RNA sequencing reveals the transcriptomic landscape and alternative splicing events induced by LGALS1 silencing in non-small cell lung cancer

Hongyang Shi[D], Yonghong Zhang[A], Juan Liu[B], Yu Wang[C], Ping Fang[E,F]

Department of Respiratory and Critical Care Medicine, The Second Affiliated Hospital of Xi'an Jiaotong University, China

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation;
D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online) *Adv Clin Exp Med*. 2024;33(1):79–90

Address for correspondence
Hongyang Shi
E-mail: shihy2003@126.com

Funding sources
This research was supported by Natural Science Basic Research Plan of Shanxi Province of China (grant No. 2020JM-410) and hospital fund of the Second Affiliated Hospital of Xi'an Jiaotong University (grant No. RC(GG)201809).

Conflict of interest
None declared

Received on June 17, 2022
Reviewed on February 8, 2023
Accepted on April 21, 2023

Published online on June 21, 2023

Abstract

Background. Non-small cell lung cancer (NSCLC) is a common clinical cancer with high mortality. The lectin galactoside-binding soluble 1 (LGALS1) is an RNA-binding protein (RBP) involved in NSCLC progression. Alternative splicing (AS) is a vital function of RBPs that contributes to tumor progression. It is unknown whether LGALS1 regulates NSCLC progression through AS events.

Objectives. To profile the transcriptomic landscape and LGALS1-regulated AS events in NSCLC.

Materials and methods. The A549 cells either with silenced LGALS1 (siLGALS1 group) or without them (siCtrl group) were subjected to RNA sequencing; differentially expressed genes (DEGs) and AS events were discovered and then the AS ratio was validated using reverse transcription-quantitative polymerase chain reaction (RT-qPCR).

Results. High LGALS1 expression indicates poor overall survival (OS), first progression (FP) and post-progression survival (PPS). A total of 225 DEGs were identified, including 81 downregulated and 144 up-regulated in the siLGALS1 group compared to the siCtrl group. Differentially expressed genes were mainly enriched in interaction-related Gene Ontology (GO) terms and involved in cGMP-protein kinase G (PKG) and calcium signaling pathways. The RT-qPCR validation showed that the expressions of *ELMO1* and *KCNJ2* were upregulated, while *HSPA6* was downregulated after LGALS1 silencing. The expressions of *KCNJ2* and *ELMO1* were upregulated to a peak at 48 h after LGALS1 knockdown, while *HSPA6* expression decreased, after which their expressions returned to baseline. The overexpression of LGALS1 rescued the elevation in *KCNJ2* and *ELMO1* expression, and decrease in *HSPA6* expression induced by siLGALS1. A total of 69,385 LGALS1-related AS events were detected, which produced 433 upregulated and 481 downregulated AS events after LGALS1 silencing. The LGALS1-related AS genes were mainly enriched in the apoptosis and ErbB signaling pathways. The LGALS1 silencing led to a decrease in the AS ratio of BCAP29 and an increase in CSNKIE and MDFIC.

Conclusions. We characterized the transcriptomic landscape and profiled AS events in A549 cells following LGALS1 silencing. Our study provides abundant candidate markers and new insights into NSCLC.

Key words: NSCLC, alternative splicing, RNA binding proteins, LGALS1 silencing, transcriptome landscape

Cite as
Shi H, Zhang Y, Liu J, Wang Y, Fang P. RNA sequencing reveals the transcriptomic landscape and alternative splicing events induced by LGALS1 silencing in non-small cell lung cancer. *Adv Clin Exp Med*. 2024;33(1):79–90. doi:10.17219/acem/163519

DOI
10.17219/acem/163519

Copyright
Copyright by Author(s)
This is an article distributed under the terms of the Creative Commons Attribution 3.0 Unported (CC BY 3.0) (https://creativecommons.org/licenses/by/3.0/)

Background

Lung cancer is a frequent malignant disease with high mortality,[1] and non-small cell lung cancer (NSCLC) accounts for the majority of lung cancer cases. This disease can be classified as lung adenocarcinoma (LUAD) or lung squamous cell carcinoma (LUSC).[2] Despite considerable progress in early diagnosis, there is still a high recurrence rate after treatment,[3] and the pathogenesis of NSCLC remains largely unknown. Thus, it is urgent to identify new driver genes and elucidate their mechanism to provide clinical advancements.

RNA-binding proteins (RBPs) play an important role in translation and gene regulation,[4,5] driving the development of many diseases and cancers.[6] The RBP lectin galactoside-binding soluble 1 (LGALS1) is a β-galactoside-binding protein that has a carbohydrate recognition domain,[7] and an extensive body of research has revealed that LGALS1 participates in the progression of multiple tumors.[8,9] For example, the upregulated expression of LGALS1 promotes the NSCLC progression by interacting with non-SMC condensin I complex subunit G (NCAPG).[10] Nevertheless, the mechanism and manner by which LGALS1 participates in the NSCLC progression are largely unknown.

Alternative splicing (AS) is a vital function of RBPs. It is reported that more than 90% of genes undergo AS to regulate gene expression, which ultimately results in proteome diversity,[11] but can also increase cancer risk.[12] The LGALS1 participates in pre-mRNA splicing, and it has been reported to regulate gene expression.[13] The LGALS1 was shown to reduce CaV1.2 calcium channel currents resulting in the regulation of vascular constriction via AS.[14] However, whether LGALS1 AS may regulate the NSCLC pathological processes remains unclear.

Objectives

We aimed to evaluate whether LGALS1 exerts its regulatory effects through the regulation of gene expression and AS to discover new prognostic and therapeutic targets for NSCLC. For this purpose, we knocked down LGALS1 in A549 cells and analyzed the differentially expressed genes (DEGs) and LGALS1-regulated AS events associated with NSCLC.

Materials and methods

Public databases

The Kaplan–Meier plotter (http://kmplot.com/analysis/index.php?p=service&cancer=lung) was used to analyze the relationship between LGALS1 expression and survival, including overall survival (OS), first progression (FP) and post-progression survival (PPS), based on the set of NSCLC samples. Data were processed based on the study by Győrffy et al.[15] The Cancer Genome Atlas (TCGA, http://cancergenome.nih.gov) was used to analyze the expression pattern of LGALS1 in NSCLC.

Cell culture and transient transfection

Human A549 (#CL-0016) and H1650 (#CL-0166) cell lines were purchased from China Procell Life Science & Technology (Wuhan, China). Normal human lung epithelial BEAS-2B cells and human H1299 cells were provided by our laboratory. Cells were cultured in a minimum essential medium (OPTI-MEM) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin, and maintained at 37°C with 5% CO_2.

All small interfering RNA (siRNA) duplexes were designed and synthesized in Suzhou Jima Gene Co. Ltd (Suzhou, China). The LGALS1 siRNA sequences (siLGALS1-1) and negative control siRNA sequences (siCtrl) are listed in Supplementary Table 1.

Silencing LGALS1 lentivirus vectors (shLGALS1) were used based on pSLenti-U6-shRNA-CMV-EGFP-2A-Puro, and blank vectors served as the control group (shCtrl).

Transient transfection of siRNA sequences and shRNA vectors was performed using Lipofectamine™ 2000 (Invitrogen, Waltham, USA), according to the manufacturer's instructions. Briefly, 1 day before transfection, A549 cells were seeded in 6-well plates (Corning, Tewksbury, USA) at a density of 5×10^5 cells per well. The siRNA sequences were mixed with OPTI-MEM at a ratio of 5 µL:45 µL. The transfection reagent Lipofectamine™ 2000 (5 µL) was also diluted with 45 µL of OPTI-MEM. Next, 50 µL of diluted siRNA and 50 µL of diluted transfection reagent were mixed and incubated for 15 min at room temperature. After that, the entire mixture was added into the A549 cell sample to incubate for another 24 h until subsequent assays were performed.

Western blot

Radioimmunoprecipitation assay (RIPA) Lysis and Extraction Buffer (#89900; Thermo Fisher Scientific, Waltham, USA) were used to lyse the A549 cells on ice for 15 min; then the product was centrifuged at 12,000 rpm for 10 min at 4°C, and the supernatant was discarded. The concentration of protein precipitate was measured with the bicinchoninic acid (BCA) method and repeated 3 times on each protein extraction. A total of 20 µg of protein was isolated using sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and transferred onto the polyvinylidene fluoride (PVDF) membrane. The transfer efficiency was checked with Ponceau red. The membrane was blocked at 25°C for 3 h with a Tris-HCl buffered salt solution with Tween-20 (TBST) containing 5%

skimmed milk powder. Next, the membrane was incubated with anti-LGALS1 (ab266850, 1:1000; Abcam, Cambridge, UK) or anti-GAPDH (60004-1-Lg, 1:1000; Proteintech, Rosemont, USA) at 4°C overnight. The next day, Goat Anti-Mouse IgG H&L (HRP) (A0216, 1:1000; Beyotime Biotechnology, Shanghai, China) and Goat Anti-Rabbit IgG H&L (HRP) (A0208, 1:1000; Beyotime Biotechnology) were added and incubated for 2 h, washed with Tris-buffered saline containing 0.1% Triton X-100 (TBST) 3 times, and then imaged using ECL chemiluminescence (Thermo Fisher Scientific, Waltham, USA) on the chemiluminescence instrument (ChemiScope series, model 6000EXP, serial No. 610005-7Q; Qinxiang, Shanghai, China).

After imaging, the expression of LGALS1 relative to GAPDH was calculated using ImageJ software (National Institutes of Health, Bethesda, USA).

RNA sequencing and data quality control

The genomic material was removed from A549 cells, and total RNA was separated using the Trizol reagent (Invitrogen) in both the siLGALS1 group (n = 3) and siCtrl group (n = 3). The RNA was purified using an RNA purification kit (Tiangen, Beijing, China) and the concentration and quality of each RNA sample was determined using SmartSpec Plus (BioRad, Hercules, USA) by calculating the absorbance at 260 nm/280 nm (A260/A280). The RNA sequencing (RNA-seq) libraries were prepared with the KAPA Stranded mRNA-seq Kit for Illumina® Platforms (#KK8401; KAPA Biosystems, Boston, USA) according to the manufacturer's instructions. Briefly, 1 μg of intact total RNA was used for library construction. Poly(A) RNA was captured with magnetic oligo-dT beads, which were then resuspended in 1X Fragment, Prime and Elute Buffer to elute the captured RNA. Next, the RNA was fragmented to the desired size by heating in the presence of Mg^{2+}. Then, first-strand cDNA was synthesized with random oligo-dT primers followed by second-strand cDNA synthesis, and double-stranded cDNA was manufactured with RNA while marking the second-strand with dUTP. Next, dAMP was added to the 3'-end of double-strand cDNA fragments to obtain the 3'-dAMP library fragments, followed by 3'-dTMP adapter ligation, and the adapter-ligated library DNA was amplified using polymerase chain reaction (PCR). Finally, library fragment size distribution was confirmed with electrophoresis, and library concentration was determined with quantitative PCR (qPCR). Then, RNA-Seq libraries were sequenced on the Illumina NovaSeq 6000 platform using paired-end 150 nt reads. Raw reads were filtered using the FASTX-Toolkit (v. 0.0.13) (http://hannonlab.cshl.edu/fastx_toolkit/index.html) to discard low-quality reads. The alignment of clean reads was then performed by tophat2[16] mapped to the GRCh38 human reference genome. Uniquely mapped results were used to obtain the read count of a gene as expression and then normalized using the FPKM algorithm.[17]

DEGs analysis

The edgeR package[18] was used to distinguish DEGs between the siLGALS1 group and the siCtrl group employing criteria of the false discovery rate (FDR) less than 0.05 and fold change ≥2 or ≤0.5. To predict functional categories of siLGALS1-related genes, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis were performed using the KOBAS 2.0 server.[19]

RBP-related AS analysis

The ABLas pipeline was applied for identifying and quantifying the AS events and regulated AS events (RASEs).[20] In short, 10 types of normative AS events were detected in 2 groups based on the splice junction reads, including cassette exons (CE), exon skipping (ES), mutual exclusive exon skipping (MXE), alternative 3' splice site (A3SS), A3SS and ES (A3SS & ES), MXE combined with an alternative polyadenylation site (3pMXE) or an alternative 5' promoter (5pMXE), alternative 5' splice sites (A5SS), A5SS and ES (A5SS & ES), and intron Retained (introR). Next, to determine RBP-related RASEs, the difference in AS events between the siLGALS1 group and the siCtrl group was evaluated using Student's t-test. The difference in AS events with a p-value cutoff corresponding to an FDR cutoff of 5% was considered RBP-related RASE.

RTqPCR verification of DEGs and AS

Primers of DEGs were designed using Primer Premier v. 6.0 (http://www.premierbiosoft.com/primerdesign/) validated with qPCR. For reverse transcription (RT)-qPCR validation (each gene was repeated 9 times), M-MLV Reverse Transcriptase (Vazyme Biotech, Nanjing, China) was used for RNA reverse transcription into cDNA. The SYBR Green PCR Reagents Kit (Yeasen Biotechnology, Shanghai, China) was used for the PCR reaction with 3 repetitions on the StepOne Realtime PCR System. The gene expression level was normalized to GAPDH using the $2^{-\Delta\Delta CT}$ method.[21]

To verify AS events (each gene was repeated 9 times), a boundary-spanning primer was used to detect alternative isoforms for the sequence containing the junction of the alternative exon (according to "model exon" to detect model splicing or "altered exon" to detect altered splicing) and constitutive exon. The sequence containing the constitutive exon was detected with an opposing primer. The primer sequences are listed in Supplementary Table 1.

Statistical analyses

GraphPad Prism v. 9.0 (GraphPad Software, San Diego, USA) was used for statistical analysis. Data were obtained from at least 3 independent samples and indicated

as means ± standard deviation (M ±SD). Considering that the sample size is less than 10, the data are all analyzed using a nonparametric test. The statistical comparison between the siLGALS1 group and the siCtrl group was performed using Mann–Whitney (M–W) test when the total sample >7; otherwise, the t-test was performed (data conformed to the normal distribution and variance homogeneity). The comparison between 3 or 4 groups was determined with Kruskal–Wallis (K–W) test followed by Dunn's multiple comparison test. A p-value <0.05 was considered statistically significant.

Results

High expression of the *LGALS1* gene corresponds to a poor prognosis of NSCLC

To determine whether *LGALS1* expression was associated with NSCLC progression, the Kaplan–Meier plotter was used for drawing OS, FP and PPS curves based on the *LGALS1* mRNA expression, which was stratified into low- and high-expression groups. The results showed that patients in the *LGALS1* high-expression group exhibited a significant decrease in the OS (log-rank p = 6.6e-13), FP (log-rank p = 0.044) and PPS (log-rank p = 0.00026) compared with the *LGALS1* low-expression group (Fig. 1), suggesting that high expression of the *LGALS1* gene correlated with poor prognosis of NSCLC. However, unexpectedly, according to TCGA, *LGALS1* expression was not statistically different in tumor and control tissues (Supplementary Fig. 1).

LGALS1 silencing alters the transcriptomic landscape of A549 cells

First, we performed RT-qPCR to determine LGALS1 expression in 4 NSCLC cell lines and normal cells. The results showed that compared with the BEAS-2B cell, LGALS1 was significantly overexpressed in 3 NSCLC cell lines, with its highest expression observed in A549 cells (K–W test, H = 9.462, p = 0.0067) (Fig. 2A). Therefore, to assess the molecular mechanisms of LGALS1 in NSCLC, we silenced LGALS1 expression in A549 cells using siRNA. The mRNA (M–W test, U = 0, p < 0.0001) and protein (t-test, t = 4.824, degrees of freedom (df) = 4, p = 0.0485) expression of LGALS1 in the siLGALS1 group were significantly lower than in the siCtrl group, suggesting that the knockdown was successful (Fig. 2B,C). Then, A549 cells containing silenced LGALS1 were subjected to transcriptome sequencing. The LGALS1 expression was also significantly reduced in the RNA-seq analysis, supporting the successful LGALS1 knockdown (t-test, p = 7.52e-14) (Fig. 2D). The Q30 and GC content of all 6 samples were larger than 94.9% and 50%, respectively (Supplementary Table 2), indicating that the quality control of transcriptome sequencing was good. Principal component analysis (PCA) showed that samples from the siLGALS1 group and the siCtrl group could be distinguished from each other (Fig. 2E), suggesting that the interference in LGALS1 expression leads to an alteration in transcriptomic profile in A549 cells.

To further investigate the genes regulated by LGALS1, DEGs (|fold change| >2 and FDR < 0.05) induced by LGALS1 silencing were characterized. A total of 225 DEGs were identified, including 144 upregulated and 81 downregulated DEGs in the siLGALS1 group compared to the siCtrl group

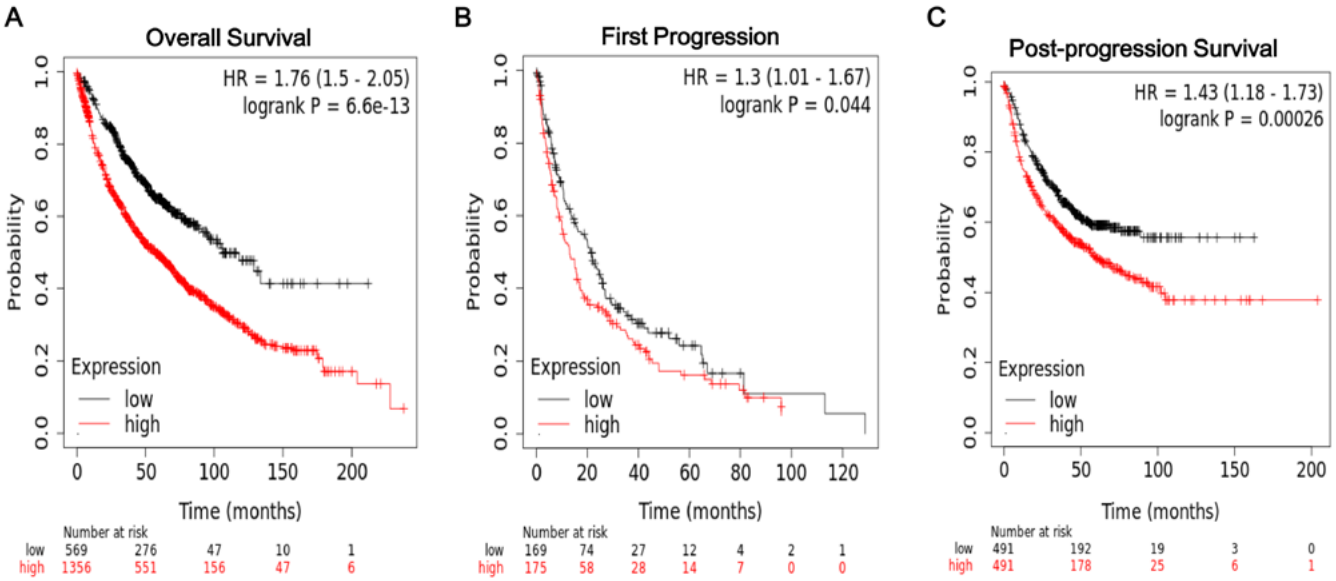

Fig. 1. High expression of lectin galactoside-binding soluble 1 (LGALS1) indicates poor prognosis in non-small cell lung cancer (NSCLC). The pattern of overall survival (OS) (A), first progression (FP) (B) and post-progression survival (PPS) (C) of NSCLC patients with high and low LGALS1 expression levels

HR – hazard ratio.

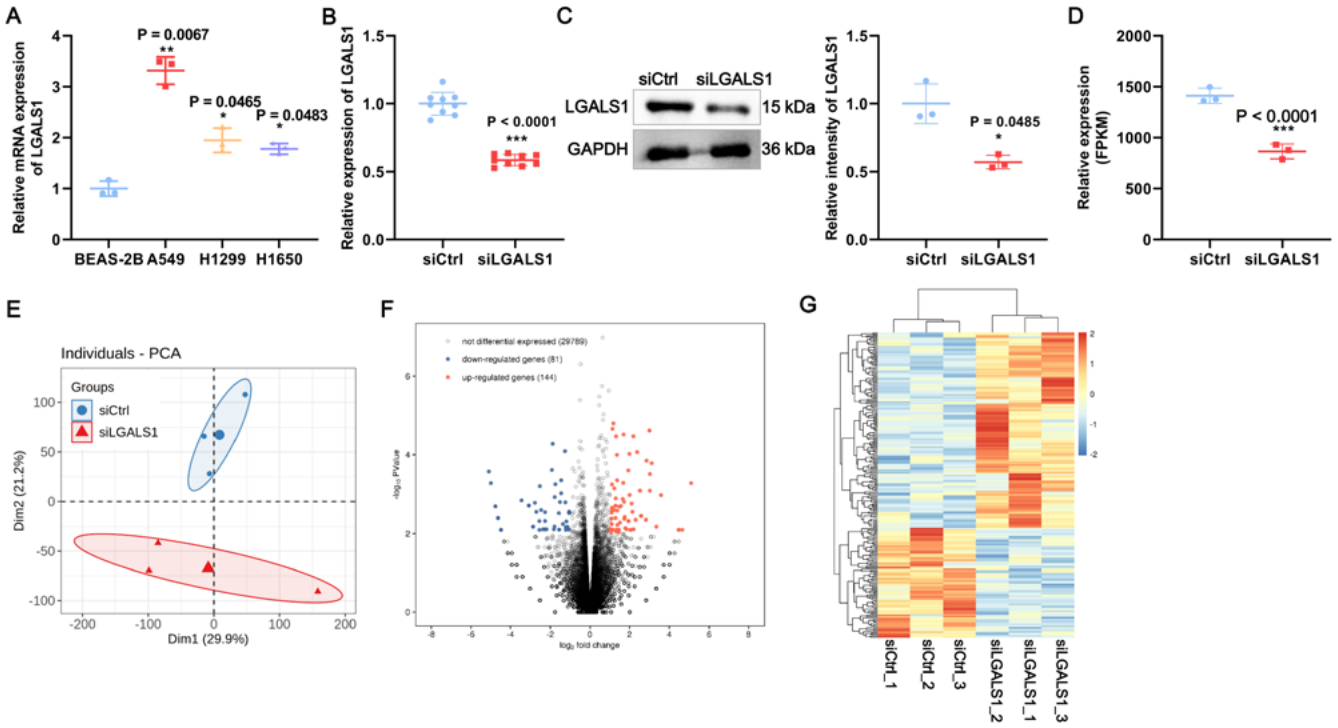

Fig. 2. Lectin galactoside-binding soluble 1 (LGALS1) silencing alters the transcriptomic landscape of A549 cells. A. LGALS1 expression in different cell lines. The silencing efficacy of LGALS1 was quantified using reverse transcription-quantitative polymerase chain reaction (RT-qPCR) (B) and western blot (C) in A549 cells; D. LGALS1 expression in RNA sequencing (RNA-seq) results exhibited by FPKM; E. Principal component analysis (PCA) plot of RNA-seq data for 3 biological replicates in the silenced LGALS1 (siLGALS1) and siCtrl groups; F. Differentially expressed genes (DEGs) when comparing siLGALS1 and siCtrl groups in A549 cells. The upregulated DEGs are marked red, and the downregulated DEGs are marked blue (siLGALS1 compared to siCtrl); G. Heatmap of DEGs

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

(Fig. 2F). After hierarchical clustering of all the DEGs, a robust distinction was observed between the 2 groups, and a high coincidence among the changes was also observed in the 3 replicates (Fig. 2G). These results suggest that LGALS1 silencing alters the transcriptomic landscape of A549 cells.

Functional clustering analysis of DEGs

Next, we used the KOBAS 2.0 server to predict the molecular functions and metabolic pathways of up- and downregulated DEGs, respectively. The results showed that upregulated DEGs only affected small molecule metabolic processes and DNA transcription, and were mainly involved in cell adhesion molecules and chemical carcinogenesis pathways (Fig. 3A,B). In addition, downregulated DEGs were enriched to interaction-related GO terms such as extracellular matrix, steroid binding, laminin binding, and ankyrin binding (Fig. 3C). The KEGG analysis found that downregulated DEGs were mainly associated with protein digestion and absorption, cGMP-protein kinase G (PKG) signaling pathway, and calcium signaling pathways (Fig. 3D). These results imply that LGALS1 silencing alters numerous physiological processes in A549 cells via DEGs.

RT-qPCR validation of LGALS1-regulated DEGs

To further verify the reliability of RNA-seq results, we selected 5 DEGs which were highly abundant during RT-qPCR validation. Importantly, these 5 genes were reported to be connected to the progression of NSCLC, including *ELMO1*,[22] *KCNJ2*,[23] *IGFBP5*,[24] *HSPA6*,[25] and *TINAGL1*.[26] The RT-qPCR results of 5 DEGs were in concordance with the RNA sequencing results (Fig. 4), with the expression of *KCNJ2* (M–W test, U = 1, p = 0.0002) and *ELMO1* (M–W test, U = 9, p = 0.0289) being upregulated after LGALS1 silencing, whereas *HSPA6* (M–W test, U = 9, p = 0.0465) was downregulated. However, the expression of TINAGL1 (M–W test, U = 33, p = 0.8148) and IGFBP5 (M–W test, U = 8, p = 0.1320) exhibited no significant difference between the siLGALS1 group and siCtrl group with RT-qPCR. These results imply that *KCNJ2*, *HSPA6* and *ELMO1* may be functional genes located downstream from LGALS1 signaling.

Given that gene regulation is time-dependent, we examined the expression of DEGs (*KCNJ2*, *HSPA6* and *ELMO1*) at different timepoints after LGALS1 knockdown to further validate the regulation of DEGs by LGALS1. Gene expression in the shCtrl group did not change with time,

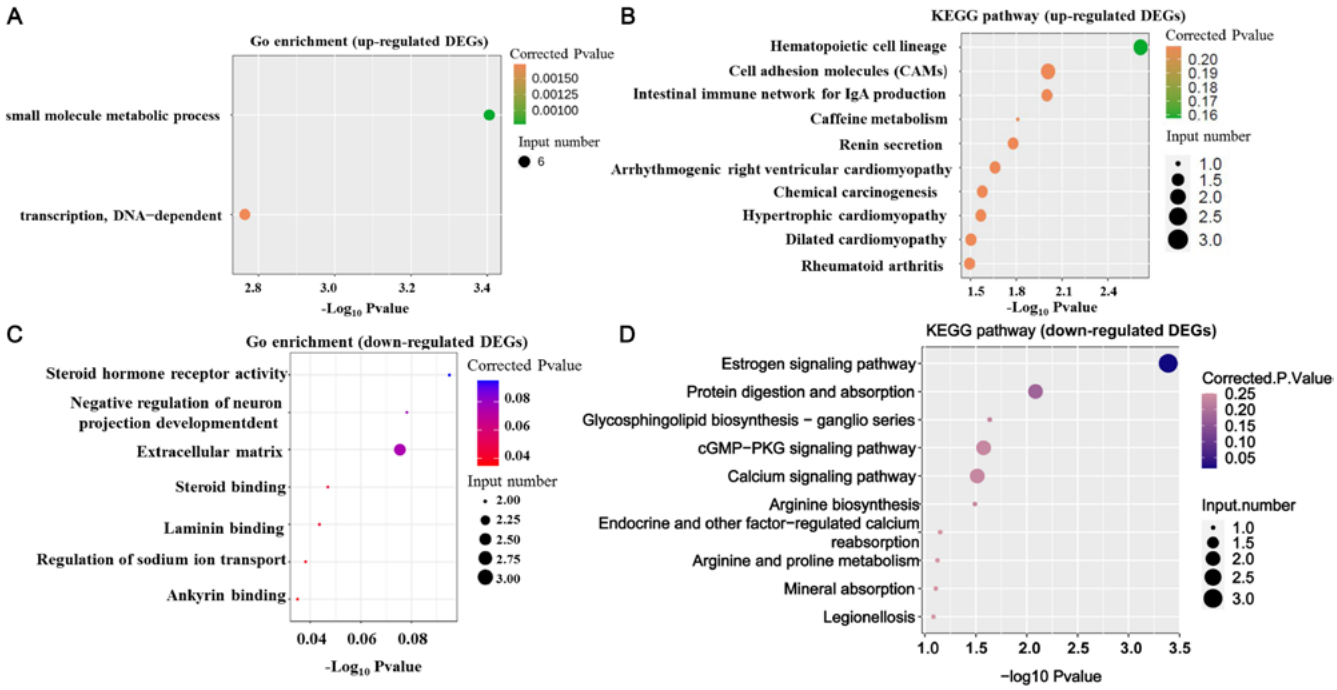

Fig. 3. Functional clustering analysis of differentially expressed genes (DEGs). A. The only 2 Gene Ontology (GO) enrichment terms of upregulated DEGs between silenced lectin galactoside-binding soluble 1 (siLGALS1) and siCtrl group; B. The top 10 enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways for upregulated DEGs between siLGALS1 and siCtrl group. The top 7 GO enrichment terms (C) and top 10 KEGG pathways (D) for downregulated DEGs between siLGALS1 and siCtrl group

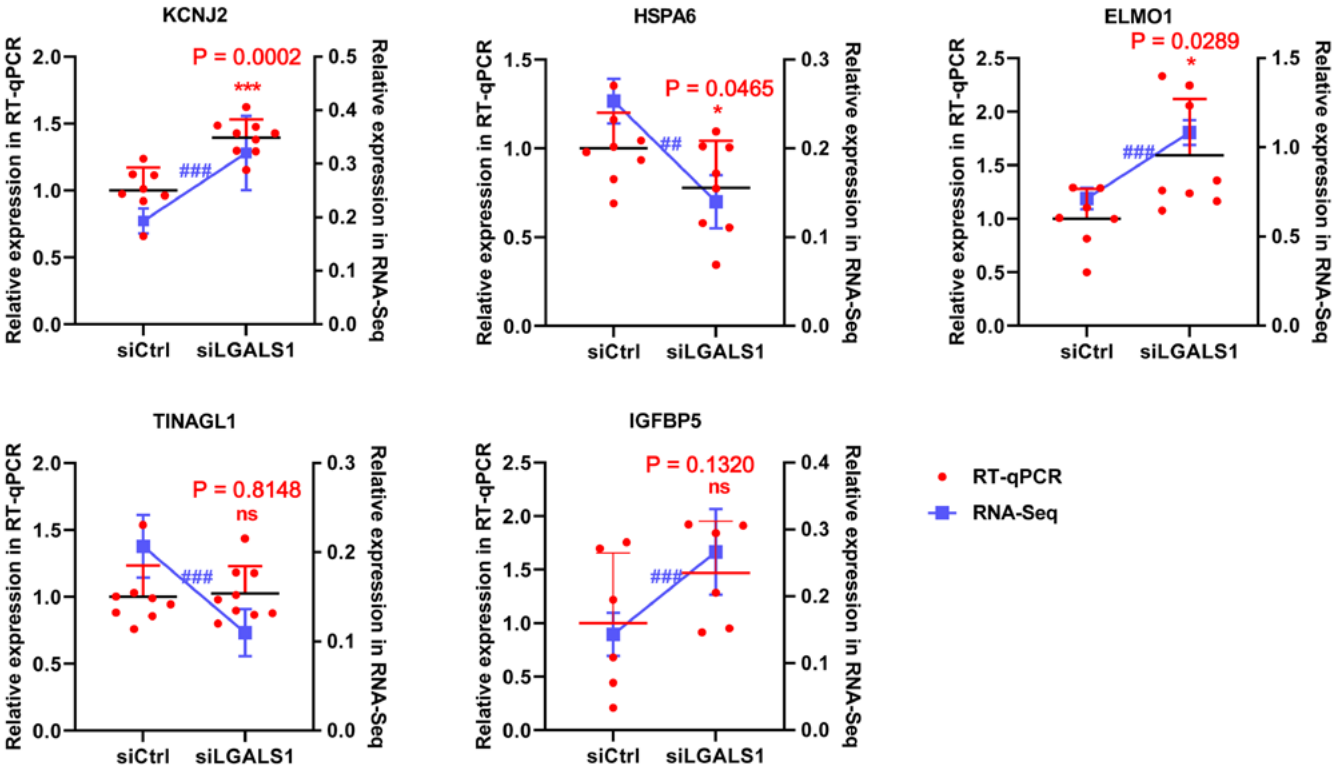

Fig. 4. Reverse transcription-quantitative polymerase chain reaction (RT-qPCR) validation of differentially expressed genes (DEGs). The RNA sequencing (RNA-seq) data are presented as the blue line charts. The RT-qPCR validation results for 5 DEGs are presented in red scatter histogram

$p < 0.001$; ### $p < 0.0001$; ns $p > 0.05$; ** $p < 0.01$; *** $p < 0.001$.

while LGALS1 mRNA expression gradually decreased in the shLGALS1 group compared with the shCtrl group, reaching a minimum at 48 h, after which the transient knockdown effect gradually disappeared (Fig. 5A). As expected, shLGALS1 regulation of DEG expression showed temporal effectiveness for 48 h. The *KCNJ2* and *ELMO1*

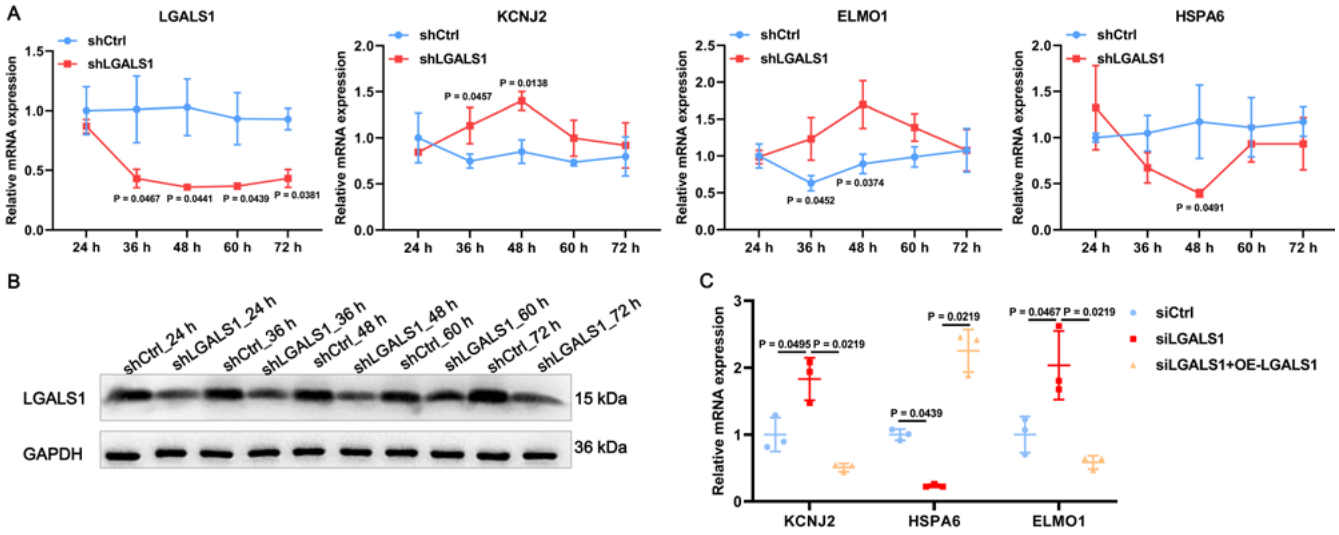

Fig. 5. Differentially expressed genes (DEGs) regulated by lectin galactoside-binding soluble 1 (LGALS1). A. Reverse transcription-quantitative polymerase chain reaction (RT-qPCR) was performed to detect mRNA expression of *LGALS1*, *KCNJ2*, *HSPA6*, and *ELMO1* at different timepoints following LGALS1 knockdown using shRNA; B. Western blot was performed to detect the protein expression of LGALS1 at different timepoints after LGALS1 knockdown using shRNA; C. A rescue experiment was used to verify the regulatory role of LGALS1 on *KCNJ2*, *HSPA6* and *ELMO1* mRNA expression

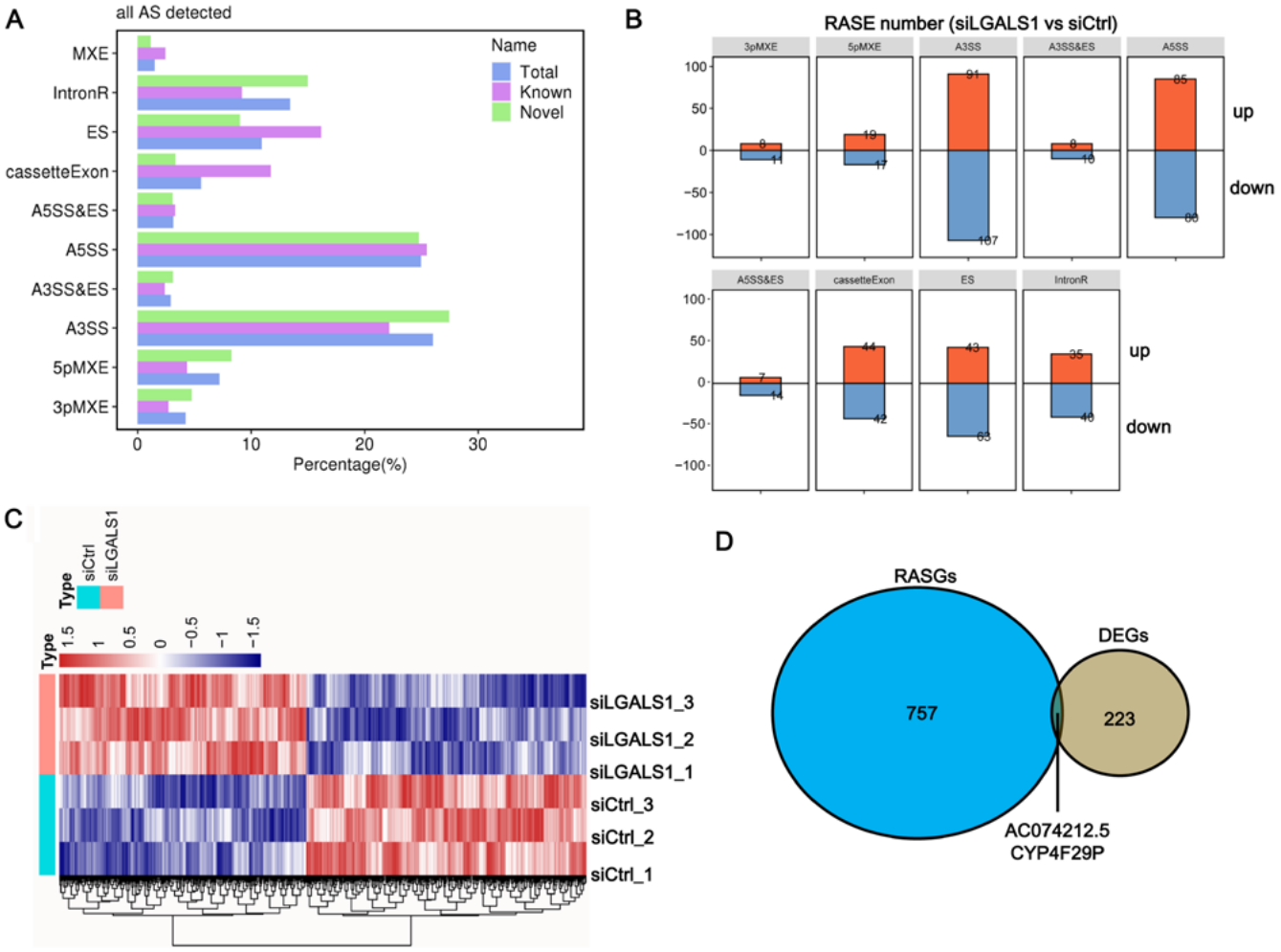

Fig. 6. Identification of lectin galactoside-binding soluble 1 (LGALS1)-mediated alternative splicing (AS) events in A549 cells. A. Classification and proportion of 10 types of AS events detected using RNA sequencing (RNA-seq); B. Classification and numbers of the significant differential regulated AS events (RASEs) between the silenced LGALS1 (siLGALS1) and siCtrl groups; C. Heatmap representation of the pattern of RASEs between the siLGALS1 and siCtrl groups; D. Venn diagram showing the overlap of LGALS1-regulated differentially expressed genes (DEGs) and regulated AS genes (RASGs)

ES – exon skipping; MXE – mutual exclusive exon skipping; A3SS – alternative 3' splice site; 3pMXE – MXE combined with an alternative polyadenylation site; 5pMXE – alternative 5' promoter; A5SS – alternative 5' splice sites; introR – intron Retained.

expressions were upregulated to a peak at 48 h after LGALS1 knockdown, while *HSPA6* expression decreased to a peak, after which their expressions all returned to nondifferential levels from the shCtrl group (Fig. 5A). The shRNA-LGALS1 lentiviral vector also inhibited LGALS1 expression at the protein level (Fig. 5B). In addition, for further validation of the genes regulated by LGALS1, a rescue experiment was conducted. We observed that overexpression of LGALS1 significantly reduced the elevation in *KCNJ2* and *ELMO1* expression and minimized the decrease in *HSPA6* expression induced by siLGALS1 alone (Fig. 5C). Taken together, these results implicate that LGALS1 can regulate the expressions of DEGs, such as *KCNJ2*, *HSPA6* and *ELMO1*.

Identification of LGALS1-mediated AS events in A549 cells

Previous studies have demonstrated that AS is involved in lung cancer development.[27] Thus, we further explored LGALS1-mediated AS events. A total of 69,385 AS events were identified, of which 18,587 were known AS, accounting for 26.79%, and 50,798 were novel AS, accounting for 73.21% (Fig. 6A, Supplementary Table 3). We found that both known and novel AS events contained 10 types (Fig. 6A, Supplementary Tables 4 and 5), of which A3SS and A5SS exhibited the largest share of the total AS events, corresponding to 26.03% and 24.95%, respectively (Fig. 6A, Supplementary Table 6).

Moreover, a total of 914 RASEs mediated by LGALS1 were identified, including 433 upregulated RASEs and 481 downregulated RASEs in the siLGALS1 group compared with the siCtrl group (Fig. 6B). These LGALS1-mediated RASEs were annotated into 9 types of AS events (Fig. 6B). Furthermore, we observed that these RASEs clustered into 2 branches, suggesting that LGALS1 silencing resulted in altered AS in A549 cells (Fig. 6C). Furthermore, LGALS1 silencing led to 914 RASEs residing in 759 genes. We overlapped genes in regulated AS genes (RASGs) and DEGs, finding that only 2 RASGs overlapped with DEG, namely AC074212.5 and CYP4F29P (Fig. 6D).

Functional enrichment analysis of RASGs

To explore the biological function of LGALS1-regulated AS, RASGs were subjected to GO and KEGG enrichment analyses. The GO analysis revealed that RASGs were preferentially enriched in molecular functions related to transcriptional regulation, such as DNA-dependent transcription, regulation of transcription from RNA polymerase, DNA-dependent regulation of transcription, and gene expression (Fig. 7A). In addition, KEGG pathway analysis found that RASGs were mainly enriched in ErbB signaling and apoptosis pathways (Fig. 7B). These results demonstrated that LGALS1 may be involved in the regulation of cell apoptosis through regulating RASGs.

Validation of the LGALS1-mediated RASGs

Given that RASGs may be associated with NSCLC progression, we validated LGALS1-mediated RASGs using qPCR. Out of 9 detected AS events, the results of 3 A3SS AS events validated using RT-qPCR were consistent with the RNA-seq data, namely BCAP29, CSNKIE and MTFP1 (Fig. 8). Among them, the AS ratio of BCAP29 (M–W test, $U = 13$, $p = 0.0142$) and MTFP1 (M–W test, $U = 11$, $p = 0.0078$) were significantly increased in the siLGALS1 group compared to the siCtrl group, whereas CSNKIE (M–W test, $U = 17$, $p = 0.0193$) was found to be opposite. The RT-qPCR validation results of the remaining 6 AS events are shown in Supplementary Fig. 2. These results further confirm the LGALS1-related AS and RASE data.

Discussion

Despite extensive research regarding NSCLC, its mortality rate is still high, and more studies on NSCLC are warranted to understand the mechanism of the disease. In the present study, we have revealed insights into the transcriptomic landscape and AS dataset in NSCLC, driven by LGALS1 through RNA-seq. Many DEGs and RASGs in this study provide new insights and abundant potential candidate markers for NSCLC that could be targeted therapeutically.

As an RBP, LGALS1 is a β-galactoside-binding protein and is alternatively known as galectin-1.[28] Many studies have demonstrated that intervention in the expression of LGALS1 plays a vital role in the progression of numerous human cancers. For instance, a recent study showed that LGALS1 is involved in NSCLC progression by interacting with NCAPG, and the knockdown of LGALS1 significantly suppresses proliferation, migration and invasion of A549 and H1299 cells.[10] In oral cancer, LGALS1 knockdown suppressed cell proliferation by arresting cells at S phase and inhibited wound healing and migration.[29] Furthermore, in certain brain tumors, silencing of LGALS1 downregulated the expression of genes involved in cell cycle progression, resulting in an accumulation of G2/M phase cells and eventually inhibiting tumor progression.[30] There is also an extensive body of literature showing that LGALS1 participates in lung cancer progression. For example, LGALS1 is overexpressed in LUAD cells and induces their growth and invasive ability.[8] The LGALS1 upregulated p38 MAPK, ERK and cyclo-oxygenase-2 expression to promote lung cancer progression.[9] In the current study, we demonstrated that silencing of LGALS1 led to the aberrant expression of 225 genes in A549 cells and involved alterations in transcription and binding-related pathways, suggesting that LGALS1 may function in NSCLC. Therefore, further research is needed on RNA-binding properties and related biological functions of LGALS1.

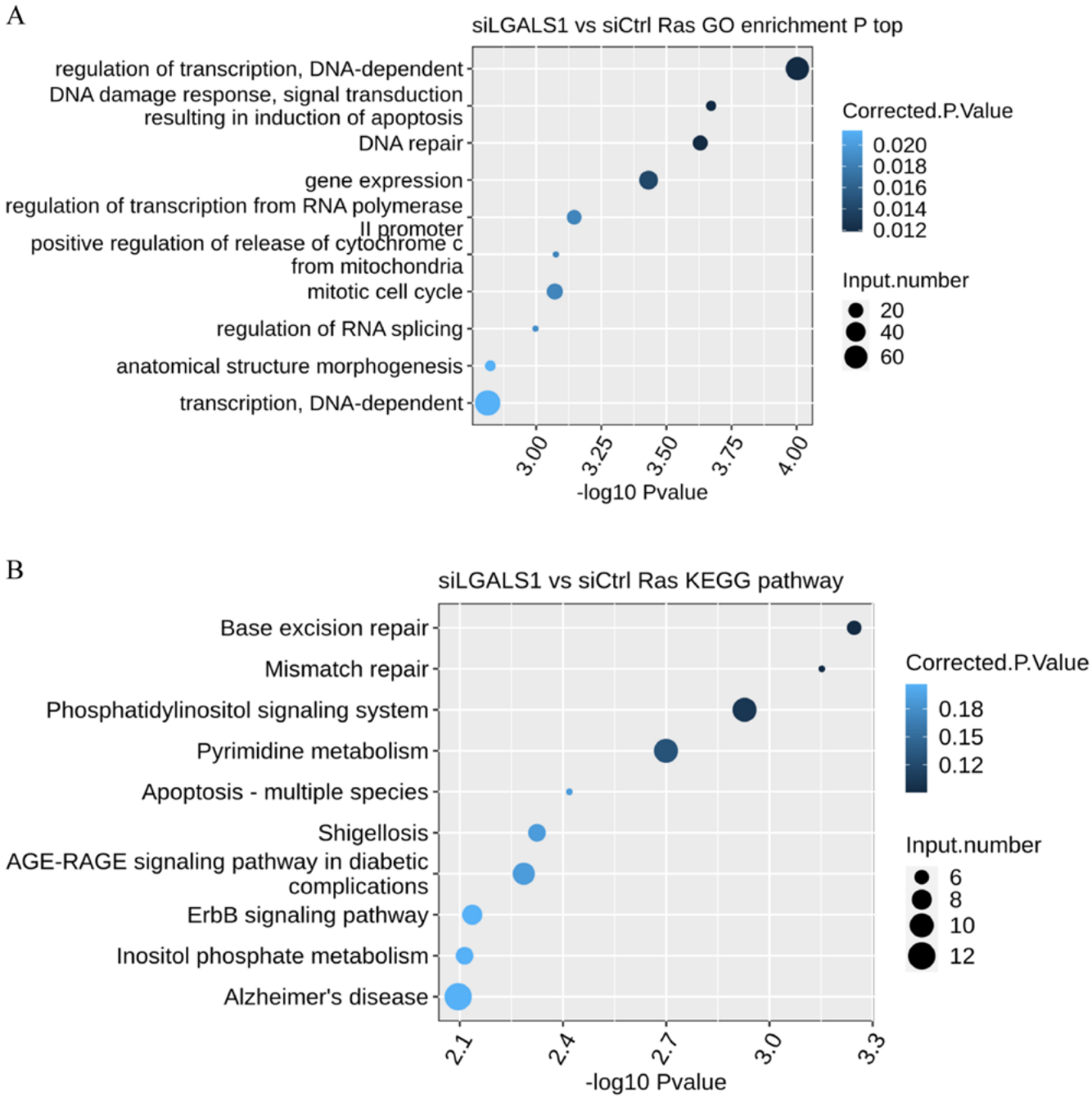

Fig. 7. Functional enrichment analysis of regulated alternative splicing genes (RASGs). A. The top 10 enriched Gene Ontology (GO) terms for RASGs; B. The top 10 enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways for RASGs

Alternative splicing events are closely associated with the function of RBPs and cancer progression,[31] with previous studies showing that AS plays a crucial role in the development of lung cancer. For instance, the AS status of BIN1 with exon 12A inclusion (the BIN1+12A isoform) could recover the activity of tumor cells through the regulation of BIN1 in NSCLC.[32] Furthermore, ESRP1 regulates the chemosensitivity of lung cancer cells through AS by inhibiting TGF-β/Smad signaling in SCLC patients.[33] Moreover, VEGFxxxb family members encoded AS in VEGF-A, and VEGF165b/VEGF165 are positively correlated with lymph node metastasis in NSCLC.[34] The current study showed that a total of 914 AS events were triggered after LGALS1 silencing, and the corresponding RASGs were implicated in ErbB signaling and apoptosis pathways. Encouragingly, some previous results support our observations. Wang et al. demonstrated that LGALS1 modulates vascular constriction by regulating AS of the cav1.2 calcium channel.[14] Therefore, LGALS1-mediated AS events play an important role in NSCLC progression.

In addition, we discovered that LGALS1-related RASGs were significantly enriched in the apoptosis and ErbB signaling pathways. Apoptosis, or programmed cell death,

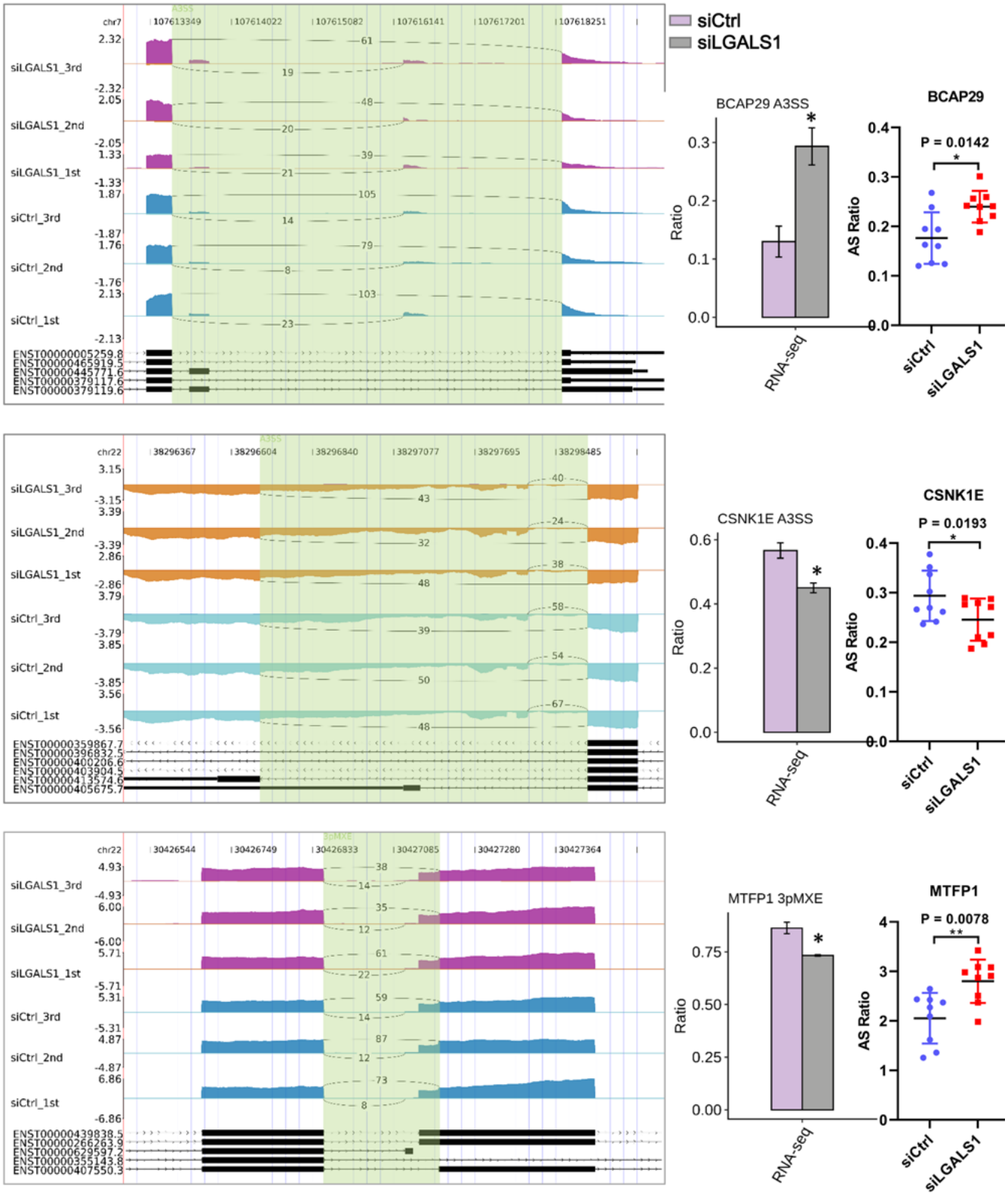

Fig. 8. Validation of the lectin galactoside-binding soluble 1 (LGALS1)-mediated regulated alternative splicing genes (RASGs). RNA-seq results and reverse transcription-quantitative polymerase chain reaction (RT-qPCR) validation results of RASGs are shown in the right panel

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

plays a critical role in cancer occurrence, progression and drug resistance. Inhibition of LGALS1 in melanoma cells can restore the viability of apoptotic T lymphocytes[35] while targeting LGALS1 by shikonin, and its derivatives were shown to induce apoptosis and autophagy in colorectal carcinoma cells.[36] Moreover, the ErbB signaling pathway is involved in cancer progression, for example, in cervical cancer.[37] Finally, EGFR is a member of the ErbB receptor

family, which regulates epithelial cell growth and survival, and its high expression or abnormal activation is correlated with drug resistance and even tumor progression in multiple cancers, such as breast cancer[38] and NSCLC.[39] Therefore, we suspect that LGALS1 may be involved in the NSCLC progression through the apoptosis and ErbB signaling pathways.

Limitations

There are some limitations of this study, one of which is that our results were not validated in animal models or clinical samples. Second, there are only 3 biological repeats for transcriptomic sequencing. Finally, no molecular experiments were performed to verify the regulatory relationship between LGALS1 on candidate target genes and AS events. Therefore, in the future, we will further explore the mechanism of LGALS1 involvement in NSCLC progression through a series of in vitro and in vivo experiments to address the above limitations.

Conclusions

In the current study, we characterized the transcriptomic landscape of A549 cells following LGALS1 silencing, identifying 225 DEGs that responded significantly to LGALS1 silencing. The expression patterns of *ELMO1*, *KCNJ2* and *HSPA6* were consistent between both RNA-seq and RT-qPCR experiments, and the expression of *ELMO1* and *KCNJ2* was upregulated, whereas *HSPA6* expression was downregulated after LGALS1 silencing. In addition, LGALS1 silencing led to 914 RASEs consisting of 10 types of AS residing in 759 genes. These LGALS1-mediated RASGs were mainly enriched in apoptosis and ErbB signaling pathways. Our results show potential novel molecular mechanisms of RBP LGALS1-mediated AS in NSCLC development.

Supplementary data

The supplementary materials are available at https://doi.org/10.5281/zenodo.8015729. The package contains the following files:

Supplementary Table 1. Primer information used in this study.

Supplementary Table 2. Results of quality control of RNA sequencing.

Supplementary Table 3. Statistics of all AS, known AS, and novel AS and their relative percentages.

Supplementary Table 4. Classification of all the detected known AS events.

Supplementary Table 5. Classification of all the detected novel AS events.

Supplementary Table 6. Distribution of each class of AS events (%).

Supplementary Fig. 1. LGALS1 expression in lung adenocarcinoma and lung squamous cell carcinoma between tumor and normal tissues in the TCGA database.

Supplementary Fig. 2. Validation of the LGALS1-mediated RASGs. RNA sequencing results and RT-qPCR validation results of RASGs were shown (^{ns} $p > 0.05$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

ORCID iDs

Hongyang Sh https://orcid.org/0000-0001-5139-970X
Yonghong Zhang https://orcid.org/0000-0003-1677-3537
Juan Liu https://orcid.org/0000-0003-3024-819X
Yu Wang https://orcid.org/0000-0003-1880-072X
Ping Fang https://orcid.org/0000-0002-9016-2323

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68(6):394–424. doi:10.3322/caac.21492
2. Ferlay J, Colombet M, Soerjomataram I, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int J Cancer*. 2019;144(8):1941–1953. doi:10.1002/ijc.31937
3. Zhou C, Wu YL, Chen G, et al. Final overall survival results from a randomised, phase III study of erlotinib versus chemotherapy as first-line treatment of EGFR mutation-positive advanced non-small-cell lung cancer (OPTIMAL, CTONG-0802). *Ann Oncol*. 2015;26(9):1877–1883. doi:10.1093/annonc/mdv276
4. Masuda K, Kuwano Y. Diverse roles of RNA-binding proteins in cancer traits and their implications in gastrointestinal cancers. *Wiley Interdiscip Rev RNA*. 2019;10(3):e1520. doi:10.1002/wrna.1520
5. Nahalka J. The role of the protein–RNA recognition code in neurodegeneration. *Cell Mol Life Sci*. 2019;76(11):2043–2058. doi:10.1007/s00018-019-03096-3
6. Pereira B, Billaud M, Almeida R. RNA-binding proteins in cancer: Old players and new actors. *Trends Cancer*. 2017;3(7):506–528. doi:10.1016/j.trecan.2017.05.003
7. Fuertes MB, Molinero LL, Toscano MA, et al. Regulated expression of galectin-1 during T-cell activation involves Lck and Fyn kinases and signaling through MEK1/ERK, p38 MAP kinase and $p70^{S6}$ kinase. *Mol Cell Biochem*. 2004;267(1/2):177–185. doi:10.1023/B:MCBI.0000049376.50242.7f
8. Qiu B, Zhang P, Xiong D, et al. CircRNA fibroblast growth factor receptor 3 promotes tumor progression in non-small cell lung cancer by regulating galectin-1-AKT/ERK1/2 signaling. *J Cell Physiol*. 2019;234(7):11256–11264. doi:10.1002/jcp.27783
9. Chung LY, Tang SJ, Sun GH, et al. Galectin-1 promotes lung cancer progression and chemoresistance by upregulating p38 MAPK, ERK, and cyclooxygenase-2. *Clin Cancer Res*. 2012;18(15):4037–4047. doi:10.1158/1078-0432.CCR-11-3348
10. Sun H, Zhang H, Yan Y, et al. NCAPG promotes the oncogenesis and progression of non-small cell lung cancer cells through upregulating LGALS1 expression. *Mol Cancer*. 2022;21(1):55. doi:10.1186/s12943-022-01533-9
11. Wu ZH, Tang Y, Zhou Y. Alternative splicing events implicated in carcinogenesis and prognosis of thyroid gland cancer. *Sci Rep*. 2021;11(1):4841. doi:10.1038/s41598-021-84403-6
12. Fu XD, Ares M. Context-dependent control of alternative splicing by RNA-binding proteins. *Nat Rev Genet*. 2014;15(10):689–701. doi:10.1038/nrg3778
13. Patterson RJ, Wang W, Wang JL. Understanding the biochemical activities of galectin-1 and galectin-3 in the nucleus. *Glycoconj J*. 2002;19(7–9):499–506. doi:10.1023/B:GLYC.0000014079.87862.c7
14. Wang J, Thio SSC, Yang SSH, et al. Splice variant specific modulation of CaV1.2 calcium channel by galectin-1 regulates arterial constriction. *Circ Res*. 2011;109(11):1250–1258. doi:10.1161/CIRCRESAHA.111.248849
15. Győrffy B, Surowiak P, Budczies J, Lánczky A. Online survival analysis software to assess the prognostic value of biomarkers using transcriptomic data in non-small-cell lung cancer. *PLoS One*. 2013;8(12):e82241. doi:10.1371/journal.pone.0082241

16. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol*. 2013;14(4):R36. doi:10.1186/gb-2013-14-4-r36
17. Trapnell C, Williams BA, Pertea G, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol*. 2010;28(5):511–515. doi:10.1038/nbt.1621
18. Robinson MD, McCarthy DJ, Smyth GK. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26(1):139–140. doi:10.1093/bioinformatics/btp616
19. Xie C, Mao X, Huang J, et al. KOBAS 2.0: A web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res*. 2011;39(Suppl 2):W316–W322. doi:10.1093/nar/gkr483
20. Xia H, Chen D, Wu Q, et al. CELF1 preferentially binds to exon-intron boundary and regulates alternative splicing in HeLa cells. *Biochim Biophys Acta Gene Regul Mech*. 2017;1860(9):911–921. doi:10.1016/j.bbagrm.2017.07.004
21. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods*. 2001;25(4):402–408. doi:10.1006/meth.2001.1262
22. Shi L, Zhang B, Sun X, et al. CC chemokine ligand 18(CCL18) promotes migration and invasion of lung cancer cells by binding to Nir1 through Nir1-ELMO1/DOC180 signaling pathway. *Mol Carcinog*. 2016;55(12):2051–2062. doi:10.1002/mc.22450
23. Lee JU, Chang HS, Jung CA, Kim RH, Park CS, Park JS. Upregulation of potassium voltage-gated channel subfamily J member 2 levels in the lungs of patients with idiopathic pulmonary fibrosis. *Can Respir J*. 2020;2020:3406530. doi:10.1155/2020/3406530
24. Li X, Noell G, Tabib T, et al. Single cell RNA sequencing identifies *IGFBP5* and *QKI* as ciliated epithelial cell genes associated with severe COPD. *Respir Res*. 2021;22(1):100. doi:10.1186/s12931-021-01675-2
25. Wang L, Hou J, Wang J, et al. Regulatory roles of HSPA6 in *Actinidia chinensis* Planch. root extract (acRoots)-inhibited lung cancer proliferation. *Clin Transl Med*. 2020;10(2):e46. doi:10.1002/ctm2.46
26. Umeyama H, Iwadate M, Taguchi YH. *TINAGL1* and *B3GALNT1* are potential therapy target genes to suppress metastasis in non-small cell lung cancer. *BMC Genomics*. 2014;15(Suppl 9):S2. doi:10.1186/1471-2164-15-S9-S2
27. Witkiewicz AK, McMillan EA, Balaji U, et al. Whole-exome sequencing of pancreatic cancer defines genetic diversity and therapeutic targets. *Nat Commun*. 2015;6(1):6744. doi:10.1038/ncomms7744
28. Barrientos G, Freitag N, Tirado-González I, et al. Involvement of galectin-1 in reproduction: Past, present and future. *Hum Reprod Update*. 2014;20(2):175–193. doi:10.1093/humupd/dmt040
29. Li JM, Tseng CW, Lin CC, et al. Upregulation of LGALS1 is associated with oral cancer metastasis. *Ther Adv Med Oncol*. 2018;10:175883591879462. doi:10.1177/1758835918794622
30. Sharanek A, Burban A, Hernandez-Corchado A, et al. Transcriptional control of brain tumor stem cells by a carbohydrate binding protein. *Cell Rep*. 2021;36(9):109647. doi:10.1016/j.celrep.2021.109647
31. Oltean S, Bates DO. Hallmarks of alternative splicing in cancer. *Oncogene*. 2014;33(46):5311–5318. doi:10.1038/onc.2013.533
32. Wang J, Liu T, Wang M, et al. SRSF1-dependent alternative splicing attenuates BIN1 expression in non-small cell lung cancer. *J Cell Biochem*. 2020;121(2):946–953. doi:10.1002/jcb.29366
33. Zheng M, Niu Y, Bu J, et al. ESRP1 regulates alternative splicing of CARM1 to sensitize small cell lung cancer cells to chemotherapy by inhibiting TGF-β/Smad signaling. *Aging*. 2021;13(3):3554–3572. doi:10.18632/aging.202295
34. Boudria A, Abou Faycal C, Jia T, et al. VEGF165b, a splice variant of VEGF-A, promotes lung tumor progression and escape from anti-angiogenic therapies through a β1 integrin/VEGFR autocrine loop. *Oncogene*. 2019;38(7):1050–1066. doi:10.1038/s41388-018-0486-7
35. Górniak P, Wasylecka-Juszczyńska M, Ługowska I, et al. BRAF inhibition curtails IFN-gamma-inducible PD-L1 expression and upregulates the immunoregulatory protein galectin-1 in melanoma cells. *Mol Oncol*. 2020;14(8):1817–1832. doi:10.1002/1878-0261.12695
36. Zhang N, Peng F, Wang Y, et al. Shikonin induces colorectal carcinoma cells apoptosis and autophagy by targeting galectin-1/JNK signaling axis. *Int J Biol Sci*. 2020;16(1):147–161. doi:10.7150/ijbs.36955
37. Zhao R, Song J, Jin Y, Liu Y. Long noncoding RNA HOXC-AS3 enhances the progression of cervical cancer via activating ErbB signaling pathway. *J Mol Histol*. 2021;52(5):991–1006. doi:10.1007/s10735-021-10007-z
38. Bolado-Carrancio A, Lee M, Ewing A, et al. ISGylation drives basal breast tumour progression by promoting EGFR recycling and Akt signalling. *Oncogene*. 2021;40(44):6235–6247. doi:10.1038/s41388-021-02017-8
39. Chiu TH, Lin CY, Hsieh MH, Lin SM, Fang YF. Prognostic factors in lung adenocarcinoma with bone metastasis treated with EGFR-TKIs. *Medicina (Kaunas)*. 2021;57(9):967. doi:10.3390/medicina57090967