Send Orders for Reprints to reprints@benthamscience.ae

 Current Genomics, 2017, 18, 360-365

REVIEW ARTICLE

Genome-wide Discovery of Circular RNAs in the Leaf and Seedling Tissues of *Arabidopsis Thaliana*

Yongchao Dou[1,#], Shengjun Li[1,#], Weilong Yang[1], Kan Liu[1], Qian Du[1], Guodong Ren[2,*], Bin Yu[1,*] and Chi Zhang[1,*]

[1]*School of Biological Sciences, University of Nebraska, Lincoln, Nebraska, USA;* [2]*State Key Laboratory of Genetic Engineering, Institute of Plant Biology and Department of Biochemistry, School of Life Sciences, Fudan University, Shanghai, China*

ARTICLE HISTORY

Received: September 23, 2016
Revised: October 12, 2016
Accepted: October 27, 2016

DOI:
10.2174/1389202918666170307161124

Abstract: ***Background:*** Recently, identification and functional studies of circular RNAs, a type of non-coding RNAs arising from a ligation of 3' and 5' ends of a linear RNA molecule, were conducted in mammalian cells with the development of RNA-seq technology.

Method: Since compared with animals, studies on circular RNAs in plants are less thorough, a genome-wide identification of circular RNA candidates in Arabidopsis was conducted with our own developed bioinformatics tool to several existing RNA-seq datasets specifically for non-coding RNAs.

Results: A total of 164 circular RNA candidates were identified from RNA-seq data, and 4 circular RNA transcripts, including both exonic and intronic circular RNAs, were experimentally validated. Interestingly, our results show that circular RNA transcripts are enriched in the photosynthesis system for the leaf tissue and correlated to the higher expression levels of their parent genes. Sixteen out of all 40 genes that have circular RNA candidates are related to the photosynthesis system, and out of the total 146 exonic circular RNA candidates, 63 are found in chloroplast.

Keywords: Circular RNAs, Non-coding RNAs, RNA-seq, Bioinformatics of non-coding RNA identification, Next-generation sequencing data analysis.

1. INTRODUCTION

Non-coding RNAs (ncRNAs) have diverse types and are functionally important [1, 2]. Some types of ncRNAs were widely studied in the recent years, such as small interfering RNAs and long noncoding RNAs [3, 4], all of which are linear RNAs. Interestingly, RNAs also exist in circular forms, which can arise from a ligation of 3' and 5' ends of a linear RNA molecule or from splicing intermediates of introns. A viroid circular RNA was first reported in 1976 [5] whereas an eukaryotic circular RNA was initially observed by electron microscopy in 1979 [6]. After that, some circular RNAs have been reported in various species. For example, similar to viroid, hepatitis D virus also has a circular RNA [7]. In archaea, circular RNAs can be formed from tRNA and rRNA introns, as well as from many other functionally unknown RNAs [8-13]. Circular RNAs exist in many eukaryotic organisms. For example, they were reported in algae [14] and yeast [15, 16], respectively. In Drosophila, muscleblind gene was found to have a circularized transcript [17]; and many circular RNAs were identified in *C. elegans* recently [18]. Furthermore, many circular RNAs were discovered in mammals. For example, the SRY gene [19] and FORMIN (FMN) gene [20] in mouse, the P450 2C24 gene [21] and the sodium-calcium exchanger NCXI [22] in rat, were reported to have circularized transcripts. Even more genes in the human genome have been reported to have circular RNAs, such as the ETS-1 gene [21], the Cytochrome P-450 2C18 gene [21], the dystrophin gene [22], the CDR-1 gene [23], and a long non-coding RNA (ncRNA) known as ANRIL [24]. Yet the number of reported circular RNAs has really exploded in recent years with high-throughput methods. Using the next-generation sequencing technology, circular RNAs can be identified genome widely and thousands of candidates have been reported in human, mouse, fruit fly, worm, and archaea [9, 25, 18, 26-28]. For example, Memczak *et al.* identified approximately 2000, 1900, and 700 circular RNAs in human, mouse, and nematode from RNA-seq data.

Circular RNAs are not short; usually, a circular RNA molecule is composed of up to 5 exons [18], and the exon regions that can be transcribed to circular RNAs are longer than the average expressed exons [25]. Circular RNAs often exist at low levels as compared to their linear counterparts [25], but a recent study with RNA-seq method showed that circular RNAs are broadly expressed and may have compa-

*Address correspondence to these authors at the School of Biological Sciences, University of Nebraska, Lincoln, Nebraska, USA; Tel/Fax: 402-472-8722; E-mail: byu3@unl.edu; czhang5@unl.edu and State Key Laboratory of Genetic Engineering, Institute of Plant Biology and Department of Biochemistry, School of Life Sciences, Fudan University, Shanghai, China; E-mail: gdren@fudan.edu.cn

# *These authors contributed equally to this work.*

1389-2029/17 $58.00+.00 ©2017 Bentham Science Publishers

rable expression levels with their linear isoforms in human [26]. A study suggested that the average ratio of expression levels between circular RNAs and their corresponding poly-A RNAs is about 1:100, and a small portion of genes even have more circular transcripts than linear isoforms [27]. Evidences show that circular RNAs, even transcribed from exons, are not poly-adenylated [26]. The DNA regions that can generate circular RNAs have different locations in the genomes, such as exon regions and intron regions. Most of them are found to be transcribed from exons, but a recent study reported the existence of many intronic circular RNAs [28]. Most circular RNAs transcribed from exons are cytoplasmic [26], while the intronic circular RNAs are mainly confined to the nucleus [28]. While the functions of most circular RNAs remain enigmatic, several of them, such as the circular RNAs from CDR-1 gene [23] and SRY gene [19], were found to work as efficient miRNA sponges [29, 18].

Plant circular RNAs have also attracted a lot of attention. Three exonic circular RNAs from Arabidopsis were originally reported, which are from exon 3 of NPY4, exons 6-7-8 of EMB2423, and exon 4-5 of CYP87A2 have been reported [30]. Following this study, additional plant circular RNAs have been identified from Arabidopsis, barley and rice [31-33]. We hypothesize that more exonic circular RNAs as well as intronic circular RNAs exist in plants including Arabidopsis. Therefore, we collected several suitable Arabidopsis RNA-seq libraries from the public domain, and identified several circular RNA candidates in both exonic and intronic regions by our own developed computational pipeline. Four computationally identified circular RNAs were then randomly selected for validation by PCR. Although plant circular RNAs have been identified, many related questions such as how plant circular RNAs are generated are not yet addressed. Our discovery of circular transcripts in Arabidopsis further proves the existence of circular RNAs in plants, and provides some clues about the biogenethesis of circular RNAs in plants.

2. MATERIALS AND METHOD

2.1. Circular RNA Identification Pipeline

A pipeline is constructed to identify circular RNAs with RNA-seq data. Circular RNAs have ligation points, and junction reads, short reads covering the ligation point of a circular RNA, can be split and aligned to two ends of the linear isoform with a special order, which is called back-splicing mapping or head-to-tail splicing mapping Fig. (**1**). Identification of circular RNA candidates is based on finding these junction-reads from back-splicing exons. In the first step of our pipeline, reads are mapped to the reference genome and all unmapped reads are collected. Pairs of anchors with the length of 20 bp are extracted from both 5' and 3' ends of each unmapped reads. These anchor pairs are mapped to the reference genome again, and all aligned positions are reported. If an anchor pair has any back-splicing mapping order, the mapping information related to the anchor pair is kept. Then, these anchors are extended until the complete read aligns. A qualified junction read is called back-splicing mapped read if it can be back-splicing mapped without no more than two mismatches. After these back-splicing mapped reads are identified, circular RNA candidates can be inferred from these mapped reads. A manual filtering step from Ref. [31], focusing on back-spliced from mis-assembly and repeats *etc*., was adopted to reduce false positive rate.

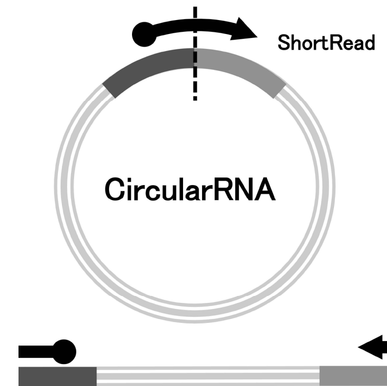

Fig. (1). A diagram to show a junction read from a back-spliced exon that is mapped to the linear reference sequence.

2.2. Intronic Circular RNA Identification

In principle, intronic circular RNAs can also be detectable by finding back-splicing mapped RNA-seq reads, but this is still an open question for intron lariat after splicing. According to our practice, only a few back-splice mapped reads were found for intronic circular RNAs, and could not be used for circular RNA identification. Therefore, all annotated intron regions were scanned and the introns with high coverage of reads were kept; high coverage indicates potential transcripts of the given introns. We also adapted a method described in Ref. [25] to identify intronic circular RNAs. An in-house script was used to find branch points on these introns. Introns with high mapped reads and branch points were treated as intronic circular RNA candidates.

2.3. Short-read Alignment

Bowtie2 (version 2.23) [34] was used to map short reads, anchors, or split reads to the reference genome. Two mismatches are allowed for reads mapping and perfect match was used for anchor mapping. Up to 40 mapped positions are allowed for the anchors mapping.

2.4. Genome Sequence and Annotations

Genome sequence and Annotations of Arabidopsis were obtained from TAIR database (http://www.arabidopsis.org/), and TAIR10 was used.

2.5. RNA-seq Data Sets

Three RNA-seq data sets were obtained from SRA, and their accession numbers are SRR352212, SRR504179, and SRR584135. The data of SRR352212 were generated to study noncoding RNAs in Arabidopsis [35]. The total RNA was depleted of rRNAs using the plant RiboMinus-Kit and 85bp strand-specific reads were sequenced. The RNA-seq data of SRR504179 were used to study long non-coding RNAs, and 55bp single end reads were sequenced [36]. The set of SRR584135 was sequenced to study regulation of alternative splicing, and 100bp single end reads were se-

quenced by Illumina Genome Analyzer II [37]. All of these data sets did not use poly-adenylation selection to process total RNA, and hence kept circular RNAs in the samples.

2.6. Validation of Circular RNA with RT-PCR Followed by Sequencing

Total RNA was isolated from inflorescence using Trizol reagent (Sigma). 5 µg total RNA was digested by 10 U RNase R (Epicenter) at 37°C for 2h. cDNA synthesis was carried out by SuperScript III (Invitrogen) with gene-specific forward primers. For detecting intronic and exonic circular RNAs, PCR amplification was performed with Ex Taq Hot Start polymerase (TaKaRa) using the primer pairs listed in the Supplementary Tables **S1**. Resulting PCR products were then cloned and sequenced. The design of a paired primer for intronic circular RNAs and exonic circular RNAs are shown in Fig. (**2A** and **B**). To have additional control, pairs of divergent and convergent primers on genomic DNAs were designed and PCR experiment was conducted. Fig. (**2D**, **E**, and **F**) shows the design of paired convergent primers and PCR results, and all primers for the control are shown in Supplementary Tables **2**.

3. RESULTS AND DISCUSSIONS

3.1. Circular RNAs Transcribed from Exon Regions are Identified in Arabidopsis

Our circular RNA identification method is applied to three RNA-seq libraries for Arabidopsis. The anchor length was set to 20bp with a minimum circular length of 100bp in finding full-length back-splicing mapped reads. The minimum number of junction reads from back-splicing exons was set to four for searching circular RNA candidates. With this setup, a number of 188 candidates were discovered, and a manual filtering step removed some candidates from repeats and mis-assembly. Most of the removed candidates were from chloroplast coded genes because of the low quality of chloroplast genome sequence. Finally, a total of 146 exonic circular RNA candidates were identified by our method. Compared with published results from human and mouse, a much smaller number of circular RNAs were identified by our method for Arabidopsis. This is likely due to temporal and spatial specific expression of circular RNAs [18]. In fact, in *C. elegans*, sequence analyses of RNAs from different developmental stages raise the number of circular RNAs from ~300 to ~1,100 [38]. In fact, five out of 188 candidates were found to be co-exist in two or three libraries that are

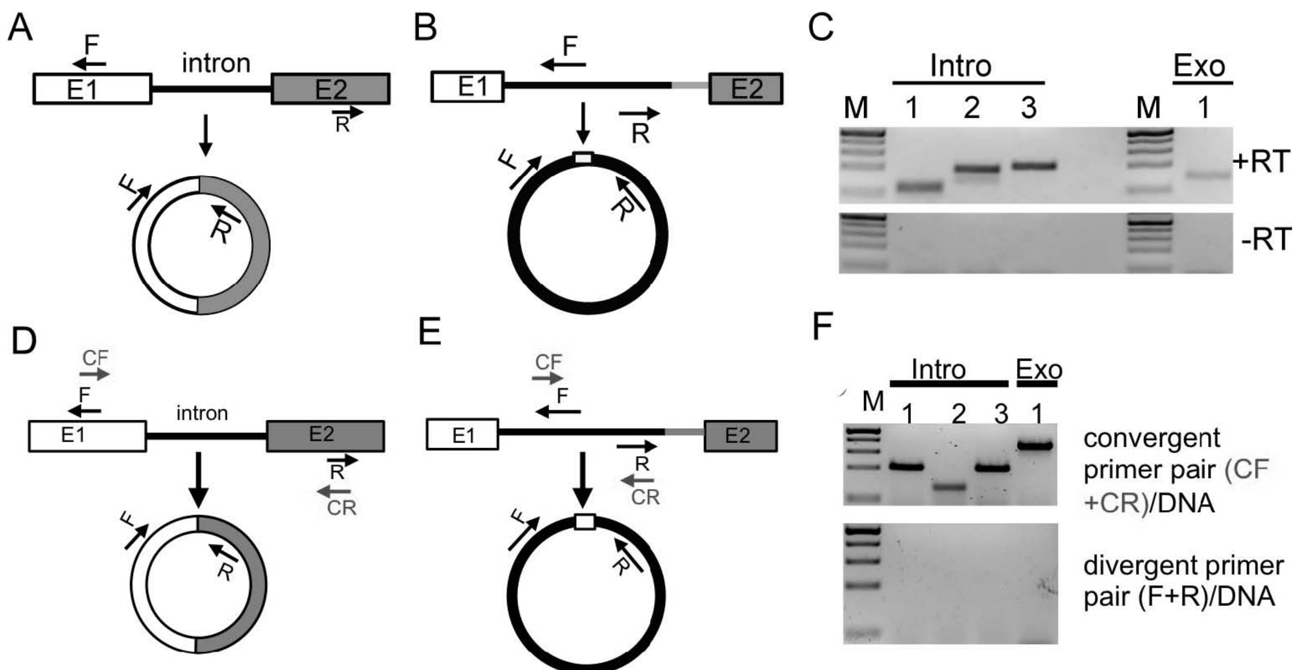

Fig. (2). Validation of predicted circular RNAs by RT-PCR. (**A**) A diagram showing the position of forward primer (**F**) and reverse primer (**R**) used for PCR amplification of an exonic circular RNA. E1: Exon 1; E2: Exon 2. (**B**) A diagram showing the position of forward primer (**F**) and reverse primer (**R**) used for detection of intronic circular RNAs. E1 and E2 indicate exons. The line (black and grey) indicates intron. The grey line indicates the portion of intron after the branch point. The open box in the circular RNAs indicates the ligation point. (**C**) RT-PCR detection of exonic RNA (EXO) and intronic RNAs (Intro). No RT (-RT) serves as negative controls. M: DNA size marker. RNAs extracted from inflorescences were treated with RNAse R and used as templates for reverse transcription using random primers. Exo 1 refers to an exonic circular RNA generated from AT2G30570. Intro 1, 2 and 3 refer to intronic circular RNAs generated from AT2G30570, AT1G31330 and AT5G35970, respectively. The resulting products were then used as templates for PCR amplification using primers indicated in (**A**) and (**B**). (**D**) and (**E**) The diagram showing the position of convergent primer pair (CF: forward; CR: reverse) and divergent primer pair (F: forward; R: reverse) used for PCR amplification of intronic and exonic circular RNAs. (**F**) PCR amplification of genomic DNAs from DNA loci generating EXO or Intro RNAs using convergent (Top panel) or divergent primer pairs (Bottom panel). Intro 1, 2 and 3 refer to the DNA fragment within At2g30570, At1g31330 and At5g35970, respectively. Exo 1 refers the DNA fragment within from At2g30570.

prepared under different condition. Another reason is that the RNA-seq libraries that we used were not specially prepared for circular RNA. All exonic circular RNA candidates, listed in Supplementary Tables **S3**, come from only 40 genes, and many genes have multiple circular transcripts. For example, the circular RNA candidate from the exon of the gene AT1G79930 (HSP91), which encodes a heat shock protein, has been identified in all three samples. We selected one circular RNA candidate from the PSBW gene (AT2G30570), which has two exons and codes a protein similar to photosystem II reaction center subunit W, to conduct experimental validation. The total RNAs from in florescence were extracted to test the circular transcripts, and the primers that we used for PCR are shown in the supplementary documents. This circular RNA was validated by RT-PCR using RNase R (degrading linear RNAs)-treated RNAs as templates (Fig. **2C**). Sequencing of the resulting products revealed that this circular RNA is 159bp long, and consists of sequences from two exons in the PSBW gene. Fig. (**3**) shows the location of four back-splicing mapped reads in this gene, which defines the range of this circular RNA. Ye *et al.*'s work reported 889 circular RNA candidates with more than five back-splicing mapped RNA-seq reads in Arabidopsis [41]. Comparing our results with 889 circular RNA candidates, we found 13 candidates appearing in both sets. The total number of our identified circular RNA candidates is less than Ye *et al.*'s work, and there is a small portion of overlap between two works. The possible reasons are the low sequencing depth and non-specifically designed RNA-seq libraries for our method. The studied tissues are also not identical.

Next, we explored the potential function for plant circular RNAs through bioinformatics analyses. In mammalian cells, some circular RNAs were found having many miRNA binding sites and working as efficient miRNA sponges [18, 29]. Therefore, the miRNA binding site prediction tool, psRNA-Target [39], was employed with the parameters described by previous works [40-42] to all predicted circular RNA candidates. Only seven out of 188 candidates have more than one putative miRNA binding sites, but all of these seven circular RNAs have multiple miRNA binding sites from two to 12. The genes covered by these seven circular RNAs are various, including D1 subunit of photosystem I reaction center, ubiquitin-conjugating enzyme 11 (UBC11), INDOLE-3-ACETIC ACID INDUCIBLE 2 (AT3G23030), and so on. Thus, it is tempting to speculate that, like animal circular RNAs, some of plant circular RNAs may function as the sponge of miRNAs to regulate the gene expression as well. On the other hand, most of plant circular RNAs do not contain any miRNA-binding sites. How do they function? It is possible that these circular RNAs may bind protein or other molecules to block or facilitate their functions. Clearly, these possibilities need further investigation.

3.2. Intronic Circular RNAs are Identified in Arabidopsis

Subsequently, we sought to identify intronic circular RNAs. After applying our method to the three RNA-seq libraries, no intronic circular RNAs are identified. The reason could be the low expression level of intronic circular RNAs in these libraries. After all, these libraries were not sequenced specifically for circular RNAs, and reads for linear RNAs dominate the whole libraries. To identify the intronic circular RNAs, we simply look for the intronic regions that have significantly high numbers of reads compared with other intronic regions as candidates, since the linear introns are likely degraded [25]. To further narrow down the candidates, those candidates having reads containing exon and intron junctions are further removed. After this step, a total of 18 high-quality candidates were identified, and all of them have relatively high expression levels in all three libraries. The detailed information of these 18 candidates was listed in the (Supplementary Tables **S4**). These intronic circular transcripts are more like a lariat RNA species, and the sequence of this circular transcript starts with "GU" and stops at "A", which is a typical branchpoint nucleotide. Three candidates were validated by RT-PCR followed by sequencing using RNase R-treated RNAs as templates, shown in Fig. (**2C**). These three circular RNAs are processed from transcripts of AT1G31330 (PSAF) encoding subunit F of photosystem I, AT5G35790 (G6PD1) encoding a plastidic glucose-6-phosphate dehydrogenase and PSBW. The introns in these three genes range from one to seven, suggesting that the generation of lariat circular RNAs is not affected by the intron numbers. Interestingly, both exonic and intronic circular RNAs are found for PSBW. (Fig. **3**) shows the profiles of aligned reads in this gene, and there is a peak in the intronic region as well. Both circular transcripts were found in the same RNA-seq library (accession number: SRR352212), suggesting that they may be co-processed from the pre-mRNAs. The biological relevance between these two different types of circular transcripts is not clear yet.

3.3. Circular RNA Biogenesis

Circular RNA biogenesis has been studied in animals, especially in human, but in plants it is just at the outset. Likely in animals, most exonic plant circular RNAs are ligation products of partial exons. Furthermore, Zhang *et al.* discovered

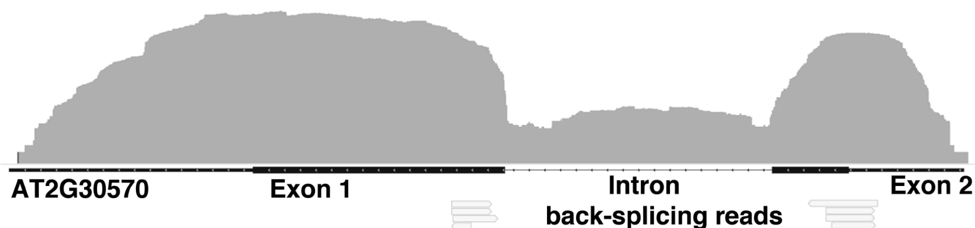

Fig. (3). The profiles of regularly and back-splicing aligned reads to AT2G30570. Three peaks above the gene annotation shows the profiles of linearly aligned short reads to the gene of AT2G30570, and there is one peak in the intron region. Four pairs of back-splicing mapped reads are shown under the gene annotation.

that, in human, exon circularization of RNAs is dependent on flanking intronic complementary sequences, and *Alu* repeat pairs can lead to alternative circularization in the same genes [43]. To test this, we collected up-stream and down-stream up-to-1kbp DNA sequences plus 30bp-end-sequences of all circular RNA candidates, and looked for complementary sequences across these sequences. We used RepeatMasker to find any repeats in the selected regions, and compared the sequence similarity between forward sequences on one side and complementary sequence on the other side of the target circular RNA regions to identify the potential complementary sequences. Interestingly, neither complementary sequences nor repeat sequences were reported in those areas. This result suggests that plant circular RNA biogenesis may require the assistance from protein factors, which may recognize specific sequence motif to forming circular RNAs in plants. However, we cannot rule out the possibility that some plant circular RNAs contain complementary sequences in their flanking regions, since the numbers of plant circular RNAs are not saturated. In addition, the fact that both exonic circular transcript and lariat transcript were detected in the same PSBW (AT2G30570) gene suggests that the intron splicing process might help the exon circulation.

3.4. Circular RNA Transcripts are Enriched in Photosynthesis System

Sixteen out of all 40 genes that have circular RNA candidates are related to the photosynthesis system, and out of the total 164 circular RNA candidates, 63 are found in chloroplast. Especially, in some regions of chloroplast, many circular RNAs are abundant and some of them are overlapped. For example, the gene ATCG00020 in chloroplast to encode a part of the photosystem II reaction center core has about 30 circular RNA candidates, and some of them were found by Ye *et al.* [33]. Interestingly, two validated circular RNA transcripts come from the PSBW (AT2G30570) gene and PSAF (AT1G31330) gene, receptively, which are also related to the photosynthesis system (both encoded subunits of photosystem reaction centers). The circular RNA transcripts relevant to the photosynthesis system were discovered in the dataset obtained from leaf tissue, and the circular RNA candidates are correlated to the higher expression levels of their corresponding genes. The biological significance of this result needs further investigation.

3.5. Plant Circular RNAs are not Poly-adenylated

Previous works indicated that circular RNAs are not poly-adenylated [26], which could be a result of their circulated structures. Therefore, we applied the circular RNA identification pipeline to some RNA-seq libraries, SRX159014, SRX150070, and SRX129192, that were prepared with poly-A(+) selection method to enrich mRNA. Not surprisingly, no circular RNA candidates were identified in these libraries. This result suggests that, like in animals [26], plant circular RNAs are not poly-adenylated.

3.6. Results of Our Pipeline on Published Data Sets

Our circular RNA identification method was applied on several published data sets to compare with the published results. The first data set includes *H. sapiens*, *C. elegans*, and mouse, which were published by Memczak *et al.* [18], and the other data set were the one used in the reference [26]. With the same setup for plant RNA-seq data, more than 50% of the outputs from our method were same to the reported results from the original work. For more details, the comparison between the original results and the results reported by our tool on the human cd19 cell data set, obtained from [18], is shown in Supplementary Tables **S5**. For this dataset, 587 circular RNA candidates were identified by our method with more than five back-splicing mapped RNA-seq reads. Out of the 587 candidates, 215 could be found from the originally reported results, and 58 reported circular RNAs could not be predicted by our method. The differences could be because our method was optimized for plants, but the low coverage of our datasets causes more false negatives predictions [44, 45].

CONCLUSION

We conducted a genome-wide identification for circular RNAs in Arabidopsis with computational methods from existing RNA-seq datasets. A total of 164 circular RNA candidates were identified, and four circular RNA transcripts, including both exonic and intronic circular RNAs, were experimentally validated. Interestingly, our results show that circular RNA transcripts are enriched in the photosynthesis system. For example, both PSBW and PSAF genes, having validated circular RNA transcripts, encode subunits of photosystem reaction center.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

YD developed the software tools and conducted data analysis. SL designed and conducted the experiment validation. WY, KL, and QD conducted bioinformatics analysis. GR, BY, and CZ initialized the project, supervised the research, and drafted the manuscript.

This work was completed utilizing the Holland Computing Center of the University of Nebraska. This project was supported by the NE Soybean Board funds.

SUPPLEMENTARY MATERIAL

Supplementary material is available on the publisher's web site along with the published article.

REFERENCES

[1] Eddy, S.R. Non-coding RNA genes and the modern RNA world. *Nat. Rev. Genet.*, **2001**, *2*(12), 919-929. DOI: 10.1038/35103511.

[2] Mattick, J.S. Makunin IV Non-coding RNA. *Hum. Mol. Genet.*, **2006**, 15 Spec No 1R17-29. DOI: 10.1093/hmg/ddl046.

[3] Batista, P.J.; Chang, H.Y. Long noncoding RNAs: cellular address codes in development and disease. *Cell*, **2013**, *152*(6), 1298-1307. DOI: 10.1016/j.cell.2013.02.012.

[4] Stefani, G.; Slack, F.J. Small non-coding RNAs in animal development. *Nat. Rev. Mol. Cell Biol.*, **2008**, *9*(3), 219-230. DOI: 10.1038/nrm2347.

[5] Sanger, H.L.; Klotz, G.; Riesner, D.; Gross, H.J.; Kleinschmidt, A.K Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures. *Proc. Natl. Acad. Sci. USA*, **1976**, *73*(11), 3852-3856.

[6] Hsu, M.T.; Coca-Prados, M. Electron microscopic evidence for the

circular form of RNA in the cytoplasm of eukaryotic cells. *Nature*, **1979**, *280*(5720), 339-340.

[7] Kos, A.; Dijkema, R.; Arnberg, A.C.; van der Meide, P.H.; Schellekens, H. The hepatitis delta (delta) virus possesses a circular RNA. *Nature*, **1986**, *323*(6088), 558-60. DOI: 10.1038/323558a0.

[8] Dalgaard, J.Z.; Garrett, R.A. Protein-coding introns from the 23S rRNA-encoding gene form stable circles in the hyperthermophilic archaeon *Pyrobaculum organotrophum*. *Gene*, **1992**, *121*(1), 103-110.

[9] Danan, M.; Schwartz, S.; Edelheit, S.; Sorek, R. Transcriptome-wide discovery of circular RNAs in Archaea. *Nucleic Acids Res.*, **2012**, *40*(7), 3131-42. DOI: 10.1093/nar/gkr1009.

[10] Lykke-Andersen, J.; Aagaard, C.; Semionenkov, M.; Garrett, R.A. Archaeal introns: splicing, intercellular mobility and evolution. *Trends Biochem. Sci.*, **1997**, *22*(9), 326-331.

[11] Salgia, S.R.; Singh, S.K.; Gurha, P.; Gupta, R. Two reactions of *Haloferax volcanii* RNA splicing enzymes: joining of exons and circularization of introns. *RNA*, **2003**, *9*(3), 319-330.

[12] Starostina, N.G.; Marshburn, S.; Johnson, L.S.; Eddy, S.R.; Terns, R.M.; & Terns, M.P Circular box C/D RNAs in *Pyrococcus furiosus*. *Proc. Natl. Acad. Sci. USA*, **2004**, *101*(39), 14097-14101. DOI: 10.1073/pnas.0403520101.

[13] Tang, T.H.; Rozhdestvensky, T.S.; d'Orval, B.C.; Bortolin, M.L.; Huber, H.; Charpentier, B.; Branlant, C.; Bachellerie, J.P.; Brosius, J.; Huttenhofer, A. RNomics in Archaea reveals a further link between splicing of archaeal introns and rRNA processing. *Nucleic Acids Res.*, **2002**, *30*(4), 921-30.

[14] Soma, A.; Onodera, A.; Sugahara, J.; Kanai, A.; Yachie, N.; Tomita, M.; Kawamura, F.; Sekine, Y. Permuted tRNA genes expressed *via* a circular RNA intermediate in *Cyanidioschyzon merolae*. *Science*, **2007**, *318*(5849), 450-3. DOI: 10.1126/science.1145718.

[15] Halbreich, A.; Pajot, P.; Foucher, M.; Grandchamp, C.; Slonimski, P. A pathway of cytochrome B mRNA processing in yeast mitochondria: specific splicing steps and an intron-derived circular DNA. *Cell*, **1980**, *19*(2), 321-329.

[16] Hensgens, L.A.; Arnberg, A.C.; Roosendaal, E.; van der Horst G.; van der Veen, R.; van Ommen, G.J.; Grivell, L.A. Variation, transcription and circular RNAs of the mitochondrial gene for subunit I of cytochrome C oxidase. *J. Mol. Biol.*, **1983**, *164*(1), 35-58.

[17] Houseley, J.M.; Garcia-Casado, Z.; Pascual, M.; Paricio, N.; O'Dell, K.M.; Monckton, D.G.; Artero, R.D. Noncanonical RNAs from transcripts of the Drosophila muscle blind gene. *J. Hered*, **2006**, *97*(3), 253-260. DOI: 10.1093/jhered/esj037.

[18] 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.; Rajewsky, N. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature*, **2013**, *495*(7441), 333-338. DOI: 10.1038/ nature11928.

[19] Capel, B.; Swain, A.; Nicolis, S.; Hacker, A.; Walter, M.; Koopman, P.; Goodfellow, P.; Lovell-Badge, R. Circular transcripts of the testis-determining gene Sry in adult mouse testis. *Cell*, **1993**, *73*(5), 1019-1030.

[20] Chao, C.W.; Chan, D.C.; Kuo, A.; Leder, P. The mouse formin (Fmn) gene: abundant circular RNA transcripts and gene-targeted deletion analysis. *Mol. Med.*, **1998**, *4*(9), 614-28.

[21] Zaphiropoulos, P.G. Circular RNAs from transcripts of the rat cytochrome P450 2C24 gene: correlation with exon skipping. *Proc. Natl. Acad. Sci. USA*, **1996**, *93*(13), 6536-65341.

[22] Surono, A.; Takeshima, Y.; Wibawa, T.; Ikezawa, M.; Nonaka, I.; Matsuo, M. Circular dystrophin RNAs consisting of exons that were skipped by alternative splicing. *Hum. Mol. Genet.*, **1999**, *8*(3), 493-500.

[23] Hansen, T.B.; Wiklund, E.D.; Bramsen, J.B.; Villadsen, S.B. Statham, A.L.; Clark, S.J.; Kjems, J. miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA. *EMBO J*, **2011**, *30*(21), 4414-22. DOI: 10.1038/emboj.2011.359.

[24] Burd, C.E.; Jeck, W.R.; Liu, Y.; Sanoff, H.K.; Wang, Z. Sharpless, NE Expression of linear and novel circular forms of an INK4/ARF-associated non-coding RNA correlates with atherosclerosis risk. *PLoS Genet.*, **2010**, *6*(12), e1001233. DOI: 10.1371/journal.pgen.1001233.

[25] Jeck, W.R.; Sorrentino, J.A.; Wang, K.; Slevin, M.K.; Burd, C.E.; Liu, J.; Marzluff, W.F.; Sharpless, N.E. Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA*, **2013**, *19*(2), 141-57. DOI: 10.1261/rna.035667.112.

[26] Salzman, J.; Gawad, C.; Wang, P.L.; Lacayo, N.; Brown, P.O. Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. *PLoS One*, **2012**, *7*(2), e30733. DOI: 10.1371/journal.pone.0030733.

[27] Salzman, J.; Chen, R.E.; Olsen, M.N.; Wang, P.L.; Brown, P.O. Cell-type specific features of circular RNA expression. *PLoS Genet.*, **2013**, *9*(9), e1003777. DOI: 10.1371/journal.pgen.1003777.

[28] Zhang, Y.; Zhang, X.O.; Chen, T.; Xiang, J.F.; Yin, Q.F.; Xing, Y.H.; Zhu, S.; Yang, L.; Chen, L.L. Circular intronic long noncoding RNAs. *Mol. Cell*, **2013**, *51*(6), 792-806. DOI: 10.1016/ j.molcel. 2013.08.017

[29] Hansen, T.B.; Jensen, T.I.; Clausen, B.H.; Bramsen, J.B.; Finsen, B.; Damgaard, C.K.; Kjems, J. Natural RNA circles function as efficient microRNA sponges. *Nature*, **2013**, *495*(7441), 384-388. DOI: 10.1038/nature11993.

[30] Wang, P.L.; Bao, Y.; Yee, M.C.; Barrett, S.P.; Hogan, G.J.; Olsen, M.N.; Dinneny, J.R.; Brown, P.O.; Salzman, J. Circular RNA is expressed across the eukaryotic tree of life. *PLoS One*, **2014**, *9*(3), e90859. DOI: 10.1371/journal.pone.0090859.

[31] 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**, 7776. DOI: 10.3389/fpls.2016.00776.

[32] 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.; Han, B. Transcriptome-wide investigation of circular RNAs in rice. *RNA*, **2015**, *21*(12), 2076-2087. DOI: 10.1261/ rna.052282.115.

[33] Ye, C.Y.; Chen, L.; Liu, C.; Zhu, Q.H.; Fan, L. Widespread noncoding circular RNAs in plants. *New Phytol.*, **2015**, *208*(1), 88-95. DOI: 10.1111/nph.13585.

[34] Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods*, **2012**, *9*(4), 357-359. DOI: 10.1038/ nmeth.1923.

[35] Hotto, A.M.; Schmitz, R.J.; Fei, Z.; Ecker, J.R.; Stern, D.B. Unexpected diversity of chloroplast noncoding RNAs as revealed by deep sequencing of the arabidopsis transcriptome. *G3 (Bethesda)*, **2011**, *1*(7), 559-570. DOI: 10.1534/g3.111.000752.

[36] Zhu, Y.; Rowley, M.J.; Bohmdorfer, G.; Wierzbicki, A.T. A SWI/SNF chromatin-remodeling complex acts in noncoding RNA-mediated transcriptional silencing. *Mol. Cell*, **2013**, *49*(2), 298-309. DOI: 10.1016/j.molcel.2012.11.011.

[37] Ruhl, C.; Stauffer, E.; Kahles, A.; Wagner, G.; Drechsel, G.; Ratsch, G.; Wachter, A. Polypyrimidine tract binding protein homologs from Arabidopsis are key regulators of alternative splicing with implications in fundamental developmental processes. *Plant Cell*, **2012**, *24*(11), 4360-75. DOI: 10.1105/tpc.112.103622.

[38] Ivanov, A.; Memczak, S.; Wyler, E.; Torti, F.; Porath, H.T.; Orejuela, M.R.; Piechotta, M.; Levanon, E.Y.; Landthaler, M.; Dieterich, C.; Rajewsky, N. Analysis of intron sequences reveals hallmarks of circular RNA biogenesis in animals. *Cell Rep.*, **2015**, *10*(2), 170-7. DOI: 10.1016/j.celrep.2014.12.019.

[39] Dai, X.; Zhao, P.X. psRNATarget: a plant small RNA target analysis server. *Nucleic Acids Res.*, **2011**, 39(Web Server issue), W155-9. DOI: 10.1093/nar/gkr319.

[40] Franco-Zorrilla, J.M.; Valli, A.; Todesco, M.; Mateos, I.; Puga, M.I.; Rubio-Somoza, I.; Leyva, A.; Weigel, D.; Garcia, J.A.; Paz-Ares, J. Target mimicry provides a new mechanism for regulation of microRNA activity. *Nat. Genet.*, **2007**, *39*(8), 1033-1037. DOI: 10.1038/ng2079.

[41] Meng, Y.; Shao, C.; Wang, H.; Jin, Y. Target mimics: an embedded layer of microRNA-involved gene regulatory networks in plants. *BMC Genom.*, **2012**, 13197. DOI: 10.1186/1471-2164-13-197.

[42] Wu, H.J.; Wang, Z.M.; Wang, M. & Wang, X.J. Widespread long noncoding RNAs as endogenous target mimics for microRNAs in plants. *Plant Physiol.*, **2013**, *161*(4), 1875-1884. DOI: 10.1104/ pp.113.215962.

[43] Zhang, X.O.; Wang, H.B.; Zhang, Y.; Lu, X.; Chen, L.L. & Yang, L. Complementary sequence-mediated exon circularization. *Cell*, **2014**, *159*(1), 134-47. DOI: 10.1016/j.cell.2014.09.001.

[44] Cocquerelle, C.; Mascrez, B.; Hetuin, D.; Bailleul, B. Mis-splicing yields circular RNA molecules. *FASEB J.*, **1993**, *7*(1), 155-60.

[45] Li, X.F.; Lytton, J. A circularized sodium-calcium exchanger exon 2 transcript. *J. Biol. Chem.*, **1999**, *274*(12), 8153-8160.