



Article

Molecular Analysis of *ZNF71 KRAB* in Non-Small-Cell Lung Cancer

Qing Ye ^{1,2}, Rehab Mohamed ¹, Duaa Dakhllallah ^{1,3,4} , Marieta Gencheva ³, Gangqing Hu ^{1,3} , Martin C. Pearce ⁵, Siva Kumar Kolluri ⁵, Clay B. Marsh ⁶, Timothy D. Eubank ^{1,3} , Alexey V. Ivanov ^{1,7} and Nancy Lan Guo ^{1,8,*}

- ¹ WVU Cancer Institute, West Virginia University, Morgantown, WV 26506, USA; qiyemix@mix.wvu.edu (Q.Y.); romohamed@mix.wvu.edu (R.M.); duaadakhllallah@hsc.wvu.edu (D.D.); michael.hu@hsc.wvu.edu (G.H.); tdeubank@hsc.wvu.edu (T.D.E.); aivanov@hsc.wvu.edu (A.V.I.)
 - ² Lane Department of Computer Science and Electrical Engineering, West Virginia University, Morgantown, WV 26506, USA
 - ³ Department of Microbiology, Immunology & Cell Biology, West Virginia University, Morgantown, WV 26506, USA; mgencheva@hsc.wvu.edu
 - ⁴ Institute of Global Health and Human Ecology, School of Sciences & Engineering, The American University of Cairo, New Cairo 11835, Egypt
 - ⁵ Cancer Research Laboratory, Department of Environmental and Molecular Toxicology, Oregon State University, Corvallis, OR 97331, USA; pearcem@oregonstate.edu (M.C.P.); Siva.Kolluri@oregonstate.edu (S.K.K.)
 - ⁶ Department of Medicine, West Virginia University, Morgantown, WV 26506, USA; cbmarsh@hsc.wvu.edu
 - ⁷ Department of Biochemistry, West Virginia University, Morgantown, WV 26506, USA
 - ⁸ Department of Occupational and Environmental Health Sciences, West Virginia University, Morgantown, WV 26506, USA
- * Correspondence: lguo@hsc.wvu.edu; Tel.: +1-304-293-6455



Citation: Ye, Q.; Mohamed, R.; Dakhllallah, 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. <https://doi.org/10.3390/ijms22073752>

Academic Editors: Elisabetta Ferretti, Zein Mersini Besharat and Agnese Po

Received: 25 February 2021

Accepted: 30 March 2021

Published: 4 April 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/>).

Abstract: Our previous study found that zinc finger protein 71 (*ZNF71*) mRNA expression was associated with chemosensitivity and its protein expression was prognostic of non-small-cell lung cancer (NSCLC). The Krüppel associated box (*KRAB*) transcriptional repression domain is commonly present in human zinc finger proteins, which are linked to imprinting, silencing of repetitive elements, proliferation, apoptosis, and cancer. This study revealed that *ZNF71 KRAB* had a significantly higher expression than the *ZNF71 KRAB*-less isoform in NSCLC tumors ($n = 197$) and cell lines ($n = 117$). Patients with higher *ZNF71 KRAB* expression had a significantly worse survival outcome than patients with lower *ZNF71 KRAB* expression (log-rank $p = 0.04$; hazard ratio (HR): 1.686 [1.026, 2.771]), whereas *ZNF71* overall and *KRAB*-less expression levels were not prognostic in the same patient cohort. *ZNF71 KRAB* expression was associated with epithelial-to-mesenchymal transition (EMT) in both patient tumors and cell lines. *ZNF71 KRAB* was overexpressed in NSCLC cell lines resistant to docetaxel and paclitaxel treatment compared to chemo-sensitive cell lines, consistent with its association with poor prognosis in patients. Therefore, *ZNF71 KRAB* isoform is a more effective prognostic factor than *ZNF71* overall and *KRAB*-less expression for NSCLC. Functional analysis using CRISPR-Cas9 and RNA interference (RNAi) screening data indicated that a knockdown/knockout of *ZNF71* did not significantly affect NSCLC cell proliferation in vitro.

Keywords: *KRAB* isoform; zinc finger protein; EMT; prognosis; chemoresponse; proliferation; CRISPR-Cas9; RNAi

1. Introduction

Lung cancer is the second most common cancer and remains the leading cause of cancer-related mortality in the U.S. Non-small-cell lung cancer (NSCLC) accounts for 84% of lung cancer cases. Major histological subtypes of NSCLC include lung adenocarcinoma, squamous cell carcinoma, and large cell carcinoma. In 2020, an estimated 228,820 adults (116,300 men and 112,520 women) in the US were diagnosed with lung cancer [1]. About

30% to 55% of the patients with NSCLC develop recurrence and die of the disease within 5 years of the surgical removal of their tumors [2]. According to the current practice guidelines, NSCLC patients with stage 2 and above receive chemotherapy, with additional radiation for stage 3A patients [3]. While adjuvant chemotherapy of stage 2 and stage 3 disease has resulted in 10–15% increase in overall survival [4], the prognosis for early-stage NSCLC remains poor [5], indicating that some patients may not benefit from it. To date, physicians do not have a precise tool to identify patients with resectable NSCLC that are likely to develop tumor recurrence or metastasis. Our previous study developed a 7-gene assay for NSCLC prognosis and prediction of chemotherapeutic benefits [6]. The ability of this gene assay to identify those at a high risk for recurrence or metastasis would potentially inform selection of specific adjuvant chemotherapy for these patients.

Among this 7-gene signature, mRNA expression of zinc finger protein 71 (*ZNF71*) was positively associated with survival in patients who received cisplatin, carboplatin, and Taxol in the studied cohorts, indicating its association with chemosensitivity [6]. Although *ZNF71* mRNA expression was not associated with NSCLC survival in the overall patient cohorts analyzed in qRT-PCR, higher *ZNF71* protein expression quantified with AQUA was associated with a more favorable survival outcome in two separate NSCLC cohorts ($n = 291$) using tissue microarrays (TMA) [6]. Zinc finger proteins (ZNFs) are involved in DNA repair, degradation of proteins, signal transductions, migration of cells, regulation of apoptosis, lipid binding, and transcription regulation [7,8]. The Krüppel associated box (KRAB) is a transcriptional repression domain and is commonly present in human zinc finger protein-based transcription factors, i.e., KRAB zinc finger proteins (KRAB-ZFPs) [9–11]. Transcriptional repression mediated by KRAB-ZFPs is linked to cell proliferation, differentiation, apoptosis, and cancer [12]. *ZNF71* (*EZF1T*) was first identified by Mataki et al. [13] as a ZFP induced by tumor necrosis factor α (TNF- α) in human umbilical vein endothelial cells. Single-nucleotide polymorphisms of *ZNF71* were found to be linked with total serum IgE in Korean asthmatics in a genome-wide association study [14]. To the best of our knowledge, there has been no report on molecular analysis of *ZNF71* KRAB isoform in cancer.

Epithelial-to-mesenchymal transition (EMT) is a highly dynamic process in which epithelial cells can convert to a mesenchymal phenotype. EMT is also reversible by the mesenchymal-to-epithelial transition (MET). Emerging evidence reveals the involvement of EMT in tumor progression, metastasis, and resistance to cancer treatment [15–18]. However, the involvement of EMT in cancer patient outcomes remains controversial [19]. Recent studies have evaluated EMT as well as stromal and immune infiltration in tumors using transcriptional profiles [19–21]. This study sought to investigate the expression of *ZNF71* KRAB and KRAB-less isoforms in NSCLC tumors and cell lines and their association with patient outcome and chemotherapeutic response. Tripartite motif containing 28 (TRIM28) protein, a universal co-factor for KRAB-ZFP transcription factors [22], contributes to EMT and might be important for tumor metastasis in lung cancer [23]. Based on the link between KRAB-ZFPs and EMT [24], this study sought to explore whether NSCLC tumors with different EMT characteristics have a distinct survival outcome and the association of *ZNF71* isoforms and EMT. Molecular functions of *ZNF71* in NSCLC proliferation were evaluated with CRISPR-Cas9 and RNAi approaches. The overall study scheme is provided in Figure 1.

Specifically, mRNA expression of *ZNF71* isoforms and their prognostic implications were analyzed with public RNA-seq data of NSCLC patient tumors ($n = 197$) [25] and cell lines ($n = 117$) [26]. A 14-gene EMT classifier was constructed to evaluate the hybrid EMT states in NSCLC tumors, and *ZEB1* expression was used to evaluate the EMT states in cell lines. This EMT classification was further validated using stromal infiltration scores computed with software ESTIMATE [20]. The association between *ZNF71* overall and isoform expression and EMT was examined in NSCLC tumors and cell lines. Functional assessment of *ZNF71* in NSCLC proliferation was evaluated with public CRISPR-Cas9 [27] and RNAi [28] screening data. Finally, the association of the *ZNF71* isoform and overall ex-

pression and chemoresponse to nine drugs commonly used to treat NSCLC was examined in cell lines (Figure 1).

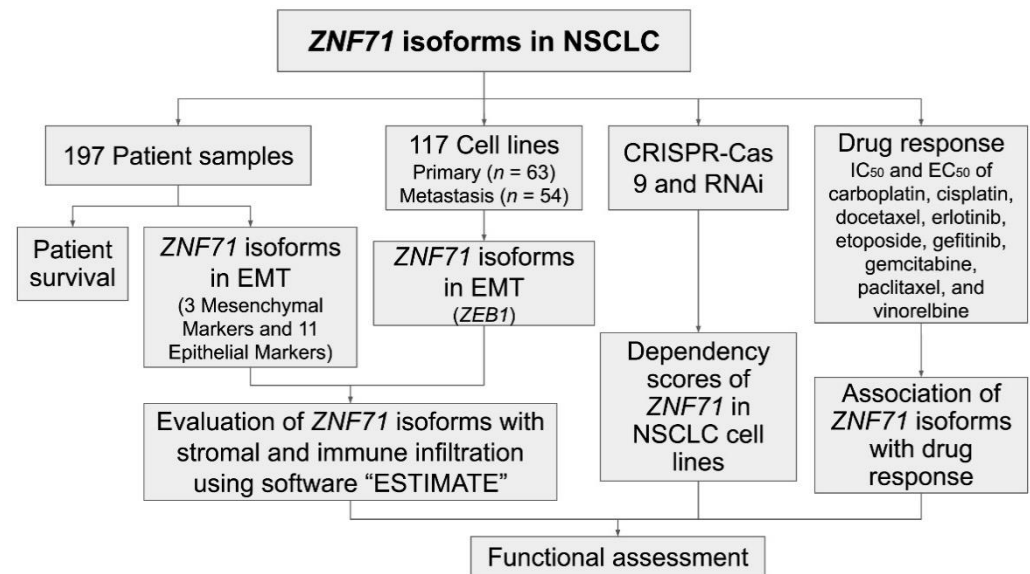


Figure 1. Overall study scheme. The arrows indicate analysis flow.

2. Results

2.1. Expression of ZNF71 Isoforms in NSCLC Tumors and Cell Lines

ZNF71 gene is comprised of four exons, where exons 1 and 2 code for 180nt 5'UTR and the first 11 amino acids of the protein. Exon 3 codes for the next 43 amino acids (aa), encompassing most of the KRAB repression domain. Exon 4, the longest, codes for the remainder of the protein, including the 13-zinc-finger putative DNA binding domain and the predicted ~4 kb 3'UTR. The third KRAB-domain-containing exon could be alternatively spliced out to produce a KRAB-less isoform that will produce a shorter in-frame protein encoded by the last exon, i.e., exon 4. Approximately half of 800 human C2H2-type ZNF genes contain the evolutionary conserved one-exon-encoded KRAB domain, which could be alternatively spiced [9–11]. We analyzed ZNF71 isoform expression in the RNA-seq dataset GSE81089 of NSCLC tumor samples ($n = 197$) [25] and correlated the analysis with patient outcomes. Patient clinical characteristics is provided in Table A1.

The ZNF71 KRAB isoform (ZNF71_203_ENST00000599599) had a significantly higher expression ($p < 0.001$, t -tests) than the KRAB-less isoform (ZNF71_201_ENST00000328070) in NSCLC patient tumor samples ($n = 197$; Figure 2A) and cell lines ($n = 117$; Figure 2B). The expression of ZNF71 KRAB and KRAB-less isoforms was significantly correlated in both patient tumor samples ($p < 8.6 \times 10^{-15}$, Pearson's correlation, Figure 2C) and cell lines ($p < 2.3 \times 10^{-8}$, Pearson's correlation, Figure 2D). The expression of overall ZNF71 and its isoforms was not significantly different among different histological subtypes in patient tumors (ANOVA tests; Figure A1). The expression of ZNF71 overall and the KRAB isoform was significantly different among histological subtypes in NSCLC cell lines ($p < 0.05$, ANOVA tests; Figure A2). In the studied cell lines, large cell carcinoma had the highest ZNF71 overall expression. Adenosquamous carcinoma had the highest ZNF71 KRAB expression. Squamous cell carcinoma had the lowest expression of both ZNF71 overall and KRAB isoform (Figure A2).

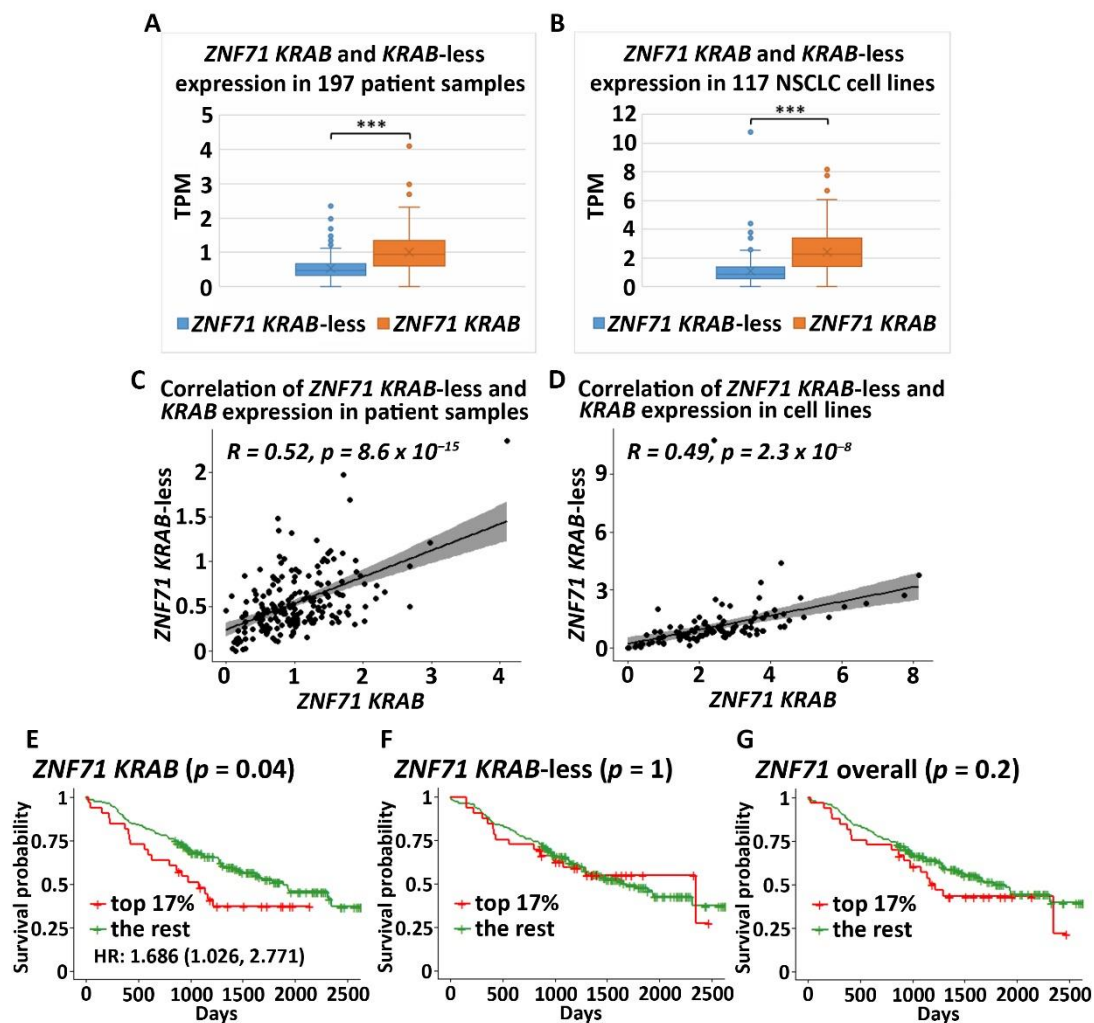


Figure 2. Zinc finger protein 71 Krüppel associated box (*ZNF71 KRAB*) and *KRAB-less* expression non-small-cell lung cancer (NSCLC) patient tumors and cell lines (A). Comparison of the expression of *ZNF71 KRAB* and *KRAB-less* isoforms in NSCLC tumors (** $p < 0.001$, two-sample *t*-tests). (B) Comparison of the expression of *ZNF71 KRAB* and *KRAB-less* isoforms in NSCLC cell lines (** $p < 0.001$, two-sample *t*-tests). (C) Pearson’s correlation of the expression of *ZNF71 KRAB* and *KRAB-less* isoforms in patient tumors. (D) Pearson’s correlation of the expression of *ZNF71 KRAB* and *KRAB-less* isoforms in cell lines. (E–G) Kaplan–Meier survival analyses of patient tumors grouped by *ZNF71 KRAB* isoform expression (E), *ZNF71 KRAB-less* isoform expression (F), and *ZNF71* overall expression (G), with top 17% expression versus the rest, respectively. When *ZNF71 KRAB* was expressed higher than transcripts per million (TPM) of 1.5 (top 17%), the patients survived for a significantly shorter length of time.

To evaluate the prognostic performance of *ZNF71 KRAB*, patient tumor samples were divided into two groups using a cutoff of *ZNF71 KRAB* expression level measured with transcripts per million (TPM) of 1.5, which corresponds to the top 17% of *ZNF71 KRAB* expression versus the rest in the patient cohort. The cutoff value of TPM of 1.5 was chosen because TPM values are commonly normalized to the value of either 1 or 2 in RNA-seq raw data processing. The results showed that when *ZNF71 KRAB* isoform was expressed higher, patients survived for a significantly shorter time (log-rank $p = 0.04$, Kaplan–Meier analysis; Figure 2E), with a hazard ratio of 1.686 [1.026, 2.771]. In contrast, *ZNF71* overall expression or *ZNF71 KRAB-less* isoform expression did not generate significant prognostic stratification in the patient cohort using the same cutoff (top 17% expression level versus the rest) in Kaplan–Meier survival analyses (Figure 2F,G). These results indicate that the *ZNF71 KRAB* isoform is a more accurate prognostic factor for NSCLC than *ZNF71* overall expression and the *KRAB-less* isoform.

2.2. Association of ZNF71 KRAB Isoform with EMT

The above RNA-seq results were confirmed in qRT-PCR assays, which showed higher ZNF71 KRAB expression than the KRAB-less isoform in NSCLC cell lines (Figure 3A). Unfortunately, the ZNF71 antibody (Abcam, Cambridge, UK; ab87250) used in our previous study [6] was discontinued. We tried several other commercially available ZNF71 antibodies, but none of them was able to detect ZNF71 protein in our panel of NSCLC cell lines (not shown). After overexpression, ZNF71 protein expression (GeneTex, Irwin, CA, USA; Cat. No. GTX116553) was observed in HEK-293T cells (Figure 3B) in Western blots. The EMT properties of the NSCLC cell lines were tested using 3 mesenchymal markers (ZEB1, VIM, and FN1) and 11 epithelial markers (CDH1, EPCAM, ESRP1, ESRP2, DDR1, CTNNB1, CD24, CLDN7, KRT8, KRT19, and RAB25) using Western blots (Figure 3B). These 14 EMT markers were used to build an EMT classifier to divide patient samples into four groups based on the rank of the average expression values of all markers (Figure 3C). The average expression of 3 mesenchymal markers and that of 11 epithelial markers were computed for each patient sample, respectively. Based on the average epithelial and mesenchymal expression rank, patient samples were categorized into four EMT phenotypes: ranked as top 50% in both mesenchymal and epithelial (named High expression overlap), ranked as top 50% in mesenchymal but bottom 50% in epithelial (named Mesenchymal), ranked as top 50% in epithelial but bottom 50% in mesenchymal (named Epithelial), and ranked as the bottom 50% in both mesenchymal and epithelial (named Low expression overlap). A heatmap of the expression of EMT markers in categorized patient samples is shown in Figure 3D.

Patients defined by the four EMT phenotypes had significantly different disease-specific survival (log-rank $p < 0.01$, Kaplan–Meier analysis). The patients in the High expression overlap group had the worst prognosis, with the shortest survival time, whereas the patients in the Epithelial group had the best prognosis, with the longest survival time (Figure 3E). The four EMT phenotypes were independent of patient cancer stage ($p = 0.8569$, chi-square tests; Table A2). The EMT phenotype was associated with ZNF71 KRAB expression ($p = 0.0099$, chi-square tests; Table A3). Within each EMT phenotype, we further performed a Kaplan–Meier survival analysis for the high ZNF71 KRAB isoform expression group (TPM ≥ 1.5) versus the low ZNF71 KRAB isoform expression group (TPM < 1.5) to assess the association of ZNF71 KRAB expression and patient survival in each EMT phenotype. In Epithelial and High expression overlap phenotypes, patients with higher KRAB expression had a significantly lower survival probability than those with lower KRAB expression (Figure 3H,I). In the Mesenchymal and Low expression overlap phenotypes, there was no significant difference in patient survival between the high and low KRAB expression groups (Figure 3F,G). In the Mesenchymal group, tumors with higher ZNF71 KRAB expression had a worse prognosis (no statistical significance; Figure 3F). In the Low expression overlap group, tumors with higher ZNF71 KRAB expression had a better prognosis (no statistical significance; Figure 3G). These results suggest that the ZNF71 KRAB isoform is a poor-prognosis marker for NSCLC tumors with high expression of epithelial markers but not for NSCLC with low expression of epithelial markers.

ZNF71 overall expression and the ZNF71 KRAB and KRAB-less expression were all significantly different among all four EMT phenotypes in NSCLC tumors ($p < 0.01$, ANOVA tests; Figure 4A–C). Furthermore, Tukey's honestly significant difference post-hoc test was performed among the EMT phenotypes. The expression of ZNF71 (overall) and its isoforms was significantly lower in the Low expression overlap patient group than in the Epithelial phenotype, and the expression of ZNF71 (overall) and the KRAB isoform was also significantly lower in the Low expression overlap patient group than in the High expression overlap patient group ($p < 0.01$, Tukey's tests; Figure 4A–C). In Epithelial and High expression overlap phenotypes with higher expression of epithelial markers, the percentage of ZNF71 KRAB isoform high expression (TPM > 1.5) was greater than that in the other two EMT phenotypes (Figure 4D).

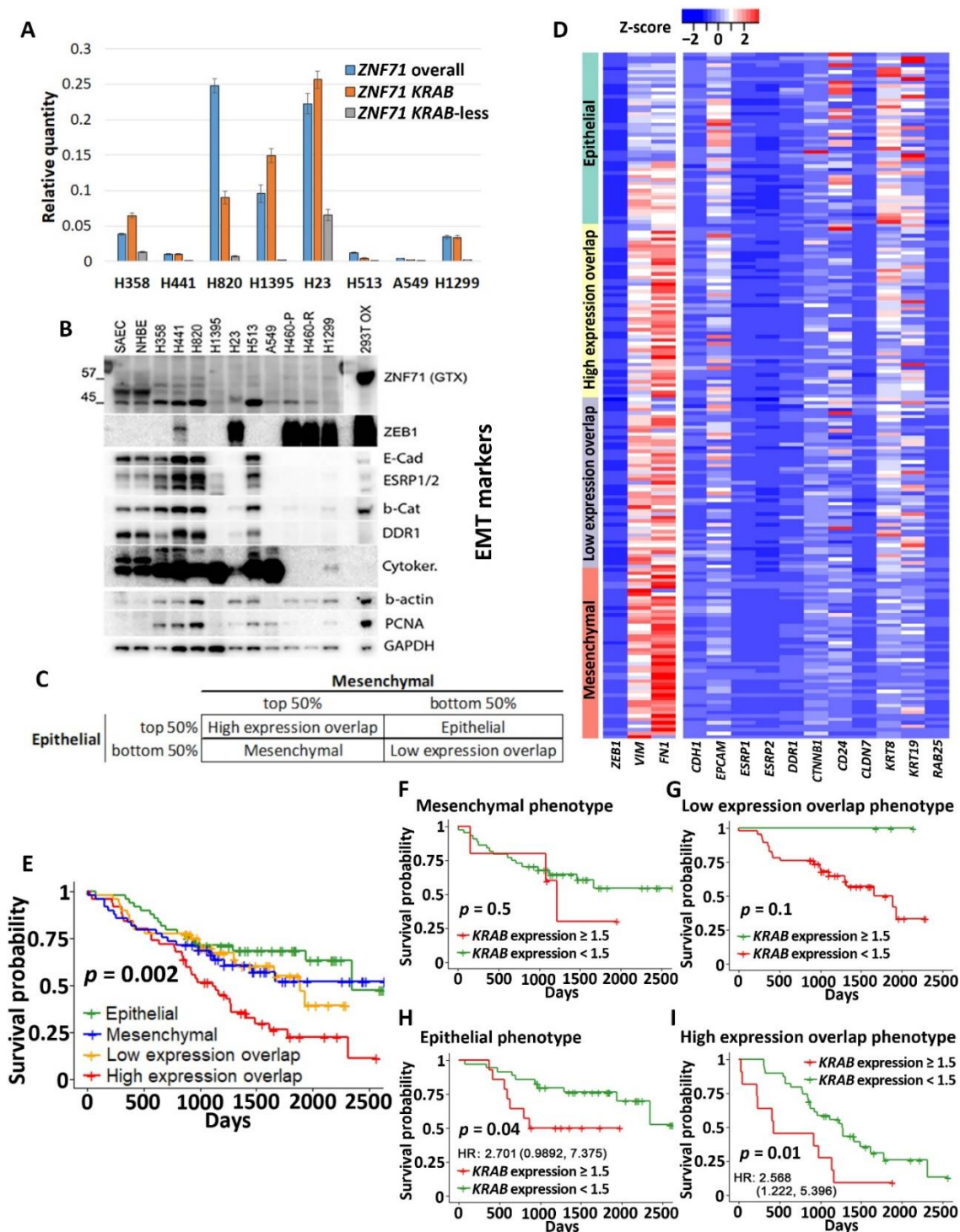


Figure 3. ZNF71 isoforms in epithelial-to-mesenchymal transition (EMT) and patient survival. (A) Relative quantity of ZNF71, KRAB, and KRAB-less in qRT-PCR assays of NSCLC cell lines. (B) Western blots of EMT markers as well as endogenous ZNF71 and overexpressed ZNF71 in HEK-293T (top lanes). (C) EMT classification of NSCLC tumors based on the average expression rank of epithelial and mesenchymal markers. The table shows the final four phenotype categorization of patient samples. (D) The expression of 14 EMT markers in patient tumors grouped by four phenotypes defined in D. (E–I) Kaplan–Meier survival analyses of patient tumors grouped by four EMT phenotypes (E) and in each phenotype (F–I). The Epithelial phenotype had the best patient survival outcome, and the High expression overlap phenotype had the worst patient survival outcome; ZNF71 KRAB is a poor prognosis marker in both phenotypes.

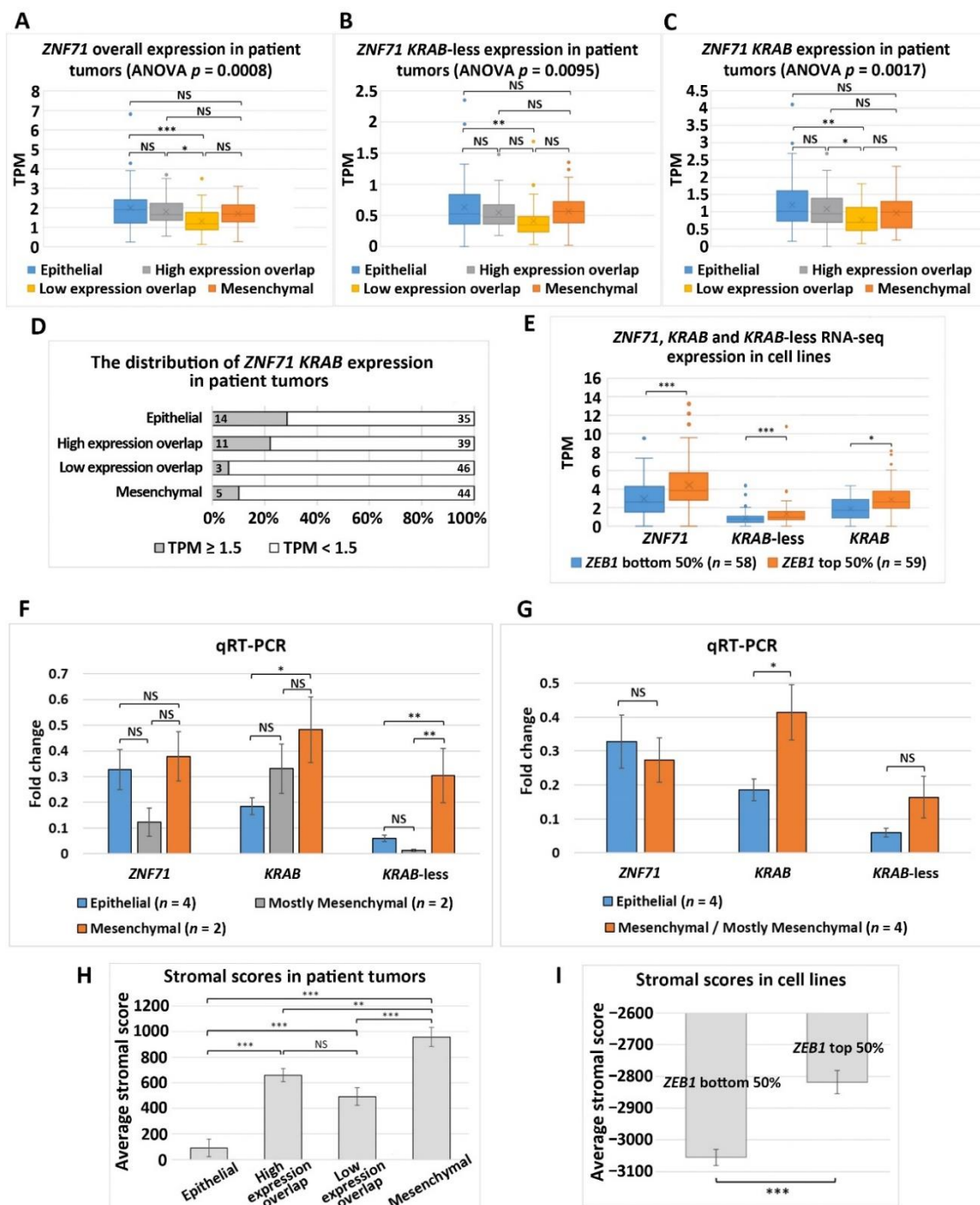


Figure 4. Association of *ZNF71* isoforms and EMT in NSCLC tumors and cell lines. (A–C) Comparison of *ZNF71* overall (A), *KRAB*-less isoform (B), and *KRAB* isoform (C) expression in patient tumors categorized by four EMT phenotypes. (D) Distribution of *ZNF71* *KRAB* expression (TPM ≥ 1.5) versus *ZNF71* *KRAB* expression (TPM < 1.5) in patient tumors defined by four EMT phenotypes. The number of patient samples in each category is provided in the figure. (E) Comparison of *ZNF71* overall, *KRAB*-less, and *KRAB* expression in NSCLC cell lines categorized by median *ZEB1* expression. (F) Fold change of *ZNF71* overall, *KRAB*, and *KRAB*-less expression in Epithelial, Mesenchymal, and Mostly Mesenchymal NSCLC cell lines in qRT-PCR analysis (individual cell line results displayed in Figure 3A). * p -Value < 0.05 , ** p -value < 0.01 , and *** p -value < 0.001 ; NS (non-significance): p -value > 0.05 in ANOVA and Tukey’s tests. (G) Fold change of *ZNF71* overall, *KRAB*, and *KRAB*-less expression in Epithelial and Mesenchymal/Mostly Mesenchymal (combined) NSCLC cell lines in qRT-PCR analysis. * p -Value < 0.05 , ** p -value < 0.01 , and *** p -value < 0.001 ; NS: p -value > 0.05 in two-sample t -tests. (H,I) EMT classifiers in both patient tumors and cell lines were validated by stromal scores quantified with ESTIMATE [20]. (H) Comparison of stromal scores of patient tumors categorized by four EMT phenotypes. (I) Comparison of stromal scores of NSCLC cell lines categorized by median *ZEB1* expression.

Solid tumors contain both epithelial and stromal cells, complicating the analysis of gene expression in patient tumor samples. On the contrary, NSCLC cell lines are devoid of any stromal components. Median *ZEB1* gene expression was used to categorize a panel of NSCLC cell lines ($n = 117$) [26] into two groups: *ZEB1* bottom 50% (more epithelial) and *ZEB1* top 50% (more mesenchymal) [19]. The NSCLC cell lines with higher *ZEB1* expression (*ZEB1* top 50%) also had significantly higher expression of *ZNF71* overall, *KRAB*, and *KRAB*-less isoforms compared to the cell lines that had lower *ZEB1* expression ($p < 0.05$, two-sample *t*-tests; Figure 4E). These more mesenchymal cell lines would NOT be comparable to patient tumors with the Epithelial phenotype. A negative association of *ZNF71 KRAB* expression and patient survival was observed in NSCLC tumors with the Epithelial phenotype (Figure 3H).

The association between *ZNF71* and EMT in RNA-seq data was further analyzed with qRT-PCR and Western blots. The cell lines included in the qRT-PCR analysis (Figure 3A) can be categorized into four groups based on the protein expression of epithelial and mesenchymal markers in Western blots (Figure 3B): Epithelial (H441, H513, H820, and H358), Mesenchymal (H23, H460, and H1299), and Mostly Mesenchymal (A549 and H1395). *ZNF71 KRAB* and *KRAB*-less expression was significantly different among the three NSCLC cell line groups: Epithelial, Mesenchymal, and Mostly Mesenchymal ($p < 0.05$, ANOVA tests). *ZNF71 KRAB* expression was significantly higher in the Mesenchymal group than in the Epithelial group ($p < 0.05$, Tukey's tests; Figure 4F). *ZNF71 KRAB*-less expression was significantly higher in the Mesenchymal group than in the Epithelial group and the Mostly Mesenchymal group ($p < 0.05$, Tukey's tests; Figure 4F). When mesenchymal and mostly mesenchymal groups were combined in the analysis, only *ZNF71 KRAB* had a significantly higher expression in Mesenchymal/Mostly Mesenchymal group than in the Epithelial group ($p < 0.05$, two-sample *t*-tests; Figure 4G). Thus, the qRT-PCR and Western blot results further substantiated the association between *ZNF71 KRAB* and EMT observed in the RNA-seq data in NSCLC patient tumors and cell lines.

Stromal infiltration of the patient samples was assessed using the stromal scores computed with software ESTIMATE [20]. The patient tumors showing the highest average stromal score were in the Mesenchymal group, followed by High expression overlap, Low expression overlap, and Epithelial groups in descending order. The stromal scores of patient tumors defined as four EMT phenotypes were significantly different from each other ($p < 0.001$, ANOVA tests; $p < 0.01$, Tukey's tests), except for the High-expression-overlap phenotype versus the Low-expression-overlap phenotype (Figure 4H). Since the NSCLC cell lines do not contain stromal cells, their average stromal scores are all negative (Figure 4I). Correlation of EMT markers and *ZNF71* isoforms with stromal and immune infiltration scores in patient tumors is provided in Table A4. These results validated our EMT classification of NSCLC tumors and cell lines.

2.3. Functional Analysis of *ZNF71* in NSCLC Cell Lines

The functional role of *ZNF71* in cell proliferation was evaluated in publicly available high-throughput CRISPR-Cas9 ($n = 78$) and RNAi ($n = 92$) screening data in NSCLC cell lines. Normalized *ZNF71* dependency scores were not significant (below the threshold of -0.5) in the NSCLC cell lines in either RNAi (processed with DEMETER2) or CRISPR-Cas9 (Figure 5), indicating that a knockdown/knockout of *ZNF71* did not significantly affect NSCLC tumor cell growth in vitro.

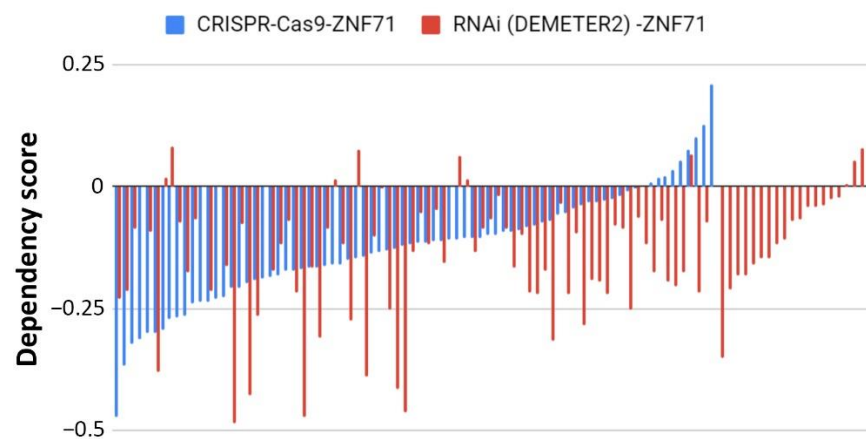


Figure 5. Functional analyses of ZNF71 in NSCLC cell lines using CRISPR-Cas9 and RNAi approaches. Dependency scores of ZNF71 in NSCLC cell lines using CRISPR-Cas9 ($n = 78$) and RNAi (normalized with DEMETER2; $n = 92$).

2.4. Association of ZNF71 Isoforms with Chemoresponse

Our previous study found that *ZNF71* overall gene expression was positively associated with prolonged survival in NSCLC patients who received cisplatin, carboplatin, and Taxol, indicating its association with chemosensitivity [6]. Drug responses of nine commonly used chemotherapeutic regimens in treating NSCLC were included in this study: carboplatin, cisplatin, paclitaxel (Taxol), docetaxel, gemcitabine, vinorelbine, etoposide, gefitinib, and erlotinib. The studied NSCLC cell lines include adenocarcinoma, squamous cell carcinoma, large cell carcinoma, and adenosquamous carcinoma (Figure A2). Due to the synergism and successful results of the combination of cisplatin–etoposide in treating small-cell lung cancer, long-term daily administration of oral etoposide in combination with cisplatin was used to treat NSCLC [29]. A systematic review showed that cisplatin–etoposide have efficacy comparable to that of carboplatin–paclitaxel when used with concurrent radiotherapy for patients with stage 3 unresectable NSCLC [30]. Gefitinib and erlotinib are widely used epidermal growth factor receptor (EGFR) tyrosine kinase inhibitors for treating advanced NSCLC with proven efficacy. A recent meta-analysis showed that gefitinib and erlotinib have comparable effects on patient survival, overall response rate, and disease control rate, with no considerable variation with regard to EGFR mutation status, ethnicity, line of treatment, and baseline brain metastasis status [31]. EGFR mutation in the studied NSCLC cell lines is provided in Table A5. EGFR mutation in sensitive and resistant cell lines in response to treatment with gefitinib and erlotinib, respectively, is provided in Table A6.

Differential expression of *ZNF71* and its isoforms was associated with chemoresponse to three drugs in the studied NSCLC cell lines. *ZNF71* overall and *KRAB* expression was significantly higher in docetaxel-resistant cell lines than docetaxel-sensitive cell lines assessed with the IC_{50} values ($p < 0.05$, two-sample t -tests; Figure 6). Docetaxel offers clinical benefits as a second-line treatment of NSCLC in patients previously treated with platinum-based chemotherapy [32]. It was recently reported that the combination of pembrolizumab (anti-PD1 immunotherapy) and docetaxel was well tolerated and substantially improved progression-free survival and overall response rate in patients with advanced NSCLC after platinum-based chemotherapy, including patients with EGFR variations [33]. *ZNF71 KRAB*-less was expressed higher in gemcitabine-sensitive cell lines than gemcitabine-resistant lines assessed with both IC_{50} and EC_{50} values ($p < 0.05$, two-sample t -tests; Figures 7 and 8). Gemcitabine, a pyrimidine nucleoside antimetabolite, has been one of the most effective agents for treating advanced NSCLC [34]. *ZNF71 KRAB* was expressed significantly higher in paclitaxel-resistant cell lines than paclitaxel-sensitive lines assessed with the IC_{50} values ($p < 0.05$, two-sample t -tests; Figure 9). Paclitaxel, a tubulin-binding agent, is commonly used to treat NSCLC in combination with a platinum-based compound [35].

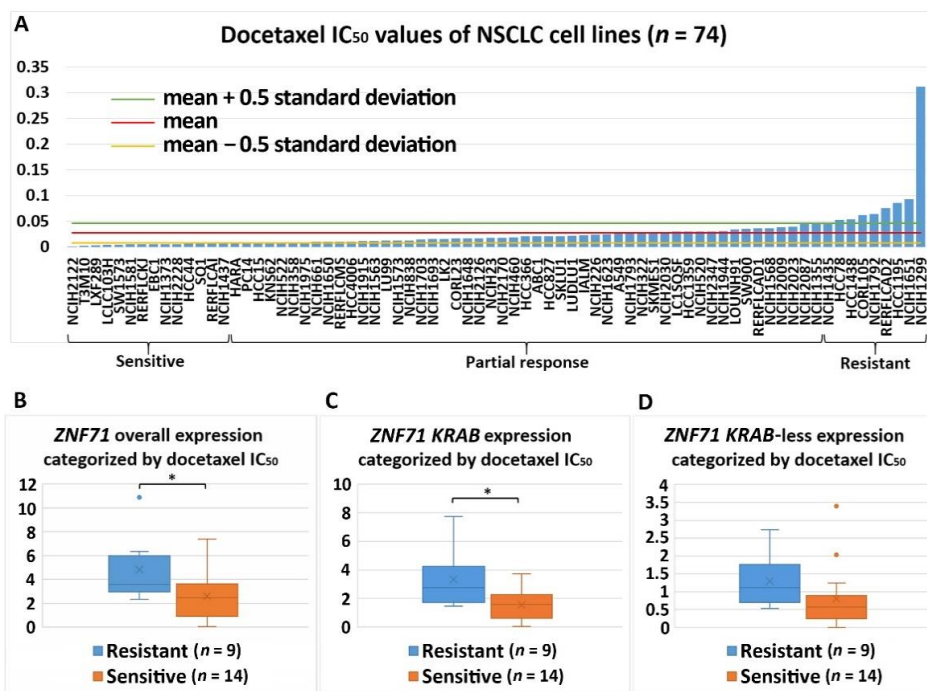


Figure 6. Association of the expression of *ZNF71* and its isoform with docetaxel drug response (IC₅₀) in NSCLC cell lines (n = 74). *ZNF71* overall and *KRAB* expression was significantly higher in docetaxel-resistant cell lines compared with docetaxel-sensitive cell lines. (A) Docetaxel IC₅₀ values of NSCLC cell lines. (B–D) Comparison of *ZNF71* overall (B), *KRAB*-less isoform (C), and *KRAB* isoform (D) expression in cell lines categorized by IC₅₀ values (resistant versus sensitive). * p < 0.05 in two-sample t-tests used in (B–D).

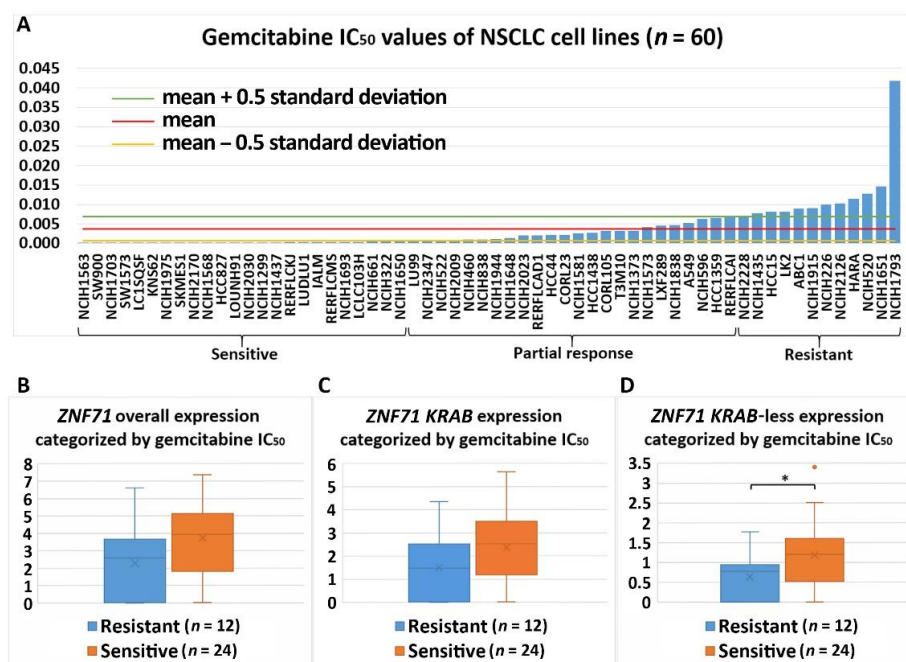


Figure 7. Association of the expression of *ZNF71* and its isoform with gemcitabine drug response (IC₅₀) in NSCLC cell lines (n = 60). *ZNF71* *KRAB*-less expression was significantly higher in gemcitabine-sensitive cell lines compared with gemcitabine-resistant cell lines. (A) Gemcitabine IC₅₀ values of NSCLC cell lines. (B–D) Comparison of *ZNF71* overall (B), *KRAB*-less isoform (C), and *KRAB* isoform (D) expression in cell lines categorized by IC₅₀ values (resistant versus sensitive). * p < 0.05 in two-sample t-tests used in (B–D).

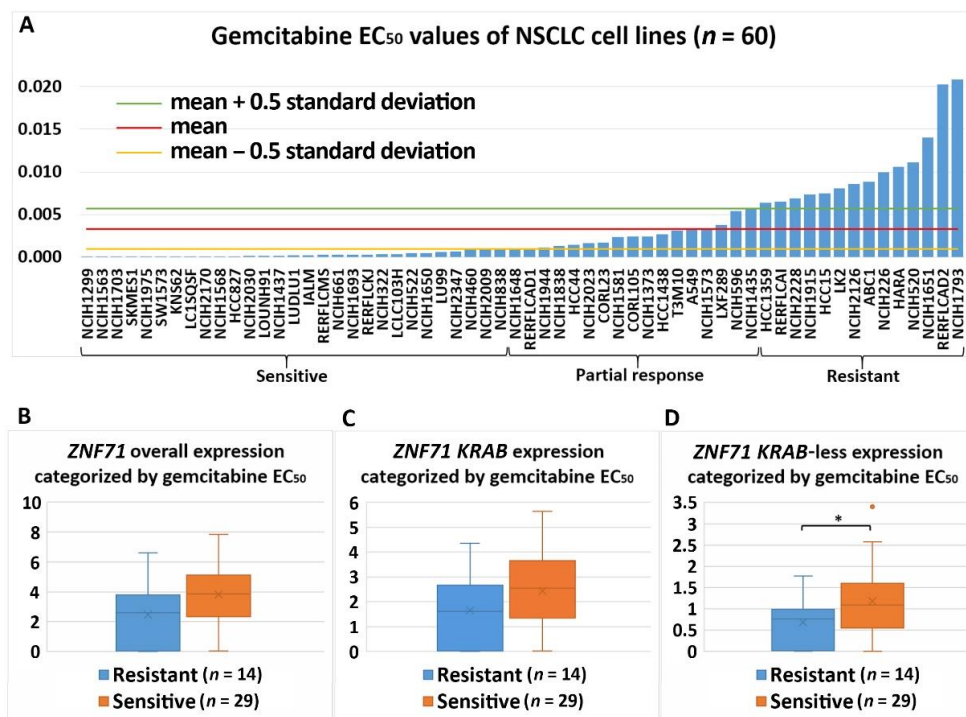


Figure 8. Association of the expression *ZNF71* and its isoform with gemcitabine drug response (EC₅₀) in NSCLC cell lines (n = 60). *ZNF71* KRAB-less expression was significantly higher in gemcitabine-sensitive cell lines compared with gemcitabine-resistant cell lines. (A) Gemcitabine EC₅₀ values of NSCLC cell lines. (B–D) Comparison of *ZNF71* overall (B), KRAB-less isoform (C), and KRAB isoform (D) expression in cell lines categorized by EC₅₀ values (resistant versus sensitive). * p < 0.05 in two-sample t-tests used in (B–D).

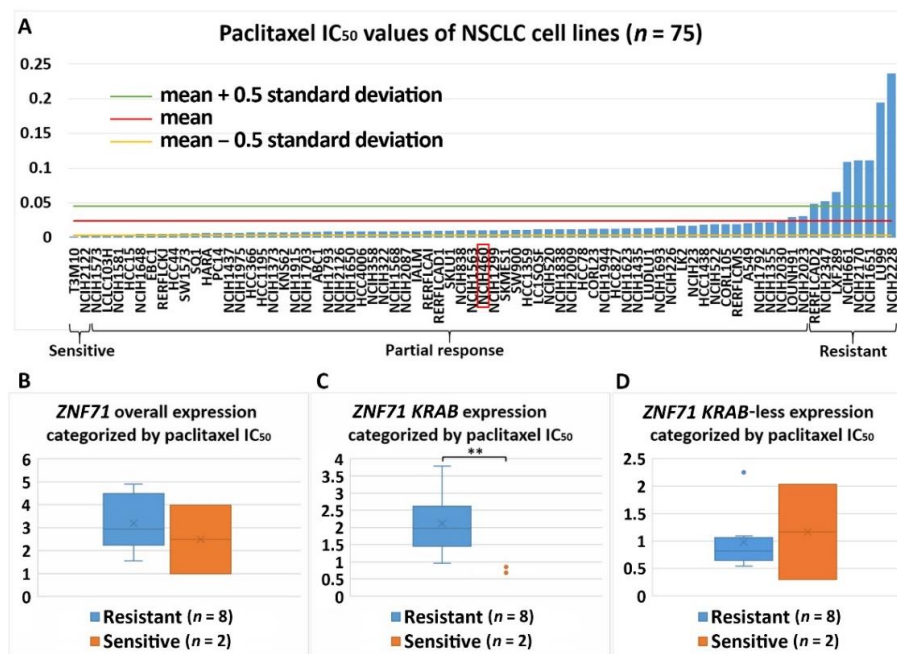


Figure 9. Association of the expression *ZNF71* and its isoform with paclitaxel drug response (IC₅₀) in NSCLC cell lines (n = 75). *ZNF71* KRAB expression was significantly higher in paclitaxel-resistant cell lines compared with paclitaxel-sensitive cell lines. (A) Paclitaxel IC₅₀ values of NSCLC cell lines. (B–D) Comparison of *ZNF71* overall (B), KRAB-less isoform (C), and KRAB isoform (D) expression in cell lines categorized by IC₅₀ values (resistant versus sensitive). ** p < 0.01 in two-sample t-tests used in (B–D).

H460 was categorized as having a partial response to paclitaxel in the studied NSCLC cell line panel (Figure 9A). Thus, the above-observed overexpression of *ZNF71* *KRAB* in paclitaxel (Taxol)-resistant versus paclitaxel (Taxol)-sensitive cell lines did not include H460. Next, the expression of *ZNF71* and its isoforms was investigated in Taxol-resistant H460 cells (H460-R) [36] and parental cells (H460-P). *ZNF71* overall expression was lower in H460-R than H460-P cells ($p < 0.05$, two-sample *t*-tests; Figure 10). *ZNF71* *KRAB* and *KRAB*-less expression had a similar differential expression pattern but was not statistically significant. These results were consistent with the observed positive correlation between *ZNF71* overall expression and prolonged survival in NSCLC patients who received cisplatin, carboplatin, and Taxol [6].

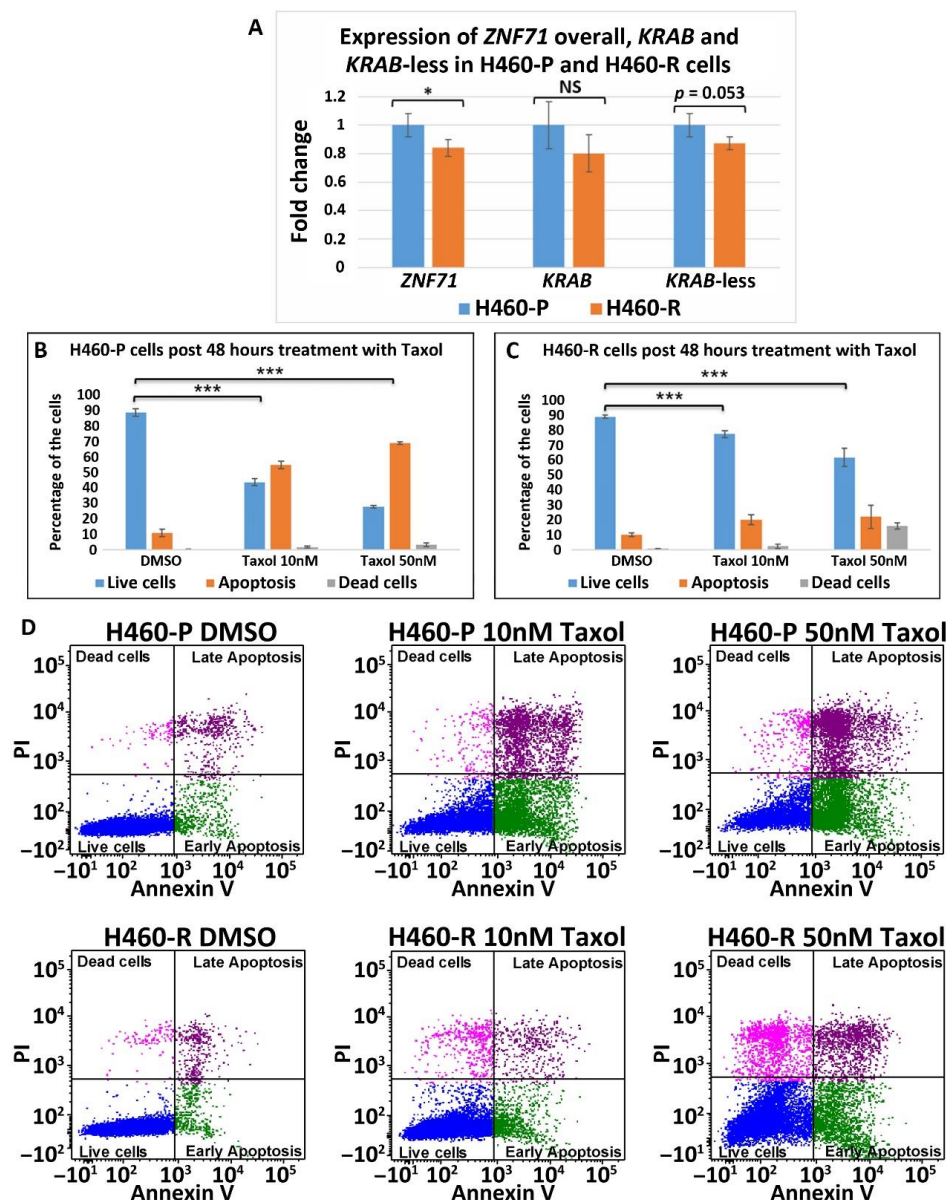


Figure 10. *ZNF71* isoforms in paclitaxel-resistant (H460-R) and parental (H460-P) cell lines. (A) Comparison of fold change of *ZNF71* overall, *KRAB*, and *KRAB*-less expression in H460-P and H460-R cell lines in qRT-PCR experiments. * $p < 0.05$, *** $p < 0.001$, NS: non-significance, two-sample *t*-tests. (B,C) H460-P (B) and H460-R (C) cells were treated either with DMSO or Taxol with the indicated doses. After 48 h, the cells were stained with Annexin V and PI and the percent of live/dead cells was determined by flow cytometry. Error bars represent standard deviation from three biological replicates. A mixed-effect model was used to assess the difference between two conditions, using R package *lme4*. The *p*-value was calculated based on asymptotic Z-distribution. (D) Flow cytometry images for H460-P and H460-R, respectively, after treatment with either DMSO or Taxol.

Overall, the observed overexpression of the *ZNF71* KRAB isoform in tumor cells resistant to docetaxel and paclitaxel treatment compared to chemo-sensitive cells is consistent with its negative association with patient survival in NSCLC.

3. Discussion

Lung cancer is difficult to manage in clinics due to its complex etiology and somatic mutations. There are currently no clinically available gene tests to predict metastasis and clinical benefits of chemotherapy for all NSCLC, stages 1 to 3A. The 7-gene assay has been validated in our previous study as prognostic and predictive of chemotherapeutic benefits in multiple U.S. hospitals and a clinical trial JBR.10. ($n = 331$) [6]. The 7-gene assay also estimates each individual patient's response to four drugs used to treat lung cancer: cisplatin, carboplatin, paclitaxel, and pemetrexed [6]. Thus, the 7-gene assay could meet the critical need in clinics to identify specific NSCLC patients who are at risk for tumor recurrence/metastasis and would benefit from receiving adjuvant chemotherapy. In addition, this capacity would reduce the use of adjuvant therapy in circumstances in which no benefit, but potentially negative side effects, would result.

Many studies have investigated molecular alterations in single genes in early-stage NSCLC and their prognostic and predictive implications [37], such as mutations in *KRAS*, *P53*, *EGFR* [38], *STK11* [39], mRNA, and protein expression of *ERCC1* [40,41]. Nevertheless, none of these genes are ready for primetime clinical applications as prognostic and predictive biomarkers for early-stage NSCLC [37]. Several prognostic/predictive gene expression signatures for early-stage NSCLC have been reported [42–45], among which the Razor 14-gene assay [44,46] (Razor Genomics, Brisbane, CA, USA) has been commercialized for the prognosis of early-stage non-squamous NSCLC and is currently recruiting patients for a clinical trial. However, the Razor 14-gene assay is not applicable for squamous cell lung carcinoma, which accounts for 26% of lung cancer cases [47]. The Myriad myPlan™ Lung Cancer and Pervenio™ Lung RS tests were commercially available in clinics for prognosis of early-stage NSCLC [48]. However, they are no longer listed in the Myriad All Products page [49]. FoundationOne CDx (Foundation Medicine, Inc, Cambridge, MA, USA) and Oncomine DX (Thermo Fisher Scientific Inc., Waltham, MA, USA) receive reimbursement coverage to match stage 4 NSCLC patients to specific proteasome inhibitor-based therapy based on their tumor genetic mutations. However, they do not predict metastasis for patients in stages 1 to 3A.

Within the 7-gene signature, mRNA of *ZNF71* overall expression measured in qRT-PCR was positively correlated with survival in patients who received cisplatin, carboplatin, and Taxol, indicating that *ZNF71* mRNA is associated with chemosensitivity. Protein expression of *ZNF71* was positively correlated with patient survival in independent tissue microarray cohort studies. However, mRNA of *ZNF71* overall expression measured in qRT-PCR was not found to be associated with NSCLC patient survival in the overall studied cohorts in our previous analysis [6]. Zinc finger proteins (ZFPs) are the largest family of transcription factors not only in human cells but also in most eukaryotes, representing about 3% of the total human genome [7,50]. A repeating motif within ZFPs contains two histidine and two cysteine amino acid residues (i.e., C2H2) coordinating zinc ion and has the ability to bind to DNA, RNA, or cellular proteins [8]. Therefore, ZFPs have a wide range of predicted functions according to their molecular structure, including DNA repair, degradation of proteins, signal transductions, migration of cells, regulation of apoptosis, lipid binding, and transcription regulation [7,8]. KRAB-ZFPs are a family of transcriptional repressors with diverse functions, most notably the silencing of transposable elements [51]. They contain an N-terminal KRAB domain and a C-terminal C2H2-type zinc finger array. The repressor action of KRAB-ZFPs requires the recruitment of KRAB-associated protein 1, also known as tripartite motif protein 28, KAP1/TRIM28. KAP1 functions as a scaffold complex composed of histone methyl transferase (SETDB1), heterochromatin protein-1 (HP-1), nucleosome remodeling and deacetylation (NuRD), and DNA methyl transferase [11]. When KRAB-ZFPs recruit KAP-1, the formed repressor complex leads to heterochromatin formation.

Our group reported that KAP1 promotes proliferation and metastatic progression of breast cancer cells [52], consistent with the observed association between KAP1 and breast cancer progression [53]. RB-associated KRAB zinc finger (BRAK) was upregulated in NSCLC and was associated with poor prognosis in patients [52]. Zinc finger protein 668 (ZNF668) was reported to suppress NSCLC invasion and migration by down-regulating Snail and upregulating E-cadherin and zonula occludens-1 [54]. Down-regulated protein expression of ZNF668 was found in NSCLC tumors compared with normal lung tissues, and there was a negative association between ZNF668 protein expression and lymph node metastasis [54].

In this study of public RNA-seq data, *ZNF71 KRAB* had a significantly higher expression than the *ZNF71 KRAB*-less isoform in NSCLC patient tumors [25] and cell lines [27]. The expression of both isoforms was significantly correlated in patient tumors and cell lines. Patients with higher *ZNF71 KRAB* expression had a significantly worse survival outcome than patients with lower *ZNF71 KRAB* expression, whereas the *ZNF71* overall expression and the *ZNF71 KRAB*-less isoform were not prognostic in the same patient cohort (Figure 2). In this study, *ZNF71* overall expression was not prognostic in RNA-seq data of NSCLC patient cohorts, which is consistent with our previous qRT-PCR results. We designed primers for *ZNF71 KRAB* and *KRAB*-less isoforms for TaqMan qRT-PCR assays, and the results in NSCLC cell lines confirmed the overall higher expression of *ZNF71 KRAB* than that of the *KRAB*-less isoform. Further investigation revealed an association between *ZNF71 KRAB* expression and EMT in RNA-seq data of NSCLC patient tumors and cell lines, which was validated in qRT-PCR and Western blot assays of cell lines (Figures 3 and 4). *ZNF71 KRAB* was overexpressed in cell lines resistant to docetaxel (Figure 6) and paclitaxel (Figure 9) treatment compared to chemo-sensitive cell lines, which was consistent with its association with poor prognosis in patients. Given the transcriptional repression role of the KRAB domain, the negative correlation between *ZNF71 KRAB* expression and NSCLC patient survival could be reasonable. These results indicate that the *ZNF71 KRAB* isoform is a more effective prognostic factor than *ZNF71* overall expression and the *ZNF71 KRAB*-less isoform for NSCLC. Further investigation is warranted to explore whether the *ZNF71 KRAB* isoform provides added prognostic and predictive value to the original 7-gene assay for NSCLC or whether *ZNF71 KRAB* could replace *ZNF71* overall expression in the 7-gene assay to achieve better prognostic performance.

The expression of overall *ZNF71* and its isoforms was not significantly different among different histological subtypes in patient tumors (Figure A1). The expression of *ZNF71* overall and the *KRAB* isoform was significantly different among histological subtypes in NSCLC cell lines (Figure A2). The observed discrepancy in *ZNF71* overall and *KRAB* expression among NSCLC histological subtypes in patient tumors and cell lines is possibly due to the fact that stromal cells are present in the NSCLC tumors but not in the studied cell lines. There was a strong association between *ZNF71* overall and *KRAB* expression and stromal infiltration in patient tumors (Table A4) but not in NSCLC cell lines (results not shown). In the future, it would be interesting to investigate whether *ZNF71* overall and *KRAB* expression could differentiate NSCLC histological subtypes in micro-dissected epithelial cells from NSCLC tumors.

The role of EMT in cancer patient outcomes is not well defined and remains controversial [19,55]. In this study, we developed an EMT classifier based on transcriptional profiles of 14 EMT markers, and NSCLC patients defined with four EMT phenotypes had distinct survival outcomes (Figure 3). Out of the 14 EMT markers, 13 (all except for *CTNNB1*) used in our classification were included in the pan-cancer EMT classifier presented in Panchy et al. [21], which also divides tumors into quadrants to characterize hybrid EMT states in tumors. Our EMT classifier was validated using stromal scores quantified with the well-established software ESTIMATE [20]. *ZNF71 KRAB* expression was prognostic in NSCLC patients with high expression of EMT markers (the Epithelial group and the High expression overlap group) but not prognostic in tumors with low expression of EMT markers (the Mesenchymal group and the Low expression overlap group). These results imply that transcriptional biomarkers might have distinct prognostic implications in tu-

mors with different EMT characteristics. The unclear role of EMT in patient outcomes is further complicated by tumor stromal and immune infiltration. For instance, the expression of *ZNF71* overall and its isoforms was not significantly different in NSCLC patient tumors separated by median *ZEB1* expression (the top 50% *ZEB1* expression group versus the bottom 50% *ZEB1* expression group, two-sample *t*-tests; results not shown). However, a strong association was found between *ZNF71 KRAB* and *ZEB1* gene expression in NSCLC epithelial cell lines (Figure 4E).

ZNF71 molecular functions have not been reported in the literature. This study evaluated the functional associations of *ZNF71* using publicly available genome-scale CRISPR-Cas9 and RNAi screening data. The knockouts and knockdowns in NSCLC cell lines were performed for overall *ZNF71*, i.e., all transcripts. They did not affect cell proliferation (Figure 5). Concordantly, overall *ZNF71* mRNA expression was not associated with patient survival. We found that *ZNF71 KRAB* was associated with patient outcome and with EMT, i.e., expressed higher in the top 50 *ZEB1* expressing cells. A mechanistic link between *ZNF71 KRAB* and EMT is currently not known. We could speculate that EMT is associated with significant changes in splicing and hence can potentially skew splicing for inclusion of the KRAB domain [56,57]. The link between other KRAB-ZFPs and EMT was reported. TRIM28 protein, a universal co-factor for KRAB-ZFP transcription factors [22], is known to participate in a wide range of aspects of cellular biology, either promoting cell proliferation [52] or mediating anti-proliferative activities [58]. TRIM28 protein is involved in cancer by regulating gene expression through heterochromatin formation, mediation of DNA damage response, inhibition of p53 activity, regulation of EMT, and maintenance of stem cell pluripotency and genome stability [59]. TRIM28 expression is induced following transforming growth factor- β (TGF- β) treatment at both protein and mRNA levels. TRIM28 deficiency impairs TGF- β -induced EMT and decreases cell migration and invasion, and the expression of TRIM28 affects the acetylation of histones on E-cadherin and N-cadherin promoters, suggesting that TRIM28 contributes to EMT and might be important for tumor metastasis in lung cancer [23]. *ZNF382 KRAB* regulates EMT and functions as a tumor suppressor in gastric cancer [24]. To gain better insight into the mechanistic link between *ZNF71 KRAB* and EMT, we are currently conducting genome-scale network analysis to identify all the genes showing a significant statistical association with *ZNF71* at gene expression and DNA copy number variation levels in NSCLC patient tumors. We have identified *ZNF71*-mediated molecular association networks in EMT using public data. A similar analysis will be conducted to identify *ZNF71 KRAB* mediated molecular networks in EMT using public RNA-seq data generated from NSCLC patient tumors. In the future knockdown/overexpression experiments, we will examine which of these EMT-relevant genes in the identified networks are affected by *ZNF71 KRAB* to investigate the potential mechanistic link between *ZNF71 KRAB* and EMT in NSCLC.

The association between *ZNF71* and its isoforms in terms of chemoresponse is also complex due to different genetic predisposition to chemotherapeutic regimens in tumor cells. In the studied NSCLC cell line panel, *ZNF71 KRAB* overexpression was observed in tumor cells that were resistant to Taxol compared to Taxol-sensitive tumor cells (Figure 9). These results were consistent with the observed negative association of *ZNF71 KRAB* expression and patient survival (Figure 3). In the H460 cell line that is defined as having a partial response to Taxol in the studied cell line panel (Figure 9A), *ZNF71* overall expression was significantly lower in H460-R than in H460-P cells (Figure 10). These results were consistent with the observed association of *ZNF71* overall expression with chemosensitivity to Taxol in NSCLC patients [6]. Significantly higher *ZNF71* overall and *KRAB* expression in docetaxel-resistant cell lines suggests their potential use in predicting predisposition to this agent or in combined immunotherapy and chemotherapy, based on the recent report of the efficacy of the combination of docetaxel and pembrolizumab in treating NSCLC [33].

ZNF71 overall and *KRAB* expression was significantly higher in a normal lung small airway epithelial cell line (SAEC) compared with that in most NSCLC cell lines analyzed in qRT-PCR. Since we cannot draw any solid conclusion based on one normal lung cell

line, we did not show these results. One possible explanation could be that SAEC cells were cultured in specialized media (Lonza, Basel, Switzerland) with added defined growth factors, while the lung cancer cell lines were cultured in standard DMEM plus 10% FBS. Many KRAB-ZFPs were reported to act as either tumor suppressors or oncogenes [60]. ZNF382 is down-regulated in multiple carcinoma types due to promoter methylation and functions as a tumor suppressor in gastric cancer [24]. ZNF23, a KRAB-containing protein, is down-regulated in human cancers and inhibits cell cycle progression [61]. RB-associated KRAB (RBAK) zinc finger is upregulated in NSCLC and promotes cell migration and invasion [62]. We are planning to carry out the following analysis of ZNF71 overall and isoform expression in our future study: (1) examine multiple normal lung epithelial cell lines in qRT-PCR, (2) analyze TCGA data for NSCLC tumors versus normal lung tissue samples, and (3) design knockdown and overexpression of ZNF71 in vitro and/or in vivo xenograft studies to examine whether it is oncogenic or tumor suppressive.

4. Materials and Methods

4.1. NSCLC Patient Samples and RNA-Seq Data

A total of 199 NSCLC patient tumors were collected in a previous study [25]. Patient clinical information is provided in Table A1. Deep RNA sequencing of the patient tumors was generated with Illumina HiSeq 2500 (raw data available at NCBI GEO with accession number GSE81089). Patients with sufficient survival information ($n = 197$) were included in this study. The raw RNA-seq data were processed with Salmon to quantify the expression of ZNF71 at the isoform level by using transcripts per million (TPM) reads [63].

4.2. Cancer Cell Line Encyclopedia (CCLE)

Gene expression data for CCLE were downloaded from DepMap 20Q2 (https://figshare.com/articles/dataset/DepMap_20Q2_Public/12280541, accessed on 1 April 2021) [26]. Gene expression data were obtained from the CCLE data portal (https://data.broadinstitute.org/ccle/CCLE_RNAseq_081117.rpkm.gct, accessed on 1 April 2021). RNA-seq data were quantified using the GTEx pipelines [64]. A total of 117 NSCLC cell lines were included in this analysis.

4.3. CRISPR-Cas9 Assays

Gene knockout effects in CCLE using CRISPR-Cas9 screens were quantified in Project Achilles [27,65]. The data were obtained from DepMap 20Q2 (https://figshare.com/articles/dataset/DepMap_20Q2_Public/12280541, accessed on 1 April 2021) [26]. The CRISPR-Cas9 data were processed with the CERES method [27]. Gene effects in each cell line were normalized such that the median non-essential gene knockout effect is 0 and the median essential gene knockout effect is -1. A gene is defined as an essential gene if it is essential to the cell growth in each line; otherwise, it is defined as a non-essential gene. There were 78 NSCLC cell lines with genome-scale CRISPR-Cas9 knockout results.

4.4. RNAi Functional Assays

Genome-scale RNAi screening data in CCLE were obtained from Project Achilles [28] (<https://depmap.org/R2-D2/>, accessed on 1 April 2021). The DEMETER2 method [28] was used to estimate average gene dependency scores in each cell line for short hairpin RNA (shRNA) libraries. Gene dependency scores were standardized with DEMETER2 such that the median of the across-cell-line average dependency scores of the positive control gene set was -1 and that of the negative control gene set was 0. There were 92 NSCLC cell lines with genome-scale RNAi screening results normalized with DEMETER2.

4.5. Drug Response

The growth inhibitory activity of 4518 drugs was quantified in 578 human cancer cell lines using the PRISM molecular barcoding and multiplexed screening method [66]. The PRISM repurposing dataset is available at the Cancer Dependency Map portal

(<https://depmap.org/portal/download/>, accessed on 1 April 2021). Drug responses of nine commonly used chemotherapeutic regimens in treating NSCLC were included in this study: carboplatin, cisplatin, paclitaxel, docetaxel, gemcitabine, vinorelbine, etoposide, gefitinib, and erlotinib. For each drug, cell lines were defined as resistant, sensitive, or partial response by using the mean \pm 0.5 standard deviation (SD) of the IC₅₀ or EC₅₀ values [67,68]. Cell lines with an IC₅₀ or EC₅₀ value greater than the mean + 0.5 SD were defined as resistant to the drug. Cell lines with an IC₅₀ or EC₅₀ value less than the mean – 0.5 SD were defined as sensitive to the drug, and those with an IC₅₀ or EC₅₀ value between the mean + 0.5 SD and the mean – 0.5 SD were defined as having a partial response to the drug. This categorization corresponds to the RECIST 1.1 system (i.e., complete response, partial response, and stable disease/disease progression) in evaluating chemotherapeutic response in solid tumors [69].

4.6. RNA Extraction, Quality Assessment, and qRT-PCR

Total RNA was extracted from multiple cell lines using the RNeasy Mini kit (Qiagen, Hilden, Germany). According to manufactures' protocol, RNA was eluted in 30 μ L of deionized water and kept at –80 °C until use. The concentration of RNA was determined with NanoDrop, and RNA purity was verified by determining the A260/A230 ratio. One microgram of total RNA was reverse-transcribed using SuperScriptIII First Strand Synthesis SuperMix for qRT-PCR (Invitrogen, Carlsbad, CA, USA). For quantitative real-time PCR, either SYBR Select Master Mix (Thermo Fisher) or a TaqMan master mix (IDT) was used, depending on the primers/probes used.

All primers were purchased from IDT. There were three sets of ZNF71 primers: Set 1 (will amplify both isoforms of the gene) primer F: 5'-CAGCACTTCAGACCTCAGTAAG-3', primer R: 5'-TTGGTGCTTTATCAGGGACG-3'; Set 2, KRAB-less ZNF71, primer F: 5'-GCCTGTCTTCCTATTCACCG-3', primer R: 5'-CATTCAGGTCTAGTCTCCAG-3', probe: /56-FAM/AGC CAT CCC /ZEN/TCT GCT GCC C/3IABkFQ/; Set 3, KRAB-ZNF71, primer F: 5'-GACGTTCAAGGATGTGACTG-3', primer R: 5'-TTCAGGTCTAGTCTCCAGTC-3', probe: /56-FAM/AGG TCC TTC /ZEN/TGG GCA GGC TC/3IABkFQ/.

The PCR reactions were loaded in a 384-well plate in triplicate. The expression level of ZNF71 was calculated using the comparative threshold ($2^{-\Delta\Delta C_t}$) method and normalized to UBC or RPL4 housekeeping genes: UBC primer set, primer F: 5'-GATTTGGTTCGCAGTTC TTC-3', primer R: 5'-CCTTATCTTGGATCTTTGCCTTG-3' and RPL4 (Cat. N PPH13915A, SABiosciences, Frederick, MD, USA). For SYBR Green, PCR melting curve analysis was performed to verify that the reactions had a single product. UBC was used as the housekeeping gene to calculate the delta Ct value for each sample in this study.

4.7. Western Blots

Cells were lysed in gel lysis buffer (GLB). GLB was composed of 50 mM Tris-HCl, pH 6.8, 2% SDS, and 10% glycerol. Total protein was quantified using the BCA assay (Pierce). An equal amount of protein lysate (20–30 μ g) was loaded and separated on 4–12% Bis-Tris gels, then transferred to polyvinylidene difluoride (PVDF) membrane (Fisher), and probed with antibodies as previously described [70]. ZNF71 antibody was purchased from GeneTex (Cat. No. GTX116553). The following antibodies were included in Western blots: ZEB1 (Sigma, St. Louis, MI, USA; HPA027524), E-cadherin (BD Biosciences, 610181), ESRP1/2 (23A7) (Novus, Centennial, CO, USA; NBP1-77971), beta-Catenin (6B3, Cell Signaling, Danvers, MA, USA; 9582), DDR1 (D1G6, Cell Signaling, 5583), pan-Cytokeratin (C11, Santa Cruz, Dallas, TX, USA; sc-8018), Actin (I-19, Santa Cruz, sc-1616), PCNA (eBioSciences, 14-6748-81), GAPDH (Millipore, Burlington, MA, USA; MAB374), and secondary horse radish peroxidase (HRP)-conjugated antibodies against mouse and rabbit (Jackson ImmunoResearch, West Grove, PA, USA). Standard chemiluminescence was used to detect the protein bands.

4.8. Responses to Taxol with Annexin V Staining for Apoptosis Using Flow Cytometry

H460 cells (denoted as H460-P) and a derivative cell line H460-R, reported to be Taxol resistant [36], were plated in 6-well plates in three biological replicates and incubated with either 10 nM or 50 nM Taxol for 48 h, with DMSO used as vehicle control. Supernatants were centrifuged to collect floating cells which were combined with trypsinized cells and subsequently stained with Annexin V and PI. The percentage of apoptotic/dead cells was determined by flow cytometry. An Annexin V binding assay was performed using the FITC Annexin V/Dead Cell Apoptosis Kit with FITC Annexin V and PI for flow cytometry (Invitrogen). PI is the standard reagent to exclude non-viable cells from flow cytometry assay, i.e., to calculate cell death. Flow cytometry was performed at the Flow Cytometry and Single Cell Core Facility at West Virginia University on the LSRFortessa machine.

4.9. Statistical Analysis

Statistical analysis was performed using Rstudio version 1.1.456 [71]. Differential gene expression between two groups was evaluated with Student's *t*-tests, and a two-sided *p*-value < 0.05 was considered statistically significant. Differences in expression among more than two groups were evaluated with ANOVA and Tukey's honestly significant difference tests, and a *p*-value < 0.05 was considered as statistically significant. Survival analysis was performed using Kaplan–Meier analysis with the *survival* package in R. Log-rank tests were used to assess the difference in survival probability from different groups in Kaplan–Meier analyses. A heatmap was generated with the *heatmap.2* function from the *gplots* package in R. Partial correlation was used to find the relationship between two variables and eliminate the variance of the third variable using functions *pcor* and *pcor.test* in the *ppcor* package in R. A mixed-effect model was used to assess the difference between two conditions, using R package *lme4*. A linear mixed-effect model was used to assess the difference between different conditions as fixed effects (biological samples as random effects), considering the correlation between technical replicates. The *p*-value was calculated based on asymptotic Z-distribution.

4.10. Assessment of Stromal Infiltration and Immune Infiltration

A gene's stromal and immune infiltration in NSCLC tumors was evaluated with the Estimation of STromal and Immune cells in MAlignant Tumours (ESTIMATE) method [20]. The function *estimateScore* was performed on gene expression data to evaluate the stromal scores and immune scores in each sample with the *estimate* package in R.

4.11. ZNF71 Overexpression

For overexpression purposes, the ZNF71 Plasmid construct was purchased from the PlasmID database (<https://plasmid.med.harvard.edu/PLASMID/OrderOverview.jsp>, accessed on 1 April 2021) maintained by Harvard University (catalog number HsCD00412101). ZNF175 plasmid construct was used as a control (catalog number HsCD00421069), and pLU-GFP vector was also used as a control. For HEK-293T cell transfection, calcium phosphate was used as a transfection reagent, which is commonly used to introduce DNA into eukaryotic cells to obtain both transient and permanent transfections. The approach is based on mixing HEPES-buffered saline containing Na₃PO₄ with CaCl₂ containing the DNA. The DNA calcium phosphate stick to the cell membrane then is taken by cellular endocytosis. For the NSCLC transfection reagent, Lipofectamine was purchased from Thermo Fisher Scientific (catalog number L3000008). Cells were cultured in 6-well plates seeding 400,000–500,000 cells per plate according to the transfection type: reverse, serum-free medium, or regular transfection. A total of 15 µg of plasmid in 50 µL of CaCl was used to transfect HEK-293 in a 35 × 10 mm cell culture dish.

5. Conclusions

This study revealed that the ZNF71 KRAB isoform is a more effective prognostic factor for NSCLC than ZNF71 overall expression and the ZNF71 KRAB-less isoform.

ZNF71 KRAB was overexpressed in NSCLC cell lines resistant to docetaxel and paclitaxel treatment compared to chemo-sensitive cell lines, consistent with its association with a poor prognosis in patients. These results suggested that the ZNF71 KRAB isoform may provide added prognostic and predictive value to the 7-gene assay developed in our previous study [6] for NSCLC, upon further evaluation using patient samples. ZNF71 KRAB expression was also associated with EMT in both NSCLC patient tumors and cell lines. The knockdown/knockout of ZNF71 did not significantly affect cell growth in NSCLC cell lines, implying that ZNF71 might not be involved in cell proliferation. The results from this study provided evidence of potential implications of the ZNF71 KRAB isoform in NSCLC prognosis and enlightened its possible mechanism in EMT for future research.

6. Patents

The 7-gene NSCLC prognostic and predictive assay is included in U.S. National Phase Patent Application No. 17/251,359 and International Non-Provisional Patent Application No. PCT/US20/23597.

Author Contributions: Conceptualization, N.L.G.; formal analysis, Q.Y.; GSE81089 data processing, G.H.; experimental validation, R.M., D.D., and M.G.; H-460 Taxol-resistant cells, M.C.P. and S.K.K.; resources, C.B.M., T.D.E., and A.V.I.; writing—original draft preparation, N.L.G., Q.Y., and R.M.; writing—review and editing, A.V.I.; supervision, N.L.G., T.D.E., and A.V.I.; funding acquisition, N.L.G. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by NIGMS/XLerateHealth LLC Ideas-to-Products Award (4UT2GM130174) to N.L.G. We acknowledge GM103434, OD016165, and GM104942 for institutional core support.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All links to publicly archived datasets are provided in the manuscript.

Acknowledgments: We thank Sijin Wen from the Department of Biostatistics and Kathleen Brundage from Flow Cytometry & Single Cell Core Facility at West Virginia University for helpful discussion and assistance in data analysis.

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

Table A1. Clinical information of the patient cohort from GSE81089.

	Clinical Variable	Count
Stage	1a	70
	1b	44
	2a	25
	2b	23
	3a	32
	4	3
Histology	Adenocarcinoma	108
	Large cell carcinoma	23
	Squamous cell carcinoma	66

Table A1. *Cont.*

Clinical Variable		Count
Gender	Female	102
	Male	95
Status	Alive	104
	Dead	93
Smoking	Current	94
	Ex > 1 year	84
	Never	19

Table A2. Chi-square tests of EMT phenotypes and cancer stage in the patient cohort. EMT phenotypes and cancer stage are independent (p -value = 0.8569).

Stage	EMT Phenotype				Total
	Epithelial	High Expression Overlap	Low Expression Overlap	Mesenchymal	
1	28	30	28	28	114
2	11	11	12	14	48
3	9	7	9	7	32
4	1	2	0	0	3
Total	49	50	49	49	197

Table A3. Chi-square tests of EMT phenotypes and ZNF71 KRAB expression in the patient cohort. EMT phenotypes and KRAB expression are associated (p -value = 0.0099).

ZNF71 KRAB Expression	EMT Phenotype				Total
	Epithelial	High Expression Overlap	Low Expression Overlap	Mesenchymal	
TPM \geq 1.5	14	11	3	5	33
TPM < 1.5	35	39	46	44	164
Total	49	50	49	49	197

Table A4. Correlation of the expression of EMT markers and ZNF71 isoforms with immune scores and stromal scores quantified with ESTIMATE in NSCLC tumors ($n = 197$).

	Correlation with Immune Scores	Correlation with Stromal Scores	Partial Correlation with Immune Scores (Removing the Effect of Stromal Scores)	Partial Correlation with Stromal Scores (Removing the Effect of Immune Scores)
ZEB1	0.13	0.55 ***	-0.235 ***	0.573 ***
VIM	0.21 **	0.3 ***	0	0.447 ***
FN1	0.11	0.34 ***	-0.29 ***	0.595 ***
CD24	-0.23 **	-0.2 **	-0.175 *	0.022
CDH1	-0.18 *	-0.12	-0.157 *	-0.05
CLDN7	-0.12	-0.12	-0.07	-0.079
CTNNB1	-0.083	0.041	-0.189 **	0.19 **
DDR1	-0.38 ***	-0.36 ***	-0.29 ***	-0.137
EPCAM	-0.24 ***	-0.25 ***	-0.167 *	-0.178 *

Table A4. Cont.

	Correlation with Immune Scores	Correlation with Stromal Scores	Partial Correlation with Immune Scores (Removing the Effect of Stromal Scores)	Partial Correlation with Stromal Scores (Removing the Effect of Immune Scores)
ESRP1	−0.56 ***	−0.42 ***	−0.404 ***	−0.184 **
ESRP2	−0.32 **	−0.25 ***	−0.117	−0.159 *
KRT8	−0.32 **	−0.19 **	−0.114	−0.103
KRT19	−0.17 *	−0.28 ***	0.043	−0.089
RAB25	−0.25 ***	−0.26 ***	−0.151 *	−0.175 *
ZNF71 overall	−0.12	−0.17 *	0.015	−0.059
ZNF71 KRAB	−0.11	−0.16 *	−0.157	−0.055
ZNF71 KRAB-less	−0.085	−0.11	−0.114	0.005

* p -value < 0.05, ** p -value < 0.01, and *** p -value < 0.001 in Pearson's correlation tests or partial correlation tests

Table A5. EGFR mutation in the studied NSCLC cell lines. The information was extracted from DepMap Portal (<https://depmap.org/portal/>, accessed on 1 April 2021).

Cell Line	Location of the EGFR Mutation	EGFR Variant Classification
HCC1359	chr7: 55087037-55087037	Missense_Mutation
HCC4006	chr7: 55242466-55242474	In_Frame_Del
HCC827	chr7: 55242466-55242480	In_Frame_Del
LOUNH91	chr7: 55242470-55242487	In_Frame_Del
	chr7: 55242494-55242494	Missense_Mutation
NCIH1355	chr7: 55273154-55273154	Missense_Mutation
NCIH1568	chr7: 55233029-55233029	Silent
	chr7: 55273056-55273056	Missense_Mutation
NCIH1650	chr7: 55242465-55242479	In_Frame_Del
NCIH1793	chr7: 55223565-55223565	Missense_Mutation
NCIH1975	chr7: 55249071-55249071	Missense_Mutation
	chr7: 55259515-55259515	Missense_Mutation
NCIH226	chr7: 55223558-55223558	Missense_Mutation
PC14	chr7: 55242465-55242479	In_Frame_Del
SKMES1	chr7: 55240751-55240751	Silent
	chr7: 55223558-55223558	Missense_Mutation

Table A6. EGFR mutation in sensitive or resistant NSCLC cell lines in response to treatment with gefitinib and erlotinib, respectively.

		Sensitive	Resistant
Gefitinib	IC ₅₀	HCC4006, HCC827, LOUNH91, NCIH1355, NCIH1568	HCC1359, NCIH1793, NCIH226, SKMES1
	EC ₅₀	HCC4006, HCC827, NCIH1355, NCIH1568, NCIH1975, NCIH226, PC14	HCC1359, SKMES1
Erlotinib	IC ₅₀	HCC4006, HCC827, NCIH1355, NCIH1650, NCIH1793	
	EC ₅₀	HCC1359, HCC4006, HCC827, NCIH1355, NCIH1650, NCIH1975, PC14	NCIH226, SKMES1

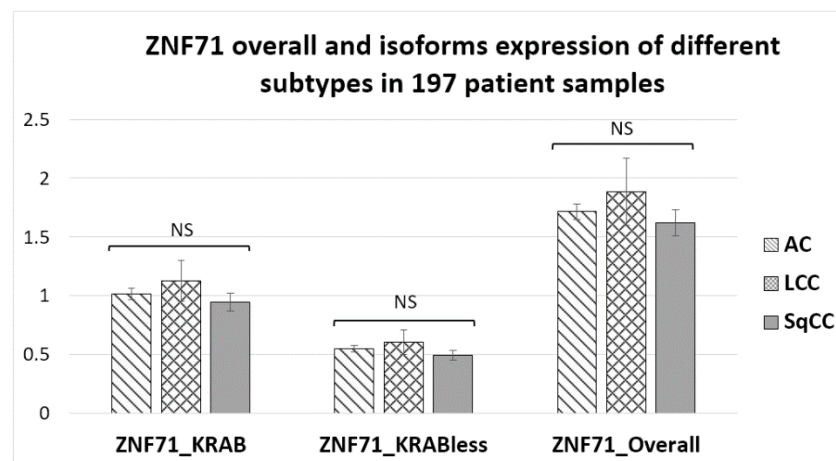


Figure A1. ZNF71 expression in different NSCLC histological subtypes in patient tumors. The expression of overall ZNF71 and its isoforms was not significantly different among different histological subtypes in patient tumors (ANOVA tests). NS: non-significance; AC: adenocarcinoma; LCC: large cell carcinoma; SqCC: squamous cell carcinoma.

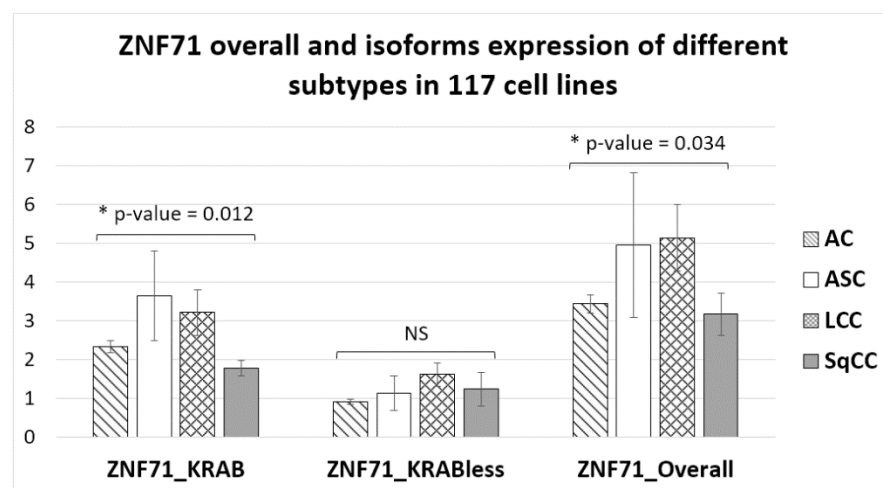


Figure A2. ZNF71 expression in different NSCLC histological subtypes in cell lines. The expression of ZNF71 overall and KRAB isoform was significantly different among histological subtypes ($p < 0.05$, ANOVA tests). Large cell carcinoma (LCC) had the highest ZNF71 overall expression. Adenosquamous carcinoma (ASC) had the highest ZNF71 KRAB expression. Squamous cell carcinoma (SqCC) had the lowest expression of both ZNF71 overall and KRAB isoform. NS: non-significance; AC: adenocarcinoma.

References

1. Lung Cancer-Non-Small Cell: Statistics. Available online: <https://www.cancer.net/cancer-types/lung-cancer-non-small-cell/statistics> and <https://www.cancer.net/cancer-types/melanoma/statistics> (accessed on 2 October 2020).
2. Hoffman, P.C.; Mauer, A.M.; Vokes, E.E. Lung cancer. *Lancet* **2000**, *355*, 479–485. [CrossRef]
3. Zheng, Y.; Jaklitsch, M.T.; Bueno, R. Neoadjuvant Therapy in Non-Small Cell Lung Cancer. *Surg. Oncol. Clin. North Am.* **2016**, *25*, 567–584. [CrossRef]
4. Crinò, L.; Weder, W.; Van Meerbeeck, J.; Felip, E. 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] [PubMed]
5. Byron, E.; Pinder-Schenck, M. Systemic and targeted therapies for early-stage lung cancer. *Cancer Control.* **2014**, *21*, 21–31. [CrossRef]

6. 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](#)]
7. 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](#)]
8. 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](#)]
9. 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](#)]
10. Bruno, M.; Mahgoub, M.; Macfarlan, T.S. The Arms Race Between KRAB-Zinc Finger Proteins and Endogenous Retroelements and Its Impact on Mammals. *Annu. Rev. Genet.* **2019**, *53*, 393–416. [[CrossRef](#)]
11. Ecco, G.; Imbeault, M.; Trono, D. KRAB zinc finger proteins. *Development* **2017**, *144*, 2719–2729. [[CrossRef](#)]
12. 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](#)]
13. Murakami, T.; Matakai, C.; Nagao, C.; Umetani, M.; Wada, Y.; Ishii, M.; Tsutsumi, S.; Kohro, T.; Saiura, A.; Aburatani, H.; et al. The Gene Expression Profile of Human Umbilical Vein Endothelial Cells Stimulated by Tumor Necrosis Factor α using DNA Microarray Analysis. *J. Atheroscler. Thromb.* **2000**, *7*, 39–44. [[CrossRef](#)] [[PubMed](#)]
14. Kim, J.-H.; Cheong, H.S.; Park, J.S.; Jang, A.-S.; Uh, S.-T.; Kim, Y.-H.; Kim, M.-K.; Choi, I.S.; Cho, S.H.; Choi, B.W.; et al. A Genome-Wide Association Study of Total Serum and Mite-Specific IgEs in Asthma Patients. *PLoS ONE* **2013**, *8*, e71958. [[CrossRef](#)]
15. Roche, J. The Epithelial-to-Mesenchymal Transition in Cancer. *Cancers* **2018**, *10*, 52. [[CrossRef](#)]
16. Nieto, M.A.; Huang, R.Y.-J.; Jackson, R.A.; Thiery, J.P. EMT: 2016. *Cell* **2016**, *166*, 21–45. [[CrossRef](#)] [[PubMed](#)]
17. Lambert, A.W.; Pattabiraman, D.R.; Weinberg, R.A. Emerging Biological Principles of Metastasis. *Cell* **2017**, *168*, 670–691. [[CrossRef](#)]
18. Moustakas, A.; De Herreros, A.G. Epithelial-mesenchymal transition in cancer. *Mol. Oncol.* **2017**, *11*, 715–717. [[CrossRef](#)] [[PubMed](#)]
19. Block, C.J.; Dyson, G.; Campeanu, I.J.; Watza, 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*, 1–14. [[CrossRef](#)] [[PubMed](#)]
20. Yoshihara, K.; Shahmoradgoli, M.; Martínez, E.; Vegesna, R.; Kim, H.; Torres-García, W.; Treviño, V.; Shen, H.; Laird, P.W.; Levine, D.A.; et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* **2013**, *4*, 2612. [[CrossRef](#)] [[PubMed](#)]
21. Panchy, N.; Azeredo-Tseng, C.; Luo, M.; Randall, N.; Hong, T. Integrative Transcriptomic Analysis Reveals a Multiphasic Epithelial-Mesenchymal Spectrum in Cancer and Non-tumorigenic Cells. *Front. Oncol.* **2020**, *9*, 1479. [[CrossRef](#)]
22. Friedman, J.R.; Fredericks, W.J.E.; Jensen, D.; Speicher, D.W.; Huang, X.P.; Neilson, E.G.; Rauscher, F.J. KAP-1, a novel corepressor for the highly conserved KRAB repression domain. *Genes Dev.* **1996**, *10*, 2067–2078. [[CrossRef](#)]
23. Chen, L.; Munoz-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](#)]
24. 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](#)]
25. Mezheyeuski, A.; Bergsland, C.H.; Backman, M.; Djureinovic, D.; Sjöblom, 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](#)] [[PubMed](#)]
26. Broad Institute. DepMap 20Q2 Public. In *DepMap*; Broad Institute: Cambridge, MA, USA, 2020.
27. 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](#)]
28. 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*, 1–13. [[CrossRef](#)] [[PubMed](#)]
29. Furuse, K. Platinum/Oral Etoposide Therapy in Non-Small Cell Lung Cancer. *Oncology* **1992**, *49*, 63–70. [[CrossRef](#)] [[PubMed](#)]
30. Steuer, C.E.; Behera, M.; Ernani, V.; Higgins, K.A.; Saba, N.F.; Shin, D.M.; Pakkala, S.; Pillai, R.N.; Owonikoko, T.K.; Curran, W.J.; et al. Comparison of Concurrent use of Thoracic Radiation with Either Carboplatin-Paclitaxel or Cisplatin-Etoposide for Patients with Stage III Non-Small-Cell Lung Cancer: A Systematic Review. *JAMA Oncol.* **2017**, *3*, 1120–1129. [[CrossRef](#)]
31. Yang, Z.; Hackshaw, A.; Feng, Q.; Fu, X.; Zhang, Y.; Mao, C.; Tang, J. Comparison of gefitinib, erlotinib and afatinib in non-small cell lung cancer: A meta-analysis. *Int. J. Cancer* **2017**, *140*, 2805–2819. [[CrossRef](#)]
32. Fossella, F.V. Docetaxel for previously treated non-small-cell lung cancer. *Oncology* **2002**, *16* (Suppl. 6), 45–51.
33. Arrieta, O.; Barrón, F.; Ramírez-Tirado, L.A.; Zatarain-Barrón, Z.L.; Cardona, A.F.; Díaz-García, D.; Yamamoto Ramos, M.; Mota-Vega, B.; Carmona, A.; Peralta Álvarez, M.P.; et al. Efficacy and Safety of Pembrolizumab Plus Docetaxel vs Docetaxel

- Alone in Patients with Previously Treated Advanced Non-Small Cell Lung Cancer: The PROLUNG Phase 2 Randomized Clinical Trial. *JAMA Oncol.* **2020**, *6*, 856–864. [CrossRef] [PubMed]
34. Hayashi, H.; Kurata, T.; Nakagawa, K. Gemcitabine: Efficacy in the Treatment of Advanced Stage Nonsquamous Non-Small Cell Lung Cancer. *Clin. Med. Insights Oncol.* **2011**, *5*. [CrossRef] [PubMed]
35. Ramalingam, S.; Belani, C.P. Paclitaxel for non-small cell lung cancer. *Expert Opin. Pharmacother.* **2004**, *5*, 1771–1780. [CrossRef] [PubMed]
36. Pearce, M.C.; Gamble, J.T.; Kopparapu, P.R.; O'Donnell, E.F.; Mueller, M.J.; Jang, H.S.; Greenwood, J.A.; Satterthwait, A.C.; Tanguay, R.L.; Zhang, X.-K.; et al. Induction of apoptosis and suppression of tumor growth by Nur77-derived Bcl-2 converting peptide in chemoresistant lung cancer cells. *Oncotarget* **2018**, *9*, 26072–26085. [CrossRef] [PubMed]
37. Carnio, S.; Novello, S.; Papotti, M.; Iacono, M.L.; Scagliotti, G.V. Prognostic and predictive biomarkers in early stage non-small cell lung cancer: Tumor based approaches including gene signatures. *Transl. Lung Cancer Res.* **2013**, *2*, 372–381. [PubMed]
38. Kosaka, T.; Yatabe, Y.; Onozato, R.; Kuwano, H.; Mitsudomi, T. Prognostic Implication of EGFR, KRAS, and TP53 Gene Mutations in a Large Cohort of Japanese Patients with Surgically Treated Lung Adenocarcinoma. *J. Thorac. Oncol.* **2009**, *4*, 22–29. [CrossRef] [PubMed]
39. Pécuchet, N.; Laurent-Puig, P.; Mansuet-Lupo, A.; Legras, A.; Alifano, M.; Pallier, K.; Didelot, A.; Gibault, L.; Danel, C.; Just, P.-A.; et al. Different prognostic impact of STK11 mutations in non-squamous non-small-cell lung cancer. *Oncotarget* **2015**, *8*, 23831–23840. [CrossRef] [PubMed]
40. Lord, R.V.N.; Brabender, J.; Gandara, D.; Alberola, V.; Camps, C.; Domine, M.; Cardenal, F.; Sánchez, J.M.; Gumerlock, P.H.; Tarón, M.; et al. Low ERCC1 expression correlates with prolonged survival after cisplatin plus gemcitabine chemotherapy in non-small cell lung cancer. *Clin. Cancer Res.* **2002**, *8*, 2286–2291.
41. Simon, G.R.; Sharma, S.; Cantor, A.; Smith, P.; Bepler, G. ERCC1 Expression Is a Predictor of Survival in Resected Patients with Non-small Cell Lung Cancer. *Chest* **2005**, *127*, 978–983. [CrossRef]
42. 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]
43. Chen, D.-T.; Hsu, Y.-L.; Fulp, W.J.; Coppola, D.; Haura, E.B.; Yeatman, T.J.; Cress, W.D. Prognostic and Predictive Value of a Malignancy-Risk Gene Signature in Early-Stage Non-Small Cell Lung Cancer. *J. Natl. Cancer Inst.* **2011**, *103*, 1859–1870. [CrossRef] [PubMed]
44. Kratz, J.R.; He, J.; Eeden, S.K.V.D.; Zhu, Z.-H.; Gao, W.; Pham, P.T.; Mulvihill, M.S.; Ziaei, F.; Zhang, H.; Su, B.; et al. A practical molecular assay to predict survival in resected non-squamous, non-small-cell lung cancer: Development and international validation studies. *Lancet* **2012**, *379*, 823–832. [CrossRef]
45. Tang, H.; Xiao, G.; Behrens, C.; Schiller, J.; Allen, J.; Chow, C.-W.; Suraokar, M.; Corvalan, A.; Mao, J.; White, M.A.; et al. A 12-Gene Set Predicts Survival Benefits from Adjuvant Chemotherapy in Non-Small Cell Lung Cancer Patients. *Clin. Cancer Res.* **2013**, *19*, 1577–1586. [CrossRef] [PubMed]
46. Woodard, G.A.; Wang, S.X.; Kratz, J.R.; Zoon-Besselink, C.T.; Chiang, C.-Y.; Gubens, M.A.; Jahan, T.M.; Blakely, C.M.; Jones, K.D.; Mann, M.J.; et al. Adjuvant Chemotherapy Guided by Molecular Profiling and Improved Outcomes in Early Stage, Non-Small-Cell Lung Cancer. *Clin. Lung Cancer* **2018**, *19*, 58–64. [CrossRef]
47. Houston, K.A.; Henley, S.J.; Li, J.; White, M.C.; Richards, T.B. Patterns in lung cancer incidence rates and trends by histologic type in the United States, 2004–2009. *Lung Cancer* **2014**, *86*, 22–28. [CrossRef] [PubMed]
48. Zheng, Y.; Bueno, R. Commercially available prognostic molecular models in early-stage lung cancer: A review of the Perenio Lung RS and Myriad myPlan Lung Cancer tests. *Expert Rev. Mol. Diagn.* **2015**, *15*, 589–596. [CrossRef]
49. Myriad. All Products. Available online: <https://myriad.com/products-services/all-products/overview> (accessed on 31 December 2020).
50. Klug, A. The Discovery of Zinc Fingers and Their Applications in Gene Regulation and Genome Manipulation. *Annu. Rev. Biochem.* **2010**, *79*, 213–231. [CrossRef]
51. Yang, P.; Wang, Y.; Macfarlan, T.S. The Role of KRAB-ZFPs in Transposable Element Repression and Mammalian Evolution. *Trends Genet.* **2017**, *33*, 871–881. [CrossRef]
52. Addison, J.B.; Koontz, C.; Fugett, J.H.; Creighton, C.J.; Chen, N.; 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.* **2014**, *75*, 344–355. [CrossRef]
53. Ho, J.; Kong, J.-W.-F.; Choong, L.-Y.; Loh, M.-C.-S.; Toy, W.; Chong, P.-K.; Wong, C.-H.; Wong, C.-Y.; Shah, N.; Lim, Y.-P. Novel Breast Cancer Metastasis-Associated Proteins. *J. Proteome Res.* **2009**, *8*, 583–594. [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. Tulchinsky, E.; Demidov, O.; Kriajevska, M.; Barlev, N.A.; Imyanitov, E. EMT: A mechanism for escape from EGFR-targeted therapy in lung cancer. *Biochim. Biophys. Acta (BBA) Bioenerg.* **2019**, *1871*, 29–39. [CrossRef]
56. 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]

57. Li, J.; Choi, P.S.; Chaffer, C.L.; Labella, K.; Hwang, J.H.O.; Giacomelli, A.; 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](#)]
58. Chen, L.; Chen, D.-T.; Kurtyka, C.; Rawal, B.; Fulp, W.J.; Haura, E.B.; Cress, W. Tripartite Motif Containing 28 (Trim28) Can Regulate Cell Proliferation by Bridging HDAC1/E2F Interactions. *J. Biol. Chem.* **2012**, *287*, 40106–40118. [[CrossRef](#)]
59. Czerwińska, P.; Mazurek, S.; Wiznerowicz, M. The complexity of TRIM28 contribution to cancer. *J. Biomed. Sci.* **2017**, *24*, 1–14. [[CrossRef](#)] [[PubMed](#)]
60. Sobocińska, J.; Molenda, S.; Machnik, M.; Oleksiewicz, U. KRAB-ZFP Transcriptional Regulators Acting as Oncogenes and Tumor Suppressors: An Overview. *Int. J. Mol. Sci.* **2021**, *22*, 2212. [[CrossRef](#)]
61. Huang, C.; Jia, Y.; Yang, S.; Chen, B.; Sun, H.; Shen, F.; Wang, Y. Characterization of ZNF23, a KRAB-containing protein that is downregulated in human cancers and inhibits cell cycle progression. *Exp. Cell Res.* **2007**, *313*, 254–263. [[CrossRef](#)] [[PubMed](#)]
62. He, B.; Wang, B.; Wang, H.; Zhang, C.; Wu, Y.; Fu, L.; Yu, G. RBAK is upregulated in non-small cell lung cancer and promotes cell migration and invasion. *Exp. Ther. Med.* **2019**, *18*, 2942–2948. [[CrossRef](#)] [[PubMed](#)]
63. 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](#)]
64. 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 March 2021).
65. 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](#)]
66. 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. Non-oncology drugs are a source of previously unappreciated anti-cancer activity. *BioRxiv* **2019**. [[CrossRef](#)]
67. Ma, Y.; Ding, Z.; Qian, Y.; Shi, X.; Castranova, V.; Harner, E.J.; Guo, L. Predicting Cancer Drug Response by Proteomic Profiling. *Clin. Cancer Res.* **2006**, *12*, 4583–4589. [[CrossRef](#)] [[PubMed](#)]
68. Guo, N.I.; Ma, Y.; Ding, Z.; Qian, Y.; Wan, Y.-W.; Tosun, K.; Shi, X.; Castranova, V.; Harner, E.J. An integrative genomic and proteomic approach to chemosensitivity prediction. *Int. J. Oncol.* **1992**, *34*, 107–115. [[CrossRef](#)]
69. Schwartz, L.H.; Litière, S.; De Vries, E.; Ford, R.; Gwyther, S.; Mandrekar, S.; Shankar, L.; Bogaerts, J.; Chen, A.; Dancy, J.; et al. RECIST 1.1—Update and clarification: From the RECIST committee. *Eur. J. Cancer* **2016**, *62*, 132–137. [[CrossRef](#)] [[PubMed](#)]
70. 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**. [[CrossRef](#)] [[PubMed](#)]
71. RS Team. *RStudio: Integrated Development for R*; RStudio Inc.: Boston, MA, USA, 2019; Available online: <http://www.rstudio.com> (accessed on 1 April 2021).