

CARACTERIZAÇÃO DO MICRORNA MIR156 ENCONTRADO EM BIBLIOTECAS DE RNA-SEQ DE FRUTOS MADUROS DE *Coffea arabica*¹

Iasminy Silva Santos²; Christiane Noronha Fernandes-Brum³; Andressa Spuri Azarias⁴; Samuel Chaves Silva⁵; Thais Cunha de Sousa Cardoso⁶; Mateus de Souza Gomes⁷; Antonio Chalfun-Junior⁸

¹Trabalho financiado pelo INCT Café

²Mestranda em Fisiologia Vegetal, Universidade Federal de Lavras, Lavras-MG, iasminysilvas@gmail.com

³Doutoranda em Biotecnologia Vegetal, Universidade Federal de Lavras, Lavras-MG, chrisbioufla@gmail.com

⁴Graduanda em Agronomia, Universidade Federal de Lavras, Andressa_spuri@hotmail.com

⁵Doutorando em Biotecnologia Vegetal, Universidade Federal de Lavras, Lavras-MG, samuelchaves21@yahoo.com.br

⁶Graduanda em Biotecnologia, Universidade Federal de Uberlândia, thaiscunhasc@gmail.com

⁷Professor Adjunto II, Universidade Federal de Uberlândia, campus Patos de Minas, Instituto de Genética e Bioquímica, matheus@ingeb.ufu.br

⁸Professor Adjunto, Universidade Federal de Lavras, Departamento de Biologia, Fisiologia Vegetal, Lavras-MG, chalfunjunior@dbi.ufla.br

RESUMO: microRNAs constituem uma classe de pequenos RNAs (20-24 nucleotídeos) endógenos não codantes envolvidos na regulação pós-transcricional da expressão gênica em eucariotos. Em plantas, mais especificamente, os miRNAs são resultados do processamento dos longos transcritos primários pelas enzimas Dicer-like e atuam predominantemente degradando mRNAs complementares. Além de plantas modelo como *Arabidopsis* e *Microtom*, os miRNAs têm sido estudados em diferentes espécies de interesse agrônomo, visando compreender suas funções regulatórias no desenvolvimento vegetativo e reprodutivo, resposta a estresses, dentre outros. Neste trabalho um pipeline específico para a busca de miRNAs homólogos em café foi aplicado em sequências advindas de bibliotecas de RNA-seq obtidas de frutos maduros de café. Como resultado 8 famílias de miRNAs foram identificadas em frutos maduros destacando-se o miRNA156 específico para a fase de maturação. Os resultados fornecem informações importantes para a compreensão do envolvimento miRNA no processo de maturação do café visando o desenvolvimento de ferramentas biotecnológicas para o melhoramento do café.

PALAVRA-CHAVE: microRNAs; café; maturação.

Characterization of micro RNA miR156 found in RNA -seq libraries of ripe fruits of *Coffea arabica*

ABSTRACT: microRNAs are a class of small RNAs (20-24 nucleotides) endogenous non-coding involved in post-transcriptional regulation of gene expression in eukaryotes. In plants, specifically, miRNA are the result of processing the primary transcripts processed by Dicer-like enzymes and act predominantly degrading complementary mRNAs. Besides in model plants like *Arabidopsis* and *Microtom*, miRNAs have been studied in different species of agronomic interest, to understand its regulatory functions in vegetative and reproductive development, stress response, among others. In this work a specific pipeline to search for miRNAs counterparts in coffee was applied to sequences resulting from RNA-seq obtained from ripe coffee fruit libraries. As a result, eight families of miRNAs have been identified in mature fruits highlighting section miRNA156 specific for maturation. The results provide important information for understanding the miRNA involvement in the coffee ripening process for the development of biotechnological tools for coffee improvement.

KEYWORDS: microRNAs, coffee, ripening.

INTRODUÇÃO

O café é a “commodity” mais comercializada no mundo depois do petróleo (CONAB 2013) e segundo a Organização Internacional do Café (OIC) o setor cafeeiro brasileiro em 2014 contribuiu com 45 milhões de sacas de 60 kg (ICO, 2015a) despontando entre os países exportadores. Quase todo café exportado pelo Brasil corresponde a espécie *C. arabica* com aproximadamente 29.3 milhões de sacas (ICO, 2015b). Diversos organismos produzem pequenos RNAs destacando-se os microRNAs (miRNAs), os quais constituem uma classe de pequenos RNAs endógenos (~ 22 nucleotídeos) envolvidos na regulação pós-transcricional da expressão gênica em eucariotos. Em plantas, resultam dos longos transcritos primários contendo imperfeitos *hairpins* que sofrem processamento pelas enzimas Rnase III tipo *Dicer*, produzindo miRNAs maduros (HUTVAGNER et al., 2001; BARTEL E BARTEL, 2003; LAI, 2003; BARTEL, 2004; GRISHOK et al., 2001; KETTING et al., 2001, LEE et al., 2003). Cada miRNA maduro deriva de porções dupla-fita do *hairpin* apresentando inicialmente um duplex contendo dois RNAs com aproximadamente 22 nt, onde um é o

miRNA maduro e o outro é conhecido como miRNA* (inseridos em braços opostos e complementares do grampo) (LAU et al., 2001; REINHART et al., 2002; SCHWARZ et al., 2003). O miRNA maduro é carregado para o complexo RISC (*RNA-induced silencing complex*, HAMMOND et al., 2000) onde funciona como um RNA guia orientando a repressão de mRNA alvos, enquanto a miRNA* é degradado (Figura 1) (SCHWARZ et al., 2003). miRNAs foram identificados em espécies de culturas cultivadas para alimentação (KULCHESKI et al. 2011), têxtil (ZHANG et al., 2007) e biocombustíveis (ZANCA et al. 2010). O estudo de miRNAs em espécies cultivadas pode contribuir para a compreensão de processos regulatórios associados com eventos morfofisiológicos, uma vez que estes pequenos RNAs regulam um grande número de genes relacionados a processos do desenvolvimento vegetal tais como controle da diferenciação dos tecidos, alternância entre fases vegetativa e reprodutiva e resposta a diferentes biótico e abiótico (XIE et al., 2007; ELDEM et al., 2012). As diversas famílias de miRNAs descritas para o cafeeiro também estão correlacionadas a diferentes processos biológicos que incluem, padrões de crescimento vegetal, processos metabólicos, respostas hormonais, respostas a estresses e rotas de sinalização celular (CHAVES, 2012). Este trabalho teve como objetivo identificar e caracterizar miRNAs encontrados em bibliotecas de RNA-seq obtidas a partir de frutos maduros de café, destacando-se o miRNAs156.

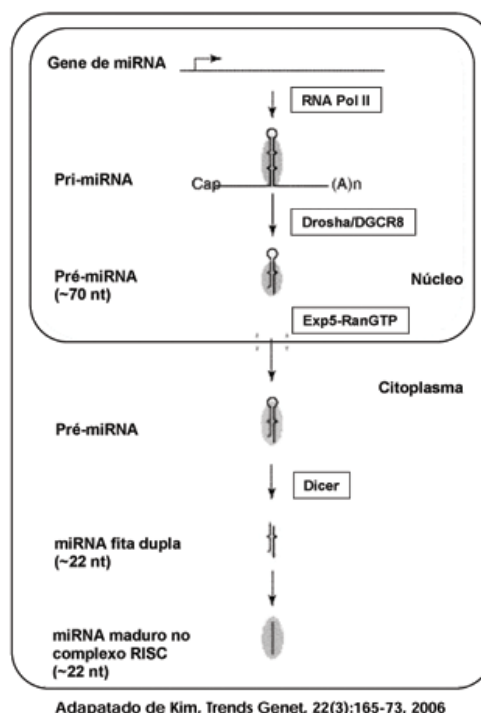


Figura 1. Via de biogênese dos miRNAs. O gene de miRNA é transcrito pela RNA polimerase II. O transcrito primário (pri-miRNA) apresenta uma estrutura *hairpin*, que é processada pela enzima RNase III, Drosha, formando o miRNA precursor (pré-miRNA) de ~70 nucleotídeos. A proteína exportina-5 leva esse produto ao citoplasma para ser processado pela RNase III, Dicer, gerando um miRNA fita dupla de ~22 nucleotídeos. Uma das fitas do duplex de miRNA é degradada enquanto a outra permanece no complexo RISC para controlar a expressão pós-transcricional de genes alvo.

MATERIAL E MÉTODOS

Bibliotecas de RNA-seq obtidas a partir frutos maduros de *C. arabica* cultivar Acauã foram utilizadas como base de dados para este estudo. O software Trinity (GRABHERR et al., 2011) (Broad Institute e Hebrew University of Jerusalem) foi usado como ferramenta para a reconstrução dos transcriptomas das bibliotecas de RNA-seq. Foi aplicado o método utilizado em *S. mansoni* (DE SOUZA GOMES et al., 2011) e adaptado para espécies *Coffea* em Chaves et al. (não publicado) em que as sequências que potencialmente formam estruturas em *hairpin* foram obtidas utilizando as ferramentas Einverted (EMBOSS) e BLASTn (NCBI). Os parâmetros para o programa Einverted foram máximo *repeat* de 336 nucleotídeos e valor *threshold* 25 (THAKUR, 2011). Foi realizada uma busca por BLASTn para encontrar sequências que potencialmente formam *hairpin* usando todas as sequências de pré-miRNAs depositadas no miRBase versão 20.0 como *query*. O valor limite de *e-value* utilizado foi 0.001, comprimento mínimo de similaridade de 25 nucleotídeos e 80% de identidade. Foram coletadas sequências entre 60 e ~400 nt. Para obter miRNAs verdadeiros, as sequências que potencialmente formam *hairpin* foram submetidas a diferentes filtros, tais como: MFE (Minimal free energy), conteúdo de GC, homologia de sequência madura, filtro de genes que codificam proteínas, filtro de RNAs não

codantes, e filtro de seqüências repetitivas. As possíveis seqüências haripin obtidas no EMBOSS e por BLASTn foram filtradas usando seus MFE determinados via RNAfold (Vienna RNA Package) com os seguintes parâmetros: energia de dobramento secundário de RNA -20 kcal/mol e opções "-p -d2 -noLP" (HOFACKER, 2009). Em seguida, as estruturas foram filtradas para o conteúdo de GC variando entre 20% a 65%. Adicionalmente, miRNAs maduros de plantas depositados no miRBase (Versão 20.0) foram alinhados às seqüências e foram aceitos no máximo 4 mismatches na estrutura madura. Outras classes de RNA não codantes (rRNA, snRNA, SL RNA, SRP, tRNAs, e RNase P) foram eliminados usando Rfam microRNA Registry (versão 11.0) (GARDNER et al., 2009). Finalmente, seqüências repetitivas foram eliminadas por um banco de dados que omite seqüências repetitivas (<http://www.repeatmasker.org/>)(RepeatMasker 4.0.2), e seqüências positivas para possíveis precursores de miRNAs reais foram separadas. Nas seqüências identificadas de pre-miRNAs de *C. arabica* foram realizadas análises de características estruturais e parâmetros termodinâmicas: Energia Mínima Livre (MFE), Energia Mínima Livre Ajustada (AMFE), Índice de Energia Mínima Livre (MFEI), tamanho, conteúdo de A, conteúdo de U, conteúdo de C, conteúdo de G, de GC e AU, razão de GC, razão de AU, Energia Mínima Livre do Conjunto (MFEE), Diversidade do Conjunto (Diversidade), e Frequência da estrutura do MFE no Conjunto (Frequência). O parâmetro de MFE ajustado (AMFE) foi definido como o MFE de 100 nucleotídeos de comprimento de seqüência e o índice de energia mínima de dobramento (MFEI), calculado pela equação: $MFEI = [(AMFE) \times 100] / (G\% + C\%)$ (ZHANG, B. et al., 2006). A diversidade, MFE e frequência do ensemble foi calculada usando RNAfold assim como o MFE das estruturas secundárias (HOFACKER, 2009). O conteúdo de GC e outras características estruturais foram calculadas usando scripts Perl.

RESULTADOS E DISCUSSÃO

Os miRNAs modelam a expressão gênica por meio do pareamento de bases específicas a um mRNA alvo, induzindo a clivagem ou o impedimento de tradução deste mRNA. Muitas das vias reguladas são fundamentais para o desenvolvimento do organismo e para a manutenção da integridade do genoma. Processos fundamentais do desenvolvimento vegetal como a determinação da identidade de células meristemáticas, da polaridade dos órgãos foliares e da transição entre os estágios vegetativo e reprodutivo, são dependentes dos mecanismos de silenciamento promovidos pelos miRNAs (CHAVES, 2012). Foram identificadas oito famílias de miRNAs em frutos maduros de café sendo analisadas as características estruturais dos seus precursores (Tabela 1). A estrutura do precursor de miR156 apresentou 102 bases em sua composição. De modo geral, a maioria das plantas (73,5%) possuem de 80 a 160 nt (ZHANG et al, 2006). Em relação à composição, o pre-miR156 apresentou conteúdo de AU (Adenina+Uracila) de 50,98% e conteúdo de GC de 49,02%. A relação GC e AU influencia diretamente o processamento destes pré-miRNAs já que maior porcentagem de AU torna-o mais instável em função das duas ligações com o hidrogênio, enquanto que miRNAs com alta porcentagem de GC apresentam maior estabilidade já que formam três ligações de hidrogênio. Logo, o miR156 pode ser processado mais facilmente em um miRNA maduro pelo complexo RISC (MI et al., 2008).

Tabela1. Características estruturais de precursores conservados de miRNAs em frutos maduros de *C. arabica*.

Amostra	miRNA ID	Tamanho	G (%)	A (%)	C (%)	U (%)	GC Content (%)	AU Content (%)
Fruto Maduro	car-miR156f	89	23,60	22,47	26,97	26,97	50,56	49,44
Fruto Maduro	car-miR828	221	22,62	27,15	22,62	27,60	45,25	54,75
Fruto Maduro	car-miR156	102	25,49	19,61	23,53	31,37	49,02	50,98
Fruto Maduro	car-miR172	163	23,93	27,61	15,34	33,13	39,26	60,74
Fruto Maduro	car-miR167	81	22,22	22,22	24,69	30,86	46,91	53,09
Fruto Maduro	car-miR167h	111	23,42	29,73	24,32	22,52	47,75	52,25
Fruto Maduro	car-miR159a	188	27,66	21,28	20,21	30,85	47,87	52,13
Fruto Maduro	car-miR319c	179	26,26	24,02	21,23	28,49	47,49	52,51
Fruto Maduro	car-miR319	178	30,34	19,10	20,22	30,34	50,56	49,44
Fruto Maduro	car-miR8130	165	25,45	25,45	18,18	30,91	43,64	56,36
Fruto Maduro	car-miR5368	83	28,92	25,30	25,30	20,48	54,22	45,78

Analisando os aspectos termodinâmicos do precursor miRNA156 selecionado é possível distinguir este miRNA de outros tipos de RNAs (Tabela 2). Através da variável MFEI constatou-se que os precursores identificados tendem a ser uma molécula de RNA silenciadora por apresentarem MFEI maiores que 0,85. Ainda para os miRNAs que apresentaram valor abaixo do referenciado, estes são significativamente mais altos que os descritos para tRNAs (0,64),

rRNAs (0,59), e mRNAs (0,62-0,66) possuem relação positiva para as demais variáveis (XIE, F. L. et al., 2007; ZHANG, B. H. et al., 2006).

Tabela 2. Características termodinâmicas de precursores conservados de miRNA156 em *C. arabica*.

Amostra	miRNA ID	MFE (kcal/mol)	MFEE (kcal/mol)	Frequência (%)	Diversidade	AMFE	MFEI
Fruto Maduro	car-miR156	-55,3	-55,3	0,3051	3,68	-54,22	-1,11

O miRNA156 foi identificado exclusivamente em frutos maduros apresentando características muito semelhantes aos ortólogos de outras espécies (Tabela 3), tais como a composição e comprimento, a qual variou 19-22 nt. O miR156 apresentou mismatches iguais a zero em relação aos ortólogos indicando elevada similaridade. A figura 2 mostra o alinhamento dos precursores de miR156 em *Coffea arabica* com seus ortólogos em outras espécies, destacando-se a região madura, que é a região que se liga ao gene alvo promovendo sua clivagem ou repressão.

Tabela 3. Característica do miRNA156 de frutos maduros de *Coffea arabica* e seus respectivos miRNAs ortólogos.

Amostra	Precurso	Nome Maduro	Sequencia Madura	Precurso inicial e final	Nome do ortólogo maduro	Nº de mismatches com ortólogos	Nº de gaps com ortólogos	Tamanho do miRNA	Tamanho do ortólogo
Fruto Maduro	car-miR156	car-miR156-3p	Ugucacucucuaucugucac	79-100	mtr-miR156b-3p	0	0	22	22
Fruto Maduro	car-miR156	car-miR156-5p	Ugacagaagagagugagacaca	05-25	bna-miR156a	0	0	21	21

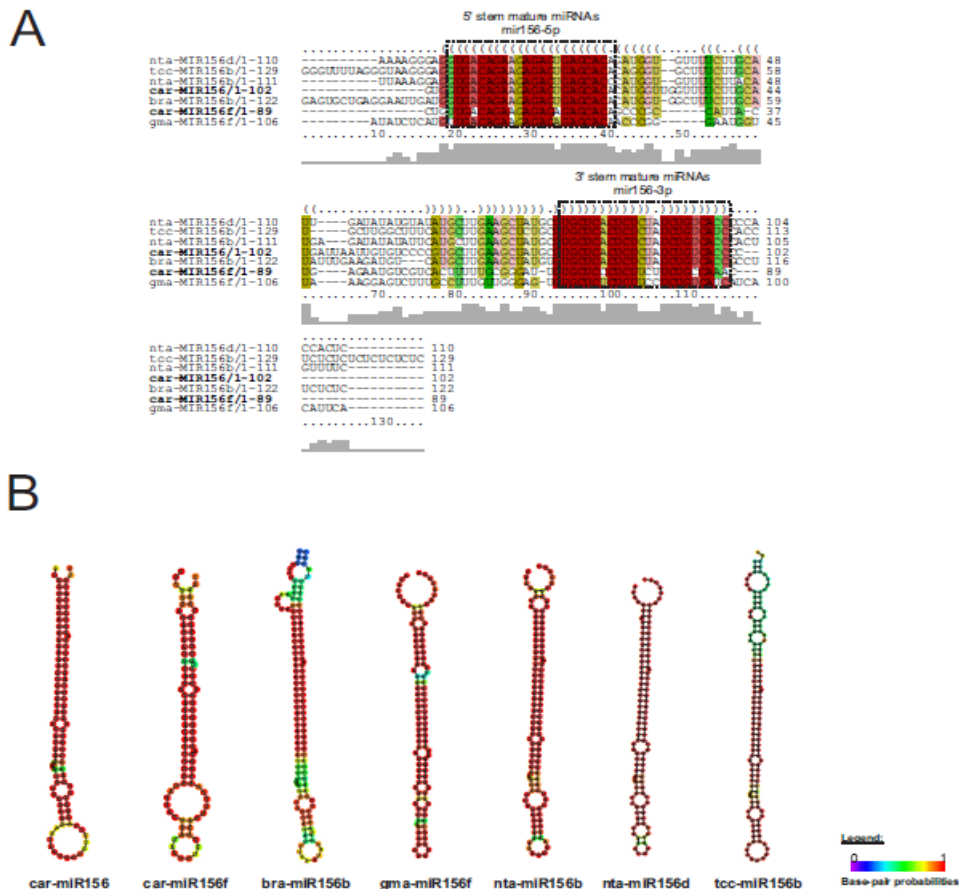


Figura 2. Alinhamento do precursor de car-MIR156 com seus ortólogos de outras espécies vegetais.

O miRNA156 foi identificado em *Solanum lycopersicum* L. estando envolvido na manutenção do estado meristemático dos tecidos vegetais dos ovários, podendo controlar etapas de desenvolvimento de frutos carnosos (FERREIRA E

SILVA et al., 2014). No presente trabalho este miRNA foi encontrado apenas em frutos maduros, evidenciando sua potencial atuação na maturação dos frutos de café. Portanto, este miRNA torna-se um importante candidato para estudos visando compreender o processo de maturação do fruto de café.

CONCLUSÕES

A partir de bibliotecas de RNA-seq foram identificadas oito famílias de miRNAs em frutos maduros, estando o miRNA156 possivelmente associado com o controle do desenvolvimento do fruto na fase de maturação. A identificação de alvos potenciais e outras análises experimentais irão fornecer informações importantes para a compreensão do envolvimento deste miRNA no processo de maturação do fruto, visando o desenvolvimento de ferramentas biotecnológicas para o melhoramento do café.

AGRADECIMENTOS

Os autores gostariam de agradecer à CAPES e ao CNPq pela concessão das bolsas de estudo e ao INCT Café pelo financiamento das pesquisas.

REFERÊNCIAS BIBLIOGRÁFICAS

- BARTEL, B.;BARTEL, D.P. (2003). MicroRNAs: at the root of plant development? *Plant Physiol.* 132, 709–717.
- BARTEL, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116, 281–297.
- CHAVES, SS. Caracterização *in silico* e análise da expressão de microRNAs e genes alvo em diferentes espécies de *Coffea*. Dissertação –programa de pós-graduação em Biotecnologia Vegetal- Universidade Federal de Lavras, 2012.
- ELDEM, V. et al. Genome-Wide Identification of miRNAs Responsive to Drought in Peach (<italic>Prunus persica</italic>) by High-Throughput Deep Sequencing. *PLoS ONE*, v. 7, n. 12, p. e50298, 2012.
- FERREIRA E SILVA, G. F. et al. microRNA156-targeted SPL/SBP box transcription factors regulate tomato ovary and fruit development. *The Plant Journal* 78, 604–618, 2014.
- 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.
- GRABHERR, M. G. et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol.* v. 29, n. 7, p. 644-52, Jul 2011.
- GRISHOK, A., PASQUINELLI, A.E., CONTE, D., LI, N., PARRISH, S., HA, I., BAILLIE, D.L., FIRE, A., RUVKUN, G., MELLO, C.C. (2001). Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell* 106, 23–34.
- HE L, HANNON GJ. MicroRNAs: small RNAs with a big role in gene regulation. *Nature Reviews Genetics.* 5:522-531.
- HOFACKER, I. L. RNA secondary structure analysis using the Vienna RNA package. *Curr Protoc Bioinformatics*, v. Chapter 12, p. Unit12.2, Jun 2009.
- ICO. BREAKDOWN OF EXPORTS OF GREEN ARABICA AND GREEN ROBUSTA FOR COUNTRIES EXPORTING SIGNIFICANT VOLUMES OF BOTH TYPES OF COFFEE FEBRUARY 2015. 2015a. Disponível em: <<http://dev.ico.org/prices/m1a.htm>>. Acesso em: 19/04.
- _____. EXPORTING COUNTRIES: TOTAL PRODUCTION CROP YEARS COMMENCING: 2009 TO 2014. 2015b. Disponível em: <<http://dev.ico.org/prices/po.htm>>. Acesso em: 19/04.
- Ketting, R.F., Fischer, S.E., Bernstein, E., Sijen, T., Hannon, G.J., and Plasterk, R.H. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev.* 15, 2654–2659.
- KULCHESKI FR, DE OLIVEIRA LFV, MOLINA LG, ALMERA MP, RODRIGUES FA, MARCOLINO J, BARBOSA JF, STOLF-MOREIRA R, NEPOMUCENO AL, MARCELINO-GUIMARAES FC, ABDELNOOR RV, NASCIMENTO LC, CARAZZOLLE MF, PEREIRA GAG, MARGIS R. Identification of novel soybean microRNAs involved in abiotic and biotic stresses. *Bmc Genomics* 12. doi:10.1186/1471-2164-12-307, 2011.
- Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S., and Kim, V.N. (2003). The nuclear RNase III Drosha initiates microRNA processing. *Nature* 425, 415–419.
- Lau, N.C., Lim, L.P., Weinstein, E.G., and Bartel, D.P. (2001). An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* 294, 858–862.
- LLAVE, C. MicroRNAs: more than a role in plant development? *Mol. Plant Pathol.*, 5, pp. 361–366, 2004. MARGIS, R. et al. The evolution and diversification of Dicers in plants. *FEBS Letters*, v. 580, n. 10, p. 2442-2450, 5/1/ 2006.
- RUEGGER, M, DEWEY, E, GRAY, WM, HOBBIE, L, TURNER, J, ESTELLE, M. The TIR1 protein of *Arabidopsis* functions in auxin response and is related to human SKP2 and yeast Grr1p. *Genes Dev* 12: pp. 198-207, 1998.
- SCHWARZ, D.S.(2003). Asymmetry in the assembly of the RNAi enzyme complex. *Cell* 115, 199–208.
- SUNKAR AND ZHU.R. SUNKAR, J.K. ZHU. Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*. *Plant Cell*, 16, pp. 2001–2019, 2004.

- THAKUR, V. et al. Characterization of statistical features for plant microRNA prediction. *BMC Genomics*, v. 12, n. 1, p. 108, 2011.
- 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.
- YAMAGUCHI, A.; ABE, M. Regulation of reproductive development by non-coding RNA in *Arabidopsis*: to flower or not to flower. *Journal of Plant Research*, v. 125, n. 6, p. 693-704, Nov 2012.
- ZANCA AS, VICENTINI R, ORTIZ-MOREA FA, DEL BEM LEV, DA SILVA MJ, VINCENTZ M, NOGUEIRA FTS. Identification and expression analysis of microRNAs and targets in the biofuel crop sugarcane. *Bmc Plant Biology* 10. doi:10.1186/1471-2229-10-260, 2010.
- 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.
- ZHANG, B. H. et al. Evidence that miRNAs are different from other RNAs. *Cellular and Molecular Life Sciences*, v. 63, n. 2, p. 246-254, Jan 2006.
- ZHANG, B. H. et al. Identification and characterization of new plant microRNAs using EST analysis. *Cell Research*, v. 15, n. 5, p. 336-360, May 2005.