Plant Biotechnology Reports (2024) 18:33–44
https://doi.org/10.1007/s11816-023-00884-z

Online ISSN 1863-5474
Print ISSN 1863-5466

ORIGINAL ARTICLE

Regulation of coconut somatic embryogenesis: decoding the role of long non-coding RNAs

A. A. Sabana[1] · Ginny Antony[1] · K. P. Gangaraj[2] · Tony Grace[1] · M. K. Rajesh[2]

Received: 29 June 2023 / Revised: 16 November 2023 / Accepted: 30 November 2023 / Published online: 19 January 2024

Abstract

Long non-coding RNAs (lncRNAs) are transcripts longer than 200 nucleotides that lack significant protein coding potential and have been shown to regulate various biological processes. This study was designed to identify lncRNAs in coconut and their role in the process of somatic embryogenesis in coconut, a crop with high recalcitrance to in vitro culture. RNA-Seq data of coconut embryogenic calli of the West Coast Tall cultivar was exploited for in silico prediction of lncRNA. From a total of 6328 transcripts, which were annotated as uncharacterised or with no homology hits with the existing database, 5110 putative lncRNAs are identified. We also studied the relationship between lncRNAs, microRNAs (miRNAs) and mRNAs and found that some of the lncRNAs act as miRNA precursors, some as potential miRNA targets and some function as endogenous target mimics (eTMs) for miRNAs. Real-time quantitative PCR confirmed that 10 selected lncRNAs showed significant differences in the expression pattern in different stages of coconut somatic embryogenesis. Our results suggest the existence of diverse lncRNAs in coconut embryogenic calli, some of which are differentially expressed. The information generated in this study could be of great value in understanding the molecular mechanisms governing somatic embryogenesis in coconut.

Keywords Coconut · Embryogenic calli · Endogenous target mimics · lncRNA · Somatic embryogenesis

Introduction

Coconut (*Cocos nucifera* L.; Arecaceae) is a ubiquitous palm in the tropics and sub-tropics and, importantly, provides livelihood sustenance for people of these regions (Arunachalam and Rajesh 2008; 2017). The mode of propagation in coconut is through nuts, and the resulting progenies are highly heterogeneous as the palms are mainly cross-pollinating. In vitro propagation via somatic embryogenesis (SE) can enable extensive production of uniform planting materials, but this propagation mode has limited success in this palm (Rajesh et al. 2016; Sabana et al. 2020).

Multiple studies implicate the role of various regulatory genes like transcription factors, microRNAs (miRNAs) and long non-coding RNAs (lncRNAs) and the interplay between these players in the conversion of somatic cells into embryogenic cells (Wójcik 2020). For instance, miR156 is known to regulate the transcriptional factor squamosal promoter binding protein-like (SPL) at various stages of SE in longan (Lin and Lai 2013), citrus (Long et al. 2018), larch (Zhang et al. 2021) and coffee (Hernández-Castellano et al. 2022). While miR160 controls embryonic development in *Arabidopsis* by regulating *ARF16* and *ARF17* (Liu et al. 2010), it is implicated in regulating the development of cotyledonary embryos in larch (Zhang et al. 2021). Scarecrow-like protein (SCL), a putative transcription factor targeted by miR171, plays a key role in asymmetric cell division, giving rise to diverse tissues in aerial parts of *Arabidopsis thaliana* (Di Laurenzio et al. 1996).

Long non-coding RNAs (lncRNAs) comprise transcripts exceeding 200 nucleotides in length that lack the potential to code for proteins, which regulate the expression of genes and genome imprinting (Kornienko et al. 2013; Böhmdorfer and Wierzbicki 2015). They can be transcribed from various regions, viz., regions around promoters, intra-genic and inter-genic regions, exonic, intronic, untranslated regions (UTRs), and sense and antisense strands (Ma et al. 2013).

✉ Ginny Antony
ginnyantony@cukerala.ac.in

✉ M. K. Rajesh
rajesh.mk@icar.gov.in

[1] Central University of Kerala, Kasaragod, Periye, Kerala 671320, India

[2] ICAR-Central Plantation Crops Research Institute (ICAR-CPCRI), Kasaragod, Kerala 671124, India

lncRNAs include natural antisense RNA (NATs) of protein-coding genes, precursors of small RNAs [miRNAs, and small nucleolar RNAs (snoRNAs)] and structural RNAs (tRNAs and rRNAs). lncRNAs are similar to mRNAs with several shared aspects in synthesis like RNA polymerases responsible for lncRNA transcription, post-transcriptional modifications such as the addition of 5’ caps, poly A tail and alternative splicing (Bhatia et al. 2017) and are involved in diverse cellular processes. In addition to their transcriptional, post-transcriptional, translational, and epigenetic functions, lncRNAs are also known to regulate variable cleavage, transcriptional interference, DNA methylation, and protein modification (Tripathi et al. 2010; Pruneski et al. 2011; Magistri et al. 2012; Hadjiargyrou and Delihas 2013; Johnsson et al. 2014). A few lncRNAs are transcribed by RNA polymerase IV and V (Wierzbicki et al. 2008). lncRNAs act as decoys for regulatory proteins or endogenous target mimics of miRNAs, the precursor for miRNAs and siRNAs (Wang and Chang 2011). There are evidences that these RNAs may play an important role in regulating plant growth and development, including reproductive development and stress response (Franco-Zorrilla et al. 2007; Yamaguchi and Abe 20122; Zhang et al. 2014; Muthusamy et al. 2015). The functional roles of a few lncRNAs have been identified in plants, such as the ‘Cold Assisted Intronic Non-coding RNA’ (COLDAIR) and the ‘Cold Induced Long Antisense Intergenic RNA’ (COOLAIR) that regulate vernalization in *Arabidopsis* (Swiezewski et al. 2009; Heo and Sung 2011). Photoperiod-sensitive male sterility in hybrid rice is regulated by the lncRNA LDMAR (Ding et al. 2012), while *Arabidopsis* lncRNA DRIR modulates drought and salt stress responses (Qin et al. 2017). lncRNA Early Nodulin 40 (ENOD40) is essential for the root nodule organogenesis in legumes and non-legume species like *Arachis hypogaea*, *Medicago truncatula*, *Zea mays* and *Oryza sativa* (Yang et al. 1993; Crespi et al. 1994; Nakamoto et al. 2010; Miao-Chih et al. 2010; Ganguly et al. 2021). In plants, lncRNAs with differential expression patterns during SE have been reported in only *Dimocarpus longan* Lour (Chen et al. 2018). A total of 7643 lncRNAs could be detected during early SE in longan; among these, the expression of 160 lncRNAs was specific to embryogenic calli, and 376 were expressed only in globular embryos. In addition, during early longan SE, most differentially expressed mRNAs (target genes) of lncRNAs were also associated with plant-pathogen interactions and plant hormone signaling. A regulatory network may be formed between miRNAs and mRNAs involving differentially expressed lncRNAs at different stages of SE (Chen et al. 2018).

The molecular intricacies underlying lncRNA-mediated regulation of genes in the process of SE are largely unknown in other plant systems, including coconut. Using computational tools in the present study, we have identified a set of lncRNAs from coconut embryogenic calli. We have explored the target genes of identified lncRNAs and the cross-talk between mRNA, miRNA and lncRNAs. The expression profiling of a subset of lncRNAs, in different stages of SE revealed their differential expression pattern during SE.

Methods

Identification of coconut embryogenic calli lncRNAs

RNA-Seq data of embryogenic calli derived from plumular explants of West Coast Tall cultivar (WCT) palms (SRX 472157) (Rajesh et al. 2016) was utilized for in silico prediction of lncRNAs. Raw reads obtained from the Hiseq2000 platform were subjected to quality check and adaptor contamination removal using TrimGalore (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). Further, the high-quality reads were aligned against the *Cocos nucifera* Chowghat Green Dwarf reference genome (NCBI sequence read archive database: BioProject no. PRJNA413280; Accession no.SRS2696501; Rajesh et al. 2020) using Hisat2 (https://github.com/DaehwanKimLab/hisat2). After mapping, Cufflinks (http://cole-trapnell-lab.github.io/cufflinks/) were utilized to assemble the uniquely mapped reads. Transcripts shorter than 200 bp were filtered out, and the remaining transcripts were analyzed further. The resultant sequences were subjected to a BLAST search against the oil palm and date palm databases and non-redundant protein database to remove any transcript matching the protein-coding genes. Subsequently, the obtained sequences were subjected to Transdecoder (https://github.com/TransDecoder/TransDecoder) to remove the transcripts with ≥ 100 amino acids, and final transcripts were subjected to the Coding Potential Calculator (CPC) tool (Kong et al. 2007) to differentiate non-coding/coding RNAs. A transcript with a CPC score > 0 was considered a potential long non-coding RNA. The lncRNAs were analyzed by aligning them to the Rfam database (http://rfam.xfam.org/).

Relationships among lncRNAs, miRNAs, and mRNAs

To identify the lncRNAs that act as miRNA precursors, we aligned the predicted lncRNAs with coconut miRNAs (Sabana et al. 2020). “psRNATarget” (Dai and Zhao 2011), with an expectation ≤ 3, was used to examine the possibility of the predicted lncRNAs being miRNAs targets. lncRNAs predicted as eTMs for miRNAs were identified using TAPIR (Bonnet et al. 2010). Using Cytoscape 3.2 (Saito et al. 2012), we plotted the interaction network of lncRNAs and miRNAs and mRNAs related to them.

Comparative analysis of coconut lncRNA with other plant species

A BLAST search was performed to predict the homology of coconut lncRNA with other plant species, ncRNA extracted from PNRD (plant non-coding RNA database) (Yi et al. 2015).

Quantitative real-time PCR

qRT-PCR was used to validate the expression profiles of 10 lncRNA of different length intervals, followed by verification of their expression patterns in different tissues, viz., leaf, plumule, zygotic embryo, embryogenic calli, endosperm calli and mature endosperm. RNA was isolated from these tissues using the Nucleospin Plant RNA kit (Cat#740949.50, Macherey–Nagel). The RT reactions were performed using a Primescript reagent kit (Cat# RR037B, Takara). A mix of 4.0 µl of 5 X PrimeScript buffer, 1 µl of 1 X PrimeScript RT enzyme mix, 1 µl of 1 X random hexamer primer, and 100 ng of RNA was used for the RT reaction. After incubation at 37 °C for 15 min, the reactions were heated to 85 °C for five seconds, then cooled to 4 °C. After cDNA synthesis, real-time PCR was carried out on an Applied Biosystems Real-Time PCR System using SYBR Green master mix (Applied Biosystems). Assays contained 2 µl of diluted cDNA, 1 µ of forward and reverse primers (2 µM), and 5 µl of SYBR Green master mix (2×). The reaction conditions were 95 ºC for 30 s, followed by 40 cycles at 95 °C for 5 s, 60 °C for 30 s and 72 °C for 30 s. β-tubulin gene was employed as the endogenous control (Rajesh et al. 2014). There were three biological replicates and three technical replicates per biological replicate. The relative expression was computed using the comparative CT method (ΔΔCT) using the formula $2^{-\Delta\Delta CT}$. Supplementary Table 1 lists the sequences of primers used for the validations. Statistical differences in the relative expression of samples were determined using the ANOVA test.

Results

Identification of coconut lncRNAs and their characteristics

The transcriptome data of coconut embryogenic calli published by Rajesh et al. (2016) was used for the analysis, and 508,399,94 raw reads were obtained from this data. After quality check analysis, 311,224,33 high-quality reads remained, of which 285,806,57 (91.38%) were uniquely mapped with the coconut CGD reference genome and generated 50,308 transcripts by Cufflink. Fasta formats of these transcripts were extracted and used for further analysis. Homology analysis of 49,511 transcripts was performed after removing shorter sequences (less than 200 bp). Annotating these transcripts to *Elaeis guineensis* and *Phoenix dactylifera* resulted in 6328 transcripts with no similarity hit with known proteins; downstream analysis of these transcripts was performed. To identify lncRNAs, transcripts with ORF length ≥ 100 amino acids were eliminated. The 5145 transcripts remaining were loaded into a coding potential calculator, and 5110 transcripts with CPC scores < 0 were selected as lncRNAs (Supplementary Table 2). Further analysis was not performed on transcripts with CPC scores greater than 0, which were considered either protein-coding or weak protein-coding sequences.

The length of the lncRNAs varied from 200 bp (CnulncRNA_22933.1) to 7095 bp (Cnu_lncRNA25122.3), the average being 626 bp. Most of the lncRNAs were < 1000 bp in length (88.3%) (Fig. 1; Supplementary Table 2).

Comparative analysis

To probe conservation between coconut lncRNAs and those from other plant species, all the predicted coconut lncRNAs were further probed against plant lncRNAs extracted from the plant non-coding RNA database (PNRD). This exercise identified 38 coconut lncRNAs showing similarity with seven *Arabidopsis* lncRNAs and one lncRNA CnulncRNA_11432.2 with *Populus tomentosa* lncRNATCONS_00039944. A BLAST search of coconut lncRNAs with the PNRD database revealed most of the coconut lncRNAs (99.3%) did not have any sequence homology with other plant lncRNAs and are therefore considered novel coconut lncRNAs.

The Blast searches of coconut lncRNA against the PNRD database showed that most of the coconut lncRNAs (99.3%) possessed no sequence homology to any of the lncRNAs of these plants and are hence considered novel lncRNAs of coconut.

Relationships between lncRNAs, miRNAs and mRNAs

To predict if the set of lncRNAs expressed in coconut embryogenic calli is a potential precursor of small RNAs, we aligned coconut novel and conserved miRNAs to the 5110 lncRNA sequences. This analysis identified 15 lncRNAs as potential precursors of 23 coconut miRNAs (12 conserved and 11 novel miRNAs) (Table 1). For example, CnulncRNA_9419.1 was predicted as a precursor of coconut conserved miRNA cnu-miR396e (Fig. 2A; Table 1). Single lncRNA may act as a precursor of different members in the same miRNA family. CnulncRNA_38259.1, acts as a precursor of five miRNAs of the family miR319

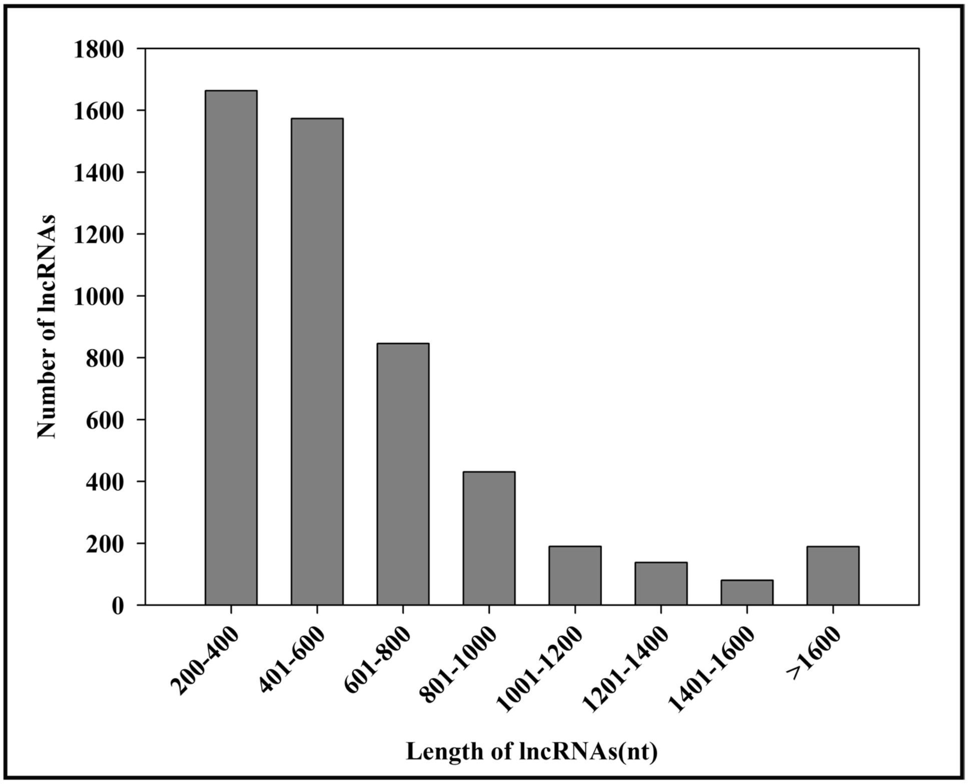
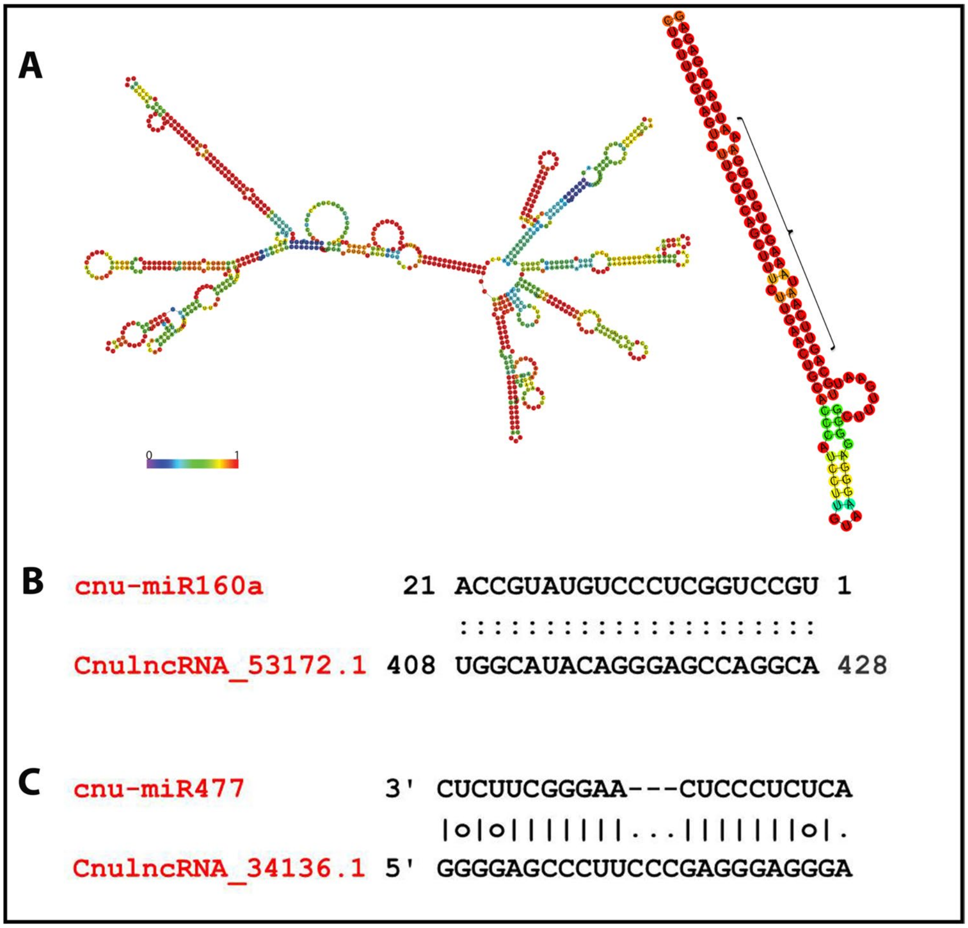


Fig. 1 Length distribution of coconut lncRNAs

Table 1 List of coconut lncRNAs which are predicted miRNA precursors

lncRNA Id	miRNA ID	Mature miRNA sequence
CnulncRNA_30665.1	cnu-miR166l	GGAAUGUUGUCUGGCUCGAGG
CnulncRNA_38259.1	cnu-miR319b	UUGGACUGAAGGGAGCUCC
	cnu-miR319d	UUGGACUGAAGGGAGCUCCC
	cnu-miR319h	UUGGACUGAAGGGAGCUCCCU
	cnu-miR319i	CUUGGACUGAAGGGAGCUCC
	cnu-miR319j	CUUGGACUGAAGGGAGCUCCC
CnulncRNA_9419.1	cnu-miR396e	GUUCAAUAAAGCUGUGGGAAA
CnulncRNA_22141.1	cnu-miR408a	UGCACUGCCUCUUCCCUGGC
	cnu-miR408b	UGCACUGCCUCUUCCCUGGCU
	cnu-miR408c	AUGCACUGCCUCUUCCCUGGC
CnulncRNA_33180.1	cnu-miR535b	UGACAACGAGAGAGAGCACGC
CnulncRNA_34042.1	cnu-miR4995	AGGCAGUGGCUUGGUUAAGGG
CnulncRNA_55394.1	cnu-miRn4	TCAAAGTTCTCTGATTGCT
CnulncRNA_23953.2	cnu-miRn9a	TTCCCGATACCTCCCATGCC
CnulncRNA_36272.1	cnu-miRn10	TTCTGTTGGAATCTAAGTCAA
CnulncRNA_468.1	cnu-miRn14-5p	TTGGGAGATGGGTTCTCGCC
	cnu-miRn14-3p	GAGGTCCCCTCTCCCAAACCT
CnulncRNA_28902.1	cnu-miRn20a	GAGCTATGAGATCTGAGGGTC
	cnu-miRn20b	AGCTATGAGATCTGAGGGTC
CnulncRNA_3825.1	cnu-miRn24	AGATCCGAAAGGAGTTCGGCT
CnulncRNA_42027.1	cnu-miRn29	TTGGATTAGAGAAGATTGTGC
CnulncRNA_52918.1	cnu-miRn50	CGGCCGAACTCCTTCGGACT
CnulncRNA_37739.1	cnu-miRn89	CGGCTGAAGGAGTTCGGC



```
cnu-miR160a         21 ACCGUAUGUCCCUCGGUCCGU 1
                       :::::::::::::::::::::
CnulncRNA_53172.1  408 UGGCAUACAGGGAGCCAGGCA 428
```

```
cnu-miR477         3' CUCUUCGGGAA---CUCCCUCUCA
                      |o|o|||||||...|||||||o|.
CnulncRNA_34136.1  5' GGGGAGCCCUUCCCGAGGGAGGGA
```

Fig. 2 Coconut lncRNA predicted to be a miRNA precursor, a target and a target mimic. **A** CnulncRNA_9419.1 predicted as a precursor of cnu-miR396e. **B** CnulncRNA_53172.1 predicted as a potential target of cnu-miR160a. **C** CnulncRNA_34136.1 act's as a target mimic for miR477

[cnu-miR319b, cnu-miR319d, cnu-miR319h, cnu-miR319i and cnu-miR319j (Table 1)].

We analyzed all coconut lncRNA sequences using the "psRNATarget" program (expectation ≤ 3) to find out whether coconut miRNAs might target these lncRNAs. A total of 73 (1.4%) lncRNAs were predicted to be targets of 72 conserved miRNAs from 29 miRNA families (Supplementary Table 3), and 115 lncRNAs were targeted by 39 coconut novel miRNAs (Supplementary Table 4). Multiple miRNAs are found to target one lncRNA. On the contrary, one miRNA is found to target multiple lncRNAs. CnulncRNA_20708.1 is predicted to be targeted by multiple members of the miRNA families' miR156, miR157 and miR529 and CnulncRNA_53172.1 is predicted as a potential target of cnu-miR160a and cnu-miR160b (Fig. 2B). The novel miRNA miRn24 was found to target 32 different lncRNAs. The regulatory network of lncRNAs and their corresponding miRNAs were constructed using Cytoscape (Fig. 3).

The endogenous target mimicry (eTM) mechanism has recently been discovered as a new regulatory mechanism for the functioning of miRNAs. Twenty-two long non-coding RNAs were predicted to function as eTMs for 21 conserved and seven novel coconut miRNAs in this study (Supplementary Table 5). For example, we identified one target mimic, CnulncRNA_34136.1, for miR477 (Fig. 2C) and three target mimic's (CnulncRNA_1763.1, CnulncRNA_490.1 and CnulncRNA_57832.1) for miR159 family members (miR159d/e/f/j) (Supplementary Table 5).

Expression analysis of lncRNAs in coconut tissues

To validate the expression pattern of coconut lncRNAs, we randomly selected 10 lncRNAs based on their length and validated their expression pattern by qRT-PCR. Among the 10 lncRNAs, three lncRNAs (viz., CnulncRNA_23928.1, CnulncRNA_55266.1 and CnulncRNA_9419.1) showed the highest expression in the embryogenic calli. Four lncRNAs, viz., CnulncRNA_53172.1, CnulncRNA_6733.1, CnulncRNA_4498.1 and CnulncRNA_29325.2 displayed highest expression in the zygotic embryo and the expression of CnulncRNA_33180.1, CnulncRNA_28901.1 and CnulncRNA_38259.1 were abundant in plumular tissues. Out of 10 lncRNAs, the expression of CnulncRNA_33180.1 was high in all the tissues compared to other lncRNAs. All the lncRNAs showed the lowest expression in endosperm calli

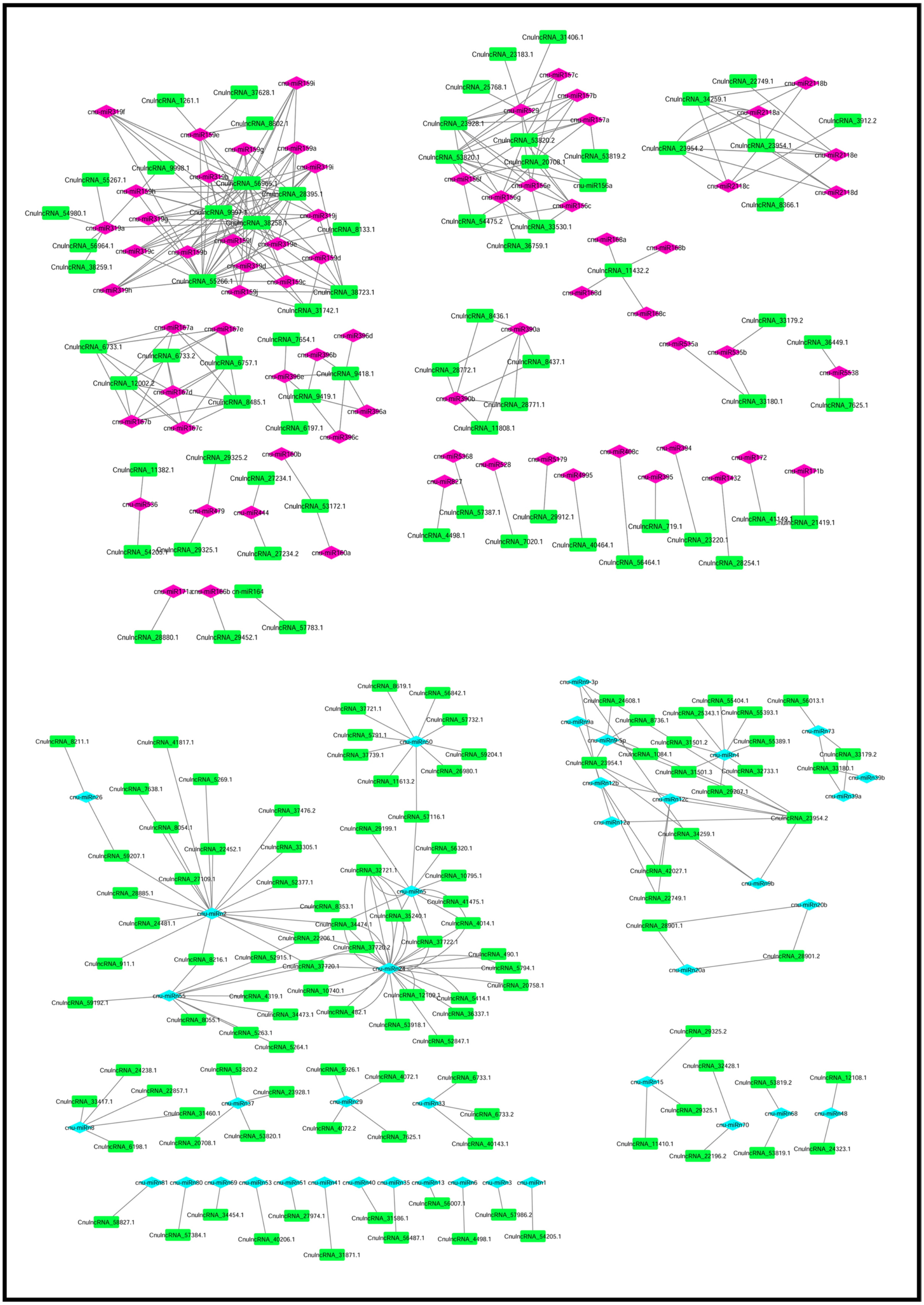


Fig. 3 Interaction network of lncRNAs and miRNA target's

and mature endosperm compared to the zygotic embryo, embryogenic calli and plumule (Fig. 4).

Discussion

Somatic embryogenesis is a highly regulated process, and though much progress has been made in deciphering the molecular process underlying SE, the information is still sketchy. The current study is the first genome-wide analysis of lncRNAs during in vitro regeneration in coconut- we have identified 5110 lncRNAs related to SE. This result is comparable with the study from SE-related lncRNAs of longan (Chen et al. 2018). The number of lncRNAs reported in longan is much less than in tea (33,400) and *Medicago truncatula* (23,324) (Varshney et al. 2019; Wang et al. 2015). Coconut lncRNAs have an average length of 626 bp, similar to reports from banana (Muthusamy et al. 2015) and chickpea (Khemka et al. 2016). This is larger than the average size of *Arabidopsis* (Liu et al. 2012) and jatropha (Yan et al. 2020) lncRNAs. Plant lncRNAs may be evolutionarily less conserved than miRNAs across plant species (Sunkar and Jagadeeswaran 2008). In this study, a BLAST search of coconut lncRNA against the PNRD database showed that only a few lncRNAs (0.74%) were conserved with other plant species. A similar observation was made in other plant species such as longan, *Populus*, cucumber, *Arabidopsis*, rice etc. (Maclntosh et al. 2001; Zhang et al. 2014; Tian et al. 2016; Chen et al. 2018).

As well documented in the literature, the processing of long RNA precursors can result in the formation of small regulatory RNAs (Wilusz et al. 2009). As a result, the set of lncRNAs expressed in coconut EC could be short RNA precursors. The first imprinted ncRNA H19 produced 23 nt miRNA. In humans and mice, it may function as a miRNA precursor to control target mRNA at the post-transcriptional stage during vertebrate development (Cai and Cullen 2007; Hung and Chang 2010; Keniry et al. 2012).

Fifteen lncRNAs were identified as coconut miRNA precursors in our study (12 conserved miRNAs belonging to six miRNA families and 11 novel miRNAs). Earlier studies in plants such as jatropha, tea, cabbage, tomato, and *Populus* have revealed that lncRNAs act as a precursor of miRNA molecules expressed in various developmental conditions (Yan et al. 2020; Varshney et al. 2019; Wang et al. 2019; Zhou et al. 2019; Chen et al. 2016). The role of lncRNAs in SE in longan was studied by Chen et al. (2018), and the results revealed that seven lncRNAs could function as precursors of miRNA family members of miR156, miR319 and miR162. These miRNAs were also connected with different developmental stages of SE in other plants (Zhang et al. 2017, 2012; Wang et al. 2012). Similar observations were also made in this study; coconut lncRNAs act as precursors of the miRNA family, including miR166, miR319, miR396, miR408, miR535 and some of the novel miRNAs. These miRNA family members are associated with the coconut SE (Sabana et al. 2020). miR166 was highly expressed in early longan SE (Xu et al. 2020), and accumulation of miR319b was observed at the cotyledonary stage of SE in lily (Zhang et al. 2017). *Arabidopsis* miR319b, on the other hand, is highly expressed in the early and late SE (Szyrajew et al. 2017).

There is a possibility that miRNAs may regulate some of the lncRNAs responding to coconut SE as targets and this study predicted that a total of 73 lncRNAs are targeted by 72 conserved miRNAs belonging to 29 miRNA families, and 115 lncRNAs are targeted by 39 coconut novel miRNAs. Thus, during coconut SE, miRNAs may regulate most of the lncRNAs, while lncRNAs may indirectly affect mRNA expression by regulating the number of miRNAs.

Endogenous target mimicry (eTM) is an important regulatory mechanism for lncRNAs in plants (Wu et al. 2013). By binding to miRNAs, lncRNAs act as endogenous miRNA traps, preventing the degradation of their true targets. Target mimicry was first observed in *Arabidopsis thaliana* in 2007 (Franco-Zorrilla et al. 2007); here, the lncRNA IPS1 (Induced by Phosphate Starvation 1) functions as an eTM of miR399, and by binding to miR399, it creates a three-nucleotide bulge in between the 10th and 11th positions of the 5’ end. This prevents from being cleaved by miR399. Later, from intergenic/non-coding gene regions, Wu et al. (2013) identified eTMs for 20 conserved miRNAs in *Arabidopsis a*nd rice. In longan, eTM regulates the expression level of the miR160a* during SE (Lin et al. 2015) and identified 40 eTMs for 15 miRNAs during the early stages of SE, proving lncRNAs LTCONS-00042843 and LTCONS-00046326 act as eTMs of miR172a and miR529d, respectively (Chen et al. 2018). In our study, we have identified 22 lncRNAs which could function as eTMs for 21 conserved and seven novel miRNAs. Interestingly, six miRNA families (miR159, miR319, miR395, miR396, miR408, miR444 and miR477) regulated lncRNAs as a target, but they were also targets of eTMs. In the network of lncRNA, miRNA and mRNA, CnulncRNA_12002.2 is predicted to target the miR167 family and act as an eTM for novel miRNA cnu-miRn33. Surprisingly, the NGS result of miRNAs (Sabana et al. 2020) showed that the miRNA family viz., cnu-miR159, cnu-miR396, cnu-miR395, cnu-miR319, cnu-miR408, cnu-miR444 and cnu-miR477 (lncRNA act as eTM for these miRNAs) are downregulated or have very low expression in the EC stage of coconut SE This indicates the possibility of an interplay between lncRNAs-miRNAs.

Expression patterns observed from the transcriptome data were confirmed by real-time PCR of selected lncRNAs, including both lncRNA acting as a precursor of miRNAs (CnulncRNA_33180.1, CnulncRNA_9419.1,

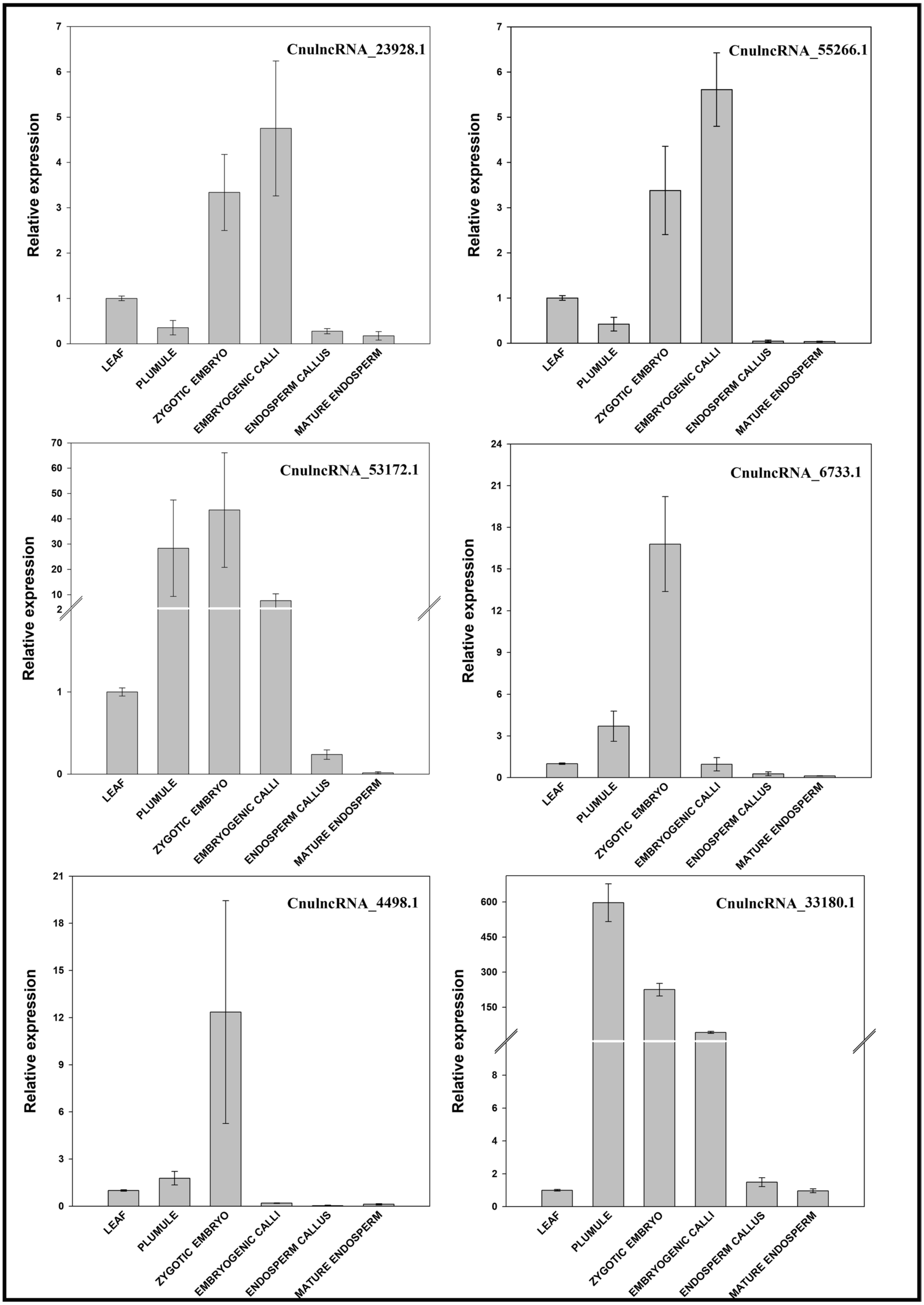


Fig. 4 qRT-PCR analysis of selected lncRNAs. The leaf represents the control sample, and the error bar indicates the mean ± SE among three biological replicates. Significant differences were calculated using ANOVA with a corrected p value ≤ 0

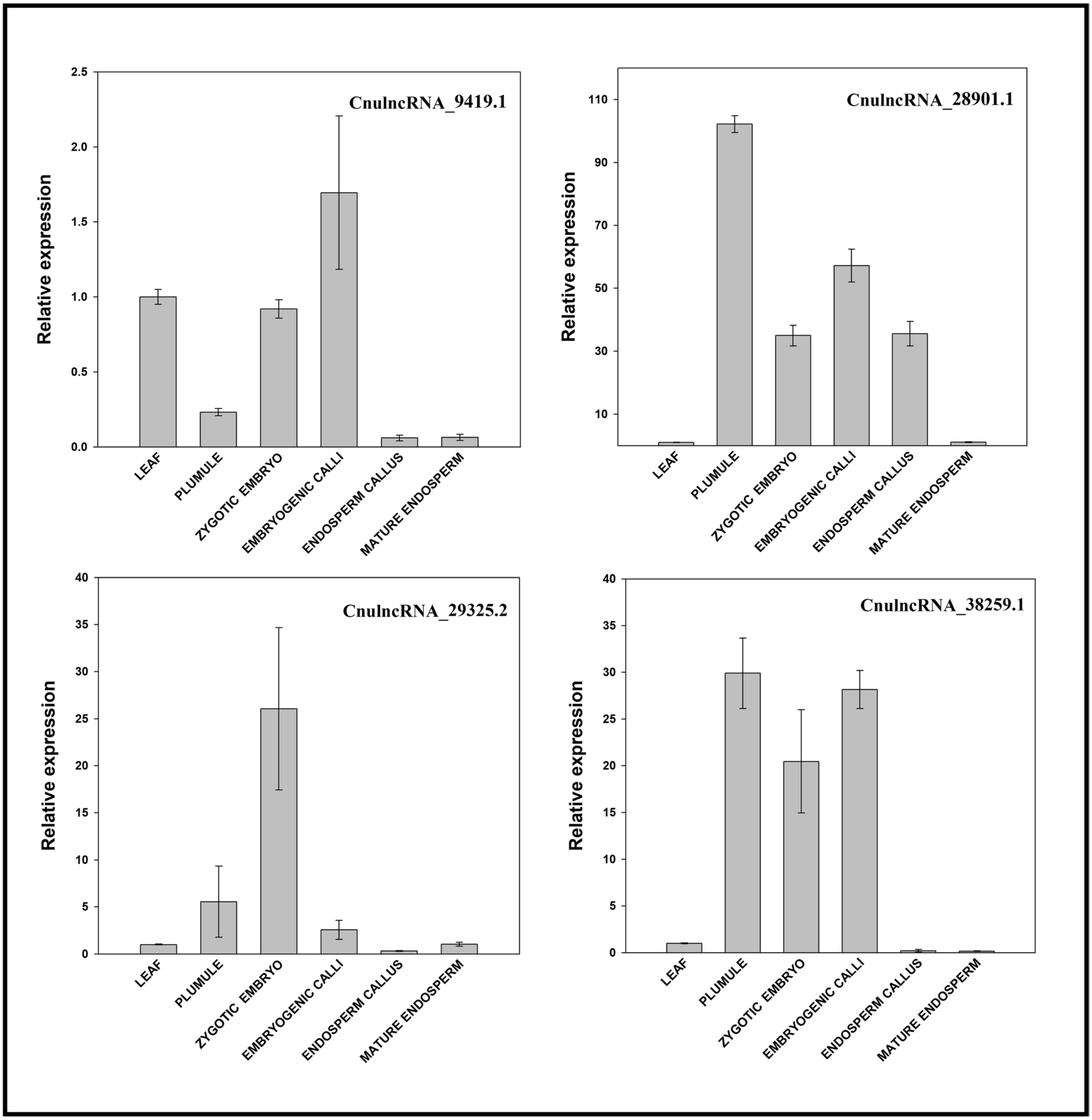


Fig. 4 (continued)

CnulncRNA_38259.1) and acting as an eTM for miRNAs (CnulncRNA_23928.1, CnulncRNA_55266.1, CnulncRNA_53172.1, CnulncRNA_6733.1, CnulncRNA_4498.1, CnulncRNA_28901.1 and CnulncRNA_29325.2) All the selected lncRNA displayed high expression in the EC, plumular explant and zygotic embryo, and low expression in the endosperm calli and its explant, mature endosperm. This suggests that lncRNAs might be regulating the embryogenic potential of coconut tissue. Endosperm calli does not have embryogenic potential to regenerate into plantlets. We provide evidence that a significant proportion of coconut lncRNAs act as miRNA precursors, targets or target mimics, suggesting a major functional role for lncRNAs as miRNA targets.

Conclusions

In conclusion, in this study, we identified 5110 lncRNAs in the EC stage of coconut SE. We also studied the relationship between lncRNAs, microRNAs (miRNAs), and mRNAs. Our findings indicate that lncRNAs, in addition to target gene regulation, regulate miRNA function as its precursor, target or endogenous target mimics. This lncRNA-miRNA-mRNA interaction may hold the secret to the regulatory mystery of coconut SE.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11816-023-00884-z.

Author contributions AAS, GA and MKR designed and executed the project; AAS, KPG, GA, TG and MKR generated and analyzed the data; AAS and GA wrote the manuscript; TG and MKR revised the manuscript. All authors have read and approved the final manuscript.

Funding This research received funding from the Science and Engineering Research Board, Government of India [CRG/2019/006573 (GA) and CRG/2020/002432 (MKR)] and Indian Council of Agricultural Research (ICAR-CPCRI Institute Project Code no. 1000761031). The first author is a recipient of a Junior Research Fellowship from the Kerala State Council for Science, Technology and Environment (KSCSTE No. 042/FSHP-LESS/2014/KSCSTE), Government of Kerala, India.

Data availability RNA-Seq data available in NCBI BioProject database under accession number PRJNA235909. Data generated or analysed during this study have been included in this manuscript.

Declarations

Conflict of interest The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Ethical approval and consent to participate Not applicable.

Human and animal ethics Not applicable.

Consent for publication All authors have provided their consent to publish this article.

References

- Arunachalam V, Rajesh MK (2008) Breeding of coconut palm (*Cocos nucifera* L.). CAB Rev. https://doi.org/10.1079/PAVSNNR20083053
- Arunachalam V, Rajesh MK (2017) Coconut genetic diversity, conservation and utilization. In: Ahuja MR, Jain SM (eds) Biodiversity and conservation of woody plants. Springer International Publishing, pp 3–36
- Bhatia G, Goyal N, Sharma S, Upadhyay SK, Singh K (2017) Present scenario of long non-coding RNAs in plants. Non-Coding RNA 3:16
- Böhmdorfer G, Wierzbicki AT (2015) Control of chromatin structure by long noncoding RNA. Trends Cell Biol 25:623–632
- Bonnet E, He Y, Billiau K, Van YDP (2010) TAPIR, a web server for the prediction of plant microRNA targets, including target mimics. Bioinformatics 26:1566–1568
- Cai X, Cullen BR (2007) The imprinted H19 non-coding RNA is a primary microRNA precursor. RNA 13(3):313–316
- Chen M, Wang C, Bao H, Chen H, Wang Y (2016) Genome-wide identification and characterization of novel lncRNAs in Populus under nitrogen deficiency. Mol Genet Genom 291:1663–1680
- Chen Y, Li X, Su L, Chen X, Zhang S, Xu X, Zhang Z, Chen Y, XuHan X, Lin Y, Lai Z (2018) Genome-wide identification and characterization of long non-coding RNAs involved in the early somatic embryogenesis in *Dimocarpus longan* Lour. BMC Genom 19:805
- Crespi MD, Jurkevitch E, Poiret M, d'Aubenton-Carafa Y, Petrovics G, Kondorosi E, Kondorosi A (1994) Enod40, a gene expressed during nodule organogenesis, codes for a non-translatable RNA involved in plant growth. EMBO J 13:5099–5112
- Dai X, Zhao PX (2011) psRNATarget: a plant small RNA target analysis server. Nucleic Acids Res. https://doi.org/10.1093/nar/gkr319
- Di Laurenzio L, Wysocka-Diller J, Malamy JE, Pysh L, Helariutta Y, Freshour G, Hahn MG, Feldman KA, Benfey PN (1996) The *SCARECROW* gene regulates an asymmetric cell division that is essential for generating the radial organization of *Arabidopsis* root. Cell 86:423–433
- Ding JH, Lu Q, Ouyang YD, Mao HL, Zhang PB, Yao JL, Xu C, Li X, Xiao J, Zhang QA (2012) long non-coding RNA regulates photoperiod-sensitive male sterility, an essential component of hybrid rice. Proc Natl Acad Sci 109:2654–2659
- Franco-Zorrilla JM, Valli A, Todesco M, Mateos I, Puga MI, Rubio-Somoza I, Leyva A, Weigel D, Garcia JA, Paz-Ares J (2007) Target mimicry provides a new mechanism for regulation of microRNA activity. Nat Genet 39:1033–1037
- Ganguly P, Roy D, Das T, Kundu A, Cartieaux F, Ghosh Z, DasGupta M (2021) The natural antisense transcript DONE40 derived from the lncRNA ENOD40 locus interacts with SET domain protein ASHR3 during inception of symbiosis in *Arachis hypogaea*. Mol Plant Microbe Interact 34:1057–1070
- Hadjiargyrou M, Delihas N (2013) The intertwining of transposable elements and non-coding RNAs. Int J Mol Sci 14(7):13307–13328
- Heo JB, Sung S (2011) Vernalization-mediated epigenetic silencing by a long intronic non-coding RNA. Science 331:76–79
- Hernandez-Castellano S, Andrade-Marcial M, Aguilar-Méndez ED, Loyola-Vargas VM, der Folter S, De-la-Pena C (2022) MiRNA expression analysis during somatic embryogenesis in *Coffea canephora*. Plant Cell Tissue Organ Cult. https://doi.org/10.1007/s11240-022-02258-9
- Hung T, Chang HY (2010) Long non-coding RNA in genome regulation: prospects and mechanisms. RNA Biol 7:582–585
- Johnsson P, Lipovich L, Grandér D, Morris KV (2014) Evolutionary conservation of long non-coding RNAs; sequence, structure, function. Biochim Biophys Acta 1840:1063–1071
- Keniry A, Oxley D, Monnier P, Kyba M, Dandolo L, Smits G, Reik W (2012) The H19 lincRNA is a developmental reservoir of miR-675 that suppresses growth and Igf1r. Nat Cell Biol 14:859–865
- Khemka N, Singh VK, Garg R, Jain M (2016) Genome-wide analysis of long intergenic non-coding RNAs in chickpea and their potential role in flower development. Sci Rep 6:1–10
- Kong L, Zhang Y, Ye ZQ, Liu XQ, Zhao SQ, Wei L, Gao G (2007) CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. Nucleic Acids Res 35:W345–W349
- Kornienko AE, Guenzl PM, Barlow DP, Pauler FM (2013) Gene regulation by the act of long non-coding RNA transcription. BMC Biol 11:59
- Lin Y, Lai Z (2013) Comparative analysis reveals dynamic changes in miRNAs and their targets and expression during somatic

embryogenesis in Longan (*Dimocarpus longan* Lour.). PLoS ONE 8:e60337

Lin Y, Lai Z, Lin L, Lai R, Tian Q, Ye W, Zhang D, Yang M, Chen Y, Zhang Z (2015) Endogenous target mimics, microRNA167, and its targets ARF6 and ARF8 during somatic embryo development in *Dimocarpus longan* Lour. Mol Breed 35:227

Liu X, Huang J, Wang Y, Khanna K, Xie Z, Owen HA, Zhao D (2010) The role of floral organs in carpels, an *Arabidopsis* loss-of-function mutation in microRNA160a, in organogenesis and the mechanism regulating its expression. Plant J 62:416–428

Liu J, Jung C, Xu J, Wang H, Deng S, Bernad L, Arenas-Huertero C, Chua NH (2012) Genome-wide analysis uncovers regulation of long intergenic non-coding RNAs in *Arabidopsis*. Plant Cell 24:4333–4345

Long JM, Liu CY, Feng MQ, Liu Y, Wu XM, Guo WW (2018) miR156-SPL modules regulate induction of somatic embryogenesis in citrus callus. J Exp Bot 69(12):2979–2993

Ma L, Bajic VB, Zhang Z (2013) On the classification of long non-coding RNAs. RNA Biol 10(6):924–933

Maclntosh GC, Wilkerson C, Green PJ (2001) Identification and analysis of *Arabidopsis* expressed sequence tags characteristic of non-coding RNAs. Plant Physiol 127:765–776

Magistri M, Faghihi MA, Rd SLG, Wahlestedt C (2012) Regulation of chromatin structure by long non-coding RNAs: focus on natural antisense transcripts. Trends Genet 28:389–396

Miao-Chih T, Ohad M, Yue W, Nima M, Wang JK, Fei L, Yang S, Eran S, Chang HY (2010) Long non-coding RNA as modular scaffold of histone modification complexes. Science 329:689–693

Muthusamy M, Uma S, Backiyarani S, Saraswathi MS (2015) Genome-wide screening for novel, drought stress-responsive long non-coding RNAs in drought-stressed leaf transcriptome of drought-tolerant and-susceptible banana (*Musa* spp) cultivars using Illumina high-throughput sequencing. Plant Biotechnol Rep 9:279–286

Nakamoto S, Tashiro K, Matsumoto A (2010) Rice ENOD40: Isolation and expression analysis in rice and transgenic soybean root nodules. Plant J 18:121–129

Pruneski JA, Hainer SJ, Petrov KO, Martens JA (2011) The Paf1 complex represses SER3 transcription in Saccharomyces cerevisiae by facilitating intergenic transcription-dependent nucleosome occupancy of the SER3 promoter. Eukaryot Cell 10:1283

Qin T, Zhao H, Cui P, Albesher N, Xiong L (2017) A nucleus-localized long non-coding RNA enhances drought and salt stress tolerance. Plant Physiol 175:1321–1336

Rajesh MK, Rachana KE, Fayas TP, Merin B, Kiran AG, Anitha K (2014) Selection and validation of reference genes for quantitative gene expression studies by real-time PCR in coconut. In: Muralidharan K, Rajesh MK, Muralikrishna KS, Jesmi V, Jayasekhar S (eds) Book of abstracts of national seminar on sustainability of coconut, arecanut and cocoa farming-technological advances and way forward. CPCRI, Kasaragod, p 34

Rajesh MK, Fayas TP, Naganeeswaran S, Rachana KE, Bhavyashree U, Sajini KK, Karun A (2016) *De novo* assembly and characterization of global transcriptome of coconut palm (*Cocos nucifera* L.) embryogenic calli using Illumina paired-end sequencing. Protoplasma 253:913–928

Rajesh MK, Chowdappa P, Behera SK, Kasaragod S, Gangaraj KP, Kotimoole CN, Nekrakalaya B, Mohanty V, Sampgod RB, Banerjee G, Prasad TSK (2020) Assembly and annotation of the nuclear and organellar genomes of a dwarf coconut (Chowghat Green Dwarf) possessing enhanced disease resistance. OMICS Int J Integr Biol 24:726–742

Sabana AA, Rajesh MK, Antony G (2020) Dynamic changes in the expression pattern of miRNAs and associated target genes during coconut somatic embryogenesis. Planta 251:1–18

Saito R, Smoot ME, Ono K, Ruscheinski J, Wang PL, Lotia S, Pico AR, Bader GD, Ideker T (2012) A travel guide to Cytoscape plugins. Nat Methods 9:1069–1076

Sunkar R, Jagadeeswaran G (2008) *In silico* identification of conserved microRNAs in large number of diverse plant species. BMC Plant Biol 8:37

Swiezewski S, Liu F, Magusin A, Dean C (2009) Cold-induced silencing by long antisense transcripts of an *Arabidopsis* Polycomb target. Nature 462:799–802

Szyrajew K, Bielewicz D, Dolata J, Wójcik AM, Nowak K, Szczygieł-Sommer A, Szweykowska-Kulinska Z, Jarmolowski A, Gaj MD (2017) MicroRNAs are intensively regulated during induction of somatic embryogenesis in *Arabidopsis*. Front Plant Sci 8:18

Tian J, Song Y, Du Q, Yang X, Ci D, Chen J, Xie J, Li B, Zhang D (2016) Population genomic analysis of gibberellin-responsive long non-coding RNAs in Populus. J Exp Bot 67:2467–2482

Tripathi V, Ellis JD, Shen Z, Song DY, Pan Q, Watt AT, Freier SM, Bennett CF, Sharma A, Bubulya PA (2010) The nuclear-retained non-coding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. Mol Cell 39:925–938

Varshney D, Rawal HC, Dubey H, Bandyopadhyay T, Bera B, Kumar PM, Singh NK, Mondal TK (2019) Tissue specific long non-coding RNAs are involved in aroma formation of black tea. Ind Crops Prod 133:79–89

Wang KC, Chang HY (2011) Molecular mechanisms of long non-coding RNAs. Mol Cell 43:904–914

Wang S, Wu K, Yuan Q, Liu X, Liu Z, Lin X, Zeng R, Zhu H, Dong G, Qian Q (2012) Control of grain size, shape and quality by OsSPL16 in rice. Nat Genet 44:950–954

Wang TZ, Liu M, Zhao MG, Chen R, Zhang WH (2015) Identification and characterization of long non-coding RNAs involved in osmotic and salt stress in *Medicago truncatula* using genome-wide high-throughput sequencing. BMC Plant Biol 15:1–13

Wang A, Hu J, Gao C, Chen G, Wang B, Lin C, Song L, Ding Y, Zhou G (2019) Genome-wide analysis of long non-coding RNAs unveils the regulatory roles in the heat tolerance of Chinese cabbage (*Brassica rapa* ssp. *chinensis*). Sci Rep 9:1–14

Wierzbicki AT, Haag JR, Pikaard CS (2008) Non-coding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. Cell 135:635–648

Wilusz JE, Sunwoo H, Spector DL (2009) Long non-coding RNAs: functional surprises from the RNA world. Genes Dev 23:1494–1504

Wójcik AM (2020) Research tools for the functional genomics of plant miRNAs during zygotic and somatic embryogenesis. Int J Mol Sci 21(14):4969

Wu HJ, Wang ZM, Wang M, Wang XJ (2013) Wide-spread long non-coding RNAs (lncRNAs) as endogenous target mimics (eTMs) for microRNAs in plants. Plant Physiol 161:1875–1884

Xu X, Chen X, Chen Y, Zhang Q, Su L, Chen X, Chen Y, Zhang Z, Lin Y, Lai Z (2020) Genome-wide identification of miRNAs and their targets during early somatic embryogenesis in *Dimocarpus longan* Lour. Sci Rep 10:1–15

Yamaguchi A, Abe M (2012) Regulation of reproductive development by non-coding RNA in *Arabidopsis*: to flower or not to flower. J Plant Res 125:693–704

Yan X, Ma L, Yang M (2020) Identification and characterization of long non-coding RNA (lncRNA) in the developing seeds of *Jatropha curcas*. Sci Rep 10:1–10

Yang WC, Katinakis P, Hendriks P, Smolders A, de Vries F, Spee J, van Kammen A, Bisseling T, Franssen H (1993) Characterization of GmENOD40, a gene showing novel patterns of cell-specific expression during soybean nodule development. Plant J 3:573–585

Yi X, Zhang Z, Ling Y, Xu W, Su Z (2015) PNRD: a plant non-coding RNA database. Nucleic Acids Res 43:D982–D989

Zhang J, Zhang S, Han S, Wu T, Li X, Li W, Qi L (2012) Genome-wide identification of microRNAs in larch and stage-specific modulation of 11 conserved microRNAs and their targets during somatic embryogenesis. Planta 236:647–657

Zhang YC, Liao JY, Li ZT, Yu Y, Zhang JP, Li QF, Qu LH, Shu WS, Chen YQ (2014) Genome-wide screening and functional analysis identify a large number of long non-coding RNAs involved in the sexual reproduction of rice. Genome Biol 15:512

Zhang J, Xue B, Gai M, Song S, Jia N, Sun H (2017) Small RNA and transcriptome sequencing reveal a potential miRNA-mediated interaction network that functions during somatic embryogenesis in *Lilium pumilum* DC. Fisch Front Plant Sci 8:566

Zhang LF, Fan YR, Lan Q, Qi LW, Han SY (2021) Expression of the SPL-like gene LaSPL9 in Japanese larch (*Larix leptolepis*) is regulated by miR156 during somatic embryogenesis. Trees 35(5):1727–1737

Zhou Y, Cho WK, Byun HS, Chavan V, Kil EJ, Lee S, Hong SW (2019) Genome-wide identification of long non-coding RNAs in tomato plants irradiated by neutrons followed by infection with Tomato yellow leaf curl virus. Peer J 7:e6286