Molecular Therapy
Nucleic Acids
Original Article

Cell-type-specific alternative polyadenylation promotes oncogenic gene expression in non-small cell lung cancer progression

Kexin Huang,[1,2,3,5] Yun Zhang,[1,5] Xiaorui Shi,[1,5] Zhiqin Yin,[1] Weiling Zhao,[3] Liyu Huang,[1] Fu Wang,[1,4] and Xiaobo Zhou[3]

[1]School of Life Science and Technology, Xidian University, Xi'an, Shaanxi 710071, China; [2]West China Biomedical Big Data Centre, West China Hospital of Sichuan University, Chengdu 610041, China; [3]Center for Computational Systems Medicine, School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, TX 77030, USA; [4]School of Pharmacy, Shaanxi Institute of International Trade and Commerce, Xianyang, Shaanxi 712046, China

Disrupted alternative polyadenylation (APA) is frequently involved in tumorigenesis and cancer progression by regulating the gene expression of oncogenes and tumor suppressors. However, limited knowledge of tumor-type- and cell-type-specific APA events may lead to novel APA events and their functions being overlooked. Here, we compared APA events across different cell types in non-small cell lung cancer (NSCLC) and normal tissues and identified functionally related APA events in NSCLC. We found several cell-specific 3′-UTR alterations that regulate gene expression changes showed prognostic value in NSCLC. We further investigated the function of APA-mediated 3′-UTR shortening through loss of microRNA (miRNA)-binding sites, and we identified and experimentally validated several oncogene-miRNA-tumor suppressor axes. According to our analyses, we found SPARC as an APA-regulated oncogene in cancer-associated fibroblasts in NSCLC. Knockdown of SPARC attenuates lung cancer cell invasion and metastasis. Moreover, we found high SPARC expression associated with resistance to several drugs except cisplatin. NSCLC patients with high SPARC expression could benefit more compared to low-SPARC-expression patients with cisplatin treatment. Overall, our comprehensive analysis of cell-specific APA events shed light on the regulatory mechanism of cell-specific oncogenes and provided opportunities for combination of APA-regulated therapeutic target and cell-specific therapy development.

INTRODUCTION

Non-small cell lung cancer (NSCLC) is one of the most frequently diagnosed cancers, with a 5-year survival rate of around 25%.[1] Since the occurrence and development of NSCLC involve complex genetic alterations, many dysregulated genes have been identified through gene expression profiling and have been applied for clinical outcome and target response prediction.[2] Numerous studies revealed that gene expression profiles can be dysregulated through a variety of mechanisms, such as exon deletion, copy number variation, and epigenetic modification.[3–5] However, the underlying mechanisms driving dysregulated gene expression in NSCLC remain incomplete. Alternative polyadenylation (APA) is an important post-transcript regulation process that can regulate gene expression by generating mRNA isoforms with different 3′-UTR lengths.[6] It is also a very common process, and more than 70% of genes in the mammalian genome undergo APA.[6] Altered 3′ UTR causes the inclusion or deletion of *cis*-regulatory elements, such as microRNA (miRNA)-binding sites, which have broad impact on mRNA expression, stability, and translation as well as cellular localization and protein diversification.[7] Recently, considerable evidence shows that dysregulated APA plays an important role in tumorigenesis and progression.[8] Proto-oncogenes tend to use the distal poly(A) sites in normal cells, which allows miRNA-guided regulation. During the tumorigenesis process, proto-oncogenes begin to use proximal poly(A) sites and generate shortened 3′ UTRs.[7] This results in the loss of miRNA-binding sites and the escape of mRNA from miRNA repression, leading to aberrantly increased expression of oncogenes. Dysregulated APA has been found in multiple cancers.[9–11] For example, APA events have been found to play a role in the regulation of CPSF1 and PABPN1 expression, with shorter 3′-UTR isoforms associated with increased expression levels in breast cancer.[12] The depletion of CPSF1 or PABPN1 weakened cancer cell proliferation. In addition, APA has been shown to regulate the expression of tumor suppressor genes, such as CFIm25, in glioblastoma.[13] Extensive evidence also shows that dysregulated APAs in NSCLC lead to altered expression of certain genes, such as PABPN1, CPEB1, E2F1, and FGF2. Lower PABPN1 expression was associated with poor prognosis in NSCLC patients.[14]

Received 10 March 2023; accepted 8 August 2023;
https://doi.org/10.1016/j.omtn.2023.08.005.

[5]These authors contributed equally

Correspondence: Liyu Huang, School of Life Science and Technology, Xidian University, Xi'an, Shaanxi 710071, China.
E-mail: huangly@mail.xidian.edu.cn
Correspondence: Fu Wang, School of Life Science and Technology, Xidian University, Xi'an, Shaanxi 710071, China.
E-mail: fwang@xidian.edu.cn
Correspondence: Xiaobo Zhou, Center for Computational Systems Medicine, School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, TX 77030, USA.
E-mail: xiaobo.zhou@uth.tmc.edu

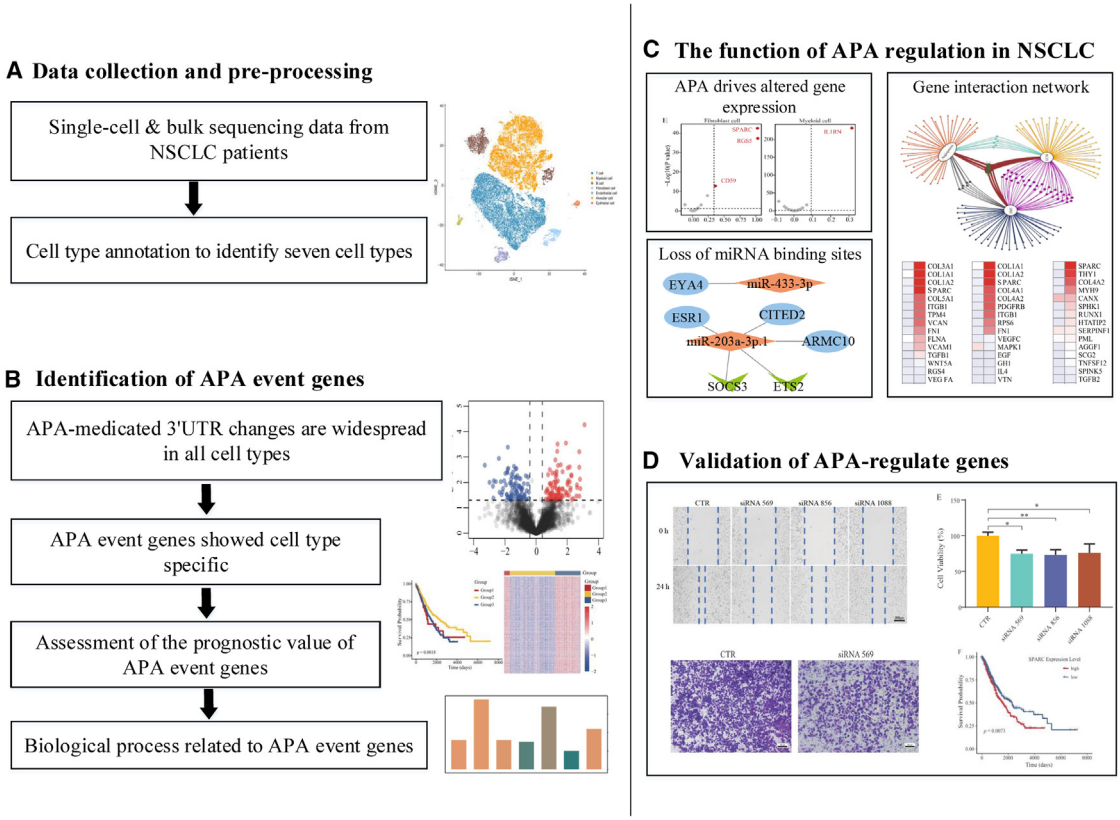


Figure 1. The workflow of the study

(A) scRNA data collection, preprocessing, and cell type annotation. (B) The identification of APA event genes in several cell types and the enrichment analysis for APA event genes. (C) The function analysis of APA regulation, including differential APA gene expression analysis, miRNA-binding sites analysis, and gene-interaction network construction analysis. (D) The cell line experiments we used to validated the function of APA gene in NSCLC.

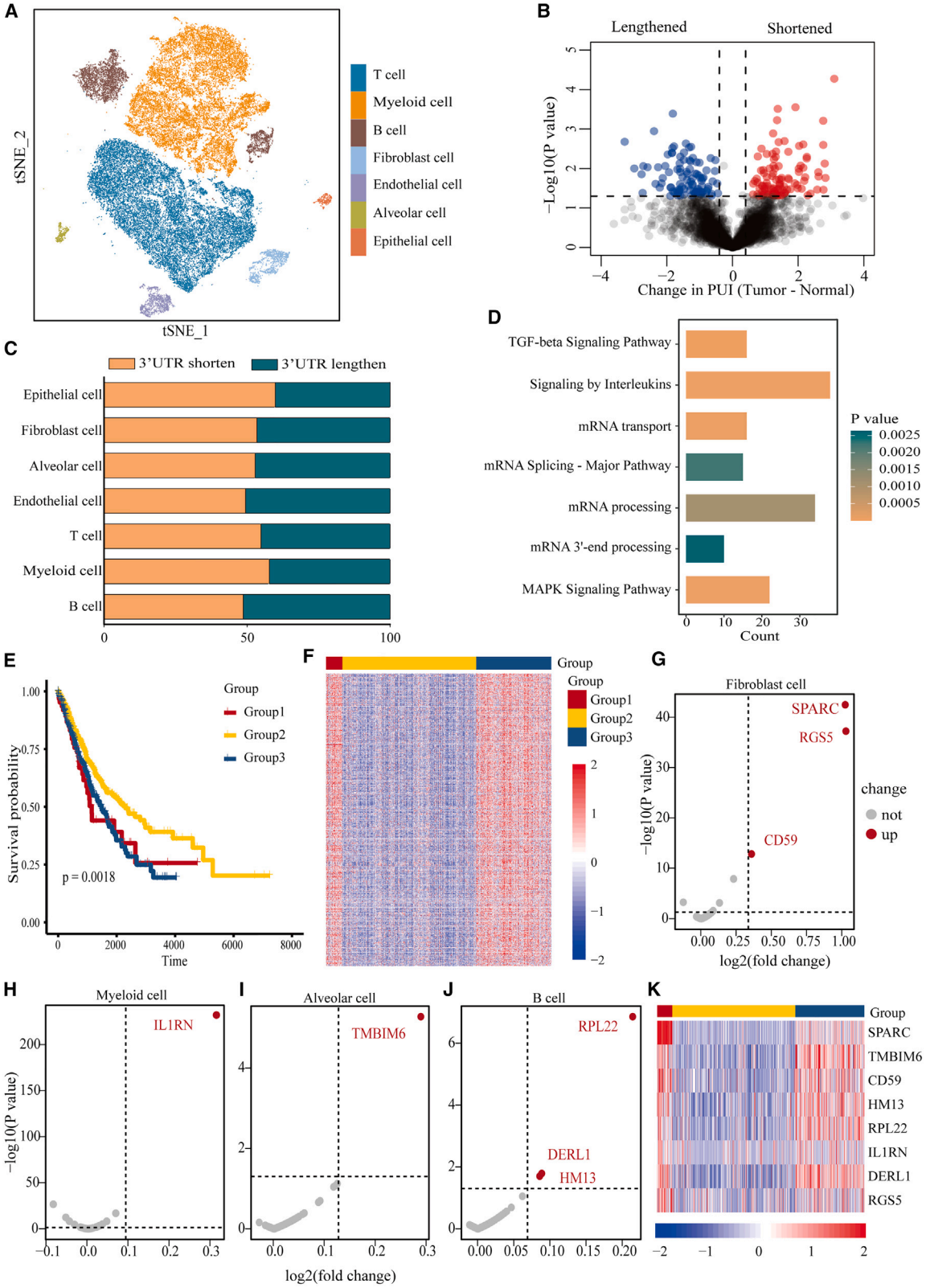
Although these studies have uncovered several APA events that are related to prognosis, our knowledge about dynamic APA regulation in different cell types remains rudimentary.[15] The development of single-cell RNA (scRNA) sequencing (scRNA-seq) provides new opportunities to explore transcriptome variability across different cell types. Previous study identified APA variations using full-length scRNA-seq data in glioblastoma, breast cancer, and renal cell carcinoma, and found that the cell-type-specific APA events showed differences across multiple cancers. For example, SET nuclear proto-oncogene was defined as a tumor cell-specific signature with a 3′-UTR-shortening APA event in the breast cancer.[16] Moreover, a prior study identified APA events in neurons that may influence cell differentiation associated with autism susceptibility.[17] The discovery of cell-type-specific APA regulation largely extends the underlying genetic alteration in diseases. To date, the landscape and the functions of diverse APA events in different cell types remain completely ambiguous in NSCLC.

In this study, we performed comprehensive analyses of APA events in NSCLC using scRNA-seq data to explore the potential role and related mechanism of APA in different cell types, as shown in Figure 1. We systematically analyzed APA events for 47,062 cells and found broadly dysregulated APA events in seven cell types. Moreover, loss of miRNA-binding sites caused by shortened 3′ UTR is associated with many human cancers. We also investigated the impact of APA-induced miRNA regulation. Our results revealed dynamic and widespread APA events in NSCLC. Most APA events are cell type specific and regulate oncogene expression such as SPARC through loss of miRNA inhibition in NSCLC. Combined with our results, shortened 3′ UTR of SPARC can lead to high expression level in cancer-associated fibroblasts (CAFs). High expression of SPARC was proved to induce the epithelial-mesenchymal transition (EMT) and promote the migration and metastasis in lung cancer cell lines. Further, cisplatin treatment was identified to have a significant effect on high-SPARC-expression patients rather than low-expression patients. This result indicated the potential of APA-regulated genes to improve stratification-based treatment strategies in NSCLC.

RESULTS

APA-regulated 3′-UTR changes in different NSCLC cell types

After data preprocessing, we obtained transcriptomic profiles of 21,340 genes for 47,062 cells in normal and NSCLC samples. Based on



(legend on next page)

previously identified biomarkers, we annotated the top 20 principal components (PCs) into seven cell types, namely T cell, myeloid cell, B cell, fibroblast cell, endothelial cell, alveolar cell, and epithelial cell (Figure 2A). Then, we identified APA events in each cell using the scAPA method. To identify the differential APA events between NSCLC and the normal tissue, we first compared poly(A) site usage index (PUI) scores of each gene for each cell type between the control and NSCLC groups using the Student's t test with $p < 0.05$. To get more reliable results, we also used ΔPUI to identify tumor-associated 3′-UTR shortening or lengthening in several cell types (Figures 2B and S2). We found 1,236 genes showed APA-regulated 3′-UTR changes, including 586 3′-UTR-lengthening events and 650 3′-UTR-shortening events. Figures S3–S9 show the landscape of APA in seven cell types in NSCLC. As shown in Figure 2C, we observed a higher number of 3′-UTR-shortening events compared to lengthening events in most cell types in NSCLC, which shows consistency with previous studies.[18] Furthermore, to identify cell-type-specific APA events, we also calculated the proportion of APA events only shown in one cell type. As shown in Figure S10, we found most APA events showed cell-type-specific characteristics. For example, of the 153 APA events in epithelial cells, 101 were epithelial-specific APA events. This result supports the prior study that APA events are highly cell-type specific. To investigate the biological relevance of the APA events, we performed pathway analysis for APA event genes (including both 3′-UTR-shortening and -lengthening event genes). Figure 2D and Table S1 show APA event genes involved in mRNA processing and splicing pathways. We also found several pathways were reported to have associations with NSCLC, such as transforming growth factor β (TGF-β) signaling pathway, mitogen-activated protein kinase (MAPK) signaling pathway, and epidermal growth factor (EGF)/EGF receptor (EGFR) signaling pathway.[19] These results indicated that the regulation of APA may have effect on the activity of NSCLC-related pathways and have a potential role in the tumorigenesis and progression.

Prognostic value of APA-regulated 3′-UTR-shortening genes

To investigate whether APA event genes have prognostic value in NSCLC patients, we performed hierarchical clustering analysis of 3′-UTR-shortening genes in The Cancer Genome Atlas (TCGA) NSCLC cohort. As shown in Figure S11, we identified three subgroups with distinct prognosis using APA event genes. As shown in Figures 2E and 2F, group 2 showed lower expression of APA event genes than group 1 and 3, and the survival analysis shows the prognosis of group 2 is significantly better than other groups ($p = 0.0018$). This result indicates the expression pattern of APA event genes may serve as a prognostic signature in NSCLC.

APA can regulate gene expression alteration

We performed differential gene expression analysis of APA-mediated 3′-UTR-shortening genes in different cell types to investigate whether APA events lead to significantly altered expression in NSCLC. As shown in Figures 2G–2J and Figure S12, eight genes in four cell types are significantly upregulated in NSCLC samples, namely SPARC, RGS5, and CD59 in fibroblast cells, IL1RN in myeloid cells, TMBIM6 in alveolar cells, and RPL22, DERL1, and HM13 in B cells. These genes are considered as the significant APA event genes. Moreover, we found these significant APA genes exhibited distinct expression patterns in three subgroups. As shown in Figure 2K, the expression of SPARC is the highest in group 1 and lowest in group 2. Our result indicated that APA-mediated 3′-UTR shortening altered gene expression in NSCLC.

APA regulates gene expression through loss of miRNA-binding sites

To investigate the patterns of APA-induced miRNA-binding site loss in the seven cell types, we used TargetScan Human to search for highly conversed miRNA-binding sites for all 3′-UTR-shortening genes across all cell types. Figure 3A shows the diagram of APA-mediated loss of miRNA-binding sites. As shown in Figure S13, the proportion of miRNA loss in different cell types ranged from approximately 45% to 61%. The results revealed that loss of miRNA-binding sites widely exists in APA event genes. For example, we found 46.7% of genes lost at least one highly conversed miRNA-binding site. It is worth noting that we found 61.3% of genes lost at least one highly conversed miRNA-binding site in CAFs, which is the highest proportion across multiple cell types.

We next focused on the eight significant APA genes to determine whether loss of specific miRNA-binding sites could lead to altered expression. As shown in Table S2, 14 miRNA-binding sites were lost due to 3′-UTR shortening. In order to identify the function of the cell-type-specific APA events on gene expression, we first performed correlation analysis between PUI values and gene expression profiles for each cell-type-specific APA event. Due to higher PUI value indicating shorter 3′ UTR and higher gene expression, we only left the genes that showed significant positive correlation for the downstream miRNA-binding analyses and experiments. In Table S2, the r means the correlation coefficient between the gene expression and PUI value of genes. Asterisk means the significance level of the correlation analysis (*$p < 0.05$ and **$p < 0.01$). The expression of CD59 shows correlation with the PUI values of CD59 in CAFs. However, no escaped miRNAs

Figure 2. APA-mediated 3′-UTR changes are widespread in several cell types and show prognostic value in NSCLC patients

(A) A t-SNE view for cells after QC, color coded by annotated cell type. (B) Identification of APA event genes in cancer-associated fibroblasts (CAFs). Log p value is plotted against ΔPUI for APA event genes. Red points represent 3′-UTR-shortening genes and blue points represent 3′-UTR-lengthening genes. Figures for other cell types are provided in Figure S2. (C) The proportion of 3′-UTR-shortening genes and 3′-UTR-lengthening genes in different cell types. (D) Significant biological pathways enriched in APA event genes (FDR $p < 0.05$). (E) K-M survival analysis for different subgroups 1 (red), subgroup 2 (yellow), and subgroup 3 (blue). (F) NSCLC patients were clustered into three subgroups based on 3′-UTR-shortening genes. (G–J) Differential expression analysis for APA event genes in different cell types. Red dots represent genes with significant gene expression alteration between the normal and the NSCLC samples (significance level, $p < 0.05$). (K) The expression of significant APA genes show distinction between three subgroups in NSCLC patients.

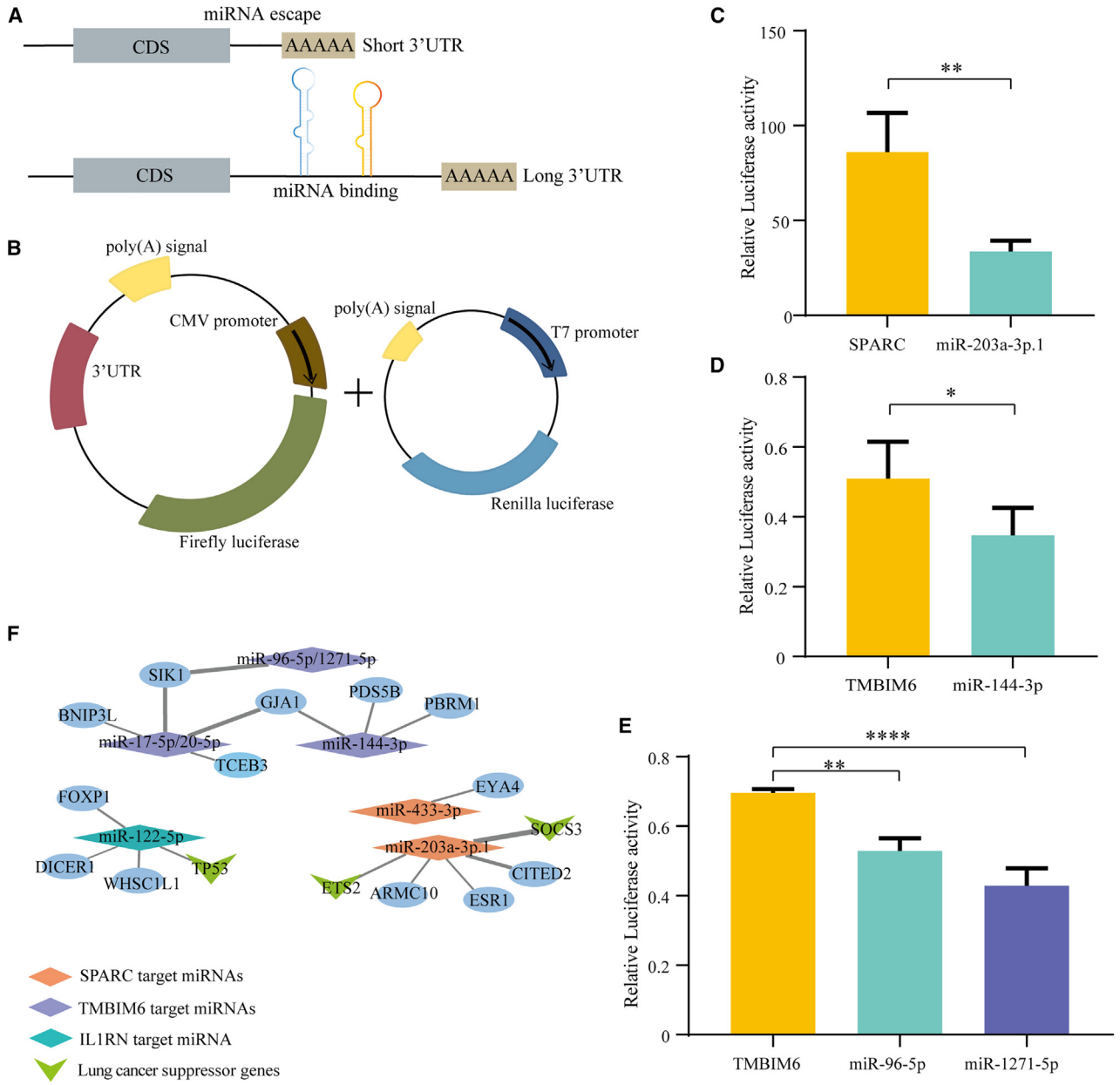
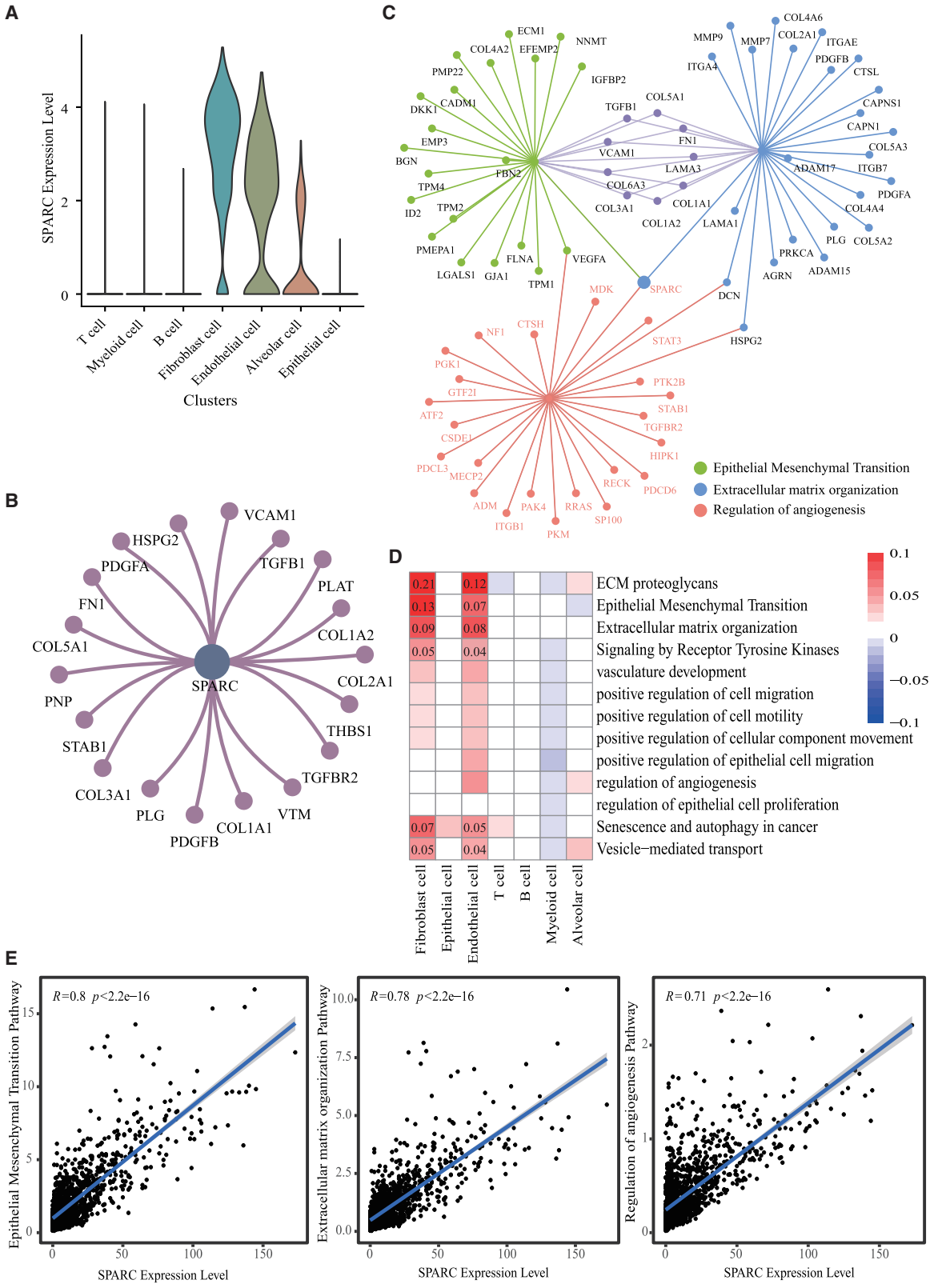


Figure 3. APA could mediate loss of miRNA-binding sites and oncogene-miRNA-tumor suppressor gene axes

(A) The diagram of APA-mediated loss of miRNA-binding sites. (B) The designation of luciferase reporter. (C–E) miRNAs that significantly inhibit luciferase activity of SPARC and TMBIM6. Results for other genes are shown in Figure S12. (F) The oncogene-miRNA-tumor suppressor gene axes in several cell types. Blue circles represent lung-cancer-related genes referred to in previous studies. Green arrows represent known NSCLC suppressor genes.

were identified on CD59 using TargetScan. To further validate the inhibition effect of these miRNAs, we then performed luciferase reporter assays for these miRNAs. Figure 3B shows the procedures of the luciferase reporter experiments. As shown in Figures 3C–3E and S14, eight miRNAs exhibited suppressive effects on the corresponding genes. For example, miR-203a-3p.1 significantly inhibited the expression of SPARC. miR-144-3p, miR-96-5p, and miR-1271-5p significantly suppress the expression of TMBIM6. miR-122-5p has a weak inhibitory effect on IL1RN ($p > 0.05$). Moreover, we also validated the miRNA inhibition with quantitative reverse transcription polymerase chain reaction experiment (qRT-PCR). As shown in Figures S15A and S15B, we verified through qRT-PCR that the mRNA expression of the relevant genes showed a decreasing trend under the action of corresponding



(legend on next page)

miRNAs. The consistency of this result with the luciferase experiments further reinforces the reliability of APA-regulated therapeutic targets at the mRNA level.

Next, we analyzed the target genes of these eight miRNAs to determine whether these miRNAs can target tumor suppressor genes. We selected target genes that were significantly negatively correlated with significant APA event genes (Table S3). Figure 3F shows the oncogene-miRNA-suppressor axes. Blue circles represent lung-cancer-related genes, and green arrows represent lung cancer suppressor genes. We found miR-203a-3p.1 inhibited tumor suppressor genes SOCS3 and ETS2. SOCS3 shows high rank (top 5%) among target genes according to cumulative weighted context++ scores (CWCSs), which indicates a high probability of miR-203a-3p.1-SOCS3 interaction. Tumor suppressor gene TP53 is one of the target genes of miR-122-5p (Figure 3E). We also analyzed the expression of target genes and found that the expression of SOCS3 was significant low in NSCLC samples (Table S4). Our results revealed that escaped miRNAs due to shortened 3′ UTRs may bind to tumor suppressor genes, thereby promoting tumor development.

Association of SPARC with extracellular matrix remodeling in CAFs

SPARC, one of the significant APA event genes, exhibited specific APA-induced 3′-UTR shortening and significant higher expression levels in CAFs ($p < 0.05$) (Figure 4A). We conducted gene-interaction network analysis to explore the potential role of SPARC in CAFs (Table S5). Figure 4B shows the genes significantly interacted with SPARC (significance level, $p < 0.05$), including COL1A1, TGFBR2, and VCAM1. Enrichment analysis shows functional pathways related to these interacted genes, including extracellular matrix (ECM) organization, EMT, vasculature development, and positive regulation of cell migration, motility, and proliferation (Table S6; Figure 4C). The expression of SPARC also shows strong significant correlation with these pathways (Figure 4E). Furthermore, most pathways had the highest activation score in CAFs across all cell types (Figure 4D), indicating that SPARC-involved signaling pathways play a more important role in CAFs. We also determined the activation level of each gene in these pathways, showing that SPARC ranks in the top 10 for most pathways (Table S7).

Requirement of SPARC in lung cancer cells

SPARC is known to involve in ECM synthesis and promotion of cell morphological changes.[20] However, its role as a tumor oncogene or suppressor is cell type and cancer type dependent.[21] Our analysis shows SPARC is highly expressed in CAFs in NSCLC. To validate the role of SPARC in lung cancer proliferation and migration, we knocked down SPARC in lung cancer cell lines with three small interfering RNAs (siRNAs). The role of SPARC in lung cancer cell migration was assessed by wound healing/scratch assay. As shown in Figures 5A–5D, compared with the control group, knocking down SPARC (siRNA1, siRNA2, siRNA3) significantly inhibited tumor cell migration 24 h post transfection. We also analyzed the impact of SPARC on cell proliferation using 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-2H-tetrazolium bromide (MTT) assay. As shown in Figure 5E, the viability of tumor cells transfected with SPARC siRNAs was reduced significantly compared with non-transfected control.

Assessment of the prognostic value of SPARC

Complementarily, we also investigate whether SPARC expression is associated with patient prognosis. We collected tumor data from 864 and 272 NSCLC patients from TCGA and the GEO database. The patients were divided into high-SPARC- and low-SPARC-expression groups according to the median SPARC expression level. We performed Kaplan-Meier (K-M) survival analysis for TCGA and GEO cohorts, respectively. As shown in Figures 5F and 5G, high-SPARC groups are associated with significantly poorer outcomes in NSCLC patients in two independent cohorts. This result suggests that high expression of SPARC may lead to poor prognosis in NSCLC.

Identification of candidate drugs that interact with SPARC

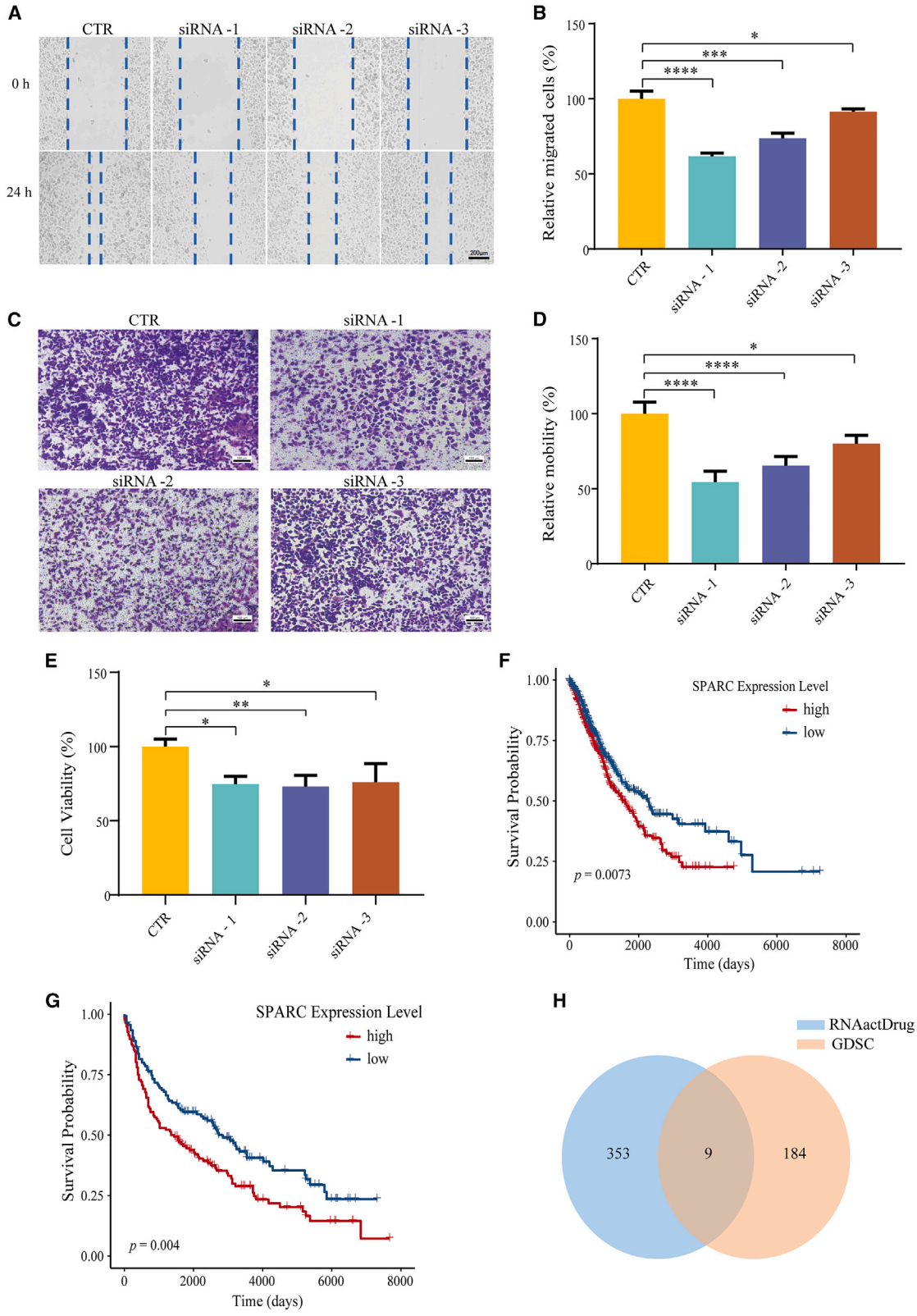
As shown in Figure 5H, we obtained nine candidate drugs that might interact with SPARC according to two databases, namely bexarotene, cisplatin, cytarabine, irinotecan, paclitaxel, temozolomide, topotecan, tretinoin, and palbociclib. Considering the activated oncogene could lead to drug resistance, we further investigated the drug sensitivity in lung cancer cell lines by using Genomics of Drug Sensitivity in Cancer (GDSC). As shown in Figures 6A–6C and S16, the positive correlation between gene expression and half-maximal inhibitory concentration (IC_{50}) indicated the activation of SPARC would lead to multiple drug resistance. Among them, the correlation between SPARC and IC_{50} of cisplatin was significantly negative, which means high SPARC expression would show sensitivity to cisplatin treatment. Therefore, the analysis based on cell lines indicated high-SPARC patients might benefit more from cisplatin treatment.

Distinct drug responses of patients with different expression of SPARC

Considering that patients in TCGA are usually treated with multiple drugs at the same time, we performed Cox regression analysis for the high-SPARC group and the low-SPARC group, respectively. To ensure statistical reliability, we only kept drugs with more than 10 treatment records. Figure 6D shows drug responses of patients in high- and low-SPARC groups. Most drugs did not show significant treatment efficiency in the two groups. Surprisingly, cisplatin had a

Figure 4. The function of SPARC in CAFs

(A) SPARC shows cell-type-specific high expression in CAFs. The significance of expression difference between CAFs and other cell types is $p < 0.05$. (B) Genes that show interactions with SPARC in CAFs. (C) SPARC is the common gene in several pathways enriched by interacted genes. (D) Activation level for SPARC-related pathways in several cell types. Red represents activation and blue represents deactivation. The numbers represent the highest activation scores of the pathway in CAFs. (E) Correlation analysis between SPARC and other genes in pathways.



(legend on next page)

significant effect on the high-SPARC group, compared to the low-SPARC group. The hazard ratios indicated that cisplatin treatment improved prognosis outcomes by 2.961-fold in the high-SPARC group compared to the non-treatment group (Figure 6D). Figure 6E shows the survival between treatment patients and non-treatment patients of high-SPARC group. We observed a significant improvement in the prognosis of patients treated with cisplatin in the high-SPARC group ($p = 0.029$). However, there was no significant improvement in the low-SPARC group ($p = 0.814$) (Figure 6F). Generally, cisplatin, which had *in vitro* and *in silico* evidence, could be considered as potential therapy in NSCLC patients with high SPARC expression level.

DISCUSSION

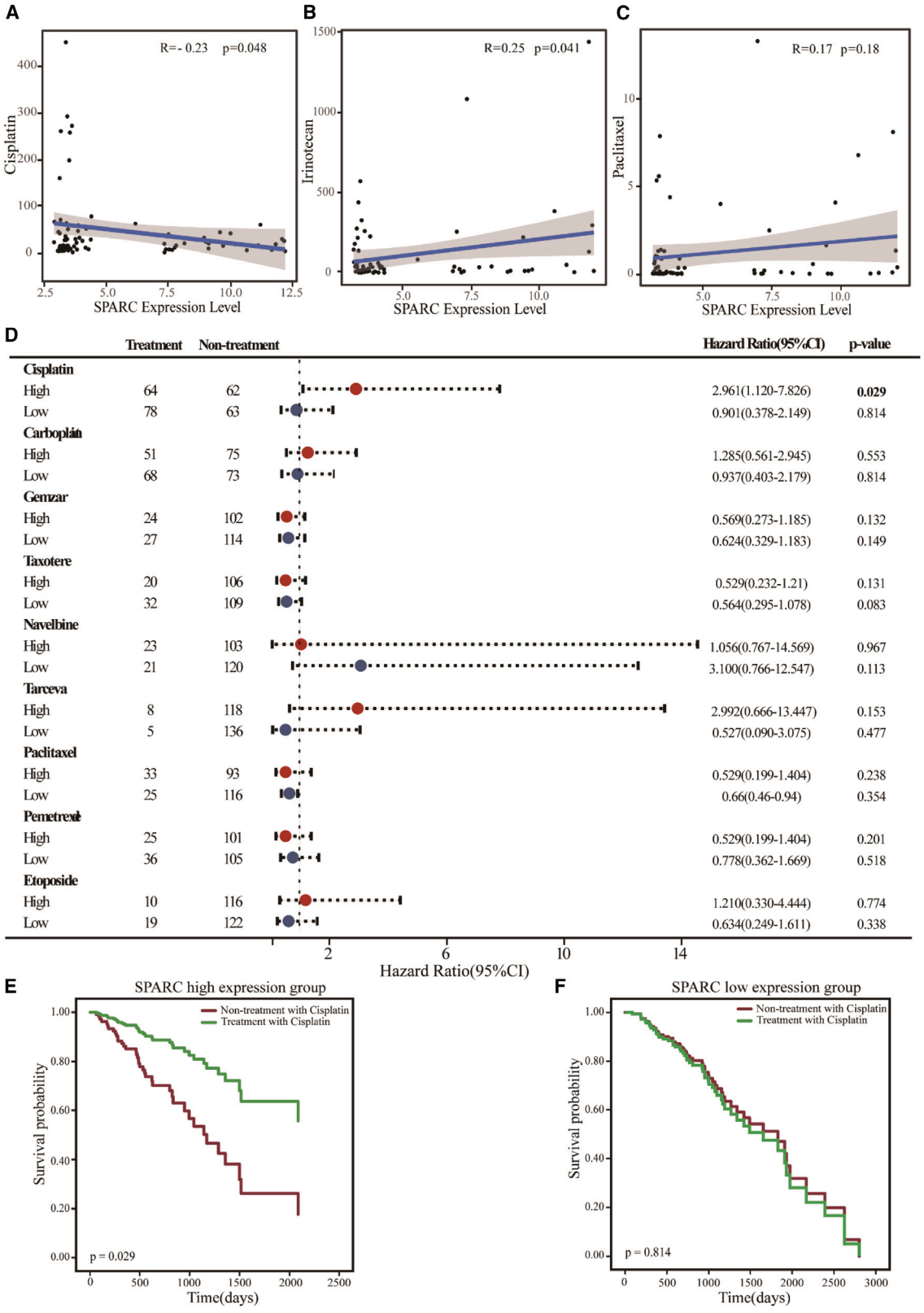
Aberrant gene expression is one of the common characteristics of cancer. However, what causes altered gene expression and whether these genes can be therapeutically targeted remains incompletely understood. APA has been identified as one of the drivers of aberrant gene expression.[8] Extensive APA events were found in multiple cancer types, which could regulate the expression of oncogenes and tumor suppressors.[18] Accumulated evidence indicates many APA events are cancer type and cell state specific.[15,22,23] To date, few attempts have been made to study APA regulation in different cell types to identify cell-specific alterations and vulnerabilities in NSCLC. In this study, we performed systematic identification and functional analysis of APA events in NSCLC using scRNA-seq data. APA events were widespread in NSCLC. We identified 1,236 APA event genes that were significant different from normal tissue in seven cell types. Most APA events were cell-type specific. Among these APA events that differed from normal tissue, there were more APA-mediated 3′-UTR-shortening events than -lengthening events in NSCLC, consistent with previous cancer studies.[10] Several NSCLC-related oncogenes were found to have significantly shortened 3′ UTRs, such as Ras-related protein Rab-14 (RAB14), FOS Like 2, AP-1 Transcription Factor Subunit (FOSL2), transmembrane BAX Inhibitor Motif Containing 6 (TMBIM6), and EH Domain Containing 1 (EHD1).[24–27] The biological pathways of these APA event genes were enriched in mRNA processing and mRNA 3′ end processing. The enriched biological processes also included TGF-β and MAPK signaling pathways.[19] Similar pathways have been reported in several cancer studies, consistent with the emergence of recurrent APA events across multiple cancer types.[18,22] APA events appear to be a potential mechanism for aberrant gene expression in NSCLC.

To explore whether APA events can drive gene expression alteration in different cell types, we analyzed APA-mediated 3′-UTR-shortening genes that are highly expressed in corresponding cell types of NSCLC. We identified several lung-cancer-related genes whose regulatory mechanisms in NSCLC were unknown previously, such as SPARC in fibroblasts, TMBIM6 in alveolar cells, and IL1RN in myeloid cells. Many studies have shown that APA-mediated 3′-UTR shortening may lead to loss of miRNA-binding sites and further increase gene expression.[6] miRNAs play critical roles in tumorigenesis and tumor progression through regulating oncogenes and tumor suppressors.[28] Thus, we investigated APA-mediated loss of miRNA-binding sites and functionally validated these miRNAs by luciferase reporter assays. Our results indicated that miR-203a-3p.1 significantly inhibited the expression of SPARC in CAFs. miR-203a-3p was identified as a tumor suppressor in gastric cancer, but its role in NSCLC remains unknown.[29] Both correlation analysis and CWCS indicated that escaped miR-203a-3p.1 can bind to tumor suppressor gene SOCS3. Previous studies show loss of SOCS3 function promotes tumor progression in NSCLC, and functional recovery of SOCS3 could be one of the platforms in NSCLC treatment.[30] Our results suggested the SPARC-miR-203a-3p.1-SOCS3 axis in CAFs supported the possibility that APA-mediated miRNA dysregulation in cancer is one of underlying mechanisms for activating oncogenes and inactivating suppressor genes.

Our in-depth analysis indicated a vital role of SPARC in CAFs in NSCLC progression and metastasis. Although SPARC was widely studied in several cancer cell lines, the role of SPARC in tumorigenesis is controversial because it exhibits highly cell-type-specific functions.[21] A prior study showed that the low expression of SPARC in lung cancer cells is due to its aberrant methylation.[31] Other studies indicate that the high SPARC expression in stromal cells is associated with lung cancer progression.[20] We demonstrated that APA-regulated 3′-UTR shortening could upregulate SPARC expression in CAFs, lead to tumorigenesis and metastasis promotion, and associate with more aggressive NSCLC phonotype and poorer prognosis. According to our analysis, SPARC is an important modulator in CAFs. SPARC was shown to interact with collagen genes, such as COL1A1, COL1A2, and COL5A1. These genes are important components in ECM organization. We speculated that high SPARC expression could promote cancer progression though inducing ECM remodeling in CAFs. Previous studies support SPARC as a collagen chaperone to modulate collagen binding.[32,33] High expression of SPARC forms a positive feedback loop with metalloproteinase cleavage sites, which in turn generates a tumor microenvironment favorable for tumor cell invasion and metastasis. Besides, our results show interactions between SPARC and TGFB1 and TGFBR2. SPARC-TGF-β has been reported to be a critical modulator of epithelial-mesenchymal transition in lung cancer.[34] SPARC has been reported to enhance the activity of TGF-β by directly interacting with the TGFBR2 complex in mesangial cells.[34] Activated TGF-β receptor

Figure 5. SPARC is required in proliferation, migration, and metastasis in lung cancer cells and could be a prognostic biomarker in NSCLC patients

(A) Wound healing experiment for SPARC knockdown cell lines. (B) The relative cell mobility of wound healing normalized with control group. (C) Transwell assay was performed to measure the cell migration abilities. (D) Relative mobility of Transwell assay normalized with control group. (E) Cells were treated with different siRNA for 48 h. Cell viabilities were evaluated by CCK-8 assay. (F and G) K-M survival analysis between groups with different SPARC level in two independent cohorts (n = 432 in both high- and low-SPARC groups in TCGA cohort; n = 136 in both high- and low-SPARC group in GEO cohort). (H) Nine candidate drugs that might interact with SPARC according to two databases.



(legend on next page)

could signal to several downstream pathways, such as MAPK and Smad pathways, and lead to epithelial program downregulation and mesenchymal program upregulation.[35] Furthermore, the interaction of SPARC-VCAM1 could also promote tumor cell metastasis through enhancing vascular leakiness.[36] Both our results and previous studies demonstrated that SPARC was included in several biological processes related to tumor cell extravasation and metastasis. Moreover, we found SPARC showed over-activation in these pathways, which indicated the central role of SPARC in the NSCLC metastasis processes.

Additionally, the other noteworthy insight from our analysis is that we found high-SPARC-expression patients showed higher sensitivity to cisplatin treatment. Our analysis on cancer cell lines revealed high SPARC expression might lead to resistance to several anti-cancer drugs, such as irinotecan and paclitaxel. However, NSCLC cell lines with high SPARC expression showed sensitivity to cisplatin treatment. To further validate the findings, we investigated the contribution of cisplatin treatment in TCGA-NSCLC cohort. The results indicated that cisplatin treatment could significantly reduce the risk of death by 2,961 times for the high-SPARC group, but this effect was not detected in the low-SPARC group. Cisplatin is one of most common treatment agents for advanced NSCLC patients.[37] Extensive studies reported significant survival advantage of patients receiving cisplatin treatment.[38,39] However, the toxicity of cisplatin could lead to several side effects, such as emesis, renal insufficiency, and neuropathy.[40] Furthermore, some of the patients were intrinsically resistant to cisplatin treatment.[41] Thus, treatment stratification is urgently needed to identify patients who may benefit from cisplatin to reduce ineffective treatment. Our study suggested better therapeutic effect of cisplatin in high-SPARC-expression patients compared with low-SPARC-expression patients, which provided new insights into stratification medicine for NSCLC.

The 3′-UTR lengthening due to APA events may lead to decreased gene expression by providing additional binding sites for regulatory factors, such as RNA-binding proteins and miRNAs. Thus, we also investigated the APA-mediated expression changes in 3′-UTR-lengthening events. We found that four genes in two cell types with 3′-UTR-lengthening events are significantly downregulated (RPL37, MLEC in endothelial cell, and THY1 and ETS2 in fibroblast cells). This result indicates that more genes show 3′-UTR-shortening-mediated upregulation in cancer samples. However, for 3′-UTR-lengthening genes with downregulation, only one gene belongs in the oncogene category according to the CancerMine database. This result shows consistency with previous studies that proto-oncogenes tend to generate shortened 3′ UTRs to avoid the inhibition of regulatory factors in tumorigenesis.[8] Therefore, we only left the 3′-UTR-shortening events for downstream analysis.

To date, numerous studies have demonstrated that APA events occurring in tumors can drive tumor cell proliferation, metastasis, and invasion. For example, APA events on CPSF1 and PABPN1 can significantly promote breast cancer progression and lead to worse prognosis of breast cancer patients. The APA-mediated 3′-UTR shortening on CDK16 lead to the impaired miR-485-5p regulation, which further causes the senescence-associated phenotypes in lung cancer cells.[42] As a key profibrotic transcript, previous study has investigated the APA mechanism of SPARC in skin fibrosis.[43] So far, there are no studies reporting the mechanism of SPARC in NSCLC. Thus, our study completes the APA-mediated oncogenic mechanisms in CAFs in NSCLC and provides new insights into the mechanisms underlying drug resistance in NSCLC patients. However, our identification of APA events on SPARC is based on computational predictions and cell line experiments, and further experimental validation at the tissue level is still required to confirm this conclusion.

MATERIALS AND METHODS

Data collection and preprocessing

NSCLC scRNA-seq data were collected from the previous study.[44] Raw scRNA-seq data were downloaded from ArrayExpress database (ArrayExpress: E-MTAB-6149 and E-MTAB-6653) and GEO database (GEO: GSM3926546). We collected 32 samples consisting of 24 NSCLC tumor samples and eight matched normal samples. The raw FASTQ sequencing files were mapped to the human reference genome *hg19* by using "cellranger count" function with the default setting (CellRanger v.3.1.0). After correcting batch effect, we performed a quality control (QC) process. According to the median number of genes and the percentage of mitochondrial genes in the lung samples, we removed the cells with <1,000 and >20,000 genes and the cells with mitochondrial gene percentage higher than 50%. A total of 47,062 cells and 21,340 genes were left for the downstream analysis. Principal-component analysis (PCA) was used for dimension reduction. Then, we performed t-distributed stochastic neighbor embedding (t-SNE) on the top 20 PCs to visualize the data. The marker genes used for cell type annotation were collected from the previous work.[36] Batch effects removal, QC, and cell type annotation were performed with R package Seurat (v.3.1.4) (R version 4.0, https://www.r-project.org/; https://satijalab.org/seurat/). The processed scRNA-seq data are used to identify the cell-type-specific APA events and the biological functions related to these APA events (including differential expression of APA genes, miRNA-binding sites, and gene-interaction network). The expression profiles of bulk sequencing datasets of NSCLC were obtained from TCGA (https://portal.gdc.cancer.gov/) and the GEO database. For the TCGA cohort, we combined the lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) cohorts with a total of 864 NSCLC patients.

Figure 6. Identification of drug resistance and sensitivity related to SPARC

(A–C) The correlation between SPARC expression and IC_{50} of drugs in lung cancer cell lines. Positive correlation indicates that the increase of gene expression would lead to drug resistance. (D) Contribution of several drugs by using Cox analysis in high-SPARC group and low-SPARC group, respectively. (E) Cox survival analysis between cisplatin treatment and non-treatment group in high-SPARC group. (F) Cox survival analysis between cisplatin treatment and non-treatment group in low-SPARC group. Significance level was set as $p < 0.05$.

The cohort collected from the GEO database (GEO: GSE30219) included 272 NSCLC patients.[45] The bulk sequencing datasets are used to investigate the clinical relationship of the cell-type-specific APA events in NSCLC, including identification and validation of the subgroups with different prognosis and survival analysis. Moreover, in order to identify the potential drug response for cell-type-specific APA genes, we collected NSCLC cell line data with different treatments from the GDSC database (https://www.cancerrxgene.org/database). GDSC provides the IC_{50} value of drugs for multiple cancer cell lines from over 75,000 experiments.

Bioinformatics analysis

Identification of genes with APA events in different cell types

APA events of genes in different cell types were identified by using the scAPA package (APA analysis based on 3′ tag scRNA-seq data) based on raw scRNA-seq data.[46] This method can quantify the relative usage of the most proximal poly(A) site within the 3′ UTR by defining the proximal PUI. For a given gene, PUI represents the proportion of proximal poly(A) site usage on this gene in one cell.[46] The range of PUI is from 0 to 1. A higher PUI score represents more usage of the proximal poly(A) site, thus indicating a shorter 3′ UTR. In order to compare the 3′-UTR alteration between NSCLC and normal samples in different cell types, we performed two-tailed student's t test for PUIs between tumor and normal group in each cell type. Furthermore, we also calculated the average PUI of all cells in each cell type in the normal and cancer groups, respectively. Then, the difference between the mean PUI values in the two groups ($\Delta PUI = meanPUI_{tumor} - meanPUI_{normal}$) was used as a measure of the difference in APA events. For a given APA event, if the ΔPUI is higher than the threshold, it indicates that the mean PUI in the tumor group is higher than the normal group, and this gene may prefer to use the proximal poly(A) site in tumor cells. Then, we identified this APA event as 3′-UTR-shortening event. Similarly, if it is lower than the threshold, we identified this APA event as the 3′-UTR-lengthening event. The threshold and the significance of differences was calculated by using permutation tests (n = 1,000) in each cell type (Figure S1) referred to in a previous study.[22] The significance level was set as $p < 0.05$. The genes with significant changes in PUIs were defined as APA event genes for downstream analyses.

Differential gene expression analysis

To investigate gene expression alterations of APA event genes in different cell types, differential gene expression analysis of APA event genes was performed based on processed scRNA expression data by using the DEseq2 (v.1.40.2) algorithm in R. For scRNA-seq data, the expression levels of various genes may vary significantly across different cell types. In this study, $p < 0.05$ and the dynamic threshold $|logFC| > (mean(abs(logFC)) + 2 \times sd(abs(logFC)))$ were used as the cutoffs to filter differentially expressed genes between tumor and normal control groups.[47] APA event genes showing significant differential expression were defined as significant APA event genes.

Assessment of the prognostic value of APA event genes

To investigate whether APA event genes provide prognostic information for NSCLC patients, we performed unsupervised clustering analysis of 3′-UTR-shortening APA event genes based on the TCGA-NSCLC cohort and GEO cohort using hierarchical agglomerative clustering.[48] The optimal number of subclusters was determined by using the graphical method the Hubert index.[49] Then, the prognosis of these subclusters was described by using K-M survival analysis. Unsupervised clustering was performed using R package Nbclust (v.3.0.1). Survival analysis was performed using the "survival" package (v.3.5–5). The log rank test was used to determine the significance of the survival difference.

miRNA-binding site loss in APA event genes

We used TargetscanHuman (v.8.0) (http://www.targetscan.org/vert_80/) to predict highly conversed miRNA-binding site loss due to APA in NSCLC. Combining with the genomic coordinates of APA events in the shortened 3′ UTRs, the loss of miRNA-binding sites with shortened 3′ UTR for each gene was obtained in each cell type, respectively. In order to investigate the global pattern of miRNA-binding site loss among APA event genes in each cell type, the percentage of genes with at least one highly conversed miRNA-binding site loss was plotted. Then, only miRNA-binding site losses on significant APA event genes were considered for downstream analysis and experimental validation. Further, the miRNAs that lost binding sites on 3′ UTRs may regulate other genes, such as tumor suppressor genes. Thus, we also investigated the target genes of these miRNAs by using TargetscanHuman. Moreover, CWCSs could estimate miRNA repression at the site. For a given miRNA, the lower CWCS indicated the site of this target gene is more favorable than other genes.[22] We selected target genes that were significant negatively correlated with significant APA event gene expression in the same cell type. These genes were further sorted by CWCS to investigate the preference of miRNA binding to target genes. The tumor suppressor gene list of NSCLC is obtained from the CancerMine database (http://bionlp.bcgsc.ca/cancermine).[50] Only the suppressors of the lung cancer will be left.

Gene-interaction analysis

In order to determine the gene function in NSCLC, we applied gene-interaction analysis for SPARC by using the CCInx package in R (https://github.com/BaderLab/CCInx). CCInx is a method to predict intercellular gene interaction based on single-cell data.[51] First, a gene-interaction network was created based on known interactions between receptor, ligand, and ECM proteins. The protein-protein interactions were obtained from iRefindex (v.14), Pathway Commons (v.8), and BioGRID (v.3.4.147).[52–54] The list of all proteins and all interactions is available on the website (https://baderlab.org/CellCellInteractions). Then, we performed enrichment analysis for genes that showed interaction with SPARC to assess the function of SPARC. All settings were held at the default values. The significance level was set as $p < 0.05$.

Gene and pathway activation analysis

To characterize the cell-type-specific functions of SPARC and its regulated genes, we applied a method called quantitative set analysis of gene expression (QuSAGE) to calculate activation scores of

enriched pathways in different cell types.[55] QuSAGE is a method that can quantify gene-set activity with a complete probability density function rather than evaluating deviation from a null hypothesis with a p value. Furthermore, we also calculated the activation scores for genes in the enriched pathways to determine the gene activation level. The activation score of a given gene can be calculated according to a t test formalism between two groups (tumor and normal group in our study). For gene i in a certain cell type, we first performed the pseudo-bulk analysis that calculated the mean expression value of the gene i across cells. For gene i, the expression profile are defined as E_i^g, where $g \in G$ are the indexes of the samples that belong to a single group N_G and G is one of tumor group or normal group. Unbiased estimates for the mean and standard deviation s_i^G within each group are then calculated, respectively, by Equations 1 and 2:

$$\overline{E}_i^G = \frac{\sum_{g \in G} E_i^g}{N_G} \quad \text{(Equation 1)}$$

$$s_i^G = \sqrt{\frac{\sum_{g \in G}\left[\left(E_i^g\right)^2 - \left(\overline{E}_i^G\right)^2\right]}{N_G - 1}} \quad \text{(Equation 2)}$$

Then, the activation score for gene i between two groups can be calculated by:

$$activation\_score_i = \frac{\overline{E}_i^{tumor} - \overline{E}_i^{normal}}{\sqrt{\frac{\left(s_i^{tumor}\right)^2}{N_{tumor}} + \frac{\left(s_i^{normal}\right)^2}{N_{normal}}}} \quad \text{(Equation 3)}$$

For a given pathway P, the activation score is quantified by the mean activation score of the genes that compose the pathway (Equation 4). N_P indicates the number of genes in pathway P.

$$activation\_score_P = \frac{\sum_{i \in P} activation\_score_i}{N_P} \quad \text{(Equation 4)}$$

For a given gene/pathway, a higher positive value means higher activation level, while a lower negative value means higher deactivation level.

Identification of potential therapeutic agents for SPARC

In order to investigate the potential therapeutic agents for SPARC, two databases were first used to collect and select potential drugs and chemical compounds, including the Comparative Toxicogenomics Database (CTD; http://ctdbase.org/) and RNAactDrug database (http://bio-bigdata.hrbmu.edu.cn/RNAactDrug/).[56,57] The overlapped drugs and chemicals of the two databases were considered candidate therapeutic agents for SPARC.

Additionally, we performed the drug sensitivity analysis for candidate therapeutic agents first on NSCLC cell lines by using data from the GDSC.[58] In this study, we evaluated the sensitivity of several candidate therapeutic drugs by using the correlation between IC_{50} values and SPARC expression. The positive correlation indicates that the increase in gene expression would lead to drug resistance. Furthermore, in order to investigate the effectiveness of the drugs in clinical treatment, we performed Cox regression analysis and survival analysis on the high-expression SPARC group and the low-expression SPARC group in the TCGA cohort. The standardization of drug names in TCGA was performed by Gene-Drug Interaction for Survival in Cancer (GDISC, https://gdisc.bme.gatech.edu/cgi-bin/gdisc/index.cgi).[59]

Enrichment analysis and statistical analysis

Enrichment analyses for APA event genes and SPARC-interaction genes were performed by using Metascape (https://metascape.org/gp/index.html).[60] The significance level was set as $p < 0.05$ with false discovery rate (FDR) correction. Correlation analysis was performed by Spearman's correlation, and comparison analysis used Student's t tests and ANOVA analysis depending on the situation. For all statistical analyses in this study, a two-tailed $p < 0.05$ was considered significant. All statistical analyses were performed using R script (version 4.0, R Core Team, R Foundation for Statistical Computing, Vienna, Austria).

Experimental methods

Cell culture and reagents

Human embryonic kidney cells (HEK293) and human alveolar adenocarcinoma cell lines A549 were obtained from American Type Culture Collection (ATCC). All the cell lines were cultured in Dulbecco's modified Eagle's medium (DMEM, HyClone) supplemented with 10% fetal bovine serum (FBS, Gibco) and 1% penicillin-streptomycin (HyClone) in a cell incubator under 5% CO_2 at 37°C. All transfections were performed with Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. Cell Counting Kit-8 (CCK-8) was purchased from Dojindo (Japan). Double-luciferase Reporter Assay kit was purchased from Transgen (China). Both miRNA mimics and siRNA were synthesized by Shanghai GenePharma.

Dual luciferase assay

About 10^5/well of HEK293 cells were plated in 24-well cell culture plates and incubated overnight. The cells were then transfected with the corresponding miRNAs using 200 ng of firefly luciferase reporter (containing the 3′ UTR of interest) and 100 ng of Renilla luciferase reporter as control. Detection was performed 24 h after transfection with the double-luciferase reporter assay kit using Glomax-20/20 Luminometer. The firefly luciferase signal was normalized to Renilla luciferase to control transfection efficiency.

Real-time qPCR

Total RNA was extracted from cells using Triquick Reagent (Solarbio, Beijing, China) according to the manufacturer's protocol. One microgram of total RNA was reverse transcribed with the Hifair II first-strand cDNA synthesis kit (YEASEN, Shanghai, China). An equal amount of cDNA was then amplified by real-time PCR using the Hieff qPCR SYBR Green Master Mix (YEASEN, Shanghai, China) and

Applied Biosystems 7300 System (Applied Biosystems, CA, USA). The PCR conditions were as follows: 95°C for 5 min, 95°C for 10 s, and 60°C for 34 s for a total of 40 cycles. Each analysis was performed in quadruplicate. Gene expression was normalized to a housekeeping gene (β-actin) and the relative expression values between the samples were calculated based on the threshold cycle (Ct) value using the 2-ΔΔCT method. Forward and reverse primer sequences for β-actin were 5′-TGCGTGACATCAAAGAGAAG-3′ and 5′-TCCATACCC AAGAAGGAAGG-3′, respectively. Forward and reverse primer sequences for SPARC were 5′-AGGTGTGTGAGCTGCACGAGA-3′ and 5′-GAAGTGGCAGGA AGAGTCGAA-3′, respectively. Forward and reverse primer sequences for TMBIM6 were 5′-AGC AGC ACC TGA AGG TC-3′ and 5′-TCA ATA TCA GGG AGC CCA AG-3′, respectively.

Cell viability assay

A549 cells (5×10^3/well) were seeded in 96-cell culture plates and incubated overnight. At 48 h after transfection with siRNAs, 10 μL of CCK-8 was added to each well and incubated for 4 h to measure the absorbance at 450 nm by a microplate reader. The cell viability under different siRNAs was calculated by the following formula: cell viability = (OD siRNA − OD blank)/(OD control − OD blank) %, where OD is optical density.

Wound scratch assay

A549 cells were seeded in 24-well plates (1×10^5 cells/well) overnight. After transfection with siRNAs, cell layers were scratched using a 10-μL sterile pipette to form a wound. Cells were continuously cultured in a serum-free medium and the wounds were recorded initially and after 24 h.

Transwell migration assay

An 8-μm Transwell chamber (Millipore) was used to study cell migration. Six-hundred microliters of complete medium was added to the lower chamber, and 200 μL of transfected A549 cells were suspended in the upper chamber with serum-free medium. After 24 h, cells in the upper layer that did not migrate were wiped off with a cotton swab, and cells in the lower layer were fixed with 4% methanol and stained with 0.1% crystal violet. The cells were observed and counted in five random areas using a microscope at 100× magnification. The average number of migrated cells was used to evaluate the capability of migration.

Experimental statistical analysis

All experiments were performed at least three repeat independent experiments, and all data were processed with GraphPad Prism 9.0 and expressed as means ± SD. Differences between data were analyzed by t test, and $p < 0.05$ was considered statistically significant.

Conclusions

In conclusion, our computational analysis and experimental validation enhanced the understanding of cell-type-specific APA-mediated 3′-UTR regulation in NSCLC. The comprehensive analysis found several oncogenic factors relating to tumorigenesis and progression that were previously uncharacterized with APA regulation. Moreover, we revealed APA-induced miRNA loss plays vital roles in the oncogene-miRNA-suppressor axis. The results of our in-depth analysis of cell-specific APA events provide interpretations that explain how APA-mediated 3′-UTR alteration can shape cell-specific gene expression changes in NSCLC.

DATA AND CODE AVAILABILITY

Raw scRNA-seq data were downloaded from ArrayExpress database with accession code E-MTAB-6149 and E-MTAB-6653 and GEO database with accession code GSM3926546.

The expression profiles of bulk sequencing datasets of NSCLC were obtained from TCGA (https://portal.gdc.cancer.gov/) and GEO database with accession code GSE30219. Batch effects removal, QC, and cell type annotation were performed with R package Seurat (R version 4.0, https://www.r-project.org/; Seurat version 3.1, https://satijalab.org/seurat/).

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.omtn.2023.08.005.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (grant number no.82227802 to L.H. and no. 32271512 to F.W.) and the Natural Science Basic Research Program of Shaanxi (grant number no. 2022JC-56 and no. 2023-JC-ZD-43 to F.W.). We appreciate the help of Dr. Yubo Wang for assistance in manuscript review prior to its submission.

AUTHOR CONTRIBUTIONS

Conceptualization, K.H., Y.Z., L.H., and X.Z.; methodology, X.S., Z.Y., and K.H.; investigation, K.H., Y.Z., and X.S.; visualization, K.H. and Z.Y.; supervision, X.Z., F.W., and L.H.; writing – original draft, K.H., X.S., and Z.Y.; writing – review & editing, W.Z. and Y.Z.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. National Cancer Institute (2020). SEER Cancer Statistics Review 1975-2017.
2. Botling, J., Edlund, K., Lohr, M., Hellwig, B., Holmberg, L., Lambe, M., Berglund, A., Ekman, S., Bergqvist, M., Pontén, F., et al. (2013). Biomarker Discovery in Non–Small Cell Lung Cancer: Integrating Gene Expression Profiling, Meta-analysis, and Tissue Microarray ValidationGene Expression–Based Biomarker Discovery in NSCLC. Clin. Cancer Res. *19*, 194–204.
3. Awad, M.M., Oxnard, G.R., Jackman, D.M., Savukoski, D.O., Hall, D., Shivdasani, P., Heng, J.C., Dahlberg, S.E., Jänne, P.A., Verma, S., et al. (2016). MET exon 14 mutations in non–small-cell lung cancer are associated with advanced age and stage-dependent MET genomic amplification and c-Met overexpression. J. Clin. Oncol. *34*, 721–730.
4. Castellanos, E., Feld, E., and Horn, L. (2017). Driven by mutations: the predictive value of mutation subtype in EGFR-mutated non–small cell lung cancer. J. Thorac. Oncol. *12*, 612–623.

5. O'Leary, K., Shia, A., and Schmid, P. (2018). Epigenetic regulation of EMT in non-small cell lung cancer. Curr. Cancer Drug Targets *18*, 89–96.
6. Di Giammartino, D.C., Nishida, K., and Manley, J.L. (2011). Mechanisms and consequences of alternative polyadenylation. Mol. Cell *43*, 853–866.
7. Mayr, C., and Bartel, D.P. (2009). Widespread shortening of 3′ UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. Cell *138*, 673–684.
8. Masamha, C.P., and Wagner, E.J. (2018). The contribution of alternative polyadenylation to the cancer phenotype. Carcinogenesis *39*, 2–10.
9. Yuan, F., Hankey, W., Wagner, E.J., Li, W., and Wang, Q. (2021). Alternative polyadenylation of mRNA and its role in cancer. Genes Dis. *8*, 61–72.
10. Xiang, Y., Ye, Y., Lou, Y., Yang, Y., Cai, C., Zhang, Z., Mills, T., Chen, N.-Y., Kim, Y., Muge Ozguc, F., et al. (2018). Comprehensive characterization of alternative polyadenylation in human cancer. J. Natl. Cancer Inst. *110*, 379–389.
11. Feng, X., Li, L., Wagner, E.J., and Li, W. (2018). TC3A: the cancer 3′ UTR atlas. Nucleic Acids Res. *46*, D1027–D1030.
12. Wang, L., Lang, G.-T., Xue, M.-Z., Yang, L., Chen, L., Yao, L., Li, X.-G., Wang, P., Hu, X., and Shao, Z.-M. (2020). Dissecting the heterogeneity of the alternative polyadenylation profiles in triple-negative breast cancers. Theranostics *10*, 10531–10547.
13. Masamha, C.P., Xia, Z., Yang, J., Albrecht, T.R., Li, M., Shyu, A.-B., Li, W., and Wagner, E.J. (2014). CFIm25 links alternative polyadenylation to glioblastoma tumour suppression. Nature *510*, 412–416.
14. Zhang, S., Zhang, X., Lei, W., Liang, J., Xu, Y., Liu, H., and Ma, S. (2019). Genome-wide profiling reveals alternative polyadenylation of mRNA in human non-small cell lung cancer. J. Transl. Med. *17*, 257–311.
15. Gruber, A.J., and Zavolan, M. (2019). Alternative cleavage and polyadenylation in health and disease. Nat. Rev. Genet. *20*, 599–614.
16. Kim, N., Chung, W., Eum, H.H., Lee, H.-O., and Park, W.-Y. (2019). Alternative polyadenylation of single cells delineates cell types and serves as a prognostic marker in early stage breast cancer. PLoS One *14*, e0217196.
17. Göpferich, M., George, N.O., Muelas, A.D., Bizyn, A., Pascual, R., Fijalkowska, D., Kalamakis, G., Müller, U., Krijgsveld, J., and Mendez, R. (2020). Single cell 3'UTR analysis identifies changes in alternative polyadenylation throughout neuronal differentiation and in autism. Preprint at bioRxiv. https://doi.org/10.1101/2020.08.12.247627.
18. Xia, Z., Donehower, L.A., Cooper, T.A., Neilson, J.R., Wheeler, D.A., Wagner, E.J., and Li, W. (2014). Dynamic analyses of alternative polyadenylation from RNA-seq reveal a 3′-UTR landscape across seven tumour types. Nat. Commun. *5*, 5274.
19. Shi, L., Wang, L., Hou, J., Zhu, B., Min, Z., Zhang, M., Song, D., Cheng, Y., and Wang, X. (2015). Targeting roles of inflammatory microenvironment in lung cancer and metastasis. Cancer Metastasis Rev. *34*, 319–331.
20. Podhajcer, O.L., Benedetti, L.G., Girotti, M.R., Prada, F., Salvatierra, E., and Llera, A.S. (2008). The role of the matricellular protein SPARC in the dynamic interaction between the tumor and the host. Cancer Metastasis Rev. *27*, 691–705.
21. Arnold, S.A., and Brekken, R.A. (2009). SPARC: a matricellular regulator of tumorigenesis. J. Cell Commun. Signal. *3*, 255–273.
22. Venkat, S., Tisdale, A.A., Schwarz, J.R., Alahmari, A.A., Maurer, H.C., Olive, K.P., Eng, K.H., and Feigin, M.E. (2020). Alternative polyadenylation drives oncogenic gene expression in pancreatic ductal adenocarcinoma. Genome Res. *30*, 347–360.
23. Yang, S.W., Li, L., Connelly, J.P., Porter, S.N., Kodali, K., Gan, H., Park, J.M., Tacer, K.F., Tillman, H., Peng, J., et al. (2020). A cancer-specific ubiquitin ligase drives mRNA alternative polyadenylation by ubiquitinating the mRNA 3′ end processing complex. Mol. Cell *77*, 1206–1221.e7.
24. Zhang, J., Zhao, X., Luan, Z., and Wang, A. (2020). Rab14 overexpression promotes proliferation and invasion through YAP signaling in non-small cell lung cancers. OncoTargets Ther. *13*, 9269–9280.
25. Wang, J., Sun, D., Wang, Y., Ren, F., Pang, S., Wang, D., and Xu, S. (2014). FOSL2 positively regulates TGF-β1 signalling in non-small cell lung cancer. PLoS One *9*, e112150.
26. Junjappa, R.P., Kim, H.-K., Park, S.Y., Bhattarai, K.R., Kim, K.-W., Soh, J.-W., Kim, H.-R., and Chae, H.-J. (2019). Expression of TMBIM6 in cancers: the involvement of Sp1 and PKC. Cancers *11*, 974.
27. Wang, X., Yin, H., Zhang, H., Hu, J., Lu, H., Li, C., Cao, M., Yan, S., and Cai, L. (2018). NF-κB-driven improvement of EHD1 contributes to erlotinib resistance in EGFR-mutant lung cancers. Cell Death Dis. *9*, 418.
28. Zhou, K., Liu, M., and Cao, Y. (2017). New insight into microRNA functions in cancer: oncogene–microRNA–tumor suppressor gene network. Front. Mol. Biosci. *4*, 46.
29. Wang, Z., Zhao, Z., Yang, Y., Luo, M., Zhang, M., Wang, X., Liu, L., Hou, N., Guo, Q., Song, T., et al. (2018). MiR-99b-5p and miR-203a-3p function as tumor suppressors by targeting IGF-1R in gastric cancer. Sci. Rep. *8*, 10119.
30. Speth, J.M., Penke, L.R., Bazzill, J.D., Park, K.S., de Rubio, R.G., Schneider, D.J., Ouchi, H., Moon, J.J., Keshamouni, V.G., Zemans, R.L., et al. (2019). Alveolar macrophage secretion of vesicular SOCS3 represents a platform for lung cancer therapeutics. JCI insight *4*, e131340.
31. Suzuki, M., Hao, C., Takahashi, T., Shigematsu, H., Shivapurkar, N., Sathyanarayana, U.G., Iizasa, T., Fujisawa, T., Hiroshima, K., and Gazdar, A.F. (2005). Aberrant methylation of SPARC in human lung cancers. Br. J. Cancer *92*, 942–948.
32. Kehlet, S.N., Manon-Jensen, T., Sun, S., Brix, S., Leeming, D.J., Karsdal, M.A., and Willumsen, N. (2018). A fragment of SPARC reflecting increased collagen affinity shows pathological relevance in lung cancer–implications of a new collagen chaperone function of SPARC. Cancer Biol. Ther. *19*, 904–912.
33. Mueller, M.M., and Fusenig, N.E. (2011). Tumor-associated Fibroblasts and Their Matrix (Springer Science & Business Media).
34. Francki, A., McClure, T.D., Brekken, R.A., Motamed, K., Murri, C., Wang, T., and Sage, E.H. (2004). SPARC regulates TGF-beta1-dependent signaling in primary glomerular mesangial cells. J. Cell. Biochem. *91*, 915–925.
35. Heldin, C.-H., Vanlandewijck, M., and Moustakas, A. (2012). Regulation of EMT by TGFβ in cancer. FEBS Lett. *586*, 1959–1970.
36. Tichet, M., Prod'Homme, V., Fenouille, N., Ambrosetti, D., Mallavialle, A., Cerezo, M., Ohanna, M., Audebert, S., Rocchi, S., Giacchero, D., et al. (2015). Tumour-derived SPARC drives vascular permeability and extravasation through endothelial VCAM1 signalling to promote metastasis. Nat. Commun. *6*, 6993.
37. Fennell, D.A., Summers, Y., Cadranel, J., Benepal, T., Christoph, D.C., Lal, R., Das, M., Maxwell, F., Visseren-Grul, C., and Ferry, D. (2016). Cisplatin in the modern era: The backbone of first-line chemotherapy for non-small cell lung cancer. Cancer Treat Rev. *44*, 42–50.
38. Douillard, J.-Y., Rosell, R., De Lena, M., Carpagnano, F., Ramlau, R., Gonzáles-Larriba, J.L., Grodzki, T., Pereira, J.R., Le Groumellec, A., Lorusso, V., et al. (2006). Adjuvant vinorelbine plus cisplatin versus observation in patients with completely resected stage IB–IIIA non-small-cell lung cancer (Adjuvant Navelbine International Trialist Association [ANITA]): a randomised controlled trial. Lancet Oncol. *7*, 719–727.
39. Pignon, J.-P., Tribodet, H., Scagliotti, G.V., Douillard, J.-Y., Shepherd, F.A., Stephens, R.J., Dunant, A., Torri, V., Rosell, R., and Seymour, L. (2008). Lung adjuvant cisplatin evaluation: a pooled analysis by the LACE Collaborative Group. J. Clin. Oncol. *26*, 3552–3559. Database of Abstracts of Reviews of Effects (DARE): Quality-Assessed Reviews [Internet]. Centre for Reviews and Dissemination (UK). https://doi.org/10.1200/JCO.2007.13.9030.
40. Santabarbara, G., Maione, P., Rossi, A., and Gridelli, C. (2016). Pharmacotherapeutic options for treating adverse effects of Cisplatin chemotherapy. Expet Opin. Pharmacother. *17*, 561–570.
41. Galluzzi, L., Vitale, I., Michels, J., Brenner, C., Szabadkai, G., Harel-Bellan, A., Castedo, M., and Kroemer, G. (2014). Systems biology of cisplatin resistance: past, present and future. Cell Death Dis. *5*, e1257.
42. Jia, Q., Xie, B., Zhao, Z., Huang, L., Wei, G., and Ni, T. (2022). Lung cancer cells expressing a shortened CDK16 3′ UTR escape senescence through impaired miR-485-5p targeting. Mol. Oncol. *16*, 1347–1364.
43. Weng, T., Huang, J., Wagner, E.J., Ko, J., Wu, M., Wareing, N.E., Xiang, Y., Chen, N.-Y., Ji, P., Molina, J.G., et al. (2020). Downregulation of CFIm25 amplifies dermal fibrosis through alternative polyadenylation. J. Exp. Med. *217*, e20181384.
44. Lambrechts, D., Wauters, E., Boeckx, B., Aibar, S., Nittner, D., Burton, O., Bassez, A., Decaluwé, H., Pircher, A., Van den Eynde, K., et al. (2018). Phenotype molding of stromal cells in the lung tumor microenvironment. Nat. Med. *24*, 1277–1289.

45. Rousseaux, S., Debernardi, A., Jacquiau, B., Vitte, A.-L., Vesin, A., Nagy-Mignotte, H., Moro-Sibilot, D., Brichon, P.-Y., Lantuejoul, S., Hainaut, P., et al. (2013). Ectopic activation of germline and placental genes identifies aggressive metastasis-prone lung cancers. Sci. Transl. Med. *5*, 186ra66.

46. Shulman, E.D., and Elkon, R. (2019). Cell-type-specific analysis of alternative polyadenylation using single-cell transcriptomics data. Nucleic Acids Res. *47*, 10027–10039.

47. Luo, Y., Zeng, G., and Wu, S. (2019). Identification of microenvironment-related prognostic genes in bladder cancer based on gene expression profile. Front. Genet. *10*, 1187.

48. Charrad, M., Ghazzali, N., Boiteau, V., and Niknafs, A. (2014). NbClust: an R package for determining the relevant number of clusters in a data set. J. Stat. Software *61*, 1–36.

49. Gustriansyah, R., Suhandi, N., and Antony, F. (2020). Clustering optimization in RFM analysis based on k-means. Indonesian Journal of Electrical Engineering and Computer Science *18*, 470–477.

50. Lever, J., Zhao, E.Y., Grewal, J., Jones, M.R., and Jones, S.J.M. (2019). CancerMine: a literature-mined resource for drivers, oncogenes and tumor suppressors in cancer. Nat. Methods *16*, 505–507.

51. Nagy, C., Maitra, M., Tanti, A., Suderman, M., Théroux, J.F., Davoli, M.A., Perlman, K., Yerko, V., Wang, Y.C., Tripathy, S.J., et al. (2020). Single-nucleus transcriptomics of the prefrontal cortex in major depressive disorder implicates oligodendrocyte precursor cells and excitatory neurons. Nat. Neurosci. *23*, 771–781.

52. Razick, S., Magklaras, G., and Donaldson, I.M. (2008). iRefIndex: a consolidated protein interaction database with provenance. BMC Bioinf. *9*, 405–419.

53. Cerami, E.G., Gross, B.E., Demir, E., Rodchenkov, I., Ö, B., Anwar, N., Schultz, N., Bader, G.D., and Sander, C. (2011). Pathway Commons, a web resource for biological pathway data. Nucleic Acids Res. *39*, D685–D690.

54. Stark, C., Breitkreutz, B.-J., Reguly, T., Boucher, L., Breitkreutz, A., and Tyers, M. (2006). BioGRID: a general repository for interaction datasets. Nucleic Acids Res. *34*, D535–D539.

55. Yaari, G., Bolen, C.R., Thakar, J., and Kleinstein, S.H. (2013). Quantitative set analysis for gene expression: a method to quantify gene set differential expression including gene-gene correlations. Nucleic Acids Res. *41*, e170.

56. Davis, A.P., Grondin, C.J., Johnson, R.J., Sciaky, D., Wiegers, J., Wiegers, T.C., and Mattingly, C.J. (2021). Comparative toxicogenomics database (CTD): update 2021. Nucleic Acids Res. *49*, D1138–D1143.

57. Dong, Q., Li, F., Xu, Y., Xiao, J., Xu, Y., Shang, D., Zhang, C., Yang, H., Tian, Z., Mi, K., et al. (2020). RNAactDrug: a comprehensive database of RNAs associated with drug sensitivity from multi-omics data. Briefings Bioinf. *21*, 2167–2174.

58. 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. (2013). Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Res. *41*, D955–D961.

59. Spainhour, J.C.G., Lim, J., and Qiu, P. (2017). GDISC: a web portal for integrative analysis of gene–drug interaction for survival in cancer. Bioinformatics *33*, 1426–1428.

60. Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A.H., Tanaseichuk, O., Benner, C., and Chanda, S.K. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat. Commun. *10*, 1523.