

## Circular RNAs remain peculiarly unclear in biogenesis and function

CHEN Liang & SHAN Ge\*

*School of Life Sciences & CAS Key Laboratory of Brain Function and Disease, University of Science and Technology of China, Hefei 230027, China*

Received April 4, 2015; accepted April 10, 2015; published online April 21, 2015

**Citation:** Chen L, Shan G. Circular RNAs remain peculiarly unclear in biogenesis and function. *Sci China Life Sci*, 2015, 58: 616–618, doi: 10.1007/s11427-015-4855-y

The known complexity of eukaryotic transcriptomes has been greatly expanded due to the identification of a large family of circular RNAs in recent years [1–3]. As a special form of RNA compared to the “common” linear form, these molecules were first described in 1976, when it was discovered that several plant viroids were single-stranded and covalently closed circular RNAs [4]. In a broad sense, the “circular” form RNAs may include viroids, genome of some RNA viruses, the loop portion of intronic lariats, as well as circular RNAs formed from back-splicing of exons. Currently in the field and also for this review, the term of circular RNA (circRNA) refers to circular molecules from back-splicing of exons unless specified.

Thousands of circRNAs are now found in human cells and in model organisms such as mice, *Drosophila*, and *C. elegans*. A question coming out immediately is how these molecules are generated in cells. Biogenesis of certain circRNAs had long been linked to the relatively long introns and the complementary repeat sequences flanking the region of circularization [5–7]. Genome-wide bioinformatic analysis actually confirmed the association of these features to circRNA biogenesis [5–7]. The exonic sequences within the circRNAs might also be important in some cases [6,7]. At current stage, these associations are linked to the biogenesis of most but not all circRNAs, and are not connected with any actual molecular mechanism. There are also disputes about whether circRNAs are generated

co-transcriptionally or post-transcriptionally [5,6]. A recent publication showed that the RNA-binding protein quaking 1 5 (QKI5) might participate in the circularization [8], although we have to keep in mind with several points before considering QKI5 as an essential player in circRNA biogenesis. First, QKI5 is a known regulatory factor of alternative splicing for selective mRNAs. Presumably many factors participating in pre-mRNA splicing or alternative splicing may also get involved in back-splicing for the formation of circRNAs “non-specifically”. One would wonder whether there is any protein “specific” for circRNA biogenesis. Secondly, circRNAs are wide-spread transcripts found essentially in all cell types examined. We would anticipate that essential factors for circRNA biogenesis are somehow ubiquitously expressed, but QKI5 is not expressed in all cells. Thirdly, flanking complementary sequences are associated with more than half of the circRNAs, and QKI5 has not been tested for biogenesis of circRNAs with flanking complementary repeats [8]. Some circRNAs showed cell type or tissue specific expression between two cells under comparison, even when their parent genes might be expressed (as mRNA) in both cells [7]. The ratio of circRNAs to mRNAs for individual genes was also dynamic among different cells under different conditions. These indicate strongly that biogenesis of circRNAs must be regulated with gene and cell type specificity, and also argue against the idea that circRNAs may be unavoidable by-products of splicing, which is an echoing point of view from early days to even today [9].

\*Corresponding author (email: shange@ustc.edu.cn)

If circRNAs are not “noise” of gene expression, they should possess crucial physiological functions. Some associations with potential regulatory roles in terms of gene expression and human diseases were previously indicated, and then the first demonstrations that two circRNAs could function as microRNA sponge came out in 2013 [2]. It was found that the circRNA CDR1as harbors multiple conserved binding sites for miR-7, and another circRNA circSry might be a miR-138 sponge. Both circRNAs localized predominantly in the cytoplasm, in consistence with their functional roles. Actually most circRNAs reside in the cytoplasm, which promotes the idea that many (if not most) circRNAs may be microRNA sponge. This point is challenged by the fact that only several out of thousands of circRNAs possess multiple computer-predicted binding sites for individual microRNAs [9]. It is also unlikely for most circRNAs to function as template for translation, as few circRNAs are associated with polyribosomes [9]. In this sense, circRNAs are most likely a large species of long noncoding RNAs, although generally the term of long noncoding RNAs (lncRNAs) refers to linear noncoding RNAs longer than 200 nt.

The recent findings of our lab have drawn people’s attention from cytoplasm to the nucleus. When investigating noncoding RNAs regulating pol II transcription, we recovered a special subclass of circRNAs [7]. Unlike most of the known circRNAs, these circRNAs predominantly localize in the nucleus. Further analysis showed that these circRNAs were also special in their sequences. The intronic sequences between the circularized exons were retained; for this reason, we termed these RNAs as exon-intron circRNA or EIciRNA. EIciRNAs, for example, EIciEIF3J and EIciPAIP2 colocalized with their parental genomic loci and the U1 snRNP. EIciRNAs interact with U1 snRNA and promote transcription of their parental genes *in cis*. These findings reveal a role for circRNAs in regulating gene expression in the nucleus and underline a regulatory strategy via specific RNA-RNA interaction between U1 snRNA and EIciRNAs. EIciRNAs are only a special subtype of the vast number of circRNAs, and it remains an open question about the physiological roles of all the other circRNAs.

Having summarized recent advances in circRNA biology, we would like to share our insights about the biogenesis and functions of circRNAs. The fact that flanking long introns and complementary sequences are associated with most backsplicing exons points to two features presumably crucial for circRNA biogenesis. First, the splicing across the flanking regions of circularization is less efficient due to the presence of long introns. Second, the two splicing sites involved in the backsplicing would be brought together by the flanking complementary sequences or other ways. Presumably protein dimerization or protein-protein interaction across the upstream and downstream sequences can also help to bring the two backsplicing sites together, which may be the actual mechanism of QKI5 in promoting circulariza-

tion [8].

We currently have no clue about how biogenesis of circRNAs is regulated for specificity in cell type and cell condition. The relative amount of circRNAs and the ratio of circRNA to mRNA from individual genes have to be regulated. These would be interesting aspects waiting for investigation. The intron retention in EIciRNAs against the intron exclusion in the other circRNAs also has to be regulated. These concerns may also relate to another unexplored subject: how circRNAs are degraded? It is the balance between biogenesis and degradation that decides the expression profile of circRNAs. Presumably circRNAs are more stable due to their inaccessibility to exonucleases. How circRNAs are degraded, is this process regulated (if yes, then how), and could degradation just like biogenesis be regulated specifically for individual circRNAs?

What are the physiological functions of circRNAs? To answer this question, we can look at the current status of linear lncRNAs. lncRNAs had been viewed by some as transcriptional noise back to about a decade ago, but are now recognized as major regulatory components with a myriad of mechanisms in eukaryotic cells. Currently we only know that a couple of circRNAs function as microRNA sponge in the cytoplasm and a special subclass of circRNAs as EIciRNA function as transcription regulator in the nucleus. Indications about the association between circRNAs and certain human diseases are reported, and we do not have any clue about the actual roles of circRNAs in etiology.

Cellular localization is often linked with physiological function and functional mechanism for biological molecules. It is proposed that circRNAs may be “localized” passively to cytoplasm due to the breakdown of nuclear membrane during mitosis, although this cannot explain the cytoplasmic localization of most circRNAs in the non-dividing cells such as neurons [10]. The phenomena that some circRNAs are cytoplasmic (e.g., CDR1as and circSry) and some are nuclear (e.g., EIciRNA) suggest strongly the existence of an active “transporting” or “sorting” mechanism for circRNAs. Interestingly, punctae of circRNAs are found in the neuronal dendrites [10], which is another indication for the involvement of special mechanism for circRNA subcellular localization.

Despite dozens of papers published about circRNAs in recent years, we are still at the dawn of this field. In particular, little is known about the molecular mechanism of their biogenesis, degradation, and cellular localization. The functions and functional mechanisms of most circRNAs are still unknown. One thing we need to keep in mind from what has already been known about circRNAs is that circRNAs may be a large family with great diversity in terms of biogenesis, degradation, cellular localization, function, and functional mechanism. Here we may make an analogy between circRNAs and linear lncRNAs, which show tremendous diversity. The field may need to start with revealing subclasses of circRNAs and further exploring each specific

subclass before we can eventually pinpoint the full picture of circRNA biology.

- 1 Salzman J, Gawad C, Wang PL, Lacayo N, Brown PO. Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. *PLoS One*, 2012, 7: e30733
- 2 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
- 3 Wang PL, Bao Y, Yee MC, Barrett SP, Hogan GJ, Olsen MN, Dinnyen JR, Brown PO, Salzman J. Circular RNA is expressed across the eukaryotic tree of life. *PLoS One*, 2014, 9: e90859
- 4 Sanger HL, Klotz G, Riesner D, Gross HJ, Kleinschmidt AK. Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures. *Proc Natl Acad Sci USA*, 1976, 73: 3852–3856
- 5 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. *Mol Cell*, 2014, 56: 55–66
- 6 Liang D, Wilusz JE. Short intronic repeat sequences facilitate circular RNA production. *Genes Dev*, 2014, 28: 2233–2247
- 7 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
- 8 Conn SJ, Pillman KA, Toubia J, Conn VM, Salmandis M, Phillips CA, Roslan S, Schreiber AW, Gregory PA, Goodall GJ. The RNA binding protein quaking regulates formation of circRNAs. *Cell*, 2015, 160: 1125–1134
- 9 Guo JU, Agarwal V, Guo H, Bartel DP. Expanded identification and characterization of mammalian circular RNAs. *Genome Biol*, 2014, 15: 409
- 10 You X, Vlatkovic I, Babic A, Will T, Epstein I, Tushev G, Akbalik G, Wang M, Glock C, Quedenau C, Wang X, Hou J, Liu H, Sun W, Sambandan S, Chen T, Schuman EM, Chen W. Neural circular RNAs are derived from synaptic genes and regulated by development and plasticity. *Nat Neurosci*, 2015, 18: 603–610

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.