

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE BIOCÊNCIAS
RENORBIO

JORGE LUÍS BANDEIRA DA SILVA FILHO

ANÁLISE DA METILAÇÃO DE DNAs EM ACESSOS DE MANDIOCA (*Manihot esculenta*
C.) SUBMETIDOS AO DÉFICIT HÍDRICO

RECIFE
2024

JORGE LUÍS BANDEIRA DA SILVA FILHO

**ANÁLISE DA METILAÇÃO DE DNAs EM ACESSOS DE MANDIOCA (*Manihot esculenta*
C.) SUBMETIDOS AO DÉFICIT HÍDRICO**

Tese apresentada ao Programa de Pós-Graduação em biotecnologia - RENORBIO, na Universidade Federal de Pernambuco, como requisito para obtenção do título de Doutor em Biotecnologia.

Área de concentração:
Biotecnologia

Orientador: Prof^o. Dr. Ederson Akio Kido

RECIFE
2024

.Catalogação de Publicação na Fonte. UFPE - Biblioteca Central

Filho, Jorge Luis Bandeira da Silva.

Análise da metilação de DNAs em acessos de mandioca (*Manihot esculenta* c.) submetidos ao déficit hídrico / Jorge Luis Bandeira da Silva Filho. - Recife, 2024.

151f.: il.

Tese (Doutorado) - Universidade Federal de Pernambuco, Centro de Biociências, Programa de Pós-Graduação em Biotecnologia, 2024.

Orientação: Ederson AKio Kido.

Inclui referências.

1. Mandioca; 2. Metilação; 3. Genética. I. Kido, Ederson AKio. II. Título.

UFPE-Biblioteca Central

JORGE LUÍS BANDEIRA DA SILVA FILHO

**ANÁLISE DA METILAÇÃO DE DNAs EM ACESSOS DE MANDIOCA (*Manihot esculenta*
C.) SUBMETIDOS AO DÉFICIT HÍDRICO**

Tese apresentada ao Programa de Pós-Graduação em biotecnologia - RENORBIO, na Universidade Federal de Pernambuco, como requisito para obtenção do título de Doutor em Biotecnologia.

Área de concentração:
Biotecnologia

Aprovada em: 28/02/2024.

BANCA EXAMINADORA

Prof^o. Dr. Valdir de Queiroz Balbino/ Universidade Federal de Pernambuco

Prof^o. Dr. Antonio Félix da Costa / Instituto Agrônômico de Pernambuco

Dr. Wilson José da Silva Junior / MGI - Complete Genomics

Prof^a Dr^a. Flavia Figueira Aburjaile / Universidade Federal de Minas Gerais

RESUMO

A metilação do DNA desempenha um papel fundamental no desenvolvimento e nas respostas das plantas a vários estresses chamando a atenção de melhoristas sobre este tema. Este trabalho teve como objetivo avaliar a dinâmica da metilação do DNA em genótipos de mandioca contrastantes para tolerância ao déficit hídrico. As variedades BRS Formosa (mais tolerante à seca) e BRS Dourada (menos tolerante à seca) foram cultivadas em casa de vegetação por 50 dias, e após isso, foi suspensa a irrigação. Os tratamentos consistiram de plantas não estressadas (controle negativo) e plantas estressadas (déficit hídrico) e três repetições. As amostras de DNA de cada variedade e tratamento foram usadas para gerar 12 bibliotecas (duas variedades de mandioca, dois tratamentos, e três réplicas) e submetidos a análise *MethylRAD-Seq*. Os dados sequenciados revelaram sítios metilados cobrindo 18 a 21% do genoma de *M. esculenta*, dependendo da variedade e tratamento. Os sítios metilados CCGG foram mapeados principalmente em regiões intergênicas, éxons e íntrons, enquanto os sítios CCNGG foram mapeados principalmente intergênicos, *upstream*, íntrons, e regiões de éxons, permanecendo menos detectados nas regiões *UTR*. Em ambos os casos, os sítios metilados nas *UTRs* foram menos detectados. A análise dos sítios diferencialmente metilados indicou que as variedades apresentaram perfis de metilação distintos, pois apenas 12% dos sítios (CCGG e CCNGG) foram metilados em ambas as variedades. Termos enriquecidos de ontologia gênica indicaram a resposta imediata da variedade tolerante à seca ao estresse, enquanto a variedade sensível parece sofrer mais danos potenciais do estresse. As redes de interação proteína-proteína previstas reforçaram tais perfis. Além disso, os genomas das duas variedades revelaram eventos *SNPs/INDELS* cobrindo genes destacados pelas redes de interação de proteínas. Desta forma, estes dados podem ser úteis em decifrar os papéis da metilação do DNA nas respostas de tolerância à seca da mandioca e na adaptação a estresses abióticos.

Palavras-chave: Estresse abiótico. Metiloma. Expressão gênica. Desmetilação. Estresse hídrico.

ABSTRACT

DNA methylation plays a fundamental role in plant development and responses to various stresses, attracting the attention of breeders on this topic. This work aimed to evaluate the dynamics of DNA methylation in cassava genotypes contrasting for water deficit tolerance. The BRS Formosa (drought-tolerant) and BRS Dourada (drought-sensitive) varieties were grown in a greenhouse for 50 days, after which irrigation was suspended. Treatments consisted of non-stressed plants (negative control) and stressed plants (water deficit) and three replications. DNA samples from each variety and treatment were used to generate 12 libraries (two cassava varieties, two treatments, and three replicates) and subjected to MethylRAD-Seq analysis. The sequenced data revealed methylated sites covering 18 to 21% of the *M. esculenta* genome, depending on variety and treatment. CCGG methylated sites were mapped mainly in intergenic, exon and intron regions, while CCNGG sites were mapped mainly in intergenic, upstream, intron, and exon regions, remaining less detected in UTR regions. In both cases, the methylated sites in the UTRs were less detected. The analysis of differentially methylated sites indicated that the varieties had different methylation profiles, as only 12% of the sites (CCGG and CCNGG) were methylated in both varieties. Gene ontology enriched terms indicated the immediate response of the drought tolerant variety to stress, while the sensitive variety seems to suffer more potential damage from stress. The predicted protein-protein interaction networks reinforced such profiles. Furthermore, the genomes of the two varieties revealed SNPs/INDELs events covering genes highlighted by the interactomes. Thus, these data may be useful in deciphering the roles of DNA methylation in cassava drought tolerance responses and adaptation to abiotic stresses.

Keywords: Abiotic stress. Methyloma. Gene expression. Demethylation. Hydric stress.

Sumário

1. Fundamentação teórica.....	9
1.1. A espécie <i>Manihot esculenta</i> Crantz	9
1.2. O melhoramento de mandioca	12
1.3. Informações gerais sobre genômica de <i>M. esculenta</i>	14
1.4. Efeitos da seca e salinização dos solos em vegetais	18
1.5. Regulação gênica - epigenética.....	21
1.6. Metodologia para análise de metilação.....	24
1.6.1. Técnica de polimorfismo de amplificação sensível à metilação (MSAP)	
24	
1.6.2. Sequenciamento de bissulfito.....	25
1.6.3. Imunoprecipitação de DNA de metila (MeDIP).....	26
1.6.4. Metodologia MethylRAD-Seq	27
1.7. Visão geral sobre estudos de metiloma em vegetais	30
2. Hipótese	33
3. Objetivos.....	33
4. Metodologia	34
4.1. Ensaio biológico: materiais vegetais, aplicação do estresse e geração das bibliotecas MethylRad-Seq.....	34
4.2. Análise de qualidade do sequenciamento e mapeamento das leituras em genoma de referência.	36
4.3. Análise da distribuição dos sítios de metilação nos elementos gênicos estruturais e funcionais.....	36
4.4. Quantificação relativa da metilação do DNA e análise de distribuição dos sítios de metilação em regiões gênicas	36
4.5. Correlação dos níveis de metilação entre diferentes amostras	37
4.6. Níveis de metilação diferencial.....	37
4.7. Comparações dos níveis de metilação entre contrastes e distribuição dos sítios diferencialmente metilados	37
4.8. Genes diferencialmente metilados exclusivos e comuns entre diferentes contrastes dos tratamentos e variedades.....	38
4.9. Análise de termos de ontologia gênica (GO).....	38
4.10. Análise funcional de potenciais redes de interação de proteínas.....	38
4.11. Análises de SNP-INDEL.....	39
5. Resultados.....	40
5.1. Os totais do sequenciamento das bibliotecas MethylRAD e os percentuais de tags mapeadas no genoma de referência de <i>M. esculenta</i>	40
5.2. As profundidades de sequenciamento das bibliotecas MethylRAD	40

5.3.	A distribuição dos sítios metilados nos elementos gênicos.....	41
5.4.	A distribuição dos sítios de metilação em genes	43
5.5.	Correlação dos níveis de metilação entre diferentes amostras	46
5.6.	Os sítios diferencialmente metilados de cada variedade nas comparações dos tratamentos com e sem o estresse aplicado.....	48
5.7.	Genes diferencialmente metilados exclusivos e comuns entre contrastes.....	51
5.8.	Análise de Termos de GO	52
5.9.	Rede de interações previstas a partir dos conjuntos de genes diferencialmente metilados de cada variedade após o estresse de seca	53
5.9.1.	Rede de interações predita para proteínas de genes hipometilados na variedade tolerante Formosa.....	53
5.9.2.	Rede de interações predita para proteínas de genes hipermetilados na variedade tolerante Formosa.....	56
5.9.3.	Rede de interações predita para proteínas de genes hipometilados na variedade sensível Dourada	59
5.9.4.	Rede de interações predita para proteínas de genes hipometilados na variedade sensível Dourada	62
5.10.	Análise SNP-INDEL.....	65
6.	Discussão.....	66
6.1.	O sequenciamento <i>MethylRad</i> em variedades de <i>M. esculenta</i>	66
6.2.	Os níveis de metilação entre as réplicas biológicas.....	66
6.3.	Os sítios metilados nos elementos genéticos estruturais do genoma de <i>M. esculenta</i>	66
6.4.	Os perfis de metilação das variedades após a aplicação do estresse	67
6.5.	Os termos de ontologia gênica dos conjuntos de genes diferencialmente metilados nas duas variedades.....	67
6.6.	Redes de interações previstas a partir dos conjuntos de genes diferencialmente metilados.....	69
6.6.1.	A rede de interações predita para proteínas de genes hipometilados na variedade tolerante Formosa.....	70
6.6.2.	A rede de interações predita para proteínas de genes hipermetilados na variedade tolerante Formosa.....	73
6.6.3.	A rede de interações predita para proteínas de genes hipometilados na variedade sensível Dourada	75
6.6.4.	A rede de interações predita para proteínas de genes hipermetilados na variedade sensível Dourada	77
7.	Conclusão.....	80
8.	Referencias bibliográficas.....	82

1. Fundamentação teórica

1.1. A espécie *Manihot esculenta* Crantz

A mandioca (*M. esculenta*.) é uma espécie da família Euphorbiaceae, subfamília Crotonoideae, tribo Manihotae, e gênero *Manihot*, com estimadas 97 espécies (ROGERS, APPAN, 1973). O gênero tem duas seções, Arborae, contendo espécies arbóreas, e Fructicosae, compreendendo arbustos de crescimento lento adaptados a pastagens de savana ou condições desérticas (JENNINGS, IGLESIAS, 2001). *Manihot esculenta* Crantz é a espécie economicamente mais importante do gênero, sendo amplamente cultivada para alimentos e aplicações industriais (NASSAR et al., 2006).

Atualmente acredita-se que o gênero é originário da América do Sul e Central, onde existem dois centros de diversidade, um no Brasil e outro no México. A mandioca foi possivelmente um dos primeiros vegetais domesticados nas Américas, fato estimado de ter ocorrido entre 5.000 e 7.000 a.C. (LATHRAP, 1970; FULLER et al, 2011). Originalmente, especulou-se que a mandioca poderia ascender e evoluir por meio da introgressão periódica de genes envolvendo várias espécies selvagens (NASSAR, ORTIZ, 2007). No entanto, mais tarde foi relatado que uma espécie selvagem, *Manihot flabellifolia* (Pohl), foi proposta como possível parente selvagem mais próximo da mandioca (ALLEM, 1994).

Muitos estudos têm apoiado a relação ancestral da mandioca cultivada moderna e a subespécie selvagem *M. esculenta* ssp. *Flabellifolia* (ROA et al., 1997; ROA et al., 2000). Embora resultados controversos tenham sido relatados, a borda sul da bacia amazônica atualmente permanece como o reconhecido centro de origem de *M. esculenta* (NASSAR et al, 2007; LÉOTARD et al, 2009). Consequentemente, esta região é o provável centro primário de diversidade da mandioca (NASSAR, 2000). O processo de domesticação envolveu seleção para tamanho de raiz, hábito de crescimento, número de hastes e capacidade de propagação clonal por meio de estacas de hastes (JENNINGS, 1976). Durante

o século XVI, comerciantes portugueses introduziram a mandioca na África e depois na Ásia também (JENNINGS, 1976).

A cultura representa um elevado valor econômico, pois, atualmente, é um produto da dieta básica para milhões de pessoas que habitam entre os trópicos (PARMAR; STURM; HENSEL, 2017). Suas raízes são uma importante fonte de alimento devido ao acúmulo de nutrientes (COCK, 1985).

A espécie é normalmente propagada vegetativamente através de estacas, embora a propagação sexual via semente possa ocorrer. As plantas possuem flores unissexuadas, monoclamídeas, sendo o único perianto existente denominado de cálice (DE MATTOS; FARIAS; FERREIRA-FILHO, 2006). Superior à inflorescência encontram-se as flores masculinas, enquanto as flores femininas são distribuídas na base. Seu caule é de modelo subarborescente e pode ser indiviso no ciclo vegetativo e ramificado no ciclo reprodutivo (DE MATTOS; FARIAS; FERREIRA-FILHO, 2006). Torna-se lenhoso no estágio adulto da planta adquirindo característica quebradiça. A espécie apresenta folhas simples, inseridas no caule, em disposição alterno-espiralada, lobada, e longamente peciolada (DE MATTOS; FARIAS; FERREIRA-FILHO, 2006). O comprimento deste pecíolo pode variar segundo o tipo de cultivar e o estágio de desenvolvimento da espécie, e a cor pode variar entre verde, vermelha, rosa ou amarelo (Figura 1) (DE MATTOS; FARIAS; FERREIRA-FILHO, 2006).

M. esculenta é uma espécie naturalmente diploide ($2n = 36$) com um tamanho de genoma de 770 Mbp (SAKURAI et al, 2013). Embora a maioria das espécies estudadas contenha 36 cromossomos e sejam consideradas diploides ($2n$), pareamento irregular na meiose e poliploidização espontânea (por exemplo, triploides ($3n$) e tetraploides ($4n$)) foram relatados em mandiocas selvagens e domesticadas (HAHN et al., 1980). Da mesma forma, a presença de três cromossomos nucleolares foi relatada na mandioca, o que é alto para verdadeiros diplóides, sugerindo que as espécies de *Manihot* são provavelmente alotetraploides segmentares com um número básico de cromossomos $x = 9$ (JENNINGS, IGLESIAS, 2001).

Figura 1. Aspectos morfológicos de *Manihot esculenta* Crantz. Caule, folhas, flores e fruto.



Fonte: Domínio público em <https://en.wikipedia.org/wiki/Cassava>

A planta é essencialmente perene, mas as raízes são colhidas anualmente ou bienalmente, sendo o sistema radicular tuberoso responsável pelo armazenamento das reservas de amido (DE MATTOS; FARIAS; FERREIRA-FILHO, 2006).

A mandioca pode ser cultivada em climas tropicais e subtropicais, e se desenvolve de forma eficiente sob exposição direta ao sol, sem tolerar alagamentos (SEBRAE/ESPM, 2008). A cultura é uma das preferidas em regiões africanas devido ao alto rendimento na produção de carboidratos em relação a área de cultivo, além da tolerar solos pobres e secos (MCCALLUM et al., 2017).

Sua importância reflete a indicação da ONU (Organização das Nações Unidas) como o alimento mais importante do século, pois é utilizada para consumo humano, animal e industrial. Na alimentação é explorada na forma de farinha de mandioca, farinha de tapioca, cozida ou frita. Na indústria é importante

em segmentos como o de papel e celulose, de panificação, têxtil, farmacêutico, de cosméticos, de fertilizantes, de petróleo e siderúrgico. Além disso, possui potencial para produção de biocombustível, pois as raízes tuberosas contêm elevado amido (cerca de 70-85% em peso seco), e estes são matéria-prima para produção de etanol. Neste processo, o amido pode ser convertido em glicose por enzimas, e a glicose seria, então, fermentada por leveduras, produzindo o álcool (REDDY et al., 2008).

Os açúcares de sua biomassa são responsáveis pela capacidade de produção de etanol em taxas viáveis, de cunho sustentável (NUWAMANYA et al., 2012). Adicionalmente, considerando as características da espécie, ela se apresenta ideal para cultivo em regiões semiáridas do Nordeste brasileiro, pois o clima se caracteriza por temperaturas médias anuais acima dos 20° C, e precipitações entre 280 a 800 mm, com ocorrência de períodos de déficit hídrico (ARAÚJO et al., 2005).

1.2. O melhoramento de mandioca

Segundo Takeda e Matsuoka (2008), no âmbito da produção de alimentos, a biotecnologia pode fornecer mecanismos para o melhoramento da produção vegetal pela aplicação de conhecimentos moleculares, desde as funções inerentes a genes específicos até o comportamento de vias e redes regulatórias envolvidas na resposta a alguma condição imposta, refletindo no desenvolvimento e crescimento da planta.

Atualmente, com o desenvolvimento da genômica e todas as possíveis ômicas derivadas (transcriptômica, proteômica, metabolômica, etc), as pesquisas para otimizar os programas de melhoramento genético em plantas estarão essencialmente atreladas ao uso de ferramentas de bioinformática. A bioinformática é um campo interdisciplinar que envolve aplicações da informática na análise de informações biológicas (LUSCOMBE et al., 2001). Ela se caracteriza por ser uma área de integração entre a biologia, a ciência da computação, a matemática, e a engenharia, em função das análises e interpretações dos dados biológicos. Desta forma, a partir de dados da biologia molecular, um pesquisador possui o potencial para “otimizar”

variedades/cultivares de uma espécie, utilizando ferramentas modernas, tais como transgenia e edição de genes (LUSCOMBE et al., 2001).

Segundo Valle et al. (2014), nos últimos 40 anos houve uma modernização no processo produtivo da mandioca, que foi adaptado às exigências da agricultura atual em relação às regiões de cultivo. Técnicas agrícolas e industriais foram adicionadas na agroindústria da mandioca, com o surgimento de uma gama de novos produtos (VALLE et al. 2014). A modernização se deu também no desenvolvimento de variedades mais resistentes a doenças e tolerantes a estresses abióticos, além de melhores produtividades (VALLE et al. 2014).

Neste processo, várias instituições colaboraram para a pesquisa e desenvolvimento de novas variedades de mandioca. Exemplos emblemáticos são representados pelo Instituto Agrônomo de Campinas (IAC) e a Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA). O IAC possui o programa de melhoramento em mandioca mais antigo, iniciado no final do século 19 (VALLE, LORENZI, 2014). Nos anos de 1960, o instituto produziu a variedade IAC 24-2 Mantiqueira, com relativo alto desempenho de produção, além de variedades como IAC-12, IAC-13 e IAC-14, que apresentavam maior resistência a estresses bióticos, e melhor produtividade em regiões específicas (VALLE, LORENZI, 2014). Por outro lado, a EMBRAPA foi responsável pelo desenvolvimento de variedades como BRS CS 01, entre os anos de 2010 e 2014, uma cultivar de mandioca para a indústria de alimentos, recomendada para cultivo nas regiões Sul/Sudeste, além da BRS 398 e da BRS 366, que têm como principais características um maior potencial produtivo, maior precocidade, maior resistência a bacterioses, e ao superalongamento, requerendo menor tempo para cozimento das raízes (EMBRAPA, 2020).

A EMBRAPA Mandioca e Fruticultura, em Cruz das Almas (BA), instituída em junho de 1975, tem como objetivo executar pesquisas para o melhoramento genético da mandioca, e de determinadas frutas tropicais (banana, abacaxi, citros, maracujá e mamão). Assim, os programas nacionais de melhoramento da mandioca representam esforços contínuos no lançamento de novos materiais economicamente viáveis nas mais variadas condições de cultivo do país.

Duas variedades de mandioca que são importantes para o cultivo no nordeste brasileiro são as variedades BRS-Formosa e BRS-Dourada desenvolvidas pela EMBRAPA. A variedade BRS-Formosa, entregue aos agricultores em 2003 (EMBRAPA, 2016), apresenta expressivo desempenho agrônômico devido à tolerância ao estresse hídrico e produção de biomassa (54 t.ha⁻¹ de raízes/ano), enquanto a variedade Dourada tem atributos competitivos para consumo, como a alta qualidade das raízes exigindo menor tempo de cozimento (EMBRAPA, 2021).

BRS Dourada é uma variedade local de pedigree desconhecido, desenvolvida especificamente para consumo *in natura*. Suas raízes apresentam coloração amarela e contém menos de 100 ppm de compostos cianogênicos, tornando-o ideal para uso culinário. Em contrapartida, a BRS Formosa é um híbrido de mandioca industrial, conhecido por sua maior concentração de compostos cianogênicos (acima de 100 ppm), o que confere sabor amargo. O progenitor feminino da BRS Formosa é o acesso BGM-0361, enquanto o progenitor masculino é desconhecido. Estas duas variedades apresentam diferenças genéticas significativas, confirmadas por análise molecular através de técnicas de genotipagem por sequenciamento (OGBONNA et al., 2021)

Com a intensificação das alterações climáticas, os estresses abióticos, incluindo a seca, deverão ser potenciados, comprometendo o rendimento das culturas e os terrenos agrícolas em todo o planeta (CHANGE, 2001), por isso o desenvolvimento de novas variedades e aperfeiçoamento das variedades correntes é de suma importância para a continuação da produção eficiente deste vegetal frente aos desafios futuros como a salinização de solos.

1.3. Informações gerais sobre genômica de *M. esculenta*

O projeto do genoma da mandioca se baseou em um piloto iniciado por meio do *DOE-JGI Community Sequencing Program* (CSP) por um consórcio de 14 membros liderado por Claude Fauquet, Joe Tohme e Pablo Rabinowicz. Este projeto piloto produziu uma cobertura um pouco abaixo de 1x de mais de 700.000 leituras Sanger usando bibliotecas de plasmídeos e fosmídeos, e forneceu informações sobre as características gerais do genoma da mandioca, tendo gerado uma fonte valiosa de sequências pareadas de *Sanger* para serem usadas

posteriormente (BREDESON et al., 2016). Em seguida, uma montagem *454-based* foi gerada (v4.1) em um projeto liderado por Steve Rounsley, Dan Rokhsar, Chinnappa Kodira e Tim Harkins. Esta montagem utilizou a plataforma *Genome Sequencer FLX Titanium* e uma estratégia *Whole Genome Shotgun* (WGS) (PROCHNIK et al., 2012). Nesta abordagem, o DNA genômico nuclear é isolado e dividido aleatoriamente em fragmentos curtos, que são então classificados em caixas de vários tamanhos. As extremidades desses fragmentos são sequenciadas e remontadas computacionalmente para reconstruir longos trechos da sequência genômica. As sequências montadas são chamadas de “*contigs*” se representarem trechos ininterruptos do genoma, e “*scaffolds*” se contiverem uma ou mais lacunas devido a sequências não amostradas ou regiões repetitivas que são difíceis de reconstruir a partir de leituras de sequências curtas. Essas lacunas são representadas como execuções do nucleotídeo desconhecido “N”, com o número de Ns correspondendo aproximadamente ao comprimento da lacuna (PROCHNIK et al., 2012). A partir da montagem v4 foi possível anotar mais de 1/3 das regiões do genoma de *M. esculenta* como repetitivo, expandindo dramaticamente o catálogo de elementos transponíveis da mandioca (PROCHNIK et al., 2012).

A versão 4.1 foi importante por expandir a gama de informações genômicas de *M. esculenta* de forma considerável. Prochhnik e colaboradores (2012) relatam que as 80.459 *ESTs* Sanger do *Genbank* que haviam até então para a espécie, foram aumentados por um novo conjunto de 2,7 milhões de leituras de bibliotecas de folhas e raízes, geradas por *454 Life Sciences* usando a plataforma *FLX Titanium*. Embora metade das leituras de *ESTs* foliares representassem transcrições de DNA ribossômico e de cloroplasto, e não úteis para anotar o genoma nuclear, o restante das 1.4 milhões de leituras foram usadas para melhorar a predição genética. Além desses dados de RNA mensageiro, sequências proteicas previstas de mamona, *Arabidopsis*, arroz, soja e *Populus* foram mapeadas na montagem da mandioca para ajudar a prever *loci* genéticos. Neste momento haviam 30.666 *loci* codificadores de proteínas na mandioca combinando *ESTs*, homologia de peptídeos com outras espécies, e os padrões de sequência estatística comuns aos genes de plantas (por exemplo, sinais de *splice* intrônicos). Mais de um terço (11.526) dos *loci* codificadores de

proteínas são apoiados por *ESTs*, que também forneciam evidências de um total de 3.485 formas alternativas de *splicing*.

A montagem do genoma da mandioca v4 foi lançado em 2009 (PROCHNIK et al., 2012) e está disponível publicamente através da plataforma V13 Phytozome (GOODSTEIN et al., 2012). Esta montagem contemplou um total de 532,5 Mb (PROCHNIK et al., 2012)

A montagem v6 é uma montagem baseada em Illumina do genótipo AM560-2. Bibliotecas de fragmentos curtos foram geradas sequenciadas com profundidade de 125X. Bibliotecas de pares de inserção de 6 kbp foram construídas e sequenciadas para sequenciamento final de fosmídeos e de longo alcance (Hi-Cos) (PHYTOZOME , 2024a).

As informações de sequência de uma ampla variedade de tamanhos de leituras foram montadas *de novo*. Os *scaffolds* resultantes foram ordenados e orientados mapeando 22.403 marcadores que haviam sido gerados anteriormente (PHYTOZOME, 2024a). A grande maioria da montagem (518,51 Mbp de 582,28 Mbp (89,0%)) ancorou-se em 18 cromossomos. Os *scaffolds* foram ancorados nos cromossomos usando um mapa genético de alta resolução (BREDESON et al., 2016). Nesta nova versão foi possível mapear 91,8% das 85.665 sequências *ESTs* disponíveis no banco de dados de nucleotídeos do NCBI, mostrando uma cobertura quase completa dos genes codificadores de proteínas na montagem (BREDESON et al., 2016). Além disto, 69.624 leituras puderam ser alinhadas e *ESTs* do sequenciamento anterior (v4.1, PROCHNIK et al., 2011) foram usadas, sendo adicionadas a outras 80.459 *ESTs* do GenBank e alinhadas ao genoma gerando 27.470 montagens alinhadas (BREDESON et al., 2016). Detalhes desta montagem podem ser conferidos na Tabela 1.

A montagem da versão 6 melhorou a contiguidade e integridade do genoma AM560-2 com *contigs* baseados em dados de sequenciamento de nova geração, ordenadas e orientadas em *scaffolds* usando bibliotecas paired-end. Em 2016 outra versão para *M. esculenta* foi lançada, a v7. Esta montagem foi construída seguindo a mesma estratégia da anterior, mas agora com sequências

subjacentes mais contíguas montadas *de novo* (PHYTOZOME, 2024b). Esta montagem foi financiada por Steve Rounsley (Univ. do Arizona), Dan Rokhsar (DOE-JGI, UC Berkeley) e Jean-Luc Jannink (Cornell Univ., NextGen Cassava) da Fundação Bill & Melinda Gates. Nesta nova montagem foram mapeados 669,4 Mbp, abrangendo 95,6% das sequências *ESTs* das cerca de 122.000 disponíveis no banco de dados de nucleotídeos do NCBI, mostrando uma cobertura quase completa dos genes codificadores de proteínas na montagem (BREDESON et al., 2016).

Atualmente as versões mais relevantes e recentes, v6.0 e v7.0, constam de 582 e 669 Mb respectivamente, sendo um pouco menor do que o tamanho estimado do genoma haploide de 750 Mb (KUON et al., 2019).

Tabela 1: Informações genéticas da montagem do genoma v6.0 de *M. esculenta* disponível no Phytozome13 (https://phytozome-next.jgi.doe.gov/info/Mesculenta_v6_1). *Scaffolds*: junção de *contigs*; N50 / L50 : métrica que mede o número de *contigs* / *scaffolds* que representam 50% da montagem.

Fonte da montagem:	JGI
Versão da montagem:	v6.0
Fonte da anotação:	JGI
Versão da anotação:	v6.1
Total do tamanho de <i>Scaffold</i> (bp):	582.281.492
Número de <i>Scaffolds</i> :	2.019
Min. número de <i>Scaffolds</i> contendo metade da montagem (L50):	10
Menor <i>Scaffold</i> de L50 (N50):	28.119.335
Tamanho total de <i>Contig</i> (bp):	495.476.809
Número de <i>Contigs</i> :	40.029
Min. Número de <i>Contigs</i> contendo metade da montagem (L50):	5.393
Menor <i>Contig</i> de L50(N50):	26.694
Número de Proteínas codificadas/transcritos:	41.381
Número de Proteínas codificadas/genes:	33.033
Porcentagem de <i>Eukaryote</i> BUSCO Genes:	97,7
Porcentagem de <i>Embryophyte</i> BUSCO Genes:	95,3
DB Xrefs:	BioProject: PRJNA394209

1.4. Efeitos da seca e salinização dos solos em vegetais

As mudanças climáticas são parte essencial da manutenção da vida no planeta, sendo fenômenos naturais que ocorrem ciclicamente, e que foram intimamente relacionadas com a evolução da vida na Terra (WANNER, 2008). Os biomas terrestres já passaram por diversas modificações em decorrência destas mudanças, e períodos de resfriamentos e aquecimentos tornaram continentes inteiros em conjuntos de habitats diferentes dos encontrados atualmente (ROJAS e FABRE, 2017). Contudo, um novo ciclo de mudanças climáticas, nunca visto na história recente do planeta, está em andamento, onde ações humanas de exploração de recursos naturais e de produção industrial massiva agem como catalisadores, adiantando em anos as mudanças climáticas que naturalmente ocorreriam em milênios (DA COSTA, 2017).

Em decorrência destas mudanças, áreas com déficit hídrico tornam-se mais comuns. Nestas regiões, mesmo as plantas mais adaptadas ao ambiente mais seco, sofrem com possíveis estresses, apresentando como consequência redução no tamanho de folhas, na extensão dos caules, na proliferação radicular, e por extensão, menor eficiência no uso da água (FAROOQ et al., 2009). Em casos mais extremos, os solos acabam salinizados, principalmente quando irrigados com água de baixa qualidade para a irrigação. A salinidade é o maior problema em todo o planeta (quase 1 bilhão de hectares de terra) e prejudica cerca de 900 milhões de hectares, ou seja, cerca de 20% do terreno útil do planeta e também cerca de metade do total de terras aráveis irrigadas globalmente (VELMURUGAN et al., 2020).

No Brasil, aspectos de solos salinos ocorrem principalmente no Rio Grande do Sul, na região do Pantanal Mato-grossense, e na região semiárida do Nordeste (RIBEIRO et al., 2003). Especialmente na região Nordeste, as extensas áreas com solos salinizados ocorreram tanto por ações climáticas, quanto pela utilização intensa das terras cultivadas (SILVA et al., 2011). A região de insuficiência hídrica nessa parte do território brasileiro cobre uma superfície

de 150 milhões de hectares. Nessa região, solos com alto teor de salinidade provocam desde a diminuição nos rendimentos das culturas até o abandono das áreas exploradas (ANDRÉO-SOUZA et al., 2010). Portanto, controlar a salinidade e desenvolver culturas mais tolerantes é muito importante para alcançar práticas agrícolas sustentáveis e atender à demanda alimentar presente e futura, em nível global (ZORB et al., 2019).

Em condições de estresse hídrico, o efeito mais comum que ocorre na planta é a desidratação celular, resultando na diminuição do potencial de pressão e no volume das células. Este efeito desencadeia fenômenos subsequentes, como o acúmulo de íons no citosol, que pode ocasionar até a morte celular por toxidez (TAIZ; ZEIGER; MURPHY, 2017). Durante o desenvolvimento das plantas, o estresse hídrico e/ou salino desencadeia principalmente efeitos como o fechamento estomático, a redução das atividades celulares/metabólicas e a diminuição da fotossíntese (TAIZ; ZEIGER; MURPHY, 2017), influenciando na produtividade e na viabilidade econômica do cultivo.

A salinidade é a medida da quantidade de sal presente no solo e na água e pode ser classificada em primária e secundária. A salinidade primária é o produto de processos naturais que depositam sais ao longo do tempo na terra e na água, como, por exemplo, o intemperismo, chuva e vento forte. A salinidade secundária é caracterizada pela ação da atividade antrópica, como desmatamento, limpeza de terras e irrigação excessiva (ARIF et al., 2020). Um solo ou água considerada salina possuem vários sais (Na^+ , Ca^{2+} , Mg^{2+} , Cl^- , SO_4^{2-} e HCO_3^-) (ARIF et al., 2020).

A salinidade tem causado um impacto negativo nas funções morfológicas e bioquímicas das plantas. Este fenômeno inibe a germinação das sementes, o crescimento, o desenvolvimento e o rendimento das plantas (ZHANG, DAI, 2019). A salinidade dificulta a maquinaria fotossintética, a transpiração e as trocas gasosas, diminuindo o conteúdo de clorofila e carotenoides, distorcendo a ultraestrutura do cloroplasto e o sistema PSII, reduzindo a condutância estomática (PAN et al., 2020).

A salinidade do solo diminui o seu potencial hídrico e o potencial hídrico foliar, perturba as relações hídricas da planta e reduz o seu turgor, o que acaba levando ao estresse osmótico (NAVADA et al., 2020). As plantas absorvem o sal do solo por meio de transportadores, esta absorção irá criar toxicidade de íons e perturbar a absorção de minerais e a homeostase. Com o acúmulo extensivo de íons (Na^+ e Cl^-) e inibição da absorção de K^+ e Ca^{2+} ocorre o desequilíbrio iônico (ISAYENKOV e MAATHUIS, 2019). A salinidade aumenta o conteúdo de espécies reativas de oxigênio (ROS) nas células vegetais e cria estresse oxidativo, influenciando na peroxidação lipídica, deterioração da membrana e danos ao DNA e às proteínas (EL GHAZALI, 2020).

A planta responde às condições de seca e salinidade por várias estratégias e abordagens como a compartimentalização, transporte de íons, adaptação osmótica, estimulação da maquinaria antioxidante e biossíntese de poliaminas (DE FREITAS et al., 2019). Isso ajuda na eliminação de ROS e estabilização da membrana, distribuição de íons, ultraestrutura e absorção de minerais (DE FREITAS et al., 2019). A planta também responde à salinidade regulando positivamente o nível de fitormônios endógenos, ou seja, ácido abscísico (ABA), auxinas, citocininas, ácido salicílico (SA), ácido jasmônico (JA), poliaminas, giberelinas e brassinosteroides (BRs) (FAHAD et al., 2015; RAO et al., 2016).

A planta regula muitos genes, fatores de transcrição, proteínas e metabólitos para propiciar a tolerância à salinidade. Abordagens ômicas como genômica, transcriptômica, proteômica e metabolômica são essenciais para o entendimento dos mecanismos da planta na adaptação à salinidade e no aumento da produtividade da cultura (D'AMELIA et al., 2018; JHA et al., 2019).

O estresse salino demonstra capacidade de suprimir o potencial hídrico, o teor de proteínas e reduz a reserva alimentar nas sementes em germinação de muitas plantas, por exemplo, em brócolis e couve-flor (WU et al., 2019). Solos salinos também já foram relatados causando diminuição na biomassa, área foliar, rendimento, caule e comprimento de raízes (ZORB et al., 2019).

A salinidade provoca mudanças rápidas e contínuas na regulação dos genes como a resposta hormonal. Os hormônios desempenham um papel vital na formação de respostas das plantas à salinidade possuindo papel multifacetado no crescimento e desenvolvimento das plantas e também na mitigação do estresse salino (ZHAO, 2021). Por exemplo: ABA causa o fechamento dos estômatos, retarda a senescência, regula vários processos metabólicos e catabólicos, além de reduzir a transpiração (CAO et al., 2017). Auxinas, por sua vez, desempenham um papel essencial no crescimento, desenvolvimento, alongamento e desenvolvimento de sistemas vasculares em condições normais e salinas (SUN et al., 2018); enquanto o etileno, desempenha um papel crítico durante vários estresses e ajuda no crescimento e desenvolvimento modulando positivamente resposta à salinidade, promovendo a absorção de íons e aumentando a atividade antioxidante (FAHAD et al., 2015).

Além das previsões sombrias decorrentes das mudanças climáticas e dos efeitos deletérios da salinização dos solos evidenciados, outro problema a ser enfrentado é o do crescimento populacional, sendo um desafio imposto o de otimizar a produção de alimento para uma população estimada de 9 bilhões de pessoas até 2050, quando o mundo viverá, provavelmente, sob a influência de três grandes crises: a escassez de petróleo, de água potável, e de alimentos. Neste cenário, a biotecnologia vegetal aplicada ao melhoramento da espécie pode ocupar um importante papel na busca pela atenuação desses problemas (CARRER; BARBOSA; RAMIRO, 2010).

1.5. Regulação gênica - epigenética

Os mecanismos envolvidos na regulação hormonal se originam nas mudanças gerais do perfil de expressão gênica, que visa a resposta imediata ao estresse (GRATIVOL, HEMERLY, FERREIRA, 2012; ZHANG et al., 2021). Em plantas que respondem ao estresse abiótico, evidências destacam mecanismos epigenéticos, como metilação do DNA e modificação de histonas, desempenhando papéis relevantes (CHAN et al., 2005; CHINNUSAMY, ZHU, 2009). A metilação do DNA, que ocorre em praticamente todos os organismos, corresponde a uma marca epigenética hereditária, mostrando papéis em muitos processos biológicos, como crescimento e desenvolvimento de plantas, indução

de floração, regulação da expressão gênica e respostas a estresses ambientais (JULLIEN et al., 2012; LI et al., 2021).

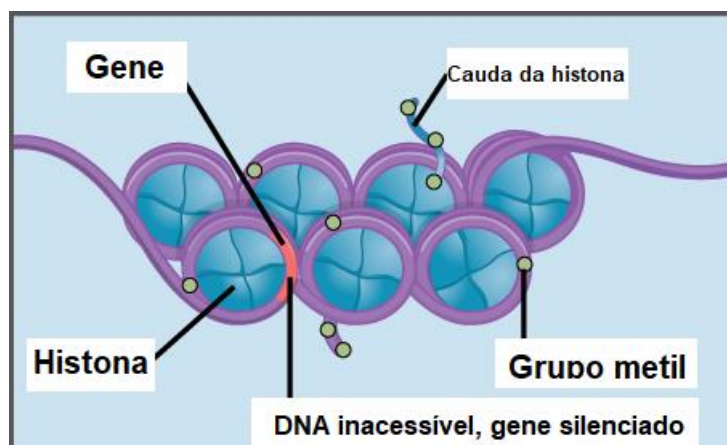
A metilação é um fator crítico na regulação da estrutura da cromatina e da expressão gênica, alterando principalmente a estabilidade e o posicionamento dos nucleossomos, as unidades fundamentais da cromatina, afetando a acessibilidade do DNA a proteínas reguladoras ou complexos proteicos envolvidos na replicação, reparo e transcrição do DNA (PIKAARD, SCHEID, 2013; VAN LIJSEBETTENS, GRASSER, 2014).

O nucleossomo compreende 145-147 pb de DNA envolto em um octâmero de proteínas histonas (duas cópias de cada uma das quatro histonas principais, H2A, H2B, H3 e H4). Tanto em plantas quanto em animais, a metilação ocorre principalmente na citosina do contexto do dinucleotídeo CG (GEHRING e HENIKOFF, 2007). No entanto, o alto nível de 5-metilcitosina em algumas espécies vegetais mostram que a metilação não está restrita ao contexto CG, isso levou à descoberta de citosina metilada em CNG e, menos abundantemente, em sequências CNN (GRUENBAUM et al., 1981).

O processo de metilação em plantas ocorre quando grupos metil são adicionados à citosina por uma das três classes de metiltransferase: i) DNA metiltransferase tipo I (MET1), que mantém a metilação GC em DNAs recém-sintetizados e são necessários para manter padrões de metilação repetitivos e únicos ao copiar sequências durante a gametogênese (SAZE, SCHEID, PASZKOWSKI, 2003); ii) CROMOMETILASE3 (CMT3), que mantém a metilação dos sítios CNG; iii) METILTRANSFERASE DE DOMÍNIOS REARRANJADOS (DRM), sendo alvo específico de pequenas moléculas de RNA (LINDQUIST, HENIKOFF, 2002).

Figura 2. Nucleossomos sendo representados por octâmeros de histonas enoveladas ao DNA. O esquema mostra a localização de grupos metil e caudas de histonas responsáveis pela

interação com proteínas que moldarão este segmento do DNA. Genes inseridos neste enovelamento são menos acessíveis aos fatores de transcrição.



Fonte: Domínio público em

https://pt.wikipedia.org/wiki/Metilação_do_ADN#/media/Ficheiro:Figure_16_03_02.png

Nas plantas, em condições normais, a proporção de citosinas metiladas para os valores totais varia de 20% a 30% (HE, CHEN e ZHU, 2011). Algumas mudanças relacionadas à metilação no genoma das plantas que respondem ao estresse são temporárias, enquanto outras podem ser estáveis e hereditárias para as gerações futuras, permitindo que as plantas descendentes lidem com esses desafios no futuro (ZHANG et al., 2018; AKHTER et al., 2021).

Análises do perfil de metilação do DNA e sua dinâmica permitem inferências sobre as respostas genéticas das plantas a estresses abióticos (BARTELS et al., 2018). Esses padrões de metilação exibem níveis dinâmicos que direcionam as respostas das plantas ao estresse ambiental (BRINKMAN et al., 2010; SONG et al., 2020). Durante as respostas ao estresse, mudanças no padrão de metilação do DNA vinculam mudanças na regulação de genes envolvidos no processo de resposta ao estresse e interações de vias reguladoras, como vias de fitormônios (WANG et al., 2014; SANTOS et al., 2020). Por exemplo, em plantas de arroz tolerantes à seca, genes associados a regiões diferencialmente metiladas (ou seja, níveis de metilação estatisticamente diferentes do controle negativo) também apresentaram associações com resposta ao estresse, morte celular programada e atividades de reservas de nutrientes. Esses aspectos destacam o envolvimento com a tolerância constitutiva ao estresse hídrico (WANG et al., 2016). Além disso, plantas

aparentemente tolerantes têm um perfil de metilação mais estável do que plantas sensíveis, apresentando menos regiões metiladas diferencialmente induzidas pela seca (WANG et al., 2016).

1.6. Metodologia para análise de metilação

A caracterização de metilomas de DNA dinâmicos em grupos filogenéticos tem atraído crescente interesse para uma melhor compreensão da evolução da metilação do DNA, bem como sua função e significado biológico em eucariotos (WANG et al., 2015). Métodos baseados em sequenciamento são promissores no cumprimento desta tarefa. Existem variadas abordagens de criação de perfil de metilação em vegetais, estas podem variar em muitos aspectos, incluindo o *input* de DNA, resolução, cobertura de região genômica e análise de bioinformática. Algumas das técnicas mais utilizadas são Polimorfismo Amplificado Sensível à Metilação (MSAP), sequenciamento de bissulfito, imunoprecipitação de DNA de metila e *MethylRad-Seq*, as quais serão abordadas a seguir.

1.6.1. Técnica de polimorfismo de amplificação sensível à metilação (MSAP)

O MSAP tem sido amplamente utilizado nas últimas três décadas para avaliar alterações na metilação do DNA em uma variedade de plantas modelo e não modelo (AGIUS et al., 2023). Essencialmente, MSAP é uma modificação da técnica de polimorfismo de comprimento de fragmento amplificado (AFLP) exceto que a enzima cortadora frequente *MseI* é substituída pelos isosquizômeros sensíveis à metilação *HpaII* e *MspI* (AGIUS et al., 2023). Assim, o método MSAP utiliza *HpaII* e *MspI*, que reconhecem o mesmo sítio alvo (5'-CCGG-3'). No entanto, a sua capacidade de clivar baseia-se no estado de metilação de resíduos específicos de citosina na sequência (AGIUS et al., 2023). Especificamente, *HpaII* cliva apenas locais com citosinas externas hemimetiladas (mCCGG), enquanto *MspI* cliva em citosinas internas hemi ou totalmente metiladas (CmCCGG). Nenhuma das enzimas cliva locais que estão totalmente metilados na citosina externa ou nas citosinas internas e externas hemi ou totalmente metiladas (AGIUS et al., 2023). Por outro lado, ambas as enzimas podem digerir sequências '5'-CCGG-3'. Estas enzimas são combinadas

com EcoRI, que é marginalmente afetada pela metilação da citosina. As digestões de DNA EcoRI/HpaII e EcoRI/MspI são ligadas a adaptadores específicos e os fragmentos ligados passam por rodadas de amplificação por PCR pré-seletiva e seletiva (AGIUS et al., 2023).

A amplificação seletiva com primers específicos marcados com fluorescência produz fragmentos de PCR que são resolvidos em sistemas de detecção de eletroforese capilar (SCHULZ et al., 2013). A matriz de dados brutos de presença e ausência de fragmentos é traduzida em uma matriz de caracteres binários (0, ausência; 1, presença). Esta informação binária está associada ao seu status de metilação (SCHULZ et al., 2013). A principal vantagem da técnica MSAP é a sua ampla viabilidade em todas as espécies, independentemente da disponibilidade do genoma de referência, a relação custo-benefício, independentemente do tamanho e complexidade do genoma, e a adequação para o perfil de grandes conjuntos de amostras. MSAP também tem algumas limitações (AGIUS et al., 2023). Ele detecta padrões de metilação do DNA em loci anônimos distribuídos aleatoriamente por todo o genoma e não pode fornecer informações sobre genes ou regiões genômicas específicas (AGIUS et al., 2023). No geral, o MSAP continua sendo uma ferramenta valiosa e de baixo custo para avaliar a metilação do DNA, tornando-o uma escolha popular em muitos laboratórios. Sua limitação está relacionada à detecção de alterações especificamente nas sequências CCGG reconhecidas por enzimas de restrição altamente sensíveis, o que limita a detecção de outros tipos de metilação (AGIUS et al., 2023).

1.6.2. Sequenciamento de bissulfito

Marcas de metilação no DNA permanecem intactas após sua extração, e tratamentos do DNA genômico com bissulfito de sódio resultam na desaminação de citosinas não metiladas em uracila, deixando as citosinas metiladas no DNA intactas (FROMMER et al., 1992). O DNA tratado é posteriormente amplificado por PCR com uma polimerase tolerante a uracila para que seja possível análise, fazendo com que a uracila se converta em timina. A metilação do DNA pode, portanto, ser lida diretamente pelo sequenciamento tradicional *Sanger* ou *Illumina* através da comparação com uma sequência de referência ou não

tratada, sendo altamente quantitativa e com resolução de pares de bases (GOUIL e KENIRY, 2019).

O sequenciamento do bissulfito informa todas as citosinas, independentemente do contexto, podendo, portanto, ser altamente benéfico (GOUIL e KENIRY, 2019). Apesar disto, existem algumas questões a serem consideradas. Em primeiro lugar, o tratamento com bissulfito é muito severo, levando à degradação do DNA tornando mais difícil a amplificação por PCR, portanto, muitas vezes são necessárias grandes quantidades de DNA de entrada (GOUIL e KENIRY, 2019). Em contrapartida, métodos que utilizam a desaminação enzimática da citosina evitam o e tratamento químico severo. Além disso, os dados da sequência de bissulfito requerem técnicas de análise bioinformática mais sofisticadas do que as necessárias para o DNA não convertido, uma vez que a sequência deve ser comparada a um genoma de referência convertido por bissulfito antes que as chamadas de metilação possam ser inferidas (GOUIL e KENIRY, 2019)..

O sequenciamento *Illumina* de DNA de bissulfito também sofre dos mesmos problemas que todos os dados de leitura curta, particularmente problemas de mapeamento para regiões repetidas ou de baixa complexidade, incluindo regiões de pertinência a 5mC, como regiões reguladoras fortemente ricas em GC e DNA repetitivo (GOUIL e KENIRY, 2019). Esses problemas são ainda agravados pela perda de diversidade de sequências devido à conversão de bissulfito (ALKAN, COE, EICHLER, 2011; ROBERTS, CARNEIRO, SCHATZ, 2017). Adicionalmente, a natureza curta das leituras reduz a probabilidade de conter um polimorfismo informativo de nucleotídeo único (SNP) (DELANEAU, 2013).

1.6.3. Imunoprecipitação de DNA de metila (MeDIP)

Usando um anticorpo de 5mC para análise de metilação, que pode ser acoplada à detecção de *array* (chip MeDIP) ou sequenciamento (MeDIP-seq). Os dados gerados por este método são consistentes com o WGBS (WARDENAAR et al., 2013). No entanto, o uso de anticorpos 5mC requer uma quantidade significativa de DNA de *input* e está associado a tendências para regiões hipermetiladas não podendo diferenciar o contexto de metilação (AGIUS

et al., 2023). Recentemente, uma abordagem Metil-seq enzimático (EM-seq) foi desenvolvida (FENG et al., 2020). Devido o tratamento com bissulfito causar danos e degradação do DNA a EM-seq usa a conversão enzimática de citosinas não metiladas em uracilas para obter o mesmo produto de sequenciamento sem afetar a integridade do DNA (FENG et al., 2020). As sequências resultantes são semelhantes às geradas pelo WGBS e podem ser analisadas da mesma forma. Bibliotecas geradas, usando EM-seq, superaram sequenciamento baseado em bissulfito em todas as medidas específicas, como cobertura, duplicação, sensibilidade e até distribuição de GC, melhores correlações entre *input* de DNA, aumento do número de CGs dentro de características genômicas e precisão das chamadas de metilação da citosina (VAISVILA et al., 2021).

1.6.4. Metodologia MethylRAD-Seq

As recentemente descobertas enzimas do tipo *Mrr* são atraentes para o desenvolvimento de novos métodos, devido à sua capacidade de coletar fragmentos de 32-bp de DNA metilados de todo o genoma para sequenciamento de alto rendimento (WANG et al., 2015). O método de análise de perfil de metilação de DNA (chamado *MethylRAD*) usando enzimas *Mrr* permite a análise de metilação *de novo* (sem referência), com utilização de concentrações de DNA extremamente baixa (por exemplo, 1 ng) e ajuste da densidade de *Tag*, tudo que ainda é inatingível para métodos de perfis de metilação amplamente utilizados como *methylation-sensitive amplification polymorphism (MSAP)*, *reduced representation bisulphite (RRBS)* e *methylated DNA immunoprecipitation - next-generation sequencing (MeDIP)* (WANG et al., 2015).

O *MethylRAD* tem várias vantagens sobre as ferramentas existentes e preenche uma lacuna do atual kit de ferramentas epigenômicas, fornecendo uma ferramenta universal que pode ser usada para diversas aplicações de pesquisa, desde espécies modelo a não modelo, amostras comuns a escassas e genomas pequenos a grandes, mas a um custo acessível (WANG, 2015).

Enzimas de restrição dependentes de metilação possuem características que as tornam atraentes para o desenvolvimento de novos métodos de estudos epigenômicos. Semelhante aos métodos baseados em afinidade, essas enzimas podem avaliar diretamente o status de metilação do DNA sem o auxílio de

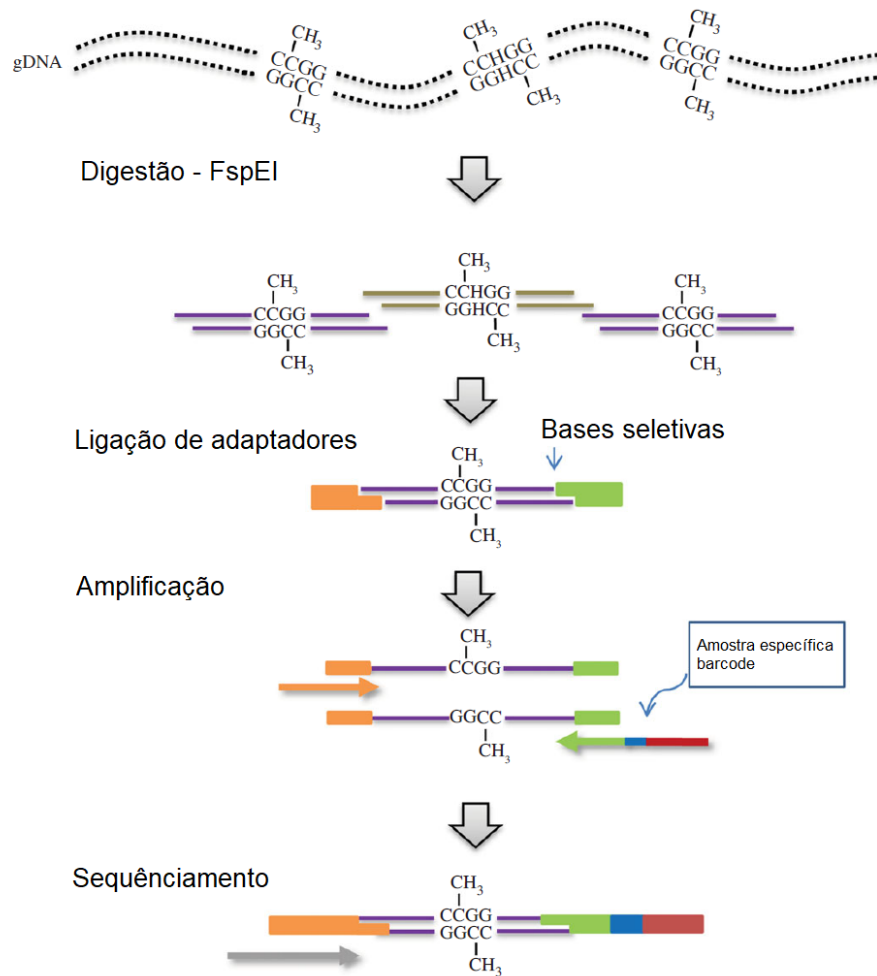
conversão química, mas com especificidade muito maior e sensibilidade (LI, TOLLEFSBOL, 2021).

MethylRAD usa uma das enzimas do tipo *Mrr* (por exemplo, *FspEI*, *MspJI*, *LpnPI*, *AspBHI*, etc.) para realizar a produção de metilomas. Por exemplo, *FspEI* pode reconhecer 5-metilcitosina (5-mC) e 5-hidroximetilcitosina (5-hmC) nos sítios C^mC e ^mCDS (na presença de um ativador; D = A ou G ou T; S = C ou G) (COHEN-KARNI et al., 2011). Como enzimas *Mrr* são bloqueadas por 5-hmC glicosilado (5ghmC) (COHEN-KARNI et al., 2011), uma discriminação adicional dos dois tipos de modificação é possível se a glicose for adicionada ao grupo hidroxila de 5-hmC (por exemplo, usando T4 b-glucosiltransferase). *FspEI* gera uma quebra de DNA de fita dupla no lado 3' da citosina modificada a uma distância fixa (N12/N16) (WANG et al., 2015). Se os sítios de destino forem simetricamente metilados, *FspEI* pode clivar bidireccionalmente para gerar fragmentos de 32 bases de comprimento com o sítio metilado no meio (WANG et al., 2015).

Esta característica enzimática permite que quase todos os locais de restrição no genoma possam ser rastreados. Além disso, ao contrário dos baseados em afinidade, métodos, *MethylRAD* podem discriminar entre CG e metilação não-CG. O procedimento de preparação da biblioteca para *MethylRAD* é relativamente simples. Os fragmentos *FspEI* têm saliências arbitrárias de quatro bases em cada extremidade 5', e a ligação do adaptador é realizada pela ligação da extremidade coesa usando adaptadores com extremidades totalmente degeneradas (50-NNNN-30) (Figura 3) (WANG et al, 2015).. Após estas etapas, a amplificação fica encarregada de finalizar o processo, assim sendo possível a verificação quantitativa dos fragmentos metilados e possibilitando análises de fragmentos diferencialmente metilados (WANG et al, 2015).

Figura 3. Visão geral esquemática do procedimento para a preparação da biblioteca *MethylRAD*. O DNA genômico é digerido com a enzima de restrição *FspEI*, produzindo fragmentos de 32 pb incluindo saliências 3' de quatro bases. Adaptadores com saliências compatíveis (NNNN) que são ligados a cada extremidade desses fragmentos. A densidade da tag pode ser ajustada usando adaptadores com saliências seletivas (por exemplo, NNNG). As construções são

amplificadas e purificadas por extração em gel. Os barcodes específicos da amostra são incorporados em cada constructo de PCR e os produtos reunidos para sequenciamento.



Adaptado de Wang et al 2015

MethylRAD exhibe alta especificidade, sensibilidade e reprodutibilidade, e permite a análise por metilação *de novo*, amostras de DNA extremamente baixas e ajuste flexível de densidade de tags. MethylRAD preenche um vazio no kit de ferramentas epigenômica atual, fornecendo uma ferramenta universal que pode ser usada de forma econômica para caracterização de metilomas de DNA dinâmicos em diversos grupos filogenéticos sem restrição de tamanho do genoma e fonte de DNA ou exigência de um genoma de referência (WANG et al, 2015).

1.7. Visão geral sobre estudos de metiloma em vegetais

Ainda que nos últimos anos a quantidade de estudos sobre perfis de metilação em vegetais tenha aumentado consideravelmente, ainda é uma área que carece de análises sistêmicas e complementares como os estudos em genômica e transcriptômica, no entanto, alguns estudos tem mostrado resultados animadores quanto ao descobrimento de informações pertinentes a epigenética de vegetais expostos a estímulos diversos.

Wang e colaboradores (2015 [B]), utilizaram a linha de introgressão de arroz IL177-103 e seu pai recorrente IR64, que mostram tolerância ao estresse salino contrastante, para caracterizar as alterações de metilação do DNA sob estresse salino e subsequente recuperação usando análise de polimorfismo amplificado sensível à metilação (MSAP). O acesso IL177-103 apresenta tolerância significativamente melhorada à salinidade, representada por maior teor relativo de água, teor de ácido abscísico endógeno, atividade de enzimas sequestradoras de espécies reativas de oxigênio e menor concentração de Na⁺ na parte aérea, em comparação com IR64. Sementes de arroz foram esterilizadas superficialmente em NaClO 0,1% (v/v) e enxaguadas várias vezes antes da germinação a 37 C em uma incubadora. Depois de 3 dias, as sementes germinadas foram transferidas para um tanque de 20 L com solução Yoshida nutritiva. As mudas foram cultivadas por 12 dias após a germinação. O tratamento de salinidade foi aplicado pela adição de NaCl para atingir uma concentração final de 125 mM. Após o tratamento de 4 dias, as plantas estressadas foram recuperadas em 5 dias por cultivo em Solução nutritiva de Yoshida sem NaCl. Os resultados do MSAP mostraram que menos de 10,5% dos locais de metilação do DNA detectados eram específicos do genótipo, de acordo com seu histórico genético semelhante. Mudanças de metilação do DNA induzidas por sal em ambos os genótipos foram detectadas principalmente nas raízes, e a maior parte das alterações de desmetilação/metilação do DNA induzidas por sal permaneceram mesmo após a recuperação, implicando sua herança na geração atual. Além disso, alguns locais com diferenças estáveis de metilação do DNA foram identificados entre IL177-103 tolerante a sal e IR64 sensível a sal, fornecendo assim marcadores epigenéticos específicos.

Abid e colaboradores (2017) mostram os efeitos da seca nas relações hídricas, trocas gasosas e alterações epigenéticas em dois genótipos de fava (*Vicia faba* L.) com níveis contrastantes de tolerância a seca. Os pesquisadores mostram que o genótipo Bachar tolerante à seca foi menos afetado pelo déficit hídrico em comparação com o F177 sensível à seca em termos de função fotossintética e eficiência no uso de água. Neste estudo, a técnica de polimorfismo de amplificação sensível à metilação (MSAP) foi usada para traçar padrões de metilação no perfil do DNA de Bachar e F177 em condições de seca e controle. Para o tratamento, sementes foram esterilizadas superficialmente em solução diluída de hipoclorito e lavadas em água destilada esterilizada. Em seguida, as sementes foram embebidas em água destilada antes da semeadura em vasos de 5 kg cheios de uma mistura de solo. A solução de Hoagland com meia-força (Hoagland e Arnon, 1950) foi usada para irrigação. Os genótipos estudados passaram por três tratamentos de água: 90% da capacidade de campo (CA) como tratamento controle, 50% FC (estresse moderado) e 30% CF (estresse severo). Os tecidos das raízes e brotos foram coletados separadamente 20 dias após o início do estresse. Usando 16 pares de primers MSAP, um total de 764 e 913 bandas foram reveladas em folhas e raízes respectivamente. Sob condições de controle nos tecidos foliares, a metilação de CCGG foi de 37,43% em Bachar e 41,23% em F177, enquanto nos tecidos radiculares, a metilação total variou entre 16,53% e 23,43% em F177 e Bachar respectivamente. Além disso, o nível geral de metilação nas folhas de ambos os genótipos de fava estudados foi consistentemente maior do que nas raízes sob todas as condições de água. Entre esses loci metilados, os locis totalmente metilados foram mais comuns que locis hemimetilado. Sob controle e tratamento de seca, o nível de metilação total em tecidos foliares foi de 37,43% e 30,62% em Bachar, 41,23% e 38,16% em F177, indicando uma queda de 6,81% e 3,07% em Bachar e F177, respectivamente, devido à exposição à seca. Os resultados revelaram que a metilação do DNA em tecidos radiculares diminuiu em 3,63% (de 23,43% para 19,80%) em Bachar e aumentou em 0,66% (de 16,53% para 17,19%) em F177 sob estresse hídrico. Também foi possível verificar regiões diferencialmente metiladas (DMRs) potencialmente relacionadas ao estresse hídrico. Cinco destas têm alta homologia com proteínas previamente identificadas ou putativas encontradas em outras plantas, incluindo lipoxigenase

(LOX), proteína quinase dependente de cálcio (CDPK), transportador ABC família (ABC), glicosil hidrolase (GH) e fosfoenolpiruvato carboxilase (PEPC). O genótipo tolerante (Bachar) apresentou nível de hipometilação que poderia levar a ativação transcricional de genes relacionados à seca. Comparação de padrões de metilação do DNA entre os genótipos considerados revelaram ainda que os tecidos foliares de Bachar exibiram um maior nível de hipometilação do que F177 (-6,81% e -3,14%). No geral, os pesquisadores concluem que as mudanças epigenética em todo o genoma são provavelmente um importante mecanismo regulatório para a resposta da fava à seca e também a outras tensões ambientais.

Guo e colaboradores (2019), estudam o efeito do estresse por frio na produtividade de plantas de arroz. Para entender as respostas transcriptômicas e a regulação da metilação do arroz em resposta ao estresse por frio, foi analisada uma variedade de arroz tolerante ao frio (*Oryza sativa* L. cv. P427). As variedades de arroz Japonica Nip e Indica 9311 foram usadas como controle. Três materiais de sementes foram embebidos em temperatura ambiente por 48 horas e depois germinados por 48 horas a 32 °C. Após o término da germinação, as sementes foram semeadas e cultivadas em incubadora de luz. As condições de cultivo foram as seguintes: dia/noite 28 °C/25 °C e 250 μM m²s⁻¹ de luz. Após o crescimento da planta até o estágio de três folhas, ela foi transferida para uma câmara climática artificial a 2–3 °C por 3 dias; então, transferidas para uma incubadora de luz por 7 dias. As mudas foram cortadas com tesoura antes do tratamento, após o tratamento e após o crescimento. Os pesquisadores descobriram que o P427 exibiu maior tolerância a baixas temperaturas, provavelmente por meio do aumento da atividade enzimática antioxidante e da promoção do acúmulo de ácido abscísico (ABA). O sequenciamento de imunoprecipitação de DNA metilado (MeDIP-seq) mostraram que o número de genes alterados por metilação foi maior em P427 (5496) e ligeiramente menor em outras duas variedades estudadas, Japonica (Nip) e 9311 (4528 e 3341, respectivamente), e apenas 2,7% (292) de genes de metilação foram detectados como genes diferencialmente metilados comuns (DMGs) relacionados à tolerância ao frio nas três variedades. Análises transcriptômicas revelaram que 1.654 genes alteraram especificamente a expressão em P427 sob estresse por

frio. Esses genes pertenciam principalmente a famílias de fatores de transcrição, como Mieloblastose (MYB), proteínas de ligação do elemento responsivo a APETALA2/etileno (AP2-EREBP), NAM-ATAF-CUC (NAC) e WRKY. Os pesquisadores demonstraram que, no geral, as diferenças nos níveis de metilação e expressão gênica de P427 em resposta ao estresse de baixa temperatura, forneceram uma base para futuras investigações da relação entre estresse ambiental, metilação do DNA, e expressão gênica em arroz.

2. Hipótese

Padrões em perfis de metilação podem ser modificados diante de respostas ao estresse biótico/abiótico ou desenvolvimento dos vegetais. Aqui, estima-se que os padrões de metilação entre duas variedades de mandioca contrastantes ao estresse hídrico (Dourada x Formosa) sejam diferentes, revelando genes importantes metilados que possam estar atrelados as características de sensibilidade ou tolerância ao estresse. A verificação da metilação diferencial entre estes genes pode revelar como o perfil de metilação em cada variedade se comporta frente aos estímulos, neste caso, o estresse hídrico. Além disso, estimamos que genes diferencialmente metilados estejam em associações diretas com vias fito-hormonais de resposta ao estresse abiótico e modulação de vias de crescimento vegetal.

3. Objetivos

Este trabalho teve como objetivo desenvolver análise dos perfis de metilação genômicos de variedades de mandioca utilizando as bibliotecas de metilação para comparação entre diferentes amostras e grupos expostos ou não ao déficit hídrico para identificação de genes relevantes em vias de resposta ao estresse que sejam úteis na caracterização e melhoramento de novas variedades de mandioca.

Para tanto, objetivos específicos foram desenvolvidos, como:

- Mapear leituras no genoma de referência de *M. esculenta* através de ferramentas de bioinformática a fim de identificar a distribuição dos sítios de metilação para cada acesso e grupo de tratamento;

- Identificar na distribuição dos sítios de metilação elementos funcionais genômicos (exons, íntrons, promotores) a fim de analisar como se distribuem os sítios de ligação nestes elementos traçando um perfil para cada amostra dos diferentes grupos de tratamento;
- Verificar e quantificar os níveis de metilação no DNA para cada variedade e cada tratamento a fim de comparar os diferentes perfis de metilação ocorrentes;
- Analisar o nível de metilação entre os diferentes tratamentos para elucidar áreas mais ou menos metiladas e correlacionar estas informações com a resposta ao estresse em diferentes acessos e tratamentos;
- Analisar o enriquecimento de ontologia gênica (GO) e vias metabólicas (KEGG) para genes relacionados aos diferentes perfis de metilação a fim de verificar suas relações com as características de cada acesso e grupo estudado;
- Analisar eventos de SNP/INDEL nos genes diferencialmente metilados para compreender os impactos destas modificações nos genes envolvidos a resposta ao estresse hídrico.

4. Metodologia

4.1. Ensaio biológico: materiais vegetais, aplicação do estresse e geração das bibliotecas MethylRad-Seq

O delineamento experimental foi em blocos casualizados, com três repetições e tratamentos sem (controle negativo) e com estresse de déficit hídrico. O ensaio foi realizado em casa de vegetação na Embrapa Mandioca e Fruticultura. O estresse foi aplicado em manivas após 50 dias de plantio, imposto pela suspensão da irrigação das plantas. Nas plantas controle, a irrigação foi mantida durante o experimento, mantendo o teor de umidade do solo próximo à capacidade de campo. O teor de umidade do solo foi monitorado diariamente (às

6:00 e 18:00 h) com aparelho TDR (*Time Domain Reflectometry*, Reflectometria no domínio do tempo), sendo a reposição hídrica das plantas controle feita até a capacidade de campo, tendo por base as leituras TDR.

A fração da água transpirável do solo (FTSW) foi calculada pela metodologia segundo Sinclair e Ludlow (1986). Foi realizado uma análise de troca gasosa foliar em folhas maduras – terceira folha, entre 9h e 10h, durante todo o experimento usando um medidor de fotossíntese portátil ADC BioScientific (modelo LCpro-SD, Ltd., Reino Unido)

Após 18 dias do início da supressão da rega as plantas apresentaram um nível de estresse severo (SEV) caracterizado por folhas completamente murchas e solo com umidade a 10% (40% da capacidade de campo). Amostras de raízes foram coletadas, congeladas em nitrogênio líquido, e armazenadas em ultrafreezer até extração de DNA total (metodologia CTAB; DOYLE, DOYLE, 1987). A qualidade e quantidade dos DNAs extraídos foram avaliadas em espectrofotômetro (NanoVue Plus, GE Healthcare, Waukesha, WI, Estados Unidos) e eletroforese em gel de agarose 1% (p/v) e os DNAs de cada variedade e tratamento (2 tratamentos, 2 variedades 3 repetições) permitiram gerar e sequenciar 12 bibliotecas MethylRAD-Seq a partir de sequenciamento Illumina HiSeq X-Ten PE 150 (CD Genomics, USA). Esta plataforma utiliza um *output Dual flow cell* para 1.6 a 1.8 Terabases. As bibliotecas foram geradas utilizando enzimas *Mrr-like* (FspE1), resumidamente a preparação da biblioteca MethylRAD começou com a digestão do gDNA com FspE1 e ligação de adaptadores. O produto de ligação foi então amplificado com PCR e codificado em barras para sequenciamento multiplex. O sequenciamento de alto rendimento de bibliotecas de *tags* permitiu acesso a informações de sequência dos locais de metilação em escala genômica. De acordo com consistência da eficiência de amplificação de *tags* isométricas, a profundidade de sequenciamento das *tags* metiladas puderam refletir o nível de metilação dos locais CCGG/CCWGG.

4.2. Análise de qualidade do sequenciamento e mapeamento das *leituras* em genoma de referência.

Ainda para efeito de entendimento, a qualidade das *leituras* sequenciadas considerou os valores *Phred* das bases, sendo as *leituras* filtradas (software PEAR v0.9.6), com eliminação daquelas com mais de 15% de bases com valor *Phred* < 30, ou com conteúdo N maior que 8%. As *leituras* denominadas *EnzymeReads*, aquelas de tamanho esperado, foram identificadas em cada biblioteca, sendo removidas aquelas sem o sítio de restrição esperado. Para identificar os locais de metilação confiáveis considerou-se os sítios CCGG e CCNGG com profundidades de sequenciamento superior a 3. As *tags* de tamanho esperado (CCGG: 32 pb; CCNGG: 31 pb) contendo as sequências de referências CCGG e CCNGG foram mapeadas (software SOAP v.2.21; parâmetros: -M 4-v 2 -r 0) no genoma de *M. esculenta* (NCBI, referência GCF_001659605.1), tendo-se desconsiderado as *leituras* repetidamente mapeadas.

4.3. Análise da distribuição dos sítios de metilação nos elementos gênicos estruturais e funcionais

Os elementos gênicos estruturais de *M. esculenta*, incluindo íntrons, éxons, *UTRs* e 2000 pb *upstream* ao TSS (*Transcription Start Site*) foram identificados dos dados do portal Phytozome para a espécie (versão 12.1; <https://phytozome.jgi.doe.gov/pz/portal.html>). Os sítios de metilação nas regiões *UTRs* foram contabilizadas a partir do software *SNPeff* (CINGOLANI et al., 2012), e nos demais elementos genéticos, a partir do software *bedtools* (v2.25.0; <https://bedtools.readthedocs.io/en/latest/>).

4.4. Quantificação relativa da metilação do DNA e análise de distribuição dos sítios de metilação em regiões gênicas

O nível de metilação dos sítios alvos CCGG e CCNGG é reflexo da consistência da eficiência de amplificação de sequências isométricas e da profundidade de sequenciamento das *tags* metiladas. Assim, a quantificação do nível de metilação foi expressa em RPMs (*leituras* por milhão), em que o nível

quantitativo do valor de metilação seria igual ao número de *leituras* que cobrem o sítio pelo total de *leituras* de qualidade na biblioteca $\times 1.000.000$.

Para os perfis de distribuição das *tags*, na análise da distribuição dos sítios de metilação em regiões gênicas, foram consideradas as regiões CDS e regiões de 2 kb *upstream* (a montante) e *downstream* (a jusante) ao TSS, bem como 2 kb a montante e a jusante ao TTS (*Transcription Terminator Site*).

4.5. Correlação dos níveis de metilação entre diferentes amostras

Gráficos de dispersão de profundidades de sequenciamento de todos os locais de metilação permitiram estimar coeficientes de correlação de Pearson. Assim, o nível de metilação de uma amostra/biblioteca na abcissa, e o nível de metilação de outra amostra na ordenada, em pares de comparações, possibilitaram estimar os coeficientes de correlação. Aqueles próximos do valor 1,00 refletem as melhores correlações entre os dois conjuntos de dados, situação ideal para verificar a consistência dos dados de repetições biológicas.

4.6. Níveis de metilação diferencial

Os níveis de metilação das variedades, considerando as comparações dos respectivos tratamentos com aplicação do estresse e sem a aplicação do estresse (contrastes FORSEV vs FORCONT; DOUSEV vs DOUCONT), permitiram classificar os sítios diferencialmente metilados em hipermetilados e hipometilados, detectados a partir do software edgeR [$p\text{-value} \leq 0,05$, e $\text{Log}_2 \text{FC} > 1$ (hipermetilação) e < -1 (hipometilação)]. Valores de $p\text{-value}$ foram ajustados pela técnica FDR (*False Discovery Rate*), para correção de erros do tipo I.

4.7. Comparações dos níveis de metilação entre contrastes e distribuição dos sítios diferencialmente metilados

Os níveis de metilação dos contrastes (FORSEV vs FORCONT; DOUSEV vs DOUCONT) foram derivados de amostras conjuntas das réplicas biológicas, formando *pools*, tendo por base as profundidades de sequenciamento de cada sítio de metilação em cada biblioteca, aplicando-se os valores de $p\text{-value}$ e de $\text{Log}_2 \text{FC}$ dos grupos implicados, a partir do software edgeR. Sítios pesquisados

estiveram presentes em ao menos um dos grupos, com profundidade de sequenciamento superior a três. Sítios diferencialmente metilados e classificados como hipermetilados ou hipometilados foram associados a genes previamente anotados, cujas expressões (dados de RPM) foram ilustradas em *heatmaps* gerados através do software Tbtools v1.09.

4.8. Genes diferencialmente metilados exclusivos e comuns entre diferentes contrastes dos tratamentos e variedades

Diagramas de Venn (<https://bioinfogp.cnb.csic.es/tools/venny/>) contemplando os conjuntos de sítios diferencialmente metilados, ao se comparar dados dos contrastes dos tratamentos com e sem o estresse de cada variedade, os quais foram associados aos identificadores dos genes de *M. esculenta*, permitiram identificar aqueles identificadores comuns ou exclusivos das respostas de Formosa e/ou Dourada ao estresse aplicado.

4.9. Análise de termos de ontologia gênica (GO)

Os conjuntos de genes portando metilações diferencialmente detectadas (hiper e hipometilados para sítios CCGG e CCNGG) nos elementos genéticos estruturais de Formosa (G01) e Dourada (G04) foram caracterizados quanto à ontologia gênica, com identificação de termos enriquecidos após testes de Fisher ($p\text{-value} < 0,01$), ao se comparar o conjunto e entrada (*input*) com os dados globais para *M. esculenta* previamente analisados no portal PlantRegMap <http://plantregmap.cbi.pku.edu.cn/>).

4.10. Análise funcional de potenciais redes de interação de proteínas

Os grupos de proteínas potencialmente codificadas por genes hiper e hipometilados das respostas de cada variedade ao estresse aplicado foram caracterizados quanto às potenciais redes de interação proteína-proteína (PPI) previstas com a ferramenta STRING (versão 11.0b; <https://string-db.org/>), tendo *A. thaliana* como modelo devido sua alta gama de informações depositadas no banco e a falta de informações relevantes nas espécies mais próximas de *M. esculenta* como *Populus trichocarpa* e *Medicago truncatula*. As interações ativas incluíram dados experimentais, de co-expressão e de co-ocorrência, com *score* mínimo de 0,7 (*high confidence*) e número de iteradores máximo de 20

(1st shell) e 10 (2nd shell). Uma análise de enriquecimento funcional, dos principais processos biológicos, funções moleculares e componentes celulares, relativos aos componentes previstos nas redes foi feita a partir do método *Aggregate Fold Change* (Yu et al., 2017). Adicionalmente, além dos termos enriquecidos de GO para as proteínas das redes PPI, potenciais informações de vias metabólicas (da base de dados KEGG), de *clusters* locais hierarquizados (STRING), e de anotações de diversas bases (Reactome, UniProt, Pfam, SMART e InterPro), foram incorporadas, quando disponíveis ou relevantes.

4.11. Análises de SNP-INDEL

O presente estudo empregou uma abordagem sistemática para analisar dados genômicos brutos disponibilizados pela Embrapa, inicialmente avaliados quanto à qualidade por meio do Fastqc (ANDREWS et al., 2010). Subsequentemente, empregou-se o software Sickle (JOSHI 2020) para filtrar os dados com base em parâmetros de qualidade (Q20) e comprimento (20). Os dados filtrados foram então utilizados para o mapeamento no genoma de referência (JGI v6.0; Identificador *BioSample*: SAMN04116956) utilizando o algoritmo BWA (LI e DURBIN, 2009). A chamada de variantes foi realizada empregando-se o GATK4 (MCKENNA et al., 2010), resultando na obtenção de arquivos VCF.

Para a filtragem das variantes, foi aplicado a ferramenta BCFTOOLS (NARASIMAH et al., 2016) com parâmetros estabelecidos em QUAL>30 e MQ>30 nos arquivos VCF. A fim de atribuir significado biológico às variantes e prever seus impactos funcionais, a anotação foi conduzida utilizando o software SNPeff (CINGOLANI et al., 2012). Para viabilizar esse processo, um banco de dados local foi criado, sendo referente à região e aos efeitos previstos das variantes, com base em um banco de dados pré-existente para o genoma de referência utilizado nas montagens. É relevante destacar que a escolha do genoma de referência se baseou em montagens prévias, garantindo uma abordagem robusta e alinhada com os padrões da comunidade científica. Essa metodologia integrada proporciona uma análise abrangente e confiável das variantes genômicas, permitindo insights valiosos para a compreensão da variabilidade genética na população estudada.

5. Resultados

5.1. Os totais do sequenciamento das bibliotecas MethylRAD e os percentuais de *tags* mapeadas no genoma de referência de *M. esculenta*

Os totais de dados MethylRAD da plataforma de sequenciamento conforme fornecidos pela empresa CD Genomics (USA) geradora dos dados (Tabela 2) indicaram que as *tags* dos sítios esperados mapearam no genoma de *M. esculenta* em percentuais que variaram de 17,90 a 21,24%, de acordo com o tratamento e variedade.

Tabela 2. Dados totais do sequenciamento MethylRAD de *leituras* de tamanho esperado com um único sítio de mapeamento no genoma de referência (*Enzyme Leituras*) de 12 amostras/tratamentos e quantidades mapeadas em genoma de *Manihot esculenta* (*Mapping Leituras*), e respectivas proporções (%). Tratamentos descritos: FORSEV (Formosa, estresse severo); FORCONT (Formosa, controle sem estresse); DOUSEV (Dourada, estresse severo); DOUCONT (Dourada, controle sem estresse).

Amostra	EnzymeReads	Leituras mapeadas	Proporção
A01-FORSEV R1	8.473.392	1.516.553	17,90%
A02-FORSEV R2	7.894.473	1.643.074	20,82%
A03-FORSEV R3	8.855.136	1.667.859	18,84%
A04-FORMCONT R1	8.703.725	1.764.280	20,27%
A05-FORMCONT R2	7.596.265	1.525.598	20,09%
A06-FORMCONT R3	6.442.813	1.240.765	19,25%
A07-DOUSEV R1	6.467.929	1.373.931	21,24%
A08-DOUSEV R2	7.530.213	1.364.509	18,12%
A09-DOUSEV R3	7.419.768	1.434.261	19,33%
A010-DOUCONT R1	6.567.305	1.220.180	18,58%
A011-DOUCONT R2	5.902.892	1.222.560	20,71%
A012-DOUCONT R3	5.095.637	1.056.007	20,72%

5.2. As profundidades de sequenciamento das bibliotecas MethylRAD

Os totais de cada sítio de metilação (CCGG ou CCNGG), bem como as profundidades médias observadas nos sequenciamentos de cada biblioteca, são vistas na **Tabela 3**. Para o sítio CCGG, as profundidades médias variaram de 15,67 (FORMCONT-R3) a 21,25% (FORMCONT-R1) e para o sítio CCNGG, estas variaram de 11,56 (DOURCONT-R2) a 14,77% (FORMCONT-R2). Em todos os tratamentos, as profundidades de sequenciamento foram maiores do

que o mínimo estabelecido de três de acordo com os parâmetros de qualidade estipulados pela empresa geradora dos dados CD Genomics.

Tabela 3. Totais de sítios metilados (com profundidade de sequenciamento maior que 3), e a profundidade média de sequenciamento observada em cada biblioteca MethylRAD. DOU: Dourada; FOR: Formosa; CONT: Controle; SEV: Estresse Severo; R1, R2 e R3: triplicatas biológicas.

Amostra	CCGG		CCNGG	
	Número de sítios	Profundidade	Número de sítios	Profundidade
A01-FORSEV-R1	40.104	17,26	33.321	14,1
A02-FORSEV-R2	41.467	20,7	28.055	14,59
A03-FORSEV-R3	42.306	17,55	34.667	15,63
A04-FORMCONT-R1	43.275	21,25	30.921	14,43
A05-FORMCONT-R2	38.463	18,1	32.124	14,77
A06-FORMCONT-R3	35.957	15,67	29.918	12,45
A07-DOUSEV-R1	36.555	19,66	25.051	13,13
A08-DOUSEV-R2	36.433	16,35	31.662	13,83
A09-DOUSEV-R3	36.084	18,17	30.887	14,16
A010-DOUCONT-R1	31.094	17,47	28.279	13,39
A011-DOUCONT-R2	34.724	17,44	27.314	11,56
A012-DOUCONT-R3	29.323	17,69	21.228	12,73

5.3. A distribuição dos sítios metilados nos elementos gênicos

A distribuição dos sítios metilados CCGG e CCNGG nos elementos gênicos analisados está ilustrada nas **Figuras 3 e 4**, respectivamente. Em todos os tratamentos, a detecção expressiva de sítios metilados CCGG ocorreu em regiões intergênicas, éxons, íntrons e *upstream* (**Figura 3**), ao passo que em sítios CCNGG, as maiores detecções ocorreram em regiões intergênicas, *upstream*, íntrons e éxons (**Figura 4**). Nos dois casos, sítios metilados em regiões de UTR foram os menos detectados.

Figura 3. Totais de sítios metilados do tipo CCGG em diferentes regiões gênicas do genoma de *M. esculenta* [upstream: região de 2.000 pb anterior ao início da transcrição]. A01 (FORSEV-R1); A02 (FORSEV-R2); A03 (FORSEV-R3); A04 (FORCONT-R1); A05 (FORCONT-R2); A06 (FORCONT-R3); A07 (DOUSEV-R1); A08 (DOUSEV-R2); A09 (DOUSEV-R3); A10 (DOUCONT-R1); A11 (DOUCONT-R2); A12 (DOUCONT-R3). DOU: Dourada; FOR: Formosa; CONT: Controle; SEV: Estresse Severo; R1, R2 e R3: triplicatas biológicas.

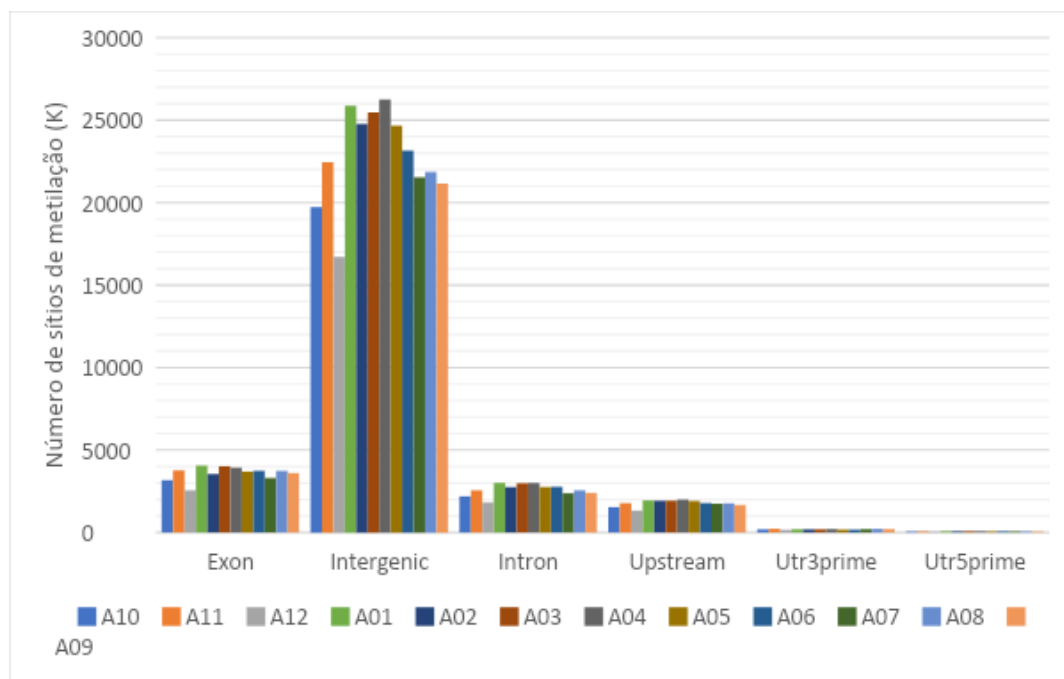
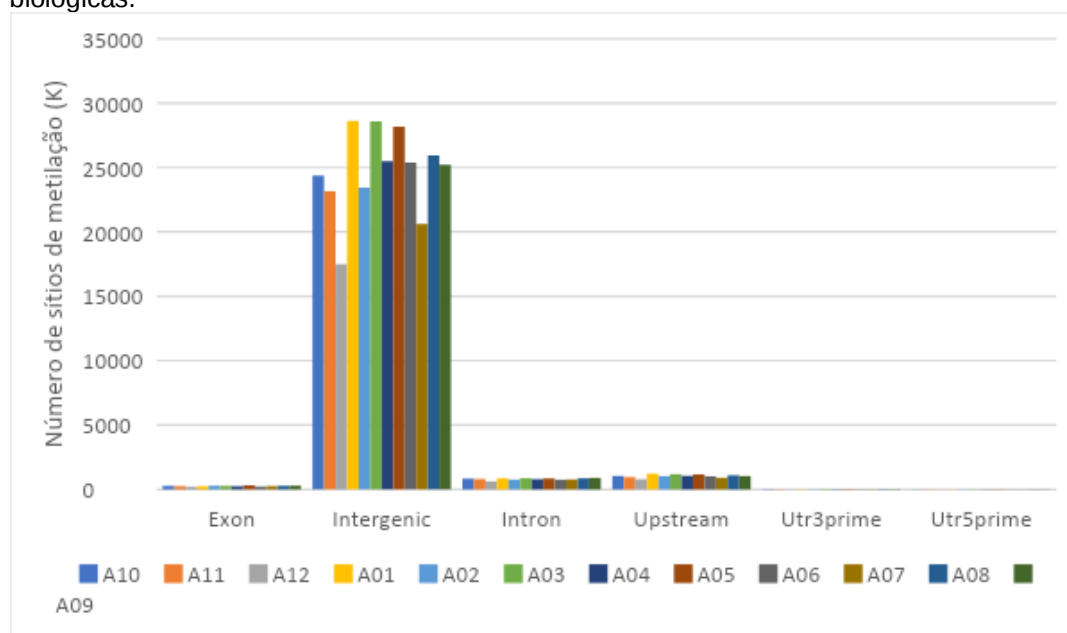


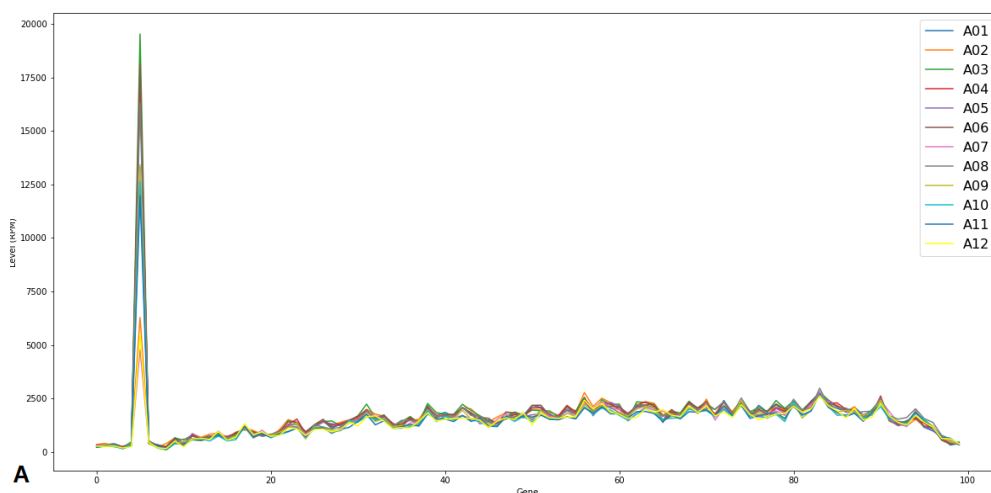
Figura 4. Totais de sítios metilados do tipo CCNGG em diferentes regiões gênicas do genoma de *M. esculenta* [upstream: região de 2.000 pb anterior ao início da transcrição]. A01 (FORSEV-R1); A02 (FORSEV-R2); A03 (FORSEV-R3); A04 (FORCONT-R1); A05 (FORCONT-R2); A06 (FORCONT-R3); A07 (DOUSEV-R1); A08 (DOUSEV-R2); A09 (DOUSEV-R3); A10 (DOUCONT-R1); A11 (DOUCONT-R2); A12 (DOUCONT-R3). DOU: Dourada; FOR: Formosa; CONT: Controle; SEV: Estresse Severo; R1, R2 e R3: triplicatas biológicas.



5.4. A distribuição dos sítios de metilação em genes

Em genes, os maiores níveis de sítios CCGG e CCNGG metilados, tendo por base dados de RPM, mostraram ocorrências nos 20% iniciais dos comprimentos dos genes (**Figuras 5A e 5B**, respectivamente). Ainda, as metilações tanto de sítios CCGG quanto de sítios CCNGG foram detectadas próximas aos inícios das transcrições (TSS; **Figuras 6A e 6B**, respectivamente), e cerca de 500 pb a montante (5') dos términos de transcrição (TTS; **Figuras 7A e 7B**), sendo adicionalmente detectada para sítios CCGG um segundo pico cerca de 1500 pb a jusante (3') aos TTS.

Figura 5. Distribuição dos sítios de metilação ao longo dos genes, com base em dados de RPM. A coordenada X representa a posição relativa no gene (isto é, 20 significa 20% das posições iniciais do gene). A: análise para sítios CCGG; B: análise para sítios CCNGG. Legenda: A01 (FORSEV-R1); A02 (FORSEV-R2); A03 (FORSEV-R3); A04 (FORCONT-R1); A05 (FORCONT-R2); A06 (FORCONT-R3); A07 (DOUSEV-R1); A08 (DOUSEV-R2); A09 (DOUSEV-R3); A10 (DOUCONT-R1); A11 (DOUCONT-R2); A12 (DOUCONT-R3). DOU: Dourada; FOR: Formosa; CONT: Controle; SEV: Estresse Severo; R1, R2 e R3: triplicatas biológicas



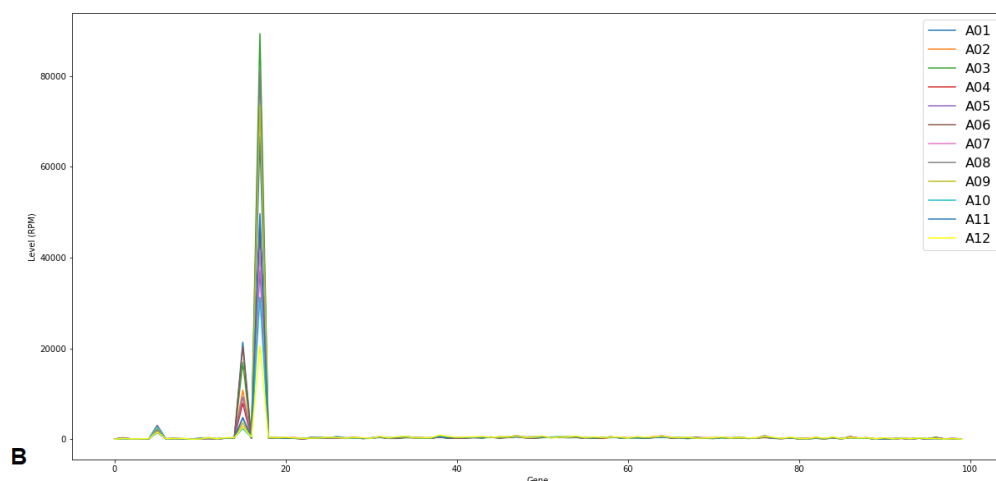
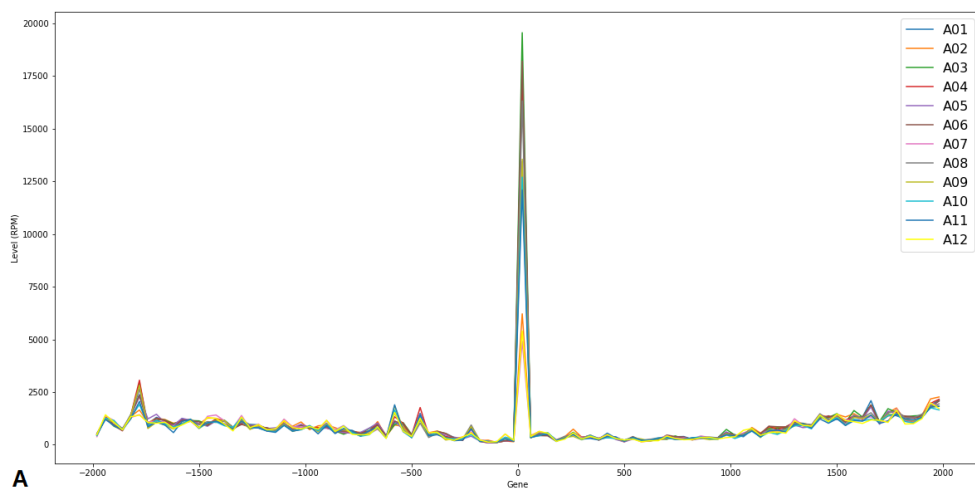
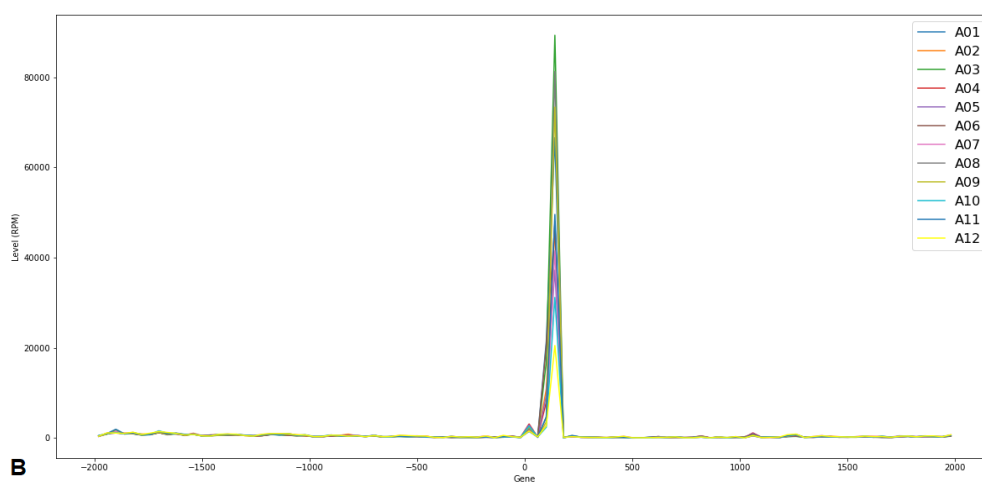


Figura 6. Distribuição dos sítios de metilação ao longo das regiões de início da transcrição (TSS). Dados referentes a contagem de RPM. A coordenada X representa a janela de 2000 pb antes e depois do início do começo da transcrição. O gráfico A representa a análise para sítio CCGG; B representa a análise para sítio CCNGG. Legenda: A01 (FORSEV-R1); A02 (FORSEV-R2); A03 (FORSEV-R3); A04 (FORCONT-R1); A05 (FORCONT-R2); A06 (FORCONT-R3); A07 (DOUSEV-R1); A08 (DOUSEV-R2); A09 (DOUSEV-R3); A10 (DOUCONT-R1); A11 (DOUCONT-R2); A12 (DOUCONT-R3). DOU: Dourada; FOR: Formosa; CONT: Controle; SEV: Estresse Severo; R1, R2 e R3: triplicatas biológicas

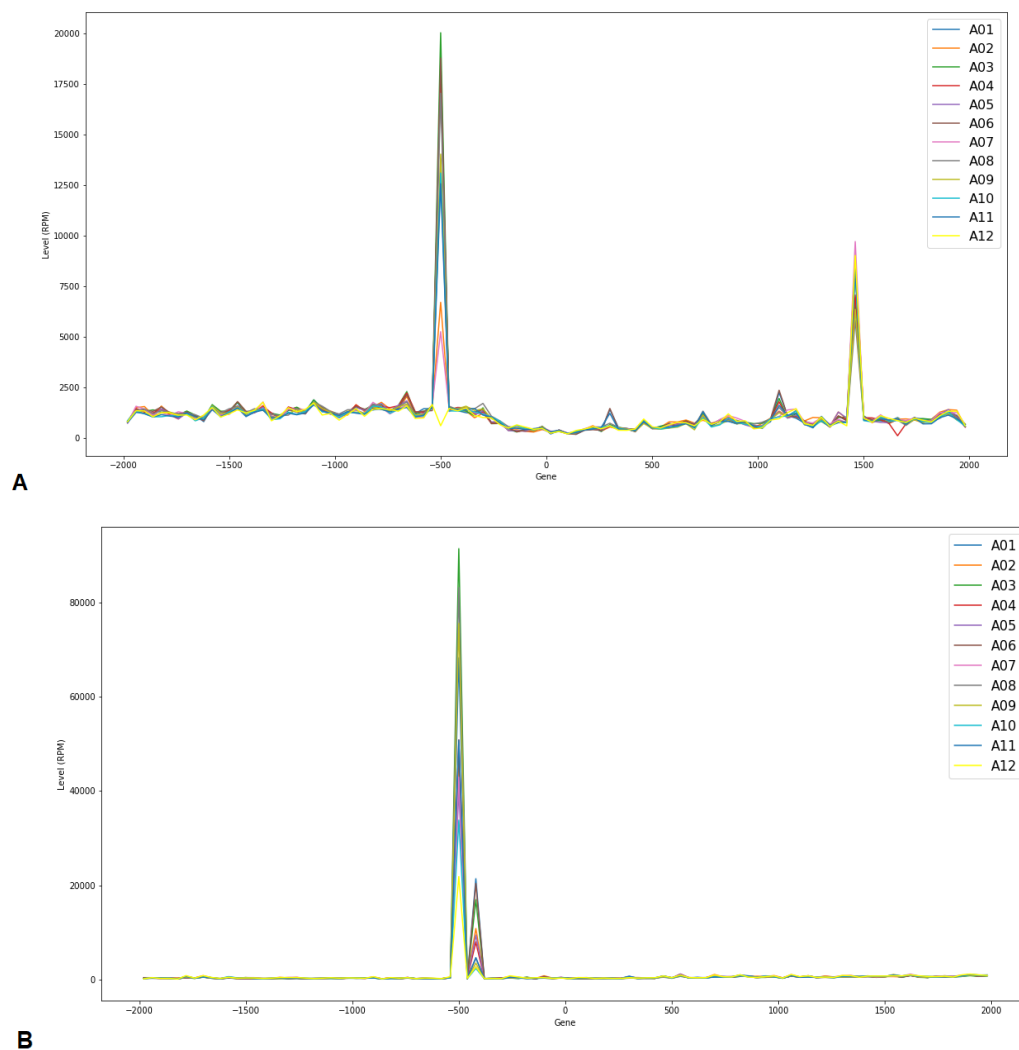


A



B

Figura 7. Distribuição dos sítios de metilação ao longo das regiões de término da transcrição (TTS). Dados referentes a contagem de RPM. A coordenada X representa a janela de 2000 pb antes e depois do início do começo da transcrição. O gráfico A representa a análise para sítio CCGG; B representa a análise para sítio CCNGG. Legenda: A01 (FORSEV-R1); A02 (FORSEV-R2); A03 (FORSEV-R3); A04 (FORCONT-R1); A05 (FORCONT-R2); A06 (FORCONT-R3); A07 (DOUSEV-R1); A08 (DOUSEV-R2); A09 (DOUSEV-R3); A10 (DOUCONT-R1); A11 (DOUCONT-R2); A12 (DOUCONT-R3). DOU: Dourada; FOR: Formosa; CONT: Controle; SEV: Estresse Severo; R1, R2 e R3: triplicatas biológicas



5.5. Correlação dos níveis de metilação entre diferentes amostras

Os níveis de metilação detectados nas diferentes bibliotecas, os quais refletem as profundidades dos sequenciamentos, são vistos na **Figura 8** (para o sítio CCGG) e na **Figura 9** (para o sítio CCNCGG). Nestas figuras, gráficos de dispersão comparando as profundidades dos sequenciamentos de duas amostras, uma na abscissa e outra na ordenada, permitiram estimar coeficientes de correlação de Pearson. Essa análise foi útil pois permitiu observar as consistências entre as réplicas biológicas de cada variedade. Nas comparações entre réplicas biológicas da variedade Dourada (sensível) na condição controle, para o sítio CCGG (**Figura 8**), os coeficientes foram $r \geq 0,86$, e na condição estresse, $r \geq 0,92$. Em ambos os casos, os dados das réplicas foram consistentes. Para Formosa e para o mesmo sítio, as réplicas biológicas na condição controle foram $r \geq 0,94$, e para a condição estresse, $r \geq 0,85$ (**Figura 8**). Para o sítio CCNCGG (**Figura 9**), as réplicas biológicas da variedade Dourada na condição controle apresentaram $r \geq 0,71$, e na condição estresse foi de $r \geq 0,82$. Para a variedade Formosa e mesmo sítio, as réplicas biológicas na condição controle apresentaram $r \geq 0,96$, e na condição estresse, $r \geq 0,88$. Logo, os dados das réplicas biológicas foram consistentes nas condições estudadas das duas variedades, permitindo fazer as devidas comparações.

Figura 8. Gráfico de correlação de Pearson para sítio CCGG de cada réplica. O triângulo superior esquerdo é o gráfico de dispersão do nível de metilação entre duas amostras que estão na diagonal da imagem, a região triangular inferior direita é o correspondente coeficiente de correlação de Pearson. Legenda: A01 (FORSEV-R1); A02 (FORSEV-R2); A03 (FORSEV-R3); A04 (FORCONT-R1); A05 (FORCONT-R2); A06 (FORCONT-R3); A07 (DOUSEV-R1); A08 (DOUSEV-R2); A09 (DOUSEV-R3); A10 (DOUCONT-R1); A11 (DOUCONT-R2); A12 (DOUCONT-R3). DOU: Dourada; FOR: Formosa; CONT: Controle; SEV: Estresse Severo; R1, R2 e R3: triplicatas biológicas.

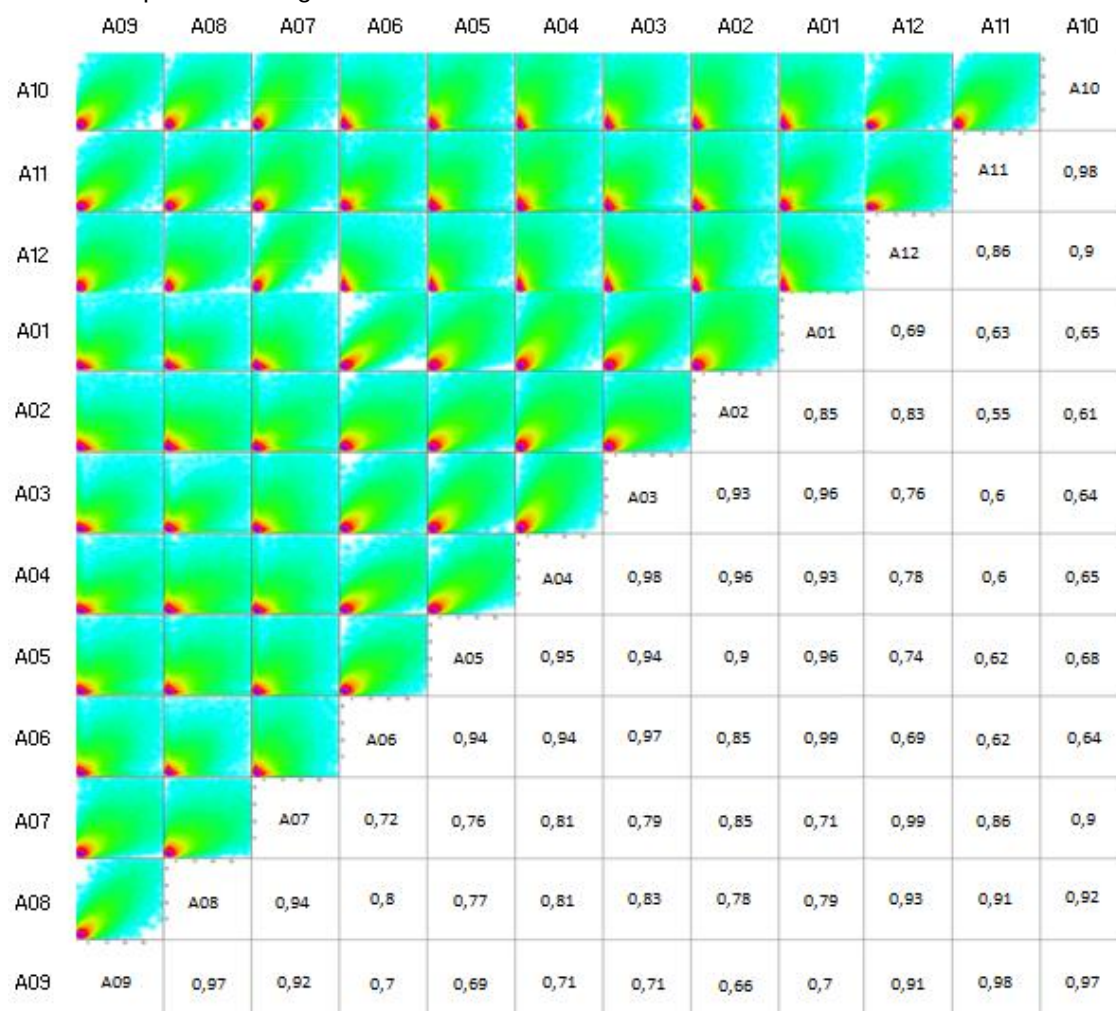
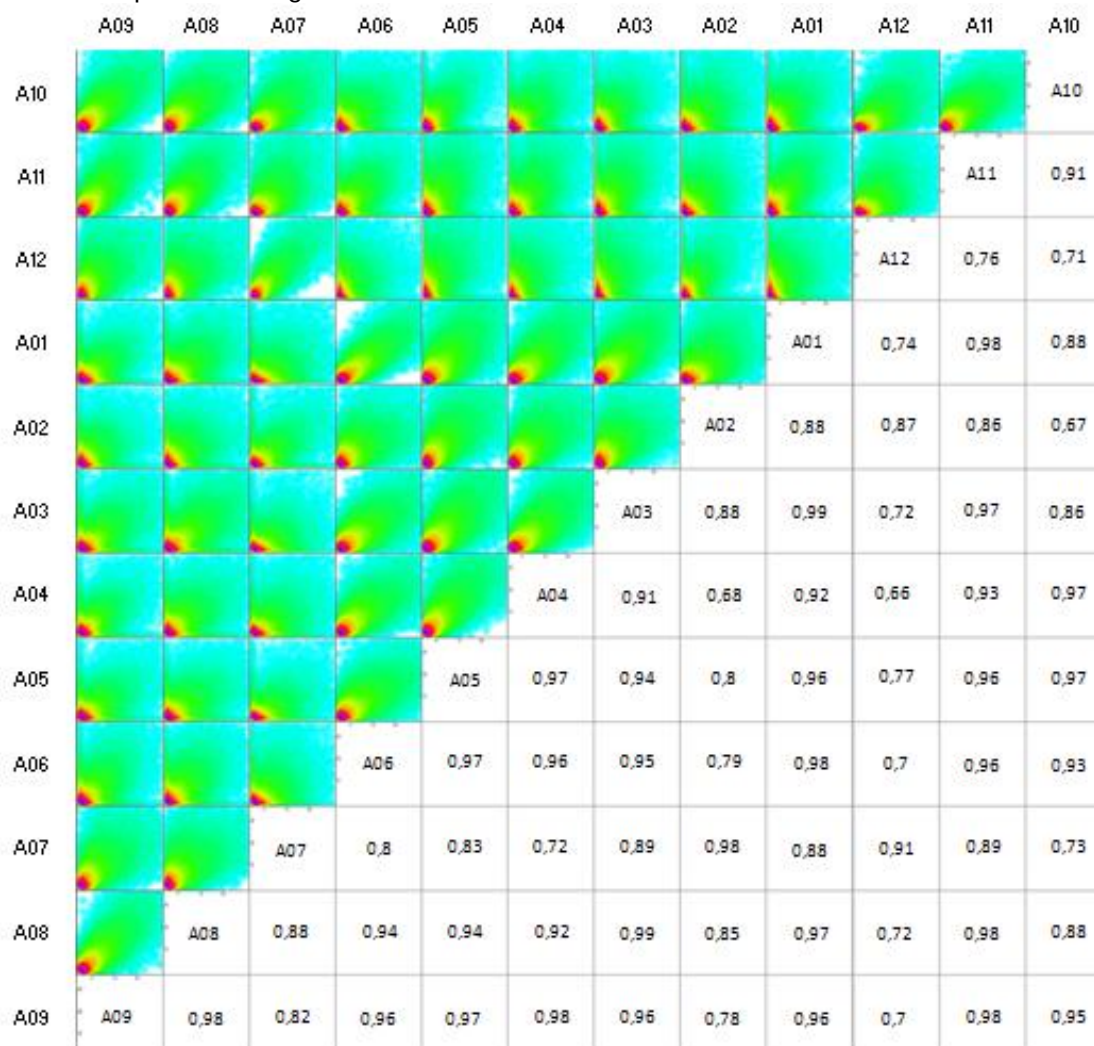


Figura 9. Gráfico de correlação de Pearson para sítio CCNGG de cada réplica. O triângulo superior esquerdo é o gráfico de dispersão do nível de metilação entre duas amostras que estão na diagonal da imagem, a região triangular inferior direita é o correspondente coeficiente de correlação de Pearson. Legenda: A01 (FORSEV-R1); A02 (FORSEV-R2); A03 (FORSEV-R3); A04 (FORCONT-R1); A05 (FORCONT-R2); A06 (FORCONT-R3); A07 (DOUSEV-R1); A08 (DOUSEV-R2); A09 (DOUSEV-R3); A10 (DOUCONT-R1); A11 (DOUCONT-R2); A12 (DOUCONT-R3). DOU: Dourada; FOR: Formosa; CONT: Controle; SEV: Estresse Severo; R1, R2 e R3: triplicatas biológicas



5.6. Os sítios diferencialmente metilados de cada variedade nas comparações dos tratamentos com e sem o estresse aplicado

Os totais de sítios (CCGG e CCNGG) metilados dos contrastes envolvendo os tratamentos com e sem o estresse de cada variedade são dados na **Tabela 4**, bem como os totais de genes diferencialmente metilados (hiper ou hipometilados). De acordo com a tabela, a variedade sensível (Dourada) apresentou mais sítios metilados, mais genes diferencialmente metilados, mais genes hipermetilados, e mais genes hipometilados do que a tolerante Formosa.

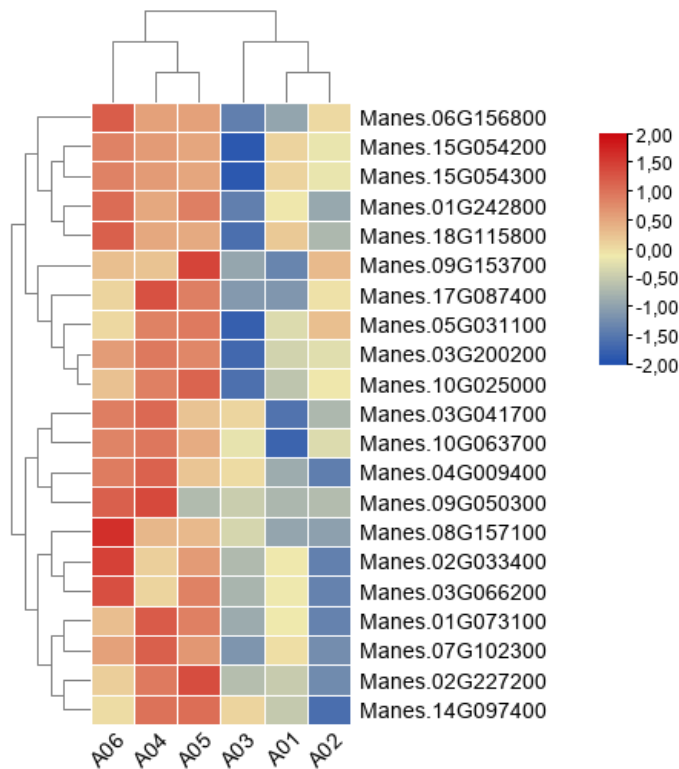
No geral, os locais nos genes diferencialmente metilados envolveram mais sítios CCGG do que CCNGG, independentemente da variedade estudada.

Tabela 4. Totais de sítios CCGG e CCNGG metilados no genoma de *Manihot esculenta* e os diferencialmente metilados (hiper e hipometilados) em genes para as variedades Formosa e Dourada ao se comparar os respectivos tratamentos com o estresse e sem o estresse de seca.

Contraste	Sítio	Total	Genes	Hipermetilados	Hipometilados
Formosa (FORSEV vs FORCONT)					
	CCGG	775	63	42	21
	CCNGG	807	14	9	5
	Subtotal	1582	77	51	26
Dourada (DOUSEV vs DOUCONT)					
	CCGG	858	90	56	34
	CCNGG	783	23	13	10
	Subtotal	1641	113	69	44
	Total	3223	203	120	70

Gráficos *heatmaps*, gerados a partir da transformação logarítmica (Log_2) das contagens RPM, mostrando os perfis de metilação para cada sítio das réplicas biológicas (R1, R2 e R3) de cada variedade, nas condições de estresse ou controle sem estresse, para os conjuntos de genes declarados diferencialmente metilados são apresentados nas **Figuras suplementares 1 a 8** no apêndice. Assim, é possível ver no exemplo abaixo, contemplando o conjunto de genes hipermetilados nos sítios CCGG, a consistência dos perfis de metilação das réplicas biológicas da variedade Dourada (sensível), na condição de estresse ou sem estresse.

Figura 10. Gráfico heatmap de genes hipometilados ($p\text{-value} \leq 0,05$ e $\log_2FC > 1$) em sítios CCGG da variedade Formosa (FOR), apresentado pelas réplicas biológicas (R1, R2 e R3) na condição de seca (SEV) ou de controle (CONT) sem estresse, tendo por base $\log_2$ das contagens RPM (leituras por milhão). Legenda: A01 (FORSEV-R1); A02 (FORSEV-R2); A03 (FORSEV-R3); A04 (FORCONT-R1); A05 (FORCONT-R2); A06 (FORCONT-R3)



Os genes diferencialmente metilados e respectivas anotações são apresentados na **Tabela Suplementar 1**. Na **Tabela 5** são visualizadas as anotações para os genes diferencialmente metilados representados no heatmap da **Figura 10**.

Tabela 5. Identificadores de genes diferencialmente hipometilados de sítios CCGG na variedade tolerante (Formosa) e suas respectivas anotações baseadas nos dados do genoma de versão 6.1 de *Manihot esculenta* da base de dados Phytozome.

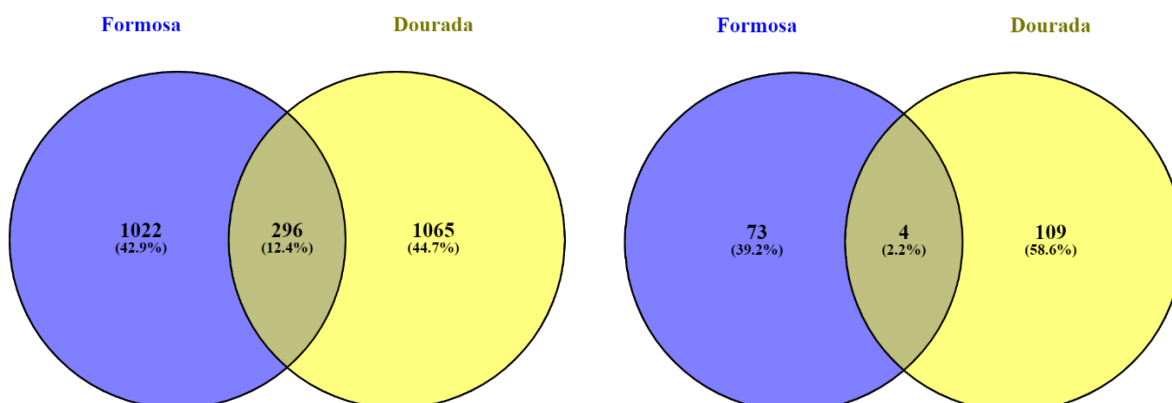
Identificador	Anotação
Manes.01G073100	DHHC-type zinc finger family protein
Manes.01G242800	P-loop containing nucleoside triphosphate hydrolases superfamily protein
Manes.02G033400	cullin 1
Manes.02G227200	cinnamate-4-hydroxylase
Manes.03G041700	inter-alpha-trypsin inhibitor heavy chain-related
Manes.03G066200	helicases;ATP-dependent helicases;nucleic acid binding;ATP binding;DNA-directed DNA polymerases;DNA binding
Manes.03G200200	Pectin lyase-like superfamily protein
Manes.04G009400	cellulose synthase-like A2
Manes.05G031100	

Manes.06G156800	manganese tracking factor for mitochondrial SOD2
Manes.07G102300	auxin response factor 6
Manes.08G157100	Major facilitator superfamily protein
Manes.09G050300	NAD(P)-binding Rossmann-fold superfamily protein
Manes.09G153700	ATPase E1-E2 type family protein / haloacid dehalogenase-like hydrolase family protein
Manes.10G025000	Glucose-1-phosphate adenylyltransferase family protein
Manes.10G063700	P-loop containing nucleoside triphosphate hydrolases superfamily protein
Manes.14G097400	NAC (No Apical Meristem) domain transcriptional regulator superfamily protein
Manes.15G054200	Carbohydrate-binding X8 domain superfamily protein
Manes.15G054300	purple acid phosphatase 15
Manes.17G087400	hydroxyproline-rich glycoprotein family protein
Manes.18G115800	O-acyltransferase (WSD1-like) family protein

5.7. Genes diferencialmente metilados exclusivos e comuns entre contrastes

Uma análise a partir dos identificadores Manes relativo ao genoma de *M. esculenta* associados aos sítios diferencialmente metilados (**Figura 11A**) evidenciou que a maioria dos locais com metilação diferencial foram exclusivos de uma ou outra variedade, com somente 12,4% dos locais metilados foram compartilhados pelas duas variedades. Os resultados indicaram que as variedades apresentaram perfis de metilação majoritariamente distintos. O mesmo padrão foi observado quando se comparou identificadores relacionados somente com genes diferencialmente metilados em cada variedade (**Figura 11B**).

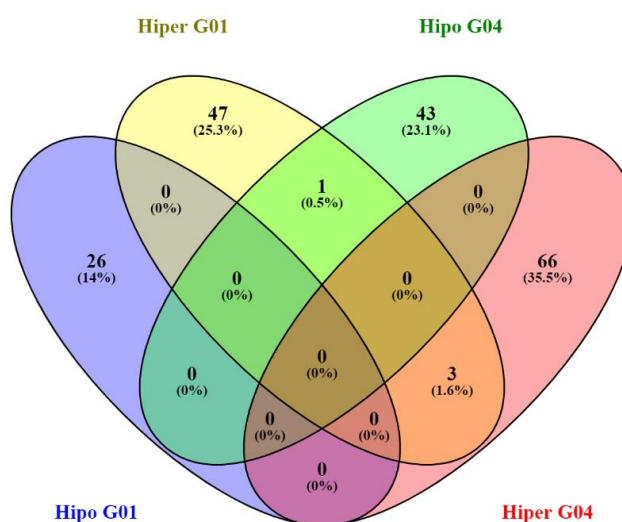
Figura 11. Diagrama de Venn utilizando identificadores Manes (*Manihot esculenta*, Phytozome v.12.1) com sítios de metilação diferencial detectados para as variedades Formosa (tolerante) e Dourada (sensível) na comparação do contraste com e sem o estresse de seca. **A:** todos os identificadores detectados com sítios metilados (CCGG e CCNGG); **B:** identificadores metilados relativos a genes.



Ao se comparar os conjuntos de genes hipermetilados e hipometilados detectados no contraste dos tratamentos com estresse e sem estresse de cada

variedade, em diagrama de Venn (**Figura 12**), observou-se apenas um gene (Manes.02G059500) hipermetilado em Formosa (tolerante) e hipometilado em Dourada (sensível) e três outros (Manes.01G181600, Manes.03G010700 e Manes.18G026700) hipermetilados em ambos, sendo os demais exclusivos de cada resposta. Assim, aparentemente, cada variedade apresentou perfis de metilação distintos após a aplicação do estresse.

Figura 12. Diagrama de Venn utilizando identificadores Manes (*Manihot esculenta*, Phytozome v.12.1) com sítios (CCGG e CCNGG) hiper e hipometilados em genes, detectados na comparação do contraste com e sem o estresse de seca para a variedade tolerante Formosa (FOR) e para a sensível Dourada (DOU).



5.8. Análise de Termos de GO

A análise de ontologia gênica dos genes hiper ou hipometilados da resposta de cada variedade, de acordo com os identificadores Manes de *M. esculenta* (v6.1, Phytozome 12.1) revelou 13 termos enriquecidos relativos a 41 genes de um total de 51 detectados hipermetilados nos dois sítios (CCGG ou CCNGG) da variedade tolerante Formosa após o estresse de seca, ao passo que sete termos enriquecidos foram associados a 20 dos 26 genes hipometilados nos dois sítios. Em relação a variedade Dourada (sensível), o enriquecimento revelou três termos enriquecidos associados a 56 genes de um total de 69 genes

hipermetilados nos dois sítios, enquanto que 18 foram associados a 32 dos 44 genes hipometilados também nos dois sítios. Os termos enriquecidos encontram-se descritos na **Tabela Suplementar 2A e 2B**.

5.9. Rede de interações previstas a partir dos conjuntos de genes diferencialmente metilados de cada variedade após o estresse de seca

5.9.1. Rede de interações predita para proteínas de genes hipometilados na variedade tolerante Formosa

A rede de interações derivado de potenciais produtos dos genes hipometilados (em sítios CCGG e CCNGG) da resposta da variedade tolerante Formosa ressaltou cinco proteínas de *A. thaliana*, consideradas ortólogas dentre 26 proteínas de *M. esculenta* relacionadas aos genes hipometilados. A rede de interações apresentou 56 nós e 95 arestas, sendo a média de nós de 3,39 e o coeficiente médio de agrupamento local de 0,495, sendo relacionadas somente as relações dos componentes decorrentes basicamente de ocorrência experimental e dados de coexpressão e/ou de coocorrência (**Figura 13**).

A caracterização através de termos de GO fornecida pela ferramenta STRING evidenciou 28 termos enriquecidos relativos a processo biológico, seis para função molecular, e 17 para componente celular (**Tabela Suplementar 3A, 3B e 3C**), ao passo que o enriquecimento associando componentes às vias metabólicas da plataforma KEGG (**Tabela Suplementar 3D**), evidenciou 10 proteínas associadas com espliceossoma, um cluster local para sinalossomo, e algumas interações previstas com a via de sinalização por auxina (**Figura 14**).

Numericamente, as proteínas do arquivo de entrada indicaram de duas a oito interações cada (**Tabela 6**).

Figura 13. Rede de interações predita para produtos de genes hipometilados (sítios CCGG e CCNGG) na variedade Formosa (Manihot esculenta) após estímulo de déficit hídrico, tendo por base proteínas ortólogas de *A. thaliana* e a ferramenta STRING. As esferas representam proteínas e as arestas, as interações (rosa: experimental; preto: coexpressão; azul: coocorrência). Proteínas do arquivo de entrada são identificadas pelo prefixo "Manes". Cores das esferas representam termos GO para processos biológicos (Azul-escuro: modificação de proteínas por conjugação ou remoção; verde-claro: vias de ativação da sinalização por auxina; amarelo: *splicing* de RNA; vermelho: resposta hormonal).

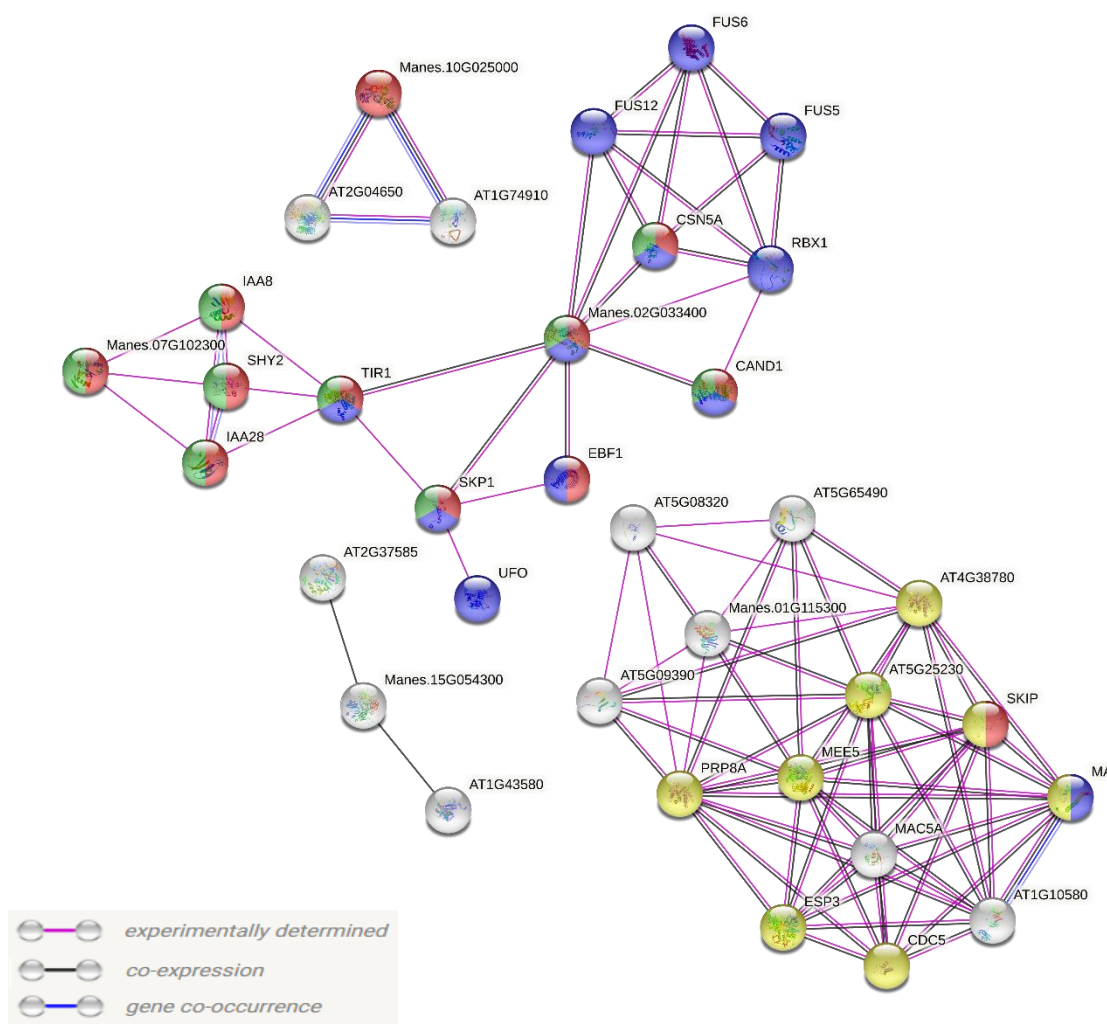


Figura 14. Rede de interações predita para produtos de genes hipometilados (sítios CCGG e CCNGG) na variedade Formosa (*Manihot esculenta*) após estímulo de déficit hídrico, tendo por base proteínas ortólogas de *A. thaliana* e a ferramenta STRING. Esferas representam proteínas e as arestas, as interações (rosa: experimental; preto: coexpressão; azul: coocorrência). Proteínas do arquivo de entrada são identificadas pelo prefixo “Manes”. Cores das esferas representam Keyword (Amarelo – vias de sinalização por auxina), vias do KEGG (vermelho – Spliceossoma), cluster local (verde – Sinalossoma).

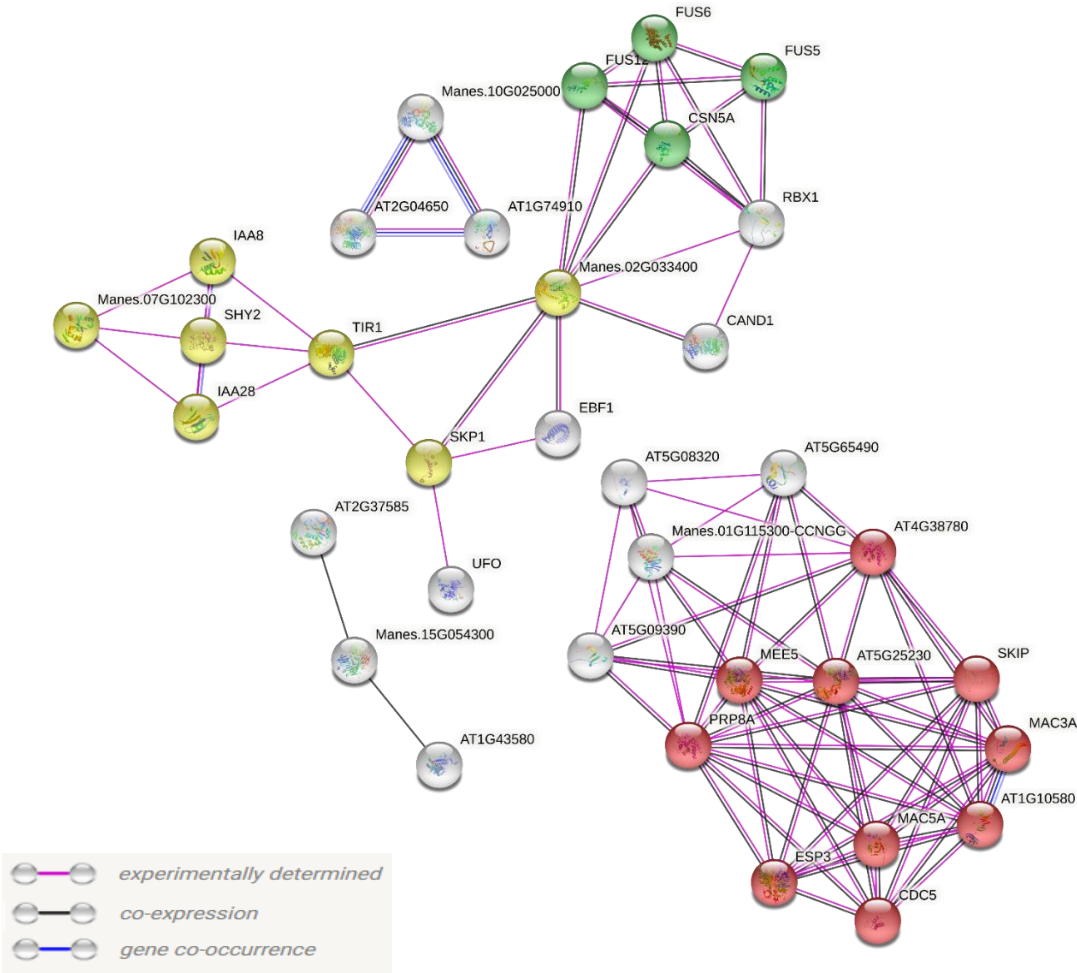


Tabela 6. Interações previstas com proteínas derivadas de genes hipometilados na variedade tolerante Formosa (*Manihot esculenta*), após um estímulo de déficit hídrico, tendo por base a ferramenta String para *Arabidopsis thaliana*.

Gene de entrada/proteína	Interações
Manes.01G115300	7
Manes.02G033400	8
Manes.15G054300	2
Manes.10G025000	2
Manes.07G102300	3

5.9.2. Rede de interações predita para proteínas de genes hipermetilados na variedade tolerante Formosa

A rede de interações derivado de potenciais produtos dos genes hipermetilados (em sítios CCGG e CCNGG) da resposta da variedade tolerante Formosa ressaltou cinco proteínas de *A. thaliana*, consideradas ortólogas dentre 47 proteínas de *M. esculenta* relacionadas aos genes hipometilados. A rede de interações apresentou 77 nós e 180 arestas, sendo a média de nós de 4,68 e o coeficiente médio de agrupamento local de 0,444, sendo relacionadas somente as relações dos componentes decorrentes basicamente de ocorrência experimental e dados de coexpressão e/ou de coocorrência (**Figura 15**). A análise de termos de GO associada a ferramenta STRING evidenciou 68 termos enriquecidos relativos a processo biológico, 13 para função molecular, e 25 para componente celular (**Tabela Suplementar 4A, 4B e 4C**), ao passo que o enriquecimento associando os componentes a potenciais vias metabólicas da plataforma KEGG (**Tabela Suplementar 4D**) evidenciou 15 proteínas interagindo com fatores de transcrição (FT) basais, 12 com proteassoma, e cinco com degradação de RNAs (**Figura 16**).

Numericamente, as proteínas apresentaram interações que variaram de uma a 13 (**Tabela 7**).

Figura 15. Rede de interação proteína-proteína predita para produtos de genes hipermetilados (sítios CCGG e CCNGG), da variedade Formosa (*Manihot esculenta*), após um estímulo de déficit hídrico, conforme modelo de *A. thaliana* e ferramenta STRING. Nós (esferas) representam proteínas, e arestas, as interações (em rosa, as experimentais; em preto, as de coexpressão; em azul, as de coocorrência). Proteínas do arquivo de entrada apresentam prefixo “Manes”. Cores das esferas representam termos GO para processos biológicos (em vermelho, processo metabólico primário; em verde-claro, regulação da expressão gênica; em amarelo, proteólise; em lilás, regulação positiva de processos biológicos; em verde-escuro, regulação positiva de processos metabólicos de macromoléculas; em azul-escuro, regulação positiva do processo metabólico de compostos de nitrogênio; em azul-claro, resposta a íons metálicos).

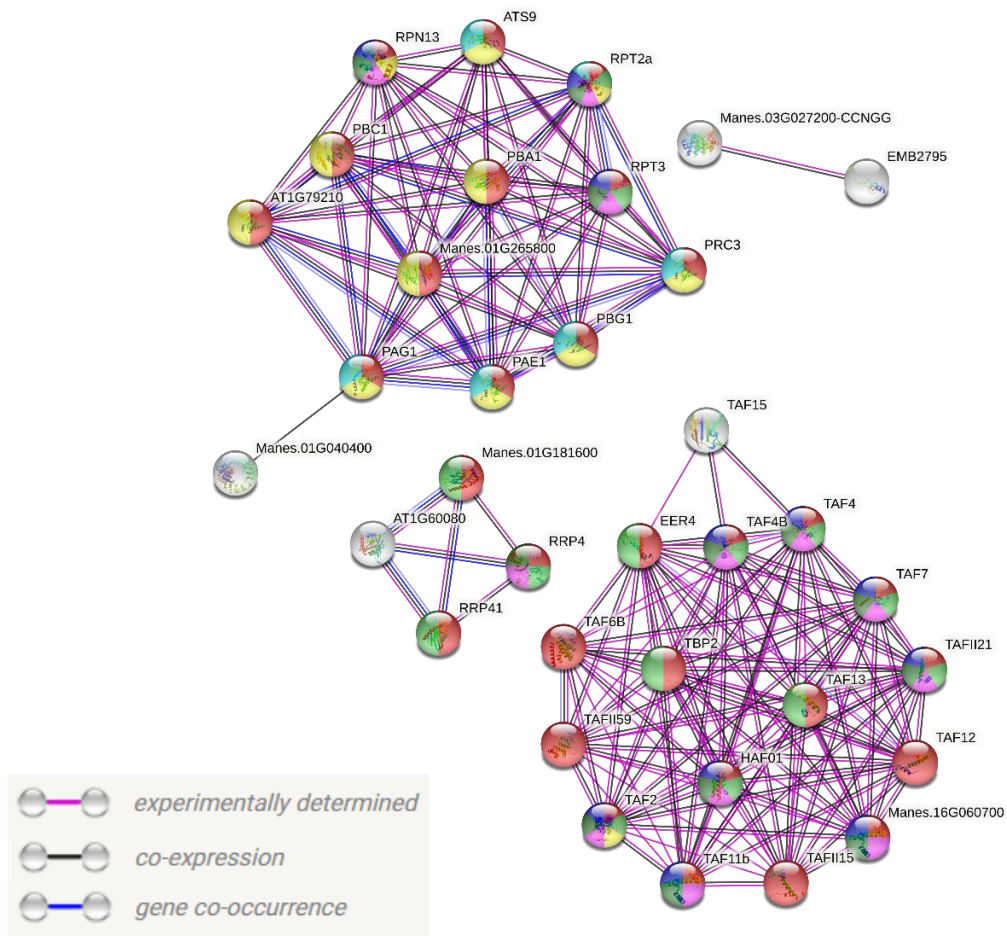


Figura 16. Rede de interação proteína-proteína obtida e visualizada pelo STRING para *A. thaliana* com base em genes hipermetilados da resposta ao estímulo de déficit hídrico em *M. esculenta* (variedade: Formosa). Na rede os nós representam proteínas e as arestas representam as interações na rede (em rosa, as experimentais; em preto, as de coexpressão; em azul, as de coocorrência). Proteínas do arquivo de entrada consistem nas proteínas de prefixo “Manes”. Cores das esferas representam termos ressaltados da base de dados KEGG (em azul, proteassoma; em verde, degradação do RNA; em vermelho, fatores de transcrição basais).

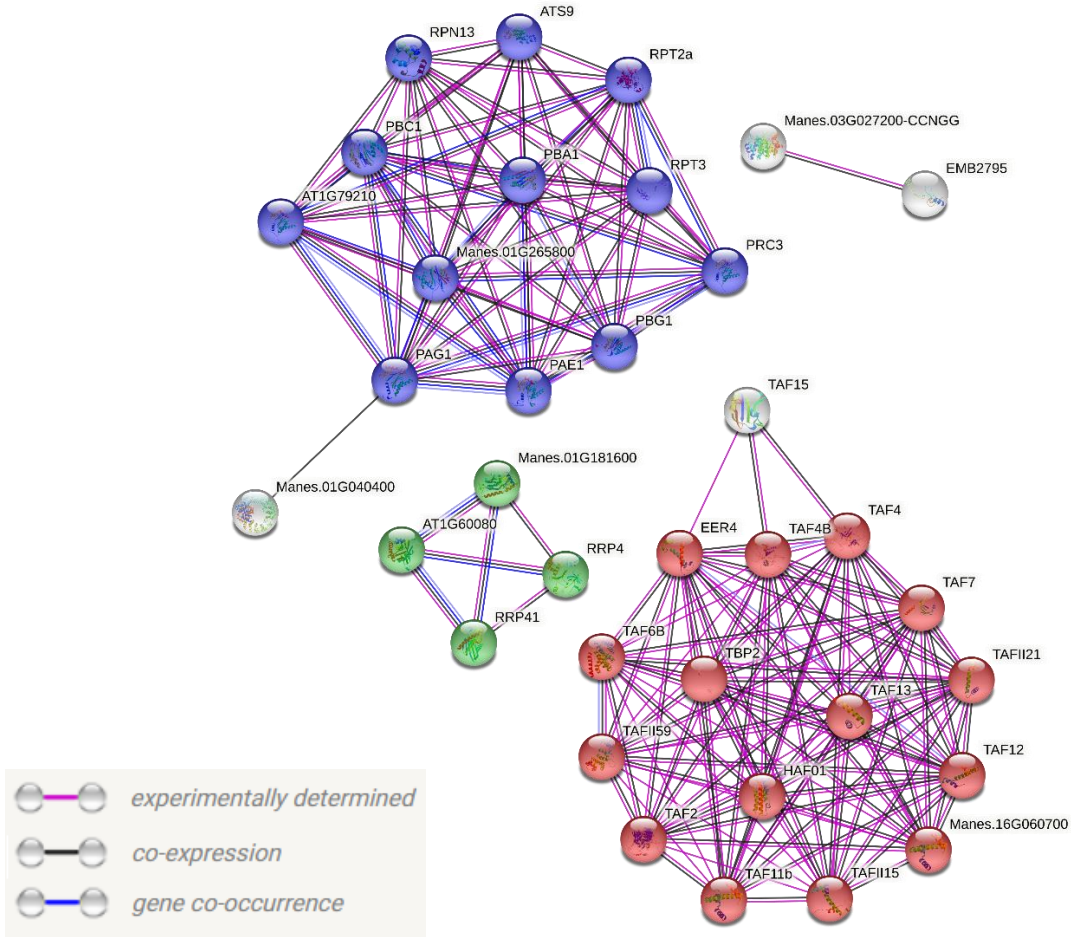


Tabela 7. Quantidade de interações previstas com proteínas da rede de interação proteína-proteína, derivada de genes hipermetilados no contraste de Formosa, após um estímulo de déficit hídrcio em relação ao controle sem estresse.

Gene de entrada/proteína	Interações
Manes.01G040400	1
Manes.01G181600	3
Manes.01G265800	11
Manes.16G060700	13
Manes.03G027200	1

5.9.3. Rede de interações predita para proteínas de genes hipometilados na variedade sensível Dourada

A rede de interações derivado de potenciais produtos dos genes hipometilados (em sítios CCGG e CCNGG) da resposta da variedade sensível Dourada ressaltou seis proteínas de *A. thaliana*, consideradas ortólogas dentre 44 proteínas de *M. esculenta* relacionadas aos genes hipometilados. A rede de interações apresentou 72 nós e 250 arestas, sendo a média de nós de 6,94 e o coeficiente médio de agrupamento local de 0,453, sendo relacionadas somente as relações dos componentes decorrentes basicamente de ocorrência experimental e dados de coexpressão e/ou de coocorrência (**Figura 17**). A análise de termos de GO associada a ferramenta STRING evidenciou 36 termos enriquecidos relativos a processo biológico, 16 para função molecular, e 28 para componente celular (**Tabela Suplementar 5A, 5B e 5C**), ao passo que o enriquecimento considerando os componentes e relações com possíveis vias metabólicas da plataforma KEGG (**Tabela Suplementar 5D**) evidenciou 13 proteínas relacionadas com ribossomos, outras cinco especificamente com a biogênese de ribossomos, e seis com o reparo de *mismatches* no material genético (**Figura 18**).

Numericamente, as proteínas apresentaram interações que variaram de duas a 23 (**Tabela 8**).

Figura 17. Rede de interação proteína-proteína predita para produtos de genes hipometilados (sítios CCGG e CCNGG), da variedade Dourada (*Manihot esculenta*), após um estímulo de déficit hídrico, conforme modelo de *A. thaliana* e ferramenta STRING. Nós (esferas) representam proteínas, e arestas, as interações (em rosa, as experimentais; em preto, as de coexpressão; em azul, as de coocorrência). Proteínas do arquivo de entrada apresentam prefixo “Manes”. As cores em cada nó representam termos GO para processos biológicos (Azul-escuro: expressão gênica; verde-claro: tradução; amarelo: biogênese ribossomal; vermelho: regulação do processo celular; lilás: reparo de Missmatch).

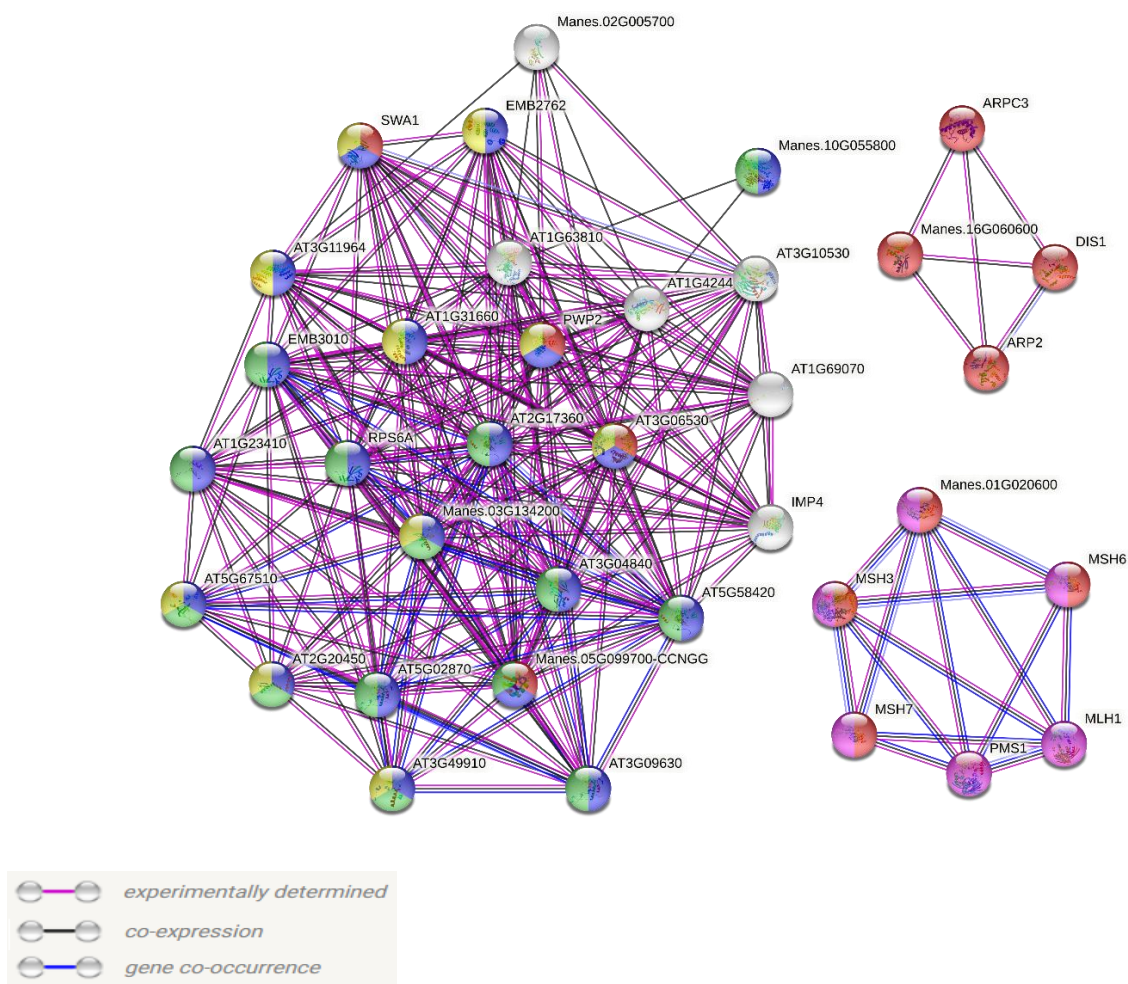


Figura 18. Rede de interação proteína-proteína obtida e visualizada pelo STRING para *A. thaliana* com base em genes hipometilados da resposta ao estímulo de déficit hídrico em *M. esculenta* (variedade: Dourada). Na rede os nós representam proteínas e as arestas representam as interações na rede (em rosa, as experimentais; em preto, as de coexpressão; em azul, as de coocorrência). Proteínas do arquivo de entrada consistem nas proteínas de prefixo “Manes”. As cores em cada nó representam termos KEGG (Laranja: Biogênese ribossomal em eucariotos; verde: reparação de *Missmatch*; azul-claro: ribossomo).

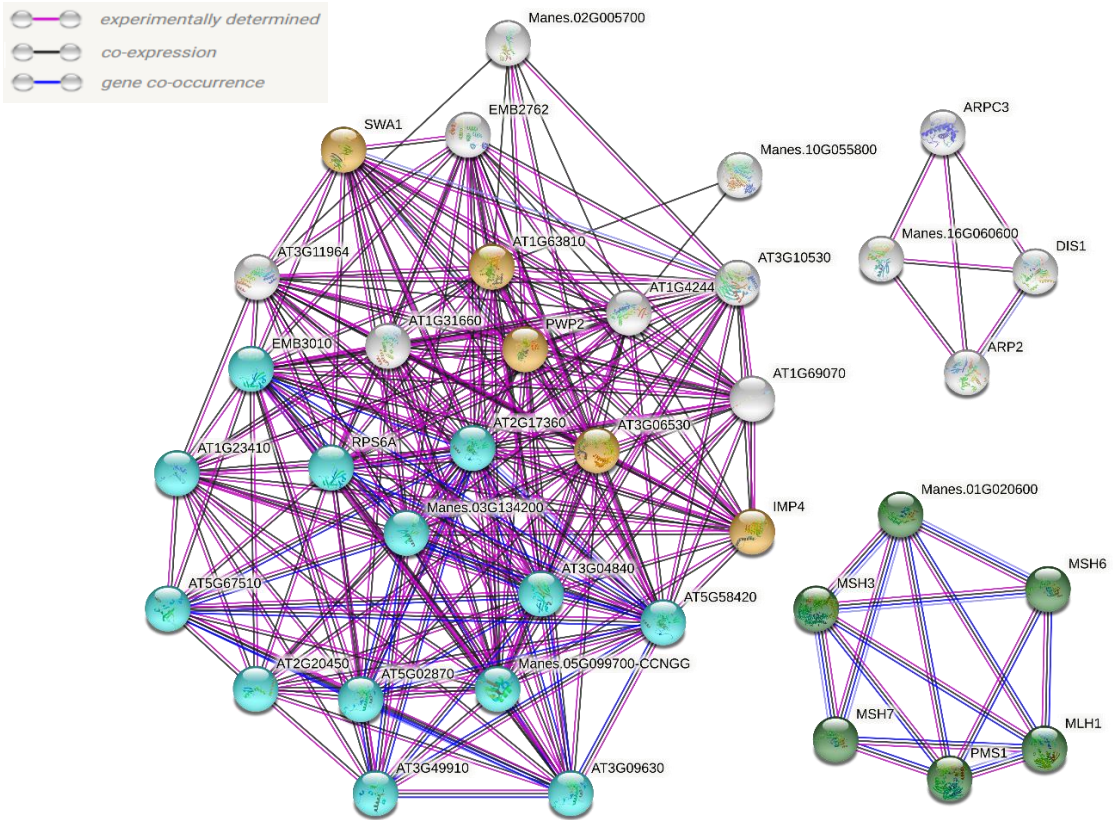


Tabela 8. Quantidade de interações previstas com proteínas da rede de interação proteína-proteína, derivada de genes hipometilados no contraste de Dourada, após um estímulo de seca em relação ao controle sem estresse.

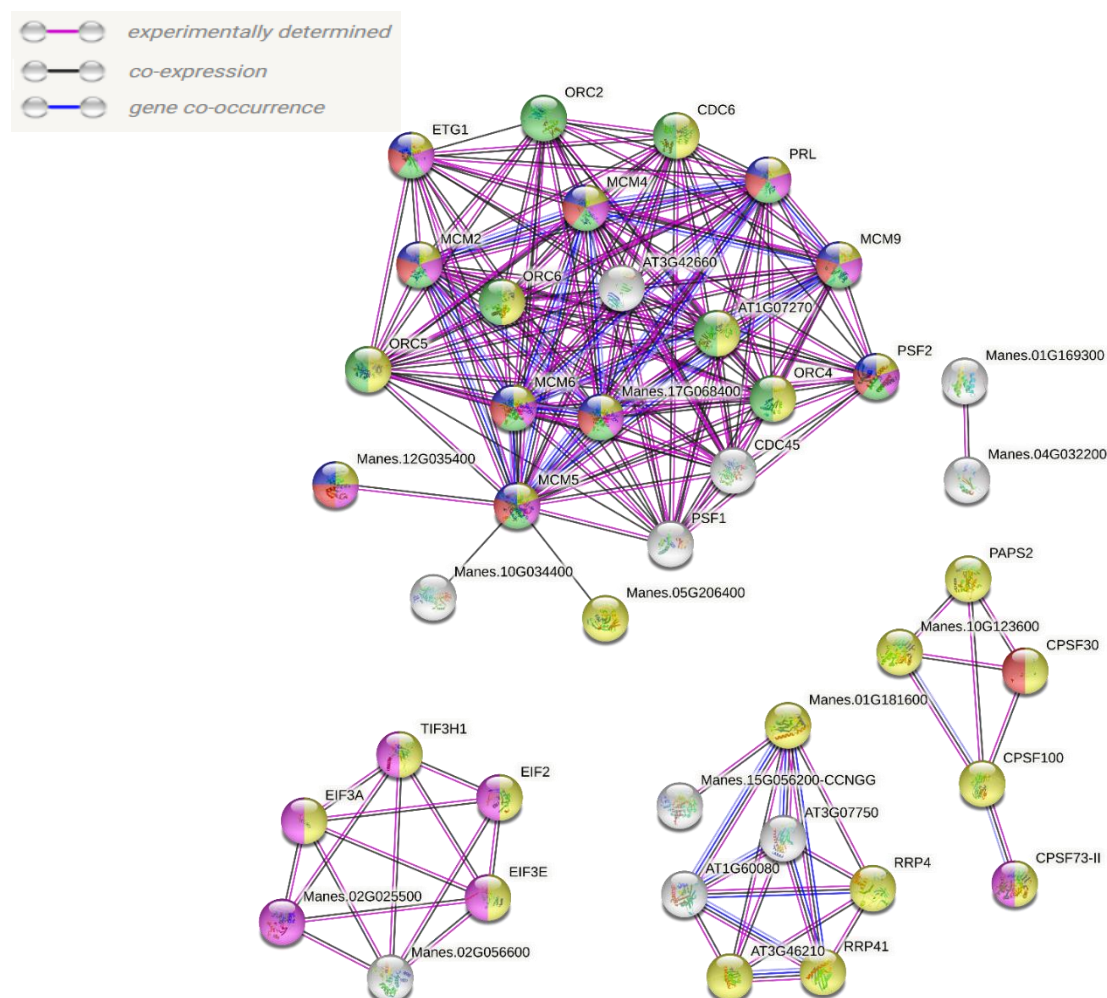
Gene de entrada/proteína	Interações
Manes.01G020600	5
Manes.02G005700	6
Manes.03G134200	23
Manes.05G099700	23
Manes.10G055800	2
Manes.16G060600	3

5.9.4. Rede de interações predita para proteínas de genes hipometilados na variedade sensível Dourada

A rede de interações derivado de potenciais produtos dos genes hipermetilados (em sítios CCGG e CCNGG) da resposta da variedade sensível Dourada ressaltou 11 proteínas de *A. thaliana*, consideradas ortólogas dentre 69 proteínas de *M. esculenta* relacionadas aos genes hipometilados. A rede de interações apresentou 95 nós e 182 arestas, sendo a média de nós de 3,83 e o coeficiente médio de agrupamento local de 0,402, sendo relacionadas somente as relações dos componentes decorrentes basicamente de ocorrência experimental e dados de coexpressão e/ou de coocorrência (**Figura 19**). A análise de termos de GO associada a ferramenta STRING evidenciou 56 termos enriquecidos relativos a processo biológico, 16 para função molecular, e 20 para componente celular (**Tabela Suplementar 6A, 6B e 6C**), ao passo que o enriquecimento considerando os componentes e relações com possíveis vias metabólicas da plataforma KEGG (**Tabela Suplementar 6D**) evidenciou seis proteínas relacionadas com a replicação do DNA, seis relacionadas com degradação, e seis relacionadas com o transporte de RNA. Na **Figura 20** é possível observar *clusters* para replicação do DNA e complexo da DNA Polimerase III (esferas vermelhas); complexo exoribonuclease (esferas azul-escuro); fator de iniciação e programação de morte celular (esferas verdes).

Numericamente, as proteínas apresentaram interações que variaram de uma a 17 (**Tabela 9**).

Figura 19. Rede de interações proteína-proteína predita para produtos de genes hipermetilados (sítios CCGG e CCNGG), da variedade Dourada (*Manihot esculenta*), após um estímulo de déficit hídrico, conforme modelo de *A. thaliana* e ferramenta STRING. Nós (esferas) representam proteínas, e arestas, as interações (em rosa, as experimentais; em preto, as de coexpressão; em azul, as de coocorrência). Proteínas do arquivo de entrada apresentam prefixo “Manes”. Cores das esferas representam termos GO para processos biológicos (amarelo: processos metabólicos primário; lilás: organização de componentes celulares; verde-claro: replicação do DNA; vermelho: resposta celular ao estresse; azul-escuro: reparação do DNA).



5.10. Análise SNP-INDEL

A análise de polimorfismos (SNP/INDEL) considerando os genes diferencialmente metilados, ao comparar as sequências genômicas de cada variedade com a referência *M. esculenta* v.6.1, um total de 15.643 alterações de bases destacaram 168 genes (dos 190 diferencialmente metilados detectados em ambas as variedades). Destas, 13.196 alterações puderam ser classificadas como modificadoras sem grandes efeitos na informação genética, enquanto 2.447 alterações foram consideradas de baixo (1.340), moderado (1.042) ou de alto impacto (65). As alterações consideradas modificadoras incluíram eventos de: variante UTR3'; íntron variante; variante UTR5'; variante genética a jusante.

Eventos de baixo impacto envolvidos: variante de ganho prematuro de códon de início UTR5'; variante da região de splice; variante da região de splice e variante do íntron; variante sinônima. Os efeitos de impacto moderado incluíram: variante *missense*; exclusão conservadora de quadro; exclusão perturbadora de quadros; inserção disruptiva no quadro; variante *missense*/variante da região de *splice*. Em contraste, efeitos de alto impacto envolveram: variante *frameshift*; variante de região de emenda; variante de início perdido; variante de *stop* perdido; variante do aceitador de *splice*/íntron; variante de *splice* e região UTR3'; variante doadora de *splice* e UTR5'. A totalização dos eventos de baixo, moderado e alto impacto detectados pelo gene candidato é dada na Tabela Suplementar S7, bem como, a posição no genoma dos eventos mais impactados, além da condição homozigótica ou heterozigótica do polimorfismo se detectado em cada variedade e o status (hiper ou hipometilação).

6. Discussão

6.1. O sequenciamento *MethyldRad* em variedades de *M. esculenta*

A partir dos parâmetros utilizados foi possível notar uma proporção de leituras mapeadas se mostram condizente com as taxas encontradas em outros mapeamentos para vegetais como em *Morus Alba* [14 – 15% (ACKAH et al., 2022)], *A. thaliana* [34 - 36% (WANG et al., 2015)] e *Zea mays* [18 – 20% (QIAN et al., 2019)]. Esta taxa de mapeamento de leituras únicas (entre 1/3 e 1/4) se dá principalmente devido a proporção de regiões repetitivas que são geralmente altamente metiladas nas plantas (COKUS, 2008). Desta forma, a quantidade de leituras usadas para as posteriores análises foram satisfatórias e confiáveis para refletir os dados de metilação exibidos pelas diferentes variedades em cada tratamento.

6.2. Os níveis de metilação entre as réplicas biológicas

O grau de correlação entre duas variáveis de escala métrica é possível pelo coeficiente de correlação de Pearson. Esses coeficientes permitiram correlacionar os dados das réplicas biológicas (R1, R2 e R3) de cada variedade, indicando serem estes consistentes, pois para sítio metilado CCGG de Formosa, as réplicas biológicas na condição controle apresentaram $r \geq 0,94$, e na condição estresse, $r \geq 0,85$, sendo que para Dourada e mesmo sítio, as réplicas biológicas na condição controle apresentaram $r \geq 0,86$, e na condição estresse, $r \geq 0,92$. Resultados semelhantes foram observados para o sítio CCNGG, que apresentou $r \geq 0,96$ e $r \geq 0,88$, na condição controle e de estresse, respectivamente, na variedade tolerante Formosa, e de $r \geq 0,71$ e $r \geq 0,82$, para a condição controle e de estresse, respectivamente, para a variedade sensível Dourada. Logo, as réplicas biológicas representaram adequadamente as variedades no estudo.

6.3. Os sítios metilados nos elementos genéticos estruturais do genoma de *M. esculenta*

A maioria dos sítios metilados (CCGG e CCNGG) ocorreu em regiões intergênicas como é esperado levando em conta as informações de perfis de metilação dispostos na literatura. A maior parte das metilações em genes

(independente do sítio) ocorreu nos 20% iniciais dos comprimentos dos genes, próximos aos sítios TSS (inícios das transcrições). Regiões TSS portando sítios de metilação nos primeiros pares de bases de genes já foram relatadas na literatura (BEWICK e SCHMITZ, 2017). A metilação nestes locais afeta sobremaneira a expressão gênica, visto que pode inibir a RNA polimerase II (Pol II) e a iniciação transcricional, o que explicaria baixos níveis de expressão desses genes metilados (LORINCZ et al., 2004; ZIBERMAN et al., 2006). Uma acentuada presença de sítios de metilação próximos ao TSS, e ao longo dos corpos gênicos (éxons e íntrons) também já foi reportada em angiospermas (ZEMACH et al., 2010). Por sua vez, na região final da transcrição, o pico de metilação para ambos os sítios se situou à montante ao TTS, ainda que um segundo pico menor para o sítio CCGG fosse detectado à jusante do TTS. Bartels et al. (2018) relatam que a metilação do DNA diminui significativamente em torno do TTS, onde a efetividade do controle da transcrição é menor em relação as regiões TSS.

6.4. Os perfis de metilação das variedades após a aplicação do estresse

A análise dos sítios diferencialmente metilados indicou que as variedades apresentaram perfis de metilação distintos, uma vez que apenas 12% dos sítios (CCGG e CCNGG) foram compartilhados. Dos poucos genes com sítios diferencialmente metilados em comum nas duas variedades, dois produtos [hidrolase (Manes.18G026700) e nuclease (Manes.01G181600)] se referiram a genes hipermetilados, cujas proteínas da manutenção basal celular estariam sendo reprimidas após o estresse. Outro gene diferencialmente metilado nas duas variedades seria hipermetilado em Formosa (tolerante) e hipometilado em Dourada (sensível), sendo o produto uma proteína RING/U-box, importante na ubiquitinação de proteínas, e na indução da senescência foliar, através da via de sinalização por ácido Jasmônico (AJ) (ZHANG et al., 2020).

6.5. Os termos de ontologia gênica dos conjuntos de genes diferencialmente metilados nas duas variedades

Dos termos de GO mais representados (enriquecidos) junto ao conjunto de genes hipometilados (sítios CCGG e CCNGG) em Formosa (tolerante) se

destacaram aqueles referentes a regulação do ciclo celular (GO:0010564), resposta a espécies reativas de oxigênio - ROS (GO:0000302), e resposta ao ácido Jasmônico (GO:0009753).

A tolerância de plantas à seca apresenta relação importante com a via de biossíntese de ácido jasmônico (SU et al., 2020). Esses autores estudaram o gene *VaNAC17* de *Vitis amurens*, constatando que o fator de transcrição aumentou a transcrição de genes responsivos ao estresse hídrico, incluindo alguns da via de biossíntese de AJ, como *LOX3*, *AOC1* e *OPR3*, sendo que as plantas de *A. thaliana* expressando o transgene *VaNAC17* aumentaram a tolerância à seca, e diminuíram o acúmulo de ROS, em comparação com plantas do tipo selvagem, nas condições de seca. Além disso, a aplicação de AJ em sementes de trigo sob estresse hídrico (simulado com 20 mL de solução de PEG a 20%, em solução de Hoagland com 1/10 de força) aumentaria a germinação de sementes em 27%, e o potencial hídrico em 60% (ILYAS et al., 2017).

Por sua vez, termos enriquecidos de GO associados ao conjunto de genes hipermetilados (sítios CCGG e CCNGG) em Formosa refletiram mais os processos metabólicos primários, em especial envolvendo hidrolases (GO:0016787), peptidases (GO:0070011, GO:0008233, GO:0004175, GO:0004180, GO:0008238, GO:0008237, GO:0004222) e proteólise (GO:0006508). Essa hipermetilação, que reprimiria genes mais relacionados ao metabolismo primário e basal ajudaria no redirecionamento energético da planta visando transcrever genes mais prontamente responsáveis e envolvidos na resposta direta ao estresse, como aqueles relacionados ao conjunto de genes hipometilados acima.

Na resposta de Dourada (variedade sensível), os termos de GO enriquecidos, tendo em conta os genes hipometilados (sítios CCGG e CCNGG) estiveram associados a funções celulares basais do organismo, com destaque para a regulação do tamanho estrutural anatômico (GO:0090066), regulação do ciclo celular (GO:0010564) e reparo do DNA (GO:0006281). A regulação do tamanho das estruturas pode ser uma estratégia importante no controle do desenvolvimento vegetal durante o estresse, desviando energias para as respostas imediatas ao minimizar o desenvolvimento normal. Por sua vez, o destaque para o reparo do DNA pode refletir danos na molécula decorrente do

estresse em Dourada, visto que estresses hídricos influenciam na quantidade de ROS intracelular gerada, podendo estes afetar as estruturas dos DNAs, requerendo os reparos. De modo contrário ao de Formosa, os termos enriquecidos de GO associados aos genes hipermetilados de Dourada estiveram relacionados somente com funções celulares basais, e, portanto, menos relacionados com as respostas imediatas ao estresse aplicado.

De forma resumida, a estratégia de Formosa (a tolerante) parece ser reprimir o metabolismo primário e o desenvolvimento vegetal, e priorizar a sinalização do estresse e as respostas mais imediatas, enquanto que a estratégia de Dourada (mais sensível) seria tentar regular o desenvolvimento, regulando o tamanho de suas estruturas, ao tempo em que dispara o sistema de reparo ao DNA, talvez em consequência de danos decorrentes do estresse aplicado. Os estudos das redes de interações de proteínas previstas podem auxiliar nessas caracterizações das estratégias das duas variedades em resposta ao estresse aplicado.

6.6. Redes de interações previstas a partir dos conjuntos de genes diferencialmente metilados

Muitos estudos mostram que as plantas tolerantes tem respostas mais efetivas devido à percepção precoce do estresse comparativamente à contraparte sensível. Isso ocorre nas primeiras horas após imposição ao estresse (PILLAY e BEYL 1990; RAMPINO et al., 2006; MOUSSA et al., 2009; GUO et al., 2009; GUO et al., 2018). Os genes diferencialmente metilados nas duas variedades aqui estudadas foram resultado de um estresse severo, em que as plantas atingiram um ponto de murcha de suas folhas e a capacidade hídrica do solo se reduziu a 10% da capacidade de campo. Este fenômeno ocorreu após dezoito dias da supressão da rega para as plantas estressada, desta forma, é esperado que as respostas metabólicas e genéticas das variedades sejam distintas.

6.6.1. A rede de interações predita para proteínas de genes hipometilados na variedade tolerante Formosa

A análise da rede de interações associada aos genes hipometilados em Formosa (tolerante) após a aplicação do estresse, evidenciado por quatro pequenos grupos PPI, destacou em uma delas, a proteína CYT1 (Manose-1-fosfato guanililtransferase 1, Manes.10G025000) interagindo com proteínas da família ADP-glicose pirofosforilase, que estão envolvidas em processos biossintéticos da planta. A CYT1 é proteína atuante no desenvolvimento da parede vegetal, sendo que mutantes de *A. thaliana* com perda de função apresentam acúmulo ectópico de calose, paredes celulares incompletas (textura e espessura irregulares, e distribuição anormal de pectinas), que afetam o turgor celular (LUKOWITZ, 2001). A pressão de turgor é um fator crítico na regulação do crescimento celular. Durante o estresse da seca, as alterações fisiológicas das plantas envolvem a perda de turgor, o ajuste osmótico e a redução do potencial de água da folha. Uma baixa pressão de turgor causada pelo estresse hídrico leva a uma redução ou cessação do crescimento, podendo afetar a adaptação do sistema hidráulico do vegetal (WOLF e GREINER, 2012; LISAR et al., 2012; BALDUCCI et al., 2014). Comparando com o genoma de referência de *M. esculenta* (v.6.1), o gene CYT1 apresentou polimorfismos SNPs/INDELs (seis de baixo impacto e dois de efeitos moderados). Um de impacto moderado (alteração missense) foi em condição heterozigótica na BRS Formosa (não polimórfica na BRS Dourada; Tabela suplementar 7). Este SNP/INDEL merece ser explorado se tal condição proporcionar uma vantagem seletiva para uma variedade tolerante à seca.

Um segundo grupo PPI destacou duas proteínas de referência (Manes.02G033400 e Manes.07G102300) relacionadas com a resposta hormonal da planta. A resposta aos sinais abióticos em plantas é, geralmente, iniciada pela perturbação da membrana plasmática (GUO et al., 2011). No enfrentamento do estresse abiótico, as plantas utilizam as vias de sinalização hormonal. O ácido abscísico (ABA) e o etileno são os estimulantes mais frequentes da resposta abiótica, podendo interagir com os diferentes fitormônios [ácido salicílico (AS), ácido jasmônico (AJ), auxina (AUX)], podendo alterar os perfis de expressão gênica das respostas adaptativas ao estresse (HIRAYAMA

e SHINOZAKI, 2010; HUAG et al., 2012). A AUX é um importante fitohormônio, com distribuição dinâmica nos tecidos vegetais, e que controla uma variedade de processos que visam adaptar o crescimento e o desenvolvimento das plantas às condições ambientais (VANNESTE, 2009).

Muitas das respostas fisiológicas e de desenvolvimento das plantas, como sinalização hormonal, divisão celular, desenvolvimento floral e manutenção do ritmo circadiano são influenciadas também pelo controle pós-traducional das proteínas, cujo turnover de proteínas por ubiquitinação e degradação pelo proteassoma 26S, é processo regulamentado e essencial em eucariotos (SMALLE e VIERSTRA, 2004). A especificidade do substrato no processo de ubiquitinação decorre das enzimas ubiquitina-ligases E3, que pertencem a diferentes grupos, de acordo com suas características mecanicistas e estruturais (PICKART, 2001). Ligases E3 compostas por subunidade de Culina, proteína RING (RBX1/Hrt1/Roc1), e receptor de substrato, que se ligam especificamente às proteínas a serem ubiquitinizadas, são denominadas CRLs (Cullin-RING Ligases). As CRLs mais bem caracterizadas são os complexos SCFs (SKP1-cullin-F-box), que incorporam Culina1 (CUL1) (LECHNER et al., 2006). A especificidade do substrato do complexo SCF é conferida por proteínas F-box, que se associam ao CUL1 através de uma proteína adaptadora SKP1. Os complexos bem caracterizados SCF^{TIR1/AFB} funcionam como receptores para a auxina (KEPINSKI e LEYSER, 2005). Ligações de SCF^{TIR1/AFB} com auxina aumentam a afinidade do complexo SCF^{TIR1/AFB} para substratos do tipo proteínas Aux/IAA, aumentando a ubiquitinação e por conseguinte o turnover das proteínas. Como resultado, fatores de resposta à auxina (ARFs), cujas atividades são normalmente reprimidas por dimerização com Aux/IAAs, são liberados, podendo, assim, atuar como reguladores transcricionais de genes responsivos a auxina (NAGPAL et al., 2005; TEALE et al., 2006; LIU et al., 2018). As mutações que diminuem a atividade de SCF^{TIR1/AFB} estabilizam as proteínas Aux/IAA, diminuindo a indução de genes regulados pela auxina, e reduzindo a sensibilidade da planta à auxina exógena (PARRY e STELLE, 2004).

Também plantas mutantes resistentes à auxina podem ser defeituosas na conjugação da proteína anti-ubiquitina RUB/NEDD8 das culinas. O sinalossoma COP9 (Constitutive Photomorphogenesis 9), conhecido como CSN, cliva RUB/NEDD8 das culinas (LYAPINA et al., 2001), sendo bioquimicamente

antagonista ao processo de conjugação de RUB. Assim, CSN atua como promotor da atividade ligase E3, ao proteger os adaptadores de substratos das CRLs de uma degradação autocatalítica, através da atividade de desrubilação (WEE et al., 2005). Subunidades de CSN foram também reportadas em vias de respostas a patógenos e respostas dependentes de AJ (LIU et al., 2002; HIND et al., 2011).

As proteínas RBX1 (componentes do complexo de ubiquitinação de SCF para resposta a AUX), CAND1 (proteína 1 associada à Culina e dissociada à nedilação) e Culina 1 (Manes.02G033400), componentes previstos na rede PPI, foram relatadas interagindo com COP9 e o complexo CSN (SINGH e CHAMOVITZ, 2019). O papel das CSN nas vias de produção de AJ também já foi reportado (HIND et al., 2011). Ainda na rede em questão, FUS12 (CSN2 - subunidade 2 do complexo do sinalossomo COP9) é um reconhecido repressor do desenvolvimento vegetal (CHAMOVITZ et al., 1996), ao passo que a proteína ARF6 (fator de resposta a auxina 6; Manes.07G102300) interagiria com proteínas IAA (como SHY2), estando o grupo diretamente relacionado com o complexo COP9.

Em outra rede, conectada por Manes.01G115300, as proteínas MEE (componente nuclear ribonucleoproteína CLO-5), MAC (fator de processamento pré-mRNA 19 homólogo 1), SKIP (proteína de interação-SNW/SKI), PRP (fator de processamento do splicing pré-mRNA), CDC (proteína do ciclo de divisão celular) e ESP (fator de pré-mRNA-splicing RNA helicase DEAH1) foram associadas pelos termos de GO com o *splicing* alternativo (AS) dos RNAs. Estresses ambientais e bióticos podem induzir transcritos alternativos (MASTRANGELO et al., 2012; HOWARD et al., 2013). Um estudo RNA-seq de *A. thaliana* infectada por *Pseudomonas syringae*, indicou que mais de 90% dos genes expressos (23.385 de 25.619) apresentaram transcritos alternativos (HOWARD et al., 2013). Por sua vez, subunidades do complexo MAC (proteínas MAC3A, MAC3B e MAC5A) foram enfatizadas no controle da transcrição em respostas imunes em *A. thaliana* (HOWARD et al., 2013), enquanto que as proteínas PRP seriam parte do complexo NTC (NineTeen Complex; KONCZ et al., 2012), o qual seria um complexo regulatório essencial para a ativação catalítica do spliceossoma (HOGG et al., 2010). Manes.01G115300 AINDA EXIBIU polimorfismos que compreendia três SNPs/INDELs; sendo um de efeito

moderado (variante *missense*) seria heterozigoto na BRS Formosa e homozigoto na BRS Dourada (Tabela Suplementar 7).

Em outro grupo PPI se destacou uma proteína PAP15 (fosfatase de ácido roxo 15, Manes.15G054300), previamente reportada proporcionando maior tolerância ao estresse osmótico e de salinidade, bem como estimulação ao crescimento e a insensibilidade ao ABA (ZHANG et al., 2008b), além de modular níveis de ácido ascórbico (AsA) ao controlar a entrada de mio-inositol na biossíntese deste composto (ZHANG et al., 2008b). Essa proteína apresentou interação com um modulador de descarga de proteínas no floema (PLM), que catalisa a biossíntese de esfingolípídios com ácidos graxos de cadeia longa (Manes.15G054300), e uma proteína beta-glucuronosyltransferase (GLCAT), envolvida na biossíntese do tipo II de arabinogalactanos, que são proteoglicanos abundantes na superfície celular, e que estão envolvidas em processos celulares, incluindo embriogênese somática, interações célula-célula e alongamento celular (YAN et al., 2019; DILOKPIMOL, GESHI, 2014). O gene supracitado apresentou seis SNPs/INDELs de baixo impacto e dois de efeitos moderados (Tabela Suplementar 7). Uma variante *missense* de efeito moderado foi heterozigótica em BRS Formosa e não polimórfica em BRS Dourada. Novamente, a condição heterozigótica foi evidenciada em BRS Formosa em detrimento dos homozigotos ou não-polimórficos no BRS Dourada.

6.6.2. A rede de interações predita para proteínas de genes hipermetilados na variedade tolerante Formosa

A rede de interações, com três pequenos grupos PPI, previsto de proteínas codificadas por genes hipermetilados após o estresse aplicado em Formosa (tolerante), mostrou em uma deles, além de componentes com múltiplas funções, várias proteínas associadas a proteassomas. Como as proteínas controlam atividades celulares e processos fisiológicos através da regulação de vias bioquímicas ou metabólicas específicas, nessa regulação, a degradação de proteínas é tão significativa quanto a síntese. Na degradação proteica, proteases e o sistema ubiquitina-proteassoma (UPS) são responsáveis pela degradação da maioria das proteínas (80 - 90%), sendo o sistema UPS uma estratégia eficiente e rápida para controlar processos celulares, removendo seletivamente

as proteínas regulatórias, e assim, desempenhando um papel relevante na regulação de processos fisiológicos (XU e XUE, 2019). Nas plantas sob estresse abiótico, proteases de várias classes são afetadas. Esse envolvimento é percebido pela indução de genes codificadores de proteases e pela mudança nos níveis de atividades proteolíticas de diferentes especificidades, sendo as cisteína-endopeptidases uma das primeiras associadas ao estímulo de seca, geralmente com perfil de indução (SIMOVA-STOILOVA et al., 2010; VASEVA et al., 2012). Apesar da rede ter apresentado componentes relacionados com a degradação de proteínas, e com o controle celular, mas sendo esta uma rede derivada de produtos de genes hipermetilados, a importância da rede seria minimizada, pois os genes tenderiam a ser reprimidos, permitindo que a planta desviasse recursos para respostas mais efetivas contra o estresse, como aquelas previstas a partir da rede de interações das proteínas de genes hipometilados em Formosa.

Ainda na rede prevista, o componente KPNB1 (importina subunidade beta-1, Manes.01G040400) seria uma proteína da família *ARM-repeat*. De acordo com Luo et al (2013) a importação de proteínas para o núcleo em resposta à seca é fundamental na reprogramação da expressão genética que levaria à tolerância à seca. A inativação de *AtKPNB1* de *A. thaliana* aumentou o fechamento estomatal, em resposta a influência do fito-hormônio ABA, reduzindo a perda de água, e aumentando a tolerância da planta à seca, o que reforça o encontro da hipermetilação do gene.

Um segundo grupo PPI destacou as subunidades do fator de iniciação da transcrição TFIID (TAF). Proteínas TAF participam da iniciação da transcrição como parte do complexo de pré-iniciação da RNA polimerase II, que é composto pela RNA pol II e um subconjunto de fatores de transcrição (TFIIA, B, D, E, F e H). O reconhecimento inicial do promotor é mediado pelo fator de transcrição basal TFIID, que inclui a proteína de ligação TATA (TBP) em um complexo com vários fatores associados à TBP (TAFs). Proteínas TAFs atuam como co-ativadores transcripcionais na montagem de complexos transcripcionais e no reconhecimento de promotores (WATERWORTH et al., 2015). Os TFs TFIIDs foram reportados como responsivos a etileno e importantes em respostas de raízes de Trigo (*Triticum aestivum*) ao estresse de seca (GRZESIAK et al., 2019).

Este poderia ser mais um mecanismo a conferir melhor resposta ao estímulo de seca na variedade Formosa, caso os genes não estivessem hipermetilados.

Um terceiro grupo PPI destacou uma proteína da família exoribonuclease 3'-5'- (Manes.01G181600) com interação com proteínas com atividade de degradação de RNA (RRP). Mutações de perda de função em *AtRRP6L1* foram associadas com a diminuição da metilação do DNA, o que pode ser devido a sua função de retenção de RNAs e acúmulo destes em associação com a cromatina (ZHANG et al., 2014). As proteínas presentes na rede seriam RRP45b e CCR4-1. RRP45b é uma subunidade do exossoma, e CCR4-1, uma proteína importante no complexo de repressão 3'-5' para degradação de RNA (LIU e CHEN, 2016). Assim, o grupo pode estar ligado a regulação epigenética e ao silenciamento gênico, bem como a regulação de RNAs.

Outra proteína representada (Manes.03G027200) mostrou apenas uma conexão com a proteína EMB2795, cuja família de proteínas foi associada com resposta a estresse biótico, conforme relato de Villajuana-Bonequi et al. (2019).

6.6.3. A rede de interações predita para proteínas de genes hipometilados na variedade sensível Dourada

A rede de interações, com três grupos PPI preditos de potenciais proteínas codificadas por genes hipometilados em resposta ao estresse aplicado em Dourada, destacou em um deles quatro proteínas de referência, sendo um transportador da família ABC (ABCF₁, Manes.02G005700), uma Lisina-tRNA ligase (Manes.10G055800) e duas proteínas ribossomais 40S (S8-2, Manes.03G134200; S9-2, Manes.05G099700), interagindo com diversas proteínas e subunidades ribossomais. Ribossomos são complexos de ribonucleoproteínas, compostos de RNAs ribossômicos (rRNAs) e proteínas ribossômicas (RPs), organizadas em uma subunidade maior e outra menor, que sintetizam a maioria dos peptídeos da célula. As RPs são essenciais para a síntese proteica, com papéis importantes no metabolismo, divisão celular e no crescimento vegetal. Mutações em genes para RPs evidenciam papéis regulatórios em processos do desenvolvimento vegetal, em que os mutantes normalmente apresentam anormalidades e redução do crescimento e da proliferação celular (BYRNE, 2009; HORIGUCHI et al., 2012). Genes para RPs

se mostraram diferencialmente regulados tanto por fatores abióticos quanto bióticos, afetando diretamente a biogênese ribossomal e o crescimento vegetal (FROMONT-RACINE et al., 2003).

Dentre os componentes previstos na rede, vários foram proteínas RPL (proteínas ribossômicas 60S: AT2G17360, AT5G58420, AT5G67510, AT2G20450 e AT3G49910), sendo proteínas RPL já relatadas como importantes para a eficiência do uso da água em vegetais (MOIN et al., 2016a). O gene *RPL10* se mostrou significativamente regulado por radiação UV-B e por estresse biótico (FERREYRA et al., 2010; CARVALHO et al., 2008; MOIN et al., 2016b), enquanto que os genes *RPL35* e *RPL32* foram regulados em resposta ao calor e ao sal, respectivamente (MUKHOPADHYAY et al., 2011), ao passo que o gene *RPL11* se mostrou regulado pelo estresse salino, tendo as plantas mutantes relacionadas (*A. thaliana*) um crescimento defeituoso e com folhas pálidas (OMIDBAKHSHFARD et al., 2012). Também mutações em *AtRPL24* afetaram negativamente o desenvolvimento das plantas, bem como a reiniciação da tradução de mRNAs codificadores dos fatores de transcrição bZIP11 e ARF, ambos importantes na resposta transcricional a estresses abióticos (ZHOU et al., 2010).

A análise dos genes relacionados presentes nos grupos acima, apresentaram diversos polimorfismos, com condição heterozigota da variedade BRS Dourada em detrimento do status homozigoto ou não-polimórfico de BRS Formosa, conforme observado para os genes que codificam Lisina -tRNA ligase (Manes.10G055800) e proteína ribossômica 40S S8-2 (Manes.03G134200), ambos mostrando um SNP/INDEL de impacto moderado (uma variante *missense*) (Tabela Suplementar 2).

Em mesma rede, além das RPs se observou também a proteína PWP [*periodic tryptophan (W) protein*] e o complexo EMB [complexo nucleolar de proteína associada (NOC4) e proteína ribossômica 40S S6-2 (RPSb)]. Proteínas PWP estão envolvidas no processamento do pré-RNA ribossomal 18S, fazendo parte da biogênese do ribossomo (MISSBACH et al., 2013), enquanto que proteínas NOC4 são importantes no processamento nucleolar de rRNA das subunidades 40S e 90S, na montagem de ribossomos e no amadurecimento de rRNA 5,8S (MISSBACH et al., 2013). Ainda, apesar do transportador ABCF₁

(Manes.02G005700), presente na rede, não ter uma relação aparente com os demais, sua importância já foi reportada em respostas de plantas ao estresse biótico e abiótico (Sun et al., 2016), sendo que plantas tratadas com ácido salicílico (AS) e etileno (ET), após estímulos abióticos (frio, salinidade e ferimentos) aumentaram a transcrição de *LrABCF₁* (*Lilium regale*) após tratamentos por AS e ET. Logo, este componente pode estar envolvido além do transporte através de membranas, também na respostas a estresses.

Um outro grupo PPI destacou proteínas do complexo SCAR e Arp2/3. Proteínas ARP do complexo Arp2/3 estão envolvidas na regulação da polimerização de actina, mediando a formação de redes ramificadas de actina. O complexo Arp2/3 tem como ativador o complexo SCAR, e estão diretamente envolvidos com o crescimento das plantas (DYACHOK et al., 2011).

Outro grupo PPI explorou a interação do componente Manes.01G020600 com proteínas envolvidas no reparo e correção de danos no DNA, ressaltando proteínas de reparo de *missmatches* (MSH, MLH e PMS1). Os componentes do sistema MMR (*mismatch repair*) formam heterodímeros MutS α (heterodímero MSH2-MSH6) que se ligam aos *missmatches* no DNA, iniciando assim o reparo. Estas proteínas corrigem incompatibilidades base-base e loops de exclusão de inserção (IDLs) resultantes da replicação de DNA, de eventuais danos ao DNA, ou de eventos de recombinação entre sequências não idênticas durante a meiose (LI et al., 2009). O reparo do DNA é importante, no caso do estresse danificar os ácidos nucleicos, para a planta manter o metabolismo funcionando de forma adequada (HUANG, TRAGANOS, DARZYNKIEWICZ 2003; MOROTOMI et al., 2018; KAMAL et al, 2021).

6.6.4. A rede de interações predita para proteínas de genes hipermetilados na variedade sensível Dourada

A rede de interações, com grupos PPI previstos de potenciais proteínas dos genes hipermetilados em Dourada (sensível), destacou em um deles, quatro das proteínas de referência, sendo a 2-oxoglutarato desidrogenase, componente E1 (Manes.10G034400), a proteína de domínio RING associada a BRCA1 (Manes.12G035400), a proteína da família de manutenção de minicromossomo (Manes.17G068400), e uma proteína não caracterizada (Manes.05G206400), as

quais interagiriam com proteínas de manutenção do minicromossomo (MCM), aos complexos da origem de replicação (ORC) e proteínas do ciclo de divisão celular (CDC).

Proteínas MCM estão envolvidas no ciclo celular e com a iniciação da replicação do DNA, na transição das fases G1 para fase S, que requer a formação e ativação de complexos pré-replicativos. A formação desses complexos se inicia na vinculação de CDC6 aos ORCs durante a fase G1, permitindo o recrutamento de complexos MCM de proteínas. A fase S é desencadeada pela ativação desse complexo por quinases dependentes de ciclina, que levam à mudança do complexo de pré-replicação para um de pós-replicação (STEVENSON et al., 2002). Além disso, o complexo ORC se associa à origem da replicação para recrutar o complexo de manutenção CDC6, CDC10 Target1 (CDT1) e MCM para formar o complexo pré-replicativo (pré-RC) na transição de fase M para G1 (WU et al., 2017). Os componentes desta rede, basicamente proteínas envolvidas na replicação do DNA e controle do ciclo celular, reuniram-se em um *cluster* local. Entretanto, tendo em conta que os produtos previstos seriam oriundos de genes hipermetilados, deduz-se, então, que a proliferação celular estaria sendo freada na variedade Dourada após o estresse aplicado.

Outro grupo PPI destacou as proteínas de referência DEAD-box ATP-dependente RNA helicase 47 (mitocondrial, Manes.02G025500) e fator de iniciação da tradução eucariótica 3 (subunidade M, Manes.02G056600), em interação com fatores de iniciação de tradução eucariótica da família eIF. Os eIFs incluem representantes de seis famílias (eIF1 - eIF6), os quais estão envolvidos na iniciação da tradução, na síntese proteica (DEVER, GREEN, 2012). Assim, eIF1, eIF1A e eIF3 formam o complexo de pré-iniciação 43S (PIC), através da ligação de subunidades ribossômicas (NANDA et al., 2013). Por sua vez, eIF4 funciona promovendo a combinação correta de PIC com a extremidade CAP-5' (LI et al., 2019). Alguns genes *eIF* já foram reportados associados a respostas de plantas a estresse abiótico. Genes como *MieIF1A-a*, *MieIF5* e *MieIF3sB* foram fortemente expressos em plantas (*Mangifera indica*) sob estresse de salinidade, seca (PEG) e frio, respectivamente (LI et al., 2019). A superexpressão de *MieIF1A-a* em plantas transgênicas de *A. thaliana* acarretou

maior tolerância a salinidade (LI et al., 2019). Novamente, sendo os componentes previstos de genes hipermetilados, possivelmente silenciados, deduz-se que a variedade Dourada não tira proveito dessa estratégia, como poderia, pois o estresse salino e o de seca compartilham muito das suas respostas. Tal fato reforça, por outro lado, a maior sensibilidade de Dourada ao estresse em questão.

Outro grupo PPI interligou componentes do complexo exossoma do RNA, entre eles RRP45B (Manes.01G181600), e proteína da superfamília de repetição ARM (Manes.15G056200). Proteínas RRP, que constituem o complexo exossoma do RNA, formam um complexo multiproteico que tem como alvo uma vasta gama de RNAs, atuando em muitas de suas fases, do amadurecimento até o controle de qualidade e rotatividade final. Aparentemente essas ações estariam sendo silenciadas na variedade Dourada. Por outro lado, há informações de relações entre os componentes atreladas ao desenvolvimento do vegetal, visto que Zhang et al. (2010) concluíram que o gene *AtRRP44A* foi necessário para o desenvolvimento de gametófitos femininos, tendo sido confirmado por Kumakura et al. (2013) que o gene *AtRRP44A* não afetaria igualmente os gametófitos masculinos.

Um outro grupo PPI interligou componentes do complexo do fator de especificação de clivagem e poliadenilação - CPSF, dentre os quais a referência Manes.10G123600, com a proteína PAP (nuclear poli(A)-polimerase). Proteínas CPSF desempenham papel fundamental na formação *pré-mRNA 3'-end*, ao interagir com a poli(A)-polimerase e outros fatores para realizar a clivagem do transcrito e a adição da cauda poli(A). Estas proteínas estariam também envolvidas no controle pós-transcricional, das respostas das plantas a estresses, especialmente de genes que mediam a tolerância ao estresse oxidativo (ZHANG et al., 2008). Neste ponto, a estratégia não estaria sendo bem explorada por Dourada, pois os genes codificadores estariam metilados e potencialmente silenciados. Ainda, as proteínas PAP adquirem especificidade através de interações com CPFS, sendo também necessárias nas reações de clivagem em locais de poliadenilação (KIEFER et al., 2009). As CPFS estão envolvidas também no controle da morte celular em vegetais, e em vias de

resposta ao ácido salicílico (AS), regulando positivamente a imunidade mediada por AS (BRUGGEMAN et al. 2014).

Outra interação prevista envolveu um fator de *splicing* com domínio PWI (Manes.01G169300) com a subunidade 4-a do complexo THO (Manes.04G03220). O domínio PWI, em fatores de *splicing*, é o de ligação a RNA/DNA, e este apresenta múltiplas funções no processamento de pré-mRNAs (SZYMCZYNA et al., 2003). O complexo THO, subunidade 4, por sua vez, participa da biogênese da partícula de ribonucleoproteína (mRNP) que atua na exportação do mRNA para fora do núcleo (JIMENO e AGUILERA, 2010).

No geral, as interações previstas nesta rede, derivada de genes hipermetilados e potencialmente silenciados em Dourada, indicaram relações diretas com o metabolismo primário, sugerindo que a variedade sensível Dourada estaria tentando minimizar o desenvolvimento vegetal, redirecionando recursos para respostas mais direcionadas ao estresse aplicado. Entretanto, a resposta esperada pode não ser tão efetiva, tendo em conta que os genes hipometilados, aqueles que estariam sendo induzidos e expressos se mostraram mais envolvidos com redução do metabolismo primário e reparo de potenciais danos ao DNA.

7. Conclusão

Em resumo, os marcadores metilados mapearam substancialmente o genoma de *M. esculenta*. A representatividade dos dados de cada variedade foi consistente com base nos coeficientes de correlação de Pearson de suas réplicas biológicas. Nos genes, a maioria dos locais metilados cobriram os 20% iniciais dos comprimentos dos genes. Os genes diferencialmente metilados mostraram mais sítios CCGG metilado do que CCNGG em ambas variedades. Além disso, as variedades BRS apresentaram perfis de metilação distintos, compartilhando poucos genes diferencialmente metilados. Com base nos perfis de hipo/hipermetilação das variedades BRS, a caracterização GO apontou uma resposta mais eficaz do tolerante BRS Formosa após o estresse. Por isso, o perfil BRS Formosa destacou-se com o desenvolvimento vegetal e metabolismo primário reprimido, com indução da sinalização hormonal de estresse,

possivelmente desencadeando genes responsivos ao estresse durante o tratamento. Por sua vez, o perfil da BRS Dourada destacou uma regulação do desenvolvimento da planta (regulando o tamanho das estruturas vegetais) e indução do sistema de reparo do DNA, provavelmente devido a danos causados pelo estresse em macromoléculas. As redes de interações de proteínas previstas reforçaram a caracterização GO descrita anteriormente. O metiloma de BRS Formosa em destacou a relevância da diminuição do metabolismo primário das plantas, controlando o crescimento e desenvolvimento, também induzindo a vias de sinalização de estresse de fitohormônios e, portanto, desencadeando uma melhor resposta ao estresse por seca. A análise também identificou genes hipermetilados que normalmente são induzidos, melhorando as respostas rápidas ao déficit hídrico, mas que estavam potencialmente reprimidos. Esses genes poderiam ser explorados em estudos que tentam melhorar a tolerância a seca em programas de melhoramento de mandioca. Em contrapartida, o metiloma da BRS Dourada refletiu sua sensibilidade à seca, induzindo o sistema de reparo do DNA, ao mesmo tempo em que tentava controlar o metabolismo primário e reduzir o desenvolvimento da planta, mas mantendo a biogênese ribossômica. Além disso, a regulação do RNA foi comprometida pela hipermetilação de genes relacionados potencialmente silenciados. Adicionalmente, a análise de SNPs/INDELs em genes que compõem as redes de interação de proteínas, além de confirmarem suas identidades também destacaram polimorfismos (em relação ao genoma de *M. esculenta* v.6.1) e seus potenciais impactos, que poderiam ser influentes no comportamento de resposta ao estresse destas variedades, uma vez que alguns genes aqui relatados tiveram impactos moderados em sua estrutura. A correlação de tais polimorfismos com a tolerância ao estresse hídrico precisa ser estudada em profundidade. Assim, o presente estudo melhora a compreensão sobre a variação do metiloma de duas variedades contrastantes de BRS após estresse de déficit hídrico, com implicações relevantes para os programas de melhoramento genético da mandioca, visando o desenvolvimento de materiais de elite tolerantes a seca para *M. esculenta*.

8. Referencias bibliográficas

ABID, Ghassen et al. Analysis of DNA methylation patterns associated with drought stress response in faba bean (*Vicia faba* L.) using methylation-sensitive amplification polymorphism (MSAP). **Environmental and Experimental Botany**, v. 142, p. 34-44, 2017.

ACKAH, Michael et al. DNA methylation changes and its associated genes in mulberry (*Morus alba* L.) Yu-711 response to drought stress using MethylRAD sequencing. **Plants**, v. 11, n. 2, p. 190, 2022.

AGIUS, Dolores Rita et al. Exploring the crop epigenome: a comparison of DNA methylation profiling techniques. **Frontiers in Plant Science**, v. 14, p. 1181039, 2023.

AKHTER, Zahida et al. In response to abiotic stress, DNA methylation confers epigenetic changes in plants. **Plants**, v. 10, n. 6, p. 1096, 2021.

ALKAN, Can; COE, Bradley P.; EICHLER, Evan E. Genome structural variation discovery and genotyping. **Nature reviews genetics**, v. 12, n. 5, p. 363-376, 2011.

ALLEM, Antonio C. The origin of *Manihot esculenta* crantz (Euphorbiaceae). **Genetic resources and crop Evolution**, v. 41, n. 3, p. 133-150, 1994.

ANDREWS, Simon et al. FastQC: a quality control tool for high throughput sequence data. 2010.

ANDRÉO-SOUZA, Y. A. R. A. et al. Efeito da salinidade na germinação de sementes e no crescimento inicial de mudas de pinhão-mansão. **Revista Brasileira de Sementes**, v. 32, n. 2, p. 83-92, 2010.

ARAÚJO, Lúcia de F. et al. Equilíbrio higroscópico da palma forrageira: Relação com a umidade ótima para fermentação sólida. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 9, n. 3, p. 379-384, 2005.

BALDUCCI, Lorena et al. How do drought and warming influence survival and wood traits of *Picea mariana* saplings?. **Journal of Experimental Botany**, v. 66, n. 1, p. 377-389, 2015.

BARTELS, Arthur et al. Dynamic DNA methylation in plant growth and development. **International journal of molecular sciences**, v. 19, n. 7, p. 2144, 2018.

BEWICK, Adam J.; SCHMITZ, Robert J. Gene body DNA methylation in plants. **Current opinion in plant biology**, v. 36, p. 103-110, 2017.

BREDESON, Jessen V. et al. Sequencing wild and cultivated cassava and related species reveals extensive interspecific hybridization and genetic diversity. **Nature biotechnology**, v. 34, n. 5, p. 562-570, 2016.

BRINKMAN, Arie B. et al. Whole-genome DNA methylation profiling using MethylCap-seq. **Methods**, v. 52, n. 3, p. 232-236, 2010.

BRUGGEMAN, Quentin et al. The polyadenylation factor subunit CLEAVAGE AND POLYADENYLATION SPECIFICITY FACTOR30: a key factor of programmed cell death and a regulator of immunity in Arabidopsis. **Plant physiology**, v. 165, n. 2, p. 732-746, 2014.

BYRNE, Mary E. A role for the ribosome in development. **Trends in plant science**, v. 14, n. 9, p. 512-519, 2009.

CAO, Da et al. A quantitative profiling method of phytohormones and other metabolites applied to barley roots subjected to salinity stress. **Frontiers in Plant Science**, v. 7, p. 2070, 2017.

CARRER, Helaine; BARBOSA, André Luiz; RAMIRO, Daniel Alves. Biotecnologia na agricultura. **Estudos avançados**, v. 24, n. 70, p. 149-164, 2010.

CARVALHO, Claudine M. et al. Regulated nuclear trafficking of rpL10A mediated by NIK1 represents a defense strategy of plant cells against virus. **PLoS pathogens**, v. 4, n. 12, p. e1000247, 2008.

CHAMOVITZ, Daniel A. et al. The COP9 complex, a novel multisubunit nuclear regulator involved in light control of a plant developmental switch. **Cell**, v. 86, n. 1, p. 115-121, 1996.

CHAN, Simon W.-L.; HENDERSON, Ian R.; JACOBSEN, Steven E. Gardening the genome: DNA methylation in Arabidopsis thaliana. **Nature Reviews Genetics**, v. 6, n. 5, p. 351-360, 2005.

CHANGE, Intergovernmental Panel On Climate. Climate change 2007: Impacts, adaptation and vulnerability. **Genebra, Suíça**, 2001.

CHINNUSAMY, Viswanathan; ZHU, Jian-Kang. Epigenetic regulation of stress responses in plants. **Current opinion in plant biology**, v. 12, n. 2, p. 133-139, 2009.

CINGOLANI, Pablo et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. **fly**, v. 6, n. 2, p. 80-92, 2012.

COCK, J. H. Cassava: New potential for a neglected crop. 191. Boulder, Colorado. 1985.

COHEN-KARNI, D. et al. Kinney, Shannon R Morey, M. Yamada-Mabuchi, S.-y. Xu, T. Davis et al. **Proceedings of the National Academy of Sciences of the United States of America**, v. 108, p. 11040-11045, 2011.

COKUS, Shawn J. et al. Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. **Nature**, v. 452, n. 7184, p. 215-219, 2008.

D'AMELIA, Luisa et al. Metabolomics for crop improvement against salinity stress. **Salinity Responses and Tolerance in Plants, Volume 2: Exploring RNAi, Genome Editing and Systems Biology**, p. 267-287, 2018.

DA COSTA, Leila Ferreira. **O desafio das mudanças climáticas: os casos Brasil e China**. Paco Editorial, 2017.

DE FREITAS, Paulo André Ferreira et al. Salt acclimation in sorghum plants by exogenous proline: physiological and biochemical changes and regulation of proline metabolism. **Plant cell reports**, v. 38, n. 3, p. 403-416, 2019.

DE MATTOS, P. L. P.; FARIAS, Alba Rejane Nunes; FERREIRA FILHO, José Raimundo. **Mandioca: o produtor pergunta, a Embrapa responde**. Brasília, DF: Embrapa Informação Tecnológica; Cruz das Almas: Embrapa Mandioca e Fruticultura Tropical, 2006.

DELANEAU, Olivier et al. Haplotype estimation using sequencing reads. **The American Journal of Human Genetics**, v. 93, n. 4, p. 687-696, 2013.

DEVER, Thomas E.; GREEN, Rachel. The elongation, termination, and recycling phases of translation in eukaryotes. **Cold Spring Harbor perspectives in biology**, v. 4, n. 7, p. a013706, 2012.

DILOKPIMOL, Adiphol; GESHI, Naomi. Arabidopsis thaliana glucuronosyltransferase in family GT14. **Plant signaling e behavior**, v. 9, n. 6, p. e28891, 2014.

DYACHOK, Julia et al. SCAR mediates light-induced root elongation in Arabidopsis through photoreceptors and proteasomes. **The Plant Cell**, v. 23, n. 10, p. 3610-3626, 2011.

DYACHOK, Julia et al. SCAR mediates light-induced root elongation in Arabidopsis through photoreceptors and proteasomes. **The Plant Cell**, v. 23, n. 10, p. 3610-3626, 2011.

EL GHAZALI, Gamal EB et al. Suaeda vermiculata Forssk. ex JF Gmel.: structural characteristics and adaptations to salinity and drought: a review. **Int J Sci**, v. 9, n. 02, p. 28-33, 2020.

EMBRAPA mandioca e fruticultura, disponível em: <https://www.embrapa.br/busca-de-solucoes-tecnologicas/-/produto-servico/3608/mandioca-brs-cs-01>, acessado em 08/07/2020

FAHAD, Shah et al. Phytohormones and plant responses to salinity stress: a review. **Plant growth regulation**, v. 75, n. 2, p. 391-404, 2015.

FAROOQ, M. et al. Plant drought stress: effects, mechanisms and management. In: Sustainable agriculture. Springer, Dordrecht, p. 153-188, 2009.

FENG, Suhua et al. Efficient and accurate determination of genome-wide DNA methylation patterns in *Arabidopsis thaliana* with enzymatic methyl sequencing. **Epigenetics & chromatin**, v. 13, n. 1, p. 1-17, 2020.

FERREYRA, María Lorena Falcone et al. Plant L10 ribosomal proteins have different roles during development and translation under ultraviolet-B stress. **Plant physiology**, v. 153, n. 4, p. 1878-1894, 2010.

FROMMER, Marianne et al. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. **Proceedings of the National Academy of Sciences**, v. 89, n. 5, p. 1827-1831, 1992. GOUILLON, Quentin; KENIRY, Andrew. Latest techniques to study DNA methylation. *Essays in biochemistry*, v. 63, n. 6, p. 639-648, 2019.

FROMONT-RACINE, Micheline et al. Ribosome assembly in eukaryotes. **Gene**, v. 313, p. 17-42, 2003.

Gehring, M., e Henikoff, S. (2007). DNA methylation dynamics in plant genomes. *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression*, 1769(5-6), 276-286.

GOODSTEIN, David M. et al. Phytozome: a comparative platform for green plant genomics. **Nucleic acids research**, v. 40, n. D1, p. D1178-D1186, 2012.

Grativol, C., Hemerly, A.S. and Ferreira, P.C.G. (2012). Genetic and epigenetic regulation of stress responses in natural plant populations. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, 1819(2), pp.176–185. doi:10.1016/j.bbagrm.2011.08.010.

GRUENBAUM, Yosef et al. Sequence specificity of methylation in higher plant DNA. **Nature**, v. 292, n. 5826, p. 860-862, 1981.

GRZESIAK, Maciej T. et al. Variation among spring wheat (*triticum aestivum* L.) genotypes in response to the drought stress. II—Root system structure. **Plants**, v. 8, n. 12, p. 584, 2019.

GUO, Hui et al. The methylation patterns and transcriptional responses to chilling stress at the seedling stage in rice. **International journal of molecular sciences**, v. 20, n. 20, p. 5089, 2019.

GUO, Lihuan et al. Cullin-RING Ubiquitin Ligase Family in Plant Abiotic Stress Pathways F. **Journal of integrative plant biology**, v. 55, n. 1, p. 21-30, 2013.

GUO, Peiguo et al. Differentially expressed genes between drought-tolerant and drought-sensitive barley genotypes in response to drought stress during the reproductive stage. **Journal of experimental botany**, v. 60, n. 12, p. 3531-3544, 2009.

GUO, Rui et al. Metabolic responses to drought stress in the tissues of drought-tolerant and drought-sensitive wheat genotype seedlings. **AoB Plants**, v. 10, n. 2, p. ply016, 2018.

HAHN, S. K. et al. Resistance of cassava clones to cassava green mite, *Mononychellus tanajoa*. **International Journal of Pest Management**, v. 26, n. 3, p. 265-267, 1980.

HE, Xin-Jian; CHEN, Taiping; ZHU, Jian-Kang. Regulation and function of DNA methylation in plants and animals. **Cell research**, v. 21, n. 3, p. 442-465, 2011.

HIND, Sarah R. et al. The COP9 signalosome controls jasmonic acid synthesis and plant responses to herbivory and pathogens. **The Plant Journal**, v. 65, n. 3, p. 480-491, 2011.

HIRAYAMA, Takashi; SHINOZAKI, Kazuo. Research on plant abiotic stress responses in the post-genome era: past, present and future. **The Plant Journal**, v. 61, n. 6, p. 1041-1052, 2010.

HOAGLAND, Dennis Robert et al. The water-culture method for growing plants without soil. **Circular. California agricultural experiment station**, v. 347, n. 2nd edit, 1950.

HOGG, Rebecca; MCGRAIL, Joanne C.; O'KEEFE, Raymond T. The function of the NineTeen Complex (NTC) in regulating spliceosome conformations and fidelity during pre-mRNA splicing. **Biochemical Society Transactions**, v. 38, n. 4, p. 1110-1115, 2010.

HORIGUCHI, Gorou et al. Ribosomes and translation in plant developmental control. **Plant Science**, v. 191, p. 24-34, 2012.

HOWARD, Brian E. et al. High-throughput RNA sequencing of pseudomonas-infected *Arabidopsis* reveals hidden transcriptome complexity and novel splice variants. **PLoS One**, v. 8, n. 10, p. e74183, 2013.

HUANG, Guo-Tao et al. Signal transduction during cold, salt, and drought stresses in plants. **Molecular biology reports**, v. 39, n. 2, p. 969-987, 2012.

HUANG, Xuan; TRAGANOS, Frank; DARZYNKIEWICZ, Zbigniew. DNA damage induced by DNA topoisomerase I-and topoisomerase II-inhibitors detected by histone H2AXphosphorylation in relation to the cell cycle phase and apoptosis. **Cell Cycle**, v. 2, n. 6, p. 613-618, 2003.

ILYAS, Noshin et al. Influence of salicylic acid and jasmonic acid on wheat under drought stress. **Communications in soil science and plant analysis**, v. 48, n. 22, p. 2715-2723, 2017.

ISAYENKOV, Stanislav V.; MAATHUIS, Frans JM. Plant salinity stress: many unanswered questions remain. **Frontiers in plant science**, v. 10, p. 80, 2019.

JENNINGS, D. L.; IGLESIAS, C. Breeding for crop improvement. In: **Cassava: biology, production and utilization**. Wallingford UK: CABI, 2001. p. 149-166.

JENNINGS, D.L. Breeding for resistance to African cassava mosaic disease: Progress and prospects. In Proceedings of the International Workshop, Muguga, Kenya, 19–22 February 1976; pp. 39–44. Available online: <http://hdl.handle.net/10625/20886> (accessed on 20 April 2022).

JHA, Uday Chand et al. Salinity stress response and ‘omics’ approaches for improving salinity stress tolerance in major grain legumes. **Plant cell reports**, v. 38, n. 3, p. 255-277, 2019.

JIMENO, Sonia; AGUILERA, Andrés. The THO complex as a key mRNP biogenesis factor in development and cell differentiation. **Journal of Biology**, v. 9, n. 1, p. 1-3, 2010.

JOSHI, N.; SICKLE, F. J. A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33). 2011. 2020.

JULLIEN, Pauline E. et al. DNA methylation dynamics during sexual reproduction in *Arabidopsis thaliana*. **Current Biology**, v. 22, n. 19, p. 1825-1830, 2012.

KAMAL, Khaled Y. et al. Modulation of cell cycle progression and chromatin dynamic as tolerance mechanisms to salinity and drought stress in maize. **Physiologia Plantarum**, v. 172, n. 2, p. 684-695, 2021.

KEPINSKI, Stefan; LEYSER, Ottoline. Plant development: auxin in loops. **Current Biology**, v. 15, n. 6, p. R208-R210, 2005.

KIEFER, Hélène et al. Inositol 1, 4, 5-triphosphate receptor-binding protein released with inositol 1, 4, 5-triphosphate (IRBIT) associates with components of the mRNA 3' processing machinery in a phosphorylation-dependent manner and inhibits polyadenylation. **Journal of Biological Chemistry**, v. 284, n. 16, p. 10694-10705, 2009.

KONCZ, Csaba et al. The spliceosome-activating complex: molecular mechanisms underlying the function of a pleiotropic regulator. **Frontiers in plant science**, v. 3, p. 9, 2012.

KUMAKURA, Naoyoshi et al. *Arabidopsis* AtRRP44A is the functional homolog of Rrp44/Dis3, an exosome component, is essential for viability and is required for RNA processing and degradation. **PLoS One**, v. 8, n. 11, p. e79219, 2013.

KUON, Joel-E. et al. Haplotype-resolved genomes of geminivirus-resistant and geminivirus-susceptible African cassava cultivars. **BMC biology**, v. 17, n. 1, p. 1-15, 2019.

LATHRAP, Donald Ward; LATHRAP, Donald W. **The upper amazon**. Southampton: Thames e Hudson, 1970.

LECHNER, Esther et al. F-box proteins everywhere. **Current opinion in plant biology**, v. 9, n. 6, p. 631-638, 2006.

LÉOTARD, Guillaume et al. Phylogeography and the origin of cassava: new insights from the northern rim of the Amazonian basin. **Molecular Phylogenetics and Evolution**, v. 53, n. 1, p. 329-334, 2009.

LI, Liangliang et al. The Arabidopsis DNA mismatch repair gene PMS1 restricts somatic recombination between homeologous sequences. **Plant Molecular Biology**, v. 69, n. 6, p. 675-684, 2009.

LI, Lishu et al. Identification and characterization of the mango eIF gene family reveals MielF1A-a, which confers tolerance to salt stress in transgenic Arabidopsis. **Scientia horticulturae**, v. 248, p. 274-281, 2019.

LI, Shizhao; TOLLEFSBOL, Trygve O. DNA methylation methods: Global DNA methylation and methylomic analyses. **Methods**, v. 187, p. 28-43, 2021.

LI, Zengqiang et al. 5-azacytidine pre-treatment alters DNA methylation levels and induces genes responsive to salt stress in kenaf (*Hibiscus cannabinus* L.). **Chemosphere**, v. 271, p. 129562, 2021.

LI, Heng; DURBIN, Richard. Fast and accurate short read alignment with Burrows–Wheeler transform. **bioinformatics**, v. 25, n. 14, p. 1754-1760, 2009.

LINDQUIST, Susan et al. Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. In: **Self-Perpetuating Structural States in Biology, Disease, and Genetics**. National Academies Press (US), 2002.

LISAR, S. Y. et al. Causes, effects and responses. **Water stress**, v. 25, n. 1, p. 33, 2012.

LIU, Lin; CHEN, Xuemei. RNA quality control as a key to suppressing RNA silencing of endogenous genes in plants. **Molecular plant**, v. 9, n. 6, p. 826-836, 2016.

LIU, Songyu et al. Tomato AUXIN RESPONSE FACTOR 5 regulates fruit set and development via the mediation of auxin and gibberellin signaling. **Scientific reports**, v. 8, n. 1, p. 1-16, 2018.

LIU, Yule et al. Tobacco Rar1, EDS1 and NPR1/NIM1 like genes are required for N-mediated resistance to tobacco mosaic virus. **The Plant Journal**, v. 30, n. 4, p. 415-429, 2002.

LORINCZ, Matthew C. et al. Intragenic DNA methylation alters chromatin structure and elongation efficiency in mammalian cells. **Nature structural & molecular biology**, v. 11, n. 11, p. 1068-1075, 2004.

LUKOWITZ, Wolfgang et al. Arabidopsis cyt1 mutants are deficient in a mannose-1-phosphate guanylyltransferase and point to a requirement of N-linked

glycosylation for cellulose biosynthesis. **Proceedings of the National Academy of Sciences**, v. 98, n. 5, p. 2262-2267, 2001.

LUO, Yanjie et al. An A rabidopsis homolog of importin β 1 is required for ABA response and drought tolerance. **The Plant Journal**, v. 75, n. 3, p. 377-389, 2013.

LUSCOMBE, Nicholas M.; GREENBAUM, Dov; GERSTEIN, Mark. What is bioinformatics? A proposed definition and overview of the field. **Methods of information in medicine**, v. 40, n. 04, p. 346-358, 2001.

LYAPINA, S. et al. COP9 signalosome promotes cleavage of NEDD8-CUL1 conjugates. **Science**, v. 292, p. 1382-1385, 2001.

MASTRANGELO, Anna M. et al. Alternative splicing: enhancing ability to cope with stress via transcriptome plasticity. **Plant Science**, v. 185, p. 40-49, 2012.

MCCALLUM, Emily J.; ANJANAPPA, Ravi B.; GRUISSEM, Wilhelm. Tackling agriculturally relevant diseases in the staple crop cassava (*Manihot esculenta*). Current opinion in plant biology, v. 38, p. 50-58, 2017.

MCKENNA, Aaron et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. **Genome research**, v. 20, n. 9, p. 1297-1303, 2010.

MISSBACH, Sandra et al. 40S ribosome biogenesis co-factors are essential for gametophyte and embryo development. **PloS one**, v. 8, n. 1, p. e54084, 2013.

MOIN, Mazahar et al. Activation tagging in indica rice identifies ribosomal proteins as potential targets for manipulation of water-use efficiency and abiotic stress tolerance in plants. **Plant, cell e environment**, v. 39, n. 11, p. 2440-2459, 2016.[A]

MOIN, Mazahar et al. Rice ribosomal protein large subunit genes and their spatio-temporal and stress regulation. **Frontiers in plant science**, v. 7, p. 1284, 2016.[B]

MOROTOMI-YANO, Keiko et al. Dynamic behavior of DNA topoisomerase II β in response to DNA double-strand breaks. **Scientific reports**, v. 8, n. 1, p. 1-14, 2018.

MOUSSA, Helal Ragab et al. Comparative response of drought tolerant and drought sensitive maize genotypes to water stress. **Australian Journal of Crop Science**, v. 1, n. 1, p. 31-36, 2008.

MUKHOPADHYAY, Pradipto et al. Transcriptional downregulation of rice rpL32 gene under abiotic stress is associated with removal of transcription factors within the promoter region. **PLoS One**, v. 6, n. 11, p. e28058, 2011.

NAGPAL, Punita et al. Auxin response factors ARF6 and ARF8 promote jasmonic acid production and flower maturation. 2005.

NAKABAYASHI, Ryo et al. Enhancement of oxidative and drought tolerance in Arabidopsis by overaccumulation of antioxidant flavonoids. **The Plant Journal**, v. 77, n. 3, p. 367-379, 2014.

NANDA, Jagpreet S. et al. Coordinated movements of eukaryotic translation initiation factors eIF1, eIF1A, and eIF5 trigger phosphate release from eIF2 in response to start codon recognition by the ribosomal preinitiation complex. **Journal of Biological Chemistry**, v. 288, n. 8, p. 5316-5329, 2013.

NARASIMHAN, Vagheesh et al. BCFtools/RoH: a hidden Markov model approach for detecting autozygosity from next-generation sequencing data. **Bioinformatics**, v. 32, n. 11, p. 1749-1751, 2016.

NASSAR, N. M. A. et al. Cassava genetic resources: wild species and indigenous cultivars and their utilization for breeding of the crop. In: **Cassava improvement to enhance livelihoods in Sub-Saharan African and Northeastern Brazil. First International Meeting on Cassava Breeding, Biotechnology and Ecology, Brasilia, Brazil, 11-15 November 2006**. Thesaurus Editora, 2007. p. 5-30.

NASSAR, N. M. A.; ORTIZ, Rodomiro. Cassava improvement: challenges and impacts. **The Journal of Agricultural Science**, v. 145, n. 2, p. 163-171, 2007.

NASSAR, Nagib. Cytogenetics and evolution of cassava (*Manihot esculenta* Crantz). **Genetics and Molecular Biology**, v. 23, p. 1003-1014, 2000.

NAVADA, Sharada et al. Biofilms remember: osmotic stress priming as a microbial management strategy for improving salinity acclimation in nitrifying biofilms. **Water Research**, v. 176, p. 115732, 2020.

NUWAMANYA, E. et al. Bio-ethanol production from non-food parts of cassava (*Manihot esculenta* Crantz). **Ambio**, v. 41, n. 3, p. 262-270, 2012.

OGBONNA, Alex C. et al. Comprehensive genotyping of a Brazilian cassava (*Manihot esculenta* Crantz) germplasm bank: insights into diversification and domestication. **Theoretical and Applied Genetics**, v. 134, p. 1343-1362, 2021.

OMIDBAKHSHFARD, Mohammad Amin et al. Effect of salt stress on genes encoding translation-associated proteins in *Arabidopsis thaliana*. **Plant signaling e behavior**, v. 7, n. 9, p. 1095-1102, 2012.

PAN, Ting et al. Non-stomatal limitation of photosynthesis by soil salinity. **Critical Reviews in Environmental Science and Technology**, v. 51, n. 8, p. 791-825, 2021.

PARMAR, Aditya; STURM, Barbara; HENSEL, Oliver. Crops that feed the world: Production and improvement of cassava for food, feed, and industrial uses. **Food Security**, v. 9, n. 5, p. 907-927, 2017.

PARRY, Geraint; ESTELLE, Mark. Regulation of cullin-based ubiquitin ligases by the Nedd8/RUB ubiquitin-like proteins. In: **Seminars in cell e developmental biology**. Academic Press, 2004. p. 221-229.

PICKART, Cecile M. Ubiquitin enters the new millennium. **Molecular cell**, v. 8, n. 3, p. 499-504, 2001.

PIKAARD, Craig S.; SCHEID, Ortrun Mittelsten. Epigenetic regulation in plants. **Cold Spring Harbor perspectives in biology**, v. 6, n. 12, p. a019315, 2014.

PILLAY, Indiren; BEYL, Caula. Early responses of drought-resistant and-susceptible tomato plants subjected to water stress. **Journal of plant growth regulation**, v. 9, p. 213-219, 1990.

PROCHNIK, Simon et al. The cassava genome: current progress, future directions. **Tropical plant biology**, v. 5, n. 1, p. 88-94, 2012.

QIAN, Yexiong et al. The Dynamics of DNA methylation in the maize (Zea mays L.) inbred line B73 response to heat stress at the seedling stage. **Biochemical and biophysical research communications**, v. 512, n. 4, p. 742-749, 2019.

RAMPINO, Patrizia et al. Drought stress response in wheat: physiological and molecular analysis of resistant and sensitive genotypes. **Plant, cell & environment**, v. 29, n. 12, p. 2143-2152, 2006.

RAO, Abdul Qayyum et al. Genomics of salinity tolerance in plants. **Plant Genomics**, v. 2016, p. 273-299, 2016.

REDDY, B. V. S. et al. Bio-fuel crops research for energy security and rural development in developing countries. **Bioenergy Research**, v. 1, n. 3-4, p. 248-258, 2008.

RIBEIRO, M. R.; FREIRE, F. J.; MONTENEGRO, AA de A. Solos halomórficos no Brasil: Ocorrência, gênese, classificação, uso e manejo sustentável. **CURI, N.; MARQUES, JJ; GUILHERME, LRG; LIMA, JM**, p. 165-208, 2003.

ROA, Ana C. et al. AFLP analysis of relationships among cassava and other Manihot species. **Theoretical and Applied Genetics**, v. 95, n. 5, p. 741-750, 1997.

ROA, Ana C. et al. Cross-species amplification of cassava (Manihot esculenta)(Euphorbiaceae) microsatellites: allelic polymorphism and degree of relationship. **American Journal of Botany**, v. 87, n. 11, p. 1647-1655, 2000.

ROBERTS, Richard J.; CARNEIRO, Mauricio O.; SCHATZ, Michael C. The advantages of SMRT sequencing. **Genome biology**, v. 14, n. 6, p. 1-4, 2013.

ROGERS, D. J.; APPAN, S. G. Manihot and Manihotoides (Euphorbiaceae): A computer assisted study. **Flora neotropica**, v. 13, 1973.

ROJAS, G. G.; FABRE, N. Agroecologia e mudanças climáticas no Trópico Semiárido. **Redes**, Santa Cruz do Sul, Universidade de Santa Cruz do Sul, v. 22, n. 2, 2017.

SAKURAI, Tetsuya et al. **Genome-wide discovery and information resource development of DNA polymorphisms in cassava**. PLoS One, v. 8, n. 9, p. e74056, 2013.

SANTOS, Ariana S. et al. Citrus scion and rootstock combinations show changes in DNA methylation profiles and ABA insensitivity under recurrent drought conditions. **Scientia Horticulturae**, v. 267, p. 109313, 2020.

SAZE, Hidetoshi; SCHEID, Ortrun Mittelsten; PASZKOWSKI, Jerzy. Maintenance of CpG methylation is essential for epigenetic inheritance during plant gametogenesis. **Nature genetics**, v. 34, n. 1, p. 65-69, 2003.

SCHULZ, Benjamin; ECKSTEIN, R. Lutz; DURKA, Walter. Scoring and analysis of methylation-sensitive amplification polymorphisms for epigenetic population studies. **Molecular ecology resources**, v. 13, n. 4, p. 642-653, 2013.

SEBRAE - SERVIÇO BRASILEIRO DE APOIO ÀS MICRO E PEQUENAS EMPRESAS. Escola Superior de Propaganda e Marketing (ESPM). Estudo de Mercado sobre a Mandioca (Farinha e Fécula). 2008. Disponível em: <http://atividaderural.com.br/artigos/5602f3e181880.pdf>. Acesso em: 13.07.2019.

SILVA, Evandro Nascimento da et al. Salt stress induced damages on the photosynthesis of physic nut young plants. **Scientia Agricola**, v. 68, n. 1, p. 62-68, 2011.

SIMOVA-STOILLOVA, Lyudmila et al. Proteolytic activity and cysteine protease expression in wheat leaves under severe soil drought and recovery. **Plant Physiology and Biochemistry**, v. 48, n. 2-3, p. 200-206, 2010.

SINGH, Amit Kumar; CHAMOVITZ, Daniel A. Role of Cop9 signalosome subunits in the environmental and hormonal balance of plant. **Biomolecules**, v. 9, n. 6, p. 224, 2019.

SMALLE, Jan; VIERSTRA, Richard D. The ubiquitin 26S proteasome proteolytic pathway. **Annual review of plant biology**, v. 55, n. 1, p. 555-590, 2004.

SONG, Xiaowen et al. Identification and functional characterization of methyl-CpG binding domain protein from *Tribolium castaneum*. **Genomics**, v. 112, n. 3, p. 2223-2232, 2020.

STEVENS, Rebecca et al. Two E2F sites in the Arabidopsis MCM3 promoter have different roles in cell cycle activation and meristematic expression. **Journal of Biological Chemistry**, v. 277, n. 36, p. 32978-32984, 2002.

SU, Lingye et al. The transcription factor VaNAC17 from grapevine (*Vitis amurensis*) enhances drought tolerance by modulating jasmonic acid biosynthesis in transgenic Arabidopsis. **Plant cell reports**, v. 39, n. 5, p. 621-634, 2020.

SUN, Daoyang et al. LrABCF1, a GCN-type ATP-binding cassette transporter from *Lilium regale*, is involved in defense responses against viral and fungal pathogens. **Planta**, v. 244, n. 6, p. 1185-1199, 2016.

SUN, Li-Jun et al. Electrochemical mapping of indole-3-acetic acid and salicylic acid in whole pea seedlings under normal conditions and salinity. **Sensors and Actuators B: Chemical**, v. 276, p. 545-551, 2018.

SZYMCZYNA, Blair R. et al. Structure and function of the PWI motif: a novel nucleic acid-binding domain that facilitates pre-mRNA processing. **Genes e development**, v. 17, n. 4, p. 461-475, 2003.

TAIZ, Lincoln et al. **Fisiologia e desenvolvimento vegetal**. Artmed Editora, 2017.

TAKEDA, S.; MATSUOKA, M. Genetic approaches to crop improvement: responding to environmental and population changes. **Nature Reviews Genetics**, v.9, p.444-57, 2008.

VAISVILA, Romualdas et al. Enzymatic methyl sequencing detects DNA methylation at single-base resolution from picograms of DNA. **Genome research**, v. 31, n. 7, p. 1280-1289, 2021.

VALLE, Teresa Losada; LORENZI, José Osmar. Variedades melhoradas de mandioca como instrumento de inovação, segurança alimentar, competitividade e sustentabilidade: contribuições do instituto agrônomo de campinas (IAC). **Cadernos de Ciência e Tecnologia**, v. 31, n. 1, p. 15-34, 2014.

VAN LIJSEBETTENS, Mieke; GRASSER, Klaus D. Transcript elongation factors: shaping transcriptomes after transcript initiation. **Trends in plant science**, v. 19, n. 11, p. 717-726, 2014..

VANNESTE, Steffen; FRIML, Jiří. Auxin: a trigger for change in plant development. **Cell**, v. 136, n. 6, p. 1005-1016, 2009.

VASEVA, I. et al. The response of plants to drought stress: the role of dehydrins, chaperones, proteases and protease inhibitors in maintaining cellular protein function. **Droughts: new research**, v. 1, p. 1-45, 2012.

VELMURUGAN, Ayyam et al. Water demand and salinity. **desalination-challenges and opportunities**, 2020.

VILLAJUANA-BONEQUI, Mitzi et al. Cell type specific transcriptional reprogramming of maize leaves during *Ustilago maydis* induced tumor formation. **Scientific reports**, v. 9, n. 1, p. 1-15, 2019.

WANG, Shi et al. MethylRAD: a simple and scalable method for genome-wide DNA methylation profiling using methylation-dependent restriction enzymes. **Open biology**, v. 5, n. 11, p. 150130, 2015.

WANG, Wensheng et al. Comparative analysis of DNA methylation changes in two rice genotypes under salt stress and subsequent recovery. **Biochemical and Biophysical Research Communications**, v. 465, n. 4, p. 790-796, 2015. [B]

WANG, Wensheng et al. Genome-wide differences in DNA methylation changes in two contrasting rice genotypes in response to drought conditions. **Frontiers in plant science**, v. 7, p. 1675, 2016.

WANG, Xiao et al. Improved tolerance to drought stress after anthesis due to priming before anthesis in wheat (*Triticum aestivum* L.) var. Vinjett. **Journal of experimental botany**, v. 65, n. 22, p. 6441-6456, 2014.

WANNER, Heinz et al. Mid-to Late Holocene climate change: an overview. **Quaternary Science Reviews**, v. 27, n. 19-20, p. 1791-1828, 2008

WARDENAAR, René et al. Evaluation of MeDIP-chip in the context of whole-genome bisulfite sequencing (WGBS-seq) in Arabidopsis. **Tiling Arrays: Methods and Protocols**, p. 203-224, 2013.

WATERWORTH, Wanda M. et al. Arabidopsis TAF 1 is an MRE 11-interacting protein required for resistance to genotoxic stress and viability of the male gametophyte. **The Plant Journal**, v. 84, n. 3, p. 545-557, 2015.

WEE, Susan et al. CSN facilitates Cullin–RING ubiquitin ligase function by counteracting autocatalytic adapter instability. **Nature cell biology**, v. 7, n. 4, p. 387-391, 2005.

WOLF, Sebastian; GREINER, Steffen. Growth control by cell wall pectins. **Protoplasma**, v. 249, n. 2, p. 169-175, 2012.

WU, Lingyun et al. Effects of solid matrix priming (SMP) and salt stress on broccoli and cauliflower seed germination and early seedling growth. **Scientia horticulturae**, v. 255, p. 161-168, 2019.

WU, Rentian et al. H3K9me3 demethylase Kdm4d facilitates the formation of pre-initiative complex and regulates DNA replication. **Nucleic acids research**, v. 45, n. 1, p. 169-180, 2017.

www.embrapa.br. (2016). BRS Formosa: variedade de mandioca recomendada para uso industrial nas microrregiões de Valença, Jequié e Santo Antônio de Jesus (BA). - Portal Embrapa. [online] Available at: <https://www.embrapa.br/busca-de-publicacoes/-/publicacao/1059201/brs-formosa-variedade-de-mandioca-recomendada-para-uso-industrial-nas-microrregioes-de-valenca-jequie-e-santo-antonio-de-jesus-ba> [Accessed 31 Aug. 2022].

www.embrapa.br. (2021). BRS Dourada e BRS Gema de Ovo: variedades de mandioca de mesa biofortificadas para plantio em Roraima. - Portal Embrapa. [online] Available at: <https://www.embrapa.br/busca-de-publicacoes/->

/publicacao/1132805/brs-dourada-e-brs-gema-de-ovo-variedades-de-mandioca-de-mesa-biofortificadas-para-plantio-em-roraima [Accessed 31 Aug. 2022].

XU, Fa-Qing; XUE, Hong-Wei. The ubiquitin-proteasome system in plant responses to environments. **Plant, cell e environment**, v. 42, n. 10, p. 2931-2944, 2019.

YAN, Dawei et al. Sphingolipid biosynthesis modulates plasmodesmal ultrastructure and phloem unloading. **Nature plants**, v. 5, n. 6, p. 604-615, 2019.

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

ZHANG, Huiming et al. Abiotic stress responses in plants. **Nature Reviews Genetics**, v. 23, n. 2, p. 104-119, 2022.

ZHANG, Huiming et al. An Rrp6-like protein positively regulates noncoding RNA levels and DNA methylation in Arabidopsis. **Molecular cell**, v. 54, n. 3, p. 418-430, 2014.

ZHANG, Jingxian et al. A polyadenylation factor subunit implicated in regulating oxidative signaling in Arabidopsis thaliana. **PloS one**, v. 3, n. 6, p. e2410, 2008.

ZHANG, Qi; DAI, Wenhao. Plant response to salinity stress. In: **Stress Physiology of Woody Plants**. CRC Press, 2019. p. 155-173.

ZHANG, Qian et al. N6-methyladenine DNA methylation in Japonica and Indica rice genomes and its association with gene expression, plant development, and stress responses. **Molecular plant**, v. 11, n. 12, p. 1492-1508, 2018.

ZHANG, Weiping; MURPHY, Caroline; SIEBURTH, Leslie E. Conserved RNaseII domain protein functions in cytoplasmic mRNA decay and suppresses Arabidopsis decapping mutant phenotypes. **Proceedings of the National Academy of Sciences**, v. 107, n. 36, p. 15981-15985, 2010.

ZHANG, Wenyan et al. An Arabidopsis purple acid phosphatase with phytase activity increases foliar ascorbate. **Plant physiology**, p. 431-440, 2008.

ZHANG, Zenglin; XU, Mengmeng; GUO, Yongfeng. Ring/U-box Protein AtUSR1 functions in promoting leaf senescence through JA signaling pathway in Arabidopsis. **Frontiers in plant science**, v. 11, p. 608589, 2020.

ZHAO, Boqing et al. Roles of phytohormones and their signaling pathways in leaf development and stress responses. **Journal of Agricultural and Food Chemistry**, v. 69, n. 12, p. 3566-3584, 2021.

ZHOU, Fujun; ROY, Bijoyita; VON ARNIM, Albrecht G. Translation reinitiation and development are compromised in similar ways by mutations in translation initiation factor eIF3h and the ribosomal protein RPL24. **BMC Plant Biology**, v. 10, n. 1, p. 1-11, 2010.

ZILBERMAN, Daniel et al. Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. **Nature genetics**, v. 39, n. 1, p. 61-69, 2007.

ZÖRB, Christian; GEILFUS, C.-M.; DIETZ, K.-J. Salinity and crop yield. **Plant biology**, v. 21, p. 31-38, 2019.

9. Apêndice

Figura suplementar 1. Gráfico heatmap de genes hipometilados ($p\text{-value} \leq 0,05$ e $\log_2\text{FC} > 1$) em sítios CCGG da variedade Formosa (FOR), apresentado pelas réplicas biológicas (R1, R2 e R3) na condição de seca (SEV) ou de controle (CONT) sem estresse, tendo por base $\log_2$ das contagens RPM (leituras por milhão). Legenda: A01 (FORSEV-R1); A02 (FORSEV-R2); A03 (FORSEV-R3); A04 (FORCONT-R1); A05 (FORCONT-R2); A06 (FORCONT-R3)

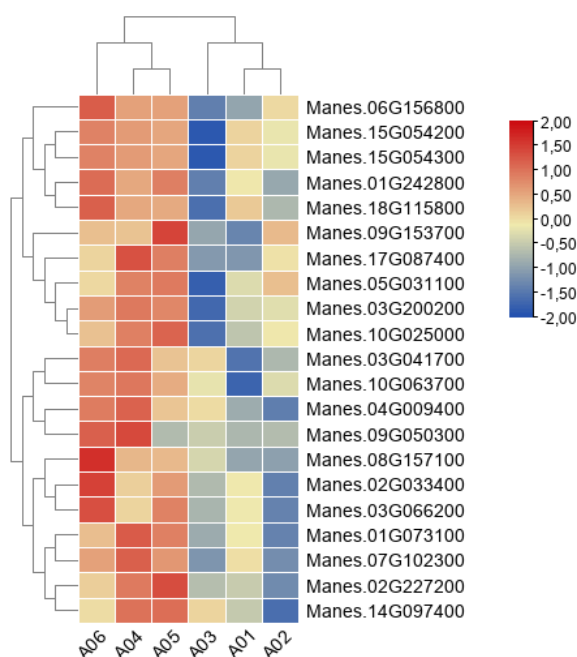


Figura suplementar 2. Gráfico heatmap de genes hipometilados ($p\text{-value} \leq 0,05$ e $\log_2\text{FC} > 1$) em sítios CCNGG da variedade Formosa (FOR), apresentado pelas réplicas biológicas (R1, R2 e R3) na condição de seca (SEV) ou de controle (CONT) sem estresse, tendo por base $\log_2$ das contagens RPM (leituras por milhão). Legenda: A01 (FORSEV-R1); A02 (FORSEV-R2); A03 (FORSEV-R3); A04 (FORCONT-R1); A05 (FORCONT-R2); A06 (FORCONT-R3)

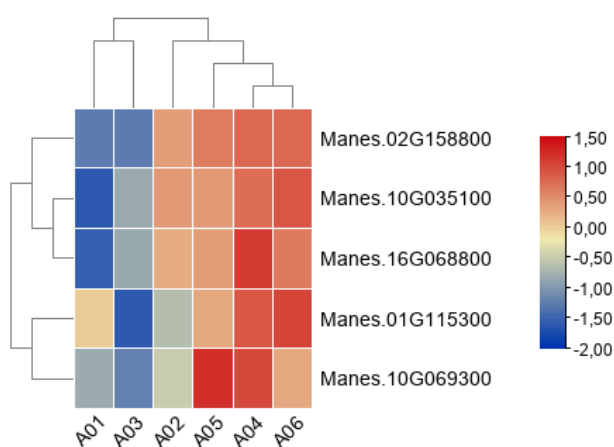


Figura suplementar 3. Gráfico heatmap de genes hipermetilados ($p\text{-value} \leq 0,05$ e $\log_2\text{FC} > 1$) em sítios CCGG da variedade Formosa (FOR), apresentado pelas réplicas biológicas (R1, R2 e R3) na condição de seca (SEV) ou de controle (CONT) sem estresse, tendo por base $\log_2$ das contagens RPM (leituras por milhão). Legenda: A01 (FORSEV-R1); A02 (FORSEV-R2); A03 (FORSEV-R3); A04 (FORCONT-R1); A05 (FORCONT-R2); A06 (FORCONT-R3).

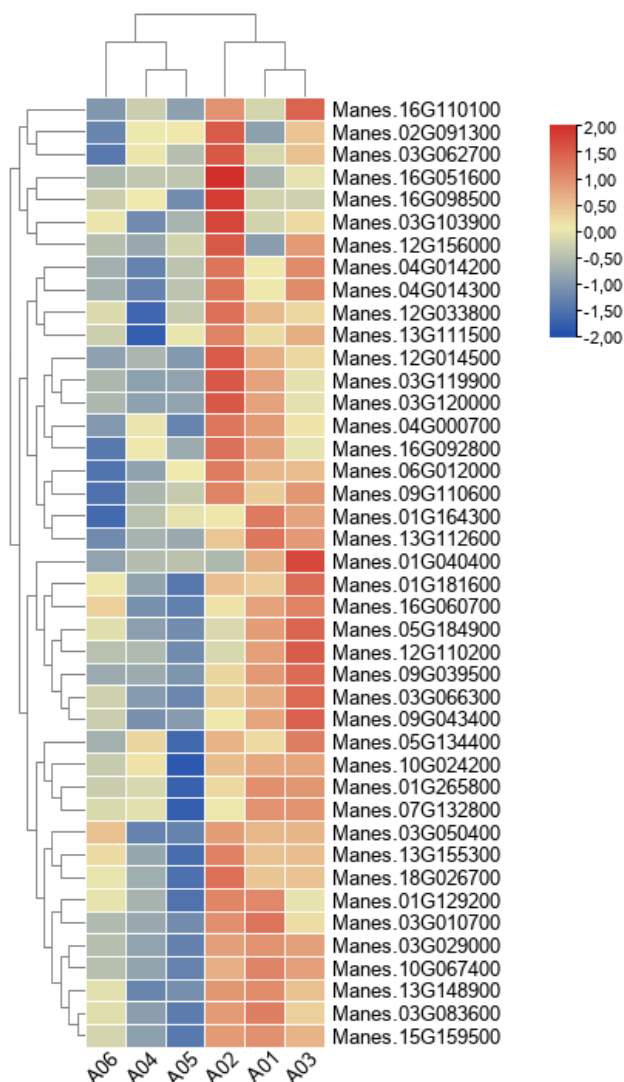


Figura suplementar 4. Gráfico heatmap de genes hipermetilados ($p\text{-value} \leq 0,05$ e $\log_2\text{FC} > 1$) em sítios CCNGG da variedade Formosa (FOR), apresentado pelas réplicas biológicas (R1, R2 e R3) na condição de seca (SEV) ou de controle (CONT) sem estresse, tendo por base $\log_2$ das contagens RPM (leituras por milhão). Legenda: A01 (FORSEV-R1); A02 (FORSEV-R2); A03 (FORSEV-R3); A04 (FORCONT-R1); A05 (FORCONT-R2); A06 (FORCONT-R3)

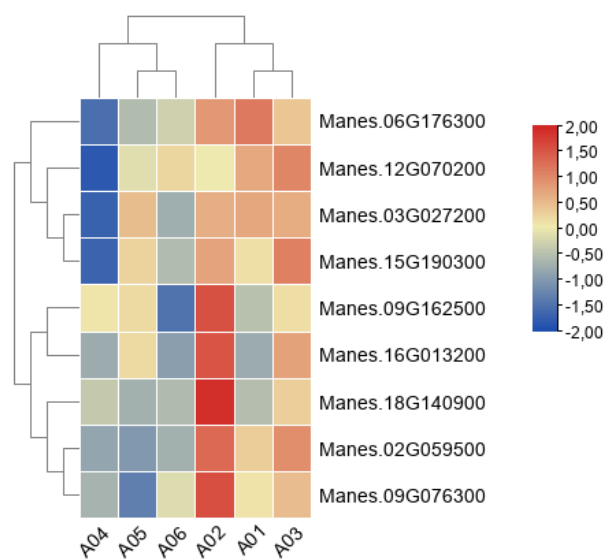


Figura suplementar 5. Gráfico heatmap de genes hipometilados ($p\text{-value} \leq 0,05$ e $\log_2\text{FC} > 1$) em sítios CCGG da variedade Dourada (DOU), apresentado pelas réplicas biológicas (R1, R2 e R3) na condição de seca (SEV) ou de controle (CONT) sem estresse, tendo por base $\log_2$ das contagens RPM (leituras por milhão). Legenda: A07 (DOUSEV-R1); A08 (DOUSEV-R2); A09 (DOUSEV-R3); A10 (DOUCONT-R1); A11 (DOUCONT-R2); A12 (DOUCONT-R3)

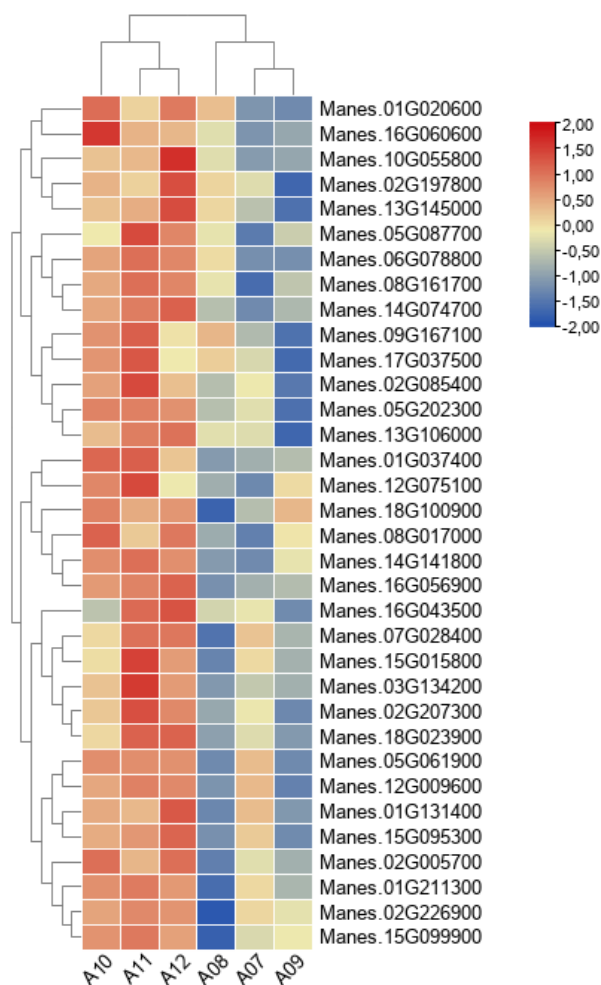


Figura suplementar 6. Gráfico heatmap de genes hipermetilados ($p\text{-value} \leq 0,05$ e $\log_2\text{FC} > 1$) em sítios CCGG da variedade Dourada (DOU), apresentado pelas réplicas biológicas (R1, R2 e R3) na condição de seca (SEV) ou de controle (CONT) sem estresse, tendo por base $\log_2$ das contagens RPM (leituras por milhão). Legenda: A07 (DOUSEV-R1); A08 (DOUSEV-R2); A09 (DOUSEV-R3); A10 (DOUCONT-R1); A11 (DOUCONT-R2); A12 (DOUCONT-R3)

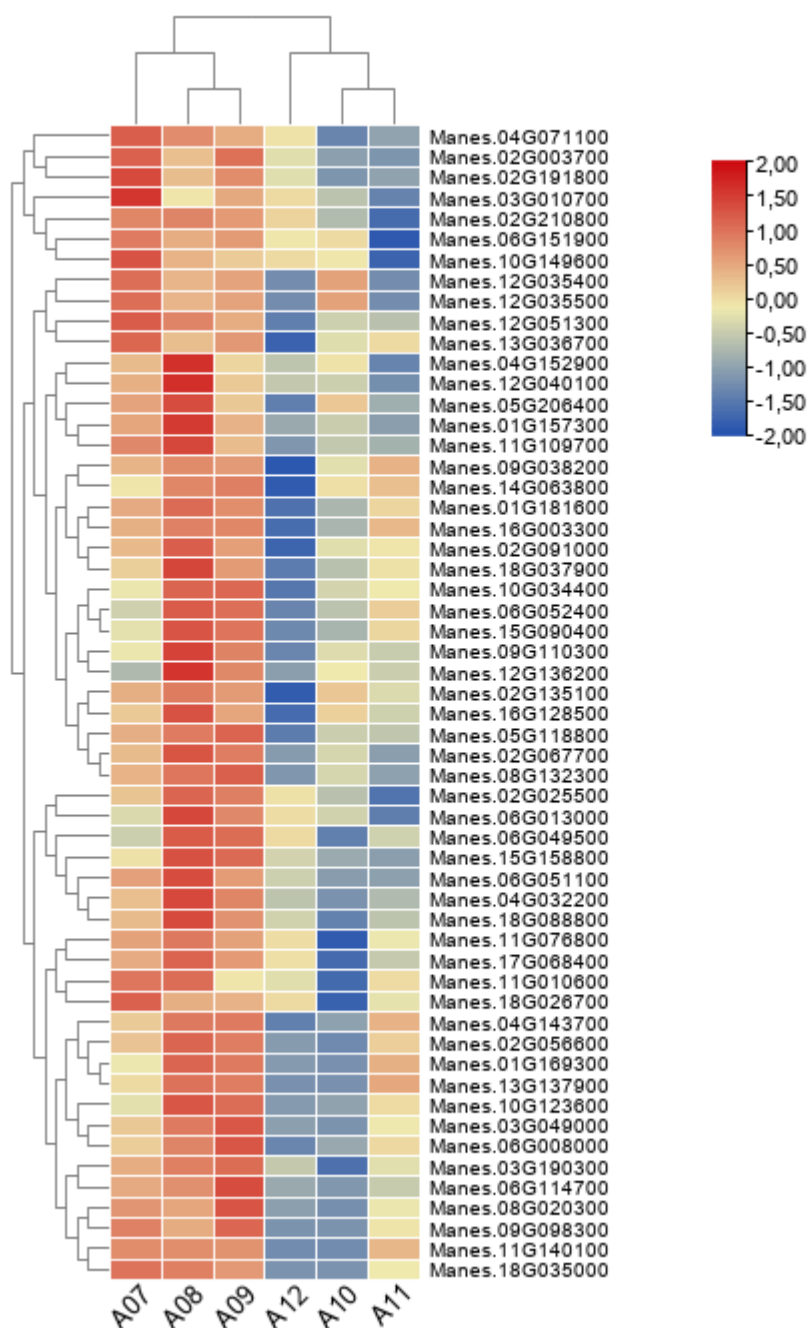


Figura suplementar 7. Gráfico heatmap de genes hipometilados ($p\text{-value} \leq 0,05$ e $\log_2\text{FC} > 1$) em sítios CCNGG da variedade Dourada (DOU), apresentado pelas réplicas biológicas (R1, R2 e R3) na condição de seca (SEV) ou de controle (CONT) sem estresse, tendo por base $\log_2$ das contagens RPM (leituras por milhão). Legenda: A07 (DOUSEV-R1); A08 (DOUSEV-R2); A09 (DOUSEV-R3); A10 (DOUCONT-R1); A11 (DOUCONT-R2); A12 (DOUCONT-R3)

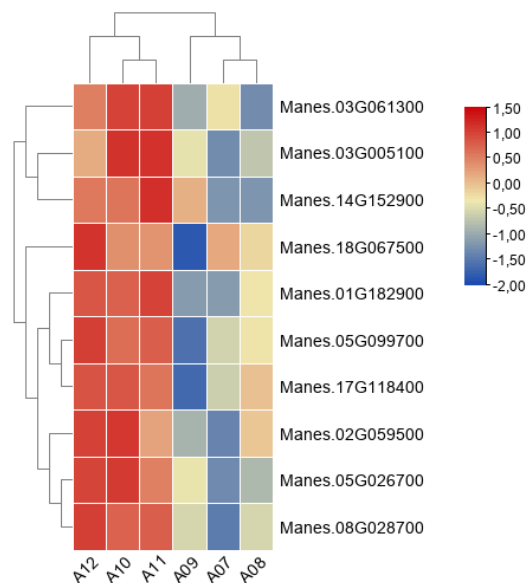


Figura suplementar 8. Gráfico heatmap de genes hipermetilados ($p\text{-value} \leq 0,05$ e $\log_2FC > 1$) em sítios CCNGG da variedade Dourada (DOU), apresentado pelas réplicas biológicas (R1, R2 e R3) na condição de seca (SEV) ou de controle (CONT) sem estresse, tendo por base $\log_2$ das contagens RPM (leituras por milhão). Legenda: A07 (DOUSEV-R1); A08 (DOUSEV-R2); A09 (DOUSEV-R3); A10 (DOUCONT-R1); A11 (DOUCONT-R2); A12 (DOUCONT-R3).

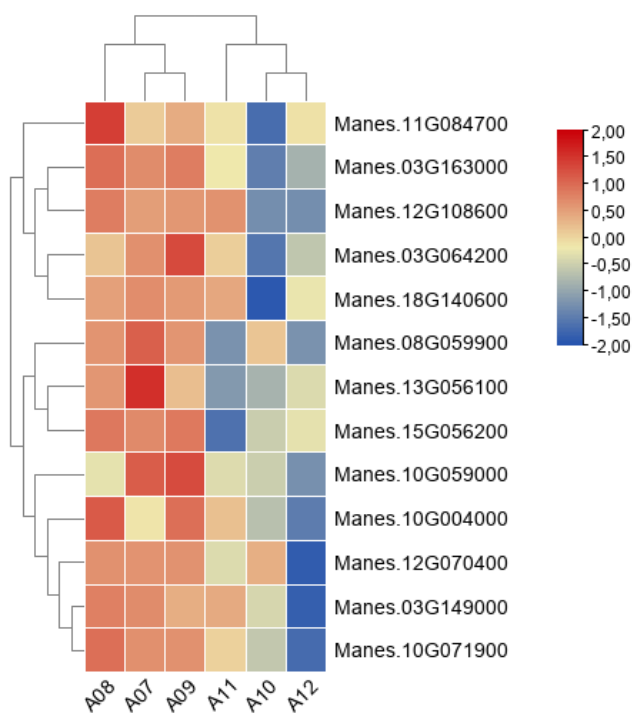


Tabela suplementar 1. Genes diferencialmente metilados. Colunas indicam, respectivamente: contraste estudado (G01 ou G04); sítio de metilação (CCGG ou CCNGG), metilação diferencialmente detectada (Hipo ou Hiper); identificador do gene, e anotação baseada no genoma v.6.1 da base de dados Phytozome.

Contraste	Sítio	Tipo de metilação	Identificador	Anotação v6.1
G01 - Formosa	CCGG	Hipo	Manes.01G073100	DHHC-type zinc finger family protein
G01 - Formosa	CCGG	Hipo	Manes.01G242800	P-loop containing nucleoside triphosphate hydrolases superfamily protein
G01 - Formosa	CCGG	Hipo	Manes.02G033400	cullin 1
G01 - Formosa	CCGG	Hipo	Manes.02G227200	cinnamate-4-hydroxylase
G01 - Formosa	CCGG	Hipo	Manes.03G041700	inter-alpha-trypsin inhibitor heavy chain-related helicases;ATP-dependent helicases;nucleic acid binding;ATP binding;DNA-directed DNA polymerases;DNA binding
G01 - Formosa	CCGG	Hipo	Manes.03G066200	
G01 - Formosa	CCGG	Hipo	Manes.03G200200	Pectin lyase-like superfamily protein
G01 - Formosa	CCGG	Hipo	Manes.04G009400	cellulose synthase-like A2
G01 - Formosa	CCGG	Hipo	Manes.05G031100	
G01 - Formosa	CCGG	Hipo	Manes.06G156800	manganese tracking factor for mitochondrial SOD2
G01 - Formosa	CCGG	Hipo	Manes.07G102300	auxin response factor 6
G01 - Formosa	CCGG	Hipo	Manes.08G157100	Major facilitator superfamily protein
G01 - Formosa	CCGG	Hipo	Manes.09G050300	NAD(P)-binding Rossmann-fold superfamily protein
G01 - Formosa	CCGG	Hipo	Manes.09G153700	ATPase E1-E2 type family protein / haloacid dehalogenase-like hydrolase family protein
G01 - Formosa	CCGG	Hipo	Manes.10G025000	Glucose-1-phosphate adenyltransferase family protein
G01 - Formosa	CCGG	Hipo	Manes.10G063700	P-loop containing nucleoside triphosphate hydrolases superfamily protein
G01 - Formosa	CCGG	Hipo	Manes.14G097400	NAC (No Apical Meristem) domain transcriptional regulator superfamily protein
G01 - Formosa	CCGG	Hipo	Manes.15G054200	Carbohydrate-binding X8 domain superfamily protein
G01 - Formosa	CCGG	Hipo	Manes.15G054300	purple acid phosphatase 15
G01 - Formosa	CCGG	Hipo	Manes.17G087400	hydroxyproline-rich glycoprotein family protein
G01 - Formosa	CCGG	Hipo	Manes.18G115800	O-acyltransferase (WSD1-like) family protein
G01 - Formosa	CCGG	Hiper	Manes.01G040400	ARM repeat superfamily protein
G01 - Formosa	CCGG	Hiper	Manes.01G129200	actin-related protein C1A
G01 - Formosa	CCGG	Hiper	Manes.01G164300	FTSH protease 9

G01 - Formosa	CCGG	Hiper	Manes.01G181600	3'-5'-exoribonuclease family protein
G01 - Formosa	CCGG	Hiper	Manes.01G265800	N-terminal nucleophile aminohydrolases (Ntn hydrolases) superfamily protein
G01 - Formosa	CCGG	Hiper	Manes.02G091300	phosphoenolpyruvate carboxylase 1
G01 - Formosa	CCGG	Hiper	Manes.03G010700	
G01 - Formosa	CCGG	Hiper	Manes.03G029000	Insulinase (Peptidase family M16) protein
G01 - Formosa	CCGG	Hiper	Manes.03G050400	Protein kinase protein with adenine nucleotide alpha hydrolases-like domain
G01 - Formosa	CCGG	Hiper	Manes.03G062700	Peroxisomal membrane 22 kDa (Mpv17/PMP22) family protein
G01 - Formosa	CCGG	Hiper	Manes.03G066300	BURP domain-containing protein
G01 - Formosa	CCGG	Hiper	Manes.03G083600	ubiquitin-specific protease 25
G01 - Formosa	CCGG	Hiper	Manes.03G103900	homogentisate 1,2-dioxygenase
G01 - Formosa	CCGG	Hiper	Manes.03G119900	
G01 - Formosa	CCGG	Hiper	Manes.03G120000	F-box/RNI-like superfamily protein
G01 - Formosa	CCGG	Hiper	Manes.04G000700	nitrilase 4
G01 - Formosa	CCGG	Hiper	Manes.04G014200	
G01 - Formosa	CCGG	Hiper	Manes.04G014300	serine carboxypeptidase-like 48
G01 - Formosa	CCGG	Hiper	Manes.05G134400	Zn-dependent exopeptidases superfamily protein
G01 - Formosa	CCGG	Hiper	Manes.05G184900	KNOTTED-like from Arabidopsis thaliana
G01 - Formosa	CCGG	Hiper	Manes.06G012000	Tetratricopeptide repeat (TPR)-like superfamily protein
G01 - Formosa	CCGG	Hiper	Manes.07G132800	Leucine-rich repeat protein kinase family protein
G01 - Formosa	CCGG	Hiper	Manes.09G039500	DNAse I-like superfamily protein
G01 - Formosa	CCGG	Hiper	Manes.09G043400	
G01 - Formosa	CCGG	Hiper	Manes.09G110600	ferritin 4
G01 - Formosa	CCGG	Hiper	Manes.10G024200	
G01 - Formosa	CCGG	Hiper	Manes.10G067400	Family of unknown function (DUF572)
G01 - Formosa	CCGG	Hiper	Manes.12G014500	Protein kinase superfamily protein
G01 - Formosa	CCGG	Hiper	Manes.12G033800	P-loop containing nucleoside triphosphate hydrolases superfamily protein
G01 - Formosa	CCGG	Hiper	Manes.12G110200	Protein kinase superfamily protein
G01 - Formosa	CCGG	Hiper	Manes.12G156000	photosystem I reaction center subunit PSI-N, chloroplast, putative / PSI-N, putative (PSAN)
G01 - Formosa	CCGG	Hiper	Manes.13G111500	Arabinanase/levansucrase/invertase

G01 - Formosa	CCGG	Hiper	Manes.13G112600	formin homolog 6
G01 - Formosa	CCGG	Hiper	Manes.13G148900	ELMO/CED-12 family protein
G01 - Formosa	CCGG	Hiper	Manes.13G155300	methyltransferase 1
G01 - Formosa	CCGG	Hiper	Manes.15G159500	uridine kinase/uracil phosphoribosyltransferase 1
G01 - Formosa	CCGG	Hiper	Manes.16G051600	Pectin lyase-like superfamily protein
G01 - Formosa	CCGG	Hiper	Manes.16G060700	TBP-associated factor 11
G01 - Formosa	CCGG	Hiper	Manes.16G092800	alpha/beta-Hydrolases superfamily protein
G01 - Formosa	CCGG	Hiper	Manes.16G098500	Myosin family protein with Dil domain
G01 - Formosa	CCGG	Hiper	Manes.16G110100	golgin candidate 1
G01 - Formosa	CCGG	Hiper	Manes.18G026700	HAD-superfamily hydrolase, subfamily IG, 5'-nucleotidase
G01 - Formosa	CCNGG	Hipo	Manes.01G115300	AAR2 protein family
G01 - Formosa	CCNGG	Hipo	Manes.02G158800	
G01 - Formosa	CCNGG	Hipo	Manes.10G035100	F-box family protein
G01 - Formosa	CCNGG	Hipo	Manes.10G069300	WWE protein-protein interaction domain protein family
G01 - Formosa	CCNGG	Hipo	Manes.16G068800	SWIM zinc finger family protein
G01 - Formosa	CCNGG	Hiper	Manes.02G059500	RING/U-box superfamily protein
G01 - Formosa	CCNGG	Hiper	Manes.03G027200	ARM repeat superfamily protein
G01 - Formosa	CCNGG	Hiper	Manes.06G176300	FTSH protease 12
G01 - Formosa	CCNGG	Hiper	Manes.09G076300	Protein of unknown function (DUF1162)
G01 - Formosa	CCNGG	Hiper	Manes.09G162500	Major facilitator superfamily protein
G01 - Formosa	CCNGG	Hiper	Manes.12G070200	Nodulin MtN21 /EamA-like transporter family protein
G01 - Formosa	CCNGG	Hiper	Manes.15G190300	beta-amylase 2
G01 - Formosa	CCNGG	Hiper	Manes.16G013200	
G01 - Formosa	CCNGG	Hiper	Manes.18G140900	
G04 - Dourada	CCGG	Hipo	Manes.01G020600	MUTS homolog 2
G04 - Dourada	CCGG	Hipo	Manes.01G037400	Major facilitator superfamily protein
G04 - Dourada	CCGG	Hipo	Manes.01G131400	Reticulon family protein
G04 - Dourada	CCGG	Hipo	Manes.01G211300	endoplasmic reticulum auxin binding protein 1
G04 - Dourada	CCGG	Hipo	Manes.02G005700	general control non-repressible 3

G04 - Dourada	CCGG	Hipo	Manes.02G085400	galacturonosyltransferase 15
G04 - Dourada	CCGG	Hipo	Manes.02G197800	
G04 - Dourada	CCGG	Hipo	Manes.02G207300	Protein kinase family protein with ARM repeat domain
G04 - Dourada	CCGG	Hipo	Manes.02G226900	Transcriptional factor B3 family protein / auxin-responsive factor AUX/IAA-related
G04 - Dourada	CCGG	Hipo	Manes.03G134200	Ribosomal protein S8e family protein
G04 - Dourada	CCGG	Hipo	Manes.05G061900	photolyase/blue-light receptor 2
G04 - Dourada	CCGG	Hipo	Manes.05G087700	poly(ADP-ribose) polymerase 2
G04 - Dourada	CCGG	Hipo	Manes.05G202300	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family
G04 - Dourada	CCGG	Hipo	Manes.06G078800	DNA glycosylase superfamily protein
G04 - Dourada	CCGG	Hipo	Manes.07G028400	dynammin-like 3
G04 - Dourada	CCGG	Hipo	Manes.08G017000	Eukaryotic aspartyl protease family protein
G04 - Dourada	CCGG	Hipo	Manes.08G161700	hydroxyproline-rich glycoprotein family protein
G04 - Dourada	CCGG	Hipo	Manes.09G167100	Domain of unknown function (DUF1995)
G04 - Dourada	CCGG	Hipo	Manes.10G055800	lysyl-tRNA synthetase 1
G04 - Dourada	CCGG	Hipo	Manes.12G009600	Calmodulin-binding transcription activator protein with CG-1 and Ankyrin domains
G04 - Dourada	CCGG	Hipo	Manes.12G075100	glutamate receptor 5
G04 - Dourada	CCGG	Hipo	Manes.13G106000	peroxisomal NAD-malate dehydrogenase 1
G04 - Dourada	CCGG	Hipo	Manes.13G145000	switch 2
G04 - Dourada	CCGG	Hipo	Manes.14G074700	CLP protease proteolytic subunit 6
G04 - Dourada	CCGG	Hipo	Manes.14G141800	response regulator 12
G04 - Dourada	CCGG	Hipo	Manes.15G015800	chromatin-remodeling protein 11
G04 - Dourada	CCGG	Hipo	Manes.15G095300	tubby like protein 7
G04 - Dourada	CCGG	Hipo	Manes.15G099900	MAP kinase 19
G04 - Dourada	CCGG	Hipo	Manes.16G043500	BES1/BZR1 homolog 4
G04 - Dourada	CCGG	Hipo	Manes.16G056900	HISTIDINE TRIAD NUCLEOTIDE-BINDING 2
G04 - Dourada	CCGG	Hipo	Manes.16G060600	Arp2/3 complex, 34 kD subunit p34-Arc
G04 - Dourada	CCGG	Hipo	Manes.17G037500	DERLIN-2.2
G04 - Dourada	CCGG	Hipo	Manes.18G023900	Major facilitator superfamily protein
G04 - Dourada	CCGG	Hipo	Manes.18G100900	SNARE associated Golgi protein family

G04 - Dourada	CCGG	Hiper	Manes.01G157300	RNA polymerase I specific transcription initiation factor RRN3 protein
G04 - Dourada	CCGG	Hiper	Manes.01G169300	splicing factor PWI domain-containing protein
G04 - Dourada	CCGG	Hiper	Manes.01G181600	3'-5'-exoribonuclease family protein
G04 - Dourada	CCGG	Hiper	Manes.02G003700	Metal-dependent protein hydrolase
G04 - Dourada	CCGG	Hiper	Manes.02G025500	P-loop containing nucleoside triphosphate hydrolases superfamily protein
G04 - Dourada	CCGG	Hiper	Manes.02G056600	Proteasome component (PCI) domain protein
G04 - Dourada	CCGG	Hiper	Manes.02G067700	Regulator of Vps4 activity in the MVB pathway protein
G04 - Dourada	CCGG	Hiper	Manes.02G091000	general regulatory factor 9
G04 - Dourada	CCGG	Hiper	Manes.02G135100	dgd1 suppressor 1
G04 - Dourada	CCGG	Hiper	Manes.02G191800	DNAJ heat shock N-terminal domain-containing protein
G04 - Dourada	CCGG	Hiper	Manes.02G210800	Kinase interacting (KIP1-like) family protein
G04 - Dourada	CCGG	Hiper	Manes.03G010700	
G04 - Dourada	CCGG	Hiper	Manes.03G049000	FTSH protease 4
G04 - Dourada	CCGG	Hiper	Manes.03G190300	DNAJ heat shock N-terminal domain-containing protein
G04 - Dourada	CCGG	Hiper	Manes.04G032200	RNA-binding (RRM/RBD/RNP motifs) family protein
G04 - Dourada	CCGG	Hiper	Manes.04G071100	Protein of unknown function (DUF616)
G04 - Dourada	CCGG	Hiper	Manes.04G143700	TRF-like 6
G04 - Dourada	CCGG	Hiper	Manes.04G152900	SU(VAR)3-9 homolog 3
G04 - Dourada	CCGG	Hiper	Manes.05G118800	CLAVATA3/ESR-RELATED 13
G04 - Dourada	CCGG	Hiper	Manes.05G206400	ubiquitin family protein
G04 - Dourada	CCGG	Hiper	Manes.06G008000	MAP kinase 9
G04 - Dourada	CCGG	Hiper	Manes.06G013000	protein kinase family protein / C-type lectin domain-containing protein
G04 - Dourada	CCGG	Hiper	Manes.06G049500	Galactose oxidase/kelch repeat superfamily protein
G04 - Dourada	CCGG	Hiper	Manes.06G051100	3-ketoacyl-CoA synthase 11
G04 - Dourada	CCGG	Hiper	Manes.06G052400	Protein of unknown function (DUF3511)
G04 - Dourada	CCGG	Hiper	Manes.06G114700	RING/U-box superfamily protein
G04 - Dourada	CCGG	Hiper	Manes.06G151900	Protein kinase superfamily protein
G04 - Dourada	CCGG	Hiper	Manes.08G020300	Zinc finger C-x8-C-x5-C-x3-H type family protein
G04 - Dourada	CCGG	Hiper	Manes.08G132300	Pseudouridine synthase family protein

G04 - Dourada	CCGG	Hiper	Manes.09G038200	P-loop containing nucleoside triphosphate hydrolases superfamily protein
G04 - Dourada	CCGG	Hiper	Manes.09G098300	proteasome inhibitor-related
G04 - Dourada	CCGG	Hiper	Manes.09G110300	
G04 - Dourada	CCGG	Hiper	Manes.10G034400	2-oxoglutarate dehydrogenase, E1 component
G04 - Dourada	CCGG	Hiper	Manes.10G123600	cleavage and polyadenylation specificity factor 73-l
G04 - Dourada	CCGG	Hiper	Manes.10G149600	beta glucosidase 13
G04 - Dourada	CCGG	Hiper	Manes.11G010600	Calcineurin-like metallo-phosphoesterase superfamily protein
G04 - Dourada	CCGG	Hiper	Manes.11G076800	
G04 - Dourada	CCGG	Hiper	Manes.11G109700	DHHC-type zinc finger family protein
G04 - Dourada	CCGG	Hiper	Manes.11G140100	vacuolar proton ATPase A1
G04 - Dourada	CCGG	Hiper	Manes.12G035400	breast cancer associated RING 1
G04 - Dourada	CCGG	Hiper	Manes.12G035500	cellulose synthase A4
G04 - Dourada	CCGG	Hiper	Manes.12G040100	Protein of unknown function (DUF3741)
G04 - Dourada	CCGG	Hiper	Manes.12G051300	plant U-box 9
G04 - Dourada	CCGG	Hiper	Manes.12G136200	Thioredoxin superfamily protein
G04 - Dourada	CCGG	Hiper	Manes.13G036700	phytochelatin synthase 1 (PCS1)
G04 - Dourada	CCGG	Hiper	Manes.13G137900	Adaptin ear-binding coat-associated protein 1 NECAP-1
G04 - Dourada	CCGG	Hiper	Manes.14G063800	embryo defective 2759
G04 - Dourada	CCGG	Hiper	Manes.15G090400	RING/U-box superfamily protein
G04 - Dourada	CCGG	Hiper	Manes.15G158800	
G04 - Dourada	CCGG	Hiper	Manes.16G003300	ERD (early-responsive to dehydration stress) family protein
G04 - Dourada	CCGG	Hiper	Manes.16G128500	
G04 - Dourada	CCGG	Hiper	Manes.17G068400	Minichromosome maintenance (MCM2/3/5) family protein
G04 - Dourada	CCGG	Hiper	Manes.18G026700	HAD-superfamily hydrolase, subfamily IG, 5'-nucleotidase
G04 - Dourada	CCGG	Hiper	Manes.18G035000	Major facilitator superfamily protein
G04 - Dourada	CCGG	Hiper	Manes.18G037900	abscisic acid responsive elements-binding factor 2
G04 - Dourada	CCNGG	Hiper	Manes.18G088800	Thioredoxin superfamily protein
G04 - Dourada	CCNGG	Hipo	Manes.01G182900	glycine-rich protein
G04 - Dourada	CCNGG	Hipo	Manes.02G059500	RING/U-box superfamily protein

G04 - Dourada	CCNGG	Hipo	Manes.03G005100	cyclic nucleotide-gated channel 14
G04 - Dourada	CCNGG	Hipo	Manes.03G061300	
G04 - Dourada	CCNGG	Hipo	Manes.05G026700	G-box binding factor 3
G04 - Dourada	CCNGG	Hipo	Manes.05G099700	Ribosomal protein S4
G04 - Dourada	CCNGG	Hipo	Manes.08G028700	Helicase protein with RING/U-box domain
G04 - Dourada	CCNGG	Hipo	Manes.14G152900	VH1-interacting kinase
G04 - Dourada	CCNGG	Hipo	Manes.17G118400	
G04 - Dourada	CCNGG	Hipo	Manes.18G067500	
G04 - Dourada	CCNGG	Hiper	Manes.03G064200	
G04 - Dourada	CCNGG	Hiper	Manes.03G149000	putative type 1 membrane protein
G04 - Dourada	CCNGG	Hiper	Manes.03G163000	COP1-interacting protein-related
G04 - Dourada	CCNGG	Hiper	Manes.08G059900	Transmembrane amino acid transporter family protein
G04 - Dourada	CCNGG	Hiper	Manes.10G004000	
G04 - Dourada	CCNGG	Hiper	Manes.10G059000	beta-hydroxyisobutyryl-CoA hydrolase 1
G04 - Dourada	CCNGG	Hiper	Manes.10G071900	fucosyltransferase 12
G04 - Dourada	CCNGG	Hiper	Manes.11G084700	Homeodomain-like superfamily protein
G04 - Dourada	CCNGG	Hiper	Manes.12G070400	dehydroquinase dehydratase, putative / shikimate dehydrogenase, putative
G04 - Dourada	CCNGG	Hiper	Manes.12G108600	heavy metal atpase 1
G04 - Dourada	CCNGG	Hiper	Manes.13G056100	
G04 - Dourada	CCNGG	Hiper	Manes.15G056200	ARM repeat superfamily protein
G04 - Dourada	CCNGG	Hiper	Manes.18G140600	

Tabela suplementar 2A. Termos GO enriquecidos a partir de genes Hipo e Hiper metilados (separadamente) para Formosa (G01). Usado como parâmetro p-value < 0,01. As colunas indicam, respectivamente: Id do termo GO; tipo de metilação do gene; descrição do termo GO; quantidade de anotações para o termo no banco de dados; quantidade de termos GO encontrados entre os genes hipometilados ou entre os genes hipermetilados inseridos na análise; parâmetros de confiança da análise (expected, p-value, q-value); aspecto do termo (P: processo biológico; C: componente celular; F: função molecular); e genes inseridos na análise que possuem o termo GO.

GO.ID	Metilação	Termo	Annotados	Quantidade	Expected	p-value	q-value	Aspecto	Genes
GO:0010193	Hipo	response to ozone regulation of cell cycle process	20	2	0.02	0.00013	6,19E+02	P	Manes.10G025000, Manes.10G069300
GO:0010564	Hipo	regulation of cell cycle process	39	2	0.03	0.00051	1,00E+03	P	Manes.03G066200, Manes.14G097400
GO:0051726	Hipo	regulation of cell cycle response to reactive oxygen species	83	2	0.07	0.00228	1,00E+03	P	Manes.03G066200, Manes.14G097400
GO:0000302	Hipo	response to jasmonic acid	119	2	0.1	0.00461	1,00E+03	P	Manes.10G025000, Manes.10G069300
GO:0051301	Hipo	cell division nucleotidyltransferase activity	147	2	0.13	0.00696	1,00E+03	P	Manes.03G066200, Manes.14G097400
GO:0016779	Hipo	activity	149	2	0.14	0.00848	1,00E+00	F	Manes.03G066200, Manes.10G025000, Manes.01G164300, Manes.01G265800, Manes.03G029000, Manes.03G083600, Manes.04G014300, Manes.05G134400, Manes.06G176300
GO:0006508	Hiper	proteolysis	604	7	1.18	0.00012	5,71E+02	P	Manes.02G091300, Manes.05G184900
GO:0045165	Hiper	cell fate commitment	42	2	0.08	0.00303	1,00E+03	P	Manes.01G129200, Manes.16G098500
GO:0015629	Hiper	actin cytoskeleton	31	2	0.07	0.0025	1,00E+00	C	Manes.01G129200, Manes.13G112600, Manes.16G098500
GO:0044430	Hiper	cytoskeletal part	148	3	0.36	0.0051	1,00E+00	C	Manes.01G129200, Manes.13G112600, Manes.16G098500
GO:0005856	Hiper	cytoskeleton	154	3	0.37	0.0057	1,00E+00	C	Manes.01G129200, Manes.13G112600, Manes.16G098500, Manes.01G164300, Manes.03G029000, Manes.05G134400, Manes.06G176300
GO:0008237	Hiper	metallopeptidase activity	84	4	0.16	1.8e-05	4,19E+01	F	Manes.01G164300, Manes.03G029000, Manes.06G176300
GO:0004222	Hiper	metalloendopeptidase activity	63	3	0.12	0.00022	1,86E+02	F	Manes.01G164300, Manes.03G029000, Manes.06G176300
GO:0070011	Hiper	peptidase activity, acting on L-amino acid peptides	487	6	0.92	0.00026	1,86E+02	F	Manes.01G164300, Manes.01G265800, Manes.03G029000, Manes.04G014300, Manes.05G134400, Manes.06G176300
GO:0008233	Hiper	peptidase activity	505	6	0.95	0.00032	1,86E+02	F	Manes.01G164300, Manes.01G181600, Manes.01G265800, Manes.03G029000, Manes.03G083600, Manes.04G000700, Manes.04G014300, Manes.05G134400, Manes.06G176300, Manes.12G033800, Manes.13G111500, Manes.15G190300, Manes.16G098500
GO:0016787	Hiper	hydrolase activity	2610	13	4.91	0.00055	2,56E+02	F	Manes.01G164300, Manes.01G265800, Manes.03G029000, Manes.06G176300
GO:0004175	Hiper	endopeptidase activity	299	4	0.56	0.00231	3,84E+02	F	Manes.04G014300, Manes.05G134400
GO:0004180	Hiper	carboxypeptidase activity	53	2	0.1	0.00446	6,49E+02	F	Manes.04G014300, Manes.05G134400
GO:0008238	Hiper	exopeptidase activity	79	2	0.15	0.00968	1,00E+03	F	Manes.04G014300, Manes.05G134400

Tabela suplementar 2B. Termos GO enriquecidos a partir de genes Hipo e Hiper metilados (separadamente) para Dourada (G04). Usado como parâmetro p-value < 0,01. As colunas indicam respectivamente: Id do termo GO; tipo de metilação do gene; descrição do termo GO; quantidade de anotações para o termo no banco de dados; quantidade de termos GO encontrados entre os genes hipometilados ou entre os genes hipermetilados inseridos na análise; parâmetros de confiança da análise (expected, p-value, q-value); aspecto do termo (P: processo biológico; C: componente celular; F: função molecular); e genes inseridos na análise que possuem o termo GO.

GO.ID	Metilação	Termo	Annotados	Quantidade	Expected	p-value	q-value	Aspecto	Genes
GO:0010564	Hipo	regulation of cell cycle process	39	2	0.07	0.0019	9.52e-01	P	Manes.01G020600, Manes.01G211300
GO:0032535	Hipo	regulation of cellular component size	42	2	0.07	0.0022	9.52e-01	P	Manes.01G211300, Manes.16G060600
GO:0090066	Hipo	regulation of anatomical structure size	42	2	0.07	0.0022	9.52e-01	P	Manes.01G211300, Manes.16G060600
GO:0051302	Hipo	regulation of cell division	47	2	0.08	0.0028	9.52e-01	P	Manes.01G020600, Manes.01G211300
GO:0051052	Hipo	regulation of DNA metabolic process	51	2	0.09	0.0032	9.52e-01	P	Manes.01G020600, Manes.01G211300
GO:0006259	Hipo	DNA metabolic process	383	4	0.64	0.0034	9.52e-01	P	Manes.01G020600, Manes.01G211300, Manes.05G087700, Manes.06G078800
GO:0006281	Hipo	DNA repair	207	3	0.35	0.0047	1.00e+00	P	Manes.01G020600, Manes.05G087700, Manes.06G078800
GO:0006974	Hipo	cellular response to DNA damage stimulus	223	3	0.37	0.0058	1.00e+00	P	Manes.01G020600, Manes.05G087700, Manes.06G078800
GO:0051726	Hipo	regulation of cell cycle	83	2	0.14	0.0084	1.00e+00	P	Manes.01G020600, Manes.01G211300
GO:0033043	Hipo	regulation of organelle organization	84	2	0.14	0.0086	1.00e+00	P	Manes.01G020600, Manes.16G060600
GO:0060089	Hipo	molecular transducer activity	121	3	0.19	0.0009	1.00E+0	F	Manes.01G211300, Manes.12G075100, Manes.15G099900
GO:0022836	Hipo	gated channel activity	56	2	0.09	0.0036	1.00E+0	F	Manes.03G005100, Manes.12G075100
GO:0004872	Hipo	receptor activity	66	2	0.11	0.0050	1.00E+0	F	Manes.01G211300, Manes.12G075100
GO:0005216	Hipo	ion channel activity	83	2	0.13	0.0078	1.00E+0	F	Manes.03G005100, Manes.12G075100
GO:0022838	Hipo	substrate-specific channel activity	92	2	0.15	0.0095	1.00E+0	F	Manes.03G005100, Manes.12G075100
GO:0003677	Hipo	DNA binding	1945	8	3.12	0.0096	1.00E+0	F	Manes.01G020600, Manes.02G226900, Manes.05G026700, Manes.05G087700, Manes.12G009600, Manes.14G141800, Manes.15G015800, Manes.15G095300

GO:0015267	Hipo	channel activity	93	2	0.15	0.0097	1,00E+0		
		passive transmembrane				3	0	F	Manes.03G005100, Manes.12G075100
GO:0022803	Hipo	transporter activity	93	2	0.15	0.0097	1,00E+0	F	Manes.03G005100, Manes.12G075100
						3	0		Manes.04G071100, Manes.10G071900, Manes.11G010600,
GO:0005802	Hiper	trans-Golgi network	179	4	0.7	0.0051	1,00E+0	C	Manes.11G140100
							0		Manes.04G071100, Manes.10G071900, Manes.11G010600,
GO:0098791	Hiper	Golgi subcompartment	188	4	0.74	0.0061	1,00E+0	C	Manes.11G140100
							0		Manes.04G071100, Manes.10G071900, Manes.11G010600,
GO:0005768	Hiper	endosome	209	4	0.82	0.0088	1,00E+0	C	Manes.11G140100
							0		

Tabela suplementar 3A. Termos de GO para processos biológicos para proteínas da rede de interações formada por sítios diferencialmente hipometilados para a variedade tolerante Formosa baseados em proteínas de *A. thaliana* órtologas às proteínas de *M. esculenta*. As colunas indicam respectivamente: Id do termo GO; descrição do termo; contagem do termo na rede de interações; background de informação para este termo no banco de dados; nível de confiabilidade; taxa de falso descobrimento; nome das proteínas que contém o termo GO

ID do termo	Descrição	Contagem de gene	Background	Strength	False discovery rate	Matching proteins nas redes
GO:0000338	protein deneddylation	4	10	2.29	0.00026	FUS5,CSN5A,FUS12,FUS6
GO:0007275	multicellular organism development	19	2260	0.61	0.00026	FUS5,MEE5,CSN5A,Manes.07G102300,UFO,Manes.10G069300,SKP1,SKIP,CAND1,FUS12,Manes.02G227200,Manes.15G054300,Manes.01G073100,FUS6,TIR1,Manes.02G033400,Manes.03G066200,Manes.04G009400,IAA28
GO:0008380	RNA splicing	8	211	1.27	0.00026	MAC3A,MEE5,CDC5,ESP3,SKIP,PRP8A,AT4G38780,AT5G25230
GO:0009733	response to auxin	9	324	1.13	0.00026	SHY2,CSN5A,Manes.07G102300,SKP1,CAND1,IAA8,TIR1,Manes.02G033400,IAA28
GO:0070647	protein modification by small protein conjugation or removal	12	701	0.92	0.00026	FUS5,MAC3A,CSN5A,UFO,SKP1,CAND1,EBF1,FUS12,FUS6,TIR1,Manes.02G033400,RBX1
GO:0000398	mRNA splicing, via spliceosome	6	113	1.41	0.00043	MAC3A,MEE5,SKIP,PRP8A,AT4G38780,AT5G25230
GO:0009734	auxin-activated signaling pathway	7	193	1.25	0.00043	SHY2,Manes.07G102300,SKP1,IAA8,TIR1,Manes.02G033400,IAA28

GO:0009791	post-embryonic development	14	1237	0.74	0.00043	MEE5,CSN5A,Manes.07G102300,UFO,Manes.10G069300,SKIP,CAND1,Manes.15G054300,FUS6,TIR1,Manes.02G033400,Manes.03G066200,Manes.04G009400,IAA28
GO:0006397	mRNA processing	8	320	1.09	0.00053	MAC3A,MEE5,CDC5,ESP3,SKIP,PRP8A,AT4G38780,AT5G25230
GO:0009725	response to hormone	13	1329	0.68	0.0023	SHY2,CSN5A,Manes.07G102300,Manes.10G069300,SKIP,SKIP,CAND1,IAA8,EBF1,Manes.10G025000,TIR1,Manes.02G033400,IAA28
GO:0007165	signal transduction	14	1610	0.63	0.0031	FUS5,SHY2,CDC5,CSN5A,Manes.07G102300,Manes.10G069300,SKIP,IAA8,EBF1,FUS12,FUS6,TIR1,Manes.02G033400,IAA28
GO:0009585	red, far-red light phototransduction	4	53	1.57	0.0047	FUS5,CSN5A,FUS12,FUS6
GO:0048608	reproductive structure development	11	1011	0.73	0.0047	MEE5,CSN5A,Manes.07G102300,UFO,Manes.10G069300,SKIP,CAND1,FUS6,TIR1,Manes.02G033400,Manes.04G009400
GO:0010033	response to organic substance	14	1719	0.6	0.0048	SHY2,CSN5A,Manes.07G102300,Manes.10G069300,SKIP,SKIP,CAND1,IAA8,EBF1,Manes.10G025000,TIR1,Manes.02G033400,Manes.06G156800,IAA28
GO:0048731	system development	13	1578	0.61	0.0083	MEE5,CSN5A,Manes.07G102300,UFO,Manes.10G069300,SKIP,CAND1,Manes.01G073100,FUS6,TIR1,Manes.02G033400,Manes.04G009400,IAA28
GO:0006508	proteolysis	9	745	0.77	0.0105	FUS5,CSN5A,UFO,SKIP,EBF1,FUS12,FUS6,Manes.02G033400,RBX1
GO:0022414	reproductive process	12	1429	0.61	0.0129	MEE5,CSN5A,Manes.07G102300,UFO,Manes.10G069300,SKIP,CAND1,Manes.15G054300,FUS6,TIR1,Manes.02G033400,Manes.04G009400
GO:0016567	protein ubiquitination	8	611	0.81	0.0156	MAC3A,UFO,SKIP,CAND1,EBF1,TIR1,Manes.02G033400,RBX1
GO:0009755	hormone-mediated signaling pathway	9	803	0.74	0.0157	SHY2,Manes.07G102300,Manes.10G069300,SKIP,IAA8,EBF1,TIR1,Manes.02G033400,IAA28
GO:0009639	response to red or far red light	5	189	1.11	0.0198	FUS5,CSN5A,FUS12,FUS6,Manes.03G066200
GO:0009873	ethylene-activated signaling pathway	5	191	1.11	0.0200	Manes.10G069300,SKIP,EBF1,TIR1,Manes.02G033400
GO:0051716	cellular response to stimulus	16	2640	0.47	0.0210	FUS5,SHY2,MAC3A,CDC5,CSN5A,Manes.14G097400,Manes.07G102300,Manes.10G069300,SKIP,IAA8,EBF1,FUS12,FUS6,TIR1,Manes.02G033400,IAA28
GO:0000244	spliceosomal tri-snRNP complex assembly	2	4	2.39	0.0225	PRP8A,AT4G38780
GO:0090697	post-embryonic plant organ morphogenesis	4	101	1.29	0.0227	CSN5A,Manes.10G069300,TIR1,IAA28
GO:0043170	macromolecule metabolic process	22	4720	0.36	0.0271	FUS5,MAC3A,MEE5,CDC5,CSN5A,UFO,ESP3,SKIP,PRP8A,CAND1,EBF1,FUS12,Manes.10G025000,Manes.01G073100,FUS6,TIR1,Manes.02G033400,Manes.03G066200,AT4G38780,RBX1,AT5G25230

GO:0050896	response to stimulus	23	5174	0.34	0.0351	FUS5,SHY2,MAC3A,MAC5A,CDC5,CSN5A,Manes.14G097400,Manes.07G102300,Manes.10G069300,SKP1,SKIP,CAND1,IAA8,EBF1,FUS12,Manes.02G227200,Manes.10G025000,FUS6,TIR1,Manes.02G033400,Manes.06G156800,Manes.03G066200,IAA28
GO:0010387	COP9 signalosome assembly	2	6	2.21	0.0375	FUS5,CSN5A
GO:0050794	regulation of cellular process	20	4214	0.37	0.0479	FUS5,SHY2,MEE5,CDC5,CSN5A,Manes.14G097400,Manes.07G102300,UFO,Manes.10G069300,SKP1,SKIP,IAA8,EBF1,FUS12,Manes.01G073100,FUS6,TIR1,Manes.02G033400,Manes.03G066200,IAA28

Tabela suplementar 3B. Termos de GO para função molecular para proteínas da rede de interações formada por sítios diferencialmente hipometilados para a variedade tolerante Formosa baseados em proteínas de *A. thaliana* órtologas às proteínas de *M. esculenta*. As colunas indicam respectivamente: Id do termo GO; descrição do termo; contagem do termo na rede de interações; background de informação para este termo no banco de dados; nível de confiabilidade; taxa de falso descobrimento; nome das proteínas que contém o termo GO

ID do termo	Descrição	Contagem de gene	Background	Strength	False discovery rate	Matching proteins nas redes
GO:0017069	snRNA binding	5	16	2.18	3.04e-06	MEE5,MAC5A,PRP8A,AT4G38780,AT5G25230
GO:0030623	U5 snRNA binding	4	4	2.69	3.04e-06	MEE5,PRP8A,AT4G38780,AT5G25230
GO:0036002	pre-mRNA binding	4	16	2.09	0.00013	MAC5A,SKIP,PRP8A,AT4G38780
GO:0017070	U6 snRNA binding	3	7	2.32	0.0012	MAC5A,PRP8A,AT4G38780
GO:0030620	U2 snRNA binding	2	3	2.51	0.0425	PRP8A,AT4G38780
GO:0097157	pre-mRNA intronic binding	2	3	2.51	0.0425	PRP8A,AT4G38780

Tabela suplementar 3C. Termos de GO para componente celular para proteínas da rede de interações formada por sítios diferencialmente hipometilados para a variedade tolerante Formosa baseados em proteínas de *A. thaliana* órtologas às proteínas de *M. esculenta*. As colunas indicam respectivamente: Id do termo GO; descrição do termo; contagem do termo na rede de interações;

background de informação para este termo no banco de dados; nível de confiabilidade; taxa de falso descobrimento; identificador das proteínas que contem o termo GO; nome das proteínas que contém o termo GO

ID do termo	Descrição	Contagem de gene	Background	Strength	False discovery rate	Matching proteins nas redes
GO:0005681	spliceosomal complex	9	112	1.59	7.22e-09	MAC3A,MEE5,MAC5A,CDC5,ESP3,SKIP,PRP8A,AT4G38780,AT5G25230
GO:0071013	catalytic step 2 spliceosome	6	28	2.02	8.61e-08	MAC3A,MEE5,MAC5A,PRP8A,AT4G38780,AT5G25230
GO:0005634	nucleus	28	4669	0.47	1.25e-05	FUS5,SHY2,MAC3A,MEE5,MAC5A,CDC5,CSN5A,Manes.14G097400,Manes.07G102300,UFO,Manes.10G069300,ESP3,SKP1,SKIP,PRP8A,CAND1,IAA8,EBF1,FUS12,Manes.10G025000,FUS6,TIR1,Manes.02G033400,AT4G38780,RBX1,AT5G25230,IAA28,AT5G65490
GO:0008180	COP9 signalosome	4	10	2.29	1.25e-05	FUS5,CSN5A,FUS12,FUS6
GO:0032991	protein-containing complex	18	2226	0.6	0.00012	FUS5,MAC3A,MEE5,MAC5A,CDC5,CSN5A,UFO,ESP3,SKIP,PRP8A,EBF1,FUS12,FUS6,TIR1,Manes.02G033400,AT4G38780,RBX1,AT5G25230
GO:0005684	U2-type spliceosomal complex	4	24	1.91	0.00013	MAC3A,MEE5,MAC5A,AT5G25230
GO:0071007	U2-type catalytic step 2 spliceosome	3	5	2.47	0.00016	MEE5,MAC5A,AT5G25230
GO:1902494	catalytic complex	11	776	0.84	0.00016	MAC3A,MEE5,MAC5A,UFO,PRP8A,EBF1,TIR1,Manes.02G033400,AT4G38780,RBX1,AT5G25230
GO:0097525	spliceosomal snRNP complex	4	36	1.74	0.00033	MEE5,PRP8A,AT4G38780,AT5G25230
GO:0000151	ubiquitin ligase complex	5	190	1.11	0.0087	UFO,EBF1,TIR1,Manes.02G033400,RBX1
GO:0043231	intracellular membrane-bounded organelle	34	9376	0.25	0.0087	FUS5,SHY2,MAC3A,MEE5,MAC5A,CDC5,CSN5A,Manes.14G097400,Manes.07G102300,UFO,Manes.10G069300,ESP3,SKP1,SKIP,PRP8A,CAND1,IAA8,EBF1,FUS12,Manes.02G227200,AT2G37585,Manes.10G025000,Manes.09G153700,Manes.01G073100,FUS6,TIR1,Manes.02G033400,Manes.06G156800,AT4G38780,RBX1,Manes.04G009400,AT5G25230,IAA28,AT5G65490
GO:0071006	U2-type catalytic step 1 spliceosome	2	5	2.29	0.0139	MAC3A,MAC5A
GO:0000974	Prp19 complex	2	8	2.09	0.0223	MAC3A,MAC5A
GO:0005682	U5 snRNP	2	8	2.09	0.0223	PRP8A,AT4G38780
GO:0031461	cullin-RING ubiquitin ligase complex	4	150	1.12	0.0311	EBF1,TIR1,Manes.02G033400,RBX1

GO:0005622	intracellular	35	10863	0.2	0.0490	FUS5,SHY2,MAC3A,MEE5,MAC5A,CDC5,CSN5A,Manes.14G097400,Manes.07G102300,UFO,Manes.10G069300,ESP3,SKP1,SKIP,PRP8A,CAND1,IAA8,EBF1,FUS12,Manes.02G227200,AT2G37585,Manes.10G025000,Manes.09G153700,Manes.01G073100,FUS6,TIR1,Manes.02G033400,Manes.06G156800,Manes.03G066200,AT4G38780,RBX1,Manes.04G009400,AT5G25230,IAA28,AT5G65490
GO:0046540	U4/U6 x U5 tri-snRNP complex	2	14	1.84	0.0492	MEE5,AT5G25230

Tabela suplementar 3D. Termos KEGG para proteínas da rede de interações oriundas de genes diferencialmente hipometilados na variedade tolerante Formosa, tendo por base proteínas de *A. thaliana* órtologas às proteínas de *M. esculenta*. As colunas indicam, respectivamente: Id do termo; descrição do termo; contagem do termo na rede de interações; background de informação para o termo no banco de dados; nível de confiabilidade; taxa de falsas descobertas; nome das proteínas que contém o termo.

ID do termo	Descrição	Contagem de gene	Background	Strength	False discovery rate	Matching proteins nas redes
ath03040	Spliceosome	10	188	1.42	3.65e-09	MAC3A,MEE5,MAC5A,CDC5,AT1G10580,ESP3,SKIP,PRP8A,AT4G38780,AT5G25230

Tabela suplementar 4A. Termos de GO para processos biológicos associados a proteínas da rede de interações, codificadas de genes diferencialmente hipermetilados na variedade tolerante Formosa, tendo por base proteínas de *A. thaliana* órtologas às proteínas de *M. esculenta*. As colunas indicam respectivamente: Id do termo GO; descrição do termo; contagem do termo na rede de interações; background de informação para este termo no banco de dados; nível de confiabilidade; taxa de falso descobrimento; nome das proteínas que contém o termo GO

ID do termo	Descrição	Contagem de gene	Background	Strength	False discovery rate	Matching proteins nas redes
GO:0006352	DNA-templated transcription, initiation	13	38	2.09	1.72e-18	TAFII59,EER4,TAF11b,TAF4B,TAFII21,TAF6B,TAF7,TBP2,TAF2,TAF12,Manes.16G060700,TAFII15,TAF4
GO:0006366	transcription by RNA polymerase II	11	52	1.88	1.90e-13	TAF13,TAFII59,EER4,TAF11b,TAFII21,TAF6B,TAF7,TBP2,TAF2,TAF12,Manes.16G060700

GO:0006351	transcription, DNA-templated	14	183	1.44	1.95e-12	TAF13,TAFII59,EER4,TAF11b,TAF4B,TAFII21,TAF6B,TAF7,TBP2,TAF2,TAF12,Manes.16G060700,TAFII15,TAF4
GO:0006367	transcription initiation from RNA polymerase II promoter	8	22	2.11	4.23e-11	TAFII59,EER4,TAF11b,TAF6B,TAF7,TAF2,TAF12,Manes.16G060700
GO:0010499	proteasomal ubiquitin-independent protein catabolic process	8	25	2.06	8.55e-11	PRC3,PBC1,PAE1,PBG1,AT1G79210,PAG1,Manes.01G265800,PBA1
GO:0051123	RNA polymerase II preinitiation complex assembly	7	13	2.28	1.68e-10	TAFII59,EER4,TAF11b,TAF6B,TAF7,TAF12,Manes.16G060700
GO:0044260	cellular macromolecule metabolic process	37	3665	0.56	4.51e-10	TAF13,RRP4,TAFII59,PRC3,EER4,TAF11b,PBC1,TAF4B,ATS9,HAF01,Manes.03G029000,PAE1,TAFII21,TAF6B,TAF7,TBP2,PBG1,TAF2,AT1G79210,Manes.06G176300-CCNGG,Manes.12G110200,RPN13,PAG1,TAF12,Manes.03G083600,Manes.04G014300,Manes.07G132800,Manes.09G039500,Manes.01G181600,Manes.01G265800,RRP41,Manes.16G060700,RPT2a,PBA1,TAFII15,TAF4,Manes.13G155300
GO:0009057	macromolecule catabolic process	19	761	0.95	5.61e-10	RRP4,PRC3,PBC1,ATS9,PAE1,PBG1,AT1G79210,RPN13,PAG1,Manes.03G083600,Manes.04G014300,Manes.09G039500,Manes.01G181600,Manes.01G265800,RRP41,RPT2a,PBA1,Manes.15G190300-CCNGG,RPT3
GO:0044265	cellular macromolecule catabolic process	17	578	1.02	8.61e-10	RRP4,PRC3,PBC1,ATS9,PAE1,PBG1,AT1G79210,RPN13,PAG1,Manes.03G083600,Manes.04G014300,Manes.09G039500,Manes.01G181600,Manes.01G265800,RRP41,RPT2a,PBA1
GO:0034654	nucleobase-containing compound biosynthetic process	15	439	1.09	2.75e-09	TAF13,TAFII59,EER4,TAF11b,TAF4B,TAFII21,TAF6B,TAF7,TBP2,TAF2,TAF12,Manes.15G159500,Manes.16G060700,TAFII15,TAF4
GO:0043632	modification-dependent macromolecule catabolic process	14	371	1.13	4.17e-09	RRP4,PRC3,PBC1,ATS9,PAE1,PBG1,AT1G79210,RPN13,PAG1,Manes.03G083600,Manes.01G181600,Manes.01G265800,RPT2a,PBA1
GO:0043170	macromolecule metabolic process	40	4720	0.48	5.15e-09	TAF13,RRP4,TAFII59,PRC3,EER4,TAF11b,PBC1,TAF4B,ATS9,HAF01,Manes.03G029000,PAE1,TAFII21,TAF6B,TAF7,TBP2,PBG1,TAF2,AT1G79210,Manes.06G176300-CCNGG,Manes.12G110200,RPN13,PAG1,TAF12,Manes.03G083600,Manes.04G014300,Manes.07G132800,Manes.09G039500,Manes.01G181600,Manes.01G265800,RRP41,Manes.16G060700,RPT2a,PBA1,TAFII15,TAF4,Manes.15G190300-CCNGG,Manes.13G155300,RPT3,Manes.01G164300
GO:0006807	nitrogen compound metabolic process	42	5335	0.45	1.06e-08	TAF13,RRP4,TAFII59,PRC3,EER4,TAF11b,PBC1,TAF4B,ATS9,HAF01,Manes.03G029000,PAE1,TAFII21,TAF6B,TAF7,TBP2,PBG1,TAF2,AT1G79210,Manes.06G176300-CCNGG,Manes.12G110200,RPN13,PAG1,TAF12,Manes.03G083600,Manes.15G159500,Manes.04G014300,Manes.07G132800,Manes.09G039500,Manes.01G181600,Manes.01G265800,RRP41,Manes.16G060700,RPT2a,PBA1,TAFII15,Manes.04G000700,TAF4,Manes.13G155300,Manes.03G103900,RPT3,Manes.01G164300
GO:0006508	proteolysis	17	745	0.91	3.02e-08	PRC3,PBC1,ATS9,Manes.03G029000,PAE1,PBG1,TAF2,AT1G79210,Manes.06G176300-CCNGG,RPN13,PAG1,Manes.03G083600,Manes.04G014300,Manes.01G265800,RPT2a,PBA1,Manes.01G164300
GO:0016070	RNA metabolic process	19	1049	0.81	7.87e-08	TAF13,RRP4,TAFII59,EER4,TAF11b,TAF4B,PAE1,TAFII21,TAF6B,TAF7,TBP2,TAF2,TAF12,Manes.09G039500,Manes.01G181600,RRP41,Manes.16G060700,TAFII15,TAF4

GO:0030163	protein catabolic process	14	484	1.01	9.33e-08	PRC3,PBC1,ATS9,PAE1,PBG1,AT1G79210,RPN13,PAG1,Manes.03G083600,Manes.04G014300,Manes.01G265800,RPT2a,PBA1,RPT3
GO:0010467	gene expression	20	1235	0.76	1.43e-07	TAF13,RRP4,TAFII59,EER4,TAF11b,TAF4B,Manes.03G029000,TAFII21,TAF6B,TAF7,TBP2,TAF2,Manes.06G176300-CCNGG,TAF12,Manes.09G039500,Manes.01G181600,RRP41,Manes.16G060700,TAFII15,TAF4
GO:0051603	proteolysis involved in cellular protein catabolic process	13	439	1.02	3.04e-07	PRC3,PBC1,ATS9,PAE1,PBG1,AT1G79210,RPN13,PAG1,Manes.03G083600,Manes.04G014300,Manes.01G265800,RPT2a,PBA1
GO:0006511	ubiquitin-dependent protein catabolic process	12	352	1.08	3.13e-07	PRC3,PBC1,ATS9,PAE1,PBG1,AT1G79210,RPN13,PAG1,Manes.03G083600,Manes.01G265800,RPT2a,PBA1
GO:0044238	primary metabolic process	43	6374	0.38	5.28e-07	TAF13,RRP4,TAFII59,PRC3,EER4,TAF11b,PBC1,TAF4B,ATS9,HAF01,Manes.03G029000,Manes.02G091300,PAE1,TAFII21,TAF6B,TAF7,TBP2,PBG1,TAF2,AT1G79210,Manes.06G176300-CCNGG,Manes.12G110200,RPN13,PAG1,TAF12,Manes.03G083600,Manes.15G159500,Manes.04G014300,Manes.07G132800,Manes.09G039500,Manes.01G181600,Manes.01G265800,RRP41,Manes.16G060700,RPT2a,PBA1,TAFII15,TAF4,Manes.15G190300-CCNGG,Manes.13G155300,Manes.03G103900,RPT3,Manes.01G164300
GO:1901575	organic substance catabolic process	20	1353	0.72	5.28e-07	RRP4,PRC3,PBC1,ATS9,PAE1,PBG1,AT1G79210,RPN13,PAG1,Manes.03G083600,Manes.04G014300,Manes.09G039500,Manes.01G181600,Manes.01G265800,RRP41,RPT2a,PBA1,Manes.15G190300-CCNGG,Manes.03G103900,RPT3
GO:1901565	organonitrogen compound catabolic process	15	699	0.88	6.84e-07	PRC3,PBC1,ATS9,PAE1,PBG1,AT1G79210,RPN13,PAG1,Manes.03G083600,Manes.04G014300,Manes.01G265800,RPT2a,PBA1,Manes.03G103900,RPT3
GO:0090304	nucleic acid metabolic process	20	1403	0.71	8.48e-07	TAF13,RRP4,TAFII59,EER4,TAF11b,TAF4B,PAE1,TAFII21,TAF6B,TAF7,TBP2,TAF2,TAF12,Manes.09G039500,Manes.01G181600,RRP41,Manes.16G060700,TAFII15,TAF4,Manes.13G155300
GO:0044248	cellular catabolic process	18	1167	0.74	1.86e-06	RRP4,PRC3,PBC1,ATS9,PAE1,PBG1,AT1G79210,RPN13,PAG1,Manes.03G083600,Manes.04G014300,Manes.09G039500,Manes.01G181600,Manes.01G265800,RRP41,RPT2a,PBA1,Manes.03G103900
GO:0071704	organic substance metabolic process	44	6962	0.35	1.92e-06	TAF13,RRP4,TAFII59,PRC3,EER4,TAF11b,PBC1,TAF4B,ATS9,HAF01,Manes.03G029000,Manes.02G091300,PAE1,TAFII21,TAF6B,TAF7,TBP2,PBG1,TAF2,AT1G79210,Manes.06G176300-CCNGG,Manes.12G110200,RPN13,PAG1,TAF12,Manes.03G083600,Manes.15G159500,Manes.04G014300,Manes.07G132800,Manes.09G039500,Manes.01G181600,Manes.01G265800,RRP41,Manes.16G060700,RPT2a,PBA1,TAFII15,Manes.04G000700,TAF4,Manes.15G190300-CCNGG,Manes.13G155300,Manes.03G103900,RPT3,Manes.01G164300
GO:0043161	proteasome-mediated ubiquitin-dependent protein catabolic process	9	185	1.24	1.98e-06	PRC3,PBC1,PAE1,PBG1,AT1G79210,PAG1,Manes.01G265800,RPT2a,PBA1
GO:0044237	cellular metabolic process	43	6729	0.36	2.34e-06	TAF13,RRP4,TAFII59,PRC3,EER4,TAF11b,PBC1,TAF4B,ATS9,HAF01,Manes.03G029000,Manes.02G091300,PAE1,TAFII21,TAF6B,TAF7,TBP2,PBG1,TAF2,AT1G79210,Manes.06G176300-CCNGG,Manes.12G110200,RPN13,PAG1,TAF12,Manes.03G083600,Manes.15G159500,Manes.04G014300,Manes.07G132800,Manes.09G039500,Manes.01G181600,Manes.01G265800,RRP41,Manes.16G060700,RPT2a,PBA1,TAFII15,Manes.09G110600,Manes.04G000700,TAF4,Manes.13G155300,Manes.03G103900,Manes.12G156000

GO:0006139	nucleobase-containing compound metabolic process	21	1736	0.63	4.11e-06	TAF13,RRP4,TAFII59,EER4,TAF11b,TAF4B,PAE1,TAFII21,TAF6B,TAF7,TBP2,TAF2,TAF12,Manes.15G159500,Manes.09G039500,Manes.01G181600,RRP41,Manes.16G060700,TAFII15,TAF4,Manes.13G155300
GO:0008152	metabolic process	46	7962	0.31	9.43e-06	TAF13,RRP4,TAFII59,PRC3,EER4,TAF11b,PBC1,TAF4B,ATS9,HAF01,Manes.03G029000,Manes.02G091300,PAE1,TAFII21,TAF6B,TAF7,TBP2,PBG1,TAF2,AT1G79210,Manes.06G176300-CCNGG,Manes.12G110200,RPN13,PAG1,TAF12,Manes.03G083600,Manes.15G159500,Manes.04G014300,Manes.07G132800,Manes.09G039500,Manes.01G181600,Manes.01G265800,RRP41,Manes.16G060700,RPT2a,PBA1,TAFII15,Manes.09G110600,Manes.04G000700,TAF4,Manes.15G190300-CCNGG,Manes.13G155300,Manes.03G103900,RPT3,Manes.01G164300,Manes.12G156000
GO:0065003	protein-containing complex assembly	11	415	0.97	1.21e-05	TAFII59,EER4,TAF11b,ATS9,Manes.02G091300,TAF6B,TAF7,Manes.06G176300-CCNGG,TAF12,Manes.16G060700,RPT2a
GO:0006725	cellular aromatic compound metabolic process	22	2137	0.56	2.67e-05	TAF13,RRP4,TAFII59,EER4,TAF11b,TAF4B,PAE1,TAFII21,TAF6B,TAF7,TBP2,TAF2,TAF12,Manes.15G159500,Manes.09G039500,Manes.01G181600,RRP41,Manes.16G060700,TAFII15,TAF4,Manes.13G155300,Manes.03G103900
GO:1901360	organic cyclic compound metabolic process	22	2229	0.55	5.19e-05	TAF13,RRP4,TAFII59,EER4,TAF11b,TAF4B,PAE1,TAFII21,TAF6B,TAF7,TBP2,TAF2,TAF12,Manes.15G159500,Manes.09G039500,Manes.01G181600,RRP41,Manes.16G060700,TAFII15,TAF4,Manes.13G155300,Manes.03G103900
GO:0034645	cellular macromolecule biosynthetic process	15	1040	0.71	7.39e-05	TAF13,TAFII59,EER4,TAF11b,TAF4B,TAFII21,TAF6B,TAF7,TBP2,TAF2,TAF12,Manes.09G039500,Manes.16G060700,TAFII15,TAF4
GO:0010604	positive regulation of macromolecule metabolic process	12	675	0.8	0.00016	RRP4,TAF11b,TAF4B,HAF01,TAFII21,TAF7,TAF2,RPN13,Manes.16G060700,RPT2a,TAF4,RPT3
GO:0006357	regulation of transcription by RNA polymerase II	9	335	0.98	0.00018	TAF13,TAF4B,HAF01,TAFII21,TAF7,TAF2,Manes.05G184900,TAF4,RPT3
GO:0034641	cellular nitrogen compound metabolic process	22	2435	0.51	0.00021	TAF13,RRP4,TAFII59,EER4,TAF11b,TAF4B,PAE1,TAFII21,TAF6B,TAF7,TBP2,TAF2,TAF12,Manes.15G159500,Manes.09G039500,Manes.01G181600,RRP41,Manes.16G060700,TAFII15,Manes.04G000700,TAF4,Manes.13G155300
GO:0034622	cellular protein-containing complex assembly	9	350	0.96	0.00024	TAFII59,EER4,TAF11b,ATS9,TAF6B,TAF7,TAF12,Manes.16G060700,RPT2a
GO:0044085	cellular component biogenesis	14	1028	0.69	0.00034	RRP4,TAFII59,EER4,TAF11b,ATS9,Manes.02G091300,TAF6B,TAF7,Manes.06G176300-CCNGG,TAF12,Manes.01G181600,RRP41,Manes.16G060700,RPT2a
GO:0034475	U4 snRNA 3-end processing	3	6	2.25	0.00055	RRP4,Manes.01G181600,RRP41
GO:0051173	positive regulation of nitrogen compound metabolic process	11	650	0.78	0.00070	TAF11b,TAF4B,HAF01,TAFII21,TAF7,TAF2,RPN13,Manes.16G060700,RPT2a,TAF4,RPT3
GO:0060255	regulation of macromolecule metabolic process	22	2675	0.47	0.00086	TAF13,RRP4,EER4,TAF11b,TAF4B,HAF01,TAFII21,TAF7,TBP2,TAF2,RPN13,Manes.15G159500,Manes.09G039500,Manes.01G181600,RRP41,Manes.05G184900,Manes.16G060700,RPT2a,TAF4,Manes.15G190300-CCNGG,Manes.13G155300,RPT3
GO:0071840	cellular component organization or biogenesis	21	2482	0.48	0.00098	RRP4,TAFII59,EER4,TAF11b,ATS9,HAF01,Manes.03G029000,Manes.02G091300,TAF6B,TAF7,Manes.06G176300-

						CCNGG,Manes.16G110100,Manes.01G129200,EMB2795,TAF12,Manes.01G181600,RRP41,Manes.16G060700,RPT2a,Manes.16G098500,Manes.13G155300
GO:0031325	positive regulation of cellular metabolic process	11	702	0.75	0.0013	TAF11b,TAF4B,HAF01,TAFI21,TAF7,TAF2,RPN13,Manes.16G060700,RPT2a,TAF4,RPT3
GO:0034427	nuclear-transcribed mRNA catabolic process,	3	9	2.07	0.0013	RRP4,Manes.01G181600,RRP41
GO:0071028	exonucleolytic, 3-5 nuclear mRNA surveillance	3	10	2.03	0.0016	RRP4,Manes.01G181600,RRP41
GO:1903508	positive regulation of nucleic acid-templated transcription	9	477	0.83	0.0023	TAF11b,TAF4B,HAF01,TAFI21,TAF7,TAF2,Manes.16G060700,TAF4,RPT3
GO:0009987	cellular process	50	11004	0.21	0.0025	TAF13,RRP4,TAFI59,PRC3,EER4,TAF11b,PBC1,TAF4B,ATS9,HAF01,Manes.03G029000,Manes.02G091300,PAE1,TAFI21,TAF6B,Manes.09G162500-CCNGG,TAF7,TBP2,PBG1,TAF2,AT1G79210,Manes.06G176300-CCNGG,Manes.16G110100,Manes.12G110200,RPN13,PAG1,Manes.01G129200,EMB2795,TAF12,Manes.03G083600,Manes.15G159500,Manes.04G014300,Manes.07G132800,Manes.09G039500,Manes.01G181600,Manes.01G265800,RRP41,Manes.05G184900,Manes.16G060700,RPT2a,PBA1,TAFI15,Manes.09G110600,Manes.16G098500,Manes.04G000700,TAF4,Manes.13G155300,Manes.01G040400,Manes.03G103900,Manes.12G156000
GO:0048522	positive regulation of cellular process	12	930	0.66	0.0028	TAF11b,TAF4B,HAF01,TAFI21,TAF7,TAF2,RPN13,Manes.01G129200,Manes.16G060700,RPT2a,TAF4,RPT3
GO:0048518	positive regulation of biological process	13	1146	0.61	0.0043	RRP4,TAF11b,TAF4B,HAF01,TAFI21,TAF7,TAF2,RPN13,Manes.01G129200,Manes.16G060700,RPT2a,TAF4,RPT3
GO:0016075	rRNA catabolic process	3	16	1.82	0.0045	RRP4,Manes.01G181600,RRP41
GO:0045944	positive regulation of transcription by RNA polymerase II	6	198	1.03	0.0052	TAF4B,HAF01,TAFI21,TAF2,TAF4,RPT3
GO:0019538	protein metabolic process	22	3075	0.41	0.0058	PRC3,EER4,PBC1,ATS9,HAF01,Manes.03G029000,PAE1,PBG1,TAF2,AT1G79210,Manes.06G176300-CCNGG,Manes.12G110200,RPN13,PAG1,Manes.03G083600,Manes.04G014300,Manes.07G132800,Manes.01G265800,RPT2a,PBA1,RPT3,Manes.01G164300
GO:0016043	cellular component organization	18	2264	0.45	0.0096	TAFI59,EER4,TAF11b,ATS9,HAF01,Manes.03G029000,Manes.02G091300,TAF6B,TAF7,Manes.06G176300-CCNGG,Manes.16G110100,Manes.01G129200,EMB2795,TAF12,Manes.16G060700,RPT2a,Manes.16G098500,Manes.13G155300
GO:0000956	nuclear-transcribed mRNA catabolic process	4	69	1.31	0.0108	RRP4,Manes.09G039500,Manes.01G181600,RRP41
GO:0010468	regulation of gene expression	18	2307	0.44	0.0117	TAF13,RRP4,EER4,TAF4B,HAF01,TAFI21,TAF7,TBP2,TAF2,Manes.09G039500,Manes.01G181600,RRP41,Manes.05G184900,RPT2a,TAF4,Manes.15G190300-CCNGG,Manes.13G155300,RPT3
GO:0000288	nuclear-transcribed mRNA catabolic	3	27	1.6	0.0160	RRP4,Manes.09G039500,Manes.01G181600

	process, deadenylation- dependent decay					
GO:0010038	response to metal ion	7	374	0.82	0.0187	PRC3,ATS9,PAE1,PBG1,PAG1,RPT2a,Manes.09G110600
GO:0090501	RNA phosphodiester bond hydrolysis	4	100	1.15	0.0398	RRP4,PAE1,Manes.09G039500,Manes.01G181600
GO:0071038	nuclear polyadenylation- dependent tRNA catabolic process	2	6	2.07	0.0399	RRP4,Manes.01G181600
GO:1901564	organonitrogen compound metabolic process	24	4053	0.32	0.0399	PRC3,EER4,PBC1,ATS9,HAF01,Manes.03G029000,PAE1,PBG1,TAF2,AT1G79210,Manes.06G176300-CCNGG,Manes.12G110200,RPN13,PAG1,Manes.03G083600,Manes.15G159500,Manes.04G014300,Manes.07G132800,Manes.01G265800,RPT2a,PBA1,Manes.03G103900,RPT3,Manes.01G164300
GO:0044267	cellular protein metabolic process	19	2809	0.38	0.0418	PRC3,EER4,PBC1,ATS9,HAF01,Manes.03G029000,PAE1,PBG1,AT1G79210,Manes.06G176300-CCNGG,Manes.12G110200,RPN13,PAG1,Manes.03G083600,Manes.04G014300,Manes.07G132800,Manes.01G265800,RPT2a,PBA1
GO:0045862	positive regulation of proteolysis	3	41	1.42	0.0460	RPN13,RPT2a,RPT3
GO:0090503	RNA phosphodiester bond hydrolysis, exonucleolytic	3	41	1.42	0.0460	RRP4,Manes.09G039500,Manes.01G181600
GO:0019439	aromatic compound catabolic process	5	198	0.95	0.0470	RRP4,Manes.09G039500,Manes.01G181600,RRP41,Manes.03G103900
GO:0000467	exonucleolytic trimming to generate mature 3-end of 5.8S rRNA from tricistronic rRNA transcript (SSU- rRNA, 5.8S rRNA, LSU-rRNA)	2	7	2.01	0.0477	RRP4,Manes.01G181600
GO:0034982	mitochondrial protein processing	2	7	2.01	0.0477	Manes.03G029000,Manes.06G176300-CCNGG
GO:0065007	biological regulation	28	5216	0.28	0.0477	TAF13,RRP4,TAFII59,EER4,TAF11b,TAF4B,HAF01,TAFII21,TAF6B,TAF7,TBP2,TAF2,Manes.12G110200,RPN13,Manes.01G129200,EMB2795,Manes.15G159500,Manes.09G039500,Manes.01G181600,RRP41,Manes.05G184900,Manes.16G060700,RPT2a,Manes.09G110600,TAF4,Manes.15G190300-CCNGG,Manes.13G155300,RPT3
GO:0071051	polyadenylation- dependent snoRNA 3- end processing	2	7	2.01	0.0477	RRP4,RRP41

Tabela suplementar 4B. Termos de GO para função molecular para proteínas da rede de interações formada por sítios diferencialmente hipermetilados para a variedade tolerante Formosa baseados em proteínas de *A. thaliana* órtologas às proteínas de *M. esculenta*. As colunas indicam respectivamente: Id do termo GO; descrição do termo; contagem do termo na rede de interações; background de informação para este termo no banco de dados; nível de confiabilidade; taxa de falso descobrimento; identificador das proteínas que contem o termo GO; nome das proteínas que contém o termo GO

ID do termo	Descrição	Contagem de gene	Background	Strength	False discovery rate	Matching proteins nas redes
GO:0004298	threonine-type endopeptidase activity	8	25	2.06	1.79e-10	PRC3,PBC1,PAE1,PBG1,AT1G79210,PAG1,Manes.01G265800,PBA1
GO:0008134	transcription factor binding	11	127	1.49	4.48e-10	EER4,TAF11b,TAF4B,HAF01,TAII21,TAF7,TBP2,TAF12,Manes.16G060700,RPT2a,TAF4
GO:0008233	peptidase activity	14	421	1.07	2.22e-08	PRC3,PBC1,Manes.03G029000,PAE1,PBG1,TAF2,AT1G79210,Manes.06G176300-CCNGG,PAG1,Manes.03G083600,Manes.04G014300,Manes.01G265800,PBA1,Manes.01G164300
GO:0004175	endopeptidase activity	12	272	1.2	2.48e-08	PRC3,PBC1,Manes.03G029000,PAE1,PBG1,AT1G79210,Manes.06G176300-CCNGG,PAG1,Manes.03G083600,Manes.01G265800,PBA1,Manes.01G164300
GO:0046982	protein heterodimerization activity	8	106	1.43	9.85e-07	TAF13,TAII59,EER4,TAF11b,TAII21,TAF6B,TAF12,Manes.16G060700
GO:0017025	TBP-class protein binding	5	16	2.05	2.22e-06	EER4,HAF01,TBP2,TAF12,RPT2a
GO:0005515	protein binding	22	1997	0.59	1.15e-05	TAF13,TAII59,EER4,TAF11b,TAF4B,HAF01,TAII21,TAF6B,TAF7,TBP2,RPN13,Manes.01G129200,EMB2795,TAF12,Manes.16G060700,RPT2a,Manes.09G110600,Manes.16G098500,TAF4,Manes.01G040400,Manes.12G156000,Manes.13G112600
GO:0016787	hydrolase activity	22	2363	0.52	0.00019	Manes.12G033800,PRC3,PBC1,Manes.03G029000,PAE1,PBG1,TAF2,AT1G79210,Manes.06G176300-CCNGG,PAG1,Manes.03G083600,Manes.04G014300,Manes.09G039500,Manes.01G181600,Manes.01G265800,RPT2a,PBA1,Manes.16G098500,Manes.04G000700,Manes.15G190300-CCNGG,RPT3,Manes.01G164300
GO:0003713	transcription coactivator activity	3	26	1.61	0.0319	TAF11b,TAF7,Manes.16G060700
GO:0008237	metallopeptidase activity	4	74	1.28	0.0319	Manes.03G029000,TAF2,Manes.06G176300-CCNGG,Manes.01G164300
GO:0016251	RNA polymerase II general transcription initiation factor activity	2	3	2.38	0.0319	TAII59,TAF6B

GO:0036402	proteasome-activating ATPase activity	2	3	2.38	0.0319	RPT2a,RPT3
GO:0140096	catalytic activity, acting on a protein	16	1898	0.48	0.0319	PRC3,PBC1,Manes.03G029000,PAE1,PBG1,TAF2,AT1G79210,Manes.06G176300-CCNGG,Manes.12G110200,PAG1,Manes.03G083600,Manes.04G014300,Manes.07G132800,Manes.01G265800,PBA1,Manes.01G164300

Tabela suplementar 4C. Termos de GO para componente celular para proteínas da rede de interações formada por sítios diferencialmente hipermetilados para a variedade tolerante Formosa baseados em proteínas de *A. thaliana* órtologas às proteínas de *M. esculenta*. As colunas indicam respectivamente: Id do termo GO; descrição do termo; contagem do termo na rede de interações; background de informação para este termo no banco de dados; nível de confiabilidade; taxa de falso descobrimento; nome das proteínas que contém o termo GO

ID do termo	Descrição	Contagem de gene	Background	Strength	False discovery rate	Matching proteins nas redes
GO:0005669	transcription factor TFIID complex	14	18	2.44	5.82e-25	TAF13,TAFII59,EER4,TAF11b,TAF4B,HAF01,TAFII21,TAF6B,TAF7,TBP2,TAF2,TAF12,Manes.16G060700,TAF4
GO:1905368	peptidase complex	17	73	1.92	5.92e-24	TAFII59,PRC3,EER4,PBC1,ATS9,PAE1,TAFII21,TAF6B,PBG1,AT1G79210,RPN13,PAG1,TAF12,Manes.01G265800,RPT2a,PBA1,RPT3
GO:1902494	catalytic complex	30	776	1.14	2.87e-23	TAF13,RRP4,TAFII59,PRC3,EER4,TAF11b,PBC1,TAF4B,ATS9,HAF01,Manes.03G029000,PAE1,TAFII21,TAF6B,TAF7,TBP2,PBG1,TAF2,AT1G79210,RPN13,PAG1,TAF12,Manes.01G181600,Manes.01G265800,RRP41,Manes.16G060700,RPT2a,PBA1,TAF4,RPT3
GO:0032991	protein-containing complex	36	2226	0.76	6.84e-17	TAF13,RRP4,TAFII59,PRC3,EER4,TAF11b,PBC1,TAF4B,ATS9,HAF01,Manes.03G029000,PAE1,TAFII21,TAF6B,TAF7,TBP2,PBG1,TAF2,AT1G79210,Manes.12G110200,RPN13,PAG1,Manes.01G129200,EMB2795,TAF12,Manes.09G039500,Manes.01G181600,Manes.01G265800,RRP41,Manes.16G060700,RPT2a,PBA1,Manes.16G098500,TAF4,RPT3,Manes.12G156000
GO:0000502	proteasome complex	12	60	1.85	2.85e-16	PRC3,PBC1,ATS9,PAE1,PBG1,AT1G79210,RPN13,PAG1,Manes.01G265800,RPT2a,PBA1,RPT3
GO:0005634	nucleus	43	4669	0.52	2.94e-12	TAF13,RRP4,TAFII59,PRC3,EER4,TAF11b,PBC1,TAF4B,ATS9,HAF01,TAF15,Manes.03G029000,Manes.02G091300,PAE1,TAFII21,TAF6B,TAF7,TBP2,PBG1,TAF2,AT1G79210,Manes.12G110200,RPN13,PAG1,Manes.01G129200,EMB2795,TAF12,Manes.03G083600,Manes.09G039500,Manes.01G181600,Manes.01G265800,RRP41,Manes.05G184900,Manes.16G060700,RPT2a,PBA1,TAFII15,TAF4,Manes.15G190300-CCNGG,Manes.13G155300,Manes.01G040400,RPT3,Manes.13G112600
GO:0005839	proteasome core complex	8	24	2.07	4.25e-12	PRC3,PBC1,PAE1,PBG1,AT1G79210,PAG1,Manes.01G265800,PBA1

GO:0005654	nucleoplasm	15	396	1.13	7.54e-11	TAF13,RRP4,TAFII59,EER4,TAF11b,TAF4B,HAF01,TAFII21,TAF6B,TAF7,TBP2,TAF2,TAF12,Manes.16G060700,TAF4
GO:0031981	nuclear lumen	19	807	0.92	1.39e-10	TAF13,RRP4,TAFII59,EER4,TAF11b,TAF4B,HAF01,TAFII21,TAF6B,TAF7,TBP2,TAF2,EMB2795,TAF12,Manes.01G181600,RRP41,Manes.16G060700,TAF4,Manes.13G112600
GO:0070013	intracellular organelle lumen	20	1004	0.85	5.91e-10	TAF13,RRP4,TAFII59,EER4,TAF11b,TAF4B,HAF01,Manes.03G029000,TAFII21,TAF6B,TAF7,TBP2,TAF2,EMB2795,TAF12,Manes.01G181600,RRP41,Manes.16G060700,TAF4,Manes.13G112600
GO:0000124	SAGA complex	5	9	2.3	3.31e-08	TAFII59,EER4,TAFII21,TAF6B,TAF12
GO:0043227	membrane-bounded organelle	53	9634	0.29	2.03e-07	TAF13,RRP4,TAFII59,PRC3,EER4,TAF11b,PBC1,TAF4B,ATS9,HAF01,TAF15,Manes.03G029000,Manes.02G091300,PAE1,TAFII21,TAF6B,TAF7,TBP2,PBG1,TAF2,AT1G79210,Manes.06G176300-CCNGG,Manes.16G110100,Manes.12G110200,RPN13,PAG1,Manes.01G129200,EMB2795,TAF12,Manes.03G083600,Manes.15G159500,Manes.04G014300,Manes.09G039500,Manes.01G181600,Manes.01G265800,RRP41,Manes.05G184900,Manes.16G060700,RPT2a,PBA1,TAFII15,Manes.09G110600,Manes.16G098500,Manes.03G066300,TAF4,Manes.15G190300-CCNGG,Manes.13G155300,Manes.01G040400,Manes.03G103900,RPT3,Manes.01G164300,Manes.12G156000,Manes.13G112600
GO:0043231	intracellular membrane-bounded organelle	52	9376	0.3	2.67e-07	TAF13,RRP4,TAFII59,PRC3,EER4,TAF11b,PBC1,TAF4B,ATS9,HAF01,TAF15,Manes.03G029000,Manes.02G091300,PAE1,TAFII21,TAF6B,TAF7,TBP2,PBG1,TAF2,AT1G79210,Manes.06G176300-CCNGG,Manes.16G110100,Manes.12G110200,RPN13,PAG1,Manes.01G129200,EMB2795,TAF12,Manes.03G083600,Manes.15G159500,Manes.04G014300,Manes.09G039500,Manes.01G181600,Manes.01G265800,RRP41,Manes.05G184900,Manes.16G060700,RPT2a,PBA1,TAFII15,Manes.09G110600,Manes.03G066300,TAF4,Manes.15G190300-CCNGG,Manes.13G155300,Manes.01G040400,Manes.03G103900,RPT3,Manes.01G164300,Manes.12G156000,Manes.13G112600
GO:0046695	SLIK (SAGA-like) complex	4	4	2.55	3.73e-07	TAFII59,EER4,TAF6B,TAF12
GO:0043229	intracellular organelle	53	9831	0.28	4.06e-07	TAF13,RRP4,TAFII59,PRC3,EER4,TAF11b,PBC1,TAF4B,ATS9,HAF01,TAF15,Manes.03G029000,Manes.02G091300,PAE1,TAFII21,TAF6B,TAF7,TBP2,PBG1,TAF2,AT1G79210,Manes.06G176300-CCNGG,Manes.16G110100,Manes.12G110200,RPN13,PAG1,Manes.01G129200,EMB2795,TAF12,Manes.03G083600,Manes.15G159500,Manes.04G014300,Manes.09G039500,Manes.01G181600,Manes.01G265800,RRP41,Manes.05G184900,Manes.16G060700,RPT2a,PBA1,TAFII15,Manes.09G110600,Manes.16G098500,Manes.03G066300,TAF4,Manes.15G190300-CCNGG,Manes.13G155300,Manes.01G040400,Manes.03G103900,RPT3,Manes.01G164300,Manes.12G156000,Manes.13G112600
GO:0005622	intracellular	55	10863	0.26	1.41e-06	TAF13,RRP4,Manes.12G033800,TAFII59,PRC3,EER4,TAF11b,PBC1,TAF4B,ATS9,HAF01,TAF15,Manes.03G029000,Manes.02G091300,PAE1,TAFII21,TAF6B,TAF7,TBP2,PBG1,TAF2,AT1G79210,Manes.06G176300-CCNGG,Manes.16G110100,Manes.12G110200,RPN13,PAG1,Manes.01G129200,EMB2795,TAF12,Manes.03G083600,Manes.15G159500,Manes.04G014300,Manes.09G039500,Manes.01G181600,Manes.01G265800,RRP41,Manes.05G184900,Manes.16G060700,RPT2a,PBA1,TAFII15,Manes.09G110600,Manes.16G098500,Manes.04G000700,Manes.03G066300,TAF4,Manes.15G190300-CCNGG,Manes.13G155300,Manes.01G040400,Manes.03G103900,RPT3,Manes.01G164300,Manes.12G156000,Manes.13G112600
GO:0019774	proteasome core complex,	4	11	2.11	5.65e-06	PBC1,PBG1,Manes.01G265800,PBA1

GO:0019773	beta-subunit complex proteasome core complex, alpha-subunit complex	4	13	2.04	9.52e-06	PRC3,PAE1,AT1G79210,PAG1
GO:0005838	proteasome regulatory particle cytoplasmic	4	31	1.66	0.00020	ATS9,RPN13,RPT2a,RPT3
GO:0000177	exosome (RNase complex)	3	9	2.07	0.00030	RRP4,Manes.01G181600,RRP41
GO:0000176	nuclear exosome (RNase complex)	3	10	2.03	0.00038	RRP4,Manes.01G181600,RRP41
GO:0110165	cellular anatomical entity	57	13776	0.17	0.0011	TAF13,RRP4,Manes.12G033800,TAFII59,PRC3,EER4,TAF11b,PBC1,TAF4B,ATS9,HAF01,TAF15,Manes.03G029000,Manes.02G091300,PAE1,TAFII21,TAF6B,Manes.09G162500-CCNGG,TAF7,TBP2,PBG1,TAF2,AT1G79210,Manes.06G176300-CCNGG,Manes.16G110100,Manes.12G110200,RPN13,PAG1,Manes.01G129200,EMB2795,TAF12,Manes.03G083600,Manes.15G159500,Manes.04G014300,Manes.07G132800,Manes.09G039500,Manes.01G181600,Manes.01G265800,RRP41,Manes.05G184900,Manes.16G060700,RPT2a,PBA1,TAFII15,Manes.09G110600,Manes.16G098500,Manes.04G000700,Manes.03G066300,TAF4,Manes.15G190300-CCNGG,Manes.13G155300,Manes.01G040400,Manes.03G103900,RPT3,Manes.01G164300,Manes.12G156000,Manes.13G112600
GO:0005829	cytosol	18	2567	0.4	0.0133	PRC3,PBC1,ATS9,Manes.02G091300,PAE1,PBG1,AT1G79210,Manes.12G110200,PAG1,Manes.03G083600,Manes.15G159500,Manes.09G039500,Manes.01G181600,Manes.01G265800,RRP41,Manes.04G000700,Manes.03G103900,RPT3
GO:0005737	cytoplasm	37	7911	0.22	0.0168	RRP4,Manes.12G033800,PRC3,PBC1,ATS9,Manes.03G029000,Manes.02G091300,PAE1,PBG1,AT1G79210,Manes.06G176300-CCNGG,Manes.16G110100,Manes.12G110200,RPN13,PAG1,Manes.01G129200,EMB2795,Manes.03G083600,Manes.15G159500,Manes.04G014300,Manes.09G039500,Manes.01G181600,Manes.01G265800,RRP41,RPT2a,PBA1,Manes.09G110600,Manes.16G098500,Manes.04G000700,Manes.03G066300,Manes.15G190300-CCNGG,Manes.01G040400,Manes.03G103900,RPT3,Manes.01G164300,Manes.12G156000,Manes.13G112600
GO:0008541	proteasome regulatory particle, lid subcomplex	2	13	1.74	0.0443	ATS9,RPN13

Tabela suplementar 4D. Termos KEGG para proteínas da rede de interações formada por sítios diferencialmente hipermetilados para a variedade tolerante Formosa baseados em proteínas de *A. thaliana* órtologas às proteínas de *M. esculenta*. As colunas indicam respectivamente: Id do termo; descrição do termo; contagem do termo na rede de interações; background de informação para este termo no banco de dados; nível de confiabilidade; taxa de falso descobrimento; nome das proteínas que contém o termo

ID do termo	Descrição	Contagem de gene	Background	Strength	False discovery rate	Matching proteins nas redes
ath03022	Basal transcription factors	15	55	1.99	3.49e-22	TAF13,TAFII59,EER4,TAF11b,TAF4B,HAF01,TAFII21,TAF6B,TAF7,TBP2,TAF2,TAF12,Manes.16G060700,TAFII15,TAF4
ath03050	Proteasome	12	58	1.87	2.02e-16	PRC3,PBC1,ATS9,PAE1,PBG1,AT1G79210,RPN13,PAG1,Manes.01G265800,RPT2a,PBA1,RPT3
ath03018	RNA degradation	6	110	1.29	0.00014	RRP4,Manes.12G033800,AT1G60080,Manes.09G039500,Manes.01G181600,RRP41

Tabela suplementar 5A. Termos de GO para processos biológicos associados a proteínas da rede de interações, codificadas por genes diferencialmente hipometilados na variedade sensível Dourada, tendo por base proteínas de *A. thaliana* órtologas às proteínas de *M. esculenta*. As colunas indicam respectivamente: Id do termo GO; descrição do termo; contagem do termo na rede de interações; background de informação para este termo no banco de dados; nível de confiabilidade; taxa de falso descobrimento; nome das proteínas que contém o termo GO

ID do termo	Descrição	Contagem de gene	Background	Strength	False discovery rate	Matching proteins nas redes
GO:0034641	cellular nitrogen compound metabolic process	30	2435	0.67	2.17e-09	Manes.13G145000,PWP2,AT1G23410,AT1G31660,Manes.16G043500,EMB2762,AT2G17360,AT2G20450,Manes.05G087700,SWA1,AT3G04840,AT3G06530,AT3G09630,Manes.10G055800,AT3G11964,Manes.01G020600,MSH7,AT3G49910,MSH6,PMS1,MLH1,Manes.01G182900-CCNGG,MSH3,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,AT5G58420,Manes.03G134200,AT5G67510
GO:0010467	gene expression	21	1235	0.81	4.45e-08	PWP2,AT1G23410,AT1G31660,Manes.16G043500,EMB2762,AT2G17360,AT2G20450,SWA1,AT3G04840,AT3G06530,AT3G09630,Manes.10G055800,AT3G11964,AT3G49910,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,AT5G58420,Manes.03G134200,AT5G67510
GO:0006412	translation	14	437	1.09	6.15e-08	AT1G23410,AT2G17360,AT2G20450,AT3G04840,AT3G09630,Manes.10G055800,AT3G49910,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,AT5G58420,Manes.03G134200,AT5G67510

GO:0006298	mismatch repair	6	18	2.1	1.19e-07	Manes.01G020600,MSH7,MSH6,PMS1,MLH1,MSH3
GO:0043170	macromolecule metabolic process	35	4720	0.45	2.14e-06	Manes.13G145000,Manes.14G074700,PWP2,AT1G23410,AT1G31660,Manes.16G043500,EMB2762,AT2G17360,AT2G20450,Manes.05G087700,Manes.05G061900,SWA1,AT3G04840,AT3G06530,AT3G09630,Manes.10G055800,AT3G11964,Manes.15G099900,Manes.01G020600,MSH7,AT3G49910,Manes.02G085400,MSH6,PMS1,Manes.17G037500,MLH1,Manes.01G182900-CCNGG,MSH3,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,AT5G58420,Manes.03G134200,AT5G67510
GO:0042254	ribosome biogenesis	10	255	1.17	3.25e-06	PWP2,AT1G31660,EMB2762,AT2G20450,SWA1,AT3G06530,AT3G11964,AT3G49910,Manes.03G134200,AT5G67510
GO:0009059	macromolecule biosynthetic process	17	1096	0.77	4.94e-06	AT1G23410,Manes.16G043500,AT2G17360,AT2G20450,AT3G04840,AT3G09630,Manes.10G055800,AT3G49910,Manes.02G085400,MSH3,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,AT5G58420,Manes.03G134200,AT5G67510
GO:0043570	maintenance of DNA repeat elements	4	6	2.4	1.21e-05	Manes.01G020600,MSH7,MSH6,MSH3
GO:0034645	cellular macromolecule biosynthetic process	16	1040	0.77	1.37e-05	AT1G23410,Manes.16G043500,AT2G17360,AT2G20450,AT3G04840,AT3G09630,Manes.10G055800,AT3G49910,MSH3,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,AT5G58420,Manes.03G134200,AT5G67510
GO:0090304	nucleic acid metabolic process	18	1403	0.69	2.11e-05	Manes.13G145000,PWP2,AT1G31660,Manes.16G043500,EMB2762,Manes.05G087700,SWA1,AT3G06530,Manes.10G055800,AT3G11964,Manes.01G020600,MSH7,MSH6,PMS1,MLH1,Manes.01G182900-CCNGG,MSH3,Manes.03G134200
GO:0034314	Arp2/3 complex-mediated actin nucleation	4	10	2.18	4.65e-05	DIS1,Manes.16G060600,ARPC3,ARP2
GO:0044260	cellular macromolecule metabolic process	28	3665	0.46	6.58e-05	Manes.13G145000,Manes.14G074700,AT1G23410,Manes.16G043500,AT2G17360,AT2G20450,Manes.05G087700,Manes.05G061900,AT3G04840,AT3G09630,Manes.10G055800,Manes.15G099900,Manes.01G020600,MSH7,AT3G49910,MSH6,PMS1,Manes.17G037500,MLH1,Manes.01G182900-CCNGG,MSH3,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,AT5G58420,Manes.03G134200,AT5G67510
GO:0006807	nitrogen compound metabolic process	34	5335	0.38	9.93e-05	Manes.13G145000,Manes.14G074700,PWP2,AT1G23410,AT1G31660,Manes.16G043500,EMB2762,AT2G17360,AT2G20450,Manes.05G087700,Manes.05G061900,SWA1,AT3G04840,AT3G06530,AT3G09630,Manes.10G055800,AT3G11964,Manes.15G099900,Manes.01G020600,MSH7,AT3G49910,MSH6,PMS1,Manes.17G037500,MLH1,Manes.01G182900-CCNGG,MSH3,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,AT5G58420,Manes.03G134200,AT5G67510
GO:0044271	cellular nitrogen compound biosynthetic process	15	1068	0.73	0.00010	AT1G23410,Manes.16G043500,AT2G17360,AT2G20450,AT3G04840,AT3G09630,Manes.10G055800,AT3G49910,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,AT5G58420,Manes.03G134200,AT5G67510
GO:0006364	rRNA processing	7	157	1.23	0.00021	PWP2,AT1G31660,EMB2762,SWA1,AT3G06530,AT3G11964,Manes.03G134200
GO:0071840	cellular component organization or biogenesis	22	2482	0.53	0.00021	Manes.07G028400,DIS1,PWP2,Manes.16G060600,AT1G31660,ARPC3,EMB2762,AT2G20450,SWA1,Manes.15G015800,AT3G06530,AT3G11964,Manes.01G020600,MSH7,ARP2,AT3G49910,Manes.02G085400,MSH6,Manes.01G211300,MSH3,Manes.03G134200,AT5G67510
GO:0006281	DNA repair	8	260	1.07	0.00036	Manes.13G145000,Manes.05G087700,Manes.01G020600,MSH7,MSH6,PMS1,MLH1,MSH3

GO:0044238	primary metabolic process	36	6374	0.33	0.00048	Manes.13G145000,Manes.14G074700,PWP2,AT1G23410,AT1G31660,Manes.16G043500,EMB2762,AT2G17360,AT2G20450,Manes.13G106000,Manes.05G087700,Manes.05G061900,SWA1,AT3G04840,AT3G06530,AT3G09630,Manes.10G055800,AT3G11964,Manes.15G099900,Manes.01G020600,MSH7,AT3G49910,Manes.02G085400,MSH6,PMS1,Manes.17G037500,MLH1,Manes.01G182900-CCNGG,MSH3,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,AT5G58420,Manes.03G134200,AT5G67510
GO:0034660	ncRNA metabolic process	8	292	1.02	0.00070	PWP2,AT1G31660,EMB2762,SWA1,AT3G06530,Manes.10G055800,AT3G11964,Manes.03G134200
GO:0000710	meiotic mismatch repair	3	6	2.28	0.00081	MSH7,MSH6,MSH3
GO:0006290	pyrimidine dimer repair	3	6	2.28	0.00081	Manes.01G020600,MSH7,MSH6
GO:0032535	regulation of cellular component size	5	77	1.39	0.0012	DIS1,Manes.16G060600,ARPC3,ARP2,Manes.01G211300
GO:0009987	cellular process	48	11004	0.22	0.0021	Manes.13G145000,Manes.07G028400,Manes.14G074700,DIS1,PWP2,Manes.02G226900,AT1G23410,Manes.16G060600,AT1G31660,ARPC3,Manes.18G023900,Manes.16G043500,EMB2762,AT2G17360,AT2G20450,Manes.13G106000,Manes.03G005100-CCNGG,Manes.05G087700,Manes.12G075100,Manes.05G061900,SWA1,AT3G04840,Manes.15G015800,AT3G06530,AT3G09630,Manes.10G055800,AT3G11964,Manes.15G099900,Manes.01G020600,MSH7,ARP2,AT3G49910,Manes.02G085400,MSH6,PMS1,Manes.01G211300,Manes.17G037500,MLH1,Manes.01G182900-CCNGG,MSH3,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,Manes.14G141800,AT5G58420,Manes.03G134200,AT5G67510
GO:0045910	negative regulation of DNA recombination	3	10	2.06	0.0021	Manes.01G020600,MSH7,MSH6
GO:0048229	gametophyte development	8	366	0.92	0.0024	PWP2,AT1G31660,Manes.15G095300,SWA1,Manes.15G015800,AT3G11964,PMS1,MLH1
GO:0071704	organic substance metabolic process	36	6962	0.29	0.0026	Manes.13G145000,Manes.14G074700,PWP2,AT1G23410,AT1G31660,Manes.16G043500,EMB2762,AT2G17360,AT2G20450,Manes.13G106000,Manes.05G087700,Manes.05G061900,SWA1,AT3G04840,AT3G06530,AT3G09630,Manes.10G055800,AT3G11964,Manes.15G099900,Manes.01G020600,MSH7,AT3G49910,Manes.02G085400,MSH6,PMS1,Manes.17G037500,MLH1,Manes.01G182900-CCNGG,MSH3,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,AT5G58420,Manes.03G134200,AT5G67510
GO:0044237	cellular metabolic process	35	6729	0.3	0.0030	Manes.13G145000,Manes.14G074700,PWP2,AT1G23410,AT1G31660,Manes.16G043500,EMB2762,AT2G17360,AT2G20450,Manes.13G106000,Manes.05G087700,Manes.05G061900,SWA1,AT3G04840,AT3G06530,AT3G09630,Manes.10G055800,AT3G11964,Manes.15G099900,Manes.01G020600,MSH7,AT3G49910,MSH6,PMS1,Manes.17G037500,MLH1,Manes.01G182900-CCNGG,MSH3,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,AT5G58420,Manes.03G134200,AT5G67510
GO:0044085	cellular component biogenesis	12	1028	0.65	0.0048	PWP2,Manes.16G060600,AT1G31660,EMB2762,AT2G20450,SWA1,Manes.15G015800,AT3G06530,AT3G11964,AT3G49910,Manes.03G134200,AT5G67510
GO:0000462	maturation of SSU-rRNA from tricistronic rRNA transcript (SSU-	3	16	1.85	0.0049	PWP2,AT3G06530,Manes.03G134200

GO:0009553	rRNA, 5.8S rRNA, LSU-rRNA) embryo sac development	5	121	1.2	0.0058	PWP2,AT1G31660,SWA1,Manes.15G015800,AT3G11964
GO:0050789	regulation of biological process	26	4671	0.33	0.0217	DIS1,PWP2,Manes.02G226900,Manes.16G060600,Manes.15G095300,ARPC3,Manes.12G009600,Manes.18G023900,Manes.16G043500,Manes.13G106000,Manes.12G075100,Manes.05G026700-CCNGG,SWA1,Manes.15G015800,AT3G06530,Manes.15G099900,Manes.01G020600,MSH7,ARP2,MSH6,Manes.01G211300,Manes.17G037500,Manes.01G182900-CCNGG,MSH3,Manes.05G099700-CCNGG,Manes.14G141800
GO:0044267	cellular protein metabolic process	19	2809	0.41	0.0229	Manes.14G074700,AT1G23410,AT2G17360,AT2G20450,Manes.05G087700,Manes.05G061900,AT3G04840,AT3G09630,Manes.10G055800,Manes.15G099900,AT3G49910,Manes.17G037500,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,AT5G58420,Manes.03G134200,AT5G67510
GO:0045943	positive regulation of transcription by RNA polymerase I	2	4	2.28	0.0252	SWA1,AT3G06530
GO:0050794	regulation of cellular process	24	4214	0.34	0.0294	DIS1,PWP2,Manes.02G226900,Manes.16G060600,Manes.15G095300,ARPC3,Manes.12G009600,Manes.16G043500,Manes.13G106000,Manes.12G075100,Manes.05G026700-CCNGG,SWA1,Manes.15G015800,AT3G06530,Manes.15G099900,Manes.01G020600,MSH7,ARP2,MSH6,Manes.01G211300,Manes.17G037500,MSH3,Manes.05G099700-CCNGG,Manes.14G141800
GO:0048522	positive regulation of cellular process	10	930	0.61	0.0373	DIS1,Manes.16G060600,ARPC3,Manes.12G009600,SWA1,Manes.15G015800,AT3G06530,ARP2,Manes.01G211300,Manes.05G099700-CCNGG
GO:0048518	positive regulation of biological process	11	1146	0.56	0.0440	DIS1,Manes.16G060600,ARPC3,Manes.12G009600,Manes.18G023900,SWA1,Manes.15G015800,AT3G06530,ARP2,Manes.01G211300,Manes.05G099700-CCNGG

Tabela suplementar 5B. Termos de GO para função molecular para proteínas da rede de interações formada por sítios diferencialmente hipometilados para a variedade sensível Dourada baseados em proteínas de *A. thaliana* órtologas às proteínas de *M. esculenta*. As colunas indicam respectivamente: Id do termo GO; descrição do termo; contagem do termo na rede de interações; background de informação para este termo no banco de dados; nível de confiabilidade; taxa de falso descobrimento; nome das proteínas que contém o termo GO

ID do termo	Descrição	Contagem de gene	Background	Strength	False discovery rate	Matching proteins nas redes
GO:0003735	structural constituent of ribosome	13	308	1.21	1.18e-08	AT1G23410,AT2G17360,AT2G20450,AT3G04840,AT3G09630,AT3G49910,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,AT5G58420,Manes.03G134200,AT5G67510
GO:0003676	nucleic acid binding	30	2854	0.6	1.29e-08	Manes.13G145000,Manes.02G226900,AT1G31660,Manes.12G009600,Manes.16G043500,AT2G17360,AT2G20450,Manes.05G087700,Manes.05G026700-

GO:0005198	structural molecule activity	14	412	1.11	1.29e-08	CCNGG,AT3G04840,Manes.15G015800,AT3G06530,AT3G09630,Manes.10G055800,AT3G11964,Manes.01G020600,MSH7,AT3G49910,MSH6,PMS1,MLH1,Manes.01G182900-CCNGG,MSH3,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,Manes.14G141800,AT5G58420,AT5G67510
GO:0030983	mismatched DNA binding	6	13	2.24	1.29e-08	AT1G23410,Manes.16G060600,AT2G17360,AT2G20450,AT3G04840,AT3G09630,AT3G49910,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,AT5G58420,Manes.03G134200,AT5G67510
GO:0003723	RNA binding	16	954	0.81	3.45e-06	Manes.01G020600,MSH7,MSH6,PMS1,MLH1,MSH3
GO:0097159	organic cyclic compound binding	36	5434	0.4	8.42e-06	AT1G31660,AT2G17360,AT2G20450,AT3G04840,AT3G06530,AT3G09630,Manes.10G055800,AT3G11964,AT3G49910,Manes.01G182900-CCNGG,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,AT5G58420,AT5G67510
GO:1901363	heterocyclic compound binding	36	5413	0.4	8.42e-06	Manes.13G145000,Manes.07G028400,DIS1,Manes.02G226900,AT1G31660,Manes.02G005700,Manes.12G009600,Manes.16G043500,AT2G17360,AT2G20450,Manes.03G005100-CCNGG,Manes.05G087700,Manes.05G026700-CCNGG,AT3G04840,Manes.15G015800,AT3G06530,AT3G09630,Manes.10G055800,AT3G11964,Manes.15G099900,Manes.01G020600,MSH7,ARP2,AT3G49910,MSH6,PMS1,MLH1,Manes.01G182900-CCNGG,MSH3,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,Manes.14G141800,AT5G58420,AT5G67510
GO:0016462	pyrophosphatase activity	12	709	0.81	0.00023	Manes.13G145000,Manes.07G028400,DIS1,Manes.02G226900,AT1G31660,Manes.02G005700,Manes.12G009600,Manes.16G043500,AT2G17360,AT2G20450,Manes.03G005100-CCNGG,Manes.05G087700,Manes.05G026700-CCNGG,AT3G04840,Manes.15G015800,AT3G06530,AT3G09630,Manes.10G055800,AT3G11964,Manes.15G099900,Manes.01G020600,MSH7,ARP2,AT3G49910,MSH6,PMS1,MLH1,Manes.01G182900-CCNGG,MSH3,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,Manes.14G141800,AT5G58420,AT5G67510
GO:0016887	ATPase activity	10	510	0.87	0.00048	Manes.13G145000,Manes.07G028400,Manes.14G074700,Manes.02G005700,Manes.15G015800,Manes.01G020600,MSH7,MSH6,PMS1,MLH1,MSH3
GO:0005488	binding	42	8453	0.28	0.00058	Manes.13G145000,Manes.07G028400,Manes.14G074700,DIS1,Manes.02G226900,AT1G23410,Manes.16G060600,AT1G31660,ARPC3,Manes.02G005700,Manes.12G009600,Manes.16G043500,AT2G17360,AT2G20450,Manes.03G005100-CCNGG,Manes.05G087700,Manes.05G026700-CCNGG,AT3G04840,Manes.15G015800,AT3G06530,AT3G09630,Manes.10G055800,AT3G11964,Manes.15G099900,Manes.01G020600,MSH7,ARP2,AT3G49910,MSH6,PMS1,Manes.01G211300,Manes.17G037500,MLH1,Manes.01G182900-CCNGG,MSH3,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,Manes.14G141800,AT5G58420,AT5G67510
GO:0017111	nucleoside-triphosphatase activity	11	666	0.8	0.00058	Manes.13G145000,Manes.07G028400,Manes.14G074700,Manes.02G005700,Manes.15G015800,Manes.01G020600,MSH7,MSH6,PMS1,MLH1,MSH3
GO:0008094	DNA-dependent ATPase activity	5	76	1.4	0.00088	Manes.15G015800,Manes.01G020600,MSH7,MSH6,MSH3
GO:0000400	four-way junction DNA binding	2	4	2.28	0.0339	MSH7,MSH6
GO:0003684	damaged DNA binding	3	31	1.57	0.0339	Manes.01G020600,MSH6,MSH3

GO:0000217	DNA secondary structure binding	3	35	1.51	0.0418	Manes.01G020600,MSH7,MSH6
GO:0003677	DNA binding	15	1956	0.47	0.0439	Manes.13G145000,Manes.02G226900,Manes.12G009600,Manes.16G043500,Manes.05G087700,Manes.05G026700-CCNGG,Manes.15G015800,Manes.10G055800,Manes.01G020600,MSH7,MSH6,PMS1,MLH1,MSH3,Manes.14G141800

Tabela suplementar 5C. Termos de GO para componente celular para proteínas da rede de interações formada por sítios diferencialmente hipometilados para a variedade sensível Dourada baseados em proteínas de *A. thaliana* órtologas às proteínas de *M. esculenta*. As colunas indicam respectivamente: Id do termo GO; descrição do termo; contagem do termo na rede de interações; background de informação para este termo no banco de dados; nível de confiabilidade; taxa de falso descobrimento; nome das proteínas que contém o termo GO

ID do termo	Descrição	Contagem de gene	Background	Strength	False discovery rate	Matching proteins nas redes
GO:0032991	protein-containing complex	31	2226	0.72	2.36e-12	Manes.14G074700,DIS1,PWP2,AT1G23410,Manes.16G060600,AT1G31660,ARPC3,EMB2762,AT2G17360,AT2G20450,SWA1,AT3G04840,AT3G06530,AT3G09630,AT3G11964,Manes.01G020600,MSH7,ARP2,AT3G49910,MSH6,PMS1,Manes.17G037500,MLH1,MSH3,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,AT5G58420,Manes.03G134200,AT5G67510
GO:0043232	intracellular non-membrane-bounded organelle	26	1407	0.85	2.36e-12	Manes.07G028400,DIS1,PWP2,AT1G23410,Manes.16G060600,AT1G31660,ARPC3,EMB2762,AT2G17360,AT2G20450,Manes.05G087700,SWA1,AT3G04840,AT3G06530,AT3G09630,AT3G11964,ARP2,AT3G49910,MLH1,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,AT5G58420,Manes.03G134200,AT5G67510
GO:1990904	ribonucleoprotein complex	18	509	1.13	2.36e-12	PWP2,AT1G23410,AT1G31660,EMB2762,AT2G17360,AT2G20450,AT3G04840,AT3G06530,AT3G09630,AT3G11964,AT3G49910,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,AT5G58420,Manes.03G134200,AT5G67510
GO:0044391	ribosomal subunit	13	262	1.28	1.56e-10	AT1G23410,AT2G17360,AT2G20450,AT3G04840,AT3G09630,AT3G49910,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,AT5G58420,Manes.03G134200,AT5G67510
GO:0022626	cytosolic ribosome mismatch repair complex	13	280	1.25	2.91e-10	AT1G23410,AT2G17360,AT2G20450,AT3G04840,AT3G09630,AT3G49910,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,AT5G58420,Manes.03G134200,AT5G67510
GO:0032300	cytosolic small ribosomal subunit	6	8	2.46	2.91e-10	Manes.01G020600,MSH7,MSH6,PMS1,MLH1,MSH3
GO:0022627		8	90	1.53	4.76e-08	AT1G23410,AT2G17360,AT3G04840,RPS6A,EMB3010,Manes.05G099700-CCNGG,AT5G58420,Manes.03G134200

GO:0005730	nucleolus	12	373	1.09	7.75e-08	PWP2,AT1G31660,EMB2762,Manes.05G087700,SWA1,AT3G06530,AT3G09630,AT3G11964,AT3G49910,RPS6A,AT5G02870,AT5G58420
GO:0042788	polysomal ribosome	7	94	1.45	1.68e-06	AT3G04840,AT3G09630,AT3G49910,RPS6A,AT5G02870,EMB3010,AT5G67510
GO:0005885	Arp2/3 protein complex	4	10	2.18	6.96e-06	DIS1,Manes.16G060600,ARPC3,ARP2
GO:0030684	preribosome	5	33	1.76	7.72e-06	PWP2,AT1G31660,EMB2762,AT3G06530,AT3G11964 Manes.07G028400,Manes.14G074700,DIS1,PWP2,Manes.02G226900,AT1G23410,Manes.16G060600,AT1G31660,Manes.15G095300,ARPC3,Manes.02G005700,Manes.12G009600,Manes.18G023900,EMB2762,AT2G17360,AT2G20450,Manes.13G106000,Manes.05G087700,Manes.05G026700-CCNGG,SWA1,AT3G04840,Manes.15G015800,AT3G06530,AT3G09630,Manes.01G131400,AT3G11964,Manes.15G099900,Manes.01G020600,MSH7,ARP2,AT3G49910,Manes.02G085400,MSH6,PM S1,Manes.01G211300,Manes.17G037500,MLH1,Manes.18G100900,Manes.01G182900-CCNGG,MSH3,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,Manes.14G141800,AT5G58420,Manes.03G134200,AT5G67510
GO:0043229	intracellular organelle	48	9831	0.27	1.61e-05	PWP2,AT1G31660,EMB2762,Manes.05G087700,SWA1,AT3G06530,AT3G09630,AT3G11964,AT3G49910,MLH1,RPS6A,AT5G02870,AT5G58420
GO:0031981	nuclear lumen	13	807	0.79	2.44e-05	PWP2,EMB2762,AT3G06530,AT3G11964
GO:0032040	small-subunit processome	4	18	1.93	3.75e-05	Manes.01G020600,MSH7,MSH6
GO:0032301	MutSalpha complex	3	3	2.58	3.95e-05	PWP2,AT1G31660,EMB2762,Manes.05G087700,SWA1,AT3G06530,AT3G09630,AT3G11964,AT3G49910,Manes.01G211300,MLH1,RPS6A,AT5G02870,AT5G58420 Manes.07G028400,DIS1,AT1G23410,Manes.15G095300,Manes.02G005700,Manes.12G009600,AT2G17360,AT2G20450,Manes.05G026700-CCNGG,AT3G04840,AT3G09630,Manes.01G131400,Manes.10G055800,AT3G49910,Manes.01G182900-CCNGG,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,AT5G58420,Manes.03G134200,AT5G67510 Manes.07G028400,Manes.14G074700,DIS1,PWP2,Manes.02G226900,AT1G23410,Manes.16G060600,AT1G31660,Manes.15G095300,ARPC3,Manes.02G005700,Manes.12G009600,Manes.18G023900,EMB2762,AT2G17360,AT2G20450,Manes.13G106000,Manes.05G087700,Manes.05G026700-CCNGG,SWA1,AT3G04840,Manes.15G015800,AT3G06530,AT3G09630,Manes.01G131400,Manes.10G055800,AT3G11964,Manes.15G099900,Manes.01G020600,MSH7,ARP2,AT3G49910,Manes.02G085400,MSH6,PMS1,Manes.01G211300,Manes.17G037500,MLH1,Manes.18G100900,Manes.01G182900-CCNGG,MSH3,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,Manes.14G141800,AT5G58420,Manes.03G134200,AT5G67510
GO:0070013	intracellular organelle lumen	14	1004	0.73	3.95e-05	PWP2,EMB2762,AT3G06530
GO:0005829	cytosol	22	2567	0.51	4.12e-05	Manes.07G028400,AT1G23410,AT2G20450,Manes.12G075100,AT3G04840,AT3G09630,Manes.10G055800,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG
GO:0005622	intracellular	49	10863	0.23	9.84e-05	PWP2,Manes.02G226900,AT1G23410,AT1G31660,Manes.15G095300,Manes.02G005700,Manes.12G009600,EMB2762,Manes.05G087700,Manes.05G026700-CCNGG,SWA1,Manes.15G015800,AT3G06530,AT3G09630,AT3G11964,Manes.15G099900,Manes.
GO:0030686	90S preribosome	3	7	2.21	0.00019	
GO:0009506	plasmodesma	11	741	0.75	0.00035	
GO:0005634	nucleus	28	4669	0.36	0.00068	

						01G020600,MSH7,AT3G49910,MSH6,PMS1,MLH1,Manes.01G182900-CCNGG,MSH3,RPS6A,AT5G02870,Manes.14G141800,AT5G58420
GO:0022625	cytosolic large ribosomal subunit	5	126	1.18	0.0018	AT2G20450,AT3G09630,AT3G49910,AT5G02870,AT5G67510
GO:0032302	MutSbeta complex	2	2	2.58	0.0028	Manes.01G020600,MSH3
GO:0032389	MutLalpha complex	2	2	2.58	0.0028	PMS1,MLH1
GO:0030688	preribosome, small subunit precursor	2	3	2.4	0.0042	AT1G31660,EMB2762
GO:0043231	intracellular membrane-bounded organelle	41	9376	0.22	0.0042	Manes.07G028400,Manes.14G074700,PWP2,Manes.02G226900,AT1G23410,AT1G31660,Manes.15G095300,Manes.02G005700,Manes.12G009600,Manes.18G023900,EMB2762,AT2G17360,AT2G20450,Manes.13G106000,Manes.05G087700,Manes.05G026700-CCNGG,SWA1,Manes.15G015800,AT3G06530,AT3G09630,Manes.01G131400,AT3G11964,Manes.15G099900,Manes.01G020600,MSH7,AT3G49910,Manes.02G085400,MSH6,PMS1,Manes.01G211300,Manes.17G037500,MLH1,Manes.18G100900,Manes.01G182900-CCNGG,MSH3,RPS6A,AT5G02870,Manes.05G099700-CCNGG,Manes.14G141800,AT5G58420,AT5G67510
GO:0110165	cellular anatomical entity	51	13776	0.15	0.0186	Manes.07G028400,Manes.14G074700,DIS1,PWP2,Manes.02G226900,AT1G23410,Manes.16G060600,AT1G31660,Manes.15G095300,ARPC3,Manes.02G005700,Manes.12G009600,Manes.18G023900,EMB2762,AT2G17360,AT2G20450,Manes.13G106000,Manes.03G005100-CCNGG,Manes.05G087700,Manes.12G075100,Manes.05G026700-CCNGG,SWA1,AT3G04840,Manes.15G015800,AT3G06530,AT3G09630,Manes.01G131400,Manes.10G055800,AT3G11964,Manes.15G099900,Manes.01G020600,MSH7,ARP2,AT3G49910,Manes.02G085400,MSH6,PMS1,Manes.01G211300,Manes.17G037500,MLH1,Manes.18G100900,Manes.01G182900-CCNGG,MSH3,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,Manes.14G141800,AT5G58420,Manes.03G134200,AT5G67510
GO:0005737	cytoplasm	34	7911	0.21	0.0429	Manes.07G028400,Manes.14G074700,DIS1,AT1G23410,Manes.16G060600,AT1G31660,Manes.15G095300,ARPC3,Manes.02G005700,Manes.12G009600,Manes.18G023900,AT2G17360,AT2G20450,Manes.13G106000,Manes.05G026700-CCNGG,AT3G04840,AT3G09630,Manes.01G131400,Manes.10G055800,Manes.15G099900,ARP2,AT3G49910,Manes.02G085400,Manes.01G211300,Manes.17G037500,Manes.18G100900,Manes.01G182900-CCNGG,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,AT5G58420,Manes.03G134200,AT5G67510

Tabela suplementar 5D. Termos KEGG para proteínas da rede de interações formada por sítios diferencialmente hipometilados para a variedade sensível Dourada baseados em proteínas de *A. thaliana* órtologas às proteínas de *M. esculenta*. As colunas indicam

respectivamente: Id do termo; descrição do termo; contagem do termo na rede de interações; background de informação para este termo no banco de dados; nível de confiabilidade; taxa de falso descobrimento; nome das proteínas que contém o termo

ID do termo	Descrição	Contagem de gene	Background	Strength	False discovery rate	Matching proteins nas redes
ath03010	Ribosome	13	318	1.19	1.51e-09	AT1G23410,AT2G17360,AT2G20450,AT3G04840,AT3G09630,AT3G49910,RPS6A,AT5G02870,EMB3010,Manes.05G099700-CCNGG,AT5G58420,Manes.03G134200,AT5G67510
ath03430	Mismatch repair	6	36	1.8	2.89e-07	Manes.01G020600,MSH7,MSH6,PMS1,MLH1,MSH3
ath03008	Ribosome biogenesis in eukaryotes	5	90	1.33	0.00078	PWP2,IMP4,AT1G63810,SWA1,AT3G06530

Tabela suplementar 6A. Termos de GO para processos biológicos, associados a proteínas da rede de interações codificadas por genes diferencialmente hipermetilados na variedade sensível Dourada, tendo por base proteínas de *A. thaliana* órtologas às proteínas de *M. esculenta*. As colunas indicam respectivamente: Id do termo GO; descrição do termo; contagem do termo na rede de interações; background de informação para este termo no banco de dados; nível de confiabilidade; taxa de falso descobrimento; nome das proteínas que contém o termo GO

ID do termo	Descrição	Contagem de gene	Background	Strength	False discovery rate	Matching proteins nas redes
GO:0006270	DNA replication initiation	12	17	2.31	1.21e-18	AT1G07270,ORC6,MCM2,ORC4,MCM5,MCM9,MCM4,CDC6,PRL,ORC5,MCM6,Manes.17G068400
GO:0006260	DNA replication	15	112	1.59	2.79e-15	AT1G07270,ORC6,MCM2,ORC4,MCM5,MCM9,MCM4,CDC6,ORC2,ETG1,PSF2,PRL,ORC5,MCM6,Manes.17G068400
GO:0006261	DNA-dependent DNA replication double-strand break repair via break-induced replication	13	86	1.64	1.20e-13	AT1G07270,ORC6,MCM2,ORC4,MCM5,MCM9,MCM4,CDC6,ETG1,PRL,ORC5,MCM6,Manes.17G068400
GO:0000727	DNA metabolic process	7	8	2.4	1.29e-10	MCM2,MCM5,MCM4,PSF2,PRL,MCM6,Manes.17G068400
GO:0006259	nucleic acid metabolic process	16	378	1.09	1.60e-09	Manes.12G035400,AT1G07270,ORC6,MCM2,ORC4,MCM5,MCM9,MCM4,CDC6,ETG1,PSF2,AT3G46210,PRL,ORC5,MCM6,Manes.17G068400
GO:0090304		26	1403	0.73	4.39e-09	RRP4,Manes.12G035400,AT1G07270,ORC6,CPSF30,MCM2,Manes.10G123600,ORC4,CPSF73-II,MCM5,MCM9,MCM4,PAPS2,CDC6,ETG1,PSF2,AT3G46210,EIF3E,Manes.01G181600,RRP41,PRL,ORC5,CPSF100,Manes.08G020300,MCM6,Manes.17G068400

GO:0006267	pre-replicative complex assembly involved in nuclear cell cycle DNA replication	6	7	2.39	6.08e-09	MCM2,MCM5,MCM4,PRL,MCM6,Manes.17G068400
GO:0000724	double-strand break repair via homologous recombination	9	67	1.59	9.78e-09	Manes.12G035400,MCM2,MCM5,MCM9,MCM4,PSF2,PRL,MCM6,Manes.17G068400
GO:0016180	snRNA processing	6	9	2.28	1.12e-08	RRP4,Manes.10G123600,CPSF73-II,AT3G46210,Manes.01G181600,RRP41
GO:0034645	cellular macromolecule biosynthetic process	22	1040	0.79	1.23e-08	AT1G07270,TIF3H1,ORC6,MCM2,ORC4,MCM5,MCM9,MCM4,CDC6,ORC2,EIF2,ETG1,PSF2,Manes.11G109700,Manes.10G071900-CCNKG,EIF3E,PRL,EIF3A,ORC5,Manes.12G035500,MCM6,Manes.17G068400
GO:0031123	RNA 3-end processing	8	50	1.66	2.89e-08	RRP4,CPSF30,Manes.10G123600,PAPS2,AT3G46210,Manes.01G181600,RRP41,CPSF100
GO:0032508	DNA duplex unwinding	8	52	1.65	3.41e-08	MCM2,MCM5,MCM9,MCM4,PSF2,PRL,MCM6,Manes.17G068400
GO:0044786	cell cycle DNA replication	7	30	1.83	4.08e-08	MCM2,MCM5,MCM4,CDC6,PRL,MCM6,Manes.17G068400
GO:0006268	DNA unwinding involved in DNA replication	6	18	1.98	1.41e-07	MCM2,MCM5,MCM4,PRL,MCM6,Manes.17G068400
GO:0034622	cellular protein-containing complex assembly	13	350	1.03	2.85e-07	Manes.12G035400,TIF3H1,MCM2,MCM5,MCM4,Manes.11G140100,EIF2,EIF3E,PRL,EIF3A,Manes.12G136200,MCM6,Manes.17G068400
GO:0034641	cellular nitrogen compound metabolic process	30	2435	0.55	3.13e-07	RRP4,Manes.12G035400,AT1G07270,TIF3H1,ORC6,CPSF30,MCM2,Manes.10G123600,ORC4,CPSF73-II,MCM5,MCM9,MCM4,PAPS2,CDC6,EIF2,ETG1,PSF2,AT3G46210,EIF3E,Manes.01G181600,RRP41,PRL,EIF3A,ORC5,CPSF100,Manes.08G020300,Manes.13G036700,MCM6,Manes.17G068400
GO:0006725	cellular aromatic compound metabolic process	28	2137	0.58	3.54e-07	RRP4,Manes.12G035400,AT1G07270,ORC6,CPSF30,MCM2,Manes.10G123600,ORC4,CPSF73-II,MCM5,MCM9,MCM4,PAPS2,CDC6,ETG1,Manes.12G070400-CCNKG,PSF2,AT3G46210,EIF3E,Manes.01G181600,RRP41,PRL,ORC5,CPSF100,Manes.08G020300,Manes.13G036700,MCM6,Manes.17G068400
GO:0046483	heterocycle metabolic process	27	2022	0.59	5.06e-07	RRP4,Manes.12G035400,AT1G07270,ORC6,CPSF30,MCM2,Manes.10G123600,ORC4,CPSF73-II,MCM5,MCM9,MCM4,PAPS2,CDC6,ETG1,PSF2,AT3G46210,EIF3E,Manes.01G181600,RRP41,PRL,ORC5,CPSF100,Manes.08G020300,Manes.13G036700,MCM6,Manes.17G068400
GO:1901360	organic cyclic compound metabolic process	28	2229	0.56	8.40e-07	RRP4,Manes.12G035400,AT1G07270,ORC6,CPSF30,MCM2,Manes.10G123600,ORC4,CPSF73-II,MCM5,MCM9,MCM4,PAPS2,CDC6,ETG1,Manes.12G070400-CCNKG,PSF2,AT3G46210,EIF3E,Manes.01G181600,RRP41,PRL,ORC5,CPSF100,Manes.08G020300,Manes.13G036700,MCM6,Manes.17G068400
GO:0043170	macromolecule metabolic process	41	4720	0.4	1.95e-06	RRP4,Manes.12G035400,AT1G07270,TIF3H1,ORC6,CPSF30,MCM2,Manes.06G013000,Manes.10G123600,ORC4,CPSF73-II,MCM5,MCM9,MCM4,PAPS2,Manes.03G049000,CDC6,ORC2,EIF2,ETG1,Manes.12G051300,PSF2,Manes.05G206400,Manes.06G008000,Manes.11G109700,Manes.10G071900-

GO:0065004	protein-DNA complex assembly	7	71	1.45	5.54e-06	CCNCGG,AT3G46210,Manes.09G098300,EIF3E,Manes.01G181600,RRP41,PRL,EIF3A,ORC5,Manes.04G152900,CPSF100,Manes.06G151900,Manes.08G020300,Manes.12G035500,MCM6,Manes.17G068400
						Manes.12G035400,MCM2,MCM5,MCM4,PRL,MCM6,Manes.17G068400
						RRP4,Manes.12G035400,AT1G07270,TIF3H1,ORC6,CPSF30,MCM2,Manes.06G013000,Manes.10G123600,ORC4,CPSF73-II,MCM5,MCM9,MCM4,PAPS2,Manes.03G049000,CDC6,EIF2,ETG1,Manes.12G070400-CCNCGG,Manes.12G051300,PSF2,Manes.05G206400,Manes.06G008000,Manes.11G109700,Manes.10G071900-
GO:0006807	nitrogen compound metabolic process	43	5335	0.37	5.93e-06	CCNCGG,AT3G46210,Manes.09G098300,EIF3E,Manes.01G181600,RRP41,PRL,EIF3A,ORC5,Manes.04G152900,CPSF100,Manes.06G151900,Manes.10G149600,Manes.08G020300,Manes.13G036700,MCM6,Manes.17G068400,Manes.10G059000-CCNCGG
						RRP4,Manes.12G035400,TIF3H1,MCM2,MCM5,MCM4,Manes.11G140100,EIF2,AT3G46210,EIF3E,Manes.01G181600,RRP41,PRL,EIF3A,Manes.12G136200,Manes.12G035500,MCM6,Manes.17G068400
GO:0044085	cellular component biogenesis	18	1028	0.7	8.36e-06	
GO:0034475	U4 snRNA 3-end processing	4	6	2.28	1.34e-05	RRP4,AT3G46210,Manes.01G181600,RRP41
GO:1902969	mitotic DNA replication	4	6	2.28	1.34e-05	MCM2,MCM4,MCM6,Manes.17G068400
GO:0006281	DNA repair	10	260	1.05	1.49e-05	Manes.12G035400,MCM2,MCM5,MCM9,MCM4,ETG1,PSF2,PRL,MCM6,Manes.17G068400
						AT1G07270,TIF3H1,ORC6,MCM2,ORC4,MCM5,MCM9,MCM4,Manes.06G051100,CDC6,ORC2,EIF2,ETG1,Manes.12G070400-CCNCGG,PSF2,Manes.11G109700,Manes.10G071900-CCNCGG,EIF3E,PRL,EIF3A,ORC5,Manes.02G135100,Manes.12G035500,Manes.13G036700,MCM6,Manes.17G068400
GO:0044249	cellular biosynthetic process	26	2275	0.52	1.79e-05	
						AT1G07270,TIF3H1,ORC6,MCM2,ORC4,MCM5,MCM9,MCM4,Manes.06G051100,CDC6,ORC2,EIF2,ETG1,Manes.12G070400-CCNCGG,PSF2,Manes.11G109700,Manes.10G071900-CCNCGG,EIF3E,PRL,EIF3A,ORC5,Manes.02G135100,Manes.12G035500,Manes.13G036700,MCM6,Manes.17G068400
GO:1901576	organic substance biosynthetic process	26	2363	0.5	3.54e-05	
						AT1G07270,TIF3H1,ORC6,MCM2,ORC4,MCM5,MCM9,MCM4,Manes.06G051100,CDC6,ORC2,EIF2,ETG1,Manes.12G070400-CCNCGG,PSF2,Manes.11G109700,Manes.10G071900-CCNCGG,EIF3E,PRL,EIF3A,ORC5,Manes.02G135100,Manes.12G035500,Manes.13G036700,MCM6,Manes.17G068400
GO:0034427	nuclear-transcribed mRNA catabolic process, exonucleolytic, 3-5	4	9	2.11	3.63e-05	RRP4,AT3G46210,Manes.01G181600,RRP41
						RRP4,Manes.12G035400,AT1G07270,TIF3H1,ORC6,MCM2,Manes.06G013000,ORC4,MCM5,MCM9,MCM4,CDC6,ORC2,EIF2,ETG1,Manes.12G051300,PSF2,Manes.06G008000,Manes.11G109700,Manes.10G071900-
GO:0044260	cellular macromolecule metabolic process	33	3665	0.41	4.23e-05	CCNCGG,AT3G46210,Manes.09G098300,EIF3E,Manes.01G181600,RRP41,PRL,EIF3A,ORC5,Manes.04G152900,Manes.06G151900,Manes.12G035500,MCM6,Manes.17G068400
GO:0071028	nuclear mRNA surveillance	4	10	2.06	4.85e-05	RRP4,AT3G46210,Manes.01G181600,RRP41
GO:0001732	formation of cytoplasmic translation initiation complex	4	12	1.98	8.33e-05	TIF3H1,EIF2,EIF3E,EIF3A

GO:0044238	primary metabolic process	45	6374	0.31	8.33e-05	RRP4,Manes.12G035400,AT1G07270,TIF3H1,ORC6,CPSF30,MCM2,Manes.06G013000,Manes.10G123600,ORC4,CPSF73-II,MCM5,MCM9,MCM4,PAPS2,Manes.03G049000,Manes.06G051100,CDC6,EIF2,ETG1,Manes.12G070400-CCNGG,Manes.12G051300,PSF2,Manes.05G206400,Manes.06G008000,Manes.11G109700,Manes.10G071900-CCNGG,AT3G46210,Manes.09G098300,EIF3E,Manes.01G181600,RRP41,PRL,EIF3A,ORC5,Manes.04G152900,Manes.02G135100,CPSF100,Manes.06G151900,Manes.10G149600,Manes.08G020300,Manes.12G035500,MCM6,Manes.17G068400,Manes.10G059000-CCNGG
GO:0071840	cellular component organization or biogenesis	26	2482	0.48	8.33e-05	RRP4,Manes.12G035400,TIF3H1,Manes.02G025500,MCM2,CPSF73-II,MCM5,MCM9,MCM4,Manes.11G140100,EIF2,ETG1,PSF2,Manes.10G071900-CCNGG,AT3G46210,EIF3E,Manes.01G181600,RRP41,PRL,EIF3A,Manes.12G136200,Manes.04G152900,Manes.12G035500,Manes.13G036700,MCM6,Manes.17G068400
GO:0022402	cell cycle process	10	333	0.94	0.00010	AT1G07270,MCM2,MCM5,MCM4,CDC6,ETG1,PRL,Manes.12G035500,MCM6,Manes.17G068400
GO:0071704	organic substance metabolic process	47	6962	0.29	0.00013	RRP4,Manes.12G035400,AT1G07270,TIF3H1,ORC6,CPSF30,MCM2,Manes.06G013000,Manes.10G123600,ORC4,CPSF73-II,MCM5,MCM9,MCM4,PAPS2,Manes.03G049000,Manes.06G051100,CDC6,ORC2,EIF2,ETG1,Manes.12G070400-CCNGG,Manes.12G051300,PSF2,Manes.05G206400,Manes.06G008000,Manes.11G109700,Manes.10G071900-CCNGG,AT3G46210,Manes.09G098300,EIF3E,Manes.01G181600,RRP41,PRL,EIF3A,ORC5,Manes.04G152900,Manes.02G135100,CPSF100,Manes.06G151900,Manes.10G149600,Manes.08G020300,Manes.12G035500,Manes.13G036700,MCM6,Manes.17G068400,Manes.10G059000-CCNGG
GO:0016075	rRNA catabolic process	4	16	1.86	0.00019	RRP4,AT3G46210,Manes.01G181600,RRP41
GO:0006378	mRNA polyadenylation	4	17	1.83	0.00023	CPSF30,Manes.10G123600,PAPS2,CPSF100
GO:1903047	mitotic cell cycle process	7	148	1.14	0.00032	AT1G07270,MCM2,MCM4,CDC6,Manes.12G035500,MCM6,Manes.17G068400
GO:0007049	cell cycle	11	489	0.81	0.00034	AT1G07270,MCM2,MCM5,MCM9,MCM4,CDC6,ETG1,PRL,Manes.12G035500,MCM6,Manes.17G068400
GO:0044237	cellular metabolic process	45	6729	0.29	0.00034	RRP4,Manes.12G035400,AT1G07270,TIF3H1,ORC6,CPSF30,MCM2,Manes.06G013000,Manes.10G123600,ORC4,CPSF73-II,MCM5,MCM9,MCM4,PAPS2,Manes.06G051100,CDC6,ORC2,EIF2,ETG1,Manes.12G070400-CCNGG,Manes.12G051300,PSF2,Manes.06G008000,Manes.11G109700,Manes.10G071900-CCNGG,AT3G46210,Manes.09G098300,EIF3E,Manes.01G181600,RRP41,PRL,EIF3A,ORC5,Manes.04G152900,Manes.02G135100,CPSF100,Manes.06G151900,Manes.10G149600,Manes.08G020300,Manes.12G035500,Manes.13G036700,MCM6,Manes.17G068400,Manes.10G059000-CCNGG
GO:0051276	chromosome organization	11	487	0.81	0.00034	Manes.12G035400,MCM2,MCM5,MCM9,MCM4,ETG1,PSF2,PRL,Manes.04G152900,MCM6,Manes.17G068400
GO:1902975	mitotic DNA replication initiation	3	4	2.34	0.00036	MCM2,MCM4,Manes.17G068400
GO:0016071	mRNA metabolic process	10	408	0.85	0.00047	RRP4,CPSF30,Manes.10G123600,CPSF73-II,PAPS2,AT3G46210,Manes.01G181600,RRP41,CPSF100,Manes.08G020300

GO:0090501	RNA phosphodiester bond hydrolysis	6	100	1.24	0.00047	RRP4,CPSF30,Manes.10G123600,AT3G46210,Manes.01G181600,CPSF100
GO:0071051	polyadenylation-dependent snoRNA 3-end processing	3	7	2.09	0.0011	RRP4,AT3G46210,RRP41
GO:0090305	nucleic acid phosphodiester bond hydrolysis	7	191	1.02	0.0013	RRP4,CPSF30,Manes.10G123600,CPSF73-II,AT3G46210,Manes.01G181600,CPSF100
GO:0016043	cellular component organization	22	2264	0.45	0.0019	Manes.12G035400,TIF3H1,Manes.02G025500,MCM2,CPSF73-II,MCM5,MCM9,MCM4,Manes.11G140100,EIF2,ETG1,PSF2,Manes.10G071900-CCNGG,EIF3E,PRL,EIF3A,Manes.12G136200,Manes.04G152900,Manes.12G035500,Manes.13G036700,MCM6,Manes.17G068400
GO:0022613	ribonucleoprotein complex biogenesis	8	295	0.89	0.0024	RRP4,TIF3H1,EIF2,AT3G46210,EIF3E,Manes.01G181600,RRP41,EIF3A
GO:0006271	DNA strand elongation involved in DNA replication	3	12	1.86	0.0039	MCM4,PRL,Manes.17G068400
GO:0006396	RNA processing	10	582	0.7	0.0080	RRP4,CPSF30,Manes.10G123600,CPSF73-II,PAPS2,AT3G46210,Manes.01G181600,RRP41,CPSF100,Manes.08G020300
GO:0006398	mRNA 3-end processing by stem-loop binding and cleavage	2	2	2.46	0.0142	Manes.10G123600,CPSF100
GO:0010467	gene expression	14	1235	0.51	0.0190	RRP4,TIF3H1,CPSF30,Manes.10G123600,CPSF73-II,PAPS2,EIF2,AT3G46210,EIF3E,Manes.01G181600,RRP41,EIF3A,CPSF100,Manes.08G020300
GO:0033554	cellular response to stress	12	956	0.56	0.0242	Manes.12G035400,CPSF30,MCM2,MCM5,MCM9,MCM4,ETG1,PSF2,PRL,Manes.13G036700,MCM6,Manes.17G068400
GO:0009987	cellular process	56	11004	0.17	0.0322	RRP4,Manes.12G035400,AT1G07270,TIF3H1,Manes.02G025500,ORC6,Manes.18G035000,CPSF30,MCM2,Manes.18G037900,Manes.06G013000,Manes.10G123600,Manes.05G118800,ORC4,CPSF73-II,MCM5,MCM9,MCM4,PAPS2,Manes.06G051100,Manes.11G140100,CDC6,ORC2,EIF2,ETG1,Manes.08G059900-CCNGG,Manes.12G070400-CCNGG,Manes.12G051300,PSF2,Manes.06G008000,Manes.11G109700,Manes.10G071900-CCNGG,AT3G46210,Manes.09G098300,EIF3E,Manes.01G181600,RRP41,PRL,EIF3A,Manes.16G03300,ORC5,Manes.12G108600-CCNGG,Manes.12G136200,Manes.09G110300,Manes.04G152900,Manes.09G038200,Manes.02G135100,CPSF100,Manes.06G151900,Manes.10G149600,Manes.08G020300,Manes.12G035500,Manes.13G036700,MCM6,Manes.17G068400,Manes.10G059000-CCNGG
GO:0033314	mitotic DNA replication checkpoint	2	5	2.06	0.0452	AT1G07270,CDC6

Tabela suplementar 6B. Termos de GO para função molecular para proteínas da rede de interações formada por sítios diferencialmente hipermetilados para a variedade sensível Dourada baseados em proteínas de *A. thaliana* órtologas às proteínas de *M. esculenta*. As colunas indicam respectivamente: Id do termo GO; descrição do termo; contagem do termo na rede de interações; background de informação para este termo no banco de dados; nível de confiabilidade; taxa de falso descobrimento; nome das proteínas que contém o termo GO

ID do termo	Descrição	Contagem de gene	Background	Strength	False discovery rate	Matching proteins nas redes
GO:0003688	DNA replication origin binding	12	17	2.31	3.62e-19	AT1G07270,MCM2,ORC4,MCM5,MCM9,MCM4,CDC6,ORC2,PRL,ORC5,MCM6,Manes.17G068400
GO:0003678	DNA helicase activity	7	40	1.7	6.87e-07	MCM2,MCM5,MCM9,MCM4,PRL,MCM6,Manes.17G068400
GO:0003690	double-stranded DNA binding	14	428	0.97	6.87e-07	AT1G07270,MCM2,ORC4,MCM5,MCM9,MCM4,CDC6,ORC2,AT3G46210,PRL,ORC5,Manes.04G152900,MCM6,Manes.17G068400
GO:0003676	nucleic acid binding	31	2854	0.5	5.64e-06	RRP4,Manes.12G035400,AT1G07270,TIF3H1,Manes.02G025500,ORC6,CPSF30,MCM2,Manes.18G037900,Manes.10G123600,ORC4,MCM5,MCM9,MCM4,PAPS2,CDC6,ORC2,EIF2,AT3G46210,EIF3E,Manes.01G181600,RRP41,PRL,EIF3A,ORC5,Manes.04G152900,CPSF100,Manes.08G020300,MCM6,Manes.17G068400,Manes.04G032200
GO:0003697	single-stranded DNA binding	7	72	1.45	9.65e-06	MCM2,MCM5,MCM9,MCM4,PRL,MCM6,Manes.17G068400
GO:0140097	catalytic activity, acting on DNA	8	135	1.23	2.37e-05	MCM2,MCM5,MCM9,MCM4,AT3G46210,PRL,MCM6,Manes.17G068400
GO:0004386	helicase activity	8	156	1.17	6.11e-05	Manes.02G025500,MCM2,MCM5,MCM9,MCM4,PRL,MCM6,Manes.17G068400
GO:0003723	RNA binding	15	954	0.66	0.00059	RRP4,TIF3H1,Manes.02G025500,CPSF30,Manes.10G123600,PAPS2,EIF2,AT3G46210,EIF3E,Manes.01G181600,RRP41,EIF3A,CPSF100,Manes.08G020300,Manes.04G032200
GO:0016887	ATPase activity	11	510	0.79	0.00090	Manes.02G025500,MCM2,MCM5,MCM9,MCM4,Manes.03G049000,PRL,Manes.12G108600-CCNKG,Manes.09G038200,MCM6,Manes.17G068400
GO:0097159	organic cyclic compound binding	38	5434	0.3	0.0021	RRP4,Manes.12G035400,AT1G07270,TIF3H1,Manes.02G025500,ORC6,CPSF30,MCM2,Manes.18G037900,Manes.06G013000,Manes.10G123600,ORC4,MCM5,MCM9,MCM4,PAPS2,Manes.03G049000,CDC6,ORC2,EIF2,Manes.12G070400-CCNKG,Manes.06G008000,AT3G46210,EIF3E,Manes.01G181600,RRP41,PRL,EIF3A,ORC5,Manes.12G108600-CCNKG,Manes.04G152900,Manes.09G038200,CPSF100,Manes.06G151900,Manes.08G020300,MCM6,Manes.17G068400,Manes.04G032200

GO:1901363	heterocyclic compound binding	38	5413	0.31	0.0021	RRP4,Manes.12G035400,AT1G07270,TIF3H1,Manes.02G025500,ORC6,CPSF30,MCM2,Manes.18G037900,Manes.06G013000,Manes.10G123600,ORC4,MCM5,MCM9,MCM4,PAPS2,Manes.03G049000,CDC6,ORC2,EIF2,Manes.12G070400-CCNGG,Manes.06G008000,AT3G46210,EIF3E,Manes.01G181600,RRP41,PRL,EIF3A,ORC5,Manes.12G108600-CCNGG,Manes.04G152900,Manes.09G038200,CPSF100,Manes.06G151900,Manes.08G020300,MCM6,Manes.17G068400,Manes.04G032200
GO:0003743	translation initiation factor activity	4	43	1.43	0.0074	TIF3H1,EIF2,EIF3E,EIF3A
GO:0003677	DNA binding	19	1956	0.45	0.0119	Manes.12G035400,AT1G07270,ORC6,CPSF30,MCM2,Manes.18G037900,ORC4,MCM5,MCM9,MCM4,CDC6,ORC2,AT3G46210,PRL,ORC5,Manes.04G152900,Manes.08G020300,MCM6,Manes.17G068400
GO:0005488	binding	47	8453	0.21	0.0290	RRP4,Manes.12G035400,AT1G07270,TIF3H1,Manes.02G025500,ORC6,CPSF30,MCM2,Manes.18G037900,Manes.06G013000,Manes.10G123600,Manes.05G118800,ORC4,MCM5,MCM9,MCM4,PAPS2,Manes.03G049000,Manes.11G140100,CDC6,ORC2,EIF2,ETG1,Manes.12G070400-CCNGG,Manes.12G051300,Manes.05G206400,Manes.06G008000,Manes.10G071900-CCNGG,AT3G46210,EIF3E,Manes.01G181600,RRP41,PRL,EIF3A,ORC5,Manes.12G108600-CCNGG,Manes.04G152900,Manes.09G038200,CPSF100,Manes.06G151900,Manes.08G020300,Manes.12G035500,Manes.13G036700,MCM6,Manes.17G068400,Manes.02G210800,Manes.04G032200
GO:0016787	hydrolase activity	20	2363	0.39	0.0356	TIF3H1,Manes.02G025500,CPSF30,MCM2,Manes.10G123600,CPSF73-II,MCM5,MCM9,MCM4,Manes.03G049000,Manes.05G206400,AT3G46210,Manes.01G181600,PRL,Manes.12G108600-CCNGG,Manes.09G038200,Manes.10G149600,MCM6,Manes.17G068400,Manes.10G059000-CCNGG

Tabela suplementar 6C. Termos de GO para componente celular para proteínas da rede de interações formada por sítios diferencialmente hipermetilados para a variedade sensível Dourada baseados em proteínas de *A. thaliana* órtologas às proteínas de *M. esculenta*. As colunas indicam respectivamente: Id do termo GO; descrição do termo; contagem do termo na rede de interações; background de informação para este termo no banco de dados; nível de confiabilidade; taxa de falso descobrimento; nome das proteínas que contém o termo GO

ID do termo	Descrição	Contagem de gene	Background	Strength	False discovery rate	Matching proteins nas redes
GO:0042555	MCM complex	8	9	2.41	8.71e-13	MCM2,MCM5,MCM9,MCM4,ETG1,PRL,MCM6,Manes.17G068400
GO:0000347	THO complex	6	14	2.09	6.62e-08	MCM2,MCM5,MCM4,PRL,MCM6,Manes.17G068400

GO:0032991	protein-containing complex	29	2226	0.58	2.18e-07	RRP4,Manes.12G035400,TIF3H1,ORC6,CPSF30,MCM2,Manes.10G123600,ORC4,MCM5,MCM9,MCM4,Manes.11G140100,ORC2,EIF2,ETG1,PSF2,AT3G46210,Manes.09G098300,EIF3E,Manes.01G181600,RRP41,PRL,EIF3A,ORC5,Manes.09G038200,CPSF100,Manes.08G020300,MCM6,Manes.17G068400
GO:0005664	nuclear origin of replication recognition complex	4	7	2.22	2.66e-05	ORC6,ORC4,ORC2,ORC5
GO:0000176	nuclear exosome (RNase complex)	4	10	2.06	3.82e-05	RRP4,AT3G46210,Manes.01G181600,RRP41
GO:0000177	cytoplasmic exosome (RNase complex)	4	9	2.11	3.82e-05	RRP4,AT3G46210,Manes.01G181600,RRP41
GO:0005852	eukaryotic translation initiation factor 3 complex	4	9	2.11	3.82e-05	TIF3H1,EIF2,EIF3E,EIF3A
GO:0016282	eukaryotic 43S preinitiation complex	4	9	2.11	3.82e-05	TIF3H1,EIF2,EIF3E,EIF3A
GO:0031981	nuclear lumen	15	807	0.73	3.82e-05	RRP4,ORC6,CPSF30,Manes.10G123600,ORC4,ORC2,ETG1,PSF2,AT3G46210,Manes.01G181600,RRP41,ORC5,CPSF100,Manes.08G020300,Manes.04G032200
GO:0033290	eukaryotic 48S preinitiation complex	4	11	2.02	3.95e-05	TIF3H1,EIF2,EIF3E,EIF3A
GO:0005634	nucleus	36	4669	0.35	0.00018	RRP4,Manes.12G035400,AT1G07270,ORC6,CPSF30,MCM2,Manes.18G037900,Manes.10G123600,ORC4,CPSF73-II,MCM5,MCM9,MCM4,PAPS2,CDC6,ORC2,EIF2,ETG1,Manes.02G091000,Manes.12G051300,PSF2,Manes.05G206400,Manes.06G008000,AT3G46210,Manes.09G098300,EIF3E,Manes.01G181600,RRP41,PRL,ORC5,Manes.04G152900,CPSF100,Manes.08G020300,MCM6,Manes.17G068400,Manes.04G032200
GO:0005847	mRNA cleavage and polyadenylation specificity factor complex	3	5	2.24	0.00033	CPSF30,Manes.10G123600,CPSF100
GO:0000228	nuclear chromosome	6	132	1.12	0.0011	ORC6,ORC4,ORC2,ETG1,PSF2,ORC5
GO:0043231	intracellular membrane-bounded organelle	53	9376	0.21	0.0016	RRP4,Manes.12G035400,AT1G07270,Manes.02G025500,ORC6,CPSF30,MCM2,Manes.18G037900,Manes.10G123600,ORC4,CPSF73-II,MCM5,MCM9,MCM4,PAPS2,Manes.03G049000,Manes.11G140100,CDC6,ORC2,EIF2,ETG1,Manes.08G059900-CCNGG,Manes.02G091000,Manes.12G070400-CCNGG,Manes.12G051300,PSF2,Manes.05G206400,Manes.06G008000,Manes.11G109700,Manes.

GO:0043229	intracellular organelle	54	9831	0.2	0.0027	10G071900-CCNGG,AT3G46210,Manes.09G098300,EIF3E,Manes.01G181600,RRP41,PRL,ORC5,Manes.12G108600-CCNGG,Manes.12G136200,Manes.09G110300,Manes.04G152900,Manes.09G038200,Manes.02G135100,CPSF100,Manes.10G149600,Manes.08G020300,Manes.12G035500,Manes.13G036700,MCM6,Manes.17G068400,Manes.02G210800,Manes.04G032200,Manes.10G059000-CCNGG RRP4,Manes.12G035400,AT1G07270,TIF3H1,Manes.02G025500,ORC6,CPSF30,MCM2,Manes.18G037900,Manes.10G123600,ORC4,CPSF73-II,MCM5,MCM9,MCM4,PAPS2,Manes.03G049000,Manes.11G140100,CDC6,ORC2,EIF2,ETG1,Manes.08G059900-CCNGG,Manes.02G091000,Manes.12G070400-CCNGG,Manes.12G051300,PSF2,Manes.05G206400,Manes.06G008000,Manes.11G109700,Manes.10G071900-CCNGG,AT3G46210,Manes.09G098300,EIF3E,Manes.01G181600,RRP41,PRL,ORC5,Manes.12G108600-CCNGG,Manes.12G136200,Manes.09G110300,Manes.04G152900,Manes.09G038200,Manes.02G135100,CPSF100,Manes.10G149600,Manes.08G020300,Manes.12G035500,Manes.13G036700,MCM6,Manes.17G068400,Manes.02G210800,Manes.04G032200,Manes.10G059000-CCNGG
GO:0071540	eukaryotic translation initiation factor 3 complex, eIF3e	2	2	2.46	0.0067	EIF3E,EIF3A
GO:0005694	chromosome	7	300	0.83	0.0088	ORC6,ORC4,ORC2,ETG1,PSF2,ORC5,Manes.04G152900
GO:0071541	eukaryotic translation initiation factor 3 complex, eIF3m	2	3	2.28	0.0103	EIF2,EIF3A
GO:0005622	intracellular	55	10863	0.16	0.0211	RRP4,Manes.12G035400,AT1G07270,TIF3H1,Manes.02G025500,ORC6,CPSF30,MCM2,Manes.18G037900,Manes.10G123600,ORC4,CPSF73-II,MCM5,MCM9,MCM4,PAPS2,Manes.03G049000,Manes.11G140100,CDC6,ORC2,EIF2,ETG1,Manes.08G059900-CCNGG,Manes.02G091000,Manes.12G070400-CCNGG,Manes.12G051300,PSF2,Manes.05G206400,Manes.06G008000,Manes.11G109700,Manes.10G071900-CCNGG,AT3G46210,Manes.09G098300,EIF3E,Manes.01G181600,RRP41,PRL,EIF3A,ORC5,Manes.12G108600-CCNGG,Manes.12G136200,Manes.09G110300,Manes.04G152900,Manes.09G038200,Manes.02G135100,CPSF100,Manes.10G149600,Manes.08G020300,Manes.12G035500,Manes.13G036700,MCM6,Manes.17G068400,Manes.02G210800,Manes.04G032200,Manes.10G059000-CCNGG
GO:0005654	nucleoplasm	7	396	0.71	0.0412	RRP4,CPSF30,Manes.10G123600,PSF2,CPSF100,Manes.08G020300,Manes.04G032200

Tabela suplementar 6D. Termos KEGG para proteínas da rede de interações formada por sítios diferencialmente hipermetilados para a variedade sensível Dourada baseados em proteínas de *A. thaliana* órtologas às proteínas de *M. esculenta*. As colunas indicam

respectivamente: Id do termo; descrição do termo; contagem do termo na rede de interações; background de informação para este termo no banco de dados; nível de confiabilidade; taxa de falso descobrimento; nome das proteínas que contém o termo

ID do termo	Descrição	Contagem de gene	Background	Strength	False discovery rate	Matching proteins nas redes
ath03030	DNA replication mRNA	6	45	1.59	1.04e-05	MCM2,MCM5,MCM4,PRL,MCM6,Manes.17G068400
ath03015	surveillance pathway	6	114	1.18	0.00072	CPSF30,Manes.10G123600,PAPS2,Manes.01G169300,CPSF100,Manes.04G032200
ath03018	RNA degradation	6	110	1.2	0.00072	RRP4,AT1G60080,AT3G07750,AT3G46210,Manes.01G181600,RRP41
ath03013	RNA transport	6	165	1.02	0.0033	TIF3H1,Manes.01G169300,EIF2,EIF3E,EIF3A,Manes.04G032200

Tabela suplementar 7. Genes diferencialmente metilados nas variedades BRS Dourada e BRS Formosa, seus identificadores e anotações, status de metilação, eventos de polimorfismos de acordo com seus impactos (baixo, moderado, alto); Posição (do polimorfismo); Ref. (nucleotídeo no genoma ref.); Polimorfismo (nucleotídeo/sequência detectada na variedade BRS); Impacto (Baixo, Moderado, Alto); Consequência (tipo polimorfismo); Status heterozigoto/homozigoto; sequências no formato FASTA.

Id	Anotação	Dourada	Formosa	Baixo	Mod.	Alto	Chr.	Posição	Ref.	Polimorfismo	Impacto	Consequência	Dourada	Formosa
Manes.01G037400	Major facilitator superfamily protein	Hypo	-	6	4	0	1	6286341	G	C	Moderate	missense variant	Heterozygous	Heterozygous
Manes.01G040400	ARM repeat superfamily protein	-	Hyper	0	1	0	1	6690829	G	A	Moderate	missense variant	Homozygous	Homozygous
Manes.01G115300	AAR2 protein family	-	Hypo	2	1	0	1	23475421	A	G	Moderate	missense variant	Homozygous	Heterozygous
Manes.01G129200	actin-related protein C1A	-	Hyper	5	6	0	1	24501443	A	G	Moderate	missense variant	Homozygous	N.P.
Manes.01G131400	Reticulon family protein	Hypo	-	12	6	0	1	24636458	T	C,G	Moderate	missense variant	Homozygous	Heterozygous
Manes.01G157300	RNA polymerase I specific transcription initiation factor	Hyper	-	1	3	0	1	26472575	T	A	Moderate	missense variant	Homozygous	Heterozygous
Manes.01G164300	RRN3 protein	-	Hyper	4	8	0	1	26958298	A	G	Moderate	missense variant	Homozygous	Homozygous
Manes.01G169300	FTSH protease 9	Hyper	-	4	3	0	1	27299476	T	G	Moderate	missense variant	N.P.	Heterozygous
Manes.01G181600	splicing factor PWI domain-containing protein	Hyper	Hyper	5	1	0	1	28066932	A	G	Moderate	missense variant	Homozygous	Heterozygous
	3\'-5\'-exoribonuclease family protein													

Manes.01G 182900	glycine-rich protein	Hypo	-	3	4	0	1	28203688	A	ATTATAATCCCTA AAATTATGAAAC TTG	Moderate	disruptive inframe insertion	Homozygous	Homozygous
Manes.01G 211300	endoplasmic reticulum auxin binding protein 1	Hypo	-	3	0	0	1	30176075	C	T	Low	synonymous variant	Heterozygous	Heterozygous
Manes.01G 242800	P-loop containing nucleoside triphosphate hydrolases superfamily protein	-	Hypo	8	34	3	1	32182976	T	C	High	stop lost and splice region variant	Homozygous	Heterozygous
Manes.02G 003700	Metal-dependent protein hydrolase	Hyper	-	2	0	0	2	480062	T	C	Low	synonymous variant	Heterozygous	Heterozygous
Manes.02G 005700	general control non-repressible 3 and ATP-binding cassette, subfamily F, member 3 (ABCF3)	Hypo	-	6	0	0	2	589555	C	A	Low	synonymous variant	Heterozygous	N.P.
Manes.02G 025500	P-loop containing nucleoside triphosphate hydrolases superfamily protein	Hyper	-	16	10	0	2	2004032	C	T	Moderate	missense variant	Homozygous	Homozygous
Manes.02G 033400	cullin 1	-	Hypo	6	0	0	2	2584201	C	T	Low	synonymous variant	Homozygous	N.P.
Manes.02G 056600	Proteasome component (PCI) domain protein	Hyper	-	2	1	0	2	4251956	G	T	Moderate	missense variant	Heterozygous	Heterozygous
Manes.02G 059500	RINGandU-box superfamily protein	Hypo	Hyper	1	0	0	2	4410319	C	T	Low	synonymous variant	Homozygous	Homozygous
Manes.02G 067700	Regulator of Vps4 activity in the MVB pathway protein	Hyper	-	7	7	0	2	5052200	C	T	Moderate	missense variant	Homozygous	Heterozygous
Manes.02G 085400	galacturonosyltransferase 15	Hypo	-	6	3	0	2	6415141	A	T	Moderate	missense variant	Heterozygous	Heterozygous
Manes.02G 135100	dgd1 suppressor 1	Hyper	-	5	4	0	2	9927127	A	C	Moderate	missense variant	N.P.	Heterozygous
Manes.02G 158800		-	Hypo	4	2	0	2	11956506	A	G	Moderate	missense variant	Heterozygous	N.P.
Manes.02G 191800	DNAJ heat shock N-terminal domain-containing protein	Hyper	-	19	7	1	2	15638170	T	C	High	splice donor variant and intron variant	Heterozygous	N.P.
Manes.02G 207300	Protein kinase family protein with ARM repeat domain	Hypo	-	26	29	2	2	20060217	G TACCT ATAGA ACCGG	GA	High	frameshift variant and start lost	N.P.	Heterozygous
Manes.02G 210800	Kinase interacting (KIP1-like) family protein	Hyper	-	2	5	1	2	22061490	T		High	splice donor variant and UTR5' variant and intron variant	Homozygous	N.P.
Manes.02G 226900	Transcriptional factor B3 family protein and auxin-responsive factor AUXandIAA-related	Hypo	-	15	7	0	2	32305236	G	A	Moderate	missense variant	Heterozygous	N.P.
Manes.02G 227200	cinnamate-4-hydroxylase	-	Hypo	12	4	0	2	32369752	G	T	Moderate	missense variant	Heterozygous	N.P.
Manes.03G 005100	cyclic nucleotide-gated channel 14	Hypo	-	3	0	0	3	420027	T TTTCTT CTTC	C	Low	synonymous variant	Homozygous	N.P.
Manes.03G 027200	ARM repeat superfamily protein	-	Hyper	23	25	0	3	2093971		T	Moderate	conservative inframe deletion	Homozygous	Heterozygous
Manes.03G 029000	Insulinase (Peptidase family M16) protein	-	Hyper	18	13	0	3	2237035	C	A	Moderate	missense variant	Homozygous	Heterozygous

Manes.03G 041700	inter-alpha-trypsin inhibitor heavy chain-related	-	Hypo	11	9	0	3	3397167	C	A	Moderate	missense variant	Heterozygous	Heterozygous
Manes.03G 049000	FTSH protease 4	Hyper	-	6	2	0	3	4182075	T	C	Moderate	missense variant	N.P.	Heterozygous
Manes.03G 050400	Protein kinase protein with adenine nucleotide alpha hydrolases-like domain	-	Hyper	1	4	0	3	4386372	A	T	Moderate	missense variant	Heterozygous	N.P.
Manes.03G 061300		Hypo	-	0	1	0	3	6257926	T	C	Moderate	missense variant	Heterozygous	Heterozygous
Manes.03G 062700	Peroxisomal membrane 22 kDa (Mpv17andPMP22) family protein	-	Hyper	6	2	0	3	6619928	A	G	Moderate	missense variant and splice region variant	N.P.	Heterozygous
Manes.03G 064200		Hyper	-	9	10	0	3	6909309	G	A	Moderate	missense variant	Heterozygous	Heterozygous
Manes.03G 066200	helicases;ATP-dependent helicases;nucleic acid binding;ATP binding;DNA- directed DNA polymerases;DNA binding	-	Hypo	23	4	0	3	7392765	C	A	Moderate	missense variant	N.P.	Heterozygous
Manes.03G 066300	BUFP domain-containing protein	-	Hyper	8	7	0	3	7405456	T	A	Moderate	missense variant	Heterozygous	Heterozygous
Manes.03G 083600	ubiquitin-specific protease 25	-	Hyper	1	4	0	3	13389675	G	T	Moderate	missense variant	Homozygous	Homozygous
Manes.03G 103900	homogentisate 1,2-dioxygenase	-	Hyper	2	3	0	3	17997370	C	T	Moderate	missense variant splice region variant and intron	Homozygous	N.P.
Manes.03G 119900		-	Hyper	1	0	0	3	21641291	C	A	Low	variant	Homozygous	Heterozygous
Manes.03G 120000	F-boxandRNI-like superfamily protein	-	Hyper	2	0	0	3	21648779	A	C	Low	synonymous variant	Homozygous	Heterozygous
Manes.03G 134200	Ribosomal protein S8e family protein	Hypo	-	3	1	0	3	23034035	A	C	Moderate	missense variant	Homozygous	Homozygous
Manes.03G 149000	putative type 1 membrane protein	Hyper	-	12	7	0	3	24400481	G	A	Moderate	missense variant	Heterozygous	N.P.
Manes.03G 163000	COP1-interacting protein- related	Hyper	-	15	19	1	3	25612847	G	T	High	stop gained	Heterozygous	N.P.
Manes.03G 190300	DNAJ heat shock N-terminal domain-containing protein	Hyper	-	3	3	0	3	27488437	G	C	Moderate	missense variant	Heterozygous	Heterozygous
Manes.03G 200200	Pectin lyase-like superfamily protein	-	Hypo	3	2	0	3	28144243	C	A	Moderate	missense variant	Homozygous	Homozygous
Manes.04G 009400	cellulose synthase-like A2	-	Hypo	3	0	0	4	1066088	A	T	Low	synonymous variant	Heterozygous	Heterozygous
Manes.04G 014300	serine carboxypeptidase-like 48 RNA-binding	-	Hyper	5	0	0	4	1698358	A	G	Low	synonymous variant UTR5' premature start codon gain	N.P.	Heterozygous
Manes.04G 032200	(RRMandRBDandRNP motifs) family protein	Hyper	-	2	0	0	4	3702843	C	T	Low	variant	Heterozygous	Heterozygous
Manes.04G 071100	Protein of unknown function (DUF616)	Hyper	-	6	6	1	4	20269779	C	T	High	splice donor variant and intron variant	Homozygous	Heterozygous

Manes.04G 143700	TRF-like 6	Hyper	-	4	3	0	4	26976821	C	G	Moderate	missense variant	Heterozygous	N.P.
Manes.04G 152900	SU(VAR)3-9 homolog 3	Hyper	-	7	4	0	4	27559194	T	A	Moderate	missense variant missense variant and splice region variant	Heterozygous	Heterozygous
Manes.05G 026700	G-box binding factor 3	Hypo	-	1	1	0	5	1956549	A	C	Moderate	missense variant	Heterozygous	Heterozygous
Manes.05G 031100		-	Hypo	1	3	0	5	2274211	A	G	Moderate	missense variant	Heterozygous	Heterozygous
Manes.05G 087700	poly(ADP-ribose) polymerase 2	Hypo	-	9	8	0	5	6882819	T	G	Moderate	missense variant synonymous variant	N.P.	Heterozygous
Manes.05G 099700	Ribosomal protein S4	Hypo	-	3	0	0	5	8428514	C	T	Low	splice donor variant and intron variant splice region variant and intron variant	Heterozygous	Homozygous
Manes.05G 134400	Zn-dependent exopeptidases superfamily protein	-	Hyper	6	3	1	5	17873097	G	GTGTGCT	High		Heterozygous	Heterozygous
Manes.05G 184900	KNOTTED-like from Arabidopsis thaliana Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family	-	Hyper	2	0	0	5	25388262	T	C	Low		Heterozygous	N.P.
Manes.05G 202300		Hypo	-	4	2	0	5	27587189	G	A	Moderate	missense variant	Heterozygous	Heterozygous
Manes.05G 206400	ubiquitin family protein	Hyper	-	5	1	0	5	28169722	G	A	Moderate	missense variant synonymous variant	N.P.	Heterozygous
Manes.06G 008000	MAP kinase 9	Hyper	-	3	0	0	6	1258988	C	T	Low	frameshift variant and splice region variant	N.P.	Heterozygous
Manes.06G 012000	Tetratricopeptide repeat (TPR)- like superfamily protein protein kinase family protein and C-type lectin domain- containing protein	-	Hyper	5	10	1	6	1754294	A	AG	High		Heterozygous	Heterozygous
Manes.06G 013000	Galactose oxidase and kelch repeat superfamily protein	Hyper	-	15	5	0	6	1991259	G	C	Moderate	missense variant synonymous variant	Homozygous	Heterozygous
Manes.06G 049500		Hyper	-	4	0	0	6	14235870	T	C	Low		N.P.	Heterozygous
Manes.06G 051100	3-ketoacyl-CoA synthase 11	Hyper	-	19	7	0	6	14435206	T	A	Moderate	missense variant	N.P.	Heterozygous
Manes.06G 052400	Protein of unknown function (DUF3511)	Hyper	-	2	2	1	6	14740302	A	G	High	stop lost and splice region variant	Heterozygous	Heterozygous
Manes.06G 078800	DNA glycosylase superfamily protein	Hypo	-	3	3	0	6	19304366	G	C	Moderate	missense variant	Homozygous	Homozygous
Manes.06G 114700	RING and U-box superfamily protein	Hyper	-	1	2	0	6	22296244	T	C	Moderate	missense variant	Heterozygous	N.P.
Manes.06G 151900	Protein kinase superfamily protein	Hyper	-	3	3	0	6	25538303	G	T	Moderate	missense variant	N.P.	Heterozygous
Manes.06G 156800	manganese tracking factor for mitochondrial SOD2	-	Hypo	6	3	0	6	25887084	G	T	Moderate	missense variant	Heterozygous	Homozygous
Manes.06G 176300	FTSH protease 12	-	Hyper	9	4	1	6	27516737	G	A	High	stop gained	N.P.	Heterozygous
Manes.07G 028400	dynammin-like 3	Hypo	-	18	11	0	7	2680614	T	A	Moderate	missense variant	N.P.	Heterozygous

Manes.07G 132800	Leucine-rich repeat protein kinase family protein	-	Hyper	6	23	1	7	25737684	C	CTT	High	frameshift variant	Heterozygous	Homozygous
Manes.08G 017000	Eukaryotic aspartyl protease family protein	Hypo	-	18	17	0	8	1593049	A	C	Moderate	missense variant	Homozygous	Heterozygous
Manes.08G 020300	Zinc finger C-x8-C-x5-C-x3-H type family protein	Hyper	-	23	6	3	8	1863081	CA	C	High	frameshift variant and start lost synonymous variant	N.P.	Heterozygous
Manes.08G 028700	Helicase protein with RINGandU-box domain	Hypo	-	1	0	0	8	2567708	A	G	Low		Homozygous	Heterozygous
Manes.08G 059900	Transmembrane amino acid transporter family protein	Hyper	-	10	8	0	8	6718506	C	A	Moderate	missense variant	Heterozygous	Homozygous
Manes.08G 132300	Pseudouridine synthase family protein	Hyper	-	0	3	0	8	29940503	C	G	Moderate	missense variant	N.P.	Heterozygous
Manes.08G 157100	Major facilitator superfamily protein	-	Hypo	10	8	0	8	32044349	G	A	Moderate	missense variant	Homozygous	Homozygous
Manes.08G 161700	hydroxyproline-rich glycoprotein family protein	Hypo	-	9	7	0	8	32348948	C	A	Moderate	missense variant	Homozygous	Homozygous
Manes.09G 038200	P-loop containing nucleoside triphosphate hydrolases superfamily protein	Hyper	-	12	3	0	9	5266100	C	T	Moderate	missense variant	Homozygous	Heterozygous
Manes.09G 039500	DNase I-like superfamily protein	-	Hyper	1	2	0	9	5391606	C	G	Moderate	missense variant	Homozygous	Homozygous
Manes.09G 043400		-	Hyper	8	7	0	9	5866242	C	G	Moderate	missense variant	Heterozygous	N.P.
Manes.09G 050300	NAD(P)-binding Rossmann-fold superfamily protein	-	Hypo	2	2	0	9	6734001	T	G	Moderate	missense variant	N.P.	Heterozygous
Manes.09G 076300	Protein of unknown function (DUF1162)	-	Hyper	3	9	0	9	12059307	G	C	Moderate	missense variant	Heterozygous	Homozygous
Manes.09G 110300		Hyper	-	1	0	0	9	23339115	C	A	Low	synonymous variant	Heterozygous	Homozygous
Manes.09G 110600	ferritin 4 ATPase E1-E2 type family protein and haloacid dehalogenase-like hydrolase	-	Hyper	3	2	0	9	23351373	AGAG	A	Moderate	disruptive inframe deletion	Heterozygous	Homozygous
Manes.09G 153700	family protein	-	Hypo	10	10	0	9	26986668	A	G	Moderate	missense variant	Heterozygous	Homozygous
Manes.09G 162500	Major facilitator superfamily protein	-	Hyper	13	12	0	9	27622128	G	A	Moderate	missense variant	N.P.	Heterozygous
Manes.09G 167100	Domain of unknown function (DUF1995)	Hypo	-	4	4	1	9	27900310	C	A	High	stop gained	Heterozygous	N.P.
Manes.10G 004000		Hyper	-	0	4	0	10	241286	T	C	Moderate	missense variant	Homozygous	Homozygous
Manes.10G 024200		-	Hyper	3	11	4	10	1928641	GATTT	G	High	frameshift variant	N.P.	Heterozygous
Manes.10G 025000	Glucose-1-phosphate adenylyltransferase family protein	-	Hypo	6	2	0	10	2042795	T	C	Moderate	missense variant	N.P.	Heterozygous
Manes.10G 034400	2-oxoglutarate dehydrogenase, E1 component	Hyper	-	6	0	0	10	2900647	A	G	Low	UTR5' premature start codon gain variant	Homozygous	Heterozygous

Manes.10G 035100	F-box family protein	-	Hypo	9	15	2	10	2990125	AT	A	High	frameshift variant	N.P.	Heterozygous
Manes.10G 055800	lysyl-tRNA synthetase 1	Hypo	-	6	2	0	10	6457960	C	T	Moderate	missense variant	Heterozygous	N.P.
Manes.10G 059000	beta-hydroxyisobutyryl-CoA hydrolase 1	Hyper	-	4	5	0	10	7486415	T	C	Moderate	missense variant	N.P.	Heterozygous
Manes.10G 063700	P-loop containing nucleoside triphosphate hydrolases superfamily protein	-	Hypo	41	59	0	10	8438957	A	G	Moderate	missense variant	Heterozygous	Heterozygous
Manes.10G 067400	Family of unknown function (DUF572)	-	Hyper	18	1	0	10	9308007	C	A	Moderate	missense variant	Heterozygous	N.P.
Manes.10G 069300	WWE protein-protein interaction domain protein family	-	Hypo	11	13	0	10	10260120	C	G	Moderate	missense variant	Heterozygous	N.P.
Manes.10G 071900	fucosyltransferase 12	Hyper	-	19	11	0	10	11570675	C	G	Moderate	missense variant	N.P.	Heterozygous
Manes.10G 123600	cleavage and polyadenylation specificity factor 73-l	Hyper	-	4	0	0	10	23544988	C	T	Low	synonymous variant	Homozygous	Homozygous
Manes.10G 149600	beta glucosidase 13	Hyper	-	8	23	1	10	26015850	G	A	High	splice donor variant and intron variant	N.P.	Homozygous
Manes.11G 010600	Calcineurin-like metallo- phosphoesterase superfamily protein	Hyper	-	0	2	0	11	1059343	T	C	Moderate	missense variant	N.P.	Heterozygous
Manes.11G 076800		Hyper	-	3	7	0	11	9047924	C	T	Moderate	missense variant	Homozygous	Homozygous
Manes.11G 084700	Homeodomain-like superfamily protein	Hyper	-	2	4	0	11	11663049	A	C	Moderate	missense variant	N.P.	Heterozygous
Manes.11G 109700	DHHC-type zinc finger family protein	Hyper	-	8	8	0	11	20291391	G	C	Moderate	missense variant	Heterozygous	Heterozygous
Manes.11G 140100		Hyper	-	13	3	0	11	25250649	C	G	Moderate	missense variant and splice region variant	N.P.	Heterozygous
Manes.12G 009600	vacuolar proton ATPase A1	Hypo	-	6	4	0	12	839231	T	C	Moderate	missense variant and splice region variant	Heterozygous	Heterozygous
Manes.12G 014500	transcription activator protein with CG-1 and Ankyrin domains	-	Hyper	9	8	0	12	1202301	G	A	Moderate	missense variant	Homozygous	Homozygous
Manes.12G 033800	Protein kinase superfamily protein	-	Hyper	2	0	0	12	2663700	T	C	Low	synonymous variant	Homozygous	Homozygous
Manes.12G 035400	P-loop containing nucleoside triphosphate hydrolases superfamily protein	Hyper	-	4	3	0	12	2818561	A	T	Moderate	missense variant	Heterozygous	N.P.
Manes.12G 035500	breast cancer associated RING 1	Hyper	-	2	1	0	12	2826617	T	C	Moderate	missense variant	Heterozygous	N.P.
Manes.12G 040100	cellulose synthase A4	Hyper	-	5	6	1	12	3358924	G	T	High	stop gained synonymous variant	Homozygous	Heterozygous
Manes.12G 051300	Protein of unknown function (DUF3741)	Hyper	-	2	0	0	12	4519506	A	G	Low		N.P.	Heterozygous
	plant U-box 9	Hyper	-											

Manes.12G 070200	Nodulin MtN21 andEamA-like transporter family protein dehydroquinase dehydratase, putative and shikimate dehydrogenase, putative	-	Hyper	1	0	0	12	6959866	G	A	Low	synonymous variant	Heterozygous	Homozygous
Manes.12G 070400		Hyper	-	11	7	1	12	6984049	C	G	High	stop lost	Heterozygous	Homozygous
Manes.12G 075100	glutamate receptor 5	Hypo	-	8	9	0	12	7740736	C	T	Moderate	missense variant	Heterozygous	Heterozygous
Manes.12G 108600	heavy metal atpase 1	Hyper	-	8	3	0	12	24972459	T	C	Moderate	missense variant	Homozygous	Homozygous
Manes.12G 110200	Protein kinase superfamily protein	-	Hyper	0	2	0	12	25582311	G	C	Moderate	missense variant	Homozygous	Homozygous
Manes.12G 136200	Thioredoxin superfamily protein photosystem I reaction center subunit PSI-N, chloroplast, putative and PSI-N, putative (PSAN)	Hyper	-	2	1	0	12	29068999	A	G	Moderate	missense variant	Heterozygous	Heterozygous
Manes.12G 156000		-	Hyper	1	0	0	12	31375042	A	G	Low	splice region variant and intron variant	Heterozygous	Homozygous
Manes.13G 036700	phytochelatin synthase 1 (PCS1)	Hyper	-	0	1	0	13	3380435	T	G	Moderate	missense variant splice acceptor variant and splice region variant and UTR3' variant and intron variant	N.P.	Heterozygous
Manes.13G 056100		Hyper	-	42	0	1	13	6070595	GCATC AA	G	High		N.P.	Heterozygous
Manes.13G 106000	peroxisomal NAD-malate dehydrogenase 1	Hypo	-	1	1	0	13	23033207	C	A	Moderate	missense variant	Heterozygous	Homozygous
Manes.13G 111500	Arabinanaseandlevansucrasean dinvertase	-	Hyper	2	3	0	13	23830928	C	A	Moderate	missense variant	Homozygous	N.P.
Manes.13G 112600	formin homolog 6	-	Hyper	11	3	0	13	23939194	T	G	Moderate	missense variant	N.P.	Heterozygous
Manes.13G 145000	switch 2	Hypo	-	8	7	0	13	27316786	C	T	Moderate	missense variant	N.P.	Heterozygous
Manes.13G 148900	ELMOandCED-12 family protein	-	Hyper	3	0	0	13	27666098	T	C	Low	synonymous variant	Heterozygous	N.P.
Manes.13G 155300	methyltransferase 1	-	Hyper	14	11	1	13	28056145	TG	T	High	frameshift variant	Homozygous	Heterozygous
Manes.14G 063800	embryo defective 2759	Hyper	-	2	1	0	14	5200500	G	A	Moderate	missense variant	N.P.	Heterozygous
Manes.14G 074700	CLP protease proteolytic subunit 6	Hypo	-	5	1	0	14	6079702	A	G	Moderate	missense variant	Homozygous	N.P.
Manes.14G 141800	response regulator 12	Hypo	-	17	5	0	14	12735613	T	C	Moderate	missense variant	N.P.	Heterozygous
Manes.14G 152900	VH1-interacting kinase	Hypo	-	2	0	2	14	16266774	GTGAG CATGC CACAG TATCA TTGAA GTCCT GC	G	High	splice donor variant and splice region variant and UTR5' variant and intron variant	N.P.	Heterozygous

Manes.15G 015800	chromatin-remodeling protein 11	Hypo	-	6	1	0	15	1290075	A	G	Moderate	missense variant	N.P.	Heterozygous
Manes.15G 054200	Carbohydrate-binding X8 domain superfamily protein	-	Hypo	2	3	0	15	4029637	C	G	Moderate	missense variant	N.P.	Heterozygous
Manes.15G 054300	purple acid phosphatase 15	-	Hypo	6	2	0	15	4033882	C	G	Moderate	missense variant	N.P.	Heterozygous
Manes.15G 056200	ARM repeat superfamily protein	Hyper	-	75	66	2	15	4191871	AT	A	High	frameshift variant	Homozygous	Homozygous
Manes.15G 090400	RINGandU-box superfamily protein	Hyper	-	14	13	0	15	6697117	A	G	Moderate	missense variant	Heterozygous	N.P.
Manes.15G 095300	tubby like protein 7	Hypo	-	3	1	0	15	7028331	T	C	Moderate	missense variant	Heterozygous	Homozygous
Manes.15G 099900	MAP kinase 19	Hypo	-	4	0	0	15	7403822	T	A	Low	synonymous variant	N.P.	Heterozygous
Manes.15G 158800		Hyper	-	12	11	1	15	12493997	G	A	High	splice donor variant and intron variant	N.P.	Homozygous
Manes.15G 159500	uridine kinaseanduracil phosphoribosyltransferase 1	-	Hyper	18	3	0	15	12578117	A	C	Moderate	missense variant	Heterozygous	N.P.
Manes.15G 190300	beta-amylase 2	-	Hyper	2	0	0	15	25033032	C	T	Low	UTR5' premature start codon gain variant	Homozygous	Homozygous
Manes.16G 003300	ERD (early-responsive to dehydration stress) family protein	Hyper	-	21	8	0	16	485314	C	G	Moderate	missense variant	N.P.	Heterozygous
Manes.16G 013200		-	Hyper	9	18	3	16	1297400	A	G	High	stop lost and splice region variant	Homozygous	N.P.
Manes.16G 043500	BES1andBZR1 homolog 4	Hypo	-	7	2	0	16	6124915	G	A	Moderate	missense variant	Heterozygous	N.P.
Manes.16G 051600	Pectin lyase-like superfamily protein	-	Hyper	14	10	0	16	11188613	A	C	Moderate	missense variant	Heterozygous	N.P.
Manes.16G 056900	HISTIDINE TRIAD NUCLEOTIDE-BINDING 2	Hypo	-	0	1	0	16	15411907	C	T	Moderate	missense variant	Homozygous	Heterozygous
Manes.16G 060600	Arp2and3 complex, 34 kD subunit p34-Arc	Hypo	-	5	1	0	16	19159330	G	A	Moderate	missense variant	N.P.	Heterozygous
Manes.16G 060700	TBP-associated factor 11	-	Hyper	1	0	0	16	19180364	A	C	Low	splice region variant	N.P.	Heterozygous
Manes.16G 068800	SWIM zinc finger family protein	-	Hypo	5	1	0	16	21967067	A	G	Moderate	missense variant	Homozygous	Heterozygous
Manes.16G 092800	alphaandbeta-Hydrolases superfamily protein	-	Hyper	10	6	0	16	24929962	T	A	Moderate	missense variant	Heterozygous	Heterozygous
Manes.16G 098500	Myosin family protein with Dil domain	-	Hyper	3	6	0	16	25453479	G	A	Moderate	missense variant	N.P.	Heterozygous
Manes.16G 110100	golgin candidate 1	-	Hyper	2	0	0	16	26566195	G	A	Low	synonymous variant	Heterozygous	N.P.
Manes.16G 128500		Hyper	-	1	4	0	16	27925240	G	A	Moderate	missense variant	N.P.	Heterozygous
Manes.17G 068400	Minichromosome maintenance (MCM2and3and5) family protein	Hyper	-	14	9	0	17	20806190	G	A	Moderate	missense variant	Heterozygous	Heterozygous

Manes.17G 087400	hydroxyproline-rich glycoprotein family protein	-	Hypo	8	5	2	17	22929312	AGC	A	High	frameshift variant splice acceptor variant and intron variant	Heterozygous	N.P.
Manes.17G 118400		Hypo	-	2	0	1	17	26118464	C	T	High	disruptive inframe insertion	Heterozygous	N.P.
Manes.18G 023900	Major facilitator superfamily protein	Hypo	-	15	9	0	18	1833025	C	CAGAGGCGAT	Moderate		Heterozygous	N.P.
Manes.18G 026700	HAD-superfamily hydrolase, subfamily 1G, 5\'-nucleotidase	Hyper	Hyper	7	9	0	18	2498787	T	C	Moderate	missense variant	Homozygous	N.P.
Manes.18G 035000	Major facilitator superfamily protein	Hyper	-	2	3	0	18	3033898	A	G	Moderate	missense variant	N.P.	Heterozygous
Manes.18G 067500		Hypo	-	7	4	0	18	5796831	T	C	Moderate	missense variant UTR5\' premature start codon gain variant	Heterozygous	Heterozygous
Manes.18G 100900	SNARE associated Golgi protein family	Hypo	-	3	0	0	18	8815416	A	G	Low		N.P.	Heterozygous
Manes.18G 115800	O-acyltransferase (WSD1-like) family protein	-	Hypo	71	71	19	18	11000690	A	C	High	splice donor variant and intron variant UTR5\' premature start codon gain variant	Homozygous	Heterozygous
Manes.18G 140600		Hyper	-	2	0	0	18	19530855	A	C	Low		N.P.	Heterozygous
Manes.18G 140900		-	Hyper	11	17	5	18	19970866	C	T	High	stop gained	N.P.	Heterozygous

Categories/events: LOW category [UTR5\' premature start codon gain variant; splice region variant; splice region variant and intron variant; synonymous variant]; MODERATE category [missense variant; conservative inframe deletion; disruptive inframe deletion; disruptive inframe insertion; missense variant and splice region variant]; HIGH category [frameshift variant; frameshift variant and splice region variant; frameshift variant and start lost; frameshift variant and stop lost; splice acceptor variant and intron variant; splice acceptor variant and splice region variant and UTR3\' variant and intron variant; splice donor variant and UTR5\' variant and intron variant; splice donor variant and intron variant; splice donor variant and splice region variant and UTR5\' variant and intron variant; stop gained; stop lost; stop lost and splice region variant].

