Review Article

The functions and mechanisms of translatable circRNAs

Chang Liu\textsuperscript{1,2}, Xinying Wu\textsuperscript{1,2}, Priyanka Gokulnath\textsuperscript{3}, Guoping Li\textsuperscript{3,\ast}, Junjie Xiao\textsuperscript{1,2,\ast}

\textsuperscript{1}Institute of Geriatrics (Shanghai University), Affiliated Nantong Hospital of Shanghai University (The Sixth People’s Hospital of Nantong), School of Medicine, Shanghai University, Nantong 226011, China

\textsuperscript{2}Cardiac Regeneration and Ageing Lab, Institute of Cardiovascular Sciences, Shanghai Engineering Research Center of Organ Repair, School of Life Science, Shanghai University, Shanghai 200444, China

\textsuperscript{3}Cardiovascular Division of the Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA.

*Correspondence should be addressed to

Dr. Junjie Xiao
Cardiac Regeneration and Ageing lab, Institute of Cardiovascular Sciences, Shanghai Engineering Research Center of Organ Repair, School of Life Science, Shanghai University, 333 Nan Chen Road, Shanghai 200444, China
Tel: +86-21-66138131
Fax: +86-21-66138131
E-mail: junjiexiao@shu.edu.cn

Dr. Guoping Li
Cardiovascular Division of the Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA
Tel: +1-617-651-6142
E-mail: gli21@mgh.harvard.edu
Running Title: A review of translation in circRNAs

Text pages: 17
Tables: 2
Figures: 3
References: 59
Words in abstract: 130
Words in introduction: 360
Words in conclusion: 133

Abbreviations
BSJ: back-splicing junction
circRNAs: Circular RNAs
eIF4G2: the eukaryotic translation initiation factor 4G2
ITAFs: IRES-transacting factors
IRES: Internal ribosomal entry sequence
LC-MS: Liquid chromatography-tandem mass spectrometry
LNP: lipid nanoparticles
LSU: large subunit
m6A: N6-methyladenosine
meRIP-seq: methylated RNA immunoprecipitation sequencing
METTL3: methyltransferase-like 3
METTL14: methyltransferase-like 14
ORF: open reading frame
RBP: RNA-binding proteins
RNC-seq: Ribosome nascent-chain complex-bound RNA sequencing
SSU: small subunit
USP28: ubiquitin specific peptidase 28
YTHDF3: YTH domain family protein 3

Section assignment: Cellular and Molecular, non-coding RNAs, translatome.

Abstract
Circular RNAs (circRNAs) are covalently closed RNA produced by back-splicing. Circ RNAs have been considered as a type of non-coding RNAs for a long time. However, recent studies have shown that circRNAs can be translated into functional proteins. Proteins specifically encoded by circRNAs have been proved to play important roles in cancer pathology. In this review, we introduce the methods commonly used to identify and validate circRNA translation in detail. We also describe the major mechanisms driving the translation of these circRNAs. In addition, we summarize the main functions of the circRNA-encoded proteins in both physiological and pathological conditions. Finally, we discuss the therapeutic potential and challenges in the usage of synthetic translatable circRNAs. This brief review highlights recent discoveries made in this field and the progress of therapy based on translatable circRNAs.

KEYWORDS: circRNAs, translation, m6A, IRES, Rolling circle translation, cap-independent

Significance Statement
Understanding the translation of circRNA could facilitate the identification of novel drug targets in various diseases. Moreover, some circRNA encoded proteins were demonstrated to have therapeutic functions in cancer. The application of synthetic
circRNAs as carriers to achieve stable protein expression *in vitro* and *in vivo* has tremendous therapeutic potential.

1. Introduction

Circular RNAs (circRNAs) are covalently closed RNA molecules generated by back-splping of precursor mRNA. They were first identified in viroids as “single-stranded and covalently closed circular RNA molecules” (Sanger et al., 1976). CircRNAs were once considered as junk RNAs produced by mis-splicing. Moreover, circRNAs could not be detected from second-generation sequencing due to their lack of polyA tail. This technical limitation largely hindered their identification. Recently, with the development of high-throughput sequencing techniques and bioinformatic analysis, many circRNAs have been identified in various species (Gruner et al., 2016) (Westholm et al., 2014). circRNAs are characterized by their stable structure, conserved sequences and cell or tissue specific expression patterns (Rybak-Wolf et al., 2015). Until now, circRNAs have been shown to be involved in a variety of human diseases, including cardiovascular disease, chronic inflammation, diabetes, neurological diseases, and cancer(Kristensen et al., 2019; Lu et al., 2021; Wang et al., 2018; Wang et al., 2020; Zhou et al., 2020).

Due to their stability and abundance in human blood, circRNAs could act as potential biomarkers for diagnosis of many diseases specifically in humans (Bei et al., 2018; Zhang et al., 2018c). The biological functions of circRNAs have been widely studied and are described as follows(Li et al., 2018; Patop et al., 2019). circRNAs such as ciRS-7 and circSry can bind to and regulate specific microRNAs(Hansen et al., 2013). Moreover, circRNAs can interact with RNA binding proteins (RBP) and modulate their biological functions (Du et al., 2017a; Du et al., 2017b; Zeng et al., 2017). circRNAs can also regulate transcription (Li et al., 2015) and interfere with
splicing (Conn et al., 2017). Although circRNAs have been considered as a kind of non-coding RNAs until recently, newly emerging evidences suggest that several circRNAs are capable of being translated into functional proteins (Legnini et al., 2017; Pamudurti et al., 2017). In this review, we summarize the most commonly observed functions of the circRNA-translated proteins. Furthermore, we describe the major mechanisms driving the translation of these circRNAs. Finally, we discuss the potential therapeutic value of translated circRNAs. Summing up, with the development of this field, the translatable circRNAs and their products would be further characterized and applied towards disease treatment.

2. Identification and validation of translatable circRNAs

Identification and validation of translatable circRNAs mainly focus on large-scale identification, sequence translation ability validation, endogenous peptide detection, translation regulation elements detection and peptide functional studies. The strategies commonly applied for the identification and validation of translatable circRNAs include: (1) Large-scale identification by ribosome profiling and polysome profiling; (2) Sequence translation ability validation through dual luciferase vector system and flag-tagged protein expression followed by western blot detection; (3) Endogenous protein detection via specific antibody and liquid chromatography-tandem mass spectrometry (LC-MS) techniques; (4) Detection of m6A modification around the start codon by m6A-RIP-seq (Table 1).

Polysome profiling, ribosome nascent-chain complex-bound RNA sequencing (RNC-Seq) and ribosome profiling (Ribo-seq) are translatomic methods that can be used to identify unknown endogenous translatable circRNAs (Li et al., 2020; Pamudurti et al., 2017; Yang et al., 2017). Polysome profiling is a technique based on
sucrose gradient ultracentrifugation. Therefore, mRNA and translatable circRNA molecules bound by more ribosomes would sediment faster in sucrose gradient and the circRNAs bound to different number of ribosomes can be separated by the sucrose gradient ultracentrifugation (Figure 1). This technology is capable of detecting large changes in circRNA translation, such as shift in the number of ribosomes bound to certain circRNAs under stress condition. The main limitation of polysome profiling is its inability to perform analysis of all translating circRNAs, due to the low concentration of circRNAs in each gradient fraction (Zhao et al., 2019b). Moreover, the high concentration of sucrose inhibits certain enzymatic reactions in RNA-seq (King and Gerber, 2016). Therefore, the total amount of ribosome-bound circRNAs recovered from sucrose gradient is limited, usually only enough for RT-PCR quantification unless a large amount of input material is used.

Compared with polysome profiling, RNC-seq shows unique advantages by resolving this problem. To perform RNC-seq, cycloheximide pre-treated cell lysate is loaded onto a 30% sucrose layer and ultracentrifugation is then applied to sediment all the translating mRNAs and circRNAs associated with ribosomes. The RNC-RNAs can be recovered from the pelleted RNC for subsequent sequencing (Zhang et al., 2018b). RNA sequencing of RNC-RNAs by this technique reveals the full-length information of translating circRNAs. Despite of this advantage, the fragility of RNC often leads to ribosome dissociation and biased analyses of RNC-RNAs (Zhao et al., 2019b). Therefore, the centrifugation procedure is critical for the separation of intact RNC and subsequent RNA sequencing.

Through ribosome profiling, one can investigate translation at a finer resolution. To perform Ribosome profiling, cell lysates are treated with a low concentration of ribonuclease (RNase) to degrade RNAs without ribosome protection, after which
ribosomes are isolated using ultracentrifugation. After rRNA depletion, the 22-35nt ribosome protected RNA fragments (also known as ribosome footprints, RFPs) are enriched and analyzed by RNA sequencing to detect ribosome binding position. The distributions and densities of ribosomes on specific transcripts enable the deduction of start codon position, translation pause and termination position and real ORFs (Baudin-Baillieu et al., 2016), which could not be investigated by other translatomic methods. The main limitation of ribosome profiling is the complex and expensive experimental system. Moreover, the short length of the RFPs makes it difficult to cover the splice junctions of circRNAs.

The endogenous circRNA-encoded specific peptides can be validated by mass spectrometry (MS) (Chen et al., 2021; Liang et al., 2021). MS allows identification of existing proteins in a certain sample. To perform protein MS, proteins are digested by proteases to form a mixture of peptides. In the mass spectrometer, the peptide mixture is ionized to form charged ions. The electric field and magnetic field of the mass spectrometer separate the ions with a specific ratio of mass to charge (M/Z). The separated ions are collected by the detector to determine the M/Z value and the identity of each peptide. Most circRNAs are produced by protein-coding genes, and putative open reading frames (ORFs) in circRNAs often overlap with ORFs in the parent mRNA sequences (Pamudurti et al., 2017). Therefore, in order to avoid confusion with the proteins translated by the parent mRNA sequences, one option is to focus on the circRNAs with ORF crossing the back-splicing junction (BSJ) (Yang et al., 2017). Alternatively, verification can be done by inserting protein tags upstream of the ORF termination codon of circRNAs to facilitate subsequent protein identification or sub-cellular localization (Sharma et al., 2021). The major limitation of MS is its sensitivity, and as a result, some low abundance proteins may not be
detected. Moreover, traditional MS detects the protein stock in cells, while nascent peptides cannot be identified by MS alone. Single-molecule imaging of nascent peptides (SINAPS) technology has been designed to visualize translating RNAs in real-time. To perform SINAPs, Sun Tag epitopes were placed in the splicing reporter, and this reporter was transiently expressed in a specific reporter cell. Therefore, as the RNA of interest is translated, the nascent protein is immediately recognized by a fluorescent nanobody expressed in the same cell (Wang et al., 2021).

Dual-luciferase analysis system is a reporting system detecting luciferase activity using luciferin as the substrate. This bioluminescent system can be used to detect the activity of translation initiation fragments with extreme sensitivity and efficiency. Eukaryotic translation of circRNAs is primarily driven by internal ribosomal entry sequence (IRES) (Pamudurti et al., 2017). To assess whether the circRNAs have been translated, a dual-luciferase vector can be used to determine the function of IRES-like component on the circRNAs (Zhang et al., 2018b). To perform dual luciferase analysis, putative IRES fragment is cloned into a tandem luciferase reporter plasmid, and the luciferase activity of firefly luciferase (Luc) relative to that of Renilla luciferase (RLuc) was measured (Luc/RLuc) (Zhang et al., 2018b). Intensity of Luc/RLuc indicates the activity of putative IRES. The advantage of dual luciferase analysis is its ability to quantitatively test the putative activity of translation initiation fragments.

Another method of assessing the circRNA translation potential is to analyze N6-methyladenosine methylation (m6A) on circRNAs. Recent studies suggest that m6A modification can promote circular RNA translation (Di Timoteo et al., 2020; Yang et al., 2017). Therefore, m6A specific RNA immunoprecipitation sequencing (m6A-RIP-seq) can be applied to analyze the translation potential of circRNAs. To perform
m6A-RIP-seq, total RNAs are extracted and immunoprecipitated by anti-m6A antibody to enrich m6A modified RNAs for subsequent RNA sequencing. After RNA sequencing, backspliced junctions specific to circRNAs are identified from the sequencing reads.

Due to the advancements in sequencing technology, many circRNAs have been identified. In spite of this, annotation databases for translatable circRNA are hardly present. One notable such database is TransCirc, which is a multi-omics circRNAs translation database designed to address this issue (Huang et al., 2021). Currently, translatable human circRNAs have been collated in TransCirc, including (1) circRNAs identified by ribosome profiling; (2) experimentally validated circRNAs containing translation initiation site (TIS), internal ribosome entry site (IRES), and N-6-methyladenosine modification site (m6A); (3) circRNAs with specific coding frame (ORF); (4) other translatable circRNAs predicted using machine learning; (5) circRNAs with translation capability verified by mass spectrometry data. In addition to TransCirc, there is another database for identifying translatable circRNAs called “riboCIRC” (Li et al., 2021). This database includes 314 studies conducted on 21 species with more than 2000 circRNAs identified with translation potential captured using Ribo-Seq/RNA-seq. Both these databases provide critical clues for the study of circRNA translation and the function of such peptide products.

3. Mechanisms driving circRNA translation

Since circRNAs do not contain 5' caps, their translation is cap-independent. Eukaryotic circRNA translation can be preliminarily classified as: (1). Internal
riboosomal entry site (IRES) mediated translation. (2). N6-methyladenosine (m6A) mediated translation. (3). Rolling circle translation (Figure 2).

3.1 IRES mediated translation

Internal ribosomal entry site is capable of recruiting ribosomes for ribosome assembly and subsequent protein production. For more than three decades, IRES sequences that initiate translation in poliovirus have been known (Pelletier and Sonenberg, 1988). Subsequent studies have found that some cap-independent mRNAs recruit ribosomes to initiate translation through IRES (Komar and Hatzoglou, 2011). As early as in 1995, it was discovered that the IRES element can initiate protein translation of synthetic circRNAs (Chen and Sarnow, 1995). Recent studies have also demonstrated that endogenous circRNA can generate proteins through IRES mediated translation (Liang et al., 2019; Yang et al., 2018; Zhang et al., 2018a). In terms of validating the mechanism of IRES mediated circRNA translation, the sequence of circRNAs can be analyzed to see if there is an IRES sequence that can be folded into a structure similar to initial tRNA and ribosome binding circRNAs can be detected by polysome profiling. In addition, since IRES can recruit ribosomes and initiate translation with the assistance of IRES-transacting factors (ITAF), RNA pull down experiments can be performed to verify the binding of ITAF with circRNA and initiation of translation. Finally, the circRNA produced protein can be verified by western blot or mass spectrometry using protein-specific antibody.

3.2 m6A-dependent initiation of translation

It was found that circRNA has recognized m6A motif enrichment, and a single m6A site is sufficient to initiate translation(Yang et al., 2017). This m6A-activated
translation requires the initiation factor eIF4G2 and the m6A recognition protein YTHDF3, which is also enhanced by methyltransferase METTL3/14, inhibited by demethylase FTO, and upregulated by heat shock (Yang et al., 2017). In this study, polysome profiling, bioinformatic prediction methods and mass spectrometry experiments have shown that m6A-driven circRNA translation is widespread. Hundreds of circRNAs with translation potential exist in vivo. Interestingly, oncogenic human papillomaviruses (HPVs) generated circE7 is m6A modified and can produce a functional E7 oncoprotein (Zhao et al., 2019a). This virus-derived protein coded by circRNA can interact with retinoblastoma protein (pRb) and promote cancer cell proliferation. In terms of verifying the mechanism of m6A-driven translation, sequence analysis can be performed to determine whether there is an m6A motif on the circRNAs. meRIP-seq can also be used to identify circRNAs with m6A modification (Zhao et al., 2019a). Ribosome binding circRNAs have been detected using polysome profiling. Also, RNA pull down can be performed on readers such as YTHDF3 and translation initiation factors such as eIF3A to verify if these factors bind to circRNAs and thus confirm the initiation of translation (Yang et al., 2017).

3.3 Rolling circle translation

CircRNAs can translate in a pattern that simulates DNA rolling circle amplification. In 2013, researchers artificially synthesized a circRNA and demonstrated that it translates proteins mimicking the DNA rolling circle amplification model of prokaryotes (Abe et al., 2013). Subsequent studies have demonstrated the protein coding ability of synthetic circRNAs in a rolling circle manner as observed in human Hela cells (Abe et al., 2015). Recent investigations have shown that the highly
expressed circRNA circ-E-Cad in glioblastoma can undergo a unique over 360-degree translation to produce the 254aa protein C-E-Cad(Gao et al., 2021). Rolling circle translation also occurs on some circRNAs without the stop codon, thereby forming an infinite Open Reading Frame (iORF). One recent study has shown that circRNA “circEGFR” formed by 14-15 exons of EGFR was significantly enriched in glioblastoma. Interestingly, the authors verified the existence of infinite ORF (iORF) in circEGFR, which was shown to be translated into a polymetric novel protein-complex, called rolling-translated EGFR (rtEGFR). The functional studies showed that interference with circEGFR significantly reduced the tumorigenicity in mouse models. Mechanistically, rtEGFR binds to and stabilizes EGFR, continuously activating the EGFR pathways (Liu et al., 2021). For the mechanism verification of rolling circle translation, the ORF sequence analysis can be carried out to confirm whether there is a multi-round ORF or iORF existing. Moreover, 3×Flag can be directly added to the N-terminal of the original iORF, and the synthetically tagged protein products could be detected using western blot. Finally, specific antibody western blotting or mass spectrometric methods could be used to detect endogenous protein products.

4. The various functions of translatable circRNAs

4.1 Role of translatable circRNAs in cancer

Although there have been many reports on circRNA translation, the functional characterization of the translated products remains unclear (Legnini et al., 2017;
Pamudurti et al., 2017; van Heesch et al., 2019; Zheng et al., 2019). Thousands of circRNAs have been predicted to contain putative open reading frame (ORF) and IRES, however, only a few circRNAs have been validated as protein-coding, including circZNF609, circMbl, circFBXW7, circ-LINCPINT, circSHPRH etc. (Legnini et al., 2017; Pamudurti et al., 2017; Yang et al., 2018; Zhang et al., 2018a; Zhang et al., 2018b), and many of them were proved to play pivotal roles in cancer pathogenesis (Table 2). It was found that about 40% of translatable circRNAs share the starting codon with their parent genes, and the peptides encoded by such circRNAs share the same N-terminal with the protein encoded by the parent genes (Pamudurti et al., 2017). Therefore, such truncated circRNA translated proteins may act as competitors of the proteins encoded by their parent genes (Figure 3). For example, AKT3-174aa can competitively bind to PDK1, thereby inhibiting the phosphorylation of PDK1 at site 308 of AKT3 protein, and thus inhibiting the PI3K/AKT signaling pathway (Xia et al., 2019). In addition, some circRNAs encoded products can also serve as protective agents against parent protein degradation, such as SHPRH-146aa, which can bind DTL and protect the full-length E3 ubiquitin protein ligase SHPRH from ubiquitination and degradation by DTL (Zhang et al., 2018a). Another circRNA protein, FBXW-185aa can competitively bind USP28 to protect the stability of FBXW7α and play an inhibitory role in human glioblastoma (Yang et al., 2018). Also, circ-beta-catenin can activate Wnt signaling pathway via antagonizing GSK3β induced β-catenin phosphorylation and degradation thereby promoting liver cancer (Liang et al., 2019). circEGFR could stabilized EGFR, sustain EGFR signaling and promote glioblastoma tumorigenicity (Liu et al., 2021). Moreover, circRNA encoded product can interact with its parent protein, such as SMO-193aa produced by circSMO (Wu et al., 2021). SMO-193aa binds to the N-terminal of SMO protein and
activate SMO by transporting cholesterol to SMO in glioblastoma.

Interestingly, some circRNA translation products function independently of the proteins encoded by their parent genes. For example, it was found that the function of circ-E-cad encoded C-E-Cad protein was completely opposite to that of its parent protein E-Cadherin. In fact, C-E-Cad has a unique 14-amino-acid carboxy terminus and can interact with EGFR to significantly promote the activation of the STAT3/AKT/ERK pathway. And therapy targeting C-E-Cad can significantly inhibit tumor progression (Gao et al., 2021). Moreover, the circRNA circ-LINCPINT generated by the back-spliced LncRNA (LINCPINT) can produce a peptide named as PINT87aa. PINT87aa has the function of binding PAF1 complex to inhibit the transcription of oncogenes (Zhang et al., 2018b).

4.2 Role of translatable circRNAs in other systems

Besides the important roles of circRNA encoded peptides in cancer pathogenesis, they were also proved to have regulatory functions in other physiological or pathological conditions (Table 2). For example, circular repeat-containing intron of C9ORF72, a circRNA composed of spliced intron with G-rich repeats, is involved in pathogenesis of amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD). This circular intron serves as the translation template of toxic dipeptide repeat (DPR) proteins (Wang et al., 2021). Interestingly, this translation is driven by repeat-associated non-AUG translation (RAN). A recent study shows that a large number of circRNAs are synthesized with increasing abundance during murine spermatogenesis (Tang et al., 2020). Interestingly, a half of these male germ cell circRNAs contain large ORFs with m6A-modified start codons in their junctions. These translatable circRNAs provide continuous supply of proteins critical for late spermiogenesis and
normal sperm functions. Translatable circRNAs may have roles in synaptic function. For example, circMbl1 and the putative circMbl1-encoded peptide are present in synaptosome fractions of Drosophila head extracts (Pamudurti et al., 2017). Further study shows that starvation and FOXO regulate the translation of circMbl, indicating its possible role in fasting regulated synaptic function. circ-ZNF609 was proved be functional in myoblast proliferation, and it can translate into a protein in myoblast (Legnini et al., 2017). Although direct relationship of this phenotype and its protein-coding capacity was not proved, this study provided another example of a protein-coding circRNA in eukaryotes.

CircRNA-derived peptides can act as intracellular regulators in response to cellular stress (Pamudurti et al., 2017; Yang et al., 2017). At present, the functions of circRNA-encoded peptides in pathogenesis of many diseases are still not fully illustrated. circRNA encoded peptides may also act as regulators in cardiovascular diseases such as ischemic cardiomyopathy, drug-induced cardiomyopathy, and other cardiac pathological conditions. In accordance with that, human cardiac translatome study has also identified certain translatable circRNAs (van Heesch et al., 2019). Whether these translated products of circRNA play regulatory role in human cardiac pathology and how they do so, requires further investigation.

5. Engineering circRNA translation for therapeutics and pharmacology: potential outcomes and challenges

One of the long-term goals of mRNA biotechnology has been stable protein production from exogenous mRNA. The structure of circRNAs allows them to avoid recognition by the innate immune system and exonuclease enzymes and to be more
stable than linear mRNAs. Due to their stability and low immunogenicity, it is highly advantageous to use circRNAs as a protein template (Wesselhoeft et al., 2019). Moreover, the folded circRNAs produce a "smaller" conformation than linear RNAs. Therefore, an increased number of circRNAs can be loaded into lipid nanoparticles (LNP) and further improve the delivery efficiency of RNA therapy. Recently, a circRNA vaccine that encodes the trimeric receptor-binding domain (RBD) of SARS-CoV-2 spike protein (circRNA\textsuperscript{RBD}) has been reported (Qu et al., 2021). LNP delivered circRNA\textsuperscript{RBD} successfully elicited sustained neutralizing antibodies and T cell responses in mice. Summing up, the synthesis of translatable circRNAs is of great value in the field of biomedicine.

However, challenges also exist. circRNAs are inefficient in circularization, difficult to purify, and their ability of protein production is weak (Wesselhoeft et al., 2018). These problems limit the potential applications of circRNAs driven protein expression. Recent studies have demonstrated improved circularization, translation efficiency and protein stability by artificially engineering circRNAs (Wesselhoeft et al., 2018). In this study, researchers have successfully engineered a self-splicing intron to efficiently circularize RNAs. In another study (Tang et al., 2020), researchers observed that different IRES elements and exon length influenced circRNA expression and translation. Moreover, to improve the efficiency of circRNA translation, m6A modification can also be considered. Interestingly, m6A modification may simultaneously inhibit innate immunity triggered by circRNAs (Chen et al., 2019). Taken together, these data provide newer insights into optimizing the design of synthetic circRNAs.

Although the immunogenicity of circRNAs is lower than that of linear mRNAs, studies have shown that cells transfected with exogenous circRNAs can activate their
disease-resistant gene mechanisms (Chen et al., 2017). The degree of innate immune activation stimulated by circRNAs may be related to the structural design and purification methods used to generate the circRNAs (Wesselhoeft et al., 2019). To further inhibit the immune response to circRNAs, nucleoside modification and other methods such as circularization by T4 RNA ligase may be incorporated in engineering circRNAs (Liu et al., in press). Moreover, their previous work show that endogenic circRNAs with 16-26 bp double-stranded structure can inhibit the abnormal activation of the immune factor PKR (Liu et al., 2019), suggesting that in vitro synthesis of non-immunogenic circRNAs with short double-stranded structure may be a novel approach to optimize circRNAs synthesis (Liu et al., in press). Another direction of optimizing circRNA therapy is to develop more efficient delivery strategies in vivo. One study has shown to achieve efficient delivery of circRNA into cells via lipid nanoparticle (LNP) both in vitro and in vivo (Wesselhoeft et al., 2019). Moreover, another recent study demonstrated AAV-mediated tissue-specific circRNA translation in vivo (Meganck et al., 2021). These studies show the possibility of applying circRNA translation in the experimental animal models. Therefore, optimizing the design of circRNA structure and development of novel circRNAs delivery technologies will largely facilitate industrialization and clinical application of therapies and pharmaceutics based on translatable circRNA in the future.

6. Conclusion

Summing up, a growing number of studies have reported that circRNA-encoded proteins are involved in regulation of the human physiology and pathology, and have been found to have a variety of biological functions with important clinical significance. At present, most of the reported translatable circRNA uses IRES driven
translation, but few studies have reported m6A-driven circRNA translation (Table 2). It is believed that with the development of RNA deep sequencing technology and algorithm, numerous proteins/peptides encoded by circRNA will be further identified, verified, and their biological functions will be revealed. In addition, with improvements in the genetic tool-box, the development and application of synthetic circRNAs as carriers to achieve stable and tissue-specific protein expression in vitro and in vivo will go on to become an important promising therapeutic strategy in the near future.

**Authorship contributions**

J.X. and G.L. participated in research design. X.W. performed the data analysis. C. L. and P. G. contributed to the writing of the manuscript. All authors read and approved the final version of the manuscript, and ensure it is the case.

**Declaration of Interests**

The authors declare no competing interests.

**References**


with minimized immunogenicity as potent PKR inhibitors. Molecular Cell.


Rybak-Wolf, A., Stottmeister, C., Glazar, P., Jens, M., Pino, N., Giusti, S., Hanan, M., Behm,


Funding Footnote

This work was supported by the grants from National Key Research and Development Project (2018YFE0113500 to J.X.), National Natural Science Foundation of China (82020108002 and 81911540486 to J.X.), Innovation Program of Shanghai Municipal Education Commission (2017-01-07-00-09-E00042 to J.X.), the grant from Science and Technology Commission of Shanghai Municipality (20DZ2255400 and 21XD1421300 to J.X.), the “Dawn” Program of Shanghai Education Commission (19SG34 to J.X.), and National Natural Science Foundation of China (82000287 to C.L.).

Figure legends

Figure 1. The major translatomic methods for circRNAs

Figure 2. The major driving mechanisms of circRNA translation

Figure 3. The functions of translatable circRNA
Table 1. Methods used to study translation of circRNAs

<table>
<thead>
<tr>
<th>Purposes</th>
<th>Technique</th>
<th>Description</th>
<th>Application</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detection</td>
<td>Ribosome profiling</td>
<td>Ribosomal isolation is followed by digestion with low concentrations of RNase T1, followed by isolation of protected RNA fragments between ribosomal monomer subunits</td>
<td>Profiles circRNAs with possible translation activity, thereby providing indirect evidence of translation potential of circRNAs</td>
<td>(van Heesch et al., 2019)</td>
</tr>
<tr>
<td>Validation</td>
<td>Mass spectrometry</td>
<td>Customized specific antibodies were used for IP and LC-MS /MS was performed after enrichment of the target band</td>
<td>Verify and confirm the corresponding peptide in vivo</td>
<td>(Chen et al., 2021; Liang et al., 2021; Tang et al., 2020)</td>
</tr>
<tr>
<td>Validation</td>
<td>Dual luciferase reporter</td>
<td>Putative IRES fragment was cloned into a tandem RLuc–Luc reporter plasmid, between RLuc and Luc, and the luciferase activity of Luc relative to that of RLuc was measured</td>
<td>To quantitatively test the putative activity of translation initiation fragments</td>
<td>(Zhang et al., 2018b)</td>
</tr>
<tr>
<td>Detection</td>
<td>Polysome profiling</td>
<td>Polysome-associated RNAs were separated with sucrose gradients. Then polysome-bound circRNAs were enriched by an RNase R treatment and identified through paired-end deep</td>
<td>Profiles circRNAs with possible translation activity and detect large changes in circRNA translation</td>
<td>(Ye et al., 2021)</td>
</tr>
</tbody>
</table>
m6A antibody was applied for RIP to enrich m6A-containing RNA, which was analyzed by high-throughput sequencing. Providing indirect evidence of translation potential of circRNAs (Yang et al., 2017)

The cell lysate is loaded onto a 30% sucrose cushion and ultracentrifugation is performed to isolate the RNC-RNAs for next-generation sequencing. Reveals the full-length information of translating circRNAs (Zhang et al., 2018b)

As the RNA is translated, the nascent protein is immediately recognized by a fluorescent nanobody expressed in the same cell. In situ visualization of translating circRNAs (Wang et al., 2021)

<table>
<thead>
<tr>
<th>Detection</th>
<th>m6A-RIP-seq</th>
<th>m6A antibody was applied for RIP to enrich m6A-containing RNA, which was analyzed by high-throughput sequencing. Providing indirect evidence of translation potential of circRNAs (Yang et al., 2017)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detection</td>
<td>RNC-seq</td>
<td>The cell lysate is loaded onto a 30% sucrose cushion and ultracentrifugation is performed to isolate the RNC-RNAs for next-generation sequencing. Reveals the full-length information of translating circRNAs (Zhang et al., 2018b)</td>
</tr>
<tr>
<td>Validation</td>
<td>SINAPS</td>
<td>As the RNA is translated, the nascent protein is immediately recognized by a fluorescent nanobody expressed in the same cell. In situ visualization of translating circRNAs (Wang et al., 2021)</td>
</tr>
</tbody>
</table>

IRES: LC-MS: Liquid chromatography-tandem mass spectrometry; Luc: Firefly luciferase; RIP: RNA immunoprecipitation; Rluc: Renilla luciferases; RNC-seq: ribosome nascent-chain complex-bound RNA sequencing; SINAPS: Single-molecule imaging of nascent peptides
<table>
<thead>
<tr>
<th>Identified in</th>
<th>Coding circRNA name</th>
<th>Identification and validation strategies</th>
<th>Translation initiation driven by</th>
<th>Functions of translation products</th>
<th>Translation products interacted with</th>
<th>Functio n related to parent gene</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drosophila heads</td>
<td>CircMbl</td>
<td>Ribosome profiling, mass spectrometry</td>
<td>IRES mediated cap-independent translation</td>
<td>Potential roles in acute fasting regulated synaptic function</td>
<td>N/A</td>
<td>N/A</td>
<td>(Pamudurti et al., 2017)</td>
</tr>
<tr>
<td>Myoblasts</td>
<td>CircZNF609</td>
<td>Mass spectrometry</td>
<td>IRES mediated cap-independent translation in a splicing-dependent manner</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>(Legnini et al., 2017)</td>
</tr>
<tr>
<td>ALS and FTD</td>
<td>Circular repeat-containing intron of C9ORF72</td>
<td>SINAPS</td>
<td>RAN translation</td>
<td>Toxic DPR proteins</td>
<td>N/A</td>
<td>N/A</td>
<td>(Wang et al., 2021)</td>
</tr>
<tr>
<td>Male germ cells</td>
<td>CircRNAs derived from Hook1 and Ranbp9</td>
<td>Detection of circRNA junction peptides using LC-MS, m6A-RIP-seq</td>
<td>m6A mediated cap-independent translation</td>
<td>Provide continuous supply of proteins critical for late spermiogenesis and normal sperm functions</td>
<td>N/A</td>
<td>N/A</td>
<td>(Tang et al., 2020)</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Glioblastoma</td>
<td>Circ-LINCPINT</td>
<td>RNC-seq, LC-MS, dual-luciferase vector system</td>
<td>IRES mediated cap-independent translation</td>
<td>Inhibit the transcriptional elongation of multiple oncogenes and inhibit proliferation of glioma cells</td>
<td>PAF1 complex</td>
<td>No</td>
<td>(Zhang et al., 2018b)</td>
</tr>
<tr>
<td>Glioblastoma</td>
<td>CircFBXW7</td>
<td>RNA sequencing, LC-MS, dual-luciferase vector system</td>
<td>IRES mediated cap-independent translation</td>
<td>Free FBXW7α to degrade c-Myc, inhibit glioma cell proliferation</td>
<td>USP28</td>
<td>Yes</td>
<td>(Yang et al., 2018)</td>
</tr>
<tr>
<td>Glioblastoma</td>
<td>CircSHPRH</td>
<td>RNA sequencing, dual-luciferase vector system,</td>
<td>IRES mediated cap-independent</td>
<td>Protect SHPRH against degradation, helps SHPRH</td>
<td>DTL</td>
<td>Yes</td>
<td>(Zhang et al., 2018a)</td>
</tr>
<tr>
<td>Tumour Type</td>
<td>CircRNA</td>
<td>RNA Sequencing, LC-MS, Dual-luciferase Vector System</td>
<td>IRES Mediated Cap-independent Translation</td>
<td>Transport Cholesterol, Activates SMO, and Promotes Tumor Development</td>
<td>SMO</td>
<td>Yes/No</td>
<td>Reference</td>
</tr>
<tr>
<td>-------------</td>
<td>---------</td>
<td>--------------------------------------------------</td>
<td>------------------------------------------</td>
<td>---------------------------------------------------------------</td>
<td>-----</td>
<td>-------</td>
<td>-----------</td>
</tr>
<tr>
<td>Glioblastoma</td>
<td>Circ-SMO</td>
<td>RNA sequencing, LC-MS</td>
<td>IRES mediated cap-independent translation</td>
<td>Transport cholesterol, activates SMO, and promote tumor development</td>
<td>SMO</td>
<td>Yes</td>
<td>(Wu et al., 2021)</td>
</tr>
<tr>
<td>Glioblastoma</td>
<td>Circ-E-Cad</td>
<td>RNA sequencing, LC-MS</td>
<td>Rolling circle translation</td>
<td>Activate EGFR–STAT3 Signaling and promote glioblastoma tumorigenicity</td>
<td>EGFR</td>
<td>No</td>
<td>(Gao et al., 2021)</td>
</tr>
<tr>
<td>Glioblastoma</td>
<td>CircEGFR</td>
<td>RNA sequencing, mass spectrometry</td>
<td>Rolling circle translation</td>
<td>Stabilized EGFR, sustain EGFR Signaling and promote glioblastoma</td>
<td>EGFR</td>
<td>Yes</td>
<td>(Liu et al., 2021)</td>
</tr>
<tr>
<td>Tumor Type</td>
<td>CircRNA</td>
<td>Molecular Profiling Methods</td>
<td>Translation Mechanism</td>
<td>Function</td>
<td>Target</td>
<td>Pub. Year</td>
<td></td>
</tr>
<tr>
<td>---------------------------</td>
<td>-------------------</td>
<td>------------------------------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>------------------</td>
<td>---------------</td>
<td></td>
</tr>
<tr>
<td>Glioblastoma</td>
<td>CircAKT3</td>
<td>High-throughput RNA sequencing, mass spectrometry</td>
<td>IRES mediated cap-independent translation</td>
<td>Inhibit the phosphorylation of AKT3 protein by PDK1, and block PI3K/AKT signaling pathway</td>
<td>PDK1</td>
<td>Yes (Xia et al., 2019)</td>
<td></td>
</tr>
<tr>
<td>Hepatocellular carcinoma</td>
<td>Circ-beta-catenin</td>
<td>RNA sequencing, dual-luciferase vector system, mass spectrometry</td>
<td>IRES mediated cap-independent translation</td>
<td>Activate Wnt signaling pathway via antagonizing GSK3β induced β-catenin phosphorylation and degradation</td>
<td>GSK3β</td>
<td>Yes (Liang et al., 2019)</td>
<td></td>
</tr>
<tr>
<td>Human colon cancer</td>
<td>CircPPP1R12A</td>
<td>LC-MS, FISH, circRNA microarray</td>
<td>IRES mediated cap-independent translation</td>
<td>Promote cell proliferation via activating Hippo-YAP signaling</td>
<td>N/A</td>
<td>N/A (Zheng et al., 2019)</td>
<td></td>
</tr>
<tr>
<td>Pathway</td>
<td>CircRNA</td>
<td>Polysome Profiling</td>
<td>m6A Mediated Cap-Independent Translation</td>
<td>Promote Cancer Cell Proliferation</td>
<td>pRb</td>
<td>Yes/No</td>
<td>Reference</td>
</tr>
<tr>
<td>---------------------------------------------</td>
<td>---------</td>
<td>--------------------</td>
<td>------------------------------------------</td>
<td>----------------------------------</td>
<td>-----</td>
<td>--------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>HPV-positive cancers</td>
<td>CircE7</td>
<td>m6A mediated cap-  independent translation</td>
<td>Promote cancer cell proliferation</td>
<td>pRb</td>
<td>Yes</td>
<td>(Zhao et al., 2019a)</td>
<td></td>
</tr>
<tr>
<td>Triple negative breast cancer</td>
<td>CircHER2</td>
<td>IRES mediated cap- independent translation</td>
<td>Activate EGFR Signaling and promote breast cancer</td>
<td>EGFR/HER 3</td>
<td>Yes</td>
<td>(Li et al., 2020)</td>
<td></td>
</tr>
</tbody>
</table>

ALS: amyotrophic lateral sclerosis; DCM: dilated cardiomyopathy; DPR: dipeptide repeat proteins; DTL: Denticelless E3 Ubiquitin Protein Ligase; EGFR: epidermal growth factor receptor; FTD: frontotemporal dementia; IRES: Internal ribosomal entry sequence; LC-MS: Liquid chromatography-tandem mass spectrometry; PAF1: polymerase associated factor 1; PDK1: pyruvate dehydrogenase kinase 1; pRb: retinoblastoma protein; RAN: Repeat-associated non-AUG translation; RNC-seq: ribosome nascent-chain complex-bound RNA sequencing; SINAPS: single-molecule imaging of nascent peptides; SMO: smoothened; USP28: ubiquitin specific peptidase 28; UTRs: the untranslated regions;
Figure 1

Cell lysis

Sucrose gradient

10%

50%

Fractionation

Free mRNAs
SSU
LSU

RT-PCR quantification

Polysome profiling

Ribosome nascent-chain complex-bound RNA sequencing

RNC pellet

30% Sucrose

RNC-mRNA isolation

Ribosome profiling

RNase digestion

rRNA depletion

Ribosome footprints (RFPs)

Translating mRNAs and circRNAs

RNC-seq

Ribo-seq
Figure 2

Back-splicing

circRNA

Nucleus

Cytoplasm

IRES-mediated translation
ITAFs
circRNA
Ribosome
Protein

m6A-mediated translation
METTL3/14
YTHDF1/2/3
Me
IRES
ATG
circRNA
Ribosome
Protein

Rolling circle translation
circRNA
ATG
Ribosome
Protein

This article has not been copyedited and formatted. The final version may differ from this version.
Figure 3

Competitors of parent protein

- PDK1
- AKT
- AKT3-174aa

Interact with its parent protein

- EGFR
- SHPRH

Protectors against degradation of parent protein

- Ub
- SHPRH-146aa

Independent of the parent protein

- PAF1
- PAF1 target gene transcription

- PINT-87aa

EGFR endocytosis and degradation