Emerging Roles of Plant Circular RNAs

Qinjie Chu[1], Enhui Shen[1], Chu-Yu Ye[1], Longjiang Fan[1], Qian-Hao Zhu[2,*]

[1] Institute of Crop Sciences and Institute of Bioinformatics, College of Agriculture and Biotechnology, Zhejiang University, Hangzhou 310058, China

[2] CSIRO Agriculture and Food, GPO Box 1700, Canberra, ACT 2601, Australia

Abstract

Circular RNAs (circRNAs) are covalently closed single-stranded loop RNA molecules with or without protein coding capability. CircRNAs were previously considered to be splicing intermediates or artifacts but are now found to be pervasively expressed in all eukaryotes studied with some demonstrated to have important molecular functions in various biological processes. CircRNA is now a hot study topic of molecular biology. In this review, we summarize the progress achieved so far on plant circRNAs, including identification and functional characterization, compare the similarities and differences of circRNAs between plants and animals, and discuss the challenges for confident detection and functional investigation of plant circRNAs. Similar to what have been found in animals, plant genomes contain a large number of circRNAs that potentially regulate a wide range of biological progresses related to plant development and biotic/abiotic responses. Despite only a few plant circRNAs have been functionally characterized, novel function/mechanism that has not been reported in animals was revealed, implying more exciting findings about plant circRNAs are expected in future studies.

Introduction

Cirular RNAs (circRNAs) are covalently closed circular single-stranded RNA molecules without free ends. The first identified circRNA was the potato spindle tuber viroid [1]. Eukaryotic genes are usually interrupted by intronic sequences, which are spliced out so that exons can be ligated to each other in the 5' to 3' order for protein translation, therefore, it came as a surprise when circular transcripts formed by ligation of exons in a non-collinear manner were later detected in eukaryotic cells [2-7]. During the last 40 years of time, circRNAs were often considered as by-products of splicing or aberrantly spliced products. However, this view on circRNAs has been challenged since the detection of a large number of such "unusual" transcripts, first in human [8] and later in all other eukaryotes studied [9,10]. With the increased number of circRNAs that were functionally characterized, it becomes clear for the biological importance of circRNAs. Until now, our knowledge on circRNAs were mainly derived from studies in animals, but we start to witness dramatically increased investigations of plant circRNAs, including functional studies. Here, we summarize the findings learned in the investigations of plant circRNAs with a focus on the unique features of plant circRNAs, discuss the challenges and questions to be addressed as we move forward with the research area of plant circRNAs.

Molecular Basis of Circular RNA Biogenesis

Based on their genomic origin, circRNAs are categorized into three types: exonic, intronic and intergenic. They have independent modes of biogenesis, but all related to the splicing mechanism of eukaryotes. Typical eukaryotic genes contain exons and introns that are spliced out from primary messenger RNAs (mRNAs) by the spliceosome machinery. A conventional collinear splicing event is fulfilled in two sequential transesterification reactions. First, the 2'-hydroxyl end of the adenosine at the intron branch-point attacks the 5' nucleotide (i.e. the 5' splice site) of the excised intron to form a lariat by a 2'-5' phosphodiester bond. This reaction generates a free 3'-hydroxyl end on the 5' exon, which then attacks the first nucleotide immediately after the excised intron (i.e. the 3' splice site) to join together two exons in a 5' to 3' orientation and release the intronic lariat consisting of a loop and a 3' tail. The intronic lariat without the 3' tail is a perfect circular RNA, i.e. intronic circRNA. Exonic circRNA is generated by a backsplice event, in which the free 3'-hydroxyl end of an exon (i.e. "tail") released at the first reaction of a conventional collinear splicing event joins to the 5' "head" of an exon that is normally upstream to form a "tail to head" RNA structure, i.e. 5' exon is shuffled downstream of 3' exon. When the 3'-tail and the 5'-head are from the same exon, a single-exon circRNA is generated, usually by a cotranscriptional mechanism. When the 5'-head is from another upstream exon, an intermediate lariat containing exons and intron(s) is formed. In this case, a circRNA with multiple exons can be produced after removing the intron(s) posttranscriptionally by the conventional, i.e. linear, splicing mechanism. Most backsplicing events are *cis*-ones that occur within the same or different exons of a single gene, but *trans*-backsplicing events are also quite common. For instance, we found that ~13% and ~34% of the identified circRNAs in *Arabidopsis thaliana* and rice (*Oryza sativa*), respectively, are potential products of *trans*-backsplicing events [11]. Involvement of spliceosome in generation of circRNAs suggests that 1) circRNAs should contain canonical splice signals, and 2) circRNA production competes with linear pre-mRNA splicing. Studies have shown that exon circularization requires canonical splice signals [12], and that majority (70%) of circularization events use canonical splice signals in animals and humans [13]. However, we found that the vast majority of rice circRNAs (92.7%) were flanked by a diverse set of non-GT/AG canonical splicing signals, suggesting that plants may use alternative mechanism(s) for circRNA biogenesis [14]. It has been reported that production of circRNAs has a positive or negative effect on expression of their parental gene [15,16]. A recent study reported that circRNA production is enhanced when core spliceosomal components are depleted and/or canonical linear mRNA splicing events are inhibited [17]. These observations suggest that the relationship between circRNAs and their parental genes is quite complicated.

Characteristics of Plant Circular RNAs

Several characteristics of animal circRNAs, such as longer flanking introns, low expression levels and tissue-specific expression profiles, are shared by plant

circRNAs [11,15,18-21], but plant circRNAs have their own unique features. A number of *cis*-elements and *trans*-factors, particularly the unique features and structures and specific binding sites of splicing factor of the flanking intronic sequences of circRNAs, have been shown to be important for regulation of biogenesis of circRNAs in animals and humans [15,22,23]. Several transcriptome-wide analyses have shown that repetitive and reverse complementary sequences, for example *Alu* repeats, are enriched in the flanking introns of circularized exons in animals and significantly correlated with exon circularization [24-26]. Compared to animals, plant exonic circRNAs are less likely to be generated from exon(s) flanked by introns containing repetitive and/or reverse complementary sequences [11,19,21,27]. For instance, of the 13 validated plant circRNAs, only two (~15%) contain >15-bp reverse complementary sequences in their flanking introns [11]. In cotton (*Gossypium spp.*), despite circRNAs seem to have more repeat sequences in their flanking introns than linear genes, only ~10% of exonic circRNAs are associated with reverse complementary intronic sequences [28]. Nevertheless, short (4-11 bp) complementary sequences found near the splice sites might be important for circRNA production in Arabidopsis and some *trans*-spliced circRNAs found in chloroplast, an organelle enriched for circRNAs, have longer complementary sequences (>20 bp) near the splice sites [27]. A recent study in maize (*Zea mays*) found that LLERCPs (reverse complementary pairs of LINE1-like elements) are significantly enriched in the 35-kb, particularly in the 5-kb, flanking regions of circRNAs [20]. The study also found that circRNAs with LLERCPs have an expression level significantly higher than those without LLERCPs nearby, indicating LLERCPs could reinforce the expression of circRNAs, although the number of LLERCPs seem not to be related to the expression level of circRNAs [20]. Because LLERCPs were found in a relatively large flanking region of circRNAs, it is of interest to know how they are related to circRNA biogenesis. It is also of interest to know whether repeat sequences located at the flanking introns of circRNAs are associated with genome complexity so that large and polyploid genomes tend to have more repeat sequences in their flanking introns of circRNAs.

Despite most fundamental features of plant circRNAs are evolutionary conserved in different plant species, some are only observed in certain plants. The majority exonic circRNAs in rice and Arabidopsis contain 1-4 exons [11,19,21]. Large parental genes with multiple shorter exons are preferentially circularized [19,20]. Multiple circRNAs can be generated from a single parental gene through alternative backsplicing and circularization [11,14,19,29]. In rice, ~6% of exonic circRNAs had their expression levels significantly positively correlated with the expression levels of their parental genes, and no negatively correlated pair was found [11]. In kiwifruit (*Actinidia spp.*), both exonic and intronic circRNAs were significantly positively correlated to parental genes [30]. This is different from the findings observed in maize and barley, in which no correlation between the accumulation of circRNAs and their corresponding parental genes was found [20,31].

Parental genes of over 700 exonic circRNAs (~15% of Arabidopsis circRNAs) are orthologues between rice and Arabidopsis [11]. Approximately 34% and 55% of circRNA-producing soybean genes are conserved orthologs in Arabidopsis and rice, respectively [21]. These results suggest conservation of circRNAs in plants. In cotton (*G. hirsutum*), ~15% of circRNAs are conserved in either Arabidopsis or rice, and ~7% are conserved in both Arabidopsis and rice [28]. However, of the 2804 maize circRNAs, only 47 (<2%) are conserved between maize and rice, and three are conserved between maize and Arabidopsis [20]. In view of the circRNA feature of tissue-specific expression, this low conservation of circRNAs could be related to different tissues used in different plants, but it could also be contributed by use of different criteria. Conservation of plant circRNAs demands more robust investigations by applying unified criteria to circRNAs from multiple plant species.

Soybean (*Glycine max*) is a partially diploidized tetraploid and about three quarters of its genes have paralogous copies. It was found that ~83% of soybean circRNAs are generated from paralogous genes but from different positions although paralogous gene pairs usually have high sequence similarity. In addition, paralogous circRNAs showed tissue-specific expression pattern [21]. In cotton, ~20% of circRNAs are conserved

between *G. hirsutum* and its two potential ancestors (*G. arboreum* and *G. raimondii*), and ~44% of circRNAs are unique to each individual species [28]. These results suggest the presence of paralog- and species-specific circRNAs. Identification and characterization of homoeologous circRNAs from other polyploids will further shed lights on whether these features are conserved in polyploid species.

Identification of Plant Circular RNAs

Presence of circRNAs in plants was first demonstrated in 2014 [32]. We then first did genome-wide identification of circRNAs in Arabidopsis and rice and found a large number of circRNAs in both species, demonstrated widespread of circRNAs in plants [11]. Since then, a number of circRNA identifications have been carried out in other plant species, including wheat (*Triticum aestivum*), barley (*Hordeum vulgare*), maize, tomato (*Solanum lycopersicum*), potato (*Solanum tuberosum*), soybean, cotton and kiwifruit (Table 1). Dozens to thousands of circRNAs were found in these species. The vast majority of these are nuclear circRNAs, but circRNAs derived from chloroplasts and mitochondria have also been found. For example, 6% and 1% of the 803 Arabidopsis circRNAs identified by Sun et al. (2016)[27] are derived from chloroplast and mitochondrion genes, respectively.

Most *de novo* identification studies used ribosome RNA depleted and RNase R [33] treated total RNA for creation of sequencing libraries with an aim to enrich circRNA species. CircRNAs were also successfully identified in publically available RNA-seq datasets that were generated using poly(A) selected mRNAs [14,34,35]. Regarding the algorithms used in circRNA identification, almost all studies adopted the pipelines or programs developed for animals and humans. That is one of the potential reasons contributing to high false positive rate of the identified plant circRNAs, particularly those from non-coding genomic regions, such as introns and intergenic regions. For example, in our experiments, 10 of the 18 exonic circRNAs were confirmed by sequencing PCR products amplified by divergent primers, but none of the 30 predicted non-exonic circRNAs was confirmed [11]. Some studies applied multiple programs to the same set of data for circRNA prediction [20,29]. Similar to what has been previously

reported [36], poor overlapping was evident. For instance, <21% of the maize circRNAs could be simultaneously identified by CIRI, CIRC_FINDER and CIRCexplorer2 [20], suggesting poor accuracy and/or sensitivity when using the programs designed for animals and humans in plants. To improve the accuracy and sensitivity of plant circRNA prediction, we developed a program termed PcircRNA_finder based on the features of plant genome and transcriptome. Based on the testing results using simulated and real RNA-seq datasets, PcircRNA_finder is much more sensitive and effective than find_circ and CIRCexplorer [37].

Exonic circRNAs are the focus of most studies in both animals and plants; however, the proportion of exonic circRNAs is significantly different in different plant species. In Arabidopsis and tomato, ~86% of circRNAs were derived from exon(s) [11,38]. In rice, only about half identified circRNAs are derived from exon(s) [11,19]. In cotton, ~60% (*G. arboreum*), ~84% (*G. raimondii*) and ~46% (*G. hirsutum*) of circRNAs are exonic [28]. In wheat, soybean and kiwifruit, only ~7%, ~16% and ~21% are exonic circRNAs, respectively, the majority are intergenic circRNAs [21,30,39]. One of the future tasks would be to find the reason(s) behind the discrepancy. Is it due to 1) species difference, 2) poor annotation of genome sequences, or 3) only a small portion of the circRNA repertories have been revealed so the circRNAs identified are biased?

Assembling full-length circRNAs is important for understanding the internal structures, alternative circularization and biogenesis of circRNAs. Considering that nearly all circRNA identification tools can only find the backsplicing site of candidate circRNAs, we developed a pipeline, circseq_cup, for assembling full-length circRNAs in plants [14]. The core of the circseq_cup algorithm is to identify and align back-spliced junction reads and their corresponding paired reads. Using this principle, we successfully assembled ~3,000 full-length circRNAs in rice. Similar rationale has also been used by the CIRI-AS algorithm for detection of the internal structures of circRNAs in human and fruit fly [40]. Based on the assembly rationale of circseq_cup and CIRI-AS, theoretically if the fragments used in sequencing library preparation are large enough and/or the length of RNA-seq reads are long enough, circRNAs of any

Plant species	Tissue/RNA used	No. of circRNAs identified	Functionally characterized circRNA[*]	Reference
Arabidopsis	Root/rRNA depleted	Several by in-house scripts		32
thaliana	Leaf/rRNA depleted and RNase R treated	6,012 by in-house scripts		11
	Multiple tissues/rRNA depleted RNA-seq data from NCBI	30,534 by CIRI, CIRCexplorer, PcircR-NA_finder, circRNA_finder, KNIFE and find_circ		59
	Seedling/rRNA depleted and RNase R treated	2,165 by CIRCexplorer2		51
	Multiple tissues/rRNA depleted and RNase R treated, and RNA-seq data from NCBI	803 by in-house scripts		27
	Leaf from different developmental stages/ RNA-seq data from NCBI	168 by in-house scripts		34
	Multiple tissues/rRNA depleted and RNase R treated	5,861 by in-house scripts		61
	Multiple tissues/RNA-seq data downloaded from NCBI	25,982 by CIRCexplorer2 and CIRI2		60
			SEPALLATA3	45
	Seedling/rRNA depleted and RNase R treated	1,534 intronic circRNAs resistant to RNase R	*At5g37720*	49,50
Oryza sativa	Root/rRNA depleted and RNase R treated	12,037 by in-house scripts		11
	Leaf and panicle/rRNA depleted and RNase R treated	2,354 by in-house scripts	*AK064900*	19
	Root/rRNA depleted and RNase R treated	3,011 by in-house scripts		14

		26,160 by CIRI, CIRCexplorer, PcircR-NA_finder, circRNA_finder, KNIFE and find_circ		59
	Multiple tissues/rRNA depleted RNA-seq data from NCBI			
Zea mays	Multiple tissues/rRNA depleted and RNase R treated, and RNA-seq data from NCBI	2,804 by CIR-Cexplorer2, CIRC_FINDER and CIRI		20
Triticum aestivum	Leaf/circRNA Enrichment Kit	88 by CIRI		39
Hordeum vulgare	Grain transfer cells/amplified RNAs	62 by CIRI		31
Glycine max	Leaf, stem and root/rRNA depleted and RNase R treated	5,372 by CIRI		21
Solanum lycopersicum	Fruit pericarps of three developmental stages/rRNA depleted and RNase R treated	9,598 by Sege-mehl and 1,018 by CIRI	*Phytoene Synthase* 1, *Phytoene Desaturase*	29
	Fruit mesocarp/rRNA depleted and RNase R treated	318 by CIRI		38
	Fruit mesocarp/rRNA depleted and RNase R treated	854 by CIRI		52
	Fruit/rRNA depleted and RNase R treated	705 by CIRI		63
Solanum tuberosum	Stem/RNA-seq data from NCBI	2,098 by CIRI2		35
Gossypium	0 and 5 dpa outer integuments/rRNA depleted	2,262 by CIRI		62
spp.	Leaf and 0 dpa ovule/rRNA depleted	4,326 by CIRI		28
Actinidia spp.	Leaf, root and stem/rRNA depleted	3,582 by CIRI		30

*Listed are the parental genes from which the functionally characterized circRNAs are derived.

full-lengths could be assembled and identified. To support this, we found that, for RNA-seq data with an insert size of 300-bp and 400-bp, the longest full-length circRNAs assembled in rice were 488 and 680 bp, respectively, although the majority of full-length circRNAs in rice are 150-180 bp long [14]. The study by Gao et al. (2016)[40] also found that a large proportion of alternative spliced circRNA exons can hardly be detected in linear mRNAs and are enriched with binding sites of splicing factors distinct from those enriched in linear mRNAs, suggesting that algorithms based solely on backsplicing junctions may not be enough for identification of all types of circRNAs.

Given the low expression levels of most circRNAs, their tissue-specific expression profile and their complexity of alternative splicing, the number of circRNAs identified so far are likely to be just the tip of the iceberg. To enhance identification of plant circRNAs, particularly full-length circRNAs, to improve the prediction accuracy and to better understand plant circRNA biogenesis, it is of great importance to develop robust novel algorithms incorporating the characteristics of plant genomes and circRNAs, and to adopt new sequencing technologies generating long reads or full-length transcripts so that the internal structures, alternative splicing events and contribution of *trans*-backsplicing can be thoroughly evaluated.

Functions of Plant Circular RNAs and their underlying Molecular Mechanisms

In animals and humans, circRNAs have been shown to be involved in a wide range of biological processes through various molecular mechanisms [9,41]. They can serve as miRNA sponges to attenuate miRNA-mediated regulation [42,43], regulate the functions of RNA-binding proteins by direct interaction in a way as they modulate miRNA activity [15,44], and regulate the transcription of their parental genes [15,16]. Given the potential large repertoire of circRNAs, their origin from genes with important biological roles, and the relationship between circRNAs and their parental genes, it is not surprise to uncover diverse biological functions of plant circRNAs. A recent study has demonstrated a role of circRNAs in development of Arabidopsis floral organs through a novel molecular mechanism [45], but until now, most plant circRNA related studies were still at the preliminary level with a focus on prediction of

potential circRNAs acting as miRNA sponges, analysis of the expression correlation between circRNAs and their parental genes and the co-expression networks involving circRNAs.

CircRNAs Function as miRNA Sponges

A number of potential plant circRNAs have been bioinformatically predicted to have miRNA sponge capacity because the presence of miRNA binding sites [38,39]. The number of circRNAs with potential miRNA binding site(s) and the number of potential miRNA binding site(s) in each individual circRNAs vary significantly in different studies and plant species. We found only a small proportion of circRNAs (6.6% and 5.0% in rice and Arabidopsis, respectively) with the potential capacity acting as target mimics of miRNAs [11]. In another rice study, although 235 (~17%) of the 1356 exonic circRNAs were found to contain putative miRNA binding sites, only 31 have two or more miRNA-binding sites, which is significantly less than that reported for the human circRNA *CDR1as* [19,43]. Similarly in maize, circRNAs have an average of 1.33 miRNA binding sites but the number of circRNAs with miRNA binding site is very low (only 15 out of 2804) [20]. These results do not provide strong evidence for the enrichment of miRNA binding sites in plant circRNAs.

To verify experimentally whether circRNAs can function as miRNA sponges, transgenic rice plants overexpressing Os08circ16564, a rice circRNA predicted to contain binding sites of miR172, was generated [19]. Overexpressing the circRNA reduced the expression level of its parental gene but without affecting the expression level of miR172, a miRNA shown to be a crucial regulator of spikelet and floral organ development in rice [46,47]. Consequently, no visible phenotypic changes in spikelet or floral organs were observed in the overexpressing rice plants [19]. These results do not support Os08circ16564 functioning as a miR172 sponge.

In fact, precautions should be taken when predicting miRNA binding site(s) of circRNAs. Unless the miRNA binding site is located at the newly formed backsplicing junctions or the circRNA-forming exon(s) are skipped during splicing of the corresponding linear mRNA, the predicted binding site exists in both the circRNA and its related linear transcripts, and its belonging is unable to be distinguished bioinformatically.

Sophisticated experiments are thus required to evaluate separately the expression changes of circRNA and its related linear transcripts, and to assess their relevant biological consequences.

CircRNA Regulates Transcription of Parental Gene by R-Loop Formation

CircRNAs are able to bind to RNA polymerase II complex by interacting with the U1 snRNA at their parental loci to regulate the transcription of their parental genes. It has been proposed that circRNAs localize or position the U1 snRNA so that U1 can stimulate RNA polymerase II activity [16]. A recent study in Arabidopsis revealed a new mechanism, by which circRNA regulates the transcription of its parental gene [45]. It was found that circRNA could bind to the DNA of its host gene through R-loop, a structure formed by hybridization of RNA and DNA, and control linear alternative splicing, and consequently the composition and expression levels of alternatively spliced variants. The Arabidopsis *SEPALLATA3* (*SEP3*) gene, an important MADS-box transcription factor involved in floral organ development, is alternatively spliced under low temperature conditions, producing a splicing variant (*SEP3.3*) without the 6th exon and a circRNA derived from the skipped exon. Transgenic Arabidopsis plants (Exon6-circRNA) overexpressing the circRNA derived from the 6th exon of *SEP3* were generated to investigate the potential functions of the circRNA. It was found that experimental overexpression of the Exon6-circRNA led to an increased level of *SEP3.3* and caused mutant floral phenotypes (reduced stamen number and increased petal number) comparable with that observed in the transgenic plants overexpressing *SEP3.3*. Dot blot experiments showed that Exon6-circRNA hybridized to DNA of the host locus by forming R-loops [45]. Together, these results suggest the ability of the skipped exon derived circRNA in stimulating exon skipping in its parental mRNA, probably by increasing pausing of RNA polymerase II due to the formation of stable R-loops. Whether this is a general function of many circRNAs in eukaryotes remains an open question.

The study by Conn et al. (2017)[45] provided evidence for the association between production of circRNAs and exon skipping in their host mRNAs. The same study also detected reliable exon-skipping events in 11 out of the 12 most abundant exonic circRNAs identified in Arabidopsis leaf RNAs [11]. These results suggest the possibility of using the abundantly expressed circRNAs as biomarkers for prediction of alternative splicing variants with skipped exon(s) [45,48].

In tomato, circRNAs derived from several genes associated with fruit ripening and coloring, including *Phytoene Synthase 1* (*PSY1*) and *Phytoene Desaturase* (*PDS*), key genes involved in carotenoid biosynthesis, were identified [29]. The expression levels of both *PSY1* and *PDS* and their circRNAs (*PSY1-circ1* and *PDS-circ1*) are gradually increased during the period of fruit ripening (color changed from green to red). When *PSY1-circ1* or *PDS-circ1* was overexpressed, transgenic fruits showed yellow instead of red. In these yellow fruits, the expression levels of *PSY1* and *PDS* are significantly decreased compared to that of wild-type. Consistent with this down-regulated levels of *PSY1* and *PDS*, the contents of lycopene and β-carotene in the yellow transgenic fruits are significantly lower than those in wild-type [29]. These results suggest that over-production of *PSY1-circ1* or *PDS-circ1* has a negative effect on transcription and/or accumulation of its parental gene, and subsequently reduces accumulation of red pigments. The molecular mechanism underlying the circRNA over-production induced repression of parental gene is yet to be uncovered.

A Role of Intronic CircRNA in Regulation of Global Gene Expression

In human, an intronic circRNA, *ci-ankrd52*, is associated with the elongating form of RNA polymerase II at its host locus and acts as a positive regulator of the transcription speed of its host gene. Knock-down *ci-ankrd52* reduces the expression level of its host gene. The stimulating role of *ci-ankrd52* could be due to its direct *cis*-interaction with its host gene or decreased rate of including downstream introns with premature stop codons in the mature linear host mRNA, which is known to reduce mRNA abundance [44]. It has also been speculated that intronic circRNAs may potentially have effects on the expression of whole transcriptome because *ci-ankrd52* can bind to genetic loci in addition to their host loci [44]. A recent study in Arabidopsis provided evidence for such a mechanism [49].

miRNAs are small non-coding RNA molecules and play an important role in plant development and stress tolerance by regulating the expression levels of their target genes via mRNA cleavage and/or translational repression. Biogenesis of miRNAs involves dicing of primary miRNAs (pri-miRNAs) by the DCL1/HYL1 complex. Intron lariats are by-products of linear mRNA splicing and are usually quickly degraded by the RNA debranching enzyme 1 (DRB1). This biological process is critical because dysfunctional DRB1 results in embryo lethal in both animals and plants. Arabidopsis non-lethal mutant *dbr1-2* accumulated a large number of RNase R resistant circular intron lariat RNAs, i.e. intronic circRNAs, and showed global reduction in miRNA accumulation due to intronic circRNA-mediated sequestration of the DCL1/HYL1 complex, which competitively inhibited the binding of the complex with pri-miRNAs and reduced accumulation of miRNAs. This observation was further confirmed by overexpression of an intronic circRNA (lariat41) derived from the first intron of At5g37720 [49]. Arabidopsis plants overexpressing the circRNA (lariat41-OE) showed curly leaves, late flowering and lower fertility [50], similar to phenotypes observed in some miRNA mutants. In addition, the expression level of *FT*, a key flowering time regulator, was significantly reduced in the lariat41-OE plants, which could also contribute to the late flowering phenotype. These results suggest that intronic circRNAs are normally by-products of splicing and degraded by DBR1, which acts as a safeguard to ensure the proper biological functions associated with the miRNA pathways, and that unusual accumulation of intronic circRNAs could be potentially used as biomarkers for certain functional defaults. These speculations and the roles of intronic circRNAs are waiting to be confirmed because visible mutant phenotypes were only observed in plants overexpressing lariat41 and not in plants overexpressing one of the other 16 intronic circRNAs that were significantly enriched in the *dbr1-2* mutant [50].

Potential Roles of circRNAs in Biotic and Abiotic Stress Responses

Similar to enrichment of intronic circRNAs in the *dbr1-2* mutant [49], differential expression of circRNAs under certain biotic or abiotic stress conditions have been investigated in several plants. CircRNAs specifically expressed at a specific developmental time point/stage or induced/repressed by certain stresses could have their roles related to certain development events at that time point/stage or in fighting against the stresses. In other words, differentially expressed circRNAs may act as important functional regulators involved in developmental-/stress-specific biological processes in plants.

We identified 27 rice exonic circRNAs differentially expressed under phosphate-sufficient or -starvation conditions [11]. In response to cold and heat treatment, 163 and 1583 circRNAs were differentially expressed in tomato and Arabidopsis, respectively [51,52]. Under heat stress conditions, the length of circRNAs increased probably due to the increased number of exons involved in circularization, suggesting that heat stress changed not only the quantity but also the quality (i.e. alternative backsplicing and circularization) of circRNAs [51]. In wheat, 62 out of the 88 identified circRNAs were found to be differentially expressed in the seedlings under dehydration stress conditions [39]. CircRNAs responsive to biotic stresses have also been reported. For instance, 584 circRNAs responsive to infection of *Pseudomonas syringae* pv. *actinidiae* (Psa), the agent for the bacterial canker disease, were found in kiwifruit [30]. In another study in potato, 429 circRNAs were found to be differentially expressed in response to infection of *Pectobacterium carotovorum* subsp. *brasiliense* (Pcb), the most important causal agent of potato blackleg and soft rot globally. Upon Pcb challenge, some circRNAs were down-regulated in cultivar susceptible to the disease but up-regulated in cultivar resistant to the disease, or vice-versa, suggesting a potential role of these circRNAs in disease response [35]. Despite differential expression of certain circRNAs in responses to various biotic or abiotic stresses, no such plant circRNA has been experimentally demonstrated to be functional.

Conclusions

CircRNAs have been found in all eukaryotes studied so far, ranging from metazoans, plants, to unicellular eukaryotes. RNA circularization is thus thought to be an evolutionarily conserved process [9,13,53]. Since finding of a large number of circRNAs in 2012 thanks to the advance of high-throughput sequencing

technologies and bioinformatics tools [8], our knowledge and understanding on circRNAs and their biogenesis have been improved very quickly and are still being constantly renewed. CircRNAs can arise from all genomic regions, including both protein-conding and non-coding genes, they can be non-protein-coding but can also be endogenously translated into peptides [54-56]. Our knowledges on circRNAs are mainly based on studies in animals and humans, which seem to be applicable in plants, but plant circRNAs do have their own unique features. In addition, the majority published studies on plant circRNAs are superficial. To deepen our understandings on plant circRNAs, attentions need to be paid to:

1) Investigate the mechanism(s) underlying biogenesis of plant circRNAs. In view of the observations that relatively few plant circRNAs are associated with repetitive and/or reverse complementary sequences in their flanking regions and that a large proportion of plant circRNAs contain non-GT/AG canonical splicing signals, it has been suggested that plants may produce circRNAs using pathways different from those found in animals. With more and more circRNAs identified in plants, we should be able to know whether lack of repetitive and reverse complementary sequences in the flanking regions of plant circRNAs is a commonality in different plant species. Meanwhile, detailed analysis of the sequences flanking all plant circRNAs can be performed to see whether they contain certain hidden common feature(s) and sequence motif(s) recognizable by RNA-binding proteins with a role in RNA splicing and/or circularization, such as homologs of Quaking and Muscleblind [15,57].

2) Develop novel algorithms specific for plant circRNA identification and improve the annotation quality of genome sequences. A major challenge of circRNA identification is to distinguish true backsplicing junctions from other types of non-collinear junctions (such as those derived from genetic rearrangements) or artefacts produced in sequencing library preparation and short read alignment [58]. The new algorithms need to address simultaneously the accuracy and sensitivity of circRNA prediction. The only program (PcircRNA_finder) specifically developed for plants in our lab has a high sensitivity but still having room for further improvement

of the prediction accuracy [37]. The quality of genome sequences and annotation is one of the major factors related to the false positive rate. A large number of plant species have been sequenced but most genome assemblies are far from satisfaction, particularly those of polyploids due to the presence of a large proportion of repeat sequences. Improving their annotation quality would significantly reduce the false positive rate of circRNA identification.

3) Enhance functional investigation of plant circRNAs. Until now, few plant circRNAs have been functionally characterized, which were done mainly by overexpressing. Acting as miRNA sponges is one of the major potential functions of plant circRNAs based on the results reported in animals. Almost every study of plant circRNA identification predicted certain number of potential circRNAs to be putative miRNA sponges. However, whether circRNAs function as miRNA sponges is yet to be analyzed by using more candidate circRNAs. In addition, study of the circRNA derived from a skipped exon of *SEP3* revealed a novel molecular mechanism by which circRNA regulates the transcription of its host gene [45], suggesting that circRNAs may have diverse regulatory modes that are yet to be discovered. Gene silencing has not been used in functional analysis of plant circRNAs. Use of this approach will be a challenge due to the sequence similarity between circRNAs and their parental genes, but it can be overcome using the CRISPR/Cas9 gene-editing technology through careful selection of the guide RNA.

In a word, despite growing number of the published papers and databases on plant circRNAs [59,60], investigations about the molecular mechanism(s) underlying biogenesis and functions of plant circRNAs are still rare, partly due to the complicated intertwining of linear and circular splicing events that can only be distinguished by rigorous controls at multiple steps of experimentation. But we believe that more exciting findings about plant circRNAs will be revealed in the coming years.

joint venture between Cotton Seed Distributors Ltd and CSIRO.

Conflict of interest

We declare that we have no conflict of interest.

References

1. Sanger, H.L., Klotz, G., Riesner, D., Gross, H.J., and Kleinschmidt, A.K. (1976) 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

2. Arnberg, A.C., Van Ommen, G.J., Grivell, L.A., Van Bruggen, E.F., and Borst, P. (1980) Some yeast mitochondrial RNAs are circular. Cell 19, 313-319

3. Capel, B., Swain, A., Nicolis, S., Hacker, A., Walter, M., Koopman, P., Goodfellow, P., and Lovell-Badge, R. (1993) Circular transcripts of the testis-determining gene *Sry* in adult mouse testis. Cell 73, 1019-1030

4. Cocquerelle, C., Daubersies, P., Majerus, M.A., Kerckaert, J.P., and Bailleul, B. (1992) Splicing with inverted order of exons occurs proximal to large introns. EMBO J 11, 1095-1098

5. Hsu, M.T., and Coca-Prados, M. (1979) Electron microscopic evidence for the circular form of RNA in the cytoplasm of eukaryotic cells. Nature 280, 339-340

6. Nigro, J.M., Cho, K.R., Fearon, E.R., Kern, S.E., Ruppert, J.M., Oliner, J.D., Kinzler, K.W., and Vogelstein, B. (1991) Scrambled exons. Cell 64, 607-613

7. Pasman, Z., Been, M.D., and Garcia-Blanco, M.A. (1996) Exon circularization in mammalian nuclear extracts. RNA 2, 603-610

8. Salzman, J., Gawad, C., Wang, P.L., Lacayo, N., and Brown, P.O. (2012) Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. PLoS One 7, e30733

9. Chen, L.L. (2016) The biogenesis and emerging roles of circular RNAs. Nat Rev Mol Cell Biol 17, 205-211

10. Salzman, J. (2016) Circular RNA expression: its potential regulation and function. Trends Genet 32, 309-316

11. Ye, C.Y., Chen, L., Liu, C., Zhu, Q-H., and Fan, L. (2015) Widespread noncoding circular RNAs in plants. New Phytol 208, 88-95

12. Starke, S., Jost, I., Rossbach, O., Schneider, T., Schreiner, S., Hung, L.H., and Bindereif, A. (2015) Exon circularization requires canonical splice signals. Cell Rep 10, 103-111

13. Holdt, L.M., Kohlmaier, A., and Teupser, D. (2017) Molecular roles and function of circular RNAs in eukaryotic cells. Cell Mol Life Sci doi: 10.1007/s00018-017-2688-5

14. Ye, C.Y., Zhang, X., Chu, Q., Liu, C., Yu, Y., Jiang, W., Zhu, Q-H., Fan, L., and Guo, L. (2017) Full-length sequence assembly reveals circular RNAs with diverse non-GT/AG splicing signals in rice. RNA Biol 14, 1055-1063

15. Ashwal-Fluss, R., Meyer, M., Pamudurti, N.R., Ivanov, A., Bartok, O., Hanan, M., Evantal, N., Memczak, S., Rajewsky, N., and Kadener, S. (2014) CircRNA biogenesis competes with pre-mRNA splicing. Mol Cell 56, 55-66

16. 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., and Shan, G. (2015) Exon-intron circular RNAs regulate transcription in the nucleus. Nat Struct Mol Biol 22, 256-264

17. Liang, D., Tatomer, D.C., Luo, Z., Wu, H., Yang, L., Chen, L.L., Cherry, S., and Wilusz, J.E. (2017) The Output of protein-coding genes shifts to circular RNAs when the pre-mRNA processing machinery is limiting. Mol Cell 68, 940-954 e943

18. Gao, Y., Wang, J., and Zhao, F. (2015) CIRI: an efficient and unbiased algorithm for *de novo* circular RNA identification. Genome Biol 16, 4

19. Lu, T., Cui, L., Zhou, Y., Zhu, C., Fan, D., Gong, H., Zhao, Q., Zhou, C., Zhao, Y., Lu, D., Luo, J., Wang, Y., Tian, Q., Feng, Q., Huang, T., and Han, B. (2015) Transcriptome-wide investigation of circular RNAs in rice. RNA 21, 2076-2087

20. Chen, L., Zhang, P., Fan, Y., Lu, Q., Li, Q., Yan, J., Muehlbauer, G.J., Schnable, P.S., Dai, M., and Li, L. (2017) Circular RNAs mediated by transposons are

associated with transcriptomic and phenotypic variation in maize. New Phytol doi: 10.1111/nph.14901

21. Zhao, W., Cheng, Y., Zhang, C., You, Q., Shen, X., Guo, W., and Jiao, Y. (2017) Genome-wide identification and characterization of circular RNAs by high throughput sequencing in soybean. Sci Rep 7, 5636

22. Jeck, W.R., and Sharpless, N.E. (2014) Detecting and characterizing circular RNAs. Nat Biotechnol 32, 453-461

23. Liang, D., and Wilusz, J.E. (2014) Short intronic repeat sequences facilitate circular RNA production. Genes Dev 28, 2233-2247

24. Jeck, W.R., Sorrentino, J.A., Wang, K., Slevin, M.K., Burd, C.E., Liu, J., Marzluff, W.F., and Sharpless, N.E. (2013) Circular RNAs are abundant, conserved, and associated with *ALU* repeats. RNA 19, 141-157

25. Zhang, X.O., Wang, H.B., Zhang, Y., Lu, X., Chen, L.L., and Yang, L. (2014) Complementary sequence-mediated exon circularization. Cell 159, 134-147

26. Ivanov, A., Memczak, S., Wyler, E., Torti, F., Porath, H.T., Orejuela, M.R., Piechotta, M., Levanon, E.Y., Landthaler, M., Dieterich, C., and Rajewsky, N. (2015) Analysis of intron sequences reveals hallmarks of circular RNA biogenesis in animals. Cell Rep 10, 170-177

27. Sun, X., Wang, L., Ding, J., Wang, Y., Wang, J., Zhang, X., Che, Y., Liu, Z., Zhang, X., Ye, J., Wang, J., Sablok, G., Deng, Z., and Zhao, H. (2016) Integrative analysis of Arabidopsis thaliana transcriptomics reveals intuitive splicing mechanism for circular RNA. FEBS Lett 590, 3510-3516

28. Zhao, T., Wang, L., Li, S., Xu, M., Guan, X., and Zhou, B. (2017) Characterization of conserved circular RNA in polyploid Gossypium species and their ancestors. FEBS Lett 591, 3660-3669

29. Tan, J., Zhou, Z., Niu, Y., Sun, X., and Deng, Z. (2017) Identification and functional characterization of tomato circRNAs derived from genes involved in fruit pigment a ccumulation. Sci Rep 7, 8594

30. Wang, Z., Liu, Y., Li, D., Li, L., Zhang, Q., Wang, S., and Huang, H. (2017) Identification of circular RNAs in kiwifruit and their species-specific response to bacterial canker pathogen invasion. Front Plant Sci 8, 413

31. Darbani, B., Noeparvar, S., and Borg, S. (2016) Identification of circular RNAs from the parental genes involved in multiple aspects of cellular metabolism in barley. Front Plant Sci 7, 776

32. Wang, P.L., Bao, Y., Yee, M.C., Barrett, S.P., Hogan, G.J., Olsen, M.N., Dinneny, J.R., Brown, P.O., and Salzman, J. (2014) Circular RNA is expressed across the eukaryotic tree of life. PLoS One 9, e90859

33. Vincent, H.A., and Deutscher, M.P. (2006) Substrate recognition and catalysis by the exoribonuclease RNase R. J Biol Chem 281, 29769-29775

34. Liu, T., Zhang, L., Chen, G., and Shi, T. (2017) Identifying and characterizing the circular RNAs during the lifespan of Arabidopsis leaves. Front Plant Sci 8, 1278

35. Zhou, R., Zhu, Y., Zhao, J., Fang, Z., Wang, S., Yin, J., Chu, Z., and Ma, D. (2017) Transcriptome-wide identification and characterization of potato circular RNAs in response to *Pectobacterium carotovorum* Subspecies *brasiliense* infection. Int J Mol Sci 19, doi: 10.3390/ijms19010071

36. Zeng, X., Lin, W., Guo, M., and Zou, Q. (2017) A comprehensive overview and evaluation of circular RNA detection tools. PLoS Comput Biol 13, e1005420

37. Chen, L., Yu, Y., Zhang, X., Liu, C., Ye, C., and Fan, L. (2016) PcircRNA_finder: a software for circRNA prediction in plants. Bioinformatics 32, 3528-3529

38. Wang, Y., Wang, Q., Gao, L., Zhu, B., Luo, Y., Deng, Z., and Zuo, J. (2017) Integrative analysis of circRNAs acting as ceRNAs involved in ethylene pathway in tomato. Physiol Plant 161, 311-321

39. Wang, Y., Yang, M., Wei, S., Qin, F., Zhao, H., and Suo, B. (2016) Identification of circular RNAs and their targets in leaves of *Triticum aestivum* L. under dehydration stress. Front Plant Sci 7, 2024

40. Gao, Y., Wang, J., Zheng, Y., Zhang, J., Chen, S., and Zhao, F. (2016) Comprehensive identification of internal structure and alternative splicing events in circular RNAs. Nat Commun 7, 12060

41. Piwecka, M., Glazar, P., Hernandez-Miranda, L.R., Memczak, S., Wolf, S.A., Rybak-Wolf, A., Filipchyk, A., Klironomos, F., Cerda Jara, C.A., Fenske, P., Trimbuch, T., Zywitza, V., Plass, M., Schreyer, L., Ayoub, S., Kocks, C., Kuhn, R., Rosenmund, C., Birchmeier, C., and Rajewsky, N. (2017) Loss of a mammalian circular RNA locus causes miRNA deregulation and affects brain function. Science 357, doi: 10.1126/science.aam8526

42. Hansen, T.B., Jensen, T.I., Clausen, B.H., Bramsen, J.B., Finsen, B., Damgaard, C.K., and Kjems, J. (2013) Natural RNA circles function as efficient microRNA sponges. Nature 495, 384-388

43. Memczak, S., Jens, M., Elefsinioti, A., Torti, F., Krueger, J., Rybak, A., Maier, L., Mackowiak, S.D., Gregersen, L.H., Munschauer, M., Loewer, A., Ziebold, U., Landthaler, M., Kocks, C., le Noble, F., and Rajewsky, N. (2013) Circular RNAs are a large class of animal RNAs with regulatory potency. Nature 495, 333-338

44. Zhang, Y., Zhang, X.O., Chen, T., Xiang, J.F., Yin, Q.F., Xing, Y.H., Zhu, S., Yang, L., and Chen, L.L. (2013) Circular intronic long noncoding RNAs. Mol Cell 51, 792-806

45. Conn, V.M., Hugouvieux, V., Nayak, A., Conos, S.A., Capovilla, G., Cildir, G., Jourdain, A., Tergaonkar, V., Schmid, M., Zubieta, C., and Conn, S.J. (2017) A circRNA from *SEPALLATA3* regulates splicing of its cognate mRNA through R-loop formation. Nat Plants 3, 17053

46. Zhu, Q-H., Upadhyaya, N.M., Gubler, F., and Helliwell, C.A. (2009) Over-expression of miR172 causes loss of spikelet determinacy and floral organ abnormalities in rice (*Oryza sativa*). BMC Plant Biol 9, 149

47. Zhu, Q-H., and Helliwell, C.A. (2011) Regulation of flowering time and floral patterning by miR172. J Exp Bot 62, 487-495

48. Kelly, S., Greenman, C., Cook, P.R., and Papantonis, A. (2015) Exon skipping is correlated with exon circularization. J Mol Biol 427, 2414-2417

49. Li, Z., Wang, S., Cheng, J., Su, C., Zhong, S., Liu, Q., Fang, Y., Yu, Y., Lv, H., Zheng, Y., and Zheng, B. (2016) Intron lariat RNA inhibits microRNA biogenesis by sequestering the dicing complex in *Arabidopsis*. PLoS Genet 12, e1006422

50. Cheng, J., Zhang, Y., Li, Z., Wang, T., Zhang, X., and Zheng, B. (2017) A lariat-derived circular RNA is required for plant development in *Arabidopsis*. Sci China Life Sci 60, doi: 10.1007/s11427-017-9182-3.

51. Pan, T., Sun, X., Liu, Y., Li, H., Deng, G., Lin, H., and Wang, S. (2017) Heat stress alters genome-wide profiles of circular RNAs in *Arabidopsis*. Plant Mol Biol 10.1007/s11103-017-0684-7

52. Zuo, J., Wang, Q., Zhu, B., Luo, Y., and Gao, L. (2016) Deciphering the roles of circRNAs on chilling injury in tomato. Biochem Biophys Res Commun 479, 132-13

53. Danan, M., Schwartz, S., Edelheit, S., and Sorek, R. (2012) Transcriptome-wide discovery of circular RNAs in *Archaea*. Nucleic Acids Res 40, 3131-3142

54. Granados-Riveron, J.T., and Aquino-Jarquin, G. (2016) The complexity of the translation ability of circRNAs. Biochim Biophys Acta 1859, 1245-1251

55. Pamudurti, N.R., Bartok, O., Jens, M., Ashwal-Fluss, R., Stottmeister, C., Ruhe, L., Hanan, M., Wyler, E., Perez-Hernandez, D., Ramberger, E., Shenzis, S., Samson, M., Dittmar, G., Landthaler, M., Chekulaeva, M., Rajewsky, N., and Kadener, S. (2017) Translation of CircRNAs. Mol Cell 66, 9-21 e27

56. Yang, Y., Fan, X., Mao, M., Song, X., Wu, P., Zhang, Y., Jin, Y., Yang, Y., Chen, L.L., Wang, Y., Wong, C.C., Xiao, X., and Wang, Z. (2017) Extensive translation of circular RNAs driven by N(6)-methyladenosine. Cell Res 27, 626-641

57. Conn, S.J., Pillman, K.A., Toubia, J., Conn, V.M., Salmanidis, M., Phillips, C.A., Roslan, S., Schreiber, A.W., Gregory, P.A., and Goodall, G.J. (2015) The RNA binding protein quaking regulates formation of circRNAs. Cell 160, 1125-1134

58. Szabo, L., and Salzman, J. (2016) Detecting circular RNAs: bioinformatic and experimental challenges. Nat Rev Genet 17, 679-692

59. Chu, Q., Zhang, X., Zhu, X., Liu, C., Mao, L., Ye, C.,

Zhu, Q-H., and Fan, L. (2017) PlantcircBase: A database for plant circular RNAs. Mol Plant 10, 1126-1128

60. Ye, J., Wang, L., Li, S., Zhang, Q., Zhang, Q., Tang, W., Wang, K., Song, K., Sablok, G., Sun, X., and Zhao, H. (2017) AtCircDB: a tissue-specific database for Arabidopsis circular RNAs. Brief Bioinform doi: 10.1093/bib/bbx089

61. Chen, G., Cui, J., Wang, L., Zhu, Y., Lu, Z., and Jin, B. (2017) Genome-wide identification of circular RNAs in *Arabidopsis thaliana*. Front Plant Sci 8, 1678

62. Hu, H., Wang, M., Ding, Y., Zhu, S., Zhao, G., Tu, L., and Zhang, X. (2017) Transcriptomic repertoires depict the initiation of lint and fuzz fibres in cotton (*Gossypium hirsutum* L.). Plant Biotechnol J doi: 10.1111/pbi.12844

63. Yin, J., Liu, M., Ma, D., Wu, J., Li, S., Zhu, Y., and Han, B. (2018) Identification of circular RNAs and their targets during tomato fruit ripening. Postharvest Biol Technol 136, 90-98