CRIT

Shao, Mengting; Hao, Shijia; Jiang, Leiming; Cai, Yujie; Zhao, Xing; Chen, Qiuyang; Gao, Xuefei; Xu, Jianzhen

Published in:
Molecular therapy - Nucleic acids

DOI:
10.1016/j.omtn.2022.10.015

Publication date:
2022

Molecular Therapy
Nucleic Acids
Original Article

CRIT: Identifying RNA-binding protein regulator in circRNA life cycle via non-negative matrix factorization

Mengting Shao,[1,6] Shijia Hao,[1,6] Leiming Jiang,[1] Yujie Cai,[1,2] Xing Zhao,[1,3] Qiuyang Chen,[1] Xuefei Gao,[4,5] and Jianzhen Xu[1,5]

[1]Computational Systems Biology Lab, Department of Bioinformatics, Shantou University Medical College (SUMC), 515041 Shantou, China; [2]Guangdong Key Laboratory of Age-Related Cardiac and Cerebral Diseases, Affiliated Hospital of Guangdong Medical University, 524000 Zhanjiang, China; [3]Department of Pathology and Medical Biology, University of Groningen, University Medical Center, Groningen, 9700 RB Groningen, the Netherlands; [4]Department of Physiology, School of Basic Medical Sciences, Southern Medical University, 510515 Guangzhou, China; [5]Guangdong Provincial Key Laboratory of Infectious Diseases and Molecular Immunopathology, 515041 Shantou, China

Circular RNAs (circRNAs) are endogenous non-coding RNAs that regulate gene expression and participate in carcinogenesis. However, the RNA-binding proteins (RBPs) involved in circRNAs biogenesis and modulation remain largely unclear. We developed the circRNA regulator identification tool (CRIT), a non-negative matrix-factorization-based pipeline to identify regulating RBPs in cancers. CRIT uncovered 73 novel regulators across thousands of samples by effectively leveraging genomics data and functional annotations. We demonstrated that known RBPs involved in circRNA control are significantly enriched in these predictions. Analysis of circRNA-RBP interactions using two large cross-linking immunoprecipitation (CLIP) databases, we validated the consistency between CRIT prediction and the CLIP experiments. Furthermore, newly discovered RBPs are functionally connected with authentic circRNA regulators by various biological associations, such as physical interaction, similar binding motifs, common transcription factor modulation, and co-expression. When analyzing RNA sequencing (RNA-seq) datasets after short hairpin RNA (shRNA)/small interfering RNA (siRNA) knockdown, we found several novel RBPs that can affect global circRNA expression, which strengthens their role in the circRNA life cycle. The above evidence provided independent confirmation that CRIT is a useful tool to capture RBPs in circRNA processing. Finally, we show that authentic regulators are more likely the core splicing proteins and peripheral factors and usually harbor more alterations in the vast majority of cancers.

INTRODUCTION

Circular RNAs (circRNAs) are endogenous non-coding RNAs produced by the circularization of specific exons. They has been widely implicated in neurodegenerative disorders, cardiovascular diseases, and cancers.[1–3] circRNAs modulate carcinogenesis via a variety of elegant mechanisms such as acting as microRNA (miRNA) sponges, translating into novel peptides, and epigenetic regulating.[3–5] Notably, circRNAs also demonstrate their diagnostic values since they can be detected both in formalin-fixed, paraffin-embedded (FFPE) tissues and human biofluids, and their expression is usually associated with cancer progression.[6–8]

Recently, several RNA-binding proteins (RBPs) have been shown to control circRNA biogenesis.[9–13] For example, the quaking (QKI) protein family, which contains the KH domain, was previously reported to play an essential role in circRNA formation during the epithelial-to-mesenchymal transition.[10] It regulates the circ-ZKSCAN1 circularization, which is associated with cancer stemness and inhibits hepatocellular carcinoma growth.[14] These RBPs bind to flanking intronic regions upstream or downstream of circ-forming exons. Except for stacking into homo/heterodimers for circRNA regulation in *trans,* several RBPs can also influence Alu pairing in *cis* (such as ADAR and DHX9). Another group of RBPs can directly bind to a full-length circRNA "body sequence," therefore regulating its more broad behaviors in cells.[15,16] These RBPs have been proposed to control circRNAs N6-methyladenosine (m6A) modification, localization, stability, and translation. For instance, ribosome-associated factors can regulate the translation of circ-ZNF609 in cap-independent manners.[17] As both groups of RBPs control circRNA metabolism and cellular functioning during the whole life cycle, we collectively defined them as circRNA "regulators" in this study.

While potential protein regulators of circRNA have continuously been identified, the progress in this field is seriously delayed, especially when

Received 11 August 2022; accepted 25 October 2022;
https://doi.org/10.1016/j.omtn.2022.10.015.

[6]These authors contributed equally

Correspondence: Jianzhen Xu, Computational Systems Biology Lab, Department of Bioinformatics, Shantou University Medical College (SUMC), 515041 Shantou, China.
E-mail: jzxu01@stu.edu.cn

compared with the rapid identification of circRNA downstream targets. This situation is mainly hampered by the fact that few computational frameworks have been rationally designed to systematically evaluate the involvement of regulating RBPs in the circRNA life cycle. As alternative approaches, researchers have to largely rely on either sequencing of immuno-precipitated RNAs after cross-linking with individual RBPs or genome-wide small interfering RNA (siRNA) screening of hundreds of RBPs, which are both time consuming and labor intensive.[18–20] Several groups have addressed a related, but not equivalent, problem, i.e., prediction of RBP-circRNA-binding sites by deep-learning architecture.[21,22] But these binding site embedding methods relied heavily on training datasets and only obtained limited RBP-specific models (from several to dozens of RBPs), which makes the extrapolation to other RBPs without canonical binding domains or novel circRNA-regulating RBPs in question. Furthermore, previous circRNA regulators identified either from the high-throughput screen or deep-learning prediction are usually derived from cells-based models or resources; whether or not they have physio-pathological impacts needs further confirmation.[19,20]

Here, we proposed a novel bioinformatics pipeline named the circRNA regulator identification tool (CRIT) to identify circRNA regulators directly from thousands of tissue samples via examining genomics, transcriptomics, and interactomics databases across 12 cancer types. CRIT integrates gene alterations, clinical significance, gene-circRNA expression correlation, and RBP-RNA-binding data into a statistical framework. CRIT can uncover both previously established circRNA regulators and novel RBPs, which are supported by multiple biological pieces of evidence.

RESULTS

Curation of various cancer omics data for human RBPs

RBPs control the entire RNA life cycle such as splicing, transportation, translation, and stability, therefore they are tightly connected with the cancer mechanism.[23,24] We reasoned that, as pathological/physiological-relevant circRNA regulators, these RBPs should have the following characteristics. (1) Since the expression of regulators could affect circRNA expression, regulators and some circRNAs should therefore have a strong expression correlation. (2) Regulators should have at least one of the following alterations, point mutation, copy-number variation, and abnormal mRNA expression, in the cancer sample. (3) The above alterations on regulators have an impact on the survival time of patients. (4) Regulators' biological role should be closely related to the Gene Ontology (GO) terms "mRNA splicing, via spliceosome," which indicate the participation in the formation of circRNA. (5) Regulators can physically interact with some circRNAs.

Cluster RBPs with non-negative matrix factorization

Based on the above biologically meaningful features, we established a CRIT pipeline to identify putative regulators by integratively analyzing various cancer omics data including mutation, copy-number variation (CNV), expression, patient follow-up information, function annotation, and protein-RNA interaction (Figure 1). The pipeline is described in detail in the materials and methods and the supplemental methods. The component genes in each cluster are listed in (Table S1).

To support the validity of CRIT's result and the analysis pipeline, we performed a comparison of regulators detected by CRIT with the results based on siRNA screens or published functional studies. Previously, Li et al. set up a circmCherry-expression screening system and identified RBPs potentially involved in circRNA regulation.[13] Among the 103 RBPs that influenced the mCherry fluorescence with >1.5-fold change, CRIT correctly identified 13 of 90 candidate regulators in cluster 23 (Fisher's exact test $p = 0.00201$). It should be noted that the above cell-based screen used artificial constructs under unphysiological conditions and may promote suspicious protein-circRNA signals. Thus, to further establish the accuracy of CRIT's result, we manually collected 19 well-established circRNA regulators (Table S2), which have been thoroughly verified by multiple assays both *in vivo* and *in vitro*.[3,25–29] CRIT correctly identified six out of the 19 established regulators (Table S2) (Fisher's exact test $p = 0.00112$). Finally, we combined the results from either vector screen or functional studies to form a set of 118 gold-standard circRNA regulators (hereinafter called "gold regulators"). Among the 90 RBPs clustered in group 23, there are 17 gold regulators (Fisher's exact test $p = 0.00012$; Table S1). Thus, we regard all RBPs in this group as putative circRNA regulators.

The agreement between the CRIT approach and the CLIP experiments is strong

Sequencing the immunoprecipitated RNAs after cross-linking with the specific RBP (CLIP-seq) can reveal the RBP-RNA interaction and binding sites. Although mRNAs are usually the primary focus of such experiments, circRNA-RBP interactions could also be recovered. For example, the PORSTAR3 database has archived 112 potential circRNA-binding RBPs derived from public CLIP-seq datasets.[30] Among them, there are 19 intersecting RBPs with the candidate regulators. Fisher's exact test shows that candidate regulators are significantly enriched with such circRNA-binding RBPs using either 1,344 RBPs (Fisher's exact test $p = 7.7 \times 10^{-5}$) or total genes (Fisher's exact test $p = 7.8 \times 10^{-30}$) as background.

Several technical difficulties in CLIP experiments often lead to suspicious circRNA-RBP interactions.[31] Therefore, the developers of the ENCORI database score each circRNA-binding RBP by the number of supported experiments.[32] This score could be regarded as a confidence of each circRNA-binding RBP. Thus, we separated the circRNA-binding RBPs from ENCORI into five sub-lists based on their stringency. The result of Fisher's exact test also shows that candidate regulators are significantly enriched under each of the different stringencies (Figure 2A). For example, at the highest confidence stringency ≥ 5, there are only 27 CLIP-proved circRNA-binding RBPs, which overlapped with two of the candidate RBPs (Figure 2A; Fisher's exact test $p = 3.3 \times 10^{-8}$ with 1,344 RBP backgrounds or 5.8×10^{-25} with total gene background). Overall, based on the above analysis of circRNA-RBP interactions, using two large CLIP databases, we validated the consistency between our CRIT prediction and the CLIP experiments.

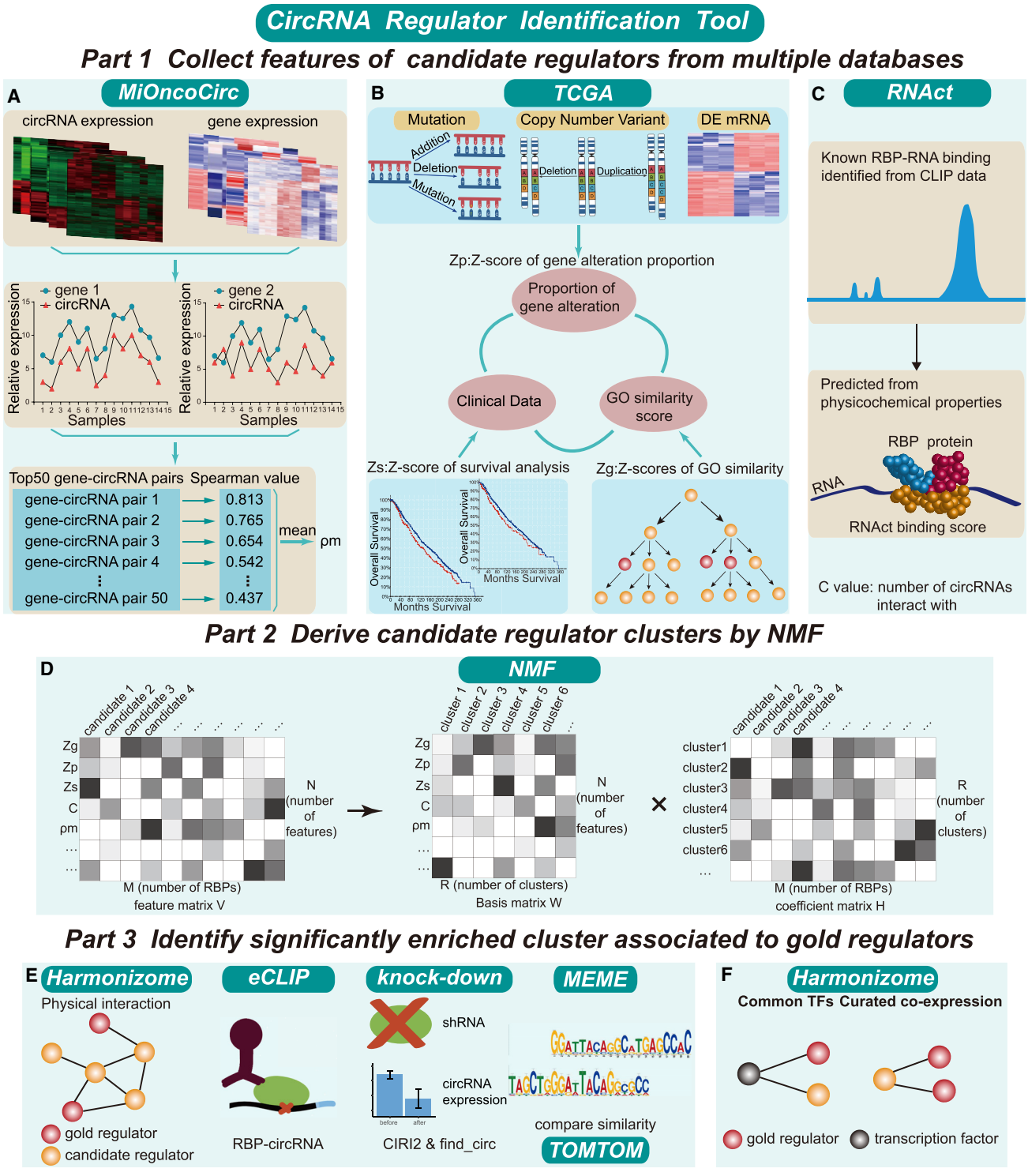
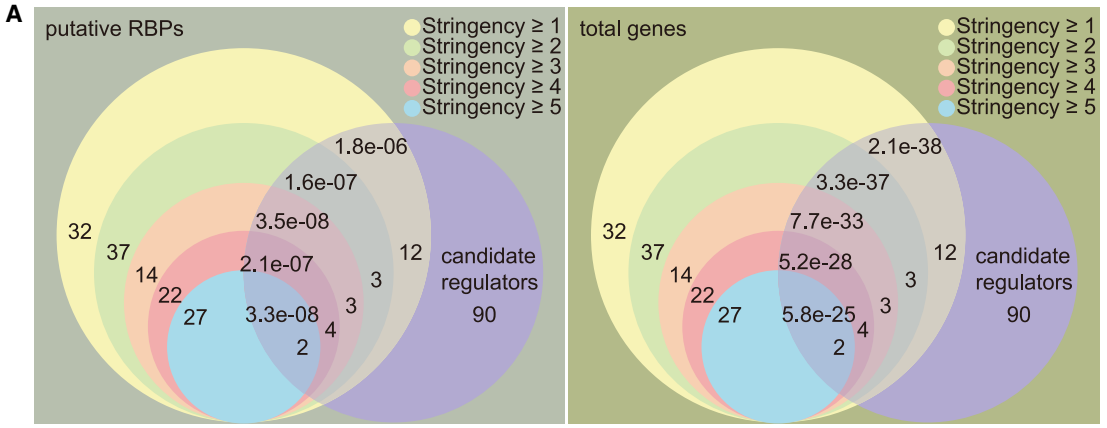


Figure 1. Construction and overview of the CRIT pipeline

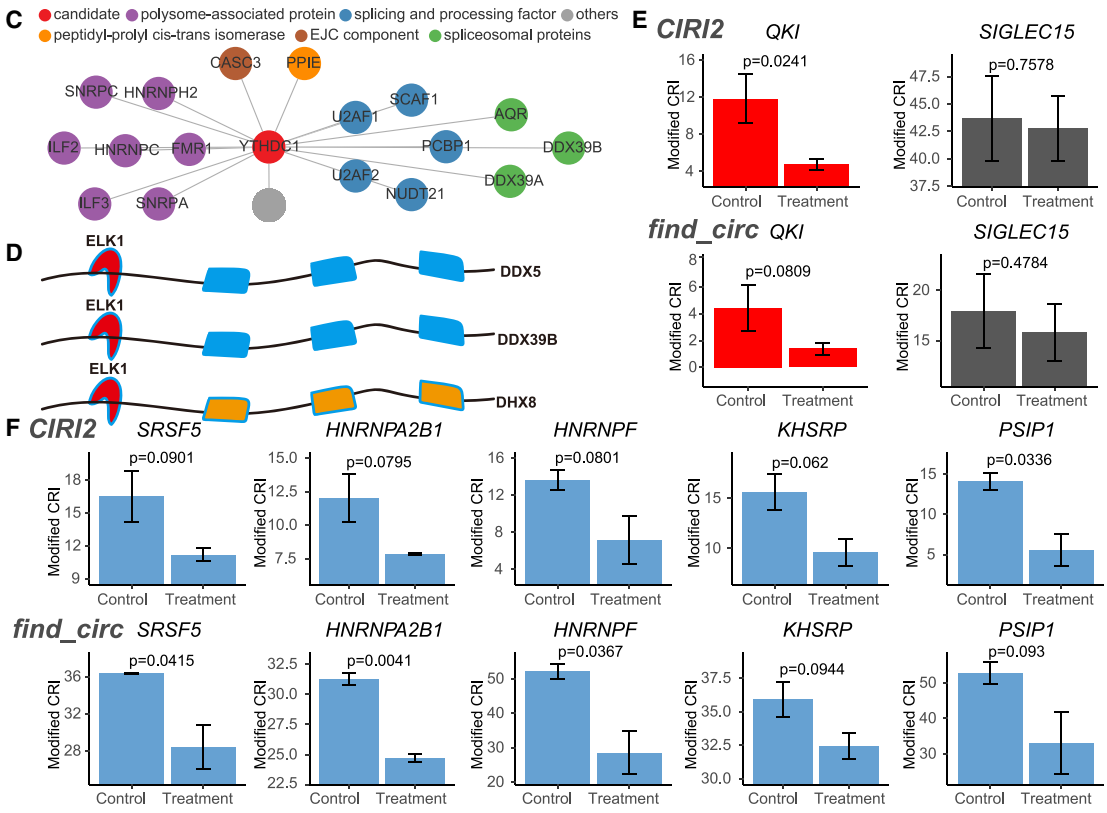
(A) A *ρm* value indicates the Spearman's rank correlation coefficients between genes and circRNAs based on the MiOncoCirc database. (B) Three weighted *Z* scores, which indicate gene alterations, clinical significance, and GO functional information, were developed to evaluate the involvement of each gene in circRNA regulation. (C) C value indicates the number of circRNAs with which each protein interacts. (D) Cluster RBPs with non-negative matrix factorization based on the feature matrix. (E) Four strong pieces of evidence were used to validate candidate regulators. (F) Two weak pieces of evidence were used to validate candidate regulators.



B

Candidate	Gold regulators	Motifs of candidate	Motifs of gold regulators	p-value	E-value	Overlap	LOGO
MFAP1	MRPL44	RCCACUGCAC UCCAGCCUGG	CCACUGCACU CCAGCCUGGG	1.29E-12	3.86E-12	19	
MFAP1	DDX42	RCCACUGCAC UCCAGCCUGG	CCACUSCACU CCAGCCUGGG	5.17E-11	1.55E-10	19	
MFAP1	DDX54	RCCACUGCAC UCCAGCCUGG	CCACUGCACU CCAGCCUGGG	1.33E-10	3.98E-10	19	
MFAP1	IGF2BP3	RCCACUGCAC UCCAGCCUGG	CCACUGCMCU CCAGCCUGGG	1.71E-10	5.14E-10	19	
MFAP1	NOVA2	RCCACUGCAC UCCAGCCUGG	CACACCACUG CACUCCAGCC	5.44E-09	1.63E-08	17	
MFAP1	U2AF2	UCCGCCCACC UCRGCCUCC	UCCUCCCACC UCAGCCUCC	8.74E-09	1.75E-08	19	

Candidate	Gold regulators	Motifs of candidate	Motifs of gold regulators	p-value	E-value	Overlap	LOGO
MFAP1	DUSP11	CCGCCCGCCUCYGCCCGCCU CGGCCUCCC	CGGCCUCCCA	2.69E-07	8.07E-07	19	
MFAP1	FMR1	CCGCCCGCCUCUCCUGCCUC CGGCCUCCC	AGCCUCCCGA	6.09E-05	0.000183	18	
MFAP1	ALKBH5	UCCGCCCACC UCRGCCUCC	CUCCUGCCUC AGCCUCCUGA	0.000418	0.001253	17	
MFAP1	ELAVL1	CCGCCCGCCUGGCGGCGGCG CGGCCUCCC	GCGGCGSSGG	0.001498	0.004494	17	
MFAP1	RBM4B	CCGCCCGCCUAAWMAAAAKB CGGCCUCCC	YWDAAHUU	0.001781	0.005344	16	
MFAP1	DDX5	CCGCCCGCCU CGGCCUCCC	ARCCCCGUYU CUAMNRAAAA	0.001962	0.005887	19	



(legend on next page)

Candidate regulators grouped by non-negative matrix factorization (NMF) are supported by other biological associations

We studied the functional associations between the 73 (excludes 17 gold regulators from 90 RBPs) candidates and the gold circRNA regulators. Among the putative regulators, we found that 57 of them have significantly more physical interactions with the gold regulators (Table S3) (false discovery rate [FDR] < 0.05). For example, the YTH domain containing 1 (YTHDC1) interactome is enriched with 17 gold circRNA regulators (Figure 2C). These interacting proteins include 7 polysome-associated proteins (FMR1, ILF2, ILF3, HNRNPH2, SNRPA, SNRPC, and HNRNPC) and several splicing and processing factors (PCBP1, NUDT21, U2AF1, U2AF2, and SCAF1).[9,33,34] In addition, YTHDC1 is also connected to multiple spliceosomal proteins, such as RNA helicases AQR, DDX39A, and DDX39B and peptidyl-prolyl *cis-trans* isomerase (PPIE). Notably, one core factor of the exon junction complex (EJC) CASC3 also binds to YTHDC1. Previous reports found that the core EJC controls RNA export, translation, and decay in mammalian cells.[35] EJCs multimerize with numerous SR proteins to form a high-order structure and coordinate mRNA processing.[36] Recently, our experiments also revealed that CASC3 directly binds to circNOL10, which inhibits breast cancer progression and metastasis.[37] Furthermore, YTHDC1 is also a well-known m6A "reader" that recognizes m6A modification by YTH domains and mediates the downstream effects such as RNA decay or translation.[38] Finally, a recent study confirmed that YTHDC1 could bind to circ-RHBDD1 and promote its intracellular localization from the nucleus to the cytoplasm.[39] All the above evidence suggested that YTHDC1 bridges methyl-selective RNA binding with a myriad of cellular processes via its interacting partners. As negative controls, we randomly sample 10,000 times from either the RBP backgrounds or the total gene background, respectively. No sampling has an equal to or greater number of RBPs that bind with the gold regulators under both backgrounds ($p < 0.0001$). A large portion of candidates (57 out of 73) have direct binding with known circRNA-binding proteins, which strongly suggested their essential roles in circRNA regulation.

Except for the 17 gold regulators and 57 candidates physically connected with gold regulators, there are only 16 additional predictions. To further discover biological associations, we used the MEME suite to search for motif similarities between these RBPs and gold regulators. As shown in Table S4, we found that all of them (ADARB1, ALKBH8, ARL6IP4, CDK11B, CELF4, CELF5, CSTF2T, GIGYF2, MFAP1, MTO1, NOL10, RBMX2, REXO4, RNF113A, ZC3H12A, ZFC3H1) have consensus motifs of gold circRNA regulators (CASC3, DDX5, DDX54, IGF2BP3, MRPL44, RBM4B, U2AF2, FMR1, SNRPA, DDX42, TARDBP, DUSP11, ELAVL1, ALKBH5, NOVA2, HNRNPC, RBM41) with statistical significance. For instance, the binding motifs of MFAP1 are similar to those of 12 gold regulators (MRPL44, DDX42, DDX54, IGF2BP3, NOVA2, U2AF2, DUSP11, FMR1, ALKBH5, ELAVL1, RBM4B, DDX5) (Figure 2B; $E < 0.01$). This observation raised the possibility that MFAP1 may also bind similar targets including circRNA sequences.

In addition, we also found two weak associations with gold regulators. We found that 14 candidate RBPs have the same transcriptional factors (TFs) as the gold circRNA regulators (Table S3; FDR < 0.05). When compared with random sampling from two background controls, it is more significant than from the total gene control ($p = 0.0001$) but not from the RBP control ($p = 0.265$). During the metabolism of circRNAs, multiple RBPs coordinate the processing of transcripts, therefore it requires an upstream signal such as a TF to finely modulate and drive the production of necessary proteins.[40] For instance, analysis of the curated TF targets from TRANSFAC found that DEAH-box helicase 8 (DHX8) shared a common transcription factor ELK1 with several RNA helicase members such as DDX39B and DDX5 (Figure 2D). In *Drosophila*, depletion of Hel25E, the homolog of DDX39B, leads to nuclear accumulation of long circRNAs, suggesting its role in circRNA transportation.[28] Previous reports indicated that both DDX39B and DDX5 are highly expressed and involved in diverse cancers.[41–43] ELK1 is an established master regulator under tumor hypoxia.[44] Although currently, little is known about the function of DHX8, it is reasonable to propose that this gene is also induced by ELK1 and plays a role via circRNA-mediated mechanism in cancers. With a cutoff $p < 0.05$, we additionally found that 5 candidate RBPs (RBBP6, RPL17, RPS12, SRSF5, and THRAP3) are co-expressed with gold regulators (Table S3). There are no random samplings greater than or equal to 5 in both background controls ($p < 0.0001$). For instance, seven gold regulator genes (EXOSC8, FUS, HNRNPA1, NOL8, RBM25, TARDBP, THOC2) are co-expressed with the putative regulator RBBP6, therefore these RBPs may interplay to constitute a circRNA-regulatory module. Overall, above findings revealed that candidate regulators are more closely connected with gold regulators, which is significant when compared with the random samplings from negative control.

Candidate regulators grouped by NMF are supported by knockdown experiments

To discover the experimental evidence for candidate regulators, we comprehensively examined publicly available RNA sequencing (RNA-seq) datasets after the depletion of the corresponding RBP. Though these experiments were not originally designed to investigate

Figure 2. Multiple evidence supported candidate regulators

(A) circRNA-binding RBPs revealed from CLIP experiments of five stringencies are used to validate candidate regulators with either RBPs or total gene background. (B) MFAP1 has consensus motifs of 12 gold regulators. Significant measures and binding motif logos were computed by the MEME suite and TOMTOM tool. (C) The protein-protein interactions (PPIs) of YTHDCC1 are enriched with 17 gold circRNA regulators. (D) Candidate DHX8 shared a common transcription factor ELK1 with gold regulators DDX5 and DDX39B. (E) Knockdown experiments revealed the global impact on circRNA expression. QKI (red) is a well-known circRNA regulator; SIGLEC15 (gray) is a transmembrane protein used as a negative control. (F) SRSF5, HNRNPA2B1, HNRNRF, KHSRP, and PSIP1(blue) are supported by knockdown experiments. Error bars represented standard deviations from the replicated experiments. P values were caculated from grouped t test.

circRNA, the raw RNA-seq data can be reanalyzed to assess the effect of RBP on the global circRNA expression. Due to the availability, only 22 out of 73 newly predicted RBPs have suitable datasets. In addition, sialic acid binding Ig like lectin 15 (SIGLEC15), a transmembrane protein, was randomly chosen as a negative control. We calculated a modified circRNA index (CRI) to compare circRNA expression between control and RBP short hairpin RNA (shRNA)/siRNA knockdown samples. It can be seen that there is no significant change in modified CRI (*m*CRI) after SIGLEC15 knockout, while *m*CRI decreased after siRNA interfering in QKI samples, which is a founder circRNA regulator (Figure 2D). These results confirmed that analyzing knockdown data can reveal whether an RBP plays a role in circRNA processing. Among the 22 putative regulators, 13 RBPs could be found with experimental evidence from knockdown data by at least one circRNA-detecting tool (Table S5). Notably, it was found that circRNAs were significantly inhibited in knockdown samples of SRSF5, HNRNPA2B1, HNRNPF, KHSRP, and PSIP1 by utilizing both tools, which strongly supports that these RBPs are true circRNA regulators (Figure 2F).

Overall, among the 90 predictions from CRIT, there are 17 already known regulators. Of the 73 newly discovered RBPs, all have at least one biological association with gold circRNA regulators. Forty-one of them have more than two supportive confirmations (Table S6), and 29 have more than two strong instances of supportive evidence including physical interactions, knockdown experiments, common motifs, and enhanced CLIP (eCLIP) binding data, which shown that NMF-based clustering can reveal both established and novel circRNA regulators at the same time.

Characteristics of authentic circRNA regulators

We compared the feature set between cirRNA-regulating proteins and other RBPs to discover the key characteristics of authentic circRNA regulators. Among the five categories of features, *Zg* and *Zp* are considered the main features that distinguish if RBPs can regulate circRNAs. *Zg* reflects if the biological function of an RBP is close to the GO term "mRNA splicing, via spliceosome." The true circRNA-regulating RBPs are more likely the core splicing protein and its peripheral factors. *Zp* reflects alterations in an RBP in cancer. Our result indicated that authentic regulators usually harbor more alterations in the vast majority of cancer datasets. There are also minor but distinguishable differences among the other three categories of features (Table S7).

DISCUSSION

Elucidating circRNA's role in cancers is a novel research focus and has rapidly developed recently. An essential step for understanding circRNA functions is to identify key regulators of circRNA in tumorigenesis. In the present study, we developed a computational framework, CRIT, to integrate multiple-omics datasets including gene alteration, clinical significance, known functional information, and transcriptome data from thousands of cancer samples and to predict a catalog of circRNA regulators across 12 cancer types. The regulations among RBPs and circRNA play a crucial role in gene post-transcriptional control and tumorigenesis.[45] How do dysregulated RBP-circRNA interactions elicit pathologic consequences and affect patient mortality and survival? An answer lies in tracing how individual regulations converge on the multi-components circuits formed by RBPs and non-coding RNAs (ncRNAs), and in the understanding of the vulnerability of the cancer system to a variety of perturbations. This highlights the importance of elucidating the pathogenesis of cancer, thus designing better therapeutic interventions through the novel lens of RBP-ncRNA circuits.[46,47]

Last but not least, the successful application of CRIT in cancers suggested it would be interesting to exploit if the strategies developed here may generalize to other diseases such as neurodegenerative disease. Indeed, we recently found that the expression pattern of circRNA inversely correlated between some cancers and Alzheimer's disease, suggesting that they may share some common circRNA-mediated mechanisms and pathogenesis.[48] Such analysis would require further fine-tuning of the mining algorithms and obtaining additional mRNA/circRNA expression profiles under specific conditions. Another potential technical extension of CRIT is to apply supervised-learning algorithms to predict circRNA-regulating RBPs based on known or validated regulators. These efforts will provide new insights into the principles of circRNA regulation and aid in the elucidation of circRNA function in the future.

MATERIALS AND METHODS

Data

Our study is based on various biological resources including TCGA (https://www.cancer.gov/tcga), MiOncoCirc,[49] GO,[50] and RNAct database.[51] First, we downloaded processed gene expression, somatic mutation, CNV, and clinical patient follow-up data from the TCGA database of 12 cancer types, which were chosen because all these data are simultaneously measured and have at least 50 samples for each data type (Table S8). Then, the circRNA and mRNAs expression data for the matching cancer types above all and the sample's annotation data are collected from the MiOncoCirc database. MiOncoCirc is an extensive cancer-centric resource of circRNAs that was constructed from cancer samples (2,000+) across a plethora of disease sites. Protein-RNA interaction data were collected from the RNAct database.[51,52] RNAct assigns an interaction score for each protein-RNA pair with the catRAPID prediction approach, which is trained by X-ray and nuclear magnetic resonance (NMR) structures. According to the original report, we threshold the RNAct binding score as 30 to obtain more convincing protein-circRNA binding sequences.[51] The R biomaRt package (v.2.48.2)[53] was used to download RNA sequence data.

CRIT pipeline

Firstly, CRIT collected a set of biologically meaningful features in circRNA modulation. Spearman's rank correlation coefficients are calculated between RBP coding genes and circRNAs based on expression data from the MiOncoCirc compendium. For each gene, the mean value of absolute correlation coefficients of the top 50 gene-circRNAs pairs, ρm, is computed since it captures the main influence

of this gene on circRNA expression. Then, from the TCGA project, we extracted an RBP-alteration matrix from thousands of cancer samples consisting of 19 batches of 12 cancer types (Table S8). We computed three *Z* scores to capture the contributions of gene alterations *Zp*, clinical significance *Zs*, and GO knowledge *Zg*. Additionally, a *C* value, which reflects the strength of a protein interacting with circRNA, was derived from RNAct protein-circRNA interaction data. This procedure collects 51 biological features used for clustering. Based on these features, we applied an NMF method to identify candidate regulator groups from 1,344 RBPs that were summarized in the previous research.[23] NMF is a well-established clustering procedure that has been successfully applied for RBP grouping.[54–56] The key issue to decompose the feature matrix with NMF is to find a proper value for the rank R (that is, the number of RBP groups). Then, we used cophenetic correlation coefficients (CPCCs) and dispersion coefficients (DCs) (Figure S1A) to quantitatively measure the clustering stability associated with each rank R based on a consensus matrix that is defined as the average connectivity matrix over 10,30,50,80,100 factorization runs. We selected the local maxima 36 for CPCCs and DCs as the potential optimized value for R, from which we calculated the consensus matrix visualizing the robustness of our clustering (Figure S1B). The CRIT algorithm is described in detail in the supplemental methods.

Evidence used to validate the putative regulators

We used various gene-gene associations from the Harmonizome database to confirm our predicted candidate regulator groups (i.e., whether the putative candidate regulator has biological associations with the known gold circRNA regulators)[57] (Figures 1E and 1F). We used the protein-protein physical interactions data from Pathway Commons, gene co-expression data from the Molecular Signatures Database, and curated TF target data from TRANSFAC for assessment.

For each candidate regulator in the predicted group, we calculated the counts of physical interaction with the gold standard regulators from previous research, then we applied Fisher's exact test to see if the known circRNA regulators are significantly enriched in their interacting partners. Multiple statistical tests were controlled by the FDR.[58] For the negative control, we perform random sampling 10,000 times from either the RBP background or the total gene background with the $p = \frac{counts}{10{,}000}$, whereas *counts* is the number of counts more than or equal to the number of candidate regulators. Similar assessments were applied to calculate the co-expression association and shared TFs among the putative regulators and gold standard regulators.

Analyzing the eCLIP datasets

To investigate the consistency between our pipeline and eCLIP experiments, we collected two RBP lists that were supported by eCLIP experiments. First, the mutual interactions among RBPs and circRNAs were extracted from ENCORI database across 2 species (human and mouse) (http://starbase.sysu.edu.cn/).[32] Each interaction has a stringency value ($\geq$1, low; $\geq$2, medium; $\geq$3, high; $\geq$5, strict), where the values indicated the supported CLIP/PARE (parallel analysis of RNA ends) experimental evidence. We obtained RBP lists ranging from 132 to 27 whose stringency value was from low to strict. We also downloaded human RBPs of the CLIPdb module in the POSTAR3 database,[30] which have annotations supported by CLIP-seq data. Then, we manually filtered a list of 112 RBPs, which bound on circRNA junction regions. To measure the agreement between the candidate regulators and the eCLIP experiments, we performed Fisher's exact test between the candidate regulators and the two RBP lists described above. Background genes have been used either as the list of 1,344 RBPs from the review[23] or the 35,371 genes from the R package "org.Hs.eg.db,"[59] respectively.

Binding motif analysis for candidate regulators

De novo motif finding was performed using circRNA sequences collected from the RNAct database. Then, we applied the MEME suite to discover binding motifs of RBP from these sequences.[60] The TOMTOM tool was used to compare the similarity of motifs of regulators among putative regulators and gold standard regulators (Figure 1F). It calculates E values based on the likelihood of seeing the observed amount of similarity between two motifs, corrected for multiple comparisons by using a position weight matrix representation of the motif.[61] Background sequences were randomly selected from the genome, which matched the GC-content distribution and length of the input sequences. The outputs of TOMTOM also include motif logos and a significance index. Here, we used E <0.01 as the significance threshold since the previous report indicated its effectiveness in finding true motif similarity.[61]

Validation of regulators by RNA-seq data after shRNA/siRNA knockdown

We comprehensively searched the GEO database (https://www.ncbi.nlm.nih.gov/gds/) and ENCODE database (https://www.encodeproject.org/) for each of the 73 predicted RBPs to see if there are public available shRNA or siRNA-RBP-seq data. Twenty-seven datasets are used in this project (IDs are listed in Table S5). Since the previous survey indicated that no single circRNA-detecting tool can outperform other methods, we reanalyzed the raw data to predict circRNA by both *CIRI2* and *find_circ* algorithms.[62,63] We calculated an *m*CRI to compare circRNA expression between control and regulator knockdown samples, thus assessing the effect of candidate regulators on the global expression of circRNAs. CRI was previously defined as the number of circRNAs with higher than its mean abundance in the given sample.[64] However, under the same conditions, different sequencing depths have a great impact on the counts of identified circRNAs. Therefore, we eliminated this effect by normalizing the CRI, i.e., $m\text{CRI} = \frac{y_m}{data\ size\ of\ each\ sample}$. After calculating the *m*CRI mean and variance of the "treatment" and "control" groups, respectively, a t test was performed on the two groups. An RBP candidate was regarded as a validated regulator if the *m*CRI was significantly changed when utilizing either of the two circRNA-detecting tools. It was highly reliable that the candidate RBP regulated circRNA if the *m*CRI was significant when utilizing both tools.

Comparison circRNA-regulating proteins with others

To investigate the main differences between cirRNAs-regulating RBPs and other proteins, we used the 90 predicted RBPs and the other 916 RBPs as comparing groups and performed t tests in the 51 features, respectively.

DATA AVAILABILITY

The source code of CRIT can be accessed at https://github.com/BioinformaticsSTU/CRIT.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.omtn.2022.10.015.

ACKNOWLEDGMENTS

This research is supported by the open fund of Guangdong Provincial Key Laboratory of Infectious Diseases and Molecular Immunopathology (GDKL202201).

AUTHOR CONTRIBUTIONS

J.X. conceived the project and designed the experiments; M.S. and S.H. developed the computational pipeline; L.J., Y.C., X.Z., Q.C., and X.G. contributed acquisition and analysis of data. J.X. and M.S. wrote the manuscript. All authors read and approved the final manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Aufiero, S., Reckman, Y.J., Pinto, Y.M., and Creemers, E.E. (2019). Circular RNAs open a new chapter in cardiovascular biology. Nat. Rev. Cardiol. *16*, 503–514. https://doi.org/10.1038/s41569-019-0185-2.

2. Hanan, M., Soreq, H., and Kadener, S. (2017). CircRNAs in the brain. RNA Biol. *14*, 1028–1034. https://doi.org/10.1080/15476286.2016.1255398.

3. Zhao, X., Cai, Y., and Xu, J. (2019). Circular RNAs: biogenesis, mechanism, and function in human cancers. Int. J. Mol. Sci. *20*, 3926. https://doi.org/10.3390/ijms20163926.

4. Kristensen, L.S., Andersen, M.S., Stagsted, L.V.W., Ebbesen, K.K., Hansen, T.B., and Kjems, J. (2019). The biogenesis, biology and characterization of circular RNAs. Nat. Rev. Genet. *20*, 675–691. https://doi.org/10.1038/s41576-019-0158-7.

5. Li, B., Zhu, L., Lu, C., Wang, C., Wang, H., Jin, H., Ma, X., Cheng, Z., Yu, C., Wang, S., et al. (2021). circNDUFB2 inhibits non-small cell lung cancer progression via destabilizing IGF2BPs and activating anti-tumor immunity. Nat. Commun. *12*, 295. https://doi.org/10.1038/s41467-020-20527-z.

6. Zhang, F., Zhao, X., Dong, H., and Xu, J. (2018). circRNA expression analysis in lung adenocarcinoma: comparison of paired fresh frozen and formalin-fixed paraffin-embedded specimens. Biochem. Biophys. Res. Commun. *500*, 738–743. https://doi.org/10.1016/j.bbrc.2018.04.145.

7. Bahn, J.H., Zhang, Q., Li, F., Chan, T.M., Lin, X., Kim, Y., Wong, D.T., and Xiao, X. (2015). The landscape of microRNA, Piwi-interacting RNA, and circular RNA in human saliva. Clin Chem. *61*, 221–230. https://doi.org/10.1373/clinchem.2014.230433.

8. Li, Y., Zheng, Q., Bao, C., Li, S., Guo, W., Zhao, J., Chen, D., Gu, J., He, X., and Huang, S. (2015). Circular RNA is enriched and stable in exosomes: a promising biomarker for cancer diagnosis. Cell Res. *25*, 981–984. https://doi.org/10.1038/cr.2015.82.

9. Aktaş, T., Avşar Ilık, İ., Maticzka, D., Bhardwaj, V., Pessoa Rodrigues, C., Mittler, G., Manke, T., Backofen, R., and Akhtar, A. (2017). DHX9 suppresses RNA processing defects originating from the Alu invasion of the human genome. Nature *544*, 115–119. https://doi.org/10.1038/nature21715.

10. Conn, S.J., Pillman, K.A., Toubia, J., Conn, V.M., Salmanidis, M., Phillips, C.A., Roslan, S., Schreiber, A.W., Gregory, P.A., and Goodall, G.J. (2015). The RNA binding protein quaking regulates formation of circRNAs. Cell *160*, 1125–1134. https://doi.org/10.1016/j.cell.2015.02.014.

11. Ashwal-Fluss, R., Meyer, M., Pamudurti, N.R., Ivanov, A., Bartok, O., Hanan, M., Evantal, N., Memczak, S., Rajewsky, N., and Kadener, S. (2014). circRNA biogenesis competes with pre-mRNA splicing. Mol. Cell *56*, 55–66. https://doi.org/10.1016/j.molcel.2014.08.019.

12. Errichelli, L., Dini Modigliani, S., Laneve, P., Colantoni, A., Legnini, I., Capauto, D., Rosa, A., De Santis, R., Scarfo, R., Peruzzi, G., et al. (2017). FUS affects circular RNA expression in murine embryonic stem cell-derived motor neurons. Nat. Commun. *8*, 14741. https://doi.org/10.1038/ncomms14741.

13. Ivanov, A., Memczak, S., Wyler, E., Torti, F., Porath, H.T., Orejuela, M.R., Piechotta, M., Levanon, E.Y., Landthaler, M., Dieterich, C., et al. (2015). Analysis of intron sequences reveals hallmarks of circular RNA biogenesis in animals. Cell Rep. *10*, 170–177. https://doi.org/10.1016/j.celrep.2014.12.019.

14. Zhu, Y.J., Zheng, B., Luo, G.J., Ma, X.K., Lu, X.Y., Lin, X.M., Yang, S., Zhao, Q., Wu, T., Li, Z.X., et al. (2019). Circular RNAs negatively regulate cancer stem cells by physically binding FMRP against CCAR1 complex in hepatocellular carcinoma. Theranostics *9*, 3526–3540. https://doi.org/10.7150/thno.32796.

15. Yang, Y., Fan, X., Mao, M., Song, X., Wu, P., Zhang, Y., Jin, Y., Yang, Y., Chen, L.L., Wang, Y., et al. (2017). Extensive translation of circular RNAs driven by N(6)-methyladenosine. Cell Res. *27*, 626–641. https://doi.org/10.1038/cr.2017.31.

16. Zhang, L., Hou, C., Chen, C., Guo, Y., Yuan, W., Yin, D., Liu, J., and Sun, Z. (2020). The role of N(6)-methyladenosine (m(6)A) modification in the regulation of circRNAs. Mol. Cancer *19*, 105. https://doi.org/10.1186/s12943-020-01224-3.

17. Legnini, I., Di Timoteo, G., Rossi, F., Morlando, M., Briganti, F., Sthandier, O., Fatica, A., Santini, T., Andronache, A., Wade, M., et al. (2017). Circ-ZNF609 is a circular RNA that can Be translated and functions in myogenesis. Mol. Cell *66*, 22–37.e9. https://doi.org/10.1016/j.molcel.2017.02.017.

18. Konig, J., Zarnack, K., Luscombe, N.M., and Ule, J. (2012). Protein-RNA interactions: new genomic technologies and perspectives. Nat. Rev. Genet. *13*, 77–83. https://doi.org/10.1038/nrg3141.

19. Li, X., Liu, C.X., Xue, W., Zhang, Y., Jiang, S., Yin, Q.F., Wei, J., Yao, R.W., Yang, L., and Chen, L.L. (2017). Coordinated circRNA biogenesis and function with NF90/NF110 in viral infection. Mol. Cell *67*, 214–227.e7. https://doi.org/10.1016/j.molcel.2017.05.023.

20. Liang, D., Tatomer, D.C., Luo, Z., Wu, H., Yang, L., Chen, L.L., Cherry, S., and Wilusz, J.E. (2017). The output of protein-coding genes shifts to circular RNAs when the pre-mRNA processing machinery is limiting. Mol. Cell *68*, 940–954.e3. https://doi.org/10.1016/j.molcel.2017.10.034.

21. Wang, Z., and Lei, X. (2021). Prediction of RBP binding sites on circRNAs using an LSTM-based deep sequence learning architecture. Brief Bioinform. https://doi.org/10.1093/bib/bbab342.

22. Yang, Y., Hou, Z., Ma, Z., Li, X., and Wong, K.C. (2021). iCircRBP-DHN: identification of circRNA-RBP interaction sites using deep hierarchical network. Brief Bioinform. 22. https://doi.org/10.1093/bib/bbaa274.

23. Neelamraju, Y., Hashemikhabir, S., and Janga, S.C. (2015). The human RBPome: from genes and proteins to human disease. J Proteomics. *127*, 61–70. https://doi.org/10.1016/j.jprot.2015.04.031.

24. Licatalosi, D.D., and Darnell, R.B. (2010). RNA processing and its regulation: global insights into biological networks. Nat. Rev. Genet. *11*, 75–87. https://doi.org/10.1038/nrg2673.

25. Patop, I.L., Wust, S., and Kadener, S. (2019). Past, present, and future of circRNAs. EMBO J. *38*, e100836. https://doi.org/10.15252/embj.2018100836.

26. Li, X., Ding, J., Wang, X., Cheng, Z., and Zhu, Q. (2019). NUDT21 regulates circRNA cyclization and ceRNA crosstalk in hepatocellular carcinoma. Oncogene. https://doi.org/10.1038/s41388-019-1030-0.

27. Yang, Z., Qu, C.B., Zhang, Y., Zhang, W.F., Wang, D.D., Gao, C.C., Ma, L., Chen, J.S., Liu, K.L., Zheng, B., et al. (2019). Dysregulation of p53-RBM25-mediated

circAMOTL1L biogenesis contributes to prostate cancer progression through the circAMOTL1L-miR-193a-5p-Pcdha pathway. Oncogene *38*, 2516–2532. https://doi.org/10.1038/s41388-018-0602-8.

28. Huang, C., Liang, D., Tatomer, D.C., and Wilusz, J.E. (2018). A length-dependent evolutionarily conserved pathway controls nuclear export of circular RNAs. Genes Dev. *32*, 639–644. https://doi.org/10.1101/gad.314856.118.

29. Hansen, T.B., Wiklund, E.D., Bramsen, J.B., Villadsen, S.B., Statham, A.L., Clark, S.J., and Kjems, J. (2011). miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA. EMBO J. *30*, 4414–4422. https://doi.org/10.1038/emboj.2011.359.

30. Zhao, W., Zhang, S., Zhu, Y., Xi, X., Bao, P., Ma, Z., Kapral, T.H., Chen, S., Zagrovic, B., Yang, Y.T., et al. (2022). POSTAR3: an updated platform for exploring post-transcriptional regulation coordinated by RNA-binding proteins. Nucleic Acids Res. *50*. D287–d294. https://doi.org/10.1093/nar/gkab702.

31. Zhang, M., Wang, T., Xiao, G., and Xie, Y. (2020). Large-Scale profiling of RBP-circRNA interactions from public CLIP-seq datasets. Genes (Basel) *11*. https://doi.org/10.3390/genes11010054.

32. Li, J.H., Liu, S., Zhou, H., Qu, L.H., and Yang, J.H. (2014). starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. Nucleic Acids Res. *42*, D92–D97. https://doi.org/10.1093/nar/gkt1248.

33. De, I., Bessonov, S., Hofele, R., dos Santos, K., Will, C.L., Urlaub, H., Lührmann, R., and Pena, V. (2015). The RNA helicase Aquarius exhibits structural adaptations mediating its recruitment to spliceosomes. Nat. Struct. Mol. Biol. *22*, 138–144. https://doi.org/10.1038/nsmb.2951.

34. Rajiv, C., and Davis, T.L. (2018). Structural and functional insights into human nuclear cyclophilins. Biomolecules *8*. https://doi.org/10.3390/biom8040161.

35. Tange, T., Nott, A., and Moore, M.J. (2004). The ever-increasing complexities of the exon junction complex. Curr. Opin. Cell Biol. *16*, 279–284. https://doi.org/10.1016/j.ceb.2004.03.012.

36. Singh, G., Kucukural, A., Cenik, C., Leszyk, J.D., Shaffer, S.A., Weng, Z., and Moore, M.J. (2012). The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus. Cell *151*, 750–764. https://doi.org/10.1016/j.cell.2012.10.007.

37. Cai, Y., Zhao, X., Chen, D., Zhang, F., Chen, Q., Shao, C.C., Ouyang, Y.X., Feng, J., Cui, L., Chen, M., et al. (2021). circ-NOL10 regulated by MTDH/CASC3 inhibits breast cancer progression and metastasis via multiple miRNAs and PDCD4. Mol. Ther. Nucleic Acids *26*, 773–786. https://doi.org/10.1016/j.omtn.2021.09.013.

38. Roundtree, I.A., Evans, M.E., Pan, T., and He, C. (2017). Dynamic RNA modifications in gene expression regulation. Cell *169*, 1187–1200. https://doi.org/10.1016/j.cell.2017.05.045.

39. Wu, A., Hu, Y., Xu, Y., Xu, J., Wang, X., Cai, A., Liu, R., Chen, L., and Wang, F. (2021). Methyltransferase-like 3-mediated m6A methylation of Hsa_circ_0058493 accelerates hepatocellular carcinoma progression by binding to YTH domain-containing protein 1. Front. Cell Dev. Biol. *9*, 762588. https://doi.org/10.3389/fcell.2021.762588.

40. Wolin, S.L., and Maquat, L.E. (2019). Cellular RNA surveillance in health and disease. Science *366*, 822–827. https://doi.org/10.1126/science.aax2957.

41. Xu, Z., Li, X., Li, H., Nie, C., Liu, W., Li, S., Liu, Z., Wang, W., and Wang, J. (2020). Suppression of DDX39B sensitizes ovarian cancer cells to DNA-damaging chemotherapeutic agents via destabilizing BRCA1 mRNA. Oncogene *39*, 7051–7062. https://doi.org/10.1038/s41388-020-01482-x.

42. He, C., Li, A., Lai, Q., Ding, J., Yan, Q., Liu, S., and Li, Q. (2021). The DDX39B/FUT3/TGFβR-I axis promotes tumor metastasis and EMT in colorectal cancer. Cell Death Dis. *12*, 74. https://doi.org/10.1038/s41419-020-03360-6.

43. Ali, M.A.M. (2021). The DEAD-box protein family of RNA helicases: sentinels for a myriad of cellular functions with emerging roles in tumorigenesis. Int. J. Clin. Oncol. *26*, 795–825. https://doi.org/10.1007/s10147-021-01892-1.

44. Aprelikova, O., Wood, M., Tackett, S., Chandramouli, G.V., and Barrett, J.C. (2006). Role of ETS transcription factors in the hypoxia-inducible factor-2 target gene selection. Cancer Res. *66*, 5641–5647. https://doi.org/10.1158/0008-5472.Can-05-3345.

45. Huang, A., Zheng, H., Wu, Z., Chen, M., and Huang, Y. (2020). Circular RNA-protein interactions: functions, mechanisms, and identification. Theranostics *10*, 3503–3517. https://doi.org/10.7150/thno.42174.

46. Jiang, L., Chen, Q., Bei, M., Shao, M., and Xu, J. (2021). Characterizing the tumor RBP-ncRNA circuits by integrating transcriptomics, interactomics and clinical data. Comput. Struct. Biotechnol. J. *19*, 5235–5245. https://doi.org/10.1016/j.csbj.2021.09.019.

47. Mohibi, S., Chen, X., and Zhang, J. (2019). Cancer the'RBP'eutics-RNA-binding proteins as therapeutic targets for cancer. Pharmacol Ther. *203*, 107390. https://doi.org/10.1016/j.pharmthera.2019.07.001.

48. Chen, D., Hao, S., and Xu, J. (2021). Revisiting the relationship between alzheimer's disease and cancer with a circRNA perspective. Front. Cell Dev. Biol. *9*, 647197. https://doi.org/10.3389/fcell.2021.647197.

49. Vo, J.N., Cieslik, M., Zhang, Y., Shukla, S., Xiao, L., Zhang, Y., Wu, Y.M., Dhanasekaran, S.M., Engelke, C.G., Cao, X., et al. (2019). The landscape of circular RNA in cancer. Cell *176*, 869–881.e13. https://doi.org/10.1016/j.cell.2018.12.021.

50. (2021). The Gene Ontology resource: enriching a GOld mine. Nucleic Acids Res. *49*. D325–d334. https://doi.org/10.1093/nar/gkaa1113.

51. Lang, B., Armaos, A., and Tartaglia, G.G. (2019). RNAct: protein-RNA interaction predictions for model organisms with supporting experimental data. Nucleic Acids Res. *47*, D601–D606. https://doi.org/10.1093/nar/gky967.

52. Armaos, A., Colantoni, A., Proietti, G., Rupert, J., and Tartaglia, G.G. (2021). catRAPID omics v2.0: going deeper and wider in the prediction of protein-RNA interactions. Nucleic Acids Res. *49*. W72–w79. https://doi.org/10.1093/nar/gkab393.

53. Durinck, S., Moreau, Y., Kasprzyk, A., Davis, S., De Moor, B., Brazma, A., and Huber, W. (2005). BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. Bioinformatics *21*, 3439–3440. https://doi.org/10.1093/bioinformatics/bti525.

54. Lee, D.D., and Seung, H.S. (1999). Learning the parts of objects by non-negative matrix factorization. Nature *401*, 788–791. https://doi.org/10.1038/44565.

55. Stražar, M., Žitnik, M., Zupan, B., Ule, J., and Curk, T. (2016). Orthogonal matrix factorization enables integrative analysis of multiple RNA binding proteins. Bioinformatics *32*, 1527–1535. https://doi.org/10.1093/bioinformatics/btw003.

56. Li, Y.E., Xiao, M., Shi, B., Yang, Y.T., Wang, D., Wang, F., Marcia, M., and Lu, Z.J. (2017). Identification of high-confidence RNA regulatory elements by combinatorial classification of RNA-protein binding sites. Genome Biol. *18*, 169. https://doi.org/10.1186/s13059-017-1298-8.

57. Rouillard, A.D., Gundersen, G.W., Fernandez, N.F., Wang, Z., Monteiro, C.D., McDermott, M.G., and Ma'ayan, A. (2016). The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins. Database *2016*. https://doi.org/10.1093/database/baw100.

58. Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. Roy. Stat. Soc. B *57*, 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

59. Carlson, M. (2019). org.Hs.eg.db: Genome wide annotation for Human. R package version 382.

60. Bailey, T.L., and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc. Int. Conf. Intell. Syst. Mol. Biol. *2*, 28–36.

61. Gupta, S., Stamatoyannopoulos, J.A., Bailey, T.L., and Noble, W.S. (2007). Quantifying similarity between motifs. Genome Biol. *8*, R24. https://doi.org/10.1186/gb-2007-8-2-r24.

62. Gao, Y., Zhang, J., and Zhao, F. (2018). Circular RNA identification based on multiple seed matching. Brief Bioinform. *19*, 803–810. https://doi.org/10.1093/bib/bbx014.

63. Memczak, S., Jens, M., Elefsinioti, A., Torti, F., Krueger, J., Rybak, A., Maier, L., Mackowiak, S.D., Gregersen, L.H., Munschauer, M., et al. (2013). Circular RNAs are a large class of animal RNAs with regulatory potency. Nature *495*, 333–338. https://doi.org/10.1038/nature11928.

64. Chen, S., Huang, V., Xu, X., Livingstone, J., Soares, F., Jeon, J., Zeng, Y., Hau, J.T., Petricca, J., Guo, H., et al. (2019). Widespread and functional RNA circularization in localized prostate cancer. Cell *176*, 831–843.e22. https://doi.org/10.1016/j.cell.2019.01.025.