BCAS Vol.28 No.4 2014

Circular RNA Signals More Complexity and Diversity in Gene Expression Regulation

By SONG Jianlan (Staff Reporter)

It is well known that when the genes of a cell are being expressed, the DNA chains unwind and transcribe their genetic information to RNA chains. One might take it for granted that the RNA chains are linear in shape, however, according to the latest results reported as an article in *Cell* on Sept. 25th by scientists at the Shanghai Institutes for Biological Sciences (SIBS), CAS, circular RNAs universally exist in eukaryotic cells and moreover, the number of them in a cell could be comparable to that of normal linear ones. This surprising discovery has stirred curiosity about possible connections between such circular RNAs and diseases.

Circular RNAs Abundant

Led by Dr. YANG Li from the CAS-MPG Partner Institute for Computational Biology (PICB) and Dr. CHEN Lingling from the Shanghai Institute of Biochemistry and Cell Biology (SIBCB) under SIBS, the joint research team successfully concentrated such circular RNAs via a special nuclease called RNase R, and identified nearly 10,000 circular RNAs in H9 human embryonic stem cells with a newly established computational pipeline. For the first time they demonstrated the detailed mechanism of exon circular RNA biogenesis and articulated a theory on the generality of "alternative circularization", indicating the gene expression regulation in mammalian cells could be more complicated than previously thought at the posttranscriptional stage.

Their work was highlighted by a "Leading Edge Preview" titled *Biogenesis of Circular RNAs* in the same issue of *Cell* by Quentin Vicens and Eric Westhof, in which the authors stated that "Zhang *et al*. precisely pinpoint the sequence requirements in the flanking introns that promote exon circularization" and that "This result is particularly significant, as it suggests a new critical role for *Alu* elements" (a type of repeated complementary sequences widely found in human genome).

In their earlier work published in Sept. 2013, CHEN and YANG's labs reported in *Molecular Cell* the systematic identification of circular RNAs originated from intronic sequences, and unveiled for the first time their biogenetic mechanism as well as molecular functioning. Through this work they also established a research system for molecular prediction as well as studies on the production, processing, and functions of circular RNAs.

Naughty RNAs

On a DNA chain in eukaryotic cells, introduces YANG, there could exist exons and introns. The former contain effective genetic information for protein coding while the latter not. After transcription, the derived RNA needs to get rid of the redundant sequences derived from introns through splicing and connect the sequences produced from exons together in a certain order to form a mature linear RNA chain for protein coding. This process, however, at higher eykaryotes could produce multiple mature RNAs originated from the same father RNA or, literally the same DNA. Biologists call this alternative splicing. "It is just like different pathways available from one place to another, say, from our institute to *Tianzifang* (a popular shopping center and place of leisure in Shanghai), given that multiple intersections exist in the between," explains YANG, using a metaphor.

"In higher eukaryotic organisms," continues YANG, "about 70% genes can produce multiple RNA products via alternative splicing. While in *homo sapiens*, this ratio could mount to 90%, and this is closely associated to normal life activity."

The existence of multiple solutions to RNA splicing greatly increases the complexity and diversity in gene expression. As the produced mature RNAs directly work as the "moulds" for casting of proteins, improper splicing of

RNAs could cause numerous important diseases, including spinal muscular atrophy, myotonic dystrophy, Alzheimer's disease and cancers.

Interestingly, recent research discovered that sometimes circular RNAs could be formed through "back splicing", where the RNA string connects nose to end to form a ring, but their detailed biogenesis has remained unclear since they were discovered in 1976. Notably, in 2013 CHEN and YANG reported in *Molecular Cell* another new type of circular RNAs derived from introns. These new circular RNAs escaped from debranching, as a result from the existence of consensus RNA elements near the right site and branchpoint for proper processing.

Now in *Cell* they further revealed in detail the molecular basis for exon circular RNA biogenesis, attributing back splicing in this case of circularization to the competition between two alternative ways of RNA pairing.

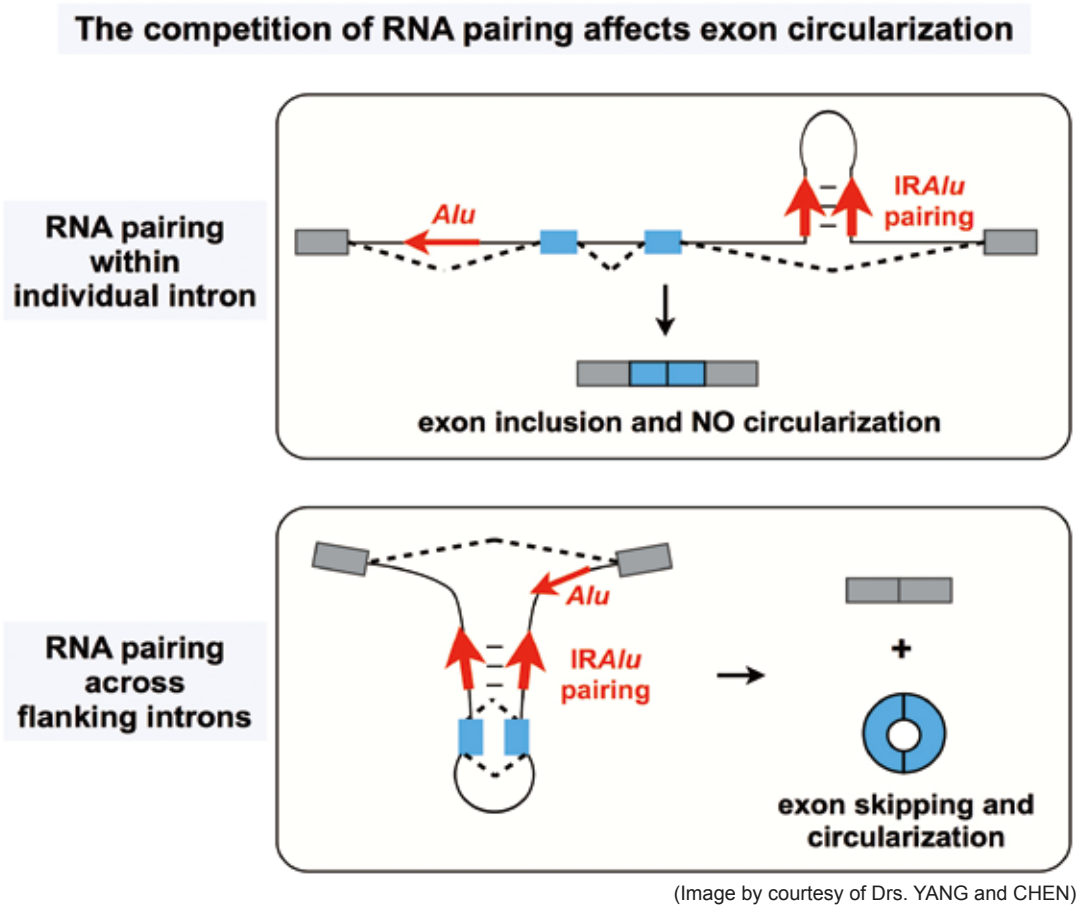
Alternative Circularization

According to the researchers, the circularization of exon RNAs is mediated by pairing of complementary sequences. The term "complementary sequences" refers to two sequences, when antiparallelly aligned, can match up to form a pair. This complementarity could occur within the same strand of RNA, resulting in interesting splicing behavior of the string: When the complementary sequences lock onto each other to form pairs, the RNA strand will splice in different ways, depending on the location of the pairing relative to the introns, either within individual introns or across introns flanking circularized exons.

"Such pairing, when occurs within the same intron, will result in a linear RNA;" explains YANG: "Whereas when it occurs across flanking introns, a circular RNA will be formed instead," he adds.

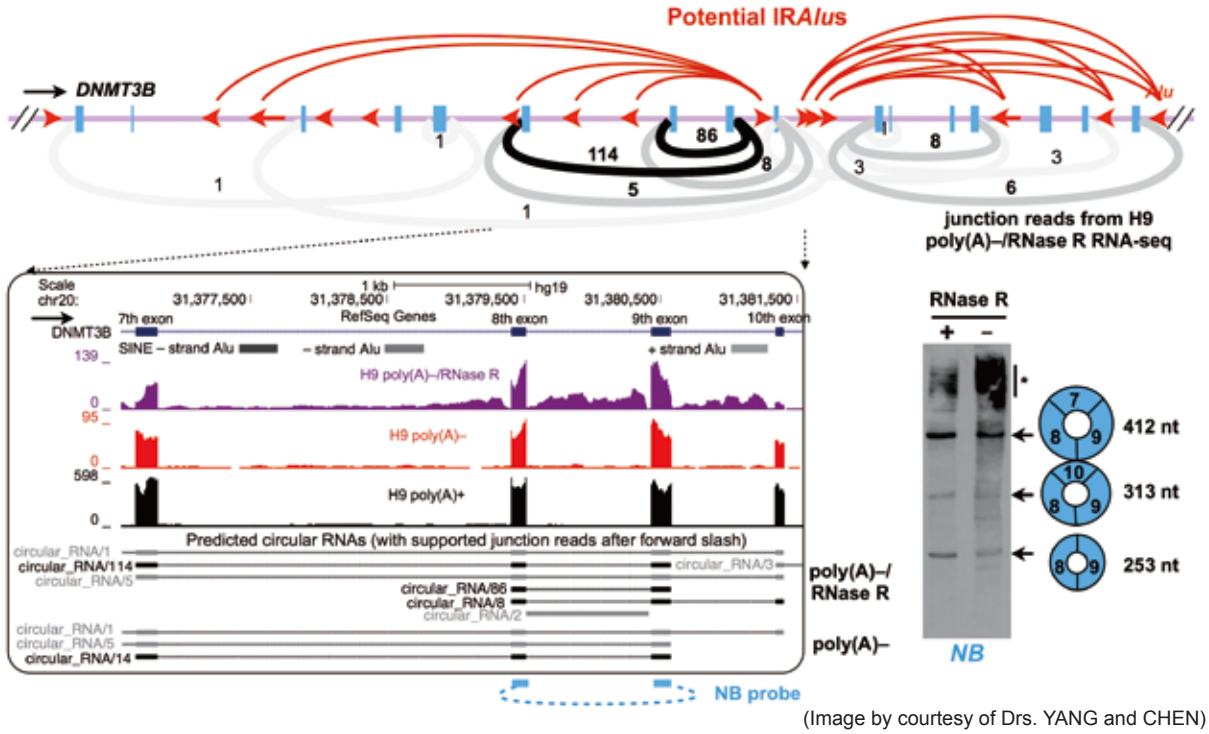
Based on these findings the joint team brought forward a theory called "alternative circularization". The competition between alternative ways of RNA pairing by complementary sequences, results in alternative exon circularization of RNA, producing multiple circular RNA transcripts derived from the same gene. The researchers named this phenomenon "alternative circularization".

Notably, YANG and CHEN say, internal complementarity frequently occurs in the intronic areas of human genome. For example, the *Alu* pairs (IR*Alu*s, inverted repeated *Alu* elements) are among such naughty sequences. The competing pairing of these many



(Image by courtesy of Drs. YANG and CHEN)

Alternative formation of inverted repeated Alu pairs and the competition between them can lead to alternative circularization, resulting in multiple circular RNA transcripts produced from a single gene



(Image by courtesy of Drs. YANG and CHEN)

complementary sequences and its dynamic regulation hence produce multiple circular RNAs from the same gene, making the consequent expression of the gene more complicated.

The circularization of RNA due to special distribution of exons on the RNA chain had been identified from many loci in mammals before the research of interest, but this was the first time the detailed mechanism of its biogenesis was revealed. YANG, CHEN and colleagues also discovered that the competition between alternative ways of pairing exhibit different patterns of combination across species, and this makes the expression of exonic circular RNAs evolutionarily dynamic.

Fog Rises

The discovery of alternative circularization is said to have further expanded the understanding of gene expression regulation. Through alternative splicing, multiple functional mRNAs (and proteins) could be produced from a single gene, and these multiple functional mRNAs were generally thought to exist only as linearized molecules. Now we know that alternative circularization coupled with alternative splicing can produce a variety of additional circular RNAs from one single gene. What could these extra circular RNAs do to the cell and globally the organism if they can also be translated into proteins? Could they lead to diseases?

This is still open to further investigation, answered CHEN, who introduces that whether these circular RNAs could be translated into proteins is still unknown. It would prove to be an exciting area to explore with promising possibility. "Life is marvelous," she comments. Obviously their work has shown that the seemingly "useless" introns are much more interesting than previously thought, as they actually increase the complexity in mammalian posttranscriptional regulation. What seems to be even more intriguing is, higher organisms, like mammals, tend to have relatively more such redundant sequences than the lower, like nematodes. Why do mammals need introns at all in their DNAs if such "invalid" sequences could get them in trouble? The answer to this question is also a big "unknown" – a far-reaching one maybe.

For more information please refer to:

Zhang XO, Wang HB, Zhang Y, Lu XH, **Chen LL**, **Yang L**., (2014) Complementary Sequence-Mediated Exon Circularization, *Cell*, 159, 134–147.
Zhang Y, Zhang XO, Chen T, Xiang JF, Yin QF, Xing YH, Zhu SS, **Yang L**, **Chen LL.**, (2013) Circular Intronic Long Noncoding RNAs, *Mol Cell*, 51:792–806.