



The Journal of Biomedical Research

Genetic variation of *circHIBADH* enhances prostate cancer risk through regulating HNRNPA1-related RNA splicing

Cheng Yifei, Shi Rongjie, Ben Shuai, Chen Silu, Li Shuwei, Xin Junyi, Wang Meilin, Cheng Gong

Cite this article as:

Cheng Yifei, Shi Rongjie, Ben Shuai, Chen Silu, Li Shuwei, Xin Junyi, Wang Meilin, Cheng Gong. Genetic variation of *circHIBADH* enhances prostate cancer risk through regulating HNRNPA1-related RNA splicing[J]. *Journal of Biomedical Research*, 2024, 38(4): 358–368. doi: 10.7555/JBR.38.20240030

View online: <https://doi.org/10.7555/JBR.38.20240030>

Articles you may be interested in

Genetic variation of the PSCA gene (rs2294008) is not associated with the risk of prostate cancer

The Journal of Biomedical Research. 2017, 31(3): 226 <https://doi.org/10.7555/JBR.31.20160072>

Association of long non-coding RNA HOTAIR and MALAT1 variants with cervical cancer risk in Han Chinese women

The Journal of Biomedical Research. 2019, 33(5): 308 <https://doi.org/10.7555/JBR.33.20180096>

A novel long non-coding RNA NFIA-AS1 is down-regulated in gastric cancer and inhibits proliferation of gastric cancer cells

The Journal of Biomedical Research. 2019, 33(6): 371 <https://doi.org/10.7555/JBR.33.20190015>

PUM1 represses CDKN1B translation and contributes to prostate cancer progression

The Journal of Biomedical Research. 2021, 35(5): 371 <https://doi.org/10.7555/JBR.35.20210067>

Microarray expression profile and functional analysis of circular RNAs in choroidal neovascularization

The Journal of Biomedical Research. 2020, 34(1): 67 <https://doi.org/10.7555/JBR.33.20190063>

RNA-seq analysis identified hormone-related genes associated with prognosis of triple negative breast cancer

The Journal of Biomedical Research. 2020, 34(2): 129 <https://doi.org/10.7555/JBR.34.20190111>



Genetic variation of *circHIBADH* enhances prostate cancer risk through regulating HNRNPA1-related RNA splicing

Yifei Cheng^{1,△}, Rongjie Shi^{2,△}, Shuai Ben¹, Silu Chen¹, Shuwei Li¹, Junyi Xin³, Meilin Wang^{4,5,#}, Gong Cheng^{2,#,✉}

¹Department of Environmental Genomics, Jiangsu Key Laboratory of Cancer Biomarkers, Prevention and Treatment, Collaborative Innovation Center for Cancer Personalized Medicine, Nanjing Medical University, Nanjing, Jiangsu 211166, China;

²Department of Urology, the First Affiliated Hospital of Nanjing Medical University, Nanjing, Jiangsu 210029, China;

³Department of Bioinformatics, School of Biomedical Engineering and Informatics, Nanjing Medical University, Nanjing, Jiangsu 211166, China;

⁴Jiangsu Cancer Hospital, Jiangsu Institute of Cancer Research, the Affiliated Cancer Hospital of Nanjing Medical University, Nanjing, Jiangsu 210009, China;

⁵The Affiliated Suzhou Hospital of Nanjing Medical University, Suzhou Municipal Hospital, Gusu School, Nanjing Medical University, Suzhou, Jiangsu 215002, China.

Abstract

The current study aimed to investigate associations of circRNAs and related genetic variants with the risk of prostate cancer (PCa) as well as to elucidate biological mechanisms underlying the associations. We first compared expression levels of circRNAs between 25 paired PCa and adjacent normal tissues to identify risk-associated circRNAs by using the MiOncoCirc database. We then used logistic regression models to evaluate associations between genetic variants in candidate circRNAs and PCa risk among 4 662 prostate cancer patients and 3 114 healthy controls, and identified *circHIBADH* rs11973492 T>C as a significant risk-associated variant (odds ratio = 1.20, 95% confidence interval: 1.08–1.34, $P = 7.06 \times 10^{-4}$) in a dominant genetic model, which altered the secondary structure of the corresponding RNA chain. In the *in silico* analysis, we found that *circHIBADH* sponged and silenced 21 RNA-binding proteins (RBPs) enriched in the RNA splicing pathway, among which HNRNPA1 was identified and validated as a hub RBP using an external RNA-sequencing data as well as the in-house (four tissue samples) and publicly available single-cell transcriptomes. Additionally, we demonstrated that HNRNPA1 influenced hallmarks including MYC target, DNA repair, and E2F target signaling pathways, thereby promoting carcinogenesis. In conclusion, genetic variants in *circHIBADH* may act as sponges and inhibitors of RNA splicing-associated RBPs including HNRNPA1, playing an oncogenic role in PCa.

Keywords: genetic variants, prostate cancer, circRNA, RNA-binding protein, RNA splicing, sing-cell RNA sequencing

[△]These authors contributed equally to this work.

[#]These authors jointly supervised this work.

✉Corresponding author: Gong Cheng, Department of Urology, the First Affiliated Hospital of Nanjing Medical University, 300 Guangzhou Road, Nanjing, Jiangsu 210029, China. E-mail: gcheng@njmu.edu.cn.

Received: 02 February 2024; Revised: 24 April 2024; Accepted: 30

April 2024; Published online: 29 May 2024

CLC number: R737.25, Document code: A

The authors reported no conflict of interests.

This is an open access article under the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited.

Introduction

Prostate cancer (PCa) is the second most common malignancy and the fifth leading cause of cancer-related deaths in men^[1], with a good prognosis and a 5-year survival rate of over 98%^[2]. However, patients with advanced PCa have a much poorer prognosis due to the lack of effective therapies^[3]. It was reported that a high heritability (approximately 57%) and an increasing risk were observed in men with a family history of PCa, implying a genetic component in the etiology of this disease^[4]. Genome-wide association studies have identified more than 200 loci associated with PCa risk^[5], but approximately 90% of these susceptibility loci are in non-coding regions. Moreover, the mechanisms underlying these associations remain unclear^[6].

Gene expression is regulated by numerous non-coding RNAs, including long non-coding RNAs, microRNAs, and circular RNAs (circRNAs), with a particular attention paid to the role of circRNAs in cancers. It is known that circRNAs are mainly formed by back-splicing events that splice an exon into an upstream one^[7]. The circRNAs may function in biological processes *via* sponging microRNAs^[8], interacting with RNA-binding proteins (RBPs) as sponges^[9], scaffolds^[10], recruiters^[11], protein function enhancers^[12], and translating proteins^[13]. Through these functions, circRNAs play either oncogenic or tumor suppressor roles in various malignancies, including cancers of the bladder^[14], stomach^[15], breast^[16], and prostate^[17–18]. Using an exome capture RNA sequencing protocol, Vo *et al*^[19] detected 160 120 circRNAs in more than 2 000 cancer samples and found a general downregulation of these circRNAs in tumor tissues, which is attributed to the regulation of cellular proliferation. Additionally, they established the MiOncoCirc database to provide a resource of circRNAs for developing diagnostic targets in cancer. Moreover, Chen *et al*^[17] discovered in several cohorts that circRNA expression levels were correlated with PCa progression. However, few studies have focused on the effects of genetic variants of circRNAs on the risk of PCa.

Therefore, the current study comprehensively investigated associations of genetic variants in circRNAs with PCa risk and survival. Furthermore, we evaluated how these genetic variants affect circRNAs and their functioning in binding with RBPs and regulating subsequent biological processes.

Subjects and methods

Study population and data processing

The current study included 4 662 PCa patients and 3 114 controls from the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial with approval from the PLCO consortium (PLCO-84), which are available from NIH-CDAS (<https://cdas.cancer.gov/datasets/plco/20/>). The baseline characteristics of the study population from the PLCO trial have been described previously^[20]. In addition, we used the MiOncoCirc database that includes 25 pairs of matched PCa tumor and adjacent normal tissues^[19] by downloading the expression matrices of circRNAs and their parental genes from the website (<https://mioncocirc.github.io/>). Furthermore, we used RNA expression profiles and clinical information from The Cancer Genome Atlas (TCGA) prostate adenocarcinoma (PRAD) as well as GSE94767 and GSE183019 datasets from the Genomic Data Commons (<https://portal.gdc.cancer.gov/>) and Gene Expression Omnibus (<https://www.ncbi.nlm.nih.gov/gds>), respectively. The mRNA expression data from TCGA were transformed to values in transcripts per million before analysis. In the validation, we used additional single-cell RNA-sequencing (scRNA-seq) data of four PCa tissues from the First Affiliated Hospital of Nanjing Medical University, along with previously published single-cell transcriptome dataset (<https://www.prostate-cellatlas.org/>)^[21].

Selection of circRNAs and single nucleotide polymorphisms (SNPs)

By comparing the expression levels of 25 paired tumor and adjacent normal tissues, we selected the risk-associated circRNAs based on the following inclusion criteria: (1) detected in circBase; (2) located on autosomal chromosomes; (3) a false discovery rate (FDR) < 0.01; (4) $|\log_2(\text{fold change [FC]})| > 1$; and (5) a call rate in normal tissues > 60%.

For SNP selection, we first extracted SNPs within the circRNA regions from the PLCO trial for further filtering. The quality control was carried out by using the following inclusion criteria: a minor allele frequency ≥ 0.05 , $P_{\text{Hardy-Weinberg equilibrium}} \geq 0.05$, and a call rate > 95%. Subsequently, the *in silico* analyses were performed for functional annotation. SNPs with a RegulomeDB (v2.1, <http://regulome.stanford.edu/>) ranking ≤ 5 or a 3DSNP score (v1.0, <https://www.omic.tech/3dsnp/>) ≥ 10 were considered functional SNPs. Tagging SNPs were obtained by conducting a

pairwise linkage disequilibrium analysis. Finally, the risk-associated analyses were performed to identify SNPs that were significantly associated with PCa risk by using a dominant genetic model.

Functional analysis of circRNAs

Based on the overview of circRNAs provided by the Cancer-Specific CircRNA Database (CSCD, <http://gb.whu.edu.cn/CSCD/>), we focused on the potential RBPs that bind to candidate circRNAs using both the circAtlas database (<http://circatlas.biols.ac.cn/>) and CSCD. The potential functions of predicted RBPs were annotated through over-representative analyses using the Kyoto Encyclopedia of Genes and Genomes pathway and gene ontology biological processes datasets. Furthermore, we applied String (<https://string-db.org/>) and Cystoscope software (version 3.7.1) to construct a protein-protein interaction (PPI) network of these RBPs.

Identification and validation of hub RBPs and key biological processes

Based on the TCGA PRAD dataset, we performed differential expression (DE) analyses to identify differences in the predicted RBPs between normal and tumor tissues. By integrating PPI network and DE analyses, we determined RBPs that were located at the hub of the network, consistently expressed among samples, and significantly differentially expressed between tumor and normal tissues as the hub RBPs.

Through correlation analysis of the hub RBPs with candidate circRNAs in MiOncoCirc, we validated the regulatory potential and correlations of circRNA to RBPs. Furthermore, we performed the DE analysis in two external bulk and two single-cell transcriptomes to verify the expression patterns of the hub RBPs. We then assessed the differences in gene expression levels (\log_2FC) between the top and the bottom 1/4 TCGA samples according to the expression levels of the hub RBPs for the Gene-Set Enrichment Analysis (GSEA) based on CancerSEA database (<http://biocc.hrbmu.edu.cn/CancerSEA/>) to discover hallmarks that might be affected by these RBPs in the tumorigenic process. Finally, we validated the activation of key biological processes by comparing the AUCell score (an algorithm that uses the area under the curve [AUC] to evaluate the enrichment of a particular gene set among the expressed genes of each cell) of the corresponding gene set of single cells between groups.

scRNA-seq

The fresh tumor tissues from 10 patients were stored in the GEXSCOPE® Tissue Preservation

Solution (Singleron, Nanjing, China) and transported on ice, which were then washed with Gibco Hanks Balanced Salt Solution (Cat. #14175095, Thermo Fisher Scientific-CN, Shanghai, China), minced and digested with GEXSCOPE® Tissue Dissociation Solution (Singleron), centrifuged and resuspended. GEXSCOPE® red blood cell lysis buffer (Singleron) was used to remove the red blood cells. Single-cell suspensions were converted to barcoded scRNA-seq libraries using the Chromium Single Cell Library, Gel Bead & Multiplex Kit (10x Genomics, Pleasanton, CA, USA), following the manufacturer's instructions. Finally, libraries were prepared using the 10x Genomics Library Kits (10x Genomics) and sequenced on Illumina Nova6000 (Illumina, San Diego, CA, USA) with a paired-end 150 bp reading strategy.

Interpretation of scRNA-seq data

Raw reads were processed to generate gene expression profiles using Cell Ranger (v.3.0.2, 10x Genomics). Reads from the 10x library were mapped to GRCh38 with ensemble version 92 gene annotation. Cells with gene counts less than 500, unique molecular identifier counts less than 1 000, or more than 15% expression levels of mitochondrial genes were filtered out, with genes expressed in less than 10 cells also excluded from the analysis.

The retained cells were then subjected to the "Seurat" program (v.4.3.0) in R for downstream analyses. After normalizing gene expression, we identified the top 2 000 variable genes that were scaled and regressed out for the expression percentage of mitochondrial genes and used for principal component analysis. The batch effect was assessed using the R package "Harmony" (v.0.1.1). We uniformed the Manifold Approximation and Projection dimensionality reduction on the Harmony dimensions to obtain two-dimensional representations for data visualization, and then computed the shared nearest neighbor to identify clusters by using Harmony dimensions. Epitheliums were labeled by a high expression of *CDH1/KRT8* and a low expression of *PTPRC*, and extracted for subsequent analyses. The annotation of epitheliums was carried out according to the epithelial subtypes identified by Henry *et al* based on the identified clusters^[22].

Statistical analysis

We used R software (version 4.3.1) and PLINK (version 1.90) for statistical analyses and performed the Wilcoxon test for all differential expression analyses. The differences in baseline and clinical

characteristics between PCa cases and the controls in the PLCO were assessed using the Student's *t*-test and χ^2 test for continuous variables and categorical variables, respectively. The goodness-of-fit test was applied to compute the Hardy-Weinberg equilibrium based on the allele frequencies of the controls. Univariate and multivariate logistic regression models were used to perform the risk-associated analyses. Differences in the expression levels of individual mRNAs and circRNAs were compared using the Wilcoxon signed-rank test and Wilcoxon rank-sum test for paired and unpaired samples, respectively. Over-representative analyses and GSEA were carried out with the R package "*clusterProfiler*" (v.4.2.2). Correlation analyses were conducted using Spearman's correlation method. AUC scores were computed using the R package "*AUCCell*" (v.1.16.0).

Results

circRNAs and SNPs selection

The selection process of circRNAs and SNPs is shown in [Fig. 1A](#). Through differential expression analyses, 128 risk-associated circRNAs were identified ([Fig. 1B](#) and [Supplementary Table 1](#) [available online]). In total, 13 197 SNPs were included for quality control, functional annotation, and linkage disequilibrium analysis, resulting in the retention of 105 risk-associated SNPs for further analyses. SNPs were then analyzed for their associations with PCa risk. As shown in [Supplementary Table 2](#) (available online) and [Fig. 1D](#), only rs11973492 in *circHIBADH* (hsa_circ_0006773 in circBase), with an adjusted odds ratio (OR) = 1.20, 95% confidence interval (CI): 1.08–1.34, and FDR = 4.30×10^{-2} , was significantly associated with PCa risk in a dominant genetic model.

Associations between rs11973492 and PCa risk

The associations of rs11973492 with PCa risk in the co-dominant, additive, dominant, and recessive genetic models are shown in [Table 1](#). The proportions of TT, TC, and CC were 45.18%, 45.24%, and 9.58% in the cases, respectively, and 49.92%, 40.95%, and 9.13% in the controls, respectively. The TC genotype was significantly associated with an increased risk of PCa, compared with the TT genotype (adjusted OR = 1.23, 95% CI: 1.10–1.37, $P = 3.07 \times 10^{-5}$). A significantly higher risk of PCa was associated with TC/CC genotypes (adjusted OR = 1.20, 95% CI: 1.08–1.34, $P = 7.06 \times 10^{-4}$) in the dominant model but not in the recessive model (adjusted OR = 1.04, 95% CI: 0.91–1.18, $P = 6.03 \times 10^{-1}$).

We further conducted stratification analyses for the association between rs11973492 and PCa by using the dominant model for clinical characteristics. As shown in [Table 2](#), TC/CC genotypes were risk factors in PCa with Gleason score ≤ 6 (adjusted OR = 1.24, 95% CI: 1.10–1.41, $P = 6.53 \times 10^{-4}$), Gleason score = 7 (adjusted OR = 1.27, 95% CI: 1.10–1.47, $P = 1.25 \times 10^{-3}$), PSA < 10 ng/mL (adjusted OR = 1.22, 95% CI: 1.09–1.37, $P = 6.88 \times 10^{-4}$), $10 \leq$ PSA < 20 ng/mL (adjusted OR = 1.33, 95% CI: 1.10–1.62, $P = 3.64 \times 10^{-3}$), and stage I/II (adjusted OR = 1.23, 95% CI: 1.10–1.37, $P = 2.57 \times 10^{-4}$). No significant risk was identified in high-risk PCa subgroups, such as Gleason ≥ 8 , PSA > 20, and stage III/IV subgroups, compared with the controls.

Annotation of SNPs and expression levels of parental genes

An overview of rs11973492 is presented in [Supplementary Fig. 1](#) (available online). Chromatin immunoprecipitation sequencing data revealed that rs11973492 was located within the binding site of CCCTC-binding factors in prostate epithelial cells ([Supplementary Fig. 1A](#)). The variant of rs11973492 might also affect the secondary structures of the corresponding RNAs ([Supplementary Fig. 1B](#) and [1C](#)). Additionally, several genes (*EVX1*, *HOTTIP*, *HOXA-AS4*, *HOXA10*, etc.) were observed to interact with rs11973492 through three-dimensional chromatin loops in PCa cell lines ([Supplementary Fig. 1D](#)).

Significantly higher expression levels of *circHIBADH* were detected in 25 adjacent normal tissues, compared with paired tumor tissues in the MiOncoCirc database ($P = 1.79 \times 10^{-7}$, [Fig. 1C](#)), while the expression level of *HIBADH*, which is the parental gene of *circHIBADH*, did not differ between adjacent normal and tumor tissues ($P = 7.71 \times 10^{-1}$, [Fig. 1E](#)). In the TCGA and GSE94767 datasets, the differences in expression levels of *HIBADH* between normal and tumor tissues were not significant ([Fig. 1F](#) and [1G](#)), suggesting that *circHIBADH* functioned independently from its parental gene.

Functional annotation of *circHIBADH*

An overview of *circHIBADH* is presented in [Fig. 2A](#). The rs11973492 SNP is located in the intron between exon 6 and exon 7 of *HIBADH*, implying that the variation of rs11973492 may reduce the generation of *circHIBADH* by altering the structure of RNA chains. In total, we predicted 21 potential RBPs of *circHIBADH* in CircAtlas, among which six (*i.e.*, eIF4A3, PTBP1, ELAVL1, HNRNPD, AGO2, and HNRNPA1) were also identified in CSCD ([Fig. 2B](#)).

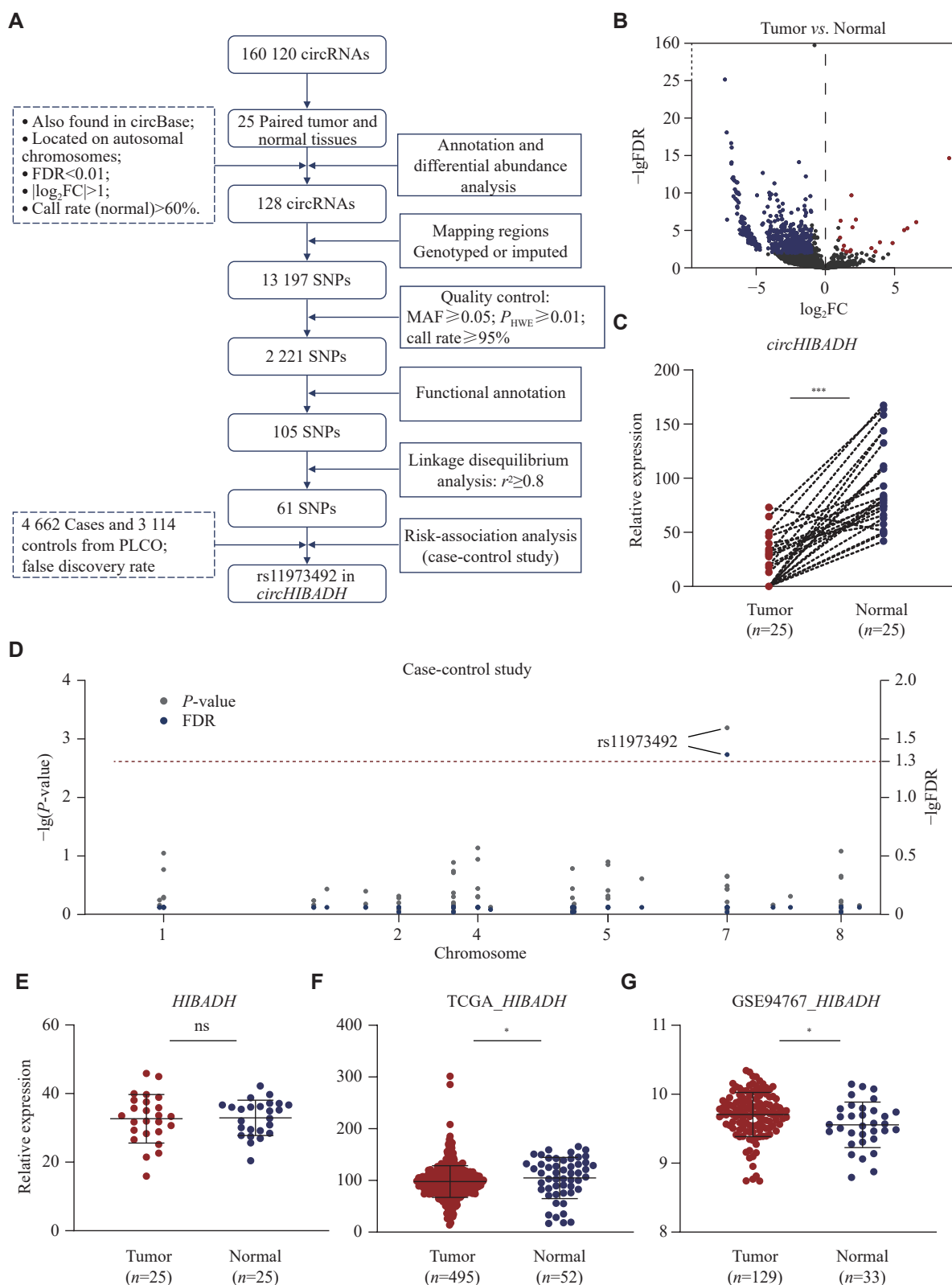


Fig. 1 SNP selection among PCA-related circRNA regions. A: The process of SNP selection. B: The volcano plot of the differential expression analysis of circRNAs between 25 paired tumor and adjacent normal tissues. C: *circHIBADH* expression difference between 25 paired tumor and adjacent normal tissues. D: The Manhattan plot exhibits the risk-association analysis of 61 SNPs after linkage disequilibrium analysis. E–G: *HIBADH* expression differences between tumor and adjacent normal tissues in MiOncoCirc, TCGA PRAD, and GSE94767 datasets. Expression levels were compared using the Wilcoxon signed-rank test and Wilcoxon rank-sum test for paired and unpaired samples, respectively. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. Abbreviations: FC, fold change; FDR, false discovery rate; SNP, single nucleotide polymorphism; MAF, minor allele frequency; HWE, Hardy-Weinberg equilibrium; PLCO, prostate, lung, colorectal, and ovarian cancer screening trials; TCGA, The Cancer Genome Atlas; ns, not significant.

Table 1 Associations between rs11973492 in *circHIBADH* and risk of prostate cancer

Genotypes	Cases (N=4 624)		Controls (N=3 079)		OR (95%CI)	P	Adjusted OR (95%CI) ^a	P ^a
	n	%	n	%				
TT	2 089	45.18	1 537	49.92	1		1	
TC (TC vs. TT)	2 092	45.24	1 261	40.95	1.22 (1.11–1.34)	4.72×10 ⁻⁵	1.23 (1.10–1.37)	3.07×10 ⁻⁵
CC (CC vs. TT)	443	9.58	281	9.13	1.16 (0.99–1.37)	7.50×10 ⁻²	1.09 (0.90–1.31)	3.95×10 ⁻¹
Additive model (CC vs. TC vs. TT)					1.13 (1.05–1.21)	5.97×10 ⁻⁴	1.11 (1.02–1.21)	1.16×10 ⁻²
Dominant model (CC & TC vs. TT)					1.21 (1.10–1.33)	4.46×10 ⁻⁵	1.20 (1.08–1.34)	7.06×10 ⁻⁴
Recessive model (CC vs. TC & TT)					1.06 (0.90–1.23)	5.04×10 ⁻¹	1.04 (0.91–1.18)	6.03×10 ⁻¹

^aAdjusted for age, smoking status, and the top ten principal components in the logistic regression model. Abbreviations: OR, odds ratio; CI, confidence interval.

Table 2 Stratification analyses of clinicopathologic variables for the association between rs11973492 and prostate cancer risk in the dominant model

Variables	Genotypes				OR (95% CI) ^a	P ^a
	TT (n)	%	TC & CC (n)	%		
Controls	1 537	49.92	1 542	50.08		
Cases	2 089	45.18	2 535	54.82	1.20 (1.08–1.34)	7.06×10 ⁻⁴
Gleason score						
≤6	1 219	45.16	1 480	54.84	1.24 (1.10–1.41)	6.53×10 ⁻⁴
=7	640	44.51	798	55.49	1.27 (1.10–1.47)	1.25×10 ⁻³
≥8	203	46.88	230	53.12	1.16 (0.94–1.43)	1.71×10 ⁻¹
PSA (ng/mL)						
<10	1 625	45.18	1 972	54.82	1.22 (1.09–1.37)	6.88×10 ⁻⁴
10–20	254	43.87	325	56.13	1.33 (1.10–1.62)	3.64×10 ⁻³
>20	114	50.89	110	49.11	0.96 (0.72–1.27)	7.58×10 ⁻¹
Stages						
Stage I / II	1 813	44.98	2 218	55.02	1.23 (1.10–1.37)	2.57×10 ⁻⁴
Stage III / IV	276	46.62	316	53.38	1.17 (0.95–1.45)	1.32×10 ⁻¹

^aAdjusted for age, smoking status and the top ten principal components in the logistic regression model. Abbreviations: OR, odds ratio; CI, confidence interval; PSA, prostate specific antigen.

Functional annotation implied that these RBPs were predominantly associated with RNA splicing processes (Fig. 2C). The PPI network also prompted that the interaction of predicted RBPs mainly enriched in the process of RNA splicing, in which proteins HNRNPA1 and HNRNPM appeared to be hubs of the network (Fig. 2D).

Identification and validation of hub RBPs

We assessed the expression levels of 21 predicted RBPs in the TCGA PRAD dataset, and found that *HNRNPA1* exhibited consistently high expression levels among the tested samples, and was significantly over-expressed in tumor tissues ($P = 3.98 \times 10^{-5}$, Fig. 3A). We also found a notably negative correlation between *circHIBADH* and *HNRNPA1* expression

levels ($R = -0.49$, $P = 3.01 \times 10^{-4}$, Fig. 3B), implying that *circHIBADH* may function as a protein sponge, thus inhibiting the HNRNPA1-induced RNA splicing process. Integrating with the aforementioned inter-protein interactions, HNRNPA1 was identified as the hub RBP in the *circHIBADH* regulating RNA splicing process. Firstly, we identified that the T allele of rs11973492 was associated with a higher expression level of *HNRNPA1*, although not statistically significant ($P = 6.79 \times 10^{-2}$, Supplementary Fig. 2 [available online]). In addition, the up-regulation of *HNRNPA1* in tumors was replicated in GSE94767 and GSE183019 datasets ($P = 1.99 \times 10^{-2}$ and 3.64×10^{-2} , respectively, Fig. 3C and 3D). As shown in Fig. 3E, the DE analysis was performed on samples of the top and the bottom 1/4 *HNRNPA1* expression levels.

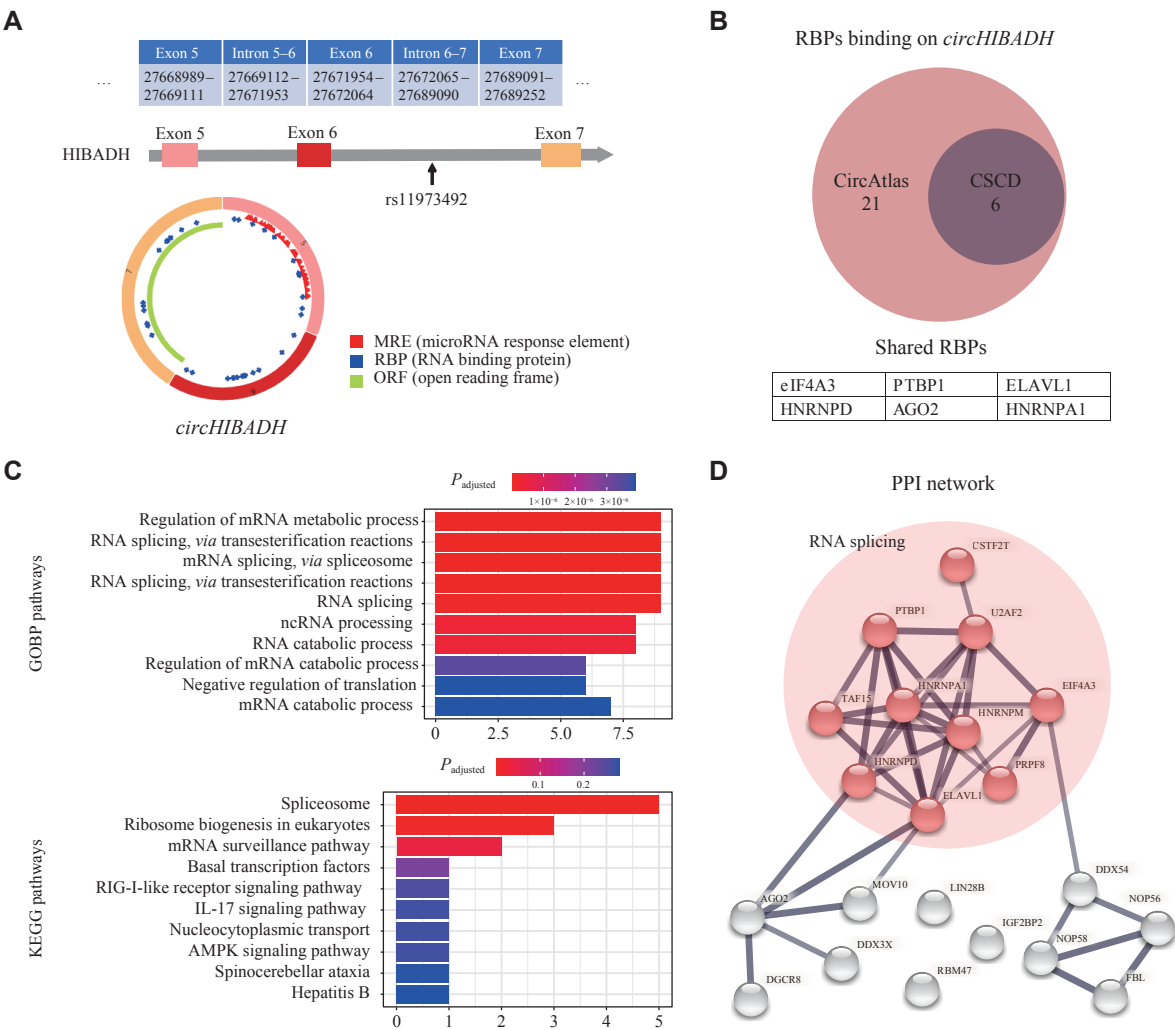


Fig. 2 Annotation of *circHIBADH*. A: The genomic loci of the *HIBADH* gene and *circHIBADH*. B: The Venn diagram of the predicted RBPs of *circHIBADH* based on the CircAtlas and CSCD databases. C: The pathway enrichment analysis of 21 predicted RBPs using the GOBP (upper) and KEGG (bottom) databases. D: The PPI network of 21 predicted RBPs. Abbreviations: RBP, RNA-binding protein; GOBP, gene ontology biological process; KEGG, Kyoto Encyclopedia of Genes and Genomes; PPI, protein-protein interaction.

Moreover, the corresponding GSEA analysis demonstrated that several carcinogenesis-associated hallmarks, *e.g.*, MYC target signaling, DNA repair, and E2F target signaling, which are known to be correlated with cell proliferation, were found to be activated in the *HNRNPA1*-high tissues (Fig. 3F). As expectedly, the expression level of *circHIBADH* was significantly negatively correlated with these downstream pathways, including RNA splicing (Supplementary Fig. 3, available online).

We then introduced two single-cell transcriptome data for additional validation, of which the cellular landscapes are shown in Supplementary Fig. 4 (available online). Epitheliums were extracted and annotated as basal epithelium (*TP63*⁺*KRT5*⁺), hillock epithelium (*KRT13*⁺*S100P*⁺), club epithelium (*SCGB3A1*⁺*WFDC2*⁺), stress response epithelium (*FOS*^{hi}*JUN*^{hi}), luminal epithelium-androgen response-high (LE_AR_hi, *ACPP*^{hi}*KLK3*^{hi}), and LE_AR-low

(LE_AR_lo, *ACPP*^{lo}*KLK4*^{hi}, Fig. 4A–4D), according to the epithelial subtypes identified by Henry *et al*[22]. As shown in Fig 4C and 4D, *PCA3*, an established PCa marker, was highly expressed in the LE_AR_lo subtype, so we assigned LE_AR_lo as tumor cells but LE-AR-hi as normal luminal cells. Consistently, the expression levels of *HNRNPA1* were significantly higher in tumor cells than in normal luminal cells (Fig. 4E and 4F). Furthermore, the gene ontology biological process "Up-regulation of RNA splicing" was notably activated in malignant cells, compared with normal luminal cells (Fig. 4G and 4H).

Discussion

In the current study, we discovered rs11973492 in *circHIBADH* as a significant risk factor for PCa. Moreover, rs11679306 appears to inhibit the generation of *circHIBADH* and suppress the sponging

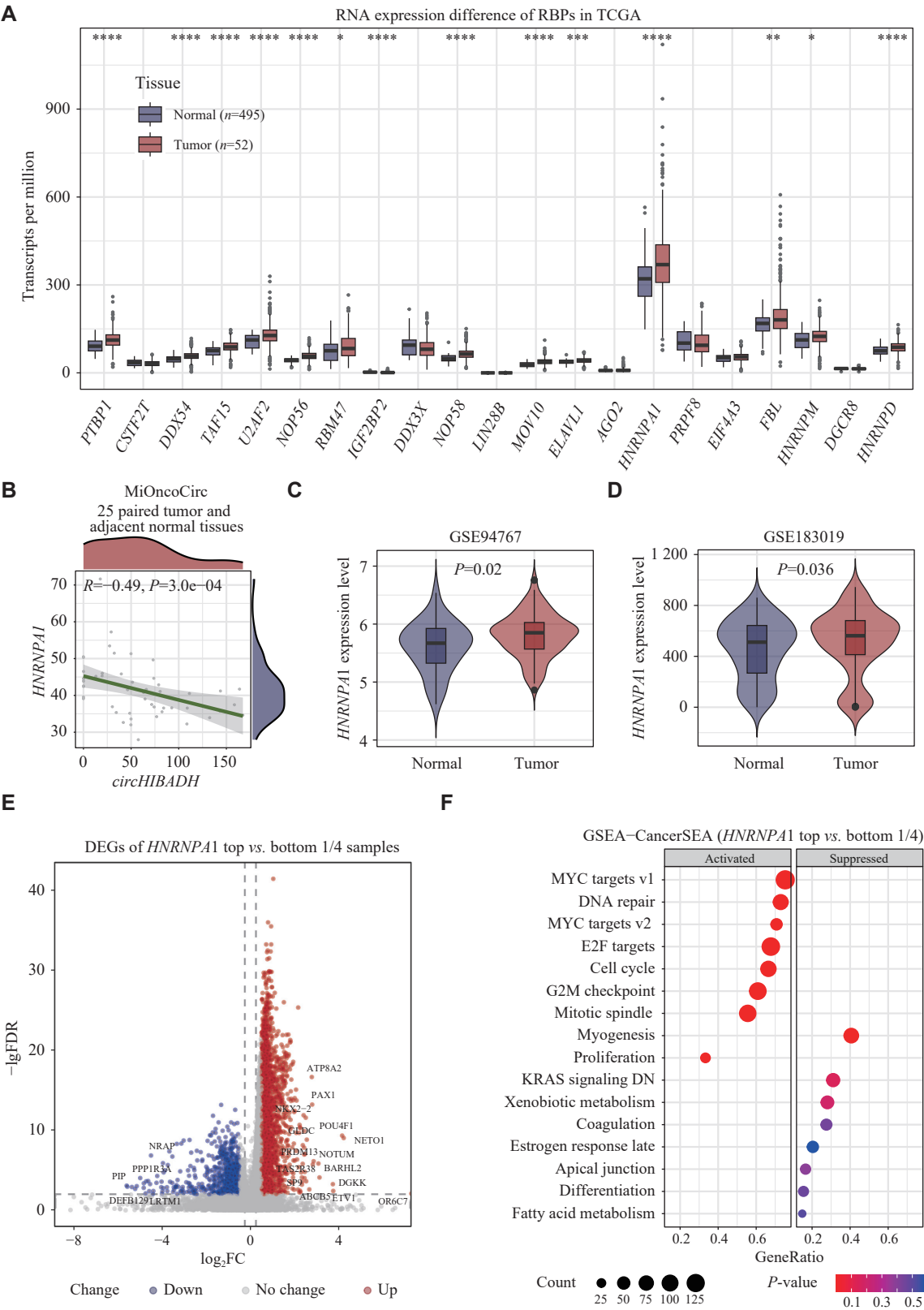


Fig. 3 Identification and validation of hub RBP. A: Differential expression analysis of 21 predicted RBPs between tumor and adjacent normal tissues in the TCGA prostate adenocarcinoma dataset. B: Correlation between the expression levels of *HNRNP A1* and *circHIBADH* in the MiOncoCirc dataset using the Spearman correlation test. C and D: Differential expression levels of *HNRNP A1* in the GSE94767 and GSE183019 datasets. E: The volcano plot of the differentially expressed genes between samples with the top and the bottom 1/4 *HNRNP A1* expression levels in the TCGA PRAD dataset. F: Gene-set enrichment analysis based on gene expression differences between samples with the top and the bottom 1/4 *HNRNP A1* expression levels using the CancerSEA database. Expression levels were compared using the Wilcoxon signed-rank test. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$. Abbreviation: RBP, RNA-binding protein.

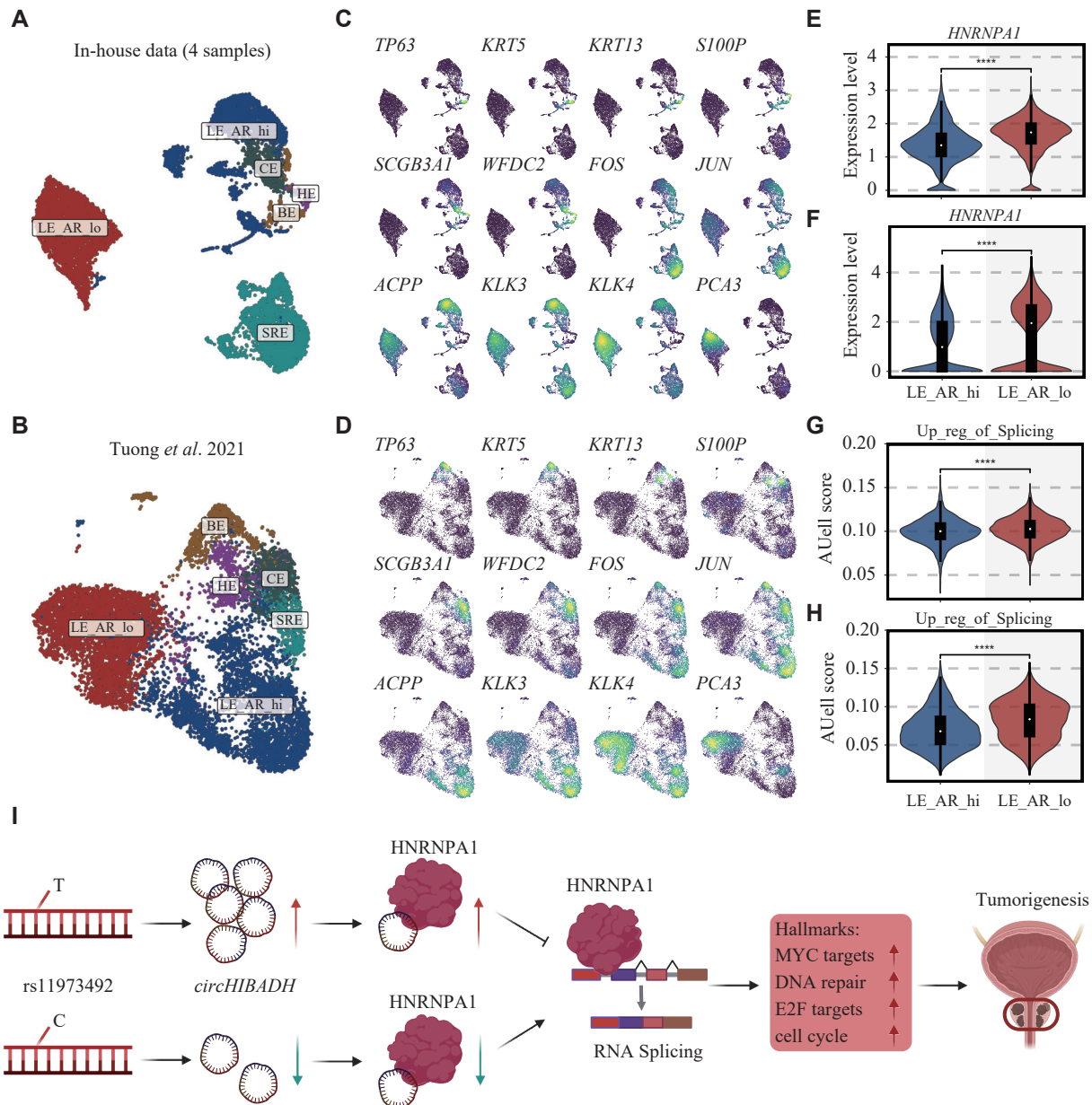


Fig. 4 Validation of hub RBP and key biological processes in single-cell transcriptomes. A and B: UMAP plot of 10 791 epitheliums in the in-house data (A) and 10 587 epitheliums in data from Tuong *et al* (B), colored by sub-cell type. C and D: The UMAP plots exhibiting the expression density of different epithelium sub-type markers in the in-house data (C) and data from Tuong *et al* (D). E and F: Differences in the expression of HNRNPA1 between LE_AR_hi and LE_AR_lo epitheliums in the in-house data (E) and 10 587 epitheliums in data from Tuong *et al* (F). G and H: Differences in the AUCell of process "up-regulation of RNA splicing" between LE_AR_hi and LE_AR_lo epitheliums in the in-house data (G) and 10 587 epitheliums in data from Tuong *et al* (H). I: The schematic diagram illustrating the mechanism of rs11973492 regulating PCa oncogenesis. Expression levels and AUCell scores were compared using the Wilcoxon signed-rank test. **** $P < 0.0001$. Abbreviations: LE_AR_lo, luminal epithelium-androgen response-low; LE_AR_hi, luminal epithelium-androgen response-high; CE, club epithelium; HE, hillock epithelium; BE, basal epithelium; SRE, stress response epithelium; RBP: RNA-binding protein; UMAP: Uniform Manifold Approximation and Projection.

of HNRNPA1 and other RBPs, thus enhancing RNA splicing-induced carcinogenesis (Fig. 4I).

circRNAs are tissue-specific and function in various biological processes in multiple malignancies^[23], including PCa. In terms of the interaction with RBPs, circRNAs may sponge RBPs to inhibit^[9], scaffold, or recruit RBPs to facilitate^[10–11] and enhance RBP expression to promote^[12] protein functions. In PCa, it

is reported that has_circ_0003258 promotes PCa metastasis through complexing with IGF2BP3^[24], and that circTFDP2 facilitated PCa progression by sponging and inhibiting PARP1^[25]. In the current study, we found that circHIBADH might suppress the RNA splicing role of HNRNPA1 and other RBPs by acting as a protein sponge.

Additionally, a significantly differential expression

of *circHIBADH*, but not the parental genes (*HIBADH*), was found between normal and tumor tissues, suggesting that circRNA functions independently from mRNAs, which may be modulated by rs11973492. Such a phenomenon has been proposed and discussed in previous studies^[19,26]. Some circRNAs located in the nucleus bind to RNA polymerase, thus regulating the transcription of parental genes^[12]. For instance, an excision of the *DOCK1* circRNA contributes to the downregulation of *DOCK1* in epithelial cells^[27]. Moreover, we found that rs11973492 was significantly associated with an early-stage PCa but not with a late-stage PCa. This disparity, on the one hand, may be attributed to the relatively smaller sample size of late-stage PCa patients. On the other hand, it suggests that rs11973492 may play a specific role in the oncogenesis of PCa rather than the progression, thus explaining its weaker association with late-stage PCa.

Furthermore, *circHIBADH* is an exonic circRNA, and rs11973492 is located in the intron between circRNA-forming exons. Studies have reported that complementary base pairing of inverted repeats in the intron flanking the exons promotes the generation of circular RNA by approaching the splicing site of circular RNA^[27–28]. Liang *et al*^[29] showed that micro-introns containing splice sites and short inverted repeats promoted the circulation of intervening exons in cells. Moreover, the secondary structure within pre-mRNAs was shown to enhance circRNA biogenesis by bringing circRNA-forming exons into a close proximity^[30]. We found that the variant of rs11973492 changed the secondary structure of the corresponding mRNA chain, which may affect the generation of circRNAs and their biological functions.

As a hub RBP, HNRNPA1 was identified to play a key role in the *circHIBADH*-related PCa risk, through the regulation of RNA splicing process. The dysregulation of RNA splicing is common in PCa, of which a notable example is the androgen receptor splice variant AR-V7^[31]. SF3B2-mediated RNA splicing was also found to drive progression in PCa^[32]. HNRNPA1, a member of heterogeneous nuclear ribonucleoproteins, is associated with pre-mRNAs in the nucleus and influences pre-mRNA processing. Likewise, HNRNPA1 promotes oncogenesis by regulating proliferation in several malignancies^[33], which is potentially implicated in enzalutamide resistance and aggressiveness in PCa^[34–36]. Additionally, HNRNPA1 has been reported as a biomarker for early biochemical recurrence of PCa^[37]. The genotype-based gene expression analysis is important to explore the function of SNPs; so, through

expression quantitative trait loci analysis, we found that the T allele of rs11973492 was associated with a higher expression level of *HNRNPA1*. Biologically, this may represent a negative feedback mechanism that occurs following the silencing of HNRNPA1 by the rs11973492-associated elevation of *circHIBADH*, resulting in the increased transcription of *HNRNPA1*.

Through bioinformatics analysis, we speculated that as a downregulated circRNA in tumors, *circHIBADH* exerted its anticancer effect through its protein sponge function. Once such a mechanism was downregulated, the silenced protein (*i.e.*, HNRNPA1) may exert its oncogenic function. The protein-binding function of circRNAs has been extensively explored: they act as sponges, scaffolds, recruiters, or function enhancers of the proteins to exert their effects. Protein sponges, similar to the classic competing endogenous RNA mechanisms, inhibit the function of the corresponding RBPs, as demonstrated in the current study. This mechanism of protein sponges has been reported in various studies involving liver cancer, head and neck squamous cell carcinoma, glioblastoma, among others^[38].

In conclusion, we found that the variant of rs11973492 obstructed the generation of *circHIBADH*, which may act as a protein sponge and inhibitor of HNRNPA1 and other RBPs, consequently enhancing RNA splicing processes and subsequent MYC targeting, DNA repair, and E2F target signaling pathways, to play an oncogenic role in PCa.

Fundings

This work was supported by the Medical Research Project of Jiangsu Commission of Health (Grant No. M2022015).

Acknowledgments

We thank Dr. Mulong Du from Nanjing Medical University for his guidance in the data interpretation.

References

- [1] Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. *CA Cancer J Clin*, 2021, 71(3): 209–249.
- [2] Wang YA, Sfakianos J, Tewari AK, et al. Molecular tracing of prostate cancer lethality[J]. *Oncogene*, 2020, 39(50): 7225–7238.
- [3] Komura K, Sweeney CJ, Inamoto T, et al. Current treatment

- strategies for advanced prostate cancer[J]. *Int J Urol*, 2018, 25(3): 220–231.
- [4] Mucci LA, Hjelmborg JB, Harris JR, et al. Familial risk and heritability of cancer among twins in Nordic countries[J]. *JAMA*, 2016, 315(1): 68–76.
 - [5] Conti DV, Darst BF, Moss LC, et al. Trans-ancestry genome-wide association meta-analysis of prostate cancer identifies new susceptibility loci and informs genetic risk prediction[J]. *Nat Genet*, 2021, 53(1): 65–75.
 - [6] Welter D, MacArthur J, Morales J, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations[J]. *Nucleic Acids Res*, 2014, 42(D1): D1001–D1006.
 - [7] Kristensen LS, Andersen MS, Stagsted LVW, et al. The biogenesis, biology and characterization of circular RNAs[J]. *Nat Rev Genet*, 2019, 20(11): 675–691.
 - [8] Zheng Q, Bao C, Guo W, et al. Circular RNA profiling reveals an abundant circHIPK3 that regulates cell growth by sponging multiple miRNAs[J]. *Nat Commun*, 2016, 7: 11215.
 - [9] Xia P, Wang S, Ye B, et al. A circular RNA protects dormant hematopoietic stem cells from DNA Sensor cGAS-mediated exhaustion[J]. *Immunity*, 2018, 48(4): 688–701.e7.
 - [10] Du WW, Fang L, Yang W, et al. Induction of tumor apoptosis through a circular RNA enhancing Foxo3 activity[J]. *Cell Death Differ*, 2017, 24(2): 357–370.
 - [11] Chen N, Zhao G, Yan X, et al. A novel *FLII* exonic circular RNA promotes metastasis in breast cancer by coordinately regulating TET1 and DNMT1[J]. *Genome Biol*, 2018, 19(1): 218.
 - [12] Li Z, Huang C, Bao C, et al. Exon-intron circular RNAs regulate transcription in the nucleus[J]. *Nat Struct Mol Biol*, 2015, 22(3): 256–264.
 - [13] Zhang M, Zhao K, Xu X, et al. A peptide encoded by circular form of *LINC-PINT* suppresses oncogenic transcriptional elongation in glioblastoma[J]. *Nat Commun*, 2018, 9(1): 4475.
 - [14] Bi J, Liu H, Dong W, et al. Circular RNA circ-ZKSCAN1 inhibits bladder cancer progression through miR-1178–3p/p21 axis and acts as a prognostic factor of recurrence[J]. *Mol Cancer*, 2019, 18(1): 133.
 - [15] Liu T, Liu S, Xu Y, et al. Circular RNA-ZFR inhibited cell proliferation and promoted apoptosis in gastric cancer by sponging miR-130a/miR-107 and modulating *PTEN*[J]. *Cancer Res Treat*, 2018, 50(4): 1396–1417.
 - [16] Kleivi Sahlberg K, Bottai G, Naume B, et al. A serum microRNA signature predicts tumor relapse and survival in triple-negative breast cancer patients[J]. *Clin Cancer Res*, 2015, 21(5): 1207–1214.
 - [17] Chen S, Huang V, Xu X, et al. Widespread and functional RNA circularization in localized prostate cancer[J]. *Cell*, 2019, 176(4): 831–843. e22.
 - [18] Shen Z, Zhou L, Zhang C, et al. Reduction of circular RNA Foxo3 promotes prostate cancer progression and chemoresistance to docetaxel[J]. *Cancer Lett*, 2020, 468: 88–101.
 - [19] Vo JN, Cieslik M, Zhang Y, et al. The landscape of circular RNA in cancer[J]. *Cell*, 2019, 176(4): 869–881. e13.
 - [20] Cheng Y, Meng Y, Li S, et al. Genetic variants in the cholesterol biosynthesis pathway genes and risk of prostate cancer[J]. *Gene*, 2021, 774: 145432.
 - [21] Tuong ZK, Loudon KW, Berry B, et al. Resolving the immune landscape of human prostate at a single-cell level in health and cancer[J]. *Cell Rep*, 2021, 37(12): 110132.
 - [22] Henry GH, Malewska A, Joseph DB, et al. A cellular anatomy of the normal adult human prostate and prostatic urethra[J]. *Cell Rep*, 2018, 25(12): 3530–3542. e5.
 - [23] Yang Q, Li F, He AT, et al. Circular RNAs: expression, localization, and therapeutic potentials[J]. *Mol Ther*, 2021, 29(5): 1683–1702.
 - [24] Yu Y, Lv D, Wang C, et al. Hsa_circ_0003258 promotes prostate cancer metastasis by complexing with IGF2BP3 and sponging miR-653–5p[J]. *Mol Cancer*, 2022, 21(1): 12.
 - [25] Ding L, Zheng Q, Lin Y, et al. Exosome-derived circTFDP2 promotes prostate cancer progression by preventing PARP1 from caspase-3-dependent cleavage[J]. *Clin Transl Med*, 2023, 13(1): e1156.
 - [26] Zhou R, Wu Y, Wang W, et al. Circular RNAs (circRNAs) in cancer[J]. *Cancer Lett*, 2018, 425: 134–142.
 - [27] Conn SJ, Pillman KA, Toubia J, et al. The RNA binding protein quaking regulates formation of circRNAs[J]. *Cell*, 2015, 160(6): 1125–1134.
 - [28] Ashwal-Fluss R, Meyer M, Pamudurti NR, et al. circRNA biogenesis competes with pre-mRNA splicing[J]. *Mol Cell*, 2014, 56(1): 55–66.
 - [29] Liang D, Wilusz JE. Short intronic repeat sequences facilitate circular RNA production[J]. *Genes Dev*, 2014, 28(20): 2233–2247.
 - [30] Zhang X, Wang H, Zhang Y, et al. Complementary sequence-mediated exon circularization[J]. *Cell*, 2014, 159(1): 134–147.
 - [31] Paschalis A, Sharp A, Welte JC, et al. Alternative splicing in prostate cancer[J]. *Nat Rev Clin Oncol*, 2018, 15(11): 663–675.
 - [32] Kawamura N, Nimura K, Saga K, et al. SF3B2-mediated RNA splicing drives human prostate cancer progression[J]. *Cancer Res*, 2019, 79(20): 5204–5217.
 - [33] Roy R, Huang Y, Seckl MJ, et al. Emerging roles of hnRNPA1 in modulating malignant transformation[J]. *WIREs RNA*, 2017, 8(6): e1431.
 - [34] Tummala R, Lou W, Gao AC, et al. Quercetin targets hnRNPA1 to overcome enzalutamide resistance in prostate cancer cells[J]. *Mol Cancer Ther*, 2017, 16(12): 2770–2779.
 - [35] Zhang M, Sun Y, Huang C, et al. Targeting the Lnc-OPHN1–5/androgen receptor/hnRNPA1 complex increases Enzalutamide sensitivity to better suppress prostate cancer progression[J]. *Cell Death Dis*, 2021, 12(10): 855.
 - [36] Lian C, Zhang C, Tian P, et al. Epigenetic reader ZMYND11 noncanonical function restricts HNRNPA1-mediated stress granule formation and oncogenic activity[EB/OL]. [2023-11-28]. <http://dx.doi.org/10.1101/2023.11.28.569112>.
 - [37] Möller K, Wecker AL, Höflmayer D, et al. Upregulation of the heterogeneous nuclear ribonucleoprotein hnRNPA1 is an independent predictor of early biochemical recurrence in TMPRSS2: ERG fusion-negative prostate cancers[J]. *Virchows Arch*, 2020, 477(5): 625–636.
 - [38] Kristensen LS, Jakobsen T, Hager H, et al. The emerging roles of circRNAs in cancer and oncology[J]. *Nat Rev Clin Oncol*, 2022, 19(3): 188–206.