RNA PROCESSING AND MODIFICATIONS

The biogenesis and emerging roles of circular RNAs

Ling-Ling Chen

Abstract | Circular RNAs (circRNAs) are produced from precursor mRNA (pre-mRNA) back-splicing of thousands of genes in eukaryotes. Although circRNAs are generally expressed at low levels, recent findings have shed new light on their cell type-specific and tissue-specific expression and on the regulation of their biogenesis. Furthermore, the data indicate that circRNAs shape gene expression by titrating microRNAs, regulating transcription and interfering with splicing, thus effectively expanding the diversity and complexity of eukaryotic transcriptomes.

Circular RNAs are covalently closed, single-stranded transcripts comprising many RNA species. One type of circular RNAs (termed circRNAs), which are produced from precursor mRNA (pre-mRNA) back-splicing in higher eukaryotes, were discovered more than 20 years ago; however, only a handful of circRNAs were found, and they were thought to be by-products of aberrant splicing with little functional potential[1–5]. Next-generation RNA sequencing (RNA-seq) of non-polyadenylated transcriptomes has recently shown that the expression of circRNAs is widespread[6–13]. For example, over 10% of expressed genes in examined cells and tissues can produce circRNAs[7,10,14].

CircRNAs are presumably produced from pre-mRNA back-splicing of exons, resulting in the exclusion of a circular RNA molecule with a 3′,5′-phosphodiester bond at the junction site (FIG. 1a). Recent studies aimed at understanding how circularization is achieved have revealed that back-splicing circularization is processed by the spliceosome and modulated by both *cis* and *trans* regulators[9,12,15–19]. Although the functions of most circRNAs are unknown, ongoing studies have shown that some circRNAs may have potentially important roles in gene regulation[8,14,20–24]. In this Progress article, I discuss recent findings on the widespread detection of circRNAs, their biogenesis and their potential functions in gene expression, such as titration of microRNAs (miRNAs), regulation of transcription and interference with pre-mRNA splicing. I also highlight the major challenges for future research.

The broad expression of circRNAs

Different types of circular RNAs, generated by distinct mechanisms, have been described (BOX 1). CircRNAs from pre-mRNA back-splicing of exons are rarely seen in polyadenylated RNA profiling studies, because they lack poly(A) tails. The recent generation of non-polyadenylated transcriptomes or the enrichment of circRNAs by treating samples with RNase R, which is an enzyme that digests linear RNAs but preserves circRNAs, followed by RNA-seq and specific computational pipelines that aim to identify reads spanning the back-splicing junctions in a reversed orientation, has uncovered over ten thousand different circRNAs in metazoans, from fruit fly and worm[11,12] to mouse and human[6–10,13,25]. CircRNAs were also found in plants such as *Oryza sativa*[26] and *Arabidopsis thaliana*[27], and in other organisms including fungi and protists[27–29].

CircRNAs are often expressed at low levels[7–10,30], arguing for the possibility that the majority of circRNAs are inert splicing by-products. However, approximately 50 genes have circRNAs that are expressed more abundantly than their linear isoforms in each examined cell line[30], and dozens of circRNAs have been experimentally confirmed to be highly expressed in a cell type-specific or tissue-specific manner[16,17]. Furthermore, thousands of circRNAs accumulate in the brain, and hundreds are upregulated during the human epithelial–mesenchymal transition (EMT) (BOX 2). Moreover, there is no clear correlation between levels of circRNAs and their corresponding mRNAs[10,16,18,30–32]. These studies indicate that the higher circRNA expression levels observed in some cells or tissues are not simply by-products of increased transcription and are therefore suggestive of functionality. It is worthwhile noting, however, that a genome-wide analysis of high-confidence circRNAs has revealed that their expression is not more cell type-specific than that of their linear counterparts[10]. Thus, to what degree expression of circRNAs is different to that of their linear transcripts in diverse cell types and tissues has not yet been fully determined. Furthermore, as circular RNAs are stable[7,8,21] and may accumulate to different levels in cells with different division rates, whether the detected steady-state levels of circRNAs truly reflect the kinetics of circRNA processing remains an intriguing topic for future investigation.

CircRNA biogenesis

CircRNAs are derived from pre-mRNAs, which are transcribed by RNA polymerase II (Pol II)[9,15–17]. The regulation of circRNA biogenesis depends, in principle, on the *cis*-regulatory elements and *trans*-acting factors that control splicing.

Spliceosome-dependent biogenesis. The processing of the back-splicing reaction seems to require the canonical spliceosomal machinery. Mutagenesis experiments of minigene vectors that express circRNAs have shown that exon circularization efficiency is dependent on the presence of canonical splice sites bracketing the exons[15,17,33]. However, although the kinetics of back-splicing have not yet been measured in cells, in general, the efficiency of back-splicing seems to be much lower than that of canonical splicing, as steady-state levels of circRNAs are often much lower than those of their linear counterparts[9,15–17]. The low back-splicing efficiency is possibly due to the fact that spliceosomes are unfavourably assembled at back-splicing

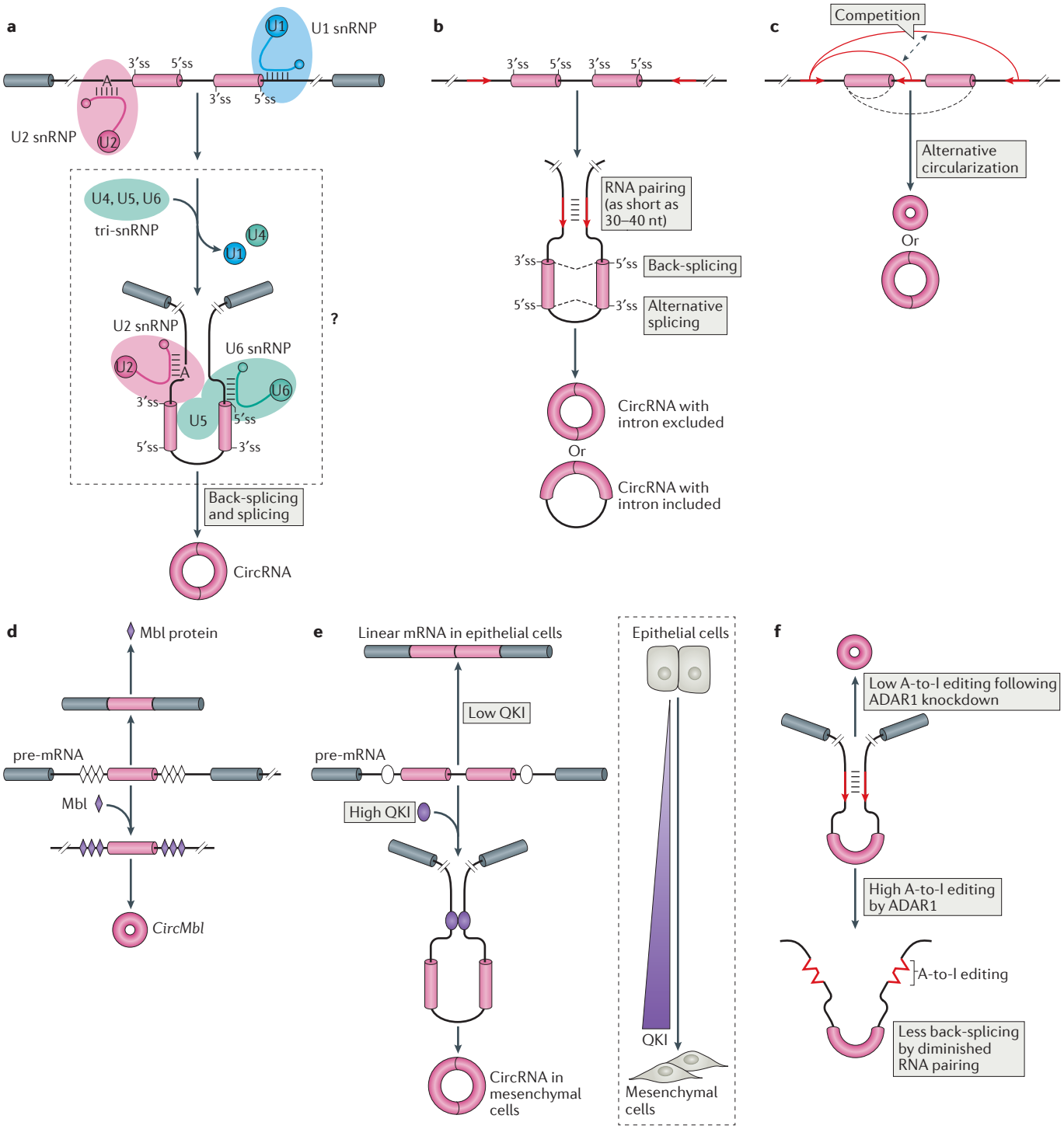

Figure 1 | **Circular RNA (circRNA) biogenesis and its regulation.** **a** | CircRNA is produced by back-splicing circularization, which is catalysed by the spliceosomal machinery[15,17,33]. Although it remains unknown exactly how the spliceosome is involved in back-splicing (dashed box), it is proposed that following the sequential assembly of spliceosomal small nuclear ribonucleoproteins (snRNPs) on the precursor mRNA (pre-mRNA), a downstream 5′ donor site of an exon (magenta bars) is joined to an upstream 3′ acceptor site to produce a circRNA. **b** | Orientation-opposite complementary sequences (red arrows) in introns flanking circularized exons promote circRNA biogenesis by pairing to form RNA duplexes that juxtapose the splice sites[9,16,19]. However, alternative splicing may occur and lead to the formation of circRNA with a retained intron (black line). **c** | Competition for RNA pairing across introns can lead to the formation of different circRNAs from a single gene[9]. **d** | Muscleblind (Mbl) protein enhances circRNA production from its own gene by binding to introns bracketing the circularized exon[15]. White diamonds denote Mbl-binding sites. **e** | A proposed model for quaking (QKI) involvement in circRNA biogenesis during the epithelial–mesenchymal transition (EMT)[18]. QKI is upregulated during EMT. Low QKI expression in epithelial cells promotes exon inclusion in the linear mRNA, whereas high QKI expression in mesenchymal cells facilitates circRNA formation by juxtaposing the circularized exons. White ovals denote QKI binding sites in QKI-regulated pre-mRNAs. **f** | A-to-I RNA editing by adenosine deaminase 1 acting on RNA (ADAR1) can suppresses circRNA formation by disrupting RNA pairing between flanking introns[12,31]. ss, splice site.

sites to catalyse the ligation of downstream 5′ donor sites with upstream 3′ acceptor sites (FIG. 1a). How exactly the spliceosome is involved in back-splicing is unclear, and future studies using *in vitro* back-splicing assays carried out in nuclear extracts may provide additional mechanistic insights into circRNA biogenesis.

Mutating the polyadenylation signal of circRNA-producing pre-mRNAs eliminated circRNA production from expression vectors[16]. This observation suggests that circularization may occur post-transcriptionally. Subsequent work revealed examples in which a polyadenylation signal was not required for circRNA production from minigene vectors[19], indicating that circRNA formation may also occur co-transcriptionally. Indeed, circRNAs could be detected in chromatin-associated RNA from fly heads, further implying the co-transcriptional nature of their processing[15]. However, it should be noted that chromatin-associated RNA is not precisely equivalent to nascent RNA, as transcripts with mature, polyadenylated 3′ ends were found in chromatin fractionations[34]. Thus, how circRNA biogenesis is linked to transcription remains to be investigated. Studying nascent circRNAs will help to reveal the processing kinetics of circRNA biogenesis and its potential coupling with other mRNA maturation steps on a genome-wide scale.

Cis-*elements facilitate circRNA formation.* Most circRNAs comprise several exons, usually two or three[9]. It seems that no specific exonic sequence besides the splice site is required for back-splicing[17]. In addition, multiple circRNAs, with the internal intron included or excluded, can be produced from the same gene by alternative splicing[9,10,22,30] (FIG. 1b). However, it is interesting to note that back-splicing may require a minimal exon length. For example, it has been shown for highly expressed circRNAs that exons of single-exon-containing circRNAs are much longer than exons of circRNAs comprising several exons[9,35].

Back-splicing frequently requires regulatory elements residing in the flanking introns of the circularized exons. Most circRNAs in mammals[7,9] and in *Caenorhabditis elegans*[12], and a few circRNAs in *Drosophila melanogaster*[19], are processed from internal exons with long flanking introns, usually containing reverse complementary sequences. Such sequences are capable of pairing to form RNA duplexes that significantly enhance back-splicing (FIG. 1b), as demonstrated in genome-wide analyses[7,9] and by circRNA expression vectors[9,16,19]. RNA pairing can occur either at repetitive elements (such as the abundant *Alu* elements in primates)[9,16,19] or from non-repetitive complementary sequences[9]. Sequences as short as 30 to 40 nucleotides are capable of promoting circRNA biogenesis[16], although stronger pairing of longer sequences considerably enhances circRNA production[9]. Back-splicing efficiency is also affected by competition of RNA pairings across the flanking introns or within individual introns (FIG. 1c). Formation of RNA pairing within an individual intron was shown to be associated with canonical splicing, whereas the formation of RNA pairing across flanking introns promoted circRNA production[9]. Moreover, a series of reverse-orientated *Alu* elements in introns could, theoretically, form different *Alu* pairs, leading to the production of different circRNAs from a single gene. This is referred as to alternative circularization[9] (FIG. 1c).

Interestingly, not all circularized exons are bracketed by complementary sequences. For example, circRNA-producing loci in *D. melanogaster* usually lack apparent flanking intronic pairing sequences but have a bias for long flanking introns[11], and only a small proportion of circularized exons are flanked with intronic repeats in *O. sativa*[26]. A circle-forming exon from the *mrps16* (mitochondrial ribosomal protein of the small subunit) gene in *Schizosaccharomyces pombe* lacking flanking complementary sequences is generated through an exon-containing lariat precursor[28]. In addition, circRNAs could also be detected from expression vectors in human cells when the inserted flanking intronic region was as short as 20 nt in length and had no inverted repeats[17]. It remains to be established whether any *cis*-elements in such intronic sequences have a meaningful effect on circRNA formation. Alternatively, short *cis*-regulatory elements in introns that can be recognized by RNA-binding proteins (RBPs) could promote exon circularization[15,18], and splicing regulatory *cis*-elements that are known to function as canonical splicing enhancers or silencers might also affect the efficiency of back-splicing[33], although the combinations of splicing regulatory elements and RBPs that are necessary for splicing and for back-splicing might be different.

Regulation by RNA-binding proteins. In addition to *cis*-elements, RBPs have been reported to regulate circRNA biogenesis[12,15,18,19,31]. Interestingly, in *D. melanogaster*, the splicing factor Muscleblind (Mbl) regulates circRNA production from its own pre-mRNA[15]. Multiple Mbl-binding sites were identified in introns flanking the circularized exon in the *Mbl* pre-mRNA. Mbl overexpression, or the addition of Mbl-binding sites in the introns flanking the circle-forming exon in an expression vector, stimulated exon circularization, suggesting that Mbl can promote circRNA production[15] (FIG. 1d). Another RBP that strikingly augments circRNA formation by pre-mRNA binding is quaking (QKI). Many circRNAs are upregulated by QKI during EMT in humans, suggesting that circRNAs can be purposefully formed in a cell type-specific manner[18]. The insertion of synthetic QKI-binding sites into introns is sufficient to produce circRNAs. As QKI forms dimers, it was proposed that QKI may bind to two flanking introns and bring the circularized exons closer together, resulting in upregulation of circRNA production[18] (FIG. 1e).

Box 1 | The different types of circular RNAs

In addition to the circular RNAs that are formed from thousands of loci in eukaryotes[6–15] by exon back-splicing circularization, different types of circular RNAs exist, which are produced by distinct mechanisms. First, circular viral RNA genomes[48,49] are ligated by host cellular enzymes to form 3′,5′-phosphodiester bonds[50] or 2′,5′-phosphodiester bonds[51]. Second, circular RNA intermediates can be generated during rRNA processing[52,53] and permuted tRNA biogenesis in Archaea and algae[54]. Third, a number of housekeeping non-coding RNAs, including small nucleolar RNAs and the ribozyme RNase P, were also identified in circular forms in Archaea[53]. Finally, various circular RNAs that are produced from spliced introns and exons have been reported. For example, excised, self-spliced group I and group II introns are ligated to form circular RNAs with either 3′,5′-phosphodiester bonds or 2′,5′-phosphodiester bonds[55–58]. Excised tRNA introns are removed by tRNA splicing enzymes and then undergo 3′,5′-phosphodiester ligation in Archaea[59,60] and Metazoa[61]. Circular RNAs derived from lariat introns have been reported in human cells and in *Xenopus tropicalis*[21,42,43].

Box 2 | **Tissue-specific and disease-associated expression of circRNAs**

Circular RNAs formed by exon back-splicing circularization (circRNAs) have a potential role in brain function, as it has been reported that they are preferentially back-spliced from neural genes[15]. Detailed analyses of RNA sequencing data sets from dissected brain tissues, neuronal-differentiated cell lines[31] and different developmental stages of mouse brain[32] indicate that thousands of circRNAs accumulate to high levels in the mammalian brain[31,32]. Many such circRNAs are upregulated during neurogenesis[31] and are significantly more enriched in synaptogenesis than their linear isoforms[32]. In *Drosophila melanogaster*, circRNAs are also upregulated in neural tissues in an age-dependent manner[11]. The expression of some brain-specific circRNAs is conserved between human and mouse and even *D. melanogaster*[31], suggesting that some of these circRNAs may indeed affect neuronal functions. As circRNAs are stable[7,8] and thus can accumulate in non-proliferating cells[62], the slow division rate of neurons may lead to passive accumulation of circRNAs. Hundreds of circRNAs are also regulated during the human epithelial–mesenchymal transition (EMT), indicating that certain circRNAs may affect EMT-related functions such as cell migration, invasion and metastasis[18]. The specific functions of circRNAs in the brain or during EMT remain largely unknown.

Some circRNAs are derived from genomic loci associated with development and diseases, implying that they may function in certain physiological and pathological conditions. For example, *circSRY* (circRNA from sex-determining region Y) has a key role in testes development in mouse[3,20]. The *INK4*/*ARF* tumour suppressor locus-associated long non-coding RNA *ANRIL* (antisense non-coding RNA in the *INK4* locus) participates in transcription repression[63], and the expression of circRNAs from *ANRIL* is positively correlated with *INK4*/*ARF* expression and atherosclerosis risk[24]. *CiRS-7* (circular RNA sponge for miR-7), one of the best known circRNAs, functions as a sponge for miR-7 in neuronal tissues[8,20] and is significantly reduced in sporadic Alzheimer disease[23], indicating that the c*iRS-7*–miR-7 axis is deregulated in individuals with Alzheimer disease. Finally, circRNA levels were found to be significantly altered during cancer-associated cell proliferation[62].

Adenosine deaminase 1 acting on RNA (ADAR1) has been shown to suppress circRNA expression[12,31]. This is associated with the adenosine-to-inosine (A-to-I) editing activity of ADAR1 on duplexed RNAs formed between sequences flanking circularized exons. A-to-I editing diminishes the complementarity and annealing of these RNA pairs and thus reduces back-splicing (FIG. 1f). However, the effect of editing by ADAR1 might be limited, as the circularization from only a few genes was affected by ADAR1 knockdown[31]. It should be noted that ADAR1 was recently shown to function not only as an editing enzyme but also as a double-stranded RNA-binding protein[36]; thus, it is possible that it may regulate circRNA formation independently of its editing activity. As there are hundreds of RBPs[37], an important future challenge will be to identify additional RBPs that are involved in circRNA biogenesis. In this regard, it will also be of interest to study the combinatorial effects of RBPs on circRNA production. A recent study has shown that back-splicing of the *D. melanogaster Laccase 2* gene is regulated by both intronic repeats and multiple hnRNP (heterogeneous nuclear ribonucleoprotein) and SR (Ser/Arg) proteins functioning in a combinatorial manner[19]. Altogether, it would seem that the regulation of circRNA production in cells is substantially more complicated than is currently appreciated.

Gene regulation by circRNAs

The generally low expression levels of circRNAs could imply that they are merely a tolerated by-product of eukaryotic transcriptomes, and by far the majority of circRNAs seem to lack functional significance. However, recent studies have revealed that some circRNAs may have important physiological functions — for example, in the brain (BOX 2) — and that they may regulate gene expression at multiple levels.

CircRNAs can function as miRNA sponges. The large majority of the annotated back-splicing circRNAs are primarily localized in the cytoplasm[6,7]. However, although circRNAs that were experimentally inserted with internal ribosome entry sites could undergo translation[18,33,38], endogenous circRNAs have not yet been reported to associate with ribosomes[7,10]. In fact, the observation that circRNAs could be targeted by RNA interference suggested that the more stable circRNAs may compete with mRNAs for miRNA binding in the cytoplasm and thus regulate gene expression[7]. The best characterized circRNA to support this model is *ciRS-7* (circular RNA sponge for miR-7), which is produced from the vertebrate cerebellar degeneration-related 1 (*CDR1*) antisense transcript (FIG. 2a). *ciRS-7* is preferentially expressed in human and mouse brain[20,39] and contains over 60 conserved miR-7 target sites[8,20]. *ciRS-7* is highly expressed, localized in the cytoplasm and has the capacity to bind to up to 20,000 miR-7 molecules per cell[8], thereby potentially influencing the availability of miR-7 to bind to its target mRNAs. Decreased expression of *ciRS-7* caused the reduced expression of mRNAs containing miR-7 binding sites, further suggesting that *ciRS-7* competes with mRNAs for miR-7 binding[8,20]. Importantly, the function of *ciRS-7* is conserved. Expression of human *ciRS-7* in zebrafish impaired midbrain development, similar to the effect of miR-7 knockdown[8]. Other than *ciRS-7*, only a few circRNAs in mammals are known to function as potential miRNA sponges. The testis-specific circRNA, sex-determining region Y (*circSRY*) contains 16 target sites for miR-138 in mouse[20] (although its human homologue has only one miR-138 target site[10]), and several circRNAs generated from zinc-finger genes contain a considerable number of target sites for miRNAs[10]. Although miRNA target sites in circRNAs need not be conserved to function as miRNA sponges, for the most part, human and mouse circRNAs still seem to infrequently contain miRNA target sites[10,32], indicating that most circRNAs might not function as miRNA sponges.

In contrast to circRNAs in mammals, *D. melanogaster* circRNAs harbour over one thousand well-conserved miRNA seed matches[11], but whether fly circRNAs generally function as miRNA sponges is yet to be determined. As a cautionary note, a recent quantitative study of miRNA function suggests that, in general, the modulation of miRNA target abundance is unlikely to have significant effects on gene expression[40].

Circular RNAs regulate transcription. Although most circRNAs are located in the cytoplasm[6,7], intron-containing circRNAs are more likely to be restricted to the nucleus in human cells[21,22]. How export or nuclear retention of circRNAs is achieved remains unclear. As many linear RNAs containing retained introns are frequently restricted to the nucleus[41], intron-containing circular RNAs may be retained in a similar manner.

Another class of nuclear circular RNAs in the cell is the circular intronic RNAs (ciRNAs), which are derived from lariat introns. To avoid debranching, their formation depends on a consensus

sequence containing a 7 nt GU-rich motif near the 5′ splice site and an 11 nt C-rich motif at the branch point site (FIG. 2b). The resulting RNA circle is covalently ligated through a 2′,5′-phosphodiester bond at the joining site and lacks the linear part stretching from the 3′ end of the intron to the branch point[21]. Hundreds of such ciRNAs were identified in human cells[21]. Stable intronic RNAs derived from lariats were also found in oocytes of *Xenopus tropicalis*[42,43]. In human cells, ciRNAs largely accumulate in the nucleus and regulate gene transcription in *cis*. Some abundantly expressed ciRNAs, such as *ci-ankrd52* and *ci-sirt7*, localize to and interact with the elongating Pol II complex. Depleting these ciRNAs decreased the transcription levels of the corresponding ankyrin repeat domain 52 (*ANKRD52*) or sirtuin 7 (*SIRT7*) genes, suggesting that ciRNAs promote Pol II transcription of their parental genes, although the mechanism of this is unknown[21] (FIG. 2b).

Pol II transcription regulation by nuclear RNA circles was further supported by another study, in which the authors applied RNA immunoprecipitation assays to study Pol II-associated non-coding RNAs[22]. They found exon–intron circRNAs (EIciRNAs), which are back-spliced circRNAs with retained introns, to be associated with Pol II. Knockdown of EIciRNAs such as *EIciEIF3J* and *EIciPAIP2* reduced the transcription of their parent genes. Mechanistically, EIciRNAs interact with the U1 snRNP (small nuclear ribonucleoprotein) and with the promoters of their encoding genes. *EIciEIF3J* and *EIciPAIP2* each contain one potential U1 snRNA-binding site at their retained introns. Blocking these RNA–RNA interactions decreased the interaction of the EIciRNAs with Pol II and also decreased the binding of EIciRNA–U1 snRNP complexes to the promoters of their encoding genes (eukaryotic translation initiation factor 3J (*EIF3J*) and poly(A)-binding protein-interacting protein 2 (*PAIP2*)), which subsequently reduced the transcription of these genes[22] (FIG. 2c).

Together, these studies suggest a role for some nuclear RNA circles in transcription regulation, although the relevant mechanisms of action are still poorly understood. In addition, it remains to be explored whether other loci, in *trans*, are affected by circRNAs. Indeed, ciRNAs are known to localize beyond their sites of synthesis[21]. It will also be interesting to determine whether nuclear-retained circular RNAs have additional roles, such as functioning as RBP decoys. For example, aggregates of TAR DNA-binding protein 43 (TDP43) accumulate in degenerating neurons of most individuals with amyotrophic lateral sclerosis (ALS). In a disease model of ALS, the intronic lariats that accumulate in the absence of the RNA lariat debranching enzyme DBR1 probably function as decoys to sequester TDP43 in the cytoplasm and suppress TDP43 toxicity[44].

The processing of circRNAs affects splicing. CircRNAs are generally derived from middle exons of protein-coding genes[9]. Thus, the processing of circRNAs could affect alternative splicing of such pre-mRNAs, potentially leading to altered gene expression. Although back-splicing is generally less favourable than splicing, it has been shown that the usage of 5′ and 3′ splice sites in circRNA biogenesis can compete with pre-mRNA splicing, resulting in lower levels of linear mRNAs that include the

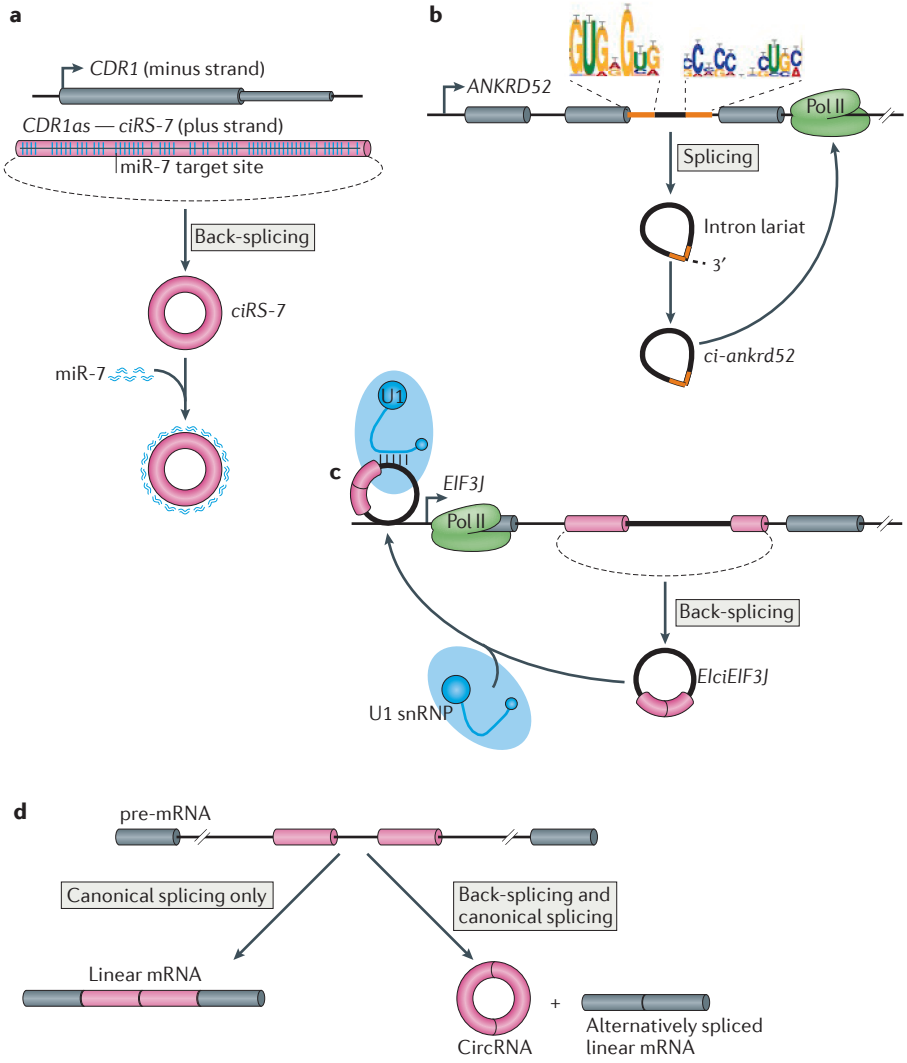

Figure 2 | **Possible roles of circular RNAs (circRNAs) in gene regulation. a** |The circRNA *ciRS-7* (circular RNA sponge for miR-7) is produced by back-splicing of the antisense transcript of cerebellar degeneration-related 1 (*CDR1*), *CDR1as*. *ciRS-7* contains over 60 conserved miR-7 target sites, is localized in the cytoplasm and functions as the miR-7 sponge[8,20]. **b** | The circular intronic RNA (ciRNA) *ci-ankrd52* is derived from the second intron of ankyrin repeat domain 52 (*ANKRD52*) and depends on consensus RNA sequences (orange bars) to avoid debranching of the lariat intron. *ci-ankrd52* mainly accumulates in the nucleus and promotes RNA polymerase II (Pol II) transcription of its encoding gene *ANKRD52* (REF. 21). **c** | *EIciEIF3J* is a back-spliced circRNA with a retained intron produced from eukaryotic translation initiation factor 3J (*EIF3J*). It interacts with the U1 snRNP (small nuclear ribonucleoprotein) and the promoter of *EIF3J* to boost *EIF3J* transcription by Pol II[22]. **d** | The processing of circRNA back-splicing affects alternative splicing of transcripts that produce circRNAs and therefore gene expression[9,14,15]. pre-mRNA, precursor mRNA.

circularized exons[9,14,15] (FIG. 2d). In principle, the more an exon is circularized, the less it will be present in the processed mRNA[14].

Exon skipping is the most common alternative splicing event in human pre-mRNAs and, theoretically, circRNA formation should be positively correlated with exon skipping in the linear mRNA (FIG. 2d). Such positive correlation between exon skipping and the production of circRNAs that contain the skipped exon has been shown in circRNA expression vectors[9] and in a genome-wide analysis of RNA-seq data sets from human endothelial cells stimulated with tumour necrosis factor-α or tumour growth factor-β[14]. However, not all skipped exons could produce circRNAs[14], suggesting that additional regulators may affect exon circularization following exon skipping.

Perspectives

To date, RNA-seq data sets and newly developed bioinformatic approaches have allowed the discovery of over ten thousand different circRNAs in various organisms. Additional studies on circRNAs are being generated at a rapid pace, yet many questions concerning circRNA biogenesis and function remain unanswered. How the spliceosome is involved in back-splicing circularization and how it is directed to perform canonical splicing or back-splicing are yet to be defined precisely. As alternative splicing is prevalent, back-splicing could be regulated by similar mechanisms. Additional *cis*- and *trans*-regulators that are involved in circRNA biogenesis, and the differential expression of circRNAs in various biological settings, remain to be fully explored. An interesting question is how the nucleocytoplasmic export of circRNAs is regulated. An untested hypothesis is that circRNA export is regulated by mechanisms similar to those that regulate mRNA export. Another intriguing question is how circular RNAs are ultimately degraded, as they are stable, probably owing to their circular structure, which is resistant to degradation by RNA exonucleolytic decay machineries.

To what extent the expression and function of circRNAs is conserved also requires a more thorough investigation. Importantly, we know little about what circRNAs do in cells. In addition to their roles in titration of miRNAs[8,20] and transcription regulation[21,22], circRNAs may function as scaffolds for the assembly of macromolecular complexes. In this way, some abundant circRNAs may regulate gene expression by serving as RBP sponges[15,44], whereas others may have unpredicted cellular functions owing to their unusual stability and structure. The differential expression of circRNAs in the brain and in human diseases (BOX 2) suggests that circRNAs may have important physiological roles. The detection of circRNAs in human saliva[45], blood[46] and exosomes[47] holds promise that circRNAs may be used as disease biomarkers.

In-depth understanding of circRNA biogenesis at the molecular level will be crucial to understanding the function of circRNAs. Owing to the fact that circRNAs are expressed with cognate linear transcripts and have unique structural features, I hope to see further improvements in methods to study these RNA circles without affecting their resident genes. Nevertheless, it is becoming clear that the pervasive existence of circRNAs and their potential roles extend the diversity and complexity of eukaryotic transcriptomes.

Ling-Ling Chen is at the State Key Laboratory of Molecular Biology, CAS Center for Excellence in Molecular Cell Science, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 320 Yueyang Road, 200031 Shanghai, China.
linglingchen@sibcb.ac.cn

doi:10.1038/nrm.2015.32
Published online 24 Feb 2016

1. Nigro, J. M. *et al.* Scrambled exons. *Cell* **64**, 607–613 (1991).
2. Cocquerelle, C., Daubersies, P., Majerus, M. A., Kerckaert, J. P. & Bailleul, B. Splicing with inverted order of exons occurs proximal to large introns. *EMBO J.* **11**, 1095–1098 (1992).
3. Capel, B. *et al.* Circular transcripts of the testis-determining gene *Sry* in adult mouse testis. *Cell* **73**, 1019–1030 (1993).
4. Cocquerelle, C., Mascrez, B., Hetuin, D. & Bailleul, B. Mis-splicing yields circular RNA molecules. *FASEB J.* **7**, 155–160 (1993).
5. Pasman, Z., Been, M. D. & Garcia-Blanco, M. A. Exon circularization in mammalian nuclear extracts. *RNA* **2**, 603–610 (1996).
6. Salzman, J., Gawad, C., Wang, P. L., Lacayo, N. & Brown, P. O. Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. *PLoS ONE* **7**, e30733 (2012).
7. Jeck, W. R. *et al.* Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA* **19**, 141–157 (2013).
8. Memczak, S. *et al.* Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* **495**, 333–338 (2013).
9. Zhang, X. O. *et al.* Complementary sequence-mediated exon circularization. *Cell* **159**, 134–147 (2014).
10. Guo, J. U., Agarwal, V., Guo, H. & Bartel, D. P. Expanded identification and characterization of mammalian circular RNAs. *Genome Biol.* **15**, 409 (2014).
11. Westholm, J. O. *et al.* Genome-wide analysis of *Drosophila* circular RNAs reveals their structural and sequence properties and age-dependent neural accumulation. *Cell Rep.* **9**, 1966–1980 (2014).
12. Ivanov, A. *et al.* Analysis of intron sequences reveals hallmarks of circular RNA biogenesis in animals. *Cell Rep.* **10**, 170–177 (2015).
13. Fan, X. *et al.* Single-cell RNA-seq transcriptome analysis of linear and circular RNAs in mouse preimplantation embryos. *Genome Biol.* **16**, 148 (2015).
14. Kelly, S., Greenman, C., Cook, P. R. & Papantonis, A. Exon skipping is correlated with exon circularization. *J. Mol. Biol.* **427**, 2414–2417 (2015).
15. Ashwal-Fluss, R. *et al.* circRNA biogenesis competes with pre-mRNA splicing. *Mol. Cell* **56**, 55–66 (2014).
16. Liang, D. & Wilusz, J. E. Short intronic repeat sequences facilitate circular RNA production. *Genes Dev.* **28**, 2233–2247 (2014).
17. Starke, S. *et al.* Exon circularization requires canonical splice signals. *Cell Rep.* **10**, 103–111 (2015).
18. Conn, S. J. *et al.* The RNA binding protein quaking regulates formation of circRNAs. *Cell* **160**, 1125–1134 (2015).
19. Kramer, M. C. *et al.* Combinatorial control of *Drosophila* circular RNA expression by intronic repeats, hnRNPs, and SR proteins. *Genes Dev.* **29**, 2168–2182 (2015).
20. Hansen, T. B. *et al.* Natural RNA circles function as efficient microRNA sponges. *Nature* **495**, 384–388 (2013).
21. Zhang, Y. *et al.* Circular intronic long noncoding RNAs. *Mol. Cell* **51**, 792–806 (2013).
22. Li, Z. *et al.* Exon-intron circular RNAs regulate transcription in the nucleus. *Nat. Struct. Mol. Biol.* **22**, 256–264 (2015).
23. Lukiw, W. J. Circular RNA (circRNA) in Alzheimer's disease (AD). *Front. Genet.* **4**, 307 (2013).
24. Burd, C. E. *et al.* Expression of linear and novel circular forms of an INK4/ARF-associated non-coding RNA correlates with atherosclerosis risk. *PLoS Genet.* **6**, e1001233 (2010).
25. Yang, L., Duff, M. O., Graveley, B. R., Carmichael, G. G. & Chen, L. L. Genome-wide characterization of non-polyadenylated RNAs. *Genome Biol.* **12**, R16 (2011).
26. Lu, T. *et al.* Transcriptome-wide investigation of circular RNAs in rice. *RNA* **21**, 2076–2087 (2015).
27. Wang, P. L. *et al.* Circular RNA is expressed across the eukaryotic tree of life. *PLoS ONE* **9**, e90859 (2014).
28. Barrett, S. P., Wang, P. L. & Salzman, J. Circular RNA biogenesis can proceed through an exon-containing lariat precursor. *eLife* **4**, e07540 (2015).
29. Broadbent, K. M. *et al.* Strand-specific RNA sequencing in *Plasmodium falciparum* malaria identifies developmentally regulated long non-coding RNA and circular RNA. *BMC Genomics* **16**, 454 (2015).
30. Salzman, J., Chen, R. E., Olsen, M. N., Wang, P. L. & Brown, P. O. Cell-type specific features of circular RNA expression. *PLoS Genet.* **9**, e1003777 (2013).
31. Rybak-Wolf, A. *et al.* Circular RNAs in the mammalian brain are highly abundant, conserved, and dynamically expressed. *Mol. Cell* **58**, 870–885 (2015).
32. You, X. *et al.* Neural circular RNAs are derived from synaptic genes and regulated by development and plasticity. *Nat. Neurosci.* **18**, 603–610 (2015).
33. Wang, Y. & Wang, Z. Efficient backsplicing produces translatable circular mRNAs. *RNA* **21**, 172–179 (2015).
34. Bentley, D. L. Coupling mRNA processing with transcription in time and space. *Nat. Rev. Genet.* **15**, 163–175 (2014).
35. Szabo, L. *et al.* Statistically based splicing detection reveals neural enrichment and tissue-specific induction of circular RNA during human fetal development. *Genome Biol.* **16**, 126 (2015).
36. Chen, T. *et al.* ADAR1 is required for differentiation and neural induction by regulating microRNA processing in a catalytically independent manner. *Cell Res.* **25**, 459–476 (2015).
37. Castello, A. *et al.* Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* **149**, 1393–1406 (2012).
38. Chen, C. Y. & Sarnow, P. Initiation of protein synthesis by the eukaryotic translational apparatus on circular RNAs. *Science* **268**, 415–417 (1995).
39. Hansen, T. B. *et al.* miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA. *EMBO J.* **30**, 4414–4422 (2011).
40. Denzler, R., Agarwal, V., Stefano, J., Bartel, D. P. & Stoffel, M. Assessing the ceRNA hypothesis with quantitative measurements of miRNA and target abundance. *Mol. Cell* **54**, 766–776 (2014).
41. Braunschweig, U. *et al.* Widespread intron retention in mammals functionally tunes transcriptomes. *Genome Res.* **24**, 1774–1786 (2014).
42. Gardner, E. J., Nizami, Z. F., Talbot, C. C. Jr & Gall, J. G. Stable intronic sequence RNA (sisRNA), a new class of noncoding RNA from the oocyte nucleus of *Xenopus tropicalis*. *Genes Dev.* **26**, 2550–2559 (2012).

43. Talhouarne, G. J. & Gall, J. G. Lariat intronic RNAs in the cytoplasm of *Xenopus tropicalis* oocytes. *RNA* **20**, 1476–1487 (2014).
44. Armakola, M. *et al.* Inhibition of RNA lariat debranching enzyme suppresses TDP-43 toxicity in ALS disease models. *Nat. Genet.* **44**, 1302–1309 (2012).
45. Bahn, J. H. *et al.* The landscape of microRNA, Piwi-interacting RNA, and circular RNA in human saliva. *Clin. Chem.* **61**, 221–230 (2015).
46. Memczak, S., Papavasileiou, P., Peters, O. & Rajewsky, N. Identification and characterization of circular RNAs as a new class of putative biomarkers in human blood. *PLoS ONE* **10**, e0141214 (2015).
47. Li, Y. *et al.* Circular RNA is enriched and stable in exosomes: a promising biomarker for cancer diagnosis. *Cell Res.* **25**, 981–984 (2015).
48. Sanger, H. L., Klotz, G., Riesner, D., Gross, H. J. & Kleinschmidt, A. K. Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures. *Proc. Natl Acad. Sci. USA* **73**, 3852–3856 (1976).
49. Kos, A., Dijkema, R., Arnberg, A. C., van der Meide, P. H. & Schellekens, H. The hepatitis delta (δ) virus possesses a circular RNA. *Nature* **323**, 558–560 (1986).
50. Flores, R. *et al.* Rolling-circle replication of viroids, viroid-like satellite RNAs and hepatitis delta virus: variations on a theme. *RNA Biol.* **8**, 200–206 (2011).
51. Cote, F. & Perreault, J. P. Peach latent mosaic viroid is locked by a 2′,5′-phosphodiester bond produced by *in vitro* self-ligation. *J. Mol. Biol.* **273**, 533–543 (1997).
52. Tang, T. H. *et al.* RNomics in Archaea reveals a further link between splicing of archaeal introns and rRNA processing. *Nucleic Acids Res.* **30**, 921–930 (2002).
53. Danan, M., Schwartz, S., Edelheit, S. & Sorek, R. Transcriptome-wide discovery of circular RNAs in Archaea. *Nucleic Acids Res.* **40**, 3131–3142 (2012).
54. Soma, A. *et al.* Permuted tRNA genes expressed via a circular RNA intermediate in *Cyanidioschyzon merolae*. *Science* **318**, 450–453 (2007).
55. Grabowski, P. J., Zaug, A. J. & Cech, T. R. The intervening sequence of the ribosomal RNA precursor is converted to a circular RNA in isolated nuclei of *Tetrahymena*. *Cell* **23**, 467–476 (1981).
56. Cech, T. R. Self-splicing of group I introns. *Annu. Rev. Biochem.* **59**, 543–568 (1990).
57. Li-Pook-Than, J. & Bonen, L. Multiple physical forms of excised group II intron RNAs in wheat mitochondria. *Nucleic Acids Res.* **34**, 2782–2790 (2006).
58. Molina-Sanchez, M. D., Martinez-Abarca, F. & Toro, N. Excision of the *Sinorhizobium meliloti* group II intron RmInt1 as circles *in vivo*. *J. Biol. Chem.* **281**, 28737–28744 (2006).
59. Abelson, J., Trotta, C. R. & Li, H. tRNA splicing. *J. Biol. Chem.* **273**, 12685–12688 (1998).
60. Salgia, S. R., Singh, S. K., Gurha, P. & Gupta, R. Two reactions of *Haloferax volcanii* RNA splicing enzymes: joining of exons and circularization of introns. *RNA* **9**, 319–330 (2003).
61. Lu, Z. *et al.* Metazoan tRNA introns generate stable circular RNAs *in vivo*. *RNA* **21**, 1554–1565 (2015).
62. Bachmayr-Heyda, A. *et al.* Correlation of circular RNA abundance with proliferation—exemplified with colorectal and ovarian cancer, idiopathic lung fibrosis, and normal human tissues. *Sci. Rep.* **5**, 8057 (2015).
63. Yap, K. L. *et al.* Molecular interplay of the noncoding RNA *ANRIL* and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of *INK4a*. *Mol. Cell* **38**, 662–674 (2010).

Acknowledgements
The author is grateful to Li Yang for discussions. This work was supported by grants 91440202 and 31322018 from the Natural Science Foundation of China (NSFC), and XDA01010206 from the Chinese Academy of Sciences (CAS).

Competing interests statement
The author declares no competing interests.