IDENTIFICACION Y CARACTERIZACION DE miRNAS INVOLUCRADOS EN LA MADURACIÓN DE FRUTOS DE *COFFEA ARABICA*

Fernandes-Brum, CN[1]; de Souza Gomes, M[2]; Amaral, LR[2]; Lima, AA[1]; Ságio, SA[1]; Lopes, AM[1]; Chalfun-Junior, A[1*]

[1] *Universidade Federal de Lavras, Lavras, Brasil*
[2] *Universidade Federal de Uberlândia, Brasil*
**chalfunjunior@dbi.ufla.br*

Introduction

Coffee is one of the most important agricultural crops in the world, posing itself as the second most traded commodity. Brazil is the largest producer with about 30% of world production, and second largest consumer of coffee. *Coffeaarabica*is responsible for 64% of world production. Unequal fruit development is one of the biggest problems for coffee productivity and quality.

miRNAs are about 20 to 22 nucleotides long and play a key role in negative control of gene expression (BARAKAT et al., 2007). By pairing specific bases, these short transcript sequences inhibit the translation process and/or promote cleavage of their target mRNAs (VOINNET, 2009). The primary transcript (pri-miRNA) is a long simple molecule of RNA with a structural conformation known as a hairpin. Pri-miRNA undergoes and then produces an intermediate precursor called pre-miRNA (VAUCHERET, H., 2006). The pre-miRNAs are transported into the cytoplasm they will be processed by DCL1 and will create a double strand product containing mature miRNA in one of their arms and in the other an additional sequence known as miRNA* (SCHWARZ et al., 2003).

The mature miRNA, is transported to the RISC *(RNA Induced Silence Complex*) protein complex while the sequence that includes miRNA* is degraded (HAMMOND et al., 2000). In the RISC the Argonaute 1 (AGO1) enzymes are guided by mature miRNA and act cleaving complemental mRNAs into two filaments that then become unable to translate a functional protein (KAWAMATA; TOMARI, 2010).

In plants, miRNAs play an important role in multiple biological and metabolic processes (LIU et al., 2008; YANG; XUE; AN, 2007). Identified miRNA functions include control of differentiation and development, switch between phases (XIE et al., 2007), signaling (MALLORY; BARTEL; BARTEL, 2005) and response to stress (SUNKAR et al., 2007).

Material and Methods

Libraries from green and cherry fruits of *C. arabica*were used as data set to further analysis in this study. Trinity software (Grabherr, Haas et al. 2011)(Broad Institute and Hebrew University of Jerusalem) was used as tool for reconstruction of transcriptomes from RNA-seq data.

We used an integrated approach to search potential conserved miRNAs (precursors and matures) in *Coffeaarabica*and *Coffeacanephora*databases: ESTs (Expressed Sequence Tag), GSS (Genome Survey Sequences) and Nucleotides deposited in NCBI (National Center for Biotechnology Information - www.ncbi.nlm.nih.gov). Briefly, we retrieved sequences that can form hairpin-like structures from Coffee databasesusing Einverted (from EMBOSS) and BLASTN tools. The parameters for Einverted program were maxrepeat 336 nucleotides and score threshold 25 (THAKUR et al., 2011). We also performed Blastn tool to search hairpin-like sequences using all pre-miRNA sequences deposited in miRBase version 20.0 as queries. The e-value threshold used was 0.001, minimal match length 25 nucleotides and 80% of identity. We collected sequences with the length between 60 and ~ 400 nt. To get real miRNAs the set of

hairpin-like sequences were filtered in the following steps i.e MFE (Minimal free energy) filter, GC content filter, mature sequence homology filter, protein coding genes filter, noncoding RNAs filter and Reapeat sequence filter. The putative hairpin-like sequences obtained from EMBOSS and BLASTn tool were filtered using MFE(s) determined via RNAfold (Vienna RNA Package) with the following parameters: RNA secondary folding energy threshold -20 kcal/mol and with the options "-p -d2 -noLP" (HOFACKER, 2009). Secondly, these structures were filtered with GC content ranging from 20% to 65%. In additional, plant mature miRNAs deposited in miRBase (Version 20.0) were aligned against the sequences and no more than 4 mismatches were accepted in whole mature miRNAs. Other classes of non-coding RNAs (i.e., rRNA, snRNA, SL RNA, SRP, tRNAs, and RNase P) were get rid of using Rfam microRNA Registry (version 11.0)(GARDNER et al., 2009). Finally, we used Repeat masker database (http://www.repeatmasker.org/)(RepeatMasker 4.0.2) to remove the repeat like-sequences separating as positive result the putative real microRNA precursors.

For further analysis a set of structural characteristics and thermodynamic parameters were selected and analyzed in the identified Coffeepre-miRNAs: Minimal Free Energy (MFE), Adjusted Minimal Free Energy (AMFE), Minimal Free Energy Index (MFEI), length, A content, U content, C content, G content, GC content, AU content, GC ratio, AU ratio, Minimal Free Energy of the thermodynamic ensemble (MFEE), Ensemble Diversity (Diversity), and frequency of the MFE structure in the ensemble (Frequency). The parameter adjusted MFE (AMFE) was defined as the MFE of a 100 nucleotide length of sequence and the minimal folding free energy index (MFEI) that was calculated by the following equation: MFEI = [(AMFE) × 100] / (G% + C%)] (ZHANG et al., 2006). The diversity, MFE and frequency of the ensemble were measured using RNAfold as well as MFE of the secondary structures (HOFACKER, 2009). The GC content and other structural characteristics were measured using Perl scripts.

Results and Discussion

In this study we identified miRNAs present in green and cherry fruits of coffee. It was possible to identify seven and eight MIR families, respectively. The mature miRNAs characteristics are very similar to the orthologs from other species (Table 1), such as composition and length, which ranged from 19-22 nt. The MIR families presented highly conserved composition and secondary structure when compared to the orthologs from other species, as demonstrated for miR159a in Figure 1. Another important feature is the identity of the first 5' nucleotide of mature miRNA, in coffee the uracil (U) was the preferred base in this position for most of the sequences characterized. Some studies have demonstrated that the first 5’ nucleotide is the major determinant for ARGONAUTE (AGO) protein association (MI et al., 2008; MONTGOMERY et al., 2008) and can directly influence the biological miRNA functioning(EBHARDT; FEDYNAK; FAHLMAN, 2010; VAUCHERET, HERVE, 2009).

Interestingly, only two MIR family werepresent in both types of fruits, car-miR159a and car-miR5368a. The other MIR families present were unique in each phase. This may indicate the importance of miRNAs in fruit ripening. Further identification of potential targets and other experimental analysis will provide important information for the comprehension of maturation process in coffee and miRNA involvement.

Table 1: Characteristic ofPutativeMature miRNA predicted from *Coffeaarabica* Green fruit and Cherry fruit and their respective ortholog miRNAs.

Sample	Precursor name	Mature Name	Mature Sequence	Start and End in precursor	Ortholog mature name	N^{o} of mismatches with Ortholog	N^{o} of gaps with Ortholog	Sizeofth emiRNA	Size of the Ortholog
Cherry fruit	car-miR156	car-miR156-3p	ugcucacucucuaucugucacc	79-100	mtr-miR156b-3p	0	0	22	22
Cherry fruit	car-miR156	car-miR156-5p	ugacagaagagagugagcaca	05-25	bna-miR156a	0	0	21	21
Cherry fruit	car-miR156f	car-miR156f-5p	uugacagaagagagagagcaca	04-25	gma-miR156f	0	0	22	22
Cherry fruit	car-miR159a	car-miR159a-3p	uuuggauugaaggagcucua	161-181	gma-miR159a-3p	0	0	21	21
Cherry fruit	car-miR159a	car-miR159a-5p	gagcuccuugaaguccaauag	08-28	gma-miR159a-5p	1	0	21	21
Cherry fruit	car-miR167	car-miR167-3p	agaucaugcgguaguuucacc	56-76	ahy-miR167-3p	2	0	21	21
Cherry fruit	car-miR167	car-miR167-5p	ugaagcugccagcaugaucuga	07-28	ccl-miR167a	0	0	22	22
Cherry fruit	car-miR167h	car-miR167h-3p	aucaugcuggcagcuucaacuacg	73-96	gma-miR167h	3	0	24	24
Cherry fruit	car-miR172	car-miR172-3p	ugagaaucuugaugaugcugcau	123-145	vvi-miR172d	0	0	23	23
Cherry fruit	car-miR172	car-miR172-5p	guagcaucaucaagauucaca	21-41	mtr-miR172c-5p	0	0	21	21
Cherry fruit	car-miR319	car-miR319-3p	uuggacugaaggguuuccuuc	154-174	stu-miR319-3p	1	0	21	21
Cherry fruit	car-miR319c	car-miR319c-3p	uuggacugaagggagcucccu	156-176	ath-miR319a	0	0	21	21
Cherry fruit	car-miR319c	car-miR319c-5p	agagcuuccuucagcccacuc	06-26	sly-miR319c-5p	0	0	21	21
Cherry fruit	car-miR5368	car-miR5368-5p	ggacagucucagguagaca	31-49	gma-miR5368	0	0	19	19

Cherry fruit	car-miR8130	car-miR8130-5p	ggguuccugguuggaagaacu	08-28	ppe-miR8130-5p	2	0	21	21
Cherry fruit	car-miR828	car-miR828-3p	agauacucauuugaacaagaug	157-178	aly-miR828-3p	5	0	22	22
Cherry fruit	car-miR828	car-miR828-5p	ucuugcucaaaugaguauucca	47-68	vvi-miR828a	0	0	22	22
Green fruit	car-miR159a	car-miR159a-3p	uuuggauugaaggagcucua	154-174	gma-miR159a-3p	0	0	21	21
Green fruit	car-miR159a	car-miR159a-5p	gagcuccuugaaguccaauag	01-21	gma-miR159a-5p	1	0	21	21
Green fruit	car-miR393a	car-miR393a-3p	aucaugcuauccccuuuggaua	147-167	ath-miR393a-3p	2	0	21	21
Green fruit	car-miR393a	car-miR393a-5p	uccaaagggaucgcauugaucc	20-41	ath-miR393a-5p	0	0	22	22
Green fruit	car-miR393b	car-miR393b-3p	aucaaugcgauccccuuuggaug	146-167	zma-miR393b-3p	2	0	22	22
Green fruit	car-miR393b	car-miR393b-5p	uccaaagggauagcaugauccc	20-41	zma-miR393b-5p	7	0	22	22
Green fruit	car-miR396a	car-miR396a-3p	guucaauaaagcugugggaug	97-117	ath-miR396a-3p	1	0	21	21
Green fruit	car-miR396a	car-miR396a-5p	uuccacagcuuucuugaacug	08-28	ath-miR396a-5p	0	0	21	21
Green fruit	car-miR5368	car-miR5368-3p	ggacagucucagguagaca	151-169	gma-miR5368	0	0	19	19
Green fruit	car-miR6118	car-miR6118-5p	uggaauuggguacuccggaaag	22-43	cca-miR6118-5p	4	0	22	22
Green fruit	car-miR6188	car-miR6188-3p	gguggaucaaugaacccagcuc	113-134	hvu-miR6188	4	0	22	22
Green fruit	car-miR8170	car-miR8170-5p	uugcuuaaagauguucuaucc	20-40	ath-miR8170-3p	3	0	21	21

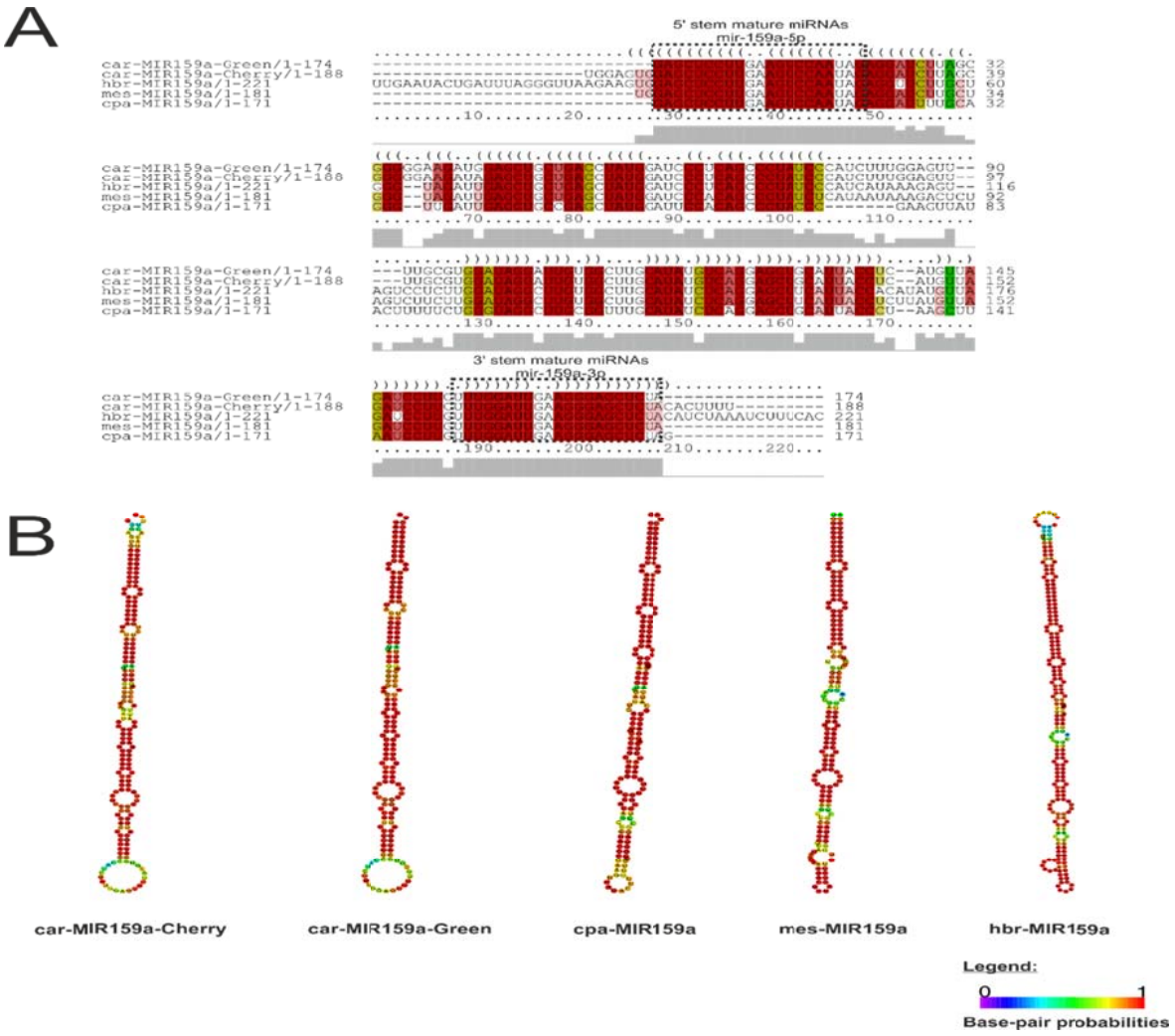

Figure 1. Alignmentof the precursorcar-miR159awithitshomologsfrom Plant species.

Conclusion

We identified seven and eight MIR families in green and cherry fruits in *Coffeaarabica*. Our data allow envisioning future studies to elucidate the maturation process in coffee and provide biotechnological tools for coffee improvement of productivity and quality.

Acknowledgements

The authors thank the Laboratory of Molecular Physiology of Plants (LFMP) of the Federal University of Lavras, the National Council for Scientific and Technological Development (CNPq) for the fellowships granted, the Minas Gerais Research Foundation (FAPEMIG) and National Institute for Science and Technology for Coffee (INCT-Café) for funding this work and the Coordination of Improvement of Higher Education (CAPES) for grants.

References

BARAKAT, A. et al. Large-scale identification of microRNAs from a basal eudicot (Eschscholzia californica) and conservation in flowering plants. **Plant Journal,** v. 51, n. 6, p. 991-1003, Sep 2007.

EBHARDT, H. A.; FEDYNAK, A.; FAHLMAN, R. P. **Naturally occurring variations in sequence length creates microRNA isoforms that differ in argonaute effector complex specificity**. Silence 1**:** 1-6 p. 2010.
GARDNER, P. P. et al. Rfam: updates to the RNA families database. **Nucleic acids research,** v. 37, n. Database issue, p. D136-40, Jan 2009.
HAMMOND, S. M. et al. **An RNA-directed nuclease mediates posttranscriptional gene silencing in Drosophila cells.** . Nature. London. 404**:** 293–296 p. 2000.
HOFACKER, I. L. RNA secondary structure analysis using the Vienna RNA package. **Current protocols in bioinformatics / editoral board, Andreas D. Baxevanis ... [et al.],** v. Chapter 12, p. Unit12 2, Jun 2009.
KAWAMATA, T.; TOMARI, Y. **Making RISC**. Trends Biochemistry Science. Amsterdan. 35**:** 68-76 p. 2010.
LIU, H. H. et al. Microarray-based analysis of stress-regulated microRNAs in Arabidopsis thaliana. **Rna-a Publication of the Rna Society,** v. 14, n. 5, p. 836-843, May 2008.
MALLORY, A. C.; BARTEL, D. P.; BARTEL, B. MicroRNA-directed regulation of Arabidopsis AUXIN RESPONSE FACTOR17 is essential for proper development and modulates expression of early auxin response genes. **Plant Cell,** v. 17, n. 5, p. 1360-1375, May 2005.
MI, S. et al. Sorting of small RNAs into Arabidopsis argonaute complexes is directed by the 5 ' terminal nucleotide. **Cell,** v. 133, n. 1, p. 116-127, Apr 4 2008.
MONTGOMERY, T. A. et al. **Specificity of ARGONAUTE7-miR390 interaction and dual functionality in TAS3 transacting siRNA formation.** . Cell 128-141**:** 128-141 p. 2008.
SCHWARZ, D. S. et al. Asymmetry in the assembly of the RNAi enzyme complex. **Cell,** v. 115, n. 2, p. 199-208, Oct 2003.
SUNKAR, R. et al. Small RNAs as big players in plant abiotic stress responses and nutrient deprivation. **Trends in Plant Science,** v. 12, n. 7, p. 301-309, Jul 2007.
THAKUR, V. et al. Characterization of statistical features for plant microRNA prediction. **BMC Genomics,** v. 12, n. 1, p. 108, 2011.
VAUCHERET, H. Post-transcriptional small RNA pathways in plants: mechanisms and regulations. **Genes & Development,** v. 20, n. 7, p. 759-771, Apr 2006.
VAUCHERET, H. AGO1 Homeostasis Involves Differential Production of 21-nt and 22-nt miR168 Species by MIR168a and MIR168b. **Plos One,** v. 4, n. 7, Jul 30 2009.
VOINNET, O. Origin, Biogenesis, and Activity of Plant MicroRNAs. **Cell,** v. 136, n. 4, p. 669-687, Feb 2009.
XIE, F. L. et al. Computational identification of novel microRNAs and targets in Brassica napus. **Febs Letters,** v. 581, n. 7, p. 1464-1474, Apr 2007.
YANG, T. W.; XUE, L. G.; AN, L. Z. Functional diversity of miRNA in plants. **Plant Science,** v. 172, n. 3, p. 423-432, Mar 2007.
ZHANG, B. et al. Conservation and divergence of plant microRNA genes. **The Plant journal : for cell and molecular biology,** v. 46, n. 2, p. 243-59, Apr 2006.