

UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
DEPARTAMENTO DE BIOLOGIA GERAL
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA



DISSERTAÇÃO DE MESTRADO

**CARACTERIZAÇÃO MOLECULAR DE UM NOVO ALELO DO GENE
ZmMATE1 QUE CONFERE TOLERÂNCIA AO ALUMÍNIO EM MILHO**

ORIENTADA: Marcella Baroni de Resende Costa

ORIENTADORA: Dra. Claudia Teixeira Guimarães

BELO HORIZONTE

Julho - 2017

Marcella Baroni de Resende Costa

**CARACTERIZAÇÃO MOLECULAR DE UM NOVO ALELO DO GENE
ZmMATE1 QUE CONFERE TOLERÂNCIA AO ALUMÍNIO EM MILHO**

Dissertação apresentada ao Programa de Pós-Graduação em Genética do Departamento de Biologia Geral do Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais, como parte das exigências para obtenção do título de Mestre em Genética.

Área de concentração: Genômica e Bioinformática

Orientadora: Dra. Claudia Teixeira Guimarães

Universidade Federal de Minas Gerais

Belo Horizonte
2017

043

Costa, Marcella Baroni de Resende.

Caracterização molecular de um novo alelo do gene ZmMATE1 que confere tolerância ao alumínio em milho [manuscrito] / Marcella Baroni de Resende Costa. - 2017.

57 f. : il. ; 29,5 cm.

Orientadora: Dra. Claudia Teixeira Guimarães.

Dissertação (mestrado) - Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas.

1. Expressão gênica - Teses. 2. Milho - Genética - Teses. 3. Alumínio - Teses. 4. Acidez do Solo. I. Guimarães, Claudia Teixeira. II. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. III. Título.

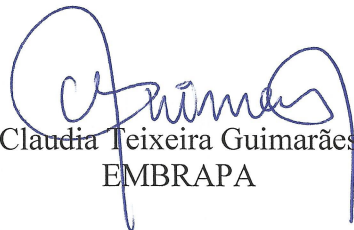
CDU: 575




"Caracterização molecular de um novo alelo do gene *ZmMATE1* que confere tolerância ao alumínio em milho "

Marcella Baroni de Resende Costa

Dissertação aprovada pela banca examinadora constituída pelos Professores:


Claudia Teixeira Guimarães
EMBRAPA


Maria Bernadete Lovato
ICB - UFMG


Sylvia Morais de Sousa
EMBRAPA


Marcel Giovanni Costa França
ICB - UFMG

Belo Horizonte, 31 de julho de 2017.

*Dedico esse trabalho aos meus
queridos pais, Margareth e Ivo.*

AGRADECIMENTOS

Aos meus pais, por todo amor, suporte, conselhos e pela dedicação. Obrigada por sempre me darem forças para seguir em frente. Ao meu irmão pelo apoio e incentivo.

Ao Danilo, pelo companheirismo, pela ajuda e por sempre me incentivar.

A minha orientadora Dra. Cláudia Teixeira Guimarães pela confiança, apoio e ensinamentos.

Aos analistas e técnicos do Núcleo de Biologia Aplicada, Ubiraci, Marcos, Célio e especialmente à Beatriz, pela valiosa ajuda e por todos os ensinamentos. À Gislene pelo suporte nos experimentos na câmara de crescimento.

Aos funcionários e estudantes do NBA, pelo carinho, incentivo, amizade e pelos momentos de descontração e risadas. Agradeço em especial à Bárbara, Laiane, Simara, Karine e Diego por todo apoio, pela ajuda e contribuições.

As minhas irmãs de república, Amanda, Gisele, Dani e Marina, pela parceria e pelo carinho.

Às Flores, Joyce, Anna, Fernanda, Lorena, Priscila e Rayssa, por sempre me apoiarem e por entenderem a minha ausência em alguns momentos.

À Embrapa Milho e Sorgo por possibilitar o desenvolvimento da dissertação.

À Universidade Federal de Minas Gerais, em especial à pós-graduação de genética, pela oportunidade.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão da bolsa.

A todos que de alguma forma contribuíram para a conclusão desse trabalho.

Muito obrigada!

SUMÁRIO

LISTA DE FIGURAS.....	VIII
LISTA DE TABELAS.....	X
LISTA DE ABREVIATURAS.....	XI
RESUMO.....	13
ABSTRACT.....	14
1.INTRODUÇÃO.....	15
2. REVISÃO BIBLIOGRÁFICA	15
2.1 O milho (<i>Zea mays</i> L.)	15
2.2 Solos ácidos	16
2.3 Mecanismos fisiológicos de tolerância ao Al	16
2.4 Genes associados com a tolerância ao Al.....	17
2.5 Regulação da expressão gênica da tolerância ao Al.....	19
2.6 Tolerância ao Al em milho	20
3.OBJETIVOS.....	22
3.1 Objetivo geral.....	22
3.2 Objetivos específicos.....	22
4. METODOLOGIA.....	23
4.1 Material genético	23
4.2 Associação do <i>ZmMATE1</i> com a tolerância ao alumínio.....	23
4.3 Caracterização do perfil de expressão do <i>ZmMATE1</i>	25
4.4 Sequenciamento e caracterização da região promotora do gene <i>ZmMATE1</i>	26
4.5 Amplificação por RACE e sequenciamento da 5' UTR.....	27
4.6 Busca por genes em milho que codificam fatores de transcrição que controlam a expressão de genes associados com tolerância ao Al	27
4.7 Caracterização do perfil de expressão dos genes candidatos que codificam possíveis fatores de transcrição em milho.....	28
5. RESULTADOS E DISCUSSÃO	29
5.1 O alelo do gene <i>ZmMATE1</i> confere tolerância ao Al	29
5.2 Padrão de Expressão Temporal e Tecidual do gene <i>ZmMATE1</i>	31
5.3 Identificação de SNPs e caracterização <i>in silico</i> da região promotora do gene <i>ZmMATE1</i>	33
5.4 Variações no tamanho e na sequência da região 5' UTR do gene <i>ZmMATE1</i>	40
5.5 Genes candidatos que codificam possíveis fatores de transcrição associados com a expressão do <i>ZmMATE1</i>	45

6. CONCLUSÕES.....	49
7. REFERÊNCIAS	50

LISTA DE FIGURAS

- Figura 1.** Modelo ilustrando os mecanismos apoplástico e simplástico de tolerância ao Al nas plantas e alguns genes associados com a tolerância. As setas vermelhas representam o fluxo de Al para dentro e no interior da célula e as setas azuis indicam o fluxo dos ácidos orgânicos (OA). Adaptado de Kochian *et al.* (2015)19
- Figura 2.** Distribuição do crescimento relativo da raiz seminal utilizado como índice tolerância ao Al na população F₂ do cruzamento entre L228-3 (linhagem tolerante ao Al) e L53 (linhagem sensível)30
- Figura 3.** Análise de regressão linear entre o crescimento relativo da raiz seminal (CRR) e o marcador molecular localizado na região promotora do gene *ZmMATE1* em 279 indivíduos F₂ derivados do cruzamento entre L228-3 (linhagem tolerante ao Al) e L53 (linhagem sensível). No eixo X são apresentadas as classes genótípicas do marcador, sendo CC o alelo derivado da L53 em homozigose; TT o alelo derivado da L228-3 em homozigose e CT os heterozigotos.....31
- Figura 4.** Perfil de expressão temporal do gene *ZmMATE1* no ápice radicular. A expressão gênica foi determinada nos tempos de 0 hora, 30 minutos, 1, 6 e 12 horas após a exposição ao Al³⁺. As barras indicam o desvio padrão da média de três repetições técnicas. Letras distintas indicam médias significativamente diferentes pelo Teste T a 5% de probabilidade.....32
- Figura 5.** Expressão tecidual do gene *ZmMATE1* após 6 horas de exposição ao Al³⁺. A expressão gênica foi determinada no ápice radicular, no resto da raiz e na parte aérea. A amostra calibradora utilizada foi o ápice da linhagem L53 após 6 horas de tratamento sem Al³⁺. As barras indicam o desvio padrão da média. Letras distintas indicam médias significativamente diferentes pelo Teste T a 5% de probabilidade33
- Figura 6.** Alinhamento múltiplo das sequências promotoras das linhagens Cateto Al237, L228-3 e L53 utilizando a ferramenta Clustal Omega. Substituições dos nucleotídeos entre as sequências são destacadas em cinza. Os números acima dos SNPs indicam a posição em pares de bases em relação ao ATG predito por Maron *et al.* (2010)37
- Figura 7.** Visão gráfica da distribuição de alguns elementos *cis*-regulatórios, descritos na legenda, ao longo de 2710 pb da região promotora do gene *ZmMATE1* da linhagem L53. Na

posição -2499, indicada pelo asterisco, está representado o SNP T/C, que corresponde ao código Y. Em **A** está representada a sequência de ligação para fatores NAC flanqueando esse SNP. Em **B** encontra-se o motivo de ligação de proteínas MYB no alelo C do genótipo L53. Letras maiúsculas dentro do motivo de ligação representam a sequência *core*, definida como a mais conservada. Letras em vermelho são as posições conservadas e de alto conteúdo informativo na região39

Figura 8. Alinhamento múltiplo das sequências 5'UTR de DNA obtidas pela PCR 5'RACE, das sequências de cDNA do *ZmMATE1* das linhagens Cateto Al237 e L53 (números de acesso: FJ015156.1 e FJ015157.1, respectivamente) e da sequência do *ZmMATE1* da linhagem PH207 (Zm00008a023530) disponível no portal Phytozome. O alinhamento foi feito pela ferramenta Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo/>). Substituições nucleotídicas entre as sequências estão marcadas em cinza. O início da tradução está destacado em preto. Sequências sublinhadas indicam o local de anelamento do primer interno específico usado na 5'RACE. As sequências apresentadas foram alinhadas somente até o *primer* interno da 5'RACE42

Figura 9. Alinhamento esquemático das sequências das 5'UTRs apresentadas na figura 8. Os asteriscos indicam a região das substituições dos aminoácidos. A reta vermelha representa o *primer* interno específico da 5'RACE e o local de alinhamento na sequência.....43

Figura 10. Topologia da proteína ZmMATE1 predita pelo programa TMHMM realçando a predição estrutural dos domínios transmembrana. A) Predição estrutural da proteína com base no cDNA do gene *ZmMATE1* da linhagem Cateto Al237 sequenciado por Maron *et al* (2010). B) Predição estrutural da proteína com base na menor sequência 5'UTR de 85 pb do gene *ZmMATE1*. A seta indica a grande alça intracelular44

Figura 11. Expressão gênica no ápice radicular do gene candidato GRMZM2G034421, similar ao *SbWRKY1*. A expressão foi determinada nos tempos de 0, 30 minutos e 6 horas após o início do tratamento, com três repetições técnicas. As barras indicam o desvio padrão da média de três repetições técnicas. Letras distintas indicam médias significativamente diferentes pelo Teste T a 5% de probabilidade47

Figura 12. Expressão gênica no ápice radicular dos genes candidatos GRMZM2G100641 e GRMZM2G179002 similares ao *SbZNF1*. A expressão foi determinada nos tempos de 0, 30 minutos e 6 horas após o início do tratamento, com três repetições técnicas. As barras indicam o desvio padrão da média de três repetições técnicas. As barras indicam o desvio padrão da

média. Letras distintas indicam médias significativamente diferentes pelo Teste T a 5% de probabilidade48

LISTA DE TABELAS

Tabela 1. Sequências dos *primers* utilizados para amplificação da região promotora do *ZmMATE1* derivado da L228-3.....26

Tabela 2. Sequências dos *primers* desenhados e selecionados para o ensaio de expressão gênica dos genes candidatos codificadores dos fatores de transcrição29

Tabela 3. Identificação, e-value, identidade e posição no cromossomo dos genes candidatos do genoma do milho similares aos fatores de transcrição ART1, STOP1, SbZNF1 e SbWRKY. Os genes GRMZM2G064853, GRMZM2G005207 e GRMZM2G324999 foram incluídos na tabela por serem diferencialmente expressos por Maron *et al.* (2008)46

LISTA DE ABREVIATURAS

ABC	<i>ATP binding cassette</i>
Al ou Al ³⁺	Alumínio
ALMT	<i>Aluminum-Activated Malate Transporter</i>
<i>Alt_{SB}</i>	Loci de tolerância ao Al em sorgo
<i>ART1</i>	<i>Al resistance transcription factor 1</i>
ASR	<i>Abscisic Acid, Stress and Ripening</i>
ALS3	<i>Aluminum Sensitive 3</i>
<i>AtMATE</i>	<i>Arabidopsis thaliana MATE</i>
AtSTOP1	<i>Arabidopsis thaliana STOP1</i>
C _{0d}	Comprimento inicial da raiz seminal
C _{1d}	Comprimento radicular com 1 dia
C _{5d}	Comprimento radicular com 5 dias
cDNA	<i>Complementary desoxyribonucleic acid</i> (DNA complementar)
CL _{+Al}	Crescimento líquido da raiz no tratamento com Al
CL _{-Al}	Crescimento líquido da raiz no tratamento sem Al
CRR	Crescimento relativo da raiz seminal
CTAB	<i>Cetyl trimethylammonium bromide</i> (Brometo de Cetil Trimetilamonio)
DREB	<i>Dehydration Responsive Element Binding</i>
EDTA	<i>Ethylenediamine tetraacetic acid</i> (ácido etilenodiamino tetra-acético)
GBS	<i>Genotyping-by-Sequencing</i> (Genotipagem por sequenciamento)
KASP	<i>Kompetitive Allele-Specific PCR</i>
MATE	<i>Multidrug and toxic compound exudation</i>
mRNA	<i>Messenger ribonucleic acid</i> (RNA mensageiro)
NAC	<u>N</u> o apical meristem (NAM), <u>A</u> rabidopsis transcription activation factor (ATAF), <u>C</u> up-shaped cotyledon (CUC)
NaCl	Cloreto de Sódio
Nramp	<i>Natural Resistance-associated Macrophage Protein</i>
Nrat1	<i>Nramp aluminum transporter 1</i>
OsART1	<i>Oriza sativa ART1</i>
pb	Pares de base
PCR	<i>Polimerase chain reaction</i> (reação em cadeia da polimerase)
RACE	<i>Rapid amplification of cDNA ends</i>
qPCR	<i>Quantitative polimerase chain reaction</i>
QTL	<i>Quantitative Trait Loci</i> (loci de herança quantitativa)
ORF	<i>Open Reading Frame</i> (quadro aberto de leitura)

R ²	Coeficiente de Determinação
5' UTR	<i>Untranslated Region</i> (região 5' não traduzida)
RIL	<i>Recombinat Inbred Line</i> (linhagem endogâmica recombinante)
<i>SbMATE</i>	<i>Sorghum bicolor</i> MATE
<i>SbWRKY1</i>	<i>Sorghum bicolor</i> WRKY1
<i>SbZNF1</i>	<i>Sorghum bicolor</i> ZNF1
SNP	<i>Single Nucleotide Polymorphism</i> (Polimorfismo de nucleotídeo único)
TSS	<i>Transcription Start Sites</i> (Sítios de Início da Transcrição)
<i>STAR1</i>	<i>sensitive to Al rhizotoxicity 1</i>
<i>STOP1</i>	<i>Sensitive to Proton Rhizotoxicity 1</i>
TaALMT1	<i>Triticum aestivum</i> ALMT
TAE	Tris-acetato EDTA
TF	Fator de transcrição
VRE	<i>VIP1 response element</i>
ZmMATE1	<i>Zea mays</i> MATE

RESUMO

Uma das principais restrições à produção de grãos em vários países tropicais é a alta toxidez de alumínio (Al) nos solos ácidos, que limita o crescimento das raízes prejudicando a absorção de água e nutrientes. Em milho (*Zea mays* L.), a tolerância ao Al é uma característica complexa que envolve vários genes, dos quais o gene *ZmMATE1* foi o único caracterizado até o momento como responsável pela tolerância ao Al. Esse gene é homólogo ao *SbMATE* de sorgo que codifica um transportador de membrana que promove a liberação de citrato nos ápices radiculares, formando compostos estáveis não tóxicos com o Al³⁺ presente na solução do solo. A ocorrência de três cópias em *tandem* do gene *ZmMATE1* foi associada com os maiores níveis de expressão desse gene e da tolerância ao Al em uma população derivada da linhagem Cateto Al237, altamente tolerante. Entretanto, a linhagem de milho L228-3 apresenta apenas uma cópia do *ZmMATE1* e altos níveis de expressão do gene, comparáveis aqueles observados em Cateto Al237. Com isso, o objetivo desse trabalho foi caracterizar as bases moleculares do alelo do gene *ZmMATE1* derivado da linhagem elite L228-3. A análise de regressão entre um marcador molecular na região promotora do *ZmMATE1* e o crescimento relativo da raiz seminal mostrou que o alelo derivado da L228-3 foi responsável por 22,17% da tolerância ao Al em uma população F₂. Uma análise de expressão detalhada mostrou que o alelo do *ZmMATE1* foi induzido pelo Al após 1 hora no ápice radicular, apresentando um perfil de expressão similar ao do alelo derivado de Cateto Al237. Por meio do sequenciamento e de predições *in silico* de uma região de -2710 pares de base em relação ao códon de iniciação do *ZmMATE1*, foi identificado um SNP (Polimorfismo de Nucleotídeo Único) entre a L53 e as linhagens tolerantes flanqueando ou dentro de regiões conservadas para ligação de fatores de transcrição, sugerindo que esse SNP pode controlar a expressão do *ZmMATE1*. A amplificação, clonagem e sequenciamento da região 5' não-traduzida (5'UTR) revelou a existência de diferentes sítios de início da transcrição para o *ZmMATE1*, uma vez que foram identificados fragmentos de diferentes tamanhos e com polimorfismos de sequência nas três linhagens. Tais resultados sugerem um provável mecanismo de controle da expressão do *ZmMATE1*. Finalmente, foi analisada a expressão de genes candidatos que codificam fatores de transcrição com similaridade de sequência aos genes *STOP1*, *ART1*, *SbZNF1* e *SbWRKY1* que atuam na regulação de genes envolvidos na tolerância ao Al em outras espécies. Um dos genes com domínio DHHC apresentou um perfil de expressão compatível com o esperado para um repressor da expressão do *ZmMATE1*. Assim, esses resultados serão alvos para estudos complementares visando elucidar os possíveis mecanismos de controle da expressão do *ZmMATE1* e da tolerância ao Al em milho.

Palavras-chaves: Transportador de citrato; ápice radicular; expressão gênica; 5'UTR.

ABSTRACT

One of the major constraints to grain yield in tropical countries is the aluminum (Al) toxicity on acid soils, which limits root growth and impairs water and mineral nutrient uptake. In maize (*Zea mays* L.), Al tolerance is a complex trait that involves several genes, of which *ZmMATE1* was the only one characterized as responsible for Al tolerance. This gene is homolog to sorghum *SbMATE*, which encodes a membrane transporter that mediates citrate exudation in root apices and forms stable non-toxic compounds when combined with Al³⁺ in the soil solution. The occurrence of three copies *in tandem* of *ZmMATE1* was associated with high levels of expression of this gene and with Al tolerance in a population derived from Cateto Al237, a highly Al-tolerant line. Nonetheless, the maize line L228-3 shows only one copy of *ZmMATE1* and high expression levels of this gene, comparable to those observed in Cateto Al237. Therefore, this study aimed to characterize the molecular basis of the *ZmMATE1* allele derived from the elite maize line L228-3. Regression analysis between a molecular marker in the *ZmMATE1* promoter region and the relative growth of the seminal root showed that the allele derived from L228-3 was responsible for 22.17% of Al tolerance in an F₂ population. A detailed expression analysis showed that this allele was induced by Al after 1 hour in the root tips and in the rest of the root, showing an expression profile similar to the allele derived from Cateto Al237. Sequencing and *in silico* predictions of -2710 base pairs to the start codon of *ZmMATE1* identified a SNP (Single Nucleotide Polymorphism) between L53 and the Al-tolerant lines flanking or co-localized to conserved binding regions of transcription factors, suggesting that this SNP may control *ZmMATE1* expression. Amplification, cloning and sequencing of the 5' untranslated region (5'UTR) revealed the presence of different transcriptional start sites for *ZmMATE1*, once fragments of different sizes and polymorphisms were identified among the three lines. These results indicate a possible control of *ZmMATE1* expression. Finally, the expression pattern of candidate genes encoding transcription factors with sequence similarity to *STOP1*, *ART1*, *SbZNF1* and *SbWRKY1* were evaluated. These transcription factors regulate the expression of genes involved in the Al tolerance in other species. One of the genes with DHHC domain presented an expression profile compatible with a putative repressor of *ZmMATE1* expression. Thus, these results are new targets for complementary studies aiming to disclose the mechanisms controlling *ZmMATE1* expression and Al tolerance in maize.

Key words: Citrate transporter; root apex; gene expression; 5'UTR.

1. INTRODUÇÃO

Os solos ácidos estão amplamente distribuídos nas regiões tropicais e subtropicais, correspondendo a aproximadamente 50% das terras cultiváveis do planeta. Devido à acidez desses solos, o alumínio encontra-se na forma de cátion trivalente (Al^{3+}) sendo altamente tóxico às raízes, o que inibe o crescimento radicular e conseqüentemente a obtenção de água e nutrientes essenciais ao desenvolvimento das plantas (Kochian, 1995).

A exsudação de ácidos orgânicos através de transportadores de membranas localizados no ápice radicular é um importante mecanismo fisiológico de tolerância ao Al em plantas. Os ácidos orgânicos formam compostos estáveis com o Al^{3+} neutralizando a toxicidade desse íon (Ma *et al.*, 2001). Os primeiros genes identificados que controlam a tolerância ao Al em planta foram aqueles que codificam transportadores de malato e citrato pertencentes às famílias ALMT e MATE, respectivamente (Sasaki *et al.*, 2004; Magalhães *et al.*, 2007; Furukawa *et al.*, 2007).

A tolerância ao Al em milho é uma característica complexa, tendo sido identificadas regiões genômicas (QTLs) explicando de 5 a 30% da variância fenotípica dessa característica em diferentes populações (Sibov *et al.*, 1999; Ninamango-Cárdenas *et al.*, 2003; Conceição *et al.*, 2009; Guimarães *et al.*, 2014). O gene *ZmMATE1*, que codifica um transportador de citrato, co-localiza com um QTL de efeito maior associado com a tolerância ao Al no cromossomo 6 de milho (Maron *et al.*, 2010). Esse gene é induzido pelo Al e possui alta expressão no ápice radicular da linhagem Cateto Al237 (tolerante) em comparação com L53, sensível ao Al (Maron *et al.*, 2010). Os altos níveis de expressão e de tolerância ao Al em Cateto Al237 foram associados com a presença de três cópias em *tandem* do gene *ZmMATE1*, combinação alélica rara em milho (Maron *et al.*, 2013). Entretanto, a linhagem tropical L228-3 apresenta expressão do *ZmMATE1* tão alta quanto em Cateto Al237 após 6 horas de tratamento com Al, mesmo apresentando apenas uma cópia do gene (Guimarães *et al.*, 2014). Assim essa linhagem, que possui características importantes para o melhoramento de milho da Embrapa, pode também apresentar um mecanismo de controle da expressão do gene *ZmMATE1* diferente da variação no número de cópias como proposto para o alelo da Cateto Al237.

2. REVISÃO BIBLIOGRÁFICA

2.1 O milho (*Zea mays* L.)

O milho é o cereal mais cultivado no mundo, sendo o Brasil o terceiro maior produtor mundial com a produção estimada em mais de 90 milhões de toneladas na safra de 2016/2017 (CONAB, 2017), o que representa grande destaque no setor agrícola nacional. Embora a produção brasileira seja alta, ela representa apenas 8% da produção mundial, superada pelos

Estados Unidos e China que, juntos, são responsáveis por mais da metade de toda a produção mundial (FAO, 2017). A maior parte dessa produção é voltada para a alimentação animal e, nos últimos anos, a demanda como matéria-prima para a geração de etanol têm crescido de forma significativa (Wallington *et al.*, 2012). Isso evidencia a importância da cadeia produtiva do milho no cenário econômico mundial e brasileiro, no qual a agricultura possui um grande componente do PIB nacional.

2.2 Solos ácidos

Os solos ácidos estão presentes em aproximadamente 50% das terras agricultáveis do mundo e se caracterizam pela alta toxidez de alumínio (Al) (von Uexküll & Mutert, 1995), o que configura uma das principais restrições à produção de grãos em vários países tropicais. Esses solos possuem valores de pH abaixo de 5,0, e sob condições de acidez, o alumínio encontra-se na forma solúvel Al^{3+} , que possui ação rizotóxica mesmo em concentrações muito baixas. O Al inibe o crescimento da raiz pela rápida inibição da expansão e do alongamento celular, cujo danos causados ao sistema radicular restringem o volume de solo explorado, resultando em prejuízos na absorção de nutrientes e de água do solo (Kochian, 1995; Kochian *et al.*, 2004) e, conseqüentemente, na produção. A toxicidade por Al é a segunda maior limitação abiótica para a produção agrícola, sendo a primeira o estresse causado pela seca (von Uexküll & Mutert, 1995). Além disso, os efeitos causados pelo Al levam a uma maior sensibilidade à seca reduzindo a produtividade das culturas.

Como estratégia agrônômica para minimizar os efeitos danosos do alumínio, aplica-se calcário nos solos ácidos, o que promove a correção do pH. Entretanto, a calagem é eficiente na correção do pH apenas nas camadas superficiais do solo (0 a 20 cm) (Foy *et al.*, 1984). Além disso, o uso do calcário implica em aumento nos custos de produção, que para alguns agricultores se torna proibitivo e, em algumas regiões, o acesso ao insumo é escasso, limitando a produção agrícola nessas condições.

2.3 Mecanismos fisiológicos de tolerância ao Al

Os mecanismos fisiológicos de tolerância ao Al podem ser divididos em duas classes distintas: apoplásticos e simplásticos (Kochian, 1995; Kochian *et al.*, 2004). O mecanismo apoplástico ou de exclusão impede que o Al^{3+} tóxico atravesse a membrana plasmática e penetre no simplasto (Ryan *et al.*, 2011) por meio da liberação de ácidos orgânicos na rizosfera principalmente pelos ápices radiculares de genótipos tolerantes (Delhaize *et al.*, 1993 a; b). Os ácidos orgânicos (malato, citrato e oxalato) agem como quelantes de íons Al^{3+} , formando compostos estáveis não tóxicos na rizosfera e conferem aos vegetais tolerância à toxidez causada por esse metal (Ma *et al.*, 2001). Esse mecanismo é o mais bem caracterizado de tolerância ao Al nas plantas, sendo descrito em um grande número de

espécies monocotiledôneas e dicotiledôneas (Kochian *et al.*, 2004). A exsudação de malato pela raiz confere tolerância ao Al em centeio (Collins *et al.*, 2008) e colza (Ligaba *et al.*, 2006). O milho (Pellet *et al.*, 1995; Maron *et al.*, 2010) e o sorgo (Magalhães *et al.*, 2007) exsudam citrato, enquanto que o trigo sarraceno e o inhame liberam o oxalato pelas raízes (Zheng *et al.*, 1998a; Ma & Miyasaka, 1998). Já outras espécies, como aveia (Zheng *et al.*, 1998b), *Arabidopsis* (Liu *et al.*, 2009), soja (Yang *et al.*, 2000; Silva *et al.*, 2001; Liang *et al.*, 2013) e trigo (Ryan *et al.*, 2009) exsudam tanto citrato quanto malato.

No mecanismo simplástico ou de tolerância interna, o Al penetra na célula, sendo imobilizado e compartimentalizado em vacúolos ou detoxificado no citosol (Ryan *et al.*, 2011). Algumas espécies de plantas como hortêncica (*Hydrangea macrophylla*) (Ma *et al.*, 1997a), trigo sarraceno (Ma & Hiradate, 2000) e *Camellia sinensis* (Morita *et al.*, 2004) são capazes de translocar e acumular Al nas folhas em concentrações acima de 1000 mg/kg, e em alguns casos excedendo 3000 mg/kg. No trigo sarraceno, o ácido oxálico participa da detoxificação do Al dentro das células das folhas formando o complexo não tóxico Al-oxalato (Ma *et al.*, 1997b). Os mecanismos de tolerância interna parecem ser comuns em plantas tropicais e subtropicais ou em regiões onde os solos ácidos são endêmicos (Metali *et al.*, 2011).

Dependendo da espécie, os dois mecanismos de tolerância podem operar sozinhos ou de forma combinada como em arroz, modulando os níveis de tolerância à toxidez de Al (Ryan *et al.*, 2011).

2.4 Genes associados com a tolerância ao Al

Vários genes associados com a tolerância ao Al já foram identificados e caracterizados em plantas (Figura 1). Em trigo (*Triticum aestivum*), foi clonado o primeiro gene de tolerância ao Al, designado *ALMT1*, pertencente à família gênica *ALMT* (*Aluminum-Activated Malate Transporter*). Esse gene codifica um transportador de malato, ativado por Al, responsável pelo efluxo de malato no ápice radicular de genótipos tolerantes ao Al (Sasaki *et al.*, 2004). O gene *TaALMT1* co-localiza com o loco principal (*Alt_{BH}*), que controla cerca de 85% da variação fenotípica da tolerância ao Al em trigo (Riede & Anderson, 1996). Ortólogos ao gene *ALMT* estão envolvidos na exsudação de malato e na tolerância ao Al em *Arabidopsis* (*AtALMT1*; Hoekenga *et al.*, 2006), colza (*BnALMT1* e *BnALMT2*; Ligaba *et al.*, 2006), centeio (*ScALMT1*; Collins *et al.*, 2008) e soja (*GmALMT1*; Liang *et al.*, 2013).

Já em sorgo (*Sorghum bicolor*), a tolerância ao Al é controlada pelo loco (*Alt_{SB}*) que explica aproximadamente 80% da variação fenotípica dessa característica (Magalhaes *et al.*, 2004). O gene que controla esse loco foi caracterizado como *SbMATE*, que codifica um transportador de membrana da família *MATE* (*multidrug and toxic compound exudation*), é ativado por Al e responsável pela exsudação de citrato no ápice radicular de linhagens tolerantes (Magalhaes *et al.*, 2007). Ortólogos desse gene foram caracterizados em plantas

de *Arabidopsis* (*AtMATE*; Liu *et al.*, 2009), trigo (*TaMATE1*; Ryan *et al.*, 2009), milho (*ZmMATE1*; Maron *et al.*, 2010), *Vigna umbellata* (*VuMATE1*; Yang *et al.*, 2011), arroz (*OsFRDL4*; Yokosho *et al.*, 2011) e couve (*BoMATE*; Wu *et al.*, 2014). Como os genes das duas famílias *ALMT* e *MATE* conferem tolerância ao Al através da liberação de ácidos orgânicos pela raiz, parece ser um exemplo marcante de coevolução das vias de tolerância ao Al por meio de transportadores que são estrutural e funcionalmente muito diferentes (Kochian *et al.*, 2015).

Transportadores do tipo ABC (*ATP binding cassette*), uma das maiores famílias de proteínas transmembrana presente em procariotos e eucariotos (Henikoff *et al.*, 1997), também estão envolvidos com a tolerância ao Al em *Arabidopsis* e em arroz. Em arroz, os genes *STAR1* e *STAR2* codificam o domínio de ligação de ATP e o domínio transmembrana, respectivamente, que são as subunidades formadoras do transportador do tipo ABC (Huang *et al.*, 2009). Esse complexo proteico é responsável pelo transporte de UDP-glicose para dentro da parede celular da raiz, que leva a modificações na parede, limitando o acúmulo de Al^{3+} e reduzindo a toxicidade por Al (Huang *et al.*, 2009). Em *Arabidopsis*, o gene *AtALS3* codifica um transportador ABC localizado na membrana plasmática que pode estar envolvido na redistribuição intracelular do Al na planta mantendo-o longe do ápice sensível da raiz (Larsen *et al.*, 2005).

Outra forma que algumas plantas desenvolveram para reduzir os níveis de Al prejudiciais à parede celular é translocar esse Al para dentro do citoplasma das células radiculares e assim transportá-lo para dentro de vacúolos através de transportadores presentes na membrana plasmática e no tonoplasto. O gene *Nrat1* de arroz codifica um transportador de membrana específico para Al^{3+} (Xia *et al.*, 2010). *Nrat1* (*Nramp aluminum transporter 1*) pertence à família *Nramp* (*Natural Resistance-associated Macrophage Protein*) e se localiza na membrana plasmática de todas as células do ápice radicular, exceto nas epidérmicas. *OsNrat1* possivelmente atua em conjunto com um transportador do tipo ABC, *OsALS1*, para remover o Al da parede celular e transportá-lo para vacúolos nas células da raiz (Xia *et al.*, 2010; Huang *et al.*, 2012).

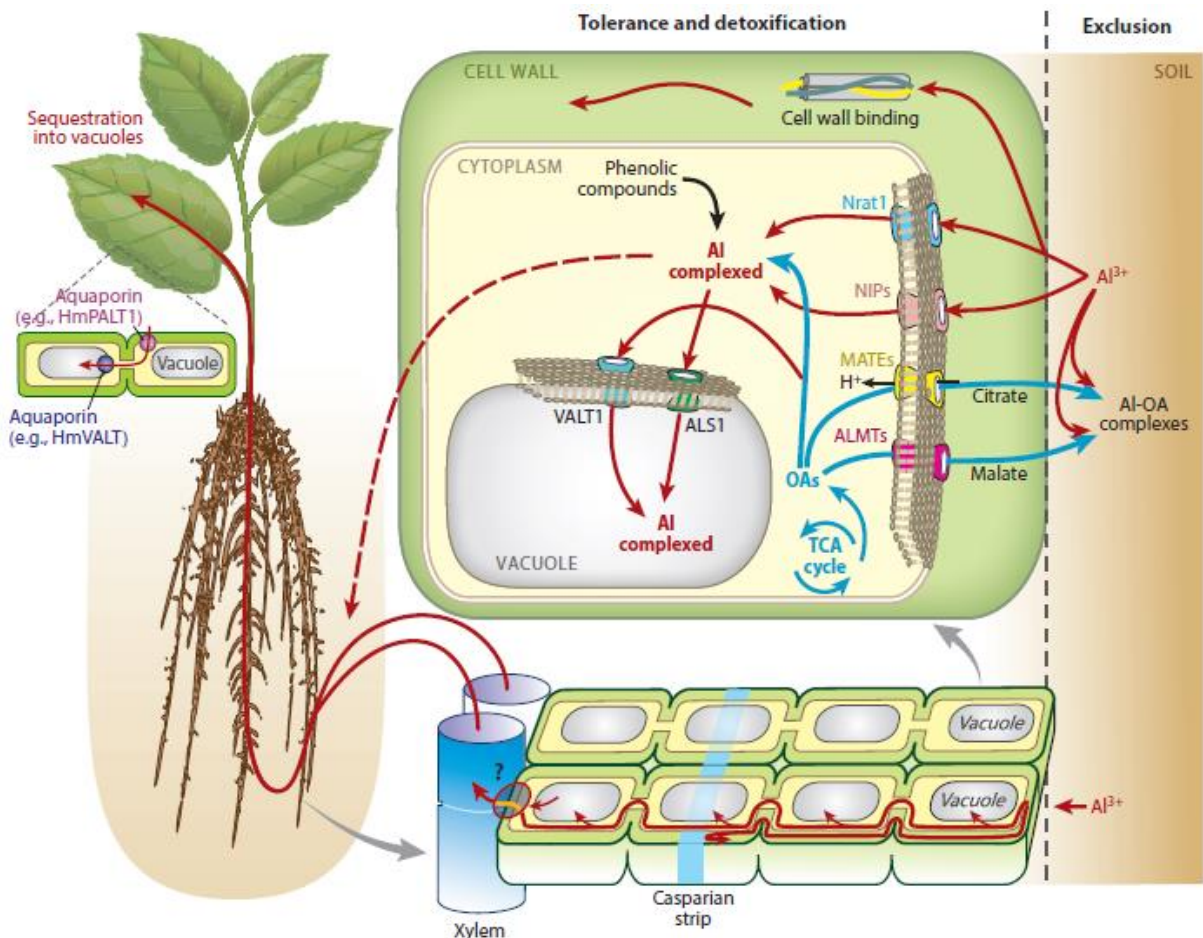


Figura 1. Modelo ilustrando os mecanismos apoplástico e simplástico de tolerância ao Al nas plantas e alguns genes associados com a tolerância. As setas vermelhas representam o fluxo de Al para dentro e no interior da célula e as setas azuis indicam o fluxo dos ácidos orgânicos (OA). Adaptado de Kochian *et al.* (2015).

2.5 Regulação da expressão gênica da tolerância ao Al

Inúmeras pesquisas investigando a expressão dos diversos genes envolvidos na tolerância ao Al identificaram elementos *cis* e fatores atuando em *trans* que regulam as vias induzidas por esse metal, além de algumas características gerais da regulação gênica. Os estudos vêm mostrando que a expressão dos genes de tolerância ao Al é maior nos genótipos tolerantes e é localizada nos ápices radiculares, região onde a toxicidade do Al causa maiores danos. Apesar dessas similaridades entre os genes em diferentes espécies, a regulação da expressão gênica pode apresentar diferenças marcantes (Kochian *et al.*, 2015). A expressão de genes *ALMT* e *MATE* é induzida pelo estresse por Al, por exemplo, em *Arabidopsis*, arroz, sorgo, milho e centeio (Kobayashi *et al.*, 2007; Liu *et al.*, 2009; Yokosho *et al.*, 2011; Magalhaes *et al.*, 2007; Maron *et al.*, 2010; Colins *et al.*, 2008; Yokosho *et al.*, 2010). Enquanto que em trigo e cevada, a expressão dos genes *TaALMT1* e *HvAACT1*, respectivamente, é constitutiva (Sasaki *et al.*, 2004; Ryan *et al.*, 2009; Furukawa *et al.*, 2007) e o efluxo de malato

e citrato ocorre independente da exposição ao Al. Já nas espécies em que a expressão é induzida por Al, o ponto máximo de expressão e de exsudação dos ácidos orgânicos ocorrem após a exposição ao estresse, podendo variar de 2 a 6 horas em *Arabidopsis* e milho (Liu *et al.*, 2009; Maron *et al.*, 2010), até de 4 a 6 dias, como ocorre em sorgo (Magalhaes *et al.*, 2007).

Em *Arabidopsis*, a exsudação de malato ativada pelo Al nas raízes é mediada pelo transportador ALMT1, que é codificado pelo gene *AtALMT1*, um homólogo do *TaALMT1* (Hoekenga *et al.*, 2006). A expressão de *AtALMT1* requer um fator de transcrição do tipo *zinc finger-C₂H₂* induzido pela presença de H⁺ e Al³⁺, codificado pelo gene *STOP1* (Iuchi *et al.*, 2007). Liu *et al.* (2009) demonstraram que o fator de transcrição STOP1 é essencial também para a expressão do *AtMATE*. Assim, apesar de os genes *AtALMT1* e *AtMATE* agirem independentemente para conferir tolerância ao alumínio em *Arabidopsis*, a expressão de ambos é controlada pelo STOP1 (Liu *et al.*, 2009).

O homólogo do *STOP1* em arroz, *ART1* (*Al resistance transcription factor 1*), também é um fator de transcrição do tipo *zinc finger-C₂H₂*, que regula a expressão de pelo menos 31 genes envolvidos na tolerância ao Al em arroz (Yamaji *et al.*, 2009), entre eles os genes *STAR1* e *STAR2* (Huang *et al.*, 2009), *Nrat1* (Xia *et al.*, 2010), *OsFRDL4* (Yokosho *et al.*, 2011) e *OsALS1* (Huang *et al.*, 2012). Um estudo no promotor do gene *OsSTAR1* identificou o elemento *cis* de ligação do ART1, que consiste na sequência GGN(T/g/a/C)V(C/A/g)S(C/G). Esse elemento *cis* está presente na região promotora de 29 genes dos 31 regulados por esse fator de transcrição (Tsutsui *et al.*, 2011).

Outra proteína que atua como um fator de transcrição pertence à família ASR (*Abscisic Acid, Stress and Ripening*), essencial para a expressão de genes relacionados à tolerância ao Al em arroz. ASR5 ativa a expressão do gene *OsSTAR1* através da interação direta com elementos *cis* no promotor do gene *STAR1*, além de regular um grande número de genes responsivos ao Al (Arenhart *et al.*, 2013; 2014). Recentemente, foi demonstrado que ASR5 atua de forma combinada e complementar com a proteína ASR1 para regular a expressão de genes envolvidos na tolerância ao Al em arroz, além de se auto-regular (Arenhart *et al.*, 2016). Em resposta ao Al, a expressão de *ASR5* aumenta por um mecanismo ainda desconhecido e regula positivamente o ASR1, que juntos ao ART1 aumentam a expressão do gene *STAR1* (Arenhart *et al.*, 2016).

2.6 Tolerância ao Al em milho

Estudos genéticos sugerem que a tolerância ao Al em milho é uma característica de herança complexa (Magnavaca *et al.*, 1987; Pandey *et al.*, 1994; Piñeros *et al.*, 2005) e o mapeamento de QTLs (*Quantitative Trait Loci*, ou locos de herança quantitativa) é uma estratégia importante para elucidar essa complexidade genética. Sibov *et al.* (1999)

identificaram dois QTLs de tolerância ao Al nos cromossomos 6 e 10 utilizando marcadores RFLP, enquanto Ninamango-Cárdenas *et al.* (2003) mapearam cinco QTLs nos cromossomos 2, 6 e 8 de milho por meio de marcadores RFLP e SSR. Posteriormente, Conceição *et al.* (2009) identificaram QTLs nos cromossomos 4, 5, 6, 8 e 10, utilizando marcadores SSR. Recentemente, Guimarães *et al.* (2014) mapearam cinco QTLs de tolerância ao Al nos cromossomos 2, 3, 5, 6 e 8, em uma população de linhagens recombinantes derivadas de Cateto Al237, utilizando alta densidade de marcadores gerados por GBS (*Genotyping-by-Sequencing*; Genotipagem por sequenciamento).

Para elucidar a base molecular da tolerância ao Al em milho, Maron *et al.* (2008) realizaram uma análise de expressão temporal de genes sob estresse com Al usando microarranjos com os genótipos sensível (L53) e tolerante (C100-6) ao Al. Um grande número de genes foram identificados exibindo padrões de expressão diferencial ao longo do tempo de exposição ao Al e entre as linhagens, sendo alguns deles validados por PCR em tempo real (Maron *et al.*, 2008). Posteriormente, dois membros da família *MATE*, diferencialmente expressos nesse experimento de microarranjos, foram co-localizados com dois QTLs de tolerância ao alumínio e submetidos a uma caracterização molecular e funcional (Maron *et al.*, 2010). Dentre eles, o gene *ZmMATE1* foi caracterizado como o gene homólogo do *SbMATE* por ativar a exsudação de citrato em ápices radiculares de *Arabidopsis* superexpressando esse gene, contribuindo para o aumento da tolerância ao Al nessas plantas (Maron *et al.*, 2010). Além disso, a sequência predita da proteína apresenta identidade com a proteína de sorgo *SbMATE* de 52% e de 64% com a proteína *AtMATE* de *Arabidopsis*, cuja posição dos domínios transmembrana é conservada em relação ao *SbMATE*. A expressão do *ZmMATE1* é concentrada nos tecidos radiculares, particularmente no ápice da raiz de milho, e consideravelmente mais alta em genótipos tolerantes ao alumínio, Cateto Al237 e C100-6, quando comparada ao genótipo sensível L53, tanto na ausência como na presença de alumínio (Maron *et al.*, 2010).

A clonagem e o sequenciamento do *ZmMATE1* entre os genótipos tolerante (Cateto Al237) e sensível (L53) revelaram seis nucleotídeos diferentes entre as duas regiões codificantes, dos quais apenas dois resultaram na substituição de aminoácidos, que não refletiam em diferenças funcionais nas proteínas codificadas (Maron *et al.*, 2010). Posteriormente, Maron *et al.* (2013) associaram a presença de três cópias em *tandem* do *ZmMATE1* na linhagem tolerante Cateto Al237 com a maior expressão desse gene e, conseqüentemente com um aumento na tolerância ao Al, quando comparado a uma única cópia do gene da linhagem sensível L53 (Maron *et al.*, 2013). A versão alélica com três cópias do *ZmMATE1* é rara em milho e está presente em apenas três das 166 linhagens avaliadas por Maron *et al.* (2013). Um trabalho recente revelou que a linhagem de milho L228-3 apresenta altos níveis de expressão do gene *ZmMATE1*, comparáveis à Cateto Al237, mas

possui apenas uma única cópia do gene (Guimarães *et al.*, 2014), podendo ser considerado como um alelo diferente do *ZmMATE1* em relação ao alelo de tolerância caracterizado na Cateto AI237.

Uma estratégia promissora para aumentar o rendimento agrícola é minimizar os efeitos da toxicidade ao AI explorando a diversidade genética, principalmente no caso do milho cuja tolerância ao AI é controlada por vários genes. Portanto, compreender as bases genéticas e moleculares da tolerância ao AI é de fundamental importância para acelerar o desenvolvimento de cultivares superiores e aumentar a produtividade agrícola em solos ácidos. Adicionalmente, os fatores genéticos responsáveis pelo controle dessa característica podem ser utilizados como ativos de inovação. Assim, o objetivo desse trabalho foi caracterizar em detalhes esse alelo do gene *ZmMATE1*, derivado da linhagem L228-3 quanto ao efeito na tolerância ao AI, padrão de expressão e possíveis mecanismos que regulam a expressão superior desse alelo, que não está associada com a variação do número de cópias do gene, como proposto para o alelo presente na linhagem Cateto AI237.

3. OBJETIVOS

3.1 Objetivo geral

Realizar a caracterização molecular do alelo do gene *ZmMATE1* derivado da linhagem elite de milho L228-3.

3.2 Objetivos específicos

- Avaliar o efeito do alelo do gene *ZmMATE1* derivado da linhagem L228-3 na resposta à tolerância ao AI em uma população F₂;
- Caracterizar o perfil de expressão temporal e tecidual do alelo do gene *ZmMATE1* da linhagem L228-3 em comparação com as linhagens contrastantes Cateto AI237 e L53;
- Identificar polimorfismos na região promotora e caracterizar a região 5'UTR do gene *ZmMATE1* nas linhagens L228-3, Cateto AI237 e L53;
- Identificar genes candidatos em milho que codifiquem fatores de transcrição possivelmente envolvidos na tolerância ao AI.

4. METODOLOGIA

4.1 Material genético

Foram usadas linhagens tropicais de milho Cateto Al237, L53 e L228-3. As linhagens Cateto Al237 e L53 são padrões de alta tolerância e alta sensibilidade ao Al, que foram utilizadas nos estudos de clonagem e caracterização do gene *ZmMATE1* (Maron *et al.*, 2010; 2013). A linhagem L228-3 será alvo do presente estudo por apresentar um alelo do gene *ZmMATE1* diferente daquele derivado da linhagem Cateto Al237, com alta expressão após 6 horas de exposição ao Al e apenas uma cópia do *ZmMATE1* (Guimarães *et al.*, 2014).

4.2 Associação do *ZmMATE1* com a tolerância ao alumínio

Para o experimento de tolerância ao Al, foi utilizada uma população de 279 indivíduos F_2 obtida pela autofecundação da geração F_1 derivada do cruzamento entre as linhagens L228-3 (tolerante) e a L53 (sensível).

O experimento foi conduzido em câmara de crescimento com temperatura diurna média de $27\pm 3^\circ\text{C}$ e noturna de $20\pm 3^\circ\text{C}$, intensidade luminosa de $330 \mu\text{mol photons m}^{-2}\text{s}^{-1}$ e 12 horas de fotoperíodo. As sementes foram desinfestadas utilizando solução de hipoclorito de sódio 0.5% (v/v) durante 5 minutos sob agitação a 160 rpm e posteriormente lavadas oito vezes com água deionizada. Em seguida, as sementes foram germinadas por quatro dias com 16 horas de fotoperíodo em temperatura de 28°C em rolos de papel de germinação umedecido. Após os quatro dias, as plântulas foram transferidas para copos de polietileno, uma plântula por copo, imersas em solução nutritiva descrita por Magnavaca *et al.* (1987) com pH ajustado para 4,0 sob aeração contínua. Após o período de aclimação por 24 horas em solução nutritiva sem Al, o comprimento inicial da raiz seminal de cada plântula (C_{0d}) foi medido manualmente com uma régua. Uma segunda medição foi realizada 24 horas após a primeira (C_{1d}) e então as plântulas foram transferidas para a solução nutritiva com $\{39\} \mu\text{M}$ de atividade de Al^{3+} e pH ajustado para 4,0. O comprimento radicular foi novamente obtido após cinco dias de exposição ao Al (C_{5d}). Como o índice foi avaliado com base em planta individual, o crescimento da raiz sem Al ($\text{CL}_{-\text{Al}}$) foi obtido pela diferença entre C_{1d} e C_{0d} multiplicada por cinco, para comparar com o crescimento durante os cinco dias com Al ($\text{CL}_{+\text{Al}}$), que foi calculado pela diferença entre C_{5d} e C_{1d} . Assim, o índice para avaliar a tolerância ao Al foi o crescimento relativo da raiz seminal (CRR), dividindo o $\text{CL}_{+\text{Al}}$ pelo $\text{CL}_{-\text{Al}}$ para cada plântula. As linhagens parentais foram utilizadas como controles utilizando as médias do crescimento da raiz de sete plântulas após cinco dias na solução com Al dividido pela média sem Al (Guimarães *et al.*, 2014).

O DNA genômico de cada plântula foi extraído a partir de folhas pelo método do CTAB descrito por Saghai-Marooof *et al.* (1984). O DNA foi quantificado por NanoDrop 1000 (Thermo

Fisher Scientific, Waltham, MA) e diluído para a concentração de uso de 10 ng/μL. A genotipagem da população F₂ foi feita utilizando um marcador SNP baseado no ensaio KASP (*Kompetitive Allele-Specific PCR*) desenvolvido pela empresa LGC Genomics (www.lgcgenomics.com). Esse SNP está localizado na região promotora do *ZmMATE1* e foi validado por Barros *et al.* (2016), sendo polimórfico entre as linhagens parentais.

A reação de amplificação foi realizada com 3 μL de Kasp Master Mix, 30 ng de DNA e 0,084 μL de Kasp Assay Mix contendo os *primers*. Os ciclos de amplificação foram: desnaturação inicial a 94°C durante 15 minutos, seguidos por 10 ciclos a 94°C durante 20 segundos, 61 °C durante 1 minuto reduzindo 0,6 °C por ciclo, seguido de mais 26 ciclos de 94°C durante 20 segundos e 55°C durante 1 minuto. A intensidade da fluorescência das amostras foi quantificada por meio do leitor de microplacas FLUOstar Omega Filter-based multi-mode microplate reader (BMG Labtech, Ortenberg, Alemanha) utilizando ROX na normalização do sinal. A genotipagem foi realizada utilizando o software KlusterCaller 1.1 (LGC Genomics, Teddington, Inglaterra).

Para avaliação da segregação genotípica da população F₂, as proporções obtidas foram comparadas com proporções esperadas (1:2:1) pelo teste de qui-quadrado (X²). A associação entre o marcador e o fenótipo (CRR) foi calculada por meio de regressão linear simples, utilizando o modelo:

$$y = \beta_0 + \beta_1 X + \varepsilon,$$

onde: y corresponde a variável resposta (CRR); β_0 é o intercepto no eixo y; β_1 equivale ao efeito; X é o valor atribuído à variável explicativa (*ZmMATE1*); e ε é o erro associado ao modelo. O valor do coeficiente de determinação (R²) indica a porcentagem explicada pela variável explicativa.

A estimativa do modo de ação gênica do *ZmMATE1* foi feita utilizando o grau médio de dominância (*d/a*), que foi obtido pela razão entre os efeitos da dominância (*d*) e aditivo (*a*), onde $d = TC - [(TT+CC)/2]$ e $a = (TT-CC)/2$. TT corresponde à média dos valores de CRR dos indivíduos homocigotos para o alelo T do parental L228-3, enquanto que CC representa a média de CRR dos indivíduos homocigotos para o alelo C derivado do parental L53, e TC denota a média dos valores de CRR dos heterocigotos da população F₂. Os valores de *d/a* adotados no presente estudo foram: recessivo ($d/a \leq -0,7$), parcialmente recessivo ($-0,7 < d/a < -0,3$), aditivo ($-0,3 \leq d/a \leq 0,3$), parcialmente dominante ($0,3 < d/a < 0,7$), dominante ($d/a \geq 0,7$).

4.3 Caracterização do perfil de expressão do *ZmMATE1*

O gene *ZmMATE1* foi caracterizado quanto ao padrão de expressão por meio de RT-PCR quantitativo (qRT-PCR) utilizando o *primer* direto 5'-TGTGAGTTTGGCGGATGTGT-3' e o *primer* reverso 5'-TCACAATCTAGGCCAGTACAACAGA-3' (Maron *et al.*, 2010; Guimarães *et al.*, 2014).

O experimento para avaliar o perfil de expressão do *ZmMATE1* foi conduzido em câmara de crescimento nas mesmas condições utilizadas para avaliar a tolerância ao Al descrito no item anterior. Após os quatro dias de germinação das sementes, as plântulas foram transferidas para copos de polietileno e imersas em solução nutritiva descrita por Magnavaca *et al.* (1987) com pH ajustado para 4,0 sob aeração contínua. Após aclimação durante 24 horas nessa solução sem Al, foram iniciados os tratamentos em solução contendo {39} µM de atividade de Al, e o controle foi realizado no tempo zero sem Al, antes das plantas serem submetidas ao estresse. A expressão temporal foi quantificada no ápice radicular no tempo zero, após 30 minutos, 1, 6 e 12 horas com e sem o tratamento com {39} µM de atividade de Al.

A coleta dos tecidos foi realizada com auxílio de um bisturi, o material foi congelado imediatamente em nitrogênio líquido e armazenado a -80°C. Cada amostra biológica, representando um genótipo em um determinado tempo, foi constituída por 21 plântulas e o experimento realizado com três repetições técnicas. A tecido-especificidade foi avaliada utilizando o primeiro centímetro do ápice radicular, o restante da raiz e a parte aérea após 6 horas de tratamento com e sem Al.

A extração do RNA total para cada amostra biológica foi feita utilizando o kit de extração RNeasy® Plant Mini Kit (Qiagen, Hilden, Alemanha), segundo as recomendações do fabricante e tratadas com DNase I por 20 minutos à temperatura ambiente.

As amostras foram quantificadas utilizando o espectrofotômetro NanoDrop 1000 (Thermo Fisher Scientific, Waltham, MA). A verificação da integridade do RNA foi feita por eletroforese em gel de agarose 2,0% com tampão Tris-acetato EDTA (TAE) 1x e corados com GelRed Nucleic Acid Stain (Biotium, Fremont, CA). A reação de transcrição reversa, para a obtenção do cDNA, foi realizada com o kit "High Capacity cDNA Reverse Transcriptase" (Applied Biosystems, Thermo Fisher Scientific, Waltham, MA), conforme as indicações do fabricante.

Os ensaios de expressão foram realizados no equipamento 7500 Fast-Real Time PCR (Applied Biosystems), seguindo-se os protocolos e recomendações do fabricante baseados na metodologia SYBR Green. O gene RNA ribossômico 18S foi utilizado como gene de referência e a expressão relativa dos genes foi obtida utilizando o método $2^{-\Delta\Delta C_t}$ (Livak & Schmittgen, 2001) com três repetições técnicas por amostra. A linhagem L53 na condição controle (tempo 0 e sem Al) foi utilizada como calibradora.

4.4 Sequenciamento e caracterização da região promotora do gene *ZmMATE1*

A região promotora contendo 2710 pb *upstream* ao ATG predito para o gene *ZmMATE1* foi sequenciada na linhagem L228-3 e comparada com as sequências da mesma região das linhagens Cateto AI237 e L53 obtidas por Maron *et al* (2013), que também foram utilizadas para o desenho dos *primers* para amplificação dos fragmentos e posterior sequenciamento. Os *primers* foram desenhados com a ferramenta Primer-Blast (<http://www.ncbi.nlm.nih.gov/>).

Tabela 1. Sequências dos *primers* utilizados para amplificação da região promotora do *ZmMATE1* derivado da L228-3

Primer	Sequência 5'- 3'*	Tamanho do produto	Localização**
P1	Direto GATGGTACCAATACGTCCTACCACACCGC'	734 pb	-2710
	Reverso AAAACCGTGGCACTTTCTTG		-1996
P2	Direto GATGGTACCTGCCCTGAAACTCGACTCAC	723 pb	-2120
	Reverso TTGTAGCAGCCCTGTGATGG		-1417
P3	Direto GACTGCACTCGCCTGTTGAT	786pb	-1302
	Reverso GTGATGGTTTCCTACTGTGCGCT		-538
P4	Direto GGCGAGATTATAATGCCACGGGA	736 pb	-736
	Reverso GATGCGGCCGCGCCACGGGGGAAGGCAATAT		-20
P5	Direto GGCATCGTCAACTATGCACTG	631 pb	-1742
	Reverso AGAAGTTTTCCGCGGCTATT		-1131

* As sequências dos *primers* marcadas em vermelho são sítios de restrição das enzimas *KpnI* (P1 e P2) e *NotI* (P4) que foram incluídas para posterior clonagem no vetor pENTR11 e estudo funcional das regiões do promotor com os SNPs.

** Localização do *primer* na região promotora em relação ao códon ATG predito por Maron *et al.* (2013).

Os amplicons foram sequenciados utilizando o Kit Big Dye Terminator v3.1 (Applied Biosystems) no sequenciador de DNA ABI PRISM 3500 (Applied Biosystems). A edição e o alinhamento das sequências foram realizados com auxílio do software Sequencher 4.1.4 (Gene Codes Corporation, Ann Arbor, MI).

Após o sequenciamento, foi feita uma análise *in silico* da região promotora, buscando sítios de ligação dos fatores de transcrição usando o programa MatInspector (<http://www.genomatix.de/matinspector.html>).

4.5 Amplificação por RACE e sequenciamento da 5' UTR

Para determinar a região 5' não traduzida (5' UTR – *Untranslated Region*) do mRNA do gene *ZmMATE1*, foi realizada uma amplificação das extremidades 5' do cDNA (PCR 5' RACE – *Rapid amplification of cDNA ends*) utilizando as amostras de RNA das linhagens Cateto Al237, L228-3 e L53 após 12 horas de estresse com Al³⁺. As reações foram feitas com o kit FirstChoice® RLM-RACE (Thermo Fisher Scientific) de acordo com as recomendações do fabricante para reações padrão. A etapa de “nested PCR” foi modificada pela utilização de 2,0 µL da reação de transcrição reversa, 12,5 µL de Master Mix GoTaq (Promega), 2,0 µL do primer específico, 2,0 µL do primer RACE 5' e 6,5 µL de água livre de nuclease. As condições de amplificação foram 94°C, 3 min; 35 ciclos de 94°C 30 seg, 55°C 30 seg, 72°C 1 min; 72°C, 7 min. Os *primers* 5' RACE específicos (reversos) interno (5'-GATCCTCATGATCTCCTGACC-3') e externo (5'-CGGCAATTGATACACCTACAGC-3') foram desenhados com a ferramenta Primer-Blast (<http://www.ncbi.nlm.nih.gov/>). A sequência molde para o desenho dos *primers* RACE-específicos utilizada foi a sequência codificante do *ZmMATE1* da linhagem Cateto Al237 disponível no GenBank com número de acesso FJ015156.1 (Maron *et al.*, 2010).

A análise dos fragmentos foi feita por eletroforese em gel de agarose 1,5% com tampão TAE (Tris 40 mM, ácido acético 20 mM e EDTA 1 mM), corados com GelRed Nucleic Acid Stain na diluição de 1:100 (Biotium). Os produtos da RACE foram purificados com Wizard® SV Gel and PCR Clean-Up System (Promega), ligados no sistema de vetor de clonagem pGEM®-T Easy (Promega) e inseridos em células de *Escherichia coli* por choque térmico (Sambrook & Russel, 2001). Para confirmação das colônias positivas foi feita PCR de colônia. A extração de DNA plasmidial dos transformantes foi realizada com Wizard®Plus SV Minipreps DNA Purification System (Promega) de acordo com as orientações do fabricante. A confirmação da clonagem foi feita por sequenciamento conforme descrito no item 4.4.

4.6 Busca por genes em milho que codificam fatores de transcrição que controlam a expressão de genes associados com tolerância ao Al

Estudos anteriores revelaram que *STOP1* (Iuchi *et al.*, 2007) e *ART1* (Yamaji *et al.*, 2009) regulam a expressão de genes envolvidos com a tolerância ao Al em Arabidopsis e arroz, respectivamente. Além desses, um estudo em andamento na Embrapa Milho e Sorgo identificou dois fatores de transcrição que regulam a expressão do gene *SbMATE* de sorgo (dados não publicados). Um dos fatores de transcrição é uma proteína *zinc finger* do tipo DHHC que aqui será referida como SbZNF1, enquanto que o outro é um WRKY, denominado aqui de SbWRKY1.

Esses fatores de transcrição foram selecionados para a busca por candidatos com elevada similaridade de sequência no genoma do milho. As buscas foram realizadas por meio

da ferramenta BLAST (Altschul *et al.*, 1997) utilizando sequências de aminoácidos das proteínas codificadas pelos genes candidatos contra as sequências proteicas preditas no genoma da linhagem de milho B73 disponíveis no Phytozome v11.0 (<https://phytozome.jgi.doe.gov/pz/portal.html>) e no MaizeGDB (<http://www.maizegdb.org/>).

Adicionalmente, foram utilizados genes diferencialmente expressos entre as linhagens de milho C100-6 e L53 sob estresse de AI obtidos pela técnica de microarranjo (Maron *et al.*, 2008), utilizando uma anotação mais recente das sondas impressas no MaizeArray (Seifert *et al.*, 2012). Esses dados foram utilizados na busca por genes que codificam fatores de transcrição do tipo DHHC e WRKY.

4.7 Caracterização do perfil de expressão dos genes candidatos que codificam possíveis fatores de transcrição em milho

O experimento para avaliar a expressão dos genes candidatos e o ensaio de expressão por PCR quantitativo (qRT-PCR) foram os mesmos realizados para caracterizar a expressão do *ZmMATE1* descrito no item 4.3.

Primers específicos para expressão desses genes foram desenhados com base na sequência codificante incluindo as regiões 5' e 3' não traduzidas disponível na plataforma Phytozome (<https://phytozome.jgi.doe.gov>), utilizando a ferramenta Primer-Blast (<http://www.ncbi.nlm.nih.gov/>), seguido por pesquisas em bancos de dados do genoma do milho para assegurar a especificidade da amplificação. A eficiência e a especificidade das reações foram determinadas por meio de curva padrão e *melting*, respectivamente. (Tabela 2).

Tabela 2. Sequências dos *primers* desenhados e selecionados para o ensaio de expressão gênica dos genes candidatos codificadores dos fatores de transcrição.

Fatores de Transcrição	Genes candidatos	Primers	
ART1/STOP1	GRMZM2G068710	Direto	5' CCCGTCCTGAAGAAGTCTCGT 3'
		Reverso	5' TCGCATTAGGCACTAGGCTC 3'
	GRMZM2G075956	Direto	5' ACAAGCTTTTTGGGCACGTT 3'
		Reverso	5' CCTGAGGTTGCTCTGATGCT 3'
	GRMZM2G129428	Direto	5' GACCTTTGGGTTGCGTGTCT 3'
		Reverso	5' AAAGGCCAGACCTCAGCAAT 3'
SbZNF1	GRMZM2G100641	Direto	5' GGTGGGAGGCTCATATTCGG 3'
		Reverso	5' GCAACGAATCCGCAGAAGAC 3'
	GRMZM2G163717	Direto	5' CCGCTGCTCATCTGTGTACG 3'
		Reverso	5' GGCGATTAGGGCGAAGGAAA 3'
GRMZM2G179002	Direto	5' GATCAAGCTGGGGTACAGAGA 3'	
	Reverso	5' GCCCCATTGATTCTGCGGTC 3'	
GRMZM2G068657	Direto	5' GGTTTCAGAGAGGAGGCGATG 3'	
	Reverso	5' TGCTATCTTCGGCGGTATGG 3'	
SbWRKY1	GRMZM2G034421	Direto	5' AGCTCATCCCGGAGCCTAA 3'
		Reverso	5' GGAGAGCGACGACTCCATAG 3'

5. RESULTADOS E DISCUSSÃO

5.1 O alelo do gene *ZmMATE1* confere tolerância ao AI

O crescimento relativo da raiz seminal (CRR), avaliado com base em plantas individuais, foi o índice utilizado para avaliar a tolerância ao AI na população F₂ obtida pelo cruzamento entre L228-3 e L53. A distribuição fenotípica do CRR apresentou desvio da normalidade pelo teste de Lilliefors ($D_{calc} = 0,09$; $D_{tab,5\%} = 0,05$), com tendência a concentração de indivíduos sensíveis ao AI, mas apresentando indivíduos mais tolerantes que o parental tolerante (Figura 2). Apesar do desvio da normalidade, os dados fenotípicos não foram transformados para permitir a comparação dos efeitos com dados anteriores.

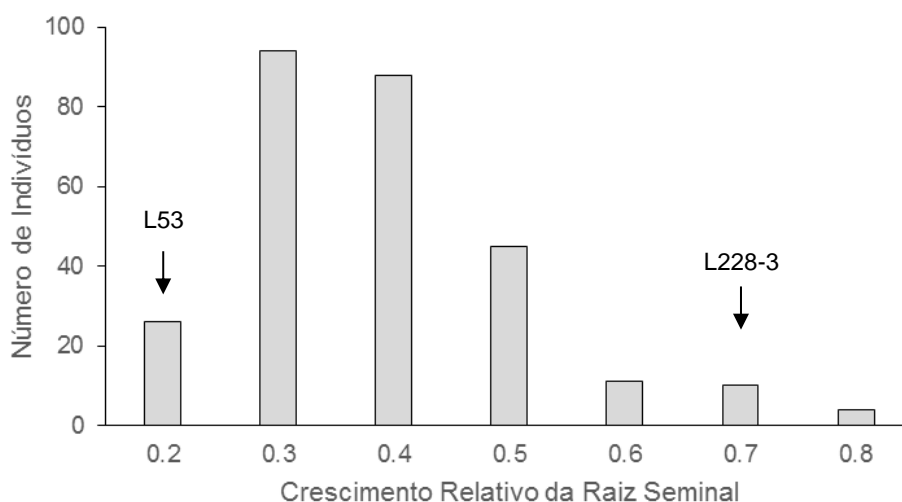


Figura 2. Distribuição do crescimento relativo da raiz seminal utilizado como índice tolerância ao AI na população F_2 do cruzamento entre L228-3 (linhagem tolerante ao AI) e L53 (linhagem sensível).

O marcador apresentou segregação esperada 1:2:1 ($\chi^2 = 3,1075$; $p = 0.2115$) e foi correlacionado significativamente com o índice CRR. O valor do coeficiente de determinação (R^2) foi de 0,2217, indicando que o gene *ZmMATE1* derivado da L228-3 explica 22,17% da tolerância ao AI, com um efeito de substituição alélica de 0.09 no crescimento relativo da raiz seminal (Figura 3). Esses valores são muito próximos ao obtido com o alelo *ZmMATE1* derivado de Cateto AI237 em uma população de linhagens recombinantes, onde o QTL (*qALT6*) co-localizado com o gene *ZmMATE1* explicou 16,2% da tolerância ao AI (Maron *et al.*, 2010) e passou a explicar 30% da tolerância ao AI quando essa população foi saturada com mais de 50.000 SNPs (Guimarães *et al.*, 2014). Nessa população, o QTL controlado pelo *ZmMATE1* teve um efeito de 0,14 no crescimento líquido relativo da raiz com base em média de famílias (Guimarães *et al.*, 2014). Cabe ressaltar que o *qALT6* foi consistentemente identificado como o QTL de grande efeito na tolerância ao AI em milho em outros trabalhos prévios (Sibov *et al.*, 1999; Ninamango-Cárdenas *et al.*, 2003; Conceição *et al.*, 2009). Adicionalmente, o valor de *d/a* foi de 0,1, indicando que o alelo do *ZmMATE1* derivado da L228-3 confere tolerância ao AI de modo aditivo.

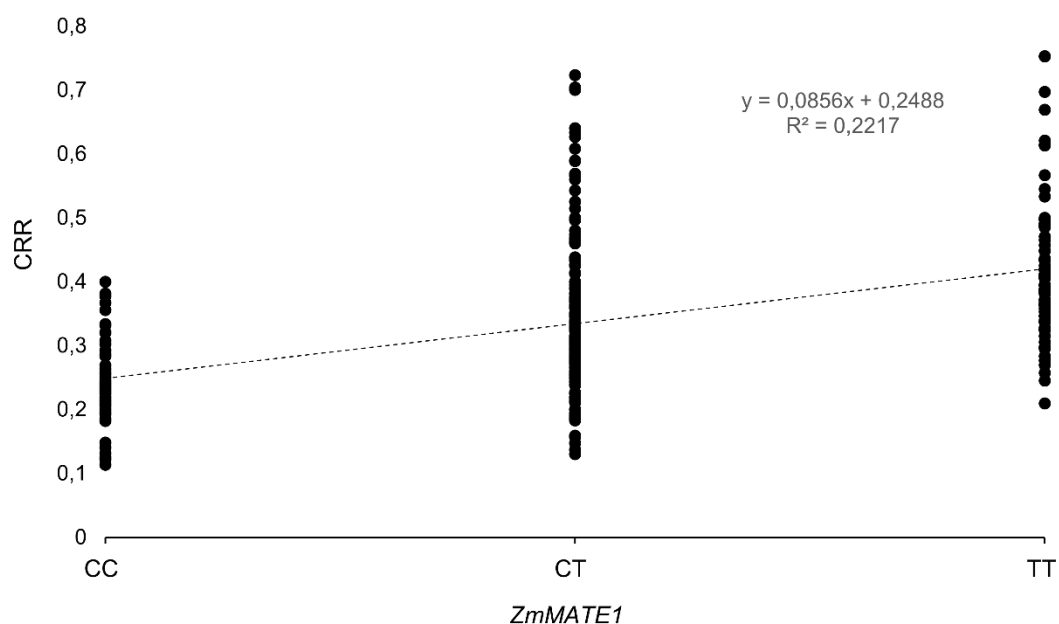


Figura 3. Análise de regressão linear entre o crescimento relativo da raiz seminal (CRR) e o marcador molecular localizado na região promotora do gene *ZmMATE1* em 279 indivíduos F_2 derivados do cruzamento entre L228-3 (linhagem tolerante ao Al) e L53 (linhagem sensível). No eixo X são apresentadas as classes genótípicas do marcador, sendo CC o alelo derivado da L53 em homozigose; TT o alelo derivado da L228-3 em homozigose e CT os heterozigotos.

5.2 Padrão de Expressão Temporal e Tecidual do gene *ZmMATE1*

A expressão do gene *ZmMATE1* foi avaliada no ápice radicular nos tempos de 0, 30 minutos, 1, 6 e 12 horas após os tratamentos com e sem alumínio (Figura 4). A expressão do *ZmMATE1* derivado da L228-3 foi similar ao padrão do alelo derivado da Cateto Al237, previamente descrito por Maron *et al.* (2010). Ambos os alelos foram induzidos pelo Al a partir de 1 hora de tratamento e apresentaram alta expressão 6 e 12 horas após estresse com Al. A expressão do *ZmMATE1* com 12 horas de tratamento com Al nas linhagens tolerantes foi aproximadamente 7 vezes maior em relação a expressão da linhagem sensível L53.

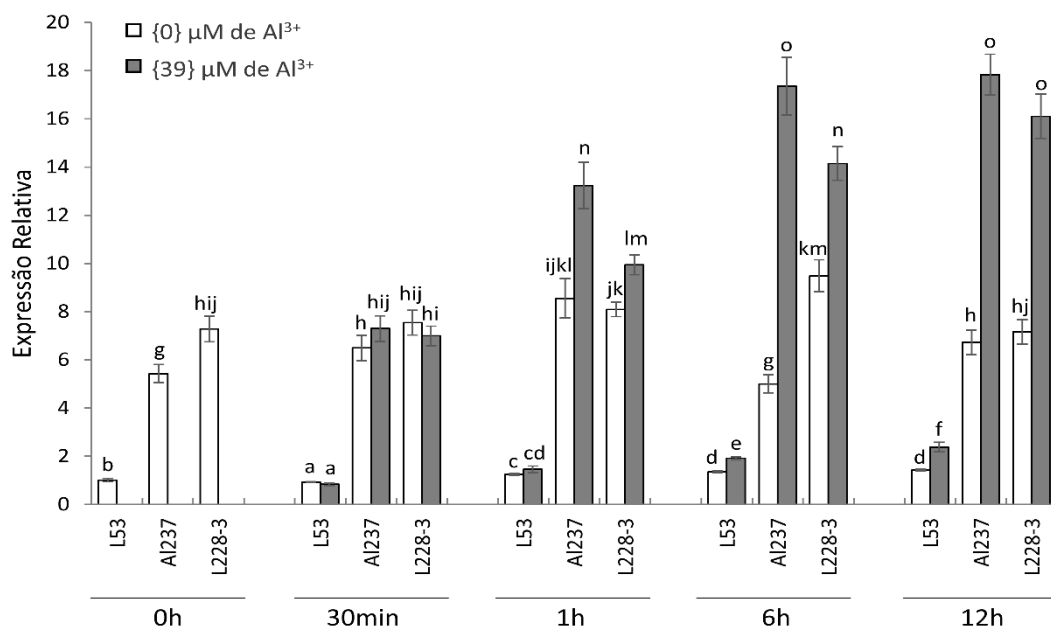


Figura 4. Perfil de expressão temporal do gene *ZmMATE1* no ápice radicular. A expressão gênica foi determinada nos tempos de 0, 30 minutos, 1, 6 e 12 horas após a exposição ao Al³⁺. As barras indicam o desvio padrão da média de três repetições técnicas. Letras distintas indicam médias significativamente diferentes pelo Teste T a 5% de probabilidade.

Considerando que após 6 horas de exposição ao Al o gene *ZmMATE1* de ambas as linhagens tolerantes apresentou altos níveis de indução, a expressão tecidual foi investigada nesse tempo (Figura 5). A expressão do *ZmMATE1* foi detectada preferencialmente nos ápices radiculares nas linhagens tolerantes, em relação ao resto da raiz e da parte aérea, confirmando os resultados obtidos por Maron *et al.* (2010). No entanto, a expressão do alelo da linhagem L228-3 foi significativamente superior no resto da raiz em comparação com o *ZmMATE1* da Cateto Al237 nos tratamentos com e sem Al, sendo que não houve diferença na expressão entre os tratamentos na L228-3. Além disso, a expressão do *ZmMATE1* no resto da raiz em L228-3 foi semelhante à expressão no ápice radicular da Cateto Al237 no tratamento com Al. Essa alta expressão no resto da raiz, principalmente sem Al, pode revelar outra função para o gene a ser investigada. Assim, uma diferença significativa no perfil de expressão do alelo derivado da L228-3 foi uma maior expressão no resto da raiz, além do ápice radicular, em comparação com o perfil de expressão do *ZmMATE1* derivado da Cateto Al237.

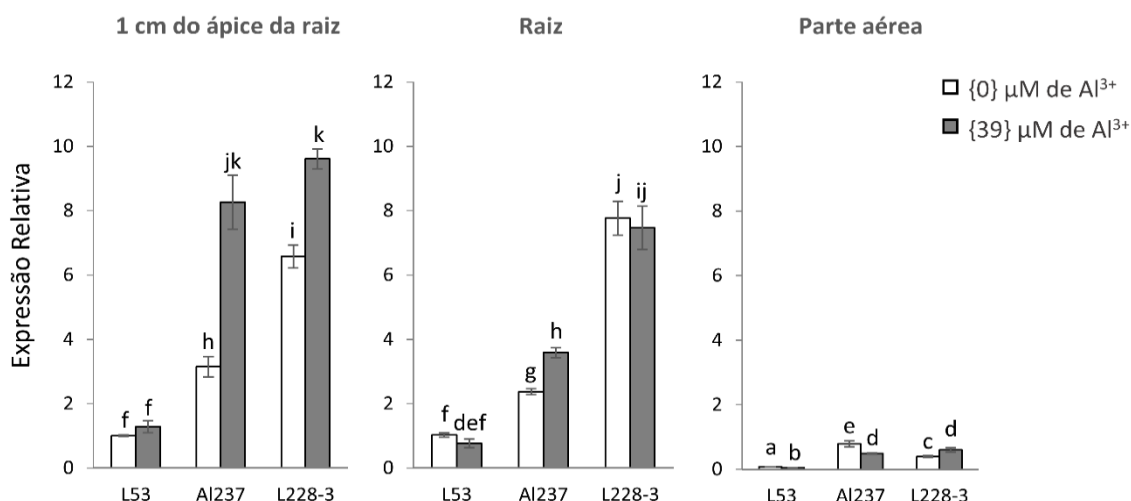


Figura 5. Expressão tecidual do gene *ZmMATE1* após 6 horas de exposição ao Al. A expressão gênica foi determinada no ápice radicular, no resto da raiz e na parte aérea. A amostra calibradora utilizada foi o ápice da linhagem L53 após 6 horas de tratamento sem Al³⁺. As barras indicam o desvio padrão da média. Letras distintas indicam médias significativamente diferentes pelo Teste T a 5% de probabilidade.

5.3 Identificação de SNPs e caracterização *in silico* da região promotora do gene *ZmMATE1*

Com o objetivo de compreender a alta expressão do gene *ZmMATE1* na linhagem L228-3 que possui apenas uma cópia do gene, similar à obtida na Cateto AI237 que possui três cópias em *tandem* do gene, uma região de 2710 pb anterior ao início da tradução do gene foi sequenciada para ser comparada com as sequências das linhagens Cateto AI237 e L53, disponibilizadas por Maron *et al.* (2013) (Figura 6). Quatro polimorfismos do tipo SNP (*Single Nucleotide Polymorphism* – Polimorfismo de Nucleotídeo Único) foram identificados entre as sequências, sendo três monomórficos entre Cateto AI237 e L228-3 e polimórficos em relação à L53, coincidentes com os polimorfismos identificados entre Cateto AI237 e L53 (Maron *et al.* 2013). Apenas um SNP na posição -2222 pb diferenciou L228-3 das linhagens Cateto AI237 e L53.

Posição	-2687	
AI237	AATACGTCCTACCACACCGCACTGTCTTCACTTCCATAGCCAACCTTCTAATATAGCAGTA	60
L228-3	AATACGTCCTACCACACCGCACTGTCTTCACTTCCATAGCCAACCTTCTAATATAGCAGTA	60
L53	AATACGTCCTACCACACCGCACTTTCTTCACTTCCATAGCCAACCTTCTAATATAGCAGTA	60

Posição	-2600	
AI237	GGATCCATCATCGAAACCTAGCAGTCAATTGAAGTTACTTCTAGTATTTCAAGCTAACTCG	120
L228-3	GGATCCATCATCGAAACCTAGCAGTCAATTGAAGTTACTTCTAGTATTTCAAGCTAACTCG	120
L53	GGATCCATCATCGAAACCTAGCAGTCAATTGAAGTTACTTCTAGTATTTCCGCTAACTCG	120

A1237	TCCATTGAACGAGCTACTTGAGTCATCATGAAACCATCTACAACATCCTCATCTGTGTAG	180
L228-3	TCCATTGAACGAGCTACTTGAGTCATCATGAAACCATCTACAACATCCTCATCTGTGTAG	180
L53	TCCATTGAACGAGCTACTTGAGTCATCATGAAACCATCTACAACATCCTCATCTGTGTAG	180

Posição	-2499	
A1237	CAGAAAGAGGTTAGTATCTTGTAGTTTCTTAGTTGTATCACTCATTTGAGACATGGCATAT	240
L228-3	CAGAAAGAGGTTAGTATCTTGTAGTTTCTTAGTTGTATCACTCATTTGAGACATGGCATAT	240
L53	CAGAAAGAGGTTAGTATCTTGTAGTTTCTTAGTTGTATCACTCATTTGAGACATGGCATAT	240

A1237	TTGAGTGCATGTGCAAGTTTTTTGTTTCATCTGCAAAAAATATATAAATAGCATCAGACTA	300
L228-3	TTGAGTGCATGTGCAAGTTTTTTGTTTCATCTGCAAAAAATATATAAATAGCATCAGACTA	300
L53	TTGAGTGCATGTGCAAGTTTTTTGTTTCATCTGCAAAAAATATATAAATAGCATCAGACTA	300

A1237	TGGAATGTGCAAGATTTTTGTTTCAGCTGTAACATGGACAGGCATATGAGGAAACCTTCT	360
L228-3	TGGAATGTGCAAGATTTTTGTTTCAGCTGTAACATGGACAGGCATATGAGGAAACCTTCT	360
L53	TGGAATGTGCAAGATTTTTGTTTCAGCTGTAACATGGACAGGCATATGAGGAAACCTTCT	360

A1237	GCTGCACTGGAAAGACAAACAACCATCTGCGAAATGAACTGAGAGGATCATGTGTACTGA	420
L228-3	GCTGCACTGGAAAGACAAACAACCATCTGCGAAATGAACTGAGAGGATCATGTGTACTGA	420
L53	GCTGCACTGGAAAGACAAACAACCATCTGCGAAATGAACTGAGAGGATCATGTGTACTGA	420

A1237	TAACGGGCTCAAGAACATAACAGCATATGAACAGAACAAGACAGCATATGCCCTCAGCT	480
L228-3	TAACGGGCTCAAGAACATAACAGCATATGAACAGAACAAGACAGCATATGCCCTCAGCT	480
L53	TAACGGGCTCAAGAACATAACAGCATATGAACAGAACAAGACAGCATATGCCCTCAGCT	480

Posição	-2222	
A1237	ATTTTTTTTAAAAAACATCATATGTAGTATACAGTTCAATAATTTCAAAGAAATCATCA	540
L228-3	ATTTTTTTTAAAAAACATCATATGTAGTATACAGTTCAATAATTTCAAAGAAATCATCA	540
L53	ATTTTTTTTAAAAAACATCATATGTAGTATACAGTTCAATAATTTCAAAGAAATCATCA	540

A1237	TAGGGCAAATCAACTATCTCATCTCAATCAATAGTTTTTCAAACAGCCATATGCCCTGAAA	600
L228-3	TAGGGCAAATCAACTATCTCATCTCAATCAATAGTTTTTCAAACAGCCATATGCCCTGAAA	600
L53	TAGGGCAAATCAACTATCTCATCTCAATCAATAGTTTTTCAAACAGCCATATGCCCTGAAA	600

A1237	CTCGACTCACATATGCACTCAAACAGCCATATGCCCTCAAAGCCATATGCCCTCAAACAG	660
L228-3	CTCGACTCACATATGCACTCAAACAGCCATATGCCCTCAAAGCCATATGCCCTCAAACAG	660
L53	CTCGACTCACATATGCACTCAAACAGCCATATGCCCTCAAAGCCATATGCCCTCAAACAG	660

A1237	CCATATGCCACGGATCCAATTGCGACGATCGCGGTGCCAATAATCTGAAGCCCCAAGAA	720
L228-3	CCATATGCCACGGATCCAATTGCGACGATCGCGGTGCCAATAATCTGAAGCCCCAAGAA	720
L53	CCATATGCCACGGATCCAATTGCGACGATCGCGGTGCCAATAATCTGAAGCCCCAAGAA	720

A1237	AGTGCCACGGTTTTAGGGCATATTTTCTGTATGATGCAACAGAAACAGCCTCCAACAGC	780
L228-3	AGTGCCACGGTTTTAGGGCATATTTTCTGTATGATGCAACAGAAACAGCCTCCAACAGC	780
L53	AGTGCCACGGTTTTAGGGCATATTTTCTGTATGATGCAACAGAAACAGCCTCCAACAGC	780

A1237	CATATGCCCTCAGTTATTTTTCAACAAAAAACAAGTGTAGTATACAGTTCACTATTTT	840
L228-3	CATATGCCCTCAGTTATTTTTCAACAAAAAACAAGTGTAGTATACAGTTCACTATTTT	840
L53	CATATGCCCTCAGTTATTTTTCAACAAAAAACAAGTGTAGTATACAGTTCACTATTTT	840

A1237	GAAAGAAAACATCACAGGACAAAGCAACTAGATCATCACACATTATCACAGCCTCATTT	900
L228-3	GAAAGAAAACATCACAGGACAAAGCAACTAGATCATCACACATTATCACAGCCTCATTT	900
L53	GAAAGAAAACATCACAGGACAAAGCAACTAGATCATCACACATTATCACAGCCTCATTT	900

A1237	CAATATTTTCAATGAAACAACAACAACAGCCTAATGCCCTCAAATGAATATAACAGGCA	960
L228-3	CAATATTTTCAATGAAACAACAACAACAGCCTAATGCCCTCAAATGAATATAACAGGCA	960
L53	CAATATTTTCAATGAAACAACAACAACAGCCTAATGCCCTCAAATGAATATAACAGGCA *****	960
A1237	TATGAAATGGCATCGTCAACTATGCACTGTCACCACAAACATCTATATGCCCTCAGTTAT	1020
L228-3	TATGAAATGGCATCGTCAACTATGCACTGTCACCACAAACATCTATATGCCCTCAGTTAT	1020
L53	TATGAAATGGCATCGTCAACTATGCACTGTCACCACAAACATCTATATGCCCTCAGTTAT *****	1020
A1237	TTTTTCAGCAAAAAAACCAAGTGTAGTATACAGTTCACATTTTTGAAAGAAAAACATCATAG	1080
L228-3	TTTTTCAGCAAAAAAACCAAGTGTAGTATACAGTTCACATTTTTGAAAGAAAAACATCATAG	1080
L53	TTTTTCAGCAAAAAAACCAAGTGTAGTATACAGTTCACATTTTTGAAAGAAAAACATCATAG *****	1080
A1237	GGCAAAGCAACTAGATCATCTCACATTATCACAAGCCTCAGTTCAATTTTTTCAATGAAA	1140
L228-3	GGCAAAGCAACTAGATCATCTCACATTATCACAAGCCTCAGTTCAATTTTTTCAATGAAA	1140
L53	GGCAAAGCAACTAGATCATCTCACATTATCACAAGCCTCAGTTCAATTTTTTCAATGAAA *****	1140
A1237	CAACAACAACAGCCTGATGCCCTTAGCTAATTTTACAAAAAGACCTAGTGTAGTCAGT	1200
L228-3	CAACAACAACAGCCTGATGCCCTTAGCTAATTTTACAAAAAGACCTAGTGTAGTCAGT	1200
L53	CAACAACAACAGCCTGATGCCCTTAGCTAATTTTACAAAAAGACCTAGTGTAGTCAGT *****	1200
A1237	TCAATATCTTCAAAGAAAAACATCACAGGGCAAAGCAAATATATCACCTCACATGTATCAC	1260
L228-3	TCAATATCTTCAAAGAAAAACATCACAGGGCAAAGCAAATATATCACCTCACATGTATCAC	1260
L53	TCAATATCTTCAAAGAAAAACATCACAGGGCAAAGCAAATATATCACCTCACATGTATCAC *****	1260
A1237	AAGCCTCAAAACAGCACTCAGGTTTCTACAGTTCATCACAGGGCTGCTACAAAGCACAT	1320
L228-3	AAGCCTCAAAACAGCACTCAGGTTTCTACAGTTCATCACAGGGCTGCTACAAAGCACAT	1320
L53	AAGCCTCAAAACAGCACTCAGGTTTCTACAGTTCATCACAGGGCTGCTACAAAGCACAT *****	1320
A1237	GGCACAGATATCTCAAGCATTGGCTCAGTAAAAATTTTCAAGAAATACAAACCTCATAGTTAT	1380
L228-3	GGCACAGATATCTCAAGCATTGGCTCAGTAAAAATTTTCAAGAAATACAAACCTCATAGTTAT	1380
L53	GGCACAGATATCTCAAGCATTGGCTCAGTAAAAATTTTCAAGAAATACAAACCTCATAGTTAT *****	1380
A1237	TGGCCAATGTCTCCTGAGCACAAACGACGACTGCACTCGCCTGTTGATAAGAACACATAG	1440
L228-3	TGGCCAATGTCTCCTGAGCACAAACGACGACTGCACTCGCCTGTTGATAAGAACACATAG	1440
L53	TGGCCAATGTCTCCTGAGCACAAACGACGACTGCACTCGCCTGTTGATAAGAACACATAG *****	1440
A1237	CACTTTTTTCCGCCTCCACAGCGAATTCAGAACATCAAGCGTAGAAACCATGGTTAC	1500
L228-3	CACTTTTTTCCGCCTCCACAGCGAATTCAGAACATCAAGCGTAGAAACCATGGTTAC	1500
L53	CACTTTTTTCCGCCTCCACAGCGAATTCAGAACATCAAGCGTAGAAACCATGGTTAC *****	1500
A1237	CTACTGATGTTGTGGCACAGAGATCACCAGCAACCACACAGGCGGCAGCAGAAGCACTCA	1560
L228-3	CTACTGATGTTGTGGCACAGAGATCACCAGCAACCACACAGGCGGCAGCAGAAGCACTCA	1560
L53	CTACTGATGTTGTGGCACAGAGATCACCAGCAACCACACAGGCGGCAGCAGAAGCACTCA *****	1560
A1237	ATGTGTCACGGGCTGCATCAATAGCCGCGGAAAACCTTCTGGACACATAGCTCAGCGGCAT	1620
L228-3	ATGTGTCACGGGCTGCATCAATAGCCGCGGAAAACCTTCTGGACACATAGCTCAGCGGCAT	1620
L53	ATGTGTCACGGGCTGCATCAATAGCCGCGGAAAACCTTCTGGACACATAGCTCAGCGGCAT *****	1620
A1237	ATTCATGTTCGGAATCCATATATGCAAACCCTAAGGTTACAGAAGAAGATAGGAGATCTC	1680
L228-3	ATTCATGTTCGGAATCCATATATGCAAACCCTAAGGTTACAGAAGAAGATAGGAGATCTC	1680
L53	ATTCATGTTCGGAATCCATATATGCAAACCCTAAGGTTACAGAAGAAGATAGGAGATCTC *****	1680

A1237	AACAGTGTCTCGAGGGGCGGGAATGGACCACGGGCGCAGGATCTGCCCTGGGCGGAACTAA	1740
L228-3	AACAGTGTCTCGAGGGGCGGGAATGGACCACGGGCGCAGGATCTGCCCTGGGCGGAACTAA	1740
L53	AACAGTGTCTCGAGGGGCGGGAATGGACCACGGGCGCAGGATCTGCCCTGGGCGGAACTAA *****	1740
A1237	ATCAGCAGGGAAGGGCGAGATCCGGTGAAGGACAGGGCAGGCAGGCACCCGAGTGC GGGA	1800
L228-3	ATCAGCAGGGAAGGGCGAGATCCGGTGAAGGACAGGGCAGGCAGGCACCCGAGTGC GGGA	1800
L53	ATCAGCAGGGAAGGGCGAGATCCGGTGAAGGACAGGGCAGGCAGGCACCCGAGTGC GGGA *****	1800
A1237	CGAACAGAGCCGAGGGAGGAAGACGACGACCAGGAAAGAGGGAAGGGAGAATGGAATAC	1860
L228-3	CGAACAGAGCCGAGGGAGGAAGACGACGACCAGGAAAGAGGGAAGGGAGAATGGAATAC	1860
L53	CGAACAGAGCCGAGGGAGGAAGACGACGACCAGGAAAGAGGGAAGGGAGAATGGAATAC *****	1860
A1237	GCTCGAAGGAGACGACGACCAGGGAAGTAATGGTGCGAACGGAGCGAGGGCAGCAGCGGC	1920
L228-3	GCTCGAAGGAGACGACGACCAGGGAAGTAATGGTGCGAACGGAGCGAGGGCAGCAGCGGC	1920
L53	GCTCGAAGGAGACGACGACCAGGGAAGTAATGGTGCGAACGGAGCGAGGGCAGCAGCGGC *****	1920
A1237	ACACGAGAAATGGTAATGACCATTAATCCAGCGTGATGCTAAATCTTCAAACGGGCGAG	1980
L228-3	ACACGAGAAATGGTAATGACCATTAATCCAGCGTGATGCTAAATCTTCAAACGGGCGAG	1980
L53	ACACGAGAAATGGTAATGACCATTAATCCAGCGTGATGCTAAATCTTCAAACGGGCGAG *****	1980
A1237	ATTATAATGCCACGGATTCAATTGCGACGATCGCGGTGCTAATAGTCTGAAGCCCGCCGA	2040
L228-3	ATTATAATGCCACGGATTCAATTGCGACGATCGCGGTGCTAATAGTCTGAAGCCCGCCGA	2040
L53	ATTATAATGCCACGGATTCAATTGCGACGATCGCGGTGCTAATAGTCTGAAGCCCGCCGA *****	2040
A1237	AGAGGGTGCTATGGTTGCAAAATTCTCCACATATACACACAAAACGGGCATATCTAGAGT	2100
L228-3	AGAGGGTGCTATGGTTGCAAAATTCTCCACATATACACACAAAACGGGCATATCTAGAGT	2100
L53	AGAGGGTGCTATGGTTGCAAAATTCTCCACATATACACACAAAACGGGCATATCTAGAGT *****	2100
A1237	CCCAAGCTCGAGCGAGCCTGCCTGCTTGAGCGCAGCCATCCAATCACATGAATATAGAAT	2160
L228-3	CCCAAGCTCGAGCGAGCCTGCCTGCTTGAGCGCAGCCATCCAATCACATGAATATAGAAT	2160
L53	CCCAAGCTCGAGCGAGCCTGCCTGCTTGAGCGCAGCCATCCAATCACATGAATATAGAAT *****	2160
A1237	CGGCCGAGTCGCAGCGACAGTAGGAAACCATCACTTTCTATATACTAACTATATGGACA	2220
L228-3	CGGCCGAGTCGCAGCGACAGTAGGAAACCATCACTTTCTATATACTAACTATATGGACA	2220
L53	CGGCCGAGTCGCAGCGACAGTAGGAAACCATCACTTTCTATATACTAACTATATGGACA *****	2220
A1237	TTTTTATTTTTCTCGGAATGTAATCTTTATATTCAGATTATTTTCACAAC TAAATTCTCG	2280
L228-3	TTTTTATTTTTCTCGGAATGTAATCTTTATATTCAGATTATTTTCACAAC TAAATTCTCG	2280
L53	TTTTTATTTTTCTCGGAATGTAATCTTTATATTCAGATTATTTTCACAAC TAAATTCTCG *****	2280
A1237	TAACAGCTAGATTCTCAGCCACTCAGAAAAAATTGAACCAAATAGCCTACCGACGACAGC	2340
L228-3	TAACAGCTAGATTCTCAGCCACTCAGAAAAAATTGAACCAAATAGCCTACCGACGACAGC	2340
L53	TAACAGCTAGATTCTCAGCCACTCAGAAAAAATTGAACCAAATAGCCTACCGACGACAGC *****	2340
A1237	GTAAGACTGGTAAAAGTTTACGTGGTGGATACACGAGGGACAAAAAAAACGTTGGAGG	2400
L228-3	GTAAGACTGGTAAAAGTTTACGTGGTGGATACACGAGGGACAAAAAAAACGTTGGAGG	2400
L53	GTAAGACTGGTAAAAGTTTACGTGGTGGATACACGAGGGACAAAAAAAACGTTGGAGG *****	2400
A1237	TAAAATGGACTCGGGGAGCCGGAGGGAAGCGGCGACCCGCGGAAGGATGCATCGCCACGG	2460
L228-3	TAAAATGGACTCGGGGAGCCGGAGGGAAGCGGCGACCCGCGGAAGGATGCATCGCCACGG	2460
L53	TAAAATGGACTCGGGGAGCCGGAGGGAAGCGGCGACCCGCGGAAGGATGCATCGCCACGG *****	2460

A1237	CGGTGGCGGGCGTAAGCGTAAGCGACGCGCACTCACGCCGACGCCCGTCTCCTTCGGCAT	2520
L228-3	CGGTGGCGGGCGTAAGCGTAAGCGACGCGCACTCACGCCGACGCCCGTCTCCTTCGGCAT	2520
L53	CGGTGGCGGGCGTAAGCGTAAGCGACGCGCACTCACGCCGACGCCCGTCTCCTTCGGCAT	2520

A1237	GGAGGACCTCTCCGCTTTGACTCTGTTTACCGCGCACAGATACCATTGCGGAAGGGACCC	2580
L228-3	GGAGGACCTCTCCGCTTTGACTCTGTTTACCGCGCACAGATACCATTGCGGAAGGGACCC	2580
L53	GGAGGACCTCTCCGCTTTGACTCTGTTTACCGCGCACAGATACCATTGCGGAAGGGACCC	2580

A1237	AGACGCAGCCTACACCTGCACGGCTGCACCGCAGGGAGGCGCCATCTGACCCGGCCCCGG	2640
L228-3	AGACGCAGCCTACACCTGCACGGCTGCACCGCAGGGAGGCGCCATCTGACCCGGCCCCGG	2640
L53	AGACGCAGCCTACACCTGCACGGCTGCACCGCAGGGAGGCGCCATCTGACCCGGCCCCGG	2640

A1237	CGGCGCTGGCTGCCCGTGCAGCCAGACCCACCCAGAGGAATCACGCAGATATTGCCTT	2700
L228-3	CGGCGCTGGCTGCCCGTGCAGCCAGACCCACCCAGAGGAATCACGCAGATATTGCCTT	2700
L53	CGGCGCTGGCTGCCCGTGCAGCCAGACCCACCCAGAGGAATCACGCAGATATTGCCTT	2700

A1237	CCCCCGTGGC ATG	2710
L228-3	CCCCCGTGGC ATG	2710
L53	CCCCCGTGGC ATG	2710

Figura 6. Alinhamento múltiplo das sequências promotoras das linhagens Cateto A1237, L228-3 e L53 utilizando a ferramenta Clustal Omega. Substituições dos nucleotídeos entre as sequências são destacadas em cinza. Os números acima dos SNPs indicam a posição em pares de bases em relação ao ATG predito por Maron *et al.* (2010).

As plantas exibem uma variedade de respostas a estresses abióticos que permitem a sobrevivência em condições adversas. Os fatores de transcrição (TFs) possuem um papel importante nas vias de sinalização que controlam essa tolerância, sendo um dos responsáveis pelas respostas aos estresses (Knight & Knight, 2001; Mahajan & Tuteja, 2005). Fatores de transcrição são proteínas que se ligam aos elementos *cis*-regulatórios presentes nas sequências de promotores, acentuadores ou repressores, que controlam a transcrição de diferentes genes. A regulação da expressão gênica envolve interações entre ativadores e repressores e a ligação ao DNA dessas proteínas de forma específica (Lee *et al.*, 2000). As sequências da região promotora do *ZmMATE1* das três linhagens foram avaliadas *in silico* visando a identificação dos potenciais sítios de ligação de fatores de transcrição. O programa MatInspector identificou 85 tipos de elementos *cis* nas regiões promotoras de Cateto A1237 e L228-3, correspondendo a 347 sítios de ligação. Devido aos SNPs, a região promotora da L53 apresentou, além dos 85 elementos *cis* preditos nos genótipos tolerantes, um sítio de ligação VRE (*VIP1 response element*) na posição de -2496 a -2499 pb e um sítio para MYB na posição -2499 pb. Assim, a região promotora da L53 apresentou 349 sítios de ligação para fatores de transcrição.

De forma geral, foram identificados sítios de ligação para fatores de transcrição DREB que respondem a desidratação e a baixas temperaturas (Agarwal *et al.*, 2006), para fatores

MADS-box que respondem a seca e a alta salinidade em tomate (Yin *et al.*, 2017) e para vários outros tipos de estresse em arroz (Lee *et al.*, 2008), para proteínas MYC que regulam a expressão de genes induzíveis por desidratação e pelo ácido abscísico (Abe *et al.*, 1997) e sítios W-box de ligação dos fatores de transcrição da família WRKY.

A sequência do elemento *cis* de ligação da proteína ART1 foi identificada por Tsutsui *et al.* (2011) como sendo GGN(T/g/a/C)V(C/A/g)S(C/G). Com base nessa sequência e considerando apenas os nucleotídeos com maior afinidade de ligação, representados pelas letras maiúsculas, foram encontrados 12 elementos *cis* para esse fator de transcrição na região promotora das três linhagens avaliadas (Figura 7). O ART1 é um fator de transcrição do tipo *zinc finger* que regula a expressão de pelo menos 31 genes envolvidos na tolerância ao Al em arroz (Yamaji *et al.*, 2009), incluindo um homólogo do *ZmMATE1*.

Dentre os polimorfismos identificados na região promotora, o SNP localizado na posição -2499 pb merece destaque, uma vez que coincide com sítios de ligação para TFs. No SNP T/C, o T está presente nas linhagens tolerantes com alto nível de expressão do *ZmMATE1* e o C na linhagem sensível com baixa expressão do gene (Barros *et al.*, 2016). Esse SNP flanqueia a região de ligação de fatores de transcrição da família NAC [No apical meristem (NAM), Arabidopsis transcription activation factor (ATAF), Cup-shaped cotyledon (CUC)], uma das maiores famílias de TFs específicos de plantas (Figura 7A). Os fatores NAC possuem funções importantes tanto no desenvolvimento vegetal, quanto em respostas a estresses abióticos como seca, alta salinidade e baixas temperaturas (Tran *et al.*, 2010; Nakashima *et al.*, 2011). O alelo C desse SNP está presente na linhagem sensível L53, coincidindo com uma região conservada para ligação dos fatores MYB (Figura 7B). Fatores de transcrição da família MYB estão envolvidos na regulação do desenvolvimento da planta, do metabolismo e em respostas a estresses bióticos e abióticos, podendo atuar como ativadores ou repressores da transcrição (Lee *et al.*, 2007; Fornalé *et al.*, 2010; Ambawat *et al.*, 2013; Huang *et al.*, 2015; Nagarajan *et al.*, 2016). Dentro desse SNP, porém na fita negativa, foi predito um sítio de ligação VRE, onde se ligam proteínas VIP1 envolvidas na infecção por *Agrobacterium* (Tzfira *et al.*, 2001) e em estresses abióticos (Tsugama *et al.*, 2012).



Figura 7. Visão gráfica da distribuição de alguns elementos *cis*-regulatórios, descritos na legenda, ao longo de 2710 pb da região promotora do gene *ZmMATE1* da linhagem L53. Na posição -2499, indicada pelo asterisco, está representado o SNP T/C, que corresponde ao código Y. **A** representa a sequência de ligação para fatores NAC flanqueando esse SNP. **B** representa o motivo de ligação de proteínas MYB no alelo C do genótipo L53. Letras maiúsculas dentro do motivo de ligação representam a sequência *core*, definida como a mais conservada. Letras em vermelho são as posições conservadas e de alto conteúdo informativo na região.

Outro elemento *cis*-regulatório do tipo MADS-box, identificado no SNP da posição -2222 pb, que diferencia a linhagem L228-3 das demais (Figura 6), está presente na fita negativa do promotor. Como o SNP é A/T e a região é rica em A e T, o sítio de ligação desse elemento *cis* está presente nas três linhagens analisadas.

O SNP mais significativo de um estudo de associação genômica ampla (do inglês, *Genome-Wide Association Study* - GWAS) para tolerância à seca em milho foi localizado no gene *ZmVPP1* que codifica um H⁺ pirofosfatase do tipo vacuolar. Uma inserção de 366 pb no promotor desse gene contendo três elementos *cis* da família de fatores de transcrição MYB é responsável pela expressão do *ZmVPP1* induzida pela seca nos genótipos tolerantes. A seleção genética do alelo do *ZmVPP1*, assim como milhos transgênicos super-expressando o gene, apresentaram maior tolerância ao déficit hídrico (Wang *et al.*, 2016). Desse modo, a localização de elementos *cis*-regulatórios relacionados a estresses ambientais próximos aos SNPs identificados pelo sequenciamento do promotor do *ZmMATE1* entre linhagens tolerantes e sensíveis ao AI tóxico pode sugerir uma influência da atividade dessas proteínas na regulação gênica que podem ser alvos para futuros estudos.

5.4 Variações no tamanho e na sequência da região 5' UTR do gene *ZmMATE1*

A sequência completa da região 5' não traduzida (5' UTR – *Untranslated Region*) do gene *ZmMATE1* foi amplificada pela técnica de PCR 5'RACE, clonada e sequenciada entre as linhagens Cateto AI237, L228-3 e L53. Fragmentos 5'UTR de diferentes tamanhos foram obtidos para todas as três linhagens. Para a linhagem Cateto AI237, foram clonados três fragmentos de 210 pb, um de 131 pb e quatro de 95 pb, para a L228-3 foram obtidos um fragmento de 210 e um de 113 pb e quatro de 95 pb. Para a L53, foi clonado um fragmento de 206 pb e oito de 95 pb. O tamanho dos fragmentos foi referente à amplificação da primeira base após o adaptador do kit 5'RACE até o *primer* específico interno considerando esse *primer* (Figuras 8 e 9).

As sequências 5'UTRs foram alinhadas com sequências do cDNA do *ZmMATE1* das linhagens Cateto AI237 (*MATE1-AI237*) e L53 (*MATE1-L53*) (Maron *et al.*, 2010), e da linhagem PH207 (Hirsch *et al.*, 2016), recentemente sequenciada (*MATE1-PH207*) que está disponível no banco de dados Phytozome, identificado como Zm00008a023530 (<https://phytozome.jgi.doe.gov/>) (Figura 8). As sequências obtidas por Maron *et al.* (2010) foram idênticas entre as linhagens L53 e Cateto AI237 (*MATE1-AI237* e *MATE1-L53*), mas foram polimórficas em relação à sequência da linhagem PH207 (*MATE1-PH207*). As 5'UTRs identificadas nesse trabalho apresentaram polimorfismos de tamanho e de nucleotídeo em relação às sequências públicas, sendo divididas em três grupos quanto ao tamanho. As linhagens Cateto AI237 e L228-3 tiveram fragmentos representantes nos três grupos, enquanto a L53 apresentou fragmentos em apenas dois grupos. Para os fragmentos maiores (Grupo 3), as linhagens Cateto AI237 e L228-3 apresentaram sequências idênticas entre si de 210 pb, com quatro polimorfismos na porção 5' em relação às sequências públicas, e o fragmento de 206 pb da L53 apresentou quatro bases a menos na porção 5'. Os fragmentos de tamanho intermediário (Grupo 2) foram obtidos apenas para as linhagens tolerantes, sendo que o fragmento de 113 pb da L228-3 apresentou 19 polimorfismos na porção inicial, enquanto o fragmento de 131 pb da Cateto AI237 apresentou sequência idêntica às sequências públicas. As 5'UTRs de menor tamanho (Grupo 1: 95 pb) foram idênticas entre as três linhagens, apresentando polimorfismos em 12 nucleotídeos no início das sequências quando comparadas com as sequências públicas.

Todas as sequências das 5'UTRs identificadas nesse trabalho foram menores que as sequências obtidas para as mesmas linhagens Cateto AI237 e L53 por Maron *et al.* (2010), que foram geradas utilizando RNA mensageiro não tratado com AI. De forma contrária, o RNA mensageiro de todas as linhagens utilizado nesse trabalho para amplificação da 5'UTR foi obtido após o tratamento de 12 horas com AI. Assim o sítio predito para o início da tradução (ATG) dessas novas 5'UTRs será posterior ao predito previamente para o gene *ZmMATE1* caracterizado funcionalmente por Maron *et al.* (2010). Cabe ressaltar que, o sítio de início da

tradução dos fragmentos maiores (Grupo 3: 210 e 206 pb) foi o mesmo predito para o gene Zm00008a023530 na linhagem PH207. Por outro lado, a classe de fragmentos intermediários (Grupo 2: 131 e 113 pb), além de ter sido identificada apenas nas linhagens que possuem alta expressão do *ZmMATE1*, os polimorfismos na porção 5' da UTR da L228-3 (fragmento de 113 pb) criaram um ATG único entre as demais sequências avaliadas. Todos os fragmentos de 95 pb apresentaram um único ATG, codificando uma proteína de menor tamanho em relação às demais.

MATE1_A1237	AGATACCATTGCGGAAGGGACCCAGACGCAGCCTACACCTGCACGGCTGCACCCGACGGGA	60
MATE1_L53	AGATACCATTGCGGAAGGGACCCAGACGCAGCCTACACCTGCACGGCTGCACCCGACGGGA	60
MATE1_PH207	-----	0
5' UTR_A1237_3	-----	0
5' UTR_L2283_3	-----	0
5' UTR_L53_3	-----	0
5' UTR_A1237_2	-----	0
5' UTR_L2283_2	-----	0
5' UTR_A1237_1	-----	0
5' UTR_L2283_1	-----	0
5' UTR_L53_1	-----	0
MATE1_A1237	GGCGCCATCTGACCCGGCCCCGGCGGCGCTGGCTGCCCGTCGCCCCAGACCCACCCAG	120
MATE1_L53	GGCGCCATCTGACCCGGCCCCGGCGGCGCTGGCTGCCCGTCGCCCCAGACCCACCCAG	120
MATE1_PH207	-----	0
5' UTR_A1237_3	-----	0
5' UTR_L2283_3	-----	0
5' UTR_L53_3	-----	0
5' UTR_A1237_2	-----	0
5' UTR_L2283_2	-----	0
5' UTR_A1237_1	-----	0
5' UTR_L2283_1	-----	0
5' UTR_L53_1	-----	0
MATE1_A1237	AGGAATCACGCAGATATTGCCTTCCCCCGTGGC ATG CACCCACCGACCCACATGCGCGCG	180
MATE1_L53	AGGAATCACGCAGATATTGCCTTCCCCCGTGGC ATG CACCCACCGACCCACATGCGCGCG	180
MATE1_PH207	-----GCGCG	5
5' UTR_A1237_3	-----	0
5' UTR_L2283_3	-----	0
5' UTR_A1237_2	-----	0
5' UTR_L2283_2	-----	0
5' UTR_L53_3	-----	0
5' UTR_A1237_1	-----	0
5' UTR_L2283_1	-----	0
5' UTR_L53_1	-----	0
MATE1_A1237	CGCGCACCCATCGGGAATTCGGACCCGATTCGCTTTCCTATTTTCATTCCACTCCACCCG	240
MATE1_L53	CGCGCACCCATCGGGAATTCGGACCCGATTCGCTTTCCTATTTTCATTCCACTCCACCCG	240
MATE1_PH207	CGCGCACCCATCGGGAATTCGGACCCGATTCGCTTTCCTATTTTCATTCCACTCCACCCG	65
5' UTR_A1237_3	-----ATTTTTCATTCCACTCCACCCG	24
5' UTR_L2283_3	-----ATTTTTCATTCCACTCCACCCG	24
5' UTR_L53_3	-----TTTCATTCCACTCCACCCG	20
5' UTR_A1237_2	-----	0
5' UTR_L2283_2	-----	0
5' UTR_A1237_1	-----	0
5' UTR_L2283_1	-----	0
5' UTR_L53_1	-----	0
MATE1_A1237	CTTTTGGGTGTGTGGACGATGCATGGGGGGAGAGAGAGGGAGAGGGAGAGTCTCCT---C	297
MATE1_L53	CTTTTGGGTGTGTGGACGATGCATGGGGGGAGAGAGAGGGAGAGGGAGAGTCTCCT---C	297
MATE1_PH207	CTTTTGGGTGTGTGGAC ATG CATGGGGGGAGAGAGAGGGAGAGGGAGAGTCTCCT CCTC	125
5' UTR_A1237_3	CTTTTGGGTGTGTGGAC ATG CATGGGGGGAGAGAGAGGGAGAGGGAGAGTCTCCT---C	81
5' UTR_L2283_3	CTTTTGGGTGTGTGGAC ATG CATGGGGGGAGAGAGAGGGAGAGGGAGAGTCTCCT---C	81
5' UTR_L53_3	CTTTTGGGTGTGTGGAC ATG CATGGGGGGAGAGAGAGGGAGAGGGAGAGTCTCCT---C	77
5' UTR_A1237_2	-----T---C	2
5' UTR_L2283_2	-----	0

5' UTR_A1237_1	-----	0
5' UTR_L2283_1	-----	0
5' UTR_L53_1	-----	0
MATE1_A1237	TTCCCTCCCCGACCCGAGCGCCGCGCCATGGAAGGGGGCGGCGAGCACCACCACCCGCTT	357
MATE1_L53	TTCCCTCCCCGACCCGAGCGCCGCGCCATGGAAGGGGGCGGCGAGCACCACCACCCGCTT	357
MATE1_PH207	TTCCCTCCCCGACCCGAGCGCCGCGCCATGGAAGGGGGCGGCGAGCACCACCACCCGCTT	185
5' UTR_A1237_3	TTCCCTCCCCGACCCGAGCGCCGCGCCATGGAAGGGGGCGGCGAGCACCACCACCCGCTT	141
5' UTR_L2283_3	TTCCCTCCCCGACCCGAGCGCCGCGCCATGGAAGGGGGCGGCGAGCACCACCACCCGCTT	141
5' UTR_L53_3	TTCCCTCCCCGACCCGAGCGCCGCGCCATGGAAGGGGGCGGCGAGCACCACCACCCGCTT	137
5' UTR_A1237_2	TTCCCTCCCCGACCCGAGCGCCGCGCCATGGAAGGGGGCGGCGAGCACCACCACCCGCTT	62
5' UTR_L2283_2	-----TATTGTGATATAACATGAGCGATGGCGAGCACCACCACCCGCTT	44
5' UTR_A1237_1	-----CACGGCAATTGATACACCACCACCCGCTT	26
5' UTR_L2283_1	-----CACGGCAATTGATACACCACCACCCGCTT	26
5' UTR_L53_1	-----CACGGCAATTGATACACCACCACCCGCTT	26
MATE1_A1237	AGCGTATTCCTCCGGGACGCGAGGCTCGCATTTCAGATGGGACGAGCTCGGTCAGGAGATC	417
MATE1_L53	AGCGTATTCCTCCGGGACGCGAGGCTCGCATTTCAGATGGGACGAGCTCGGTCAGGAGATC	417
MATE1_PH207	AGCGTATTCCTCCGGGACGCGAGGCTCGCATTTCAGATGGGACGAGCTCGGTCAGGAGATC	245
5' UTR_A1237_3	AGCGTATTCCTCCGGGACGCGAGGCTCGCATTTCAGATGGGACGAGCTCGGTCAGGAGATC	201
5' UTR_L2283_3	AGCGTATTCCTCCGGGACGCGAGGCTCGCATTTCAGATGGGACGAGCTCGGTCAGGAGATC	201
5' UTR_L53_3	AGCGTATTCCTCCGGGACGCGAGGCTCGCATTTCAGATGGGACGAGCTCGGTCAGGAGATC	197
5' UTR_A1237_2	AGCGTATTCCTCCGGGACGCGAGGCTCGCATTTCAGATGGGACGAGCTCGGTCAGGAGATC	122
5' UTR_L2283_2	AGCGTATTCCTCCGGGACGCGAGGCTCGCATTTCAGATGGGACGAGCTCGGTCAGGAGATC	104
5' UTR_A1237_1	AGCGTATTCCTCCGGGACGCGAGGCTCGCATTTCAGATGGGACGAGCTCGGTCAGGAGATC	86
5' UTR_L2283_1	AGCGTATTCCTCCGGGACGCGAGGCTCGCATTTCAGATGGGACGAGCTCGGTCAGGAGATC	86
5' UTR_L53_1	AGCGTATTCCTCCGGGACGCGAGGCTCGCATTTCAGATGGGACGAGCTCGGTCAGGAGATC	86
MATE1_A1237	<u>ATGAGGATC</u>	426
MATE1_L53	<u>ATGAGGATC</u>	426
MATE1_PH207	<u>ATGAGGATC</u>	254
5' UTR_A1237_3	<u>ATGAGGATC</u>	210
5' UTR_L2283_3	<u>ATGAGGATC</u>	210
5' UTR_L53_3	<u>ATGAGGATC</u>	206
5' UTR_A1237_2	<u>ATGAGGATC</u>	131
5' UTR_L2283_2	<u>ATGAGGATC</u>	113
5' UTR_A1237_1	<u>ATGAGGATC</u>	95
5' UTR_L2283_1	<u>ATGAGGATC</u>	95
5' UTR_L53_1	<u>ATGAGGATC</u>	95

Figura 8. Alinhamento múltiplo das sequências 5'UTR de DNA obtidas pela PCR 5'RACE, das sequências de cDNA do *ZmMATE1* das linhagens Cateto A1237 e L53 (números de acesso: FJ015156.1 e FJ015157.1, respectivamente) e da sequência do *ZmMATE1* da linhagem PH207 (Zm00008a023530) disponível no portal Phytozome. O alinhamento foi feito pela ferramenta Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo/>). Substituições nucleotídicas entre as sequências estão marcadas em cinza. O início da tradução está destacado em preto. Sequências sublinhadas indicam o local de anelamento do primer interno específico usado na 5'RACE. As sequências apresentadas foram alinhadas somente até o *primer* interno da 5'RACE.

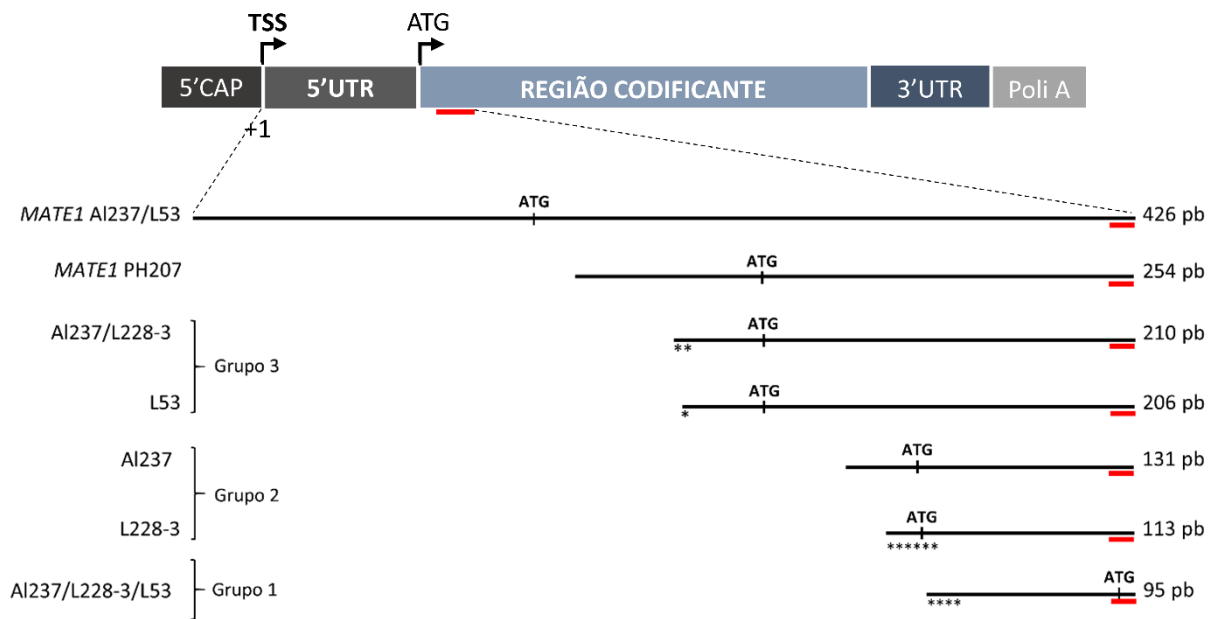


Figura 9. Alinhamento esquemático das sequências das 5'UTRs apresentadas na figura 8. Os asteriscos indicam a região das substituições dos aminoácidos. A reta vermelha representa o *primer* interno específico da 5'RACE e o local de alinhamento na sequência.

Para verificar se os polimorfismos da região 5' não traduzida poderiam alterar o quadro aberto de leitura (ORF – *Open Reading Frame*) do gene, as sequências foram analisadas no programa ORF Finder (<https://www.ncbi.nlm.nih.gov/orffinder/>). Os polimorfismos nas 5'UTRs não alteraram a sequência de aminoácidos nem causaram parada prematura da tradução, sendo coincidente com a ORF da proteína codificada pelo gene *ZmMATE1*, caracterizado por Maron *et al.* (2010). No entanto, os polimorfismos de tamanho das 5'UTRs sequenciadas geraram novos sítios de início da tradução, alterando o início da proteína em termos de número de aminoácidos. Diante disso, a predição estrutural das proteínas foi analisada pelo programa TMHMM Server v. 2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>) e não foram observadas diferenças de topologia entre elas. As novas proteínas preditas, apesar de menores, mantêm os dez domínios transmembrana e uma grande alça intracelular, presentes nas proteínas codificadas a partir das sequências funcionais preditas por Maron *et al.* (2010) (Figura 10). Essa alça é uma característica apresentada pelas proteínas da família *MATE* responsáveis pelo transporte de citrato (Liu *et al.*, 2009).

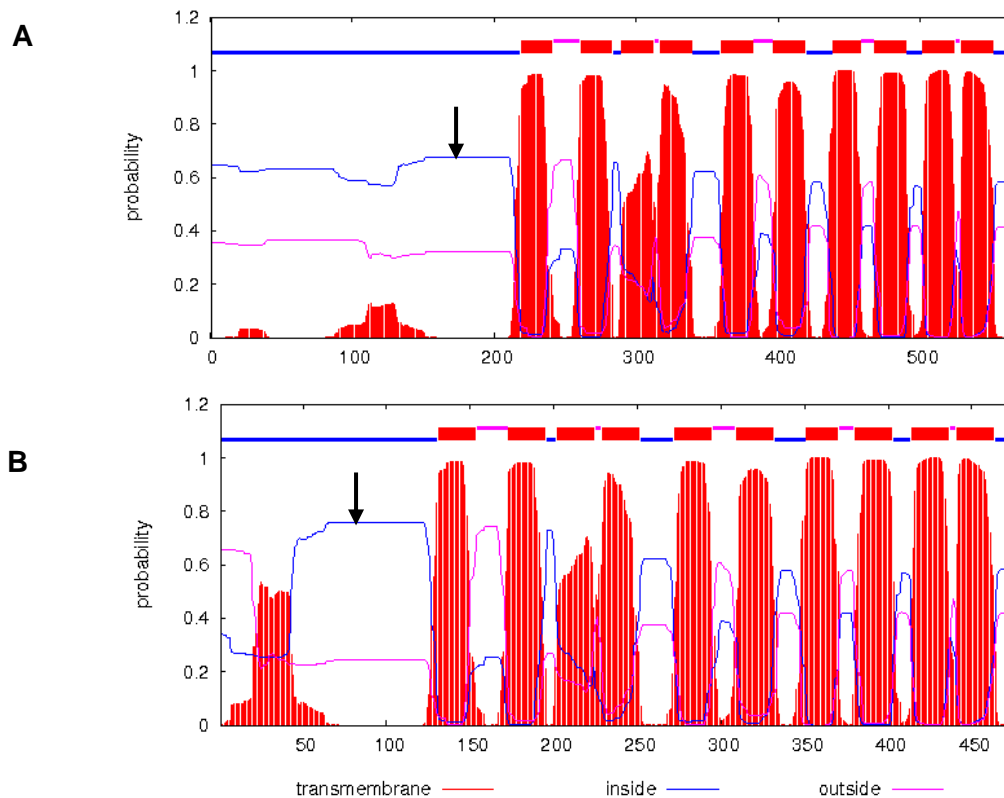


Figura 10. Topologia da proteína ZmMATE1 predita pelo programa TMHMM realizando a predição estrutural dos domínios transmembrana. A) Predição estrutural da proteína com base no cDNA do gene *ZmMATE1* da linhagem Cateto A1237 sequenciado por Maron *et al* (2010). B) Predição estrutural da proteína com base na menor sequência 5'UTR de 85 pb do gene *ZmMATE1*. A seta indica a grande alça intracelular.

A obtenção de um cDNA completo nas extremidades 5' e 3' garante que toda a região transcrita gere informações necessárias sobre o sítio de iniciação da transcrição. As regiões não traduzidas podem influenciar a expressão gênica modulando a estabilidade do mRNA ou a eficiência da tradução. Sendo assim, a síntese de mRNA com distintas UTRs pode definir diferentes propriedades que permitem uma regulação mais refinada da expressão gênica (Hughes, 2006). Um estudo detectou inícios alternativos da transcrição em locos mapeados de *Arabidopsis* e de arroz com tendência a preservar a ORF, e ainda revelou que um número significativo de genes possui *splicing* alternativo dentro da 5'UTR o que sugere o envolvimento em mecanismos regulatórios (Nagasaki *et al.*, 2005).

O promotor do gene *AtALMT1*, que confere tolerância ao Al em *Arabidopsis*, apresenta três sítios de início da transcrição (TSSs), dois dos quais foram associados com possíveis TATA box (Tokizawa *et al.*, 2015). A quantificação de cada transcrito revelou que aqueles TSS de menor tamanho estavam associados com a presença do TATA box tiveram sua proporção relativa aumentada após 6 e 24 horas de tratamento com Al, sendo que o menor transcrito foi o mais abundante. Em contraste, a proporção do transcrito maior, cujo TSS não estava

associado com o TATA box diminuiu em resposta ao tratamento com AI (Tokizawa *et al.*, 2015). Esses resultados sugerem que mudanças no TSS modularam a expressão do *AtALMT1* induzida por AI e corroboram com o fato de que 5'UTRs de diferentes tamanhos estão envolvidos no controle da expressão gênica (Yamamoto *et al.*, 2011). Apesar de ainda pouco conclusivo, dentre os três grupos de fragmentos da 5'UTR identificados no presente trabalho, os TSSs de menor tamanho tiveram uma maior proporção de clones sequenciados em todas as linhagens, sugerindo uma maior proporção em relação aos demais fragmentos.

Como foram identificados três grupos das 5'UTRs de diferentes tamanhos nas linhagens Cateto AI237, L228-3 e L53, existem pelo menos três transcritos menores que aqueles preditos para as linhagens Cateto AI237 e L53 (Maron *et al.*, 2010), e que podem estar envolvidos na regulação da expressão diferencial do *ZmMATE1* entre as linhagens. A região promotora da linhagem L228-3 foi sequenciada com base na sequência disponível nos bancos de dados de Cateto AI237 e L53, que terminou no ATG predito para essas linhagens (Figura 8). Assim, uma pequena região entre esse ATG e o início das 5'UTRs obtidas não foi sequenciada na L228-3. Como alguns polimorfismos de sequência foram identificados nas 5'UTRs clonadas, torna-se importante confirmar tais sequências nas regiões próximas aos TSSs, principalmente porque a L228-3 apresentou vários polimorfismos nos fragmentos 5'UTR. Apesar de não haver domínios TATA na sequência disponível para Cateto AI237 e L53, não podemos afirmar que o mesmo seja verdade para L228-3.

5.5 Genes candidatos que codificam possíveis fatores de transcrição associados com a expressão do *ZmMATE1*

Os genes *STOP1* de *Arabidopsis* (Iuchi *et al.*, 2007) e *ART1* de arroz (Yamaji *et al.*, 2009) foram utilizados nas buscas por candidatos no genoma do milho por controlarem a expressão de genes associados com a tolerância ao AI, incluindo genes da família *MATE* nessas espécies. Dois outros genes foram selecionados com base em dados recentes e ainda não publicados do nosso grupo de pesquisa. Esses genes codificam proteínas com domínios WRKY (*SbWRKY1*) e zinc finger-DHHC (*SbZNF1*), que controlam a expressão do *SbMATE* em sorgo.

Candidatos similares ao STOP1 e ART1

Foram identificados três genes candidatos no genoma do milho codificando proteínas com identidade de sequência acima de 45% em relação ao *STOP1* e *ART1* (Tabela 3). Esse índice de identidade foi escolhido considerando que esses dois fatores compartilham identidade de 41,2% entre as respectivas espécies *Arabidopsis* e arroz. A expressão desses genes candidatos foi avaliada somente no tempo 0 e 30 minutos de exposição ao AI, cujos resultados não apresentaram diferenças consideráveis (dados não apresentados).

O gene *OsART1* possui expressão constitutiva nas raízes de arroz, não sendo afetada pelo alumínio (Yamaji *et al.*, 2009), assim como o gene *AtSTOP1* também não apresenta mudanças significativas na expressão quando avaliado em baixo pH e em diferentes tratamentos com Al (Iuchi *et al.*, 2007). Esses resultados corroboram com os dados do experimento de microarranjos utilizando ápices radiculares de genótipos contrastantes de milho sob estresse de Al (Maron *et al.*, 2008), onde não foram identificados genes que codificam proteínas *zinc finger* do tipo C₂H₂ com padrão de expressão induzido pelo Al no genótipo tolerante. Dentre os genes que codificam proteínas com domínio *zinc finger* do tipo C₂H₂ foram identificados três genes diferencialmente expressos no tempo 0 sem Al e outros dois genes induzidos pelo Al na linhagem sensível (L53), um após 6 horas e outro após 24 horas do tratamento.

Assim, apesar de ser necessário avaliar a expressão desses genes candidatos similares ao *STOP1* e *ART1* em tempos mais longos de tratamento com Al, 6 e 12 horas, esses genes podem não ser diferencialmente expressos, como ocorre com os possíveis homólogos em *Arabidopsis* e arroz.

Tabela 3. Identificação, e-value, identidade e posição no cromossomo dos genes candidatos do genoma do milho similares aos fatores de transcrição ART1, STOP1, SbWRKY e SbZNF1. Os genes GRMZM2G005207, GRMZM2G324999 e GRMZM2G064853 foram incluídos na tabela por serem diferencialmente expressos no experimento de microarranjo conduzido por Maron *et al.* (2008).

Query	Genes Candidatos	e-value	Identidade (%)		Crom	Posição (Mbp)
			ART1	STOP1		
ART1/STOP1 41,2%	GRMZM2G068710_T07	1,1e-118	59,43	47,07	10	10,146
	GRMZM2G075956_T02	2,1e-78	46,96	52,55	8	165,627
	GRMZM2G129428_T01	5,9e-78	45,09	50,75	3	174,504
SbWRKY1	GRMZM2G034421_T01	3,1e-105	80,09		8	118,496
	GRMZM2G005207_T01	-	28,42		10	3,996
	GRMZM2G324999_T01	-	26,09		1	216,633
SbZNF1	GRMZM2G100641_T01	0,0	89,49		6	150,497
	GRMZM2G163717_T02	8,5e-142	55,97		4	195,964
	GRMZM2G179002_T03	6,8e-123	49,27		2	193,069
	GRMZM2G068657_T01	7,8e-114	47,85		1	136,049
	GRMZM2G064853_T01	-	27,10		3	180,550

Candidatos similares ao *SbWRKY1*

Um gene candidato em milho apresentou identidade da sequência proteica de 80% com o fator de transcrição *SbWRKY1* de sorgo. Adicionalmente, dois genes GRMZM2G005207 e GRMZM2G324999 codificando fatores de transcrição da família WRKY foram diferencialmente expressos no ápice radicular sob estresse de Al utilizando microarranjos (Maron *et al.*, 2008) (Tabela 3).

A expressão do gene GRMZM2G034421 foi induzida após 6 horas sem estresse de Al nas linhagens L228-3 (tolerante) e L53 (sensível), sendo reprimida pelo Al nesse tempo em ambas as linhagens (Figura 11). Entre os genes identificados no experimento de microarranjos, o GRMZM2G005207 foi mais expresso na linhagem tolerante C100-6 em comparação com a L53 após 2, 6 e 24 h de tratamento com Al, enquanto o GRMZM2G324999 foi induzido pelo Al na L53 após 2 horas de estresse por Al e em ambas as linhagens após 6 e 24 horas de tratamento com Al em comparação com a condição controle (Maron *et al.*, 2008). O perfil de expressão desses genes candidatos não foi compatível com a expressão do *ZmMATE1* nem com a expressão do *SbWRKY1*, que é um ativador transcricional do *SbMATE* (dados não publicados).

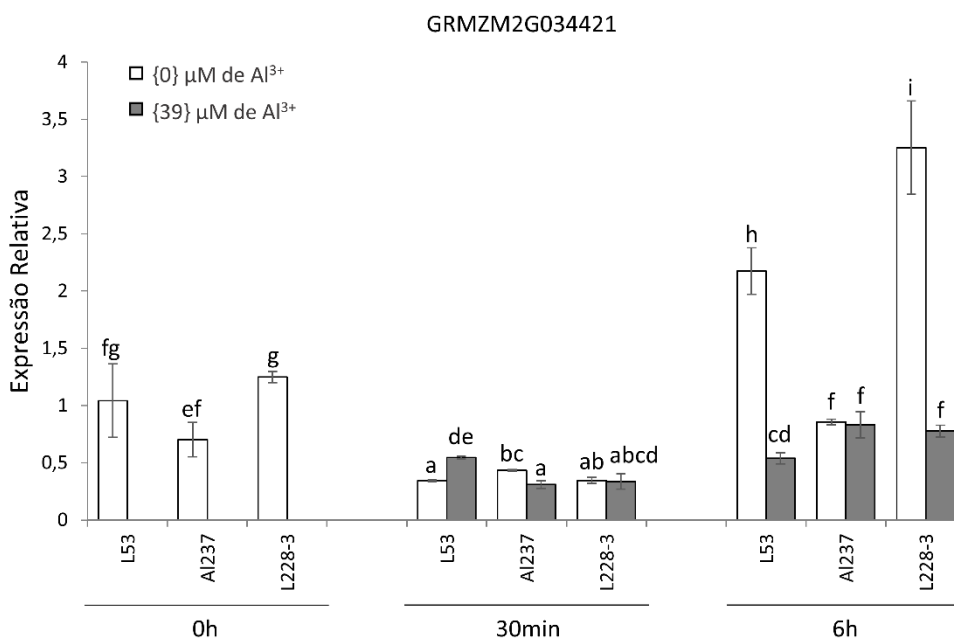


Figura 11. Expressão gênica no ápice radicular do gene candidato GRMZM2G034421, similar ao *SbWRKY1*. A expressão foi determinada nos tempos de 0, 30 minutos e 6 horas após o início do tratamento, com três repetições técnicas. As barras indicam o desvio padrão da média de três repetições técnicas. Letras distintas indicam médias significativamente diferentes pelo Teste T a 5% de probabilidade.

Candidatos similares ao *SbZNF1*

Foram identificados quatro genes candidatos em milho, cujas proteínas preditas apresentaram identidade de sequência superior a 47% com o fator de transcrição *zinc finger-DHHC* de sorgo (*SbZNF1*) (Tabela 3). Dentre eles, os genes candidatos GRMZM2G100641 e GRMZM2G179002 tiveram os perfis de expressão avaliados nos tempos de 0, 30 minutos e 6 horas com e sem Al (Figura 12).

A expressão do gene GRMZM2G100641 foi induzida nas linhagens L228-3 (tolerante) e L53 (sensível) após 6 horas com e sem Al, comparada com a condição controle (tempo 0 e sem Al). No entanto, o tratamento com Al após 6 horas reprimiu a expressão desse gene em ambas as linhagens (Figura 12), que não parece ser compatível com o padrão de expressão do *ZmMATE1* nessas linhagens.

O gene GRMZM2G179002 foi mais expresso no ápice radicular da L53 e menos expresso nas linhagens tolerantes em todos os tempos. No entanto, após 6 horas de estresse com Al ocorreu a máxima indução na L53 e a máxima repressão na L228-3 (Figura 12). Como o gene *ZmMATE1* possui alta indução pelo Al no tempo de 6 horas (Figura 3), podemos considerar que o perfil de expressão do GRMZM2G179002 é compatível com a expressão de um repressor do *ZmMATE1*. Como essa é uma primeira evidência, mais estudos deverão ser realizados, bem como incluir os outros genes candidatos, GRMZM2G163717 e GRMZM2G068657.

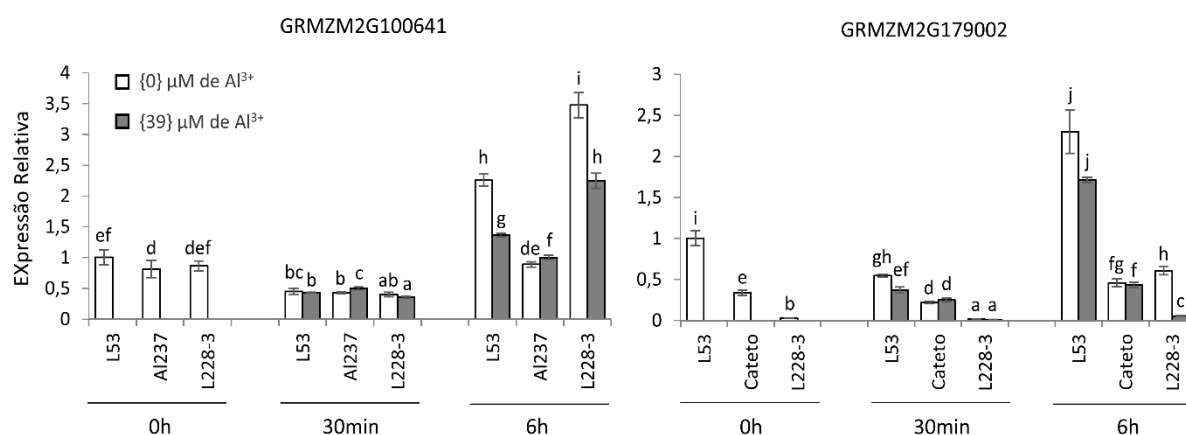


Figura 12. Expressão gênica no ápice radicular dos genes candidatos GRMZM2G100641 e GRMZM2G179002 similares ao *SbZNF1*. A expressão foi determinada nos tempos de 0, 30 minutos e 6 horas após o início do tratamento, com três repetições técnicas. As barras indicam o desvio padrão da média de três repetições técnicas. As barras indicam o desvio padrão da média. Letras distintas indicam médias significativamente diferentes pelo Teste T a 5% de probabilidade.

Além dos genes candidatos selecionados com base em similaridade de sequência, foi identificado o gene GRMZM2G064853, que codifica uma proteína com domínio *zinc finger* do tipo DHHC, diferencialmente expresso sob estresse de AI no estudo com microarranjos (Maron *et al.*, 2008) (Tabela 3). Esse gene candidato também apresentou maior expressão na linhagem L53 quando comparado com a C100-6 após 2, 6 e 24 horas de exposição ao AI. Os experimentos de microarranjos e de PCR quantitativo foram consistentes em mostrar que esses genes codificando proteínas com domínio DHHC foram mais expressos na linhagem sensível em comparação com as linhagens tolerantes ao AI sob estresse de AI.

Um repressor transcricional do gene *ALMT1* em *Arabidopsis* foi caracterizado, sendo ele codificado pelo gene *AtWRKY46* (Ding *et al.*, 2013). O gene *AtWRKY46* foi reprimido pelo AI a partir de 3 horas de estresse, em resposta inversa ao aumento na expressão do *AtALMT1* a partir desse tempo de tratamento até 24 horas, indicando que o padrão de expressão é uma característica a ser considerada na identificação de fatores transcricionais. Assim, apesar da proteína codificada pelo gene candidato GRMZM2G179002 possuir um domínio *zinc finger*-DHHC similar ao *SbZNF1* que atua como um ativador transcricional do *SbMATE*, podemos sugerir que esse fator de transcrição possa atuar como um repressor do *ZmMATE1*, de maneira semelhante ao *AtWRKY46* no gene *AtALMT1*.

6. CONCLUSÕES

O gene *ZmMATE1* derivado da linhagem L228-3 confere tolerância ao AI de modo aditivo e explica 22,17% dessa tolerância em uma população F₂. A expressão do gene *ZmMATE1* da linhagem L228-3 foi induzida a partir de 1 hora após o estresse por AI, sendo semelhante ao padrão do alelo derivado da Cateto AI237. O alelo da L228-3 apresentou uma maior expressão no ápice radicular, além de uma expressão significativamente superior no resto da raiz em comparação com o *ZmMATE1* derivado de Cateto AI237.

Um polimorfismo do tipo SNP na posição -2499 pb diferenciou as linhagens tolerantes (Cateto AI237 e L228-3), que possuem alta expressão do gene, da linhagem sensível ao AI (L53), que possui baixa expressão do gene. Esse polimorfismo coincide com sequências preditas para ligação dos fatores de transcrição MYB e VRE, além de flanquear um possível sítio de ligação de transfatores NAC. Tais características tornam essa região um possível alvo para estudos adicionais sobre fatores envolvidos na regulação da expressão do *ZmMATE1*.

Polimorfismos de sequência e de tamanho na 5'UTR do gene *ZmMATE1* nas linhagens Cateto AI237, L228-3 e L53 indicam a existência de diferentes sítios de início da transcrição, que pode ser um mecanismo de controle da expressão do *ZmMATE1*. Adicionalmente, a expressão de um gene que codifica um possível fator de transcrição com domínio *zinc finger* DHHC é compatível com a função de um possível repressor da expressão do *ZmMATE1*.

Os resultados obtidos nesse trabalho apresentam novos caminhos a serem investigados para o melhor entendimento dos mecanismos moleculares que controlam a expressão do gene *ZmMATE1* derivado da linhagem L228-3. Tais conhecimentos serão úteis para elucidar os mecanismos de tolerância ao Al em milho, com grande aplicabilidade para aumentar a produção agrícola em solos ácidos.

7. REFERÊNCIAS

- Abe, H., Yamaguchi-Shinozaki, K., Urao, T., Iwasaki, T., Hosokawa, D., Shinozaki, K. (1997). Role of Arabidopsis MYC and MYB homologs in drought-and abscisic acid-regulated gene expression. *The Plant Cell*, 9(10), 1859-1868.
- Agarwal, P. K., Agarwal, P., Reddy, M. K., Sopory, S. K. (2006). Role of DREB transcription factors in abiotic and biotic stress tolerance in plants. *Plant cell reports*, 25(12), 1263-1274.
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research*, 25(17), 3389-3402.
- Ambawat, S., Sharma, P., Yadav, N. R., Yadav, R. C. (2013). MYB transcription factor genes as regulators for plant responses: an overview. *Physiology and Molecular Biology of Plants* 19, 307–321.
- Arenhart, R. A., Bai, Y., de Oliveira, L. F. V., Neto, L. B., Schunemann, M., dos Santos Maraschin, F., Wang, Z. Y. (2014). New insights into aluminum tolerance in rice: the ASR5 protein binds the STAR1 promoter and other aluminum-responsive genes. *Molecular plant*, 7(4), 709-721.
- Arenhart, R. A., de, L., César, J., Pedron, M., Carvalho, F. E., da Silveira, J. A. G., Margis, R. (2013). Involvement of ASR genes in aluminium tolerance mechanisms in rice. *Plant, cell & environment*, 36(1), 52-67.
- Arenhart, R. A., Schunemann, M., Bucker Neto, L., Margis, R., Wang, Z. Y., Margis-Pinheiro, M. (2016). Rice ASR1 and ASR5 are complementary transcription factors regulating aluminium responsive genes. *Plant, cell & environment*, 39(3), 645-651.
- Barros, B. A., Mitre, L. K., Pinto, M.O., Magalhaes, J. V., Guimaraes, L. J. M., Guimaraes, C. T. Marcador Alelo-Específico Associado com Níveis de Expressão do Gene *ZmMATE1* em Milho. In: *XXXI Congresso Brasileiro de Milho e Sorgo, 2016, Bento Gonçalves*. XXXI Congresso Brasileiro de Milho e Sorgo, 2016.

- Collins, N. C., Shirley, N. J., Saeed, M., Pallotta, M., & Gustafson, J. P. (2008). An ALMT1 gene cluster controlling aluminum tolerance at the Alt4 locus of rye (*Secale cereale* L.). *Genetics*, 179(1), 669-682.
- CONAB - Companhia Nacional de Abastecimento. 10º Levantamento de Grãos Safra 2016/2017. Julho 2017. [<http://www.conab.gov.br/conteudos.php?a=1253&t=2> acessado em 11/07/2017].
- Conceição, L. D. H. C. S., Tessele, C., Barbosa Neto, J. F. (2009). Diallel analysis and mapping of aluminum tolerance in corn inbred lines. *Maydica*, 54: 55-61.
- Delhaize, E., Craig, S., Beaton, C. D., Bennet, R. J., Jagadish, V. C., Randall, P. J. (1993a). Uptake and distribution of aluminum in root apices. *Plant Physiology* 103:685-693.
- Delhaize, E., Ryan, P. R., Randall, P. J. (1993b). Aluminum tolerance in wheat (*Triticum aestivum* L.) (II. Aluminum-stimulated excretion of malic acid from root apices). *Plant Physiology* 103:695-702.
- Delhaize, E., Hebb, D. M., Ryan, P. R. (2001). Expression of a *Pseudomonas aeruginosa* citrate synthase gene in tobacco is not associated with either enhanced citrate accumulation or efflux. *Plant Physiology*, 125(4), 2059-2067.
- Ding, Z. J., Yan, J. Y., Xu, X. Y., Li, G. X., Zheng, S. J. (2013). WRKY46 functions as a transcriptional repressor of ALMT1, regulating aluminum-induced malate secretion in *Arabidopsis*. *The Plant Journal*, 76(5), 825-835.
- FAO Food outlook: biannual report on global food markets, junho 2016. Disponível em <<http://www.fao.org>> Acesso em: junho de 2017.
- Fornalé, S., Shi, X., Chai, C., Encina, A., Irar, S., Capellades, M., *et al* (2010). ZmMYB31 directly represses maize lignin genes and redirects the phenylpropanoid metabolic flux. *Plant J.*, 64, 633–644.
- Foy, C
. D. (1984). Physiological effects of hydrogen, aluminum and manganese toxicities in acid soil. In: Adams, F. (Ed.) Soil acidity and liming. 2. ed. Madison: Soil Science Society American, 57-97.
- Furukawa, J., Yamaji, N., Wang, H., Mitani, N., Murata, Y., Sato, K., Ma, J. F. (2007). An aluminum-activated citrate transporter in barley. *Plant and Cell Physiology*, 48(8), 1081-1091.
- Goodwin, S. B., Sutter, T. R. (2009). Microarray analysis of *Arabidopsis* genome response to aluminum stress. *Biologia Plantarum*, 53(1), 85-99.
- Gruber, B. D., Ryan, P. R., Richardson, A. E., Tyerman, S. D., *et al*. (2010). HvALMT1 from barley is involved in the transport of organic anions. *J. Exp. Bot.*, 61: 1455-1467.
- Guimarães, C. T.; Simões, C. C.; Pastina, M. M.; Maron, L. G.; Magalhães, J. V.; Vasconcellos, R. C. C.; Guimarães, L. J. M.; Lana, U. G. P.; Tinoco, C. F. S.; Noda, R, W.; Jardim-Belicuas,

- S. N.; Kochian, L. V.; Alves, V. M. C.; Parentoni, S. N. (2014). Genetic dissection of Al tolerance QTLs in the maize genome by high density SNP scan. *BMC genomics*, 15(1), 153.
- Henikoff, S., Greene, E. A., Pietrokovski, S., Bork, P., Attwood, T. K., Hood, L. (1997). Gene families: the taxonomy of protein paralogs and chimeras. *Science*, 278(5338), 609-614.
- Hirsch, C., Hirsch, C. D., Brohammer, A. B., Bowman, M. J., Soifer, I., Barad, O., *et al.* (2016). Draft assembly of elite inbred line PH207 provides insights into genomic and transcriptome diversity in maize. *The Plant Cell Online*, tpc-00353.
- Hoekenga, O. A., Maron, L. G., Piñeros, M. A., Cançado, G. M. A., *et al.* (2006). *AtALMT1*, which encodes a malate transporter, is identified as one of several genes critical for aluminum tolerance in Arabidopsis. *Proc. Natl. Acad. Sci. USA*, 103: 9738-9743.
- Huang C. K., Lo, P. C., Huang, L. F., Wu, S. J., Yeh, C. H., Lu, C. (2015). A single-repeat MYB transcription repressor, MYBH, participates in regulation of leaf senescence in Arabidopsis. *Plant Mol Biol*, 88: 269–86
- 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. *The Plant Journal*, 69(5), 857-867.
- 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*, 21(2), 655-667.
- Hughes, T. A. (2006). Regulation of gene expression by alternative untranslated regions. *Trends in Genetics*, 22(3), 119-122.
- 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. *Proceedings of the National Academy of Sciences*, 104(23), 9900-9905.
- Jiang, Y., & Deyholos, M. K. (2006). Comprehensive transcriptional profiling of NaCl-stressed Arabidopsis roots reveals novel classes of responsive genes. *BMC Plant Biology*, 6(1), 25.
- Jiang, Y., Deyholos, M. K. (2009). Functional characterization of Arabidopsis NaCl-inducible WRKY25 and WRKY33 transcription factors in abiotic stresses. *Plant molecular biology*, 69(1-2), 91-105.
- Kidd, P. S., Llugany, M., Poschenrieder, C. H., Gunse, B., Barcelo, J. (2001). The role of root exudates in aluminium resistance and silicon-induced amelioration of aluminium toxicity in three varieties of maize (*Zea mays* L.). *Journal of Experimental Botany*, 52(359): 1339-1352.
- Knight, H., Knight, M. R. (2001). Abiotic stress signalling pathways: specificity and cross-talk. *Trends in plant science*, 6(6), 262-267.

- Kobayashi, Y., Hoekenga, O. A., Itoh, H., Nakashima, M., Saito, S., Shaff, J. E., Koyama, H. (2007). Characterization of AtALMT1 expression in aluminum-inducible malate release and its role for rhizotoxic stress tolerance in Arabidopsis. *Plant Physiology*, 145(3), 843-852.
- Kochian, L. V. (1995). Cellular mechanisms of aluminum toxicity and resistance in plants. *Annual review of plant biology*, 46(1), 237-260.
- Kochian, L. V., Hoekenga, O. A., Piñeros, M. A. (2004). How do crop plants tolerate acid soils? Mechanisms of aluminum tolerance and phosphorous efficiency. *Annu. Rev. Plant Biol.*, 55, 459-493.
- Kochian, L. V., Piñeros, M. A., Liu, J., Magalhaes, J. V. (2015). Plant adaptation to acid soils: the molecular basis for crop aluminum resistance. *Annual Review of Plant Biology*, 66, 571-598.
- Larsen, P. B., Geisler, M. J., Jones, C. A., 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(3), 353-363.
- Lee, S., Woo, Y. M., Ryu, S. I., Shin, Y. D., Kim, W. T., Park, K. Y., An, G. (2008). Further characterization of a rice AGL12 group MADS-box gene, OsMADS26. *Plant physiology*, 147(1), 156-168.
- Lee, T. G., Jang, C. S., Kim, J. Y., Kim, D. S., Park, J. H., Kim, D. Y., Seo, Y. W. (2007). A Myb transcription factor (*TaMyb1*) from wheat roots is expressed during hypoxia: roles in response to the oxygen concentration in root environment and abiotic stresses. *Physiologia Plantarum*, 129: 375–385
- Lee, T. I., Young, R. A. (2000). Transcription of eukaryotic protein-coding genes. *Annual review of genetics*, 34(1), 77-137.
- Li, S., Fu, Q., Chen, L., Huang, W., Yu, D. (2011). Arabidopsis thaliana WRKY25, WRKY26, and WRKY33 coordinate induction of plant thermotolerance. *Planta*, 233(6), 1237-1252.
- Liang, C., Piñeros, M. A., Tian, J., Yao, Z., Sun, L., Liu, J., Liao, H. (2013). Low pH, aluminum, and phosphorus coordinately regulate malate exudation through GmALMT1 to improve soybean adaptation to acid soils. *Plant Physiology*, 161(3), 1347-1361.
- Ligaba, A., Katsuhara, M., Ryan, P. R., Shibasaki, M., Matsumoto, H. (2006). The BnALMT1 and BnALMT2 genes from rape encode aluminum-activated malate transporters that enhance the aluminum resistance of plant cells. *Plant Physiology*, 142(3), 1294-1303.
- Liu, J., Magalhaes, J. V., Shaff, J., Kochian, L. V. (2009). Aluminum-activated citrate and malate transporters from the MATE and ALMT families function independently to confer Arabidopsis aluminum tolerance. *Plant J.*, 57: 389-399.
- Livak, K. J., Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. *methods*, 25(4), 402-408.

- Ma, J. F., Hiradate, S. (2000). Form of aluminium for uptake and translocation in buckwheat (*Fagopyrum esculentum* Moench). *Planta*, 211(3), 355-360.
- Ma, J. F., Hiradate, S., Nomoto, K., Iwashita, T., Matsumoto, H. (1997a). Internal detoxification mechanism of Al in hydrangea (identification of Al form in the leaves). *Plant Physiology*, 113(4), 1033-1039.
- Ma, J. F., Ryan, P. R., Delhaize, E. (2001). Aluminium tolerance in plants and the complexing role of organic acids. *Trends Plant Sci.*, 6: 273-278.
- Ma, J. F., Zheng, S. J., Matsumoto, H., Hiradate, S. (1997b). Detoxifying aluminium with buckwheat. *Nature*, 390(6660), 569-570.
- Ma, Z., Miyasaka, S. C. (1998). Oxalate exudation by taro in response to Al. *Plant Physiology*, 118(3), 861-865.
- Magalhaes, 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(4), 1905-1914.
- Magalhães, J. V., Liu, J., Guimaraes, C. T., Lana, U. G. P., Alves, V. M. C., Wang, Y., 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 Genet.*, 39:1156-1161.
- Magnacava, R., Gardner, C. O., Clark, R. B. (1987) Evaluation of inbred maize lines for aluminum in nutrient solution. In: H.W. Gabelman & B.C. Loughman (Eds.), *Genetics Aspects of Plant Mineral Nutrition*, pp. 255–265. Martinus Nijhoff Publishers, Dordrecht
- Mahajan, S., Tuteja, N. (2005). Cold, salinity and drought stresses: an overview. *Archives of biochemistry and biophysics*, 444(2), 139-158.
- Maron, L. G., Guimarães, C. T., Kirst, M., Albert, P. S., Birchler, J. A., Bradbury, P. J., ... Magalhaes, J. V. (2013). Aluminum tolerance in maize is associated with higher MATE1 gene copy number. *Proceedings of the National Academy of Sciences*, 110(13), 5241-5246.
- Maron, L. G., Kirst, M., Mao, C., Milner, M. J., Menossi, M., Kochian, L. V. (2008). Transcriptional profiling of aluminum toxicity and tolerance responses in maize roots. *New Phytologist*, 179(1), 116-128.
- Maron, L. G., Piñeros, M. A., Guimarães, C. T., Magalhaes, J. V., Pleiman, J. K., Mao, C., Shaff, J., Belicuas, S. N. J., Kochian, L. V. (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(5), 728-740.
- Metali, F., Salim, K. A., Burslem, D. F. (2012). Evidence of foliar aluminium accumulation in local, regional and global datasets of wild plants. *New Phytologist*, 193(3), 637-649.

- Morita, A., Horie, H., Fujii, Y., Takatsu, S., Watanabe, N., Yagi, A., Yokota, H. (2004). Chemical forms of aluminum in xylem sap of tea plants (*Camellia sinensis* L.). *Phytochemistry*, 65(20), 2775-2780.
- Nagarajan, V. K., Satheesh, V., Poling, M. D., Raghothama, K. G., Jain, A. (2016). Arabidopsis MYB-Related HHO2 Exerts a Regulatory Influence on a Subset of Root Traits and Genes Governing Phosphate Homeostasis, *Plant Cell Physiol.*, 57(6), 1142–1152.
- Nagasaki, H., Arita, M., Nishizawa, T., Suwa, M., Gotoh, O. (2005). Species-specific variation of alternative splicing and transcriptional initiation in six eukaryotes. *Gene*, 364, 53-62.
- Nakashima, K., Takasaki, H., Mizoi, J., Shinozaki, K., Yamaguchi-Shinozaki, K. (2012). NAC transcription factors in plant abiotic stress responses. *Biochim. Biophys. Acta*, 1819: 97–103.
- Ninamango-Cardenas, F., Guimarães, C.T., Martins, P., Parentoni, S. N., Carneiro, N. P., et al. (2003). Mapping QTLs for aluminum tolerance in maize. *Euphytica*, 130: 223–232.
- Pandey, S., Ceballos, H., Magnavaca, R., Bahia Filho, A. F. C., Duque-Vargas, J., & Vinasco, L. E. (1994). Genetics of tolerance to soil acidity in tropical maize. *Crop Science*, 34(6), 1511-1514.
- Piñeros, M. A., Shaff, J. E., Manslank, H.S., Alves, V. M. C., Kochian, L. V. (2005). Aluminum resistance in maize cannot be solely explained by root organic acid exudation. A comparative physiological study. *Plant Physiology*, 137:321-241.
- Riede, C. R., Anderson, J. A. (1996). Linkage of RFLP markers to an aluminum tolerance gene in wheat. *Crop Science*, 36:905–9.
- Rushton, P. J., Somssich, I. E., Ringler, P., Shen, Q. J. (2010). WRKY transcription factors. *Trends in plant science*, 15(5), 247-258.
- Ryan, P. R., Raman, H., Gupta, S., Horst, W. J., Delhaize, E. (2009). A second mechanism for aluminum resistance in wheat relies on the constitutive efflux of citrate from roots. *Plant Physiology*, 149(1), 340-351.
- Ryan, P. R., Tyerman, S. D., Sasaki, T., Furuichi, T., Yamamoto, Y., Zhang, W. H., Delhaize, E. (2011). The identification of aluminium-resistance genes provides opportunities for enhancing crop production on acid soils. *Journal of Experimental Botany*, 62: 9-20.
- Saghai-Marouf, M. A., Soliman, K. M., Jorgensen, R. A., Allard, R. W. (1984). Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. *Proceedings of the National Academy of Sciences*, 81(24), 8014-8018.
- Sambrook, J., Russel, D. (2001). Molecular cloning - a laboratory manual. *New York Cold Spring Harbor*.

- Sasaki, T., 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.
- Seifert, F., Thiemann, A., Pospisil, H., Scholten, S. (2012). Re-annotation of the maize oligonucleotide array. *Maydica*, 57(1), 49-55.
- Sibov, S. T., Gaspar, M. J., Ottoboni, L. M. M., Arruda, P., Souza, A. P. (1999). Two genes controlling aluminum tolerance in maize: genetic and molecular mapping analyses. *Genome*, 42:475-482.
- Silva, I. R., Smyth, T. J., Raper, C. D., Carter, T. E., Ruffy, T. W. (2001). Differential aluminum tolerance in soybean: an evaluation of the role of organic acids. *Physiologia Plantarum*, 112(2), 200-210.
- Tokizawa, M., Kobayashi, Y., Saito, T., Kobayashi, M., Iuchi, S., Nomoto, M., Koyama, H. (2015). Sensitive to proton rhizotoxicity1, calmodulin binding transcription activator2, and other transcription factors are involved in aluminum-activated malate transporter1 expression. *Plant physiology*, 167(3), 991-1003.
- Tran, L. S., Nishiyama, R., Yamaguchi-Shinozaki, K., Shinozaki, K. (2010). Potential utilization of NAC transcription factors to enhance abiotic stress tolerance in plants by biotechnological approach. *GM Crops.*, 1(1), 32–9.
- Tsugama, D., Liu, S., Takano, T. (2012). A bZIP protein, VIP1, is a regulator of osmosensory signaling in Arabidopsis. *Plant physiology*, 159(1), 144-155.
- Tsutsui, T., Yamaji, N., Ma, J. F. (2011). Identification of a cis-acting element of ART1, a C2H2-type zinc-finger transcription factor for aluminum tolerance in rice. *Plant Physiol.*, 156: 925–931.
- Tzfira, T., Vaidya, M., Citovsky, V. (2001). VIP1, an Arabidopsis protein that interacts with Agrobacterium VirE2, is involved in VirE2 nuclear import and Agrobacterium infectivity. *The EMBO Journal*, 20(13), 3596-3607.
- von Uexküll, H. R., Mutert, E. (1995). Global extent, development and economic impact of acid soils. In: Plant-Soil Interactions at Low pH: Principles and Management. RA Date, NJ Grundon, GE Raymet, and ME Probert (Eds.). Kluwer Academic Publishers, Dordrecht, Netherlands, p. 5-19.
- Wallington, T. J., Anderson, J. E., Mueller, S. A., Kolinski Morris, E., Winkler, S. L., Ginder, J. M., Nielsen, O. J. (2012). Corn ethanol production, food exports, and indirect land use change. *Environmental science & technology*, 46(11), 6379-6384.
- Wang, X., Wang, H., Liu, S., Ferjani, A., Li, J., Yan, J., Yang, X., Qin, F. (2016). Genetic variation in ZmVPP1 contributes to drought tolerance in maize seedlings. *Nature genetics*, 48(10), 1233-1241.

- Wu, X., Li, R., Shi, J., Wang, J., Sun, Q., Zhang, H., Xing, Y., Qi, Y., Zhang, N., Guo, Y. D. (2014). Brassica oleracea MATE encodes a citrate transporter and enhances aluminum tolerance in Arabidopsis thaliana. *Plant and Cell Physiology*, 55(8), 1426-1436.
- Xia, J., Yamaji, N., Kasai, T., Ma, J. F. (2010). Plasma membrane-localized transporter for aluminum in rice. *Proc Natl Sci U S A* 107: 18381-18385.
- Yamaji, N., Huang, C. F., Nagao, S., Yano, M., Sato, Y., Nagamura, Y., & Ma, J. F. (2009). A zinc finger transcription factor ART1 regulates multiple genes implicated in aluminum tolerance in rice. *The Plant Cell*, 21(10), 3339-3349.
- Yamamoto, Y. Y., Yoshioka, Y., Hyakumachi, M., Obokata, J. (2011). Characteristics of core promoter types with respect to gene structure and expression in Arabidopsis thaliana. *DNA research*, 18(5), 333-342.
- Yang, Z. M., Sivaguru, M., Horst, W. J., Matsumoto, H. (2000). Aluminium tolerance is achieved by exudation of citric acid from roots of soybean (*Glycine max*). *Physiologia Plantarum*, 110(1), 72-77.
- Yin, W., Hu, Z., Hu, J., Zhu, Z., Yu, X., Cui, B., Chen, G. (2017). Tomato (*Solanum lycopersicum*) MADS-box transcription factor SIMBP8 regulates drought, salt tolerance and stress-related genes. *Plant Growth Regulation*, 1-14.
- 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.
- Zheng, S. J., Ma, J. F., Matsumoto, H. (1998). Continuous secretion of organic acids is related to aluminium resistance during relatively long-term exposure to aluminium stress. *Physiologia Plantarum*, 103(2), 209-214.