CARACTERIZAÇÃO DE GENES ENVOLVIDOS NA REGULAÇÃO EPIGENÉTICA E ANÁLISE DE SEUS EFEITOS EM MONOCOTILEDÔNEAS

WÁLACI DA SILVA SANTOS

Orientadora: Clícia Grativol Gaspar de Matos

UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE
CAMPOS DOS GOYTACAZES

CAMPOS DOS GOYTACAZES – RJ
FEVEREIRO/2020

FICHA CATALOGRÁFICA
UENF - Bibliotecas
Elaborada com os dados fornecidos pelo autor.

S237 Santos, Walaci da Silva.

CARACTERIZAÇÃO DE GENES ENVOLVIDOS NA REGULAÇÃO EPIGENÉTICA E ANÁLISE DE SEUS EFEITOS EM MONOCOTILEDÔNEAS / Walaci da Silva Santos. - Campos dos Goytacazes, RJ, 2020.

98 f. : il.
Bibliografia: 90 - 95.

Dissertação (Mestrado em Biotecnologia Vegetal) - Universidade Estadual do Norte Fluminense Darcy Ribeiro, Centro de Biociências e Biotecnologia, 2020.
Orientadora: Clicia Grativol Gaspar.

1. MONOCOTILEDÔNEAS. 2. EPIGÉNETICA. 3. EXPRESSÃO GÊNICA. I. Universidade Estadual do Norte Fluminense Darcy Ribeiro. II. Título.

CDD - 660.6

CARACTERIZAÇÃO DE GENES ENVOLVIDOS NA REGULAÇÃO EPIGENETICA E ANÁLISE DE SEUS EFEITOS EM MONOCOTILEDÔNEAS

WÁLACI DA SILVA SANTOS

Dissertação de Mestrado apresentada ao Centro de Biociências e Biotecnologia da Universidade Estadual do Norte Fluminense Darcy Ribeiro, como parte das exigências para obtenção do título de Mestre em Biotecnologia Vegetal

Aprovada em 14 de fevereiro de 2020

Banca examinadora:

Dr. Gonçalo Apolinário de Souza Filho (UENF)

Dr. Thiago Motta Venâncio (UENF)

Dr. Geraldo Felipe Ferreira e Silva (CTC)

Drª. Clicia Grativol Gaspar de Matos (UENF)
(orientadora)

Este trabalho foi realizado no Laboratório de Química e Função de Proteínas e Peptídeos (LQFPP) da Universidade Estadual do Norte Fluminense Darcy Ribeiro (UENF), sob orientação da Drª Clícia Grativol Gaspar de Matos e financiado pela FAPERJ e UENF.

Agradecimentos

Agradeço a Deus pela proteção, saúde e força para conduzir esse trabalho, me tornando uma pessoa melhor.

A meu pai Jamir, minha mãe Terezinha e meu irmão William, que mesmo distantes, me deram todo suporte para a realização desse trabalho.

A minha orientadora Clícia, por me tornar uma pessoa melhor através da ciência, por saber conduzir a orientação desse trabalho, puxando as orelhas e sendo compreensível no momento certo, por despertar em mim a curiosidade científica e por ser mais do que uma orientadora, uma amiga.

A minha namorada Paula, por todo carinho, compreensão e por estar sempre ao meu lado, me dando todo suporte.

Aos meus amigos e amigas de laboratório, Fernanda, Sara, Leandro, Paula, Juliana, Jacymara, Maria Luiza, Cida, Jaqueline e Sara Nália, pelas conversas, pela distração necessária e pela ajuda na realização desse trabalho.

Aos meus amigos de república, Douglas Terra, Thiago, Douglas, Matheus e Luis, pela amizade e apoio de todos os momentos.

Aos laboratórios LQFPP, LFBM. LBT e LBCT pela disponibilização dos equipamentos utilizados durante esse trabalho.

A UENF, por me possibilitar realizar um mestrado gratuito e de qualidade.

Ao Programa de Pós Graduação em Biotecnologia Vegetal, por todo suporte que foi dado.

Aos professores do PGBV, por todo conhecimento compartilhado, fundamental para me tornar um pesquisador.

Dedicatória

Dedico esse trabalho a meu pai Jamir da Silva Santos e minha mãe Terezinha de Jesus da Silva Santos, que não tiveram a oportunidade de realizar nem o Ensino Fundamental, mas que não mediram esforços para que eu pudesse chegar até aqui.

Resumo

Mecanismos epigenéticos que atuam na regulação da expressão gênica têm sido cada vez mais estudados em diversas plantas. Em *Arabidopsis thaliana* os efetores da metilação do DNA e da via de pequenos RNAs estão bem caracterizados. Em monocotiledôneas esses mecanismos ainda não foram bem elucidados, como em plantas de cana-de-açúcar e milho que possuem grande importância econômica para o PIB brasileiro. O estudo dos mecanismos de regulação da expressão gênica dessas plantas tornou-se uma importante ferramenta para obtenção de cultivares resistentes a estresses bióticos e abióticos. Assim, o objetivo deste trabalho foi caracterizar genes de metilação e desmetilação do DNA e genes envolvidos na via de pequenos RNAs em 12 monocotiledôneas e analisar o efeito regulatório da desmetilação na expressão desses genes em plântulas de milho e cana-de-açúcar. Primeiramente, foram utilizados genes de metilação (MET1, CMT3 e DRM2), desmetilação do DNA (ROS1, DME, DML2 e DML3) e os genes da via de pequenos RNAs (DCL1- 4 e AGO1-10) de Arabidopsis para a anotação em *Amborela trichopoda* e outras 12 plantas monocotiledôneas, usando sequências disponíveis no Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html). As sequências de Arabidopsis foram obtidas no TAIR (https://www.arabidopsis.org) e as sequências de cana-de-açúcar foram obtidas do genoma publicado de *Saccharum spontaneum*. No total, foram anotados 251 genes. Além da anotação, utilizamos sequências proteicas para reconstruir a filogenia, que revelou 3 grupos de metilases e 4 de Dicer-like. Além disso, observamos a presença de 7 grupos de Argonautas em monocotiledôneas, diferentemente de Arabidopsis que possui 10 grupos e a presença de 2 grupos de desmetilases em monocotiledôneas, diferindo dos 4 grupos existentes em Arabidopsis. Através de análise *in silico*, utilizando o site (https://mpss.danforthcenter.org/) identificamos também, possíveis regiões reguladoras de desmetilases – MEMS - em milho e arroz. A utilização do inibidor de metilação em plântulas de milho e cana-de-açúcar revelou a importância da metilação no processo de desenvolvimento vegetal. Através de qPCR, comprovamos o efeito da ausência de metilação nos níveis de expressão de ROS1 de milho. Esse trabalho gerou informações importantes para o conhecimento evolutivo e funcional dos genes de metilação e desmetilação do DNA e das vias de pequenos RNAs, em plantas monocotiledôneas, além disso, fornecemos informações que podem ser utilizadas visando aplicações biotecnológicas, que visem o desenvolvimento de novos cultivares resistentes a pragas e estresses ambientais.

Palavras chaves: Monocotiledôneas, Epigenética, Expressão gênica.

Abstract

Epigenetic mechanisms that act in the regulation of gene expression have been increasingly studied in several plants. In *Arabidopsis thaliana* the effectors of DNA methylation and the pathway of small RNAs are well characterized. In monocotyledonous plants these mechanisms have not yet been elucidated. Some monocotyledons, such as sugarcane and corn, have great economic importance for the Brazilian GDP. The study of gene expression regulation of these plants has become an important tool for obtaining cultivars resistant to biotic and abiotic stresses. Thus, the objective of this work was to characterize DNA methylation and demethylation genes and those involved in the pathway of small RNAs in 12 monocots and, to analyze the regulatory effect of demethylation on the expression of these genes in corn and sugarcane seedlings. First, methylation genes (MET1, CMT3 and DRM2), DNA demethylation (ROS1, DME, DML2 and DML3) and the genes of the small RNA pathway (DCL1- 4 and AGO1-10) of Arabidopsis were used for the annotation in *Amborela trichopoda* and 12 other monocotyledonous plants, using sequences available at Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html). The sequences of Arabidopsis were obtained from TAIR (https://www.arabidopsis.org) and the sugarcane sequences were obtained from the published genome of *Saccharum spontaneum*. In total, 251 genes were annotated. In addition to the annotation, we used protein sequences to reconstruct the phylogeny, which revealed 3 groups of methylases and 4 groups of Dicer-like. In addition, we observed the presence of 7 groups of Argonauts in monocots, unlike Arabidopsis which has 10 groups, and the presence of 2 groups of demethylases in monocots, differing from the 4 groups existing in Arabidopsis. Through *in silico* analysis, using the website (https://mpss.danforthcenter.org/) we also identified possible regulatory regions for demethylases - MEMS - in corn and rice. The use of methylation inhibitor in corn and sugarcane seedlings revealed the importance of methylation in the plant development process. Using qPCR, we showed the effect of the absence of methylation on the expression levels of ROS1 of maize. This work generated important information for the evolutionary and functional knowledge of DNA methylation and demethylation genes and of small RNA pathways in monocotyledonous plants. In addition, we provide information that can be used in resulting works aimed at biotechnological applications, aimed at development of new cultivars resistant to pests and environmental stresses.

Key words: Monocotyledons, Epigenetics, Gene expression.

Lista de figuras

Figura 1: Classificação filogenética das *Poaceae*. 4

Figura 2: Distribuição da produção de cana-de-açúcar no mundo. 6

Figura 3: Comparação entre a área revestida com plantação de milho e a produção nos três maiores produtores mundiais do grão.. 8

Figura 4: A metilação atua como um sensor epigenético na regulação de ROS. 10

Figura 5: Pipeline da metodologia utilizada para anotaçao gênica.. 16

Figura 6: Pipeline das análises computacionais para identificação de regiões MEMS e a interação de siRNAs com genes de desmetilação em plantas monocotiledôneas. .. 17

Figura 7: Fluxograma das metodologias utilizadas para realização de análises fenotípicas e moleculares de plântulas de milho em condições controle e tratado..... 20

Figura 8: Apresentação do número de genes anotados em cada grupo e o total de genes anotados. 23

Figura 9: Filogenia de proteínas desmetilases de plantas monocotiledôneas.. 28

Figura 10: Filogenia de proteínas metilases em monocotiledôneas. 35

Figura 11: Filogenia de proteínas Dice-like das familias de 1 a 4 em plantas monocotiledôneas. 40

Figura 12: Filogenia de proteínas Argonautas de familias de 1 a 10 em diferentes plantas monocotiledôneas. 55

Figura 13: Representação de estruturas de genes 57

Figura 14: Identificação de possíveis regiões MEMS no genoma de milho..... 61

Figura 15: Identificação de possíveis regiões MEMs no genoma de arroz..... 65

Figura 16: Avaliação fenotípica do desenvolvimento de plântulas de milho com inibidor de metilação (5-Azac) e controles DMSO e H_2O). 66

Figura 17: Análise do desenvolvimento da parte aérea de plântulas de milho sob condições de controle e tratadas com inibidor de metilação 5-Azac. 67

Figura 18: Análise do desenvolvimento de raiz de plântulas de milho sob condições controle e tratamento com inibidor de metilação 5-AzaC..... 68

Figura 19: Expressão do gene ROS1 através de RT-qPCR em parte aérea de plântulas de milho. 69

Figura 20: Avaliação fenotípica do desenvolvimento de plântulas de cana-de-açúcar com inibidor de metilação (5-Azac) e H_2O. 82

Figura 21: : Representação estatísticas do desenvolvimento da parte aérea de plântulas de cana-de-açúcar sob condições de controle e tratamento com inibidor de metilação 5-AzaC.. 83

Figura 22: Representação estatísticas do desenvolvimento de raízes de plântulas de cana-de-açúcar sob condições de controle e tratamento com inibidor de metilação 5-AzaC. 84

Lista de tabelas

Tabela 1: Classificação das subfamílias de *Poacea*. 3

Tabela 2: Localização cromossômica dos genes de desmetilação de DNA em milho e arroz. 18

Tabela 3: Anotação des genes de desmetilação em diferentes plantas monocotiledôneas. 25

Tabela 4: Anotação de genes de metilação do DNA em diferentes plantas monocotiledôneas. 30

Tabela 5: Anotação de genes Dicer-like de famílias de 1 a 4 em diferentes plantas monocotiledôneas 37

Tabela 6: Anotação de genes Argonautas de familias de 1 a 10 em diferentes plantas monocotiledôneas. 42

Tabela 7: Tabela de comparação de genes anotados pelo nosso grupo e por Pei et al 2019. 71

Lista de abreviaturas

AGO – Argonauta

5-Azac – azacitidina

BOP -- Bambusoideae, Oryzoideae e Poideae

CMT3 – Chromometylase 3

CONAB -- Companhia Nacional de Abastecimento

DCL – Dicer-Like

DME – Demetilase

DML – Demeter-like

DRM2 – Domains Rearranged Methyltransferase 2

dsRNA – Double Strand RNA

FAO -- Food and Agricultura Organization of the United Nations

FAOSTAT -- Food and Agriculture Organization Corporate Statistical Database

JTT -- Jones, Taylor e Thornton

MEMS -- Methylation Monitoring Sequence

MET1 -- DNA Methyltransferase 1

MF-- Massa Fresca

MS -- Massa Seca

MITES -- Miniature Inverted-repeat Transposable Elements

PACMAD -- Panicoideae, Aristidoideae, Chloridoideae, Micrairoideae, Arundinoideae e Danthonioideae.

RdDM -- RNA-directed DNA methylation

RDR2 – RNA - Dependet de RNA Polymerase 2

RNAi – RNA interference

ROS1 – Repressor of Silencing 1 siRNA – Small interfering RNA

sRNA – Small RNA

TAIR - The Arabidopsis Information Resource

TE -- Transposable Elements

Sumário

1. Introdução 1
1.1. Família *Poaceae* 1
1.2. Filogenia e evolução de monocotiledôneas 1
1.3. Cana-de-açúcar 5
1.4. Milho 6
1.5. Epigenética 8
1.5.1. Metilação do DNA 9
1.5.2. Pequenos RNAs 11
1.5.3. Modificação de histonas 13
2. Justificativa 13
3. Objetivos 14
3.1. Objetivo geral 14
3.2. Objetivos específicos 14
4. Metodologia 15
4.1. Caracterização de genes envolvidos na regulação epigenética de plantas monocotiledôneas 15
4.1.1. Identificação e anotação de genes 15
4.1.2. Filogenia 16
4.1.3. Estrutura gênica 17
4.2. Identificação e caracterização de regiões MEMS 17
4.2.1. Localização dos genes e das regiões MEMS 17
4.2.2. Análise da abundância de siRNAs e expressão de genes de desmetilação 18
4.3. Análise do efeito da ausência de metilação do DNA durante o desenvolvimento de plântulas de milho e cana-de-açúcar 19
4.3.1. Crescimento de plântulas de milho *in vitro* 19
4.3.1.1. Material vegetal e tratamento com 5-Azac 19
4.3.1.2. Avaliação do desenvolvimento vegetal 20
4.3.2. Análise Molecular 20
4.3.2.1. Extração de RNA e qPCR 21
4.3.3. Crescimento de plântulas de cana-de-açúcar *in vitro* 22
4.3.3.1. Material vegetal e tratamento com 5-Azac 22
4.3.3.2. Avaliação do desenvolvimento vegetal 22
4. Resultados 23

4.1. Anotação, caracterização e filogenia dos genes envolvidos na regulação epigenética 23
4.1.1. Anotação e filogenia 24
4.1.2. Estrutura gênica 56
4.2. Identificação e caracterização de regiões MEMS 58
4.2.1. Milho 58
4.2.2. Arroz 62
4.3. Análise do efeito do inibidor de metilação na expressão da ROS1 e no desenvolvimento vegetal 66
5. Discussão 70
6. Conclusão 75
7. Referências 76
Apêndice A 82

1. Introdução

1.1. Família *Poaceae*

A importância econômica e a complexidade genômica que as *Poaceae* possuem são os principais motivos pelos quais os estudos com plantas dessa família têm sido enfoque de diversos laboratórios de pesquisas. Apesar desse grupo não ser o maior em número de espécies, as *Poaceae*, popularmente conhecidas como gramíneas apresentam-se como o principal grupo de plantas de importância econômica na produção mundial de grãos. Segundo os dados da FAO (Food and Agricultura Organization of the United Nations), no ano de 2017 a produção mundial de grãos ficou em torno de 3,41 bilhões de toneladas, sendo que 82,7% desse valor é atribuído a soma das produções de trigo, arroz, milho, cevada e sorgo. Estes últimos são os principais representantes agrícolas da família *Poaceae*, cujo soma de suas produções teve um total de 2,82 bilhões de toneladas.

Além das plantas citadas, as gramíneas possuem representantes que não são produtoras de grãos, mas que também compreendem grande importância econômica, como é o caso das diversas espécies de capins utilizados nas pastagens que alimentam o gado, influenciando diretamente nas produções de leite e carne (MUIR; PITMAN; FOSTER, 2011). Esse também é o caso da cana-de-açúcar, que é uma planta de grande importância no setor bioenergético, sendo a principal fonte de produção de açúcar e álcool do mundo (JAISWAL et al., 2017). Apenas no ano de 2017, a produção mundial de cana-de-açúcar ficou em torno de 1,8 bilhões de toneladas (FAO, 2017). Estima-se que em todo o mundo cerca de 26 milhões de hectares, distribuídos entre Brasil, Índia, China, Tailândia, Paquistão, México, Austrália, Colômbia, Guatemala e Estados Unidos, encontrem-se ocupados com o plantio de cana-de-açúcar (FAO, 2017).

1.2. Filogenia e evolução de monocotiledôneas

Análises filogenéticas realizadas com DNA mitocondrial e cloropastidial mostram que monocotiledôneas e dicotiledôneas divergiram-se a cerca de 200 milhões de anos (WOLFE et al, 1989; LAROCHE et al, 1995). O grupo das monocotiledôneas é dividido em oito ordens, dentre elas a ordem das *Poales,* que compreende 17 famílias, onde se insere a família das *Poaceae.* As *Poaceae* surgiram

a mais de 70 milhões de anos atrás (GAUT, 2002) e algumas características dessas primeiras gramíneas ainda são conservadas. Originalmente as gramíneas eram plantas que viviam as margens de florestas e locais de sombra, essas características ainda são observadas, por exemplo, nas espécies de bambus. No decorrer da história evolutiva, as gramíneas adquiriram a capacidade de sobreviverem em locais secos e abertos, sendo essa uma característica de extrema importância para fins agrícolas (KELLOGG, 2001). As monocotiledôneas teriam se originado de um ancestral comum que possuía apenas cinco cromossomos, porém, sucessivos eventos de poliploidia fizeram com que a família *Poaceae,* tenha se tornado uma das famílias das plantas com maior variação de número de cromossomos por espécie (GAUT, 2002). Como exemplo, o genoma do trigo possui seis cópias para cada cromossomo e aproximadamente 15 milhões de bases (ZIMIN et al., 2017) e o arroz que possui 65,7% de seu genoma duplicado (CALCAGNO et al., 2008).

A classificação filogenética atual das gramíneas divide 11.506 espécies em 12 subfamílias, que são divididas em tribos, subtribos, gêneros e espécies, como é mostrado na ordem decrescente por número de espécies na tabela 1.

Tabela 1: Classificação das subfamílias de *Poacea*.

Subfamílias	Tribos	Subtribos	Gêneros	Espécies
Poideae	15	30	202	3.968
Panicoideae	13	19	247	3.241
Bambusoideae	3	15	125	1.670
Chloridoideae	5	26	124	1.602
Aristidoideae	1	--	3	367
Danthonioideae	1	--	19	292
Micraroideae	3	--	8	184
Oryzoideae	4	2	19	115
Arundinoideae	2	2	14	40
Pharoideae	1	--	3	12
Puelioideae	2	--	2	11
Anomocloideae	2	--	2	4

As subfamílias *Poideae, Panicoideae, Bambusoideae* e *Clhorodoideae* juntas correspondem a 90% das espécies de gramíneas. No entanto a filogenia das *Poaceae*, representada na figura 1, mostra a existência de dois grandes clados, chamado BOP e PACMAD. O clado BOP é composto pelas subfamílias *Poideae*, *Bambusoideae* e *Oryzoideae*. Já o clado PACMAD é integrado pelas subfamílias *Panicoideae, Chloridoideae, Aristidoideae, Danthonioideae, Micrairoideae* e *Arundinoideae*. As subfamílias *Pharoideae, Pueloideae* e *Anomocloideae* são consideradas tribos isoladas e não se enquadram nos clados BOP e PACMAD. O clado BOP engloba plantas de importância econômica como trigo, cevada, aveia e arroz, já o clado PACMAD é onde se encontram o milho, sorgo e cana-de-açúcar. Esses clados divergiram-se a cerca de 50 milhões de anos. Dentro do clado PACMAD, sorgo e milho divergiram das setárias a cerca de 28 milhões de anos atrás e entre si a cerca de nove milhões de anos atrás (GAUT, 2002; SORENG et al., 2017).

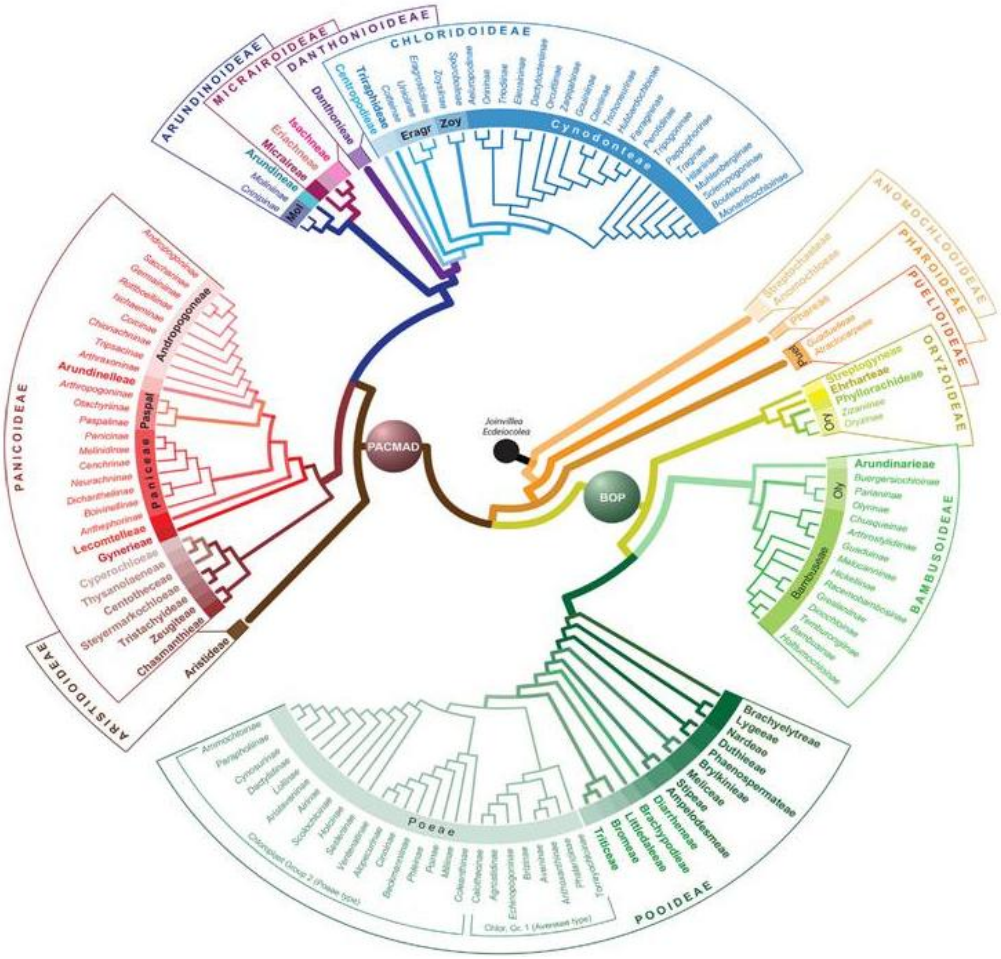


Figura 1: Classificação filogenética das *Poaceae*, incluindo 12 subfamílias, 53 tribos e 90 subtribos evoluindo no sentido horário a partir de Joinvillea e Ecdeiocolea (*outgroups*). BOP = Bambusoideae, Oryzoideae e Poideae; PACMAD = Panicoideae, Aristidoideae, Choridoideae, Micrairoideae, Arundinoidea e Danthonioideae. Ramos grossos representam as divisões das subfamílias; meio médios representam as divisões das tribos e ramos finos representam as divisões das subfamílias (modificado de Sorent ET AL, 2017).

Dentre todas as famílias e subfamílias existentes no grupo das monocotiledôneas, as subfamílias *Poideae* e *Panicoideae* são as que possuem maior relevância, tanto em número de espécies, que somadas ultrapassam 7.000, quanto a importância econômica, pois é onde estão localizadas plantas como *Pannicum*, *Miscanthus*, *Zea mays, Saccharum spontaneum, Soghum bicolor* entre outras (BURKE et al., 2016).

1.3. Cana-de-açúcar

A cana-de-açúcar (*Saccharum spp*.) é uma gramínea alta, podendo chegar até seis metros de altura dependendo da variedade, predominantemente plantada em países de clima tropical, como pode ser visto na figura 2. Pertence à família *Poaceae* e a subfamília *Panicodeae*, mesma família de outras plantas com importância econômica como arroz, milho, sorgo e trigo (CHEAVEGATTI-GIANOTTO et al., 2011). O gênero *Saccharum* surgiu antes que os continentes tivessem as formas atuais, portanto, das 35 – 40 espécies descritas desse gênero, algumas são consideradas nativas da Ásia/África e outras são consideradas nativas das Américas. As espécies *Saccharum barberi* e *Saccharum sinense* foram as primeiras a serem cultivadas na Índia e na China, porém as espécies *Saccharum spontaneum* e *Saccharum officinarum,* deram origem as cultivares produzidas atualmente. *S. officinarum* que possui um genótipo 2n = 80, é conhecida por ser uma cana nobre pela alta produção de açúcar, já *S. spontaneum* possui grande variabilidade genética sendo 2n = 40 a 128. A hibridização proporcionou grande avanço na produção de cana-de-açúcar, resolvendo problemas de doenças, aumento de produtividade e melhorando a capacidade de adaptação para o crescimento em condições de estresses. Estudos mostram que os genomas dos híbridos modernos são constituídos de 10-20% de cromossomos de origem de S. spontaneum, 17% são de uma combinação entre cromossomos de *S. spontaneum* com *S. officinarum* e o restante é composto por cromossomos de origem de *S. officinarum* (D'HONT, 1996). Apesar da complexidade genômica, experimentos utilizando cana-de-açúcar, tem gerado bons resultados, incluindo sequenciamentos genômicos de algumas cultivares, como a R570 (GARSMEUR et al., 2018), S. *spontaneum* AP85-441(ZHANG et al., 2018) e a cultivar SP80-3280 (SOUZA et al., 2019).

As primeiras mudas de cana-de-açúcar foram trazidas ao Brasil pelo português Martim Afonso de Souza em 1517 e plantadas na capitania de São Vicente (hoje São Paulo), onde também foi construído o primeiro engenho de açúcar. Porém, foi no Nordeste, principalmente em Pernambuco, que tivemos as primeiras grandes produções de cana-de-açúcar (PARANHOS, 2008). Em 2017, dos 1,8 milhões de toneladas de cana-de-açúcar colhidas no mundo, cerca de 40% foi colhido no Brasil, que teve uma safra com mais de 750 mil toneladas. Esses números colocam o Brasil em primeiro lugar na produção mundial de cana-de-açúcar, seguido pela Índia em segundo lugar e a China que ocupa o terceiro lugar. Com relação aos principais derivados, o Brasil é o maior produtor mundial de açúcar e o segundo maior produtor

mundial de etanol, ficando atrás de Estados Unidos e à frente de países como a China e Canadá (FAOSTAT, 2017). A cana-de-açúcar é uma opção de energia sustentável, podendo diminuir consideravelmente o consumo de petróleo fóssil e a emissão de CO_2 na atmosfera. Segundo a Agência Internacional de Energia, até 2045 a produção de bioenergia através da cana-de-açúcar, pode deslocar até 13,7% do consumo de petróleo bruto e 5,6% da emissão de CO_2, comparados aos dados de 2014 (JAISWAL et al., 2017).

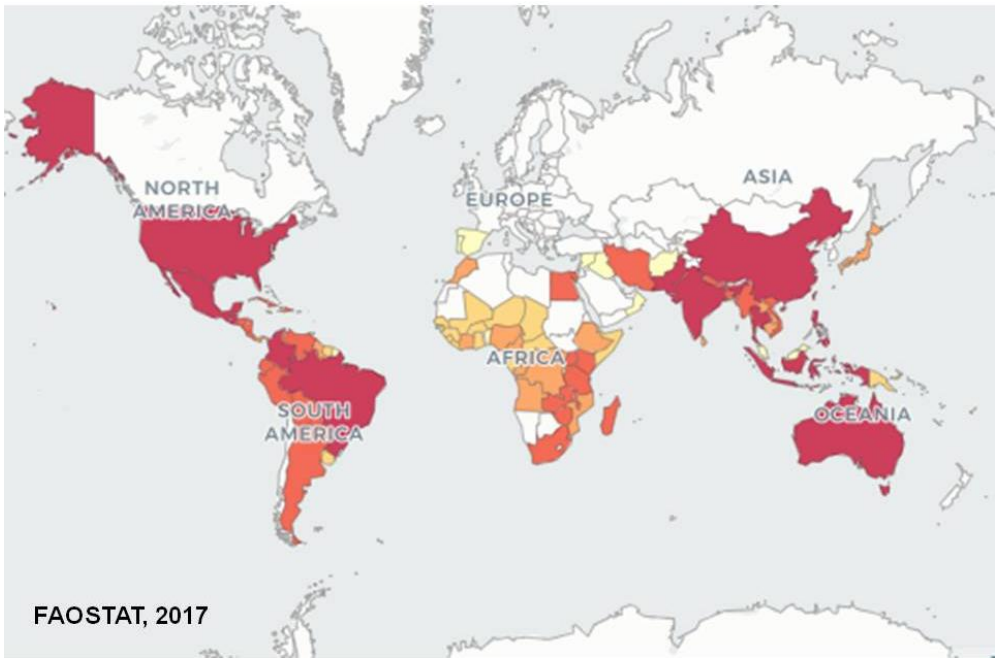


Figura 2: Distribuição da produção de cana-de-açúcar no mundo. Países coloridos em vermelho escuro são os maiores produtores.

1.4. Milho

O milho (*Zea mays ssp*.), assim como a cana-de-açúcar, pertence a família *Poaceae*, a subfamilia Panicoideae e a tribo Andropogoneae (GAUT, 2002). O milho que conhecemos hoje é uma planta que foi domesticada a cerca de nove mil anos atrás, a partir de um ancestral ainda vivente, conhecido como teosinto e que tem sua origem na América Central (KISTLER et al., 2018). O teosinto, conhecido como milho selvagem é citologicamente muito parecido com o milho domesticado, podendo inclusive formar híbridos entre eles, porém, morfologicamente possuem muitas diferenças. O teosinto possui ramos laterais alongados com pendão na ponta, já o

milho atual possui apenas um pendão apical e os ramos laterais dão origem às espigas. As espigas também possuem grandes diferenças, sendo que a espiga do milho é maior e possui várias fileiras de grãos, diferentemente do teosinto que possui uma espiga menor e com apenas duas fileiras de grãos (DOEBLEY, 2004). O milho chegou ao Brasil ainda em processo de domesticação, a cerca de seis mil anos atrás, através de tribos indígenas da Amazônia (KISTLER et al., 2018). Hoje existem diversas variedades e linhagens de milho que são cultivadas em todo o mundo. Os híbridos formados a partir das linhagens puras podem ser classificados em híbrido simples, duplo híbrido ou triplo híbrido, dependendo do cruzamento que for realizado. Uma linhagem pura é formada a partir de sementes crioulas, ou seja, sementes ancestrais que são autofecundadas por pelo menos cinco gerações. Quando falamos de variedades de milho, é importante ressaltar a cultivar B73, pois teve grande relevância para a ciência, sendo o organismo base para importantes descobertas como os elementos transponíveis (TE) (MCCLINTOCK, 1951), a família dos MITES (FESCHOTTE; JIANG; WESSLER, 2002) e os retrotranspons em plantas (LAFAYETTE, 1999).

Atualmente o milho é um dos cereais mais produzidos no mundo, utilizado na alimentação humana e na ração de diversos animais, além de ser fonte para produção de etanol. Em 2017 foram colhidos mais de 1,1 bilhões de toneladas de milho no mundo, sendo os Estados Unidos o maior produtor, seguido de Brasil no segundo lugar e da China no terceiro lugar. Porém o Brasil vem crescendo muito na produção de milho. De 2015 a 2017, a área ocupada com plantações de milho nos Estados Unidos e na China foi reduzida, enquanto no Brasil houve uma significante expansão. Em 2017 o Brasil colheu 97,7 milhões de toneladas de milho, um número ainda distante dos 370 milhões de toneladas colhidas pelos Estados Unidos, mas as perspectivas são boas, como mostram a figura 3, onde a FAO compara a área revestida com plantação de milho e a quantidade colhida (FAOSTAT, 2017).

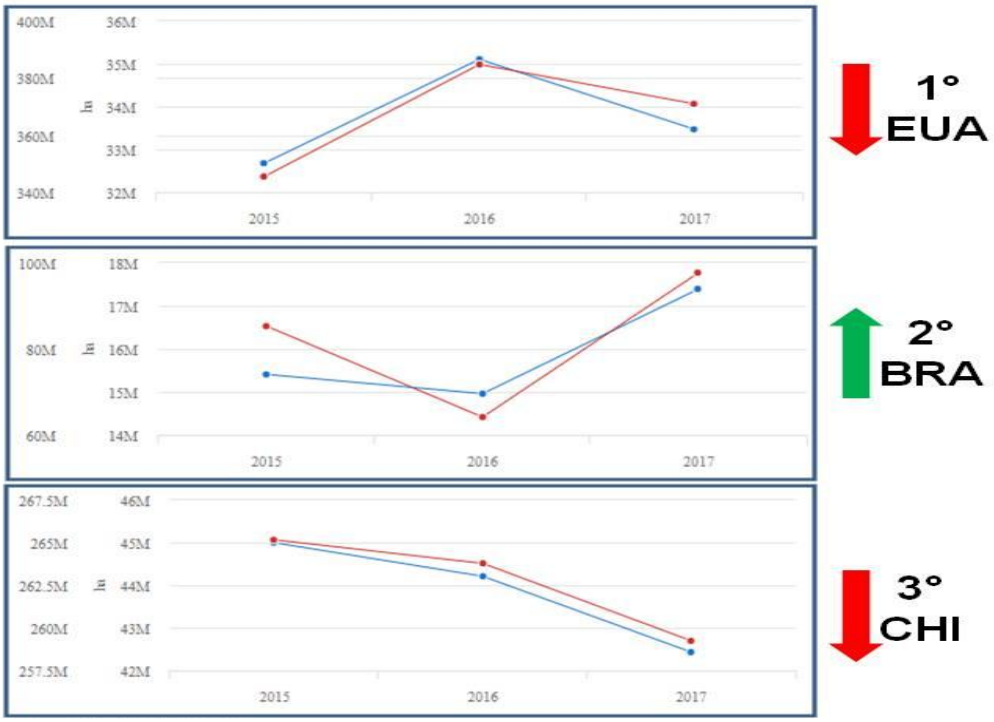


Figura 3: Comparação entre a área revestida com plantação de milho e a produção nos três maiores produtores mundiais do grão. Linha vermelha representa a quantidade produzida em toneladas e a linha azul representa o número de hectares plantados com milho (FAOSTAT, 2019).

O milho é cultivado em duas épocas distintas do ano, a primeira safra, também chamada de safrinha, é realizada entre abril e junho e corresponde a uma safra menor devido aos períodos de seca que afetam os estágios de desenvolvimento do grão, a segunda safra é plantada no final do ano e colhida no início do ano seguinte. O milho é cultivado em todas as regiões brasileiras e se divide entre produção empresarial e na agricultura familiar, contudo, o estado do Mato Grosso é o maior produtor seguido pelo estado do Paraná (CONAB, 2018).

1.5. Epigenética

A epigenética é uma ciência que estuda modificações que ocorrem nos genomas que não alteram a sequência de nucleotídeos do DNA (WEINHOLD, 2006).

Essas modificações são importantes para o processo de adaptação dos organismos em situações adversas como, por exemplo, estresses bióticos e abióticos. Em plantas, que são organismos sésseis, os mecanismos epigenéticos podem atuar para que as mesmas consigam se aclimatar nas diferentes condições que lhes são impostas. A epigenética atua fortemente nas vias de regulação da expressão gênica (PIKAARD; SCHEID, 2014).

As mudanças epigenéticas podem ser hereditárias e atuam principalmente silenciando genes, transcricionalmente ou pós-transcricionalmente. Dentre alguns mecanismos já descritos, destacam-se as modificações de histonas e a metilação do DNA, que são processos que regulam a transcrição do DNA. Outro mecanismo epigenético muito importante envolve a participação de pequenos RNAs e atua regulando a transcrição de genes e também na regulação pós-transcricional (JIN; LI; ROBERTSON, 2011).

1.5.1. Metilação do DNA

O processo de metilação do DNA consiste em adicionar um grupamento metil no carbono 5' das citosinas do gene alvo. O processo de metilação em plantas pode ocorrer na região promotora ou no corpo do gene e quase sempre atua silenciando a sua transcrição, inibindo a ligação de ativadores, promovendo a ligação de repressores ou atuando indiretamente na modificação de histonas (DOMCKE et al., 2015). Em plantas, a citosina pode ser metilada no contexto CG, CHG ou CHH (ZEMACH et al., 2010) A inserção do grupo metil ocorre através de enzimas com caráter específico para tal função. Em plantas, a DNA METHYLTRANFERASE 1 (MET1), a CHROMOMETYLASE 3 (CMT3) e a DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2) são enzimas que adicionam grupamento metil no DNA. A manutenção da metilação em plantas é realizada pela MET1, que é um ortólogo da DNMT1 de mamíferos (KANKEL et al., 2003). A MET1 possui um domínio chamado de CXXC rico em cisteína, que auxilia na metilação no contexto CG (SONG et al., 2011). A metilação no contexto CHG é realizado pela CMT3 que é uma proteína especifica de plantas, já no contexto CHH a metilação é realizada pela DRM2, que atua preferencialmente em regiões de transposons (THAO; HARMER; ZILBERMAN, 2014). Em plantas o processo de desmetilação ocorre de forma ativa, onde a citosina metilada é simplesmente removida para dar lugar a outra citosina. As plantas possuem pelo menos quatro famílias de enzimas que realizam esse processo, DME, ROS1, DML2 e DML3 (AGIUS; KAPOOR; ZHU, 2006). DME atua preferencialmente em

células gaméticas tanto masculinas quanto femininas. Já a ROS1, DML2 e DML3 atuam em todos os tipos de células e tecidos, agindo preferencialmente em transposons próximos aos genes (HUH et al., 2008). A ROS1 realiza um trabalho antagônico a RdDM (*RNA-directed DNA methylation*), para evitar a hipermetilação do genoma (DU et al., 2014; LEI et al., 2015). Além disso, ROS1 funciona como um indicador da atividade da MET1 e da via de RdDM. Esse comportamento da ROS1 só é possível porque na região promotora de seu gene, existe uma sequência específica de 39 pb, que funciona como um metilstato, medindo os níveis de metilação no DNA e assim controlando a expressão da ROS1 (WILLIAMS et al., 2015). O mecanismo RdDM controla diretamente a expressão de ROS1. Ao contrário do que acontece nos promotores dos outros genes, a metilação na região MEMS causa um efeito de regulação positiva da ROS1 (figura 4) que, sendo expressa, efetuará a desmetilação das regiões metiladas por RdDM, ou seja, a regulação da metilação da região MEMS é efetuada por RdDM e ROS1 (ZHANG; LANG; ZHU, 2018).

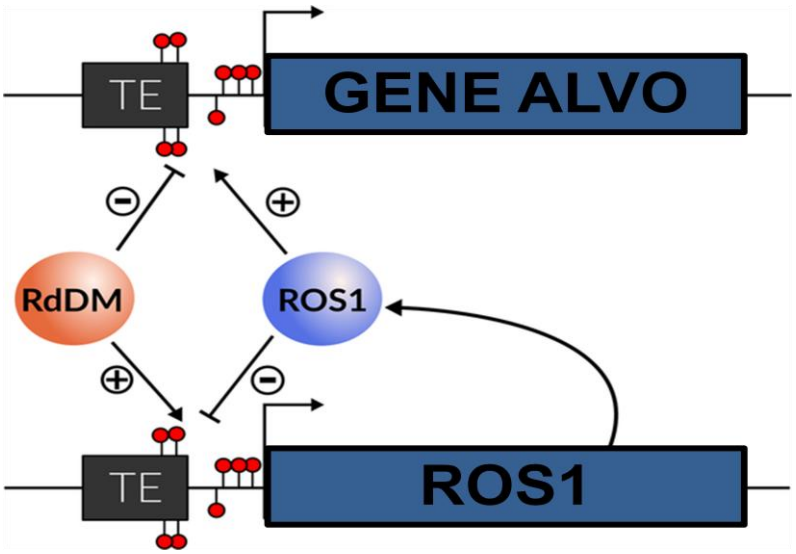


Figura 4: A metilação atua como um sensor epigenético na regulação de ROS. Modificado de (WILLIAMS et al., 2015)

Estudos com plantas mutante para MET1, consequentemente hipometiladas no contexto CG, mostram que as gerações sucessivas possuem um aumento nos níveis de metilação no contexto CHH, como resultado da ação de RdDM e a falta de expressão de ROS1 (MATHIEU et al., 2007), o que corrobora a ideia de que ROS1 pode ser um indicador da atividade de MET1.

Durante o desenvolvimento das plantas a metilação é controlada nos diferentes tecidos. Na gametogênese tanto masculina quanto feminina, a MET1 é reprimida e ocorre uma grande atuação da DME (LEI et al., 2015) . A ação da DME faz com que transposons sejam desmetilados e gere transcritos, formando siRNAs que atuaram na metilação do DNA direcionada por RNA (RdDM) (IBARRA et al., 2012; SLOTKIN et al., 2010). A metilação do DNA também está envolvida no processo de amadurecimento dos frutos, como já foi demonstrado em tomate e em maçã (HE et al., 2009; LANG et al., 2017; TELIAS et al., 2011). Fatores ambientais também podem alterar a metilação do DNA em plantas. Em *Medicago truncatula* foi observada uma alteração na expressão de DME em plantas infectadas com microrganismos (NAGYMIHÁLY et al., 2017; SATGÉ et al., 2016). Em soja, plantas infectadas por nematoides possuem uma hipometilação em raízes (RAMBANI et al., 2015). Alguns estudos mostram que plantas mutantes para os efetores da metilação ou da desmetilação, podem adquirir maior resistência a patógenos. Por exemplo, em Arabidopsis, mutantes para ROS1 possuem maior desenvolvimento de vasos condutores quando em contato com o patógeno Hyoloperonospora (ANA, 2016). Um mutante para ROS1, DML2 e DML3 de Arabidopsis demonstrou maior susceptibilidade ao fungo *Fusarium oxysporium*, indicando que a regulação da metilação do DNA em plantas submetidas a patógenos, está envolvida com o sistema de defesa (LE et al., 2014).

Estudos da metilação do DNA em plantas sujeitas a seca, salinidade, frio, radiação e deficiência nutricional tem sido cada vez mais abordados em diversas plantas como Arabidopsis, milho, arroz, cevada, trigo e outras (SECCO et al., 2015). Por exemplo, em tomate o frio causa uma baixa expressão da DML2, responsável por regular genes que atuam na síntese do sabor do tomate, fazendo com que em condição de frio, os tomates percam sabor (ZHANG et al., 2016).

1.5.2. Pequenos RNAs

Pequenos RNAs não-codificadores (sRNAs), que consistem em 20-24 nucleotídeos (nt), têm sido cada vez mais investigados como importantes reguladores da expressão gênica. Na década de 90, trabalhos com *Caenorhabditis elegans*, identificaram RNAs em formato de dupla fita (dsRNA), e descobriram que essas moléculas eram capazes de interferir na expressão de genes importantes no fenótipo (FIRE et al, 1998). A partir daí, com o avanço da tecnologia, começaram a surgir diversos trabalhos relacionados à caracterização de pequenos RNAs em diversos organismos. Atualmente a técnica de interferência por RNA (RNAi), é caracterizada em

animais, plantas e leveduras, e possui grande importância no conhecimento da organização do genoma desses organismos (SLEDZ; WILLIAMS, 2005) .

Trabalhos com plantas ocorreram ainda na década de 1990, utilizando petúnias, já indicava a possível existência desses pequenos RNAs Interferentes. Atualmente sabemos que em plantas, duas categorias principais de sRNAs reguladores são distinguidos com base na sua biogênese e função: microRNAs (miRNAs) e RNAs de interferência (siRNAs).

A biogênese de um microRNA pode ter processos diferentes entre os organismos. Em plantas, um microRNA emerge a partir de um transcrito nascente da RNA Polimerase 2. Esse transcrito, logo após a sua síntese, adquire um formato de dupla fita na forma de um *hairpin*. A partir desse transcrito, chamado pri-microRNA, até o microRNA maduro, que é a forma como ele atua, ocorrem processamentos na molécula por diversas enzimas. Uma enzima muito importante nesse processo de maturação do microRNA é a Dicer-like (DCL), que é responsável pela clivagem do pri-microRNA em pedaços de dsRNA menores, com aproximadamente 21 nucleotídeos (DUNOYER; VOINNET, 2005). Outra enzima importante que também participa desse processo de maturação do microRNA é a RNA-DEPENDENT DE RNA POLYMERASE 2 (RDR2) que faz a replicação de transcritos de regiões transponíveis (XIE et al., 2004; ZHAO et al., 2008). Logo após o processamento, outras enzimas atuam na separação das fitas do microRNA, e uma das fitas se liga a um complexo enzimático denominado RISC/ARGONAUTA, que irá conduzir o microRNA até o gene alvo (CHENDRIMADA et al., 2005).

Os siRNAs são pequenos RNAs dupla fita que, em geral, são produzidos a partir da transcrição de elementos repetitivos do genoma, como os transposons. Uma gama de enzimas é responsável pela geração da dupla fita de 24-nt de tamanho, que poderá atuar silenciando a classe de transposons que o gerou (GRATIVOL et al., 2012). Neste mecanismo de silenciamento participam as enzimas RDR2, DCL3 e AGO4 e compõem o mecanismo conhecido como RdDM, no qual siRNAs induzem a metilação do DNA em regiões específicas. Esse mecanismo possui uma importante atuação em plantas submetidas a diferentes estresses. Estudos mostram que ocorre uma variação de expressão dos genes Dicer-like, Argonautas e RDR em arroz submetido ao frio, seca e salinidade (KAPOOR et al., 2008) e em tomate submetido a patógenos (BAI et al., 2012). As Dicer-like e as Argonautas são ribonucleases com múltiplos domínios. As Dicers possuem seis domínios importantes, DExD-helicase, helicases-C, DUF283, PAZ, RNasIII e dsRB (MARGIS et al., 2006). EM plantas, Dicer-like 1 e 4 clivam pequenos RNAs de 24 nucleotídeos, Dicer-like 2 cliva pequenos RNAs de 22 nucleotídeos e Dicer-like 3 cliva pequenos RNAs de 24 nucleotídeos (BASSO et al., 2019). O domínio PAZ

presente nas Dicer-like, também está presente nas Argonautas, isso porque ambas proteínas interagem com siRNAs (LINGEL et al., 2003; SONG et al., 2003). O domínio Paz interage com a extremidade 3' do siRNA. Além do PAZ, as Argonautas possuem o domínio PIWI (HALL; CAROLINA, 2005).

1.5.3. Modificação de histonas

A organização da dupla fita de DNA é dada através da cromatina, onde se encontram nucleossomos que se enovelam a dupla fita e causam a compactação do DNA. As regiões onde o DNA está mais compactado são chamadas de heterocromatina e as regiões onde o DNA está menos compactado são chamadas de eucromatina. Quando um determinado gene está acessível para transcrição, dizemos que ele está ativo, e quando a heterocromatina está formada na região deste gene, dizemos que ele está inativo ou silenciado. Para entender melhor esse processo de movimentação dos nucleossomos, precisamos entender que ele é composto por oito histonas, que são proteínas que possuem uma cauda rica em arginina e lisina que é submetida a processos de acetilação, fosforilação, ubiquitinização e metilação (LIU et al., 2014; TARIQ; PASZKOWSKI, 2004). Esses processos alteram a conformação e a carga das histonas e causam o movimento do nucleossomo e a remodelagem da cromatina. Em plantas, essas modificações de histonas estão diretamente associadas à regulação da expressão gênica (BERR; SHA; SHEN, 2011).

2. Justificativa

Nosso planeta vem passando por alterações climáticas, causadas inclusive, por ações humanas, e uma delas é a grande quantidade de dióxido de carbono lançado na atmosfera (NOHARA et al., 2013). Este problema faz com que o efeito estufa traga consequências como o aumento da temperatura global e consequentemente trazendo problemas para a agricultura (NELSON et al., 2014). Outro problema mundial e que também afeta a agricultura, é a superpopulação. Em 1960 o planeta era habitado por quatro bilhões de seres humanos, em apenas 60 anos esse número já está quase o dobrado, totalizando 7,7 bilhões de habitantes e a expectativa é que em 2050 o planeta tenha cerca de 10 bilhões de habitantes (SCHÜTZ; JANSEN; VERHOFF, 2011) e uma das preocupações é como alimentar todo mundo. Nesse aspecto, a ciência tem contribuído de maneira significativa para o aumento da produção agrícola,

através de pesquisas que contribuem para o melhoramento vegetal a fim de obter cultivares mais produtivas, resistentes a pragas e doenças e capazes de se adaptarem a diferentes climas. Esse melhoramento pode ser realizado de maneira clássica, fazendo cruzamentos de variedades diferentes a fim de selecionar as melhores características até que se tenha uma nova cultivar com as melhores características de cada ancestral (RONALD, 2011). Outra ferramenta importante nesse processo de melhoramento vegetal é a engenharia genética, principalmente através da transgenia, que possibilita uma rapidez na seleção das características desejáveis para cada cultura. Grande parte dos alimentos produzidos hoje em todo o planeta são produtos de plantas geneticamente modificadas (HALLMANN, 2007).

Diante da responsabilidade que a ciência possui nessas questões citadas acima, é de extrema importância à investigação molecular, a fim de desvendar diversos mecanismos que possam contribuir para tais estudos de melhoramento e obtenção de plantas resistentes a doenças e pragas. Os mecanismos de regulação da expressão gênica podem ser essenciais nessas conquistas e os estudos que envolvem mecanismos epigenéticos ainda possuem um grande campo a ser explorado. Sendo assim, a proposta deste trabalho é contribuir para o conhecimento molecular que envolve a regulação epigenética de plantas monocotiledôneas.

3. Objetivos

3.1. Objetivo geral

O objetivo desse trabalho foi caracterizar genes envolvidos nas vias de metilação e desmetilação e, vias de pequenos RNAs em plantas monocotiledôneas e o efeito regulatório da desmetilação na expressão desses genes.

3.2. Objetivos específicos

A. Identificar e caracterizar os genes de metilação e desmetilação do DNA: MET1, CMT3, DRM2, DME, ROS1, DML2 e DML3 e da via de pequenos RNAs: Dicer-like, RDR2 e Argonautas em genomas de plantas monocotiledôneas.

B. Identificar regiões de regulação da expressão de genes de desmetilação do DNA em milho e arroz.

C. Analisar o efeito da desmetilação do DNA no fenótipo de plântulas de milho e cana-de-açúcar e na regulação de ROS1.

4. Metodologia

4.1. Caracterização de genes envolvidos na regulação epigenética de plantas monocotiledôneas

4.1.1. Identificação e anotação de genes

Sequências dos genes de metilação e desmetilação do DNA, MET1, CMT3, DRM2, DME, DML2, DML3, ROS1 e dos genes que estão envolvidos nas vias de pequenos RNAs e pertencem as famílias Dicer-like 1- 4 e Argonautas 1 – 10, de *Arabidopsis thaliana*, foram obtidos a partir do banco de dados TAIR (https://www.arabidopsis.org/) (Figura 5). A partir dessas sequências, foi realizado um alinhamento no banco de dados Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html) para identificação desses genes nas plantas monocotiledôneas: *Oryza sativa, Zea mays, Sorghum bicolor, Brachypodium distachyon, Hordeum vulgare, Panicum hallii, Setaria itálica, Setaria viridis, Triticum aestivum, Musa acuminata, Miscanthus sinensis.* Além disso, foram obtidas sequências de *Amborela trichopoda*, uma angiosperma basal que possui um ancestral comum a todas outras citadas acima. Foram consideradas apenas as sequências que apresentaram o valor de *e-value* igual a zero. Sequências de cana-de-açúcar foram obtidas a partir do alinhamento de sequências de sorgo com o genoma de *S. spontaneum* publicado por (ZHANG et al., 2018). Através desse alinhamento foram selecionadas as sequências com melhores valores de *e-value* e similaridade. Sequências redundantes foram identificadas e eliminadas. A identificação de cada gene foi anotada, assim como o número de nucleotídeos dos genes e dos transcritos e o número de aminoácidos das proteínas. Foram anotadas também o número de éxons, a região alinhada, os valores de *e-value* e similaridade de cada alinhamento.

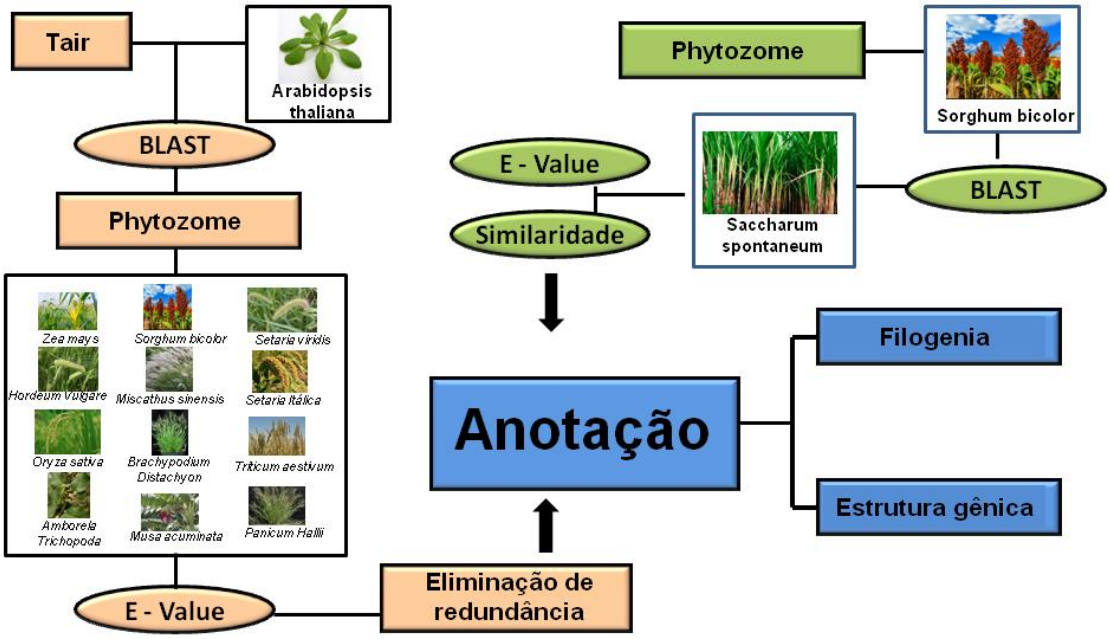


Figura 5: Pipeline da metodologia utilizada para anotação gênica. Boxes em verde apresentam a anotação das sequências de cana-de-açúcar e os boxes em rosa apresentam a anotação das sequências dos genes das demais plantas.

4.1.2. Filogenia

As sequências de aminoácidos dos efetores de metilação e desmetilação do DNA e das vias de pequenos RNAs anotadas de quatorze plantas diferentes, foram utilizadas para determinação da filogenia, utilizando o software MEGA 7 (KUMAR; STECHER; TAMURA, 2016). Primeiramente, as sequências foram separadas em quatro grupos diferentes, Metilases, Desmetilases, Dicer-likes e Argonautas. Cada grupo foi alinhado separadamente utilizando a ferramenta Muscle e posteriormente foram submetidas a uma análise para obtenção do melhor modelo evolutivo a ser aplicado durante construção da arvore filogenética. O modelo proposto por Jones, Taylor e Thornton em 1992 (JTT) foi aplicado para todos os grupos analisados. Nesse modelo, as sequências são agrupadas com base em um peptídeo comum, assim os pares de sequências com aproximadamente 85% de similaridade são agrupados e a troca de aminoácidos entre essas sequências são contadas (JONES; TAYLOR; THORNTON, 1992). Para os grupos de Metilases e Argonautas, foi acrescentado ainda o fator gama (G) que avalia a probabilidade dos sítios pertencerem a categoria em que estão inseridos, além disso, as frequências são calculadas com base nos dados disponíveis (F), com isso a fórmula aplicada para esses dois grupos foi (JTT+G+F). Para os grupos de Desmetilases e Dicer-likes, o mesmo modelo foi

aplicado com a adição de um fator que calcula a proporção de sítios invariantes (I) e, portanto, a fórmula utilizada para esses dois grupos foi (JTT+G+I+F).

4.1.3. Estrutura gênica

A predição da estrutura gênica dos genes de metilação e desmetilação do DNA e de vias de pequenos RNAs em milho e sorgo foi realizada utilizando sequências gênicas obtidas no Phytozome. Foi realizada a contagem do número de nucleotídeos das regiões 5'UTR, CDS e 3'UTR de cada gene.. As informações foram anotadas em arquivo de texto. O site (http://gsds.cbi.pku.edu.cn/index.php) foi utilizado para realizar a predição da estrutura de cada gene.

4.2. Identificação e caracterização de regiões MEMS

4.2.1. Localização dos genes e das regiões MEMS

As localizações dos genes de desmetilação do DNA de milho e arroz foram obtidas a partir do banco de dados Phytozome (Figura 6). Foram anotados seus respectivos IDs, o cromossomo em que cada gene se encontra bem como a localização exata de início e fim do gene, além do sentido em que se encontra. Essas informações estão disponíveis na tabela 2.

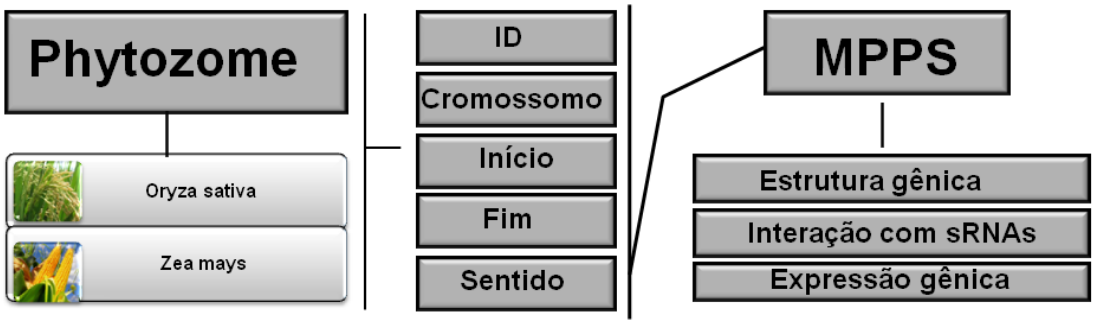


Figura 6: Pipeline das análises computacionais para identificação de regiões MEMS e a interação de siRNAs com genes de desmetilação em plantas monocotiledôneas.

Tabela 2: Localização cromossômica dos genes de desmetilação de DNA em milho e arroz.

Planta	ID	Cromossomo	Início	Fim	Sentido
Milho	GRMZM2G131756_T01	6	150018275	150029231	Reverso
	GRMZM2G422464_T02	4	218670827	218688798	Reverso
	GRMZM2G123587_T01	5	161900129	161946558	Reverso
Arroz	LOC_Os01g11900.1	1	6444319	6456076	Direto
	LOC_Os05g37410.1	5	21886293	21896908	Reverso
	LOC_Os05g37350.1	5	21843770	21856444	Reverso
	LOC_Os02g29230.1	2	17354840	17362939	Direto

4.2.2. Análise da abundância de siRNAs e expressão de genes de desmetilação

As informações obtidas no Phytozome foram inseridas no site (https://mpss.danforthcenter.org/) (NAKANO et al., 2019) para obtenção de dados de expressão de pequenos RNAS nos genes de desmetilação do DNA. Imagens detalhadas de regiões com 3.000 pares de bases que antecedem o início de cada gene foram demonstradas em *O. sativa*, além das imagens demonstrativas, o site possibilitou que fosse feita uma análise quantitativa desses pequenos RNAs em diferentes tecidos da planta. A mesma metodologia foi aplicada aos genes de desmetilação de *Z. mays*, com o acréscimo de uma análise de expressão do gene ROS1 em diferentes tecidos na planta. As informações de expressão gênica e abundância de pequenos RNAs foram dispostas em tabelas e utilizadas para construções de gráficos no programa PRISM.

4.3. Análise do efeito da ausência de metilação do DNA durante o desenvolvimento de plântulas de milho e cana-de-açúcar

4.3.1. Crescimento de plântulas de milho *in vitro*

4.3.1.1. Material vegetal e tratamento com 5-Azac

Foram utilizadas sementes de milho da cultivar Dekaulb (DKB) 177, que passaram por processo de desinfestação por 1 minuto com álcool 70% e dez minutos com hipoclorito, seguido de lavagem com água Mili-Q (Figura 7). Para quebra de dormência, as sementes foram mantidas imersas por 5 horas em água Mili-Q.

O meio de cultura foi preparado utilizando sacarose, MS 1/10, Agar 8%, a solução foi diluída em água destilada, misturada com auxílio de um agitador magnético e o pH foi calibrado para 5,8 utilizando uma solução de KOH. Após ser autoclavado, o meio foi vertido em tubos de ensaio com volume de 20 ml cada.

Para condição de tratamento, foi adicionado o inibidor de metilação 5 - Azacitidina (5-Azac) diluído em DMSO com concentração final de 100µM em cada tubo. Para condições de controle, foram preparados tubos de meio de cultura com o mesmo volume de DMSO e tubos com o mesmo volume de água. Os tubos de ensaio contendo as sementes de milho foram mantidos por sete dias em câmara de vegetação do tipo BOD, em temperatura de 28°C e foto período de 12h luz/12h escuro.

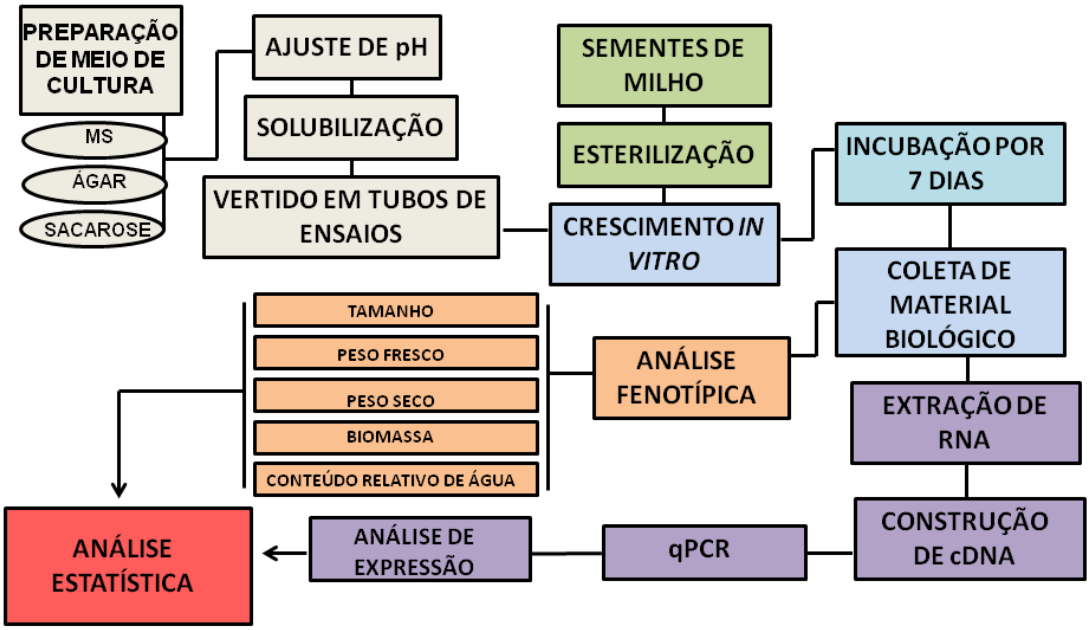


Figura 7: Fluxograma das metodologias utilizadas para realização de análises fenotípicas e moleculares de plântulas de milho em condições controle e tratado.

4.3.1.2. Avaliação do desenvolvimento vegetal

Foram consideradas duas réplicas para cada condição, sendo que na condição de controle com H_2O, cada réplica foi composta de *pool* de seis plântulas e nas condições de controle com DMSO e tratado com 5-Azac cada réplica foi composta por *pool* de cinco plântulas. As plântulas de sete dias tiveram parte aérea e raiz medidas com auxílio de um paquímetro digital., Parte aérea e raiz foram pesadas separadamente em uma balança de precisão, para obtenção de massa fresca e depois foram incubadas em estufa de aquecimento e mantidas por 24 horas à 60 °C e pesadas novamente para obtenção de massa seca. Com os valores de massa fresca e massa seca, foi realizado o cálculo para obtenção dos valores de biomassa através da fórmula (MS/MF.100) e dos valores de conteúdo relativo de água através da fórmula (MF-MS/MF).100, onde MS= massa seca e MF=massa fresca. Os valores de tamanho, massa fresca e massa seca, biomassa e conteúdo relativo de água, foram submetidos ao programa PRISM para avaliações estatísticas.

4.3.2. Análise Molecular

4.3.2.1. Extração de RNA e qPCR

Amostras de raiz e parte aérea de plântulas de milho foram maceradas em nitrogênio líquido e o RNA total foi extraído seguindo o protocolo de Trizol-Thermo Fisher. O pellet foi eluído em água DEPC e a quantificação de RNA total nas amostras foi realizado no espectofotômetro NANODROP 2000 - Thermo Fisher. Foram quantificados 10 µg de RNA de cada amostra com volume final de 50µl. O preparo do mix de DNAse, para cada amostra, foi realizado utilizando 25 ul de tampão (200mM Tris-Cl pH8,3; 500mM KCL; 25mM MgCl2; DEPC 0,1%), 1µl de DNAse e 24µl de água, totalizando 50µl. Em cada amostra foi aliquotado 50µl do mix, totalizando 100µl em cada eppendorf. As amostras foram incubadas por 10 minutos à 37°. Em seguida foi adicionado 500 ul de água DEPC e 600 ul de Fenol. As amostras foram misturadas por inversão manual e centrifugadas a 10.000 rpm a 4° por 10 minutos. Após a centrifugação foi coletado o sobrenadante e adicionado o volume recuperado de clorofórmio: álcool isoamílico (24:1). Após mistura por inversão, as amostras foram novamente centrifugadas por10.000 rpm a 4° por 10 minutos e o sobrenadante foi coletado. Adicionamos 1 ul de Glycoblue, 1/10 (do volume coletado) de Acetato de Sódio (3M NAOAc pH 5,2) e 2 volumes de etanol 100% frio. Posteriormente as amostras foram misturadas por inversão e incubar no freezer -70° por 30 minutos. Após a retirada do freezer, houve uma nova centrifugação a 10.000 rpm a 4° por 10 minutos e o sobrenadante foi descartado. Foi adicionado um volume de etanol e centrifugado por 5.000 rpm a 4° por 5 minutos. O sobrenadante foi descartado e as amostras foram secas na capela, e finalmente foram ressuspendidas em 10ul de água tratada com 0,1% DEPC e autoclavada.

Foi utilizado 1µg de RNA tratado para síntese de cDNA utilizando o kit GoScript™ Reverse (Promega), com volume final de 20ul, que foram diluídos em 100ul de água ultrapura.

A expressão do gene ROS1 foi analisada por PCR quantitativo, utilizando o protocolo de Syber Green, com volume final de 15ul em placas de 48 poços. Foram sintetizados os primers (Foward e Reverse) DNG101, DNG103 e ZmEF1a já publicados por (WILLIAMS et al., 2015). Os primers de Alfa-tubulina e Beta-tubulina foram utilizados como normalizadores de acordo com (LIN et al., 2014). Os dados gerados equipamento Step One System - Thermo Fisher foram analisados estatisticamente no programa PRISM.

4.3.3. Crescimento de plântulas de cana-de-açúcar *in vitro*

4.3.3.1. Material vegetal e tratamento com 5-Azac

As plântulas de cana-de-açúcar foram obtidas através micropropagação utilizando *explants* da cultivar SP803280 crescidas em casa de vegetação. Os *explants* passaram por assepsia, foram cultivados em meio de cultivo contendo MS (4.43g), sacarose (30g) e fitagel (2g) em tubos de ensaio e mantidos em sala escura a temperatura de 25°C por aproximadamente 45 dias. Após esse período, os calos embriogênicos foram replicados e transferidos para frascos com 40 ml de meio MS tratado com o inibidor de metilação 5-Azac na concentração final de 100µM e controle utilizando água.

4.3.3.2. Avaliação do desenvolvimento vegetal

Foram consideradas quatro réplicas para as condições de controle e tratado, com *pool* de três plantas em cada réplica. As plantas com 21 dias de tratamento foram retiradas dos frascos para obtenção de imagens para avaliações de diferenças fenotípicas entre plantas crescidas em condições controle com H_2O e tratamento com 5-Azac. Com um bisturi, parte aérea e raiz de cada planta foram separadas e analisadas separadamente. Parte aérea de cada planta foi medida com auxílio de uma régua e depois foi feito o cálculo médio do tamanho de cada réplica. As raízes de cada réplica foram mensuradas através do deslocamento de volume, utilizando uma proveta com 10 ml de água. Parte aérea e raiz foram pesadas separadamente em uma balança de precisão para obtenção de massa fresca e, depois, foram incubadas em estufa de aquecimento, mantidas por 24 horas à 60° e pesadas novamente para obtenção de massa seca. Com os valores de massa fresca e massa seca, foi feito o cálculo para obtenção dos valores de biomassa através da fórmula (MS/MF.100) e dos valores de conteúdo relativo de água através da fórmula (MF-MS/MF).100. Os valores de tamanho, massa fresca e massa seca, biomassa e conteúdo relativo de água, foram submetidos ao programa PRISM para avaliações estatísticas.

4. Resultados

4.1. Anotação, caracterização e filogenia dos genes envolvidos na regulação epigenética

A anotação dos genes de metilação e desmetilação, bem como os genes envolvidos na via de pequenos RNAs, foi realizada no banco de dados Phytozome, utilizando sequências de Arabidopsis como referência. Como cana-de-açúcar ainda possui genoma disponível no Phytozome, foi realizado alinhamento com proteínas de *S. bicolor* para a identificação dos genes em *S. spontaneum*, por se tratar de uma planta filogeneticamente mais próxima.

Além de Arabidopsis e as outras doze plantas monocotiledôneas, optamos por utilizar também sequências de *A. trichopoda*, por ser uma gimnosperma basal, e atuar como grupo irmão nas análises filogenéticas. No total, 451 genes foram anotados em quatro grupos distintos, como mostra a representação na figura 8.

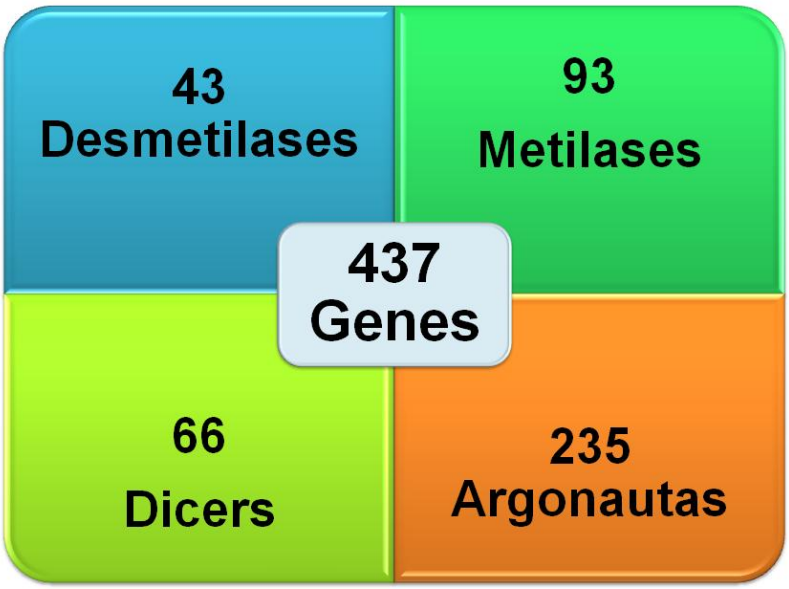


Figura 8: Apresentação do número de genes anotados em cada grupo e o total de genes anotados.

4.1.1. Anotação e filogenia

Para realizar a anotação dos genes de desmetilação, utilizamos sequências proteicas de quatro desmetilases de Arabidopsis, ROS1, DML2, DML3 e DME. Na anotação, apenas *B. distachyon e O. sativa* apresentaram o mesmo número de desmetilases que Arabidopsis. *Hordeum vulgare* apresentou apenas uma desmetilase, enquanto *A. trichopoda, Panicum halli, Musa acuminata e S. spontaneum* apresentaram duas. Já *Z. mays, S. bicolor, Setaria italica e Setaria viridis* apresentaram três desmetilases e por fim *Triticum aestivum* e *Miscanthus sinensis* apresentaram sete desmetilases. Demais informações sobre o tamanho do gene e do transcrito, além do número de éxon de cada gene e os valores referentes ao alinhamento, podem ser consultados na tabela 3.

A anotação revelou que as proteínas de desmetilação possuem grande variação com relação ao tamanho, podendo possuir até mais de dois mil aminoácidos, como é o caso de uma desmetilase de *M. acuminata* ou ser bem pequena, com pouco mais de 700 aminoácidos, como uma desmetilase de *M. sinensis* (tabela 3). Da mesma forma, os genes também apresentaram tamanhos variados, mas a maioria possui entre 10.000 e 15.000 nucleotídeos. O maior gene anotado foi o gene GRMZM2G422464_T02, que apresentou 46.430 nucleotídeos, já o menor gene anotado foi Sspon.07G0024030-1B que apresentou apenas 2.601 nucleotídeos. A tabela 3 também mostra que as desmetilases apresentaram de 15 a 20 éxons. As desmetilases identificadas em *S. spontaneum,* obtiveram boa similaridade com proteínas *de S. bicolor,* principalmente com relação ao tamanho das proteínas. Porém, os genes de *S. bicolor* são maiores que os de *S. spontaneum.*

A filogenia mostrada na figura 9 revela que em monocotiledôneas, as desmetilases se dividem em dois grupos principais, diferentemente de Arabidopsis. Além disso, proteínas de *A. trichopoda*, são mais próximas de Arabidopsis do que de monocotiledôneas.

Tabela 3: Anotação de genes de desmetilação em diferentes plantas monocotiledôneas.

Planta/ Abreviação	Gene	Id	Tamanho do transcrito	Tamanho da proteína	Tamanho do gene	Número de éxons	Tamanho da região alinhada	e-Value
Arabidopsis thaliana (Ath)	ROS1	AT2G36490.1	4713	1393	6787	20	1-1393	0.0
	DML2	AT3G10010.1	4184	1332	6382	20	1-1332	0.0
	DML3	AT4G34060.1	3308	1044	5424	20	1-1044	0.0
	DME	AT5G04560.2	6475	1987	9200	19	1-1.987	0.0
Sorghum bicolor (Sbi)	DMLa	Sobic.008G085300.1	6438	1878	12825	17	1186-1724 / 649 - 798	0.0
	DMLb	Sobic.004G149800.1	6585	1914	15972	17	857-1385 / 461-627	0.0
	DMLc	Sobic.009G155900.1	6220	1789	10862	19	857-138 / 530-640	0.0
Amborela trichopoda (Atr)	DMLa	scaffold00040.230	6399	2132	15493	18	857-1385 / 487-627	0.0
	DMLb	scaffold00088.93	5931	1976	11995	18	857-1385 / 474-661	0.0
Brachypodium distachyon (Bdi)	DMLa	Bradi3g43692.1	6913	1946	17681	17	1186-1728 / 646-799	0.0
	DMLb	Bradi4g08870.1	6757	1974	15116	15	856-1385/474-627	0.0
	DMLc	Bradi2g23797.4	6250	1807	10227	20	498-1028/347-445	7.1E-149
	DMLd	Bradi4g16620.6	5985	1744	15458	19	1192-1724 / 673-797	0.0

Hordeum vulgare (Hvu)	DML	HORVU1Hr1G073540.1	6241	1747	17532	15	1192-1724 / 673-797	0.0
Panicum hallii (Pha)	DMLa	Pahal.D01680.1	7170	1963	13741	17	1186-1728 / 672-798	0.0
	DMLb	Pahal.C02800.1	7295	1708	17407	19	1179-1695 / 673-789	0.0
Setaria itálica (Sit)	DMLa	Seita.1G164200.1	6792	1954	15750	17	857-1393/457-627	0.0
	DMLb	Seita.2G152900.1	6767	2082	14290	18	857-1385/472-627	0.0
	DMLc	Seita.3G223900.1	4992	1663	9801	20	857-1382/530-658	0.0
Setaria viridis (Svi)	DMLa	Sevir.1G166200.1	6657	2033	14125	18	1186-1728 / 649-799	0.0
	DMLb	Sevir.2G158400.3	6792	2082	14316	18	857-1385 / 472-627	0.0
	DMLc	Sevir.3G228900.1	5982	1696	11123	19	857-1382/ 530-658	0.0
Zea mays (Zma)	DMLa	GRMZM2G131756_T01	6183	1904	17972	17	1186-1728 / 649-799	0.0
	DMLb	GRMZM2G422464_T02	5721	1906	46430	14	1186-1728 / 641-798	0.0
	DMLc	GRMZM2G123587_T01	3270	909	10957	16	1192-1721	0.0
Oryza sativa (Osa)	DMLa	LOC_Os01g11900.1	6474	1953	11758	17	790-1318/446-597	8.8E-180
	DMLb	LOC_Os05g37410.1	5490	1829	10616	21	789-1330/481-618	8.7E-171
	DMLc	LOC_Os05g37350.1	5544	1847	12675	20	789-1329/481-618	1.6E-159
	DMLd	LOC_Os02g29230.1	4911	1636	8100	15	1219-1628/742-885/3-139	2,00E-135
Triticum aestivum (Tae)	DMLa	Traes_1DL_F0F402904.2	3339	969	9283	18	857-1385/530-641	0.0
	DMLb	Traes_1BL_05EB7AD97.1	4771	1510	7961	17	988-1505/568-733/178-226	0.0

	DMLc	Traes_1BL_263DE6AA9.1	5917	1769	10060	20	1192-1724 / 673-797	0.0
	DMLd	Traes_1AL_5F65C4004.1	5859	1765	10844	19	1192-1724 / 673-797	0.0
	DMLe	Traes_5BL_8F11FD1A0.2	6643	1944	13665	15	1189-1696 / 649-798	0.0
	DMLf	Traes_5DL_85C8862E8.2	6952	1981	14088	17	1189-1728 / 649-798	0.0
	DMLg	Traes_5AL_37D1F4150.1	6923	1946	11709	15	857-1362/474-627	0.0
Musa acuminata (Mac)	DMLa	GSMUA_Achr10T15480_001	3465	1154	9029	22	1169-1728 / 649-824	0.0
	DMLb	GSMUA_Achr3T22820_001	7022	2222	19650	34	1062-1722 / 672-908	0.0
Miscanthus sinensis (Msi)	DMLa	Misin08G138700.1	6705	1916	15714	17	1186-1728 / 649-798	0.0
	DMLb	Misin14G079500.1	6212	1886	12704	17	1186-1728 / 560-798	0.0
	DMLc	Misin07G374100.1	3972	1323	11075	16	462-1037	6.7E-150
	DMLd	Misin16G151400.1	5046	1326	6754	6	584-709/1103-1185	4,00E-28
	DMLe	Misin17G159200.1	5544	1496	13256	15	1121-1307/ 584-709	1.9E-70
	DMLf	Misin15G097800.1	2220	739	10409	17	1186-1728	0.0
	DMLg	Misin15G097700.1	3154	933	18291	17	1078-1629/647-709	6.6E-103
Saccharum spontaneum (Spp)	DMLa	Sspon.04G0026000-1P	ND	1702	7666	ND	9-1697	0.0
	DMLb	Sspon.07G0024030-1B	ND	459	2601	ND	1-455	0.0

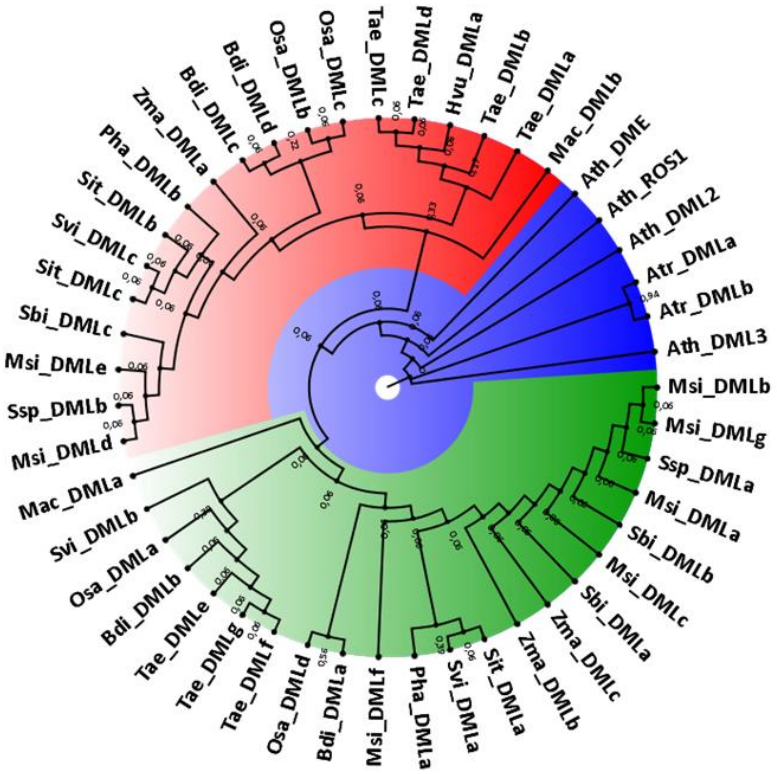


Figura 9: Filogenia de proteínas desmetilases de plantas monocotiledôneas. Em azul, representação das desmetilases de Arabidopsis e *A. trichopoda*. Em verde e vermelho, representação da divisão de subfamílias de desmetilases em monocotiledôneas.

Em relação aos genes de metilação do DNA, MET1, CMT3 e DRM2, foram anotados 93 genes. *A. trichopoda* apresentou quatro genes de metilação do DNA, enquanto *Z. mays, H. vulgare, P. hallii, O. sativa* e *S. spontaneum* apresentaram cinco. *S. bicolor* e *S. viridis* apresentaram seis metilases, *S. Itálica* e *M. acuminata* apresentaram sete, enquanto *B. distachyon* apresentou nove. O maior número foi de *T. aestivum*, que apresentou 21 genes de metilação do DNA. Com isso, observamos que o gene DRM2 é conservado em monocotiledôneas, com exceção de *T. aestivum*, que apresentou várias cópias. Já os genes MET1 e CMT3 aparecem pelo menos duplicados em todas as monocotiledôneas.

A anotação revelou que o grupo das monocotiledôneas possui maior número de metilases do tipo MET1 e CMT3, que são envolvidas na manutenção da metilação, do que DRM2 que está envolvida na metilação *de novo* (Tabela 4). Em todas as plantas, foi observado que o tamanho das proteínas de MET1 e CMT3 é superior ao

tamanho de DRM2. Além disso, a tabela 5 também revela que os genes de metilação são maiores em monocotiledôneas do que em Arabidopsis e que os genes de CMT3 são maiores, possuindo entre sete mil e nove mil nucleotídeos e em média 20 éxons. A grande exceção é o gene LOC_Os03g12570.1 de *O. sativa* que possui mais de 111 mil nucleotídeos e 30 éxons. As MET1 possuem tamanhos parecidos com CMT3, porém apresentam em média 10 éxons. Já as DRM2 são menores, com média de cinco mil nucleotídeos e sete éxons. A identificação das metilases em *S. spontaneum*, mostrou que as proteínas MET1 são muito parecidas com as de *S. bicolor*, tanto no tamanho dos genes quanto o tamanho das proteínas.

Tabela 4: Anotação de genes de metilação do DNA em diferentes plantas monocotiledôneas.

Planta/ Abreviação	Gene	Id	Tamanho do transcrito	Tamanho da proteína	Tamanho do gene	Número de éxons	Tamanho da região alinhada	e-Valle
Arabidopsis thaliana (Ath)	MET1	AT5G49160.1	4045	1534	6140	11	1 - 1534	0.0
	CMT3	AT1G69770.1	2764	839	5268	21	1- 839	0.0
	DRM2	AT5G14620.1	2183	626	3452	10	1 - 626	0.0
Sorghum bicolor (Sbi)	MET1a	Sobic.002G056000.1	4959	1552	10551	11	37 - 1528	0.0
	MET1b	Sobic.001G055700.1	5308	1525	10959	11	37 - 1527	0.0
	CMT3a	Sobic.004G197400.1	3389	918	8846	20	46 - 833	0.0
	CMT3b	Sobic.006G214000.1	3281	913	9842	24	46 - 833	0.0
	CMT3c	Sobic.009G083900.2	4519	1277	18435	21	50 - 829	0.0
	DRM2	Sobic.001G535800.4	2346	604	7164	7	106 - 624	0.0
Amborela trichopoda (Atr)	MET1	scaffold00017.270	4701	1566	14801	10	37 - 1534	0.0
	CMT3a	scaffold00024.163	2655	884	20403	21	48 - 827	0.0
	CMT3b	scaffold00048.135	3552	1183	56578	19	48 - 832	0.0
	DRM2	scaffold00096.44	1509	502	1971	4	282 - 623 / 204 - 238	1.2E-164
Brachypodium distachyon (Bdi)	MET1a	Bradi1g05380.1	4593	1530	7047	11	37 - 1527	0.0
	MET1b	Bradi1g55287.2	5129	1537	7938	11	37 - 1528	0.0
	CMT3a	Bradi1g68985.2	4862	1400	9390	27	50 - 827	0.0
	CMT3b	Bradi3g21450.2	3149	894	7653	20	50 - 834	0.0
	CMT3c	Bradi3g39050.1	3037	897	6338	20	50 - 827	0.0
	CMT3d	Bradi1g66167.2	4181	1095	13113	21	50 - 831	0.0
	DRM2a	Bradi4g05680.2	1764	587	4000	7	115-625	1.2E-176
	DRM2b	Bradi1g77873.1	1830	609	4246	7	197-625	1.5E-155
	DRM2c	Bradi1g11420.2	1554	517	3959	10	107-623	1.1E-131
Hordeum	MET1a	HORVU7Hr1G11710	5069	1529	10140	14	37 - 1527	0.0

vulgare (Hvu)		0.1						
	MET1b	HORVU0Hr1G03904 0.9	3963	1320	5232	11	232 - 1528	0.0
	MET1c	HORVU5Hr1G10653 0.1	1978	328	3164	5	1204 - 1529	0.0
	CMT3	HORVU6Hr1G08925 0.20	2924	922	5624	21	50 - 827	0.0
	DRM2	HORVU0Hr1G03879 0.4	2352	445	7683	4	192-623	1.4E-144
Panicum hallii (Pha)	MET1a	Pahal.B00904.2	5576	1542	12232	12	37 - 1527	0.0
	MET1b	Pahal.I01406.1	4386	1461	5978	14	37 - 1528	0.0
	CMT3a	Pahal.I03682.1	4413	1470	10274	22	46 - 827	0.0
	CMT3b	Pahal.C00877.1	4445	1289	9599	21	50 - 829	0.0
	DRM2	Pahal.I00079.1	2401	602	4288	7	106-623	0.0
Setaria itálica (Sit)	MET1a	Seita.2G059200.1	5408	1542	11908	12	37 - 2528	0.0
	MET1b	Seita.9G054000.1	5221	1527	9791	12	38 - 1528	0.0
	MET1c	Seita.2G058100.1	1139	342	1502	5	1198 - 1528	0.0
	CMT3a	Seita.9G479200.1	3495	983	6800	21	47 - 828	0.0
	CMT3b	Seita.9G300300.1	3254	915	7130	20	46 - 827	0.0
	CMT3c	Seita.3G080000.2	4372	1295	9355	21	50 - 829	0.0
	DRM2	Seita.9G571500.1	2281	602	4872	7	107-623	0.0
Setaria viridis (Svi)	MET1a	Sevir.2G062300.1	6048	1542	11350	11	37 - 1528	0.0
	MET1b	Sevir.9G053500.1	5163	1527	9722	12	38 - 1528	0.0
	CMT3a	Sevir.9G482900.1	3317	983	6621	21	47 - 828	0.0
	CMT3b	Sevir.9G306200.1	3272	915	7134	20	46 - 827	0.0
	CMT3c	Sevir.3G081800.2	4011	1295	8994	21	50 - 829	0.0
	DRM2	Sevir.9G574800.1	2319	602	4884	7	107-623	0.0
Zea mays (Zma)	MET1a	GRMZM2G333916_T01	5028	1536	6398	11	37 - 1524	0.0
	MET1b	GRMZM2G334041_T01	5028	1536	6012	11	37 - 1524	0.0

	CMT3a	GRMZM2G025592_T01	3809	915	7755	20	46 - 827	0.0
	CMT3b	GRMZM2G005310_T02	3152	915	9516	19	46 - 833	0.0
	DRM2	GRMZM2G092497_T03	2922	602	24414	7	106-623	0.0
Oryza sativa (Osa)	MET1a	LOC_Os07g08500.1	1605	377	2298	6	1153 - 1528	0.0
	MET1b	LOC_Os03g58400.1	1384	376	2245	6	1153 - 1528	0.0
	CMT3a	LOC_Os03g12570.1	5862	1787	111118	30	50 - 829	0.0
	CMT3b	LOC_Os05g13780.1	4224	1407	13681	22	50 - 829	0.0
	DRM2	LOC_Os03g02010.4	2275	507	3700	4	140-623	0.0
Triticum aestivum (Tae)	MET1a	Traes_2AS_9B3DEDDAB.2	5235	1531	9956	11	37 - 1528	0.0
	MET1b	Traes_2BS_173E86712.1	5322	1531	9626	11	37 - 1528	0.0
	MET1c	Traes_2DS_D3601814B.2	4596	1446	5636	11	63 - 1528	0.0
	MET1d	Traes_7DL_E374EE2DF.1	4590	1455	6179	11	69 - 1527	0.0
	MET1e	Traes_7BL_BCE8CA9A6.1	4935	1455	6416	11	58 - 1527	0.0
	MET1f	Traes_7AL_5E4FB486C.2	4587	1455	6141	11	58 - 1524	0.0
	MET1g	Traes_5BL_9774FCB1A.1	4428	1475	5484	10	38 - 1485	0.0
	CMT3a	Traes_4AS_0194CD713.1	3101	930	6089	21	50 - 828	0.0
	CMT3b	Traes_4DL_A74B9AF77.1	2847	846	5840	21	50 - 828	0.0
	CMT3c	Traes_6DL_7626A87FF.1	2577	733	5007	19	93 - 833	0.0
	CMT3d	Traes_6DS_93F72DE87.1	2423	750	5691	19	59 - 827	0.0
	CMT3e	Traes_7DL_20BFD8	2034	678	4903	15	87 - 779	0.0

		A19.1						
	CMT3f	Traes_6AS_A52CF52D3.2	2617	705	5785	17	122 - 827	0.0
	CMT3g	Traes_6BL_64DEE7723.2	1998	635	3780	16	209 - 833	0.0
	CMT3h	Traes_6AL_9E0046317.1	2366	627	4292	15	209 - 833	0.0
	DRM2a	Traes_5BL_169D5FB79.3	1785	532	3868	8	121-623	2.2E-178
	DRM2b	Traes_5DL_3A965D7A6.2	2307	601	5623	7	115-623	7.3E-177
	DRM2c	Traes_4BL_E9B610B59.1	1812	575	3840	7	102-625	1.1E-175
	DRM2d	Traes_5AL_92DEAA86D.2	1812	575	3835	7	102-625	1.3E-175
	DRM2e	Traes_4DL_B25F54064.1	1812	575	3801	7	102-625	6,00E-172
	DRM2f	Traes_5AL_BA3074F36.1	2043	618	5376	8	115-623	1.9E-177
Musa acuminata (Mac)	MET1a	GSMUA7T00440_001	4275	1424	8525	13	311-1528 / 37-151 / 80-252	0.0
	MET1b	GSMUA4T22500_001	4059	1352	6724	13	302-1529 / 37-163 / 80-244	0.0
	CMT3a	GSMUA2T21720_001	2640	879	6335	21	50 - 827	0.0
	CMT3b	GSMUA3T27650_001	2547	848	19777	22	50 - 830	0.0
	DRM2a	GSMUA8T33680_001	1530	509	5032	8	115-623	7.2E-152
	DRM2b	GSMUA3T23940_001	1806	601	10481	12	96-602	5.1E-149
	DRM2c	GSMUA1T25520_001	1386	461	2292	7	121-623	6.6E-115
Miscanthus sinensis (Msi)	MET1a	Misin03G049000.1	5589	1543	13928	12	37 - 1528	0.0
	MET1b	Misin04G057200.1	5588	1545	15872	12	37 - 1528	0.0

	MET1c	Misin01G048100.3	4920	1491	6243	11	37 - 1528	0.0
	CMT3a	Misin01G259900.1	3353	915	7900	20	46 - 827	0.0
	CMT3b	Misin07G387900.1	3224	920	8600	20	46 - 833	0.0
	CMT3c	Misin02G252000.1	3095	911	9324	20	46 - 827	0.0
	CMT3d	Misin17G095300.2	4863	1421	12743	20	50 - 829	0.0
	DRM2	Misin02G507500.1	2265	607	4802	8	85-624	0.0
Saccharum spontaneum (Spp)	MET1	Sspon.01G0048600-2C	ND	1598	11831	ND	83-1598	0.0
		Sspon.02G0023350-1A	ND	1459	5376	ND	1-1459	0.0
	CMT3	Sspon.03G0029610-1B	ND	962	5971	ND	6-942	0.0
		Sspon.01G0015710-2B	ND	776	8959	ND	280-776	0.0
	DRM2	Sspon.03G0016990-1P	ND	610	4177	ND	1-604	0.0

Diferentemente das desmetilases, as proteínas de metilação apresentaram boa similaridade com as proteínas de Arabidopsis, como pode ser visto na filogenia representada na figura 10. Na figura é possível perceber as metilases se dividem em três grupos (MET1, CMT3 e DRM2) e que cada grupo possui boa similaridade entre si, mesmo sendo de plantas diferentes. Na figura 10 podemos observar também a aproximação das proteínas da subfamília Panicodeae*, Z. mays, S. bicolor* e *S. spontaneum.*

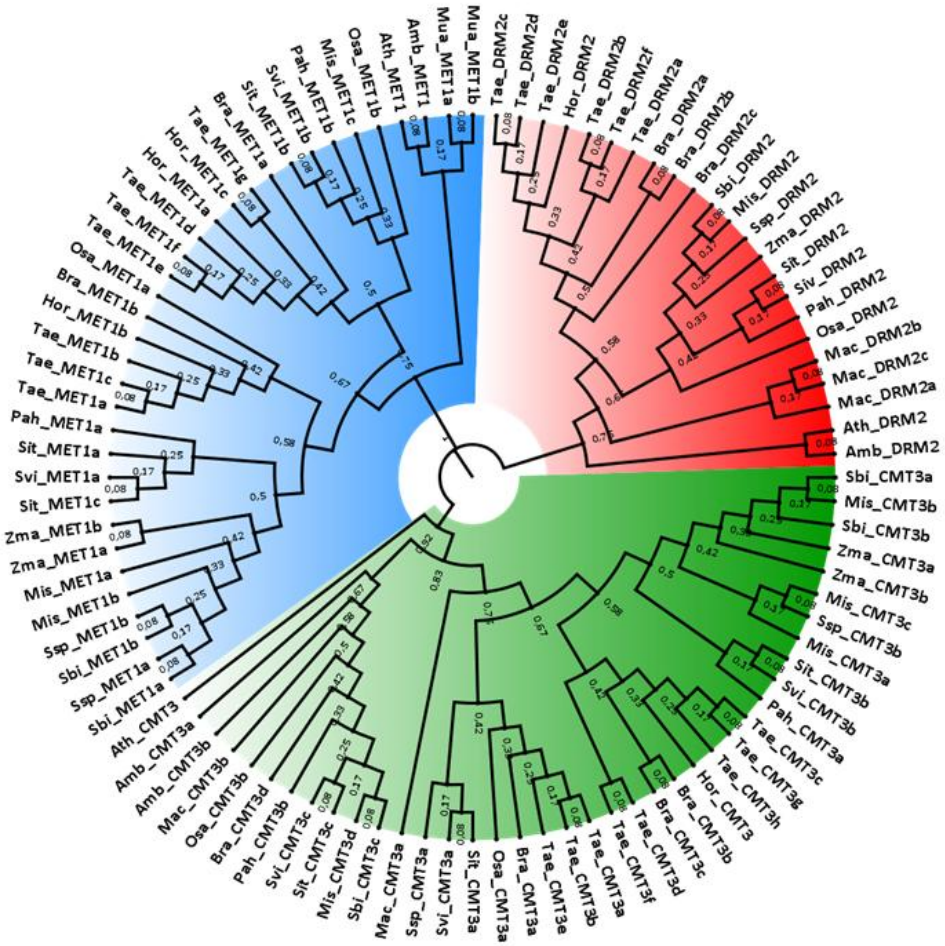


Figura 10: Filogenia de proteínas metilases em monocotiledôneas. Em azul, proteínas do tipo METILTRANSFERASE 1 (MET1); em vermelho, proteínas do tipo DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2) e em verde, proteínas do tipo CHROMOMETYLASE 3 (CMT3).

A anotação dos genes Dicer-like, foi realizada utilizando um gene representante de cada família, de 1 a 4 de Arabidopsis. Foram anotados 66 genes Dicer-like, incluindo as quatro de Arabidopsis. Em *A. trichopoda*, foram anotadas apenas duas Dicer-like e em *M. acuminata* apenas três. Em *Z. mays*, os alinhamentos mostraram que há um representante de cada família de Dicer-like, totalizando quatro genes anotados. *S. spontaneum* teve seis, *B. distachyon* teve sete e *T. aestivum* teve nove genes Dicer-like anotados, as demais plantas tiveram cinco genes Dicer-like anotados em cada uma. As outras informações da anotação estão disponíveis na tabela 5. A anotação realizada utilizando sequências de *S. bicolor* para identificar Dicer-like em *S. spontaneum*, revelou que cana-de-açúcar possui uma DCL1, duas DCL2, uma DCL3 e duas DCL4, indicando uma duplicação nas famílias DCL2 e DCL4. A identidade dessas proteínas com as de *S. bicolor* ficou entre entre 82% a 95%.

As Dicer-like são proteínas com tamanho bem conservado entre as plantas monocotiledôneas, havendo pouca variação entre 1400 a 1600 aminoácidos. Porém, há algumas exceções como, por exemplo, DCL1 de *H. vulgare*, que possui 1971 aminoácidos e DCL2b *de S. spontaneum* com 584 aminoácidos, são respectivamente a maior e a menor DCL anotada.

A filogenia das Dicer-like mostrada na figura 11 foi realizada utilizando o modelo evolutivo JTT + G +I +F, com 1000 de *bootstrap*. O resultado revelou que apesar de possuírem funções parecidas, cada família de Dicer-like de 1 a 4 se agrupa, mostrando grande conservação dentro do grupo das monocotiledôneas, repetindo o que é visto em Arabidopsis. A filogenia mostra também que as proteínas da família DCL3 tem maior semelhança com as proteínas da família DCL1 e do mesmo modo, a família DCL2 é filogeneticamente mais próxima da família de DCL4 (Figura 11).

Tabela 5: Anotação de genes Dicer-like de famílias de 1 a 4 em diferentes plantas monocotiledôneas

Planta/ Abreviação	Gene	Id	Tamanho do transcrito	Tamanho da proteína	Tamanho do gene	Número de éxons	Tamanho da região alinhada	e-Value
Arabidopsis thaliana (Ath)	DCL1	AT1G01040.2	5877	1910	7705	20	1-1910	0.0
	DCL2	AT3G03300.1	4824	1388	8289	22	12-1371	0.0
	DCL3	AT3G43920.1	4596	1531	7283	24	1-1580	0.0
	DCL4	AT5G20320.1	5660	1702	10049	25	1-1702	0.0
Sorghum bicolor (Sbi)	DCL1	Sobic.001G527800.1	6777	1929	11921	19	177-1909	0.0
	DCL2	Sobic.001G180000.2	4874	1386	16839	20	20-1383	0.0
	DCL3a	Sobic.003G396700.1	5670	1657	10392	25	47-1580	0.0
	DCL3b	Sobic.001G214000.1	5904	1637	11819	26	46-1580	0.0
	DCL4	Sobic.006G142775.1	5190	1665	25396	25	118-1701	0.0
Amborela trichopoda (Atr)	DCL2	scaffold00016.358	4275	1424	35781	23	20-1383	0.0
	DCL4	scaffold00026.16	4116	1371	66380	18	396-1701	0.0
Brachypodium distachyon (Bdi)	DCL1	Bradi1g77087.1	6602	1930	12125	19	185-1909	0.0
	DCL2a	Bradi1g15440.1	4800	1402	23374	20	12-1383	0.0
	DCL2b	Bradi1g21030.3	4032	1343	12934	20	27-1386	0.0
	DCL3a	Bradi2g23187.3	5084	1491	11446	26	639-1580	0.0
	DCL3b	Bradi2g58270.4	2855	882	4116	11	706-1580	0.0
	DCL3c	Bradi3g29287.1	5460	1621	12107	26	47-1577	0.0
	DCL4	Bradi5g15337.1	5559	1625	22090	24	106-1701	0.0
Hordeum vulgare (Hvu)	DCL1	HORVU4Hr1G084890.1	8341	1971	12070	18	185-1909	0.0
	DCL2	HORVU5Hr1G019300.9	4702	1473	42949	20	20-1366	0.0
	DCL3a	HORVU3Hr1G092430.9	5342	1685	9507	26	47-1580	0.0
	DCL3b	HORVU1Hr1G042710.5	5333	1683	18509	26	47-1577	0.0
	DCL4	HORVU2Hr1G088270.2	5463	1612	49263	25	107-1701	0.0
Panicum hallii (Pha)	DCL1	Pahal.I00170.1	6559	1928	11531	19	149-1909	0.0
	DCL2	Pahal.I02663.1	4635	1394	18980	20	20-1383	0.0

	DCL3a	Pahal.E00491.1	4491	1496	7446	21	46-1513	0.0
	DCL3b	Pahal.J01444.1	5580	1641	12466	26	46-1577	0.0
	DCL4	Pahal.G01022.1	4511	1570	23332	24	158-1701	0.0
Setaria itálica (Sit)	DCL1	Seita.9G562200.1	6506	1933	12318	19	170-1909	0.0
	DCL2	Seita.9G179200.1	5034	1544	14392	21	23-1383	0.0
	DCL3a	Seita.5G422300.1	4938	1645	8618	27	47-1580	0.0
	DCL3b	Seita.9G215300.1	5046	1567	11111	25	90-1577	0.0
	DCL4	Seita.7G162200.1	5377	1632	16606	25	118-1701	0.0
Setaria viridis (Svi)	DCL1	Sevir.9G565900.1	6668	1933	12326	19	170-1909	0.0
	DCL2	Sevir.9G177900.1	5114	1580	14382	20	23-1383	0.0
	DCL3a	Sevir.5G427800.1	5446	1593	11054	25	47-1580	0.0
	DCL3b	Sevir.9G214200.1	4997	1593	11054	26	639-1577 / 90-579	0.0
	DCL4	Sevir.7G171000.1	5486	1632	16691	25	118-1701	0.0
Zea mays (Zma)	DCL1	GRMZM2G040762_T01	4615	1307	7323	15	613-1909	0.0
	DCL2	GRMZM2G301405_T01	3061	707	11215	8	608-1294	5.3E-172
	DCL3	GRMZM2G413853_T01	4715	1517	17307	24	46-1513	0.0
	DCL4	GRMZM2G024466_T01	1727	511	2484	7	400-906	6.7E-123
Oryza sativa (Osa)	DCL1	LOC_Os03g02970.1	6583	1883	9765	18	185-1909	0.0
	DCL2	LOC_Os03g38740.1	5009	1410	14894	20	23-1382	0.0
	DCL3a	LOC_Os01g68120.1	5199	1597	10037	24	190-1580/ 47-134	0.0
	DCL3b	LOC_Os10g34430.1	5198	1571	12108	25	639-1579/ 83-545	0.0
	DCL4	LOC_Os04g43050.1	4935	1591	14911	24	118-1652	0.0
Triticum aestivum (Tae)	DCL1a	Traes_4BL_B3A1B8342.2	5976	1779	10416	19	225-1909	0.0
	DCL1b	Traes_5AL_72A7552B9.3	6338	1614	10540	18	201-1909	0.0
	DCL3a	Traes_3AL_562D6614F.1	4761	1586	8968	24	639-1580/63-579	0.0
	DCL3b	Traes_3DL_2DC78B18A.1	4761	1586	8347	21	639-1580 / 63-579	0.0

	DCL3c	Traes_1AL_E7144546E.1	5165	1627	12181	26	47-1577	0.0
	DCL3d	Traes_1DL_C646B6990.1	4958	1494	11897	21	145-1577	0.0
	DCL4a	Traes_2DL_E96DCDCB4.2	4641	1392	14743	19	382-1701	0.0
	DCL4b	Tres_2BL_393ED78A5.2	2970	990	6968	11	548-1529	0.0
	DCL4c	Traes_2AL_54C78A86E.1	2970	990	7936	11	548-1529	0.0
Musa acuminata (Mac)	DCL1	GSMUA8T12350_001	5694	1818	14058	24	15-1124	0.0
	DCL3	GSMUA5T10880_001	4976	1598	15156	29	639-1580 / 58-584	0.0
	DCL4	GSMUA9T22460_001	5124	1363	14559	18	454-1699	0.0
Miscanthus sinensis (Msi)	DCL1	Misin02G493200.1	7352	1925	13502	19	149-1909	0.0
	DCL2	Misin01G162600.1	4593	1118	18358	16	295-1383	0.0
	DCL3a	MisinT163500.1	5668	1658	12033	25	47-1580	0.0
	DCL3b	Misin02G194800.1	5416	1643	10938	26	46-1579	0.0
	DCL4	Misin12G144700.1	5419	1627	24402	25	118-1701	0.0
Saccharum spontaneum (Spp)	DCL1a	Sspon.01G0001230-1A	ND	1590	7361	ND	13-1590	0.0
	DCL2a	Sspon.01G0055150-1C	ND	854	12250	ND	7-646	0.0
	DCL2b	Sspon.01G0022370-1A	ND	1233	16217	ND	1-583	0.0
	DCL3	Sspon.03G0000660-1A	ND	954	7218	ND	1-954	0.0
	DCL4a	Sspon.05G0024630-1T	ND	627	8751	ND	1-627	0.0
	DCL4b	Sspon.05G0024640-2C	ND	602	5544	ND	49-602	0.0

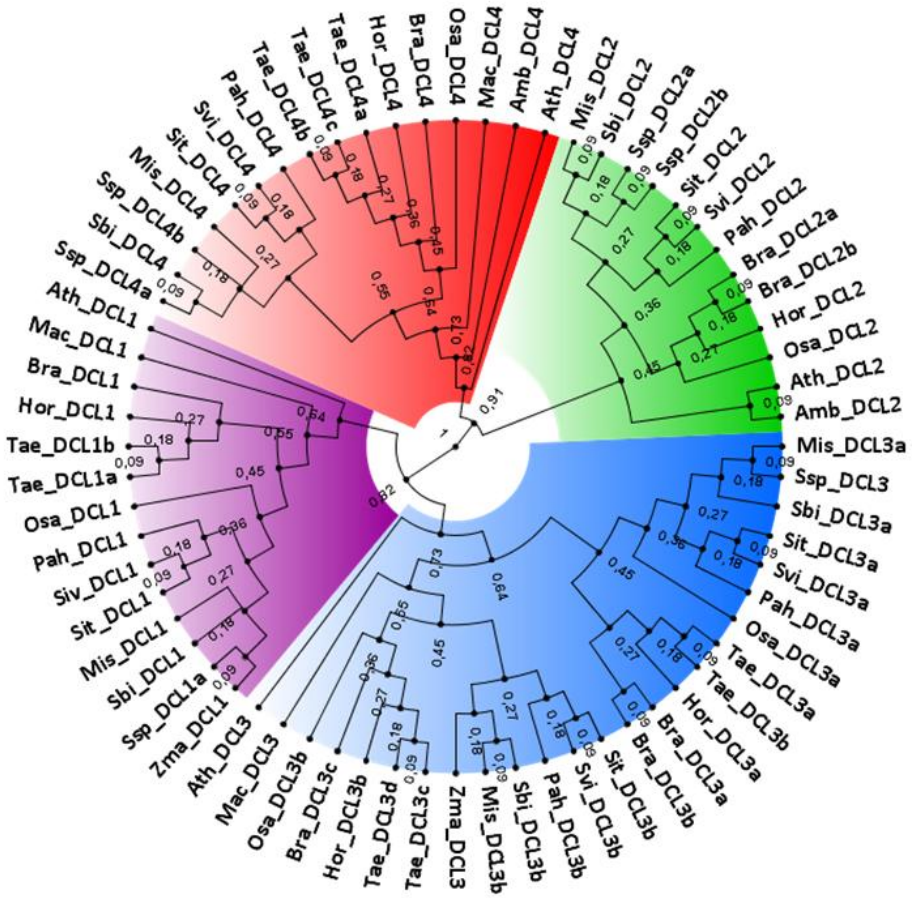


Figura 11: Filogenia de proteínas Dicer-like das familias de 1 a 4 em plantas monocotiledôneas. Em roxo, Dicer-like 1; em vermelho, Dicer-like 4; em verde, Dicer-like 2 e em azul, Dicer-like 3.

A anotação dos genes Argonauta foi a mais extensa, por se tratar de um grupo muito grande, com dez famílias descritas em Arabidopsis. Utilizamos um representante de cada uma dessas famílias para fazer a anotação nas demais plantas. No total, foram anotados 245 genes Argonauta (Tabela 9). Além do dez representantes de Arabidopsis, foram anotadas nove Argonauta em *A. trichopoda*, 13 em *H. vulgari* e em *M. acuminata*, 14 em *S. spontaneum,* 15 em *B. distachyon, S. itálica e S. viridis,* em *O sativa* e *P. hallii,* 19 em *S. bicolor*, 20 em *Z. mays,* 31 em *M. sinensis* e 39 em *T. aestivum*. Apesar de S. bicolor possuir 19 genes Argonauta, o genoma de *S. spontaneum* apresenta apenas 14, como mostra a tabela 10, onde mostramos as informações sobre a identificação de genes Argonauta em cana-de-açúcar. As proteínas Argonauta são de tamanho relativamente grande, principalmente as das

famílias 1 e 10, que quase sempre se apresentam com mais de 1000 aminoácidos (Tabela 6).

Tabela 6: Anotação de genes Argonautas de famílias de 1 a 10 em diferentes plantas monocotiledôneas.

Planta/ Abreviação	Gene	Id	Tamanho do transcrito	Tamanho da proteína	Tamanho do gene	Número de éxons	Tamanho da região alinhada	e-Value
Arabidopsis thaliana (Ath)	AGO1	AT1G48410.1	3501	1048	6489	21	115-1048	0.0
	AGO2	AT1G31280.1	3343	1014	3634	3	142-1014	0.0
	AGO3	AT1G31290.1	3585	1194	4025	3	242-1194	0.0
	AGO4	AT2G27040.2	3189	924	5824	22	1-924	0.0
	AGO5	AT2G27880.1	3176	997	5407	20	33-997	0.0
	AGO6	AT2G32940.1	2997	878	5480	22	1-878	0.0
	AGO7	AT1G69440.1	3161	990	3640	3	1-990	0.0
	AGO8	AT5G21030.1	2553	850	4381	21	1-850	0.0
	AGO9	AT5G21150.1	3018	896	6086	22	1-896	0.0
	AGO10	AT5G43810.1	3330	988	5826	19	175-1048	0.0
Sorghum bicolor (Sbi)	AGO1/10a	Sobic.006G181000.1	3812	1077	10505	22	115-1048	0.0
	AGO1/10b	Sobic.009G003700.1	3554	1109	6241	22	115-1048	0.0
	AGO1/10c	Sobic.004G357700.5	4285	1028	10519	24	115-1048	0.0
	AGO1/10d	Sobic.010G276600.1	3463	1048	6781	21	115-1048	0.0
	AGO1/10e	Sobic.010G180800.1	4456	981	13795	21	175-1048	0.0
	AGO2/3a	Sobic.006G215000.1	4329	1092	6336	3	142- 967	0.0
	AGO4/8/9a	Sobic.003G129500.3	4382	900	9345	22	48-924	0.0
	AGO4/8/9b	Sobic.009G259900.1	3203	909	7163	22	52-924	0.0

	AGO4/8/9c	Sobic.003G129200.1	2942	898	12671	22	50-924	0.0
	AGO5a	Sobic.001G136000.4	2918	884	6066	22	136-997	0.0
	AGO5b	Sobic.001G136150.1	2930	743	6084	20	266-997	0.0
	AGO5c	Sobic.002G061800.2	2834	881	7596	21	141-997	0.0
	AGO5d	Sobic.001G053700.2	2936	849	8765	21	161-997	0.0
	AGO5e	Sobic.002G299200.1	3675	1044	6777	21	144-997	0.0
	AGO5f	Sobic.001G153600.1	3106	935	5638	22	146-997	0.0
	AGO6a	Sobic.008G087200.1	3183	881	8503	22	33-878	0.0
	AGO7a	Sobic.001G331800.1	3300	1033	4284	3	182-1048	0.0
	AGOa	Sobic.002G299200.1	3675	1044	6777	21	179-1048	0.0
	AGOb	Sobic.001G153600.1	3106	935	5638	22	179-1048	0.0
Amborela trichopoda (Atr)	AGO1/10a	scaffold00044.24	3327	1108	10962	22	116-1048	0.0
	AGO1/10b	scaffold00002.566	2808	935	10186	22	175-1048	0.0
	AGO1/10c	scaffold00122.61	2907	968	9332	22	175-1048	0.0
	AGO2/3a	scaffold00032.76	3006	1001	4377	2	161-966	0.0
	AGO4/8/9	scaffold00008.121	2358	785	107665	21	177-924	0.0
	AGO5a	scaffold00024.271	3105	1034	11584	22	139-997	0.0
	AGO5b	scaffold00002.501	2865	954	8974	22	118-997	0.0
	AGO5c	scaffold00169.41	2919	972	15754	22	124-997	0.0
	AGO7a	scaffold00058.31	3042	1013	4418	3	144-990	0.0
Brachypodium distachyon (Bdi)	AGO1/10a	Bradi5g18540.1	4090	1094	8361	22	115-1048	0.0
	AGO1/10b	Bradi3g51077.1	4058	1060	8222	22	115-1048	0.0
	AGO1/10c	Bradi3g60697.2	3476	1043	7755	22	175-1048	0.0
	AGO1/10d	Bradi1g29577.1	3740	1044	8249	22	115-1048	0.0

	AGO1/10e	Bradi1g36907.2	3501	953	12086	22	175-1048	0.0
	AGO2/3a	Bradi5g21800.1	3123	1040	4440	5	142-1014	0.0
	AGO4/8/9a	Bradi2g10370.1	3262	924	8352	22	45-924	0.0
	AGO4/8/9b	Bradi2g14147.1	3389	914	6686	22	52-924	0.0
	AGO4/8/9c	Bradi2g10360.2	2259	621	4466	18	200-857	0.0
	AGO5a	Bradi1g12431.2	2595	864	5936	23	137-997	0.0
	AGO5b	Bradi1g54977.1	3401	1032	7075	20	130-997	0.0
	AGO5c	Bradi1g05162.2	3310	846	8397	21	161-997	0.0
	AGO6a	Bradi4g08587.1	3121	888	8959	22	28-878	0.0
	AGO7a	Bradi1g16060.2	3587	1038	5165	3	150-990	0.0
	AGO	Bradi1g28260.3	3243	1001	6748	23	146-997	0.0
Hordeum vulgare (Hvu)	AGO1/10a	HORVU7Hr1G007000.21	3261	1086	6390	9	115-1048	0.0
	AGO1/10b	HORVU7Hr1G120600.1	3530	963	8895	22	178-1048	0.0
	AGO1/10c	HORVU6Hr1G036840.2	3675	949	12518	21	175-1048	0.0
	AGO2/3a	HORVU2Hr1G098650.1	3688	1047	6253	3	142-966	0.0
	AGO4/8/9a	HORVU3Hr1G038830.1	3388	922	7793	21	52-924	0.0
	AGO4/8/9b	HORVU1Hr1G095300.3	2850	886	7444	20	57-887	0.0
	AGO4/8/9c	HORVU7Hr1G107770.4	2194	510	3862	15	419-924	0.0
	AGO5a	HORVU4Hr1G004030.1	3255	1033	9003	17	136-997	0.0
	AGO5b	HORVU4Hr1G004010.8	2297	765	5118	17	136-873	0.0
	AGO5c	HORVU2Hr1G031130.	2834	843	6532	22	155-997	0.0

		2						
	AGO6a	HORVU0Hr1G012490.1	3179	883	10921	22	28-878	0.0
	AGO7a	HORVU0Hr1G005350.1	3326	1012	4217	3	150-990	0.0
	AGO	HORVU3Hr1G021290.1	3378	1051	6654	21	146-997	0.0
Panicum hallii (Pha)	AGO1/10a	Pahal.G01956.1	3800	1097	14818	22	115-1048	0.0
	AGO1/10b	Pahal.C00991.1	3809	1110	7549	22	115-1048	0.0
	AGO1/10c	Pahal.A04013.1	5604	1034	9566	22	115-1048	0.0
	AGO1/10d	Pahal.D00250.1	4748	1050	10026	21	115-1048	0.0
	AGO1/10e	Pahal.D00982.1	3904	963	14378	22	175-1048	0.0
	AGO2/3a	Pahal.G02287.1	3586	1012	5325	3	142-1014	0.0
	AGO4/8/9a	Pahal.E03348.1	5602	902	8818	22	52-924	0.0
	AGO4/8/9b	Pahal.C01666.1	3382	906	6813	22	57-924	0.0
	AGO4/8/9c	Pahal.E03353.1	2013	671	3610	16	54-924	0.0
	AGO5a	Pahal.I02233.1	3520	1080	6924	22	117-997	0.0
	AGO5b	Pahal.B00968.1	3042	1013	8955	22	141-997	0.0
	AGO5c	Pahal.I01423.1	3562	1058	9126	22	137-997	0.0
	AGO5d	Pahal.I01378.1	4049	888	9774	35	137-902	0.0
	AGO5e	Pahal.I01385.1	4451	850	11578	32	137-902	0.0
	AGO7a	Pahal.B00488.1	3797	1151	4897	3	142-990	0.0
	AGO	Pahal.B03884.1	3680	1048	6915	21	131-1048	0.0
Setaria itálica (Sit)	AGO1/10a	Seita.7G201100.1	3800	1104	13863	22	158-990	0.0
	AGO1/10b	Seita.3G070600.1	3823	1102	8888	22	158-990	0.0
	AGO1/10c	Seita.1G378700.1	5489	1023	9974	21	158-990	0.0

	AGO1/10d	Seita.4G288700.1	3126	1041	6534	22	158-990	0.0
	AGO1/10e	Seita.4G225900.1	4181	965	7800	22	147-990	0.0
	AGO2/3a	Seita.7G236800.1	3376	1024	6129	3	142-1014	0.0
	AGO4/8/9a	Seita.5G043300.1	3317	902	7799	22	52-924	0.0
	AGO4/8/9b	Seita.3G117800.1	3238	910	6726	22	52-924	0.0
	AGO5a	Seita.9G138800.1	3447	1148	7497	23	123-997	0.0
	AGO5b	Seita.2G066500.1	3483	1041	7569	21	141-997	0.0
	AGO5c	Seita.9G052000.1	3519	1060	9352	22	137-997	0.0
	AGO6a	Seita.5G185000.1	2807	862	7183	23	28-878	0.0
	AGO6b	Seita.2G148000.1	3751	882	10797	22	28-878	0.0
	AGO7a	Seita.9G359200.1	3637	1030	4892	3	132-990	0.0
	AGO	Seita.2G310800.1	3120	954	5696	21	146-997	0.0
Setaria viridis (Svi)	AGO1/10a	Sevir.7G213000.1	3611	1092	6701	22	158-990	0.0
	AGO1/10b	Sevir.3G072400.1	3926	1102	8989	22	158-990	0.0
	AGO1/10c	Sevir.1G385400.1	3970	1023	8357	22	158-990	0.0
	AGO1/10d	Sevir.4G301000.1	3581	1050	6678	21	158-990	0.0
	AGO1/10e	Sevir.4G237200.1	4335	965	7954	22	147-990	0.0
	AGO2/3a	Sevir.7G248200.1	3500	1024	5280	3	142-1014	0.0
	AGO4/8/9a	Sevir.5G040900.1	3447	902	7931	22	52-924	0.0
	AGO4/8/9b	Sevir.3G120100.1	3717	910	7288	22	52-924	0.0
	AGO5a	Sevir.9G137300.1	3651	1160	7245	23	123-997	0.0
	AGO5b	Sevir.2G069400.1	2983	878	7008	21	141-997	0.0
	AGO6a	Sevir.5G186400.1	2595	852	7104	23	28-878	0.0
	AGO6b	Sevir.2G154500.1	3479	882	10520	22	28-878	0.0

	AGO7a	Sevir.9G365300.1	3565	1030	4343	3	132-990	0.0
	AGOa	Sevir.2G321800.1	2943	957	5075	21	144-928	0.0
	AGOb	Sevir.2G321900.1	3253	954	5788	22	146-997	0.0
Zea mays (Zma)	AGO1/10a	AC209206.3_FGT011	3279	1092	6539	22	115-1048	0.0
	AGO1/10b	GRMZM2G039455_T01	6337	1078	9856	22	115-1048	0.0
	AGO1/10c	GRMZM2G441583_T01	5178	1102	8451	22	115-1048	0.0
	AGO1/10d	GRMZM2G162525_T01	3635	554	5406	15	494-1048	0.0
	AGO1/10e	GRMZM2G120019_T01	2207	356	3092	12	629-982	0.0
	AGO1/10f	GRMZM2G361518_T01	3336	1027	5729	21	178-1048	0.0
	AGO1/10f	GRMZM2G079080_T02	3326	966	6819	22	175-1048	0.0
	AGO1/10g	AC189879.3_FGT003	3279	1092	6539	22	175-1048	0.0
	AGO2/3a	GRMZM2G354867_T01	3120	1039	5380	4	142-967	0.0
	AGO2/3b	GRMZM2G007791_T01	3495	1033	7525	3	142-1014	0.0
	AGO4/8/9a	GRMZM2G089743_T02	1777	471	11325	15	453-924	0.0
	AGO4/8/9b	GRMZM2G589579_T01	4051	898	7346	22	48-924	0.0
	AGO4/8/9c	GRMZM2G141818_T03	5508	910	7990	22	52-924	0.0
	AGO5a	GRMZM2G461936_T01	2030	819	4038	11	137-739	0.0
	AGO5b	GRMZM2G059033_T01	3162	1013	6924	21	138-997	0.0
	AGO5c	GRMZM2G123063_T01	2826	941	8252	19	139-997	0.0

	AGO6a	GRMZM2G347402_T01	5160	635	9005	15	33-641	0.0
	AGO7a	GRMZM5G892991_T01	3388	1032	4346	3	147-990	0.0
	AGOa	GRMZM2G105250_T01	3408	1053	5558	21	147-997	0.0
	AGOb	GRMZM2G457370_T01	3205	919	5549	22	146-997	0.0
Oryza sativa (Osa)	AGO1/10a	LOC_Os04g47870.1	3018	1118	15105	22	115-1048	0.0
	AGO1/10b	LOC_Os02g45070.1	3883	1082	11693	22	115-1048	0.0
	AGO1/10c	LOC_Os02g07310.1	3151	876	7130	21	178-1048	0.0
	AGO1/10d	LOC_Os02g58490.1	3647	1011	7275	22	142-1048	0.0
	AGO1/10e	LOC_Os06g51310.1	3759	1038	8406	22	115-1048	0.0
	AGO1/10f	LOC_Os06g39640.1	3193	973	6418	22	175-1048	0.0
	AGO2/3a	LOC_Os04g52540.1	3478	1034	4249	3	149-1014	0.0
	AGO4/8/9a	LOC_Os01g16870.1	3128	904	7606	22	52-924	0.0
	AGO4/8/9b	LOC_Os04g06770.1	3374	911	7698	22	52-924	0.0
	AGO5a	LOC_Os03g47830.1	2679	892	12542	21	136-997	0.0
	AGO5b	LOC_Os03g47820.1	3206	1058	7977	20	143-997	0.0
	AGO5c	LOC_Os07g09020.1	3528	1052	9655	22	141-997	0.0
	AGO5d	LOC_Os03g58600.1	3452	1058	8963	22	137-997	0.0
	AGO5e	LOC_Os03g57560.1	4028	633	9001	23	320-953	0.0
	AGO7a	LOC_Os03g33650.1	3685	1048	4605	3	149-990	0.0
	AGO	LOC_Os07g28850.1	3538	1088	6576	21	135-997	0.0
Triticum aestivum (Tae)	AGO1/10a	Traes_2AL_2512A7F91.1	3408	1117	6344	22	115-1048	0.0
	AGO1/10b	Traes_2BL_93099ACF4.1	2235	645	4863	17	378-1048	0.0

AGO1/10c	Traes_4AL_A118C6C84.1	3247	1077	6321	22	115-1045	0.0
AGO1/10d	Traes_7AS_56569A5AC.1	3258	1080	3243	22	115-1048	0.0
AGO1/10e	Traes_7DS_4D01B6175.1	3188	868	5859	22	186-1048	0.0
AGO1/10f	Traes_6AL_317133B3F.1	3817	1068	8183	22	115-1048	0.0
AGO1/10g	Traes_6DL_804FB7F75.1	5065	1067	8341	22	115-1048	0.0
AGO1/10h	Traes_6BL_9CFA54D4A.1	3910	1068	8173	22	115-1048	0.0
AGO1/10i	Traes_6AL_616161AAB.1	7364	978	11580	22	175-1048	0.0
AGO1/10j	Traes_6DL_58620B158.1	2724	877	6024	17	175-1048	0.0
AGO1/10k	Traes_7DL_C255A109C.1	2373	717	4969	14	178-870	0.0
AGO1/10l	Traes_7AL_E7C586D38.1	1907	490	3151	13	540-997	0.0
AGO1/10m	Traes_6AS_FBB2AFAAB.1	2847	948	7048	21	175-1048	0.0
AGO1/10n	Traes_6DS_9DD64BD48.1	4279	883	7934	3	314-1048	0.0
AGO1/10o	Traes_7DL_C538856D4.1	2469	664	4538	16	370-1048	0.0
AGO1/10p	Traes_7AL_96766587F.1	2121	582	4373	16	459-1048	0.0
AGO2/3a	Traes_2BL_7713B3533.2	3042	845	4841	2	162-966	0.0
AGO2/3b	Traes_2AL_DFE4C65F6.2	3221	952	4665	2	142-966	0.0
AGO4/8/9a	Traes_3DS_57EA31670.1	2536	799	5052	15	43-752	0.0
AGO4/8/9b	Traes_3B_F4E4667F8.1	2766	918	5556	21	43-924	0.0

AGO4/8/9c	Traes_3AS_8EE711E2C.2	3420	999	7119	22	43-924	0.0
AGO4/8/9d	Traes_1DL_64B330BBB.1	3943	922	9989	22	57-924	0.0
AGO4/8/9e	Traes_1AL_095416BC0.1	3445	925	8551	21	57-924	0.0
AGO4/8/9f	Traes_1BL_7C037D478.2	3288	927	9147	22	57-924	0.0
AGO4/8/9g	Traes_7AL_1BAB53DCE.1	2706	901	4825	22	52-924	0.0
AGO4/8/9h	Traes_7AL_D88450A3C.2	2393	768	4670	18	112-855	0.0
AGO5a	Traes_5BL_F505BF164.1	2451	732	5570	20	299-1048	0.0
AGO5b	Traes_4AL_7CC35DF1D.1	2127	705	5470	21	175-914	0.0
AGO5c	Traes_4DS_88D2821C6.1	2297	751	5472	21	175-923	0.0
AGO5d	Traes_2BS_8368F6B5D.1	2758	795	6346	21	193-1002	0.0
AGO5e	Traes_2DS_4CC8FD7E3.1	3053	838	6638	21	193-1048	0.0
AGO6a	Traes_5DL_672EE3605.1	3117	882	10088	22	28-878	0.0
AGO6b	Traes_5BL_F611D65E0.1	3235	883	10306	22	28-878	0.0
AGO6c	Traes_5AL_07EFD5712.1	2652	883	7500	22	28-878	0.0
AGO6d	Traes_1BL_05F7B7DFA.1	2725	893	6505	20	28-762	0.0
AGO7a	Traes_2BL_24111235C.1	3637	1023	4536	3	150-990	0.0
AGO7b	Traes_2AL_3F3117458.1	3125	934	4200	3	150-990	0.0
AGOa	Traes_3AS_3F8424E4E.1	2690	832	5109	19	179-1048	0.0

	AGOb	Traes_3B_CA99AB66C.1	2842	841	5204	21	179-1048	0.0
Musa acuminata (Mac)	AGO1/10a	GSMUA_Achr3T13970_001	3432	1060	8103	24	116-1048	0.0
	AGO1/10b	GSMUA_Achr1T17950_001	3305	1055	8047	24	115-1048	0.0
	AGO1/10c	GSMUA_Achr3T27070_001	3430	1076	10186	25	115-1048	0.0
	AGO1/10d	GSMUA_AchrUn_randomT09260_001	3361	1064	10776	24	153-1048	0.0
	AGO1/10e	GSMUA_Achr1T07120_001	3327	1057	8334	22	151-1048	0.0
	AGO1/10f	GSMUA_Achr4T30720_001	2802	933	7161	22	175-1048	0.0
	AGO1/10g	GSMUA_Achr6T05310_001	2983	928	7295	22	175-1048	0.0
	AGO1/10h	GSMUA_Achr7T13220_001	2841	946	7377	24	175-1048	0.0
	AGO4/8/9	GSMUA_Achr9T05150_001	3142	915	5793	22	52-924	0.0
	AGO5a	GSMUA_Achr6T04510_001	4065	1237	15822	30	134-997	0.0
	AGO6a	GSMUA_Achr9T27810_001	2616	871	4389	22	24-878	0.0
	AGO6b	GSMUA_Achr9T22630_001	2691	896	7028	22	28-878	0.0
	AGO7a	GSMUA_Achr7T26400_001	1914	637	2841	8	144-990	0.0
Miscanthus sinensis (Msi)	AGO1/10a	Misin11G197400.1	9177	1076	12213	22	115-1048	0.0
	AGO1/10b	Misin12G187600.1	2120	582	7555	16	467-1048	0.0
	AGO1/10c	Misin17G003900.1	3819	1110	7565	22	115-1048	0.0
	AGO1/10d	MisinT076700.1	3556	1104	6324	22	115-1048	0.0
	AGO1/10e	Misin08G136800.1	2406	801	11078	22	186-1048	0.0

AGO1/10f	MisinT017800.1	3898	1039	8238	22	115-1048	0.0
AGO1/10g	MisinT016500.1	3413	1015	8689	23	115-1048	0.0
AGO1/10h	Misin18G279300.1	3748	1048	7975	21	115-1048	0.0
AGO1/10i	Misin19G274500.1	3883	1049	7976	21	115-1048	0.0
AGO1/10j	Misin18G181500.1	1809	602	9238	19	410-1048	0.0
AGO1/10k	Misin19G169800.1	3463	968	6752	22	175-1048	0.0
AGO2/3a	Misin11G211100.1	3668	1014	5084	3	142-967	0.0
AGO2/3b	Misin12G205400.1	3619	1051	5022	3	142-1014	0.0
AGO4/8/9a	Misin06G110100.1	3832	901	9364	22	48-924	0.0
AGO4/8/9b	Misin05G127800.1	3408	900	8156	22	48-924	0.0
AGO4/8/9c	Misin16G263600.1	3380	909	7071	22	52-924	0.0
AGO4/8/9d	Misin17G266500.1	3631	909	11038	22	52-924	0.0
AGO4/8/9e	Misin06G109800.1	2023	656	18744	18	307-924	0.0
AGO4/8/9f	Misin05G127200.1	2712	903	12037	23	52-924	0.0
AGO5a	Misin02G123400.1	1702	485	3292	16	558-1048	0.0
AGO5b	MisinT157800.1	2415	792	7391	23	213-1048	0.0
AGO5c	Misin03G053300.1	3346	1034	7413	21	175-1048	0.0
AGO5d	Misin02G030700.1	3378	730	6368	20	314-1048	0.0
AGO6a	Misin15G098400.1	3402	881	10804	23	33-878	0.0
AGO6b	Misin14G080300.1	3861	856	12694	21	33-857	0.0
AGO7a	Misin01G314600.1	2740	812	4546	3	537-990	0.0
AGO7b	Misin02G304100.1	4126	1036	5719	3	142-990	0.0
AGOa	Misin03G248400.1	3536	1010	6576	21	179-1048	0.0
AGOb	Misin04G281500.1	1809	602	9238	19	360-988	0.0

	AGOc	Misin03G248300.1	2726	747	14477	23	190-988	0.0
	AGOd	Misin02G113300.1	2673	890	6765	23	128-988	0.0
Saccharum spontaneum (Spp)	AGO1/10a	Sspon.05G0005190-1A	ND	1082	8940	ND	19-1082	0.0
	AGO1/10b	Sspon.08G0006580-1A	ND	942	6373	ND	103-942	0.0
	AGO1/10c	Sspon.07G0016240-1A	ND	1146	6133	ND	1-976	0.0
	AGO1/10d	Sspon.08G0000380-1A	ND	1033	11620	ND	5-1033	0.0
	AGO2/3a	Sspon.05G0030960-1C	ND	1015	5040	ND	2-1015	0.0
	AGO2/3b	Sspon.05G0030950-1P	ND	1011	4496	ND	7-1011	0.0
	AGO4/8/9a	Sspon.02G0016200-2D	ND	946	11123	ND	1-890	0.0
	AGO4/8/9b	Sspon.03G0016720-2B	ND	805	5451	ND	1-770	0.0
	AGO4/8/9c	Sspon.03G0042850-1C	ND	840	11356	ND	1762	0.0
	AGO5a	Sspon.01G0030290-2B	ND	931	13611	ND	94-918	0.0
	AGO5b	Sspon.01G0028360-2B	ND	1054	6622	ND	209-1054	0.0
	AGO6a	Sspon.07G0020900-1A	ND	634	3694	ND	1-634	0.0
	AGO6b	Sspon.07G0020880-2C	ND	507	2835	ND	1-497	0.0
	AGO7a	Sspon.01G0012750-2C	ND	1029	4063	ND	1-1029	0.0

A filogenia das Argonautas mostrada na figura 12 revelou uma redução no número de famílias em monocotiledôneas em relação a Arabidopsis. Na árvore filogenética, as Argonauta se agrupam em sete clados diferentes, indicando que possivelmente em monocotiledôneas existam sete grupos de Argonauta, ao invés de dez famílias como em Arabidopsis. Essa redução no número de famílias de Argonauta, sugere uma sobreposição de função dessas proteínas em monocotiledôneas. Com relação aos nomes de cada grupo, as proteínas foram renomeadas de acordo com a sequência de referência que se agrupou no mesmo clado que as demais de monocotiledôneas. Um dos clados não teve nenhuma sequência de Arabidopsis, o que sugere que essa seja uma classe de Argonauta exclusiva de monocotiledôneas, sendo chamada de AGO. Os grupos de AGO5, AGO6 e AGO7 se agruparam separadamente e com respectiva AGO de Arabidopsis no mesmo clado (Figura 13). Já AGO1 e AGO10, se agruparam no mesmo clado e, por isso, as proteínas de monocotiledôneas foram chamadas de AGO1/10. O mesmo procedimento ocorreu com AGO2 e AGO3, que em monocotiledôneas foram chamadas de AGO2/3 e com as proteínas AGO4, AGO8 e AGO9, que em monocotiledôneas foram chamadas de AGO4/8/9.

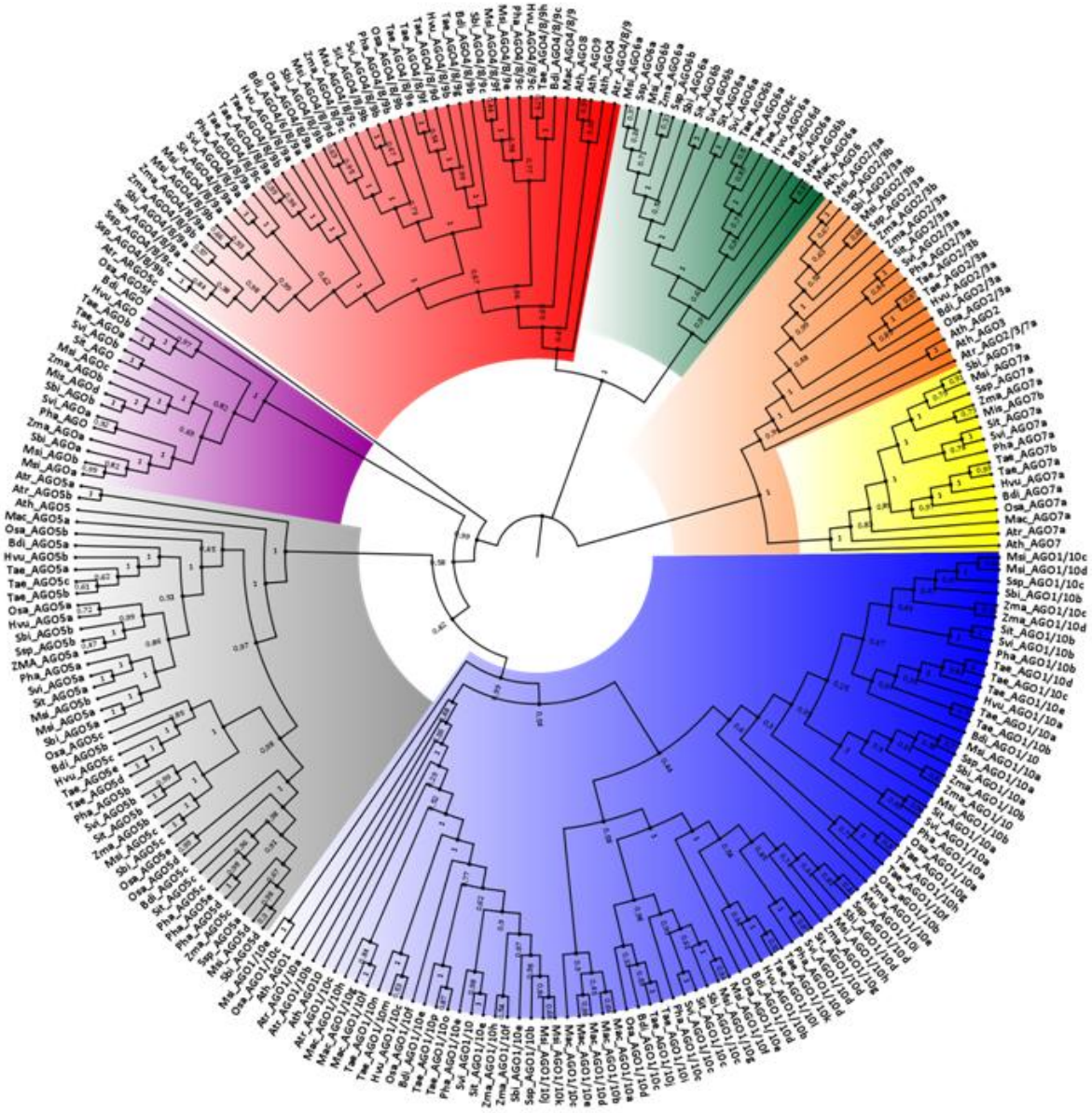
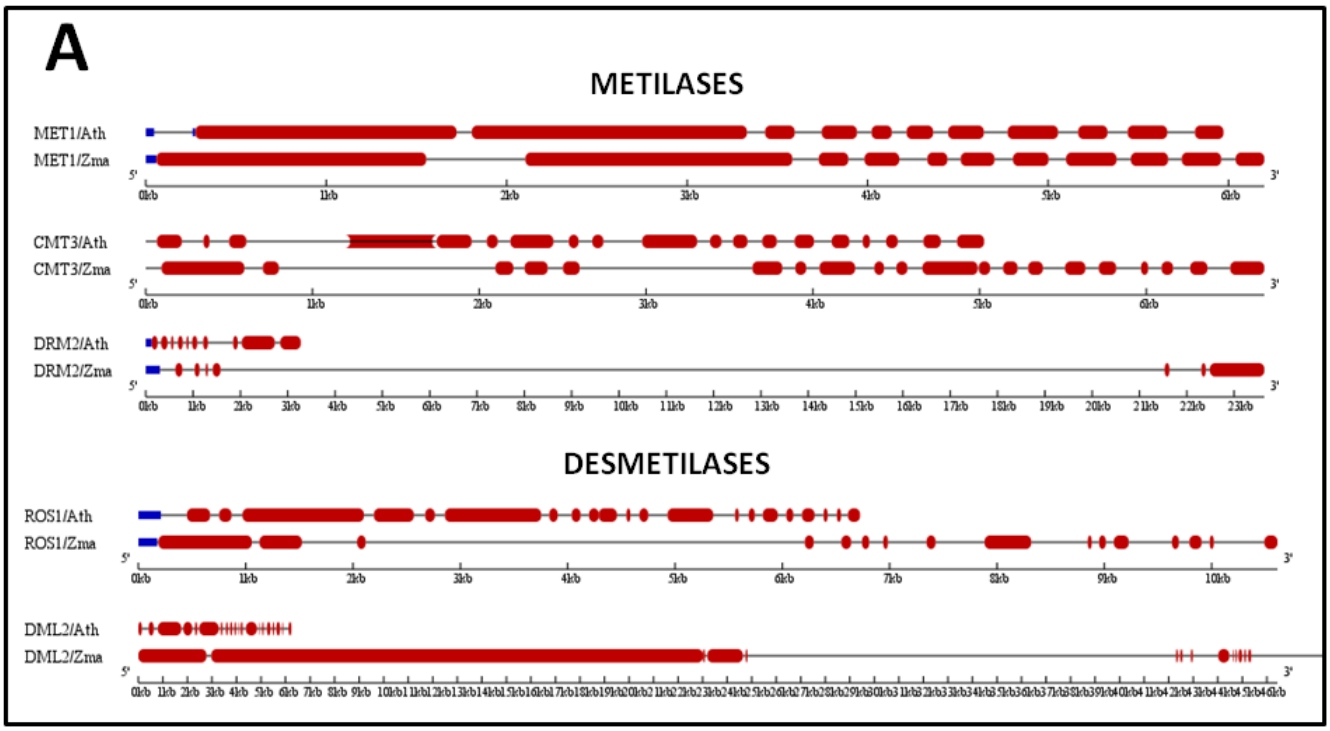
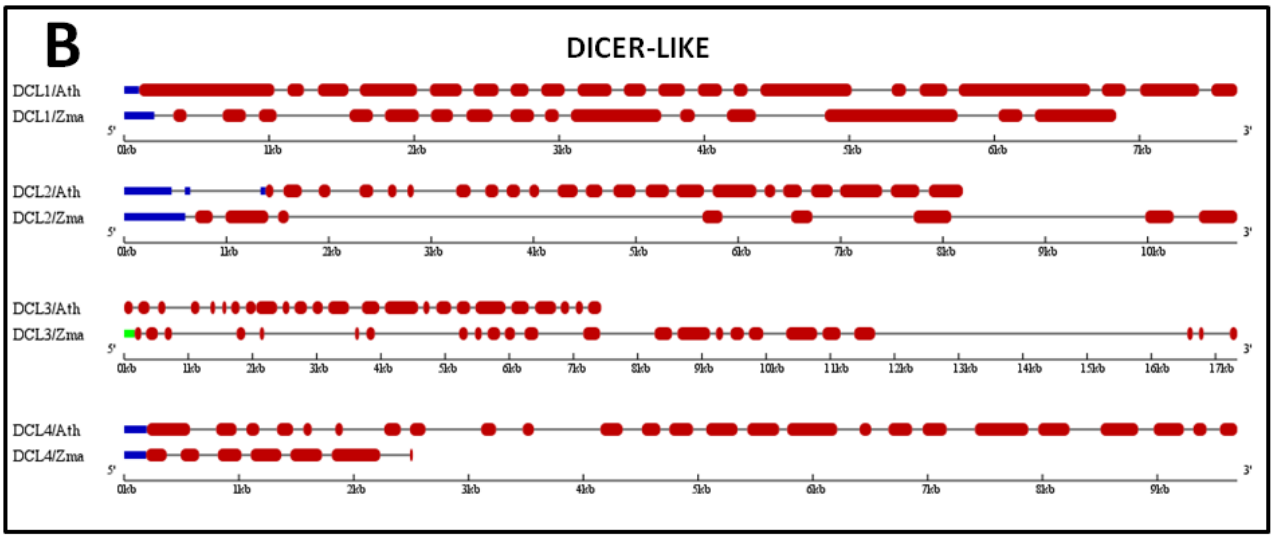


Figura 12: Filogenia de proteínas Argonautas de famílias de 1 a 10 em diferentes plantas monocotiledôneas. Em azul, Argonautas 1/10; em cinza, Argonautas 5; em amarelo Argonautas 7; em laranja, Argonautas 2/3; em verde, Argonautas 6; em vermelho, Argonautas 4/8/9 e em roxo, grupo de Argonautas exclusivo de monocotiledôneas

4.1.2. Estrutura gênica

Utilizando sequências genômicas disponíveis no Phytozome, foi realizada a anotação estrutural de genes metilação, desmetilação, Dicer-like e argonautas de Arabidopsis *e Z. mays*, a fim de analisar a conservação desses genes. A figura 13A revela uma boa conservação do gene MET1 de Arabidopsis em *Z. mays* com éxons longos, se comparada a CMT3, que também possuem boa conservação de éxons, em relação ao gene de Arabidopsis. Além disso, observamos que o gene DRM2 possui menos éxons e que em *Z. mays* houve uma grande expansão dos íntrons. A expansão dos íntrons pode ser observada também nas estruturas de ROS1 e DML2 de *Z. mays*, na mesma figura. Na figura 13B observamos uma boa conservação de DCL1, mas uma expansão no tamanho de DCL2 e DCL3 de *Z. mays*, em relação aos genes de Arabidopsis. Já na análise de DCL4, observamos um gene menor, com menos éxons em *Z. mays*, que possivelmente revela que o gene está fragmentado. Na figura 13C, observamos que o número de éxons de argonauta é conservado, com expansão dos íntrons em AGO1/10, AGO2/3, AGO4/8/9 e AGO6 de *Z. mays*. Os genes AGO2 e AGO7 possuem poucos e longos éxons, tanto em Arabidopsis quanto em *Z. mays.*





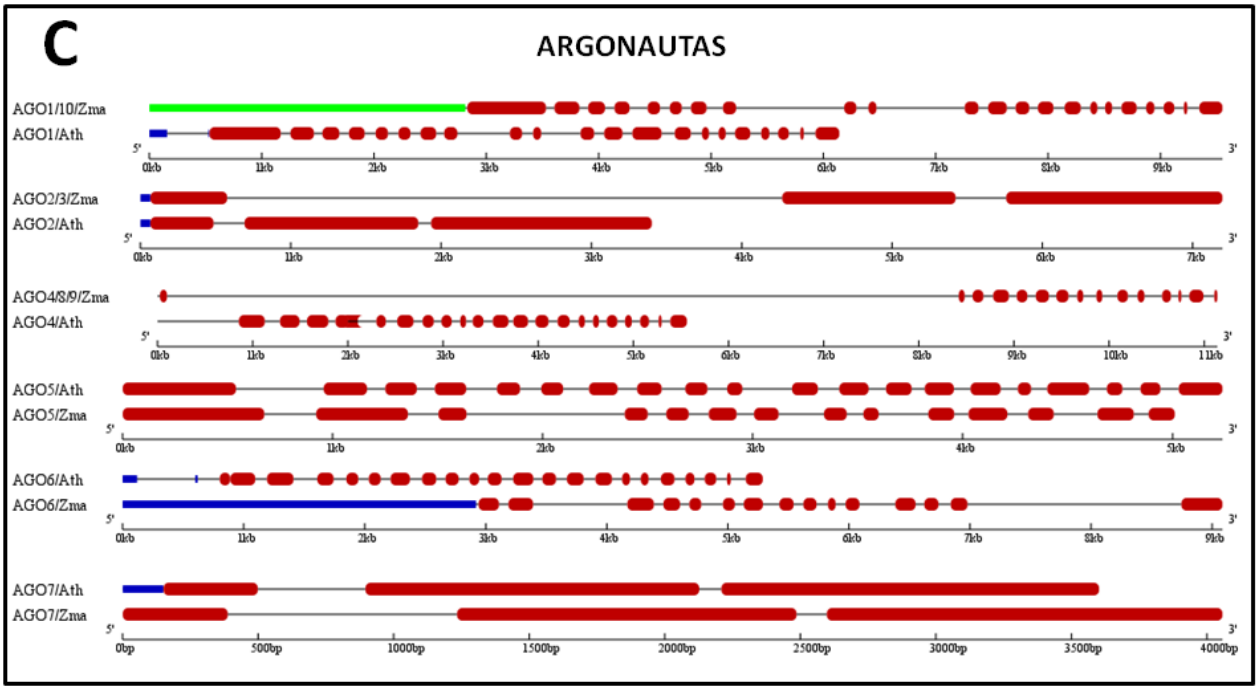


Figura 13: Representação de estruturas de genes de metilação e desmetilação (A), Dicer-like (B) e argonautas (C) em Arabidopsis thaliana e Zea mays. Os boxes em vermelho representam os exóns, boxes em azul representam a região 5'UTR, boxes em verde claro representam a região 3'UTR. Entre os boxes, as linhas representam os íntrons.

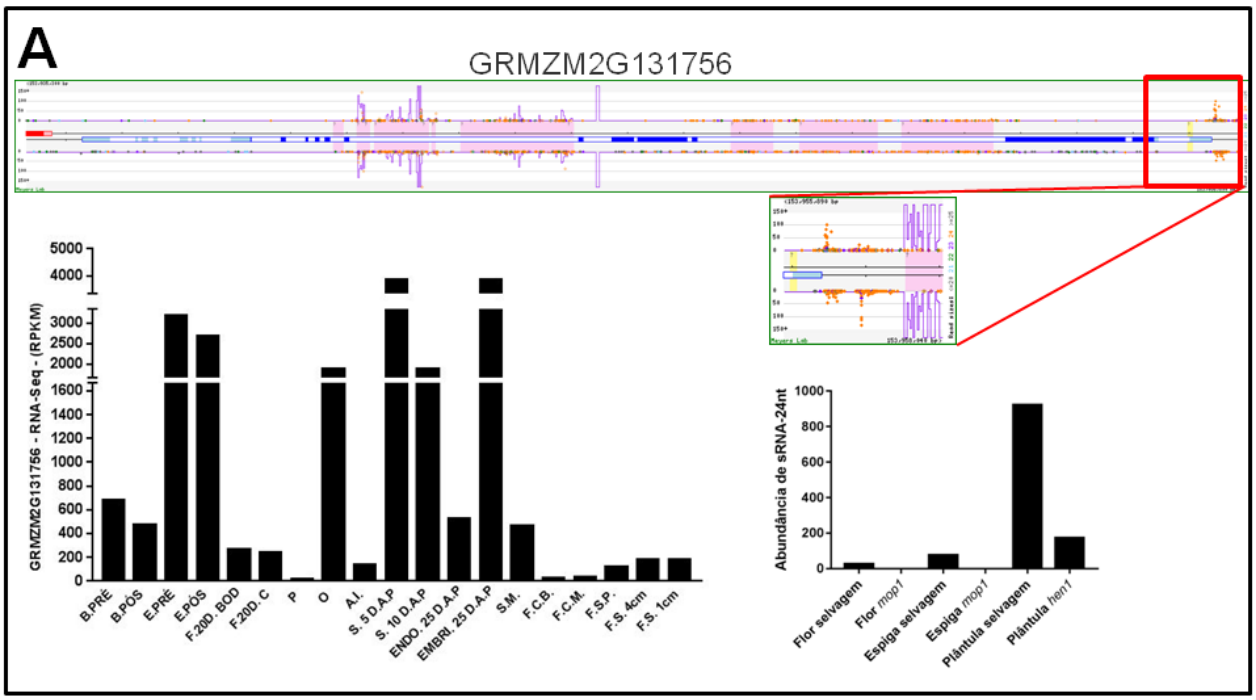
4.2. Identificação e caracterização de regiões MEMS

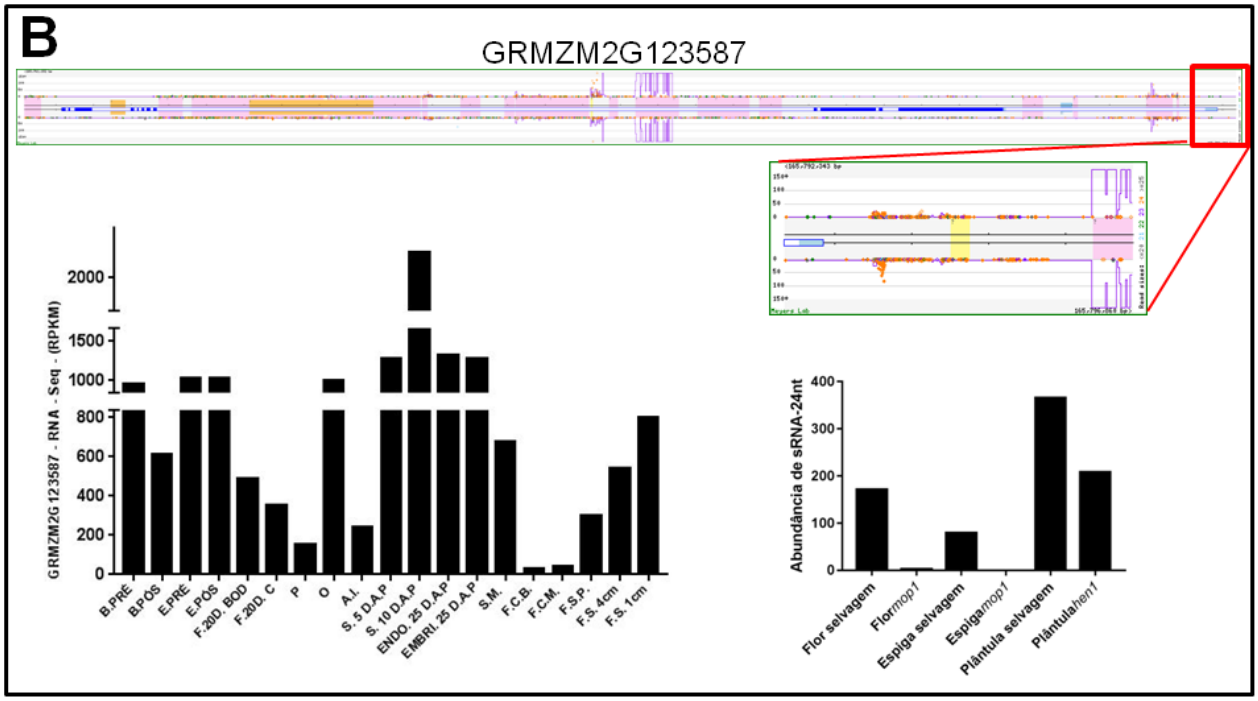
Utilizando o site https://mpss.danforthcenter.org/, foi realizada uma busca *in silico* de regiões descritas na literatura como "*DNA methylation monitoring sequence*" (MEMS) que atua regulando a expressão da desmetilase ROS1 em Arabidopsis. Essa região se caracteriza pelo acúmulo de sRNAs de 24 nucleotídeos, indicando uma hipermetilação efetuada pela via RdDM no genoma. ROS1 é uma desmetilase que atua no processo de desmetilação efetuado especificamente por RdDM. Outra característica de ROS1 é o alto nível de expressão em tecidos germinativos, ativando principalmente, genes envolvidos no processo de desenvolvimento. Em nosso trabalho, foram anotadas três desmetilases em *Z. mays*, quatro desmetilases em *O. sativa*.

4.2.1. Milho

Os genes de desmetilação de *Z. mays*, GRMZM2G131756_T01, GRMZM2G422464_T02 e GRMZM2G123587_T01, foram analisados. A figura 14 mostra os genes inteiros no sentido reverso, com destaque na região com acúmulo de sRNAs. Na figura, é possível ver o gráfico de abundância de sRNAs de 24 nucleotídeos nessa região, em diferentes bibliotecas de pequenos RNAs de milho. Além disso, utilizamos também, dados de RNA-Seq para demonstrar a expressão de cada gene em diferentes tecidos da planta. No gene GRMZM2G131756_T01, os níveis de sRNAs diminuíram na biblioteca mutante de *hen1*, que é uma metiltransferase de RNA envolvida na via de siRNA (YANG et al., 2006). Em *Z. mays*, o processo de paramutação controla a expressão de alelos de genes heterozigotos (CHANDLER, 2007), esse processo é efetuado por sRNA gerados a partir de uma proteína MOP (DORWEILER et al., 2000) que atua como um ortólogo de RDR2 em Arabidopsis (NOBUTA et al., 2008). Em mutantes para *mop,* tanto em flor quanto em espiga a expressão de sRNAs diminuiu quando comparados ao selvagem (Figura 14A). O nível de expressão de GRMZM2G131756_T01 em tecidos como espiga pré e pós-emergida (E.PRE e E.POS), em sementes com cinco e dez dias após a polinização (S.5.D.A.P. e S.10.D.A.P) e em embrião de 25 dias após a polinização (EMBRI. 25 D.A.P.) foi maior.

Na figura 14B, o gene GRMZM2G422464_T02 aparece com maior expressão nos diversos tecidos da planta, exceto no pólen (P), anteras inteiras (A.I.), folha células da bainha da folha (F.C.B.) e células mesófilas da folha (F.C.M.). No gráfico de sRNAs, foi observada uma redução na abundância de sRNAs nos mutantes de *hen1* e *mop flor e mop* espiga, em relação aos selvagens. Na figura 14C, o gene GRMZM2G123587_T01, que também se encontra no sentido reverso apresenta um padrão de expressão semelhante ao GRMZM2G422464_T02 da figura 14B, apresentando apenas a diferença de ser menos expresso em folhas com 20 dias de idade - câmara de crescimento (F.20D.BOD). No gráfico de sRNAs, observamos um padrão de abundância de sRNAs semelhante ao observado na figura 14A, principalmente na biblioteca de mutantes para HEN1.




B
GRMZM2G123587
GRMZM2G 123587 - RNA - Seq - (RPKM)
2000
1500
1000
800
600
400
200
0
B.PRÉ
B.PÓS
E.PRÉ
E.PÓS
F.20D. BOD
F.20D. C
P
O
A.I.
S. 5 D.A.P.
S. 10 D.A.P.
ENDO. 25 D.A.P.
EMBRI. 25 D.A.P.
S.M.
F.C.B.
F.C.M.
F.S.P.
F.S. 4cm
F.S. 1cm
Abundância de sRNA-24nt
400
300
200
100
0
Flor selvagem
Flor*mop1*
Espiga selvagem
Espiga*mop1*
Plântula selvagem
Plântula*hen1*

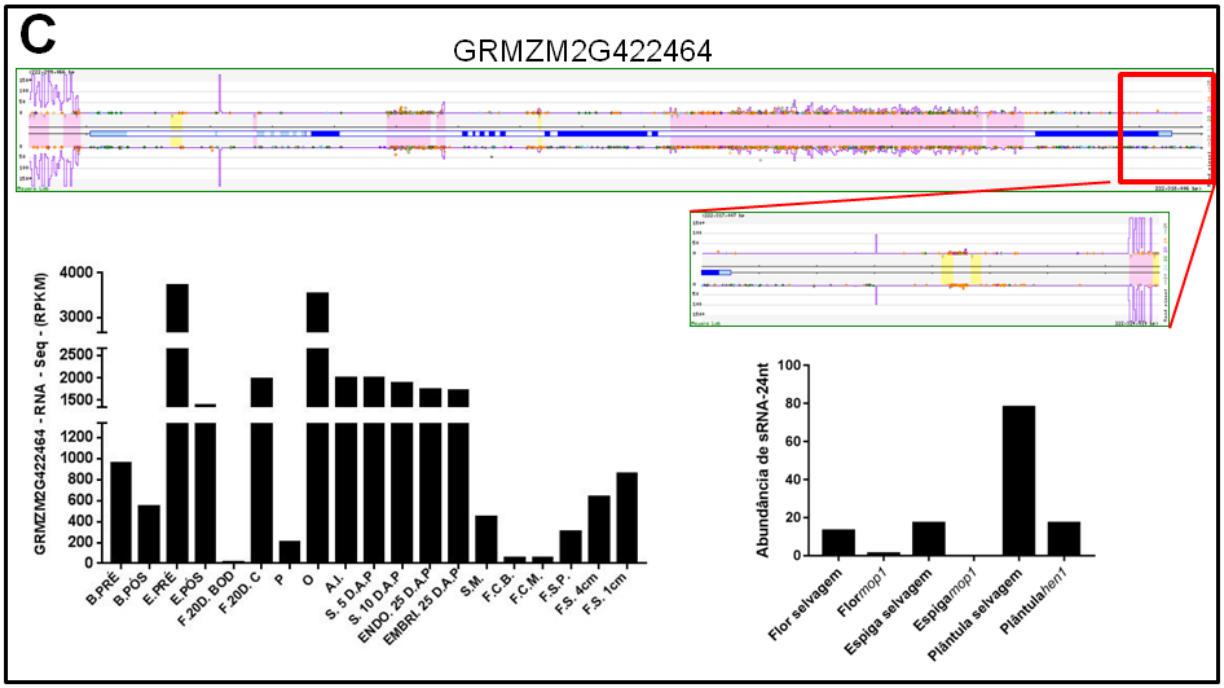


Figura 14: Identificação de possíveis regiões MEMS no genoma de milho. Os genes de desmetilases encontram-se no sentido reverso e as regiões de éxons estão representadas em azul. O boxe vermelho destaca a possível localização das regiões MEMS, devido ao acúmulo de sRNAs de 24 nucleotídeos, representados por pontos laranjas. Os demais pontos coloridos representam pequenos RNAs de outros tamanhos. O gráfico à esquerda representa a expressão dos genes de desmetilação nos seguintes tecidos: B.PRE (Borla pré-emergida), B. PÓS (Borla pós-emergida), E.PRE (Espiga pré-emergida), E.PÓS (Espiga pós emergida), F.20D.BOD (Folhas com 20 dias de idade - câmara de crescimento), F. 20D.C (Folhas de 20 dias de mudas - cultivadas em campo), P (Pólen), O (Óvulo), A.I. (Anteras inteiras), S.5 D.A.P. (Semente 5 dias após a polinização), S. 10 D.A.P. (Semente 10 dias após a polinização), ENDO. 25 D.A.P. (Endosperma 25 dias após a polinização), EMBRI. 25 D.A.P. (Embrião 25 dias após a polinização), S.M. (Sedas maduras), F.C.B. (Folha de milho, células da bainha), F.C.M. (Folha de milho, células mesófilas), F.S.P. (Folha de milho, seção de ponta), F.S. 4 cm (Folha de milho, seção de 4 cm) e F.S. 1 cm (Folha de milho, seção de 1 cm). O gráfico à direita representa a abundância de sRNAs de 24 nucleotídeos

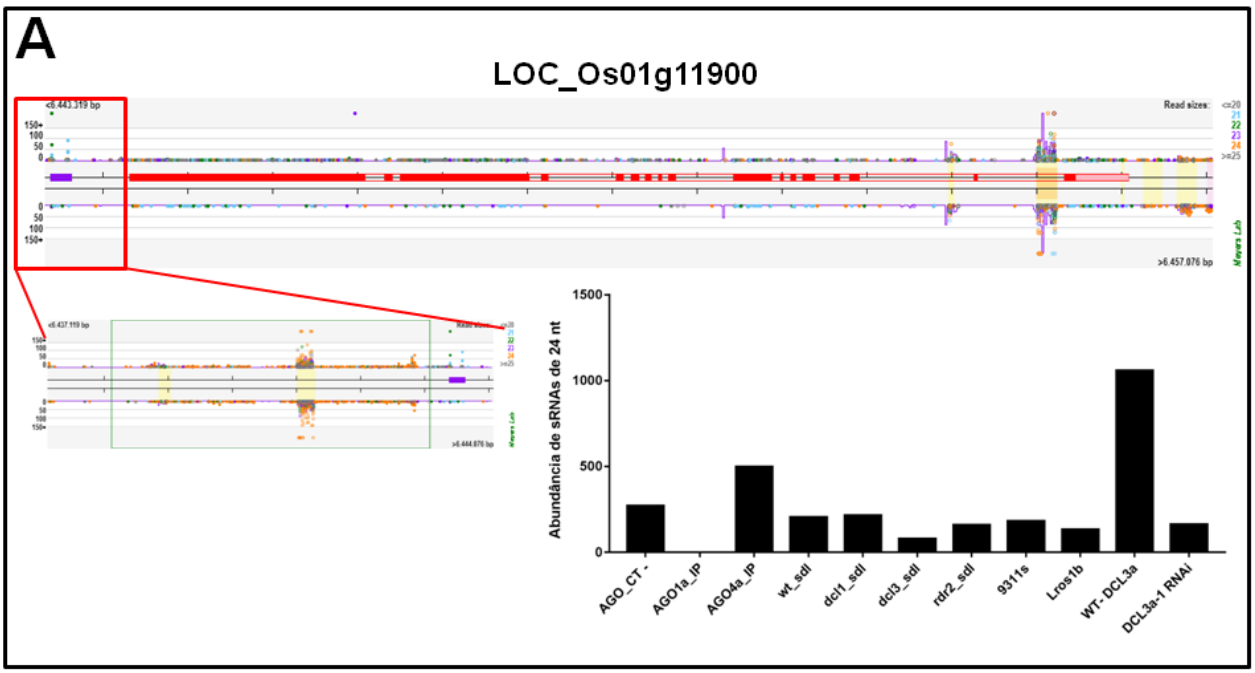
4.2.2. Arroz

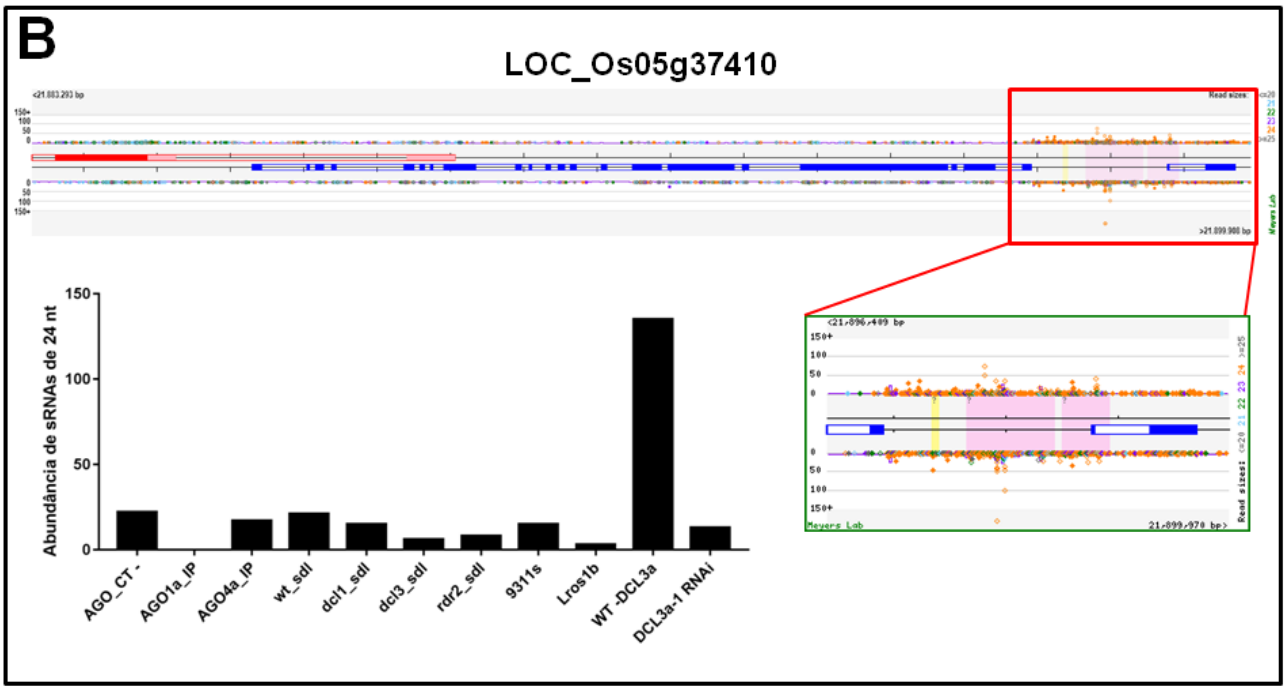
Nas análises com *O. sativa*, utilizamos os genes de desmetilação LOC_Os01g11900.1, LOC_Os05g37410.1, LOC_Os05g37350.1 e LOC_Os02g29230.1, que são mostrados na figura 15. Na figura 15A, observamos o gene LOC_Os01g11900.1 em vermelho, no sentido direto. É possível observar que próximo à região promotora desse gene, existe um gene MIR osa-Mir2096 que pode influenciar na associação de siRNAs nessa região. A região que possivelmente abriga o MEMS foi observada anteriormente a esse elemento transponível e é demonstrado no boxe ampliado. Ainda na figura 15A, o gráfico de sRNAs mostra grande abundância destes em praticamente todas bibliotecas analisadas, exceto para AGO1-IP (sRNAs ligados a AGO1a em mudas com 3 semanas de idade). Outro fator interessante mostrado na figura 15A é a diminuição dos sRNAs de DCL3–RNAi comparada ao controle de DCL3, ressaltando que DCL3 participa ativamente da via de síntese destes pequenos RNAs.

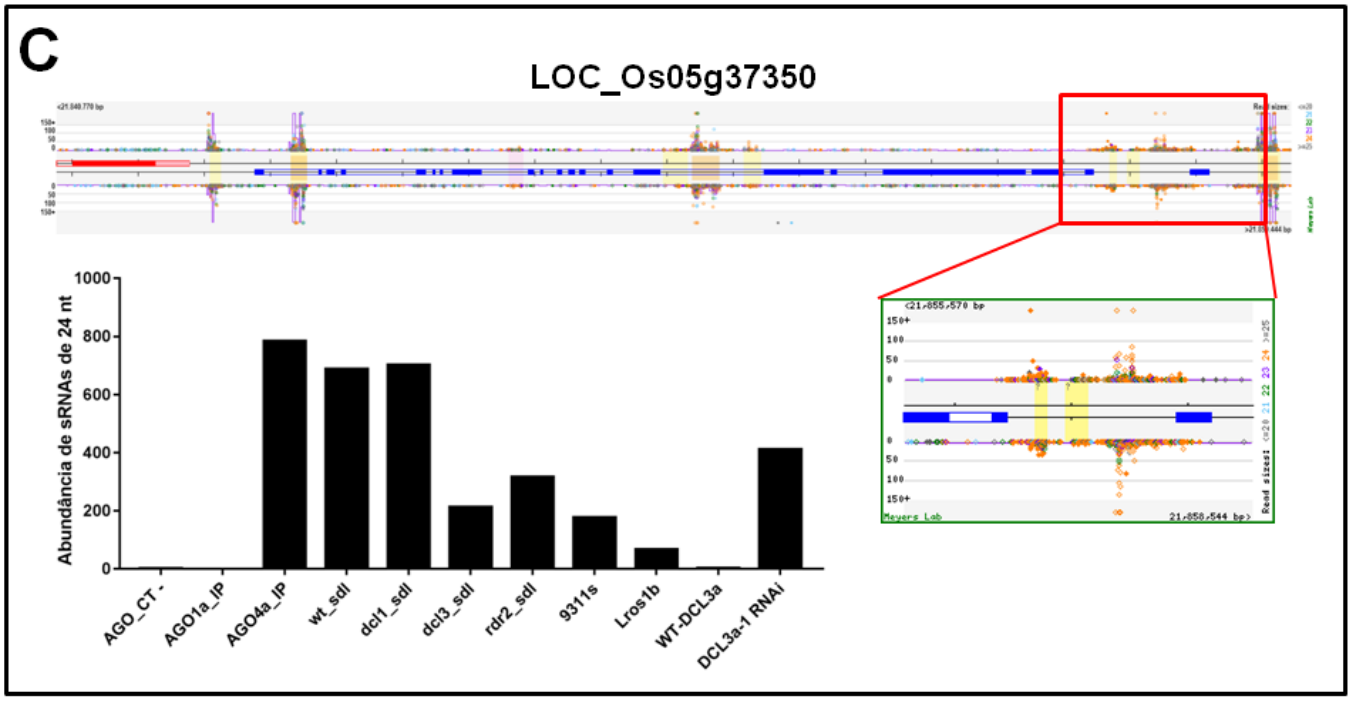
Na figura 15B, o gene LOC_Os05g37410.1, o gráfico de abundância de sRNAs, apresenta o mesmo comportamento em relação as bibliotecas de DCL3, porém em menor proporção como pode ser visto na escala do gráfico. No geral, essa região apresentou uma menor abundância de pequenos RNAs em praticamente todas as bibliotecas analisadas, ressaltando que as bibliotecas mutantes de AGO1 praticamente não mostraram presença de sRNAs.

Na figura 15C, observamos o gene LOC_Os05g37350.1 no sentido reverso com a região MEMS próxima a um pequeno gene de apenas um éxon, ampliada logo abaixo. O gráfico de abundância demonstra um padrão de sRNAs diferentes dos mostrado nas figuras 15A e 15B, isso porque as maiores abundâncias de sRNAs foram observadas nas bibliotecas de AGO4a_IP – (sRNA ligados a AGO4a , extraído de mudas com 3 semanas de idade), *dcl1*_sdl – (sRNA de mutantes para dcl1, extraído de mudas com 3 semanas de idade), wt_sdl –(sRNA de tipo selvagem, extraído de mudas com 3 semanas de idade) e as bibliotecas de DCL3 apresentaram pequena abundancia de sRNAs de 24 nucleotídeos.

Na figura 15D, observamos o gene LOC_Os02g29230.1 em vermelho, no sentido direto com a região MEMS ampliada logo abaixo. O gráfico de abundância de siRNAs segue o padrão de abundância observado nas figuras 15A e 15B, com destaque para a diminuição de sRNAs em bibliotecas de DCL3 – RNAi e o aumento de sRNAs em biblioteca de AGO4a, sendo que esses dois genes estão envolvidos na via de siRNAs.


A
LOC_Os01g11900
Abundância de sRNAs de 24 nt
1500
1000
500
0
AGO_CT -
AGO1a_IP
AGO4a_IP
wt_sdl
dcl1_sdl
dcl3_sdl
rdr2_sdl
9311s
Lros1b
WT-DCL3a
DCL3a-1 RNAi


B
LOC_Os05g37410
Abundância de sRNAs de 24 nt
150
100
50
0
AGO_CT -
AGO1a_IP
AGO4a_IP
wt_sdl
dcl1_sdl
dcl3_sdl
rdr2_sdl
9311s
Lros1b
WT -DCL3a
DCL3a-1 RNAi


C
LOC_Os05g37350
Abundância de sRNAs de 24 nt
1000
800
600
400
200
0
AGO_CT -
AGO1a_IP
AGO4a_IP
wt_sdl
dcl1_sdl
dcl3_sdl
rdr2_sdl
9311s
Lros1b
WT-DCL3a
DCL3a-1 RNAi

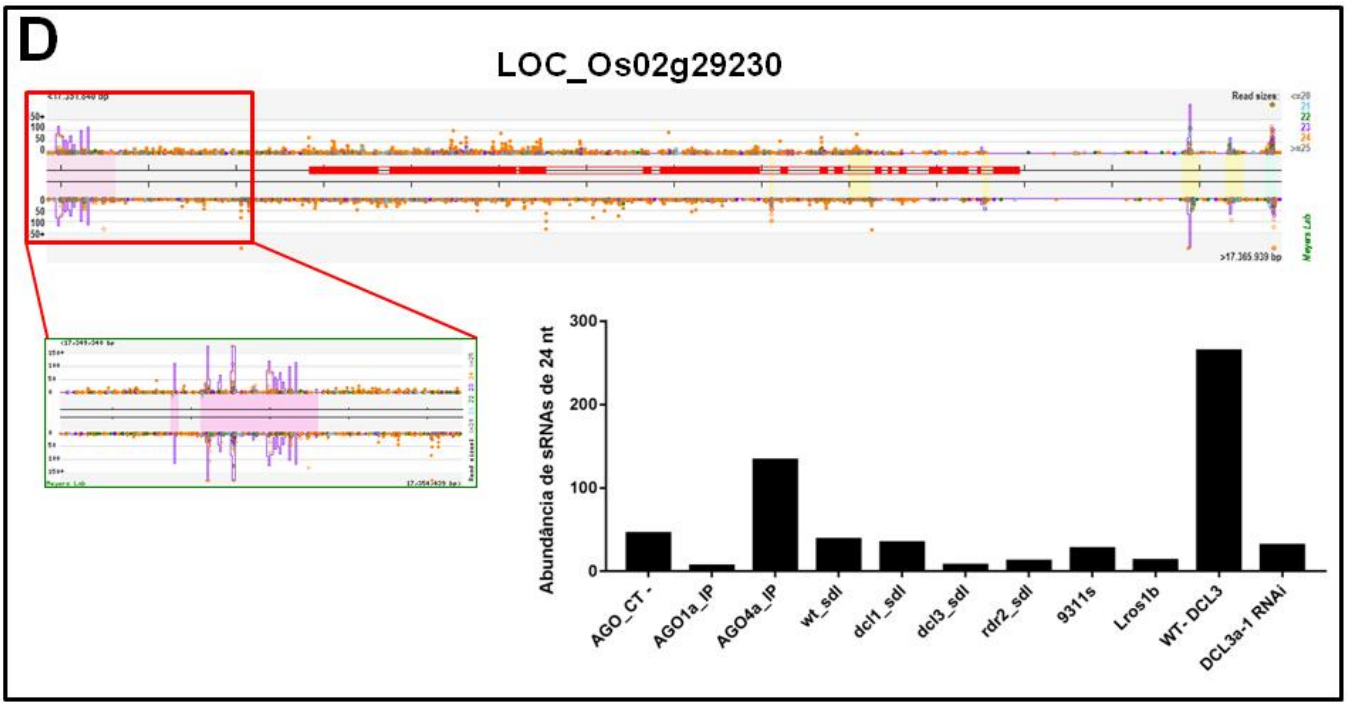


Figura 15: Identificação de possíveis regiões MEMs no genoma de arroz. Os genes de desmetilases que se encontram no sentido reverse estão representados com suas regiões de éxons em azul, os genes de desmetilases que encontram-se no forward estão representados com suas regiões de éxons em vermelho. O boxe vermelho destaca a possível localização das regiões MEMs, devido ao acúmulo de siRNA de 24 nucleotídeos, representados por pontos laranjas. Outros pontos coloridos representam pequenos RNAs de outros tamanhos. O gráfico representa a abundância de siRNAs de 24 nucleotídeos nas regiões que antecedem os genes de desmetilação nas seguintes bibliotecas de transcriptoma de arroz: AGO-CT (controle sRNA para mutantes de AGO) AGO1a_IP – (sRNAs de AGO1a extraídos de mudas com 3 semanas de idade), AGO_CT – (controle de sRNA para AGO1a), AGO4a_IP – (sRNA de AGO4a , extraído de mudas com 3 semanas de idade), wt_sdl – (sRNA de tipo selvagem, extraído de mudas com 3 semanas de idade) dcl1_sdl – (sRNA de mutantes para dcl1, extraído de mudas com 3 semanas de idade), dcl3_sdl – (sRNA do os dcl3, extraído de mudas com 3 semanas de idade), rdr2_sdl – (sRNA de mutantes de rdr2, extraído de mudas com 3 semanas de idade), 9311s – (mudas de 6 dias do tipo selvagem 93-11) Lros1b – (Folhas, mutantes para ros1), , WT – DCL3 – (Arroz de tipo selvagem), DCL3a-1 RNAi – (Arroz mutante para DCL3a-1)

4.3. Análise do efeito do inibidor de metilação na expressão da ROS1 e no desenvolvimento vegetal

Para análise do efeito da desmetilação na expressão dos genes ROS1, plântulas de milho foram tratadas com o inibidor de metilação 5-Azac. A figura 16 mostra o fenótipo dessas plântulas nos diferentes tratamentos. As plântulas que receberam o tratamento com DMSO apresentaram pequenas diferenças fenotípicas, como parte aérea e raízes menores em relação ao tratamento com H_2O. Porém, as plântulas que receberam o tratamento com 5-Azac, apresentaram parte aérea e raízes muito reduzidas em relação às plântulas tratadas com H_2O e as tratadas com DMSO.

Figura 16: Avaliação fenotípica do desenvolvimento de plântulas de milho com inibidor de metilação (5-Azac) e controles DMSO e H_2O).

A figura 17 mostra os valores das medições de comprimento, massa fresca, massa seca, biomassa e conteúdo relativo de água da parte aérea das plântulas nas três condições. Em todos os gráficos, foi aplicado o TEST T, para avaliar significância estatística entre as amostras. Os gráficos A, B e C mostraram diferença estatística nos valores observados nas plântulas tratadas com 5-Azac, já os gráfico D e E não apresentaram diferenças significativas.

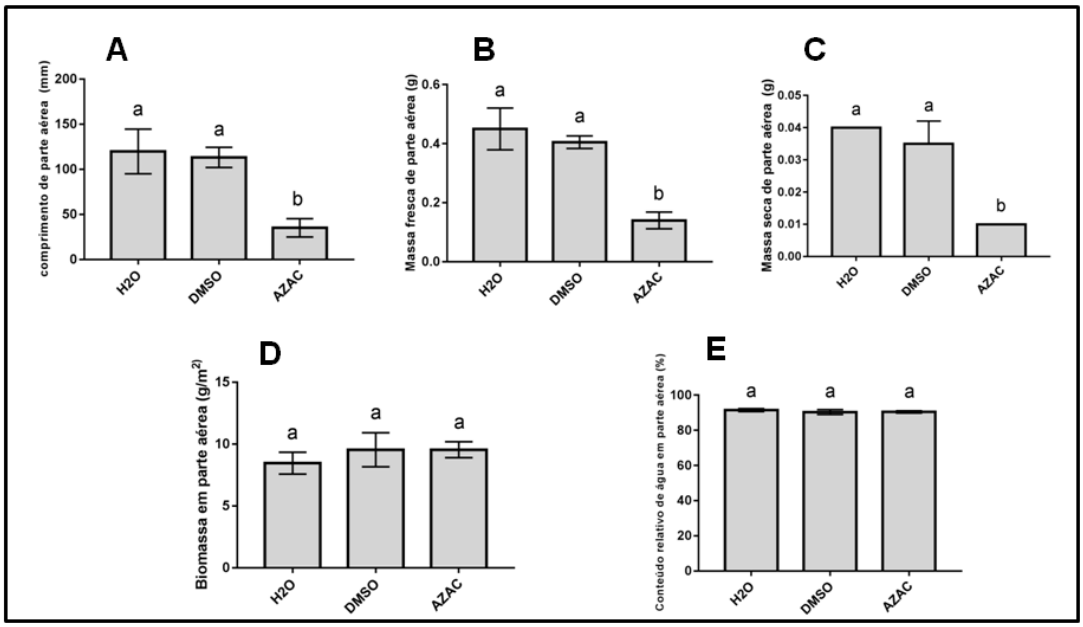


Figura 17: Análise do desenvolvimento da parte aérea de plântulas de milho sob condições de controle e tratadas com inibidor de metilação 5-Azac. Em (A) comprimento em (mm), em (B) massa fresca em gramas, em (C) massa seca em gramas, em (D) biomassa em (g/m^2) e em (E) conteúdo relativo de água em (%). Análise de significância com *p-value* <0,05.

Na figura 18, observamos os dados referentes aos valores estatísticos de raiz das plântulas de milho cultivadas nas três condições diferentes. Semelhante ao que foi observado em parte aérea, os gráficos A, B e C demonstraram diferenças significativas entre plântulas de milho tratadas com 5-Azac em relação aos demais tratamentos, tanto no comprimento, na massa fresca e na massa seca de raízes. O gráfico D, mostra que houve um aumento de biomassa em raízes tratadas com 5-Azac, em relação aos demais tratamentos, porem não houve diferença significativa nesses valores. O gráfico E também não apresentou diferença significativa entre as amostras.

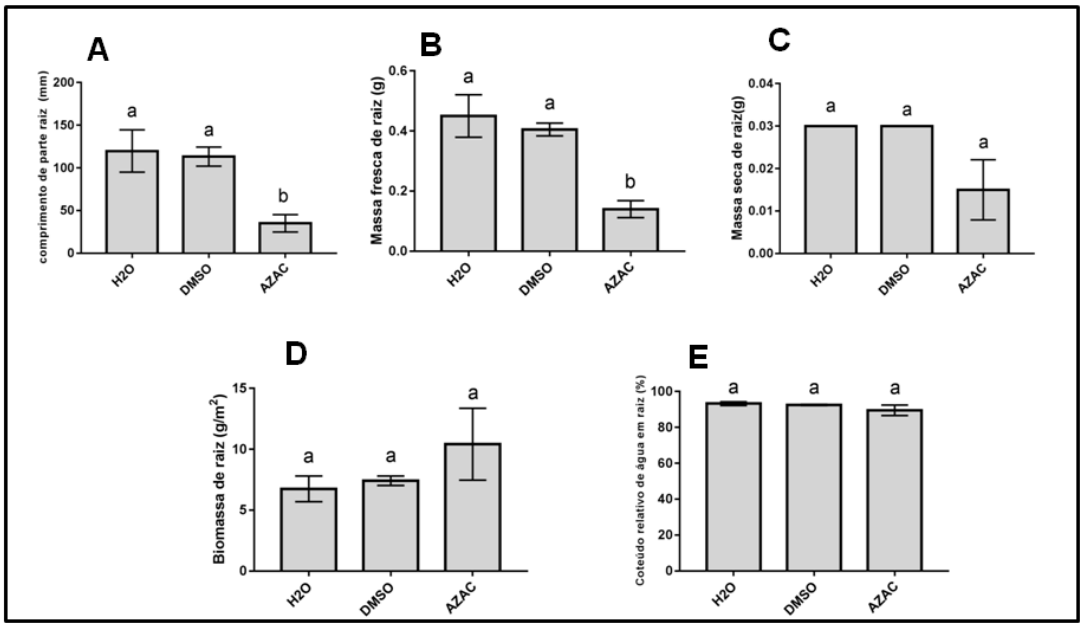


Figura 18: Análise do desenvolvimento de raiz de plântulas de milho sob condições controle e tratamento com inibidor de metilação 5-AzaC. Em (A) comprimento em (mm), em (B) massa fresca em gramas, em (C) massa seca em gramas, em (D) biomassa em (g/m^2) e em (E) conteúdo relativo de água em (%). Análise de significância com *p-value* <0,05.

Resultados semelhantes, porém não significativamente diferentes foram observados em *S. spontaneum* (Apêndice A), que mostrou um menor desenvolvimento de parte aérea de plântulas.

A fim de comprovar o efeito do 5-Azac sob a expressão de ROS1 em plântulas de milho, foi realizada a extração de RNA dessas plântulas, seguido de tratamento com DNAse, construção de cDNA e RT-qPCR. A desmetilase ROS1 atua de maneira antagônica a metiltransferases, podendo assim, ser utilizada como marcador da presença de metilação no genoma de milho. A figura 19A mostra uma redução significativa da expressão de ROS1 no tratamento com 5-Azac. Na figura 19B, foi utilizado primer para outro gene ROS1, e corrobora o resultado observado no gráfico A, onde ROS1 é significativamente menos expressa no tratamento com 5-Azac.

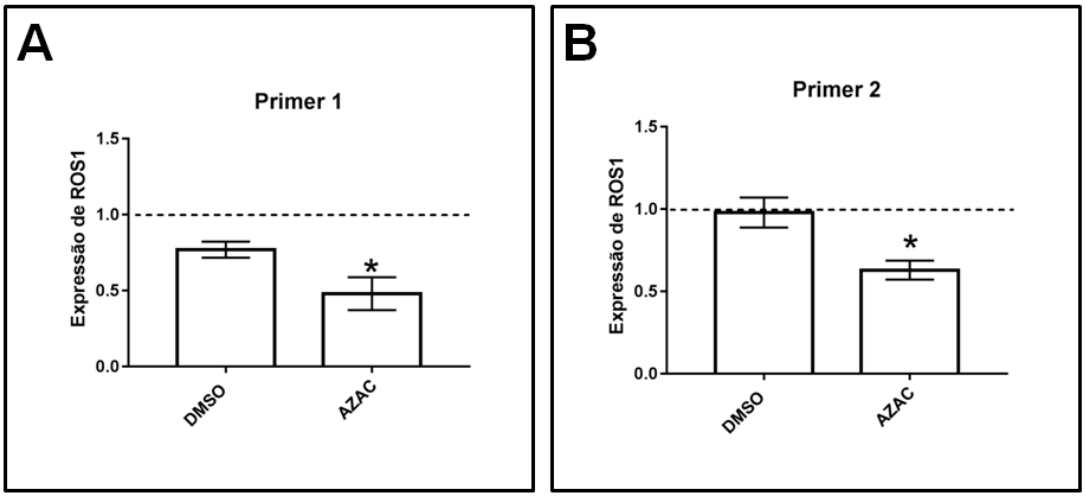


Figura 19: Expressão do gene ROS1 através de RT-qPCR em parte aérea de plântulas de milho. A linha tracejada representa a expressão de ROS1 em situação controle utilizando apenas H_2O no tratamento. Os asteriscos representam que houve significância estatística em relação ao controle ao *p-value* <0,05. Em A, o primer ROS1 - DNG103 e em B primer ROS1 - DNG101.

5. Discussão

As incríveis descobertas cientificas do século passado e a chegada da tecnologia de sequenciamento de DNA, abriram as portas para o surgimento de um novo ramo científico, as ciências genômicas, que engloba conhecimentos da informática, biologia molecular, bioquímica e genética para compreender e desvendar importantes mecanismos que atuam na regulação funcional do genoma. Nesse contexto, a epigenética surge com grande força para explicar a regulação da expressão gênica. Atualmente, vivemos na era pós genômica e nosso maior desafio é compreender a integração dos mecanismos epigenéticos e como eles reagem em situações diferentes, principalmente em condições de estresses. Nesse processo, é importante que os efetores desses mecanismos sejam bem conhecidos e caracterizados.

Em nossa análise *in silico*, buscamos anotar e caracterizar genes envolvidos na metilação e desmetilação do DNA em diferentes plantas monocotiledôneas. Além disso, realizamos a identificação desses genes em *S. spontaneum*, utilizando as informações do genoma publicado por (ZHANG et al., 2018). No total, foram anotados 437 IDs de genes envolvidos nas vias de metilação e desmetilação do DNA, e via de pequenos RNAs, além dos 21 genes de Arabidopsis utilizados como referência para nossa anotação. Um de nossos objetivos foi analisar a evolução desses genes dentro do grupo das monocotiledôneas e, portanto, foram realizadas análises filogenéticas.

Segundo PEI et al., 2019 os genes envolvidos na via de manutenção da metilação, que é o caso de MET1 e CMT3, evoluíram precocemente enquanto os genes envolvidos na via de metilação *de novo* surgiram mais tarde. Eventos de poliploidia também fizeram com que houvesse uma expansão desses genes em algumas plantas (PEI et al., 2019). As monocotiledôneas possuem vários representantes com genoma duplicado, como *O. sativa, S, spontaneum, B. distachyon, M. acuminata, M. sinensis e T. aestivum,* entre outras. Esse fato pode ser observado em nossas anotações, por exemplo, enquanto Arabidopsis e *A. trichopoda* possuem respectivamente três e quatro metilases, as monocotiledôneas não apresentam menos que cinco, com algumas plantas apresentando um número bem elevado, como é o caso de *M. sinensis* com oito, *B. distachyon* com nove e *T. aestivum* com 21. Já com as desmetilases, ocorre uma redução do número de genes em monocotiledôneas, como mostrado na figura 9, que mostra a presença de dois grandes grupos de desmetilases em monocotiledôneas. Isto sugere que as mesmas possuem funções menos específicas, por exemplo, em Arabidopsis, DME é relatada como uma

desmetilase que possui maior atuação em sementes (CHOI et al., 2002). Arabidopsis possui quatro desmetilases, esse número se repete em *O. sativa* e *B. distachyon* e é maior apenas em *S. spontaneum,* e *T. aestivum.* Nas demais plantas, o número de desmetilases anotadas foram de duas ou três em cada.

Em relação ao grupo das Dicer-like, observamos uma expansão no número de representantes nas monocotiledôneas, principalmente DCL3, que aparece duplicada em quase todas as plantas anotadas. A exceção fica com *M. acuminata* e *S. spontaneun*, que apresentaram apenas uma DCL3. Já *B. distachyon* e *T. aestivum* apresentaram três e quatro DCL3, respectivamente.

No grupo das Argonautas, a análise filogenética mostrou que as Argonauta anotadas para monocotiledôneas, formam apenas sete clados (Figura 12), sugerindo que existem apenas sete famílias, ao invés de dez, como em Arabidopsis. Porém, apesar da filogenia revelar um número menor de famílias Argonauta em monocotiledôneas, essas possuem um número maior de membros. A redução do número de famílias, sugere uma sobreposição de função das Argonauta em monocotiledôneas. Assim, as classes de Argonauta foram chamadas de AGO1/10, AGO2/3, AGO4/8/9, AGO5, AGO6, AGO7 e AGO, sendo essa última sem representante Argonauta de Arabidopsis no mesmo clado, sugerindo que o mesmo seja um grupo exclusivo de monocotiledôneas.

Enquanto a anotação teve como objetivo demonstrar a evolução desses genes especificamente dentro do grupo das monocotiledôneas, (PEI et al., 2019) anotou 33 genes envolvidos na via de metilação em 77 espécies diferentes a fim de investigar a origem e evolução desses genes no reino vegetal. O número de genes anotados como metilases e desmetilases, assim como de DCL3 e AGO4 de ambos os trabalhos, podem ser conferidos na tabela 7.

Tabela 7: Tabela de comparação de genes anotados pelo nosso grupo e por Pei et al 2019.

	Metilases		**Desmetilases**		**DCL3**		**AGO4**	
Planta/ Abreviação	**Nosso grupo**	**Pei et al**	**Nosso grupo**	**Pei et al**	**Nosso grupo**	**Pei et al**	**Nosso grupo**	**Pei et al**
Arabidopsis thaliana (Ath)	3	3	4	4	1	1	1	1
Sorghum bicolor (Sbi)	6	7	3	5	2	2	3	2
Amborela trichopoda (Atr)	4	3	2	2	0	1	1	1
Brachypodium distachyon (Bdi)	9	7	4	5	3	2	3	2
Hordeum	5	4	1	3	2	2	3	3

vulgare (Hvu)								
Panicum hallii (Pha)	5	-	2	-	2	-	3	-
Setaria itálica (Sit)	7	6	3	4	2	2	2	3
Setaria viridis (Svi)	6	-	3	-	2	-	2	-
Zea mays (Zma)	5	6	3	3	1	1	3	2
Oryza sativa (Osa)	5	6	4	6	2	2	2	2
Triticum aestivum (Tae)	21	30	7	12	4	4	8	12
Musa acuminata (Mac)	7	5	2	3	1	2	1	2
Miscanthus sinensis (Msi)	8	-	2	-	2	-	6	-
Saccharum spontaneum (Ssp)	5	-	7	-	1	-	3	-

Em geral, existem concordâncias e discordâncias nas anotações, possivelmente por utilização de metodologias diferentes, já que Pei utilizou diferentes bancos de dados e admitiu sequências com 30% de similaridade com as sequências proteicas de Arabidopsis. Em nossas análises, admitimos apenas sequências acima de 50% de similaridade. Pei e colaboradores também relataram ter observado na estrutura de vários genes a presença de íntrons mais longos, o que pode ser observado nas nossas análises, quando comparamos as estruturas dos genes de *Z. mays* com Arabidopsis, sendo que o número de éxons é conservado (Figura 13).

Na análise *in silico*, para buscar regiões denominadas MEMS utilizando o site https://mpss.danforthcenter.org/ (NAKANO et al., 2019). As regiões MEMS são caracterizadas por abrigarem siRNAs de 24 nucleotídeos, como uma forma de sensor de metilação do genoma. Esses siRNAs se acumulam numa posição estratégica, antes da região promotora do gene ROS1, uma desmetilase que atua de maneira antagônica ao processo RdDM (LEI et al., 2015), controlando os níveis de metilação do genoma. Portanto a expressão do gene ROS1 é controlada por MEMS e a atuação de MEMS depende do nível de metilação realizada por RDdM (LU et al., 2020).

Na literatura não há muitos relatos de identificação dessa região MEMS em plantas monocotiledôneas. Portanto, baseados em informações obtidas por estudos com Arabidopsis (LU et al., 2020), buscamos identificar essa região em *Z. mays e O. sativa.* Em *Z. mays*, observamos na figura 14A, uma maior expressão do gene GRMZM2G131756 em tecidos germinativos, fato bem característico de genes de desmetilação (LAW; JACOBSEN, 2010). Nas bibliotecas de sRNA, foi vista uma baixa

abundância de sRNAs em bibliotecas de flor e espiga *mop*, em relação ao selvagem, e no mutante *hen1*. Essa variação de sRNAs em bibliotecas de mop e *hen1* de selvagens e mutantes podem ser observadas também nas figuras 14B e 14C. As análises com *O. sativa*, mostraram na figura 15A, 15C e 15D, uma precipitação de sRNAs de 24 nucleotídeos nas bibliotecas para AGO4-IP, o que pode ser explicado pelo fato de siRNAs e AGO4 participarem em conjunto do silenciamento gênico no complexo RISC, pelo mecanismo RdDM (XIE; YU, 2015). Nas figuras 15A, 15B e 15D, observamos menor abundância de sRNAs nas bibliotecas de mutantes para DCL3, confirmando que DCL3 participa ativamente da via de síntese destes sRNAs (SUNG et al., 2008).

Em relação à análise realizada utilizando o inibidor de metilação 5-Azac, foi observado um menor desenvolvimento de parte aérea tanto em plântulas de milho, como de cana-de-açúcar (Apêndice A). Isso foi observado em raízes de milho (Figura 20). A diferença na biomassa das plântulas tratadas com 5-Azac em relação ao tratamento com H_2O, não foi significativa.

Os resultados fenotípicos observados tanto no experimento utilizando milho (Figura 17), como no experimento utilizando cana-de-açúcar (Apêndice A, Figura 21) vão de acordo com outros trabalhos relatados na literatura. Em Arabidopsis, o tratamento com 5-Azac alterou o tamanho de raízes (VIRDI et al., 2015) Em *S. bicolor* um menor desenvolvimento de raízes e parte aérea em plântulas de quatro dias de crescimento também foi observado (TURCO et al., 2017) . (LU et al., 2016) observou um menor desenvolvimento de plântulas de arroz tratadas com 5-Azac, mas que a combinação de 5-Azac com o herbicida Atrazina, promoveu a recuperação de plântulas tratadas apenas com o herbicida. Outros autores relatam a alteração na metilação global do genoma, causado por 5-Azac. YANG et al., 2010 relata que a alteração global da metilação, causou decréscimo de metilação na H4K9 e um aumento de acetilação na histona quatro (H4), durante a mitose de células meristemáticas de raízes de milho. (GRZYBKOWSKA et al., 2018) observou que 5-Azac, afeta a expressão de genes envolvidos na embriogênese somática de Arabidopsis, um resultado semelhante ao obtido por (SOLÍS et al., 2015) trabalhando com embriões de cevada e por (BETEKHTIN et al., 2018) trabalhando com calos embriogênicos de *B. distachyon*.

Em relação a expressão de ROS1 mostrada na figura 20, a menor expressão de ROS1 no tratamento com 5-Azac pode ser explicada por (LEI et al., 2015), que diz que ROS1 atua desfazendo a metilação realizada por MET1. Sendo assim, o resultado visto no qPCR, corrobora a ideia de que não havendo atividade de MET1, também haverá baixos níveis de expressão de ROS1 em milho.

6. Conclusão

- Foram anotados 251 genes envolvidos na via de metilação e desmetilação do DNA e vias de pequenos RNAs em 14 plantas diferentes, sendo que 27 genes foram ineditamente identificados em *S. spontaneum.*
- A filogenia demonstrou que os grupos das desmetilases e das Argonautas sofreram uma redução do número de famílias em monocotiledôneas em comparação a Arabidopsis. Já as metilases tiveram duplicações de seus efetores em quase todas as plantas analisadas. Dentro do grupo das Dicer-likes, foi observada uma duplicação de DCL3 em monocotiledôneas.
- A análise de estrutura gênica revelou que os genes MET1, ROS1, DCL3 e AGO4 de milho e sorgo são maiores, com íntrons mais longos, comparados com os genes de Arabidopsis.
- Foram identificadas possíveis regiões MEMS reguladoras de ROS1 nos genomas de *Z. mays e O. sativa*, através de análise *in sílico*.
- O experimento utilizando o inibidor de metilação 5-Azac demonstrou que a metilação do DNA possui papel primordial no desenvolvimento de plântulas de milho e cana-de-açúcar e que ROS1 pode ser utilizada com um marcador de metilação global no genoma de milho.

7. Referências

AGIUS, F.; KAPOOR, A.; ZHU, J. Role of the Arabidopsis DNA glycosylase ¯lyase ROS1 in active DNA demethylation. **PNAS**, v. 103, n. 31, 2006.

ANA, L. The role of DNA (de) methylation in immune responsiveness of Arabidopsis. **The plant journal**, p. 361–374, 2016.

BAI, M. et al. Genome-wide identi fi cation of Dicer-like , Argonaute and RNA-dependent RNA polymerase gene families and their expression analyses in response to viral infection and abiotic stresses in Solanum lycopersicum. **Gene**, v. 501, n. 1, p. 52–62, 2012.

BASSO, M. F. et al. MicroRNAs and new biotechnological tools for its modulation and improving stress tolerance in plants. **Plant biotechnology journal**, p. 1–19, 2019.

BERR, A.; SHA, S.; SHEN, W. Biochimica et Biophysica Acta Histone modi fi cations in transcriptional activation during plant development. **BBA - Gene Regulatory Mechanisms**, v. 1809, p. 567–576, 2011.

BETEKHTIN, A. et al. 5-Azacitidine induces cell death in a tissue culture of brachypodium distachyon. **International Journal of Molecular Sciences**, v. 19, n. 6, 2018.

BURKE, S. V. et al. Evolutionary relationships in Panicoid grasses based on plastome phylogenomics (Panicoideae; Poaceae). **BMC Plant Biology**, v. 16, n. 1, p. 1–11, 2016.

CALCAGNO, T. et al. Identification and Characterization of Shared Duplications between Rice and Wheat Provide New Insight into Grass Genome Evolution. **Plant cell**, v. 20, n. January, p. 11–24, 2008.

CHANDLER, V. L. Paramutation: From Maize to Mice. **Cell**, v. 128, n. 4, p. 641–645, 2007.

CHEAVEGATTI-GIANOTTO, A. et al. Sugarcane (Saccharum X officinarum): A Reference Study for the Regulation of Genetically Modified Cultivars in Brazil. **Tropical Plant Biology**, p. 62–89, 2011.

CHEMICALS, R. letters to nature. v. 391, n. February, p. 806–811, 1998.

CHENDRIMADA, T. P. et al. TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. **Nature**, v. 436, n. 7051, p. 740–744, 2005.

CHOI, Y. et al. DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in Arabidopsis. **Cell**, v. 110, n. 1, p. 33–42, 2002.

CONAB. Acompanhamento da safra brasileira. **Companhia Nacional de Abastecimento**, 2018.

D'HONT. Characterisation of the double genome structure of modern sugarcane cultivars (Saccharum spp .) by molecular cytogenetics. **Molecular and General Genetics**, p. 405–413, 1996.

DOEBLEY, J. T g m e. **Annual Review of Genetics**, p. 37–59, 2004.

DOMCKE, S. et al. transcription factors determines binding of NRF1. **Nature**, 2015.

DORWEILER, J. E. et al. Mediator of paramutation 1 is required for establishment and maintenance of paramutation at multiple maize loci. **Plant Cell**, v. 12, n. 11, p. 2101–2118, 2000.

DU, J. et al. Mechanism of DNA methylation-directed histone methylation by KRYPTONITE. **Molecular Cell**, v. 55, n. 3, p. 495–504, 2014.

DUNOYER, P.; VOINNET, O. The complex interplay between plant viruses and host RNA-silencing pathways. **Current Opinion in Plant Biology**, v. 8, n. 4, p. 415–423, 2005.

FAOSTAT. **Mundo 1995-2016**.

FESCHOTTE, C.; JIANG, N.; WESSLER, S. R. Plant transposable elements: where genetics meets genomics. **Nature Reviews Genetics**, v. 3, n. May, p. 329–341, 2002.

GARSMEUR, O. et al. A mosaic monoploid reference sequence for the highly complex genome of sugarcane. **Nature Communications**, v. 9, n. 1, 2018.

GAUT, B. S. Evolutionary dynamics of grass genomes. **New phytologist**, n. 132, p. 15–28, 2002.

GRATIVOL, C. et al. Biochimica et Biophysica Acta Genetic and epigenetic regulation of stress responses in natural plant populations ☆. **BBA - Gene Regulatory Mechanisms**, v. 1819, n. 2, p. 176–185, 2012.

GRZYBKOWSKA, D. et al. Azacitidine (5-AzaC)-treatment and mutations in DNA methylase genes affect embryogenic response and expression of the genes that are involved in somatic embryogenesis in Arabidopsis. **Plant Growth Regulation**, v. 85, n. 2, p. 243–256, 2018.

HALL, T. M. T.; CAROLINA, N. Structure and Function of Argonaute Proteins. **Structure**, v. 13, p. 1403–1408, 2005.

HALLMANN, A. Transgenic Plant Journal ©2007 Global Science Books Algal Transgenics and Biotechnology. **Transgenic Plant Journal**, v. 1, n. 1, p. 81–98, 2007.

HE, X. et al. An Effector of RNA-Directed DNA Methylation in Arabidopsis Is an ARGONAUTE 4- and RNA-Binding Protein. **Cell**, v. 137, n. 3, p. 498–508, 2009.

HUH, J. H. et al. Review Cellular Programming of Plant Gene Imprinting. **Cell**, v. 1, p. 735–744, 2008.

IBARRA, C. A. et al. Active DNA Demethylation in Plant Companion Cells Reinforces Transposon Methylation in Gametes. **Science**, v. 337, n. 6100, p. 1360–1364, 2012.

JAISWAL, D. et al. Brazilian sugarcane ethanol as an expandable green alternative to crude oil use. **Nature Climate Change**, v. 7, n. 11, p. 788–792, 2017.

JIN, B.; LI, Y.; ROBERTSON, K. D. DNA methylation: Superior or subordinate in the epigenetic hierarchy? **Genes and Cancer**, v. 2, n. 6, p. 607–617, 2011.

JONES, D. T.; TAYLOR, W. R.; THORNTON, J. M. The rapid generation of mutation data matrices. **Bioinformatics**, v. 8, n. 3, p. 275–282, 1992.

KANKEL, M. W. et al. Arabidopsis MET1 Cytosine Methyltransferase Mutants.

Genetics, v. 1122, n. March, p. 1109–1122, 2003.

KAPOOR, M. et al. Genome-wide identification, organization and phylogenetic analysis of Dicer-like, Argonaute and RNA-dependent RNA Polymerase gene families and their expression analysis during reproductive development and stress in rice. **BMC Genomics**, v. 17, p. 1–17, 2008.

KELLOGG, E. A. Update on Evolution Evolutionary History of the Grasses 1. **Plant Physiology**, v. 125, n. 3, p. 1198, 2001.

KENNETH H. WOLFE, MANOLO GOUY, YAU-WEN YANG, PAUL M. SHARP, A. W.-H. L. Date of the monocot-dicot divergence estimated from chloroplast DNA sequence data. **PNAS**, v. 86, n. August, p. 6201–6205, 1989.

KISTLER, L. et al. Multiproxy evidence highlights a complex evolutionary legacy of maize in South America. **Science**, v. 1313, n. 6420, p. 1309–1313, 2018.

KUMAR, S.; STECHER, G.; TAMURA, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. **Molecular biology and evolution**, v. 33, n. 7, p. 1870–1874, 2016.

LAFAYETTE, W. Lant etrotransposons. **Annual Review of Genetics**, v. 33, p. 479–532, 1999.

LANG, Z. et al. Critical roles of DNA demethylation in the activation of ripening-induced genes and inhibition of ripening-repressed genes in tomato fruit. **PNAS**, v. 114, p. 4511–4519, 2017.

LAROCHE ET AL. **mitochondrial dna and monocotdicot divergence time**, 1995.

LAW, J. A.; JACOBSEN, S. E. Patterns in Plants and Animals. **Nature Reviews Genetics**, v. 11, n. 3, p. 204–220, 2010.

LE, T. et al. DNA demethylases target promoter transposable elements to positively regulate stress responsive genes in Arabidopsis. **Genome Biology**, v. 15, n. 458, p. 1–18, 2014.

LEI, M. et al. Regulatory link between DNA methylation and active demethylation in Arabidopsis. **PNAS**, v. 112, n. 11, p. 3553–3557, 2015.

LIN, Y. et al. Validation of potential reference genes for qPCR in maize across abiotic stresses, hormone treatments, and tissue types. **PLoS ONE**, v. 9, n. 5, 2014.

LINGEL, A. et al. Structure and nucleic-acid binding of the Drosophila Argonaute 2 PAZ domain. **Nature**, v. 426, p. 465–469, 2003.

LIU, X. et al. Identification of RNA silencing components in soybean and sorghum. **BMC Bioinformatics**, v. 15, n. 1, p. 1–13, 2014.

LU, Y. et al. Involvement of MEM1 in DNA demethylation in Arabidopsis. **Plant Molecular Biology**, n. 0123456789, 2020.

LU, Y. C. et al. Genome-wide identification of DNA methylation provides insights into the association of gene expression in rice exposed to pesticide atrazine. **Scientific Reports**, v. 6, n. August 2015, p. 1–15, 2016.

MARGIS, R. et al. The evolution and diversification of Dicers in plants. **Febs letters**, v. 580, n. 10, p. 2442–2450, 2006.

MATHIEU, O. et al. Transgenerational Stability of the Arabidopsis Epigenome Is Coordinated by CG Methylation. **Cell**, p. 851–862, 2007.

MCCLINTOCK, B. CHROMOSOME ORGANIZATION AND GENIC EXPRESSION. **CSH Symposia**, p. 13–47, 1951.

MUIR, J. P.; PITMAN, W. D.; FOSTER, J. L. Sustainable, low-input, warm-season, grass-legume grassland mixtures: Mission (nearly) impossible? **Grass and Forage Science**, v. 66, n. 3, p. 301–315, 2011.

NAGYMIHÁLY, M. et al. Ploidy-dependent changes in the epigenome of symbiotic cells correlate with specific patterns of gene expression. **PNAS**, v. 114, n. 17, p. 4543–4548, 2017.

NAKANO, M. et al. Next-Gen Sequence Databases : RNA and Genomic Informatics Resources for Plants. **Plant Physiology**, 2019.

NELSON, G. C. et al. Climate change effects on agriculture: Economic responses to biophysical shocks. **Proceedings of the National Academy of Sciences of the United States of America**, v. 111, n. 9, p. 3274–3279, 2014.

NOBUTA, K. et al. Distinct size distribution of endogenous siRNAs in maize: Evidence from deep sequencing in the mop1-1 mutant. **Proceedings of the National Academy of Sciences of the United States of America**, v. 105, n. 39, p. 14958–14963, 2008.

NOHARA, D. et al. Dependency of climate change and carbon cycle on CO2 emission pathways. **Environmental Research Letters**, v. 8, n. 1, 2013.

PARANHOS, P. O Açúcar No Norte Fluminense. **Arquivo público do estado de São Paulo**, 2008.

PEI, L. et al. Tracing the origin and evolution history of methylation-related genes in plants. **BMC Plant Biology**, v. 19, n. 1, p. 1–13, 2019.

PIKAARD, C. S.; SCHEID, O. M. Epigenetic Regulation in Plants. **CSH Perspectives in Biology**, p. 1–31, 2014.

RAMBANI, A. et al. The Methylome of Soybean Roots during the Compatible Interaction with the Soybean Cyst Nematode 1. **Plant Physiology**, v. 168, n. August, p. 1364–1377, 2015.

RONALD, P. Plant genetics, sustainable agriculture and global food security. **Genetics**, v. 188, n. 1, p. 11–20, 2011.

SATGÉ, C. et al. Reprogramming of DNA methylation is critical for nodule development in Medicago truncatula. **Nature Plants**, v. 2, n. October, 2016.

SCHÜTZ, H.; JANSEN, M.; VERHOFF, M. A. How to Feed the World in 2050. **FAO**, v. 228, n. 5–6, p. 151–159, 2011.

SECCO, D. et al. Stress induced gene expression drives transient DNA methylation changes at adjacent repetitive elements. **Plant Biology**, p. 1–26, 2015.

SLEDZ, C. A.; WILLIAMS, B. R. G. RNA interference in biology and disease. **Blood**, v. 106, n. 3, p. 787–795, 2005.

SLOTKIN, R. K. et al. Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. **Cell**, v. 136, n. 3, p. 461–472, 2010.

SOLÍS, M. T. et al. 5-azacytidine promotes microspore embryogenesis initiation by decreasing global DNA methylation, but prevents subsequent embryo development in rapeseed and barley. **Frontiers in Plant Science**, v. 6, n. JUNE, p. 1–17, 2015.

SONG, J. et al. The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. **Nature structural e molecular biology**, v. 10, n. 12, p. 1026–1032, 2003.

SONG, J. et al. Structure of DNMT1-DNA Complex Reveals a Role for Autoinhibition in Maintenance DNA Methylation. **Science**, v. 331, n. 6020, p. 1036–1040, 2011.

SORENG, R. J. et al. Review A worldwide phylogenetic classi fi cation of the Poaceae (Gramineae) II : An update and a comparison of two 2015 classi fi cations. **Journal of systematics and evolution**, 2017.

SOUZA, G. M. et al. Assembly of the 373k gene space of the polyploid sugarcane genome reveals reservoirs of functional diversity in the world ' s leading biomass crop. **Giga science**, v. 8, n. 12, p. 1–18, 2019.

SUNG, H. C. et al. Physcomitrella patens DCL3 is required for 22-24 nt siRNA accumulation, suppression of retrotransposon-derived transcripts, and normal development. **PLoS Genetics**, v. 4, n. 12, 2008.

TARIQ, M.; PASZKOWSKI, J. DNA and histone methylation in plants. **Trends in Genetics**, v. 20, n. 6, 2004.

TELIAS, A. et al. Apple skin patterning is associated with differential expression of MYB10 Apple skin patterning is associated with differential expression of MYB10. **BMC Plant Biology**, v. 93, n. May, 2011.

THAO, K.; HARMER, S. L.; ZILBERMAN, D. The Arabidopsis Nucleosome Remodeler DDM1 Allows DNA Methyltransferases to Access H1-Containing Heterochromatin. **USDA**, v. 153, n. 1, p. 193–205, 2014.

TURCO, G. M. et al. DNA methylation and gene expression regulation associated with vascularization in Sorghum bicolor. **New Phytologist**, v. 214, n. 3, p. 1213–1229, 2017.

VIRDI, K. S. et al. Arabidopsis MSH1 mutation alters the epigenome and produces heritable changes in plant growth. **Nature Communications**, v. 6, p. 1–9, 2015.

WEINHOLD, B. Epigenetics: the science of change. **Environmental health perspectives.**, v. 114, n. 3, p. 160–167, 2006.

WILLIAMS, B. P. et al. Methylation-Sensitive Expression of a DNA Demethylase Gene Serves As an Epigenetic Rheostat. **PLoS Genetics**, v. 11, n. 3, p. 1–18, 2015.

XIE, M.; YU, B. siRNA-directed DNA Methylation in Plants. **Current Genomics**, v. 16, n. 1, p. 23–31, 2015.

XIE, Z. et al. Genetic and Functional Diversification of Small RNA Pathways in Plants. **Plant Biology**, v. 2, n. 5, p. 642–652, 2004.

YANG, F. et al. Trichostatin A and 5-azacytidine both cause an increase in global histone H4 acetylation and a decrease in global DNA and H3K9 methylation during mitosis in maize. **Plant Biology**, v. 10, 2010.

YANG, Z. et al. HEN1 recognizes 21 – 24 nt small RNA duplexes and deposits a

methyl group onto the 2 0 OH of the 3 0 terminal nucleotide. **Nucleic acids research**, v. 34, n. 2, p. 667–675, 2006.

ZEMACH, A. et al. Genome-wide evolutionary analysis of eukaryotic DNA methylation. **Science**, v. 328, n. 5980, p. 916–919, 2010.

ZHANG, B. et al. Chilling-induced tomato flavor loss is associated with altered volatile synthesis and transient changes in DNA methylation. **PNAS**, v. 113, n. 44, 2016.

ZHANG, H.; LANG, Z.; ZHU, J. Dynamics and function of DNA methylation in plants. **Nature Reviews Molecular Cell Biology**, 2018.

ZHANG, J. et al. Allele-defined genome of the autopolyploid sugarcane Saccharum spontaneum L. **Nature Genetics**, v. 50, n. 11, p. 1565–1573, 2018.

ZHAO, Y. et al. Analysis of DNA Methylation in Cotton Hybrids and Their Parents 1. **Molecular Biology**, v. 42, n. 2, p. 169–178, 2008.

ZIMIN, A. V et al. The first near-complete assembly of the hexaploid bread wheat genome , Triticum aestivum. **Giga science**, v. 6, n. 11, p. 1–7, 2017.

Apêndice A

As plântulas de cana-de-açúcar foram obtidas de *explants* da cultivar SP803280, por meio de cultura de tecido e crescidas em casa de vegetação por aproximadamente 45 dias. Posteriormente, os calos embriogênicos foram replicados e transferidos para frascos com 40 ml de meio MS tratado com o inibidor de metilação 5-Azac na concentração final de 100µM e controle utilizando água. A figura 20 mostra o fenótipo dessas plântulas nos diferentes tratamentos. As plântulas que receberam o tratamento com 5-Azac apresentaram parte aérea reduzida e um menor desenvolvimento das raízes, mas não foram dados significativos.

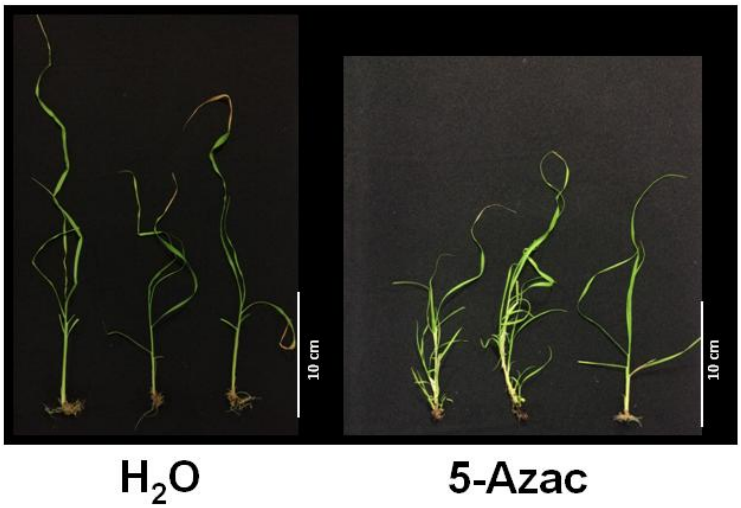


Figura 20: Avaliação fenotípica do desenvolvimento de plântulas de cana-de-açúcar com inibidor de metilação (5-Azac) e H_2O.

A figura 22 mostra os valores estatísticos que as plântulas das duas condições apresentaram em relação à parte aérea. Comprimento (A), massa fresca (B), massa seca (C), biomassa, (D) e conteúdo relativo de água (E). Em todos os gráficos, foi aplicado o TEST T, para avaliar significância estatística entre as amostras. Nenhum gráfico apresentou diferença estatística.

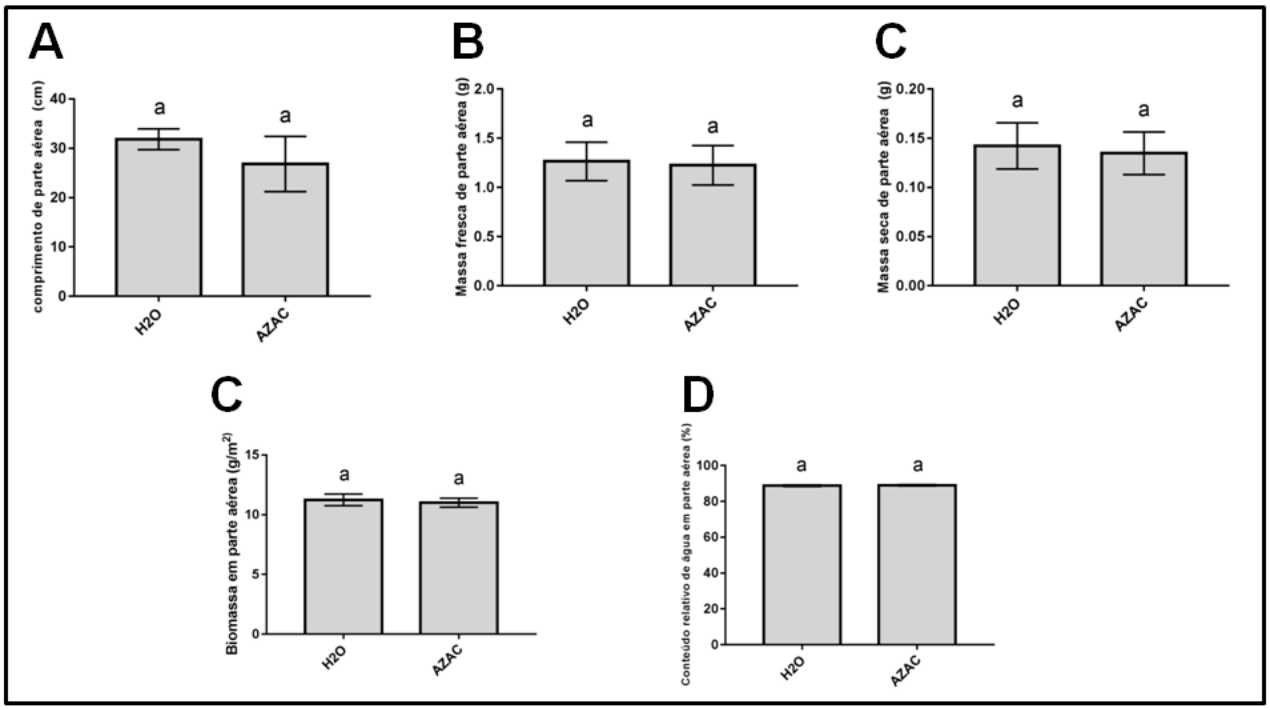


Figura 21: : Representação estatísticas do desenvolvimento da parte aérea de plântulas de cana-de-açúcar sob condições de controle e tratamento com inibidor de metilação 5-AzaC. Em (A) comprimento em (cm), em (B) massa fresca em (g), em (C) massa seca em (g), em (D) biomassa em (g/m^2) e em (E) conteúdo relativo de água em (%). O símbolo (a) significa que os gráficos não possuem diferença estatística.

A figura 22 mostra os valores estatísticos que as plântulas das duas condições apresentaram em raízes. Deslocamento de volume a partir de 10 ml (A), massa fresca (B), massa seca (C), biomassa, (D) e conteúdo relativo de água (E). Em todos os gráficos, foi aplicado o TEST T, para avaliar significância estatística entre as amostras. Nenhum gráfico apresentou diferença estatística.

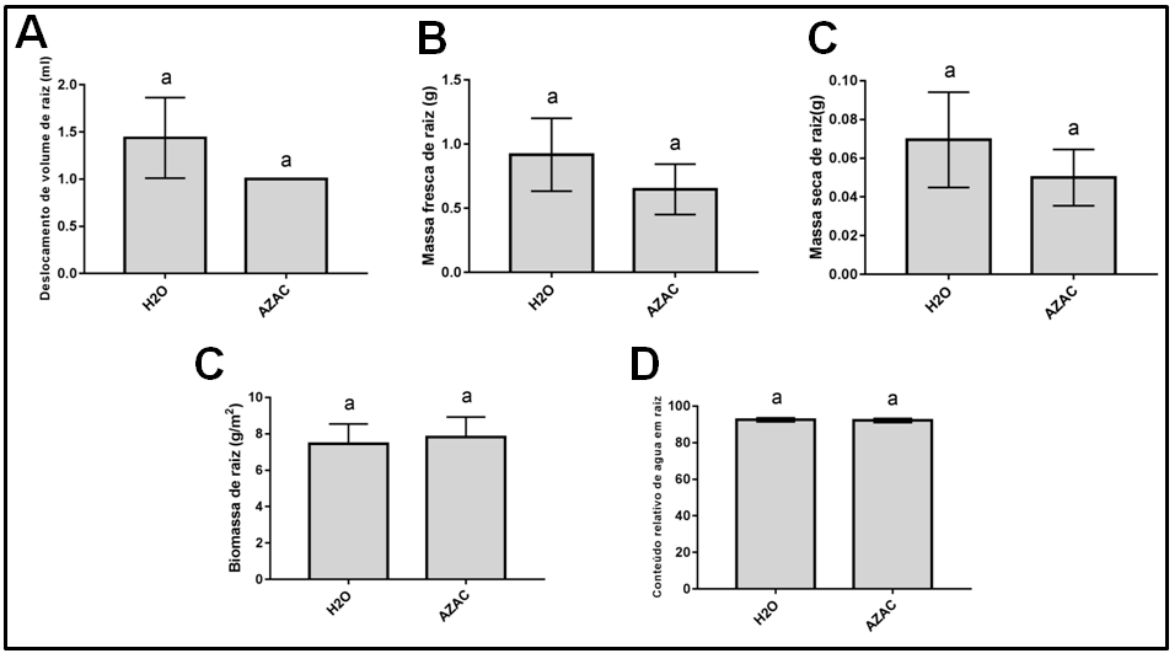


Figura 22: Representação estatísticas do desenvolvimento de raízes de plântulas de cana-de-açúcar sob condições de controle e tratamento com inibidor de metilação 5-AzaC. Em (A) Deslocamento de volume de raiz (ml), em (B) massa fresca em (g), em (C) massa seca em (g), em (D) biomassa em (g/m^2) e em (E) conteúdo relativo de água em (%). O símbolo (a) significa que os gráficos não possuem diferença estatística.