M.T.; Bueno, R. Neoadjuvant Therapy in Non-Small Cell Lung Cancer. *Surg. Oncol. Clin. N. Am.* **2016**, *25*, 567–584. [[CrossRef](#)] [[PubMed](#)]
- Crino, L.; Weder, W.; van Meerbeeck, J.; Felip, E.; Group, E.G.W. Early stage and locally advanced (non-metastatic) non-small-cell lung cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* **2010**, *21*, v103–v115. [[CrossRef](#)]
- Byron, E.; Pinder-Schenck, M. Systemic and Targeted Therapies for Early-Stage Lung Cancer. *Cancer Control* **2014**, *21*, 21–31. [[CrossRef](#)] [[PubMed](#)]
- FDA Approves Atezolizumab as Adjuvant Treatment for Non-Small Cell Lung Cancer. Available online: <https://www.fda.gov/drugs/resources-information-approved-drugs/fda-approves-atezolizumab-adjuvant-treatment-non-small-cell-lung-cancer> (accessed on 8 October 2021).
- Guo, N.L.; Dowlati, A.; Raese, R.A.; Dong, C.; Chen, G.; Beer, D.G.; Shaffer, J.; Singh, S.; Bokhary, U.; Liu, L.; et al. A Predictive 7-Gene Assay and Prognostic Protein Biomarkers for Non-small Cell Lung Cancer. *EBioMedicine* **2018**, *32*, 102–110. [[CrossRef](#)] [[PubMed](#)]

11. Huntley, S.; Baggott, D.M.; Hamilton, A.T.; Tran-Gyamfi, M.; Yang, S.; Kim, J.; Gordon, L.; Branscomb, E.; Stubbs, L. A comprehensive catalog of human KRAB-associated zinc finger genes: Insights into the evolutionary history of a large family of transcriptional repressors. *Genome Res.* **2006**, *16*, 669–677. [[CrossRef](#)]
12. Bruno, M.; Mahgoub, M.; Macfarlan, T.S. The Arms Race Between KRAB-Zinc Finger Proteins and Endogenous Retroelements and Its Impact on Mammals. *Ann. Rev. Genet.* **2019**, *53*, 393–416. [[CrossRef](#)] [[PubMed](#)]
13. Ecco, G.; Imbeault, M.; Trono, D. KRAB zinc finger proteins. *Development* **2017**, *144*, 2719–2729. [[CrossRef](#)]
14. Lupo, A.; Cesaro, E.; Montano, G.; Zurlo, D.; Izzo, P.; Costanzo, P. KRAB-Zinc Finger Proteins: A Repressor Family Displaying Multiple Biological Functions. *Curr. Genom.* **2013**, *14*, 268–278. [[CrossRef](#)] [[PubMed](#)]
15. Roche, J. The Epithelial-to-Mesenchymal Transition in Cancer. *Cancers* **2018**, *10*, 52. [[CrossRef](#)] [[PubMed](#)]
16. Lambert, A.W.; Pattabiraman, D.R.; Weinberg, R.A. Emerging Biological Principles of Metastasis. *Cell* **2017**, *168*, 670–691. [[CrossRef](#)] [[PubMed](#)]
17. Nieto, M.A.; Huang, R.Y.; Jackson, R.A.; Thiery, J.P. EMT: 2016. *Cell* **2016**, *166*, 21–45. [[CrossRef](#)]
18. Moustakas, A.; de Herreros, A.G. Epithelial-mesenchymal transition in cancer. *Mol. Oncol.* **2017**, *11*, 715–717. [[CrossRef](#)]
19. Ye, Q.; Mohamed, R.; Dukhlallah, D.; Gencheva, M.; Hu, G.; Pearce, M.C.; Kolluri, S.K.; Marsh, C.B.; Eubank, T.D.; Ivanov, A.V.; et al. Molecular Analysis of ZNF71 KRAB in Non-Small-Cell Lung Cancer. *Int. J. Mol. Sci.* **2021**, *22*, 3752. [[CrossRef](#)]
20. Chen, J.; Yang, H.; Teo, A.S.M.; Amer, L.B.; Sherbaf, F.G.; Tan, C.Q.; Alvarez, J.J.S.; Lu, B.; Lim, J.Q.; Takano, A.; et al. Genomic landscape of lung adenocarcinoma in East Asians. *Nat. Genet.* **2020**, *52*, 177–186. [[CrossRef](#)]
21. Block, C.J.; Dyson, G.; Campeanu, I.J.; Watzka, D.; Ratnam, M.; Wu, G. A stroma-corrected ZEB1 transcriptional signature is inversely associated with antitumor immune activity in breast cancer. *Sci. Rep.* **2019**, *9*, 17807. [[CrossRef](#)]
22. Tulchinsky, E.; Demidov, O.; Kriajevska, M.; Barlev, N.A.; Imyanitor, E. EMT: A mechanism for escape from EGFR-targeted therapy in lung cancer. *Biochim. Biophys. Acta Rev. Cancer* **2019**, *1871*, 29–39. [[CrossRef](#)]
23. Pei, L.; He, X.; Li, S.; Sun, R.; Xiang, Q.; Ren, G.; Xiang, T. KRAB zinc-finger protein 382 regulates epithelial-mesenchymal transition and functions as a tumor suppressor, but is silenced by CpG methylation in gastric cancer. *Int. J. Oncol.* **2018**, *53*, 961–972. [[CrossRef](#)]
24. Friedman, J.R.; Fredericks, W.J.; Jensen, D.E.; Speicher, D.W.; Huang, X.P.; Neilson, E.G.; Rauscher, F.J., 3rd. KAP-1, a novel corepressor for the highly conserved KRAB repression domain. *Genes Develop.* **1996**, *10*, 2067–2078. [[CrossRef](#)]
25. Chen, L.; Muñoz-Antonia, T.; Cress, W.D. Trim28 contributes to EMT via regulation of E-cadherin and N-cadherin in lung cancer cell lines. *PLoS ONE* **2014**, *9*, e101040. [[CrossRef](#)]
26. Subramanian, A.; Narayan, R.; Corsello, S.M.; Peck, D.D.; Natoli, T.E.; Lu, X.; Gould, J.; Davis, J.F.; Tubelli, A.A.; Asiedu, J.K.; et al. A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell* **2017**, *171*, 1437–1452. [[CrossRef](#)]
27. Lamb, J.; Crawford, E.D.; Peck, D.; Modell, J.W.; Blat, I.C.; Wrobel, M.J.; Lerner, J.; Brunet, J.P.; Subramanian, A.; Ross, K.N.; et al. The Connectivity Map: Using gene-expression signatures to connect small molecules, genes, and disease. *Science* **2006**, *313*, 1929–1935. [[CrossRef](#)]
28. Ishikawa, K.; Ishikawa, A.; Shoji, Y.; Imai, T. A genotoxic stress-responsive miRNA, miR-574-3p, delays cell growth by suppressing the enhancer of rudimentary homolog gene in vitro. *Int. J. Mol. Sci.* **2014**, *15*, 2971–2990. [[CrossRef](#)]
29. Zhang, D.; Wang, H.; He, H.; Niu, H.; Li, Y. Interferon induced transmembrane protein 3 regulates the growth and invasion of human lung adenocarcinoma. *Thorac. Cancer* **2017**, *8*, 337–343. [[CrossRef](#)] [[PubMed](#)]
30. Huggins, D.N.; LaRue, R.S.; Wang, Y.; Knutson, T.P.; Xu, Y.; Williams, J.W.; Schwertfeger, K.L. Characterizing Macrophage Diversity in Metastasis-Bearing Lungs Reveals a Lipid-Associated Macrophage Subset. *Cancer Res.* **2021**, *81*, 5284–5295. [[CrossRef](#)]
31. Alvarez, S.W.; Sviderskiy, V.O.; Terzi, E.M.; Papagiannakopoulos, T.; Moreira, A.L.; Adams, S.; Sabatini, D.M.; Birsoy, K.; Possemato, R. NFS1 undergoes positive selection in lung tumours and protects cells from ferroptosis. *Nature* **2017**, *551*, 639–643. [[CrossRef](#)] [[PubMed](#)]
32. Liu, J.; Liu, L.; Cao, L.; Wen, Q. Keratin 17 Promotes Lung Adenocarcinoma Progression by Enhancing Cell Proliferation and Invasion. *Med. Sci. Monit.* **2018**, *24*, 4782–4790. [[CrossRef](#)] [[PubMed](#)]
33. Quinn, J.J.; Jones, M.G.; Okimoto, R.A.; Nanjo, S.; Chan, M.M.; Yosef, N.; Bivona, T.G.; Weissman, J.S. Single-cell lineages reveal the rates, routes, and drivers of metastasis in cancer xenografts. *Science* **2021**, *371*. [[CrossRef](#)]
34. Li, T.; Fan, J.; Wang, B.; Traugh, N.; Chen, Q.; Liu, J.S.; Li, B.; Liu, X.S. TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. *Cancer Res.* **2017**, *77*, e108–e110. [[CrossRef](#)]
35. Li, T.; Fu, J.; Zeng, Z.; Cohen, D.; Li, J.; Chen, Q.; Li, B.; Liu, X.S. TIMER2.0 for analysis of tumor-infiltrating immune cells. *Nucleic Acids Res.* **2020**, *48*, W509–W514. [[CrossRef](#)] [[PubMed](#)]
36. Li, B.; Severson, E.; Pignon, J.C.; Zhao, H.; Li, T.; Novak, J.; Jiang, P.; Shen, H.; Aster, J.C.; Rodig, S.; et al. Comprehensive analyses of tumor immunity: Implications for cancer immunotherapy. *Genome Biol.* **2016**, *17*, 174. [[CrossRef](#)] [[PubMed](#)]
37. Nusinow, D.P.; Szpyt, J.; Ghandi, M.; Rose, C.M.; McDonald, E.R., 3rd; Kalocsay, M.; Jané-Valbuena, J.; Gelfand, E.; Schweppe, D.K.; Jedrychowski, M.; et al. Quantitative Proteomics of the Cancer Cell Line Encyclopedia. *Cell* **2020**, *180*, 387–402. [[CrossRef](#)]
38. Mezheyeuski, A.; Bergsland, C.H.; Backman, M.; Djureinovic, D.; Sjoblom, T.; Bruun, J.; Micke, P. Multispectral imaging for quantitative and compartment-specific immune infiltrates reveals distinct immune profiles that classify lung cancer patients. *J. Pathol.* **2018**, *244*, 421–431. [[CrossRef](#)]

39. Corsello, S.M.; Nagari, R.T.; Spangler, R.D.; Rossen, J.; Kocak, M.; Bryan, J.G.; Humeidi, R.; Peck, D.; Wu, X.; Tang, A.A.; et al. Discovering the anti-cancer potential of non-oncology drugs by systematic viability profiling. *Nat. Cancer* **2020**, *1*, 235–248. [[CrossRef](#)] [[PubMed](#)]
40. Wang, Z.Y.; Tung, S.R.; Strichartz, G.R.; Håkanson, R. Investigation of the specificity of FK 888 as a tachykinin NK1 receptor antagonist. *Br. J. Pharmacol.* **1994**, *111*, 1342–1346. [[CrossRef](#)]
41. Shabbir, M.; Stuart, R. Lestaurtinib, a multitargeted tyrosine kinase inhibitor: From bench to bedside. *Expert Opin. Investig. Drugs* **2010**, *19*, 427–436. [[CrossRef](#)]
42. Minturn, J.E.; Evans, A.E.; Villablanca, J.G.; Yanik, G.A.; Park, J.R.; Shusterman, S.; Groshen, S.; Hellriegel, E.T.; Bensen-Kennedy, D.; Matthay, K.K.; et al. Phase I trial of lestaurtinib for children with refractory neuroblastoma: A new approaches to neuroblastoma therapy consortium study. *Cancer Chemother. Pharmacol.* **2011**, *68*, 1057–1065. [[CrossRef](#)]
43. Pareek, S.; Huang, Y.; Nath, A.; Huang, R.S. The success story of drug repurposing in breast cancer. *Drug Repurposing Cancer Ther.* **2020**, 173–190. [[CrossRef](#)]
44. Cancer Stat Facts: Lung and Bronchus Cancer. Available online: <https://seer.cancer.gov/statfacts/html/lungb.html> (accessed on 26 May 2021).
45. Hachey, K.J.; Colson, Y.L. Current innovations in sentinel lymph node mapping for the staging and treatment of resectable lung cancer. *Semin. Thorac. Cardiovasc. Surg.* **2014**, *26*, 201–209. [[CrossRef](#)]
46. Zhu, C.Q.; Ding, K.; Strumpf, D.; Weir, B.A.; Meyerson, M.; Pennell, N.; Thomas, R.K.; Naoki, K.; Ladd-Acosta, C.; Liu, N.; et al. Prognostic and predictive gene signature for adjuvant chemotherapy in resected non-small-cell lung cancer. *J. Clin. Oncol.* **2010**, *28*, 4417–4424. [[CrossRef](#)]
47. Klug, A. The discovery of zinc fingers and their applications in gene regulation and genome manipulation. *Ann. Rev. Biochem.* **2010**, *79*, 213–231. [[CrossRef](#)]
48. Laity, J.H.; Lee, B.M.; Wright, P.E. Zinc finger proteins: New insights into structural and functional diversity. *Curr. Opin. Struct. Biol.* **2001**, *11*, 39–46. [[CrossRef](#)]
49. Cassandri, M.; Smirnov, A.; Novelli, F.; Pitolli, C.; Agostini, M.; Malewicz, M.; Melino, G.; Raschella, G. Zinc-finger proteins in health and disease. *Cell Death Discov.* **2017**, *3*, 17071. [[CrossRef](#)] [[PubMed](#)]
50. Yang, P.; Wang, Y.; Macfarlan, T.S. The Role of KRAB-ZFPs in Transposable Element Repression and Mammalian Evolution. *Trends Genet. TIG* **2017**, *33*, 871–881. [[CrossRef](#)]
51. Addison, J.B.; Koontz, C.; Fugett, J.H.; Creighton, C.J.; Chen, D.; Farrugia, M.K.; Padon, R.R.; Voronkova, M.A.; McLaughlin, S.L.; Livengood, R.H.; et al. KAP1 promotes proliferation and metastatic progression of breast cancer cells. *Cancer Res.* **2015**, *75*, 344–355. [[CrossRef](#)]
52. Chen, L.; Chen, D.T.; Kurtyka, C.; Rawal, B.; Fulp, W.J.; Haura, E.B.; Cress, W.D. Tripartite motif containing 28 (Trim28) can regulate cell proliferation by bridging HDAC1/E2F interactions. *J. Biol. Chem.* **2012**, *287*, 40106–40118. [[CrossRef](#)] [[PubMed](#)]
53. Czerwińska, P.; Mazurek, S.; Wiznerowicz, M. The complexity of TRIM28 contribution to cancer. *J. Biomed. Sci.* **2017**, *24*, 63. [[CrossRef](#)]
54. Zhang, X.; Jiang, G.; Wu, J.; Zhou, H.; Zhang, Y.; Miao, Y.; Feng, Y.; Yu, J. Zinc finger protein 668 suppresses non-small cell lung cancer invasion and migration by downregulating Snail and upregulating E-cadherin and zonula occludens-1. *Oncol. Lett.* **2018**, *15*, 3806–3813. [[CrossRef](#)]
55. Warzecha, C.C.; Carstens, R.P. Complex changes in alternative pre-mRNA splicing play a central role in the epithelial-to-mesenchymal transition (EMT). *Semin. Cancer Biol.* **2012**, *22*, 417–427. [[CrossRef](#)]
56. Li, J.; Choi, P.S.; Chaffer, C.L.; Labella, K.; Hwang, J.H.; Giacomelli, A.O.; Kim, J.W.; Ilic, N.; Doench, J.G.; Ly, S.H.; et al. An alternative splicing switch in FLNB promotes the mesenchymal cell state in human breast cancer. *eLife* **2018**, *7*. [[CrossRef](#)]
57. Waldman, A.D.; Fritz, J.M.; Lenardo, M.J. A guide to cancer immunotherapy: From T cell basic science to clinical practice. *Nat. Rev. Immunol.* **2020**, *20*, 651–668. [[CrossRef](#)]
58. Buchan, S.L.; Rogel, A.; Al-Shamkhani, A. The immunobiology of CD27 and OX40 and their potential as targets for cancer immunotherapy. *Blood* **2017**, *131*, 39–48. [[CrossRef](#)]
59. Marin-Acevedo, J.A.; Dholaria, B.; Soyano, A.E.; Knutson, K.L.; Chumsri, S.; Lou, Y. Next generation of immune checkpoint therapy in cancer: New developments and challenges. *J. Hematol. Oncol.* **2018**, *11*, 39. [[CrossRef](#)]
60. Burris, H.A.; Infante, J.R.; Ansell, S.M.; Nemunaitis, J.J.; Weiss, G.R.; Villalobos, V.M.; Sikic, B.I.; Taylor, M.H.; Northfelt, D.W.; Carson, W.E., 3rd; et al. Safety and Activity of Varlilumab, a Novel and First-in-Class Agonist Anti-CD27 Antibody, in Patients With Advanced Solid Tumors. *J. Clin. Oncol.* **2017**, *35*, 2028–2036. [[CrossRef](#)] [[PubMed](#)]
61. Ahrends, T.; Bąbała, N.; Xiao, Y.; Yagita, H.; van Eenennaam, H.; Borst, J. CD27 Agonism Plus PD-1 Blockade Recapitulates CD4+ T-cell Help in Therapeutic Anticancer Vaccination. *Cancer Res.* **2016**, *76*, 2921–2931. [[CrossRef](#)] [[PubMed](#)]
62. Buchan, S.L.; Fallatah, M.; Thirdborough, S.M.; Taraban, V.Y.; Rogel, A.; Thomas, L.J.; Penfold, C.A.; He, L.Z.; Curran, M.A.; Keler, T.; et al. PD-1 Blockade and CD27 Stimulation Activate Distinct Transcriptional Programs That Synergize for CD8(+) T-Cell-Driven Antitumor Immunity. *Clin. Cancer Res.* **2018**, *24*, 2383–2394. [[CrossRef](#)]
63. Forde, P.M.; Chaft, J.E.; Smith, K.N.; Anagnostou, V.; Cottrell, T.R.; Hellmann, M.D.; Zahurak, M.; Yang, S.C.; Jones, D.R.; Broderick, S.; et al. Neoadjuvant PD-1 Blockade in Resectable Lung Cancer. *N. Engl. J. Med.* **2018**, 378. [[CrossRef](#)] [[PubMed](#)]
64. Camidge, D.R.; Doebele, R.C.; Kerr, K.M. Comparing and contrasting predictive biomarkers for immunotherapy and targeted therapy of NSCLC. *Nat. Rev. Clin. Oncol.* **2019**, *16*, 341–355. [[CrossRef](#)]

65. Ye, Q.; Singh, S.; Qian, P.R.; Guo, N.L. Immune-Omics Networks of CD27, PD1, and PDL1 in Non-Small Cell Lung Cancer. *Cancers* **2021**, *13*, 4296. [[CrossRef](#)]
66. Andersen, M.E.; Yang, R.S.; French, C.T.; Chubb, L.S.; Dennison, J.E. Molecular circuits, biological switches, and nonlinear dose-response relationships. *Environ. Health Perspect.* **2002**, *110*, 971–978. [[CrossRef](#)]
67. Kim, E.H.; Park, A.K.; Dong, S.M.; Ahn, J.H.; Park, W.Y. Global analysis of CpG methylation reveals epigenetic control of the radiosensitivity in lung cancer cell lines. *Oncogene* **2010**, *29*, 4725–4731. [[CrossRef](#)] [[PubMed](#)]
68. Lee, M.N.; Tseng, R.C.; Hsu, H.S.; Chen, J.Y.; Tzao, C.; Ho, W.L.; Wang, Y.C. Epigenetic inactivation of the chromosomal stability control genes BRCA1, BRCA2, and XRCC5 in non-small cell lung cancer. *Clin. Cancer Res.* **2007**, *13*, 832–838. [[CrossRef](#)] [[PubMed](#)]
69. Adams, M.N.; Burgess, J.T.; He, Y.; Gately, K.; Snell, C.; Zhang, S.D.; Hooper, J.D.; Richard, D.J.; O’Byrne, K.J. Expression of CDCA3 Is a Prognostic Biomarker and Potential Therapeutic Target in Non-Small Cell Lung Cancer. *J. Thorac. Oncol.* **2017**, *12*, 1071–1084. [[CrossRef](#)] [[PubMed](#)]
70. Maier, T.; Güell, M.; Serrano, L. Correlation of mRNA and protein in complex biological samples. *FEBS Lett.* **2009**, *583*, 3966–3973. [[CrossRef](#)]
71. Pushpakom, S.; Iorio, F.; Eyers, P.A.; Escott, K.J.; Hopper, S.; Wells, A.; Doig, A.; Guilliams, T.; Latimer, J.; McNamee, C.; et al. Drug repurposing: Progress, challenges and recommendations. *Nat. Rev. Drug Discov.* **2019**, *18*, 41–58. [[CrossRef](#)] [[PubMed](#)]
72. Dimitrakopoulos, F.D.; Kottorou, A.E.; Kalofonou, M.; Kalofonos, H.P. The Fire Within: NF- κ B Involvement in Non-Small Cell Lung Cancer. *Cancer Res.* **2020**, *80*, 4025–4036. [[CrossRef](#)] [[PubMed](#)]
73. Cai, Z.; Tchou-Wong, K.M.; Rom, W.N. NF-kappaB in lung tumorigenesis. *Cancers* **2011**, *3*, 4258–4268. [[CrossRef](#)] [[PubMed](#)]
74. Hopewell, E.L.; Zhao, W.; Fulp, W.J.; Bronk, C.C.; Lopez, A.S.; Massengill, M.; Antonia, S.; Celis, E.; Haura, E.B.; Enkemann, S.A.; et al. Lung tumor NF- κ B signaling promotes T cell-mediated immune surveillance. *J. Clin. Investig.* **2013**, *123*, 2509–2522. [[CrossRef](#)] [[PubMed](#)]
75. DiDonato, J.A.; Mercurio, F.; Karin, M. NF- κ B and the link between inflammation and cancer. *Immunol. Rev.* **2012**, *246*, 379–400. [[CrossRef](#)] [[PubMed](#)]
76. Hinz, M.; Krappmann, D.; Eichten, A.; Heder, A.; Scheidereit, C.; Strauss, M. NF-kappaB function in growth control: Regulation of cyclin D1 expression and G0/G1-to-S-phase transition. *Mol. Cell. Biol.* **1999**, *19*, 2690–2698. [[CrossRef](#)] [[PubMed](#)]
77. Ho, Y.S.; Chen, C.H.; Wang, Y.J.; Pestell, R.G.; Albanese, C.; Chen, R.J.; Chang, M.C.; Jeng, J.H.; Lin, S.Y.; Liang, Y.C.; et al. Tobacco-specific carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) induces cell proliferation in normal human bronchial epithelial cells through NFkappaB activation and cyclin D1 up-regulation. *Toxicol. Appl. Pharmacol.* **2005**, *205*, 133–148. [[CrossRef](#)] [[PubMed](#)]
78. Kumar, M.; Allison, D.F.; Baranova, N.N.; Wamsley, J.J.; Katz, A.J.; Bekiranov, S.; Jones, D.R.; Mayo, M.W. NF- κ B regulates mesenchymal transition for the induction of non-small cell lung cancer initiating cells. *PLoS ONE* **2013**, *8*, e68597. [[CrossRef](#)]
79. Lu, Q.; Lu, S.; Huang, L.; Wang, T.; Wan, Y.; Zhou, C.X.; Zhang, C.; Zhang, Z.; Li, X. The expression of V-ATPase is associated with drug resistance and pathology of non-small-cell lung cancer. *Diagn. Pathol.* **2013**, *8*, 145. [[CrossRef](#)]
80. Vasaikar, S.V.; Straub, P.; Wang, J.; Zhang, B. LinkedOmics: Analyzing multi-omics data within and across 32 cancer types. *Nucleic Acids Res.* **2018**, *46*, D956–D963. [[CrossRef](#)] [[PubMed](#)]
81. Starczynowski, D.T.; Lockwood, W.W.; Delehouzee, S.; Chari, R.; Wegrzyn, J.; Fuller, M.; Tsao, M.S.; Lam, S.; Gazdar, A.F.; Lam, W.L.; et al. TRAF6 is an amplified oncogene bridging the RAS and NF-kappaB pathways in human lung cancer. *J. Clin. Investig.* **2011**, *121*, 4095–4105. [[CrossRef](#)]
82. Micke, P.; Edlund, K.; Holmberg, L.; Kultima, H.G.; Mansouri, L.; Ekman, S.; Bergqvist, M.; Scheibenflug, L.; Lamberg, K.; Myrdal, G.; et al. Gene copy number aberrations are associated with survival in histologic subgroups of non-small cell lung cancer. *J. Thorac. Oncol.* **2011**, *6*, 1833–1840. [[CrossRef](#)] [[PubMed](#)]
83. Jabs, V.; Edlund, K.; Konig, H.; Grinberg, M.; Madjar, K.; Rahnenfuhrer, J.; Ekman, S.; Bergkvist, M.; Holmberg, L.; Ickstadt, K.; et al. Integrative analysis of genome-wide gene copy number changes and gene expression in non-small cell lung cancer. *PLoS ONE* **2017**, *12*, e0187246. [[CrossRef](#)]
84. Patro, R.; Duggal, G.; Love, M.I.; Irizarry, R.A.; Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **2017**, *14*, 417–419. [[CrossRef](#)]
85. Zhan, C.; Zhang, Y.; Ma, J.; Wang, L.; Jiang, W.; Shi, Y.; Wang, Q. Identification of reference genes for qRT-PCR in human lung squamous-cell carcinoma by RNA-Seq. *Acta Biochim. Biophys. Sin.* **2014**, *46*, 330–337. [[CrossRef](#)]
86. Walter, R.F.; Werner, R.; Vollbrecht, C.; Hager, T.; Flom, E.; Christoph, D.C.; Schmeller, J.; Schmid, K.W.; Wohlschlaeger, J.; Mairinger, F.D. ACTB, CDKN1B, GAPDH, GRB2, RHOA and SDCBP Were Identified as Reference Genes in Neuroendocrine Lung Cancer via the nCounter Technology. *PLoS ONE* **2016**, *11*, e0165181. [[CrossRef](#)] [[PubMed](#)]
87. Saviozzi, S.; Cordero, F.; Lo Iacono, M.; Novello, S.; Scagliotti, G.V.; Calogero, R.A. Selection of suitable reference genes for accurate normalization of gene expression profile studies in non-small cell lung cancer. *BMC Cancer* **2006**, *6*, 200. [[CrossRef](#)]
88. Chang, Y.C.; Ding, Y.; Dong, L.; Zhu, L.J.; Jensen, R.V.; Hsiao, L.L. Differential expression patterns of housekeeping genes increase diagnostic and prognostic value in lung cancer. *PeerJ* **2018**, *6*, e4719. [[CrossRef](#)] [[PubMed](#)]
89. Yang, X.; Amgad, M.; Cooper, L.A.D.; Du, Y.; Fu, H.; Ivanov, A.A. High expression of MKK3 is associated with worse clinical outcomes in African American breast cancer patients. *J. Translat. Med.* **2020**, *18*. [[CrossRef](#)] [[PubMed](#)]
90. Hildebrand, D.K.; Lang, J.D.; Rosenthal, H. *Prediction Analysis of Cross Classifications*; John Wiley & Sons: New York, NY, USA, 1977.

91. Guo, N.L.; Wan, Y.W.; Bose, S.; Denvir, J.; Kashon, M.L.; Andrew, M.E. A novel network model identified a 13-gene lung cancer prognostic signature. *Int. J. Comput. Biol. Drug Des.* **2011**, *4*, 19–39. [[CrossRef](#)] [[PubMed](#)]
92. Guo, N.L.; Wan, Y.W. Pathway-based identification of a smoking associated 6-gene signature predictive of lung cancer risk and survival. *Artif. Intell. Med.* **2012**, *55*, 97–105. [[CrossRef](#)]
93. DepMap. DepMap 20Q2 Public. 2020. Available online: https://figshare.com/articles/dataset/DepMap_20Q2_Public/12280541 (accessed on 1 April 2021).
94. Aguet, F. TOPMed RNA-Seq Pipeline Harmonization Summary. Available online: https://github.com/broadinstitute/gtex-pipeline/blob/master/TOPMed_RNAseq_pipeline.md (accessed on 24 December 2021).
95. Meyers, R.M.; Bryan, J.G.; McFarland, J.M.; Weir, B.A.; Sizemore, A.E.; Xu, H.; Dharia, N.V.; Montgomery, P.G.; Cowley, G.S.; Pantel, S.; et al. Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nat. Genet.* **2017**, *49*, 1779–1784. [[CrossRef](#)] [[PubMed](#)]
96. Dempster, J.M.; Rossen, J.; Kazachkova, M.; Pan, J.; Kugener, G.; Root, D.E.; Tsherniak, A. Extracting Biological Insights from the Project Achilles Genome-Scale CRISPR Screens in Cancer Cell Lines. *bioRxiv* **2019**. [[CrossRef](#)]
97. McFarland, J.M.; Ho, Z.V.; Kugener, G.; Dempster, J.M.; Montgomery, P.G.; Bryan, J.G.; Krill-Burger, J.M.; Green, T.M.; Vazquez, F.; Boehm, J.S.; et al. Improved estimation of cancer dependencies from large-scale RNAi screens using model-based normalization and data integration. *Nat. Commun.* **2018**, *9*, 4610. [[CrossRef](#)] [[PubMed](#)]
98. Yang, W.; Soares, J.; Greninger, P.; Edelman, E.J.; Lightfoot, H.; Forbes, S.; Bindal, N.; Beare, D.; Smith, J.A.; Thompson, I.R.; et al. Genomics of Drug Sensitivity in Cancer (GDSC): A resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res.* **2013**, *41*, D955–D961. [[CrossRef](#)] [[PubMed](#)]
99. Addison, J.B.; Voronkova, M.A.; Fugett, J.H.; Lin, C.C.; Linville, N.C.; Trinh, B.; Livengood, R.H.; Smolkin, M.B.; Schaller, M.D.; Ruppert, J.M.; et al. Functional Hierarchy and Cooperation of EMT Master Transcription Factors in Breast Cancer Metastasis. *Mol. Cancer Res.* **2021**, *19*, 784–798. [[CrossRef](#)] [[PubMed](#)]
100. Team, R. *RStudio: Integrated Development Environment for R, 1.4.1106*; RStudio, PBC: Boston, MA, USA.