Circular RNA translation – new discovery and challenges

Qingqing Miao [1], **Bing Ni** [Corresp., 2], **Jun Tang** [1]

[1] Clinical College of People's Liberation Army Affiliated to Anhui Medical University, Hefei, Anhui, China

[2] The Third Military Medical University, Chongqing, Chongqing, China

Corresponding Author: Bing Ni
Email address: nibing@tmmu.edu.cn

The circular (circ)RNAs are a newly recognized group of noncoding (nc)RNAs. Research to characterize the functional features of circRNAs has uncovered distinctive profiles of conservation, stability, specificity and complexity. However, a new line of evidence has indicated that although circRNAs can function as ncRNAs, such as in the role of miRNA sponges, they are also capable of coding proteins. To date, several circRNAs have been verified to be able to translate proteins or peptides with functions that mainly influence the functions of their maternal genes. These findings greatly broaden our research approach and the knowledge of ncRNAs, meanwhile these findings also raise questions about whether circRNA is still classified as non-coding RNA. Here, we systematically summarize the history and evidence for the translation of circRNAs, including the evolution implications, molecular structures, regulation and mechanism, experimental validation and computational prediction for the coding ability of circRNAs.

Circular RNA translation – new discovery and challenges

Qingqing Miao[1,2], Bing Ni[2,*] , Jun Tang[1,*]

[1] Department of Dermatology, Clinical College of People's Liberation Army Affiliated to Anhui Medical University, Hefei 230001, China

[2] Department of Pathophysiology, College of High Altitude Military Medicine, Third Military Medical University, Chongqing 400038, China

* Corresponding authors:

Bing Ni, E-mail: nibing@tmmu.edu.cn; Tel: +86-23-68772228; Fax: +86-23-68772228

Jun Tang, E-mail: jiangzhuyan@stu.ahmu.edu.cn; Tel: +86-551-65966263; Fax: +86-551-65966263

Abstract

The circular (circ)RNAs are a newly recognized group of noncoding (nc)RNAs. Research to characterize the functional features of circRNAs has uncovered distinctive profiles of conservation, stability, specificity and complexity. However, a new line of evidence has indicated that although circRNAs can function as ncRNAs, such as in the role of miRNA sponges, they are also capable of coding proteins. To date, several circRNAs have been verified to be able to translate proteins or peptides with functions that mainly influence the functions of their maternal genes. These findings greatly broaden our research approach and the knowledge of ncRNAs, meanwhile these findings also raise questions about whether circRNA is still classified as non-coding RNA. Here, we systematically summarize the history and evidence for the translation of circRNAs, including the evolution implications, molecular structures, regulation and mechanism, experimental validation and computational prediction for the coding ability of circRNAs.

Keywords: ncRNAs; circRNAs; translation; mRNA; miRNA

1. Introduction

The classic "central dogma of molecular biology" suggests that the DNA constituent of our chromosomes is transcribed into RNA and subsequently translated into proteins. High-throughput sequencing technology has not only verified the dynamic complexity of gene expression but also revealed the existence of delicate regulatory processes at the RNA level[1]. The RNA form of genetic information serves as the intermediary between DNA and its protein products[2]; as such, it is believed that levels of RNA are at the core of life's complex functions[3]. At the turn of the century, whole-genome sequencing indicated that while approximately 93% of the DNA in the human genome is transcribed into RNA, only approximately 2% of the DNA sequences encode proteins[4]. This finding suggested that there are large amounts of noncoding (nc)RNAs in mammalian cells.

Although the newly discovered ncRNAs were at first largely dismissed as "transcriptional noise", focused investigations began to reveal functional roles in cell biology and many disease types. Researchers' attention has now turned towards defining the roles of ncRNAs in regulating and modulating host gene expression[5, 6]. The current collective data have allowed the two major groups of ncRNAs—the long (l)ncRNAs and small RNAs, grossly stratified according to size—to be further categorized according to function; these functional subcategories include ribosomal (r)RNA, transfer (t)RNA, small nuclear (sn)RNA, small nucleolar (sno)RNA, PIWI-interacting (pi)RNA, micro (mi)RNA, lncRNA, circular (circ)RNA and transcription initiation (ti)RNA[7, 8]. Among these, the miRNAs and lncRNAs have been extensively studied and confirmed to function in gene transcription through pivotal activities in a versatile regulation network[9, 10].

The circRNAs have shown particular stability and functional versatility *in vivo*, e.g., acting as miRNA sponges in both physiological and pathological processes[11]. Furthermore, circRNAs have been demonstrated as capable of translating directly into protein, indicating an intriguing potential to directly function in many processes of life. In this review, we will discuss the most recent progress of the research into the translational capacity of circRNAs and towards defining the underlying mechanisms.

2. Survey methodology

This paper was based on review and research articles in reputable peer-reviewed journals and government websites. The research was conducted using PubMed, Google Scholar and reports. The words "noncoding RNAs", "circRNAs", "miRNAs", "lncRNAs", "transcription", "translation", "protein", "coding" and a combination of those were used to retrieve literature from the databases.

3. CircRNA biology

CircRNAs are single-stranded covalently closed circular (CCC) RNA molecules generated from a broad array of genomic regions, ranging from intergenic, intronic and coding sequences to 5′- or 3′-untranslational sequences[12, 13]. Two models of circRNA biosynthesis have been proposed, both involving back-splicing catalyzed by the spliceosomal machinery. The first of the two, the "exon skipping" model, begins with

classical splicing to generate linear RNA. The downstream exon links to the upstream exon, with one or more exons being skipped; the skipped exons then further back-splice to form precursor circRNAs, which undergo further processing to become mature circRNAs. The second of the two models, the "direct back-splicing" circularization model, is related mostly to complementary motifs; in this, the complementary pairing RNA back-splices to produce a precursor circRNA together with an exon-intron(s)-exon intermediate, and the latter is further processed to produce a linear RNA with skipped exons or which is targeted for degradation[14-16] (**Figure 1**).

To date, three functions have been defined for the circRNAs. First, circRNAs harbor miRNA complementary sequences, facilitating their combination with and ability to adjust the biological function of a large number of miRNAs by functioning as molecular sponges. A specific example of this is the circMTO1, which acts as the sponge of miR-9 to suppress hepatocellular carcinoma progression[17]. Furthermore, one circRNA may combine with several kinds of miRNAs; for instance, circHIPK3 has been reported to combine with 9 miRNAs (miR-29a, miR-29b, miR-124, miR-152, miR-193a, miR-338, miR-379, miR-584 and miR-654) to synergistically inhibit cell proliferation. Second, circRNAs can directly regulate transcription, splicing and expression of a parental gene. The exon-intron circRNAs (EIciRNAs) are examples of this regulation, interacting with RNA polymerase II and enhancing transcription of their parental genes[18]. Third, circRNAs directly interact with proteins, such as the ternary complex circ-Foxo3-p21-CDK2, which serves to arrest the function of CDK2 and interrupt cell cycle progression[19]. However, studies indicate that one circRNA might simultaneously harbor more than one of the above functions, which is evidenced by the finding that circ-Amotl1 can act both as a sponge for miR-17 to promote cell proliferation, migration and wound healing and as a target for protein binding (c-Myc, Akt1 and PDK1) to promote the proliferation of tumor cells and enhancement of cardiac repair[20-22]. The detailed information on the biological consequences of circRNA has been reviewed elsewhere[23-25].

More interestingly, the latest research findings are providing hints towards a potential fourth function of circRNAs—translation (**Figure 2**), which opens a new field for researchers to explore the biological functions of circRNA-derived proteins.

4. CircRNA translation potential: a controversial issue explored unceasingly

It is commonly believed that mRNAs represent the primary controller of cells, carrying out the necessary functions for life. Since the endogenous circRNAs appear to not be associated with polysomes, they presumably lack the potential for translation[26, 27]. Although this notion has not been definitively disproven, it still attracts scientists' interests in exploring the unknown, hoping to advance the field of research into circRNA translational potency forward, from theory to practical knowledge.

4.1 Theoretical basis for direct translation of endogenous circRNAs

4.1.1 Evolutionary perspective and abundance of circRNAs in mammalian cells

Dong et al[28] reported on the use of complementary sequence index tagging to analyze short interspersed nuclear repetitive DNA elements (referred to as 'SINEs', which contribute to circRNA formation) and, particularly, to explore the complexity of circRNA expression patterns during species evolution. The finding that circRNAs in lower organisms can translate into proteins supports the possibility that those in more advanced organisms may retain a similar function.

Actually, high-throughput sequencing technology has revealed that circRNAs are abundant in human cells, tissues and body fluids, even in exosomes[13]. Detected throughout the cell, the highest amounts are found in cytoplasm. Although the nucleus harbors lower amounts of circRNAs, their degradation rate in this subcellular compartment is lower than that of mRNA[29]. The half-lives of some cytoplasmic circRNAs are more than 48 h[27], and if numerous circRNAs accumulate, their levels could surpass even those of their corresponding linear mRNAs[30]. Therefore, it can be speculated that circRNAs containing translational elements and open reading frames (ORFs) may combine with ribosomes, thereby initiating translation.

4.1.2 Analogous to similar ncRNAs

Recent studies demonstrate that many lncRNAs are able to translate into functional polypeptides. In 2013, Magny et al[31] found the putative ncRNA 003 in 2L (pncr003:2L), including two potentially functional small ORFs in the fly's heart, which could translate into bioactive peptides and synergistically regulate cardiac calcium uptake. In 2015, Anderson et al[32] discovered an annotated lncRNA that translates a conserved micropeptide, myoregulin (MLN), that functions as a regulator of skeletal muscle physiology. One year afterward, Nelson et al[33] found that a peptide named 'dwarf open reading frame' (DWORF) is encoded by a putative lncRNA. This peptide acts in a mutually exclusive manner, apart from the other three inhibitors of phospholamban, sarcolipin, and MLN, to competitively combine with the sarco/endoplasmic reticulum Ca^{2+} ATPase pump to adjust the reuptake of the Ca^{2+} in muscle. Last year, Matsumoto et al[34] identified another functional novel polypeptide encoded by a lncRNA. This peptide can negatively regulate mTORC1 activation by interacting with the lysosomal v-ATPase in the late endosome/lysosome. In the subsequent research, increasingly more lncRNAs with the capacity for translating proteins (peptides) will be explored. The discovery of this special type of lncRNA gives reason to speculate that the full spectrum of biological significance of the coding ability of circRNAs remains to be uncovered.

4.1.3 Molecular structure

Internal ribosome entry site (IRES).

It is well known that there are two translation modes, cap-dependent translation and cap-independent translation. The traditional cap-dependent translation accounts for a basal level of protein synthesis under normal growth conditions. In contrast, cap-independent translation contributes to cell proliferation or cellular adaptation/survival when traditional protein synthesis is severely inhibited; this second mode is mediated by the IRES. IRES-mediated translation, therefore, serves as an urgent breakdown maintenance mechanism

during cell stress, ensuring basic protein needs are met[35, 36]; as such, this mechanism is often triggered in conditions of viral invasion, tumor or other human diseases[37-39]. Thus, it is not surprising that the IRES itself was originally identified by researchers studying the virus parasitic mechanism[40]. Since then, comparative sequencing analysis has led to the identification of IRES components throughout the human genome[41]. Functional studies have characterized the IRES in mRNA as dependent upon the molecule's special structure, allowing the 40S subunit to avoid assembling directly at the 5'-untranslated sequence[38].

In 2016, Li et al[42] established a circRNA database, circRNADb (http://reprod.njmu.edu.cn/ circrnadb), the first of its kind, summarizing circRNA-encoded protein information based upon 32,914 human exonic circRNAs. Interestingly, their initial explorations of this dataset found ORFs in about half of the circRNAs, and IRESs in about half of those; as such, those 7,170 circRNA sequences were considered to fit the characteristic requirements for protein translation capabilities. To date, four types of virus IRES structures are classified as having the functional ability to hijack eukaryotic translation machinery, and all work with a common mechanical principle, leading to 80S ribosomal assembly and extension[43]. However, in eukaryotic mRNAs and circRNAs, the IRES-mediated ribosome assembly mechanism is less well known; only isolated examples of IRESs with known IRES-transacting factor binding sites or resolved three-dimensional structure are available[35,36].

RNA modification.

Statistical analyses have estimated that RNA molecules may contain more than 100 distinct modifications[44]. Approximately 16 species of modifications in mRNA have been recognized to date, and the vast majority of these involve the m^6A, ψ and m^5C chemical modifications[45]. The m^6A modification is related to mRNA stability, splicing processing, polypeptide translation and miRNA processing and is correlated with stem cell fate and biological rhythms[46-48]. The ψ (pseudouridylation) modification serves three main functions, namely, changing the codon, enhancing the transcript stability and regulating the stress response. To date, only the m^6A modification has been verified in circRNAs, wherein it plays a role in promoting translation[49]. However, research on the m^5C modification on ncRNAs has been very limited, though the ncRNA and mRNA have been found to hold thousands of m^5C modification sites in recent years[46-48]. Therefore, it is speculated that more modification types will be found in both circRNAs and mRNAs with continued research. Such modifications will likely function not only in terms of translation but also in adjusting the functions of circRNAs as ncRNAs.

4.2 Experimental exploration for endogenous circRNA translation in eukaryotic cells

4.2.1 Early exploration findings

The first indications of a translational role for circRNAs emerged from studies of virus nucleic acids. One of the first observations of a circRNA behaving as a translational template was made with the single-stranded circular RNA genome of the hepatitis δ virus, a satellite virus of the hepatitis B virus; encapsulation of the former by hepatitis B virions was found to result in the production of a single viral protein of 122 amino acids,

in a noncanonical manner[50]. In 1995, Chen et al[51] demonstrated that synthetic circRNAs containing IRES elements were able to correctly translate into polypeptides in rabbit reticulocyte lysate, but those without IRES could not. Furthermore, the authors speculated that this type of RNA can translate along the RNA circles for multiple consecutive rounds. In 1998, Perriman et al[52] used plasmids for creating RNA cyclase ribozymes to produce desired circRNAs that were inserted into the green fluorescent protein (GFP) ORF (finite GFP encoding) and stop codon-devoid GFP reading frame (infinite GFP encoding). The authors showed that both circRNAs can directly translate along with GFP in *Escherichia coli* strains and, in the meantime, the infinite GFP-encoding RNA could be translated into an extremely long repeating poly-GFP. These findings validated Chen's previous prediction in 1995[51]. In 1999, Li et al[53] reported that a circRNA containing exon 2 of the Na/Ca exchanger gene NCX1 might translate for a protein. It is a pity that they could not detect a protein corresponding exactly to what they predicted from the circular transcript; however, when the circRNAs were made into linear RNAs and transfected into HEK-293 cells, the linear versions of circRNAs were shown to result in proteins of the expected size of ~70 kDa, and the transfected cells possessed Na/Ca exchange activity.

Over a decade later, Wang et al[54] reported on their construction of an efficient back-splicing circRNA, which could be translated into functional GFP proteins in human and Drosophila cell lines. Furthermore, due to the nuclease resistance characteristics of circRNAs, when the cell was transfected with circRNA, protein production was prolonged for several days. In the same year, Abe et al[55] provided evidence that circRNAs were translated into infinite FLAG proteins in rabbit reticulocyte lysate and in HeLa cells with an infinite ORF in the absence of any particular translation initiation element such as a poly-A tail, IRES, or a cap structure. This series of experiments proves that artificial circRNAs with stop codon mutations have a rolling circle amplification mechanism to code for long repeating poly proteins. In 2014, Haidar et al[56] reported a small new virusoid with CCC RNA (220 nt) associated with rice yellow mottle virus that could translate into a 16-kDa highly basic protein. This example is the only one that codes proteins among all known viroids and virusoids. This unique natural supercompact “nano genome” even overlaps its initiation and termination codons to UGAUGA[56].

Nevertheless, all these scattered reports, however, are limited to viruses, bacteria, or synthetic circRNA[57] (**Table 1**), and the translation ability of endogenous circRNAs still requires further exploration.

4.2.2 Solid evidence for endogenous circRNA direct translation

In 2013, Jeck et al[27] reported that circRNAs are abundant, conserved and associated with ALU repeats, but there are no detectable levels of exonic circRNAs in the ribosome-bound fraction (via ribosome profiling). One year later, Dawood et al[58] raised doubts about this conclusion when they reported their findings from a bioinformatic analysis; IRES regions in circRNAs represented predicted binding sites for RNA binding proteins, including some known to modulate IRES-driven translation.

In 2017, it was finally proved that endogenous circRNAs are capable of directly translating into proteins. By using ribosome footprinting and immunoprecipitation of Drosophila brain tissues, Pamudurti et al[59]

210 demonstrated that circRNA sequences could be bound by ribosomes, including the termination codon. The
211 study focused on circ-Mbl from the Mbl gene among all of the ribo-circRNAs and repeatedly verified that circ-
212 Mbl could translate into protein. Through the construction of an overexpression vector, the substitution of the
213 ORF with a split Cherry molecule was made and target mass spectrometry confirmed the Drosophila brain
214 circ-Mbl was immunoprecipitated. In the same year, through a screening study of circRNAs related to human,
215 mouse (C2,C12) and a Duchenne muscular dystrophy disease model, Legnini et al[57] showed that circ-
216 ZNF609 was combined with ribosomes and that its encoded protein may be involved in the myoblast
217 growth process; however, the circ-ZNF609 was found to be translated at almost two orders of magnitude lower
218 efficiency than that of the linear form.
219 Thereafter, Yun et al[49] explored circRNA translation ability by the same approach and found that
220 control sequences without IRES were also capable of translating the target protein. These unexpected circRNA
221 translation events were initiated by eIF4G2 and eIF3A and associated with the m^6A modification. When the
222 m^6A modifications were “erased”, the target protein translation activity was substantially affected, to the point
223 that it completely disappeared. Ribosome spectrum analysis confirmed that a multitude of endogenous
224 circRNAs were bound by ribosomes, but whether these circRNAs harbored any IRESs was not examined.
225 Finally, high-throughput sequencing analysis determined that approximately 13% of the total circRNAs carried
226 the m^6A modification. Months later, another independent study showed that circRNAs carry extensive m^6A
227 modifications and are expressed in cell type-specific patterns[60]. The writing and reading machinery of these
228 m^6A modifications were found to be similar to those of mRNAs (i.e. involving the METTL3/14 and YTH
229 proteins) but were distinctive in their location patterns; the data also suggested that the m^6A modification did
230 not appear to promote degradation of circRNAs as it does for mRNAs. Ultimately, interpretation of these
231 findings indicates that switching the state of m^6A modifications may allow for functional control of circRNAs.
232 Most recently, Yang et al[61] reported that the circ-FBXW7 can translate for a new protein FBXW7-
233 185aa during glioma tumorigenesis. Intriguingly, this protein cooperates with FBXW7, which is encoded in
234 their parental genes, to control c-Myc stability and repress cell cycle acceleration and the consequent
235 proliferation. This is the first study to provide definitive evidence of protein translation via circRNA synergy
236 with the protein expression by parental genes and joint function of the proteins. Zhang et al[57] further reported
237 that circ-SHPRH, a circRNA containing an IRES-driven ORF, translates into a functional protein. For this
238 process, circ-SHPRH utilizes overlapping genetic codes to create a UGA stop codon, causing translation of the
239 SHPRH-146aa protein. The translated SHPRH-146aa functions as a protector of the full-length SHPRH
240 protein, guarding against degradation by the ubiquitin proteasome and consequently inhibiting cell
241 proliferation and tumorigenicity in human glioblastoma. Detailed information for the published coding
242 circRNAs is summarized in **Table 1**.
243

244 **5. Methodology for the prediction of circRNA translation capacity**

245 The phenomenon of circRNA translation continues to attract increasing attention from scientists. This

year has seen the publication of methodological articles, indicating the overall momentum for technical advancement of the field. For example, Yun et al[65] described a minigene reporter system to measure IRES-mediated translation in circRNAs, and Bartsch et al[66] detailed a sucrose gradient-based method to pinpoint association of a given circRNA with distinct ribosomal fractions. The latter method allows the evaluation of the coding potential of candidate circRNAs and its association with the translation machinery. In addition, ribosome imprinting (FRP or Ribo-seq) and a series of improved methods previously used to understand the intracellular translation machinery can also be used in the investigation of circRNA translation. This method first extracts RNA, digests it by RNA enzymes and then deep sequences it. After those steps, only fragments that are protected by binding to ribosomes are preserved, which directly shows the translation of RNA[67, 68].

Though these experimental methods are very helpful for the investigation of the translational capacity of circRNAs, bioinformatics analysis might be the easiest and fastest way to answer whether circRNAs have encoding ability before the beginning of a study. Researchers can obtain useful information from the websites simply by importing the circRNA sequence. As mentioned above, circRNADb (http://reprod.njmu.edu.cn/circrnadb) is the first website to record the coding ability of circRNAs in detail, including but not limited to the IRES and ORF starting and ending sites[42]. Meng et al[69] has also provided an integrated tool (http://bis.zju.edu.cn/CircPro) capable of detecting circRNAs with translation potential from high-throughput sequencing data.

In addition, circRNAs are actually a kind of lncRNA, so software that predicts sequence coding capabilities can be borrowed and is not limited to circRNA-specific websites. For instance, ORFfinder (https://www.ncbi.nlm.nih.gov/orffinder/) is a tool that can calculate all possible ORFs and allows you to choose different start codons to soften the search terms, while IRESite (http://www.iresite.org/) compiles those IRES sequences from viruses to higher-order cells that have been found, as well as the secondary structure of the predicted virus IRES. Other tools include PhyloCSF (https://github.com/mlin/PhyloCSF/wiki) which uses a comparative genomics method to search for protein coding regions[70], and CPC (http://cpc.cbi.pku.edu.cn) which assesses the protein-coding potential of transcripts by using sequence features and a support vector machine[71]. In addition, the Coding Potential Assessment Tool (CPAT; http://lilab.research.bcm.edu/cpat/index.php) can be used, which applies a novel alignment-free method to rapidly distinguish coding and noncoding regions from a large pool of candidates[72].

It is worth noting that most of the circRNAs translation ability predictions obtained from websites or software are based on exons and ORFs that are included in the circRNA of interest; therefore, potential translation possibilities may be missed for circRNAs containing introns. This issue will be an intriguing and insightful addition to the circRNA research field.

6. Challenges and perspectives

The field of RNA research has continually emphasized the structural and functional versatility of RNA molecules. This versatility has in turn inspired translational and clinical researchers to explore the utility of

RNA-based therapeutic agents for a wide variety of medical applications. Several RNA therapeutics with diverse modes of action are currently being evaluated in large late-stage clinical trials, and many more are in the early clinical development stage, including strategies to modulate target gene expression, such as mRNA, siRNA and miRNA[73]. For instance, mRNA-modified dendritic cells have shown promising and efficient results in clinical trials[74], and siRNA-based therapeutic agents such as pegaptanib (Macugen) and bevacizumab (Avastin; off-label use) have shown success for the treatment of wet, age-related macular degeneration in clinical testing[75].

The circRNAs may regulate gene expression through different mechanisms, including direct translation[76]. Therefore, considering their stability and specific expression features, the circRNAs with translation potential could represent strong candidates for development as clinical tools to therapeutically manipulate a wide variety of physiologic and pathologic processes. Interestingly, Wesselhoeft et al[62] engineered exogenous circRNA for robust and stable protein expression in eukaryotic cells, showing it to be a promising alternative to linear mRNA according to its exceptional protein production qualities, in terms of both quantity of protein produced and stability of production. However, the real-life application of circRNA as a clinical tool to treat disease remains a remarkable challenge, requiring extensive and in-depth preclinical research.

In terms of the discovery and exploration of endogenous circRNA-translated proteins, we speculate that there will be a large number of circRNAs with translational function that will be discovered as the research field matures. The significance of the function of those proteins or peptides translated by circRNAs will also need to be explored thoroughly. The current studies on the functions of circRNA-translated proteins are focused mainly on their effects on maternal genes; for instance, it has been found that FBXW7-185aa, the translational product of human circ-FBXW7, cooperates with the full-length FBXW7 to reduce the half-life of c-Myc by antagonizing USP28-induced c-Myc stabilization[61], and SHPRH-146aa, the translational product of human circ-SHPRH, protects full-length SHPRH from degradation by the ubiquitin proteasome, leading to inhibited cell proliferation and tumorigenicity[64]. However, whether the circRNA-encoded proteins have other functions besides influencing their maternal genes, as well as the molecular processes of those functions, represent equally important focuses of the long-term research.

The collective evidence to date implies that the translation of endogenous circular RNA into proteins or peptides may be a widespread phenomenon, though the coding potential of circRNAs previously had been largely disregarded. Therefore, further studies on the translational capacity of circRNAs should be encouraged and, for the immediate future, should focus on the functions and detailed mechanisms of the circRNA modifications themselves, such as the 5’ cap-independent translation of circRNAs and circRNA-derived protein or peptides. The resulting insights will also be helpful towards furthering our understanding of ncRNA functions in general.

Acknowledgements:

This work was supported by grants from the National Key Research and Development Project (No. 2016YFA0502203), the National Foundation of China (Nos. 81502728 and 81670534), and the Anhui Provincial Natural Science Foundation (No. 1408085MH149). The funders had no role in the study design, data analysis, or decision to publish.

Conflicts of interests:

The authors declare no conflicts of interest related to this publication.

References

1 Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ: Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat Genet 2008;40:1413-1415.

2 Crick F: Central dogma of molecular biology. Nature 1970;227:561-563.

3 Licatalosi DD, Darnell RB: Rna processing and its regulation: Global insights into biological networks. Nat Rev Genet 2010;11:75-87.

4 Consortium EP: An integrated encyclopedia of DNA elements in the human genome. Nature 2012;489:57-74.

5 Peschansky VJ, Wahlestedt C: Non-coding rnas as direct and indirect modulators of epigenetic regulation. Epigenetics 2014;9:3-12.

6 Meller VH, Joshi SS, Deshpande N: Modulation of chromatin by noncoding rna. Annu Rev Genet 2015;49:673-695.

7 Wright MW, Bruford EA: Naming 'junk': Human non-protein coding rna (ncrna) gene nomenclature. Human genomics 2011;5:90-98.

8 Cech TR, Steitz JA: The noncoding rna revolution-trashing old rules to forge new ones. Cell 2014;157:77-94.

9 Mondal T, Kanduri C: Maintenance of epigenetic information: A noncoding rna perspective. Chromosome Res 2013;21:615-625.

10 Guil S, Esteller M: Rna-rna interactions in gene regulation: The coding and noncoding players. Trends Biochem Sci 2015;40:248-256.

11 Han C, Seebacher NA, Hornicek FJ, Kan Q, Duan Z: Regulation of micrornas function by circular rnas in human cancer. Oncotarget 2017;8:64622-64637.

12 LL C: Regulation of circrnabiogenesis. RNABioI 2015

13 Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, Maier L, Mackowiak SD, Gregersen LH, Munschauer M, Loewer A, Ziebold U, Landthaler M, Kocks C, le Noble F, Rajewsky N: Circular rnas are a large class of animal rnas with regulatory potency. Nature 2013;495:333-338.

14 Jeck WR, Sharpless NE: Detecting and characterizing circular rnas. Nature biotechnology 2014;32:453-461.

15 Lasda E, Parker R: Circular rnas: Diversity of form and function. Rna 2014;20:1829-1842.

16 Ashwal-Fluss R, Meyer M, Pamudurti NR, Ivanov A, Bartok O, Hanan M, Evantal N, Memczak S, Rajewsky N, Kadener S: Circrna biogenesis competes with pre-mrna splicing. Molecular cell 2014;56:55-66.

17 Han D, Li J, Wang H, Su X, Hou J, Gu Y, Qian C, Lin Y, Liu X, Huang M, Li N, Zhou W, Yu Y, Cao X: Circular rna circmto1 acts as the sponge of microrna-9 to suppress hepatocellular carcinoma progression. Hepatology 2017;66:1151-1164.

18 Li Z, Huang C, Bao C, Chen L, Lin M, Wang X, Zhong G, Yu B, Hu W, Dai L, Zhu P, Chang Z, Wu Q, Zhao Y, Jia Y, Xu P, Liu H, Shan G: Exon-intron circular rnas regulate transcription in the nucleus. Nat Struct Mol Biol 2015;22:256-264.

19 Du WW, Yang W, Liu E, Yang Z, Dhaliwal P, Yang BB: Foxo3 circular rna retards cell cycle progression via forming ternary complexes with p21 and cdk2. Nucleic Acids Res 2016;44:2846-2858.
20 Yang ZG, Awan FM, Du WW, Zeng Y, Lyu J, Wu, Gupta S, Yang W, Yang BB: The circular rna interacts with stat3, increasing its nuclear translocation and wound repair by modulating dnmt3a and mir-17 function. Mol Ther 2017;25:2062-2074.
21 Yang Q, Du WW, Wu N, Yang W, Awan FM, Fang L, Ma J, Li X, Zeng Y, Yang Z, Dong J, Khorshidi A, Yang BB: A circular rna promotes tumorigenesis by inducing c-myc nuclear translocation. Cell Death Differ 2017;24:1609-1620.
22 Zeng Y, Du WW, Wu Y, Yang Z, Awan FM, Li X, Yang W, Zhang C, Yang Q, Yee A, Chen Y, Yang F, Sun H, Huang R, Yee AJ, Li RK, Wu Z, Backx PH, Yang BB: A circular rna binds to and activates akt phosphorylation and nuclear localization reducing apoptosis and enhancing cardiac repair. Theranostics 2017;7:3842-3855.
23 Li X, Yang L, Chen LL: The biogenesis, functions, and challenges of circular rnas. Molecular cell 2018;71:428-442.
24 Barrett SP, Salzman J: Circular rnas: Analysis, expression and potential functions. Development 2016;143:1838-1847.
25 Chen I, Chen CY, Chuang TJ: Biogenesis, identification, and function of exonic circular rnas. Wiley interdisciplinary reviews RNA 2015;6:563-579.
26 Guo JU, Agarwal V, Guo H, Bartel DP: Expanded identification and characterization of mammalian circular rnas. Genome biology 2014;15:409.
27 Jeck WR, Sorrentino JA, Wang K, Slevin MK, Burd CE, Liu J, Marzluff WF, Sharpless NE: Circular rnas are abundant, conserved, and associated with alu repeats. RNA 2013;19:141-157.
28 Dong R, Ma XK, Chen LL, Yang L: Increased complexity of circrna expression during species evolution. RNA Biol 2017;14:1064-1074.
29 Panda AC, Grammatikakis I, Munk R, Gorospe M, Abdelmohsen K: Emerging roles and context of circular rnas. Wiley interdisciplinary reviews RNA 2017;8
30 Wilusz JE: Circular rnas: Unexpected outputs of many protein-coding genes. RNA Biol 2017;14:1007-1017.
31 Magny EG, Pueyo JI, Pearl FM, Cespedes MA, Niven JE, Bishop SA, Couso JP: Conserved regulation of cardiac calcium uptake by peptides encoded in small open reading frames. Science 2013;341:1116-1120.
32 Anderson DM, Anderson KM, Chang CL, Makarewich CA, Nelson BR, McAnally JR, Kasaragod P, Shelton JM, Liou J, Bassel-Duby R, Olson EN: A micropeptide encoded by a putative long noncoding rna regulates muscle performance. Cell 2015;160:595-606.
33 Nelson BR, Makarewich CA, Anderson DM, Winders BR, Troupes CD, Wu F, Reese AL, McAnally JR, Chen X, Kavalali ET, Cannon SC, Houser SR, Bassel-Duby R, Olson EN: A peptide encoded by a transcript annotated as long noncoding rna enhances serca activity in muscle. Science 2016;351:271-275.
34 Matsumoto A, Pasut A, Matsumoto M, Yamashita R, Fung J, Monteleone E, Saghatelian A, Nakayama KI, Clohessy JG, Pandolfi PP: Mtorc1 and muscle regeneration are regulated by the linc00961-encoded spar polypeptide. Nature 2017;541:228-232.
35 Lang KJ, Kappel A, Goodall GJ: Hypoxia-inducible factor-1alpha mrna contains an internal ribosome entry site that allows efficient translation during normoxia and hypoxia. Mol Biol Cell 2002;13:1792-1801.
36 Riley A, Jordan LE, Holcik M: Distinct 5' utrs regulate xiap expression under normal growth conditions and during cellular stress. Nucleic Acids Res 2010;38:4665-4674.
37 Holcik M, Sonenberg N: Translational control in stress and apoptosis. Nat Rev Mol Cell Biol 2005;6:318-327.
38 Sonenberg N, Hinnebusch AG: Regulation of translation initiation in eukaryotes: Mechanisms and biological targets. Cell 2009;136:731-745.
39 Faye MD, Holcik M: The role of ires trans-acting factors in carcinogenesis. Biochim Biophys Acta 2015;1849:887-897.

399 40 Baird SD, Turcotte M, Korneluk RG, Holcik M: Searching for ires. RNA 2006;12:1755-1785.
400 41 Weingarten-Gabbay S, Elias-Kirma S, Nir R, Gritsenko AA, Stern-Ginossar N, Yakhini Z, Weinberger A, Segal E:
401 Comparative genetics. Systematic discovery of cap-independent translation sequences in human and viral genomes. Science
402 2016;351
403 42 Chen X, Han P, Zhou T, Guo X, Song X, Li Y: Circrnadb: A comprehensive database for human circular rnas with protein-
404 coding annotations. Sci Rep 2016;6:34985.
405 43 Yamamoto H, Unbehaun A, Spahn CMT: Ribosomal chamber music: Toward an understanding of ires mechanisms. Trends
406 Biochem Sci 2017;42:655-668.
407 44 Gilbert WV, Bell TA, Schaening C: Messenger rna modifications: Form, distribution, and function. Science
408 2016;352:1408-1412.
409 45 Cantara WA, Crain PF, Rozenski J, McCloskey JA, Harris KA, Zhang X, Vendeix FA, Fabris D, Agris PF: The rna
410 modification database, rnamdb: 2011 update. Nucleic Acids Res 2011;39:D195-201.
411 46 Squires JE, Patel HR, Nousch M, Sibbritt T, Humphreys DT, Parker BJ, Suter CM, Preiss T: Widespread occurrence of 5-
412 methylcytosine in human coding and non-coding rna. Nucleic Acids Res 2012;40:5023-5033.
413 47 Hoernes TP, Huttenhofer A, Erlacher MD: Mrna modifications: Dynamic regulators of gene expression? RNA Biol
414 2016;13:760-765.
415 48 Roundtree IA, Evans ME, Pan T, He C: Dynamic rna modifications in gene expression regulation. Cell 2017;169:1187-
416 1200.
417 49 Yang Y, Fan X, Mao M, Song X, Wu P, Zhang Y, Jin Y, Yang Y, Chen LL, Wang Y, Wong CC, Xiao X, Wang Z:
418 Extensive translation of circular rnas driven by n6-methyladenosine. Cell Res 2017;27:626-641.
419 50 Kos A, Dijkema R, Arnberg AC, van der Meide PH, Schellekens H: The hepatitis delta (delta) virus possesses a circular rna.
420 Nature 1986;323:558-560.
421 51 Chen CY, Sarnow P: Initiation of protein synthesis by the eukaryotic translational apparatus on circular rnas. Science
422 1995;268:415-417.
423 52 Perriman R, Ares M, Jr.: Circular mrna can direct translation of extremely long repeating-sequence proteins in vivo. RNA
424 1998;4:1047-1054.
425 53 Li XF, Lytton J: A circularized sodium-calcium exchanger exon 2 transcript. J Biol Chem 1999;274:8153-8160.
426 54 Wang Y, Wang Z: Efficient backsplicing produces translatable circular mrnas. RNA 2015;21:172-179.
427 55 Abe N, Matsumoto K, Nishihara M, Nakano Y, Shibata A, Maruyama H, Shuto S, Matsuda A, Yoshida M, Ito Y, Abe H:
428 Rolling circle translation of circular rna in living human cells. Sci Rep 2015;5:16435.
429 56 AbouHaidar MG, Venkataraman S, Golshani A, Liu B, Ahmad T: Novel coding, translation, and gene expression of a
430 replicating covalently closed circular rna of 220 nt. Proc Natl Acad Sci U S A 2014;111:14542-14547.
431 57 Granados-Riveron JT, Aquino-Jarquin G: The complexity of the translation ability of circrnas. Biochim Biophys Acta
432 2016;1859:1245-1251.
433 58 Dudekula DB, Panda AC, Grammatikakis I, De S, Abdelmohsen K, Gorospe M: Circinteractome: A web tool for exploring
434 circular rnas and their interacting proteins and microrna s. RNA Biol 2016;13:34-42.
435 59 Pamudurti NR, Bartok O, Jens M, Ashwal-Fluss R, Stottmeister C, Ruhe L, Hanan M, Wyler E, Perez-Hernandez D,
436 Ramberger E, Shenzis S, Samson M, Dittmar G, Landthaler M, Chekulaeva M, Rajewsky N, Kadener S: Translation of circrnas.
437 Molecular cell 2017;66:9-21 e27.
438 60 Zhou C, Molinie B, Daneshvar K, Pondick JV, Wang J, Van Wittenberghe N, Xing Y, Giallourakis CC, Mullen AC:
439 Genome-wide maps of m6a circrnas identify widespread and cell-type-specific methylation patterns that are distinct from mrnas.

Cell Rep 2017;20:2262-2276.
61 Yang Y, Gao X, Zhang M, Yan S, Sun C, Xiao F, Huang N, Yang X, Zhao K, Zhou H, Huang S, Xie B, Zhang N: Novel role of fbxw7 circular rna in repressing glioma tumorigenesis. J Natl Cancer Inst 2018;110
62 Wesselhoeft RA, Kowalski PS, Anderson DG: Engineering circular rna for potent and stable translation in eukaryotic cells. Nature communications 2018;9:2629.
63 Legnini I, Di Timoteo G, Rossi F, Morlando M, Briganti F, Sthandier O, Fatica A, Santini T, Andronache A, Wade M, Laneve P, Rajewsky N, Bozzoni I: Circ-znf609 is a circular rna that can be translated and functions in myogenesis. Molecular cell 2017;66:22-37 e29.
64 Zhang M, Huang N, Yang X, Luo J, Yan S, Xiao F, Chen W, Gao X, Zhao K, Zhou H, Li Z, Ming L, Xie B, Zhang N: A novel protein encoded by the circular form of the shprh gene suppresses glioma tumorigenesis. Oncogene 2018
65 Yang Y, Wang Z: Constructing gfp-based reporter to study back splicing and translation of circular rna. Methods Mol Biol 2018;1724:107-118.
66 Bartsch D, Zirkel A, Kurian L: Characterization of circular rnas (circrna) associated with the translation machinery. Methods Mol Biol 2018;1724:159-166.
67 Ingolia NT, Ghaemmaghami S, Newman JR, Weissman JS: Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science 2009;324:218-223.
68 Ingolia NT, Lareau LF, Weissman JS: Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. Cell 2011;147:789-802.
69 Meng X, Chen Q, Zhang P, Chen M: Circpro: An integrated tool for the identification of circrnas with protein-coding potential. Bioinformatics 2017;33:3314-3316.
70 Lin MF, Jungreis I, Kellis M: Phylocsf: A comparative genomics method to distinguish protein coding and non-coding regions. Bioinformatics 2011;27:i275-282.
71 Kong L, Zhang Y, Ye ZQ, Liu XQ, Zhao SQ, Wei L, Gao G: Cpc: Assess the protein-coding potential of transcripts using sequence features and support vector machine. Nucleic Acids Res 2007;35:W345-349.
72 Wang L, Park HJ, Dasari S, Wang S, Kocher JP, Li W: Cpat: Coding-potential assessment tool using an alignment-free logistic regression model. Nucleic Acids Res 2013;41:e74.
73 Sullenger BA, Nair S: From the rna world to the clinic. Science 2016;352:1417-1420.
74 Benteyn D, Heirman C, Bonehill A, Thielemans K, Breckpot K: Mrna-based dendritic cell vaccines. Expert Rev Vaccines 2015;14:161-176.
75 Garba AO, Mousa SA: Bevasiranib for the treatment of wet, age-related macular degeneration. Ophthalmol Eye Dis 2010;2:75-83.
76 Lyu D, Huang S: The emerging role and clinical implication of human exonic circular rna. RNA Biol 2017;14:1000-1006.

Figure legends

Figure 1. Proposed circRNA formation models.

(A) The “exon skipping” model. (B) The “direct back-splicing” model. Black thin lines represent intron sequence; colored thick lines represent different exon sequences.

Figure 2. Functions of circRNAs.

(1) Molecular sponge for miRNA; (2) Regulation of transcription, splicing and expression of parental gene by binding to Pol II; (3) Interaction with proteins; (4) Direct translation of circRNAs.

Table 1(on next page)

Table 1. The published circRNAs with translation potential

1

2 **Table 1. The published circRNAs with translation potential**

CircRNA source	Research model	Translation product	Functions	Reference
CircRNAs in viruses and bacteria	Hepatitis δ virus	Protein of 122 amino acids	Hepatitis delta antigen (HDAg)	[50]
	A virusoid associated with rice yellow mottle virus	16-kDa highly basic protein	RNA-binding activity	[56]
	Escherichia coli: 795-nt circular mRNA	GFP	GFP reporter	[52]
Artificial circRNAs or synthetic modified RNA	HEK-293 cells	GFP	GFP reporter	[54]
	Rabbit reticulocyte lysate and HeLa cells	FLAG protein (EGF, IGF-1, IGF-2)	FALG reporter	[55]
	Rabbit reticulocyte lysate	23-kDa product	Not determined	[51]
	HEK293 cells	GFP, firefly luciferase, human erythropoietin	GFP reporter	[62]
Endogenous circRNAs	Drosophila: circMbl1	10-kDa protein	Enriched in synaptosomes and modulated by starvation and FOXO, with the exact function not yet determined	[59]
	Human: circ-ZNF609	circ-ZNF609-encoded protein	Specifically controlling myoblast proliferation	[63]
	Human: circ-FBXW7	FBXW7-185aa	Cooperating with FBXW7 to reduce the half-life of c-Myc by antagonizing USP28-induced c-Myc stabilization	[61]
	Human: circ-SHPRH	SHPRH-146aa	Protecting full-length SHPRH from degradation by the ubiquitin proteasome, leading to inhibited cell proliferation and tumorigenicity	[64]

3

4

Figure 1

Figure 1. Proposed circRNA formation models.

(A) The “exon skipping” model. (B) The “direct back-splicing” model. Black thin lines represent intron sequence; colored thick lines represent different exon sequences.

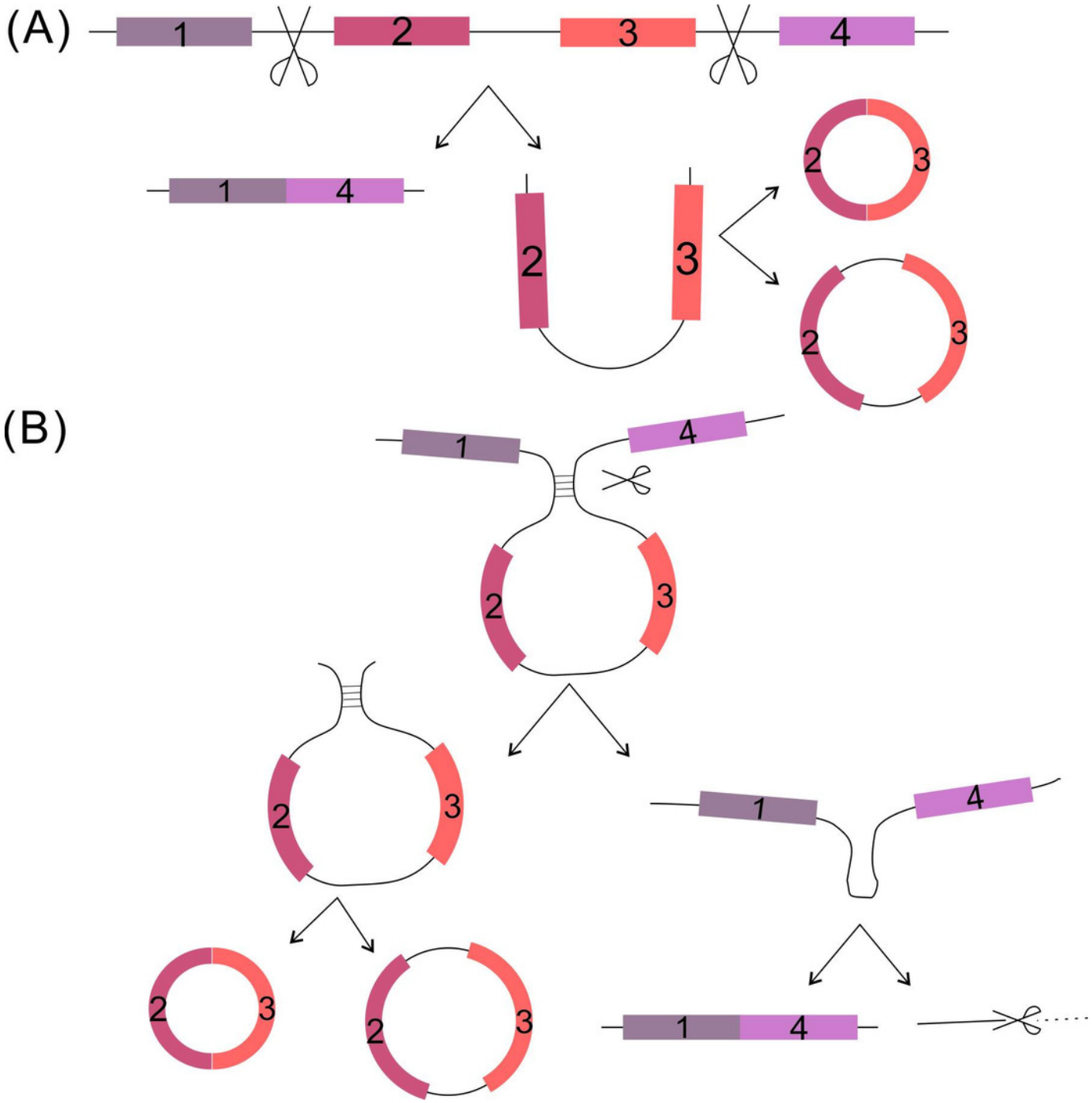
(A)
1
2
3
4
1
4
2
3
2
3
2
3
(B)
1
4
2
3
2
3
2
3
1
4
1
4

Figure 2

Figure 2. Functions of circRNAs.

(1) Molecular sponge for miRNA; (2) Regulation of transcription, splicing and expression of parental gene by binding to Pol II; (3) Interaction with proteins; (4) Direct translation of circRNAs.