
Review

Why circRNAs are Abnormally Expressed in Tumor? – A Mini Review for the circRNA Biogenesis in Cancer

Ying Huang¹, Qubo Zhu¹ *

¹ Xiangya School of Pharmaceutical Sciences in Central South University, Changsha 410013, Hunan, China;

* Correspondence: qubozhu@csu.edu.cn (Qubo Zhu). Tel.: +86-18670318650.

Simple Summary: Circular RNAs (circRNAs) are circular non-coding RNA molecules without 5' cap and 3' poly(A) tail structure, which play an important role in tumor development, invasion and metastasis, *etc.* However, the mechanism of circRNA dysregulation in cancer remains unclear. Different from the classic splicing of linear RNA, circRNA is formed by back-splicing and is regulated by many cis-acting elements and trans-acting proteins. Exploring how the dysregulation of cis-regulatory elements and trans-acting proteins in tumor cells affects the biogenesis of circRNA, which in turn affects the development and prognosis of cancer, is of great significance for circRNA to become a cancer biomarker and therapeutic target.

Abstract: Circular RNAs (circRNAs), which are a class of non-coding RNA with covalently closed loops, play important roles in epigenetics regulation of gene expression at both the transcriptional and post-transcriptional level. Accumulating evidence demonstrated that numerous circRNAs were abnormally expressed in tumors and their dysregulation was involved in the tumorigenesis and metastasis of cancer. Although the functional mechanisms of many circRNAs have been revealed, why circRNAs are dysregulated in cancer remains elusive. CircRNAs are generated by a “backsplicing” process, which is regulated by different cis-regulatory elements and trans-acting proteins. Therefore, how these cis- and trans-elements change during tumorigenesis and how they regulate the biogenesis of circRNAs in cancer are two questions that interest us. In this review, we summarized the pathways for the biogenesis of circRNAs; and then illustrated why circRNAs dysregulated in cancer by discussing the changes of cis-regulatory elements and trans-acting proteins that related to circRNA splicing and maturation in cancer.

Keywords: circular RNAs (circRNAs); biogenesis; trans-acting proteins; cis-regulatory elements

1. Introduction:

Circular RNAs (circRNAs) are a class of endogenous noncoding RNAs (ncRNA), which form a covalently closed continuous loop structure without 5'-3' polarity nor polyA tail (1). As an evolutionarily conserved ncRNA across phyla, circRNA plays essential roles in epigenetics regulation of gene expression at both the transcriptional and post-transcriptional level (2, 3), such as acting as a 'sponge' through competitive binding of miRNA (4), forming a RNA-protein complexes that regulate gene transcription (5, 6), interacting with RNA-binding proteins as protein scaffolds (7), and even encoding proteins (8). Growing evidence has confirmed that circRNAs involve in the tumorigenesis and metastasis of cancer (9). Numerous circRNAs are considered as promising biomarkers for diagnosis and prognosis in different cancers; and some of them are even developed as new therapeutic targets for the cancer treatment (10, 11). Mechanism studies have also gradually revealed the important roles of circRNA in cancer, including cell proliferation and apoptosis, angiogenesis, epithelial to mesenchymal transition (EMT), tumor microenvironment and drug resistance (9). However, why circRNAs are abnormally expressed and how the circRNAs are regulated in tumor remain unknown.

Different from the classic canonical splicing of linear RNA, the biogenesis of circRNAs *via* backsplicing reversely connects the downstream (3') splice donor site to an upstream (5') splice acceptor site covalently (12). Two hypotheses of backsplicing are highly recognized until now: the "lariat-driven circularization" and the "direct back-splicing". The former model starts with canonical splicing for a linear RNA with skipped exons and a long lariat containing circRNA sequence, which is then further back-spliced to form a circRNA, while the latter model starts with back-splicing directly for a circRNA together with an exon-intron(s)-exon intermediate (13). These two processes both require spliceosomal machinery, and are elaborately regulated by both *cis*-regulatory elements and *trans*-acting factors (14). The *cis*-regulatory elements include the splice sites, enhancers, and silencers, especially the intronic complementary sequences (ICSs) near the junction points, such as the inverted *Alu* repeats (15); and the *trans*-acting factors include spliceosome factors, cleavage factors RNA helicases, RNA-binding proteins (RBP) (16). Just like the alternative splicing (AS) and alternative polyadenylation (17) of linear RNA occur frequently with the disordered expression of splicing and maturation factors in cancer (18), aberrant biogenesis of circRNA is always observed in cancer with the dysregulation of spliceosomal machinery, and accumulating evidences indicate that the expression of circRNA is commonly altered in varies malignancies (19).

In this review, we summarized the pathways for the biogenesis of circRNAs; and then illustrated why circRNAs dysregulated in cancer by discussing the changes of *cis*-regulatory elements and *trans*-acting proteins that related to circRNA splicing and maturation in cancer.

2. The biogenesis of circRNAs:

The biogenesis of circRNAs via backsplicing is generally coupled with but different from the canonical splicing of linear RNAs (20). Moreover, it is also distinct from the formation of other types of circular RNAs, such as the intron lariat escaped from debranching (21), the rRNA intermediates (22), and the direct ligation product of single strand RNA (23). Most highly expressed circRNAs are composed of multiple internal exons of pre-mRNAs and contain few introns, indicating that back-splicing and canonical splicing are coupled (20). However, the co-expression of circRNAs and the corresponding putative linear RNAs with exon exclusion is not detected, possibly due to the fast degradation of untranslatable linear RNAs (24). Recent advances even showed that backsplicing and canonical splicing competed each other for the same transcripts (25).

Most scientists believe backsplicing and canonical splicing are coupled process that are both required for the formation of circRNA. According to which step comes first, the processes of circRNA biogenesis can be classified into 2 groups. Once canonical splicing happens first, the RNA precursor will generate a linear RNA with skipped exons and a long lariat containing introns and exons, which is then back-spliced to generate a circRNA (Fig. 1A). We called this process as “exon skipping” or “lariat-driven circularization” model. On the other hand, if the back-splicing happens first, the RNA precursor will generate a circRNA directly together with an intermediate containing both introns and exons, which is further processed to produce a linear RNA (Fig. 1B). This process is referred as “direct back-splicing”. Current researches demonstrate that these two proposed models are both effective in vivo (13, 21). According to the difference of circularization mechanism, “direct back-splicing” process can be divided into 2 groups, “intron-pairing driven circularization” and “RBP mediated circularization”. The former one induces circularization due to proximity between the splice sites by pairing between complementary motifs within the introns flanking the exons (26); while the later one promotes formation of the circle forming junction by bringing the circle-forming exons within proximity of each other through the interaction between RBPs bound with the intronic sequence motifs (25, 27).

In summary, the circRNA biogenesis is a complex process involved spliceosomal machinery, complementary intron pairing, RBP binding, and so on. How these elements regulate circRNA expression as well as how they change during tumorigenesis are of great interest.

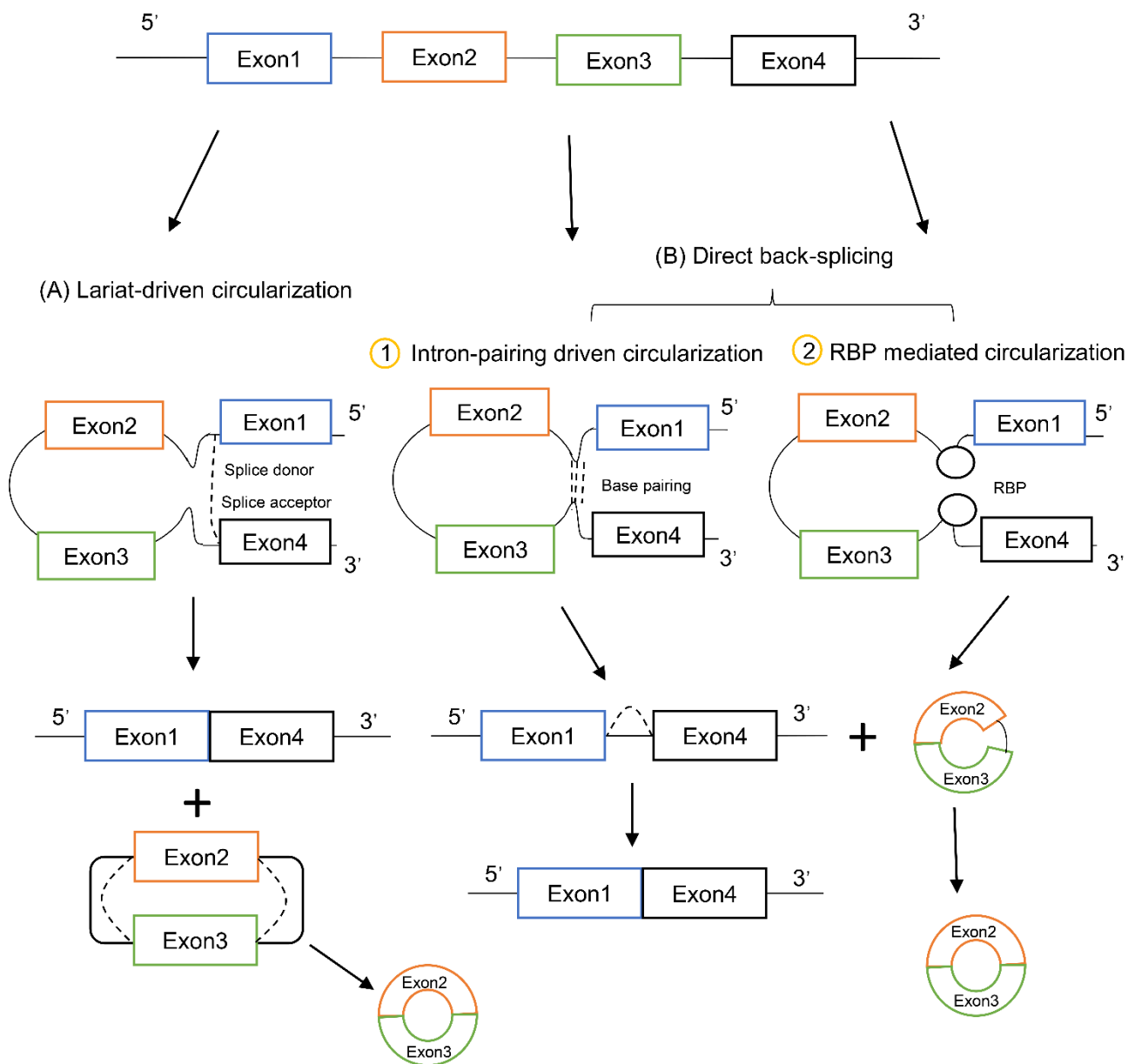


Figure 1. Biogenesis of circRNA. **(A)** Canonical splicing occurs first, exon skipping forms an mRNA composed of exons 1 and 4, and a lariat structure composed of exons 2 and 3, and then back-splicing produces circRNA. **(B)** Back-splicing occurs first, firstly, circRNA containing exons and introns is generated through back-splicing, and then further processed to generate linear mRNA. It can be divided into two types, one is (1) "intron-pairing driven circularization", the complementary sequences of the flanking introns pair with each other to bring the splice sites closer to promote circularization, and the other is (2) "RBP mediated circularization", RBP interacts with the intron binding sites to make the splice sites closer to induce circularization.

3. Trans-acting proteins and RNAs:

A variety of trans-acting proteins have been reported to be involved in the biogenesis of circRNAs. By regulating the splicing process of pre-mRNA, it affects the production and expression of circRNA at the post-transcriptional level, and further affects the proliferation and migration of tumor cells and the prognosis of cancer patients. According to the different mechanisms of trans-acting factors, we divide them into four categories: spliceosome factors, cleavage factors, RNA helicases, and RNA-binding proteins; and explain separately to illustrate the biogenesis process of circRNA and its influence on cancer.

3.1. Spliceosome factors:

The spliceosome is a large complex composed of 5 small nuclear RNAs (snRNAs), which combine with proteins to form particles, called small nuclear ribonucleoproteins (snRNPs). In eukaryotes, it is composed of U1, U2, U5 and U4, U6 snRNPs and a large number of proteins (28, 29). Both the pre-mRNA exons and their flanking introns contain special snRNP protein binding sites, and snRNP regulates the production of circRNA by specific interaction with the binding sites (30). Spliceosomes can catalyze the splicing events of most transcripts and remove the most common type of introns (31).

U2 auxiliary factor (U2AF) is a spliceosome factor and a non-snRNP protein, which is necessary for the binding of U2 snRNP to pre-mRNA branch sites (32). U2AF2 acts as an oncogene in several cancers, such as glioma, primary non-small cell lung cancer, and melanoma (33-35). In glioma, the expressions of U2AF2 and circRNA ARF1 (cARF1) are both up-regulated, and they are associated with decreased survival rates of patients. Mechanism studies show that U2AF2 promotes the proliferation, invasion and angiogenesis of human brain microvascular endothelial cells (HBMECs) by up-regulating the expression of cARF1 in gliomas (36).

Studies have found that changes in the expression levels of many core components of the spliceosome can have a profound impact on the level of circRNA production. For example, an important part of the spliceosome U2 snRNP (SF3B1 or SF3A1), snRNP-U1-70K and snRNP-U1-C involve in 5' splice site recognition, and the largest protein component of the spliceosome cross-links with U5 and U6 snRNAs (prp8), etc. When they are exhausted, the splicing products of pre-mRNA will shift to the direction of circRNA production, and the production of linear RNA will decrease (16). Among them, the mutation of SF3B1 is considered to be a common driving factor of hematological malignancies (37). Mutations in SF3B1 are found in most patients with myelodysplastic syndrome (MDS) (38, 39). The mutation of SF3B1 has also become an important factor in a variety of pigmented tumors, such as uveal melanoma (UM) (40), leptomeningeal melanoma (41), and blue nevus-like cutaneous melanoma (42). In MDS and UM, compared with SF3B1 wild-type patients, patients with SF3B1 mutations have a higher 5-year survival rate and a better cancer prognosis (39, 40). However, there is no direct evidence that the spliceosomes affect the procession of cancer by affecting the production of circRNA, and it will be the direction of continued research and exploration in the future.

3.2. Cleavage factors:

According to the reports, several cleavage factors can directly cut the adjacent intron sequences of exons forming circRNA to promote the formation of circRNA (43).

For example, the cleavage factor ESRP1 can evenly target circRNA GGT-rich sequences to accelerate the circRNA biogenesis, such as circANKS1B and circUHRF1. Clinically, in breast cancer, the expression of ESRP1 is positively correlated with circANKS1B. The high expression of circANKS1B is closely related to the clinical stage and is an independent risk factor for the overall survival rate of breast cancer patients. At the same time, breast cancer patients with high expression of ESRP1 also have a poor overall survival (44). In oral squamous cell carcinoma (OSCC), circUHRF1 is significantly up-regulated. The overexpression of circUHRF1 is closely related to the poor prognosis of OSCC patients, and promotes tumor proliferation, migration, invasion and epithelial mesenchymal transformation (EMT) (45).

The cleavage factor NUDT21 is a part of the cleavage factor Im (CFIm) complex. In the form of a dimer, it combines with two UGUA sequences upstream of poly(A) to form a loop to prevent the splicing of pre-mRNA by the CPSF family, thereby promoting the formation of circRNA (46). In hepatocellular carcinoma (HCC), the expression level of circRNA is generally down-regulated, and the level of NUDT21 is also significantly down-regulated, HCC patients with low levels of NUDT21 have poor overall survival (47). In short, the down-regulation of NUDT21 in HCC inhibits the production of circRNA containing UGUA sequence, which in turn leads to the imbalance of the proliferation of HCC cells (48).

As the core component of the 3'-end splicing complex, Cleavage and Polyadenylation Specific Factor 4 (CPSF4) also has a regulatory effect on the formation of circular RNA. It can specifically bind to the PAS regulatory elements on the precursor mRNA, and mediates the cleavage of the CPSF family complex, thereby inhibiting the formation of circular RNA and affecting the expression of the formed circular RNA. The high expression of CPSF4 in HCC also corresponds to the overall down-regulation of circular RNA. Further studies demonstrated that overexpression of CPSF4 could promote the proliferation of HCC cells, and was closely related to the clinicopathological characteristics of HCC patients (49).

3.3. RNA helicases:

RNA helicases, a family of proteins that promote unwinding of RNA during splicing and translation, play important roles in circRNA biogenesis, because unwinding of the paired RNA sequence is a critical process for RNA circulation.

DExH-box helicase 9 (DHX9) is a nuclear RNA helicase that can interact with inverted complementary sequences and unwind RNA pairs flanking circularized exons to prevent the production of circRNAs (50). In lung adenocarcinoma (LUAD), overexpressed CircDCUN1D4 significantly inhibits tumor metastasis, reduces glycolysis, and can be used as an independent prognostic factor for LUAD patients. Researchers found that DHX9 interacted with AluJo/AluSc binding sites to reduce the expression of circDCUN1D4. The expression of DHX9 was negatively correlated with the expression of circDCUN1D4 in LUAD tissues, and lower circDCUN1D4 level in LUAD was due to the up-regulation of DHX9 (51). In addition, studies have found that DHX9 can negatively

regulates circPICALM levels by binding to its inverted repeats Alu (IRAlus). CircPICALM was downregulated in bladder cancer, and low circPICALM expression was related to advanced T stage, high grade and poor overall survival rate (52). Unlike circDCUN1D4 and circPICALM, DHX9 post-transcriptionally regulated the biogenesis of Circular RNA cSMARCA5 by binding to other nonrepetitive but inverted complementary sequences instead of inverted repeated Alu pairs, such as I14RC (reverse complementary sequences in intron 14) and I16RC (reverse complementary sequences in intron 16). DHX9 was significantly upregulated in HCC and the expression of cSMARCA5 was negatively correlated with the histochemical score of DHX9 in 40 HCC tissues. Furthermore, the tumor suppressor cSMARCA5 plays an important role in the growth and metastasis of HCC, and can be used as a biomarker for the prognosis of patients undergoing liver cancer resection (53).

Eukaryotic initiation factor 4A3 (EIF4A3, DDX48) is a member of the DEAD box family, which is one of the largest helicase families characterized by an Asp-Glu-Ala-Asp (DEAD) motif (54). EIF4A3 does not show helicase activity by itself (55); only when the other 2 core components of exon junction complex (EJC) MLN51 and Magoh/Y14 bind with EIF4A3 together, the ATPase and helicase activity of EIF4A3 are stimulated (56), and the pre-mRNA splicing process starts. The expressions of circMMP9 and circASAP1 are significantly upregulated in glioblastoma multiforme (GBM), and they promote the proliferation, migration, and invasion of GMB. Through the circinteractome database, the author found that EIF4A3 has four binding sites in the upstream region of circMMP9 mRNA transcript, which can promote the circularization of exons 12 and 13, and then promote the expression of circMMP9 (57). Similar to circMMP9, EIF4A3 is also positively correlated with circASAP1. However, EIF4A3 increases the expression of circASAP1 by binding with the downstream flanking sequence of circASAP1 instead of the upstream sequence (58). In triple-negative breast cancer (TNBC), circSEPT9 is significantly up-regulated, and is closely related to the clinical stage and poor prognosis of TNBC patients. Recent researches demonstrated that EIF4A3 increases the expression of circSEPT9 by binding to the SEPT9 pre-mRNA, and modulates cell cycle (59).

ADAR1 is an RNA editing enzyme that can catalyze adenosine (A) in the reverse complementation region of introns to inosine (I), untie the pairing structure, and prevent the splicing site from approaching for reverse splicing, thereby strongly inhibiting the generation of circRNA (60, 61). Studies have shown that ADAR1 is closely related to the poor prognosis of liver cancer. ADAR1 inhibits the production of CircARSP91 in HCC and further promotes the proliferation of HCC (61). ADAR1 overexpression is also one of the reasons for the down-regulation of hsa_circ_0004872 expression in gastric cancer tissues. Has_circ_0004872 acts as a "miRNA sponge" to inhibit the proliferation, invasion and migration of gastric cancer cells, and can serve as a diagnostic and prognostic marker for gastric cancer patients (62). In pancreatic ductal adenocarcinoma (PDAC), circNEIL3 plays a role of oncogene and promotes the progression and metastasis of PDAC. ADAR1 can inhibit the formation of circNEIL3 by catalyzing the A of NEIL3 complementary sequence to G, and affect the clinical stage and survival of PDAC patients (63).

3.4. RNA-binding proteins

Except for the spliceosome factors, cleavage factors and RNA helicases, several RNA-Binding Proteins (RBPs) can shorten the distance between the receptor site and the donor site through interact with the flanking introns of exons, which promotes the production of corresponding circRNAs (64)., such as Heterogeneous nuclear ribonucleoprotein L (HNRNPL), QKI, FUS etc.

HNRNPL can bind to the flanking introns of exons to promote the production of circRNA, especially when it exists on both sides (65). In Gastric cancer (50), the expression of circLMO7 is significantly up-regulated. TCGA database shows that HNRNPL is highly expressed in gastric cancer tissues and is positively correlated with the expression level of circLMO7. Deep exploration shows HNRNPL promotes the production of circLMO7 by interacting with the four binding sites of the flanking introns of the circLMO7 exon (66).

The Quaking (QKI) belongs to the evolutionarily conserved signal transduction and activator of RNA (67) family and contain a single heterogeneous nuclear ribonucleoprotein K homology (KH) RNA-binding domain (68). QKI dimers can bind with two separate RNA recognition elements on a single RNA molecule (69). Therefore, QKI may promote the biogenesis of circRNA by binding with pre-mRNA to bring circ-forming exons closer to each other (70). CircZKSCAN, which is an independent and significant factor for HCC, is positively correlated with the survival rate of patients; and the expression of QKI is positively correlated with the expression of circZKSCAN. The abnormal down-regulation of QKI in HCC tissue may lead to the down-regulation of circZKSCAN1 and accelerated growth of cancer cells(71).

FUS, a well-known proto-oncogene, is reported to regulate the expression of circRNA in murine embryonic stem cell-derived motor neurons (17). In primary colorectal cancer tissues, the expression of CircLONP2 is significantly up-regulated, and the expression of CircLONP2 is positively correlated with the mRNA level of FUS. High expression of FUS is responsible for the low overall survival rate of patients with colorectal cancer, and poor clinicopathological characteristics. Furthermore, the reverse complementary matches in intron 7 of CircLONP2 are composed of repeated "GUUG" or "ACUU" regions, which are just between the binding motifs of FUS. Mechanism studies show that FUS regulates the metastasis of colorectal cancer cells by interaction with the reverse complementary matches of circLONP2 to regulate the biogenesis of circLONP2 (72). In lung cancer, CircZNF609, which functions as an onco-circRNA, is highly expressed, and promotes the proliferation and migration of cancer cells. FUS RNA-binding protein can bind to the upstream of ZNF609 pre-mRNA to induce the formation of CircZNF609 and increase its level in lung cancer (73). In TNBC, circHIF1A is significantly overexpressed and is associated with metastasis, poor prognosis and TNBC subtypes. The RNA-binding protein FUS regulates the biogenesis of circHIF1A by interacting with flanking introns, and significantly promoted TNBC growth and metastasis (74).

4. cis-regulatory elements

In addition to trans-acting factors that regulate circRNA biogenesis, many cis-regulatory elements can also influence the cyclization of RNA. In particular, reverse complementary sequences (including repeated complementary sequences and non-repeated complementary sequences, such as the Alu element) and direct-action sites of some splicing factors (such as the PAS site, UGUA, GGT-rich) play important roles in the generation of circRNA.

4.1. intronic complementary sequences

Bioinformatics analysis shows that there are long introns at both ends of the reverse spliced circRNA exons. According to calculation, the presence of Alu in the flanking introns is highly correlated with the formation of circRNA (26). The pairing of the inverted repeat sequence of the Alu element makes the downstream splice donor and upstream splice acceptor sites close to each other, thereby promoting the occurrence of reverse splicing events (26). Alu elements are inverted repeats on the introns flanking exons, which can play a key role in the formation of circRNA together with trans-acting proteins. For example, DHX9 can combine with the Alu inverted repeats on circDCUN1D4 (51) and CircPICALM (52) to inhibit the generation of circRNA, thereby inhibiting the proliferation, metastasis and invasion of lung cancer and colon cancer respectively, which is of great significance to the prognosis of cancer patients.

The formation of circRNA depends on the pairing ability of complementary sequences, which can be repetitive sequences such as Alu elements, or non-repetitive but complementary sequences. For example, intron 14 and intron 17 of CircLONP2 have a highly matched sequence. Intron 17 is composed of several repeated "GUUG" or "ACUU" regions, which are located exactly within the binding motif between the proto-oncogene FUS. FUS binds to the flanking introns of CircLONP2 to promote its generation (72). In addition, colorectal cancer cells with up-regulated expression of circLONP2 have significantly enhanced migration and invasion capabilities. DHX9 can also bind to the reverse complementary sequence flanking the circSMARCA5 intron to reduce its expression, which is of great significance for inhibiting the proliferation and metastasis of HCC (53). In addition, the longer flanking introns are not necessary for circular RNA formation, but the extended length can introduce more Alu elements that, in turn, promote exon circularization. (20). The distance and competition between the complementary sequences in the flanking of introns will affect alternative splicing, thereby affecting circRNA circularization, and may also lead to the occurrence of alternative reverse splicing at the same locus, resulting in a variety of circRNAs (75).

But not all intron complementary sequences can promote pre-mRNA reverse splicing and exon circularization, the hairpin structure caused by base pairing may inhibit the biogenesis of circRNA (76). For example, the 321 to 343 nucleotide of EPHB4 locus has a low-complexity sequence of poly(A). The binding of the Poly(A) binding protein (PABP) to this region can spatially interfere with its hairpin formation ability, thereby promoting the formation of EPHB4 circRNA. In addition, studies have shown that circEphB4 is up-regulated in glioma cells and promotes the proliferation and glycolysis of glioma cells, which is closely related to the poor prognosis of patients (77).

4.2. splice sites

Both the RBPs and the mutual pairing sequences regulate circRNA generation by bringing upstream splicing receptors and downstream splicing donors close to each other. While, the splicing factor can directly cleave the splice site on the intron to regulate the biogenesis of circRNA.

The polyadenylation site (PAS), which plays an important role in the process of pre-mRNA 3'-UTR cleavage and maturation, plays an important role in regulating the expression of circular RNA. Among them, AAUAAA is the classic PAS motif (78), and other variants include AUAAAA, AUUAAA, CAAUAA, AAAAAA, AUAAAAG, etc (79). The PAS site is usually located 10-30 nucleotides upstream of the splice site. The PAS element can be recognized by the cleavage factors to promote the cleavage of the circRNA and reduce its expression. Recent studies have shown that CPSF4 can inhibit the formation of circRNA by recognizing and cutting the AAUAAA motif on pre-mRNA, thereby inhibiting the proliferation of HCC cells, which is of great significance to the prognosis of HCC patients(49).

The two UGUA elements located upstream of the PAS site can also regulate RNA cyclization. In HCC, the UGUA sequence located at the proximal end of the 3'-UTR is considered to be the splicing and polyadenylation site in the alternative splicing and alternative polyadenylation (17) process, which can bind to the RNA splicing complex protein to regulate the length of the 3'-UTR. And then regulate the APA process(80). Recent studies have shown that the UGUA sequence in the pre-mRNA in HCC is an important cis-regulatory element in the production of circRNA. It can be recognized by the cleavage complex protein and combined with it to form a circular structure, thereby It avoids the shearing of pre-mRNA by the shearing protein and promotes the looping process of circular RNA(48). For example, in HCC tissues, NUDT21 combines with the two "UGUA" sequences upstream of poly(A) in the form of dimers to form a loop, which prevents the cleavage factor from cleavage pre-mRNA and promotes the production of circRNA. So as to inhibit the proliferation of HCC(48, 81).

In addition, studies have shown that the ggt-rich sequence on the flanking intron of the pre-mRNA exon is the binding sequence of the splicing factor ESRP1. ESRP1 can directly bind to the ggt-rich sequence to cut the circRNA, thereby accelerating the circRNA (such as circUHRF1 and circANKS1B) biogenesis, which in turn promotes tumor growth and metastasis, is closely related to the poor prognosis of cancer patients(44, 45).

5. Conclusions

circRNA has been proven to be an important biomarker of cancer in many studies. In this review, starting from the biogenesis mechanism of circRNA, we further explored the causes of circRNA dysregulation in cancer from two aspects of trans-acting factors and cis-acting elements, as well as the relationship with the occurrence, development and prognosis of cancer. For the trans-acting elements, spliceosome binds to the binding site and facilitates the pre-mRNA splicing, cleavage factors cut the pre-mRNA directly, RNA helicases unwrap the paired complementary sequences, and RBPs make the intron complementary sequences close to each other. There are different ways to affect the circularization of circRNA. Among the cis-acting elements, the pairing of intron complementary sequences (Alu elements) and the binding of splice sites (PAS, UGUA, ggt-rich) are essential for the formation of circRNA. Mutations of trans-acting factors and

cis-acting elements may affect the expression of circRNA, which in turn affects the progression of a variety of cancers. However, the current research on the formation mechanism of circRNA and its potential relationship with cancer is still in its infancy. It is believed that with the continuous in-depth study of circRNA regulation mechanism, circRNA is not only a biomarker, but may also become a potential therapeutic target for cancer.

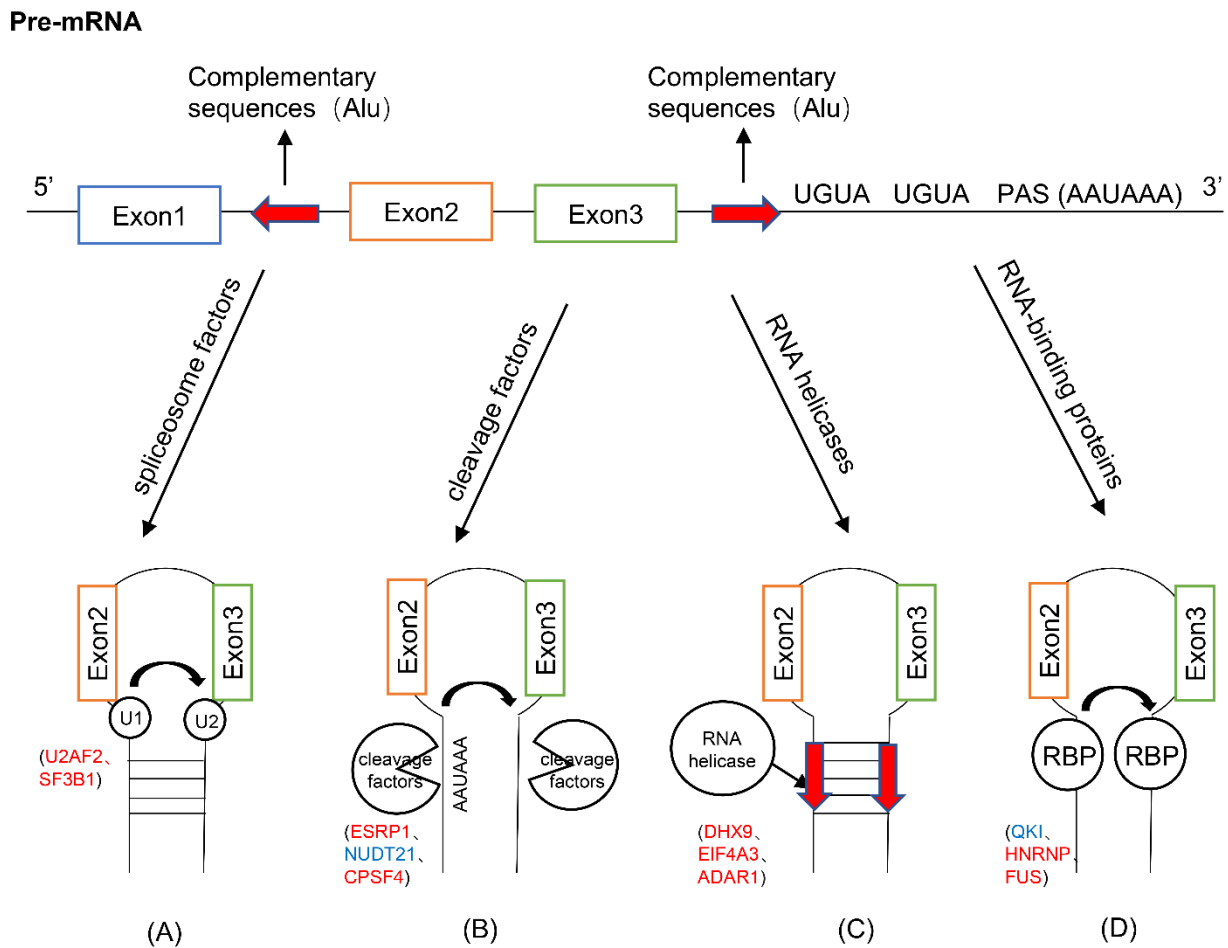


Figure 2. The roles of cis-acting elements and trans-acting factors in circRNA biogenesis. **(A)** The combination of spliceosome factors (U2AF2, SF3B1) with pre-mRNA binding sites. **(B)** Cleavage factors (ESRP1, NUDT21, CPSF4) directly cleave the cis-acting elements (PAS, UGUA) on pre-mRNA. **(C)** RNA helicases (DHX9, EIF4A3, ADAR1) unwind or interfere with intron complementation Sequence (Alu). **(D)** RNA-binding proteins (QKI, HNRNP, FUS) bind to the pre-mRNA binding site to make the splice sites close. For the trans-acting factors, the ones labeled in red indicate increased expression in cancer, and blue ones indicate decreased expression in cancer.

Table 1. Trans-acting proteins and RNAs for RNA circulation

	Proteins/RNA	Target circRNAs	Correlation	cancer	ref.
spliceosome factors	U2AF	cARF1	Positive	Glioma	(36)
cleavage factors	ESRP1	circANKS1B	Positive	Breast cancer	(44)
		circUHRF1	Positive	Oral squamous cell carcinoma	(45)
	NUDT21	circRNAs with UGUA	Positive	Hepatocellular carcinoma	(48)
	CPSF4	circRNAs with PAS	Negative	Hepatocellular carcinoma	(49)
RNA helicases	DHX9	circDCUN1D4	Negative	Lung adenocarcinoma	(51)
		circPICALM	Negative	Bladder cancer	(52)
		cSMARCA5	Negative	Hepatocellular carcinoma	(53)
	EIF4A3	circMMP9	Positive	Glioblastoma multiforme	(57)
		circASAP1	Positive	Glioblastoma multiforme	(58)
		circSEPT9	Positive	Triple-negative breast cancer	(59)
	ADAR1	CircARSP91	Negative	Hepatocellular carcinoma	(61)
		hsa_circ_0004872	Negative	gastric cancer	(62)
RNA-binding proteins	HNRNPL	circNEIL3	Negative	pancreatic ductal adenocarcinoma	(63)
		circLMO7	Positive	Gastric cancer	(66)
	QKI	circZKSCAN	Positive	Hepatocellular carcinoma	(71)
		CircLONP2	Positive	Primary colorectal cancer	(72)
	FUS	CircZNF609	Positive	Lung cancer	(73)
		circHIF1A	Positive	Triple-negative breast cancer	(74)

Funding: This research was funded by Hunan Provincial Natural Science Foundation of China (2018JJ2493), and Independent Exploration and innovation program for Graduate students of Central South University (1053320211988).

Conflicts of Interest: The authors declare no conflict of interest.

Reference:

1. Meng S, Zhou H, Feng Z, Xu Z, Tang Y, Li P, et al. CircRNA: functions and properties of a novel potential biomarker for cancer. *Mol Cancer*. 2017;16(1):94.
2. Qu S, Zhong Y, Shang R, Zhang X, Song W, Kjems J, et al. The emerging landscape of circular RNA in life processes. *RNA Biol*. 2017;14(8):992-9.
3. Ding J, Zhou W, Li X, Sun M, Ding J, Zhu Q. Tandem DNAzyme for double digestion: a new tool for circRNA suppression. *Biol Chem*. 2019;400(2):247-53.
4. Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK, et al. Natural RNA circles function as efficient microRNA sponges. *Nature*. 2013;495(7441):384-8.
5. Wilusz JE, Sharp PA. Molecular biology. A circuitous route to noncoding RNA. *Science*. 2013;340(6131):440-1.
6. Zhang Y, Zhang XO, Chen T, Xiang JF, Yin QF, Xing YH, et al. Circular intronic long noncoding RNAs. *Mol Cell*. 2013;51(6):792-806.
7. Chen L, Shan G. CircRNA in cancer: Fundamental mechanism and clinical potential. *Cancer letters*. 2021;505:49-57.
8. Meng X, Chen Q, Zhang P, Chen M. CircPro: an integrated tool for the identification of circRNAs with protein-coding potential. *Bioinformatics*. 2017;33(20):3314-6.
9. Wang X, Li HJ, Lu YJ, Cheng LM. Circular RNAs in Human Cancer. *Front Oncol*. 2021;10.
10. Zhang HD, Jiang LH, Sun DW, Hou JC, Ji ZL. CircRNA: a novel type of biomarker for cancer. *Breast Cancer*. 2018;25(1):1-7.
11. Wei G, Zhu J, Hu HB, Liu JQ. Circular RNAs: Promising biomarkers for cancer diagnosis and prognosis. *Gene*. 2021;771:145365.
12. Qu S, Yang X, Li X, Wang J, Gao Y, Shang R, et al. Circular RNA: A new star of noncoding RNAs. *Cancer letters*. 2015;365(2):141-8.
13. Chen LL, Yang L. Regulation of circRNA biogenesis. *RNA Biol*. 2015;12(4):381-8.
14. Li X, Yang L, Chen LL. The Biogenesis, Functions, and Challenges of Circular RNAs. *Mol Cell*. 2018;71(3):428-42.
15. Pervouchine DD. Circular exonic RNAs: When RNA structure meets topology. *Bba-Gene Regul Mech*. 2019;1862(11-12).
16. Liang D, Tatomer DC, Luo Z, Wu H, Yang L, Chen LL, et al. The Output of Protein-Coding Genes Shifts to Circular RNAs When the Pre-mRNA Processing Machinery Is Limiting. *Mol Cell*. 2017;68(5):940-54 e3.
17. Lorenzo E, Stefano DM, Pietro L, Alessio C, Ivano L, Davide C, et al. FUS affects circular RNA expression in murine embryonic stem cell-derived motor neurons. *Nature communications*. 2017;8.
18. Gruber AJ, Zavolan M. Alternative cleavage and polyadenylation in health and disease. *Nat Rev Genet*. 2019;20(10):599-614.
19. Li J, Sun D, Pu W, Wang J, Peng Y. Circular RNAs in Cancer: Biogenesis, Function, and Clinical Significance. *Trends Cancer*. 2020;6(4):319-36.
20. Zhang XO, Wang HB, Zhang Y, Lu X, Chen LL, Yang L. Complementary sequence-mediated exon circularization. *Cell*. 2014;159(1):134-47.
21. Lasda E, Parker R. Circular RNAs: diversity of form and function. *RNA (New York, NY)*. 2014;20(12):1829-42.
22. Tang TH, Rozhdestvensky TS, d'Orval BC, Bortolin ML, Huber H, Charpentier B, et al. RNomics in Archaea reveals a further link between splicing of archaeal introns and rRNA processing. *Nucleic acids research*. 2002;30(4):921-30.
23. Flores R, Navarro JA, de la Pena M, Navarro B, Ambros S, Vera A. Viroids with hammerhead ribozymes: Some unique structural and functional aspects with respect to other members of the group. *Biological Chemistry*. 1999;380(7-8):849-54.

24. Jeck WR, Sharpless NE. Detecting and characterizing circular RNAs. *Nat Biotechnol.* 2014;32(5):453-61.
25. Ashwal-Fluss R, Meyer M, Pamudurti NR, Ivanov A, Bartok O, Hanan M, et al. circRNA biogenesis competes with pre-mRNA splicing. *Mol Cell.* 2014;56(1):55-66.
26. Jeck WR, Sorrentino JA, Wang K, Slevin MK, Burd CE, Liu J, et al. Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA (New York, NY).* 2013;19(2):141-57.
27. Conn SJ, Pillman KA, Toubia J, Conn VM, Salmanidis M, Phillips CA, et al. The RNA binding protein quaking regulates formation of circRNAs. *Cell.* 2015;160(6):1125-34.
28. Sahebi M, Hanafi MM, van Wijnen AJ, Azizi P, Abiri R, Ashkani S, et al. Towards understanding pre-mRNA splicing mechanisms and the role of SR proteins. *Gene.* 2016;587(2):107-19.
29. Chen HC, Cheng SC. Functional roles of protein splicing factors. *Bioscience reports.* 2012;32(4):345-59.
30. C KM, Dongming L, C TD, Beth G, M MZ, Sara C, et al. Combinatorial control of Drosophila circular RNA expression by intronic repeats, hnRNPs, and SR proteins. *%J Genes & development.* 2015;29(20).
31. Ishikawa H, Nobe Y, Izumikawa K, Taoka M, Yamauchi Y, Nakayama H, et al. Truncated forms of U2 snRNA (U2-tfs) are shunted toward a novel uridylylation pathway that differs from the degradation pathway for U1-tfs. *RNA Biol.* 2018;15(2):261-8.
32. Zhang P, Feng S, Liu G, Wang H, Fu A, Zhu H, et al. CD82 suppresses CD44 alternative splicing-dependent melanoma metastasis by mediating U2AF2 ubiquitination and degradation *%J Oncogene.* 2016;35(44).
33. Cheng D ZM, Yu H, Pan Z, Liu L, Geng Q, Pan H, Yan M, Yao M. OTUB2 stabilizes U2AF2 to promote the Warburg effect and tumorigenesis via the AKT/mTOR signaling pathway in non-small cell lung cancer *%J Theranostics.* 2019;9(1).
34. Chul-Kee P, Inho P, Seungmook L, Choong-Hyun S, Youngil K, Sung-Hye P, et al. Genomic dynamics associated with malignant transformation in IDH1 mutated gliomas. *%J Oncotarget.* 2015;6(41).
35. Marzese DM, Liu M, Huynh JL, Hirose H, Donovan NC, Huynh KT, et al. Brain metastasis is predetermined in early stages of cutaneous melanoma by CD44v6 expression through epigenetic regulation of the spliceosome *%J Pigment Cell & Melanoma Research.* 2015;28(1).
36. Yang J, Jinpeng Z, Junshuang Z, Haiying Z, Long L, Hao L, et al. The U2AF2 /circRNA ARF1/miR-342-3p/ISL2 feedback loop regulates angiogenesis in glioma stem cells. *%J Journal of experimental & clinical cancer research : CR.* 2020;39(1).
37. Zhou Z, Gong Q, Wang Y, Li M, Wang L, Ding H, et al. The biological function and clinical significance of SF3B1 mutations in cancer. *Biomarker research.* 2020;8:38.
38. Song J, Hussaini M, Qin D, Zhang X, Shao H, Zhang L, et al. Comparison of SF3B1/DNMT3A Comutations With DNMT3A or SF3B1 Mutation Alone in Myelodysplastic Syndrome and Clonal Cytopenia of Undetermined Significance. *American journal of clinical pathology.* 2020;154(1):48-56.
39. Bergot T, Lippert E, Douet-Guilbert N, Commet S, Corcos L, Bernard DG. Human Cancer-Associated Mutations of SF3B1 Lead to a Splicing Modification of Its Own RNA. *Cancers.* 2020;12(3).
40. Hou C, Xiao L, Ren X, Tang F, Guo B, Zeng W, et al. Mutations of GNAQ, GNA11, SF3B1, EIF1AX, PLCB4 and CYSLTR in Uveal Melanoma in Chinese Patients. *Ophthalmic research.* 2020;63(3):358-68.
41. Küsters-Vandeveldel HV, Creytens D, van Engen-van Grunsven AC, Jeunink M, Winnepenninckx V, Groenen PJ, et al. SF3B1 and EIF1AX mutations occur in primary leptomeningeal melanocytic neoplasms; yet another similarity to uveal melanomas. *Acta neuropathologica communications.* 2016;4:5.
42. Kong Y, Krauthammer M, Halaban R. Rare SF3B1 R625 mutations in cutaneous melanoma. *Melanoma research.* 2014;24(4):332-4.
43. Conn SJ, Pillman KA, Toubia J, Conn VM, Salmanidis M, Phillips CA, et al. The RNA Binding Protein Quaking Regulates Formation of circRNAs *%J Cell.* 2015;160(6).
44. Kaixuan Z, Bangshun H, B YB, Tao X, Xiaoxiang C, Mu X, et al. The pro-metastasis effect of circANKS1B in breast cancer. *%J Molecular cancer.* 2018;17(1).
45. Zhao W, Cui Y, Liu L, Qi X, Liu J, Ma S, et al. Splicing factor derived circular RNA circUHRF1 accelerates oral squamous cell carcinoma tumorigenesis via feedback loop *%J Cell Death & Differentiation: Official journal of the ADMC Associazione Differenziamento e Morte Cellulare.* 2020;27(7).

46. Qin Y, Molly C, M GG, Sylvie D. Crystal structure of a human cleavage factor CFI(m)25/CFI(m)68/RNA complex provides an insight into poly(A) site recognition and RNA looping. %J Structure (London, England : 1993). 2011;19(3).
47. Tan S, Li H, Zhang W, Shao Y, Liu Y, Guan H, et al. NUDT21 negatively regulates PSMB2 and CXXC5 by alternative polyadenylation and contributes to hepatocellular carcinoma suppression %J Oncogene. 2018;37(35).
48. Xiaojing L, Junyao D, Xueying W, Zeneng C, Qubo Z. NUDT21 regulates circRNA cyclization and ceRNA crosstalk in hepatocellular carcinoma. %J Oncogene. 2020;39(4).
49. Wang X, Dong J, Li X, Cheng Z, Zhu Q. CPSF4 regulates circRNA formation and microRNA mediated gene silencing in hepatocellular carcinoma. Oncogene. 2021.
50. Aktaş T, Ilik İA, Maticzka D, Bhardwaj V, Rodrigues CP, Mittler G, et al. DHX9 suppresses RNA processing defects originating from the Alu invasion of the human genome %J Nature: International weekly journal of science. 2017;544(7648).
51. Yingkuan L, Hui W, Bing C, Qixing M, Wenjie X, Te Z, et al. circDCUN1D4 suppresses tumor metastasis and glycolysis in lung adenocarcinoma by stabilizing TXNIP expression %J Molecular Therapy - Nucleic Acids. 2021;23.
52. Yan D, Dong W, He Q, Yang M, Huang L, Kong J, et al. Circular RNA circPICALM sponges miR-1265 to inhibit bladder cancer metastasis and influence FAK phosphorylation %J EBioMedicine. 2019;48.
53. Jian Y, Qing-Guo X, Zhen-Guang W, Yuan Y, Ling Z, Jin-Zhao M, et al. Circular RNA cSMARCA5 inhibits growth and metastasis in hepatocellular carcinoma. %J Journal of hepatology. 2018;68(6).
54. Linder P, Jankowsky E. From unwinding to clamping — the DEAD box RNA helicase family %J Nature Reviews Molecular Cell Biology. 2011;12(8).
55. Noble CG, Song H. MLN51 Stimulates the RNA-Helicase Activity of eIF4AIII %J PLOS ONE. 2007;2(3).
56. Ballut L, Marchadier B, Baguet A, Tomasetto C, Séraphin B, Hir HL. The exon junction core complex is locked onto RNA by inhibition of eIF4AIII ATPase activity %J Nature Structural & Molecular Biology. 2005;12(10).
57. Wang R, Zhang S, Chen X, Li N, Li J, Jia R, et al. EIF4A3-induced circular RNA MMP9 (circMMP9) acts as a sponge of miR-124 and promotes glioblastoma multiforme cell tumorigenesis %J Molecular Cancer. 2018;17(3).
58. Yutian W, Chenfei L, Peng Z, Lin Z, Xiao L, Jianxing Y, et al. EIF4A3-induced circular RNA ASAP1 promotes tumorigenesis and temozolomide resistance of glioblastoma via NRAS/MEK1/ERK1–2 signaling %J Neuro-Oncology. 2021;23(4).
59. Zheng X, Huang M, Xing L, Yang R, Wang X, Jiang R, et al. The circRNA circSEPT9 mediated by E2F1 and EIF4A3 facilitates the carcinogenesis and development of triple-negative breast cancer %J Molecular Cancer. 2020;19(19).
60. Ivanov A, Memczak S, Wyler E, Torti F, Porath HT, Orejuela MR, et al. Analysis of intron sequences reveals hallmarks of circular RNA biogenesis in animals. Cell reports. 2015;10(2):170-7.
61. Shi L, Yan P, Liang Y, Sun Y, Shen J, Zhou S, et al. Circular RNA expression is suppressed by androgen receptor (AR)-regulated adenosine deaminase that acts on RNA (ADAR1) in human hepatocellular carcinoma. Cell death & disease. 2017;8(11):e3171.
62. Ma C, Wang X, Yang F, Zang Y, Liu J, Wang X, et al. Circular RNA hsa_circ_0004872 inhibits gastric cancer progression via the miR-224/Smad4/ADAR1 successive regulatory circuit. Mol Cancer. 2020;19(1):157.
63. Shen P, Yang T, Chen Q, Yuan H, Wu P, Cai B, et al. CircNEIL3 regulatory loop promotes pancreatic ductal adenocarcinoma progression via miRNA sponging and A-to-I RNA-editing. Mol Cancer. 2021;20(1):51.
64. Zhao X, Cai Y, Xu J. Circular RNAs: Biogenesis, Mechanism, and Function in Human Cancers %J International Journal of Molecular Sciences. 2019;20(16).
65. Fei T, Chen Y, Xiao T, Li W, Cato L, Zhang P, et al. Genome-wide CRISPR screen identifies HNRNPL as a prostate cancer dependency regulating RNA splicing %J Proceedings of the National Academy of Sciences. 2017;114(26).
66. Jiacheng C, Xing Z, Penghui X, Haixiao W, Sen W, Lu Z, et al. Circular RNA circLMO7 acts as a microRNA-30a-3p sponge to promote gastric cancer progression via the WNT2/ β -catenin pathway %J Journal of Experimental & Clinical Cancer Research. 2021;40(1).
67. Stark H, Dube P, Lührmann R, Kastner B. Arrangement of RNA and proteins in the spliceosomal U1 small nuclear ribonucleoprotein particle %J Nature: International weekly journal of science. 2001;409(6819).
68. Wu JI, Reed RB, Grabowski PJ, Artzt K. Function of quaking in Myelination: Regulation of Alternative Splicing %J Proceedings of the National Academy of Sciences of the United States of America. 2002;99(7).

69. Teplova M, Hafner M, Teplov D, Essig K, Tuschl T, Patel DJ. Structure–function studies of STAR family Quaking proteins bound to their in vivo RNA target sites %J Cold Spring Harbor Laboratory Press. 2013;27(8).
70. J CS, A PK, John T, M CV, Marika S, A PC, et al. The RNA binding protein quaking regulates formation of circRNAs. %J Cell. 2015;160(6).
71. Yan-Jing Z, Bo Z, Gui-Juan L, Xu-Kai M, Xin-Yuan L, Xi-Meng L, et al. Circular RNAs negatively regulate cancer stem cells by physically binding FMRP against CCAR1 complex in hepatocellular carcinoma. %J Theranostics. 2019;9(12).
72. Han K, Wang F-W, Cao C-H, Ling H, Chen J-W, Chen R-X, et al. CircLONP2 enhances colorectal carcinoma invasion and metastasis through modulating the maturation and exosomal dissemination of microRNA-17 %J Molecular Cancer. 2020;19(6).
73. Shaoxia L, Ningning Y, Xu J, Jian W, Junqiang D, Yonghua G. FUS-induced circular RNA ZNF609 promotes tumorigenesis and progression via sponging miR-142-3p in lung cancer %J Journal of Cellular Physiology. 2020;236(1).
74. Tong C, Xiaolong W, Chen L, Hanwen Z, Ying L, Dianwen H, et al. CircHIF1A regulated by FUS accelerates triple-negative breast cancer progression by modulating NFIB expression and translocation. %J Oncogene. 2021.
75. Zhang XO, Dong R, Zhang Y, Zhang JL, Luo Z, Zhang J, et al. Diverse alternative back-splicing and alternative splicing landscape of circular RNAs. Genome research. 2016;26(9):1277-87.
76. Liang D, Wilusz JE. Short intronic repeat sequences facilitate circular RNA production. Genes & development. 2014;28(20):2233-47.
77. Jin C, Zhao J, Zhang ZP, Wu M, Li J, Liu B, et al. CircRNA EPHB4 modulates stem properties and proliferation of gliomas via sponging miR-637 and up-regulating SOX10. Molecular oncology. 2021;15(2):596-622.
78. Proudfoot NJ. Ending the message: poly(A) signals then and now. Genes & development. 2011;25(17):1770-82.
79. Millevoi S, Vagner S. Molecular mechanisms of eukaryotic pre-mRNA 3' end processing regulation. Nucleic acids research. 2010;38(9):2757-74.
80. Sun M, Ding J, Li D, Yang G, Cheng Z, Zhu Q. NUDT21 regulates 3'-UTR length and microRNA-mediated gene silencing in hepatocellular carcinoma. Cancer letters. 2017;410:158-68.
81. Yang Q, Coseno M, Gilmartin GM, Doublé S. Crystal structure of a human cleavage factor CFI(m)25/CFI(m)68/RNA complex provides an insight into poly(A) site recognition and RNA looping. Structure (London, England : 1993). 2011;19(3):368-77.