Briefings in Bioinformatics, 00(00), 2018, 1–9

doi: 10.1093/bib/bby111
Advance Access Publication Date: 15 November 2018
Review article

Characteristics of plant circular RNAs

Qinjie Chu, Panpan Bai, Xintian Zhu, Xingchen Zhang, Lingfeng Mao, Qian-Hao Zhu, Longjiang Fan, Chu-Yu Ye

Corresponding author: Chu-Yu Ye, Institute of Crop Science, Zhejiang University, Hangzhou 310058, China. Tel.: +86(571)88982730; Fax: +86(571)88982730; E-mail: yecy@zju.edu.cn

Abstract

Circular RNA (circRNA) is a kind of covalently closed single-stranded RNA molecules that have been proved to play important roles in transcriptional regulation of genes in diverse species. With the rapid development of bioinformatics tools, a huge number (95 143) of circRNAs have been identified from different plant species, providing an opportunity for uncovering the overall characteristics of plant circRNAs. Here, based on publicly available circRNAs, we comprehensively analyzed characteristics of plant circRNAs with the help of various bioinformatics tools as well as in-house scripts and workflows, including the percentage of coding genes generating circRNAs, the frequency of alternative splicing events of circRNAs, the non-canonical splicing signals of circRNAs and the networks involving circRNAs, miRNAs and mRNAs. All this information has been integrated into an upgraded online database, PlantcircBase 3.0 (http://ibi.zju.edu.cn/plantcircbase/). In this database, we provided browse, search and visualization tools as well as a web-based blast tool, BLASTcirc, for prediction of circRNAs from query sequences based on searching against plant genomes and transcriptomes.

Key words: plant circRNAs; PlantcircBase; BLASTcirc; non-canonical splicing signals

Introduction

Circular RNAs (circRNAs) are covalently closed, single-stranded RNA molecules generated by back-splicing events, and used to be considered as by-products of splicing. The first circRNAs, viroids that propagate in plants tomato and *Gynura*, were detected by Sanger *et al.* [1] in the 1970s. With the significant progress of high-throughput sequencing technology and the rapid development of bioinformatics tools and resources [2, 3], circRNAs have been recently identified on a large scale, first in humans [4] and then in many other species, from fruit fly [5, 6], worm [7], diverse mammalians [6, 8, 9] to plants [10–12]. Tens of thousands of circRNAs have been identified by various bioinformatics prediction tools, such as CIRCexplorer [13, 14], circRNA_finder [5], CIRI [15, 16], find_circ [6], KNIFE [17] and Segemehl [18]. Each tool has its own pros and cons regarding the precision and sensitivity in circRNA prediction [19–22], but all have been used to identify numerous circRNAs in various species. The number of circRNAs identified in different human samples has reached 183 000 as of July 2018 according to CIRCpedia [14]. A total of over 95 000 circRNAs have also been identified in 12 plant species (Table 1) by different research groups through June of this year.

Qinjie Chu is a PhD student of the Institute of Crop Science, Institute of Bioinformatics, Zhejiang University. Her research interest is on plant circRNAs.
Panpan Bai is an MSc student of the Institute of Crop Science, Institute of Bioinformatics, Zhejiang University. Her research interest is on plant circRNAs.
Xintian Zhu is an MSc student of the Institute of Crop Science, Institute of Bioinformatics, Zhejiang University. Her research interest is on plant non-coding RNAs.
Xingchen Zhang is a graduated MSc student of the Institute of Crop Science, Institute of Bioinformatics, Zhejiang University.
Lingfeng Mao is a PhD student of the Institute of Crop Science, Institute of Bioinformatics, Zhejiang University. His research interest is on plant genomes.
Qian-Hao Zhu is a professor of the CSIRO Agriculture and Food. His research interests are on plant functional genomics and non-coding RNAs.
Longjiang Fan is a professor of the Institute of Crop Science, Institute of Bioinformatics, Zhejiang University. His research interests are on plant genomes and non-coding RNAs.
Chu-Yu Ye is an associate professor of the Institute of Crop Science, Institute of Bioinformatics, Zhejiang University. His research interests are on plant genomes and non-coding RNAs.
Submitted: 4 September 2018; **Received (in revised form):** 25 September 2018

Table 1. Number of circRNAs predicted in plants (through June 2018)

Organisms	Total number of circRNAs[a]	References
Arabidopsis thaliana	38 938	[10, 11, 23–27]
Gossypium arboretum	1041	[28]
Gossypium hirsutum	499	[28]
Glycine max	5323	[29]
Gossypium raimondii	1478	[28]
Hordeum vulgare	39	[30]
Oryza sativa	40 311	[11, 12, 31]
Poncirus trifoliata	556	[32]
Solanum lycopersicum	1904	[33–35]
Solanum tuberosum	1728	[36]
Triticum aestivum	88	[37]
Zea mays	3238	[38]
Total	95 143	–

[a]circRNAs identified by different groups have been combined together based on their positions provided by different publications (Genome versions or other details are provided in Supplementary data).

Studies in animals and humans have shown that the biogenesis of circRNAs is regulated by *cis*-regulatory elements such as repetitive sequences or non-repetitive reverse complementary sequences flanking the circularized exons, and *trans*-acting factors such as RNA binding proteins [39, 40]. In contrast, the flanking sequences of plant circRNAs do not seem to be enriched for repetitive elements or reverse complementary sequences [11, 12], indicating that circularization of plant circRNAs might be regulated by alternative mechanisms that are yet to be found.

Recent studies have shown that circRNAs in animals play an important regulatory role in a diverse of biological processes, such as functioning as miRNA sponges [6, 41], promoting transcription of their parental coding genes [42, 43] and affecting the expression of parental genes by competing with linear splicing [44, 45]. However, the biological functions of plant circRNAs still remain largely unclear although the pervasiveness of circRNAs in plants has been confirmed by genome-wide investigation in rice and *Arabidopsis* [10–12] and other plants (Table 1). A recent study [46] demonstrated that an *Arabidopsis* circRNA generated from the sixth exon of *SEPALLATA3* (*SEP3*) regulates the expression of its parental gene (*SEP3*) by forming an R-loop (an RNA:DNA hybrid), representing the first achievement towards understanding the molecular mechanisms circRNAs in plants. More recently, a lariat-derived circRNA, generated from the first intron of AT5G37720, was discovered to regulate gene expression and influence development of *Arabidopsis* [47]. Rapid progress of research on circRNAs in both animals and plants will enhance our knowledge and understanding of how circRNAs function in regulating a diverse of biological processes.

Most plant circRNAs are generated from annotated genes

To clearly identify from where circRNAs are generated, circRNAs were classified into ten types (Figure 1A) based on the annotated genomic features, at which the two back-splicing sites of a certain circRNA is located. The ten types are e-circRNA, ei-circRNA, i-circRNA, ie-circRNA, u-circRNA, ue-circRNA, ui-circRNA, ig-circRNA, igg-circRNA and ag-circRNA (e, i, u, g, ig and ag represent exon, intron, UTR region, genic region, intergenic region and across-genic region, respectively). Apart from ig-circRNAs (No.8 in Figure 1A, green bar in Figure 1B), all other types of circRNAs are fully or partly overlapped with genic region of the genome. We found that ig-circRNAs account for a small proportion (less than 20%) of all identified circRNAs in most plant species (9 of 12), except *Gossypium hirsutum* (22%), *Solanum tuberosum* (28%), and *Zea mays* (63%) (Figure 1B, Table S1). In *Arabidopsis thaliana* and *Oryza sativa*, 86% (33,394) and 92% (37,033) of circRNAs are derived from protein coding genes, respectively. Taken together, the results revealed that most identified plant circRNAs are generated from annotated genes, including both exonic and intronic regions, and thus might play roles in influencing the expression of their host genes directly or indirectly.

We also set up a generic system for nomenclature of circRNAs based on their parental genes and classified types, and have used the system in the current plant circRNA database (PlantcircBase 3.0). In this system, each circRNA name contains four parts with the formula 'X_circ_Y.Z'. X and Y represents the parental gene and the type of the circRNA, respectively. Z is a number showing that it is the Zth circRNA from the locus. 'circ' represents circRNA. For simplicity, instead of ten types, circRNAs were divided into four types in the nomenclature system, i.e. the value of Y will be selected only from 'g', 'ig', 'igg' and 'ag', which represent genic, intergenic, intergenic-genic and across-genic circRNAs, respectively. When using this system, e-circRNA, ei-circRNA, i-circRNA, ie-circRNA, u-circRNA, ue-circRNA and ui-circRNA were all considered as genic circRNAs. For example, 'AT1G01010_circ_g.1' represents the first genic circRNA derived from the parental gene AT1G01010.

Widespread alternative splicing of plant circRNAs

Alternative splicing is a common and complex splicing pattern in higher eukaryotes [48], and has also been found in back-splicing events in humans and fruit flies [14, 49]. For circRNAs, alternative splicing may happen at back-splicing sites, which is called alternative back-splicing, or in internal sequences, which is similar to normal alternative splicing found in linear mRNAs and is called alternative splicing too [14]. More than 10 circRNA isoforms from a single locus have been experimentally validated in rice [31], suggesting that alternative splicing events could be also exists in plant circRNAs.

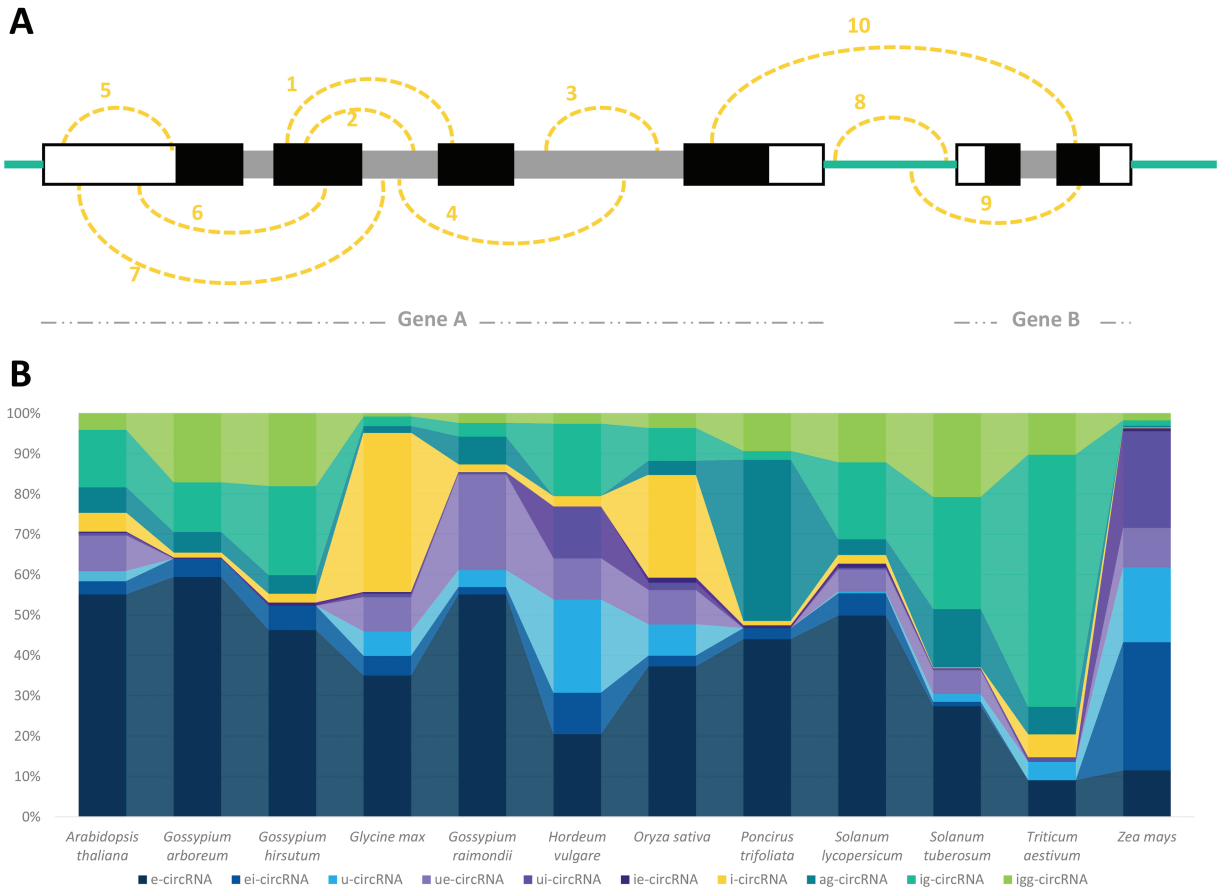

Figure 1. Types of circRNAs and their proportions in different plant species. **(A)** Ten types of circRNAs. (1) e-circRNA, two back-splicing sites of a circRNA are both at exons; (2) ei-circRNA, one back-splicing site of a circRNA is at exon while the other is at intron; (3) i-circRNA, two back-splicing sites of a circRNA are both at a single intron; (4) ie-circRNA, two back-splicing sites of a circRNA are at two different introns across one or several exons; (5) u-circRNA, two back-splicing sites of a circRNA are both at UTRs; (6) ue-circRNA, one back-splicing site of a circRNA is at UTR while the other is at exon; (7) ui-circRNA, one back-splicing site of a circRNA is at UTR while the other is at intron; (8) ig-circRNA, two back-splicing sites of a circRNA are both at a single intergenic region; (9) igg-circRNA, one back-splicing site of a circRNA is at intergenic region while the other is at genic region; (10) ag-circRNA, two back-splicing sites of a circRNA are at two different genes. The black, gray and blank bars represent exons, introns and UTRs, respectively. The green lines represent intergenic region of the genomes. **(B)** Proportion of each type of circRNA in each plant species. Different colors represent different types as shown in the bottom of the panel.

We investigated the alternative splicing events of plant circRNAs and demonstrated that alternative back-splicing are widespread. Identified circRNAs with overlapping genomic positions were defined as isoforms from the circRNA locus. The number of circRNA loci identified so far in 12 plant species were shown in Figure 2A. Up to over 30% of circRNA loci have alternative back-splicing events in *A. thaliana*, *O. sativa* and *S. tuberosum*. Approximately 10% of circRNA loci in the other plant species except *Hordeum vulgare* are alternatively spliced (Figure 2B, Table S2). No detectable alternative back-splicing event in *H. vulgare* is probably because only very limited number of circRNAs have been identified (39 in total) in this species (Figure 2A and B). Most circRNA loci generate 2–4 alternative back-splicing isoforms (Figure 2C; only those loci with ≤10 isoforms are shown). In *A. thaliana*, we found 281 circRNA loci being able to generate >10 isoforms (from 11 to 5547). The corresponding circRNA loci having more than 10 isoforms in *O. sativa* and *S. tuberosum* are 295 (with 11–205 isoforms) and 17 (with 11–168 isoforms), respectively. In view of the finding that some protein coding genes have thousands of alternatively spliced transcripts (e.g. the *DSCAM* gene in *Drosophila* could theoretically generate at most 38 016 different transcripts by alternative splicing) [50, 51], it is not surprising that a circRNA locus has hundreds or thousands of isoforms.

Diverse non-canonical splicing signals of plant circRNAs

The U2-dependent spliceosome (major spliceosome or canonical spliceosome) involved in catalyzing pre-mRNA splicing has the splicing signals of GU and AG at the 5′ and 3′ splice sites, respectively, and is the most abundant ribonucleoprotein complex in higher eukaryotes [52]. Due to the splicing mechanism of major spliceosome, most circRNA identification tools have a criterion to filter out circRNAs without the GU/AG splicing signals [53]. However, our recent study [31] has shown that both predicted and validated circRNAs in rice had diverse non-canonical splicing signals, and provided evidence for the widespread non-canonical splicing signals in plant circRNAs. By analyzing circRNAs identified in 12 plant species, we found that most circRNAs have canonical splicing signals of GT/AG or CT/AC (Figure 3A, Table S3), but the proportion of circRNAs with non-canonical splicing signals is variable in different species. The irregular distribution of canonical and non-canonical splicing signals among different species might be caused by different circRNA prediction tools, different filtering criteria and different RNA-Seq strategies used in identification of circRNAs. All of these could have a profound effect on the number and characteristics of the identified circRNAs.

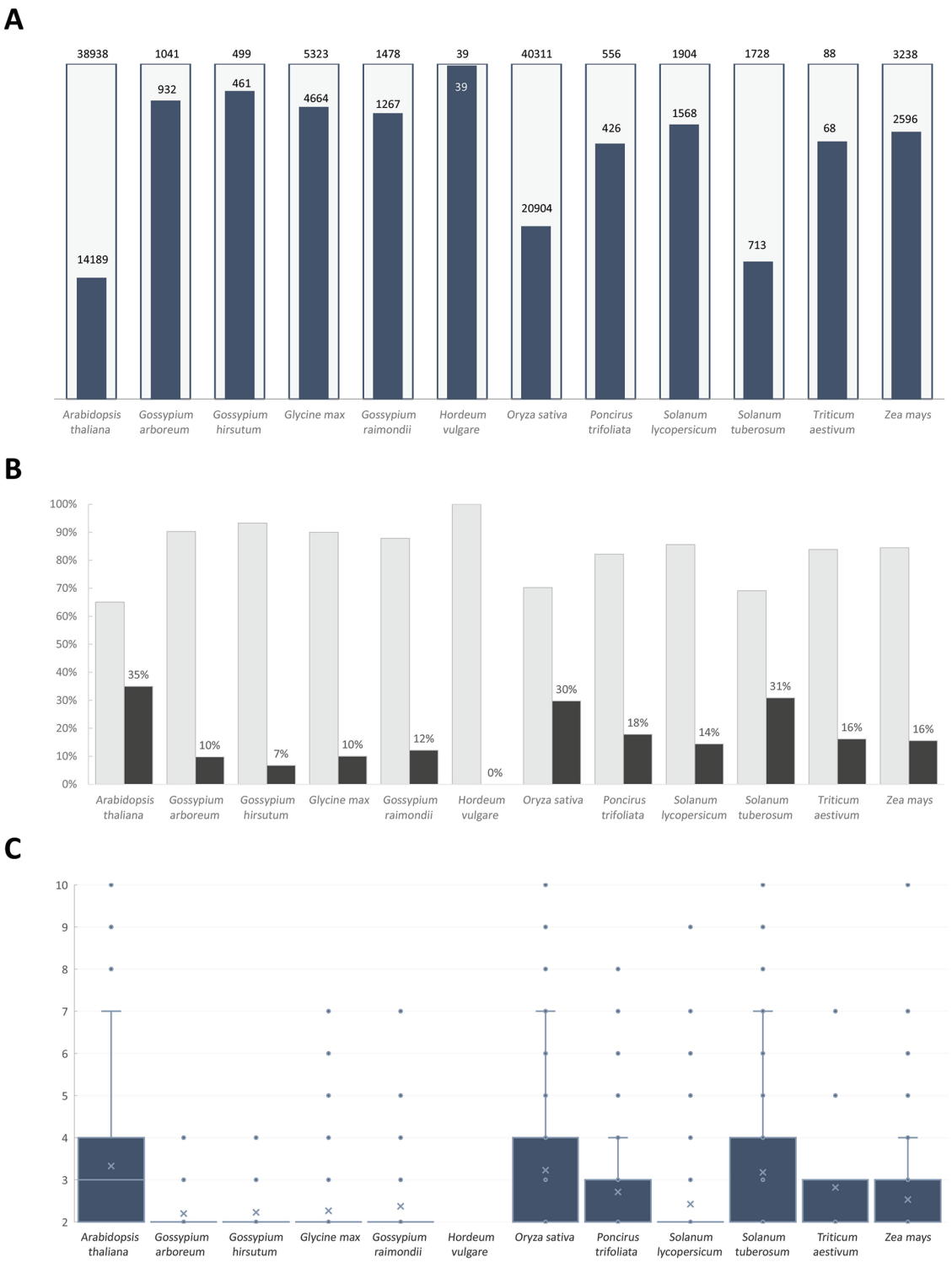

Figure 2. Alternative back-splicing events in plant circRNAs. **(A)** Number of circRNA loci in each plant species. External bars represent the total number of identified circRNAs (circRNA isoforms) and internal bars represent the number of circRNA loci. **(B)** Percentage of circRNA loci that generate single isoform (light gray bars) and multiple isoforms (black bars) in each species. **(C)** Number of circRNA isoforms generated by loci with alternative back-splicing events. The cross marks represent average values of isoform numbers. Dots represent discrete points while drawing boxplots. It is noteworthy that isoform numbers larger than 10 in *A. thaliana* (281 loci), *Glycine max* (two loci, 44 and 29 isoforms, respectively), *O. sativa* (295 loci), *Solanum lycopersicum* (one locus, 17 isoforms), *S. tuberosum* (17 loci) and *Z. mays* (two loci, 20 and 17 isoforms, respectively) were not drawn in this figure.

In order to systematically investigate the pattern of splicing signals in plant circRNAs, we identified circRNAs with all types of splicing signals (permutation and combination of four nucleobases) in *O. sativa* and *A. thaliana* using CIRI [16] with our modifications (see details in Supplementary data). The results (Figure 3B) showed that high confident circRNAs with

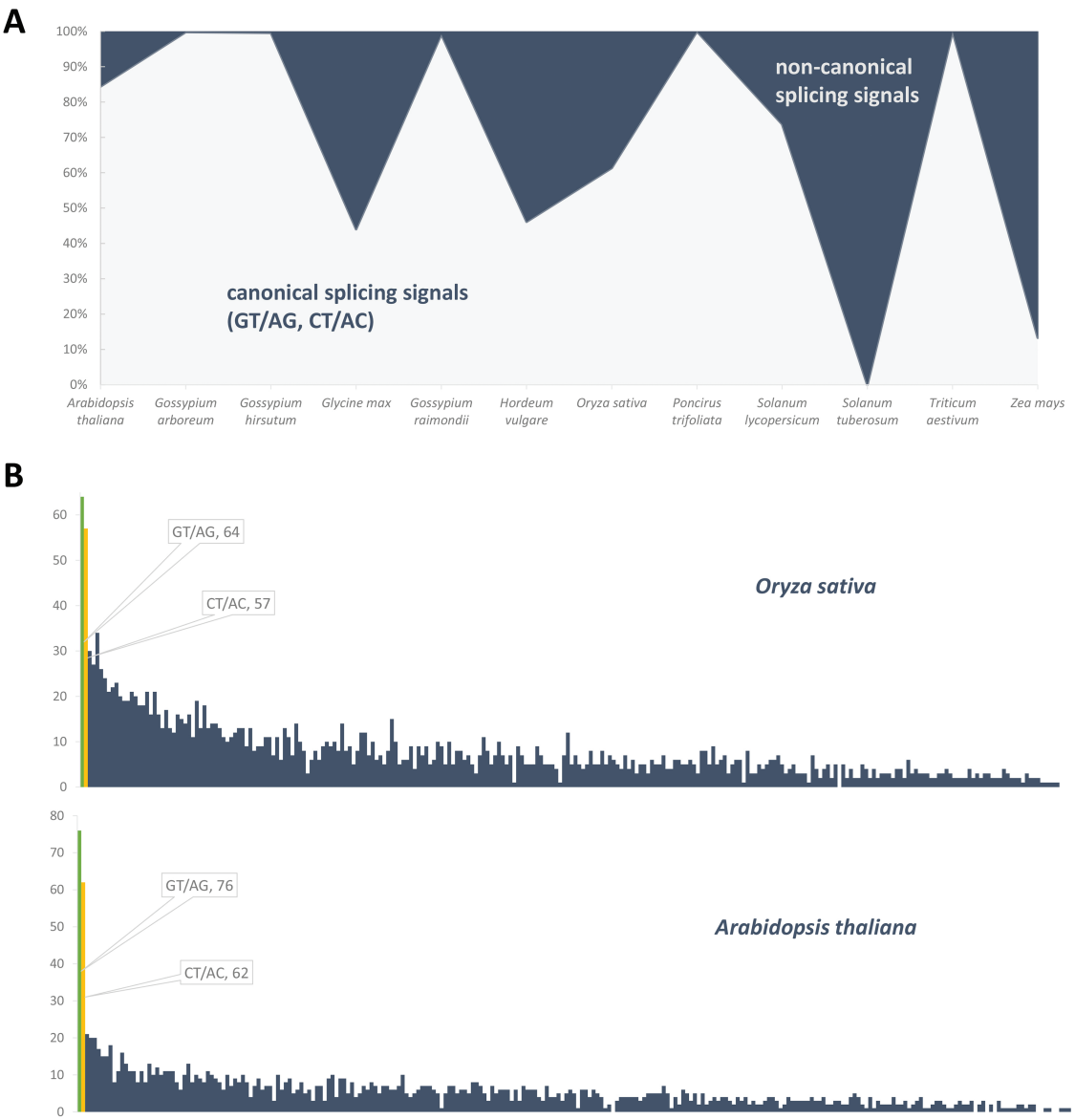

Figure 3. Splicing signals of plant circRNAs. **(A)** Percentage of canonical (light gray area) and non-canonical (dark blue area) splicing signals in plant circRNAs collected from publications (95 143 circRNAs from 12 plant species were used). **(B)** Distribution of diverse splicing signals of circRNAs in *O. sativa* and *A. thaliana* identified by the modified CIRI. RNA-Seq data used in this analysis were SRR1618548 (*O. sativa*) and SRR5591021 (*A. thaliana*). Horizontal and vertical axes represent different types of splicing signals and number of identified circRNAs, respectively. Bars filled with dark blue represent circRNAs with relatively higher confidence, and canonical splicing signals were highlighted in green and yellow.

the canonical splicing signals were not in the majority (only account for 6% and 9% of all identified circRNAs in *O. sativa* and *A. thaliana*, respectively), indicating that non-canonical splicing mechanisms might be involved in biosynthesis of plant circRNAs. This is consistent with our previous result [31] obtained in rice using an independent circRNA identification tool named circseq_cup.

Potential circRNA–miRNA–mRNA networks in plants

To investigate the potential functions of plant circRNAs, competing endogenous RNA networks associated with circRNAs have been previously predicted, for example the circRNA–miRNA interaction network (thousands of circRNAs and 92 miRNAs) in soybean [29], the circRNA-centered subnetwork among co-expressed circRNAs, lincRNAs and coding genes in potato [36], the possible networks among differentially expressed circRNAs, coding genes and miRNAs in tomato [35] and the circRNA–miRNA–mRNA networks at different developmental stages in *A. thaliana* [23, 25, 26]. Due to the lack of experimental evidences, majority of these predicted networks were insufficiently credible. However, a recent study in mice [54] has successfully elucidated a sophisticated non-coding RNA regulatory network involving an lncRNA, a circRNA and two miRNAs by gene editing, provided compelling proof for the significant regulatory function by networks involving circRNAs.

Here, we estimated the interactions between circRNAs and miRNAs (circRNA serves as miRNA sponge), miRNAs and mRNAs

Table 2. Statistics of potential networks involving circRNAs, miRNAs and parental mRNAs

Organisms	Networks	circRNAs	miRNAs	mRNAs
A. thaliana	8322/59	31 507	183	9369
G. arboretum	599/0	682	0	599
G. hirsutum	255/5	277	5	255
G. max	3736/50	5126	117	3 876
G. raimondii	1089/25	1322	34	1088
H. vulgare	31/1	31	1	31
O. sativa	10 448/90	35 088	272	11 177
P. trifoliata	236/2	272	2	235
S. lycopersicum	989/15	1245	16	985
S. tuberosum	249/14	915	97	260
T. aestivum	19/1	26	15	18
Z. mays	2208/26	3126	26	2209
Total	28 181/288	79 617	758	30 102

Note: X/Y in the column of 'Networks' represents X networks only involving circRNAs and their parental genes, and Y networks involving circRNAs, miRNAs, and mRNAs. Numbers in the columns of 'circRNAs', 'miRNAs' and 'mRNAs' represent the number of RNAs involved in the interaction networks.

(miRNA targets mRNA) and the relationships between circRNAs and their parental genes (details in Supplementary data), and constructed a total of 288 potential circRNA–miRNA–mRNA networks in the 12 plant species (Table 2). Most predicted networks (over 98%) only contained circRNAs and their parental genes, suggesting that miRNA sponges might not be the major functional mechanism of plant circRNAs.

Release 3.0 of PlantcircBase

A functional database that integrates circRNA repertories and visualization tools would significantly facilitate studies on circRNAs. To this end, diverse circRNA databases have been established, including extensively used comprehensive databases for animals such as circBase [2], CIRCpedia [14], circRNAdb [3] as well as plant circRNA databases PlantcircBase [55] and AtCircDB [56]. AtCircDB is a tissue-specific circRNA database for *A. thaliana*. PlantcircBase collected all plant circRNAs from publicly available publications by us. In order to keep up with the rapid progress of plant circRNA research and provide a more comprehensive and user-friendly online resource for researchers, we have further upgraded the database.

The current version (Release 3.0) of PlantcircBase not only is a repository of plant circRNAs that collected a total of 95 143 circRNAs from 12 plant species (through June 2018), but also has evolved as a comprehensive database with a set of bioinformatics tools to serve the needs of circRNA investigations, including browse, search, visualization and BLASTcirc. In the section of browse, circRNAs from each species were listed in tables with general information of each circRNA. By clicking the name of a circRNA, a page with more detailed circRNA information will be opened, including normal information such as name, organism, aliases, genomic position, etc. and advanced information such as sequence at back-splicing site, coding capacity (details in Supplementary data), potential amino acid sequence and other characteristic information of the circRNA. In the section of search, collected circRNAs can be searched by keywords such as the name of circRNA or its parental gene, or by sequences. In the section of visualization, genomic structures of circRNAs could be drawn based on the genomic positions of those circRNAs. BLASTcirc is a web-based tool for determining whether the query sequences contain back-splicing sites based on basic local alignment, provide the statistical significance of the predicted circRNAs using a global alignment algorithm and optimize a composite score for each alignment by taking into account all potential genomic locations of back-splicing sites (details in Supplementary data). In addition, genome browse, database statistics, download and visualization of networks involving circRNAs, miRNAs and mRNAs, and so on, are all available at PlantcircBase 3.0, which could be found at http://ibi.zju.edu.cn/plantcircbase/.

Future perspectives

False positive circRNAs identified by bioinformatics tools

For the huge number of circRNAs identified in diverse species, there is no doubt that some of them are not genuine circRNAs. Previous review [21] proposed that both experimental strategies for RNA-Seq and bioinformatics tools for identification could cause false positive while detecting circRNAs. For instance, while preparing RNA-Seq libraries, two RNAs containing common sequences could be joined together by reverse transcriptase, or two cDNAs could be ligated during adapter ligation, both might generate non-negligible circRNA artifacts. When detecting circRNA candidates by bioinformatics tools, false alignments could happen at potential back-splicing sites due to homologous sequences and/or compatibility of several unaligned bases. It is a tremendous challenge to overcome these problems for both experimental and bioinformatics researchers when identifying circRNAs.

In plants, progress has been made to predict more reliable circRNAs by developing a plant-specific circRNA prediction tool, PcircRNA_finder, based on specific characteristics of plant genomes [57]. Although PcircRNA_finder provided a more precise and sensitive prediction result in plants compared to other methods [57], it still has some shortcomings, such as using multiple software for catching back-splicing sites (which is not convenient for users) and only exonic circRNAs could be found (as genome annotations of some plants are very poor). It is thus urgently required to further develop more tools for circRNA prediction in plants because, except PcircRNA_finder, there is no other alternative bioinformatics tool specially designed for analyzing plant circRNAs. We hope that the characteristics of plant circRNAs reported in this study will provide clues for developing such tools.

Non-canonical splicing signals of plant circRNAs

Candidate circRNAs with non-canonical splicing signals (by U12 or minor spliceosome) used to be considered as false positive,

and were often filtered out by most prediction tools. Using a statistically based tool, KNIFE [17], which was thought to be more reliable than other circRNA identification tools that predict circRNAs based on only back-splicing reads [17, 58], researches have demonstrated the presence of circRNAs with non-canonical splicing signals in humans. In plants, a number of non-canonical circRNAs have also been identified and some were validated experimentally [31, 33], providing evidences for the existence of non-canonical circRNAs in plants. The flanking sequences of humans and plant circRNAs have different characteristics. Most plant circRNAs were not flanked by long repeat sequences or reverse complementary sequences [11, 12], but flanked by short reverse complementary sequences or repeat sequences. It is unclear whether these short sequences are related to biogenesis of plant circRNAs [27, 38] and the formation of non-canonical circRNAs. In view the number of non-canonical circRNAs in plants, non-canonical spliceosome mechanisms could be the major pathway involved in biogenesis circRNAs in plants, but the exact mechanism is yet to be uncovered.

Full-length/internal structures of plant circRNAs

Full-length sequences of circRNAs are essential for understanding the internal structures and alternative splicing events of circRNAs, and will be helpful for investigating the potential functions of circRNAs. However, the vast majority of the reported circRNAs do not have their full-length sequences because most circRNA prediction tools used only reads that support back-splicing sites. Despite the importance of full-length circRNAs, studies on strategies and methodology of full-length circRNA assembly are still rare in both animals and plants. Two tools, CIRI-full [59] and circseq_cup [31], are now available for prediction of full-length circRNAs. Both use RNA-Seq reads across back-splicing sites and their corresponding paired reads. Based on the assembly principle of CIRI-full and circseq_cup, theoretically if there are long-enough RNA-Seq reads and large-enough insert length, circRNAs of any lengths could be assembled. Apart from developing novel tools for full-length circRNA prediction, improving genome assembly of plant species is also critical because the quality of genome annotation determines the accuracy of read alignment and finally the accuracy of the assembled full-length circRNAs [60].

Functional investigations of plant circRNAs

Functional characterization of circRNAs is still in its infancy stage. Even for circRNAs from humans and animals, only a few have been experimentally proved to be functional, such as affecting splicing of their parental genes, regulating gene transcription, acting as miRNA sponges, being translated into polypeptides or working as biomarkers (reviewed by [40]). In plants, functional mechanism of only one circRNA from *Arabidopsis* has been revealed [46], although a large number of plant circRNA have been identified. The majority of plant circRNAs were generated from protein coding genes and some of them are conserved among different species. We expect that at least the conserved circRNAs if not all would be functional and play a certain role in plant development and/or response to environmental changes. One of the great challenges ahead for experimental functional exploration of plant circRNAs is the difficulty to distinguish circular transcripts from their linear parental mRNAs due to their overlapping sequences, which makes it difficult to knock down circRNAs by the traditional RNAi approach. Gene editing would be the choice for functional investigation of circRNAs because it requires only a 20 nt long specific target sequence.

Authors' contributions

C.Y., L.F. and Q.C. conceived the structures of this manuscript. Q.C. wrote the original manuscript. Q.C., P.B., X.T.Z., X.C.Z. and L.M. collected circRNA information. Q.C. and P.B. analyzed the characteristics of circRNAs. Q.C., P.B. and L.M. analyzed the networks involved circRNAs. X.C.Z. established the algorithm of BLASTcirc. Q.C., Q.H.Z., L.F. and C.Y. revised the manuscript. All of the authors read and approved the manuscript.

Key Points

- A total of 95 143 circRNAs were identified in 12 plant species.
- Most plant circRNAs were generated from coding genes.
- Alternative splicing is pervasive in plant circRNAs.
- A large number of circRNAs possess non-canonical splicing signals.

Supplementary Data

Supplementary data are available online at https://academic.oup.com/bib.

Funding

National Science Foundation of China (91740108 to L.F.); the Natural Science Foundation of Zhejiang Province (LY17C130004 to C.Y.); and the Jiangsu Collaborative Innovation Center for Modern Crop Production (JCIC-MCP).

References

1. Sanger HL, Klotz G, Riesner D, *et al*. Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures. *Proc Natl Acad Sci* 1976;**73**(11):3852-6.
2. Glazar P, Papavasileiou P, Rajewsky N, *et al*. circBase: a database for circular RNAs. *RNA* 2014;**20**:1666–70.
3. Chen X, Han P, Zhou T, *et al*. circRNADb: a comprehensive database for human circular RNAs with protein-coding annotations. *Sci Rep* 2016;**6**:34985.
4. Salzman J, Gawad C, Wang PL, *et al*. Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. *PLoS One* 2012;**7**(2):e30733.
5. Westholm JO, Miura P, Olson S, *et al*. Genome-wide analysis of Drosophila circular RNAs reveals their structural and sequence properties and age-dependent neural accumulation. *Cell Rep* 2014;**9**:1966–81.
6. Memczak S, Jens M, Elefsinioti A, *et al*. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* 2013;**495**:333–8.
7. Ivanov A, Memczak S, Wyler E, *et al*. Analysis of intron sequences reveals hallmarks of circular RNA biogenesis in animals. *Cell Rep* 2015;**10**:170–7.
8. Fan X, Zhang X, Wu X, *et al*. Single-cell RNA-seq transcriptome analysis of linear and circular RNAs in mouse preimplantation embryos. *Genome Biol* 2015;**16**:148.

9. Liang G, Yang Y, Niu G, *et al*. Genome-wide profiling of Sus scrofa circular RNAs across nine organs and three developmental stages. *DNA Res* 2017;**24**(5):523-35.
10. Wang PL, Yun B, Muh-Ching Y, *et al*. Circular RNA is expressed across the eukaryotic tree of life. *PLoS One* 2014;**9**:e90859.
11. Ye CY, Chen L, Liu C, *et al*. Widespread noncoding circular RNAs in plants. *New Phytol* 2015;**208**:88–95.
12. Lu TT, Cui LL, Zhou Y, *et al*. Transcriptome-wide investigation of circular RNAs in rice. *RNA* 2015;**21**: 2076–87.
13. Zhang XO, Wang HB, Zhang Y, *et al*. Complementary sequence-mediated exon circularization. *Cell* 2014;**159**: 134–47.
14. Zhang XO, Dong R, Zhang Y, *et al*. Diverse alternative back-splicing and alternative splicing landscape of circular RNAs. *Genome Res* 2016;**26**:1277–87.
15. Gao Y, Wang J, Zhao F. CIRI: an efficient and unbiased algorithm for de novo circular RNA identification. *Genome Biol* 2015;**16**:4.
16. Gao Y, Zhang J, Zhao F. Circular RNA identification based on multiple seed matching. *Brief Bioinform* 2018;**19**(5):803-10.
17. Szabo L, Morey R, Palpant NJ, *et al*. Statistically based splicing detection reveals neural enrichment and tissue-specific induction of circular RNA during human fetal development. *Genome Biol* 2015;**16**:126.
18. Hoffmann S, Otto C, Doose G, *et al*. A multi-split mapping algorithm for circular RNA, splicing, trans-splicing and fusion detection. *Genome Biol* 2014;**15**:R34.
19. Jeck WR, Sharpless NE. Detecting and characterizing circular RNAs. *Nat Biotechnol* 2014;**32**:453–61.
20. Chen I, Chen CY, Chuang TJ. Biogenesis, identification, and function of exonic circular RNAs. *Wiley Interdiscip Rev RNA* 2015;**6**:563–79.
21. Szabo L, Salzman J. Detecting circular RNAs: bioinformatic and experimental challenges. *Nat Rev Genet* 2016;**17**:679–92.
22. Zeng X, Lin W, Guo M, *et al*. A comprehensive overview and evaluation of circular RNA detection tools. *PLoS Comput Biol* 2017;**13**(6):e1005420.
23. Chen G, Cui J, Wang L, *et al*. Genome-wide identification of circular RNAs in Arabidopsis thaliana. *Front Plant Sci* 2017;**8**:1–12.
24. Dou Y, Li S, Yang W, *et al*. Genome-wide discovery of circular RNAs in the leaf and seedling tissues of Arabidopsis thaliana. *Curr Genomics* 2017;**18**(4):360-5.
25. Liu T, Zhang L, Chen G, *et al*. Identifying and characterizing the circular RNAs during the lifespan of Arabidopsis leaves. *Front Plant Sci* 2017;**8**:1–9.
26. Pan T, Sun X, Liu Y, *et al*. Heat stress alters genome-wide profiles of circular RNAs in Arabidopsis. *Plant Mol Biol* 2018;**96**:217–29.
27. Sun X, Wang L, Ding J, *et al*. Integrative analysis of Arabidopsis thaliana transcriptomics reveals intuitive splicing mechanism for circular RNA. *FEBS Lett* 2016;**590**:3510–6.
28. Zhao T, Wang L, Li S, *et al*. Characterization of conserved circular RNA in polyploid Gossypium species and their ancestors. *FEBS Lett* 2017;**591**:3660–9.
29. Zhao W, Cheng Y, Zhang C, *et al*. Genome-wide identification and characterization of circular RNAs by high throughput sequencing in soybean. *Sci Rep* 2017;**7**: 5636.
30. Darbani B, Noeparvar S, Borg S. Identification of circular RNAs from the parental genes involved in multiple aspects of cellular metabolism in barley. *Front Plant Sci* 2016;**7**:776.
31. Ye C-Y, Zhang X, Chu Q, *et al*. Full-length sequence assembly reveals circular RNAs with diverse non-GT/AG splicing signals in rice. *RNA Biol* 2017;**14**(8):1055–63.
32. Zeng RF, Zhou JJ, Hu CG, *et al*. Transcriptome-wide identification and functional prediction of novel and flowering-related circular RNAs from trifoliate orange (Poncirus trifoliata L. Raf.). *Planta* 2018;**247**: 1191–202.
33. Tan J, Zhou Z, Niu Y, *et al*. Identification and functional characterization of tomato circRNAs derived from genes involved in fruit pigment accumulation. *Sci Rep* 2017;**7**: 1–9.
34. Zuo J, Wang Q, Zhu B, *et al*. Deciphering the roles of circRNAs on chilling injury in tomato. *Biochem Biophys Res Commun* 2016;**479**:132–8.
35. Yin J, Liu M, Ma D, *et al*. Identification of circular RNAs and their targets during tomato fruit ripening. *Postharvest Biol Technol* 2018;**136**:90–8.
36. Zhou R, Zhu Y, Zhao J, *et al*. Transcriptome-wide identification and characterization of potato circular RNAs in response to pectobacterium carotovorum subspecies brasiliense infection. *Int J Mol Sci* 2018;**19**(1):71.
37. Wang Y, Yang M, Wei S, *et al*. Identification of circular RNAs and their targets in leaves of Triticum aestivum L. under dehydration stress. Front. *Plant Sci* 2017;**7**:2024.
38. Chen L, Zhang P, Fan Y, *et al*. Circular RNAs mediated by transposons are associated with transcriptomic and phenotypic variation in maize. *New Phytol* 2018;**217**(3):1292-306.
39. Chen LL. The biogenesis and emerging roles of circular RNAs. *Nat Rev Mol Cell Biol* 2016;**17**(4):205-11.
40. Li X, Yang L, Chen L-L. The Biogenesis, functions, and challenges of circular RNAs. *Mol Cell* 2018;**71**:428–42.
41. Hansen TB, Jensen TI, Clausen BH, *et al*. Natural RNA circles function as efficient microRNA sponges. *Nature* 2013; **495**:384–8.
42. Zhang Y, Zhang XO, Chen T, *et al*. Circular intronic long noncoding RNAs. *Mol Cell* 2013;**51**:792–806.
43. Li Z, Huang C, Bao C, *et al*. Exon-intron circular RNAs regulate transcription in the nucleus. *Nat Struct Mol Biol* 2015;**22**: 256–64.
44. Ashwal-Fluss R, Meyer M, Pamudurti NR, *et al*. CircRNA biogenesis competes with pre-mRNA splicing. *Mol Cell* 2014;**56**:55–66.
45. Kelly S, Greenman C, Cook PR, *et al*. Exon skipping is correlated with exon circularization. *J Mol Biol* 2015;**427**(15): 2414-17.
46. Conn VM, Hugouvieux V, Nayak A, *et al*. A circRNA from SEPALLATA3 regulates splicing of its cognate mRNA through R-loop formation. *Nat Plants* 2017;**3**:17053.
47. Cheng J, Zhang Y, Li Z, *et al*. A lariat-derived circular RNA is required for plant development in Arabidopsis. *Sci China Life Sci* 2018;**61**:204–13.
48. Black DL. Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem* 2003;**72**:291-336.
49. Gao Y, Wang J, Zheng Y, *et al*. Comprehensive identification of internal structure and alternative splicing events in circular RNAs. *Nat Commun* 2016;**7**:12060.
50. Black DL. Protein diversity from alternative splicing: a challenge for bioinformatics and post-genome biology. *Cell* 2000;**103**(3):367-70.
51. Graveley BR. Alternative splicing: increasing diversity in the proteomic world. *Trends Genet* 2001;**17**(2):100-7.
52. Will CL, Lührmann R. Spliceosome structure and function. *Biol Cold Spring Harb Perspect* 2011;**3**(7):a003707.

53. Gao Y, Zhao F. Computational strategies for exploring circular RNAs. *Trends Genet* 2018;**34**:389–400.
54. Kleaveland B, Shi CY, Stefano J, *et al*. A network of noncoding regulatory RNAs acts in the mammalian brain. *Cell* 2018;**174**(2):350-62 e17.
55. Chu Q, Zhang X, Zhu X, *et al*. PlantcircBase: a database for plant circular RNAs. *Mol Plant* 2017;**10**:1126-8.
56. Ye J, Wang L, Li S, *et al*. AtCircDB: a tissue-specific database for Arabidopsis circular RNAs. *Brief Bioinform* 2017;1–8.
57. Chen L, Yu Y, Zhang X, *et al*. PcircRNA_finder: a software for circRNA prediction in plants. *Bioinformatics* 2016;**32**(22): 3528-9.
58. Szabo L, Salzman J. Detecting circular RNAs: bioinformatic and experimental challenges. *Nat Rev Genet* 2016;**17**:679–92.
59. Zheng Y, Zhao F. Detection and reconstruction of circular RNAs from transcriptomic data. *Circ RNAs* 1724;**2018**:8–12.
60. Chu Q, Shen E, Ye C-Y, *et al*. Emerging Roles of Plant Circular RNAs. *J Plant Cel* 2018;**1**:1–14.