

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

CENTRO DE BIOTECNOLOGIA

PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA CELULAR E MOLECULAR

Bibliotecas de AGO-IP de flores de *Arabidopsis thaliana* contêm RNAs
circulares que podem atuar como esponjas de miRNAs

Autor: Érika Frydrych Capelari

Dissertação submetida ao Programa de Pós-
Graduação em Biologia Celular e Molecular do
Centro de Biotecnologia da UFRGS como requisito
parcial para a obtenção do título de Mestre;

Orientação: Rogerio Margis

Co-orientação: Franceli Rodrigues Kulcheski

Porto Alegre

Julho de 2018

Instituições e Fontes Financiadoras:

UFRGS: Universidade Federal do Rio Grande do Sul

CBiot/UFRGS: Centro de Biotecnologia da UFRGS

CNPq: Conselho Nacional de Desenvolvimento Científico e Tecnológico

CAPES: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

Agradecimentos

Primeiramente à Deus, por me amar de tal maneira que me concedeu saúde, paciência, sabedoria e força para concluir mais essa etapa em minha vida.

Ao amor da minha vida, Gian. Obrigada pelo teu amor e cuidado dia-a-dia. Obrigada pelas palavras de incentivo e ânimo sempre que foi preciso. Obrigada por ser meu confidente e melhor amigo.

Aos meus pais, Everson e Tânia. Sempre vou ser grata a vocês por todo o amor, carinho, preocupação e apoio.

Ao meu orientador, Rogerio Margis. Obrigada pela oportunidade de trabalhar e aprender contigo por todos estes anos. Obrigada pelo exemplo que és.

Aos membros do LGPP: Nureyev, Isabel, Cordenonsi, Pabulo, Henrique, Débora e Frank. Conviver diariamente com vocês é um privilégio e uma alegria. Obrigada por somarem não só na minha vida profissional, mas também na pessoal.

Lista de Figuras.....	6
Lista de Tabelas.....	7
Lista de abreviaturas	8
Resumo.....	10
1. Revisão Bibliográfica	12
1.1 Classes de RNA	12
1.2 Interferência por RNA.....	12
1.3 Biogênese de miRNAs em plantas.....	14
1.4 Regulação gênica mediada por miRNAs.....	18
1.5 RNAs endógenos competidores.....	20
1.6 Biogênese de circRNA em plantas	23
1.7 Características de circRNAs	27
1.7.1 Identificação em diferentes espécies.....	27
1.7.2 Conservação do circRNA	28
1.7.3 Estrutura do circRNA	28
1.7.4 Localização dos circRNAs	29
1.7.5 Expressão de circRNAs	30
1.8 Regulação gênica mediada por RNA circular (circRNA).....	31
1.9 circRNAs como esponja de miRNAs	31
1.10 Predição de RNAs circulares	34
1.10.1 O método CirComPara.....	35
3. Objetivos	38
3.1 Objetivo geral	38
3.2 Objetivos específicos	38
4. Material e métodos.....	39
4.1 Dados experimentais de imunoprecipitação de Argonautas (AGO-IP)	39
4.2 Bibliotecas de RNAseq e de sRNAs	39
4.3 Identificação de circRNAs nas bibliotecas de RNAseq a partir de AGO-IP	39
4.4 Contagem de reads de miRNAs e análises de mRNAs-alvo em bibliotecas AGO-IP	40
4.5 Material vegetal.....	41
4.6 Extração de RNA total e síntese de cDNA	41
4.7 Desenho de primers	42
4.8 Detecção de circRNAs e mRNAs parentais	43

4.9	Purificação de produtos de PCR.....	43
4.10	Sequenciamento.....	43
4.11	Análises estatísticas.....	43
5.	Resultados e discussão.....	45
5.1	Identificação de circRNAs em bibliotecas AGO-IP.....	45
5.2	circRNAs com sítios de ligação para miRNAs.....	47
5.3	mRNAs-alvo de miRNAs estão enriquecidos em bibliotecas AGO-IP.....	48
5.4	Validação de circRNAs por RT-PCR e sequenciamento.....	50
6.	Perspectivas.....	54
7.	Referências.....	55

Lista de Figuras

Figura 1. Biogênese de miRNAs e modo de ação	15
Figura 2. Localização dos domínios em Dicer.....	17
Figura 3. Estrutura cristalizada da proteína Argonata de <i>Pyrococcus furiosus</i>	18
Figura 4. RNAs endógenos competidores (ceRNAs)	21
Figura 5. Diferentes formas de <i>splicing</i>	23
Figura 6. Modelo de formação de ecircRNA.....	24
Figura 7. Remoção de íntrons do pre-mRNA pelo spliceossomo.....	25
Figura 8. Principais funções dos circRNAs	32
Figura 9. Atuação dos RNA circulares (circRNAs) na regulação gênica.....	33
Figura 10. Fluxograma do <i>software</i> CirComPara.....	36
Figura 11. Primers para detectar circRNAs e seus mRNAs parentais.....	43
Figura 12. Fluxograma para identificação de circRNAs, miRNAs e mRNAs-alvo em bibliotecas AGO-IP e controle.....	44
Figura 13. Confirmação por RT-PCR do circ_At2g35940.....	50
Figura 14. Confirmação por RT-PCR do circ_At1g12080.....	51
Figura 15. Confirmação por sequenciamento do circ_At1g12080.....	

Lista de Tabelas

Tabela 1. CircRNAs identificados em bibliotecas AGO-IP por 5 métodos diferentes.....	45
Tabela 2. Descrição dos 12 circRNAs putativos.....	46
Tabela 3. miRNAs potencialmente associados a circRNAs nas bibliotecas AGO-IP.....	47
Tabela 4. Descrição de mRNAs-alvo de miRNAs presentes em circRNAs.....	48
Tabela 5. <i>Primers</i> utilizados para RT-PCR e RT-qPCR.....	49

Lista de abreviaturas

AGO: Argonauta

AGO-IP: imunoprecipitação de Argonauta

Asp: ácido aspártico

cDNA: DNA complementar

ceRNA: RNA competidor endógeno (do inglês, *competing endogenous RNA*)

circRNA: RNA circular (do inglês, *circular RNA*)

DCL: *Dicer-Like*

DRB: proteína de ligação a RNA de fita dupla (do inglês, *double-stranded RNA binding protein*)

dsRNA: RNA de fita dupla (do inglês, *double-stranded RNA*)

EcircRNA: RNA circular exônico (do inglês, *exonic circular RNA*)

EicircRNA: RNA circular exônico-intrônico (do inglês, *exonic-intronic circular RNA*)

gDNA: DNA genômico

His: histidina

IcircRNA: RNA circular intrônico (do inglês, *intronic circular RNA*)

IP: imunoprecipitação

IRES: sitio interno de entrada do ribossomo (do inglês, *internal ribosome entry site*)

lncRNA: RNA longo não-codificador (do inglês, *long non-coding RNAs*)

MRE: elemento de resposta a miRNA (do inglês, *miRNA responsive element*)

ncRNA: RNA não-codificador (do inglês, *non-coding RNAs*)

NGS: sequenciamento de nova geração (do inglês, *next generation sequencing*)

nt: nucleotídeos

pb: pares de bases

pre-miRNA: miRNA precursor

pri-miRNA: miRNA primário

PTGS: Silenciamento gênico pós-transcricional (do inglês, *post-transcriptional gene silencing*)

RBP: proteína de ligação ao RNA (do inglês, *RNA binding protein*)

RNAi: RNA de interferência

rRNA: RNA ribossômico

shRNA: pequeno RNA hairpin (do inglês, *short hairpin RNA*)

siRNA: pequeno RNA de interferência (do inglês, *small interfering RNA*)

sncRNA: pequeno RNA não-codificador (do inglês, *small non-coding RNAs*)

snoRNA: pequeno RNA nucleolar (do inglês, *small nucleolar RNAs*)

sRNA: pequeno RNA (do inglês, *small RNAs*)

TGS: Silenciamento gênico transcricional (do inglês, *transcriptional gene silencing*)

tRNA: RNA transportador

Resumo

RNAs endógenos competitivos (ceRNAs) são transcritos naturais que atuam como esponjas endógenas de miRNAs, modulando a ação de miRNAs sobre mRNAs-alvo. RNA circular (circRNA) é uma dentre as várias classes de ceRNA. circRNAs são produzidos a partir de um processo chamado *backsplicing*. CircRNAs já foram identificados em diversos eucariotos; no entanto, nas plantas, ainda não foi demonstrado se estes são capazes de atuar como esponjas eficazes de miRNAs. Dados públicos de RNAseq de bibliotecas de imunoprecipitação de Argonauta (AGO-IP) a partir de flores de *Arabidopsis thaliana* foram utilizados em um método multi-comparativo de análises de bioinformática para identificar potenciais RNAs circulares. Utilizando cinco diferentes métodos, foram preditos de 15 a 27812 circRNAs com pelo menos dois *reads* na junção do *backsplicing*. Posteriormente, a plataforma psRNAtarget foi utilizada para discriminar os circRNAs com sítios de ligação de miRNAs, que potencialmente possam estar atuando como esponjas. Como a AGO forma um complexo ternário com miRNA e mRNA-alvo, também foram comparadas as contagens de alvos em bibliotecas AGO-IP e controle, demonstrando que os mRNAs-alvo desses miRNAs também estão enriquecidos nas bibliotecas AGO-IP. Neste trabalho, foram encontrados dois prováveis circRNAs que podem potencialmente funcionar como esponjas de miRNA. Esse achado contribui para um melhor entendimento desta nova forma de regulação transcricional e pós-transcricional em plantas, ampliando o conhecimento e aplicações do tema.

Abstract

Competing endogenous RNAs (ceRNAs) are natural transcripts that act as endogenous sponges of miRNAs, modulating the action of miRNAs on mRNA targets. Circular RNA (circRNA) is one among the various classes of ceRNA. They are produced from a process called backsplicing. CircRNAs have been identified in several eukaryotes; however, in plants, it has not yet been demonstrated whether they are able to act as effective sponges for miRNAs. Public RNAseq data from Argonaute immunoprecipitation libraries (AGO-IP) from *Arabidopsis thaliana* flowers were used with a multi-comparative method of bioinformatics analyzes to identify putative circular RNAs. Using five different methods, 15 to 27812 circRNAs with at least two reads at the backsplicing junction were predicted. Subsequently, the psRNAtarget platform was used to discriminate circRNAs harboring miRNA binding sites, which could potentially be acting as sponges. Identities and amounts of mRNAs present in AGO-IP and control libraries were quantified, as AGO can also form a ternary complex among miRNA and their target mRNAs. It showed that target mRNAs from circRNA:miRNA pairs were also enriched in the AGO-IP libraries. Two circRNAs were positively identified, corroborating their action as potential miRNA sponges. This finding leads to a better understanding of a new form of transcription and post-transcription regulation in plants, increasing the knowledge and applications of the topic.

1. Revisão Bibliográfica

1.1 Classes de RNA

Duas importantes classes de RNA que participam da expressão gênica, e foram descobertos nas décadas de 50 e 60 são os RNAs codificadores, tendo como principal representante o RNA mensageiros (mRNA) e os RNAs não-codificadores (ncRNAs), como RNA transportador (tRNA) e RNA ribossômico (rRNA). Por vários anos, acreditou-se que os ncRNAs não possuíam outras funções, se não a de participar na síntese proteica. Porém, a partir da década de 80, outros tipos de RNAs começaram a ser descobertos e acrescentados à classe de ncRNAs. Como por exemplo, os pequenos RNAs nucleolares (snRNAs), que surgiram como possíveis reguladores na excisão de íntrons, e também os micro RNAs (miRNAs). Além disso, em 1993, os laboratórios de Ambros e de Ruvkun anunciaram a descoberta de um pequeno RNA que era capaz de controlar o tempo de transições de desenvolvimento em *C. elegans* por pareamento de bases em sequências parcialmente complementares na região 3' UTR do mRNA-alvo (LEE; FEINBAUM; AMBROS, 1993). Inicialmente, este RNA foi considerado uma “estranheza”, mas em seguida foram relatados centenas de RNAs com aproximadamente 22 nucleotídeos semelhantes, agora chamados miRNAs, em *Drosophila*, células humanas, plantas, entre outros (LAGOS-QUINTANA, 2001; LAU et al., 2001; LEE; AMBROS, 2001; RUVKUN et al., 2000).

A revolução dos ncRNAs ganhou enorme impulso com a descoberta de miRNAs e seus muitos congêneres, ressaltando sua importância nos eventos pós-transcricionais na expressão gênica, particularmente em organismos eucarióticos. Hoje, sabe-se que este mundo dos ncRNAs envolve todos os organismos vivos, uma vez que o sequenciamento em grande escala tem descoberto a existência de milhares de longos ncRNAs (lncRNAs) com uma variedade de papéis impressionantes na expressão de genes e na remodelação do genoma eucariótico.

1.2 Interferência por RNA

A interferência por RNA (RNAi) é um mecanismo celular responsável pelo silenciamento gênico pós-transcricional (*Post Transcription Gene Silencing* - PTGS), que atua sobre o mRNA. Na essência deste mecanismo está uma molécula de RNA de fita dupla (dsRNA – *double stranded* RNA) que, ao ser incorporada a um complexo intracitoplasmático, se liga a uma sequência de nucleotídeos complementares localizada no mRNA-alvo, ocasionando assim o silenciamento, por

inibição da tradução e/ou degradação do mRNA. Há evidências de que dsRNAs estejam também envolvidos com a manutenção da condensação de regiões da cromatina e com a supressão da transcrição nas proximidades destas regiões. Entretanto a associação entre o silenciamento transcricional (*Transcriptional Gene Silencing* – TGS) e RNAi ainda não está completamente esclarecida. Em fungos, a deleção de genes relacionados à via de RNAi causa perda do silenciamento gênico e desestruturação da heterocromatina. Embora esse fenômeno ainda não seja bem compreendido em nível molecular, parece ter grande importância para o funcionamento adequado de genes e para a manutenção da integridade genômica (PETERSEN, C.P., DOENCH, J.G., GRISHOK, A. AND SHARP, 2005).

O RNAi foi descrito pela primeira vez em petúnias, no início da década de 1990 (NAPOLI, 1990), onde plantas transgênicas que superexpressavam genes para produção de pigmentos apresentavam flores brancas devido à inibição da síntese do pigmento, por silenciamento do transgene e do gene endógeno. Esse fenômeno, conhecido como “co-supressão”, também foi observado em outras espécies de plantas, fungos e outros organismos. Entretanto, o mecanismo que levava ao silenciamento gênico ainda era desconhecido (COGONI et al., 1996).

Em 1998, os pesquisadores norte-americanos, Andrew Z. Fire e Craig C. Mello elucidaram o silenciamento gênico por RNAi (FIRE et al., 1998). A partir de então, esse mecanismo foi sendo descrito com ocorrência em vários organismos eucariotos (TIJSTERMAN; KETTING; PLASTERK, 2002; ULLU; TSCHUDI; CHAKRABORTY, 2004). Os dsRNAs envolvidos no RNAi podem ser classificados de acordo com sua origem e função em pelo menos três categorias: siRNAs (*short interfering RNAs*), shRNAs (*short hairpin RNAs*) e miRNAs (microRNAs) (MEISTER; TUSCHL, 2004).

Os siRNA são moléculas, tanto endógenas quanto sintéticas, de dupla fita de RNA de 19 a 30 pb, que atuam por meio de pareamento a sequências complementares ao mRNA-alvo, causando sua degradação e, portanto, silenciamento específico do gene. O silenciamento do gene-alvo possibilita a obtenção de informações sobre a função exercida na célula pela proteína que esse gene codifica (MORRIS; CHAN; JACOBSEN, 2004).

shRNAs são RNAs de dupla fita construídos para apresentarem uma estrutura semelhante aos miRNAs. Podem ser sintetizados e introduzidos diretamente na célula ou transcritos dentro da célula a partir de vetores que codificam o shRNA junto a um promotor da RNA polimerase III.

Nesse caso, o transcrito é processado pela Dicer da mesma forma que os miRNAs (PADDISON et al., 2002).

Os miRNAs representam pequenos RNAs endógenos de fita simples, com aproximadamente 22 nucleotídeos, cuja principal função é atuar como silenciadores pós-transcricionais, seja pela clivagem do mRNA alvo ou pela inibição de sua tradução em proteína. Sua descoberta data de pouco mais de duas décadas, em estudos realizados com *Caenorhabditis elegans*, e são hoje reconhecidos como reguladores fundamentais da expressão gênica em plantas e animais. Em 2002 foi descrito o primeiro miRNA em *Arabidopsis thaliana* (RHOADES et al., 2002), e atualmente já foram identificadas 329 sequências nesta mesma espécie, de acordo com a versão de 2018 do miRBase, que é um banco de dados público (KOZOMARA; GRIFFITHS-JONES, 2014).

1.3 Biogênese de miRNAs em plantas

A biogênese de um miRNA diz respeito a todos os eventos, desde a transcrição do gene, como sua maturação, até chegar ao miRNA maduro integrado ao complexo RISC (Figura 1). miRNAs são RNAs de 20-24 nucleotídeos produzidos a partir de precursores de RNA em forma de grampo (*hairpin*), que através de pareamento com sítios totalmente ou parcialmente complementares em regiões 3' não traduzidas (UTRs) do mRNA, medeiam o silenciamento pós-transcricional de genes codificadores de proteína.

Os genes de miRNAs são transcritos, na maioria das vezes, pela RNA polimerase II (JONES-RHOADES; BARTEL; BARTEL, 2006), e geram um longo transcrito de miRNA-primário (pri-miRNA), que pode incluir outros miRNAs e também éxons codificadores de proteínas. Alguns miRNAs são relatados como sendo processados a partir de ncRNAs bem caracterizados, como os tRNAs ou snoRNAs.

O pri-miRNA apresenta uma estrutura de grampo ou *stem-loop*, ou seja, um tronco (*stem*) que apresenta 33-35 pares de bases, um terminal em forma de alça (*loop*) e segmentos de fita simples (braços) de RNA tanto no lado 5' como no lado 3'. O grampo do pri-miRNA é caracterizado por um eixo de dsRNA com alguns nucleotídeos não pareados e conectados por uma região terminal de RNA fita simples. Além disso, os pri-miRNAs possuem uma 7-metilguanossina

(CAP) na região 5' terminal, são poliadenilados na região 3' e apresentam tamanho variado, podendo ser constituídos por mais de mil nucleotídeos (AXTELL; WESTHOLM; LAI, 2011).

O processamento de pri-miRNAs ocorre no núcleo, por enzimas contendo domínios de RNaseIII, e pode ser realizado de forma simultânea com o processo de transcrição (KIM; KIM, 2007; MORLANDO et al., 2008). A clivagem do pri-miRNA é feita por uma RNase III chamada *Dicer-like* que associada com outras proteínas, como a proteína HYL1, também chamada de DRB1 (*Double-Stranded RNA Binding protein*), resulta no miRNA precursor, chamado pre-miRNA. O pre-miRNA possui de 70 a 100 pares de bases e contém um trecho de fita dupla e uma alça de fita simples, formando uma estrutura denominada *hairpin*, e quando clivado pela Dicer dá origem a um duplex de miRNA, com tamanho variando entre 19 e 22 pares de bases. Uma das fitas deste duplex é escolhida seletivamente para ser incorporada ao Complexo de Silenciamento Induzido por RNA (RISC), sendo denominada “fita guia” ou “miRNA maduro”, enquanto que a outra fita é degradada, sendo denominada “passageira” ou miRNA*. Geralmente, a fita do duplex com menor estabilidade na região 5' é quem origina o miRNA maduro. Após as extremidades 3' não pareadas serem metiladas pela enzima HUA ENHANCER1 (HEN1), este duplex é exportado para o citoplasma através da exportina-5, onde uma das fitas é degradada e a outra é incorporada à proteína Argonata e inserida no RISC.

Tanto os precursores quanto os miRNAs maduros estão sujeitos a degradação regulada (AMERES; ZAMORE, 2013), envolvendo frequentemente a adição ou modificação de nucleotídeos na região terminal 3', seja pela falta da metilação nas extremidades não pareadas do duplex, ou pela adição de U, como sinal para degradação pelo proteassomo.

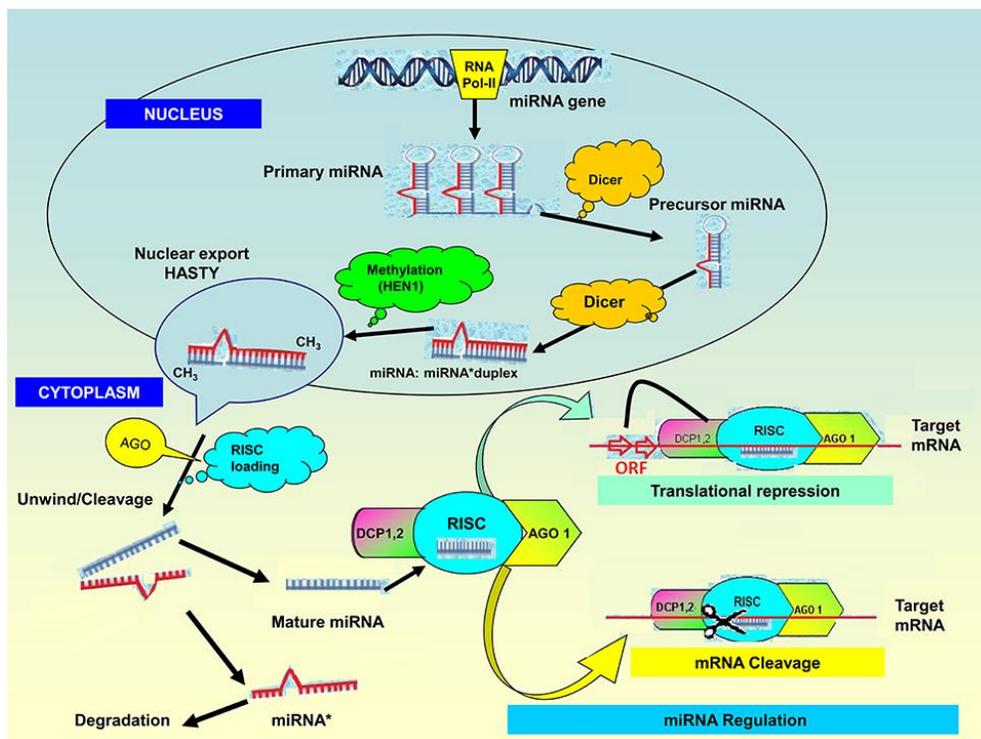


Figura 1. Biogênese de miRNAs e modo de ação. A biogênese do miRNA começa no núcleo, onde o gene miRNA é geralmente transcrito pela RNA polimerase II em um longo transcrito primário (pri-miRNA). Em seguida, é clivado pela enzima nuclear semelhante a RNase III, chamada DICER-LIKE (DCL1) gerando um mRNA precursor (pre-miRNA). Novamente ocorre uma clivagem mediada pela DLC em associação com outras proteínas, como HYL1, originando um duplex miRNA:miRNA*. Após ter suas extremidades 3' livres metiladas pela RNA metiltransferase HEN1, o duplex é exportado para o citoplasma por um poro nuclear, onde vai ser incorporado ao complexo de silenciamento induzido por RNA (RISC). O miRNA* é liberado e degradado, enquanto que o miRNA maduro é carregado pela proteína ARGONAUTE (AGO) para o RISC. Para o miRNA exercer sua função na regulação da expressão gênica, seja inibindo a transcrição ou bloqueando a tradução do mRNA em proteína, é necessário que haja o emparelhamento de bases entre o miRNA e o mRNA alvo. Fonte: (DJAMI-TCHATCHOU et al., 2017).

Em *Arabidopsis* existem quatro proteínas RNase III DICER-LIKE (DCL) e tem sido proposto que cada uma requer a associação com uma proteína DRB, a fim de exercer sua função

(HIRAGURI et al., 2005; MARGIS et al., 2006; NAKAZAWA et al., 2007). A DCL1, localizada no núcleo, reconhece moléculas de RNA que têm a capacidade de formar estruturas em forma de grampo e clivá-las, liberando miRNAs de 21 nt que posteriormente serão usados no complexo RISC para silenciar a expressão de mRNAs-alvo (KURIHARA, 2005; XIE; KASSCHAU; CARRINGTON, 2003). Alguns estudos sugerem que a DCL2 pode competir com a DCL1 pela ligação à DRB1, limitando assim, a produção de alguns miRNAs (HAN et al., 2004). A principal atividade da DCL3 é processar longos dsRNAs de regiões da heterocromatina do genoma em siRNAs de 24 nt (DAXINGER et al., 2009). Já a DCL4, de forma semelhante a DCL1, participa do processamento de pequenos RNAs de 21 nt.

Proteínas Dicer-like são caracterizadas por possuírem oito domínios (MARGIS et al., 2006) (Figura 2). Um domínio DExD (domínio helicase), um domínio Helicase-C (com a região carboxi-terminal conservada), um domínio Duf283, um domínio PAZ (PIWI/Argonatura/Zwille), dois domínios de RNase III (que se ligam diretamente ao domínio PIWI de proteínas Argonauta) e dois domínios dsRB (domínio de ligação a dsRNA), sendo que estes três últimos domínios, PAZ, RNase III e dsRB são os responsáveis pela ligação e clivagem do dsRNA.

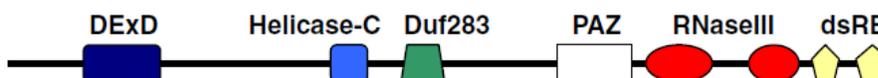


Figura 2. Localização dos domínios em Dicer-like. Disposição linear de domínios tipicamente encontrados em proteínas Dicer-like. DExD, domínio helicase; Helicase-C, domínio helicase com carboxi-terminal conservado; Duf283, domínio de função desconhecida; PAZ, domínios Piwi/Argonatura/Zwille; RNaseIII, dois domínios RNase III; dsRB, domínio de ligação a RNA fita dupla. (Adaptado de MARGIS et al., 2006).

A Dicer é uma RNase III essencial no processo de RNAi e apresenta um domínio de ligação ao RNA, chamado PAZ, que determina o tamanho do dsRNA gerado. Além disso, a Dicer também está envolvida na incorporação de uma das fitas do miRNA a um grande complexo de proteínas efetoras, chamado Complexo de Silenciamento Induzido por RNA (*RNA Induced Silencing Complex* – RISC).

1.4 Regulação gênica mediada por miRNAs

Plantas utilizam uma sofisticada maquinaria de silenciamento de RNA para controlar a expressão de genes relacionados aos mais diversos processos moleculares que afetam o desenvolvimento e a fisiologia das mesmas. Uma das principais vias de silenciamento que ocorre pós-transcricionalmente é aquela mediada pelos miRNAs, sendo que aproximadamente dois terços dos miRNAs de plantas controlam a expressão de fatores de transcrição reguladores do desenvolvimento (JONES-RHOADES; BARTEL, 2004; RHOADES et al., 2002).

A regulação gênica mediada por miRNA ocorre por meio de diversos e distintos mecanismos, como supressão da tradução, desestabilização do mRNA e clivagem do mRNA (WINTER et al., 2009), sendo que o envolvimento de sRNAs em fenômenos de silenciamento transcricional é melhor caracterizado em plantas e leveduras (VOLPE, 2002). Cada miRNA tem como alvo múltiplos mRNAs e grande parte dos mRNAs são regulados por múltiplos miRNAs (CHI et al., 2009). Dessa forma, as interações entre miRNAs ligado aos mesmos mRNAs contribuem de forma importante para as redes regulatórias de genes (GURTAN; SHARP, 2013).

Assim como proteínas reguladoras da expressão gênica, os miRNA contribuem de forma importante para o controle dos processos de desenvolvimento, de diferenciação e de doenças (GURTAN; SHARP, 2013). É interessante destacar que certos vírus aproveitam os miRNAs de hospedeiros específicos para suas funções virais, como a replicação do genoma (JOPLING, 2005), enquanto outros adquiriram e manipularam genes de miRNA de hospedeiros para aumentar sua infecção (SKALSKY; CULLEN, 2010). Além disso, miRNAs atuam em uma série de processos biológicos, como desenvolvimento (LELANDAIS-BRIERE et al., 2010), sinalização hormonal (LIU; CHEN, 2009), floração e determinação sexual (CHUCK; CANDELA; HAKE, 2009) e adaptação a estresses diversos (CHEN; CHEN; CHUANG, 2015; LU et al., 2008; MALLORY; VAUCHERET, 2006; SHUKLA; CHINNUSAMY; SUNKAR, 2008).

O complexo RISC permite o pareamento entre a fita do miRNA incorporada e a região homóloga do mRNA-alvo por complementaridade de bases. Normalmente, quando a complementaridade é total, ocorre a clivagem do mRNA e quando parcial, principalmente na região *seed* do miRNA, entre as bases 10-11, ocorre repressão da tradução e posterior degradação do mRNA (RIVAS et al., 2005). Além da Dicer, outras proteínas fazem parte do complexo RISC, como a HYL1 (*HYPONASTIC LEAVE 1*), também chamada de DRB1 (do inglês *dsRNA Binding*

Protein), a proteína de ligação a dsRNA TGH (*TOUGH*) (REN et al., 2012), a fosfatase CPL1 (*C-TERMINAL DOMAIN PHOSPHATASE-LIKE1*) (MANAVELLA et al., 2012), a SE (*SERRATE*) e a AGO (*ARGONAUTE*) (RICARTE FILHO; KIMURA, 2006), entre outras.

Plantas possuem 10 ou mais proteínas AGO, que de acordo com comparação filogenética, são divididas em três clados principais. E *Arabidopsis*, o clado 1 engloba as proteínas AGO1, AGO5 e AGO10. O clado 2, AGO2, AGO3 e AGO7. Já as proteínas AGO4, AGO6, AGO8 e AGO9 compõem o clado 3 (MOREL et al., 2002; VAUCHERET, 2008). No entanto, esta classificação baseia-se exclusivamente na similaridade proteica e não indica necessariamente semelhança funcional. As proteínas Argonautas apresentam atividade de endonuclease dirigida contra a fita de mRNA complementar ao miRNA e são também responsáveis pela seleção da fita do miRNA que será incorporada ao RISC (HÖCK; MEISTER, 2008; RAND et al., 2005).

Argonautas são definidas por possuírem três domínios funcionais principais conservados, que participam em conjunto para posicionar corretamente a sequência do miRNA em relação ao mRNA alvo. São eles: os domínios PAZ, MID e PIWI (Figura 3). A região N-terminal é uma região variável, que pode facilitar a separação do duplex miRNA:mRNA-alvo, e possui o domínio PAZ, que foi mostrado por ancorar na região 3' do pequeno RNA vinculado. A região C-terminal é composta pelos domínios MID e PIWI, que, devido a uma ligação na junção desses domínios, ancora à extremidade 5' do RNA (PARKER, 2010; PATEL et al., 2011).

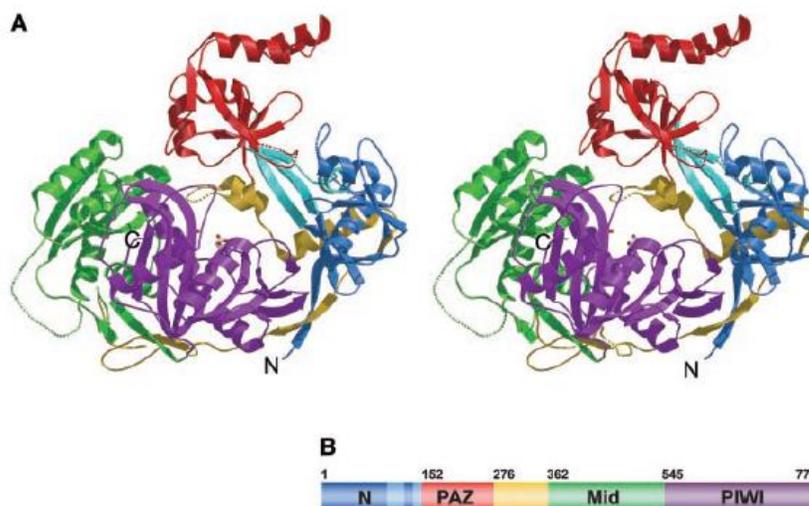


Figura 3. Estrutura cristalizada da proteína Argonauta de *Pyrococcus furiosus*. (A) Visão geral da proteína Argonauta mostrando o domínio N-terminal (azul), o domínio *stalk* (azul claro),

o domínio PAZ (vermelho), o domínio médio (verde), o domínio PIWI (roxo) e o conector interdomínio (amarelo). (B) Diagrama esquemático das bordas do domínio. Fonte: Adaptado de (SONG, 2004).

O domínio PIWI das AGOs de *Arabidopsis thaliana* contém um núcleo catalítico composto pelos aminoácidos Asp-Asp-His [DDH] ou Asp-Asp-Asp [DDD] e sua estrutura e local catalítico se assemelham bastante à de uma proteína contendo domínio RNase H de *Bacillus holodurans* (SONG, 2004). Através de análises mutacionais que desativam o núcleo catalítico DDH ou DDD, foi demonstrado que o domínio PIWI das AGOs é fundamental para executar o corte de alvos de RNA (CARBONELL et al., 2012).

Tanto em animais como em plantas, a imunoprecipitação de proteínas AGO seguida de sequenciamento de RNAs associados a estas proteínas imunoprecipitadas revela que diversas proteínas AGO interagem com diversos sRNAs que variam muito em sequência. A escolha de sRNAs associados às AGOs é influenciada pelo duplex de sRNAs. Além disso, a imunoprecipitação de diversas Argonautas de *Arabidopsis* indica que a identidade do nucleotídeo na região 5', juntamente com o tamanho do sRNA, contribuem para a triagem de sRNAs nessas proteínas (Kim, 2008). Por exemplo, a maioria dos sRNAs associados com AGO1 tem um U no 5' e tem 21 ou 22 nucleotídeos de comprimento (FRANCO-ZORRILLA et al., 2007; MI et al., 2008), enquanto a maioria dos sRNAs associados a AGO2 tem um A no 5' e 21 nucleotídeos de comprimento (MI et al., 2008; MONTGOMERY et al., 2008; TAKEDA et al., 2008).

1.5 RNAs endógenos competidores

RNAs endógenos competidores (ceRNAs – do inglês *competing endogenous RNAs*) são transcritos naturais de genes codificadores e não codificadores de proteínas (Figura 4). Os ceRNAs são caracterizados por possuírem elementos de respostas a miRNAs (MREs – do inglês *microRNAs Responsive Elements*) podendo, portanto, se ligar competitivamente aos miRNAs, sequestrando-os de seus mRNAs-alvo primários. Além disso, ceRNAs contém MREs para uma combinação de diferentes miRNAs. Desta forma, atuam na regulação da expressão gênica, impedindo a ação repressora dos miRNAs sobre seus alvos, através de uma diminuição da disponibilidade de

miRNAs no citoplasma (PARKER, 2010; SALMENA et al., 2011; TAY; RINN; PANDOLFI, 2014).

Um dos primeiros exemplos de ceRNA descrito em plantas demonstrou que o RNA não codificador IPS1 (*Induced by Phosphate Starvation 1*) possui um efeito sequestrador efetivo sobre o miR-399, sendo capaz de mimetizar o MRE de seu mRNA-alvo PHO2 (*Putative Ubiquitin-conjugating Enzyme E2 24*). Esse mecanismo foi chamado de *target mimicry* ou mimetismo de alvo (FRANCO-ZORRILLA et al., 2007). Em 2011, foi levantada a hipótese de que existe uma rede de interação entre os ceRNAs, na qual eles estariam se comunicando e co-regulando seus níveis citoplasmáticos através da competição por um conjunto limitado de miRNA (SALMENA et al., 2011). Ou seja, todos os transcritos que contenham MREs em comum podem se comunicar e regular uns aos outros, competindo especificamente por miRNAs compartilhados (Figura 4).

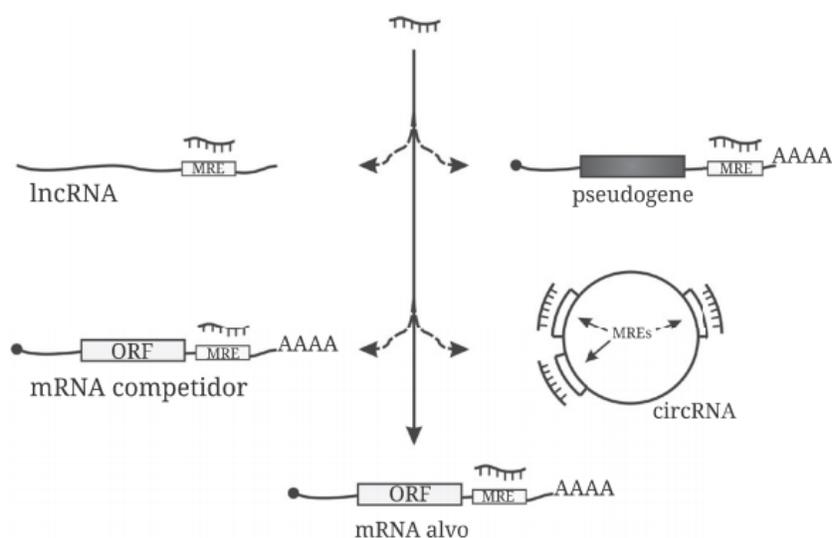


Figura 4. RNAs endógenos competidores (ceRNAs). Diversos RNAs podem atuar como ceRNAs e sequestrar os miRNAs de seus mRNAs-alvo. Dentre eles: *long non-coding* RNAs (lncRNAs), outros mRNAs competidores, pseudogenes e também RNAs circulares (circRNAs). ORF: Quadro aberto de leitura (Open Reading Frame); AAAA: cauda poli-A; MRE: Elemento de resposta a miRNA (miRNA Responsive Element) (PEREIRA, 2015).

Recentemente, uma classe adicional de ceRNA foi identificada, correspondendo a RNAs circulares (circRNA, do inglês *circular RNA*) (MEMCZAK et al., 2013). CircRNAs são uma classe

distinta de RNAs não-codificadores endógenos que resultam de uma forma não-canônica de *splicing* alternativo (HOUSELEY et al., 2006; WILUSZ; SHARP, 2013) e que estão envolvidos na regulação da expressão em níveis de transcrição e pós-transcrição (Hansen, T.B. et al., 2013). A primeira evidência sobre a existência de circRNA foi demonstrada em 1976, quando Sanger e colaboradores demonstraram que alguns viroides de plantas são compostos de moléculas de RNA de fita simples, covalentemente fechadas (SANGER et al., 1976). Alguns anos depois, também foi observado que o vírus da hepatite delta (HDV), um vírus animal, tem um genoma de circRNA (KOS et al., 1986). Posteriormente, quando outros pesquisadores investigavam um gene supressor de tumor em células humanas, observaram que alguns transcritos apresentavam éxons em uma ordem diferente daquela que estava presente no DNA genômico, o que sugeria a presença de circRNAs (COCQUERELLE et al., 1992).

Embora já houvesse certo conhecimento sobre a existência de transcritos circulares há pelo menos vinte anos, por muito tempo essas moléculas foram consideradas como artefatos provenientes do processamento de RNA ou como transcritos com ordem de éxons embaralhada devido a um processo de *splicing* incomum (CAPEL et al., 1993; COCQUERELLE et al., 1992; NIGRO et al., 1991). Já que não podiam ser sequenciados e mapeados nos genomas, os circRNAs eram previamente descartados como ruído transcricional nos eucariotos. Há pouco tempo que os circRNAs foram adicionados à crescente lista de RNAs não-codificadores (ncRNAs) e atualmente têm despertado grande atenção na comunidade científica por dois principais motivos.

O primeiro deles refere-se às evidências de que uma grande quantidade de circRNAs endógenos são expressos nas mais diversas espécies, dentre elas mamíferos (MEMCZAK et al., 2013; RYBAK-WOLF et al., 2015), nematoides (IVANOV et al., 2015), *archaea* (DANAN et al., 2012), plantas (LU et al., 2015a; WANG et al., 2017; YE et al., 2015; ZHAO et al., 2017), entre outras. Além disso, alguns circRNAs parecem ser conservados entre diferentes espécies (JECK; SHARPLESS, 2014; MEMCZAK et al., 2013; SALZMAN et al., 2013).

O segundo motivo refere-se às principais funções propostas para os circRNAs:

- (i) regulação do processamento alternativo de RNAs (ASHWAL-FLUSS et al., 2014; ZHANG et al., 2014b);
- (ii) regulação da transcrição (CHAO et al., 1998);

(iii) interação com proteínas que se ligam a RNA e interação com complexos ribonucleoproteicos (HENTZE; PREISS, 2013; ROMEO, 1998);

(iv) esponjas de miRNAs (HANSEN et al., 2013; MEMCZAK et al., 2013).

1.6 Biogênese de circRNA em plantas

Os circRNAs podem ser originados a partir de éxons (circRNA exônico, EcircRNA) (JECK et al., 2013; NIGRO et al., 1991), de íntrons (circRNA intrônico, IcircRNA) (TALHOUARNE; GALL, 2014; ZHANG et al., 2013) ou de ambos (circRNA exônico-intrônico, EircRNA) (LI et al., 2015). Até mesmo íntrons de tRNA podem formar RNAs circulares estáveis (TcircRNA) (LU et al., 2015b). Entretanto, a maioria das moléculas de RNAs circulares eucarióticas são oriundas de éxons de genes codificadores de proteínas durante o *splicing*, que é um processo celular geralmente catalisado pela maquinaria spliceossomal ou pelas ribozimas dos grupos I e II (SUZUKI et al., 2006) (Figura 5).

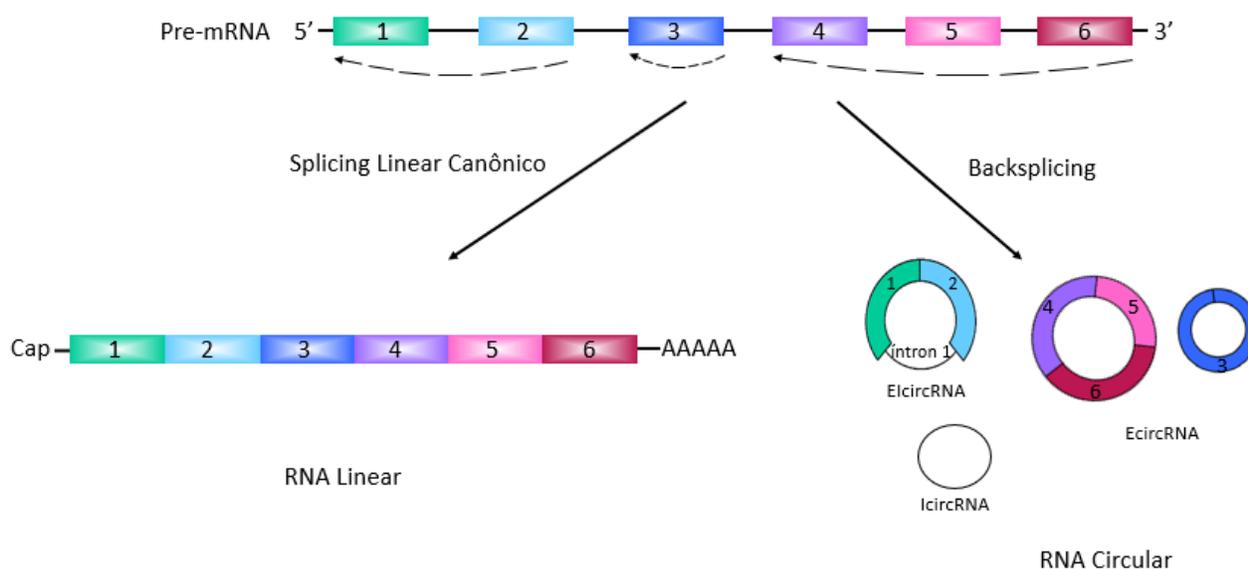


Figura 5. Diferentes formas de *splicing*. A partir de um pre-mRNA é possível a ocorrência de *splicing* linear canônico, dando origem a um RNA linear terminado com um Cap na região 5' e uma sequência poliadenilada na região 3'. Também é possível a ocorrência de *backsplicing*, originando um RNA circular com as extremidades 5' e 3' unidas, sendo que este RNA circular pode ser formado por um ou mais éxons, por íntron e ainda por éxon e íntron.

Embora possam ser originados de maneiras diferentes (ANDREEVA; COOPER, 2015; CHEN; CHEN; CHUANG, 2015; CHEN; YANG, 2015), duas principais hipóteses de biogênese de circRNA são fortemente aceitas:

A primeira baseia-se em um rearranjo de éxons devido ao *splicing* diferencial sendo denominado "de cabeça para cauda", ou *backsplicing* (HANSEN; KJEMS; DAMGAARD, 2013; JECK; SHARPLESS, 2014; MEMCZAK et al., 2013). No *backsplicing* direto, éxons sofrem *splicing* em ordem não-canônica. Nesse caso, o sítio de *splice* 5' (doador) a jusante é unido a um sítio 3' (aceptor) a montante, gerando assim uma transcrição circular (Figura 6).

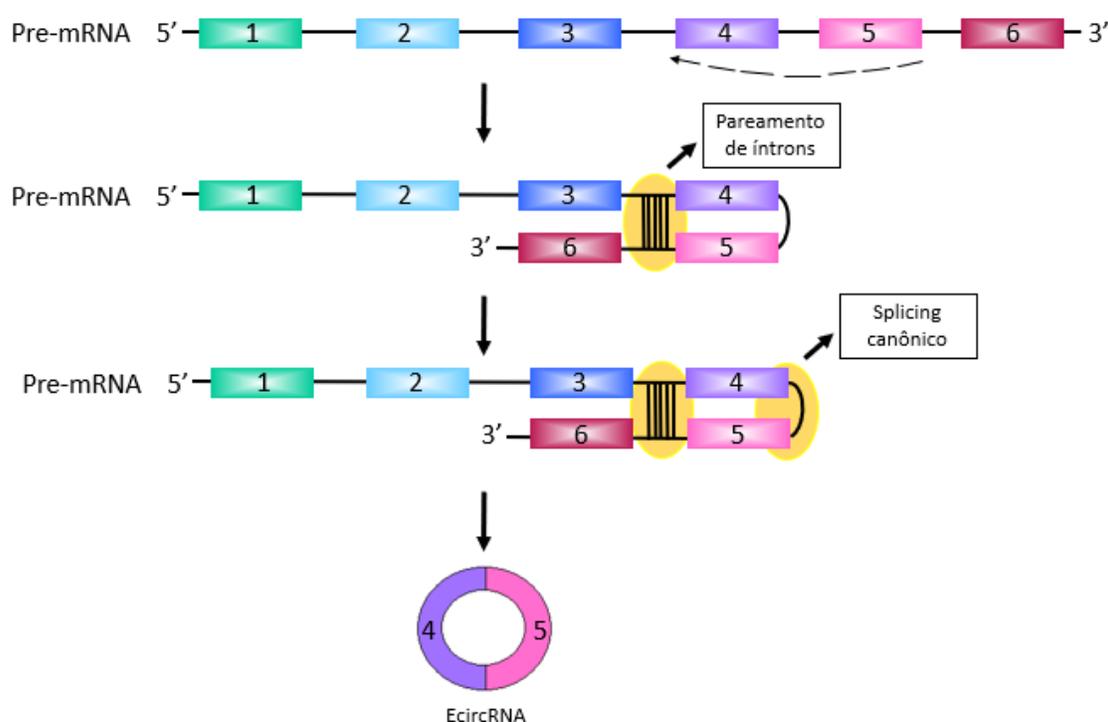


Figura 6. Modelo de formação do ecircRNA por *backsplicing*. No *backsplicing* direto ocorre o pareamento de íntrons ou pareamento através de proteína de ligação ao RNA. Inicialmente o ponto de derivação 5' a jusante ataca o sítio doador 3' a montante. A seguir, o sítio doador 3' ataca o sítio aceitador 5'. Então, as duas extremidades dos éxons liberados são conectadas após a remoção do íntron.

O segundo modelo geralmente aceito descreve uma ação do spliceossomo em dois passos (LIN et al., 2008). No primeiro, o ponto de derivação, geralmente uma adenosina, ataca o sítio de

splicing 5' através do seu grupo 2'-hidroxila, formando uma ligação fosfodiéster 2'-5' e uma extremidade 3'-hidroxila livre no éxon 5'. No segundo passo, o 3'-hidroxila recém-gerado ataca o local de união de 3', e uma ligação fosfodiéster 3'-5' é formada para ligar os dois éxons. Nesse mecanismo ocorre a formação de dois éxons unidos e o íntron *lariat* com uma ligação 2'-5' é excisado (ZAPHIROPOULOS, 1996) (Figura 7).

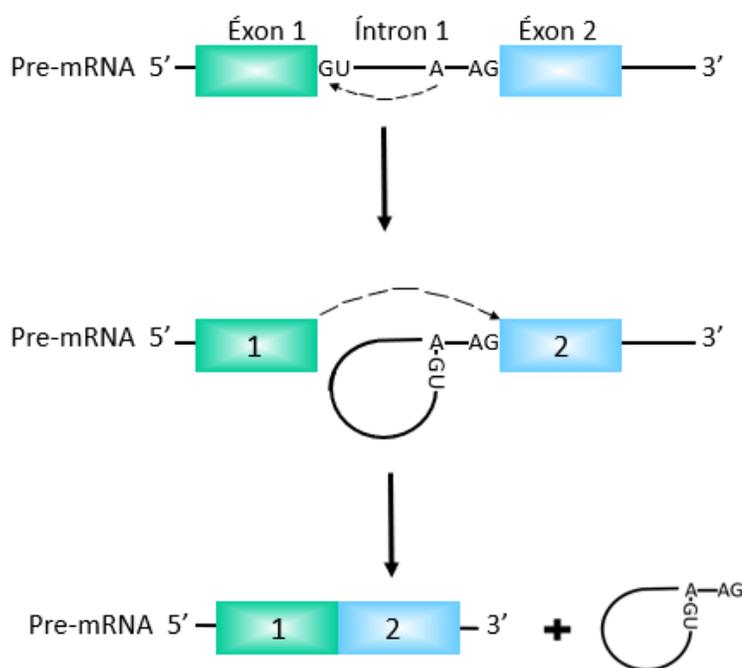


Figura 7. Remoção de íntrons do pré-mRNA pelo spliceossomo. O radical de ataque 2'-OH, de um co-fator adenosínico específico do íntron, ataca o sítio de *splicing* 5', formando uma ligação fosfodiéster 2'-5' e uma extremidade 3'-hidroxila livre no éxon 5'. A seguir, o 3'-hidroxila gerado ataca o sítio de splicing 5', dando origem a uma ligação fosfodiéster 3'-5'. Como resultado, os dois éxons são unidos e o íntron *lariat* com uma ligação 2'-5' é excisado.

Ambos os mecanismos funcionam *in vivo* e estão correlacionados com o spliceossomo canônico, mas o *backsplicing* direto ocorre com mais frequência do que a exclusão de éxon na biogênese do EcircRNA (JECK; SHARPLESS, 2014). Acredita-se também que o pareamento de RNA, que pode ser formado por elementos repetitivos (JECK et al., 2013) ou por sequências não repetitivas, mas complementares, através de íntrons flanqueadores (ZHANG et al., 2014b), está positivamente associado à circularização do éxon. Além disso, proteínas de ligação ao RNA

(RBPs, do inglês *RNA Binding Protein*), como ADAR (do inglês *Adenosine Deaminases Acting on RNA*), MBL (do inglês, *Mannose-Binding Lectin*) e QKI (do inglês, *Quaking*), também podem estar envolvidas neste processo (ASHWAL-FLUSS et al., 2014; CONN et al., 2015; IVANOV et al., 2015).

Para garantir a precisão dos produtos do mRNA e os processos consequentes é preciso um controle estrito na escolha dos sítios de *splice*. Os circRNAs tipicamente compreendem sequências exônicas que sofrem *splicing* em sítios canônicos e acredita-se que o spliceossomo esteja envolvido na formação desses sinais (ASHWAL-FLUSS et al., 2014). Entretanto, a maquinaria de *splicing* também apresenta certos graus de flexibilidade na escolha deste local e essa variância nos sítios de pareamento do *splicing* alternativo pode criar isoformas de transcritos variáveis. Dessa forma, circRNAs competem com o *splicing* do mRNA pela maquinaria spliceossomal (YE et al., 2015).

Sabe-se também que, nem sempre, os níveis de expressão de circRNAs e de seus mRNAs lineares parentais apresentam uma correlação (SALZMAN et al., 2013), sugerindo que este processo é regulado e que o spliceossomo pode distinguir entre o *backsplicing* e o *splicing* alternativo. Entretanto o mecanismo detalhado e a identificação das proteínas envolvidas ainda precisam ser esclarecidos.

Como dito anteriormente, circRNAs são predominantemente derivados de éxons, cujas extremidades 5' e 3' estão ligadas covalentemente para formar uma junção "cabeça-cauda". Esta circularização aumenta a estabilidade do circRNA, tornando-o resistente à RNase R, uma forte exoribonuclease de 3' a 5' que degrada eficientemente RNAs lineares (SUZUKI; TSUKAHARA, 2014). Além disso, a maioria dos circRNAs compreende geralmente dois ou três éxons (ZHANG et al., 2014b). Acredita-se que não seja necessária uma sequência exônica específica, além do sítio de *splicing*, para o *backsplicing* (STARKE et al., 2015). Além disso, vários circRNAs, com o íntron interno incluído ou excluído, podem ser produzidos a partir do mesmo gene por *splicing* alternativo, gerando isoformas de um mesmo circRNA (GUO et al., 2014; SALZMAN et al., 2013; ZHANG et al., 2014b). No entanto, é interessante destacar que o *backsplicing* pode exigir um comprimento mínimo de éxon. Por exemplo, RNAs circulares que possuem um único éxon tendem a ter este éxon com um comprimento muito maior do que os éxons presentes em RNAs circulares composto por vários éxons (SZABO et al., 2015; ZHANG et al., 2014b).

Por terem as extremidades 3' e 5' unidas, as técnicas moleculares tradicionais utilizadas na análise de RNA não eram capazes de detectar circRNAs, por não apresentarem extremidades 3' ou 5' livres ou mesmo um final de RNA poliadenilado. Os avanços no sequenciamento de RNA de alto rendimento têm possibilitado novas oportunidades para investigar transcriptomas, revelando a existência de uma grande quantidade de circRNAs que não eram possíveis ser identificados anteriormente, especialmente os ecircRNAs (SALZMAN et al., 2012). Análises de dados de RNAseq revelam que ecircRNAs são abundantes em transcriptomas de mamíferos, e alguns deles são evolutivamente conservados em termos de sequência e expressão (JECK et al., 2013; RYBAK-WOLF et al., 2015; SALZMAN et al., 2012), sugerindo que eles possuem função celular específica.

Geralmente, um evento de ecircRNA pode ser detectado alinhando sequências expressas (por exemplo, *reads* de RNAseq) a um genoma de referência. As bibliotecas de RNAseq mais adequadas para a identificação de circRNAs são aquelas tratadas com uma etapa de remoção de rRNAs previamente ao sequenciamento, já que a célula possui uma enorme quantidade de rRNA, e isso dificultaria a detecção de circRNAs, por estarem presentes em pequenas quantidades. Além disso, os métodos tradicionais de preparação de bibliotecas de RNAseq envolvem uma etapa anterior ao sequenciamento, em que os mRNAs são enriquecidos utilizando oligonucleotídeos contendo somente desoxitimidinas (Oligo[dT]). Desta forma os oligos dT hibridizam-se com a cauda poli(A) dos mRNAs, e, potencialmente, os circRNAs são removidos da amostra (YE et al., 2015). Após o alinhamento dos *reads* de RNAseq ao genoma de referência, as junções não-colineares (NCL) são determinadas com base na presença de dois éxons conectados em que a ordem do éxon é inconsistente com o genoma de referência.

1.7 Características de circRNAs

1.7.1 Identificação em diferentes espécies

RNAs circulares têm sido amplamente descobertos e descritos em diversas espécies, como humanos (IVANOV et al., 2015; MEMCZAK et al., 2013), *C. elegans* (IVANOV et al., 2015), archaea (DANAN et al., 2012), dentre outros.

Pouco se sabe sobre as características e funções dos circRNAs em plantas. Estas moléculas apresentam características similares aos circRNAs de animais com relação a conservação, expressão diferencial tecido-específica e presença de longos íntrons flanqueando a região onde é realizado o *backsplicing*, mas em geral não apresentam sequências repetitivas ou complementares nessas regiões (YE et al., 2015). Curiosamente, em plantas uma parcela significativa dos circRNAs realizam o *backsplicing* em sequências localizadas dentro de éxons e não em sua extremidade onde está localizada a junção canônica de *splicing* (LU et al., 2015a).

1.7.2 Conservação do circRNA

Os circRNAs exibem certa conservação de sequência. Por exemplo, 457 de 2121 circRNAs em seres humanos foram encontrados com ortólogos circulares em murinos (MEMCZAK et al., 2013) e aproximadamente 4% de genes ortólogos de humanos e ratos podem gerar circRNAs (SALZMAN et al., 2013). Esta conservação de circRNAs entre espécies também implica que os circRNAs não são subprodutos de *splicing* de pré-mRNA. No entanto, outro estudo revelou que os éxons dentro dos circRNAs ortólogos humano-rato não são, necessariamente, mais conservados do que os seus éxons lineares vizinhos (GUO et al., 2014).

1.7.3 Estrutura do circRNA

Normalmente, os circRNAs são compostos por um ou três éxons (ZHANG et al., 2014b). Às vezes, um íntron interno pode permanecer no circRNA ou pode ser excluído, produzindo isoformas de circRNAs do mesmo gene por *splicing* alternativo. Outro fator importante a ser ressaltado é que o *backsplicing* pode requerer um comprimento de éxon mínimo. Esta característica foi observada em circRNAs altamente expressos (CHEN; YANG, 2015; SZABO et al., 2015; ZHANG et al., 2014b).

Sabe-se que íntrons que flanqueiam éxons circularizados são geralmente muito mais longos do que a média e não causam a formação do circRNA por si só. No entanto, íntrons mais longos tendem a conter sequências complementares reversas, e esta é uma característica conservada em animais e plantas que contribui para a formação de circRNAs (IVANOV et al., 2015; YE et al., 2015). Além disso, os circRNAs apresentam uma deleção de polimorfismos nos sítios de ligação

ao miRNA, indicando que os locais de ligação ao miRNA em circRNAs são submetidos a pressão seletiva similar àqueles em mRNAs lineares (THOMAS; SÆTROM, 2014).

1.7.4 Localização dos circRNAs

Vários estudos têm mostrado que EcircRNAs são enriquecidos em amostras citoplasmáticas (GUO et al., 2014; JECK et al., 2013; MEMCZAK et al., 2013; SALZMAN et al., 2012). Enquanto que os EircRNAs e IcircRNAs estão localizados predominantemente no núcleo (LI et al., 2015; ZHANG et al., 2013), sugerindo que eles podem regular a transcrição gênica. Entretanto, o mecanismo de transporte ainda não é claro.

O complexo formado pela junção éxon-éxon pode recrutar fatores de exportação de mRNA durante a formação do EcircRNA pela maquinaria spliceossomal e, em seguida, os EcircRNAs são transportados pelo sistema de exportação nuclear (LE HIR et al., 2001). Como os EcircRNAs são gerados pela maquinaria spliceossomal no núcleo e podem ser encontrados em conjuntos de RNA vinculados à cromatina, eles provavelmente são transportados pelo sistema de exportação nuclear, ou escapam do núcleo durante a divisão celular.

A localização citosólica também pode suportar a função pós-transcrição de EcircRNAs. Os exemplos mais representativos são CDR1as/ciRS-7 e o circRNA Sry, que são predominantemente localizados no citoplasma e funcionam como esponjas de miRNA quando o miRNA específico (miR-7 para CDR1as/ciRS-7 e miR138 para circRNA Sry) está presente (HANSEN; KJEMS; DAMGAARD, 2013; MEMCZAK et al., 2013). No entanto, um outro estudo mostrou que alguns RNAs circulares exônicos com segmentos intrônicos retidos entre os éxons estavam predominantemente localizados no núcleo, onde eles estariam regulando seus genes parentais através de interações específicas de RNA-RNA (LI et al., 2015).

Essas observações indicam que diferentes EcircRNAs podem diferir nas suas localizações subcelulares preferenciais, sugerindo também que podem possuir funções variadas.

1.7.5 Expressão de circRNAs

RNAs circulares são expressos ao longo da árvore eucariótica da vida e acredita-se que exista uma correlação fraca entre os níveis de circRNAs e suas contrapartes lineares (SALZMAN et al., 2013). Geralmente, os circRNAs são menos abundantes e apresentam níveis de expressão mais baixos, em média < 3%, em relação a abundância do transcrito canônico de seus parentais lineares (ENUKA et al., 2016; JECK et al., 2013). No entanto, em alguns casos, eles podem superar mais de dez vezes os níveis de suas isoformas lineares (MEMCZAK et al., 2013).

circRNAs amplamente expressos em vários tipos celulares mostraram um nível de expressão significativamente maior do que os expressos de forma restrita (GAO; WANG; ZHAO, 2015). Além disso, circRNAs geralmente exibem expressão tecido- e desenvolvimento-específica (SALZMAN et al., 2013; SZABO et al., 2015; VENØ et al., 2015). Por exemplo, em moscas, os circRNAs tendem a ter um maior nível de expressão nos tecidos neurais (WESTHOLM et al., 2014), enquanto em seres humanos e em ratos, descobriu-se que circRNAs estão enriquecidos no cérebro (RYBAK-WOLF et al., 2015; YOU et al., 2015) e nas plaquetas (ALHASAN et al., 2016). No entanto, os mecanismos e funções relacionadas por trás da mudança na abundância de circRNA, tanto em animais quanto em plantas, aguardam uma investigação mais aprofundada.

Em 2015, Ye e colaboradores identificaram 12.037 e 6.012 circRNAs em *Oryza sativa* e *Arabidopsis thaliana*, respectivamente, sendo que em arroz, alguns circRNAs foram diferencialmente expressos sob as condições contrastantes de concentração de fosfato inorgânico (Pi), sugerindo que circRNAs podem desempenhar um papel em resposta ao estresse por escassez de Pi (YE et al., 2015). Paralelamente, no mesmo ano Lu e colaboradores também relataram 2.354 circRNAs em *Oryza sativa* (LU et al., 2015a). Nos últimos anos, foram identificados 854 circRNAs em tomate (ZUO et al., 2016), 5372 circRNAs em soja (ZHAO et al., 2017) e 88 circRNAs em trigo (WANG et al., 2017)

Recentemente, Pan e colaboradores descreveram que o estresse térmico também tem grandes impactos na biogênese dos circRNAs, e que os circRNAs induzidos por este estresse biótico podem participar resposta da planta ao estresse através das redes de ceRNA mediadas por circRNA (PAN et al., 2018).

1.8 Regulação gênica mediada por RNA circular (circRNA)

Com o advento da tecnologia de NGS (do inglês, *Next Generation Sequencing*), associado ao desenvolvimento de algoritmos de predição de circRNAs, a identificação, a biogênese e a função de circRNAs têm sido amplamente relatadas em animais, como humanos, ratos e *Drosophila*. Estes achados trouxeram os circRNAs de volta à atenção com sua detecção e investigação em vários organismos e situações. Desde então, milhares de RNAs circulares endógenos têm sido identificados e caracterizados recentemente em células eucarióticas. Assim como miRNAs e RNAs longos e não-codificadores (lncRNAs), circRNAs estão enriquecendo o mundo dos RNA.

Inicialmente, circRNAs endógenos não eram relatados por se associarem com ribossomos para tradução, indicando que eles poderiam ter uma tendência a funcionar como uma nova classe de RNAs longos não-codificadores (lncRNAs). No entanto, um estudo recente demonstrou *in vivo* que um único éxon pode ser eficientemente transcrito para gerar um circRNA e que este circRNA pode ser traduzido a partir de uma IRES (do inglês, *Internal Ribosome Entry Site*) para gerar uma proteína GFP funcional, mediada pela modificação de base de RNA mais abundante, N⁶-metiladenosina (m⁶A) (YANG et al., 2017). Dessa forma, é possível que RNAs circulares podem realmente funcionar como mRNA para conduzir a tradução dentro das células.

1.9 circRNAs como esponja de miRNAs

Nos últimos anos, diversos grupos de pesquisa têm publicado resultados interessantes a respeito de circRNAs, com destaque à biogênese dos circRNAs e os possíveis mecanismos que os envolvem. Algumas destas descobertas têm mostrado que circRNAs são muito estáveis, abundantes e apresentam um padrão de expressão tecido-específico (DANAN et al., 2012; HANSEN et al., 2013; MEMCZAK et al., 2013).

Grande parte dos circRNAs são formados através de *backsplicing* (JECK et al., 2013; ZHANG et al., 2014a) e são enriquecidos com sítios de ligação de miRNA e de outras proteínas de ligação ao RNA (Figura 8). Devido a sua capacidade de atrair a ligação de miRNAs e outras proteínas, ultimamente tem se usado o termo “esponjas” para estes RNAs circulares. Um estudo de bioinformática mostrou que os sítios de miRNA em circRNAs são livres de polimorfismos, o que sugere um papel importante dos circRNAs na regulação das atividades de miRNA (THOMAS;

SÆTROM, 2014). Embora esses resultados indiquem que alguns circRNAs funcionam como esponjas endógenas de miRNA, outros estudos sugerem que, em mamíferos, a maioria dos circRNAs não atuam desta maneira, já que a grande maioria não tem mais sítios de ligação a miRNA do que mRNAs co-lineares (GUO et al., 2014; YOU et al., 2015).

Como mencionado anteriormente, a circularização de éxons geralmente implica a excisão do íntron, mas em alguns casos raros eles podem ser mantidos entre os éxons (JECK et al., 2013). No entanto, estes circRNAs com íntrons retidos não foram observados por funcionar como esponjas miRNA.

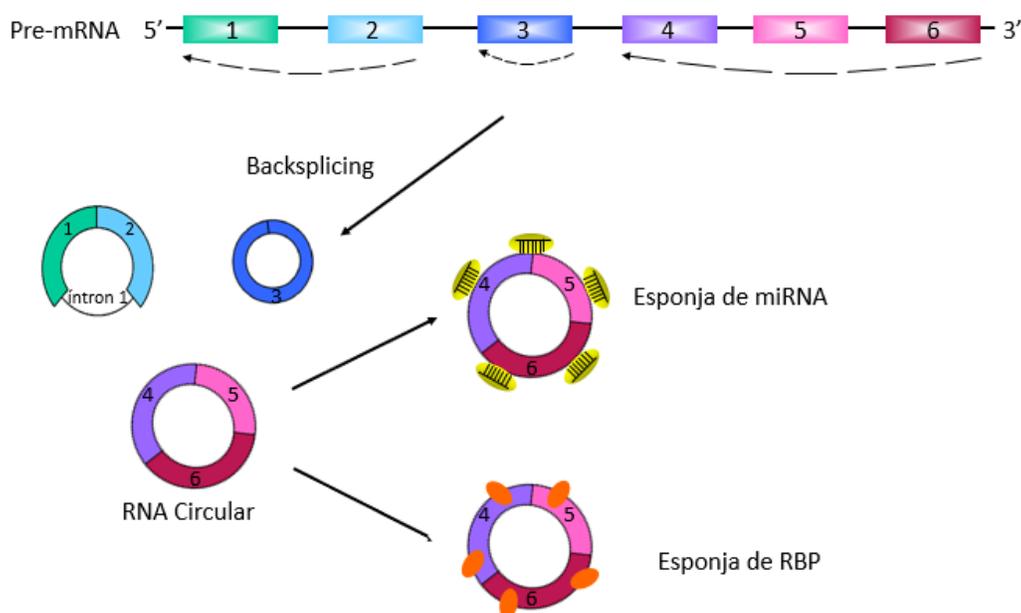


Figura 8. Principais funções dos circRNAs. Após a circularização via *backsplicing*, os RNAs circulares podem exercer diferentes funções. Entre elas, atuam como esponjas de miRNA e como esponjas de proteínas de ligação ao RNA (do inglês, *RBP* – *RNA Binding Proteins*).

Observações de que os circRNAs poderiam ser direcionados por interferência de RNA sugerem que circRNAs mais estáveis podem competir com mRNAs para a ligação de miRNA no citoplasma e, portanto, regular a expressão gênica. O circRNA mais bem caracterizado que dá suporte a este modelo é o ciRS-7, esponja para miR-7, que é produzido a partir da transcrição antisense e está relacionado à degeneração cerebelar de vertebrados 1 (CDR1) (Figura 9). Este

circRNA contém mais de 70 sítios de ligação de miR-7 convencionais, é expresso preferencialmente no cérebro humano e de rato, se localiza no citoplasma e pode aumentar o nível de expressão de genes-alvo de miR-7, reprimindo fortemente a atividade de miR-7. Dessa forma, a diminuição da expressão de ciRS-7 causa a redução da expressão de mRNAs contendo sítios de ligação a miR-7, sugerindo ainda que o ciRS-7 compete com mRNAs para ligação de miR-7 (CHEN, 2016). Além disso, é importante ressaltar que a função de ciRS-7 é conservada.

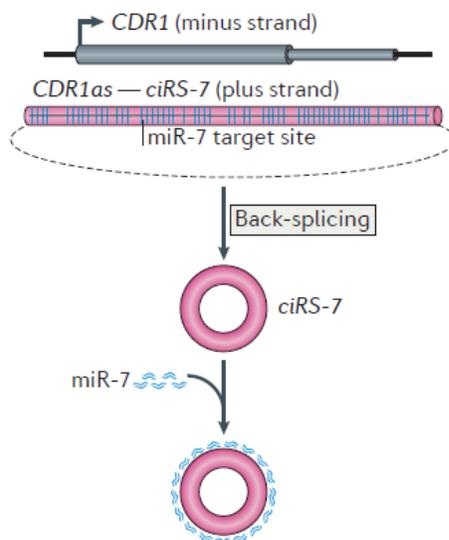


Figura 9. Atuação dos RNA circulares (circRNAs) na regulação gênica. O circRNA ciRS-7 (RNA circular esponja para mir-7) é produzido por *backsplicing* do transcrito antisenso do gene CDR1 (degeneração cerebelar de vertebrados 1), CDR1as. ciRS-7 contém mais de 70 sítios conservados para miR-7, localiza-se no citoplasma e funciona como esponja para o miR-7. Adaptado de (CHEN, 2016).

Outro circRNA já validado por ser altamente associado com a proteína efetora de miRNA, Argonata, é o circSry, que possui 16 sítios de ligação ao miR-138 (HANSEN et al., 2013). A ligação de miRNA não desestabiliza estes EcircRNA. Em vez disso, concorre com a ligação entre o miRNA e seus genes codificadores-alvo e, assim, reduz o efeito da repressão pós-transcrição mediada por miRNA. Assim, circRNAs que contêm sítios de ligação de miRNA são potenciais reguladores pós-transcricionais, que podem participar de diversos processos biológicos ao interagir com miRNAs.

1.10 Predição de RNAs circulares

As bibliotecas de RNA mais utilizadas para se fazer a predição de circRNAs são aquelas as quais passaram pelo processo de remoção de RNA ribossomal, seguido de RNAseq. Além disso, bibliotecas de imunoprecipitação de Argonauta (AGO-IP) seguidas de RNAseq também são muito utilizadas para a detecção de alvos de miRNA.

Nos experimentos de AGO-IP, primeiramente é realizada a imunoprecipitação da Argonauta com um anticorpo específico para uma proteína AGO em extrato celular. Dessa forma, pode-se isolar a AGO e as moléculas associadas a ela, como os miRNAs e seus alvos. Se considerarmos que circRNAs podem atuar como “esponjas” de miRNAs eles também poderiam ser detectados juntamente aos alvos de mRNA nestas bibliotecas.

Já a detecção de circRNAs a partir de dados de RNAseq baseia-se na identificação de *reads* que abrangem as junções de *backsplicing* que são geradas durante a biogênese dos circRNAs. Os *reads* referentes à região do *backsplicing* irão alinhar ao genoma em uma ordem cruzada. Ou seja, dois segmentos de um único *read* alinham separadamente em uma ordem inversa. Assim, a detecção de circRNAs a partir de *reads* de RNAseq requer métodos apropriados para alinhamento e análise de *reads* não-colineares.

Diversos métodos computacionais para a detecção de eventos de *backsplicing* a partir de dados de RNAseq têm sido desenvolvidos, como *find_circ* (MEMCZAK et al., 2013), *CircExplorer* (ZHANG et al., 2014b), *circRNA_finder* (WESTHOLM et al., 2014), *testrealign* (HOFFMANN et al., 2014), *CIRI* (GAO; WANG; ZHAO, 2015), *KNIFE* (SZABO et al., 2015), *UROBORUS* (SONG et al., 2016), *NCLscan* (CHUANG et al., 2016), *PTESFinder* (IZUOGU et al., 2016), e *Acfs* (YOU; CONRAD, 2016). Cada um destes *softwares* usa diferentes estratégias para identificar circRNAs, utilizando diferentes alinhadores de *reads*, com *inputs* variáveis, como genoma e anotação gênica, e fornecem *outputs* em termos de anotação de junção de *backsplicing* previstas, com taxas de validação variáveis entre si.

Os circRNAs preditos mais abundantes não são realmente positivos em todos os casos. Em geral, circRNAs identificados por um único método carecem de confiabilidade. Dessa forma, a combinação de pelo menos dois métodos pode ser uma boa prática para aumentar a robustez de detecção de circRNA, levando a uma diminuição no número de previsões falso-positivas.

1.10.1 O método CirComPara

Em 2017, Gaffo e colaboradores desenvolveram uma *pipeline* de bioinformática para detectar, quantificar e anotar circRNAs a partir de monócitos e macrófagos de uma paciente com haploinsuficiência na produção de QKI (proteína de ligação ao RNA) e seu irmão saudável (GAFFO et al., 2017) (Figura 10). O CirComPara é um *software* que utiliza diferentes algoritmos de predição de circRNAs, através da identificação de regiões de *backsplicing*. Sendo eles: FindCirc, TestRealign e CircExplorer2, que trabalha com três diferentes ancoradores (Segemhel, Star e Tophat).

O *input* necessário para o *pipeline* do CirComPara são *reads* de RNAseq proveniente de sequenciamento Illumina (tanto *reads single-end* como *pair-end* são aceitos), o genoma de referência no formato multi FASTA e a anotação gene/transcrito no formato GTF, proveniente de bancos de dados como Ensembl. Para cada amostra, CirComPara realiza primeiro um pré-processamento dos *reads* brutos para reter apenas fragmentos de alta qualidade que serão usados na análise posteriormente. Em seguida, um alinhamento preliminar do *read* ao genoma de referência (linear) é realizado com critérios rigorosos, especialmente importantes para os *reads pair-end*: *reads* não pareados e/ou pareados de maneira errada são considerados não mapeados e, portanto, separados dos *read* lineares alinhados. Os alinhamentos são usados para detectar e quantificar o transcrito linear e a expressão gênica, que são calculados para cada amostra como contagens brutas e valores normalizados (fragmentos por kilobase por milhão de *reads* mapeados; FPKM).

Os *reads* que não são mapeados ao genoma de referência são aqueles utilizados como candidatos para a detecção de junção de *backsplicing*.

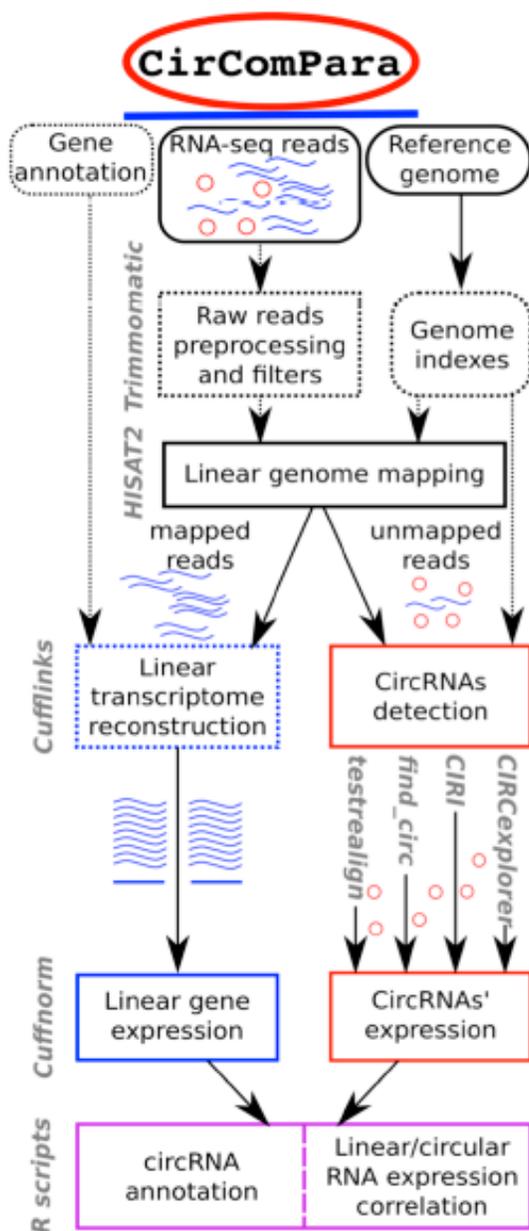


Figura 10. Fluxograma do *software* CirComPara. Caixas com cantos arredondados representam os dados de entrada requeridos pelo *software*. As ferramentas de bioinformática utilizadas estão representadas com cor cinza, ao lado do nível do *pipeline* relativo; linhas pontilhadas representam funções opcionais.

2. Justificativa

Estudos recentes identificaram abundância de circRNAs em mamíferos e demonstraram que os circRNAs endógenos são reguladores importantes em animais, que podem atuar como esponjas de miRNAs (HANSEN; KJEMS; DAMGAARD, 2013; JECK; SHARPLESS, 2014; SALZMAN et al., 2012). No entanto, em comparação com os circRNAs animais, o conhecimento dos circRNAs de plantas ainda é limitado (LU et al., 2015a; SABLOK; ZHAO; SUN, 2016; WANG et al., 2017; YE et al., 2015; ZUO et al., 2016).

Além disso, sabendo da importância da proteína Argonauta e da existência e importância dos RNAs circulares, seria de grande valor para o conhecimento científico descobrir se estas duas moléculas podem interagir uma com a outra e se os RNAs circulares podem atuar como esponjas de miRNAs também em plantas, como já é sabido em animais.

Dessa maneira, o conhecimento e aplicações a respeito deste tema, que está em crescente evolução e descobertas, serão ampliados.

3. Objetivos

3.1 Objetivo geral

Identificar RNAs circulares de *Arabidopsis* associados à proteína Argonata e com potencial de atuarem como esponjas de miRNA.

3.2 Objetivos específicos

- Identificar e confirmar a presença de RNAs circulares em bibliotecas de RNAseq de AGO-IP a partir flores de *Arabidopsis thaliana*;
- Identificar circRNAs com sítios para ligação de miRNAs;
- Procurar por alvos destes miRNAs nas bibliotecas de RNAseq de AGO-IP;
- Validar os transcritos de circRNAs;
- Projetar *primers* divergentes para detectar os RNAs circulares (circRNAs), e primers para detectar RNAs parentais, microRNAs (miRNAs) e alvos dos miRNAs;
- Analisar os níveis de expressão dos circRNAs, mRNAs e alvos dos miRNAs por PCR quantitativo (qPCR).

4. Material e métodos

4.1 Dados experimentais de imunoprecipitação de Argonautas (AGO-IP)

Para realizar as identificações computacionais de circRNAs, miRNAs e genes-alvo do presente trabalho, foram utilizadas sequências obtidas em um trabalho anterior em que foi demonstrado que a desativação de algum dos núcleos catalíticos do domínio PIWI faz com que esse domínio perca sua atividade de clivar o mRNA-alvo (CARBONELL et al., 2012). Duas linhagens de *Arabidopsis thaliana*, uma superexpressando o domínio PIWI selvagem (DDH) e a outra superexpressando o domínio PIWI com mutação no terceiro resíduo da tríade catalítica (DAH), foram utilizadas para a obtenção das bibliotecas de AGO-IP. Inflorescências destas duas linhagens transgênicas foram divididas em duas frações, sendo uma a fração que foi imunoprecipitada com a Argonauta e outra a fração controle ou *input*, ou seja, que não passou por imunoprecipitação. Para a imunoprecipitação, as células lisadas foram incubadas com um anticorpo anti-HA. Ambas as frações foram utilizadas para gerar tanto bibliotecas de sRNAs quanto biblioteca de RNA total, utilizando o sequenciador HiSeq2000 (Illumina).

4.2 Bibliotecas de RNAseq e de sRNAs

As bibliotecas de RNAseq e as bibliotecas de pequenos sRNAs provenientes do experimento de AGO-IP analisadas foram: AGO1-DDH (SRR546147), AGO1-DDH-replicata (SRR546148), AGO1-DAH (SRR546149) e AGO1-DAH-replicata (SRR546150), publicadas no GEO (do inglês, *Gene Expression Omnibus*), com o número de acesso GSE40259. A qualidade e a presença de adaptadores nestas bibliotecas foram visualizadas usando o *software* FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). O *trimming* de qualidade e a remoção de adaptadores das bibliotecas de RNAseq e sRNAs foram realizadas usando os programas Cutadapt/Sickle (<https://github.com/najoshi/sickle>) e Trimmomatic (BOLGER; LOHSE; USADEL, 2014), respectivamente.

4.3 Identificação de circRNAs nas bibliotecas de RNAseq a partir de AGO-IP

As bibliotecas de AGO-IP SRR546147, SRR546148, SRR546149 e SRR546150 foram escolhidas para identificar, quantificar e anotar potenciais circRNAs, utilizando o *software*

CirComPara (GAFFO et al., 2017), que requer como dados de entrada *reads* de RNAseq a partir de sequenciamento Illumina, um genoma de referência, e anotação gênica, geralmente proveniente do banco de dados Ensembl. Ambos os métodos de detecção de RNAs circulares fazem a identificação de *backsplicing* em cada biblioteca com um mínimo de dois *reads*. Como genoma de referência foi utilizado o genoma de *Arabidopsis thaliana* TAIR versão 10 e a anotação gênica foi obtida a partir do Ensembl Plants (<https://plants.ensembl.org/index.html>). A fim de filtrar e selecionar somente os fragmentos de alta qualidade, foi realizado um pré-processamento. Estes *reads* de alta qualidade foram então alinhados ao genoma de referência e a partir deste alinhamento foram divididos em dois grupos. Um grupo dos *reads* pareados perfeitamente ao genoma e outro grupo dos *reads* não pareados ou com discordância no pareamento, sendo que os pertencentes ao segundo grupo são os candidatos para a detecção de circRNAs.

4.4 Contagem de *reads* de miRNAs e análises de mRNAs-alvo em bibliotecas AGO-IP

Para identificar miRNAs que poderiam estar interagindo com circRNAs foi utilizada a ferramenta psRNAtarget (DAI; ZHAO, 2011). Todos os miRNAs maduros de *Arabidopsis*, listados na versão 21 do miRBase (KOZOMARA; GRIFFITHS-JONES, 2014), que é um banco de dados público, foram utilizados nesta análise. Aqueles miRNAs que apresentaram um *expectation value* (e-value) de no máximo cinco foram selecionados para análises subsequentes.

Para realizar a contagem de *reads* de miRNA, as bibliotecas AGO-IP de sRNAs AGO1-DDH e AGO1-DDH-replicata foram consideradas como uma única biblioteca, bem como AGO1-DAH e AGO1-DAH-replicata foram consideradas como outra biblioteca. Dessa forma, para a contagem geral de miRNAs foram utilizadas três bibliotecas (AGO-IP-DDH, AGO-IP-DAH e *input*). Para obter os valores de contagem dos *reads*, as sequências dos miRNAs foram alinhadas com as sequências de sRNAs de cada biblioteca, utilizando o algoritmo Bowtie 1 (LANGMEAD et al., 2009). Foram utilizados os parâmetros-padrão para o alinhamento, não sendo permitidos *mismatches*. A contagem dos *reads* foi normalizada de acordo com o tamanho das bibliotecas.

A fim de avaliar se os mRNAs-alvo daqueles miRNAs considerados enriquecidos também estavam enriquecidos nas bibliotecas AGO-IP, os mesmos dados utilizados para a identificação de circRNAs foram utilizados. Para obter a contagem de *reads*, os *reads* de cada biblioteca de RNAseq foram mapeados ao transcriptoma de *Arabidopsis thaliana* TAIR versão 10, obtido a

partir do banco de dados Phytozome (<https://phytozome.jgi.doe.gov/pz/portal.html>), utilizando o algoritmo Bowtie2 (LANGMEAD; SALZBERG, 2012). Foram considerados os parâmetros-padrão e um *Expectation value* máximo de 3.

4.5 Material vegetal

Para a análise da expressão dos RNAs circulares que potencialmente funcionam como esponjas de miRNAs, plantas de *Arabidopsis thaliana*, ecótipo Columbia-0 (Col-0), foram crescidas em solo turfado sob temperatura de 22 °C, com fotoperíodo de dia longo (16 h de luz/ 8 h de escuro) em salas de crescimento com 65% de umidade relativa do ar, durante seis semanas. Amostras de folhas, flores, eixo e siliquis foram coletadas e armazenadas em nitrogênio líquido para subsequente armazenamento à -80 °C.

4.6 Extração de RNA total e síntese de cDNA

A extração de RNA total foi feita utilizando o reagente TRIZOL, de acordo com o protocolo do fabricante. Para cada 100 mg de tecido macerado 1 mL de TRIZOL foi adicionado. A solução foi incubada por 5 minutos à temperatura ambiente e em seguida foram adicionados 100 µL de clorofórmio. Após misturar bem e incubar por mais 3 minutos à temperatura ambiente, a solução foi centrifugada por 15 minutos, a 12.000 rpm, a 4 °C. A fase aquosa foi transferida para outro tubo de micro centrífuga de 1,5 mL e então foi adicionado 500 µL de álcool isopropílico gelado. Depois de incubado por 10 minutos à temperatura ambiente, foi feita a centrifugação por 15 minutos, a 12.000 rpm, a 4 °C e o sobrenadante removido. Foram realizadas três lavagens com etanol 75% gelado e depois de evaporado o excesso, o *pellet* foi solubilizado em 50 µL de água destilada. Por fim, a solução foi incubada à temperatura de 60 °C por 10 minutos e armazenada em freezer. Após a extração, a análise da integridade do RNA pode ser realizada através de eletroforese em gel de agarose 1%, onde foi visualizada sob luz UV e imagem digital gerada por meio do sistema Gel-Doc (Bio-Rad).

Para a síntese dos DNAs complementares (cDNAs) foram utilizadas sequências iniciadoras de sentido reverso de cada circRNA e a Transcriptase Reversa M-MLV (Promega), seguindo o protocolo do fabricante.

4.7 Desenho de *primers*

Os *primers* foram projetados usando a ferramenta Primer3 (ROZEN; SKALETSKY, 2000). Para validar os circRNAs identificados por bioinformática foram projetados *primers* divergentes. Isso significa que o *primer Reverse* (PUR, de Primer Universal Reverso) anela-se à extremidade 5' do circRNA e o *primer Foward* (PCF, de Primer Circular Foward) se anela à extremidade 3'. Para a detecção do mRNA parental foi projetado um conjunto de *primers* convergentes, tanto para o cDNA quanto para o DNA genômico (gDNA). Neste caso, o *primer Foward* (PPF, de Primer Parental Foward) é projetado para ficar *upstream* ao *primer Reverse* (Figura 11). O Primer Universal Reverso (PUR) utilizado foi o mesmo tanto para a detecção do circRNA como a do seu mRNA parental. O tamanho dos produtos de PCR utilizando-se os pares de *primers* para a detecção dos circRNAs é de 134 pb para o circRNA At2g35940, 95 pb para o circRNA At1g12080, 128 pb para o circRNA At5g2770, 129 pb para o circRNA At2g42170 e 143 pb para o circRNA At3g13990. Já para a detecção do cDNA e do gDNA dos mRNAs parentais destes circRNAs, utilizando os pares de *primers* apropriados, espera-se obter, respectivamente, fragmentos de 149 pb e 540 pb para o circRNA At2g35940, 158 pb e 341 pb para o circRNA At1g12080, 166 pb e 517 pb para o circRNA At5g2770, 106 pb e 511 pb para o circRNA At2g42170 e 112 pb e 855 pb para o circRNA At3g13990.

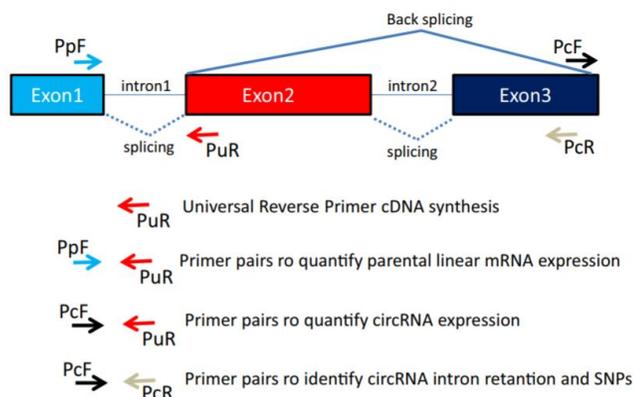


Figura 11. Primers para detectar circRNAs e seus mRNAs parentais. Representação genérica de um gene composto por três éxons e dois íntrons apresentada com as posições relativas de hibridação dos *primers* específicos: PuR - reverso universal, PpF - parental Forward, PcF - circular Forward e PcR - circular reverso. O splicing canônico é representado pelos pontilhados contínuos enquanto que o *backsplicing* é representado pela linha contínua. (Adaptado de (KULCHESKI; CHRISTOFF; MARGIS, 2016)).

4.8 Detecção de circRNAs e mRNAs parentais

A expressão de cada circRNA foi avaliada por PCR convencional, utilizando *primers* divergentes, como mencionado na sessão 4.7. A Reação em Cadeia da Polimerase (PCR) foi realizada utilizando a enzima Taq DNA polimerase Platinum (Invitrogen). Todas as reações de foram realizadas no aparelho Applied Biosystems Veriti. As reações foram conduzidas em um volume de 20 μL contendo 10 μL do cDNA diluído (1:100), 0,4 mM de dNTPs, 10x Tampão, 3 mM de MgCl_2 , 0,25 U de Platinum Taq DNA polimerase (Invitrogen) e 0,1 μM de cada oligonucleotídeo. As condições da PCR foram: um passo inicial de 5 min a 95 °C seguido de 40 ciclos 10 segundos a 95 °C de desnaturação da fita dupla, 15 segundos a 60 °C para anelamento das sequências iniciadoras e 15 segundos a 72 °C para extensão. A confirmação dos fragmentos foi realizada por eletroforese em gel de agarose 3%.

4.9 Purificação de produtos de PCR

A fim de confirmar por sequenciamento os circRNAs preditos por bioinformática e confirmados por PCR, foi realizada a purificação dos produtos de PCR utilizando o kit de purificação (Ludwig), de acordo com as recomendações do fabricante. Para cada um volume de produto de PCR foram utilizados cinco volumes do tampão PB, e em seguida as etapas de centrifugação na coluna de purificação são iniciadas. Após, adicionou-se o tampão WB e novamente foi feita centrifugação. Por fim, o produto de PCR purificado foi eluído em 30 μL do tampão EB.

4.10 Sequenciamento

Para as reações de sequenciamento pelo método Sanger foi utilizada a quantidade suficiente de produto de PCR purificado de forma que a concentração final da reação fosse em torno de 4,5 pmol/ μL .

4.11 Análises estatísticas

A contagem dos miRNAs foi normalizada de acordo com o tamanho das bibliotecas. Foram considerados enriquecidos nas bibliotecas AGO-IP aqueles miRNAs em que a contagem de *reads*

normalizados foi maior nas duas bibliotecas AGO-IP quando comparados com o controle (vetor vazio).

Para obter os mRNAs-alvo significativamente mais frequentes nas quatro bibliotecas AGO-IP comparados aos mRNAs controles presentes no *input* total, foi utilizado o pacote DESeq do *software* R (ANDERS; HUBER, 2010), considerando um *p-value* de 0,05.

5. Resultados e discussão

5.1 Identificação de circRNAs em bibliotecas AGO-IP

Para identificar RNAs circulares que potencialmente atuam como esponjas de miRNAs, foram utilizadas bibliotecas de Argonauta imunoprecipitada (AGO-IP), seguida de RNAseq (Figura 12). Estas bibliotecas foram obtidas a partir de flores de *Arabidopsis thaliana*, tanto a partir de uma linhagem superexpressando Argonauta de tipo selvagem (DDH) quanto a partir de uma linhagem mutante que não possui capacidade de clivar (DAH).

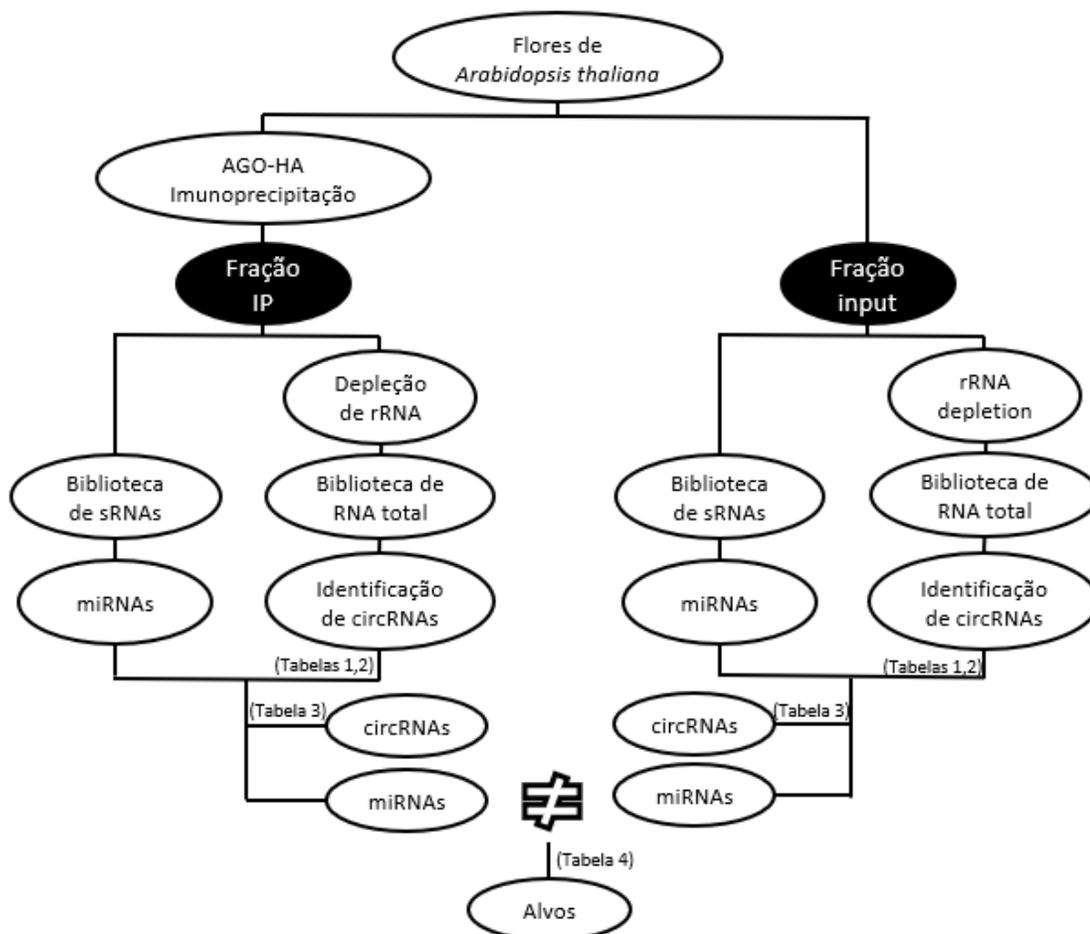


Figura 12. Fluxograma para identificação de circRNAs, miRNAs e mRNAs-alvo em bibliotecas AGO-IP e controle. A identificação dos circRNAs potencialmente esponjas para miRNAs, bem como a identificação dos mRNAs-alvo destes miRNAs foi realizada de acordo com a figura. O resultado obtido em cada etapa está identificado entre parênteses, com o número da tabela correspondente.

Utilizando o *software* CirComPara, foram preditos 29.358 circRNAs em bibliotecas AGO-IP RNAseq (Tabela 1). Utilizando o CircExplorer2 com o ancorador Segemehl foram identificados 86 possíveis circRNAs, enquanto que com o ancorador Star foram identificados 15 e com o ancorador Tophat 23. O número de circRNAs putativos identificados pelo algoritmo FindCirc foi de 1422 e pelo algoritmo TestRealign esse número foi 27812.

É interessante ressaltar que o grande número de circRNAs detectados pelo método TestRealign e sua diferença em relação a outros métodos (quase 20 vezes maior que o segundo método que mais detectou circRNAs), provavelmente se deve ao fato de que o TestRealign não realiza nenhuma etapa de pós-processamento de alinhamento específico para a identificação de circRNA, como os outros métodos realizam.

Tabela 1. CircRNAs identificados em bibliotecas AGO-IP por 5 métodos diferentes

		CircExplorer2			FindCirc	TestRealign
		Segemehl	Star	Tophat		
CircExplorer2	Segemehl	<u>86</u>	9	12	10	26
	Star		<u>15</u>	7	3	3
	Tophat			<u>23</u>	7	7
FindCirc					<u>1422</u>	198
TestRealign						<u>27812</u>

Para aumentar a confiabilidade da detecção destes circulares, foram selecionados para as demais análises somente os RNAs circulares detectados por pelo menos três dos cinco métodos. Dessa forma, o número de circRNAs identificados foi reduzido para 12 (Tabela 2). Em geral, estes 12 circRNAs apresentam uma ampla distribuição cromossomal, já que foram identificados nos cromossomos 1, 2, 3 e 5 de *Arabidopsis thaliana*, bem como uma grande diversidade nas funções dos genes parentais que dão origem a estes. Além disso, pode-se observar ainda uma variação em relação ao comprimento, desde 49 nt (At5g16880) até 1840 nt (At2g42170), bem como na quantidade de éxons que formam os circRNAs, de um até quatro. É importante destacar que, de acordo com o output do *software* CirComPara, o tamanho dos circRNAs descritos na Tabela 2 corresponde à região que pode conter, ou não, íntrons. Sendo assim, o tamanho do circRNA gerado pode não ser exatamente aquele descrito. Então, se considerarmos, por exemplo, o circ_At2g35940, separando somente a região codificadora, espera-se obter um circRNA de 134

nt. De acordo com estes dados preliminares, a ideia proposta em relação à razão tamanho do circRNA X número de éxons não pode ser considerada como totalmente verdadeira, como fora descrito anteriormente (SZABO et al., 2015; ZHANG et al., 2014b). Além disso, a grande maioria destes circRNAs é de origem exônica e ainda, alguns podem ser originados a partir de *backsplicing* imperfeito.

Tabela 2. Descrição dos 12 circRNAs putativos

Biblioteca	Gene_id	Circ_id	Função do gene parental	Origem	Exons	Comprimento (nt)	Métodos
DDH-IP	At1g02560	circ_At1g02560	Protease de CLP codificada no núcleo	exônico	2	224	5
DDH-IP	At1g12080	circ_At1g12080 **	Proteína vacuolar de ligação ao cálcio	exônico*	1	95	4
DDH-IP	At1g31810	circ_At1g31810	Homologia com formina 14	exônico	1	50	3
DDH-IP	At1g52360	circ_At1g52360	Subunidade beta do Coatomer	intrônico	1	224	3
DDH-IP	At2g02410	circ_At2g02410	Proteína não caracterizada (K06962)	exônico	1	71	4
DDH-IP	At2g35940	circ_At2g35940 **	Homeodomínio tipo BEL1 1	exônico	1	929	4
DDH-IP	At2g42170	circ_At2g42170 **	Proteína da família actina	exônico	4	1840	5
DDH-IP	At5g16880	circ_At5g16880	Alvo da proteína Myb 1	exônico*	1	49	4
DDH-IP	At5g56950	circ_At5g56950	Proteína de montagem do Nucleossoma NAP-1	intrônico*	1	118	4
DAH-IP	At3g01800	circ_At3g01800	Fator de reciclagem de ribossomo	exônico	1	68	4
DAH-IP	At3g13990	circ_At3g13990 **	Proteína relacionada à quinase de função desconhecida (DUF1296)	exônico	3	1061	3
DAH-IP	At5g27720	circ_At5g27720 **	Proteína da família de ribonucleoproteína nuclear pequena	exônico	4	984	5

* circRNA originados a partir de backsplicing imperfeito

** circRNAs com sítios de ligação para miRNAs

5.2 circRNAs com sítios de ligação para miRNAs

A interação entre o complexo miRNA-AGO e o mRNA-alvo forma um complexo ternário e leva à repressão de transcritos pela clivagem do mRNA ou pela repressão da tradução. Outra forma possível de complexo ternário pode ser através da interação entre complexo miRNA-AGO e circRNAs. Para demonstrar que circRNAs de plantas podem funcionar como esponjas de miRNA, impedindo a regulação do mRNA-alvo através do miRNA, selecionamos os cinco circRNAs que possuíam sítios de ligação para miRNAs (Tabela 3). São eles: circ_At1g12080, circ_At2g35940, circ_At2g42170, circ_At3g13990 e circ_At5g27720. A quantidade de diferentes miRNAs foi variável entre os cinco circRNAs, bem como a quantidade de *reads* detectada em cada biblioteca. Além disso, alguns atuam na repressão da transcrição enquanto que outros na inibição da tradução.

Tabela 3. microRNA potencialmente associados a circRNAs nas bibliotecas AGO-IP

circRNAs	miRNAs	Contagem de reads de miRNA			Inibição por
		AGO1-DDH	AGO1-DAH	Input	
circ_At1g12080	miR837-5p *	14	2	0	Tradução
	miR397a-5p *	1014	520	29	Tradução
	miR5654-3p	2	0	0	Clivagem
circ_At2g35940	miR8182-5p *	2	9	0	Tradução
	miR830-3p *	29	13	0	Clivagem
	miR833a-5p *	50	35	17	Tradução
circ_At2g42170	miR159c-3p	95375	110327	117402	Tradução
	miR4221-5p *	631	313	7	Tradução
	miR5024-3p	19	13	13	Tradução
	miR5645bd-5p	2	2	0	Clivagem
	miR831-3p *	81	11	3	Tradução
circ_At3g13990	miR4239-5p *	17	7	0	Tradução
	miR838-3p *	226	59	41	Clivagem
circ_At5g27720	miR846-5p *	693	183	0	Tradução
	miR861-3p	2	0	0	Tradução

* miRNA considerados enriquecidos em bibliotecas AGO-IP

A partir dos 15 miRNAs, foram considerados enriquecidos em bibliotecas AGO-IP aqueles que apresentaram uma contagem de *reads* de miRNA mais alta quando comparada com o *input* controle. Dessa forma, 10 miRNAs foram selecionados: miR837-5p, miR397a-5p, miR8182-5p, miR830-3p, miR833a-5p, miR4221-5p, miR831-3p, miR4239-5p, miR838-3p e miR846-5p.

5.3 mRNAs-alvo de miRNAs estão enriquecidos em bibliotecas AGO-IP

Os miRNAs enriquecidos em bibliotecas AGO-IP foram selecionados para avaliar se os seus mRNAs-alvo também se apresentavam enriquecidos nas bibliotecas AGO-IP. No total, 260 alvos foram identificados, com o valor de E variando entre 0,5 e 3. Dos 10 miRNAs considerados enriquecidos (Tabela 3), oito apresentaram mRNAs alvos significativamente mais frequentes nas bibliotecas AGO-IP do que na biblioteca controle. Dessa forma, o número de alvos foi reduzindo para 64 (Tabela 4).

Como ambas a linhagens transgênicas apresentaram alvos de miRNAs enriquecidos nas bibliotecas de imunoprecipitação, elas foram mantidas juntas nas análises de bioinformática.

Tabela 4. Descrição de mRNAs alvos de miRNAs presentes em circRNAs

Acesso do alvo	miRNA	Expectation	Inibição por	Comprimento (nt)	Contagem de reads de alvos (por biblioteca)						Contagem de reads de alvos (média)				
					AgoIP (norm)			Input (norm)			AgoIP (norm)	Input (norm)			
* At2g38080.1		1	Clivagem	2021	81	56	46	48	22	12	25	25	58	21	
** At5g60020.1		1	Clivagem	2049	38	37	32	27	22	11	17	21	33	18	
* At3g06040.1	miR397a-5p	3	Clivagem	864	29	38	24	27	13	16	16	12	29	14	
* At3g06040.2		3	Clivagem	1253	29	38	24	27	13	18	17	13	29	15	
* At3g06040.3		3	Clivagem	657	29	35	22	25	12	15	16	12	28	13	
* At3g06470.1		3	Clivagem	1092	47	50	108	98	4	2	8	4	75	4	
* At3g54170.1		2.5	Clivagem	1262	28	28	17	16	12	5	9	13	22	10	
* At4g13070.1	2.5	Clivagem	1775	12	5	6	9	3	1	3	2	8	2		
* At5g60040.1	2.5	Clivagem	4582	77	60	57	55	22	28	14	23	62	22		
* At5g60040.2	2.5	Clivagem	4428	57	39	34	27	20	25	11	22	39	20		
* At1g13350.1	3	Clivagem	2454	220	177	90	81	28	17	26	26	142	24		
* At1g13350.2	3	Clivagem	2535	221	177	90	82	28	17	27	26	142	24		
** At1g77660.1	miR4221-5p	3	Clivagem	1765	24	18	23	23	9	10	17	13	22	12	
* At2g33240.1		3	Clivagem	5313	58	45	28	14	12	6	11	17	36	12	
* At3g02170.1		3	Clivagem	3300	407	398	251	220	161	93	204	163	319	155	
* At4g14510.1		3	Clivagem	2940	82	69	34	45	25	22	18	24	57	22	
** At1g22230.1		3	Tradução	1262	2	3	2	0	0	0	1	0	2	0	0
* At1g31650.1	3	Tradução	2255	128	143	177	209	36	19	20	36	164	28	28	
* At2g38610.1	3	Tradução	1452	57	58	63	47	21	21	31	30	56	26	26	
* At2g38610.2	3	Tradução	1450	58	59	64	47	21	21	31	30	57	26	26	
* At2g35160.1	miRNA8182-5p	3	Clivagem	2798	26	23	18	12	14	6	8	9	20	9	
* At4g22580.1		3	Clivagem	1628	47	50	36	25	9	3	14	14	39	10	
* At1g23400.1	3	Clivagem	1822	68	76	95	87	23	22	19	32	81	24	24	
* At1g49880.1	2.5	Tradução	803	42	37	61	58	1	0	1	4	50	2	2	
* At3g46060.1	miR831-3p	3	Tradução	1132	76	72	76	77	36	31	54	43	75	41	41
* At3g46060.2		3	Tradução	1142	68	63	72	73	35	29	51	41	69	39	39
* At3g46060.3		3	Tradução	1061	68	62	69	71	35	29	51	41	68	39	39
** At2g36890.1	miR833a-5p	2.5	Clivagem	971	7	6	3	6	2	0	2	1	6	1	1
* At1g67230.1		2.5	Clivagem	3981	488	582	211	189	132	113	143	146	368	134	134
** At4g15410.1	2.5	Clivagem	1542	52	55	43	38	22	24	25	37	47	27	27	
* At1g01590.1	3	Clivagem	2115	20	19	13	17	6	3	9	8	17	6	6	
* At1g52500.1	3	Clivagem	1721	1	0	0	0	0	1	0	0	0	0	0	
* At1g52500.2	3	Clivagem	1491	40	30	36	22	10	11	11	15	32	12	12	
* At3g02950.1	miR837-5p	3	Clivagem	1135	15	19	25	17	10	8	9	9	19	9	9
* At3g55860.1		3	Clivagem	1366	182	169	348	320	23	5	28	16	255	18	18
* At4g00850.1		3	Clivagem	1062	24	20	23	28	4	2	1	6	24	3	3
** At5g48600.1		2	Tradução	3796	62	67	42	31	29	21	27	33	50	28	28
* At5g48600.2		2	Tradução	3965	68	69	45	35	31	22	27	34	55	29	29
* At1g12430.1	3	Tradução	3216	72	105	83	70	45	39	50	42	82	44	44	
* At1g12430.2	3	Tradução	3224	71	103	81	69	44	39	50	42	81	44	44	
* At3g12380.1	2.5	Clivagem	2323	36	41	32	23	18	14	12	22	33	16	16	
* At3g12380.2	2.5	Clivagem	2313	36	40	32	23	18	14	12	22	33	16	16	
* At1g21740.1	3	Clivagem	2862	72	97	44	38	16	32	18	24	63	22	22	
* At1g64180.1	3	Clivagem	2072	15	13	13	12	4	0	4	5	13	3	3	
* At1g70470.1	3	Clivagem	765	17	18	21	13	5	3	4	4	17	4	4	
* At4g01080.1	3	Clivagem	1583	137	127	70	56	35	20	39	40	98	33	33	
* At5g09460.1	3	Clivagem	2546	118	153	113	111	56	26	34	47	124	41	41	
* At5g09461.1	3	Clivagem	2546	118	153	113	111	56	26	34	47	124	41	41	
* At5g09462.1	3	Clivagem	2546	118	153	113	111	56	26	34	47	124	41	41	
* At5g09463.1	3	Clivagem	2546	118	153	113	111	56	26	34	47	124	41	41	
* At5g20110.1	3	Clivagem	778	32	28	34	20	3	1	3	2	28	2	2	
* At5g46030.1	miR838-3p	2	Tradução	732	26	18	24	35	14	16	10	8	26	12	12
* At2g44430.1		2.5	Tradução	2196	134	139	66	52	16	20	19	21	98	19	19
* At5g22640.1		2.5	Tradução	2814	268	242	233	244	126	102	104	129	247	115	115
* At5g22640.2		2.5	Tradução	2457	199	181	165	177	106	83	86	107	181	95	95
* At5g40340.1		2.5	Tradução	3096	895	826	448	327	74	49	94	83	624	75	75
* At5g56210.1		2.5	Tradução	2004	22	21	25	19	6	6	2	5	22	5	5
* At5g62390.1		2.5	Tradução	1859	409	441	297	250	143	142	147	175	349	152	152
* At5g17910.1		3	Tradução	4532	231	220	145	118	76	56	88	87	178	77	77
* At5g17910.2		3	Tradução	4757	233	221	147	120	76	56	89	87	180	77	77
** At5g41960.1		3	Tradução	874	7	8	10	11	3	2	3	6	9	4	4
* At5g57790.1	3	Tradução	1407	40	41	20	17	14	11	11	11	29	12	12	
* At5g57790.2	3	Tradução	916	37	38	19	14	14	8	8	9	27	10	10	
* At5g01950.1	miR846-5p	2	Clivagem	3387	63	53	54	44	39	21	28	29	53	29	29

* padj < 0.05

** pvalue < 0.05

5.4 Validação de circRNAs por RT-PCR e sequenciamento

Para validar os circRNAs preditos pela bioinformática, foram realizadas reações de PCR com amostras de folhas e flores, utilizando as sequências iniciadoras com sentido divergentes, em que a sequência iniciadora inversa estava na extremidade 5' do circular e a sequência iniciadora direta estava na sua região 3' (Tabela 5).

Tabela 5. Primers utilizados para RT-PCR

Nome	Sequência (5' → 3')
PCF_At2g42170	GCAGCCGCTAATATGAGG
PUR_At2g42170	GTTTTCGCCCTCTACTTCC
PPF_At2g42170	GAGCTGTCAGCAGCATTCC
PCF_At1g12080	AATTGTAGAAACAAACGAAGAAGT
PUR_At1g12080	GGTGCTTCGGATTTTTGTTC
PPF_At1g12080	GAAATGGCCACCGTTGAG
PCF_At2g35940	GGCACGGCAGAGAGAC
PUR_At2g35940	CGACGTCGTTTTCTTTGATTC
PPF_At2g35940	GCAATCACACACATCTCCAC
PCF_At5g27720	AAGAGGACGTGGTGTGG
PUR_At5g27720	TTCACCAAATGCCCATTTGTA
PPF_At5g27720	GTTACCTTCTTGCCGATC
PCF_At3g13990	ATTCTTCTGCCCAGTCCTC
PUR_At3g13990	TCTTTATGTTGGCAGGCTC
PPF_At3g13990	GCTAACGAGGCTGTGGA

PCR: Primer Circular Forward

PUR: Primer Universal Reverse

PPF: Primer Parental Forward

Dos cinco circRNAs selecionados, dois apresentaram padrão de bandas de acordo com o esperado. Para o circ_At2g35940, o tamanho esperado para o fragmento de cDNA do parental, utilizando *primers* convergentes, era 149 pb, enquanto que o esperado para o gDNA era de 540 pb. Ambos corresponderam ao proposto. Entretanto, como pode-se observar, outras duas bandas apareceram para o cDNA do mesmo circRNA, com aproximadamente 183 pb e 290 pb (Figura 13). Isto é um indicativo de que este circRNA pode apresentar outras isoformas, mas que curiosamente não foram detectadas na análise de bioinformática realizada pelo *software* CircComPara.

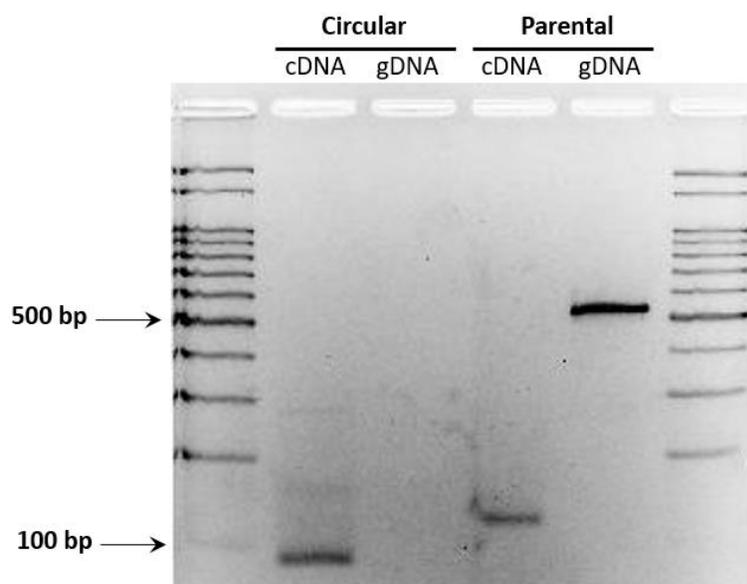


Figura 13. Confirmação por RT-PCR do circRNA_At2g35940. Após a reação de PCR, as amostras foram aplicadas em gel de agarose 3% para visualização do padrão de bandas. O marcador de peso molecular utilizado foi de 100 pb. cDNA: DNA complementar; gDNA: DNA genômico.

Após repetir as reações de PCR com os outros quatro circRNAs, purificar o produto e analisar em gel de agarose 3%, o circ_At1g12080, assim como o circ_At2g35940, apresentou um fragmento de acordo com o tamanho esperado, de 74 pb, mas também apresentou outros dois fragmentos.

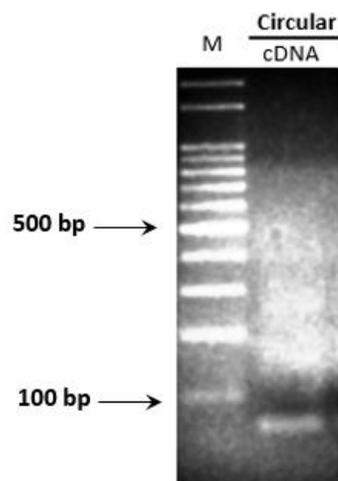


Figura 14. Confirmação por RT-PCR do circRNA_At1g12080. Após a purificação dos produtos de PCR, o padrão de bandas apresentado foi similar ao esperado. M: marcador de peso molecular de 100 pb, cDNA: At1g12080.

Após purificar o produto da PCR, que utilizou os *primers* divergentes para amplificar o fragmento referente ao circ_At1g12080, o resultado foi sequenciado e analisado (Figura 15). De acordo com a predição de bioinformática realizada pelo *software* CirComPara, o tamanho do circRNA gerado pelo parental At1g12080 é de 74 pb. Entretanto, de acordo com o sequenciamento, utilizando os mesmo *primers*, este circRNA é formado por 252 pb. Corroborando os resultados obtidos por RT-PCR e sequenciamento, podemos afirmar que o circ_At1g12080 possui duas isoformas. Entretanto, é curioso destacar que a isoforma proveniente do sequenciamento não foi identificada pelos métodos de bioinformática.



Figura 15. Confirmação por sequenciamento do circRNA_At1g12080. Alinhamento entre a sequência do gene parental (primeira linha), com a sequência do circRNA proveniente das análises de informática (segunda linha) e a sequência obtida através do sequenciamento.

Em relação aos outros três circRNAs circ_At2g42170, circ_At3g13990 e circ_At5g27720, estes não apresentaram produtos de amplificação por PCR. Este é um resultado inesperado, visto que o circ_At2g42170 e o circ_At5g27720 foram identificados pelos cinco métodos propostos neste trabalho.

6. Perspectivas

Este trabalho ainda tem muitos pontos que precisam ser aprofundados e esclarecidos, a fim de que seja possível um maior conhecimento a respeito dessa classe de ncRNA que, em plantas, até então não se tem muitas informações.

Os próximos passos serão avaliar o nível de expressão dos miRNAs que possuem sítios de ligação aos circRNAs que foram considerados enriquecidos nas bibliotecas AGO-IP e avaliar também o nível de expressão dos mRNAs alvo destes miRNAs.

Dessa forma, poderemos confirmar que, assim como em animais, RNAs circulares funcionam como esponjas de miRNAs em plantas, e são capazes de sequestrar o miRNA, impedindo a regulação dos mRNAs-alvo.

Além disso, outra abordagem para avaliar a meia-vida estrutural e funcional dos RNAs circulares será realizada através do tratamento com oligonucleotídeos com sequência antisense (ASOs) à junção *backsplicing* do circRNA.

7. Referências

- ALHASAN, A. A. et al. Circular RNA enrichment in platelets is a signature of transcriptome degradation. **Blood**, v. 127, n. 9, p. e1–e11, 3 mar. 2016.
- AMERES, S. L.; ZAMORE, P. D. Diversifying microRNA sequence and function. **Nature reviews. Molecular cell biology**, v. 14, n. 8, p. 475–88, ago. 2013.
- ANDERS, S.; HUBER, W. Differential expression analysis for sequence count data. **Genome Biology**, v. 11, n. 10, p. R106, 2010.
- ANDREEVA, K.; COOPER, N. G. F. Circular RNAs: New players in gene regulation. **Advances in Bioscience and Biotechnology**, v. 06, n. 06, p. 433–441, 2015.
- ASHWAL-FLUSS, R. et al. circRNA biogenesis competes with pre-mRNA splicing. **Molecular Cell**, v. 56, n. 1, p. 55–66, out. 2014.
- AXTELL, M. J.; WESTHOLM, J. O.; LAI, E. C. Vive la différence: biogenesis and evolution of microRNAs in plants and animals. **Genome biology**, v. 12, n. 4, p. 221, 2011.
- BOLGER, A. M.; LOHSE, M.; USADEL, B. Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics**, v. 30, n. 15, p. 2114–2120, 1 ago. 2014.
- CAPEL, B. et al. Circular transcripts of the testis-determining gene Sry in adult mouse testis. **Cell**, v. 73, n. 5, p. 1019–1030, jun. 1993.
- CARBONELL, A. et al. Functional analysis of three Arabidopsis ARGONAUTES using Slicer-defective mutants. **The Plant Cell**, v. 24, n. 9, p. 3613–3629, 1 set. 2012.
- CHAO, C. W. et al. The mouse formin (Fmn) gene: abundant circular RNA transcripts and gene-targeted deletion analysis. **Molecular medicine (Cambridge, Mass.)**, v. 4, n. 9, p. 614–28, set. 1998.
- CHEN, I.; CHEN, C.-Y.; CHUANG, T.-J. Biogenesis, identification, and function of exonic circular RNAs. **Wiley interdisciplinary reviews. RNA**, v. 6, n. 5, p. 563–79, 2015.
- CHEN, L.-L. The biogenesis and emerging roles of circular RNAs. **Nature reviews. Molecular cell biology**, v. 17, n. 4, p. 205–11, 2016.
- CHEN, L.-L.; YANG, L. Regulation of circRNA biogenesis. **RNA Biology**, v. 12, n. 4, p. 381–388, 3 abr. 2015.
- CHI, S. W. et al. Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. **Nature**, v. 460, n.

7254, p. 479–86, 23 jul. 2009.

CHUANG, T.-J. et al. NCLscan: accurate identification of non-co-linear transcripts (fusion, trans-splicing and circular RNA) with a good balance between sensitivity and precision. **Nucleic Acids Research**, v. 44, n. 3, p. e29–e29, 18 fev. 2016.

CHUCK, G.; CANDELA, H.; HAKE, S. Big impacts by small RNAs in plant development. **Current Opinion in Plant Biology**, v. 12, n. 1, p. 81–86, fev. 2009.

COCQUERELLE, C. et al. Splicing with inverted order of exons occurs proximal to large introns. **The EMBO journal**, v. 11, n. 3, p. 1095–8, mar. 1992.

COGONI, C. et al. Transgene silencing of the *ai-1* gene in vegetative cells of *Neurospora* is mediated by a cytoplasmic effector and does not depend on DNA-DNA interactions or DNA methylation. **The EMBO journal**, v. 15, n. 12, p. 3153–63, 17 jun. 1996.

CONN, S. J. et al. The RNA binding protein quaking regulates formation of circRNAs. **Cell**, v. 160, n. 6, p. 1125–34, 12 mar. 2015.

DAI, X.; ZHAO, P. X. psRNATarget: a plant small RNA target analysis server. **Nucleic Acids Research**, v. 39, n. suppl_2, p. W155–W159, 1 jul. 2011.

DANAN, M. et al. Transcriptome-wide discovery of circular RNAs in Archaea. **Nucleic acids research**, v. 40, n. 7, p. 3131–42, abr. 2012.

DAXINGER, L. et al. A stepwise pathway for biogenesis of 24-nt secondary siRNAs and spreading of DNA methylation. **The EMBO Journal**, v. 28, n. 1, p. 48–57, 7 jan. 2009.

DJAMI-TCHATCHOU, A. T. et al. Functional roles of microRNAs in agronomically important plants—Potential as targets for crop improvement and protection. **Frontiers in Plant Science**, v. 8, p. 378, 22 mar. 2017.

ENUKA, Y. et al. Circular RNAs are long-lived and display only minimal early alterations in response to a growth factor. **Nucleic Acids Research**, v. 44, n. 3, p. 1370–1383, 18 fev. 2016.

FIRE, A. et al. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. **Nature**, v. 391, n. 6669, p. 806–11, 19 fev. 1998.

FRANCO-ZORRILLA, J. M. et al. Target mimicry provides a new mechanism for regulation of microRNA

activity. **Nature genetics**, v. 39, n. 8, p. 1033–7, ago. 2007.

GAFFO, E. et al. CirComPara: A multi-method comparative bioinformatics pipeline to detect and study circRNAs from RNA-seq data. **Non-Coding RNA**, v. 3, n. 1, p. 8, 10 fev. 2017.

GAO, Y.; WANG, J.; ZHAO, F. CIRI: an efficient and unbiased algorithm for de novo circular RNA identification. **Genome Biology**, v. 16, n. 1, p. 4, 2015.

GUO, J. U. et al. Expanded identification and characterization of mammalian circular RNAs. **Genome Biology**, v. 15, n. 7, p. 409, 29 jul. 2014.

GURTAN, A. M.; SHARP, P. A. The role of miRNAs in regulating gene expression networks. **Journal of molecular biology**, v. 425, n. 19, p. 3582–600, 9 out. 2013.

HAN, M. et al. The Arabidopsis double-stranded RNA-binding protein HYL1 plays a role in microRNA-mediated gene regulation. **Proceedings of the National Academy of Sciences of the United States of America**, v. 101, n. 4, p. 1093–8, 27 jan. 2004.

HANSEN, T. B. et al. miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA. **The EMBO Journal**, v. 30, n. 21, p. 4414–4422, 2 nov. 2011.

HANSEN, T. B. et al. Natural RNA circles function as efficient microRNA sponges. **Nature**, v. 495, n. 7441, p. 384–388, 27 mar. 2013.

HANSEN, T. B.; KJEMS, J.; DAMGAARD, C. K. Circular RNA and miR-7 in cancer. **Cancer research**, v. 73, n. 18, p. 5609–12, 15 set. 2013.

HENTZE, M. W.; PREISS, T. Circular RNAs: splicing's enigma variations. **The EMBO Journal**, v. 32, n. 7, p. 923–925, 5 mar. 2013.

HIRAGURI, A. et al. Specific interactions between Dicer-like proteins and HYL1/DRB- family dsRNA-binding proteins in Arabidopsis thaliana. **Plant Molecular Biology**, v. 57, n. 2, p. 173–188, jan. 2005.

HÖCK, J.; MEISTER, G. The Argonaute protein family. **Genome biology**, v. 9, n. 2, p. 210, 2008.

HOFFMANN, S. et al. A multi-split mapping algorithm for circular RNA, splicing, trans-splicing and fusion detection. **Genome Biology**, v. 15, n. 2, p. R34, 10 fev. 2014.

HOUSELEY, J. M. et al. Noncanonical RNAs from transcripts of the Drosophila muscleblind gene. **Journal of Heredity**, v. 97, n. 3, p. 253–260, 1 maio 2006.

- IVANOV, A. et al. Analysis of intron sequences reveals hallmarks of circular RNA biogenesis in animals. **Cell reports**, v. 10, n. 2, p. 170–7, 13 jan. 2015.
- IZUOGU, O. G. et al. PTESFinder: a computational method to identify post-transcriptional exon shuffling (PTES) events. **BMC Bioinformatics**, v. 17, n. 1, p. 31, 13 dez. 2016.
- JECK, W. R. et al. Circular RNAs are abundant, conserved, and associated with ALU repeats. **RNA**, v. 19, n. 2, p. 141–157, 1 fev. 2013.
- JECK, W. R.; SHARPLESS, N. E. Detecting and characterizing circular RNAs. **Nature Biotechnology**, v. 32, n. 5, p. 453–461, 1 maio 2014.
- JONES-RHOADES, M. W.; BARTEL, D. P. Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. **Molecular Cell**, v. 14, n. 6, p. 787–799, 18 jun. 2004.
- JONES-RHOADES, M. W.; BARTEL, D. P.; BARTEL, B. MicroRNAs and their regulatory roles in plants. **Annual Review of Plant Biology**, v. 57, n. 1, p. 19–53, jun. 2006.
- JOPLING, C. L. Modulation of Hepatitis C Virus RNA abundance by a liver-specific microRNA. **Science**, v. 309, n. 5740, p. 1577–1581, 2 set. 2005.
- KIM, Y.-K.; KIM, V. N. Processing of intronic microRNAs. **The EMBO Journal**, v. 26, n. 3, p. 775–783, 7 fev. 2007.
- KOS, A. et al. The hepatitis delta (delta) virus possesses a circular RNA. **Nature**, v. 323, n. 6088, p. 558–60, 1986.
- KOZOMARA, A.; GRIFFITHS-JONES, S. miRBase: annotating high confidence microRNAs using deep sequencing data. **Nucleic Acids Research**, v. 42, n. D1, p. D68–D73, jan. 2014.
- KURIHARA, Y. The interaction between DCL1 and HYL1 is important for efficient and precise processing of pri-miRNA in plant microRNA biogenesis. **RNA**, v. 12, n. 2, p. 206–212, 22 dez. 2005.
- LAGOS-QUINTANA, M. Identification of novel genes coding for small expressed RNAs. **Science**, v. 294, n. 5543, p. 853–858, 26 out. 2001.
- LANGMEAD, B. et al. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. **Genome Biology**, v. 10, n. 3, p. R25, 2009.
- LANGMEAD, B.; SALZBERG, S. L. Fast gapped-read alignment with Bowtie 2. **Nature Methods**, v. 9, n. 4,

p. 357–359, 4 abr. 2012.

LAU, N. C. et al. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. **Science**, v. 294, n. October, p. 858–862, 2001.

LE HIR, H. et al. The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. **The EMBO journal**, v. 20, n. 17, p. 4987–97, 3 set. 2001.

LEE, R. C.; AMBROS, V. An Extensive class of small RNAs in *Caenorhabditis elegans*. **Science**, v. 294, n. OCTOBER, p. 862–864, 2001.

LEE, R. C.; FEINBAUM, R. L.; AMBROS, V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. **Cell**, v. 75, n. 5, p. 843–54, 3 dez. 1993.

LELANDAIS-BRIERE, C. et al. Small RNA diversity in plants and its impact in development. **Current Genomics**, v. 11, n. 1, p. 14–23, 1 mar. 2010.

LI, Z. et al. Exon-intron circular RNAs regulate transcription in the nucleus. **Nature Structural & Molecular Biology**, v. 22, n. 3, p. 256–264, 9 mar. 2015.

LIN, W. et al. Functional connectivity MR imaging reveals cortical functional connectivity in the developing brain. **American Journal of Neuroradiology**, v. 29, n. 10, p. 1883–1889, nov. 2008.

LIU, Q.; CHEN, Y.-Q. Insights into the mechanism of plant development: Interactions of miRNAs pathway with phytohormone response. **Biochemical and Biophysical Research Communications**, v. 384, n. 1, p. 1–5, 19 jun. 2009.

LU, T. et al. Transcriptome-wide investigation of circular RNAs in rice. **RNA (New York, N.Y.)**, v. 21, n. 12, p. 2076–87, dez. 2015a.

LU, Y. et al. Roles of microRNA in plant defense and virus offense interaction. **Plant Cell Reports**, v. 27, n. 10, p. 1571–1579, 15 out. 2008.

LU, Z. et al. Metazoan tRNA introns generate stable circular RNAs in vivo. **RNA**, v. 21, n. 9, p. 1554–1565, set. 2015b.

MALLORY, A. C.; VAUCHERET, H. Functions of microRNAs and related small RNAs in plants. **Nature Genetics**, v. 38, n. 6s, p. S31–S36, jun. 2006.

- MARGIS, R. et al. The evolution and diversification of Dicers in plants. **FEBS letters**, v. 580, n. 10, p. 2442–50, 1 maio 2006.
- MEISTER, G.; TUSCHL, T. Mechanisms of gene silencing by double-stranded RNA. **Nature**, v. 431, n. 7006, p. 343–349, 16 set. 2004.
- MEMCZAK, S. et al. Circular RNAs are a large class of animal RNAs with regulatory potency. **Nature**, v. 495, n. 7441, p. 333–338, 27 mar. 2013.
- 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, abr. 2008.
- MONTGOMERY, T. A. et al. Specificity of ARGONAUTE7-miR390 interaction and dual functionality in TAS3 trans-acting siRNA formation. **Cell**, v. 133, n. 1, p. 128–41, 4 abr. 2008.
- MOREL, J. et al. Fertile hypomorphic ARGONAUTE (ago1) mutants impaired in post-transcriptional gene silencing and virus resistance. **The Plant Cell**, v. 14, n. 3, p. 629–639, mar. 2002.
- MORLANDO, M. et al. Primary microRNA transcripts are processed co-transcriptionally. **Nature Structural & Molecular Biology**, v. 15, n. 9, p. 902–909, set. 2008.
- MORRIS, K. V.; CHAN, S. W.; JACOBSEN, S. E. Small interfering RNA –induced transcriptional gene silencing in human cells. **Science**, v. 305, n. August, p. 1289–1293, 2004.
- NAKAZAWA, Y. et al. The dsRNA-binding protein DRB4 interacts with the Dicer-like protein DCL4 in vivo and functions in the trans-acting siRNA pathway. **Plant Molecular Biology**, v. 63, n. 6, p. 777–785, 9 mar. 2007.
- NAPOLI, C. Introduction of a chimeric chalcone synthase gene into petunia results in reversible co-suppression of homologous genes in trans. **THE PLANT CELL ONLINE**, v. 2, n. 4, p. 279–289, 1 abr. 1990.
- NIGRO, J. M. et al. Scrambled exons. **Cell**, v. 64, n. 3, p. 607–613, fev. 1991.
- PADDISON, P. J. et al. Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. **Genes & development**, v. 16, n. 8, p. 948–58, 15 abr. 2002.
- PAN, T. et al. Heat stress alters genome-wide profiles of circular RNAs in Arabidopsis. **Plant Molecular Biology**, v. 96, n. 3, p. 217–229, 24 fev. 2018.
- PARKER, J. S. How to slice: Snapshots of Argonaute in action. **Silence**, v. 1, n. 1, p. 1–10, 2010.

- PATEL, A. J. et al. Quantifying density fluctuations in volumes of all shapes and sizes using indirect umbrella sampling. **Journal of Statistical Physics**, v. 145, n. 2, p. 265–275, 10 out. 2011.
- PEREIRA, T.C., e colaboradores. Introdução ao mundo dos miRNAs. SBG, 2015.
- PETERSEN, C.P., DOENCH, J.G., GRISHOK, A. AND SHARP, P. The biology of short RNAs. In: **The RNA World**. 3th. ed. [s.l.] Cold Spring Harbor Laboratory Press; 3 edition, 2005.
- RAND, T. A. et al. Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation. **Cell**, v. 123, n. 4, p. 621–9, 18 nov. 2005.
- RHOADES, M. W. et al. Prediction of plant microRNA targets. **Cell**, v. 110, n. 4, p. 513–20, 23 ago. 2002.
- RICARTE FILHO, J. C. M.; KIMURA, E. T. MicroRNAs: nova classe de reguladores gênicos envolvidos na função endócrina e câncer. **Arquivos Brasileiros de Endocrinologia & Metabologia**, v. 50, n. 6, p. 1102–1107, dez. 2006.
- RIVAS, F. V. et al. Purified Argonaute2 and an siRNA form recombinant human RISC. **Nature Structural & Molecular Biology**, v. 12, n. 4, p. 340–349, 30 abr. 2005.
- ROMEO, T. Global regulation by the small RNA-binding protein CsrA and the non-coding RNA molecule CsrB. **Molecular Microbiology**, v. 29, n. 6, p. 1321–1330, set. 1998.
- ROZEN, S.; SKALETSKY, H. Primer3 on the WWW for general users and for biologist programmers. **Methods in molecular biology (Clifton, N.J.)**, v. 132, p. 365–86, 2000.
- RUVKUN, G. et al. Conservation of the sequence and temporal expression of let - 7 heterochronic regulatory RNA. **Nature**, v. 408, n. 6808, p. 86–89, 2 nov. 2000.
- RYBAK-WOLF, A. et al. Circular RNAs in the mammalian brain are highly abundant, conserved, and dynamically expressed. **Molecular Cell**, v. 58, n. 5, p. 870–885, jun. 2015.
- SABLOK, G.; ZHAO, H.; SUN, X. Plant Circular RNAs (circRNAs): Transcriptional regulation beyond miRNAs in plants. **Molecular plant**, v. 9, n. 2, p. 192–194, 1 fev. 2016.
- SALMENA, L. et al. A ceRNA Hypothesis: The Rosetta Stone of a hidden RNA language? **Cell**, v. 146, n. 3, p. 353–358, ago. 2011.
- SALZMAN, J. et al. Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. **PLoS ONE**, v. 7, n. 2, p. e30733, 1 fev. 2012.

SALZMAN, J. et al. Cell- type specific features of circular RNA expression. **PLoS Genetics**, v. 9, n. 9, p. e1003777, 5 set. 2013.

SANGER, H. L. et al. Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures. **Proceedings of the National Academy of Sciences of the United States of America**, v. 73, n. 11, p. 3852–6, nov. 1976.

SHUKLA, L.; CHINNUSAMY, V.; SUNKAR, R. The role of microRNAs and other endogenous small RNAs in plant stress responses. **Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms**, v. 1779, n. 11, p. 743–748, nov. 2008.

SKALSKY, R. L.; CULLEN, B. R. Viruses, microRNAs, and host interactions. **Annual review of microbiology**, v. 64, n. Pol II, p. 123–41, 2010.

SONG, J.-J. Crystal structure of Argonaute and its implications for RISC Slicer activity. **Science**, v. 305, n. 5689, p. 1434–1437, 3 set. 2004.

SONG, X. et al. Circular RNA profile in gliomas revealed by identification tool UROBORUS. **Nucleic Acids Research**, v. 44, n. 9, p. e87–e87, 19 maio 2016.

STARKE, S. et al. Exon circularization requires canonical splice signals. **Cell reports**, v. 10, n. 1, p. 103–11, 6 jan. 2015.

SUZUKI, H. et al. Characterization of RNase R-digested cellular RNA source that consists of lariat and circular RNAs from pre-mRNA splicing. **Nucleic acids research**, v. 34, n. 8, p. e63, 8 maio 2006.

SZABO, L. et al. Statistically based splicing detection reveals neural enrichment and tissue-specific induction of circular RNA during human fetal development. **Genome biology**, v. 16, n. 1, p. 126, 16 jun. 2015.

TAKEDA, A. et al. The mechanism selecting the guide strand from small RNA duplexes is different among argonaute proteins. **Plant & cell physiology**, v. 49, n. 4, p. 493–500, abr. 2008.

TALHOUARNE, G. J. S.; GALL, J. G. Lariat intronic RNAs in the cytoplasm of *Xenopus tropicalis* oocytes. **RNA**, v. 20, n. 9, p. 1476–1487, set. 2014.

TAY, Y.; RINN, J.; PANDOLFI, P. P. The multilayered complexity of ceRNA crosstalk and competition. **Nature**, v. 505, n. 7483, p. 344–352, 16 jan. 2014.

THOMAS, L. F.; SÆTROM, P. Circular RNAs are depleted of polymorphisms at microRNA binding sites. **Bioinformatics**, v. 30, n. 16, p. 2243–2246, 15 ago. 2014.

TIJSTERMAN, M.; KETTING, R. F.; PLASTERK, R. H. A. The genetics of RNA silencing. **Annual review of genetics**, v. 36, n. 1, p. 489–519, 2002.

ULLU, E.; TSCHUDI, C.; CHAKRABORTY, T. RNA interference in protozoan parasites. **Cellular Microbiology**, v. 6, n. 6, p. 509–519, 15 abr. 2004.

VAUCHERET, H. Plant ARGONAUTES. **Trends in plant science**, v. 13, n. 7, p. 350–8, jul. 2008.

VENØ, M. T. et al. Spatio-temporal regulation of circular RNA expression during porcine embryonic brain development. **Genome biology**, v. 16, n. 1, p. 245, 5 nov. 2015.

VOLPE, T. A. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. **Science**, v. 297, n. 5588, p. 1833–1837, 13 set. 2002.

WANG, Y. et al. Identification of circular RNAs and their targets in leaves of *Triticum aestivum* L. under dehydration stress. **Frontiers in Plant Science**, v. 7, n. January, p. 1–10, 5 jan. 2017.

WESTHOLM, J. O. et al. Genome-wide analysis of *Drosophila* circular RNAs reveals their structural and sequence properties and age-dependent neural accumulation. **Cell Reports**, v. 9, n. 5, p. 1966–1980, dez. 2014.

WILUSZ, J. E.; SHARP, P. A. Molecular biology. A circuitous route to noncoding RNA. **Science (New York, N.Y.)**, v. 340, n. 6131, p. 440–1, 26 abr. 2013.

WINTER, J. et al. Many roads to maturity: microRNA biogenesis pathways and their regulation. **Nature Cell Biology**, v. 11, n. 3, p. 228–234, 1 mar. 2009.

XIE, Z.; KASSCHAU, K. D.; CARRINGTON, J. C. Negative feedback regulation of Dicer-Like1 in *Arabidopsis* by microRNA-guided mRNA degradation. **Current biology : CB**, v. 13, n. 9, p. 784–9, 29 abr. 2003.

YANG, Y. et al. Extensive translation of circular RNAs driven by N6-methyladenosine. **Cell Research**, v. 27, n. 5, p. 626–641, 10 maio 2017.

YE, C.-Y. et al. Widespread noncoding circular RNAs in plants. **The New phytologist**, v. 208, n. 1, p. 88–95, out. 2015.

YOU, X. et al. Neural circular RNAs are derived from synaptic genes and regulated by development and

plasticity. **Nature Neuroscience**, v. 18, n. 4, p. 603–610, 1 abr. 2015.

YOU, X.; CONRAD, T. O. Acfs: accurate circRNA identification and quantification from RNA-Seq data. **Scientific Reports**, v. 6, n. 1, p. 38820, 8 dez. 2016.

ZAPHIROPOULOS, P. G. Circular RNAs from transcripts of the rat cytochrome P450 2C24 gene: correlation with exon skipping. **Proceedings of the National Academy of Sciences**, v. 93, n. 13, p. 6536–6541, 25 jun. 1996.

ZHANG, H. et al. MiR-7, inhibited indirectly by lincRNA HOTAIR, directly inhibits SETDB1 and reverses the EMT of breast cancer stem cells by downregulating the STAT3 pathway. **Stem cells (Dayton, Ohio)**, v. 32, n. 11, p. 2858–68, nov. 2014a.

ZHANG, X.-O. et al. Complementary sequence-mediated exon circularization. **Cell**, v. 159, n. 1, p. 134–147, set. 2014b.

ZHANG, Y. et al. Circular intronic long noncoding RNAs. **Molecular Cell**, v. 51, n. 6, p. 792–806, set. 2013.

ZHAO, W. et al. Genome-wide identification and characterization of circular RNAs by high throughput sequencing in soybean. **Scientific Reports**, v. 7, n. 1, p. 5636, 17 dez. 2017.

ZUO, J. et al. Deciphering the roles of circRNAs on chilling injury in tomato. **Biochemical and Biophysical Research Communications**, v. 479, n. 2, p. 132–138, out. 2016.