

LAURA GONÇALVES COSTA MARTINS

**IDENTIFICAÇÃO DE FATORES TRANSCRICIONAIS QUE CONTROLAM A
EXPRESSÃO DO GENE *SbMATE***

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Magister Scientiae*.

VIÇOSA
MINAS GERAIS – BRASIL
2016

**Ficha catalográfica preparada pela Biblioteca Central da Universidade
Federal de Viçosa - Câmpus Viçosa**

T

M386i
2016
Martins, Laura Gonçalves Costa, 1992-
Identificação de fatores transcricionais que controlam a
expressão do gene *SbMATE* / Laura Gonçalves Costa Martins. –
Viçosa, MG, 2016.
ix, 54f. : il. (algumas color.) ; 29 cm.

Inclui anexos.

Inclui apêndices.

Orientador: Elizabeth Pacheco Batista Fontes.

Dissertação (mestrado) - Universidade Federal de Viçosa.

Referências bibliográficas: f.40-46.

1. Sorgo - Genética molecular. 2. Genética vegetal.
3. Expressão gênica. I. Universidade Federal de Viçosa.
Departamento de Bioquímica e Biologia Molecular. Programa de
Pós-graduação em Genética e Melhoramento. II. Título.

CDD 22. ed. 581.35

LAURA GONÇALVES COSTA MARTINS

**IDENTIFICAÇÃO DE FATORES TRANSCRICIONAIS QUE CONTROLAM A
EXPRESSION DO GENE *SbMATE***

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Magister Scientiae*.

APROVADA: 24 de fevereiro de 2016.

Jurandir Vieira de Magalhaes
(coorientador)

Maiana Reis Pimenta
(coorientadora)

Maximiller Dal-Bianco Lamas Costa

Elizabeth Pacheco Batista Fontes
(Orientadora)

AGRADECIMENTOS

Agradeço primeiramente a Deus, por todas as oportunidades vivenciadas, por me guiar durante esses anos e permitir que eu chegasse até aqui. Sou grata também por ter colocado tantas pessoas maravilhosas a meu lado, que me possibilitaram a conclusão de mais esse ciclo de minha vida.

A minha mãe Maria e ao meu padrasto Renato, pelo apoio diário, todo o amor e dedicação, que foram fundamentais para que eu seguisse nessa caminhada e pudesse chegar até o final.

Aos meus irmãos Frederico e Ana Luisa, pelo carinho e incentivo desde o início. A todos os familiares, pela paciência, apoio e por acreditarem em meu potencial.

À Professora e orientadora Elisabeth Pacheco Batista Fontes, pela dedicação e excelente orientação neste trabalho, por todos os conhecimentos transmitidos, por ter acreditado e apoiado este projeto e por contribuir para o meu futuro profissional, sendo um exemplo a ser seguido.

Ao professor Jurandir Magalhães, pela participação efetiva em meu projeto, sempre me auxiliando com boas ideias e sugestões, que possibilitaram que este trabalho se tornasse completo.

À Maiana Reis Pimenta, peça fundamental nessa pesquisa, por toda a ajuda, paciência, amizade e principalmente excelentes ideias disponibilizadas. Seu conhecimento foi primordial para o desenvolvimento deste trabalho.

Ao Maximiller Dal-Bianco Lamas Costa por ter aceitado prontamente o convite para fazer parte da banca de defesa

A todos os grandes amigos do Laboratório de Biologia Molecular de Plantas, desde os pós-graduandos aos ICS. Márcia, Paola, Bianca, João Paulo, Cleydson, Otávio, Pedro, meu muito obrigado pela ajuda, amizade e ideias compartilhadas. Ao Virgílio, por estar sempre ao meu lado, sendo minha companhia de laboratório, agradeço também pelas conversas, apoio e amizade. À Claudinha pela grande ajuda nos momentos de dificuldade, pela paciência, disponibilidade e por toda a amizade dedicada a mim. Ao Marco Aurélio, pela prontidão a ajudar desde o início dessa etapa. Ao meu amigo Luís, por me auxiliar nos experimentos e por todo o carinho de sempre. Aos vizinhos de bancada, Hanna, Bruno, Otto e Valdeir por toda paciência para comigo. Aos demais integrantes, minha gratidão por toda a ajuda, pelo ótimo convívio estabelecido e principalmente pela recepção ao laboratório.

Aos meus amigos de uma vida toda, agradeço pela força, risadas e por estarem sempre comigo, me escutando, apoiando e incentivando. À Denise e Michele, amigas de infância. À Patrícia, amiga de laboratório antigo, sempre predisposta a ceder seus conhecimentos. Ao Wilson, amigo de todas as horas e para todos momentos, meu muito obrigado. À Gil, juh, Dani, Gabis e outras grandes amizades que a graduação me proporcionou. Aos demais amigos, que estiveram juntos de mim nesse momento, obrigada pelo carinho.

À Universidade Federal de Viçosa, por estes seis anos de recepção. Ao Departamento de Biologia e ao Programa de Genética e melhoramento, por todas as oportunidades e pelo conhecimento adquirido.

À Marlene e Glaucia funcionárias do Laboratório, ao Sr Paulo e demais funcionários do Bioagro, meus agradecimentos pela disposição em ajudar e pela amizade cedida.

À CNPQ, pelo financiamento durante esse um ano e meio de realização deste projeto.

Enfim, a todos meu muito obrigado !

SUMÁRIO

RESUMO	vi
ABSTRACT	viii
1. Introdução	1
2. Revisão de literatura.....	4
2.1 Mecanismo de toxidez por alumínio	4
2.2 Mecanismos de tolerância ao alumínio	5
2.3 Regulação da expressão de genes de tolerância ao alumínio.	7
3. Material e métodos	10
3.1 Ensaio de interação dos putativos transfatores (genes candidatos) com o promotor de SbMATE em leveduras.....	10
3.1.1 Clonagem da região promotora de SbMATE em vetor de expressão em levedura.....	10
3.1.2 Clonagem de cDNAs sintéticos no vetor de expressão de leveduras para ensaios de mono-híbrido	11
3.1.3 Transformação de leveduras	14
3.1.4 Ensaio de Mono-híbrido em leveduras	14
3.1.5 Determinação da concentração de 3-amino-1,2,4-triazole (3AT)	15
3.2 Análise in sílico dos transfatores candidatos.....	15
3.3 Ensaio de transativação do promotor SbMATE em protoplastos de Arabidopsis.....	16
3.3.1 Construção dos diferentes fragmentos do promotor de SbMATE fusionados ao gene repórter GUS.....	16
3.3.2 Transformação de Arabidopsis	18
3.3.3 Clonagem dos transfatores putativos em vetor de expressão em plantas	18
3.3.4 Ensaio de transativação em protoplasto de plantas	20
4. Resultados e discussão	22
4.1..Identificação in sílico de fatores transcricionais candidatos que controlam a expressão do gene SbMATE	22
4.2 A proteína WRKY-like, mas não NFY-like, interage estavelmente com o promotor SbMATE em leveduras.....	25
4.3 Os transfatores WRKY41-like e ZincFinger-like transativam a região proximal do promotor em Arabidopsis thaliana.....	29
4.4. Os sítios de ligação ao DNA dos fatores transcricionais WRKY-like e ZincFinger-like estão contidos em uma região de 146 bp, posições -2386 a -1975, do promotor SbMATE.	31
4.5 Os genes WRKY-like e ZincFnger-like de Sorghum bicolor são amplamente distribuídos no reino vegetal.....	34

5. Conclusões e Perspectivas.....	38
6. Referencias bibliográficas	40
7. Suplementares	47

RESUMO

MARTINS, Laura Gonçalves Costa, M.Sc., Universidade Federal de Viçosa, fevereiro de 2016. **Identificação de fatores transcricionais que controlam a expressão do gene *sbMATE***. Orientadora: Elizabeth Pacheco Batista Fontes. Coorientadores: Jurandir Vieira de Magalhaes e Maiana Reis Pimenta.

O gene *SbMATE*, que confere tolerância a alumínio em sorgo, é altamente expresso no ápice da radícula e codifica um transportador de membrana pertencente à família MATE (*multi drug and toxic compound extrusion family*) que é responsável pelo efluxo de citrato ativado por alumínio. A identificação de elementos cis-regulatórios no promotor do gene *SbMATE* que confere tolerância ao alumínio e de fatores transcricionais que controlam a expressão do referido gene constituem etapas fundamentais para o entendimento do mecanismo de resistência a Al mediado por *SbMATE*. O objetivo do presente trabalho foi a identificação e caracterização de fatores transcricionais que interagem de uma maneira específica com o promotor para controlar a expressão do gene *SbMATE*. Entre 22 possíveis fatores transcricionais mapeados em um eQTL que influencia a expressão de *SbMATE*, quatro deles, NF-YB-Like (Sorbic.009g166200), WRKY-like (Sorbic.009g174300), ZincFinger-like (Sorbic.009g151400) e SORBIDRAFT_09g022540, foram selecionados para análise de atividade de regulação de transcrição específica de promotores. Ensaio de mono-híbrido em leveduras demonstraram que a proteína WRKY41-like, mas não NFY-like, interage com o promotor *SbMATE* isolado tanto de linhagens tolerantes quanto de susceptíveis. A interação ocorre na região proximal (posição -2102 ao ATG) comum a todos os promotores *SbMATE* testados. Estes resultados foram confirmados em ensaios de transativação de promotores fusionados a GUS em protoplastos de *Arabidopsis*. Para isso, *Arabidopsis thaliana* ecotipo Col-0 foi inicialmente transformada com construções de DNA, contendo a região proximal ou de fragmentos truncados do promotor *SbMATE* fusionados à GUS. As folhas das linhagens transgênicas foram utilizadas para o preparo de protoplastos para expressão transiente dos quatro transfatores selecionados. Tanto WRKY-like quanto ZincFinger-like transativaram especificamente o promotor de *SbMATE*, enquanto que NF-YB-like e SORBIDRAFT_09g022540 não foram capazes de ativar a expressão do gene repórter. Por meio de ensaios de deleção do promotor, os sítios de interação de WRKY-like e ZincFinger-like foram mapeados em um domínio cis-regulatório de 127 pb, delimitado pelas posições

-2102 a -1975 no promotor *SbMATE*. Coletivamente, estes resultados indicam que os transfatores WRKY-like e ZincFinger-like estão envolvidos no controle da expressão do gene *SbMATE*.

ABSTRACT

MARTINS, Laura Gonçalves Costa, M.Sc., Universidade Federal de Viçosa, February , 2016. **Identification of transcription factors that control the expression of the gene *SbMATE***. Adviser: Elizabeth Pacheco Batista Fontes. Co-Advisers: Jurandir Vieira de Magalhães and Maiana Reis Pimenta .

The *SbMATE* gene, which confers tolerance aluminum in sorghum, is highly expressed at radicle apex and encodes a membrane transporter belonging to the family MATE (multi drug and toxic compound extrusion family) that is responsible for the efflux of citrate activated by aluminum. The identification of cis-regulatory elements in *SbMATE* gene promoter that confer tolerance to aluminum and transcription factors that control the expression of such gene are key steps for understanding the *SbMATE*- mediated resistance mechanism.. The objective of this study was the identification and characterization of transcription factors that interact specifically with the promoter to control the expression of the gene *SbMATE*. Among 22 transcriptional factors mapped in a eQTL that influences the expression of *SbMATE*, four of them, NF-YB-Like (Sorbic.009g166200), WRKY-like (Sorbic.009g174300), ZincFinger-like (Sorbic.009g151400) and SORBIDRAFT_09g022540 were selected for analysis of transcriptional regularion activity of specific promoters. Mono-hybrid assays in yeast demonstrated that the WRKY41-like protein, but not NFY-like, interacts with the promoter *SbMATE* isolated from both tolerant and susceptible varieties. The interaction occurred in the proximal region (position -2102 to the ATG) common to all *SbMATE* promoters tested. These results were confirmed in transactivation assays of GUS-fused promoters in Arabidopsis protoplasts. For this, Arabidopsis thaliana ecotype Col-0 was initially transformed with DNA constructs, containing the proximal region or truncated fragments of the *SbMATE* promoter fused to GUS. The leaves of transgenic lines were used for the preparation of protoplasts for transient expression of the four selected transactors. Both WRKY-like and ZincFinger-like transactivated specifically the *SbMATE* promoter, while NF-YB-like and SORBIDRAFT_09g022540 were unable to activate the expression of the reporter gene. Through promoter deletion assays, WRKY-like interaction and ZincFinger-like sites were mapped in a cis-regulatory domain of 127 bp, delimited by positions -2102 to -1975 in the

SbMATE promoter. Collectively, these results indicate that the WRKY-like and ZincFinger-like transactors are involved in the control of the *SbMATE* gene expression.

1. Introdução

Plantas estão constantemente expostas a condições ambientais adversas, e os estresse abiótico, tais como: seca, salinidade, variações de temperatura e acidez do solo, constituem a principal causa de prejuízos na agricultura em todo o mundo. A acidez do solo destaca-se como um tipo de estresse importante, uma vez que afeta 40% das terras aráveis. Estresse por acidez do solo é especialmente comum em regiões tropicais e subtropicais devido a maior pluviosidade (Von Uexkull and Mutert, 1995).

O alumínio é o metal mais abundante e o terceiro elemento mais comum na crosta terrestre. Em valores de pH levemente ácido ou neutro, o alumínio se encontra principalmente na forma de aluminos silicatos insolúveis ou óxidos. Nestes valores de pH ácido, formas fitotóxicas do alumínio são liberadas na solução do solo em níveis que afetam o crescimento da raiz (Delhaize *et al*, 2012; Sade *et al*, 2016). Toxidez por Al provoca rápida inibição do crescimento radicular, o que resulta em um sistema radicular pouco desenvolvido, incapaz de explorar camadas mais profundas de solo (Kochian, 1995). O alumínio também causa impactos na membrana plasmática, na deposição de calose e danos ao DNA (Sade *et al*, 2016). Como alternativa de defesa, as plantas utilizam mecanismos de proteção que podem ser classificados em tolerância interna ao alumínio no simplasma e exclusão do Al do simplasma (Ma, 2001).

Como alternativa de defesa, as plantas utilizam mecanismos de proteção que podem ser classificados em tolerância interna ao alumínio no simplasma e exclusão do Al do simplasma.

O sorgo é o sexto cereal mais cultivado e o quinto mais produzido no mundo (FAO 2010), sendo que seus grãos são utilizados na alimentação humana e, no Brasil, são destinados predominantemente à alimentação animal. Nos últimos anos, a produção de sorgo no Brasil vem demonstrando aumentos expressivos, decorrentes à evolução do sistema de cultivo e da disponibilidade de cultivares com maior produção e adaptabilidade às diversas regiões do país (CONAB, 2012). Devido ao seu valor agrônomico e nutritivo, pode ser comparado ao milho (Rodrigues *et al.*, 2002), aparecendo como alternativa viável, uma vez que seu cultivo pode atender uma parte da demanda brasileira de grãos (Coelho *et al.*, 2002). Sendo assim, o sorgo destaca-se como uma cultura economicamente importante.

Existe uma grande diversidade quanto ao uso da cultura do sorgo, aproximadamente 500 milhões de pessoas nos países em desenvolvimento dependem deste como principal alimento básico (Mutisya, *et al.*, 2009), onde é usado em diferentes preparações como pães,

mingaus, cuscuz e tortilhas (Dicko, 2006; Queiroz, 2009; Queiroz, 2011). Sendo assim, o sorgo é considerado uma espécie versátil que pode ser produzida como fonte de grãos, forragem, açúcares e biomassa. Devido ao pequeno genoma (735 Mb), o sorgo tem sido utilizado como um modelo atrativo para avanços no entendimento da estrutura, função e evolução dos genomas dos cereais. Caracterizado como uma cultura economicamente importante, o desenvolvimento de cultivares superiores mais produtivas tem sido o motivo de atração e interesse cada vez mais crescente de melhoristas e biólogos moleculares de plantas (Ribas, 2003).

A toxidez causada por alumínio (Al) é um dos estresse abióticos que mais limitam a produção agrícola em solos ácidos, representando em cerca de 40% das terras cultiváveis no planeta (Ma et al.; 2001). A herança da tolerância ao Al em sorgo foi estudada em uma população F2:3 derivada do cruzamento das linhagens de sorgo SC283 (tolerante) e BR007 (sensível). Os resultados mostraram tratar-se de uma herança monogênica, sendo o gene de tolerância ao Al, localizado no loco *Alt_{SB}* (Magalhães et al., 2004). Posteriormente, o gene *SbMATE*, foi isolado pela técnica de clonagem posicional (Magalhães et al., 2007).

O gene *SbMATE* possui maior expressão no ápice radicular e codifica um transportador de membrana pertencente à família multigênica MATE (*Multi drug and Toxic Compound Extrusion family*), responsável pelo efluxo de citrato ativado pelo tratamento com Al. Na linhagem tolerante, os níveis de expressão do gene *SbMATE* aumentam ao longo do tempo de exposição ao Al. O citrato liberado em resposta ao Al forma um complexo estável e insolúvel com os íons Al (atóxico), prevenindo assim a sua absorção pela planta (Magalhães et al., 2007).

Além dos genes que conferem resistência ao alumínio (Al), a expressão dos genes *MATE* e *ALMT* em *Arabidopsis* também sofrem influências de fatores regulatórios. O gene *sensitive to próton rhizotoxicity 1 (STOP1)* e seu homólogo *Al resistance transcription fator 1 (ART1)* codificam fatores de transcrição do tipo dedo de zinco e estão envolvidos com a tolerância ao alumínio em *Arabidopsis* e arroz respectivamente (Iuchi et al., 2007; Liu et al., 2009; Sawaki et al., 2009; Yamaji et al., 2009).

A região codificadora do gene de tolerância ao alumínio, em sorgo *SbMATE*, é idêntica entre cultivares tolerantes e sensíveis. Além do polimorfismo encontrado no segundo íntron, elementos transponíveis do tipo MITE foram detectados na região promotora, sendo o número de repetições MITE positivamente relacionada com a tolerância (Magalhães et al., 2007).

Objetivando-se estudar a atuação de fatores de transcrição no controle da expressão do gene *SbMATE*, foram realizados ensaios de transativação em levedura, para análise de uma possível interação entre os transfatores selecionados, ao promotor do gene *SbMATE*. Ensaios de transativação em protoplastos também foram realizados, a fim de se analisar de forma mais específica a possível região responsável por transativar o promotor. Para isso, foram gerados uma série de 6 fragmentos adicionais de regiões truncadas do promotor da variedade tolerante SC283.

A identificação e caracterização de interações moleculares que controlam expressão gênica constituem um pré-requisito para definição de estratégias com a finalidade de alterar geneticamente a expressão e acúmulo de proteínas celulares específicas. As cinéticas de indução do gene *SbMATE* e seu padrão de expressão aumentado em linhagens tolerantes à toxidez do alumínio estão coerentes com um modelo no qual a potencialização da expressão do gene alvo confere tolerância a condições adversas de alumínio no solo. Desse modo, o presente trabalho, abordou aspectos relacionados com a regulação do gene *SbMATE*, transativação do promotor do gene por proteínas que apresentam alta expressão genica na presença do Al e ensaios *in silico* dos possíveis transativadores.

2. Revisão de literatura

2.1 Mecanismo de toxidez por alumínio

Diferentes classes de espécies de Al podem ser categorizadas, referente à fitotoxicidade desempenhada por este. No que diz respeito às classes presentes na solução que banha a raiz, estas incluem formas livres ou formas mononucleares de Al, Al polinuclear e Al como complexos de baixo peso molecular. No citoplasma celular, podem ser considerados também complexos macromoleculares reversíveis ou irreversíveis.

Em soluções ácidas ($\text{pH} < 5.0$), o Al solubiliza-se passando à forma de octaedro hexahidratado ($\text{Al}(\text{H}_2\text{O})_6^{3+}$), que por convenção é chamado Al^{3+} . Esta espécie é fitotóxica para as plantas, mesmo em concentrações micromolares (Delhaize e Ryan, 1995), interferindo em grande variedade de processos fisiológicos e moleculares (Kochian *et al.*, 1995). Além disso, causa danos no sistema radicular das plantas, cuja extensão depende do grau de susceptibilidade do genótipo afetado e do nível de saturação de alumínio no solo (Magalhães, 2008). O aumento do pH causa desprotonações sucessivas formando $\text{Al}(\text{OH})_2^+$ e $\text{Al}(\text{OH})^{+2}$. Em pH neutro, o alumínio apresenta-se na forma de $\text{Al}(\text{OH})_3$ (gibbsite), composto que não é tóxico às raízes das plantas, sendo esta a base química da neutralização do efeito fitotóxico do Al^{3+} pela prática da calagem. À medida que o pH aumenta para os valores comumente encontrados no citoplasma, ($\text{pH}=7.4$), o íon aluminato $\text{Al}(\text{OH})_4^-$ domina a especiação (Kochian *et al.*, 1995), reduzindo assim, a fluidez das membranas e aumentando a densidade de empacotamento dos lipídios (Chen *et al.*, 1991).

Além disso, o Al pode induzir a despolarização da membrana por inibir a H^+ -ATPase da membrana plasmática, impedindo a formação e manutenção do gradiente de H^+ (Ahnet *et al.*, 2001; Kochian, 2005). O Al interfere no transporte secundário de íons, causando indiretamente alteração na homeostase dos íons nas células da raiz (Kochian, 2005). Assim também, inibe rápida e eficazmente o influxo de Ca^{2+} para dentro das células vegetais (Rengel, 2003), modulando a atividade de transportadores através de mudanças no potencial de membrana (Kochian, 2005). No citoplasma celular, pode interromper a dinâmica do citoesqueleto, interagindo com os microtúbulos e filamentos de actina (Kochian, 2004), além de poder afetar as funções da respiração mitocondrial (Panda, 2008) e induzir o estresse oxidativo (Richards, 1998; Sade *et al.*, 2016).

Embora, a maior fração do Al associado ao sistema radicular seja encontrado no apoplasma (Kue, 2008), parte significativa do cátion consegue penetrar rapidamente e

interagir com moléculas e estruturas subcelulares do simplasma (Lazof, 1996; Taylor, 2000). Intracelularmente, o cátion é capaz de se ligar ao DNA, afetando negativamente sua atividade molde e a estrutura da cromatina, resultando em modificação do processo de divisão celular (Barceló, 2002; Kochian, 2005).

Plantas afetadas pelo Al apresentam ainda sintomas de deficiência de nutrientes, tais como P e o Ca^{2+} , que são complexados ao alumínio ainda no solo, tornando-o indisponível para as plantas (Huang *et al.*, 1996; Pereira, 2008)

2.2 Mecanismos de tolerância ao alumínio

Nas últimas décadas, muitos laboratórios ao redor do mundo têm concentrado seus esforços em identificar e caracterizar os mecanismos de tolerância das plantas a níveis tóxicos de Al em solos ácidos (Ma, 2001; Pana, 2009). A compreensão destes mecanismos é de fundamental importância para o desenvolvimento de procedimentos para a seleção rápida de plantas resistentes ao Al e com bom desempenho em solos ácidos (Barceló, 2002).

Existe uma complexidade de mecanismo aos quais as plantas utilizam para responder ao estresse por alumínio, selecionados durante a evolução, como forma de contornarem seus efeitos tóxicos (Hunag, 2009). Tais mecanismos parecem ser resultado de duas estratégias de desintoxicação do Al utilizadas pelas plantas e podem ser classificados em: tolerância interna ao alumínio no simplasma e tolerância externa. Na tolerância interna, a planta seria capaz de crescer e se desenvolver mesmo após a entrada de Al no simplasma das células. Já na tolerância externa, ocorre a exclusão do Al do ápice radicular por diferentes processos incluindo sua imobilização nas paredes celulares, a permeabilidade seletiva das membranas, a formação de barreira de pH induzida pela planta na rizosfera.

Em plantas que operam o mecanismo interno de desintoxicação, o Al é quelado no citosol por ácidos orgânicos, proteínas, ou outros ligantes orgânicos formando complexos não-tóxicos e, em seguida, estes complexos são compartimentalizados nos vacúolos. Deste modo, o Al modifica pouco ou nada o metabolismo da planta (Kochian, 2005; Ma, 2000; Ma, 2001. Ryan, 2010). Este mecanismo de tolerância é encontrado principalmente em espécies endêmicas de regiões com solos ácidos, onde a capacidade de lidar com a toxicidade ao Al é pré-requisito para sua sobrevivência (Ryan, 2010).

Alguns estudos demonstram a existência de um mecanismo, que envolve genes que codificam proteínas similares às proteínas do tipo transportadores ABC (ATP binding cassette). Dois genes que codificam para transportadores ABC: *ALS3* e *ALS1*, associados à

tolerância ao alumínio em *Arabidopsis* (Larsen, 2005; 2007), têm sido funcionalmente caracterizados, como responsáveis por desintoxicar substâncias orgânicas e inorgânicas (Shulz, 2006). *ALSI* é expresso no ápice das raízes e em tecidos vasculares, sendo expresso na membrana vacuolar (Larsen, 2007). *ALS3* é expresso em diferentes órgãos e está localizado na membrana plasmática (Larsen *et al.*, 2005).

Os genes *STAR1* (*Sensitive to Al rhizotoxicity1*) e *STAR2* foram identificados por sua indução ao Al. *STAR1* codifica uma proteína que contém um domínio de ligação à nucleotídeos, enquanto *STAR2* codifica uma proteína com domínio transmembrana. Ambos os genes são expressos principalmente nas raízes e são especificamente induzidos pela exposição ao Al. *STAR1* interage com *STAR2* para formar um complexo que funciona como um único transportador ABC. O complexo transporta UDP-Glc e é necessário para tolerância ao Al em arroz (Huang, 2009). *STAR2* é homólogo à *ASL3* embora apresentem padrões de expressão diferentes. *STAR2* é expresso apenas em raízes, enquanto *ASL3* é expresso em raízes, folhas, caules e flores.

Todavia, o mecanismo predominante de tolerância ao alumínio em plantas envolve a ativação de transportadores de membrana, que intermedeiam a exudação de ácidos orgânicos pela raiz (Kochian, 2004). Através do mecanismo de exclusão do alumínio, as plantas são capazes de impedir que seja absorvido pela raiz e se acumule no simplasma (Ryan, 2010). Isto se dá por meio da exudação pela planta de compostos fenólicos (Ofei-Manu, 2001), formação de mucilagem (Miyasaka, 2001), aumento do pH da rizosfera (Degenhardt, 1998) e exudação de ácidos orgânicos como o malato, citrato e oxalato, capazes de formar complexos com Al no meio externo; realizando a alcalinização da rizosfera; a redistribuição e, ou efluxo do alumínio (Ma, 2007).

Embora a exudação de ácidos orgânicos já fosse previamente conhecida, o gene que controla a resistência ao alumínio, através da exudação de ácidos orgânicos, só foi isolado há onze anos. Sasaki *et al.* (2004), isolaram e caracterizaram, pela primeira vez, um gene que confere um efluxo de malato ativado por alumínio em trigo (*Triticum aestivum*). *TaALMT* (*Triticum aestivum*aluminium-activated malate transporter) codifica um membro da família ALMT que consiste em proteínas transmembranas (Delhaize, 2007), a partir de um transportador de malato ativado por alumínio que é capaz de conferir tolerância ao alumínio em células vegetais. A família ALMT está localizado na membrana plasmática e funciona como um canal de ânion ativado pelo alumínio, exudando malato pelas células da raiz (Yamaguchi, 2005; Zhang, 2008).

Outros importantes genes de resistência ao alumínio foram também descritos; nesses casos, o ácido orgânico exudado é o citrato, sendo a proteína envolvida pertencente à família MATE (*multi drug and toxic compound extrusion*) (Furukawa, 2007). A família de proteínas transportadoras MATE corresponde a um grande e diversificado grupo de proteínas presentes em procariotos e eucariotos. Genes MATE foram relacionados à resistência ao alumínio em sorgo (Magalhães, 2007), soja (*GmALMT*; Liang *et al.*, 2013) cevada (*HvAACT1*; Furukawa, 2007), trigo (*TaMATE1*; Ryan, 2009), *Arabidopsis* (*AtMATE*; Liu *et al.*, 2009), centeio (*ScFRDL2*; Yokosho, 2010), milho (Maron, 2010) e arroz (*OsFRDL4*; Yokosho, 2011).

Embora os membros destas famílias de proteínas sejam diferentes em sua sequência e estrutura, conferem resistência ao Al^{3+} de uma forma similar, estão envolvidas na exudação de malato e citrato, podendo, em alguns casos, coexistirem na mesma espécie (Liu *et al.*, 2009). Acredita-se que o efluxo de oxalato e de outras espécies de ácidos orgânicos possa ser controlado por uma terceira família de genes, ainda não identificada.

2.3 Regulação da expressão de genes de tolerância ao alumínio.

Em *Arabidopsis*, a expressão de ambos os genes, *AtALMT* e *AtMATE*, é regulada por um fator de transcrição do tipo dedo de Zinco Cys₂His₂-type, designado *STOP1* (*Sensitive to proton rhizotoxicity 1*), que é também associado com a tolerância à baixos pHs (Iuchi, 2007).

Homólogo à *STOP1*, o gene *ART1* (*Aluminium resistance transcription factor1*), também pertence à família de fatores de transcrição do tipo dedo de Zinco Cys₂His₂-type, interage com o cis-elemento [GGN(T/g/a/C)V(C/A/g)S(C/G)] e regula a expressão de genes envolvidos na tolerância ao alumínio (TSUTSUI, 2011), como *STAR1* e *STAR2*, pertencentes à família de transportadores ABC (Huang, 2009), *Nrat1* (Xia, 2010), *OsFRDL* (Yokosho, 2011) e *OsALS1* (Huang, 2012).

OsFRDL4, representante da família MATE, em arroz, é um transportador de citrato com baixo nível de expressão nas raízes, na ausência de Al, mas induzido após exposição ao metal (Yokosho, 2011). *OsALS1* é um transportador ABC envolvido na tolerância ao alumínio em arroz (Huang, 2012).

Em *Sorghum bicolor*, através de clonagem posicional, Magalhães *et al.*, (2007) fizeram o mapeamento genético de linhagens tolerantes e sensíveis ao alumínio e

localizaram o lócus de *SbMATE*, AltSB, numa região de 24, 6 Kpb do BAC 181g10, mapeado na região terminal do terceiro cromossomo.

Com relação ao gene *SbMATE*, foram encontradas regiões polimórficas apenas em íntrons e regiões não codificantes quando comparadas aos éxons. A mais divergente dentre elas, apresenta um indel de 728-pbupstream ao TATA box predito para *SbMATE* e abriga um elemento transponível do tipo MITE. Magalhães *et al.*, (2007) amplificaram a região de inserção MITE e verificaram correlação positiva e significativa entre a variação no tamanho dessa região e a tolerância ao alumínio. A análise da região MITE em acessos de sorgo selecionados também apresentou padrão altamente estruturado e repetido além de seqüências que flanqueavam o local de inserção MITE. Esta estrutura pode ser única ou repetida três, quatro e cinco vezes nas diferentes variedades BR007, BR012 e SC283 respectivamente.

Polimorfismos na região codificadora são mais raros, tendo sido observados em duas linhagens Tx642 e SC566 (Melo *et al.*, 2013). Na primeira, sensível ao alumínio, inserções, deleções e SNPs promovem a formação de um stop codon prematuro, gerando uma proteína *SbMATE* truncada. Na linhagem SC566, que é resistente ao alumínio e mostra altas taxas de exudação de citrato pelas raízes, o único SNP, identificado no primeiro éxon, causa uma mudança não conservativa do aminoácido leucina para histidina. Estes mesmos autores mostraram que linhagens isogênicas feitas pela introgressão de alelos de AltSB apresentaram grande variação na tolerância ao alumínio quando comparadas aos seus parentais. Além disso, foi demonstrado que, esta transferência incompleta de genes responsáveis pela tolerância ao alumínio em linhagens isogênicas é associada à redução na expressão de *SbMATE* e indica a utilização de fatores regulatórios agindo *in trans*. Nesse estudo, foi demonstrado que a baixa expressão do gene *SbMATE* em linhagens de sorgo, sensíveis ao Al é devida predominantemente à perda ou redução de transfatores em *background* genéticos sensíveis ao Al. A complementação da baixa expressão do gene *SbMATE*, devido a deficiência funcional dos transfatores, pode ocorrer em estoques genéticos híbridos, gerados por cruzamento de um parental tolerante que apresente altos níveis de expressão de *SbMATE*, responsáveis em restaurar a expressão dos alelos sensíveis.

A partir da detecção de um eQTL de efeito maior, co-localizado com o loco *Alt_{SB}*, se confirmou a dominância dos efeitos em *cis* no controle da expressão do gene *SbMATE*. Entretanto, um outro eQTL foi detectado fora do loco *Alt_{SB}*, no qual possivelmente estão localizados fatores em *trans* que afetam a expressão do gene *SbMATE*.

Devido ao fato das regiões físicas dos eQTLs serem extensas, compreendendo vários genes candidatos, foi realizado o estudo de mapeamento associativo (*genome-wide*

association – GWAS) para mapeamento de alta resolução dos transfatores, por meio da qual foi possível encontrar SNPs associados à tolerância ao Al e à expressão do gene *SbMATE* co-localizados na região dos eQTLs, o que possibilitou a identificação dos 22 transfatores candidatos (dados não publicados). Em estudos posteriores, foram selecionados, entre os 22 possíveis, 5 transfatores putativos, com base na maior indução ao Alumínio, aos quais serão objeto de estudo desse projeto referencial .

Levando-se em consideração que a região codificadora de *SbMATE* apresenta alta porcentagem de identidade entre os alelos sensíveis e tolerantes, e que, a maior divergência entre as linhagens apresenta-se no número de cópias repetidas de inserções MITE, acredita-se que as mesmas estejam localizadas em regiões regulatórias de *SbMATE*, e que, ajam aumentando a expressão do gene no ápice da raiz. Coletivamente, estes resultados sugerem que a induzibilidade do gene de resistência *SbMATE* por Al é controlada por sequências regulatórias fora da região codificadora, que são alvos de fatores de transcrição (Magalhaes et al; 2007). Entretanto, elementos cis-regulatórios ainda não foram precisamente mapeados no promotor do gene *SbMATE* e, tampouco, foram isolados fatores de transcrição que controlam de uma maneira específica a expressão do referido gene, o que constituem os objetivos principais da presente investigação.

3. Material e métodos

3.1 Ensaios de interação dos putativos transfatores (genes candidatos) com o promotor de *SbMATE* em leveduras.

Células de levedura *Saccharomyces cerevisiae* da linhagem W303, previamente transformadas com o vetor pHIS2.1, contendo diferentes extensões de sequências 5' que flaqueiam o gene *SbMATE*, foram individualmente co-transformadas com cinco construções de cDNA de sorgo clonadas em pGADT7-AD. A interação dos possíveis transfatores com a região promotora do gene foi monitorada pela ativação do gene repórter *HIS3*.

3.1.1 Clonagem da região promotora de *SbMATE* em vetor de expressão em levedura

Promotores completos de *SbMATE* das variedades SC283 (tolerante ao Alumínio), Tx430 e BR012 (sensíveis) foram sintetizados comercialmente, ligados ao vetor, pBlueScript II SK (pBSK) e propagados em *Escherichia coli*, estirpe DH5 α .

Colônias positivas foram selecionadas em meio contendo ampicilina (10 mg/L), e os clones resultantes armazenados em estoques em glicerol (Tabela 1). Na nomenclatura dos clones, o número inicial representa a posição de extensão da sequência promotora a partir do possível ATG do gene *SbMATE*, p denota promotor, seguido pelo número de repetições MITEs contidas e a descrição da variedade da qual se derivou a referida sequência promotora.

Tabela 1: Clones gerados pela inserção de sequências promotoras do gene *SbMATE* de diferentes variedades de *Sorghum bicolor* (SC283, Tx430, BR007 e BR012) em pBlueScript II SK.

Plasmídeos*	Catálogo de laboratório
-5684pSbMATE Fm5miteatg-SC283	pUFV2371
-5299pSbMATE Fm4miteatg-BR012	pUFV2374
-4214pSbMATE Fm1miteatg-Tx430	pUFV2372
--4940pSbMATE Fm3miteatg-BR007	pUFV2373
-3935pSbMATE 5miteatg	pUFV2502
-2028 pSbMATE 0miteatg	pUFV2503

*O número inicial representa a extensão da sequência promotora a partir do ATG do gene *SbMATE* da variedade especificada.

As regiões promotoras sintetizadas foram liberadas do respectivo plasmídeo recombinante, conforme Tabela 1, por digestão com as enzimas *EcoRI* e *SmaI*, purificados de géis de agarose utilizando-se o kit Gel Extraction Kit (Qiagen). Em seguida, os fragmentos purificados foram ligados ao vetor pHIS2.1 (Clontech) de expressão em leveduras, previamente digerido com as mesmas enzimas, utilizando 1µL da enzima T4 DNA ligase (400U/ml), adicionalmente à 1 µL do tampão de ligação (10X) da enzima, 6µl de inserto em 2 µL de vetor, (proporção de 3:1 da relação fragmento /plasmídeo) para um volume final de reação de 10 µL, foram então incubadas “over night” a 4°C. As reações de ligação foram utilizadas na transformação de células competentes DH5α, pelo método do choque térmico, e os transformantes positivos foram selecionados em meio contendo o antibiótico kamicina (10 mg/L). Os clones confirmados por sequenciamento foram estocados em -80°C (Tabela 2) e contêm o gene repórter HIS3 sob o controle do fragmento de promotor especificado na nomenclatura do clone, sendo específico para o sistema de mono-híbrido para o escrutineio de cDNAs de fatores de transcrição.

Tabela 2: Plasmídeos recombinantes derivados de pHIS2.1, contendo fragmentos do promotor de SbMATE utilizados no sistema mono híbrido de leveduras .

Plasmídeos *	Catálogo de laboratório
-5684pSbMATE Fm5miteatg-SC283HIS3	pUFV2376
-5299pSbMATE Fm4miteatg-Tx430HIS3	pUFV2377
-4214pSbMATE Fm1miteatg-BR012HIS3	pUFV2379
-3935pSbMATE 5miteHIS3	pUFV2529
-2028pSbMATE 0miteHIS3	pUFV2530
-2102(32) pSbMATE -SC283HIS3	pUFV2044

* O numero inicial indica a extensão do fragmento do promotor de *SbMATE* da variedade de sorgo especificada ate o primeiro ATG..

3.1.2 Clonagem de cDNAs sintéticos no vetor de expressão de leveduras para ensaios de mono-híbrido

Entre os vinte e dois (22) cDNAs de transfatores candidatos a interagirem com o promotor de SC283, identificados por mapeamento associativo a partir de eQTL ligado à tolerância a Al em *trans*, quatro(4) foram selecionados de acordo com o nível de expressão

gênica na presença de Alumínio. Estes transfadores, sintetizados comercialmente no vetor pBlueScript II SK (pBSK) flanqueados pela metade do sítio de recombinação AttB1 (*Foward*) e AttB2 (*Reverse*), foram utilizados como template na reação de PCR utilizando oligonucleotídeos específicos descritos na Tabela 3, contendo sítios das enzimas de restrição BamHI (*Foward*) e XhoI (*reverse*). Para que a sequência aberta de leitura fosse mantida; para Sorbic.009G151400 (16) não foram adicionados sítios de enzimas de restrição. A reação de PCR (reação em cadeia da polimerase) continha 1,5 µL do DNA template, 5 µL do tampão 10X da enzima *Taq-platinum* DNA polimerase, 4,0 µL da mistura 2,5 mM de dNTPs (desoxirribonucleotídeos trifosfatos), 2µL de cada um dos 443 oligonucleotídeos (*Foward e Reverse* 10 uM cada), 2 µl de MgSO₄ 50mM e 2,5 U DNA polimerase (*Taq-Platinum*, invitrogem) completados com 33.3µL de água deionizada estéril para 50µL de volume final. O tubo foi submetido a PCR no termociclador programável, nas seguintes condições: um ciclo inicial de desnaturação de 95°C/3min, seguidos de 30 ciclos de 95°C /45 seg, 55°C /45 seg, 68°C /1,5 min e extensão final de 68°C /7 min. Os produtos amplificados foram submetidos à eletroforese em gel de agarose a 1% (p/v), contendo Brometo de Etídio (0,25µg/ml) e visualizados sob luz ultravioleta (UV) purificados do gel (o kit Gel Extraction Kit (Qiagen) e clonados em pCR-TOPO 2.1 utilizando T4 ligase (400U/ml), adicionalmente à 1 µL do tampão de ligação (10X) da enzima, incubadas “over night” a 4°C. As reações de ligação foram utilizadas na transformação de células competentes DH5α, pelo método do choque térmico, e os transformantes positivos foram selecionados em meio contendo o antibiótico ampicilina (10 mg/L). Os clones confirmados por sequenciamento foram estocados em -80°C (Tabela 4).

Tabela 3: Oligonucleotídeos usados para amplificação dos cDNAs sintéticos a partir de insertos em pBSK.

Transfator	Primer	Sequencia
Sorbic.009G166200 (13)	Sb09g022810BamHIFwd	ggatcccgATGGCGGACGCTCCG
Sorbic.009G166200 (13)	Sb09g022810XhoIRvs	ctcgagTTA GTT TGA GAT ATC CCC ATT
Sorbic.009G174300 (14)	Sb09g023500BamHIFwd	ggatccgtATG ACG CTG TCT CCA CCG CAG
Sorbic.009G174300 (14)	Sb09g023500XhoIRvs	ctcgagCTA GTA CCC ATC ATA CGA AAA
Sorbic.009G151400 (16)	Sb09g021530Fwd	ATG GCG CGT CCG CAG CGA GTG

Sorbic.009G151400 (16)	Sb09g021530Rvs	CTA GTG AGT CTC TGC TAG AGA
SORBIDRAFT_09g022540 (19)	Sb09g022540BamHIFwd	ggatccgtATG CCG CCG CTG GCG
SORBIDRAFT_09g022540 (19)	Sb09g022540XhoIRvs	ctcgagTCA TTC CTT CTT CCC TAT GCC

*Letras minúsculas na sequência dos oligonucleotídeos indicam os sítios de restrição incluídos no produto amplificado.

Para clonagem dos cDNAs sintéticos no vetor de expressão de leveduras pGADT7-AD, as regiões codificadoras dos transfadores Sorbic.009G166200 (13), Sorbic.009G174300 (14) e SORBIDRAFT_09g022540 (19) foram liberados de TOPO 2.1 utilizando as enzimas de restrição *BamHI* (Promega) e *XhoI* (Promega). A digestão ocorreu por 2 horas à 37°C utilizando-se 10ng/μl de DNA, 2u *BamHI*, 2u *XhoI*, 5 μL de tampão “D” (Promega) e 36 μL de água miliQ autoclavada. Para Sorbic.009G151400 (16) foi utilizada apenas *EcoRI* (Invitrogen). A digestão ocorreu por 2 horas à 37°C utilizando-se 10ng/μL de DNA, 2u *EcoRI*, 5 μL de tampão “React 3” (Invitrogen) e 38 μL de água miliQ autoclavada. As reações de restrição foram corridas em gel de agarose a 1% (p/v), purificadas e clonadas com T4 ligase, em pGADT7-AD previamente digeridos com as mesmas enzimas. O vetor pGADT7-AD foi tratado com 2U de fosfatase alcalina (CIAP-Calf Intestinal Alkaline Phosphatase) responsável por manter o plasmídeo na sua forma linear, a reação foi incubada a 37°C por 30 minutos.

As ligações foram utilizadas na transformação de células competentes de *Escherichia coli* da estirpe DH5α, pelo método do choque térmico, os transformantes positivos foram selecionados com ampicilina (10 mg/L) e diagnosticados por sequenciamento. Os clones resultantes (Tabela 4) foram armazenados em glicerol 50% e mantidos em freezer -80°C. Nos clones resultantes, a região codificadora do transfator putativo está fusionada “in frame” ao terminal carboxílico do domínio de ativação de GAL4 e de uma etiqueta HA, sob o controle do promotor e de sinais de terminação de transcrição do gene *ADHI*.

Tabela 4: Clones obtidos pela inserção da região codificadora dos genes selecionadas em diferentes vetores de clonagem e expressão em leveduras.

Transfator candidato	Tamanho (pb)	Clones em pBSK	Clones em PCR2.1-TOPO	Clones em pGADT7-AD
Sorbic.009G166200 (13)	604	pUFV2458	pUFV2594	pUFV2597(pNF-Y-like)

Sorbic.009G174300 (14)	769	pUFV2459	pUFV2595	pUFV2598(pWRKY41-like)
Sorbic.009G151400 (16)	1354	pUFV2461	pUFV2596	pUFV2659(pZincfinger-like)
SORBIDRAFT_09g022540 (19)	589	pUFV2464	pUFV2637	pUFV2648 (pSb09g022540)

3.1.3 Transformação de leveduras

Células de levedura *Saccharomyces cerevisiae* da linhagem W303 (MATa/MAT α leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 [phi⁺]), que é deficiente na produção de leucina, triptofano, uracila e histidina (Trp-, Leu-, Ura-, His-), foram crescidas em meio YPD-A por 16 horas, sendo então submetidas à centrifugação de 13000 g por 30 segundos e re-suspensas com 1 mL de água deionizada. Após re-centrifugação a 13000 g por 30 segundos foram obtidas as células pré-lavadas utilizadas para transformação com DNA plasmidial, utilizando-se o método de transformação com acetato de lítio/polietilenoglicol (PEG), segundo instruções da Clonteh. Os transformantes obtidos foram plaqueados em meio seletivo e crescidos por três a quatro dias a 28°C.

3.1.4 Ensaio de Mono-híbrido em leveduras

A interação dos possíveis transfatores com a região promotora de *SbMATE* foi monitorada pela ativação do gene repórter HIS3. Para isso, células de levedura *Saccharomyces cerevisiae* da linhagem W303, previamente transformadas com os clones do promotor pUFV2376 (-5684pSbMATE-SC283pHIS2.1), pUFV2379 (-5299pSbMATE-BR012pHIS2.1), pUFV2377 (-4214pSbMATE-Tx430pHIS2.1) e pUFV2044 (-2102pSbMATEpHIS2.1) foram co-trasformadas com os clones pUFV2597 (pNF-YB-like), pUFV2598 (pWRKY41-like), pUFV2659 (pZincfinger-like) e pUFV2648, selecionados a partir dos 22 transfatores putativos identificados por eQTL. Posteriormente, as leveduras transformadas foram plaqueadas em SD-Leu-Trp, e mantidas em BOD 28°C por 36 horas, até atingir OD de 1,0, para monitoramento da co-transformação com os dois plasmídeos. Para análise de interação dos transfatores com os fragmentos de promotores utilizados, os

transformantes foram crescidos em meio deficiente de Leu, Trp e His e suplementado com diferentes concentrações de 3-Amino-1,2,4-triazole (3-AT)

3.1.5 Determinação da concentração de 3-amino-1,2,4-triazole (3AT)

Na ausência de ativação, a expressão constitutiva de HIS3 é muito baixa. Durante a triagem, a expressão basal é controlada pela adição de 3-amino-1,2,4-triazole (3-AT) ao meio. Para se determinar a concentração necessária para suprimir totalmente a expressão de HIS3 para os clones a serem utilizados, células de levedura *Saccharomyces cerevisiae* da linhagem W303 transformadas com os plasmídeos recombinantes foram plaqueadas em meio SD-Leu-Trip-His suplementado com 3-AT, nas concentrações de 2,5 mM; 5,0 mM; 7,5 mM e 10 mM.

3.2 Análise in silico dos transfatores candidatos

As sequências completas dos fatores de transcrição putativos foram obtidas nos bancos de dados do Phytozome (<http://www.phytozome.net>), e submetidas ao programa WolfPsort (<http://www.genscript.com/wolf-psort.html>), para a análise da localização subcelular. A busca de ortólogos foi feita por meio do programa eggNOG, (<http://eggnogdb.embl.de>). Para a busca heurística local com algoritmo BLASTp (Basic Local Alignment Search Tools for protein), sequências de aminoácidos de genes das espécies *Arabidopsis thaliana*, *Cariacica papaya*, *Cucumis sativus*, *Citrus sinensis*, *Glicine max*, *Manihot esculenta*, *Oriza sativa*, *Phaseolus vulgaris*, *Sorgum bicolor*, *Solanum lycopersicum*, *Solanum tuberosum*, *Triticum aestivum*, *Vitis vinifera* e *Zea mays* foram recuperadas do sistema on-line Phytozome V.10.3 (<https://phytozome.jgi.doe.gov>). Juntamente, os genes Sobic.009G174300 (WRKY DNA-binding domain) e Sobic.009G151400 (DHHC-type Zn-finger proteins) de *Sorgum bicolor* foram utilizados como molde contra as sequências de aminoácidos das espécies selecionadas.

Utilizando comparação par a par das sequências de aminoácidos o algoritmo BLAST foi ajustado com parâmetro do *E-value* determinado em 10E-10. Baseando-se nos critérios de maior identidade, similaridade e cobertura da sequência suportados pelo índice de confiança *E-value*, os cinco genes melhor ranqueados em cada espécie foram selecionados. Sendo assim, um conjunto de dados em formato FASTA contendo 70 sequências de

aminoácidos foi criado para cada gene template. Para as análises filogenéticas, as sequências de aminoácidos destes genes foram alinhadas com auxílio do software MUSCLE (Edgar, 2004), utilizando parâmetros default. Deste modo, as árvores foram construídas utilizando o método da máxima verossimilhança, com 10.000 réplicas de bootstraps, ajustado com modelo Jones-Taylor-Thorton utilizando matriz de substituição de nucleotídeos BLOSUM45 por meio do software FastTree Version 2.1.4 SSE3 (Price et all 2010) e visualizadas com auxílio do software Figtree v1.4 (Rambaut, 2015).

3.3 Ensaio de transativação do promotor *SbMATE* em protoplastos de *Arabidopsis*

3.3.1 Construção dos diferentes fragmentos do promotor de *SbMATE* fusionados ao gene repórter *GUS*

DNA total da variedade tolerante SC283 foi utilizado como molde para isolamento de fragmentos do promotor *SbMATE* utilizando oligonucleotídeos específicos que anelam em diferentes extensões do referido promotor e contem sítios para as enzimas de restrição EcoRI e HindIII, gerando fragmentos de diferentes tamanhos (Tabela 5). Para reação de PCR, foi utilizado 1µl do DNA molde (10ng/µl), 1µl de cada oligonucleotídeo (*Foward e Reverse* a 10µM cada), 0,1µl da enzima de alta processividade DNA polimerase (*Taq platinum* invitrogen), 5µl de DMSO 20%, em 2µl do tampão 10X da enzima *Taq platinum* DNA polimerase, 1µl de dNTPs (desoxirribonucleotídeos trifosfatos) 2,5 mM, 0,8 µl MgSO₄ 50mM, completados em 6,1 µl de água mili-Q, para um volume final de reação 20 µl. As condições de amplificação foram: um ciclo inicial de desnaturação de 95°C/2min, seguidos de 40 ciclos de 94°C /30 seg, 55°C /45 seg, 68°C /2,5 min e extensão final de 68°C /5 min. Os produtos amplificados foram submetidos a eletroforese em gel de agarose a 1%(p/v), contendo brometo de etídio (0,25µg/ml) e visualizados sob luz ultravioleta (UV).

Tabela 5 : Oligonucleotídeos usados para amplificação dos fragmentos do promotor de SC283.

Oligonucleotídeo	Sequência 5´-3´	Posição
PALTSB1975-Fwd*	ccccgaattcGCATGGGGCCATCAACGGTG	-1975
PALTSB1-Rvs**	ccccaagcttCATGTCCGGCCTAGCTACAAACC	1
PATTSB2838-Fwd*	ccccgaattcGGATCCAGTGAGCTACCGTACC	-2102
PALTSB536-Fwd*	ccccgaattcCCCCGTCTTGCTCTTCAGCTCT	-536

PALTSB1360-Fwd*	ccccgaattcGTCGTCACGTCGACCAAACCTTTC	-1360
PALTSB536-Rvs ***	CcccaagcttAGAGCTGAAGAGCAAGACGGG	-536

*Todos os oligonucleotídeos forward contêm sítio de EcoRI.

**O oligonucleotídeo reverse SC283 contém o ATG putativo, códon inicial da sequência do gene *SbMATE*, seguido do sítio de HindIII.

*** O oligonucleotídeo reverso contendo o sítio de HindIII

Os fragmentos amplificados foram clonados em pCR®8/GW/TOPO, utilizando TOPO clonase (Life Technology), de acordo com as especificações do fabricante (Tabela 6) confirmados por sequenciamento, armazenados em glicerol 50% e mantidos em freezer - 80°C .

Tabela 6: Clones gerados pela inserção de fragmentos do promotor de *SbMATE* da variedade SC283 em TOPO8.

Clones obtidos*	Catálogo do laboratório
-2102 pSbMATE- SC283	pUFV 2003
-1975 pSbMATE- SC283	pUFV 2007
-1360 pSbMATE- SC283	pUFV 2005
-536 pSbMATE- SC283	pUFV 2008
-1975 to-536 pSbMATE-SC283	pUFV 2006
-1360 to-536 pSbMATE-SC283	pUFV 2004

*A enumeração inicial indica a extensão do fragmento do promotor clonado em relação ao possível ATG, a letra p denota promotor, seguido do nome do gene *SbMATE* e da variedade que originou o fragmento do promotor.

Os diferentes fragmentos do promotor foram, em seguida, transferidos por recombinação para pMDC162, utilizando a enzima LR clonase a 25°C por um período de 2 horas. Após o tempo de incubação, as reações de recombinação foram tratadas com proteinase K (2µg/µl), por 15 min a 37°C e utilizadas na transformação de células competentes de *Escherichia coli* da estirpe DH5α, pelo método do choque térmico. Os transformantes positivos foram selecionados com o antibiótico Higromicina (10mg/L), diagnosticados por sequenciamento e armazenados em glicerol 50% e mantidos em freezer -

80°C. Os clones resultantes estão descritos na Tabela 7 e contem o gene repórter GUS sob controle do fragmento do promotor *SbMATE* indicado.

Tabela 7: Clones gerados pela clonagem de fragmentos do promotor de *SbMATE* da variedade SC283 em pMDC162.

Clones obtidos	Catologação do laboratório
-2102 pSbMATE-SC283::GUS	pUFV2028
-1975 pSbMATE-SC283::GUS	pUFV2032
-1360 pSbMATE- SC283::GUS	pUFV2030
-536 pSbMATE- SC283::GUS	pUFV2149
-1975 to-536pSbMATE-SC283::GUS	pUFV2031
-1360 to-536SbMATE-SC283::GUS	pUFV2029

3.3.2 Transformação de *Arabidopsis*

As construções de DNA descritas na Tabela 7 foram utilizadas para transformar *Agrobacterium tumefaciens* estirpe GV3101. As colônias positivas foram diagnosticadas via PCR, utilizando os oligonucleotídeos específicos para os fragmentos dos promotores de *SbMATE* e, em seguida, utilizadas para a transformação de *Arabidopsis thaliana* v. Col 0 pela técnica do mergulhamento floral (Bent, 2002). Após o período necessário para o seu desenvolvimento, as sementes foram coletadas e armazenadas. Os transformantes primários foram selecionados em meio com higromicina 10 mg/L. Na fase de seleção de transformantes, as sementes passaram por um processo de desinfestação em 1 mL de solução de desinfestação (12,5% de hipoclorito de sódio: 37% de água: 50 de etanol 70%), por 15 minutos, seguidos de três lavagens com etanol 100%.

3.3.3 Clonagem dos transfatores putativos em vetor de expressão em plantas

Os quatro (4) transfatores candidatos a interagirem com o promotor de SC283, foram sintetizados comercialmente contendo a metade da região de recombinação AttB1 e AttB2,

clonados em pBSK. Para a reestruturação da região de recombinação AttB inteira, os cDNAs dos genes candidatos foram amplificados por PCR utilizando oligonucleotídeos específicos para a região de recombinação, AttB1 (GGGGACAAGTTTGTACA)-*Foward* e AttB2 (GGGGACCACTTTGTACA)-*Reverse*.

O produto de PCR purificado foi transferido por recombinação (GW) ao vetor de entrada comercial pDONR201, utilizando-se a enzima BP clonase (Invitrogen), segundo especificações do fabricante. O produto de recombinação foi utilizado na transformação de DH5 α . As colônias crescidas em meio com antibiótico kanamicina (10mg/L) foram confirmadas por sequenciamento e geraram os clones listados na Tabela 8. Por possuir a mesma marca de seleção do vetor de destino, os clones em pDONR201 foram digeridos com a enzimas de restrição *ApaI* (Boehringer Mannheim), utilizando 2 μ l desta a 1U, 5 μ l do tampão da enzima 10X, 5 μ l do DNA 1 μ g e completados em 38 μ l de água, para um volume final de 50 μ l, a reação foi realizada a 37°C por 2 horas. O vetor aberto foi purificado com Kit comercial (QIAGEN) utilizados na recombinação com vetor de destino pEarlyGate104, utilizando LR clonase a 25°C *over night*. Após o tempo de incubação, as reações de recombinação foram tratadas com proteinase K (2 μ g/ μ l), por 15 min a 37°C e utilizadas na transformação de células competentes de *Escherichia coli* da estirpe DH5 α , pelo método do choque térmico. Os transformantes positivos foram selecionados com o antibiótico kanamicina (10mg/L), diagnosticados por sequenciamento e armazenados em glicerol 50% e mantidos em freezer -80°C. Os clones resultantes estão descritos na Tabela 8.

Tabela 8: Clones obtidos pela inserção da região codificadora dos genes selecionadas em diferentes vetores de clonagem e expressão em plantas.

Clones obtidos	Clones em pBSK	Clones em pDONR201	Clones em pEarlyGate104
Sorbic.009G166200 (13)	pUFV2458	pUFV2497	pUFV 2603 (p35S:NFY-like)
Sorbic.009G174300 (14)	pUFV2459	pUFV2498	pUFV 2640 (p35S: p WRKY41-like)
Sorbic.009G151400 (16)	pUFV2461	pUFV2499	pUFV 2641 (p35S: p Zincfinger-like)
SORBIDRAFT_09g022540(19)	pUFV2464	pUFV2522	pUFV2647 (p35S:SORBIC09g022540)

3.3.4 Ensaio de transativação em protoplasto de plantas

Com a finalidade de confirmar a atividade de transativação dos transfatores identificados com construções do promotor do gene *SbMATE*, foram realizados ensaios de β -glucuronidase em protoplastos de plantas de *Arabidopsis*, transformadas com os referidos promotores do gene *SbMATE*.

3.3.4.1. Obtenção de protoplastos

Para o isolamento de protoplastos, foram utilizadas *Arabidopsis thaliana* com 15 dias de idade como descrito por Niyogi (2005) com modificações. As plantas foram cortadas com bisturi e incubadas em solução enzimática (Cellulase 1%, Macerozyme 0.2%, Mannitol 0.8M, KCl 1M e MES 100 mM pH 5.7) 10- 15 ml, sob vácuo por 30 segundos e mantidas sob agitação por seis horas com aproximadamente 80 rpm a 30°C. A cada 30 min, a digestão era monitorada examinando as células por microscopia. As folhas digeridas foram filtradas com filtro de nylon de 65 μ M de espessura e centrifugadas a 200 x g por 2 min. O sobrenadante foi descartado e o *pellet* lavado duas vezes em 10 mL de solução de lavagem (NaCl 5M, CaCl₂ 1M, KCl 1M e MES 100 mM), centrifugado a 200 x g por 2 min e ressuspendidos em 1 mL de tampão de eletroporação (Mannitol 0.8 mM, MgCl₂ 1 mM e MES 100 mM).

3.3.4.2 Transfeccção em protoplastos

Para o ensaio de expressão transiente, 2×10^5 a 5×10^6 de protoplastos foram eletroporados a 250 V, 250 μ F com 10 μ g do DNA de interesse, 30 μ g de DNA de esperma de salmão em um volume final de 0.8 mL. Após 36 h de incubação, sendo 12 horas no claro e 24 horas no escuro, os protoplastos foram centrifugados e congelados em nitrogênio líquido.

3.3.4.3 Ensaio da atividade GUS

O ensaio fluorimétrico para avaliar a atividade da β -glucuronidase foi conduzido segundo o protocolo de Jefferson *et al.*, (1987), utilizando-se methylumbelliferone (MU) como padrão. Amostras de três transformantes independentes e dos respectivos controles

foram trituradas e homogeneizadas em 100 µl de tampão GUS [NaH₂PO₄ pH 7.0, 100 mM, EDTA 10 mM, sarcosil 0,1% (p/v), Triton X-100 0,1% (v/v)]. Para a reação de atividade da enzima, 25 µl do extrato foram adicionados a 25 µl do tampão GUS contendo como substrato 2 mM de 4-methylumbeliferil-β-D-glucuronideo (MUG). A mistura então foi incubada a 37°C por 30 minutos e a atividade de GUS foi mensurada utilizando o Leitor Microplacas InfiniteM200pro (TECAN). A concentração de proteínas totais foi determinada pelo método de Bradford (Bradford, 1976).

4. Resultados e discussão

4.1 Identificação *in silico* de fatores transcricionais candidatos que controlam a expressão do gene *SbMATE*

A localização de um eQTL fora do loco *Alt_{SB}* confirma a possibilidade da existência de possíveis fatores localizados em *trans* que afetam a expressão do gene *SbMATE* (dados não publicados). O mapeamento por intervalos múltiplos MIM (*mutiple interval mapping*-MIM, Kao *et al* ., 1999) também foi realizado e observou-se a ocorrência de interação epistática, compatível com um modelo de interação entre transfatores e a região promotora do gene *SbMATE* .

Devido ao fato das regiões físicas dos eQTLs serem extensas, compreendendo vários genes candidatos, foi realizado um estudo de mapeamento associativo (*genome-wide association* – GWAS) para um mapeamento de alta resolução dos transfatores, sendo possível encontrar SNPs associados à tolerância ao Al e à expressão do gene *SbMATE*, co-localizados na região dos eQTLs. Estas análises possibilitaram a identificação prévia de vinte e dois (22) transfatores candidatos, localizados no cromossomo 9, dos quais quatro (Sb09g022810, Sb09g023500, Sb09g021530 e Sb09g022540) foram selecionados para o presente estudo, por demonstrarem maior expressão gênica na presença do alumínio.

As sequencias de nucleotídeos e de aminoácidos correspondentes aos genes candidatos foram inicialmente obtidas no banco de dados do phytozome (anotação do genoma v2.1) e posteriormente re-analisadas na versão mais recente do genoma anotado (v3.1) (Figuras suplementares S1, S2 e S3). O gene Sb09g022540 (Figura suplementar S4), com 528 pb, não foi encontrado na versão mais atual do genoma de *Sorghum bicolor*. Segundo informações disponibilizadas no sistema online phytozome, os *loci* da anotação v2.2 foram provisoriamente mapeados na anotação v3.1 pelo algoritmo BLAT. Os *loci* v2.1 e v3.1 foram comparados par a par, e as proteínas de identidade igual ou superior a 70%, permaneceram com a mesma anotação na versão V3.1. Este critério permitiu o mapeamento de 88% de *loci* v2.1 com v3.1. Aparentemente, o gene Sb09g022540 não se enquadrava no critério de identidade superior a 70% e, desta forma, não foi encontrada na versão mais recente. Entretanto, uma busca no banco de dados por sequencias com alta similaridade identificou a proteína Sorbic.009G163400, como possível Sb09g022540, embora muito maior que a sequencia investigada (1461 pb) (Figura Suplementar S5). O alinhamento entre as duas sequencias evidencia que a proteína de interesse Sb09g022540 provavelmente

representa apenas uma parte da proteína original. Tal fato poderia estar relacionado a um *splicing* alternativo ou outro fator ainda desconhecido (Figura Suplementar S6).

O gene de interesse Sorbic.009g166200 (Sinônimo: Sb09g022810) possui 543pb e codifica uma proteína predita com 180 resíduos de aminoácidos na qual pode ser predito um domínio conservado, similar ao NF-YB (Nuclear transcription factor Y subunit B, Histone-like transcription fator fold), encontrado entre os aminoácidos 68 a 84 [CVSEFISFITSEASDKC], sendo assim designada proteína NF-YB-like (Figura 1).

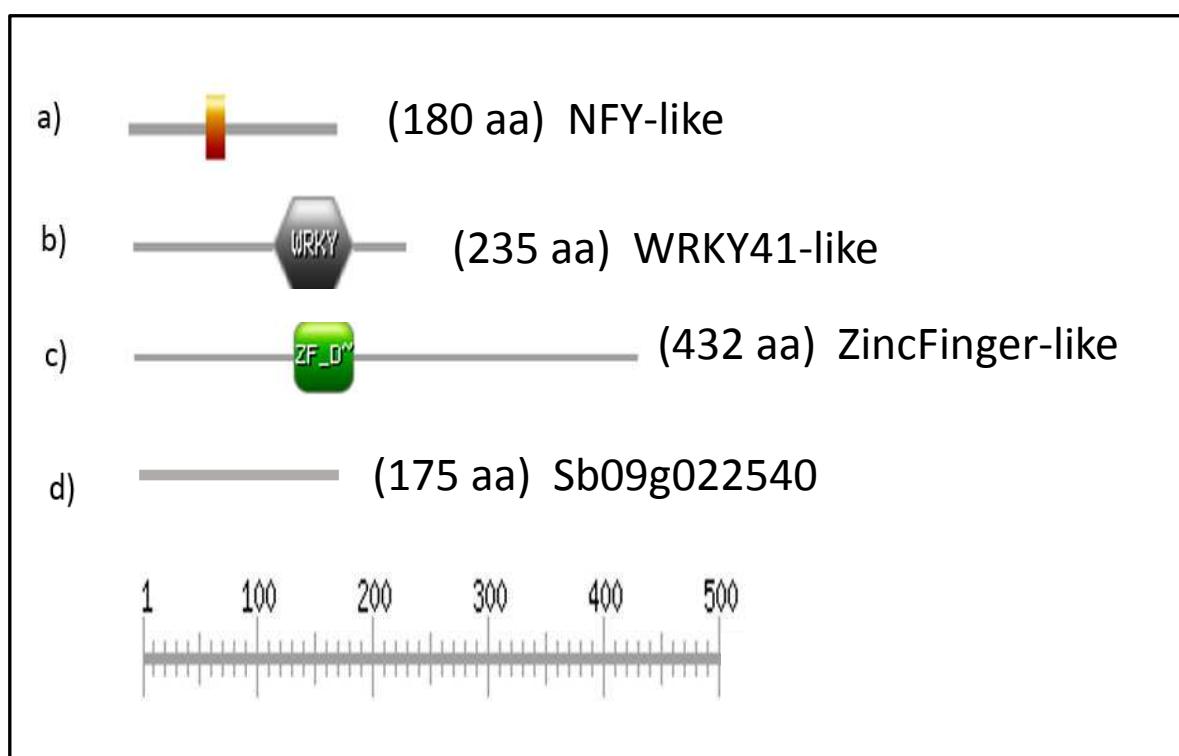


Figura 1: Representação esquemática da estrutura primária dos transfatores preditos, designados na figura. O tamanho das proteínas preditas está representado em escala e o correspondente número de aminoácidos é indicado em parêntesis. Os domínios conservados são esquematizados em cores. Em vermelho, N domínio de histone-like fold ou NF-YB entre os aminoácidos 68-84, em preto, domínio WRKY localizado entre os aminoácidos 122 a 191 e em verde, domínio ZincFinger, posições 138 a 188.

Consistente com uma possível função de fator transcricional, a proteína NF-YB-like possui um sinal de localização nuclear predito. O complexo NF-Y estimula a transcrição de vários genes por meio de interação com os elementos cis-regulatórios CCAAT de promotores de genes em plantas. Como fator transcricional geral de promotores, a perfilação de expressão por RNA-seq de NF-YB-like em sorgo revela uma expressão relativamente alta

em todos os tecidos analisados e condições de tratamento como patógenos, ABA, NAOH e PEG (Figura 2A).

O gene de interesse Sorbic.009g174300 (14) (Sb09g023500), de 706 pb, codifica uma proteína predita de 235 resíduos de aminoácidos, contendo um domínio de ligação ao DNA do tipo WRKY, que compreende os resíduos de aminoácidos da posição 122 a 191 [LTEAPENDHFHWRKYGEKNILYAEYPRLYYKCGysDDHKCPAKKYVQQqSNTYPPLFLVTLNEHTCDTL], além de uma sinal de localização nuclear predito na região delimitado pelos aminoácidos posições 14-58. A proteína codificada por Sorbic.009g174300 foi designada WRKY-like (Figura 1). Interessantemente e consistente com um possível papel regulador na expressão do gene *SbMATE*, o gene WRKY-like é expresso relativamente em baixo nível e exclusivamente em raízes de plântulas em condições normais de crescimento e tratadas com ABA, PEG e NAOH (Figura 2B).

Além desses, o terceiro gene de interesse Sorbic.009G151400 (16) (Sb09g021530) possui 1299 pb e codifica uma proteína de 432 resíduos de aminoácidos onde se pode prever um domínio “Zinc finger do tipo DHHC, localizado entre os resíduos de aminoácido nas posições 138 e 188 [KYCDTCMIYRPPRCSHCNVCVERFDHHCWPVGGQCIGQRNYRYFFMFVS]), designado ZincFinger-like (Figura 1). A proteína ZincFinger-like possui um sinal de localização nuclear predito localizado entre aminoácidos 14-58 e segmentos transmembranas nas posições 35-56, 56-88, 183-205, 228-250. A perfilação de expressão do gene ZincFinger-like revela um padrão de expressão relativamente alto e uniforme em todos os órgãos e condições analisadas (Figura 2C). Para SORBIDRAFT 09g022540, que contém 523 pb, não foi possível encontrar domínios conservados, exibindo uma expressão baixa e específica de anteras e pétalas (Figura 1; Figura suplementar S4, Figura 2D).

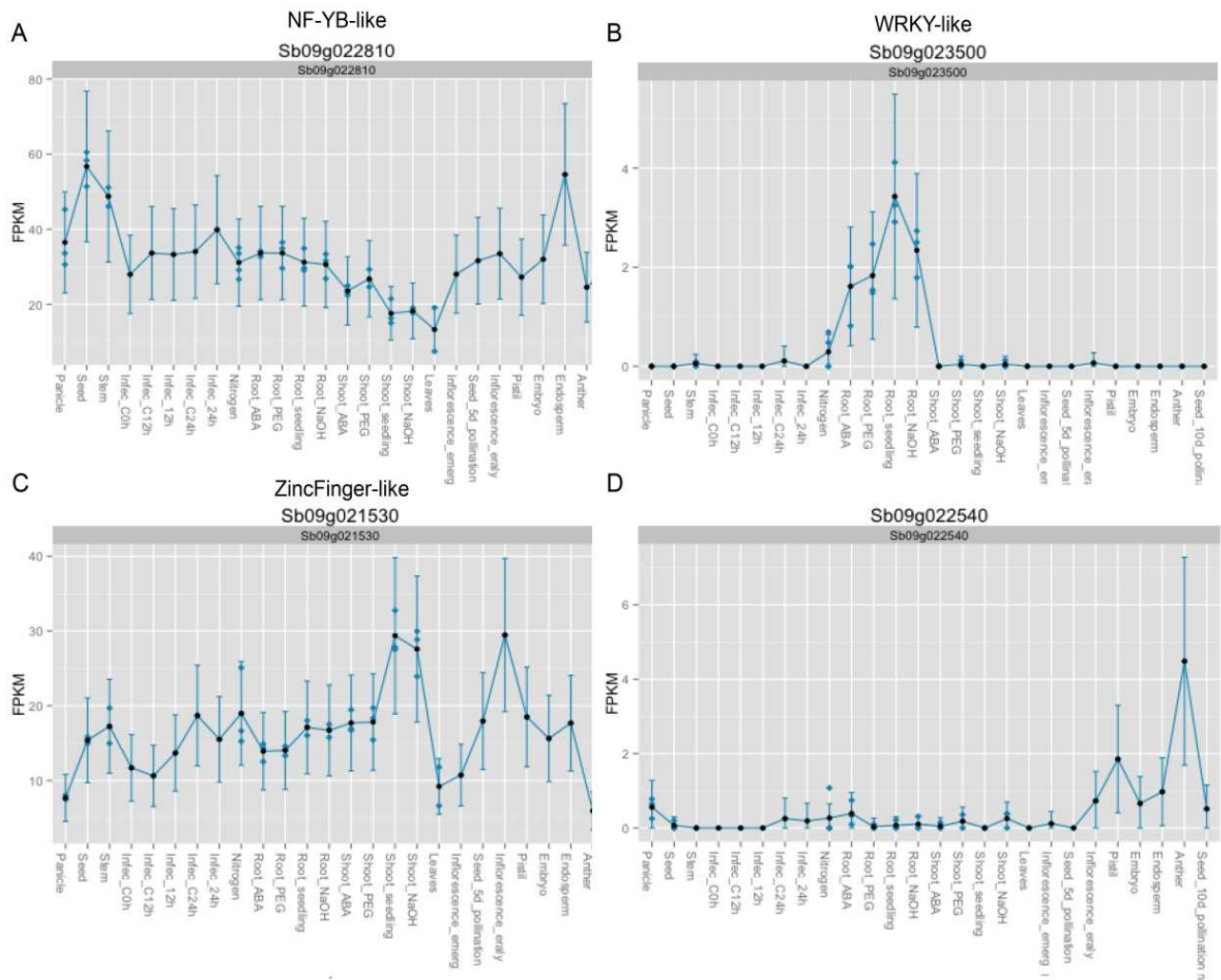


Figura 2: Análise in silico da expressão dos transatores putativos. Os gráficos representam a perfilação de expressão dos putativos transatores, como indicado na figura, em diferentes órgãos de sorgo e condições, conforme indicado na figura e determinado por RNA-seq.

4.2 A proteína WRKY-like, mas não NFY-like, interage estavelmente com o promotor *SbMATE* em leveduras.

Para se identificar possíveis transatores que controlam a expressão do gene *SbMATE*, foram utilizados promotores inteiros do gene *SbMATE*, derivados das variedades SC283, tolerante ao alumínio, Tx430 e BR012, sensíveis ao alumínio (Figura 3), como alvos dos putativos transatores NF-YB-like e WRKY-like por meio do sistema mono-híbrido de leveduras.

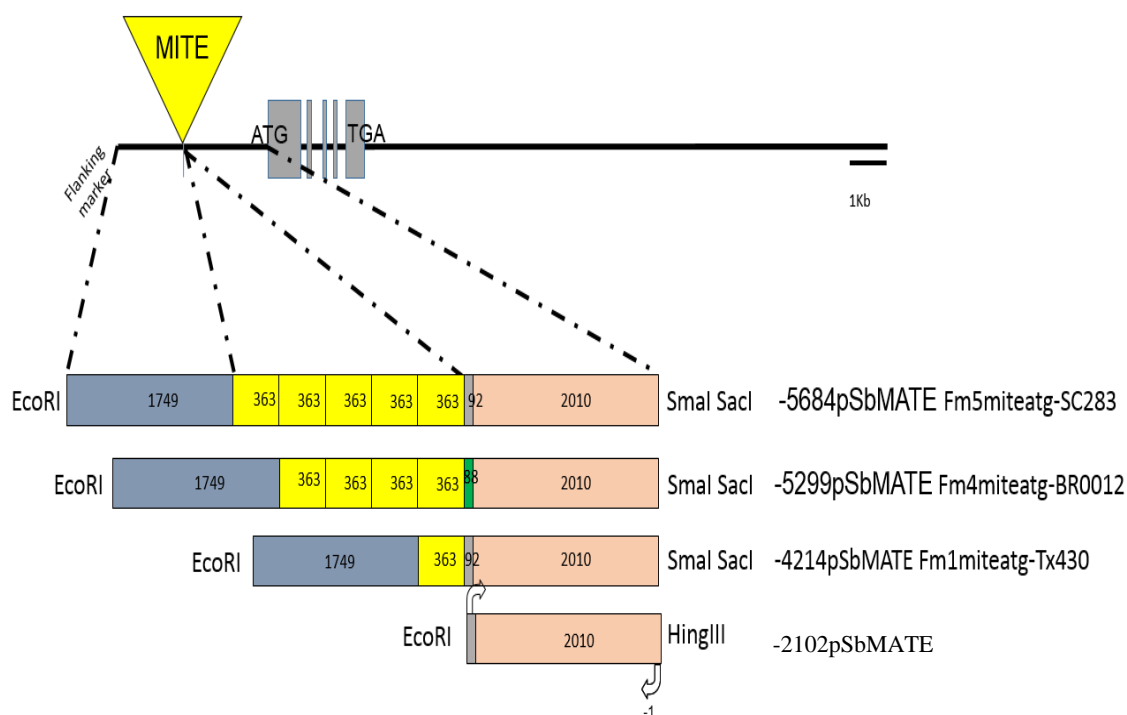


Figura 3. Representação esquemática dos promotores inteiros de *SbMATE* de diferentes variedades de Sorghum, tolerante (SC283) ou sensíveis ao alumínio (BR012 e Tx430), sintetizados comercialmente. Os números representam a extensão das sequências de DNA em pb. As sequências MITE estão indicadas em amarela. A construção truncada é derivada do promotor de *SbMATE* da variedade SC283.

As leveduras W303 foram transformadas com as diferentes construções do promotor fusionadas ao gene HIS3 (pSbMATE-HIS3) e a ativação transcricional do gene repórter HIS3 foi monitorada diante da expressão da proteína NFY-like fusionada ao domínio de ativação de GAL4 (Figura 4). Transformantes duplos independentes foram incubados em meio SD deficiente de Leu e Trp (SD-Leu-Trp, Figura 4) para assegurar a presença de ambos os plasmídeos (pSBMATE-HIS3 e pNF-YB-like-GAL4). Embora, a co-transformação de leveduras tenha ocorrida com alta eficiência, NF-YB-like não foi capaz de ativar o gene repórter, uma vez que as leveduras co-transformadas não foram capazes de crescer em meio deficiente de histidina em qualquer concentração de 3-AT (Figura 4, SD-Leu-Trp-His). Estes resultados indicam que o transfator NFY-like não interage com os promotores testados pelo ensaio mono-híbrido em leveduras (Figura 4).

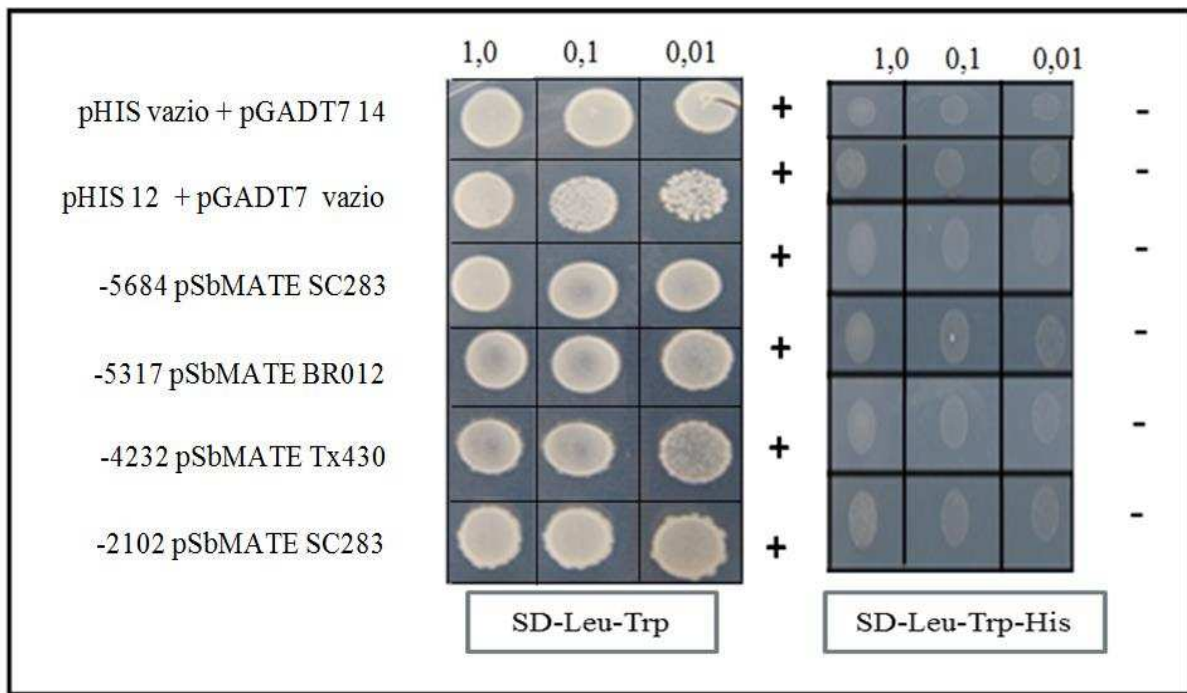


Figura 4: NF-YB-like não interage com promotores de *SbMATE* em leveduras. Ensaio mono-híbrido de leveduras utilizando os promotores de *SbMATE*, conforme esquematizado na Figura 3, e o plasmídeo NF-YB-like::AD). Todas as leveduras co-trasformadas foram crescidas em meio SD deficiente de leucina e triptofano (SD-Leu-Trp) para confirmar a presença de ambas as construções dos plasmídeos recombinantes nas células. A interação entre o promotor indicado à esquerda da figura e o transfator analisado NF-YB-like foi avaliada pelo crescimento de leveduras em meio SD deficiente de leucina, triptofano e histidina (SD-Leu-Trp-His) e suplementado com 3-AT. As duas primeiras linhas são controles negativos do ensaio, vetor do promotor vazio + vetor de expressão contendo o gene NF-YB-like e vice-versa. Diferentes diluições dos transformantes são indicados no topo das figuras.

Similarmente, foi avaliado a capacidade de WRKY-like de interagir com os promotores *SbMATE* em leveduras (Figura 5). Em contraste aos resultados de NF-YB-like, WRKY-like fusionada ao domínio de ativação de GAL4 foi capaz de ativar a autotrofia a histidina/adenina em leveduras co-trasformadas com a construção de todos os promotores *SbMATE* testados. Estes resultados indicam que WRKY-like interage com os promotores *SbMATE* isolados de variedades tolerantes e sensíveis a alumínio e que se diferenciam principalmente pelo número de repetições MITE. Uma vez que o número de repetições MITE tem sido correlacionado com o nível de expressão do gene *SbMATE* em resposta a Al, a possibilidade de que o alvo cis-regulatório de WRKY-like resida na região MITE foi avaliada em ensaios de deleção do promotor. Assim sendo, foi preparada um fragmento truncado do promotor *SbMATE* da variedade tolerante que se estende até a posição -2102 e

representa uma região proximal comum a todos os promotores *SbMATE* testados, mas sem o domínio cis-regulatório MITE.

A expressão de WRKY-like em leveduras co-transformadas com a construção do promotor truncado promoveu o crescimento da levedura modificada na ausência de histidina e presença de 3-AT (Figura 5). A interação entre WRKY-like e a região proximal do promotor *SbMATE* é específica porque WRKY-like não foi capaz de promover autotrofia a His em leveduras co-transformadas com o vetor pHIS3 vazio (Figura 5). Estes resultados indicam que o sitio de ligação ao DNA de WRKY41-like reside na região de 2102 bp, comum a todos os promotores analisados e designada de região proximal.

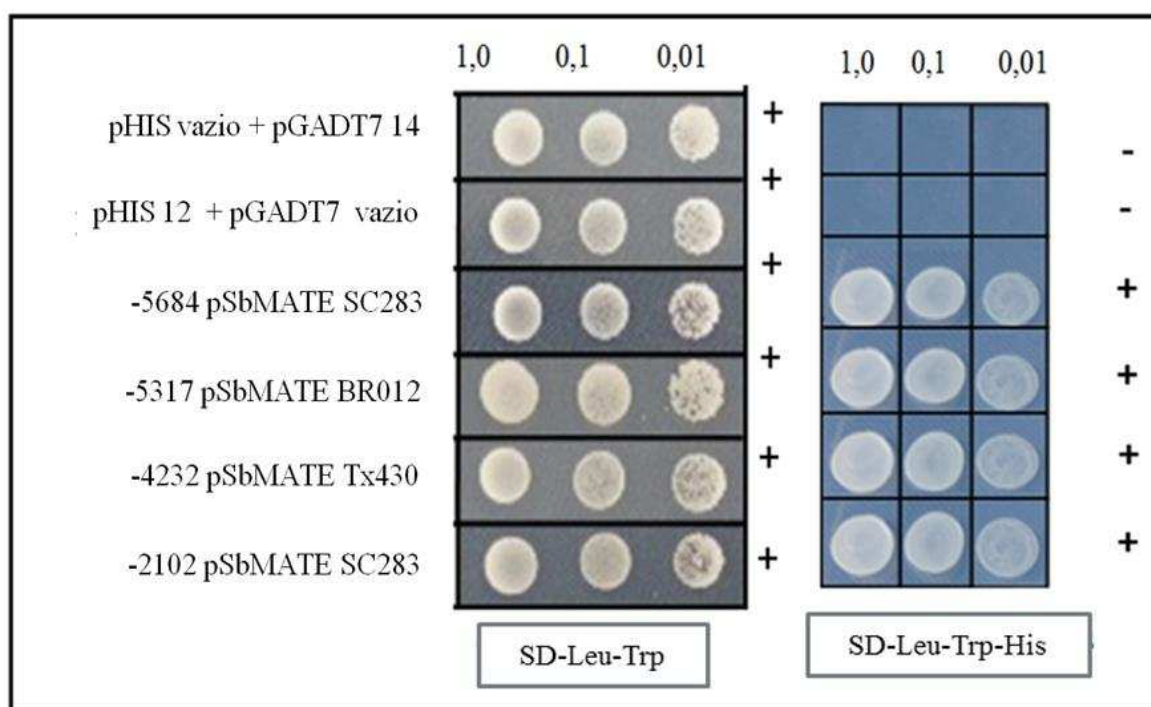


Figura 5: WRKY-like interage com os promotores de *SbMATE* em leveduras. Ensaio mono-híbrido de leveduras utilizando os promotores de *SbMATE*, conforme esquematizado na Figura 3, e o plasmídeo NF-YB-like::AD. Todas as leveduras co-trasformadas foram crescidas em meio SD deficiente de leucina e triptofano (SD-Leu-Trp) para confirmar a presença de ambas as construções dos plasmídeos recombinantes nas células. Crescimento de leveduras em meio SD deficiente de leucina, triptofano e histidina (SD-Leu-Trp-His) e suplementado com 3-AT demonstrou a interação entre o promotor indicado à esquerda da figura e o transfator analisado WRKY-like. As duas primeiras linhas são controles negativos do ensaio, vetor do promotor vazio + vetor de expressão contendo o gene WRKY-like e vice-versa. Diferentes diluições dos transformantes são indicados no topo das figuras.

4.3 Os transfatores WRKY41-like e ZincFinger-like transativam a região proximal do promotor em Arabidopsis thaliana

Os resultados de mono-híbrido em leveduras permitiram mapear o provável sítio de ligação do transfator WRKY-like na região proximal do promotor *SbMATE*. A fim de fornecer evidências complementares de que o gene *SbMATE* é regulado por WRKY-like, foi realizado ensaios de transativação do gene repórter glucuronidase (GUS) em protoplastos de *Arabidopsis*, usando a região proximal do promotor fusionado a GUS.

Estes experimentos permitem avaliar a atividade de ativação transcricional do transfator em plantas. Além do gene WRKY-like, foram incluídos no ensaio os putativos transfatores ZincFinger-like e Sb09g022540. NF-YB-like foi incluído como possível controle negativo, além do vetor de expressão em plantas, pEARLY, vazio. Os protoplastos foram preparados a partir de folhas de *Arabidopsis* estavelmente transformada com o gene GUS sob o controle da região proximal do promotor *SbMATE* (-2396pSbMATE-SC283:GUS). Os resultados de transativação são mostrados pela atividade de GUS, normalizada contra protoplastos de plantas selvagens não transformadas (Figura 6).

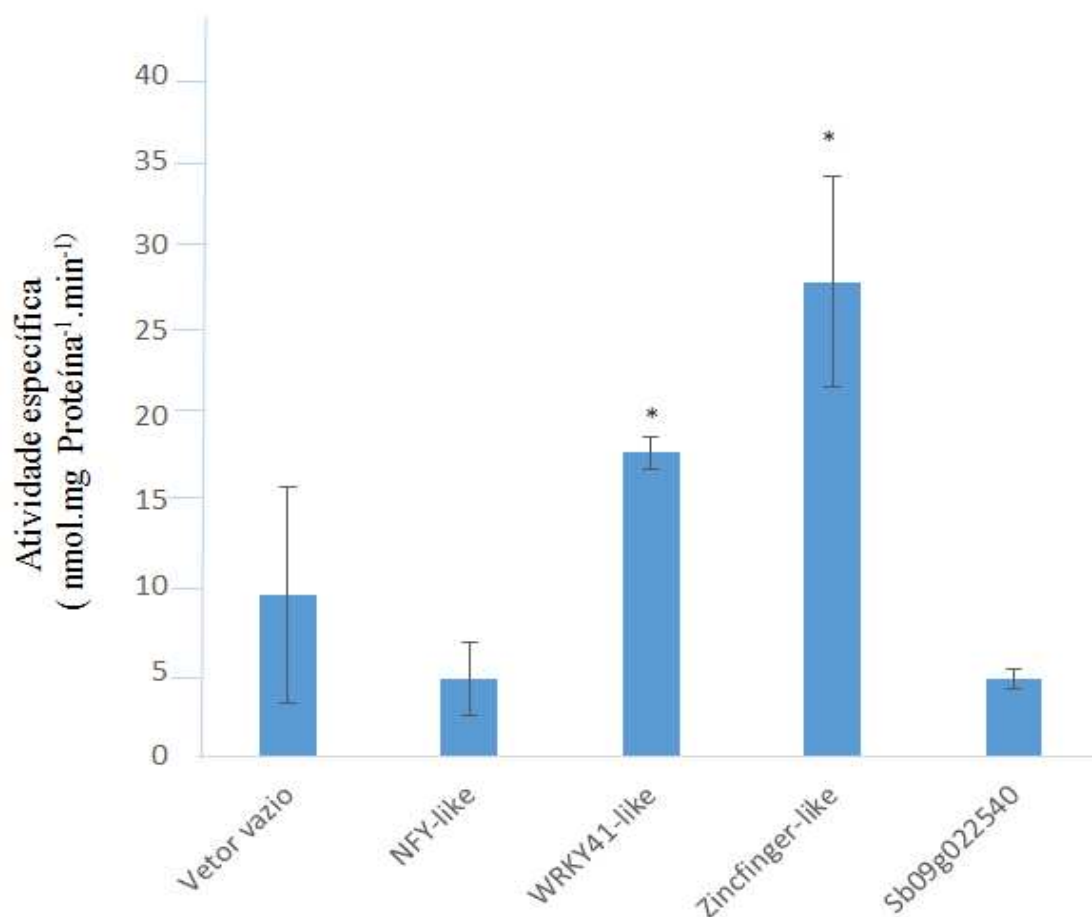


Figura 6. Expressão transiente de WRKY-like e ZincFinger-like em protoplastos de Arabidopsis transativa o promotor *SbMATE* (região proximal). Protoplastos de Arabidopsis estavelmente transformados com -2102pSbMATE-SC23::GUS foram eletroporados com plasmídeos contendo NF-YB-like, WRKY-like, ZincFinger-like, SB09g022540 ou o vetor pEarlyGate104 vazio. Após 36 horas, a atividade de β -glucuronidase foi determinada a partir de extratos proteicos isolados das células transfetadas. As barras indicam o desvio padrão da média (n=3). Os asteriscos indicam médias significativamente diferentes do controle pelo teste Scott-Knott, $p < 0.05$.

Consistente com os resultados de ligação ao promotor em leveduras, WRKY-like, mas não NF-YB-like, especificamente transativou a região proximal do promotor *SbMATE* em planta. Similarmente, o putativo transfator ZincFinger-like foi capaz de especificamente transativar a região proximal do promotor *SbMATE*, enquanto que a expressão de Sb09g022540 não causou ativação do promotor *SbMATE*. A falha de NF-YB-like, que possui características de transfatores gerais, em transativar o promotor *SbMATE* substantia o argumento que os fatores transcricionais ZincFinger-like e WRKY-like controlam o promotor *SbMATE* de uma forma específica.

4.4. Os sítios de ligação ao DNA dos fatores transcrpcionais WRKY-like e ZincFinger-like estão contidos em uma região de 146 bp, posições -2386 a -1975, do promotor *SbMATE*.

Para mapear o sítio de ligação ao DNA de WRKY-like e ZincFinger-like, foram feitas deleções na extremidade 5' (trunicações até -1975, -1360 e -536) e ampliações internas (-1975 a -536; -1360 a -536) da região proximal do promotor *SbMATE* gerando 6 fragmentos truncados do promotor da variedade tolerante SC283 que foram fusionados ao gene repórter GUS (Figura 7) e utilizados para transformar *Arabidopsis*.

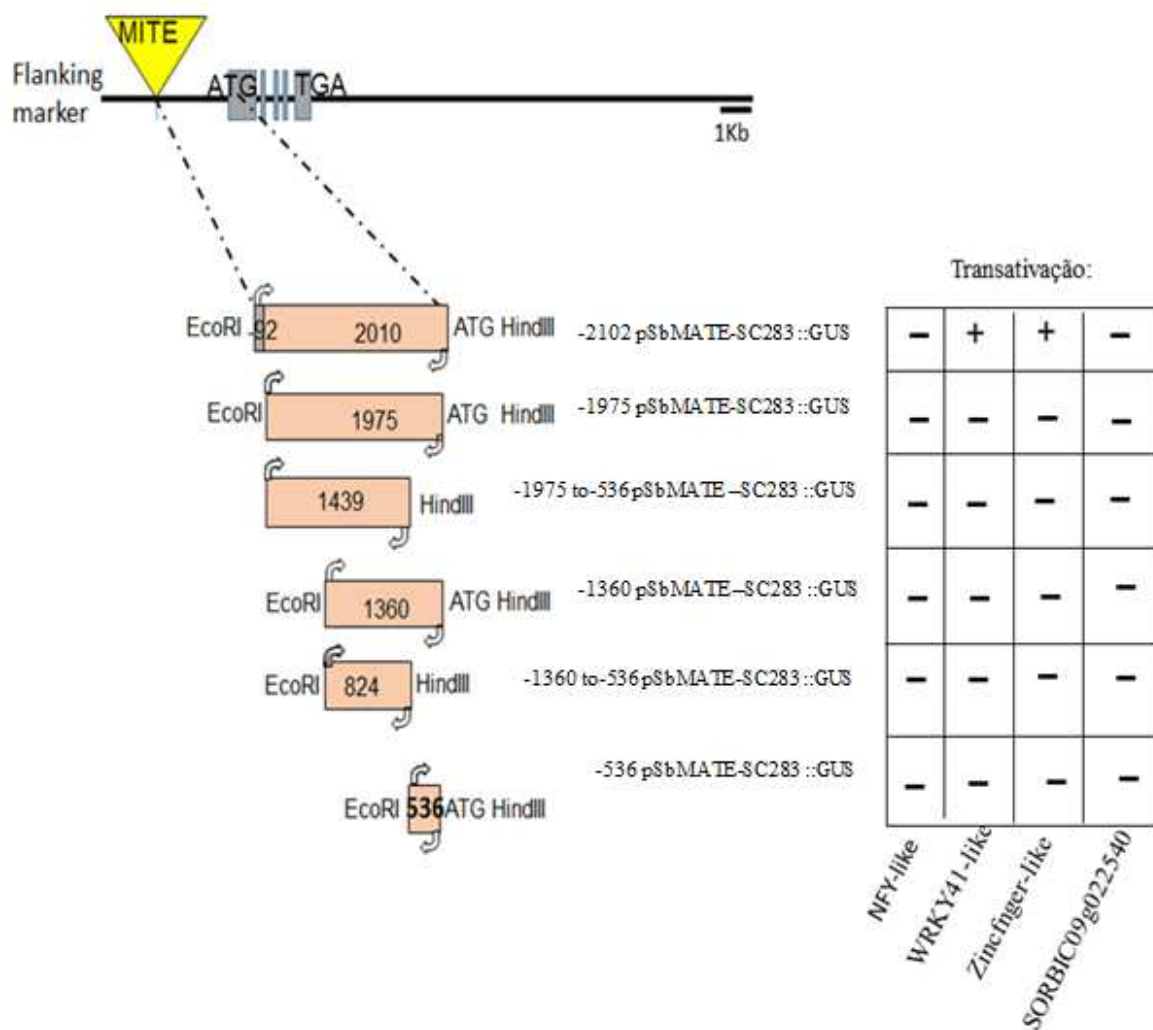


Figura 7: Representação esquemática de deleções da região proximal do promotor *SbMATE* fusionadas a GUS e usadas para transformação de *Arabidopsis*. O quadro à direita sumariza os resultados de transativação dos referidos fragmentos de promotores pelos transfatores indicados na figura.

A capacidade dos fatores transcripcionais WRKY-like e ZincFinger-like em transativar o gene repórter foi avaliada em ensaios de transativação em protoplastos de *Arabidopsis* estavelmente transformadas com as construções do promotor (Figura 8). Os resultados também são sumarizados na Figura 7.

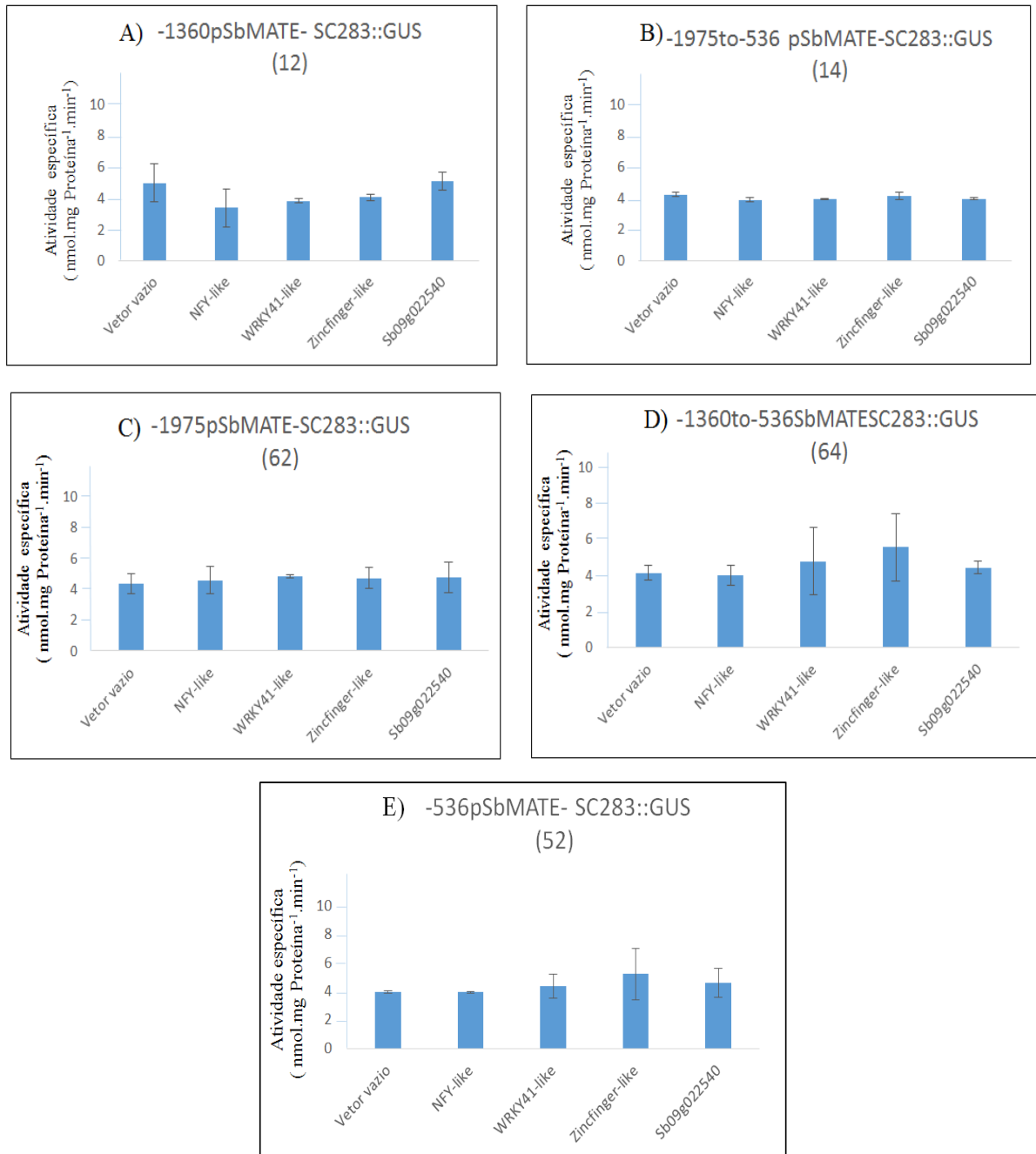


Figura 8. Análise de transativação de diferentes fragmentos do promotor *SbMATE* pelos transfatores indicados na figura. Protoplastos de *Arabidopsis* estavelmente transformados com as construções do promotor fusionadas a GUS, conforme indicado na figura, foram eletroporados com plasmídeos contendo NF-YB-like, WRKY-like, ZincFinger-like, SB09g022540 ou o vetor pEarlyGate104

vazio. Após 36 horas, a atividade de β -glucuronidase foi determinada a partir de extratos proteicos isolados das células transfectadas. As barras indicam o desvio padrão da média (n=3).

A deleção da região proximal até a posição -1975 aboliu completamente a transativação do promotor *SbMATE* tanto por WRKY-like quanto por ZincFinger-like. Estes resultados estabelecem que o sítio de ligação ao DNA dos dois fatores transcricionais está contido na região deletada de 127 bp, delimitada pelas posições -2102 a -1975. Consistente com esta interpretação nenhum outro fragmento do promotor foi transativado pelos transfatores WRKY-like e ZincFinger-like (Figuras 7 e 8).

Estes resultados indicam que o sítio provável de interação dos transfatores é delimitado pela posição -2102 a -1975. Essa sequência de 127 pb foi submetida à análise para a presença de cis-elementos no programa Plant CARE (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>). Após análise, foram encontradas potenciais cis elementos regulatórios que respondem a variações do meio ambiente e cis-elementos gerais de transcrição como CAAT box e TATA box (Figura 9; Tabela Suplementar S1).

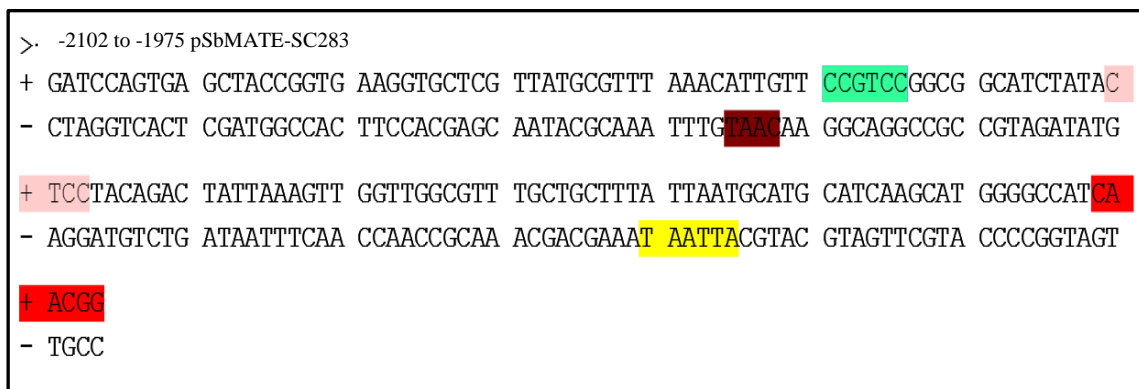


Figura 9: Análise dos cis- elementos regulatórios da sequência de 127pb presente no fragmento -2102 to -1975 pSbMATE-SC283::GUS (32). Os cis-elementos estão identificados em diferentes colorações, podendo contabilizar um total de 6 elementos preditos.

Um sítio *A-box* (em verde), definido como cis -acting elemento regulador associado com P e L -box envolvidos na indução da atividade transcricional, requerido mas suficiente para conferir capacidade de resposta a luz, esta localizado na posição 51. Além desse elemento de resposta a estresse, destaca-se um *Box 4* (em amarelo), que representa um cis-

elemento descrito como parte de uma matriz conservada de DNA (pal- CMA1) envolvido com a capacidade de resposta luz, localizado na posição -110. Destacam-se ainda um CAAT *box* (em vinho), localizado na posição -45 comum ao elemento cis em promotor e *enhancer* e um CCAAT *box* (em vermelho), na posição +139, descrito como possível local de ligação de MYBHv1 que está de acordo com a sequência alvo consenso pyAAC(G;T)G. Como representante de cis-elementos que confere tecido-especificidade, observa-se a presença da sequência consenso CCGTCC na posição + 51 (em verde), descrita como elemento regulador cis -acting relacionado com ativação do meristema (Tabela Suplementar S1). Além desses, mas sem importância biológica já que é dependente de posição em genes eucarióticos (posição -30), diversos TATA *boxes* foram identificados. Estudos de mutagênese dirigida in situ no domínio cis-regulatório de 127 bp deverão ser conduzidos para precisar os sítios de ligação dos transfatores WRKY-like e ZincFinger-like.

4.5 Os genes WRKY-like e ZincFinger-like de *Sorghum bicolor* são amplamente distribuídos no reino vegetal.

As sequências dos transfatores WRKY-like e ZincFinger-like de *Sorghum bicolor* foram usadas como protótipos para identificação de genes ortólogos no genoma de *Arabidopsis thaliana*, *Carica papaya*, *Citrus sinensis*, *Cucumis sativus*, *Glycine max*, *Manihot esculenta*, *Phaseolus vulgaris*, *Solanum lycopersicum*, *Solanum tuberosum*, *Triticum aestivum*, *Oryza sativa* and *Zea mays*. Foram selecionados os cinco genes mais relacionados de cada espécie de plantas para construir árvores filogenéticas (Figuras 10 e 11).

Com base na genealogia, ambos os transfatores WRKY-like e ZincFinger-like formaram clados independentes com genes de monocotiledôneas (Figuras 10 e 11). O gene WRKY-like é mais relacionado com o ortólogo de milho GRMZM2G034421_P01, com o qual compartilha 82% de identidade e 85% de similaridade de sequência dentro de uma cobertura de sequências de 85%, seguido pelo gene de trigo Traes_1AL_309623B48.1 e o gene de arroz LOC_Os05g40070.1 (Tabela Suplementar S2).

GRMZM2G100641_P01 de milho (91% de identidade, 94% de similaridade de sequencias e 96% de cobertura), seguido pelo gene LOC_Os05g36090.1 de arroz (82% de identidade, 89% de similaridade e 96% de cobertura) e pelo gene Traes_1BL_19E07DA0F.1 de trigo (80% de identidade, 89% de similaridade e 96% de cobertura). Este alto grau de conservação entre ortólogos de diferente espécies sugere que os transfatores WRKY-like e ZincFinger-like exerce funções conservadas em diferentes espécies de planta.

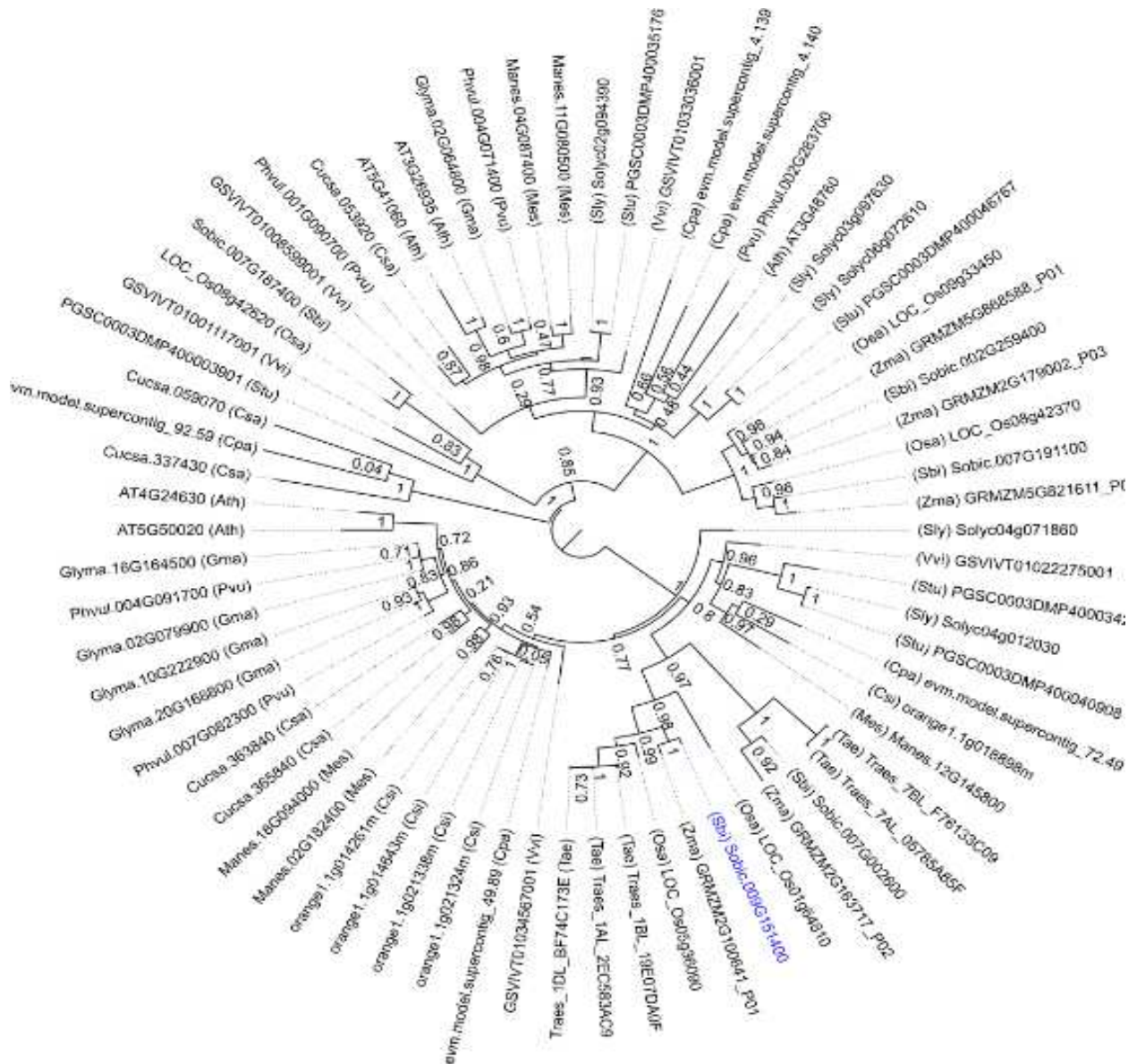


Figura 11: Análise filogenética do transfaktor ZincFinger –like (Sorbic09g151400(16)). A sequencia de aminoácidos dos cinco genes mais relacionados das espécies indicadas foram recuperadas do banco de dados TAIR (<http://arabidopsis.org/>) e Phytozome v10.3 e alinhadas usando MUSCLE. As arvores filogenéticas foram construídas utilizando método de máxima verossimilhança, com 10.000 réplicas de bootstraps. O gene ZincFinger-like (Sorbic09g151400(16)) está destacado em azul. Os nomes das espécies de plantas estão indicados em parêntesis.

As proteínas contendo domínio WRKY formam uma família ampla de fatores de transcrição específicos de plantas, sendo que tem sido relacionadas com controle de expressão gênica como defesa e em resposta a estresses. Nenhuma informação funcional adicional existe com relação aos ortólogos de WRKY-like do sorgo. Sendo assim, a caracterização funcional de WRKY-like deverá ser conduzida por meio de abordagens bioquímicas e genéticas. Similarmente, características estruturais relevantes do gene ZincFinger-like foram identificadas, como, por exemplo, a presença de domínios transmembranas. A ancoragem de fatores de transcrição em membranas representa um aspecto regulatório eficiente utilizado pelas células eucarióticas para manter o gene inativo na ausência de estímulos. Nestes casos, a presença do estímulo promove a liberação do fator de transcrição da membrana e subsequente translocação para o núcleo onde ativa a expressão um conjunto seletivo de genes contendo seu elemento cis-regulatório. Assim sendo, a determinação da localização subcelular do transfator ZincFinger em condições normais e em resposta a estímulos deverá ser conduzida como etapa fundamental para se entenderem os mecanismos regulatórios que controlam o gene *SbMATE*.

5. Conclusões e Perspectivas

O gene *SbMATE* confere resistência a Alumínio, exibe uma expressão tecido-específica, predominantemente no ápice das raízes e é induzido por Al em sorgo. Apesar da relevância desse gene como alvo potencial para desenvolvimento de estratégias moleculares visando à obtenção de cultivares tolerantes à Al, as bases moleculares que controlam a expressão do referido gene não foram ainda elucidadas. Nesta investigação, foram fornecidas diversas linhas de evidência que conclusivamente demonstram que os fatores transcricionais WRKY-like (Sorbic.009g174300) e ZincFinger-like (Sorbic.009g151400) contribuem positivamente para o controle da expressão do gene *SbMATE*.

Inicialmente, foi demonstrado que o transfator WRKY-like interage com os promotores do gene *SbMATE* isolados de variedades tolerantes e suscetíveis ao Al em leveduras. A ligação de WRKY-like ao promotor *SbMATE* é específica, porque o transfator NFY-B-like e o vetor de expressão pEarlyGate104 vazio não interagem com nenhum fragmento do promotor em leveduras. A interação ocorre em uma região do promotor comum a todos os promotores *SbMATE* testados que se estende do primeiro códon ATG até a posição -2102, designada de região proximal do promotor.

Consistente com um papel no controle da expressão do gene *SbMATE*, o gene WRKY-like é expresso exclusivamente na raiz. Além disso, em ensaios de transativação do gene repórter GUS em protoplastos de *Arabidopsis* transformados com a região proximal do promotor *SbMATE* fusionados a GUS, foi demonstrado que, não somente WRKY-like, mas também ZincFinger-like ativam o promotor de uma maneira específica. Em contraste, os transfatores não relacionados, como NF-YB-like e SORBIDRAFT_09g022540 não foram capazes de se ligarem ao promotor em leveduras e ativar a expressão do gene repórter em protoplastos de *Arabidopsis*. Por meio de ensaios de transativação em protoplastos com regiões truncadas do promotor *SbMATE*, os possíveis sítios de ligações dos transfatores WRKY-like e ZincFinger-like foram mapeados em um domínio cis-regulatório de 127 pb, delimitado pelas posições -2102 a -1974.

A relativa proximidade dos sítios de ligação ao DNA dos transfatores WRKY-like e ZincFinger-like no domínio cis-regulatório de 127 bp do promotor *SbMATE* pode ser um indicativo de que os dois transfatores formam dímeros para ativação da expressão do gene *SbMATE*. Ensaios de interação de proteínas, como duplo híbrido em leveduras, co-immunoprecipitação e/ou de complementação de fluorescência bimolecular (BiFC), juntamente

com ensaios de transativação de genes repórteres com a expressão simultânea dos dois transfatores, poderão testar esta hipótese. Além disso, os elementos cis-regulatórios alvos de WRKY-like e ZincFinger-like poderão ser precisamente identificados por meio de ensaios de mutagêneses dirigida in situ do domínio cis-regulatório.

6. Referencias bibliográficas

- Ahn S J, Sivaguru M, Osawa H, Chung GC, Matsumoto H** (2001). Aluminum inhibits the H⁺-ATPase activity by permanently altering the plasma membrane surface potentials in squash roots. *Plant Physiology*, **126**, 1381-1390.
- Alscher RG, Erturk N, Heath LS** (2002) Role of superoxide dismutases (SODs) in controlling oxidative stress in plants. *Journal of Experimental Botany*, **53**: 1331-1341.
- Apel K, Hirt H** (2004). REACTIVE OXYGEN SPECIES: Metabolism, Oxidative Stress, and Signal Transduction *Annu. Rev. Plant Biol.* 2004. **55**:373–99.
- Barceló J, Poschemrieder C** (2002). Fast root growth responses, root exudates, and internal detoxification as clues to the mechanisms of aluminium toxicity and resistance: a review. *Environmental and Experimental Botany*, **48**, 75-92
- Calbo A G, Cambraia J** (1980) Efeito do alumínio sobre a composição mineral de dois cultivares de sorgo (*Sorghum bicolor* L. Moench). *Revista Ceres*, 27:369-378.
- Chang Y C, Ma J F, Matsumoto H** (1998) Mechanisms of Al-induced iron chlorosis in wheat (*Triticum aestivum*) Al-inhibited biosynthesis and secretion of phytosiderophore. *Physiologia Plantarum*, **102**:9-15.
- Chen J, Sucoff E I, Stadelmann E J** (1991). Aluminum and temperature alteration of cell membrane permeability of *Quercus rubra*. *Plant Physiol*, **96**:644-649.
- Coelho A M, Waquil J M, Karan D, Casela C R, Ribas P M** (2002). Seja o doutor do seu sorgo. *Informacoes Agronomicas*, Piracicaba, n.100, dez. 2002. Arquivo do Agronomo, Piracicaba, n.14, 24p. dez. Encarte.
- Degenhardt, j.; Jarsen, p.b.; Howell, s.h. and Kochian, l.v. (1998)**. Aluminum resistance in the *Arabidopsis* mutant alr-104 is caused by an aluminum-induced increase in rhizosphere pH *Plant Physiol.* 117:19-27.
- Delhaize E, Ryan P R** (1995). Aluminum toxicity and tolerance in plants. *Plant Physiology*, **107**, 315-321.
- Delhaize E, Gruber B D, Ryan P R** (2007). The roles of organic anion permeases in aluminium resistance and mineral nutrition. *FEBS Letters* 581, 2255–2262 doi:10.1016/j.febslet.2007.03.057
- Dicko M H, Gruppen H, Traoré A S, Voragen A G J, van Berkel W J H** (2006). Review Sorghum grain as human food in Africa: relevance of content of starch and amylase activities *African Journal of Biotechnology* Vol. **5** (5), pp. 384-395, 1 March.

Edgar RC (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32:792-1797.

Freitas F A, Kopp M M, Sousa R O, Zimmer P D, Carvalho F I F, Oliveira A C (2006). Absorção de P, Mg, Ca e K e tolerância de genótipos de arroz submetidos a estresse por alumínio em sistemas hidropônicos. *Ciência Rural* 36:72-79.

Furukawa J, Yamaji N, Wang H, Mitani N, Murata Y, Sato K, Katsuhara M, Takeda K, Ma JF. (2007). An aluminum-activated citrate transporter in barley. *Plant and Cell Physiology* 48, 1081–1091.

Huang J W, Pellet D M, Papernik L A (1996) Aluminum interactions with voltage-dependent calcium transport in plasma membrane vesicles isolated from roots of aluminum-sensitive and -resistant wheat cultivars. *Plant Physiology*, v.110, n.2, p.561-569, 1996.

Huang C F, Yamaji N, Mitani N, Yano M, Nagamura Y, Ma J F (2009). A bacterial-type ABC transporter is involved in aluminum tolerance in rice. *The Plant Cell*, Vol. 21, 655-667.

Huang C F, Yamaji N, Chen Z, Ma J F (2012) A tonoplast-localized half-size ABC transporter is required for internal detoxification of aluminum in rice. *Plant J.* 69, 857–867.

Iuchi S, Koyama H, Iuchi A, Kobayashi Y, Kitabayashi S, Kobayashi Y, Ikka T, Hirayama T, Shinozaki K, Kobayashi M. (2007) Zinc finger protein STOP1 is critical for proton tolerance in Arabidopsis and coregulates a key gene in aluminum tolerance. *Proc. Natl. Acad. Sci. U.S.A.* 104, 9900–9905.

Jefferson, R. A., Kavanagh, T. A., & Bevan, M. W. (1987). GUS fusions: β -glucuronidase. *The EMBO Journal*, 6(13), 3901–3907.

Jones D L, Kochian L V (1997). Aluminum interaction with plasma membrane lipids and enzyme metal binding sites and its potential role in Al cytotoxicity. *FEBS Letters* 400 (1997) 51-57.

Kochian L V (1995). Mechanisms of aluminium toxicity and resistance in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 46:237-60

Kochian L V, Hoekenga O A, Pineros M A (2004). How do crop plants tolerate acid soils? Mechanisms of aluminum tolerance and phosphorous efficiency. *Annual Review Plant Biology*, 55, 459-493.

- Kochian L V, Pineros M A, Hoekenga O A** (2005). The physiology, genetics and molecular biology of plant aluminum resistance and toxicity. *Plant and Soil*, **274**, 175-195. DOI 10.1007/s11104-004-1158-7.
- Larsen P B, Geiseler M J B, Jones CA, Williams K M, Cancel J D** (2005) ALS3 encodes a phloem-localized ABC transporter-like protein that is required for aluminum tolerance in *Arabidopsis*. *The Plant Journal*, **41**: 353-363.
- Larsen P B, Cancel J D, Rounds M, Ochoa V** (2007) Arabidopsis ALS1 encodes a root tip and stele localized half type ABC transporter required for root growth in an aluminium toxic environment. *Planta*, **225**:1447-1458
- Lazof D B, Goldsmith J G, Ruffy T W, Linton R W** (1996) The early entry of Al into cells of intact soybean roots (a comparison of three developmental root regions using secondary ion mass spectrometry imaging). *Plant Physiology*, **112**, 1289-1300.
- Liang, C., M. A. Pineros, J. Tian, Z. Yao, L. Sun** (2013) Low pH, aluminum, and phosphorus coordinately regulate malate exudation through GmALMT1 to improve soybean adaptation to acid soils. *Plant Physiol.* **161**: 1347–1361.
- Liu JP, Magalhaes JV, Shaff J, Kochian LV** (2009). Aluminum-activated citrate and malate transporters from the MATE and ALMT families function independently to confer *Arabidopsis* aluminum tolerance. *The Plant Journal* **57**, 389–399.
- Ma J F** (2000) Mini Review: Role of organic acids in detoxification of aluminum in higher plants. *Plant & Cell Physiology*, **41**(4), 383-390.
- Ma J F, Ryan P R, Delhaize E** (2001) Aluminium tolerance in plants and the complexing role of organic acids. *Trends in Plant Science*, **6**, 273-278.
- Ma J F** (2007) Syndrome of aluminum toxicity and diversity of aluminum resistance in higher plants. *International Review of Cytology*, **264**, 225-253.
- Magalhães J V, Garvin D F, Wang Y, Sorrells M E, Klein P E, Schaffert R E, Li L, Kochian L V** (2004) Comparative mapping of a major aluminum tolerance gene in sorghum and other species in the Poaceae. *Genetics* **167**: 1905-1914.
- Magalhães J V, Liu J, Guimarães C T, Lana U G P, Alves V M C, Wang Y-H, Schaffert R E, Hoekenga O A, Piñeros M A, Shaff J E, Klein P E, Carneiro N P., Coelho C M, Trick H N, Kochian L V** (2007) A gene in the multidrug and toxic compound extrusion (MATE) family confers aluminum tolerance in sorghum. *Nature Genetics*, **39**, 1156-1161.

- Magalhães JvV** (2010) Review: Part Of A Highlight Section On Plant–Soil Interactions At Low Ph How a microbial drug transporter became essential for crop cultivation on acid soils: aluminium tolerance conferred by the multidrug and toxic compound extrusion (MATE) family. *Annals of Botany* **106**: 199–203, 2010 doi:10.1093/aob/mcq115.
- Maron LG, Pineros MA, Guimaraes CT, Magalhaes JV, Pleiman JK, Mao CZ, Shaff J, Belicuas SNJ, Kochian LV.** (2010). Two functionally distinct members of the MATE (multi-drug and toxic compound extrusion) family of transporters potentially underlie two major aluminum tolerance QTLs in maize. *The Plant Journal* **61**, 728–740.
- Matsumoto H, Hirasawa F, Torinkai H** (1976) Localization absorbed aluminium in peã root and its binding to nuclei acid. *Plant Cell Physiology* **17**:627-631.
- Melo J O, Lana U G P, Pineros M A, Alves V M C, Guimarães G T, Liu J, Zheng Y, Zhong S, Fei Z, Maron L G, Schaffert R E, Kochian L V, Magalhães J V** (2013) Incomplete transfer of accessory loci influencing SbMATE expression underlies genetic background effects for aluminum tolerance in sorghum *The Plant Journal* **73**, 276–288 doi: 10.1111/tpj.12029.
- Mendonça R J, Cambraia J, Oliveira J A, Oliva M A** (2003) Efeito do alumínio na absorção e na utilização de macronutrientes em duas cultivares de arroz. *Pesquisa Agropecuaria Brasileira*, **38**:843-848.
- Moldes C A** (2006). Respostas de enzimas antioxidantes a aplicação do herbicida glifosato em variedades de soja transgênicas e não transgênicas- Tese de doutorado.
- Mutisya J, Sun C, Rosenquist S, Baguma Y, Jansson C** (2009) Diurnal oscillation of SBE expression in sorghum endosperm. *Journal of Plant Physiology*, Stuttgart, v. **166**(4) p. 428-434.
- Miyasaka, s. And Hawes, m.c.** (2001). Possible role of root border cells in detection and avoidance of aluminum toxicity. *Plant Physiol.* **125**:1978-1987.
- Nguyen N T, Hiep N D, Fujita K** (2005) Iron enhances aluminum-induced leaf necrosis and plant growth inhibition in *Eucalyptus camaldulensis*. *Plant and Soil* **277**:139-152.
- Niyogi, K.K.** (2005) Arabidopsis protoplast isolation and transformation. bsw3.naist.jp/plantglobal/protocol/protoplast.pdf
- Ofei-Manu P, Watsunga T, Ishiwaka** (2001) The plasma membrane strength of the root-tip cells and phenolic compounds are correlated with AL tolerance in several common woody plants. *Soil Sci Plant Nutrition* **47**:3559-375.
- Ohki K** (1987). Aluminum stress on Sorghum growth and nutrient relationship. *Plant Soil* **98**:195-202.

- Panda S K, Yamamoto Y, Kondo H, Matsumoto H.**(2008) Mitochondrial alterations related to programmed cell death in tobacco cells under aluminium stress . *C. R. Biologies* **331** 597–610.
- Panda S K, Baluska F, Matsumoto H** (2009) Aluminum stress signaling in plants. *Plant Signaling & Behavior*, **4**, 592-597.
- Peixoto P H P, Cambraia J, Sant'anna R, Mosquim P R, Moreira M A** (2001) Aluminum effects on fatty acid composition and lipid peroxidation of a purified plasma membrane fraction of root apices of two sorghum cultivars. *Journal of Plant Nutrition*, **24**, 1061-1070.
- Pereira J M, Cambraia J, Meira F J E, Ribeiro C** (2008) Efeito do alumínio sobre a absorção, o acúmulo e o fracionamento do fósforo em sorgo Bragantia: revista de ciencias agronómicas, Vol. 67, Núm. 4, octubre-diciembre, pp. 961-967.
- Peter R. Ryan*, Harsh Raman, Sanjay Gupta 2 , Walter J. Horst, And Emmanuel Delhaize** (2009). A second mechanism for aluminum resistance in wheat relies on the constitutive efflux of citrate from roots; *Plant Physiology*, january, vol. 149, pp. 340–351, www.plantphysiol.org 2008 american society of plant biologists.
- Price, Morgan N., Paramvir S. Dehal, and Adam P. Arkin** (2010):. "FastTree 2—approximately maximum-likelihood trees for large alignments." *PloS one* **5.3** e9490.
- Queiroz V A V, Vizzotto M, Carvalho C W P, Martino H S D** (2009)O Sorgo na Alimentação Humana Circular Técnica **133** Embrapa ISSN 0100-9915.
- Queiroz V A V, Gonçalves A C A, Rodrigues J A S, Schaffert R E** (2011)Como Utilizar Grãos de Sorgo no Preparo de Produtos Caseiros Circular Técnica **190** Embrapa ISSN 1679-0162.
- Rambaut A** (2015) FigTree v1.4: Tree figure drawing tool. Website: <http://tree.bio.ed.ac.uk/software/figtree/>
- Rengel Z, Zhang W H** (2003). Role of dynamics of intracellular calcium in aluminium-toxicity syndrome. *New Phytologist*, **159**, 295-314.
- Richards, K.D.; Schott, E.J.; Sharma, Y.K.; Davis, K.R. and Gardner, R.C.**(1998)Aluminum Induces Oxidative Stress Genes in *Arabidopisthaliana* *Plant Physiol.* (1998) **116**: 409–418.
- Ribas P M** (2003) Sorgo: Introdução e Importância Econômica. Documentos **26** Embrapa ISSN 1518-4277.

- Rodrigues P H M, Senatore A L., Lucci C S, Andrade S J T, Lima F R, Melotti L**(2002)Valor nutritivo da silagem de sorgo tratada com inoculantes enzimo-microbianos. *Acta Scientiarum, Londrina*, v. **24**, p.1141-1145.
- Rout G R, Samantaray S, Das P** (2001) Aluminium toxicity in plants: A review *Agronomie* **21**:3-21.
- Ryan P R, Delhaize E** (2010)The convergent evolution of aluminium resistance in plants exploits a convenient currency. *Functional Plant Biology*, **37**, 275–284.
- Sade H, Meriga B, Surapu V, Gadi J, Sunita M. S. L, Suravajhala P, Kavi Kishor PB** (2016) Toxicity and tolerance of aluminum in plants: tailoring plants to suit to acid soils *Biometals* **29**:187–210
- Sasaki T S, Yamamoto Y, Ezaki B, Katsuhara, M, Ahn S J, Ryan P R, Delhaize E, Matsumoto, H**(2004) A wheat gene encoding an aluminum-activated malate transporter. *The Plant Journal***37**, 645-653.
- Scandalios J G** (1993) Oxigen stress and superoxide dismutase. *Plant Physiology* v.**101**, p. 7-12.
- Schulz B, Kolukisaoglu H** (2006)Genomics of plant ABC transporters: the alphabet of photosynthetic life forms or just holes in membranes? *FEBS Lett.* 580, 1010–1016.
- Sivaguru M, Paliwal K**(1993) Differential aluminium tolerance in some tropical rice cultivars. II. Mechanism of aluminium tolerance **Journal of Plants Nutrition**, **16**: 1717-1732.
- Taylor G J, McDonald-Stephens J L, Hunter D B, Bertsch P M, Elmore D, Rengel Z, Reid R J** (2000). Direct measurement of aluminum uptake and distribution in single cells of *Chara corallina*. *Plant Physiology*, **123**, 987-96.
- Von Uexkull H R, Mutert E** (1995). Global extent, development and economic impact of acid soils. *Plant and Soil, Dordrecht*, v.**171**, p.1-15.
- Yamaguchi M, Sasaki T, Sivaguru M, Yamamoto Y, Osawa H, Ahn SJ, Matsumoto H** (2005) Evidence for the plasma membrane localization of Al-activated malate transporter (ALMT1). *Plant & Cell Physiology* **46**, 812–816. doi:10.1093/pcp/pci083
- Yamamoto Y, Kobayashi Y, Davi S R, Rikiishi S, Matsumoto H** (2002). Aluminium toxicity is associated with mitochondrial dysfunction and the production of reactive oxygem species in plant cells. *Plant Physiology*vol. **128**:63-72.

- Yamamoto Y, Kobayashi Y, Matsumoto H** (2001) Lipid peroxidation is as early symptom triggered by aluminium, but not primary cause of elongation inhibition in pea roots. *Plant Physiology* v. **125**: 199-208.
- Yokosho, K.; Yamaji, N. And Ma, J.F.** (2010). Isolations and characterization of the MATE genes in rye. *Funct. YOKOSHO, K.; YAMAJI, N. AND MA, J.F.* (2010). Isolations and characterization of the MATE genes in rye. *Funct. Plant. Biol.* 37:296-303.*Plant. Biol.* 37:296-303.
- Yokosho K, Yamaji N, Ma J F** (2011). An Al-inducible MATE gene is involved in external detoxification of Al in rice. *Plant J.* **68**:1061-1069.
- Wagatsuma T, Ishikawa S, Obata H, Tawaraya K, KatohATOHDA S.** (1995)Plasma membrane of younger and outer cells in the primary specific site for aluminium toxicity in roots. *Plant Soil* **171**: 105-112.
- Xia J, Yamaji N, Kasai T, Ma J F** (2010) Plasma membrane-localized transporter for aluminum in rice. *Proc. Natl Acad. Sci. USA*, **107**, 18381–18385.
- Zhang W, Ryan PR, Sasaki T, Yamamoto Y, Sullivan W, Tyerman SD** (2008) Characterization of the TaALMT1 protein as an Al³⁺-activated anion channel in transformed tobacco (*Nicotiana tabacum* L.) cells. *Plant & Cell Physiology* 49, 1316–1330. doi:10.1093/pcp/pcn10

7. Suplementares

```
>Sobic.009G166200.1
GCACCCCTCGCGGGCCGCTTCTCTTCTCCAGCGTCCGATCTCCCACTCGCCTCCCTCACCGCAGCTCTCT
CTGCTCGGTCGCTTCGCCCCCCTCCGCCCTCTCCCTGCGCGCGGCTCGCCACCTCGCTCTCCCTCC
CTCCACGTCGCCCCGTGCCCGCGCTTATATAAGTGCAGGAGGAGCTCATGGCGGACGCTCCGGCGAGCCCT
GGGGCGGGCGGGAGCCACGAGAGCGGGAGCCCCAGGGGCGGGAGGTGGAGGCGGTGGCAGCGTCA
GGGAGCAGGACAGGTTCTGCCATCGCCAACATCAGTCGCATCATGAAGAAGGCCATCCCGGCTAACGG
GAAGATCGCCAAGGACGCCAAGGAGACCGTGCAGGAGTGCCTCTCAGAGTTCATCTCTTCATCACTAGC
GAAGCGAGTGACAAGTGCCAGAGGGAGAAGCGGAAGACCATCAACGGCGACGACCTGCTGTGGGCCATGG
CCACATTGGGGTTTGGAGACTACATTGAACCACTCAAGGTGTACCTGCAGAAGTACAGAGAGATGGAGGG
TGATAGTAAGTTAACTGCAAAAACCGGTGATGGCTCAATTAAAAAGGATGCCCTTGGTTCATGTGGGAGGA
AGTAGCTCAGCTGCACAAGGGATGGGCCAACAAAGGAGCATACAACCAAGGAATGGGTTATATGCAACCTC
AGTACCATAATGGGGATATCTCAAACATAATGAAGATCTGGACCTTTTCTGCGACAGCTGCTCTTCCCTGA
GGTGCATATTTGGTCATAGTTATTTACTGAGACACCTTGCGGTGACCATTAAAGAGTAACCAGTCACCCC
CGGTAGGTCCAAGAAGTGAAGGTCCAAGAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
TGGGAATA

>Sobic.009G166200.1
MADAPASPGGGGSHESGSPRGGGGGGGVSREQDRFLPIANI SRIMKKAIPANGKIAKDAKETVQECVS
EFISFITSEASDKCQREKRKTINGDLLWAMATLGFEDYIEPLKVYLQKYREMEGDSKLTAKTGDGSIKKDALGHVG
GSSSAAQGMGQQGAYNQGMGYMQPQYHNGDISN*
```

Figura suplementar S1: Sequencia de nucleorídeos e de aminoácidos do transfator candidato NFY-like. Na sequencia de nucleotídeos, a região 5' ÚTR está destacada em verde, em azul escuro, os exons, em azul claro, os introns e em rosa, a região 3'UTR não transcrita. Na sequencia de aminoácidos, o domínio conservado está marcado com letras vermelhas

```
>Sobic.009G174300.1
AAGTCATGTCTCTGTCACAACTTCTCAGTTCTCACCCCGCCGGCGCCGGCTCCAGTCCCAGCCAGCTTC
CACATTTTTCTTTCTTACCTTTTGAAAAGGTGTTCCCCCTTTGTGACGCTGCTCCAACCTCCATGACGC
TGTCCTCACCCGAGCCGCCGCCCTCCTCCAGCCCAAGGCACGCAGCGATCCAGGAGCTCAGGAGAGGTAC
CCAGCTGGCGGAGCTGCTCAGGCAGCAGGTAGAGCTCATCCCGGAGCCTAACCGCCGGCAAGCTGCAGTG
GTCAACGTGGGCGAAATATCCATGGCTATGGAGTCGTCGCTCTCCATTCTCCAGTCTGAGATGGAGCACC
CCTTCGTCTCCGAGGTATGGCTGCGCCCACTGCCTACTCTGACGGAGGCAGCACCAGCAGGGAAAGGAA
TGGTCCCGTGGCCCGTACAAGAAGGTTGAGGCACCGGCGAGGCAGAGATGGAGCTGAACTCCCGATCAAG
GAGATACTGACTGAGGCACCAGAAAACGATCATTTCCACTGGAGGAAATATGGTGAAAAGAATATCCTCT
ATGCTGAATATCCAAGGTTATACTACAAGTGCGGTTACAGCGATGACCACAAGTGCCCGGCAAAGAAATA
CGTGACGAGCAAAGCAACACCTACCCTCCACTTTTCTTGGTCAACCCTGATCAACGAGCATACTTGGGAC
ACCTTGTTCGGGATGAGCCCAGCTCAAGCAGCAGTGGTTCGCAGGTTCTCGACTTACAAAGGCATCGC
TTTCTCCTGAAGAAGACAGCAGCATGCCTGTGTCTATGCACAGATATTCGTTTTCGTATGATGGGTAATA
GTCGTCTTTCGTTGCCAGGATCGATGTCGGCAGATGGATTCCAGATCATCAGGTAGATTTATCCTTGGGA
CCTGGGTGGTAGCGCACTCGCTCAGGCAAGCAGACTCACGCGGCAAGAGTTGTAGCTCGGCGTGCACCTG
CCTTTTGATTAAGGATTTCAATGTACCTGCCAGATCCAAACCCACGCATTTAATTTATTTCTCATTTCAC
GGATCACACAAGCTGGTCTCGGTGGGTGTGGTTTGTATGGAAGGGCAAGGCTTTTTTCTCTTAAGAGCAA
GGTCAATAATAGAGCCAAGTGCTTAGCTATAAGCCAAGTGATTAGAGCCA

>Sobic.009G174300.1
MTLSPPQPPSSSPRHAAIQELRRGTQLAELLRQQVELIPEPNRRQAAVVNVGEISMAMESLSILQSEM
EHPFVSEVMAAPTAYSDDGGSTSRERNPVARTRVRHRRGRDGAELPIKEILTEAPENDHFHWKRYGEKNILY
```

```
AEYPRLYYKCGYSDDHKCPAKKYVQQSNTYPPLFLVTLINHEHTCDTLFRDEPSSSSSSGSQVLDFTKASLS  
PEEDSSMPVSMHRYSFSYDGY
```

Figura suplementar S2: Sequencia nucleotídica e proteica do transfator candidato WRKY41-like. Na sequencia de nucleotídeos, a região 5' ÚTR está destacada em verde, em azul escuro, os exons, em azul claro, os introns e em rosa, a região 3'UTR não transcrita. Na sequencia de aminoácidos, o domínio conservado está marcado com letras vermelha.

```
>Sobic.009G151400.2  
TTACACCTCACCTCCCCGTCCCCCCCACACGCACAAAGGCGACGCGACGCGACGCCAGCAGCGCTTGCTT  
TCCTCTCCCTGCTCTGCTCGGCCCTCGCCGCTTCCCCCTCCTCTACCCCTGCCTCTGCCTGCCTGCC  
TCAACGATTCCTCCGCTTTTCCTTTTTCAAATCGCCCTTTTCTCGTGCGCGTGTGACTCCTTCCTGGT  
TCCTGCTCGCCGGCGGGTCTGGGCTGGGGGACGGACGCTGCCTCGGATGAGGCCGCGCTGACCCGAGCC  
GGACCACCCGGCGCCACCCATGGCGCGTCCGACGAGTGTACGAGGCTTGGAAAGGGAACAATGTAAA  
GATCCTGCCTTCTCTTTCCTCTGCCCTCCTCTCGCTTCATCTTGTTAGTTCTTGAGCCGGGAATCTGGGT  
TGCACGTTCGGTTCCTCCGTGGGGAATTCCGCTTCTTGCTGCTGTATGGGATCGTTTTGGTAGAGTTGG  
TTTGCGCCTGACTTATGCGGCCAGTGTGCGGTTACTACTACTTTACTGTTCTCGGCGATTCCAGAAAT  
TCTTCTTTGGTGGGAGGCTCATATTCGGACCTGACGCCAAGTCACTGATCCTCTCTGTTTCGCTCATTGT  
GATCCCGGATTGGTCTTCTGCGGATTTGTTGCTCGGCATCTCCGCCACCATTCCCGGATTACAATGCA  
GGATATGCGATTTCCTGCGGTGGCTGTGGTGTATGATTTACGTGCTAGTCCTGCTCTTGTTACATCAG  
CGCAGGATCCTGGGATCGTTCGCCGCGCGGCCACCCGCCGGAGGAAGAGTTTTCTTATGGGAACGCCTT  
GTCTGGGGGCACGCCTGGGAGGCTGCAATTCGCCGTGTGAAGGAAGTCATGGTTAAGGGGATGCCTGTG  
AAAGTGAAGTATTGTGACACCTGCATGATTTATCGGCCCTCCTCGATGCTCACACTGTTTCGATATGCAACA  
ATTGTGTAGAGCGCTTTGACCATCATTGTCCCTGGGTGGACAATGCATTGGACAGCGCAATTACCGGTA  
CTTTTTCATGTTTGTCTTCTCCTCGACTCTTCTCTGCATTTATGTATTTGCAATGTCAGCCTTGTACATC  
AAGTTTCTCATGGACGAGGGTTATCCAACAGTATGGAAGGCTTCAAACTCTCCCGCTTCTTTGGGGC  
TTTTGATATATTGTTTCATTGCTCTCTGGTTTGTGCGGTGGACTCACTGGATTTCAATTGTACCTCATTAG  
CACCACCAGACCACGTACGAGAATTTTCGATACAGATCAGACAGCAGGCCTAATATCTATAGTCAAGGA  
TGCTGAATAACTTTCTAGAAGTCTTCTGTAGCAAGACAAAACCTTCTAAACACAAGTTTCGGGCATATG  
CACAAGAGGAAGTGCACCTCCAACAGTTAGCTTTGGTAGGGATGTTGAAGATGAACAGTTGGCGGTCC  
TCGGTCGAAAGTTGAAGATGATCTTGAATTTGGTAGTGATCTTCTGAAGATCTCTCAGCGGCCAATTAT  
GAGGATGTTAATGTAGAAGTGGGAGGCTCCTCACAGGCTCAGCAGTGAGATGGAAGGCATAGCCAATGCCA  
AACTTGTCATGTTTCAGAATCACAGATCCCTGCAGTTGGGAGCGACGTAAGGGTGCACACTCAAGCTG  
GGATCGGAGAAGTGGGAACCTGGGACATGTCCATCAGATGTACTTACAAGAAGCGCATCAGATGTAATTGAA  
CGTAGCGTGTGGCCACGGAGGCTGCACATCCATCTTAGCAGAGACTCACTAGTGAATTGTTTGTGCAT  
TCAGAACAATGTGTGTGCAGTTTGGATTGCAAGGCCATTTAAGTTAACTCATGTTGGCTCACCCAGACAG  
CAAGGAAGAAAACGCCTGCATGATTTGACTTTGACTGCCTTGATTGGGTTTTCTTCCGACCCCTTGT  
CAATTGTCATGTATGAGCAACTATTTAGTCAAAGAATTAGCCCCATTGTGCCATTATATTCGTGTGTTG  
TGCCATAGTCTCAATCTGACGCTGTGTTGCGTATGGTCTCTGTAAATTGCAAGTGCAGGATTCTCAA  
TTTCTCATCCAATTGTAATTCTTCTTAGGCTATTTCTTTGAGAATGTTATCTCGCCTGGCCCT
```

```
>Sobic.009G151400.2  
MRPVCGLLLLYCSRFRQKFFFGRLIFGPDAKSLILSVSLIVPVLVFCGFVARHLRHHFPDYNAGYAIPAVAVVFMIV  
LVLLLVTSAQDPGIVPRAAHPPEEEFSYGNALSSGTPGRLQFPRVKEVMVKGMPVKVKYCDTCMIYRPPR  
CSHCSIENNVERFDHHC PWVGQCI GQRNYRYFFMFVSSSTLLCIYVFAMSALYIKFLMDEGYPTVWKAFKHPAS  
LGLLIYCFIALWVFGGLTGFLYLISTNQTTYENFRYRSDSRPNIIYSQGCLNNFLEVFCSKTKPSKHKFRAYAQE  
EVRPPTVSFGRDVEDEPVGGPRSKVEDDLEIGSDLLKISQRRNYEDVNVEVGGPHRLSSEMEGIANAKLV  
ACSESQIPAVGSDVRVRHSSWDRRSGNWDMSDVLTRSASDVIERSVFATEAAHPSLAETH*
```

Figura suplementar S3: Sequencia nucleotídica e proteica do transfator candidato ZincFinger-like. Na sequencia de nucleotídeos, a região 5' ÚTR está destacada em verde, em azul escuro, os exons, em azul claro, os introns e em rosa, a região 3'UTR não transcrita. Na sequencia de aminoácidos o domínio conservado marcado com letras vermelhas.

>Sb09g022540 CDS

```
ATGCAGGGGCATGCGTACAGCCGCCTGGGCAGCTTCGGCTCCGCGGCGAGCGCGCCGTCCCCCGCCGCC
GCCGTCGCCGTCGTACAGGTGCCGGCTGGGGCCGGCGCCGGCGCCGGCGCTGGCGGGGCGAGCCGGACG
CCCGCAAGGTGGGTTTCGGCGAGGGGCATCCCGGGGGCCGCGGCGTCGGCAACGGCGGCGGGCGGGCG
GGGGTCGCCCACCGTGGCGGCGGGCGGCGGCGGCGGAGGGTGGCCAGGGCGGTGCTCGCGACTG
CTGCGGAGGCAGGCGGTGTTCTTTCGCGCCGCTGCTGTACGTCGCCGCGATGCTGCTGTACATGGGCTCG
CTCCCGCTCGACGCCGTGCCGCGGATCATCGCGGCCAGCCGCTGGGTGCGTGTACCGCAGCCCGCAACTG
TACGCGCGGCTCCGCGCCGACATGGACGCCGACAACCTCACCGGCGCGGCTTGCAACGTGGCGCAAATGTT
TGGCCGTATTGATCACAGAATGGTTGAAGCATGA
```

>Sb09g022540

```
MQGHAYSRLGSFGSAASAPSPPPPPSPSSQVPAGAGAGAGAGGGSRTPAKVGSARGIPGAAASA
TAAARAGVAHRGGGGGAARRVARAVLATLLRRQAVFLFAPLLYVAAMLLYMGSPLDAVPRIIARQP
PGSVYRSPQLYARLRADMDADNSTGAACNVAQMFGRIDHRMVEA
```

Figura suplementar S4: Sequencia nucleotídica e proteica do transfator candidato Sb09g022540.
*Sequência obtida na versão anterior (v2.1) do Phytozome e indisponível na versão atual v3.1

>Sobic.009G163400.2

```
GACACTGCTGCGGAGGCAGGCGGTGTTCTTTCGCGCCGCTGCTGTACGTCGCCGCGATGCTGCTGTAC
ATGGGCTCGCTCCCGCTCGACGCCGTGCCGCGGATCATCGCGCGCCAGCCGCTGGGTGCGGTGTACCGCA
GCCCAGCAACTGTACGCGCGGCTCCGCGCCGACATGGACGCCGACAACCTCACCGGCGCGCTAGCAACTGT
ATGGAGGCACACTTACAAAGGTGGCACATGGCGGCCCTGCATAAAAAACATGACAAATGGTTTACCTGAA
TCAAATGGCTACATATACGTTGAGGCAAATGGTGGTTTAAATCAACAAAGGACATCGATATGCAATGCAG
TCGCCATTGCTGGTTTTCTGAATGCTACTCTTATTATCCCAAATTTCCATTTCCATAGCATTGGAGGGA
TCCTAGCAAATTCAGCGATATCTATGATAAAGATCAATTTGTACAACGTCTACAAAATGATGTTAGAGTA
GTCGACAAAATTCGGATTTTATAATGGAGCGGTTTGGTCATAATCTCAGCAATGTGTTCAATTTCAAGA
TAAAGGCTGGGCACGCATTCAATATTACAAAGATGTTGTTCTTCCAAAATTAGTTGAAGAAGATTTCAT
CAGGATTTCTCCTTTTGCCAACCGACTCTCATTTGATGCACCTTCTGCTGTTCAACGGTTGAGATGCTTG
GCAAACCTTTGAAGCCTTAAAAATCTCTAAACCAATTGTGTCTTTGTCTGAAACCTTAGTTTCTCGAATGA
GAGAGAAAAGTGTGCAAGTGATGGAAAATATATCTCAGTGCATCTTCGTTTTGAGGAGGATATGATTGC
CTTCTCATGTTGTGTATATGATGGTGGTGATGAGGAGAAGAAGGAAATGGATGCAGCCAGAGAAATAGGT
TGGAGAGGGAAATTTACTAAGAGAGAGGACGGGTAATAAGACCAGGAGTGATCAGAATGACTGGAAAATGTC
CTTTACACCTTTGGAGGTTGGGTTAATGCTTCGTGGAATGGGTTTCAGCAATAAGACTGCAATCTTTTT
GGCATCTGGAAAGATTTACAGAGCAGAGAAGAACATGGCTTCTCTTCTTGAATGTTTCTCTCCTACAG
ACAAAAGAGACACTGGCATCAGAAGAGGAACCTTGCTCCATTCAAGAATTTCTCCTCAAGGATGGCTGCTG
TTGACTATAGCGTTTGTGCTCAAAGTGAAGTTTTTGTGACCACACAAGGCGGAAATTTCCCGCATTTCCT
TATGGGTCACAGGAGATACTTGTATGGTGGGCATTCAAAAACAATTAACCAGACAAAAGAAGGCTTGCT
GTACTTTTTGATAATCCACGTATAGGGTGAAGGCATTGAAGCGGCACCTTCTTAATATGAGAGCACACA
GCGATGCCAAGGGGATTGAAATGAAAAGACCAAACGAATCCATATATACCTTTCCATGTCTGACTGCAT
GTGCCGCTTGAACAGAACAGAACACTCGAAATCCAAACACAGTAGATAGTATCTCGTAGTTGCATTCCAG
ACTAACCTTATTCTTTCTGACATCCAGTTTAGTTTGTTCATTCTTCATTTGCTGTGCGGCTTTTCATTAC
TTCATCTGCAAGCATACTTGCCAGCTGCCACAGCATTGAAAGCAATCTATTCTCCCTATGTAGCGGTTTT
TCCTTGTATAGCATATGCCAGCATGGAACCCAGGTTCTCCATGGAGTATTAGGCCAATTGTTATGCT
AACTATACATCAGCTGACTATAATCTCAACAACAGATCTGACTCCAGTAGGA
```

>Sobic.009G163400.2

MLLYMGSLPLDAVPRI IARQPPGSVYRSPQLYARLRADMDADNSTGALATVWRHTYKGGTWRPCIKNMTNG
 LPESNGYIYVEANGGLNQQRTSICNAVAIAGFLNATLIIPNFHHSIWRDPSKFSIDIYDKDQFVQRLQNDVRVVD
 KIPDFIMERFGHNLNSVFNFKIKAWARIQYYKDVVLPKLVEERFIRISPFANRLSFDAPSAVQRLRCLANFEALKF
 SKPIVLSLSETLVSRMREKSVASDGKYSVHLRFEEDMIAFSCCVYDGGDEEKKEMDAAREIGWRGKFTKRGRVIRP
 GVIRMTGKCPLTPLEVGLMLRGMGFSNKTAIFLASGKIYRAEKNMASLLEMFLLQTKETLASEEELAPFKNFSSRM
 AAVDYSVCAQSEVFVTTQGGNFPHFLMGHRRYLYGGHSKTIKPKDKRRLAVLFDNPRI GWKALKRHLNMRA
 HSDAKGIEMKRPNESIYTFPCDCMCLNRTEHSKSKHSR*

Figura suplementar S5: Sequência de nucleotídeos e de aminoácidos de Sobic.009g163400, encontrado por similaridade com Sb09g022540. Na sequencia de nucleotídeos, a região 5' ÚTR está destacada em verde, em azul escuro, os éxons, em azul claro, os introns e em rosa, a região 3'UTR não transcrita.

```

CLUSTAL O(1.2.1) multiple sequence alignment

Sb09g022540.1      ATGCAGGGGCATGCGTACAGCCGCCTGGGCAGTTCGGCTCCGCGGCAGCGCGCGCTCC
Sobic.009G163400.2  -----

Sb09g022540.1      CCCCCCGCCGCCGCTCGCCGTCGTCACAGGTGCCGGTGGGGCCGGCCGGCGCCGGC
Sobic.009G163400.2  -----

Sb09g022540.1      GCTGGCGGGGGCAGCCGGACGCCCGGAAGGTGGGTTCGGCGAGGGGCATCCCGGGGCC
Sobic.009G163400.2  -----

Sb09g022540.1      GCGGCGTCGGCAACGGCGCGCGGGCGGGGGTGCACCACCGTGGCGGGCGGGCGGGC
Sobic.009G163400.2  -----

Sb09g022540.1      GCGGCGGGAGGGTGGCCAGGGCGGTGCTCGCGACACTGTCGCGAGGCAGGCGGTGTT
Sobic.009G163400.2  -----

Sb09g022540.1      CTCCTCGCGCCGCTGCTGTACGTCGCCCGCATGCTGCTGTACATGGGCTCGCTCCCGTC
Sobic.009G163400.2  -----ATGCTGCTGTACATGGGCTCGCTCCCGCTC
                        *****

Sb09g022540.1      GAGCCGCTGCCCGGATCATCGCGGCCAGCCGCCTGGGTCCGTTACCGCAGCCCGCAA
Sobic.009G163400.2  GAGCCGCTGCCCGGATCATCGCGGCCAGCCGCCTGGGTCCGTTACCGCAGCCCGCAA
                        *****

Sb09g022540.1      CTGTACGCGCGGCTCCGCGCCGACATGGACGCCGACAACCCACCGCGCGGTTGCAAC
Sobic.009G163400.2  CTGTACGCGCGGCTCCGCGCCGACATGGACGCCGACAACCCACCGCGCGGTTGCAAC
                        *****

Sb09g022540.1      GTGGCGCAAATGTTGGCCGTATTGATCACAGAATG-GTTGAA-----GCATGA-----
Sobic.009G163400.2  GTATGGAGGCACACTTACAAAGGTGGCACATGGCGGCCCTGCATAAAAACATGACAAAT
                        ** * * * * * * * * * *

Sb09g022540.1      -----
Sobic.009G163400.2  GGTTTACCTGAATCAATGGCTACATATACGTTGAGGCAAATGGTGGTTTAAATCAACAA

Sb09g022540.1      -----
Sobic.009G163400.2  AGGACATCGATATGCAATGCAGTCGCCATTGCTGTTTTCTGAATGCTACTCTTATTATC

Sb09g022540.1      -----
Sobic.009G163400.2  CCAAATTTCCATTTCCATAGCATTGGAGGGATCCTAGCAAATTCAGCGATATCTATGAT

Sb09g022540.1      -----
Sobic.009G163400.2  AAAGATCAATTTGTACAACGCTCTACAAATGATGTTAGAGTAGTCGACAAAATCCGGAT

Sb09g022540.1      -----
Sobic.009G163400.2  TTTATAATGGAGCGGTTTGGTCATAATCTCAGCAATGTGTTCAATTTCAAGATAAAGGCT

Sb09g022540.1      -----
Sobic.009G163400.2  TGGGCACGCATTCAATATTACAAAGATGTTGTCTTCCAAAATTAGTTGAAGAAAGATT

Sb09g022540.1      -----
Sobic.009G163400.2  ATCAGGATTTCTCCTTTTGCCAACCGACTCTCATTGATGACCTTCTGCTGTTCAACGG

Sb09g022540.1      -----
Sobic.009G163400.2  TTGAGATGCTTGGCAAACCTTGAAGCCTTAAATTTCTCTAAACCAATTGTGCTTTGTCT
  
```

```

Sb09g022540.1 -----
Sobic.009G163400.2  GAAACCTTAGTTTCTCGAATGAGAGAGAAAAGTGTGCAAGTGATGGAAAATATATCTCA

Sb09g022540.1 -----
Sobic.009G163400.2  GTGCATCTTCGTTTTGAGGAGGATATGATTGCCTTCTCATGTTGTGTATATGATGGTGGT

Sb09g022540.1 -----
Sobic.009G163400.2  GATGAGGAGAAAGGAAATGGATGCAGCCAGAGAAATAGGTGGAGAGGAAATTTACT

Sb09g022540.1 -----
Sobic.009G163400.2  AAGAGAGGACGGGTAATAAGACCAGGAGTGATCAGAATGACTGGAAAATGTCCTCTTACA

Sb09g022540.1 -----
Sobic.009G163400.2  CCTTTGGAGTTGGGTTAATGCTTCGTGGAATGGGTTTCAGCAATAAGACTGCAATCTTT

Sb09g022540.1 -----
Sobic.009G163400.2  TTGGCATCTGGAAAGATTTACAGAGCAGAGAAGAACATGGCTTCTCTTGAATGTTT

Sb09g022540.1 -----
Sobic.009G163400.2  CCTCTCCTACAGACAAAAGAGACACTGGCATCAGAAGAGGAAGTTCCTCCATTCAGAAAT

Sb09g022540.1 -----
Sobic.009G163400.2  TTCTCCTCAAGGATGGCTGCTGTTGACTATAGCGTTTGTGCTCAAAGTGAAGTTTTGTG

Sb09g022540.1 -----
Sobic.009G163400.2  ACCACACAAGGCCGAAATTTCCCGCATTTCTTATGGGTCACAGGAGATACTTGTATGGT

Sb09g022540.1 -----
Sobic.009G163400.2  GGGCATTCAAAAACAATTAACCAGACAAAAGAAGGCTTGCTGTACTTTTTGATAATCCA

Sb09g022540.1 -----
Sobic.009G163400.2  CGTATAGGGTGAAGGCATTGAAGCGGCACTTGCTTAATATGAGAGCACACAGCGATGCC

Sb09g022540.1 -----
Sobic.009G163400.2  AAGGGGATTGAAATGAAAAGACCAACGAATCCATATATACCTTTCCATGTCCTGACTGC

Sb09g022540.1 -----
Sobic.009G163400.2  ATGTGCCGCTTGAACAGAACAGAACACTCGAAATCCAACACAGTAGATAG

```

Figura suplementar S6: Alinhamento entre a sequência Sb09g022540 obtida por experimento de eQTL e Sobic.009g163400 obtido por similaridade. Foi utilizado o programa Clustal Omega.

Tabela suplementar S1: Identificação dos elementos cis-regulatórios preditos para a região proximal do promotor de SC283, de 146 pb. É apresentada a posição, sequência e função dos elementos encontrados






Site Name	Organism	Position	Strand	Matrix score.	sequence	Function
+ A-box 						
A-box	<i>Petroselinum crispum</i>	51	+	6	CCGTCC	Indução da atividade transcricional
+ Box 4 						
Box 4	<i>Petroselinum crispum</i>	110	-	6	ATTAAT	Parte de uma matriz conservada de DNA
+ CAAT-box 						
CAAT-box	<i>Hordeum vulgare</i>	45	-	4	CAAT	comum ao elemento cis em promotor e regiões <i>enhancer</i>
+ CCAAT-box 						
CCAAT-box	<i>Hordeum vulgare</i>	139	+	6	CAACGG	possível local de ligação de MYBHv1
+ CCGTCC-box 						
CCGTCC-box	<i>Arabidopsis thaliana</i>	51	+	6	CCGTCC	elemento regulador cis-acting relacionado com ativação do meristema

Tabela Suplementar S2 : Genes ortólogos deWRKY-like em diferentes espécies de plantas

Gene name (Template)	Template description	Best hit	Best hit description	E-value	Identity	Similarity	Coverage
Sobic.009G174300	WRKY DNA -binding domain	AT3G56400.1	WRKY dna-binding protein 70	3E-15	30,27	48,65	73
Sobic.009G174300	WRKY DNA -binding domain	evm.model.supercontig_19.45	WRKY transcription factor 38-related	3E-15	44,16	63,64	33
Sobic.009G174300	WRKY DNA -binding domain	Cucsa.177060.1	WRKY dna-binding domain (wrky)	1E-16	30,53	47,37	78
Sobic.009G174300	WRKY DNA -binding domain	orange1.1g018407m	WRKY transcription factor 38-related	1E-14	44,19	59,3	35
Sobic.009G174300	WRKY DNA -binding domain	Glyma.09G274000.1.p	WRKY dna-binding domain (wrky)	3E-17	33,7	46,96	75
Sobic.009G174300	WRKY DNA -binding domain	Manes.05G030900.1.p	WRKY dna-binding domain (wrky)	3E-16	28,57	47,77	91
Sobic.009G174300	WRKY DNA -binding domain	LOC_Os05g40070.1	WRKY84, expressed	2E-57	52,71	67,49	85
Sobic.009G174300	WRKY DNA -binding domain	Phvul.008G081800.1	WRKY dna-binding domain (WRKY)	6E-18	33,52	49,45	77
Sobic.009G174300	WRKY DNA -binding domain	Sobic.009G174300.1.p	similar to Putative uncharacterized protein OSJNBa0018K15.11	4E-176	100	100	100
Sobic.009G174300	WRKY DNA -binding domain	Solyc09g015770.2.1	WRKY DNA -binding domain (WRKY)	3E-16	31,84	46,37	76
Sobic.009G174300	WRKY DNA -binding domain	PGSC0003DMP400050864	WRKY transcription factor 6	3E-15	31,89	45,41	76
Sobic.009G174300	WRKY DNA -binding domain	Traes_1AL_309623B48.1	WRKY DNA -binding domain (WRKY)	2E-87	58,65	72,15	93
Sobic.009G174300	WRKY DNA -binding domain	GSVIVT01030174001	WRKY DNA -binding domain (WRKY)	4E-18	29,17	47,22	88
Sobic.009G174300	WRKY DNA -binding domain	GRMZM2G034421_P01	WRKY DNA -binding domain (WRKY)	2E-112	82,55	85,85	85

Tabela Suplementar S3- Genes ortólogos de ZincFinger-like em diferentes espécies de plantas

Gene name (Template)	Template description	Best hit	Best hit description	E-value	Identity	Similarity	Coverage
Sobic.009G151400	Zinc Finger Dhhc Domain Containing Protein	AT4G24630.1	Dhhc-Type Zinc Finger Family Protein	8E-170	59,81	70,94	94
Sobic.009G151400	Zinc Finger Dhhc Domain Containing Protein	evm.model.supercontig_49.89	Protein S-Acyltransferase 8	0	65,52	78,57	91
Sobic.009G151400	Zinc Finger Dhhc Domain Containing Protein	Cuca.363840.1	Protein S-Acyltransferase 8	0	64,37	76,66	91
Sobic.009G151400	Zinc Finger Dhhc Domain Containing Protein	orange1.1g014643m	Protein S-Acyltransferase 8	0	65,19	79,01	91
Sobic.009G151400	Zinc Finger Dhhc Domain Containing Protein	Glyma.16G164500.1.p	Protein S-Acyltransferase 9-Related	2E-176	63,21	77,04	91
Sobic.009G151400	Zinc Finger Dhhc Domain Containing Protein	Manes.18G094000.1.p	Protein S-Acyltransferase 8	0	64,18	77,4	94
Sobic.009G151400	Zinc Finger Dhhc Domain Containing Protein	LOC_Os05g36090.1	Zinc Finger Dhhc Domain-Containing Protein, Putative, Expressed	0	82,69	89,42	96
Sobic.009G151400	Zinc Finger Dhhc Domain Containing Protein	Phvu.004G091700.1	Protein S-Acyltransferase 9-Related	0	63,52	78,66	91
Sobic.009G151400	Zinc Finger Dhhc Domain Containing Protein	Sobic.009G151400.2.p	Similar To Os05g0436900 Protein	0	100	100	100
Sobic.009G151400	Zinc Finger Dhhc Domain Containing Protein	Solyc04g071860.2.1	Protein S-Acyltransferase 8	2E-178	63,26	75,67	93
Sobic.009G151400	Zinc Finger Dhhc Domain Containing Protein	PGSC0003DMP400034244	Zinc Finger Protein	2E-152	62,16	78,38	77
Sobic.009G151400	Zinc Finger Dhhc Domain Containing Protein	Traes_1BL_19E07DA0F.1	Dhhc-Type Zn-Finger Proteins	0	80,77	89,18	96
Sobic.009G151400	Zinc Finger Dhhc Domain Containing Protein	GSVIVT01034567001	Protein S-Acyltransferase 8	7E-180	61,32	73,11	95
Sobic.009G151400	Zinc Finger Dhhc Domain Containing Protein	GRMZM2G100641_P01	Protein S-Acyltransferase 8	0	91,15	94,74	96