Nucleic Acids Research, 2025, **53**, gkaf505
https://doi.org/10.1093/nar/gkaf505
Gene regulation, Chromatin and Epigenetics

Super-enhancer-mediated circRNAs exhibit high splicing circularization diversity and transcriptional activity

Shaodong Huang[1,2,†], Yulong Han[3,4,†], Yiran Liu[1,2,†], Lina Bu[1,2,†], Chenyang Wu[1,2], Ziyan Rao[1,2], Chuan Ye[1,2], Hansen Cheng[1,2], Yunxi Liao[1,2], Yunan Sun[1,2], Yushu Zou[1,2], Zixu Wang[1,2], Bryan-Yu Sun[1,2], Shu Meng[3,4,*], Dongyu Zhao[1,2,*]

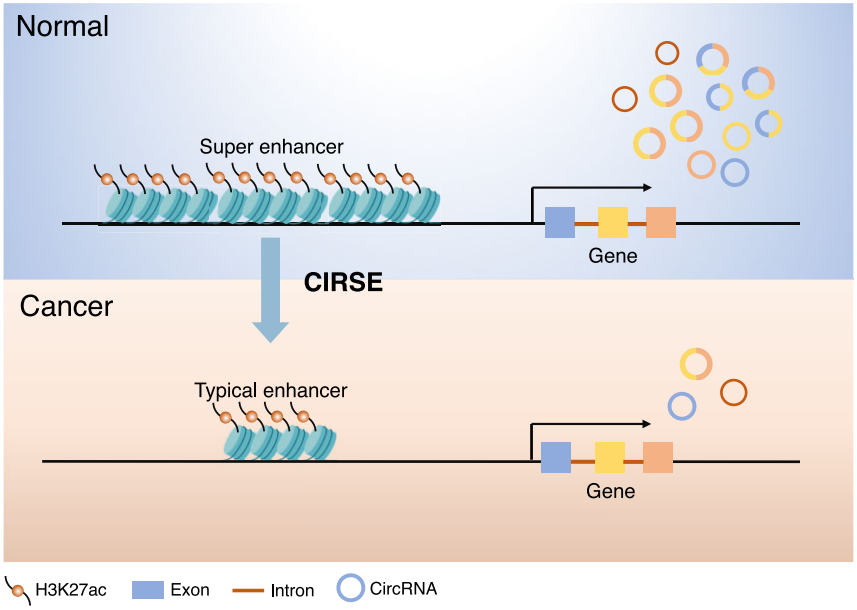
[1]Department of Biomedical Informatics, School of Basic Medical Sciences, Peking University, Beijing 100191, China
[2]State Key Laboratory of Vascular Homeostasis and Remodeling, Peking University, Beijing 100191, China
[3]Department of Basic Science Research, Guangzhou National Laboratory, Guangzhou 510005, China
[4]State Key Laboratory of Respiratory Disease, The First Affiliated Hospital of Guangzhou Medical University, Guangzhou Medical University, Guangzhou, Guangdong 510120, China

[*]To whom correspondence should be addressed. Email: zhao.dongyu@pku.edu.cn
Correspondence may also be addressed to Shu Meng. Email: meng_shu@gzlab.ac.cn
[†]The first four authors should be regarded as Joint First Authors.

Abstract

Circular RNAs (circRNAs), an emerging subclass of noncoding RNAs, have been increasingly recognized as critical regulators in diverse biological functions and cellular processes. Despite their functional significance, the epigenetic mechanisms governing circRNA biogenesis remain poorly understood. Our study reveals that H3K27ac-marked super-enhancers (SEs) significantly enhance both circRNA splicing circularization diversity and transcriptional activation of their host genes. Intriguingly, other histone modifications—including H3K4me3, H3K36me3, H3K27me3, and H3K9me3—exhibit distinct regulatory effects on circRNA transcriptional activity. Through comprehensive analysis of 195 transcriptomic profiles, we identified a pan-cancer epigenomic tumor-suppressor signature termed CircRNA Isoform Reduction for Shortened Enhancers in cancer (CIRSE). Notably, CIRSE demonstrates strong prognostic potential in lung adenocarcinoma, as validated by comprehensive survival analyses. Combining Nanopore sequencing with CLIP-Seq approaches, we further elucidated the dual regulatory mechanism involving circRNA stability maintenance and back-splicing junction selection mediated by specific RNA-binding proteins. Functional validation confirmed that CIRSE-defined tumor-suppressive circRNAs are essential for maintaining malignant phenotypes in cancer models. Our findings not only provide mechanistic insights into the epigenetic regulation of circRNAs, but also pave the way for mutation-agnostic discovery of tumor-suppressive circRNAs in precision oncology applications.

Graphical abstract



Received: February 10, 2025. **Revised:** April 27, 2025. **Editorial Decision:** May 24, 2025. **Accepted:** May 28, 2025

Introduction

Circular RNAs (circRNAs) have attracted increasing attention as a class of single-stranded, covalently closed, noncoding RNA molecules [1, 2]. CircRNAs are derived from pre-mRNAs through a noncanonical splicing event called back-splicing, in which a downstream 5′ splice donor is joined with an upstream 3′ splice acceptor [3–6]. Increasing evidences has shown that circRNAs play diverse biological roles [7–9], including acting as circRNA sponges [10–12], interacting with proteins [13], translating protein [14], and suppressing host gene (genes expressing circRNAs) expression by competing with canonical splicing [15, 16]. Numerous cancer driver circRNAs have been identified, such as circ0001946 (CDR1as), whose loss promotes melanoma invasion and metastasis [17], and circ0001445 (circSMARCA5), which enhances drug sensitivity in breast cancer cells [16]. These findings support the potential of circRNAs as promising therapeutic targets and clinical biomarkers [18, 19]. However, most studies still focus on individual circRNAs, a systematic approach to screen cancer driver circRNAs remains limited.

An additional layer of complexity lies in the remarkable isoform diversity of circRNAs, mainly driven by alternative splicing and circularization, enabling a host gene to express multiple circRNA isoforms [20–23]. While emerging studies have documented specific instances of circRNA isoform variation through advanced sequencing approaches, the regulatory circuitry governing their biogenesis and the functional implications of this isoform diversification remain poorly understood [21].

As cancer is increasingly recognized as a highly epigenetics-related process, considerable attention has been paid to epigenetic alteration during the tumorigenesis [24]. As a type of genomic *cis*-regulatory element [25, 26], enhancers dynamically facilitate gene expression in different tissues and can be identified in many ways, including H3K27ac and H3K4me1 ChIP-Seq [27–29]. Among them, H3K27ac is widely used in the identification of active enhancers [30]. A subset of enhancers, termed super-enhancers (SEs), are characterized by dense occupancy of master transcription factors and broad H3K27ac enrichment, often spanning several kilobases. SEs are known to control genes that define cell identity and are frequently involved in disease or cancer development [31–36]. However, although SEs have been widely studied in the context of coding gene regulation and disease pathogenesis [29, 32, 34, 37], their functional impact on noncoding RNA biogenesis, particularly circRNAs, remains unclear [37].

In light of the limited understanding of this area, we profiled a landscape of circRNAs and enhancers across multiple cell types. Moreover, the concept of splicing circularization diversity was introduced, which referred to the circRNA diversity of genes. It was measured by the circRNA isoform count (the number of distinct circRNA isoforms derived from a single gene). By integrating epigenomic and transcriptomic data, we revealed the relationship between circRNA splicing circularization diversity and SEs, and investigated the functions of host genes targeted by different enhancers. We further analyzed the alterations in enhancer and circRNA profiles between cancer and normal cells. Through this analysis, we identified a tumor-suppressive signature termed CircRNA Isoform Reduction for Shortened Enhancers in cancer (CIRSE), characterized by a reduction in circRNA isoform diversity linked to enhancer shortening in cancer. Host genes showing this signature were defined as CIRSE genes, and circRNAs derived from these genes, whose expression was downregulated in cancer, were defined as CIRSE circRNAs. Finally, based on the CIRSE signature, we developed a novel screening pipeline to identify potential tumor-suppressive circRNAs, offering a more systematic and efficient approach than conventional experimental methods such as cell viability, wound healing, and transwell assays.

Materials and methods

Cell culture

The human NSCLC cell line H1975 (ATCC, Manassas, VA, USA) was cultured in RPMI-1640 medium (Gibco, Gaithersburg, MD, USA), fortified with 10% fetal bovine serum (FBS) at 37°C with 5% CO_2. The cells reached the exponential stage and then underwent 0.25% trypsin digestion (Thermo Fisher, China) and routine passaging every 2–3 days. After three passages, proliferating H1975 cells were harvested for subsequent research.

Cytoplasmic and nuclear separation

Cytoplasmic and nuclear fractions were separated. A Cytoplasmic and Nuclear RNA Purification Kit (AmyJet Scientific Incorporation) was used to detect circRNA expression *in vitro*. H1975 cells were collected for RNA extraction. Real-time quantitative polymerase chain reaction (RT-qPCR) was subsequently used to verify circRNA expression across varying localities.

RT-qPCR analysis and semi-quantitative analysis

Total RNA was isolated using TRIzol reagent (Invitrogen) according to the manufacturer's instructions. Total RNA was digested using RNase R to enrich the circRNAs, and the RNase R-treated RNA was reverse-transcribed using random primers. GoScript Reverse Transcriptase (Promega) was used to synthesize the first-strand complementary DNA. RT-qPCR was performed using the gene-specific primers listed in Supplementary Table S1. The polymerase chain reaction products were tested using 2% agarose gel electrophoresis. Separation was performed at 120 V for 30 min, and the products were visualized under ultraviolet light. A DNA Marker 2000 (TSINGKE, China) was used as the size reference. Semi-quantitative analysis was conducted using ImageJ software.

Cell proliferation assay

Cell viability was assessed using CCK-8 kits (Dojindo, Japan). The H1975 cells were plated into 96-well plates (2000 cells/well) for varying treatments. Posttreatment (12, 24, 36, 48, and 72 h) wells were supplemented with CCK-8 and incubated for an additional 2 h prior to measuring 450 nm absorbance using an automated microplate reader (Bio-Rad, USA).

Scratch test assay

H1975 cellular migration was assessed using cell scratch assays. Trypsin-detached cells were seeded at 1×10^5 cells per well of a six-well plate and transfected with circRNA small interfering RNA (siRNA) or negative control siRNA. Upon attaining 90% confluence, a sterile pipette tip was used to uniformly etch a transverse plate lesion, after which incubation

took place for 24 and 48 h in separate groups. Subsequently, wound healing was examined microscopically (Olympus) at 0, 24, and 48 h, and the images were captured.

Transwell assays

H1975 cell invasion was assessed using transwell assays. Transwell chambers of an 8-μm pore size (Corning, Corning, NY, USA) were placed in 24-well plates, and the apical chamber was covered (to detect invasion) with 80 μl of Matrigel (Becton Dickinson). Approximately 1×10^4 H1975 cells were suspended in the apical compartment after transfection with circRNA siRNA or negative control siRNA. The basal compartment received 600 μl of complete medium supplemented with 10% FBS. After incubation at room temperature for 24 h, the chambers were removed, and the noninvading cells in the apical chamber were wiped off with cotton swabs. The invading cells in the basal chamber were fixed for 30 min in 4% paraformaldehyde (PFA). The cells were stained for 30 min with 0.1% crystal violet and observed under a microscope (Olympus). The number of invaded H1975 cells was quantified.

Actinomycin D assays

Seeded IMR90 or A549 cells in 12-well plates were treated with 2.5 μg/ml actinomycin D. All samples were then collected at designated intervals and subjected to RT-qPCR analysis to determine the stability of circRNA.

General bioinformatics analysis

We used a one-tailed Wilcoxon test to determine the significance of the difference in the boxplot and to obtain the differential circRNAs and enhancers in the pan-cancer analysis because of the many samples [38]. We used edgeR in lung cancer and B-cell lymphoma [39]. Fisher's exact test and quantile normalization were performed using Python script. To ensure comparable distributions, we performed stratified sampling based on gene (or circRNA) length or expression levels. Specifically, genes from the SE, typical enhancers (TE), and non-enhancer (NE) groups were ranked and partitioned into bins according to either length or expression [e.g. using a bin size of 100 TPM (transcripts per kilobase per million mapped reads) for expression]. Within each bin, we randomly selected an equal number of genes from each group, matched to the smallest group size, to construct the final subsets.

Back-splicing junction site sequence detection

For circRNAs annotated in CircBase [40], back-splicing junction (BSJ) sequences were directly retrieved from the database. For the circRNAs *de novo* identified by CIRIquant [41], BSJ sequences were extracted from the pseudo-circular reference generated by concatenating two full-length sequences of the BSJ region. The pseudo-circular reference was generated using CIRIquant [41].

CircRNA and mRNA identification and quantification

RNA-Seq reads were mapped to the human genome (hg19) using BWA (v0.7.17-r1188) and Hisat2 (v2.2.1) [42, 43]. Expression levels of circRNAs and messenger RNAs (mRNAs) were quantified by CIRIquant and CIRCexplorer3 algorithm with default parameters [41, 44]. Ensembl (release 87) GTF was used to annotate the genes and circRNAs [45]. BSJ read count per million mapped reads (CPM) and the circRNA junction ratio were used for quantification of circRNA expression. The gene expression level was measured in TPM.

ChIP-Seq data analysis

ChIP-Seq reads were mapped to the human genome version hg19 using the bowtie version 1.2 [46]. We then submitted the mapped reads to the Dregion function in DANPOS version 2.2.2 (https://sites.google.com/site/danposdoc/) to calculate the ChIP-Seq signal (read density) at each base pair of the genome, subtract the background input signal, normalize the read number, and define the individual enrichment peaks [47]. The Dregion function stored the signal values at each base pair in a Wiggle format file. The files were converted into BigWig using WigToBigWig (https://www.encodeproject.org/software/wigtobigwig/). The BigWig files were submitted to integrative genomics viewer (IGV) to visualize the ChIP-Seq signal at each base pair [48–50].

Identification of SEs

SEs were identified using Rank Ordering of Super-Enhancers (ROSE) algorithm, with H3K27ac peaks and corresponding BAM files as input [29, 51]. The stitching size was 12.5 kb. The transcription start site (TSS) exclusion zone size was 2 kb. The identified enhancers were annotated using the GREAT algorithm with default parameters. The used annotation gene set was GREAT gene, which includes protein-coding genes located on assembled chromosomes and associated with at least one meaningful annotation [52]. For genes linked to multiple enhancers, SEs were considered to exert dominant regulatory influence. In such cases, the widest enhancer was assigned as the primary regulatory element.

CircRNA expression imputation

The expression levels of circRNAs and mRNAs in each sample were calculated to fill the technical zero value. Then we aggregated them together and generated an expression matrix. *K*-nearest neighbor smoothing (KNN-smoothing) algorithm was performed to fill the zero value by aggregating information from similar neighbors in the expression matrix [53]. The number of neighbors used for smoothing was 10 (parameter -k 10). The number of principal components used to identify neighbors was 10 (parameter -d 10). The seed for pseudo-random number generator was set as 0 (parameter -s 0). The expression matrix after imputation across various cell lines was used to calculate the circRNA isoforms of each gene and perform subsequent analysis.

Analysis of CRISPRi-Screen

Data from the CRISPRi-Screen data of OVCAR-3 were collected from public data [54] (Supplementary Table S2). As the data were single-end RNA-Seq, we performed CIRCexplorer3 to detect and quantify circRNAs with default parameters [44]. We gathered and compared the relative expression (quantified by CIRCscore) and circRNA isoform counts of the circRNAs derived from the gene loci targeted by SEs in the SE-silence group with those in the control group.

Analysis of BSJ-related RBP

CLIP-Seq data of various RNA-binding proteins (RBPs) were downloaded from CLIPdb [55]. We compared the proportion of BSJ bound by RBPs to forward-splicing junction (FSJ). *P*-values were determined by proportion *Z*-test. The absolute value of significance score was $-\log_{10}P$. If the RBP binding proportion of BSJ is higher than that of FSJ, the significance score is positive. Otherwise, it was negative. The final significance scores for each RBPs were calculated as the average of their scores across different cell lines (Supplementary Table S3). To avoid potential impacts of transcriptional activity differences, the CLIP-Seq signal density curves for RBPs were derived from the genomic loci of BSJ sites in SE, TE, and NE circRNAs with similar host gene expression levels.

CircRNA identification and quantification in Nanopore RNA-Seq data

The Nanopore RNA-Seq data in the IMR-90 and A549 cell lines, with two replication samples in each cell line, were integrally processed using the CIRI-long tool (https://github.com/bioinfo-biols/CIRI-long) [56]. We successfully performed circRNA identification and the isoform collapse steps, which returned the candidate circRNA FASTA files and multiple matrices about the circRNA information. Finally, we obtained the circRNA expression level and isoform count in each cell line based on Nanopore full-length sequencing.

CircRNA secondary structure prediction and minimal free energy calculation

The full-length circRNA FASTA file from CIRI-long was processed using the RNAfold function from the ViennaRNA package (https://www.tbi.univie.ac.at/RNA/). It predicted the minimal free energy structure and the corresponding minimal free energy with default parameters [57]. The circRNAs for comparison had a length and sequence similarity index >0.7. Length similarity index was the ratio of shorter circRNA length to the longer circRNA length. Sequence similarity index was scaled alignment score that was quantified through global sequence alignment with Biopython pairwise2 module [58].

Analysis of DRB/TT_{chem}-seq

The raw reads were mapped to the target (*Homo sapiens* hg19) and spike-in (*Saccharomyces cerevisiae* sacCer3) genomes with the -quantMode GeneCounts option [59]. The scale factors were calculated using the read count from the yeast spike-in. We created scaled, strand-specific BigWig files by first using SAMtools to split the BAM file into forward strand and reverse strand [60], and then applying deep bamCoverage function with the -scaleFactor to convert BAM file to scaled BigWig files [61]. The pol II speed was calculated by R script (https://github.com/crickbabs/DRB_TT-seq/releases/tag/v1.2) [62].

Identification of reverse complementary matches in flanking introns

Exons were initially excluded from the hg19 GTF file to generate a comprehensive intronic BED file. For each circRNA, the closet upstream and downstream introns were identified as flanking introns by closest function in BEDTools [63]. The regions were used for reverse complementary match (RCM) identification. We used MMseqs2 to perform sequence alignment between upstream introns and downstream introns [64]. Intron pairs with *E*-value <0.01 and exhibited reverse complementary pairing were considered as RCMs.

Statistical analysis

All experiments were performed at least twice. Tukey's multiple comparison test and one-way analysis of variance were used for statistical analyses. Unless otherwise specified, Student's *t*-test was used to compare data conforming to a normal distribution between the two groups. In cases in which data from the two groups exhibited a non-normal distribution or heterogeneity of variance, the Mann–Whitney *U* test was applied. For pairwise comparisons between groups, Tukey's multiple comparison test was deemed appropriate. Data conforming to a normal distribution and exhibiting homogeneity of variance from multiple groups were analyzed using Dunnett's *t*-test, unless otherwise stated. If the data from multiple groups had a non-normal distribution or showed heterogeneity of variance, the Kruskal–Wallis *H* test was used instead. The mean value ± SD is presented in all graphs.

Results

SE-regulated genes tend to express diverse circRNA isoforms

To establish a comprehensive circRNA landscape, we analyzed 195 rRNA-depleted RNA-Seq datasets spanning 61 distinct cell lines and tissue types from ENCODE and GEO databases. This collection comprised 36 normal cell types (100 samples) and 25 cancer cell types (95 samples), representing diverse biological contexts. Through rigorous computational analysis using CIRIquant [41], we systematically identified 435 946 circRNAs (Supplementary Fig. S1A). The detected circRNAs quantified as BSJ read CPM exhibited a broad dynamic range of expression levels (0.0052–37.7 CPM). The vast majority of circRNAs exhibited low abundance, with over 99.3% showing expression levels below 1 CPM (Supplementary Fig. S1B). We also classified circRNAs by their genomic loci of BSJ [41]. Regional distribution analysis revealed distinct compartmentalization: exonic-derived circRNAs constituted the predominant subgroup (58.1%), followed by intronic-derived subgroup (26.2%). The remaining populations consisted of antisense circRNAs (12.8%) and intergenic circRNAs, which represented the smallest proportion at merely 2.9% (Supplementary Fig. S1C), consistent with previous studies [65, 66]. Consequently, our investigation specifically targeted circRNAs originating exclusively from genes. Through alternative splicing and circularization mechanisms, individual gene loci were shown to be capable of generating multiple distinct circRNA isoforms [67]. Our analysis revealed that 78.9% of host genes produced at least two circRNA isoforms, with a notable 40.2% generating ≥10 isoforms (Supplementary Fig. S1D). To ensure analytical rigor, we cross-validated our findings using an independent computational approach (CIRCexplorer3), which demonstrated strong concordance with our primary results [44] (Supplementary Fig. S1E).

Intriguingly, we initially observed that genes modified by broad H3K27ac peaks in IMR-90 cells commonly expressed abundant circRNAs with multiple isoforms, such as

MYH9, *CHD2*, and *YWHAE* (Fig. 1A and Supplementary Table S4). Validation via RT-qPCR confirmed the expression of most of identified circRNAs (Supplementary Fig. S2 and Supplementary Table S4). Conversely, genes exhibiting diminished or absent H3K27ac signals demonstrated significantly reduced circRNA expression. Notably, loci such as *FBXO8*, *PPIE*, and *TTC12* produced only sparse circRNA transcripts, accompanied by a marked restriction in isoform diversity (Fig. 1B, Supplementary Fig. S2, and Supplementary Table S4). SEs, which are essentially clusters of enhancers, are usually characterized by wide H3K27ac peaks [29, 30]. This motivated us to specifically investigate the regulatory mechanism between SEs and circRNAs biogenesis. We first performed SE identification in IMR-90 cells using the ROSE algorithm [29, 51]. This analysis revealed 1423 SEs and 25 747 TEs (Fig. 1C), consistent with established genomic enhancer profiling methodologies. Through comprehensive enhancer annotation, we stratified genes into three distinct regulatory categories: (i) SE genes (associated with SEs), (ii) TE genes (linked to TEs), and (iii) NE-regulated genes. Subsequent analysis demonstrated significant differences in circRNA production across these groups. SE genes exhibited the highest splicing circularization diversity, followed by TE genes, whereas NE genes showed minimal circRNA isoform count. This enhancer-dependent stratification revealed a strong positive correlation between enhancer strength and circRNA isoform complexity. This phenomenon was reproducibly observed across both Illumina and Nanopore sequencing (Fig. 1D and Supplementary Fig. S3A). Notably, cumulative abundance of circRNAs derived from SE host genes significantly exceeded that of circRNAs originating from TE and NE host genes (Fig. 1E and Supplementary Fig. S3B). Furthermore, SE genes demonstrated significant enrichment for genes generating multiple circRNA isoforms compared to other gene categories (Fig. 1F). To further explore this trend, we quantified the proportion of genes exhibiting high circRNA isoform diversity among SE, TE, and NE genes, calculating respective ratios relative to NE host genes. Strikingly, SE host genes showed a markedly higher ratio than TE host genes, with this disparity becoming more pronounced under increasingly stringent cutoff thresholds (Fig. 1G). Moreover, KEGG pathway analysis showed that SE host genes were significantly enriched in cancer-related pathways, including pathways in cancer, microRNAs in cancer, and non-small cell lung cancer pathways (Supplementary Fig. S3C). Gene set enrichment analysis (GSEA) also revealed that host genes regulated by wide enhancers were enriched in cancer-associated pathways (Supplementary Fig. S3D–G).

High splicing circularization diversity in SE host genes is independent of gene length and expression levels

Further analysis revealed that SE genes exhibited significantly greater gene length or higher expression level compared to TE and NE genes (Fig. 2A and B). These findings prompted us to investigate whether the enhanced splicing circularization diversity observed in SE genes could be principally attributed to their gene length or expression level [23, 29, 68, 69]. Therefore, we strategically regrouped SE, TE, and NE genes into subsets with similar gene lengths or expression levels (Fig. 2C and D). Although the three gene sets (SE, TE, and NE) were carefully balanced for gene length and expression levels, SE genes consistently demonstrated enhanced circRNA isoform production compared to TE and NE genes (Fig. 2E and F). This phenomenon was exemplified by broad H3K27ac-modified genes such as *G3BP1*, *DOCK10*, and *HSPA8*, which generated diverse circRNA isoforms. In contrast, genes with comparable gene lengths but regulated by narrow H3K27ac peaks, including *DNAJC12*, *CCDC148*, and *CORO1B*, exhibited markedly fewer circRNA isoforms (Fig. 2G). Consistently, genes such as *MEIS1*, *MICAL2*, and *EXT1* produced a higher number of circRNA isoforms compared to *NQO1*, *RPS27A*, and *TSPO*, although exhibiting comparable mRNA expression levels (Fig. 2H). Furthermore, to establish causal evidence for SE-mediated regulation of circularization events, we performed functional validation using CRISPR interference (CRISPRi)-mediated SE perturbation in OVCAR-3 cellular models [54]. Following systematic SE perturbation, we observed marked downregulation of circRNA expression levels (Supplementary Fig. S4A). The count of circRNAs detected around the SEs also exhibited a substantial reduction (Supplementary Fig. S4B). The results indicated that SEs play a key role in the regulation of splicing circularization diversity, independent of gene length and expression level.

SEs enhance circRNA diversity cross cell types

Given that SE genes exhibited a higher propensity to express circRNAs in IMR-90, we sought to determine whether this association existed robustly across diverse cell lines. Enhancer annotations across 32 cell lines, including 17 cancer and 15 normal cell types, were collected from GEO, ENCODE, and SEdb 2.0 [70]. We observed the proportion of host genes to access their trends to express circRNAs. SE genes exhibited the strongest tendency to express circRNA and this phenomenon was highly consistent across multiple cell types (Fig. 3A and Supplementary Fig. S5A). A similar pattern was also observed in splicing circularization diversity, where the SE host genes produced significantly more circRNA isoform than TE and NE host genes in almost each cell type (Fig. 3B and Supplementary Table S5). However, although the regulation relationship between SEs and splicing circularization diversity exhibited broad conservation, the identity of SE host genes was cell-type specific. Counting the frequency of genes identified as SE, TE, or NE host genes across 32 cell lines, we observed that, in most cases, genes were simultaneously identified as the SE host genes in only a few cell lines, but as TE or NE host genes in more cell types. Intriguingly, TE host genes were more conserved than NE host genes. We speculated that TE host genes and their circRNAs may function like housekeeping genes (Fig. 3C).

Building upon our findings of distinct enhancer regulatory patterns, we conducted a study of other histone modifications. Similarly, genes were categorized by the width of histone modification coverage in the genomic regions. Genes with broad H3K4me3 and broad H3K36me3 exhibited a high likelihood of expressing circRNAs, while genes with narrow peaks or absent histone modification showed a low tendency to express circRNAs (Fig. 3D and E). In contrast, broad H3K27me3 and broad H3K9me3 tended to inhibit circRNA expression (Fig. 3F and G). In addition to these four common histone modifications, other histone modifications or variants, including H3K79me2, H3K9ac, H3K4me2, H4K20me1, H3K4me1, and H2AFZ, also revealed a potential relationship with circRNA expression (Supplementary Fig. S5B).

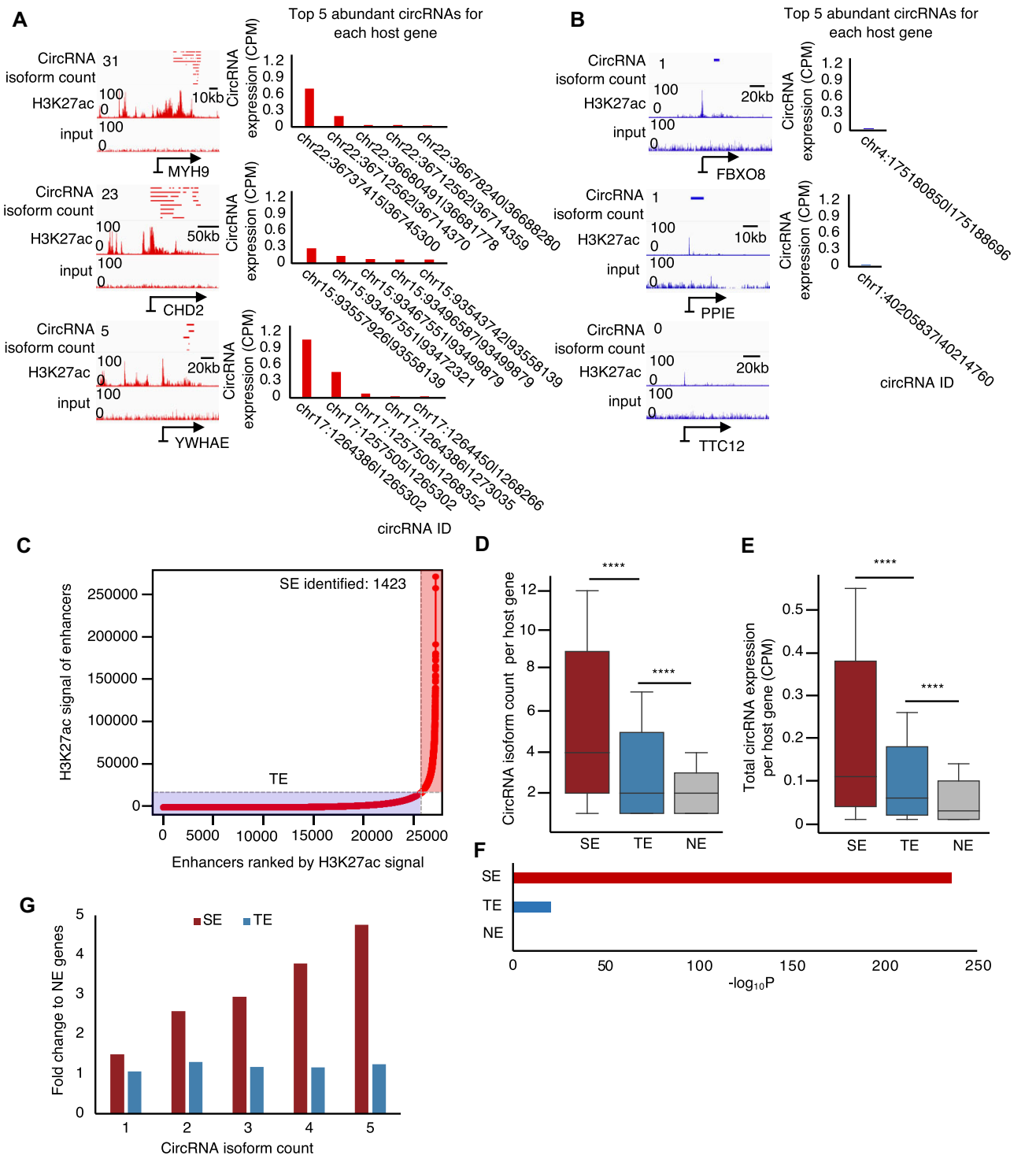


Figure 1. Genes regulated by SEs express multiple circRNA isoforms. (**A**, **B**) H3K27ac ChIP-Seq signal of the gene. The line above the ChIP-Seq signal denotes a circRNA isoform. The relative positions to the genes show the genomic loci of the BSJ reads (left). The number on the top left corner represents the number of circRNA isoforms detected. Expression of the top 5 abundant circRNA isoforms derived from the same host genes (right). CPM, back-spliced junction read CPM. CircRNA ID, genomic loci for BSJ of circRNA *de novo* identified. (**C**) Distribution of the H3K27ac signal across enhancers identified in IMR-90. The uneven distribution of the signal allowed for the identification of 1423 SEs. Boxplot showing the circRNA isoform count (**D**) and the sum of the total circRNA abundance (**E**) of host genes in different groups. SE, host genes regulated by SEs. TE, host genes regulated by TEs. NE, host genes without regulating enhancers. *P*-values determined by one-tailed Wilcoxon test. CircRNA abundance unit, CPM. (**F**) Enrichment level of host genes with a circRNA isoform count ≥4 in different groups. *P*-values determined by one-tailed Fisher's exact test. (**G**) Fold change in the occupation of high circRNA isoform count genes in SE and TE genes to that of NE genes. The *X*-axis represents the cutoff for high splicing circularization diversity genes. ****$P < .0001$.

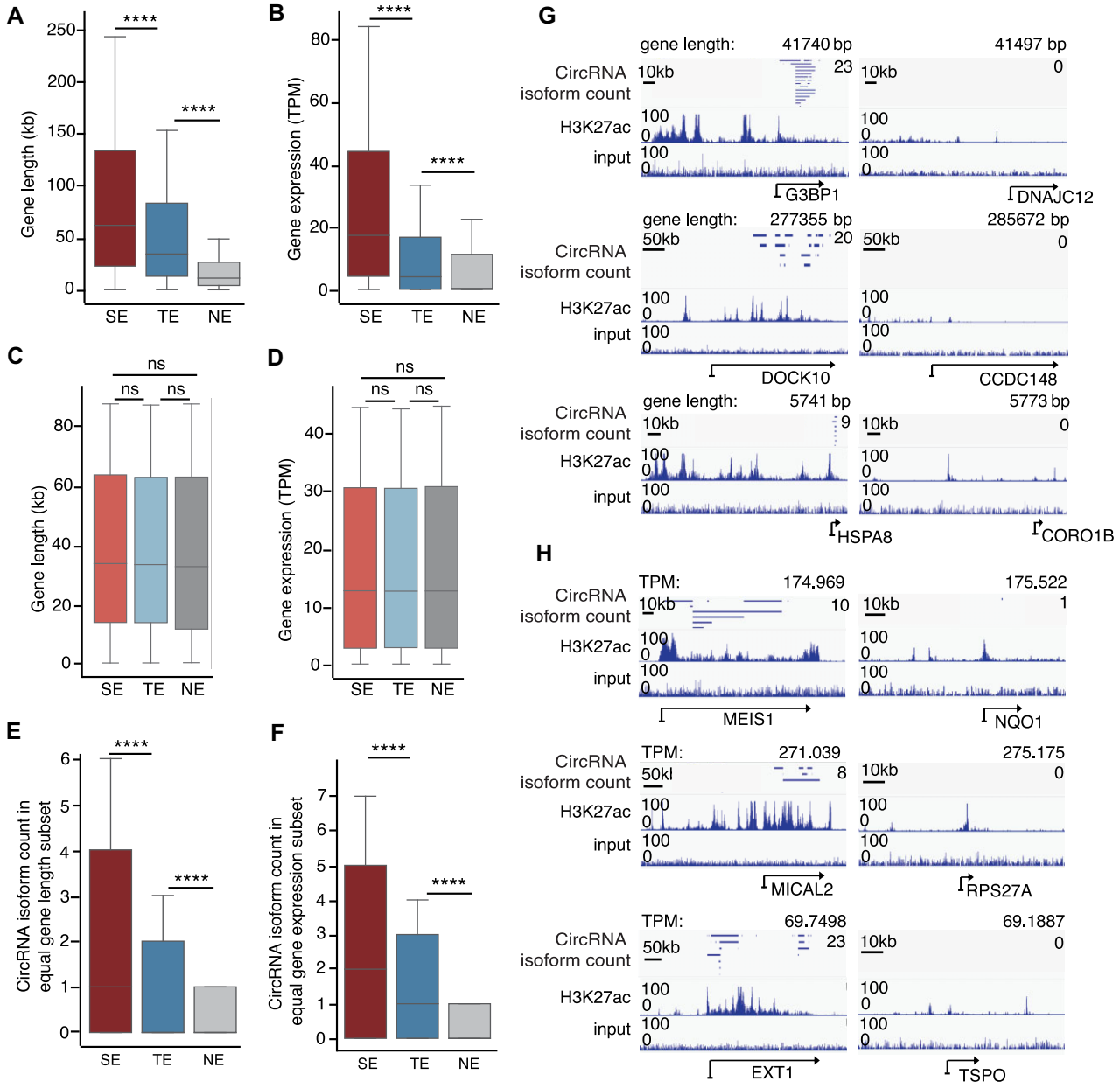


Figure 2. SE, TE, and NE genes expressed different circRNA isoform levels despite balancing their gene lengths or expression levels. Boxplot showing the gene length (**A**; kilobase pair, kb) and expression (**B**, TPM) of genes. Boxplot showing the gene length (**C**; kb) and expression (**D**; TPM) of genes in the subsets. The subsets were selected through an identical distribution from the SE, TE, and NE groups. Boxplot showing the circRNA isoform count of genes in the equal gene length subset (**E**) and in the equal gene expression subset (**F**). (G, H) H3K27ac ChIP-Seq signal in the sample genes. The line above the ChIP-Seq signal denotes a circRNA isoform. The relative positions to the genes show the BSJ position. The top right corner number denotes the number of circRNA isoforms detected. The numbers above the ChIP-Seq signal represent the gene length (**G**) or expression (**H**) of the sample gene. *P*-values determined by one-tailed Wilcoxon test. ****$P < .0001$. ns, not significant.

Collectively, all of these results revealed a complex association between epigenomic elements and circRNA biogenesis.

SE circRNAs show enriched BSJ-related RBP binding signal

Numerous RBPs have been reported to play a role in circRNA biogenesis [71–74]. For instance, NUDT21 has been reported to elevate the formation of circRNA by regulating its circularization [73]. In light of these findings, we collected CLIP-Seq data of various RBPs from CLIPdb database and subsequently identified their binding sites in the genomic regions [55]. Calculating the binding proportion of RBPs at BSJ and FSJ sites, we found that 101 RBPs demonstrated a significantly higher likelihood of binding to BSJ sites compared to FSJ sites. These RBPs were thus designated as BSJ-related RBPs. Among the BSJ-related RBPs, many of them have been previously reported to be involved in circRNA biogenesis or to perform functional roles in association with circRNAs [71–79]. Conversely, 55 RBPs exhibited a preference for binding to FSJ sites and were referred to as FSJ-related RBPs (Supplementary Fig. S6A and Supplementary Table S3). BSJ-related RBPs had

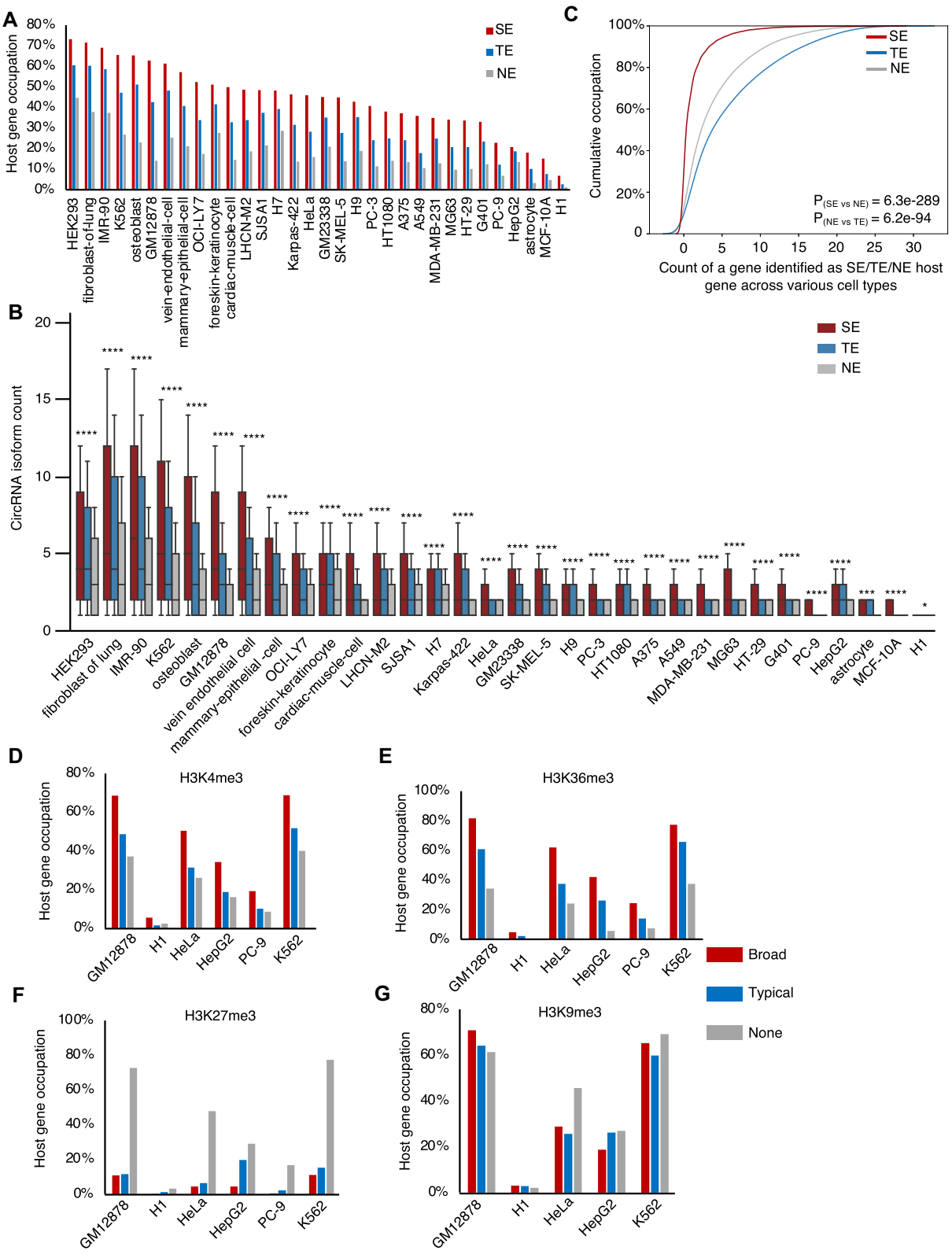


Figure 3. SEs promote circRNA splicing circularization diversity, and various histone modifications are associated with circRNA transcriptional activity across multiple cells. (**A**) Proportion of genes regulated by SE, TE, and NE (not regulated by enhancers) that express circRNAs in different cell lines and tissues. (**B**) Boxplot showing the circRNA isoform counts in different cell types. Significance among SE, TE, and NE genes as well as circRNA isoform count of each cell line is recorded in Supplementary Table S5. (**C**) Cumulative cure of the frequency of genes identified as SE/TE/NE host genes in different cell types. Proportion of host genes in genes covered by different histone modifications (**D**, H3K4me3; **E**, H3K36me3; **F**, H3K27me3; **G**, H3K9me3). Broad, genes regulated by the top 15%-width histone modifications. Typical, genes regulated by other width histone modifications. None, genes without histone modification. $*P < .05$, $***P < .001$, and $****P < .0001$. *P*-values determined by one-tailed Wilcoxon test.

significantly more signal binding peaks within BSJ regions than FSJ-related RBPs (Supplementary Fig. S6B). Intriguingly, in SE circRNAs, the number of binding peaks for BSJ-related RBPs at BSJ sites was also greater than that in TE and NE circRNAs. However, no such significant difference was observed for FSJ-related RBPs (Supplementary Fig. S6C and D). Similarly, SE circRNAs showed a higher binding signal density of BSJ-related RBPs compared to that in TE and NE circRNAs, while FSJ-related RBPs did not exhibit such differences (Supplementary Fig. S6E–J), demonstrating that BSJ-related RBPs play a key role in the high circRNA isoform count of SE genes.

In addition to RBPs, RNA polymerase II (pol II) speed has been reported to be highly correlated with splicing efficiency and circRNA biogenesis [80, 81]. To elucidate the relationship between splicing circularization diversity and pol II speed, we performed DRB/TT_{chem}-Seq analysis in HEK293 cell line. Our results revealed accelerated pol II elongation rate across host genes (Supplementary Fig. S7A). Moreover, SE host genes also exhibited markedly higher pol II speed compared to TE and NE host genes (Supplementary Fig. S7B), suggesting that pol II speed may also contribute to SE-mediated high splicing circularization diversity.

Reduced splicing circularization diversity in CIRSE genes links to cancer progression

Since most of circRNAs were identified through BSJ reads [82], a substantial number of lowly expressed circRNAs could not be detected robustly. It was observed that the majority of circRNAs were detected in a few samples (Supplementary Fig. S8A), leading to low sparsity of expression matrix. To address this challenge, KNN-smoothing was performed to impute the circRNA expression levels by observed expression levels of circRNAs and mRNAs [53] (Supplementary Fig. S8B). The circRNA imputation remarkably improved the frequency of the same circRNAs detected across various samples (Supplementary Fig. S8C). The proportion of the nonzero values increased from 1.81% to 10.37% (Supplementary Fig. S8D), effectively improving the data density.

To explore the role of circRNA splicing circularization diversity in tumorigenesis from a global perspective, we undertook a comparative analysis in pan-cancer cell lines. In sum, host genes exhibited a marked reduction in circRNA isoform counts in tumors compared to normal cells (Supplementary Fig. S8E). The circRNA expression level in tumors was also lower than that in normal cells [73] (Supplementary Fig. S8F). Compared with normal cells, a widespread and biased alteration of circRNA isoform count was observed in cancer cells, with 4020 (including 2736 GREAT genes) downregulated genes and 1683 (including 1251 GREAT genes) upregulated genes ($P < .01$) (Fig. 4A and Supplementary Fig. S9A). Here, the genes whose circRNA isoform count was significantly downregulated or upregulated in cancer cells were defined as isoform-decrease or isoform-increase genes, respectively. In general, there were more isoform-decrease genes than isoform-increase genes. This trend was maintained when we defined the genes using different cutoffs, and the number of isoform-decrease genes was 1.9–2.6 times that of isoform-increase genes (Fig. 4B). GO pathway enrichment analysis revealed that the isoform-decrease genes were enriched in cancer-associated pathways, primarily involving angiogenesis, migration, and proliferation (Fig. 4C), whereas the isoform-increase genes were associated with pathways such as nucleosome organization, mRNA processing, and mRNA modification (Supplementary Fig. S9B).

Given the close correlation between enhancers and circRNA biogenesis, we also systematically analyzed enhancer distribution information across 32 cell lines involving 17 cancer cells and 15 normal cells. Due to the cell specificity of SEs [29], only a few genes were simultaneously regulated by SEs across several cells. To better profile the relationship between the changes of enhancer intensity and circRNA splicing circularization diversity during cancer development, we selected genes that were regulated by SEs in at least two cancer or normal cells, with a significantly longer or shorter enhancer width in pan-cancer. The approach identified 421 genes regulated by SEs in at least two cancer cells with lengthened enhancers (enhancer-lengthen genes), and 358 genes regulated by SEs in at least two normal cells with shortened enhancers (enhancer-shorten genes) in cancer (Supplementary Fig. S9C). There were 98 isoform-decrease genes that overlapped with enhancer-shorten genes and 48 isoform-increase genes that overlapped with enhancer-lengthen genes. Significant enrichment was observed in both gene sets, revealing a consistent trend of enhancer width and the circRNA isoform count (Fig. 4D and E). Similar to the definition of the CIRSE gene and CIRSE circRNA, the overlapping of isoform-increase genes and enhancer-lengthen genes was defined as CircRNA Isoform Increasing for Lengthened Enhancer in cancer (CIILE) genes. CircRNAs derived from CIILE genes and whose expression level was upregulated in cancer were defined as CIILE circRNAs.

According to the changes in circRNA isoform count and enhancer width, we got four gene sets, including CIRSE and CIILE genes. Functional pathway enrichment analysis revealed that only the CIRSE gene set was significantly enriched in cancer pathways, such as positive regulation of angiogenesis and positive regulation of cell population proliferation, while the other gene sets were not enriched (Supplementary Fig. S9D). Therefore, we decided to focus on these 98 CIRSE genes in the follow-up study (Supplementary Table S6). To explore their tumor-suppressive potential, we integrated public full-length scRNA-Seq data, and quantified cumulative circRNA isoform count of CIRSE genes as PSS in each cell. Remarkably, cancer cells exhibited a significantly lower PSS compared to normal counterparts in breast and renal cancers (Fig. 4F and Supplementary Fig. S10), implicating CIRSE as a potential tumor-suppressive signature.

CIRSE pattern acts as tumor suppressors in lung cancer

To more precisely elucidate the role of the CIRSE gene and CIRSE circRNA in tumorigenesis, we conducted a comprehensive analysis in lung and renal cancers. In lung cancer, comparing A549 (lung cancer cell) with IMR-90 (lung normal cell) cells, we identified 3173 isoform-decrease genes and 538 isoform-increase genes (Fig. 5A). There were 470 SE genes that transformed into TE or NE genes (SE-to-TE/NE) with shortened enhancers and 339 genes that transformed into SE genes from TE or NE genes (TE/NE-to-SE) with lengthened enhancers (Fig. 5B). Consistent with pan-cancer findings, isoform-decrease and isoform-increase genes were enriched in SE-to-TE/NE and TE/NE-to-SE genes, with 234 and 39 overlapping genes, respectively (Fig. 5C

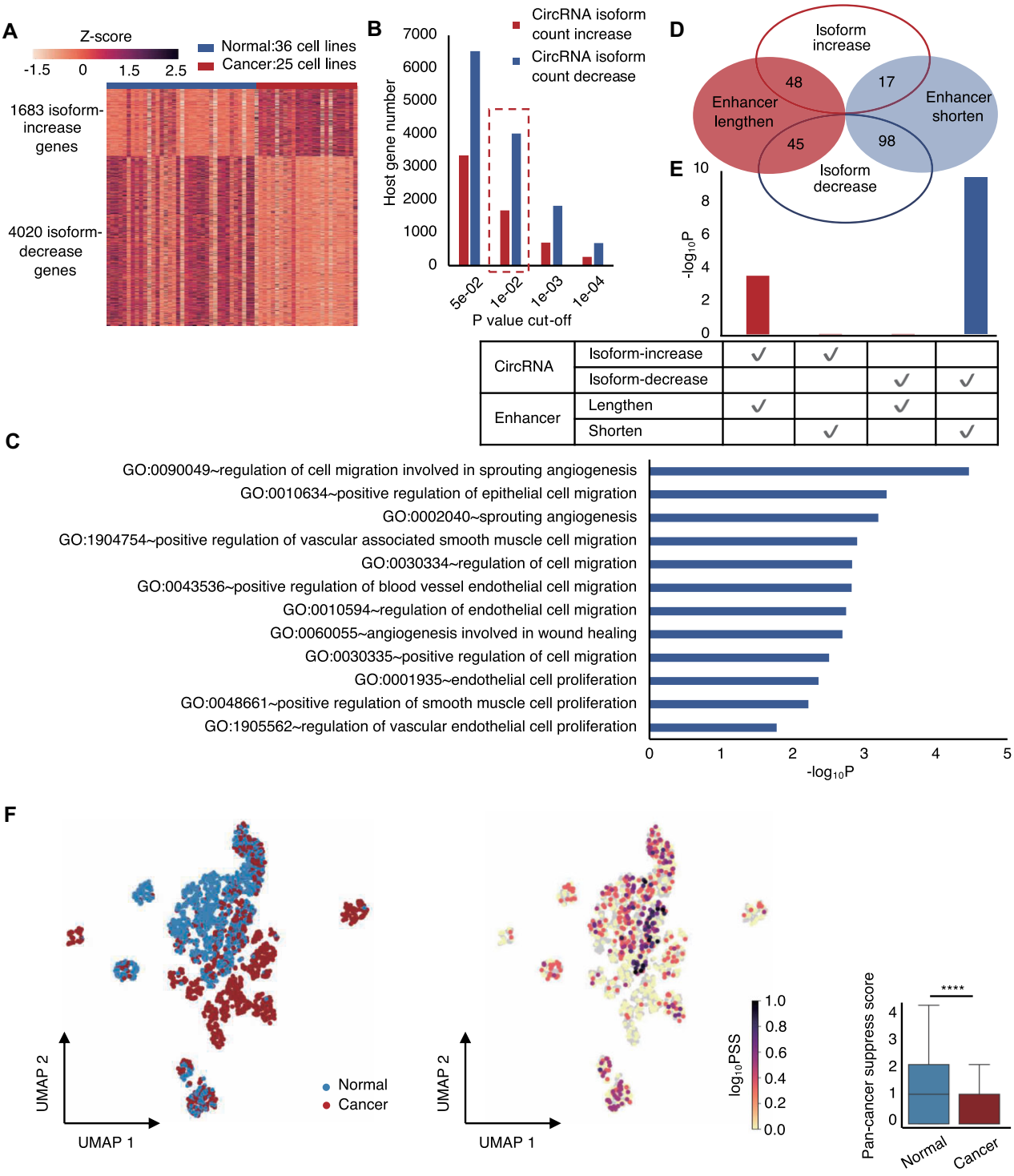


Figure 4. Systematic pan-cancer analysis revealed that CIRSE is associated with tumorigenesis. (**A**) Heatmap showing the circRNA isoform count of host genes whose circRNA isoform count increased (isoform-increase genes) or decreased (isoform-decrease genes) in cancer cells. The *z*-score represents the normalized circRNA isoform count. (**B**) The number of isoform-increase genes and isoform-decrease genes when defined with individual *P*-value cutoffs. (**C**) Enrichment level of the GO analysis of isoform-decrease genes. (**D**) Venn plot showing the overlapping of isoform-increase genes, isoform-decrease genes, enhancer-lengthen genes, and enhancer-shorten genes. (**E**) Enrichment level of isoform-increase genes and isoform-decrease genes in enhancer-lengthen genes and enhancer-shorten genes. *P*-values determined by Fisher's exact test. (**F**) UMAP plot of 2084 breast cancer single cells, colored by cell types (left) and log-normalized pan-cancer suppress scores (PSSs). All cells were divided into normal and tumor cells using CopyKAT. Boxplot (right) showing the PSS in normal and tumor cells (right). *P*-values determined by one-tailed Wilcoxon test. $^{****}P < .0001$.

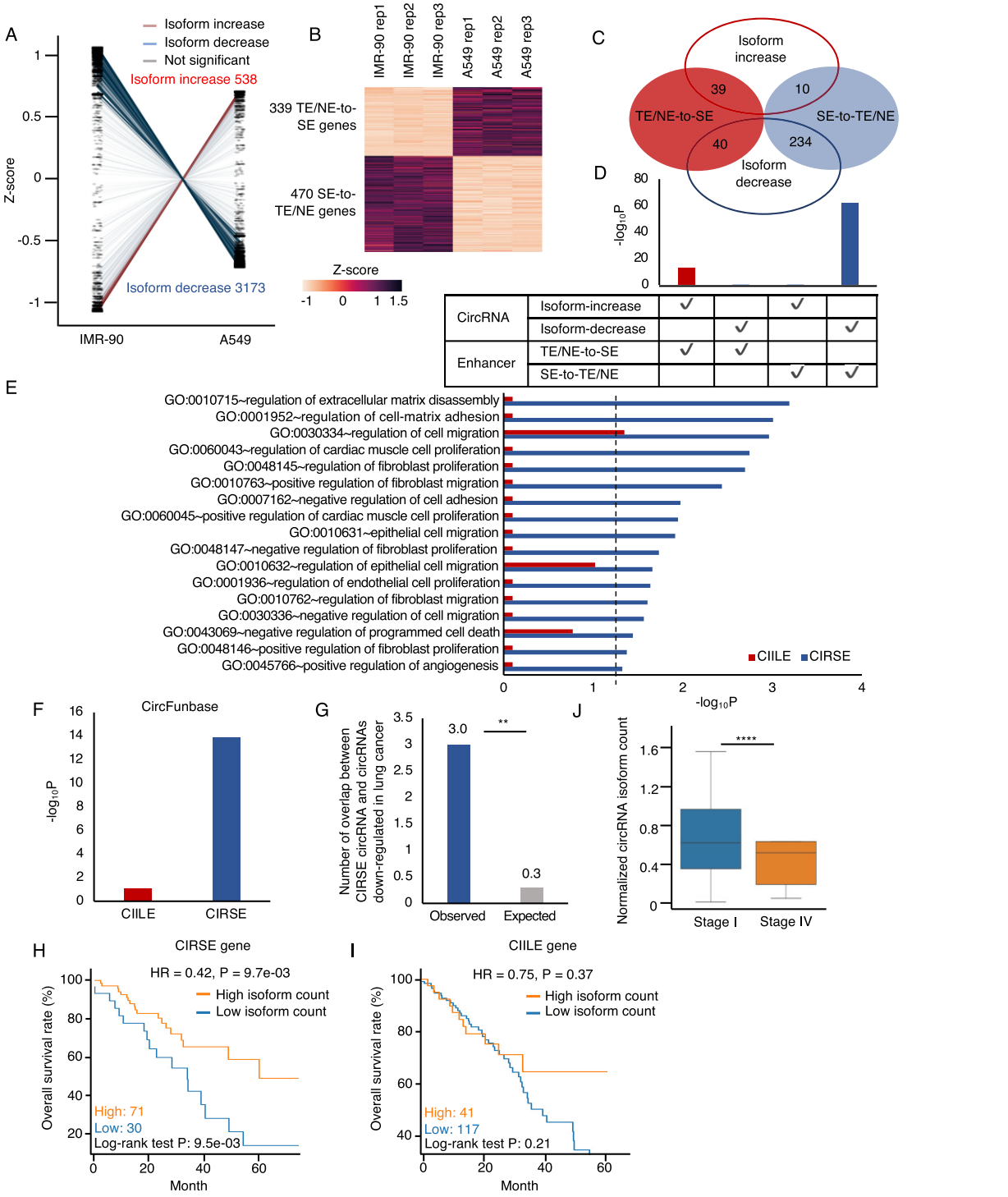


Figure 5. CIRSE shows a tumor-suppressing tendency in lung cancer. (**A**) Changes of circRNA isoform counts in A549 (lung cancer cell line) compared with IMR-90 (lung normal cell line). *Y*-axis represents *z*-score of the circRNA isoform count. Red lines, genes whose circRNA isoform count significantly increased in A549 (isoform-increase genes). Blue lines, genes whose circRNA isoform count significantly decreased in A549 (isoform-decrease genes). Gray lines, without significant difference. *Z*-score, normalized circRNA isoform count. (**B**) Heatmap showing the enhancer width (*z*-score) of genes that transformed between SE genes and other genes, with a significant ($P < .05$) change in enhancer width, in the IMR-90 and A549 cell lines. (**C**) Venn plot showing the overlapping of isoform-increase genes, isoform-decrease genes, TE/NE-to-SE genes, and SE-to-TE/NE genes. (**D**) Enrichment level of isoform-increase genes and isoform-decrease genes in TE/NE-to-SE genes and SE-to-TE/NE genes. (**E**) Enrichment level of GO analysis of host genes. Red box, CIILE genes; blue box, CIRSE gene; the black dotted line, *P*-value <.05. (**F**) Enrichment levels of CIRSE genes and CIILE genes in host genes of lung cancer-associated circRNA. (**G**) Observed and expected number of overlap between CIRSE circRNAs annotated in CircBase and validated circRNAs who were downregulated in lung cancer. (**H, I**) Hazard ratio (HR) and the Kaplan–Meier plot of the overall survival (OS) rate of LUAD patients with a high or low cumulative circRNA isoform count level of the lung cancer CIRSE genes (or CIILE genes) signatures in the TCGA data. High, high circRNA isoform count patients. Low, low circRNA isoform count patients. *P*-values of HR determined by Cox proportion hazards model. *P*-values of the survival curve determined by log-rank test. The number cutoff of patients was calculated by the lowest log-rank test *P*-value. (**J**) Boxplot showing the normalized circRNA isoform count of CIRSE genes for patients in different tumor stages. Clinical data were obtained from TCGA. $^{**}P < .01$ and $^{****}P < .0001$.

and D). GO pathway analysis further revealed the enrichment of CIRSE genes in the pathways related to angiogenesis, migration, and proliferation (Supplementary Table S7), while CIILE genes were not enriched in these pathways (Fig. 5E). For renal cancer, we compared G401 (renal cancer cell) with the HEK293 (renal normal cell) cell line and identified 1158 isoform-decrease and 2415 isoform-increase genes (Supplementary Fig. S11A). The count of genes with a transformation of SE-to-TE/NE and TE/NE-to-SE was 332 and 995, respectively (Supplementary Fig. S11B). Consistent with lung cancer, isoform-decrease genes were also enriched in SE-to-TE/NE genes (Supplementary Fig. S11C and D). Renal cancer CIRSE genes were enriched in cancer-related pathways, such as the regulation of angiogenesis, epithelial to mesenchymal transition, and cell proliferation (Supplementary Fig. S11E and Supplementary Table S8). The results indicated that CIRSE genes also exhibited cancer-associated functions in lung and renal cancers.

We further gathered known cancer-associated circRNAs and their host genes from databases. It was revealed that CIRSE genes were enriched in the host genes of cancer-associated circRNAs recorded in CircFunBase as well as the host genes whose circRNAs were downregulated in lung cancer from Lnc2Cancer [83, 84] (Fig. 5F and Supplementary Fig. S12A). Moreover, we collected 110 downregulated circRNAs in cancer, 18 of which were lung cancer-associated circRNAs. Lung cancer CIRSE circRNAs showed dual enrichment in these cancer-related circRNAs and lung cancer-associated circRNAs with statistical significance (Fig. 5G, Supplementary Fig. S12B, and Supplementary Table S9), suggesting the high tumor-suppressing potential of CIRSE circRNAs.

Given CIRSE genes established biomarker potential in pan-cancer, to better capture the clinical value of CIRSE genes in lung cancer, we evaluated the CIRSE circRNA isoform count in different lung adenocarcinoma (LUAD) patients. LUAD patients with lower lung CIRSE circRNA isoform counts demonstrated worse overall survival and disease-free survival, whereas the circRNA isoform count of CIILE genes did not show a significant distinguishing effect [85] (Fig. 5H and I, and Supplementary Fig. S12C and D). The HR also showed that a high CIRSE circRNA isoform count indicated lower risk, while patients with lower CIRSE circRNA isoforms had a poor prognosis (Fig. 5H and Supplementary Fig. S12C). Furthermore, LUAD patients with an advanced pathological state or distant tumor metastasis also showed a marked CIRSE circRNA isoform count reduction signature (Fig. 5J and Supplementary Fig. S12E).

Structural instability of CIRSE circRNAs in tumors

To investigate the features of cancer driver circRNAs beyond their BSJs, we performed Nanopore sequencing in IMR-90 and A549 cell lines. A total of 8720 and 7940 circRNAs were detected in Nanopore sequencing from IMR-90 and A549. The length of the identified circRNAs predominantly spanned 128–1024 bp (Fig. 6A), with similar circRNA lengths between normal and cancer cell lines (Supplementary Fig. S13). Consistent with the finding of pan-cancer, the genes in normal cells (IMR-90) tended to express more circRNA isoforms than those in cancer cells (A549) (Fig. 6B). The proportion of genes producing more circRNA isoforms was also higher in IMR-90 (normal cell line) than in A549 (cancer cell line) (Fig. 6C). This trend extended to circRNA expression levels, where circRNAs in IMR-90 also exhibited a higher abundance than A549 cell line (Fig. 6D). Notably, lung cancer CIRSE and CIILE genes identified through Illumina sequencing retained their characteristics within Nanopore sequencing. It was observed that the CIRSE genes expressed more circRNA isoforms in IMR-90 than in A549, whereas CIILE genes produced fewer circRNA isoforms (Fig. 6E), showing consistent Illumina and Nanopore sequencing results.

Given that mRNA with a stable secondary structure contributes to a long half-life [86], we considered stability to be a potential factor in affecting circRNA splicing circularization diversity. We detected stability using the MFE of circRNAs and found that CIRSE circRNAs contained lower MFE in IMR-90 (Fig. 6F), indicating a more stable structure. Comparing the circRNAs (both length and sequence similarity index were >0.7) derived from the same gene locus in two cells, we found that CIRSE circRNAs from IMR-90 showed a more stable structure (Fig. 6G, Supplementary Fig. S14A–C, and Supplementary Table S10), while CIILE circRNAs from A549 showed higher stability (Fig. 6H). The results were also confirmed by experimental validation (Fig. 6G and H, and Supplementary Fig. S14A–C), which was consistent with the difference in splicing circularization diversity between cancer and normal cells. Although SE genes tended to be longer than TE genes in IMR-90 (Fig. 2A), the length of SE circRNAs was comparable to that of TE circRNAs in both IMR-90 and A549, but they both exhibited greater length than NE circRNAs (Supplementary Fig. S15A and B). To fairly compare the stability of the SE, TE, and NE circRNAs, we regrouped the circRNAs to eliminate the length difference (Fig. 6I and J). It was observed that SE circRNAs had more stable structures than TE and NE circRNAs (Fig. 6K and L), suggesting a potential mechanism of the SEs mediating high splicing circularization diversity by regulating circRNA degradation.

CIRSE analysis for discovery of tumor-suppressive circRNAs

Based on the findings above, a tumor-suppressive circRNA (CIRSE circRNA) screening pipeline was established, and applied to the A549 and IMR-90: (i) selecting SEs that were converted to TEs or lost entirely, with significant shortening (SE-to-TE/NE), (ii) identifying host genes with downregulated circRNA splicing circularization diversity (isoform-decrease), and (iii) detecting circRNAs with decreased abundance in cancer (Supplementary Fig. S16). The pipeline identified 1551 candidate tumor-suppressing circRNAs in lung cancer. For experimental validation, we randomly selected 10 candidates, manually inspected the H3K27ac signal pattern to confirm enhancer shortening in each candidate, and searched the literature to confirm that the candidate circRNA has not been previously reported as a tumor-suppressing circRNA in lung cancer (Supplementary Table S11).

To detect the cellular distribution of these 10 candidates, we isolated the cytoplasmic and nuclear fractions of the cells and measured their expression levels using RT-qPCR. The circRNAs were all predominantly located in the cytoplasm (Fig. 7A). They were knocked down using siRNAs. Due to the low abundance of candidate circRNAs in A549 cell line, H1975 was used as a substitute. RT-qPCR was performed to verify the silence efficiency. Nine out of 10 circRNAs were significantly downregulated, while 1 candidate circRNA with low abundance failed to be knocked down significantly

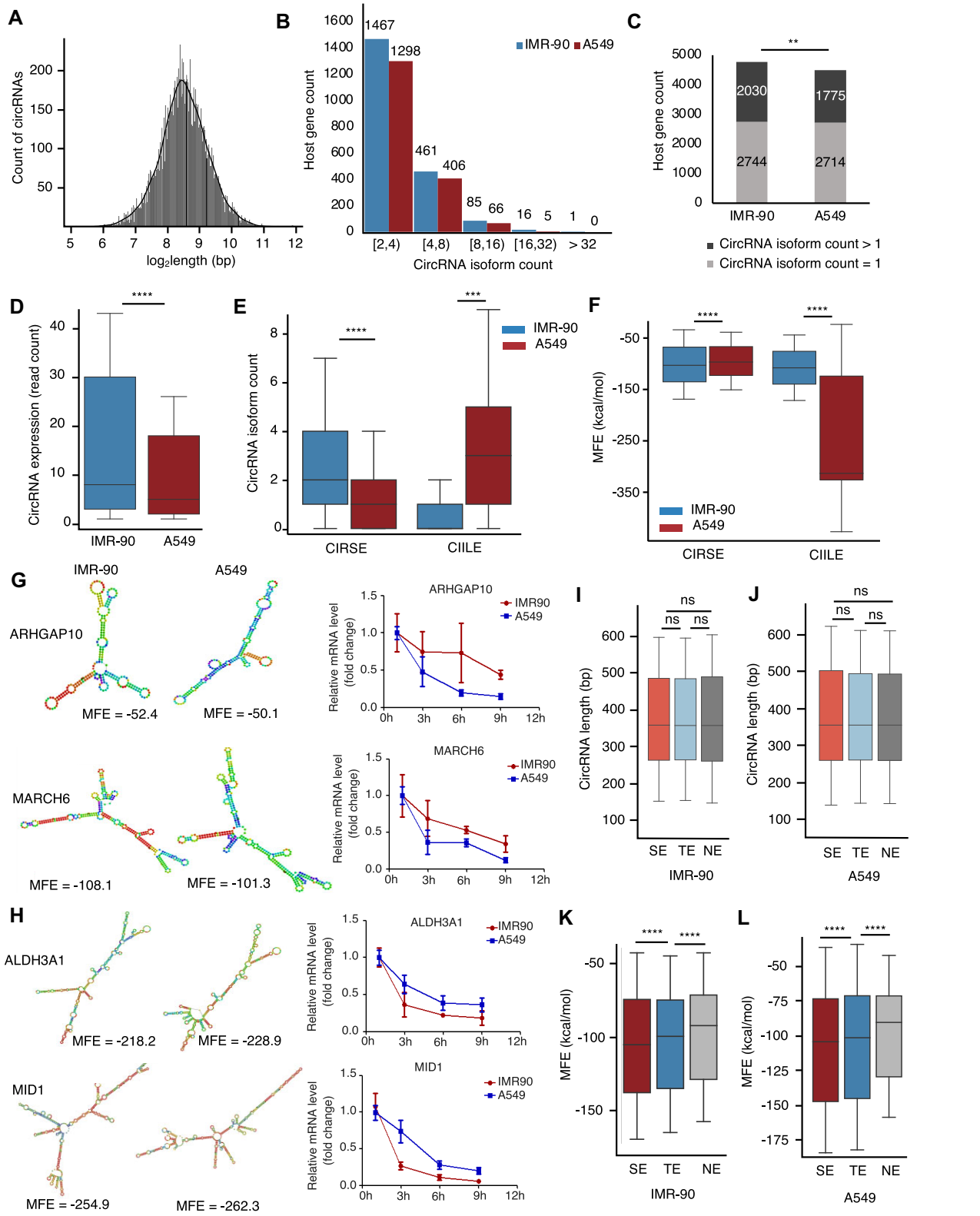


Figure 6. Structural stability plays a key role in regulating splicing circularization diversity. (**A**) Distribution of detected circRNA length. (**B**) Counts of host genes expressing different circRNA isoform counts in IMR-90 and A549 cells. (**C**) Bar plot showing the count of host genes with only one/more than one circRNA isoforms. *P*-values determined by *Z*-test. (**D**) Boxplot showing the circRNA expression levels in IMR-90 and A549. Read count, counts of total supported reads. Boxplot showing the circRNA isoform count (**E**) and MFE (**F**) of the gene sets in IMR-90 and A549. MFE, minimum free energy. Secondary structure of circRNAs from the same CIRSE (**G**) or CIILE (**H**) gene (left) and the stability of them assessed by RT-qPCR in IMR90 and A549 cell lines following actinomycin D treatment (right) in IMR-90 and A549. Unit of MFE, kcal/mol. Boxplot showing the circRNA length of the circRNA subset after equalizing the length in IMR-90 (**I**) and A549 (**J**). (**K**, **L**) Boxplot showing the MFE of SE/TE/NE circRNAs. The subsets were selected using an identical circRNA length distribution (I, J). *P*-values determined by one-tailed Wilcoxon test. $^{**}P < .01$, $^{***}P < .001$, and $^{****}P < .0001$.

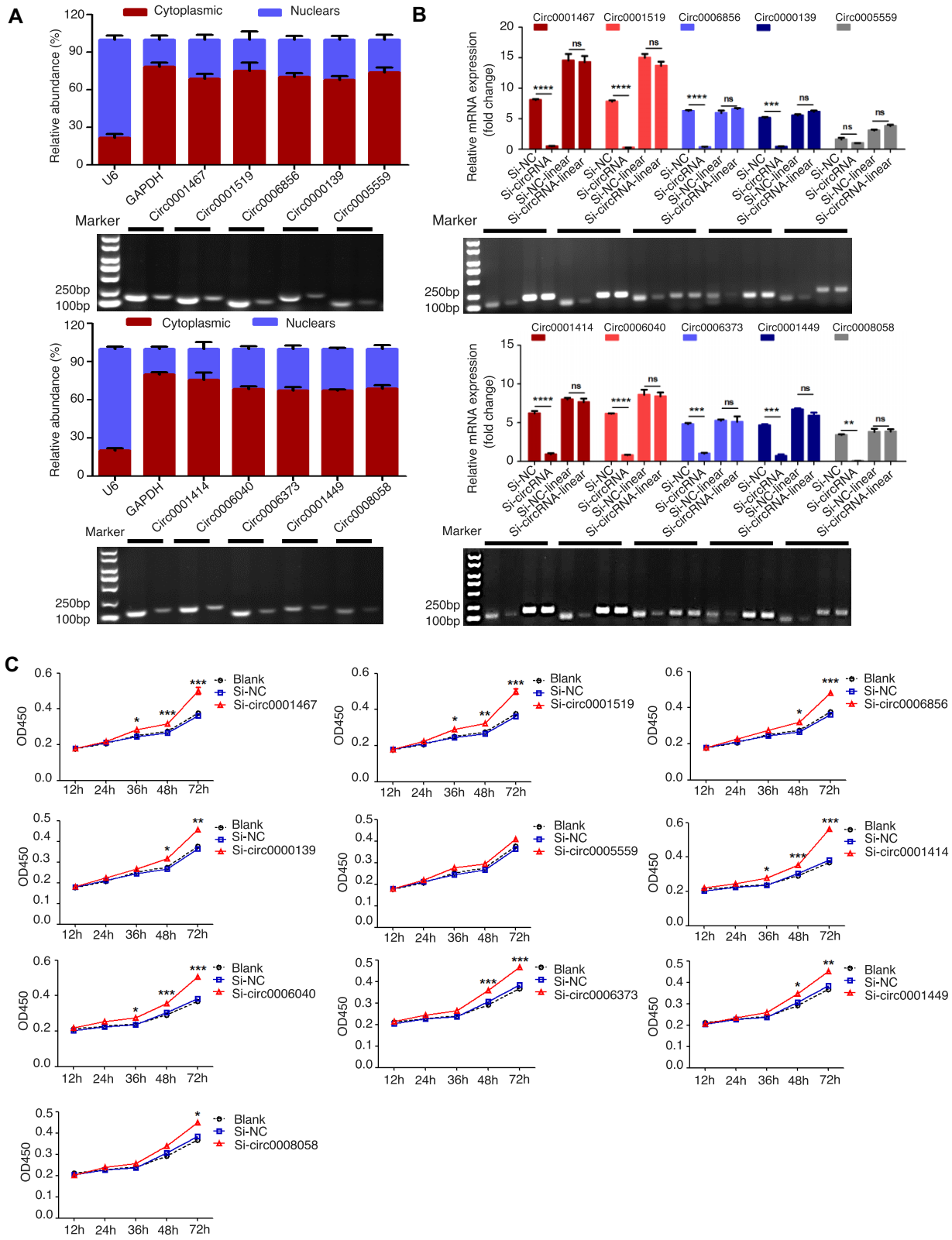


Figure 7. The expression distribution of different circRNAs and knockdown circRNAs significantly promoted their proliferation in H1975 cells. (**A**) Expression of circRNAs observed in H1975 cells. Analysis using agarose gel electrophoresis and nuclear–cytoplasmic fractionation assay revealed the predominant cytoplasmic localization of circRNA in H1975 cells. (**B**) Transfected H1975 with three different circRNA siRNAs with different targeting sequences. The combination of all three siRNAs exhibited the most efficient knockdown (KD) efficacy. The KD efficiency of various circRNAs by siRNA was assessed using RT-qPCR. Agarose gel electrophoresis analysis and RT-qPCR results revealed significantly higher expression levels in the control groups compared with the KD circRNA groups. (**C**) CCK-8 assay demonstrated enhanced cell proliferation in H1975 cells upon circRNA knockdown compared with the circRNA-NC group. The experiments were performed in triplicate. NC, nontarget control. The error bars represent the mean ± SD, $n = 3$. ns, not significant. $^{*}P < .05$, $^{**}P < .01$, $^{***}P < .001$, and $^{****}P < .0001$.

(Fig. 7B and Supplementary Fig. S17A). Their linear mRNA (control) remained unchanged, indicating that siRNA specifically targeted circRNA. A significant increase in cell proliferation (Fig. 7C), migration (Supplementary Fig. S17B and D), and invasion was observed in cells with successful circRNA knockdown (Supplementary Fig. S17C and E). In contrast, for the circRNAs used as negative controls, their knockdown had no impact on cell proliferation, migration, or invasion (Supplementary Fig. S18A–G). Furthermore, we found that the group with high silence efficiency showed a large increase in cell proliferation, migration, and invasion. For example, the circ0001467 group was found to induce greater cell proliferation in an early stage and exhibited a stronger migration and invasion ability than the other groups. This finding suggested a dose effect in these tumor-suppressing circRNAs. Overall, our results confirmed that CIRSE served as a useful signature for tumor suppressor discovery.

Discussion

Several studies reported that SE-associated circRNA modulated many biological processes, such as cell pyroptosis, cancer malignancy, and cardiac regeneration [87–89]. However, current research predominantly reported association between SEs and individual circRNAs, while systematical investigations into the relationship between epigenetic regulation and circRNA genesis remain scarce [90]. Here, we presented a landscape of circRNAs combined with enhancers across various cell types and revealed a widespread epigenetic regulation mechanism of splicing circularization diversity.

Notably, our study presents a novel analytical framework that extends beyond conventional expression-level analysis to investigate circRNA biology. Multiple circRNA isoforms are derived from the same gene locus [20], and they are the main contributors to splicing circularization diversity. With regard to the reduction of splicing circularization diversity, it not only shows a global decrease in the expression level of circRNAs from a single gene but also exhibits a disruption of circRNA diversity. To systematically assess the splicing circularization diversity of host gene, we introduced a novel metric termed "circRNA isoform count" to quantify the diversity of alternative splicing-derived circularization events. Furthermore, our analysis revealed a progressive decline in both circRNA diversification potential and transcriptional output across SE-, TE-, and NE-associated genes, correlating with diminishing enhancer regions.

In addition to circRNAs, we observed that the protein-coding genes regulated by SEs exhibited high diversity of linear transcripts using StringTie [91]. We found that host genes expressed more variety of transcript isoforms (Supplementary Fig. S19A). SE genes also exhibited a significantly higher count of different transcript isoforms than TE and NE genes (Supplementary Fig. S19B). Furthermore, we found that the transcripts of host genes experienced more alternative splicing events (Supplementary Fig. S19C), and the frequency of alternative splicing events in SE genes was also significantly higher than TE and NE genes (Supplementary Fig. S19D), revealing a high diversity of alternative splicing.

Some RBPs regulate circRNA biogenesis. QKI was reported to help circRNA formation through binding to intronic motifs, while DHX9 could inhibit circRNA genesis by resolving RNA duplexes [72, 92]. Here, we got 101 BSJ-related RPBs, and found they tended to bind SE circRNAs more than TE and NE circRNAs. Many BSJ-related RBPs have been reported to be involved in circRNA biogenesis [71–74]. Among the top 10 BSJ-related RPBs ranked by significance score, most were closely associated with circRNA, including NUDT21 and KHSRP, which have been reported to enhance circRNA expression [73–79]. Targeting these RBPs may provide a strategy to make influence on cancer progression via circRNAs.

Notably, SE circRNAs demonstrated enhanced structural stability relative to TE and NE circRNAs, which correlated with extended molecular longevity and increased retention of diverse circRNA isoforms. Several studies have demonstrated that increased stability of circRNAs may contribute to their functional persistence. For example, circHIPK3 and circFOXO3 could promote tumor progression by acting as microRNA (miRNA) sponges and regulated oncogenic pathways [93, 94]. This suggested that enhanced circRNA stability may be a key factor in the regulation of splicing circularization diversity.

RCM is a kind of sequence that has been reported to contribute in circRNA biogenesis [95], the frequency of which in the genome may be another factor that regulates the changes of splicing circularization diversity during tumorigenesis. Thus, we explored the relationship between RCMs and flanking introns of SE circRNAs. However, unlike the consistent trend of occupation of host genes or circRNA isoform count of SE, TE, and NE genes (Fig. 3A and B), the ranking of proportion of SE, TE, and NE circRNAs with RCMs varied in different cell types (Supplementary Fig. S20A). Moreover, we calculated the proportion of circRNAs with RCMs among circRNAs derived from the same host genes. No specific pattern was observed in how this proportion varied across different enhancers, which contrasted with the consistent observation that SE enhanced splicing circularization diversity across multiple cell lines (Supplementary Fig. S20B).

Unlike CIRSE, CIILE genes showed enrichment in epigenomic-related pathways, such as chromatin remodeling and histone modification. This suggests that they may perform housekeeping functions in biological processes. In this study, we also found some isoform-decrease genes whose enhancers lengthened or did not change in cancer, which suggests that there exist other regulatory mechanisms for circRNA splicing circularization diversity.

Current detection approaches for circRNAs predominantly rely on quantifying reads mapped to BSJ regions [41, 44, 96]. However, this methodology fails to account for reads mapped to distal regions of circRNAs, which consequently remain undetected. These limitations fundamentally compromise circRNA detection sensitivity, resulting in sparse circRNA expression matrices that ultimately lead to inaccurate downstream analyses. A similar challenge exists in scRNA-Seq data interpretation, where the KNN-smoothing approach has been validated as an effective solution for imputing technical zero values [97–99]. In this study, we employ this established methodology to identify nearest neighbor samples through integration of both mRNA and circRNA expression profiles, thereby enabling systematic imputation of zero-count entries to enhance the efficacy of subsequent analytical processes.

The functional annotation of circRNAs represents a significant challenge in current research. Host genes have been extensively employed as primary references for predicting circRNA functions [41, 100]. This approach is supported by substantial evidence from multiple studies demonstrating that circRNAs frequently exhibit strong functional associations with

their host genes [15, 16]. For example, circRNAs could regulate host genes as miRNA sponges or affect their mRNA stability [101]. CircRNA biogenesis could compete with pre-mRNA splicing, thereby affecting the expression of host genes [15]. Multiple circRNAs derived from cancer driver genes (oncogene or tumor suppressor gene) have been proven to take part in tumorigenesis, such as *TP53*, *PTEN*, *MYC*, and *EGFR* [102–105].

In summary, this study presents an alternative paradigm for circRNA investigation that extends beyond conventional expression-level analyses. Through integrated multi-omics approaches, we established a comprehensive association between epigenetic modifications and circRNA biogenesis patterns. Furthermore, we uncovered a novel regulatory mechanism through which SEs orchestrate diversity in circRNA splicing and circularization processes. Building upon these mechanistic insights, we developed an innovative screening pipeline that improves the identification method of tumor-suppressive circRNAs, thereby providing a robust framework for functional noncoding RNA discovery in oncological research.

Acknowledgements

Author contributions: D.Z. and S.H. conceptualized and designed the study. S.H., Y.L. and L.B. conducted a data analysis. S.M. and Y.H. conceptualized and validated the experiment. D.Z., S.H., and Y.L. wrote and edited the paper. C.W., Z.R., C.Y., H.C, Y.L., Y.S., Y.Z., Z.W. and B.S. gave important suggestions. All authors read and approved the final manuscript. Fangqing Zhao kindly provided guidance for our work.

Supplementary data

Supplementary data is available at NAR online.

Conflict of interest

None declared.

Funding

National Natural Science Foundation of China (NSFC, 32270603, 82170512); Fundamental Research Funds for the Central Universities (BMU2021YJ057); Major Project of Guangzhou National Laboratory (Grant no. GZNL2023A02009 to S.M.); General Program of Beijing Natural Science Foundation of China (5242010).

Data availability

The raw data of Nanopore RNA-Seq were deposited to Gene Expression Omnibus (GEO), with accession number GSE262357. Public datasets rRNA-depleted RNA-Seq, CLIP-Seq, and ChIP-Seq of various histone modification data were downloaded from GEO and the ENCODE project website (Supplementary Table S2). Files containing SE and TE elements of various samples were downloaded from the public database SEdb 2.0 (http://www.licpathway.net/sedb). The GREAT gene list was used for peak annotation and was downloaded from the website of the GREAT bioinformatics tool (http://great.stanford.edu/public/html/). The names and sequences of the validated circRNAs were obtained from CircBase. Validated cancer-associated circRNAs were downloaded from CircFunBase and Lnc2Cancer. CircRNA expression array of LUAD patients was obtained from the public database CircNet. The clinical data of LUAD patients were downloaded from TCGA (https://www.cancer.gov/ccg/research/genome-sequencing/tcga/history).

References

1. Wu W, Zhao F, Zhang J. circAtlas 3.0: a gateway to 3 million curated vertebrate circular RNAs based on a standardized nomenclature scheme. *Nucleic Acids Res* 2024;**52**:D52–60. https://doi.org/10.1093/nar/gkad770
2. Dong R, Ma XK, Li GW *et al.* CIRCpedia v2: an updated database for comprehensive circular RNA annotation and expression comparison. *Genomics Proteomics Bioinformatics* 2018;**16**:226–33. https://doi.org/10.1016/j.gpb.2018.08.001
3. Cape& B, Swain A, Nicolis S *et al.* Circular transcripts of the testis-determining gene *Sry* in adult mouse testis. *Cell* 1993;**3**:1019–30.
4. Vo JN, Cieslik M, Zhang Y *et al.* The landscape of circular RNA in cancer. *Cell* 2019;**176**:869–81. https://doi.org/10.1016/j.cell.2018.12.021
5. Salzman J, Gawad C, Wang PL *et al.* Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. *PLoS One* 2012;**7**:e30733. https://doi.org/10.1371/journal.pone.0030733
6. Salzman J, Chen RE, Olsen MN *et al.* Cell-type specific features of circular RNA expression. *PLoS Genet* 2013;**9**:e1003777. https://doi.org/10.1371/annotation/f782282b-eefa-4c8d-985c-b1484e845855
7. Teng M, Guo J, Xu X *et al.* Circular RMST cooperates with lineage-driving transcription factors to govern neuroendocrine transdifferentiation. *Cancer Cell* 2025;**43**:891–904. https://doi.org/10.1016/j.ccell.2025.03.027
8. Niu M, Wang C, Chen Y *et al.* Identification, characterization and expression analysis of circRNA encoded by SARS-CoV-1 and SARS-CoV-2. *Brief Bioinform* 2024;**25**:bbad537. https://doi.org/10.1093/bib/bbad537
9. Zeng X, Zhong Y, Lin W *et al.* Predicting disease-associated circular RNAs using deep forests combined with positive-unlabeled learning methods. *Brief Bioinform* 2020;**21**:1425–36. https://doi.org/10.1093/bib/bbz080
10. Hansen TB, Jensen TI, Clausen BH *et al.* Natural RNA circles function as efficient microRNA sponges. *Nature* 2013;**495**:384–8. https://doi.org/10.1038/nature11993
11. Qu S, Yang X, Li X *et al.* Circular RNA: a new star of noncoding RNAs. *Cancer Lett* 2015;**365**:141–8. https://doi.org/10.1016/j.canlet.2015.06.003
12. Kristensen LS, Andersen MS, Stagsted LVW *et al.* The biogenesis, biology and characterization of circular RNAs. *Nat Rev Genet* 2019;**20**:675–91. https://doi.org/10.1038/s41576-019-0158-7
13. Abdelmohsen K, Panda AC, Munk R *et al.* Identification of HuR target circular RNAs uncovers suppression of PABPN1 translation by CircPABPN1. *RNA Biology* 2017;**14**:361–9. https://doi.org/10.1080/15476286.2017.1279788
14. Pamudurti NR, Bartok O, Jens M *et al.* Translation of circRNAs. *Mol Cell* 2017;**66**:9–21. https://doi.org/10.1016/j.molcel.2017.02.021
15. Ashwal-Fluss R, Meyer M, Pamudurti NR *et al.* CircRNA biogenesis competes with pre-mRNA splicing. *Mol Cell* 2014;**56**:55–66. https://doi.org/10.1016/j.molcel.2014.08.019
16. Xu X, Zhang J, Tian Y *et al.* CircRNA inhibits DNA damage repair by interacting with host gene. *Mol Cancer* 2020;**19**:128. https://doi.org/10.1186/s12943-020-01246-x
17. Hanniford D, Ulloa-Morales A, Karz A *et al.* Epigenetic silencing of CDR1as drives IGF2BP3-mediated melanoma invasion and

metastasis. *Cancer Cell* 2020;**37**:55–70. https://doi.org/10.1016/j.ccell.2019.12.007
18. Chen LL. The expanding regulatory mechanisms and cellular functions of circular RNAs. *Nat Rev Mol Cell Biol* 2020;**21**:475–90. https://doi.org/10.1038/s41580-020-0243-y
19. Wang S, Zhang K, Tan S *et al.* Circular RNAs in body fluids as cancer biomarkers: the new frontier of liquid biopsies. *Mol Cancer* 2021;**20**:13. https://doi.org/10.1186/s12943-020-01298-z
20. Gao Y, Wang J, Zheng Y *et al.* Comprehensive identification of internal structure and alternative splicing events in circular RNAs. *Nat Commun* 2016;**7**:12060. https://doi.org/10.1038/ncomms12060
21. Zhang XO, Dong R, Zhang Y *et al.* Diverse alternative back-splicing and alternative splicing landscape of circular RNAs. *Genome Res*. 2016;**26**:1277–87. https://doi.org/10.1101/gr.202895.115
22. Chen I, Chen CY, Chuang TJ. Biogenesis, identification, and function of exonic circular RNAs. *WIREs RNA* 2015;**6**:563–79. https://doi.org/10.1002/wrna.1294
23. Chen LL, Yang L. Regulation of circRNA biogenesis. *RNA Biol* 2015;**12**:381–8. https://doi.org/10.1080/15476286.2015.1020271
24. Jiang Y, Jiang YY, Lin DC. Super-enhancer-mediated core regulatory circuitry in human cancer. *Comput Struct Biotechnol J* 2021;**19**:2790–5. https://doi.org/10.1016/j.csbj.2021.05.006
25. Heinz S, Romanoski CE, Benner C *et al.* The selection and function of cell type-specific enhancers. *Nat Rev Mol Cell Biol* 2015;**16**:144–54. https://doi.org/10.1038/nrm3949
26. Jiang H, Shukla A, Wang X *et al.* Erratum: role for Dpy-30 in ES cell-fate specification by regulation of h3k4 methylation within bivalent domains. *Cell* 2011;**144**:825. https://doi.org/10.1016/j.cell.2011.02.036
27. Sengupta S, George RE. Super-enhancer-driven transcriptional dependencies in cancer. *Trends Cancer* 2017;**3**:269–81. https://doi.org/10.1016/j.trecan.2017.03.006
28. Ma H, Qu J, Pang Z *et al.* Super-enhancer omics in stem cell. *Mol Cancer* 2024;**23**: 153. https://doi.org/10.1186/s12943-024-02066-z
29. Whyte WA, Orlando DA, Hnisz D *et al.* Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* 2013;**153**:307–19. https://doi.org/10.1016/j.cell.2013.03.035
30. Creyghton MP, Cheng AW, Welstead GG *et al.* Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci USA* 2010;**107**:21931–6. https://doi.org/10.1073/pnas.1016071107
31. Jiang YY, Lin DC, Mayakonda A *et al.* Targeting super-enhancer-associated oncogenes in oesophageal squamous cell carcinoma. *Gut* 2017;**66**:1358–68. https://doi.org/10.1136/gutjnl-2016-311818
32. Ooi WF, Xing M, Xu C *et al.* Epigenomic profiling of primary gastric adenocarcinoma reveals super-enhancer heterogeneity. *Nat Commun* 2016;**7**:12983. https://doi.org/10.1038/ncomms12983
33. Zhou B, Wang L, Zhang S *et al.* INO80 governs superenhancer-mediated oncogenic transcription and tumor growth in melanoma. *Genes Dev* 2016; **30**:1440–53. https://doi.org/10.1101/gad.277178.115.
34. Oldridge DA, Wood AC, Weichert-Leahey N *et al.* Genetic predisposition to neuroblastoma mediated by a LMO1 super-enhancer polymorphism. *Nature* 2015;**528**:418–21. https://doi.org/10.1038/nature15540
35. Chapuy B, McKeown MR, Lin CY *et al.* Discovery and characterization of super-enhancer-associated dependencies in diffuse large B cell lymphoma. *Cancer Cell* 2013;**24**:777–90. https://doi.org/10.1016/j.ccr.2013.11.00
36. Hnisz D, Abraham BJ, Lee TI *et al.* XSuper-enhancers in the control of cell identity and disease. *Cell* 2013;**155**:934–47. https://doi.org/10.1016/j.cell.2013.09.053
37. Suzuki HI, Young RA, Sharp PA. Super-enhancer-mediated RNA processing revealed by integrative microRNA network analysis. *Cell* 2017;**168**:1000–14. https://doi.org/10.1016/j.cell.2017.02.015
38. Li Y, Ge X, Peng F *et al.* Exaggerated false positives by popular differential expression methods when analyzing human population samples. *Genome Biol* 2022;**23**:79. https://doi.org/10.1186/s13059-022-02648-4
39. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010;**26**:139–40. https://doi.org/10.1093/bioinformatics/btp616
40. Glažar P, Papavasileiou P, Rajewsky N. CircBase: a database for circular RNAs. *RNA* 2014;**20**:1666–70. https://doi.org/10.1261/rna.043687.113
41. Zhang J, Chen S, Yang J *et al.* Accurate quantification of circular RNAs identifies extensive circular isoform switching events. *Nat Commun* 2020;**11**:90.
42. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* 2015;**12**:357–60. https://doi.org/10.1038/nmeth.3317
43. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 2009;**25**:1754–60. https://doi.org/10.1093/bioinformatics/btp324
44. Ma XK, Wang MR, Liu CX *et al.* CIRCexplorer3: a CLEAR pipeline for direct comparison of circular and linear RNA expression. *Genomics Proteomics Bioinformatics* 2019;**17**:511–21. https://doi.org/10.1016/j.gpb.2019.11.004
45. Cunningham F, Allen JE, Allen J *et al.* Ensembl 2022. *Nucleic Acids Res* 2022;**50**:D988–95. https://doi.org/10.1093/nar/gkab1049
46. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;**9**:357–9. https://doi.org/10.1038/nmeth.1923
47. Chen K, Xi Y, Pan X *et al.* DANPOS: dynamic analysis of nucleosome position and occupancy by sequencing. *Genome Res*. 2013;**23**:341–51. https://doi.org/10.1101/gr.142067.112
48. Robinson JT, Thorvaldsdottir H, Turner D *et al.* igv.Js: an embeddable JavaScript implementation of the Integrative Genomics Viewer (IGV). *Bioinformatics* 2023;**39**:btac830. https://doi.org/10.1093/bioinformatics/btac830
49. Robinson JT, Thorvaldsdóttir H, Winckler W *et al.* Integrative genomics viewer. *Nat Biotechnol* 2011;**29**:24-6.
50. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings Bioinf* 2013;**14**:178–92. https://doi.org/10.1093/bib/bbs017
51. Lovén J, Hoke HA, Lin CY *et al.* Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* 2013;**153**:320–34. https://doi.org/10.1016/j.cell.2013.03.036
52. McLean CY, Bristor D, Hiller M *et al.* GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol* 2010;**28**:495–501. https://doi.org/10.1038/nbt.1630
53. Wagner F, Yan Y, Yanai I. K-nearest neighbor smoothing for high-throughput single-cell RNA-seq data. bioRxiv, https://doi.org/10.1101/217737, 9 April 2018, preprint: not peer reviewed.
54. Kelly MR, Wisniewska K, Regner MJ *et al.* A multi-omic dissection of super-enhancer driven oncogenic gene expression programs in ovarian cancer. *Nat Commun* 2022;**13**:4247. https://doi.org/10.1038/s41467-022-31919-8
55. Yang YCT, Di C, Hu B *et al.* CLIPdb: a CLIP-seq database for protein-RNA interactions. *BMC Genomics* 2015;**16**:51. https://doi.org/10.1186/s12864-015-1273-2

56. Zhang J, Hou L, Zuo Z *et al.* Comprehensive profiling of circular RNAs with nanopore sequencing and CIRI-long. *Nat Biotechnol* 2021;**7**:893. https://doi.org/10.1038/s41587-021-00934-3
57. Lorenz R, Bernhart SH, Höner Zu Siederdissen C *et al.* ViennaRNA Package 2.0. *Algorithms Mol Biol* 2011;**6**:26.
58. Cock PJA, Antao T, Chang JT *et al.* Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* 2009;**25**:1422–3. https://doi.org/10.1093/bioinformatics/btp163
59. Dobin A, Davis CA, Schlesinger F *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 2013;**29**:15–21. https://doi.org/10.1093/bioinformatics/bts635
60. Li H, Handsaker B, Wysoker A *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* 2009;**25**:2078–9. https://doi.org/10.1093/bioinformatics/btp352
61. Ramírez F, Dündar F, Diehl S *et al.* DeepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* 2014;**42**:W187–91. https://doi.org/10.1093/nar/gku365
62. Gregersen LH, Mitter R, Svejstrup JQ. Using TTchem-seq for profiling nascent transcription and measuring transcript elongation. *Nat Protoc* 2020;**15**:604–27. https://doi.org/10.1038/s41596-019-0262-3
63. Quinlan AR. BEDTools: the Swiss-Army tool for genome feature analysis. *Curr Protoc Bioinformatics* 2014;**2014**:11.12.1–34.
64. Mirdita M, Steinegger M, Breitwieser F *et al.* Fast and sensitive taxonomic assignment to metagenomic contigs. *Bioinformatics* 2021;**37**:3029–31. https://doi.org/10.1093/bioinformatics/btab184
65. Long F, Li L, Xie C *et al.* Intergenic CircRNA Circ_0007379 inhibits colorectal cancer progression by modulating miR-320a biogenesis in a KSRP-dependent manner. *Int J Biol Sci.* 2023;**19**:3781–803. https://doi.org/10.7150/ijbs.85063
66. Chen X, Chen RX, Wei WS *et al.* PRMT5 circular RNA promotes metastasis of urothelial carcinoma of the bladder through sponging miR-30c to induce epithelial–mesenchymal transition. *Clin Cancer Res* 2018;**24**:6319–30. https://doi.org/10.1158/1078-0432.CCR-18-1270
67. Wang C, Liu WR, Tan S *et al.* Characterization of distinct circular RNA signatures in solid tumors. *Mol Cancer* 2022;**21**:63. https://doi.org/10.1186/s12943-022-01546-4
68. Li X, Yang L, Chen LL. The biogenesis, functions, and challenges of circular RNAs. *Mol Cell* 2018;**71**:428–42. https://doi.org/10.1016/j.molcel.2018.06.034
69. Cocquerelle C, Mascrez B, Hétuin D *et al.* Mis-splicing yields circular RNA molecules. *FASEB J* 1993;**7**:155–60. https://doi.org/10.1096/fasebj.7.1.7678559
70. Wang Y, Song C, Zhao J *et al.* SEdb 2.0: a comprehensive super-enhancer database of human and mouse. *Nucleic Acids Res* 2023;**51**:D280–90. https://doi.org/10.1093/nar/gkac968
71. Errichelli L, Dini Modigliani S, Laneve P *et al.* FUS affects circular RNA expression in murine embryonic stem cell-derived motor neurons. *Nat Commun* 2017;**8**:14741. https://doi.org/10.1038/ncomms14741
72. Conn SJ, Pillman KA, Toubia J *et al.* The RNA binding protein quaking regulates formation of circRNAs. *Cell* 2015;**160**:1125–34. https://doi.org/10.1016/j.cell.2015.02.014
73. Li X, Ding J, Wang X *et al.* NUDT21 regulates circRNA cyclization and ceRNA crosstalk in hepatocellular carcinoma. *Oncogene* 2020;**39**:891–904. https://doi.org/10.1038/s41388-019-1030-0
74. Okholm TLH, Sathe S, Park SS *et al.* Transcriptome-wide profiles of circular RNA and RNA-binding protein interactions reveal effects on circular RNA biogenesis and cancer pathway expression. *Genome Med* 2020;**12**:112. https://doi.org/10.1186/s13073-020-00812-8
75. Conn VM, Gabryelska M, Marri S *et al.* Srrm4 expands the repertoire of circular rnas by regulating microexon inclusion. *Cells* 2020;**9**:2488. https://doi.org/10.3390/cells9112488
76. Ma S, Xu Y, Qin X *et al.* RUNX1, FUS, and ELAVL1-induced circPTPN22 promote gastric cancer cell proliferation, migration, and invasion through miR-6788-5p/PAK1 axis-mediated autophagy. *Cell Mol Biol Lett* 2024;**29**:95. https://doi.org/10.1186/s11658-024-00610-9
77. Jiang Y, Wang Z, Ying C *et al.* FMR1/circCHAF1A/miR-211-5p/HOXC8 feedback loop regulates proliferation and tumorigenesis via MDM2-dependent p53 signaling in GSCs. *Oncogene* 2021;**40**:4094–110. https://doi.org/10.1038/s41388-021-01833-2
78. Zheng Z, Zeng X, Zhu Y *et al.* CircPPAP2B controls metastasis of clear cell renal cell carcinoma via HNRNPC-dependent alternative splicing and targeting the miR-182-5p/CYP1B1 axis. *Mol Cancer* 2024;**23**:4. https://doi.org/10.1186/s12943-023-01912-w
79. Chen S, Cao X, Zhang J *et al.* circVAMP3 Drives CAPRIN1 phase separation and inhibits hepatocellular carcinoma by suppressing c-myc translation. *Adv Sci* 2022;**9**:e2103817.
80. Zhang Y, Xue W, Li X *et al.* The biogenesis of nascent circular RNAs. *Cell Rep* 2016;**15**:611–24. https://doi.org/10.1016/j.celrep.2016.03.058
81. Moehle EA, Braberg H, Krogan NJ *et al.* Adventures in time and space: splicing efficiency and RNA polymerase II elongation rate. *RNA Biology* 2014;**11**:313–9. https://doi.org/10.4161/rna.28646
82. Vromman M, Anckaert J, Bortoluzzi S *et al.* Large-scale benchmarking of circRNA detection tools reveals large differences in sensitivity but not in precision. *Nat Methods* 2023;**20**:1159–69. https://doi.org/10.1038/s41592-023-01944-6
83. Meng X, Hu D, Zhang P *et al.* CircFunBase: a database for functional circular RNAs. *Database* 2019;**2019**:baz003. https://doi.org/10.1093/database/baz003
84. Gao Y, Shang S, Guo S *et al.* Lnc2Cancer 3.0: an updated resource for experimentally supported lncRNA/circRNA cancer associations and web tools based on RNA-seq and scRNA-seq data. *Nucleic Acids Res* 2021;**49**:D1251–8. https://doi.org/10.1093/nar/gkaa1006
85. Chen Y, Yao L, Tang Y *et al.* CircNet 2.0: an updated database for exploring circular RNA regulatory networks in cancers. *Nucleic Acids Res* 2022;**50**:D93–D101. https://doi.org/10.1093/nar/gkab1036
86. Mauger DM, Joseph Cabral B, Presnyak V *et al.* mRNA structure regulates protein expression through changes in functional half-life. *Proc Natl Acad Sci USA* 2019;**116**:24075–83. https://doi.org/10.1073/pnas.1908052116
87. Liu H, Jiang Y, Shi R *et al.* Super enhancer-associated circRNA-circLrch3 regulates hypoxia-induced pulmonary arterial smooth muscle cells pyroptosis by formation of R-loop with host gene. *Int J Biol Macromol* 2024;**268**:130853. https://doi.org/10.1016/j.ijbiomac.2024.130853
88. Chen Y, Du C, Tang J *et al.* Super-enhancer-associated circPVT1 promotes malignancy of hepatocellular carcinoma via YBX1-mediated RRM2 activation. *Cancer Lett* 2025;**611**:217395. https://doi.org/10.1016/j.canlet.2024.217395
89. Huang S, Li X, Zheng H *et al.* Loss of super-enhancer-regulated circRNA nfix induces cardiac regeneration after myocardial infarction in adult mice. *Circulation* 2019;**139**:2857–76. https://doi.org/10.1161/CIRCULATIONAHA.118.038361
90. Li XC, Tang ZD, Peng L *et al.* Integrative epigenomic analysis of transcriptional regulation of human circRNAs. *Front Genet* 2021;**11**:590672. https://doi.org/10.3389/fgene.2020.590672
91. Pertea M, Pertea GM, Antonescu CM *et al.* StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol* 2015;**33**:290–5. https://doi.org/10.1038/nbt.3122
92. Aktaş T, Ilik IA, Maticzka D *et al.* DHX9 suppresses RNA processing defects originating from the Alu invasion of the human genome. *Nature* 2017;**544**:115–9. https://doi.org/10.1038/nature21715

93. Kong Z, Wan X, Lu Y *et al.* Circular RNA circFOXO3 promotes prostate cancer progression through sponging miR-29a-3p. *J Cel Mol Med* 2020;**24**:799–813. https://doi.org/10.1111/jcmm.14791
94. Zheng Q, Bao C, Guo W *et al.* Circular RNA profiling reveals an abundant circHIPK3 that regulates cell growth by sponging multiple miRNAs. *Nat Commun* 2016;**7**:11215. https://doi.org/10.1038/ncomms11215
95. Ivanov A, Memczak S, Wyler E *et al.* Analysis of intron sequences reveals hallmarks of circular RNA biogenesis in animals. *Cell Rep* 2015;**10**:170–7. https://doi.org/10.1016/j.celrep.2014.12.019
96. Zeng X, Lin W, Guo M *et al.* A comprehensive overview and evaluation of circular RNA detection tools. *PLoS Comput Biol* 2017;**13**:e1005420. https://doi.org/10.1371/journal.pcbi.1005420
97. Wang M, Gan J, Han C *et al.* Imputation methods for scRNA sequencing data. *Appl Sci* 2022;**12**:10684.
98. Gong W, Kwak IY, Pota P *et al.* DrImpute: imputing dropout events in single cell RNA sequencing data. *BMC Bioinformatics* 2018;**19**:220. https://doi.org/10.1186/s12859-018-2226-y
99. Xu C, Cai L, Gao J. An efficient scRNA-seq dropout imputation method using graph attention network. *BMC Bioinformatics* 2021;**22**:582. https://doi.org/10.1186/s12859-021-04493-x
100. Zhang M, Xu K, Fu L *et al.* Revealing epigenetic factors of circRNA expression by machine learning in various cellular contexts. *iScience* 2020;**23**:101842. https://doi.org/10.1016/j.isci.2020.101842
101. Wang X, Li H, Lu Y *et al.* Regulatory effects of circular RNAs on host genes in human cancer. *Front Oncol*. 2021;**10**:58616. https://doi.org/10.3389/fonc.2020.586163
102. Yan S, Wei H, Li Q *et al.* CircTP53 promotes colorectal cancer by acting as a miR-876-3p sponge to increase cyclin-dependent kinase-like 3 expression. *Cell Signal* 2021;**78**:109845. https://doi.org/10.1016/j.cellsig.2020.109845
103. Fu P, Lin L, Zhou H *et al.* Circular rna circegfr regulates tumor progression via the mir-106a-5p/ddx5 axis in colorectal cancer. *Braz J Med Biol Res* 2021;**54**:e10940. https://doi.org/10.1590/1414-431x2020e10940
104. Jin C, Dong D, Yang Z *et al.* CircMYC regulates glycolysis and cell proliferation in melanoma. *Cell Biochem Biophys* 2020;**78**:77–88. https://doi.org/10.1007/s12013-019-00895-0
105. Li C, Li X. circPTEN suppresses colorectal cancer progression through regulating PTEN/AKT pathway. *Mol Ther Nucleic Acids* 2021;**26**:1418–32. https://doi.org/10.1016/j.omtn.2021.05.018

Received: February 10, 2025. **Revised:** April 27, 2025. **Editorial Decision:** May 24, 2025. **Accepted:** May 28, 2025