

Bárbara Hufnagel Maciel

**Mineração e Caracterização de Alelos do Gene de Tolerância ao
Alumínio *Alt_{SB}* em Sorgo**

Dissertação apresentada à Universidade Federal de Minas Gerais, como parte das exigências do Programa de Pós-Graduação em Genética, para obtenção do título de Mestre em Genética.

Orientador: Jurandir Vieira de Magalhães

Co-Orientadora: Cláudia Teixeira Guimarães

Belo Horizonte
Universidade Federal de Minas Gerais
Instituto de Ciências Biológicas
2009

"A mente que se abre a uma nova idéia jamais voltará ao seu tamanho original."

Albert Einstein

**Aos meus amados pais,
Vânia e Marcos,
dedico.**

Agradecimentos

A Deus por cada dia de vida.

Ao programa de Pós-Graduação em Genética da UFMG pela oportunidade.

Ao Generation Challenge Programme e Fundação McKnight pelo suporte financeiro.

À Embrapa Milho e Sorgo pelo apoio proporcionado para a realização desse trabalho.

Aos meus orientadores, Dr. Jurandir Vieira de Magalhães e Dra. Cláudia Teixeira Guimarães, pela confiança em mim depositada, pelo apoio, orientação e ensinamentos.

Aos meus colegas do Curso de Pós-Graduação.

Aos amigos Janaína, Carlos, Uiraci, Belkiss, Luisa e Fernanda pela grande ajuda na condução dos experimentos, pelo incentivo e por tornarem os dias mais agradáveis.

À Gislene Cristeli, pela ajuda constante e grande amizade.

À Dra. Maria Lúcia Simeone pelo apoio e incentivo.

Ao Dr. Jon Shaff pelos ensinamentos e análises de ICP e exsudação de citrato realizadas.

Ao Dr. Robert Eugene Schafertt pelas sugestões.

Aos meus pais, Vânia e Marcos, meus exemplos de vida, pelo amor e apoio incondicional.

Ao meu querido irmão, Thiago, por partilhar tantos anos de convivência feliz.

Às minhas amigas, Danny, Flávia, Kívia, Tainá e Rejane, por estarem sempre presentes, mesmo distantes.

Às minhas amigas de república, Kátia, Maíra e Mayara, pelos momentos de descontração.

Aos amigos do NBA pela agradável convivência e apoio, em especial ao Miguel, Marília, Célio, Ivana, Fabiane, Sr. Joaquim e Raimundo.

A todos que direta ou indiretamente contribuíram para a conclusão dessa importante etapa da minha vida.

Muito Obrigada!

Sumário

Lista de Figuras	VII
Lista de Tabelas	x
Lista de Abreviaturas	XII
Resumo	15
Abstract	16
1. Introdução.....	17
1.1. Mecanismos Fisiológicos da Tolerância ao Alumínio	18
1.2. Genética da Tolerância ao Alumínio em Plantas	19
1.3. Bases Moleculares da Tolerância ao Alumínio	20
1.4. Variabilidade Alélica no Loco <i>Alt_{SB}</i>	23
1.5. Mineração de Alelos	24
2. Objetivos.....	26
2.1. Objetivo Principal.....	26
2.2. Objetivos Específicos	26
3. Material e Métodos	27
3.1. Desenho e Teste dos <i>Primers</i>	27
3.2. Estratégia de <i>Pooling</i>	30
3.3. Mineração de alelos do loco <i>Alt_{SB}</i>	31
3.4. Análise dos Haplótipos do Loco <i>Alt_{SB}</i>	34
4. Resultados e Discussão	37
4.1. Geração de Marcadores Haplótipos-específicos.....	37
4.2. Identificação de haplótipos com os marcadores do loco <i>Alt_{SB}</i>	47
4.3. Estratégia de <i>pooling</i> para os marcadores do loco <i>Alt_{SB}</i>	48
4.4. Mineração de alelos do Loco <i>Alt_{SB}</i>	50
4.5. Caracterização fisiológica e molecular dos acessos recuperados com os marcadores haplótipo-específicos do loco <i>Alt_{SB}</i>	59
5. Conclusão.....	66
6. Referencias Bibliográficas	67
Anexo 1	77

Lista de Figuras

- Figura 1 - Representação do loco *Alt_{SB}* indicando as seis regiões onde foram identificados os polimorfismos entre as linhagens parentais. O gene *SbMATE* (ORF7, exons representados por caixas cinza e introns por linhas pretas), o gene similar a uma proteína hipotética (ORF8), e o gene similar a sacarose fosfato sintase (ORF9) foram flanqueados por um marcador correspondente a um indel de 7 pb entre os parentais SC283 e BR007 na ORF6 e um SNP R(G/A) na ORF10. 23
- Figura 2 - Esquema de amplificação dos *primers* para os polimorfismos do tipo indel. As setas representam os *primers* e as cores, azul e vermelho, os sítios de anelamento. INS= Inserção. DEL=deleção. O *primer forward* é específico para inserção. 27
- Figura 3 - Ilustração esquemática do desenho de *primers* na técnica ARMS-PCR (adaptado de Ye et al., 2001). Nesse exemplo foi utilizado, a título de exemplificação, um SNP R (G/A). Quatro *primers*, dois alelos-específicos, internos, e dois *primers* externos são utilizados em uma única reação de PCR. Dois produtos alelos específicos, um para o alelo G e outro para o alelo A, são amplificados por dois pares de *primers*. O primeiro corresponde à amplificação do *primer A* alelo-específico com o *primer* externo reverso e o outro com o *primer G* alelo-específico com o *primer* externo *forward*. Um *mismatch*, representado pelo asterisco *, é adicionado na antepenúltima base do final 3' do *primer*. Os dois *primers* externos são desenhados de tal forma que os fragmentos alelo-específicos amplificados possuam tamanhos diferentes e possam ser resolvidos em gel de agarose. 28
- Figura 4 - Sistema ARMS-PCR utilizando três *primers*. Nesse caso, um *primer* alelo-específico, interno, e dois *primers* externos são utilizados em cada reação de PCR, sendo necessárias, portanto, duas reações de PCR separadas para a amplificação dos alelos. Um *mismatch*, representado pelo asterisco *, é adicionado na antepenúltima base do final 3' do *primer* interno, para garantir a especificidade da reação. As setas representam os *primers* e as suas cores o fragmento ao qual eles correspondem. 29
- Figura 5 - Representação do loco *Alt_{SB}* com as seis regiões amostradas para o mapeamento associativo com as respectivas posições em pares de bases (pb). Estão representados o gene *SbMATE* (ORF7, exons representados por caixas cinza e introns por linhas pretas), ORF8 e ORF9. 37
- Figura 6 - Esquema de amplificação do marcador *SbMATE_E1_5519pb*. Os padrões de amplificação são apresentados para dois indivíduos contendo os haplótipos contrastantes, SC566 (tolerante ao AI) e BR007 (sensível), e para o indivíduo heterozigoto. Foram feitas duas reações diferentes, sendo uma para cada alelo. Na primeira canaleta está o marcador de tamanho molecular. 39
- Figura 7 – Esquema de amplificação do marcador *SbMATE_I2_5985pb*. Os padrões de amplificação são apresentados para dois indivíduos contendo os haplótipos contrastantes, SC283 (tolerante ao AI) e BR007 (sensível), e para o indivíduo heterozigoto. Foram feitas duas reações diferentes, sendo uma para cada alelo. Na primeira canaleta está o marcador de tamanho molecular. 40

Figura 8 - Esquema de amplificação do marcador <i>SbMATE_I2_6083pb</i> . Os padrões de amplificação são apresentados para dois indivíduos com haplótipos contrastantes, SC283 (tolerante ao Al) e BR007 (sensível) e para o indivíduo heterozigoto. Na primeira canaleta está o marcador de tamanho molecular.	41
Figura 9 - Esquema de amplificação do marcador <i>SbMATE_I2_6094pb</i> . Os padrões de amplificação são apresentados para dois indivíduos contendo os haplótipos contrastantes, SC283 (tolerante ao Al) e BR007 (sensível), e para o indivíduo heterozigoto. Foram feitas duas reações diferentes, sendo uma para cada alelo. Na primeira canaleta está o marcador de tamanho molecular.	42
Figura 10 - Esquema de amplificação do marcador ORF8_8365pb. Os padrões de amplificação são apresentados para dois indivíduos com haplótipos contrastantes, SC283 (tolerante ao Al) e BR007 (sensível) e para o indivíduo heterozigoto. Na primeira canaleta está o marcador de tamanho molecular.	42
Figura 11 - Esquema de amplificação do marcador ORF8_8424pb. Os padrões de amplificação são apresentados para dois indivíduos com haplótipos contrastantes, SC283 (tolerante ao Al) e BR007 (sensível) e para o indivíduo heterozigoto. Na primeira canaleta está o marcador de tamanho molecular.	43
Figura 12 - Esquema de amplificação do marcador ORF9_12487pb. Os padrões de amplificação são apresentados para dois indivíduos com haplótipos contrastantes, SC283 (tolerante ao Al) e BR007 (sensível). Na primeira canaleta está o marcador de tamanho molecular.	43
Figura 13 - DNA <i>pooling</i> para o marcador <i>SbMATE_I2_6083pb</i> . Cada canaleta representa uma mistura de DNA genômico entre os indivíduos BR007 e SC283 nas proporções que estão devidamente identificadas na figura (proporção BR007:SC283). M= marcador de tamanho molecular 1Kb.	48
Figura 14 – Crescimento Líquido Relativo para os 30 acessos do painel do IGD que foram selecionados com base nos marcadores desenhados para os polimorfismos do loco <i>Alt_{SB}</i> . Esses acessos foram expostos à {27} µM de Al por cinco dias em solução nutritiva pH 4.0. As barras na cor cinza representam os controles negativos e com a cor branca o controle positivo, SC283.	53
Figura 15 - Expressão relativa do gene <i>SbMATE</i> em linhagens selecionadas de sorgo, tomando-se como calibrador a linhagem BR007. As linhagens de sorgo foram expostas a, {27} µM de Al por cinco dias em pH 4,0. Os valores representam médias de três repetições e as barras verticais indicam o desvio padrão. As letras abaixo do genótipo referem-se às classes de tolerância ao Al previamente caracterizados: (S) sensível, (T) tolerante e (AT) altamente tolerante.	61
Figura 16 - Acumulação de Al nos ápices radiculares de linhagens de sorgo submetidas à {27} µM de Al por 5 dias em solução nutritiva pH 4,0. SC566 [0 Al] e BR007 [0 Al] são os controles negativos do experimento e estão destacados em cinza claro. Os valores representam a média de três repetições e as barras verticais indicam o desvio padrão. As letras abaixo do genótipo referem-se às classes de tolerância ao Al, previamente	

caracterizados por Caniato (2009): (S) sensível, (T) tolerante, (AT) altamente tolerante e (C) controles. 63

Figura 17 – Exsudação de citrato pelos ápices radiculares de linhagens de sorgo submetidas a {0} μM (barras pretas) e {27} μM (barras cinza) de Al por 5 dias em solução nutritiva pH 4,0. Os valores representam a média de três repetições e as barras verticais indicam o erro padrão. As letras abaixo do genótipo referem-se às classes de tolerância ao Al, previamente caracterizados: (S) sensível, (T) tolerante e (AT) altamente tolerante..... 64

Lista de Tabelas

Tabela 1 – Polimorfismos do loco *Alt_{SB}*. Para cada polimorfismo estão representados o tipo de polimorfismo, indel ou SNP, com o alelo favorável indicado em negrito, a região em que o polimorfismo está localizado (vide Figura 5), o valor de significância de associação do polimorfismo à característica (p-valor) e a porcentagem da variância fenotípica explicada para CLR_{5d} (r^2). Fonte: Caniato et al. (2009). 37

Tabela 2 – Descrição dos marcadores haplótipo-específicos para o loco *Alt_{SB}*. Para cada polimorfismo está indicado o tipo do polimorfismo de acordo com a nomenclatura IUPAC, o tipo de sistema ARMS-PCR que foi desenvolvido para cada loco e a combinação de *primers* para amplificar cada fragmento específico com os tamanhos de fragmentos em pares de bases (pb)..... 38

Tabela 3 – Descrição completa dos sistemas de marcadores, indicando o tipo do polimorfismo, o sistema ARMS-PCR em que foi baseado o marcador, a temperatura de anelamento para cada conjunto o identificador dos *primers* (Id), o tipo do *primer*, a temperatura de dissociação (T_m , *melting temperature*) assim como a sequência de cada *primer*. Os alelos favoráveis estão destacados em negrito. 45

Tabela 4 - Resultados dos cálculos para determinação do tamanho ideal do *pool* de DNA. Os valores correspondem à eficiência da estratégia (1-F) para diferentes tamanhos de população (n) e diferentes números de indivíduos no pool (n), considerando a frequência do alelo de interesse previamente determinada. Os valores máximos para cada condição estão destacados em negrito..... 49

Tabela 5 - Haplótipos dos 14 acessos tolerantes do painel ICRISAT/CIRAD para o loco *Alt_{SB}*. Estão apresentados os polimorfismos que foram utilizados para o desenvolvimento dos marcadores ARMS-PCR. Os alelos favoráveis para cada polimorfismo estão destacados em negrito. 51

Tabela 6 - Haplótipos do loco *Alt_{SB}* presentes no painel do IGD. Foram considerados os polimorfismos utilizados na genotipagem do painel com quatro marcadores ARMS-PCR. A localização dos polimorfismos está representada na Figura 4. As células em negrito indicam a presença do alelo favorável. Nas duas primeiras linhas estão descritos os haplótipos das linhagens BR007 e SC283, padrões de sensibilidade e tolerância ao AI, respectivamente.. 52

Tabela 7 - Acessos de sorgo do painel do IGD recuperados em três classes de tolerância, altamente tolerante (AT, CLR_{5d} > 80%), tolerante (T, 30%<CLR_{5d}<80%) e sensível (CLR_{5d}<30%) pela genotipagem com quatro marcadores no loco *Alt_{SB}*. A posição dos marcadores no contexto no loco *Alt_{SB}* está mostrada na Figura 4. %Total denota a porcentagem de indivíduos recuperados por cada marcador em relação ao total de indivíduos recuperados pela genotipagem. %Marcador denota a porcentagem de indivíduos de cada classe em relação ao número total de indivíduos recuperados por cada marcador.54

Tabela 8 - Eficiência da estratégia de *pooling* quanto à redução do número de genotipagens no painel INRAN. Nº de *pools* positivos denota o número de *pools* que apresentaram o alelo de interesse. Nº de reações necessárias para abertura dos *pools* denota o número de reações que são necessárias ao se genotipar todos os *pools* que apresentaram o alelo de

interesse. Total indica o total de reações para genotipar o painel com cada marcador (os 41 *pools* iniciais mais os *pools* positivos). Eficiência (1-F) indica a eficiência da metodologia de *pooling*, sendo F a proporção entre o número de reações realizadas e o número de reações necessárias sem a metodologia *pooling*. 55

Tabela 9 - Haplótipos do loco *Alt_{SB}* presentes no painel do INRAN. Foram considerados os polimorfismos utilizados na genotipagem do painel com seis marcadores ARMS-PCR. A localização dos polimorfismos está representada na Figura 4. As células em negrito indicam a presença do alelo favorável. Nas duas primeiras linhas estão descritos os haplótipos das linhagens BR007 e SC283, padrões de sensibilidade e tolerância ao Al, respectivamente.. 56

Tabela 10 – Avaliação dos acessos do painel do INRAN para tolerância em {27} µM de Al solução nutritiva. CRR_{3D} e CRR_{5D} denotam o crescimento radicular relativo aos 3 e 5 dias de estresse de Al, respectivamente; CRR_{5D}MIN denota o valor mínimo do CRR_{5D} encontrado para aquele acesso; CRR_{5D}MAX denota o valor máximo do CRR_{5D} encontrado para aquele acesso; IRV Média denota a injúria radicular visual média das raízes dos acessos e N° Plantas Tolerantes denota o número de plantas com IRV>3 que foram classificadas como tolerantes para cada acesso. 57

Tabela 11 - Descrição dos nove haplótipos do loco *Alt_{SB}* que apresentaram acessos tolerantes. Cada coluna apresenta um loco polimórfico associado com a tolerância ao Al e o número de linhagens que apresentam o haplótipo. A linhagem BR007, padrão de sensibilidades, possui o haplótipo *Alt_{SB}_2*, o segundo mais frequente do painel e presente principalmente em linhagens sensíveis. Já a linhagem SC283, padrão de tolerância, possui o haplótipo *Alt_{SB}_22*, presente em linhagens altamente tolerantes e tolerantes. Esses dois haplótipos estão destacados em negrito. 59

Tabela 12 - Acessos de sorgo selecionados para a caracterização fisiológica e molecular dos haplótipos. Nome dos acessos, haplótipo do loco *Alt_{SB}*, número de acessos no painel de associação ICRISAT/CIRAD que apresentam esse haplótipo, classe de tolerância ao Al, país de origem do acesso e as variáveis fenotípicas CLR_{3d}, CLR_{5d} e IRV. Os valores representam médias de duas repetições e sete plantas por repetição obtidos por Caniato (2009). 60

Tabela 13 - Correlação entre as variáveis estudadas para as 16 linhagens representantes dos haplótipos do loco *Alt_{SB}*. Para cada correlação estão representados os coeficientes de correlação de Pearson, r, na primeira linha, seguidos dos p-valores. Exsudação denota a exsudação de citrato pelas raízes por 6 horas em 5 dias na presença de {27} µM de Al. CLR_{3d} e CLR_{5d} denotam o crescimento líquido relativos das raízes 3 e 5 dias na presença de {27} µM de Al, respectivamente. Expressão denota a expressão relativa do gene *SbMATE* após 5 dias na presença de {27} µM de Al. 'Ac Al' denota a acumulação de Al nos ápices radiculares após um período de 5 dias na presença de {27} µM de Al. 62

Lista de Abreviaturas

°C: Grau Centígrado

ABC: *ATP Binding Cassette* (domínio de ligação de ATP)

Al: Alumínio

ALMT1: Aluminum-activated Malate Transporter (transportador de malato ativado por alumínio)

Alt1: loco de tolerância ao Al em centeio

Alt2: loco de tolerância ao Al em centeio\

Alt3: loco de tolerância ao Al em centeio

Alt4: loco de tolerância ao Al em centeio

Alt_{BH}: loco de tolerância ao Al em trigo

Alt_{SB}: gene de tolerância ao alumínio em sorgo

AOs: ácidos orgânicos

ARMS: *Amplification Refractory Mutation System* (sistema de amplificação refratária de mutação)

AtALMT1: gene do transportador de malato de arabidopsis

AtMATE: gene do transportador de citrato de arabidopsis

ATP: adenosina trifosfato

BnALMT1: gene do transportador de malato de canola

BnALMT1: gene do transportador de malato de canola

BnALMT2: gene do transportador de malato de canola

C_{0d}: comprimento radicular inicial

C_{3d}: comprimento radicular aos 3 dias de crescimento

C_{5d}: comprimento radicular aos 5 dias de crescimento

CAPS: *Cleaved Amplified Polymorphic Sequences* (clivagem de sequências polimórficas amplificadas)

CIRAD: *Centre de Cooperation Internationale em Recherche Agronomique pour Le Développement* (França)

CLR: Crescimento Líquido Relativo

cM: centiMorgan

CRR: Crescimento Radicular Relativo

DNA: ácido dexidoribonucleico

dNTP: desoxido nucleotídeo trifosfato

Frd3: gene do transportador de citrato de citrato

ICRISAT: *International Crops Research Institute for the Semi-Arid Tropics* (India)

IGD: *Institute for Genomic Diversity* (EUA)

INRAN: *Institute National de la Recherche Agronomique du Niger* (Niger)

Indel: inserção/deleção

IRV: Injúria Radicular Visual

MATE: *Multi-drug And Toxin Efflux* (transportador de membrana)

μL: microlitro

μM: micromolar

mL: mililitro

mM: milimolar

MITE: *Miniature Inverted Repeated Transposable Elements* (elementos curtos de transposição)

Mpb: mega pares de bases

ng: nanogramo

ORF: *Open Read Frame* (quadro de leitura aberta)

PCR: Polymerase Chain Reaction (reação em cadeia da polimerase)

pH: potencial hidrogeniônico

QTL: *Quantitative Trait Loci* (loco de característica quantitativa)

QTN: *Quantitative Trait Nucleotide* (nucleotídeo de característica quantitativa)

RFLP: *Restriction Fragment Length Polymorphism* (polimorfismo de comprimento de fragmento de restrição)

RIL: *Recombinant Inbred Lines* (linhagens recombinantes)

RNA: ácido ribonucleic

SAM: seleção assistida por marcadores

SbMATE: gene do transportador de citrato de sorgo

ScALMT1: transportador de malato de centeio

SNP: *Single Nucleotide Polimorphisms* (polimorfismo de base única)

SSR: *Single Sequence Repeats* (repetição de sequências simples)

STS: *Sequence-Tagged Sites* (sítios marcados por sequências)

TaALMT1: gene do transportador de malato de trigo

TaMATE1: gene do transportador de citrato de trigo

T_m: *Temperature of melting* (temperatura de fusão)

U: unidade enzimática

Resumo

HUFNAGEL, Bárbara, M.Sc., Universidade Federal de Minas Gerais, Dezembro 2009.
Mineração e Caracterização de Alelos do Gene de Tolerância ao Alumínio *Alt_{SB}* em Sorgo. Orientador: Dr. Jurandir Vieira de Magalhães. Co-orientadora: Dra. Cláudia Teixeira Guimarães.

A toxidez causada pelo alumínio é um dos principais fatores que afetam o crescimento e o desenvolvimento das culturas em solos ácidos, que ocupam grandes extensões de terras agricultáveis no mundo. Um gene de efeito maior na tolerância ao alumínio em sorgo, *Alt_{SB}*, foi identificado via clonagem posicional. Esse gene codifica um transportador de citrato localizado na membrana plasmática das células dos ápices radiculares, e confere tolerância por meio de um mecanismo fisiológico baseado na exsudação de citrato ativada pelo alumínio na rizosfera. A tolerância ao Al é uma característica relativamente rara em sorgo, ocorrendo em uma frequência de aproximadamente 5%, o que dificulta a identificação de novas fontes de tolerância ao Al pelo melhorista. Em trabalhos prévios, o gene *Alt_{SB}* foi submetido a uma análise associativa que revelou 10 polimorfismos associados a características relacionadas à tolerância ao alumínio. No presente trabalho, sete desses polimorfismos foram convertidos em marcadores do tipo *Amplification Refractory Mutation System* (ARMS), que são marcadores de fácil genotipagem e de baixo-custo, e que foram utilizados para a identificação de alelos do gene *Alt_{SB}* que conferem tolerância ao Al. Em dois painéis de sorgo, os marcadores permitiram a recuperação de 7 e 9 acessos tolerantes. Na genotipagem de um painel de melhoramento foi aplicada uma estratégia de *pooling* de DNA que aumentou a eficiência do processo, resultando em uma economia de 54% das reações de PCR necessárias para genotipar todo o painel. Estudos posteriores confirmaram que parte dos acessos identificados possui alelos funcionais do gene *Alt_{SB}*, sendo, consequentemente, tolerantes ao Al. Entretanto, análises de expressão do gene *Alt_{SB}*, de acúmulo de alumínio nos ápices radiculares e de exsudação de citrato pela raiz indicaram que um dos acessos identificados possui gene(s) de tolerância distinto(s) do gene *Alt_{SB}*. Esse(s) gene(s) controlam, possivelmente, um mecanismo de detoxificação interna de Al, mecanismo esse totalmente distinto dos já descritos em sorgo até o momento. Com os marcadores gerados nesse trabalho pode-se agora buscar na diversidade genética da espécie, de forma eficiente, acessos de sorgo altamente tolerantes ao Al.

Abstract

HUFNAGEL, Bárbara, M.Sc., Universidade Federal de Minas Gerais, December 2009. Allele **Mining and Characterization of haplotypes of the aluminum tolerance gene, Alt_{SB} , in Sorghum**. Adviser: Dr. Jurandir Vieira de Magalhães. Co-adviser: Dra. Cláudia Teixeira Guimarães.

Aluminum toxicity is one of the major factors impairing crop growth and development on acid soils, which occupy large areas of potentially arable lands worldwide. A major Al tolerance gene in sorghum, Alt_{SB} , was isolated via positional cloning. This gene encodes a root citrate efflux transporter located in the plasma membrane of root apical cells, and underlies a physiological mechanism based on Al-induced citrate release into the rhizosphere. Al tolerance is a relatively rare trait in sorghum, occurring at a frequency of 5%, which makes it difficult for the plant breeder to identify sources of Al tolerance in the breeding population. Within previous studies, the Alt_{SB} gene was subjected to an association analysis that revealed 10 polymorphisms associated with traits related to aluminum tolerance. In the present work, seven of these polymorphisms were converted into *Amplification Refractory Mutation System* (ARMS) markers, which are easy to use and amenable to low-cost genotyping, to identify Alt_{SB} alleles that confer Al tolerance. These markers recovered 7 and 9 Al-tolerant accessions in two different sorghum panels. A DNA pooling strategy was applied allowing for greater efficiency in the genotyping process with a 54% economy of the PCR reactions needed to genotype all individuals in the panel. Additional studies confirmed that some of the accessions that were identified have functional Alt_{SB} alleles being thus Al tolerant. However, analysis of Alt_{SB} expression, Al accumulation and citrate exudation indicated that one of the accessions rely on Al tolerance gene(s) distinct from Alt_{SB} to express its tolerance. These genes possibly control a tolerance mechanism based on internal detoxification, which has not been described in sorghum. With the markers generated in this work, it is now possible to efficiently identify Al tolerant accessions in the species genetic diversity.

1. Introdução

Solos ácidos ocorrem em áreas extensivas nas zonas tropicais e temperadas, afetando 3,95 bilhões de hectares e correspondendo a mais de 30% das terras cultiváveis no mundo (von Uexkull e Mutert, 1995). Esses solos se caracterizam por valores de pH <5,0, baixa disponibilidade de nutrientes essenciais ao desenvolvimento das plantas e elevada toxicidade mineral, em especial de aluminosilicatos. No Brasil, em torno de 127 milhões de hectares aptos para a utilização agrícola são compostos por solos ácidos (Rocha, 1997) que, devido ao excesso de manganês, à deficiência de fósforo e, principalmente, à toxidez causada pelo alumínio, apresentam limitações à produção vegetal.

O alumínio (Al) é o metal mais abundante na crosta terrestre e está presente no solo em formas minerais como aluminosilicatos, óxidos, fosfatos e sulfatos. Em valores de pH neutro, o Al encontra-se, predominantemente, na forma de $\text{Al}(\text{OH})_3$ que não é tóxica para as plantas (Driscoll e Schecher, 1990). Entretanto, em condições ácidas (pH <5,5), íons Al^{3+} , que estão na forma de cátion trivalente hexa-hidratado ($\text{Al}(\text{H}_2\text{O})_6^{3+}$), sofrem hidrólises sucessivas, dando origem às formas $\text{Al}(\text{OH})^{2+}$ e $\text{Al}(\text{OH})_2^+$ (Martin, 1992), que possuem efeitos fitotóxicos. A exposição ao Al dos ápices radiculares de plantas sensíveis resulta em uma rápida inibição do crescimento das raízes (Ryan et al. 1992; Sivaguru et al., 1999). O Al atinge várias estruturas celulares, incluindo a membrana plasmática e a parede celular, e afeta processos celulares como vias de transdução de sinais e a homeostase do cálcio (Jones et al., 1998; Zhang e Rengel, 1999; Ma et al, 2002). A toxidez desse metal leva a mudanças acentuadas na morfologia da raiz, causando o inchamento do ápice e o comprometimento da formação de raízes secundárias (Ciamporova, 2002). Por limitar o aprofundamento radicular, o alumínio dificulta a absorção de nutrientes e água, aumentando a sensibilidade das plantas a vários estresses, especialmente o estresse de seca (Foy et al., 1993).

A severidade da toxidez do alumínio depende ainda da atividade livre do Al, da duração da exposição e das condições ambientais (Kochian et al., 2005; Delhaize et al., 2007) bem como da sensibilidade ao metal. Existe uma grande variação na tolerância ao alumínio tanto entre espécies de plantas bem como entre indivíduos dentro de uma mesma espécie, que é resultado de mecanismos complexos, internos ou externos, para combater a toxidez do alumínio (Kochian, 1995).

O sorgo (*Sorghum bicolor* L. Moench) é o quinto cereal mais produzido no mundo, depois do trigo, arroz, milho e cevada (FAO, 2009), e vem sendo cultivado para a produção de biomassa para a fabricação de biocombustível e alimentação animal, além da produção de grãos e bebida para o consumo humano. Esse cereal, que é anual e predominantemente

de auto-fecundação, apresenta uma notável tolerância à seca e a outros estresses ambientais, tornando-se uma excelente alternativa para a produção agrícola em áreas marginais (Amaral et al., 2003). O sorgo apresenta várias características interessantes para pesquisa genômica em gramíneas e, com a disponibilização da sequência completa de seu genoma (<http://www.phytozome.net/sorghum>), há um enorme potencial para a sua exploração visando à descoberta de genes de interesse agrônomo.

Sorghum é um gênero muito diverso, que compreende espécies cultivadas e espontâneas (Schertz e Dalton, 1988). O sorgo cultivado, *Sorghum bicolor*, é uma espécie diplóide cujo genoma apresenta conteúdo haplóide de 10 cromossomos e tamanho físico de 697 mega pares de bases (Mpb). O sorgo é dividido em cinco raças morfológicas básicas, bicolor, caudatum, durra, guinea e kafir, e em outras dez raças intermediárias, que são combinações entre as raças básicas (Harlan e De Wet, 1972). Acredita-se que esse cereal tenha sido domesticado no nordeste da África entre 5000 e 6000 anos atrás, a partir da espécie selvagem *Sorghum verticilliflorum*, que deu origem, posteriormente, a raça bicolor considerada primordial (Deu et al., 2008; Shewayyrga et al., 2006). Estudos demonstram que os padrões de diversidade genética em sorgo refletem fatores como origem geográfica e racial (Aldrich et al., 1992; Deu et al., 2006; 2008; Cui et al., 1995; Djè et al., 2000; Barnaud et al., 2007).

1.1. Mecanismos Fisiológicos da Tolerância ao Alumínio

Os mecanismos de tolerância ao alumínio propostos na literatura podem ser divididos em dois grupos: mecanismos simplásticos, que envolvem a detoxificação do Al a partir da sua imobilização ou neutralização dentro da célula, e mecanismos de exclusão ou apoplásticos, que decorrem da neutralização ou imobilização do Al externamente à célula. A detoxificação interna do alumínio resulta da formação de complexos de alumínio com ácidos orgânicos ou outros quelantes e o sequestro desses complexos nos vacúolos (Ma et al., 1997). Acredita-se que, em algumas espécies tolerantes, o Al penetra no simplasto, onde são formados complexos com ácidos orgânicos, sendo posteriormente transportados através do sistema vascular para as partes aéreas da planta, onde são acumulados dentro de vacúolos, como em arábida (Larsen et al., 2005) e hortênsia (Takeda et al., 1985), ou em outros tecidos menos sensíveis ao Al.

O mecanismo de exclusão é o mais aceito e comprovado cientificamente para a tolerância ao alumínio em plantas (Kochian et al., 2005; Delhaize et al., 2007), ocorrendo em diversas espécies como o milho (Piñeros et al., 2002), feijão (Miyasaka et al. 1991), trigo (Delhaize et al., 1993), arábida (Hoekenga et al., 2003; Liu et al., 2009), triticale (Ma et

al., 2000), sorgo (Magalhães et al., 2007), dentre outras. Nesse mecanismo, ácidos di e tricarboxílicos de baixo peso molecular são liberados pelos ápices radiculares, formando complexos estáveis e não-tóxicos com os íons Al^{3+} presentes na rizosfera. Tais complexos são incapazes de atravessar a membrana plasmática, reduzindo ou mesmo anulando os efeitos fitotóxicos do alumínio (Kochian et al., 2005).

Os ácidos orgânicos (AOs) mais comumente liberados dos ápices das raízes são o malato e o citrato, sendo que o citrato possui uma capacidade estimada de 6 a 8 vezes maior para quelar o Al do que o malato, formando complexos mais estáveis com esse metal (Hue et al., 1986). Em muitas espécies de plantas, a exposição ao Al ativa rapidamente a exsudação de ácidos orgânicos e o tipo de ácido orgânico exsudado é, aparentemente, espécie-específico. Em trigo (*Triticum aestivum*), o alumínio ativa a exsudação de malato quase que instantaneamente e a taxa de liberação deste AO permanece constante ao longo do tempo de exposição ao metal (Osawa e Matsumoto, 2001). Já em outras espécies como *Secale cereale* (Li et al., 2000), *Cassia tora* (Ishikawa et al., 2000), *Brassica napus* (Ligaba et al., 2006), *Arabidopsis thaliana* (Kobayashi et al., 2007) e *Sorghum bicolor* (Magalhães et al., 2007) ocorre um período de indução para a liberação de AOs após a exposição ao alumínio.

1.2. Genética da Tolerância ao Alumínio em Plantas

O controle genético da tolerância ao Al tem sido bastante estudado em gramíneas, especialmente em membros da tribo Triticeae, representada pelo trigo, pelo centeio e pela cevada. Em alguns cultivares de trigo, muitos genes com efeitos aditivos parecem estar envolvidos (Aniol e Gustafson, 1984; Aniol, 1991), sendo a tolerância controlada por um gene de efeito maior em outros cultivares (Aniol e Gustafson, 1984; Larkin, 1987). Delhaize et al. (1993) demonstraram que o loco *Alt1* foi responsável pela tolerância e sensibilidade de duas linhagens isogênicas de trigo contrastantes para essa característica. Esse parece ser o mesmo loco identificado como *Alt2* por Luo e Dvorak (1996). Riede e Anderson (1996) mapearam um gene de tolerância de efeito maior, denominado *Alt_{BH}*, a 1,1 cM do marcador RFLP bcd1230 no cromossomo 4DL de trigo. Esse gene explicou 85% da variação fenotípica para a tolerância ao Al em uma população de *Recombined Inbred Lines* (RILs), derivadas do cruzamento entre as linhagens BH1146 e Anahuac, tolerante e sensível ao Al, respectivamente. O primeiro gene de tolerância ao Al clonado em plantas foi o *ALMT1* em trigo (Sasaki et al., 2004), que provavelmente corresponde ao gene *Alt_{BH}*.

Em centeio foram descritos quatro locos relacionados com a tolerância ao Al: *Alt1*, localizado no cromossomo 6RS; *Alt2*, localizado no cromossomo 3RS; *Alt3*, no cromossomo

4RL e *Alt4* no cromossomo 7RS (Aniol e Gustafson, 1984; Galego et al., 1998; Matos et al., 2005). Quatro *Quantitative Trait Loci* (QTLs) de tolerância ao Al foram descritos em aveia (*Avena strigosa* Schreb.), explicando 55% da variação fenotípica da tolerância (Wight et al., 2006). A tolerância ao Al em arroz apresenta herança complexa e QTLs associados a essa característica foram mapeados em todos os 12 cromossomos de arroz (Wu et al., 2000; Nguyen et al., 2001, 2002, 2003). Ao todo foram identificados 40 QTLs em quatro populações diferentes, sendo que alguns QTLs foram consistentemente identificados em mais de uma população (Wu et al., 2000; Nguyen et al., 2001, 2002, 2003).

Em sorgo, a herança da tolerância ao Al em uma população $F_{2:3}$ derivada do cruzamento da linhagem tolerante SC283 e da sensível BR007 foi compatível com a presença de um gene de tolerância de efeito maior, denominado *Alt_{SB}*, que apresentou dominância parcial nas condições adotadas para a avaliação do fenótipo (Magalhães et al., 2004). Os estudos de genômica comparativa relatados por esses autores indicaram que o loco *Alt_{SB}* em sorgo e o loco *Alt_{BH}* em trigo não são ortólogos, enquanto que um QTL maior para tolerância ao Al detectado no cromossomo 1 de arroz em vários estudos (Wu et al. 2000; Nguyen et al. 2001, 2002, 2003) encontra-se em posição sintênica aquela do loco *Alt_{SB}* em sorgo. A herança da tolerância ao Al em sorgo foi estudada por Caniato et al. (2007) por meio de uma análise genética em 12 populações F_2 com marcadores do tipo STS (*Sequence-Tagged Sites*) que flanqueiam o loco *Alt_{SB}*. Os autores relataram que a grande variabilidade fenotípica para a tolerância ao Al observada entre 12 linhagens de sorgo pôde ser explicada por uma séria alélica no loco *Alt_{SB}* e pela presença de outros genes distintos ao gene *Alt_{SB}*.

1.3. Bases Moleculares da Tolerância ao Alumínio

Os genes que controlam a tolerância ao Al até então identificados codificam proteínas transportadoras de membrana, pertencentes às famílias *Al-activated Malate Transporter* (ALMT) e *Multi-drug And Toxin Efflux* (MATE), que transportam malato e citrato, respectivamente (Sasaki et al., 2004; Magalhães et al., 2007). A família de genes que codificam proteínas ALMT foi a primeira a ser identificada. O gene *TaALMT1*, de trigo, codifica uma proteína hidrofóbica contendo de cinco a sete domínios transmembrana, que media o efluxo de malato para a rizosfera. Foram identificados dois haplótipos desse gene, diferenciados por seis *single nucleotides polymorphism* (SNPs) na região codante, que resultam em uma substituição aminoacídica. No entanto, ambos os haplótipos ocorrem tanto em genótipos tolerantes quanto nos sensíveis, indicando que essa substituição não é o fator causativo da tolerância ao Al em trigo, que, por sua vez, parece ser modulada pelo nível de

expressão desses alelos (Sasaki et al., 2004). O gene *TaALMT1* é expresso constitutivamente no ápice radicular de linhagens tolerantes e os níveis de expressão estão positivamente correlacionados com a tolerância ao alumínio e a exsudação de malato (Raman et al., 2005). Outros experimentos mostraram que plantas de tabaco e de cevada superexpressando o gene *TaALMT1* apresentaram maior efluxo de malato ativado pelo Al e uma tolerância ao alumínio superior quando comparada aos controles (Sasaki et al., 2004; Delhaize et al., 2004).

Genes homólogos ao gene *TaALMT1* conferem tolerância ao alumínio em outras espécies de plantas. O loco *Alt₄* de centeio (*Secale cereale*) possui um agrupamento de genes homólogos ao *TaALMT1* (Collins et al., 2008). O gene *ScALMT1* codifica uma proteína com 86% de identidade com aquela codificada pelo *TaALMT1*, cuja expressão é induzida pelo Al e é maior nos cultivares tolerantes quando comparados aos sensíveis, sendo responsável pela tolerância ao Al em centeio (Collins et al., 2008). *Brassica napus* possui dois genes da família ALMT, o *BnALMT1* e *BnALMT2* (Ligaba et al., 2006), que também são induzidos pelo Al e ativam a exsudação de malato pelos ápices radiculares. O gene *AtALMT1*, que confere tolerância ao alumínio em arabidopsis, possui 44% de similaridade com o gene *TaALMT1*, é induzido em presença de Al e ativa a liberação de malato (Hoekenga et al., 2006).

Utilizando a técnica de clonagem posicional em sorgo, Magalhães et al. (2007) localizaram o loco *Alt_{SB}* em um fragmento de 24,6 Kpb (kilo pares de bases), incluindo um membro da família MATE, *SbMATE*, cuja expressão nos ápices radiculares de uma linhagem isogênica tolerante foi significativamente maior do que na linhagem sensível. A expressão do gene *SbMATE* na linhagem tolerante foi induzida pelo tratamento com Al, ao longo do tempo de exposição ao metal, sendo seguida por um aumento na exsudação de citrato pelos ápices radiculares. Já a linhagem isogênica sensível apresentou inibição drástica do crescimento radicular causada pelo Al, taxa de exsudação de citrato extremamente reduzida e expressão do gene *SbMATE* abaixo do limite de detecção. Adicionalmente, plantas transgênicas de arabidopsis superexpressando *SbMATE* apresentaram maior tolerância ao Al e exsudação de citrato quando comparadas com os controles. Assim, concluiu-se que o gene *SbMATE* codifica um transportador de citrato induzido pelo Al que confere tolerância ao Al no loco *Alt_{SB}* (Magalhães et al., 2007)

As proteínas da família MATE estão presentes em quase todos os procariotos e eucariotos, tendo sido caracterizadas primeiramente em bactérias por sua capacidade de conferir resistência às drogas pelo mecanismo de efluxo de pequenos compostos tóxicos e metabólitos secundários (Omote et al., 2006). O primeiro gene da família MATE identificado em plantas foi o gene *Frd3*, em arabidopsis, que codifica uma proteína que transporta citrato

para dentro do xilema na forma de um complexo Fe:citrato. Além disso, a expressão ectópica desse gene causou uma exsudação contínua de citrato nas raízes de *arabidopsis*, associada a um aumento da tolerância ao alumínio (Durrett et al., 2007). O trigo também apresenta um mecanismo secundário de tolerância ao alumínio nos genótipos altamente tolerantes, baseado na exsudação de citrato, modulada pelo gene *TaMATE1* (Ryan et al., 2009). No entanto, ao contrário de outros genes *MATE* como o *SbMATE* em sorgo (Magalhães et al., 2007), o gene *TaMATE1* apresenta padrão de expressão constitutiva em trigo. Já o gene *AtMATE*, de *arabidopsis*, é homólogo ao gene *SbMATE* e sua indução na presença de alumínio ativa a exsudação de citrato pelo ápice radicular (Liu et al., 2009).

Recentemente, outros genes têm sido implicados na tolerância ao Al em plantas. O gene *Stop1*, presente em *arabidopsis*, codifica uma proteína do tipo *zinc-finger* (Cys_2Hys_2ZF) que é um fator de transcrição localizado no núcleo (Iuchi et al., 2007). Uma mutação nesse gene, que resulta na substituição de um resíduo de histidina por uma tirosina, reprime a expressão dos genes *AtALMT1* (Iuchi et al., 2007) e *AtMATE* (Liu et al., 2009), resultando na perda do mecanismo de liberação de malato e de citrato e, conseqüentemente, da tolerância ao alumínio. Esses resultados indicam que o gene *Stop1* é essencial na indução da expressão desses dois genes, sendo um fator chave nas vias de transdução de sinal que regulam os mecanismos de tolerância ao alumínio (Sawaki et al., 2009). Esse gene regula também um mecanismo de tolerância a protons H^+ e de combate à infecção por patógenos. Assim, o gene *Stop1* parece ter uma ação pleiotrópica nos mecanismos de defesa a estresses abióticos e bióticos (Sawaki et al., 2009; Liu et al., 2009).

Há ainda os genes envolvidos na tolerância ao alumínio, ligados a mecanismos de detoxificação interna desse metal, que são os transportadores ABC (Huang et al., 2009; Larsen et al., 2005). Esses transportadores possuem a habilidade de transportar moléculas através das membranas mediante a hidrólise de ATP, cujas proteínas, em geral, possuem dois domínios relativamente bem conservados para ligação de ATP (*ATP Binding Cassette* – ABC), além de dois domínios transmembrana (Schulz e Kolukisaoglu, 2006). Os transportadores ABC representam uma grande família de proteínas em plantas, com mais de 120 membros em *arabidopsis* e arroz. Dos transportadores ABC de plantas que foram funcionalmente caracterizados, alguns agem sequestrando substâncias tóxicas orgânicas e inorgânicas para dentro de vacúolos (Larsen et al., 2005; Huang et al., 2009). Em arroz, foram identificados dois genes, *Star1* e *Star2*, responsáveis pela tolerância ao Al (Huang et al., 2009). Tais autores observaram que esses genes são expressos principalmente nas raízes e são induzidos pelo alumínio, e que as proteínas codificadas por eles participam do complexo do transportador ABC.

1.4. Variabilidade Alélica no Loco *Alt_{SB}*

Variações fenotípicas para a tolerância ao AI em sorgo são condicionadas por uma série alélica no loco *Alt_{SB}* (Caniato et al., 2007). Com o trabalho de clonagem posicional do gene *Alt_{SB}* foi determinada uma região de 24,6 kpb no cromossomo 3 na qual o loco de tolerância está localizado (Magalhães et al. 2007). Essa região, flanqueada por marcadores moleculares do tipo STS, contém três quadros de leitura aberta (*open-reading frame*, ORF), sendo que uma dessas ORFs, a ORF7 (Figura 1) corresponde ao gene *SbMATE*, que é responsável pela tolerância ao AI em sorgo. Um gene similar a uma proteína hipotética (ORF8) e um outro similar ao gene da sacarose fosfato sintase (ORF9, Figura 1) foram também encontrados na região. Os polimorfismos na região foram determinados a partir do sequenciamento desse fragmento nos parentais da população de mapeamento BR007, sensível ao AI, e SC283, tolerante ao AI. Entre os dois parentais foram identificados polimorfismos do tipo SNP, inserção/deleção (indel) e uma inserção de transposon do tipo MITE entre as linhagens SC283 e BR007 (Figura 1).

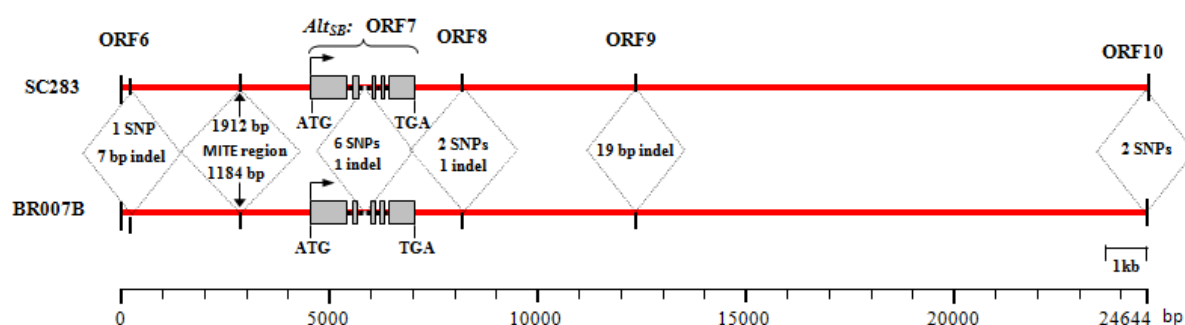


Figura 1 - Representação do loco *Alt_{SB}* indicando as seis regiões onde foram identificados os polimorfismos entre as linhagens parentais. O gene *SbMATE* (ORF7, exons representados por caixas cinza e introns por linhas pretas), o gene similar a uma proteína hipotética (ORF8), e o gene similar a sacarose fosfato sintase (ORF9) foram flanqueados por um marcador correspondente a um indel de 7 pb entre os parentais SC283 e BR007 na ORF6 e um SNP R(G/A) na ORF10.

Em um trabalho posterior, esses polimorfismos foram avaliados quanto ao seu papel na tolerância ao AI em um painel composto de 254 linhagens de sorgo. Por meio de mapeamento associativo (Yu et al., 2006), foram identificados 10 polimorfismos significativamente associados com a tolerância ao AI em sorgo. Os haplótipos elite do gene *Alt_{SB}*, isto é, os haplótipos pertencentes aos materiais mais tolerantes, estão presentes em frequência extremamente baixa (0,047) dentro do painel de associação. Isso ilustra a dificuldade na identificação de fontes de tolerância ao AI para serem utilizados nos programas de melhoramento. Entretanto, polimorfismos associados com a tolerância ao

alumínio podem ser utilizados na mineração de alelos superiores, isto é, alelos que confirmam alta tolerância ao Al em sorgo. A identificação e validação desses alelos por meio de marcadores haplótipo-específicos em um germoplasma diverso constituiu a hipótese desse estudo.

1.5. Mineração de Alelos

A mineração de alelos busca a identificação alelos úteis em coleções de ampla variabilidade genética, como aquelas presentes nos bancos de germoplasma. A utilização de tais materiais no melhoramento de plantas depende do conhecimento do desempenho agrônomo e de características que sejam de interesse. Normalmente, a diversidade desses painéis é explorada pela sua avaliação fenotípica, que, geralmente, se mostra inviável para um grande número de linhagens (Gebhardt et al., 2004; Prada, 2009). Assim, grande parte da diversidade existente nesses painéis permanece inexplorada devido à falta de ferramentas rápidas e eficientes para se identificar alelos úteis ao melhoramento vegetal (Kaour et al., 2008).

Com o desenvolvimento da biotecnologia nas últimas décadas, grandes avanços nos estudos da genômica têm permitido a identificação e clonagem de genes de interesses agrônômicos como aqueles responsáveis pela tolerância ao Al (Huang et al., 2003; Sasaki et al., 2004; Magalhães et al., 2007). Entretanto, um dos maiores desafios na pesquisa agrônômica atual é transformar o conhecimento gerado em tecnologias que melhoristas possam usar rotineiramente em seus programas.

Várias técnicas têm sido utilizadas na mineração de alelos, mas a maioria é baseada em técnicas de detecção de *Single Nucleotide Polymorphisms*, SNPs (Batley et al., 2003; Latha et al., 2004; Chagnè et al., 2008). Os SNPs são os polimorfismos mais abundantes nos genomas das espécies e, juntamente com os polimorfismos do tipo inserção/deleção (indels), têm sido frequentemente utilizados como marcadores moleculares (Batley et al., 2003; Bui e Liu, 2009). Esses polimorfismos são adaptados à genotipagem em larga escala, discriminação de alelos e isolamento de genes utilizando mapeamento genético por ligação ou por associação. No entanto, ensaios para a genotipagem eficiente e robusta são também essenciais para a identificação de mutações que estejam fortemente associadas com o fenótipo (Bui e Liu, 2009).

Muitos métodos têm sido desenvolvidos para genotipagem de SNPs, baseados em diferentes tecnologias como extensão de *primers*, hibridização, ligação, amplificação por PCR, digestão por enzima de restrição e re-sequenciamento. Entretanto, algumas dessas estratégias dependem de equipamentos caros e de difícil manuseio (Bundock et al., 2006).

Metodologias simples e de baixo-custo para a utilização de SNPs como marcadores podem ampliar a utilização desses SNPs no melhoramento assistido e na mineração de alelos.

Uma técnica simples para detecção de SNPs que tem sido muito utilizada em plantas é a clivagem de sequências polimórficas amplificadas (Cleaved Amplified Polymorphic Sequences, CAPS - Konieczny e Ausubel, 1993), a qual requer *primers* loco ou gene específicos para amplificar a região de interesse, seguida de uma clivagem por enzimas de restrição e separação dos fragmentos por eletroforese. No entanto, essa técnica, por utilizar um passo de digestão enzimática, acaba por elevar o custo e o tempo necessário para a análise (Bui e Liu, 2009).

Uma metodologia alternativa é a ARMS-PCR (Amplification Refractory Mutation System PCR; Ye et al., 2001), baseada em um método simples e econômico que envolve uma única reação de PCR seguida de eletroforese em gel. Nessa técnica são utilizados três (Bundock et al., 2006) ou quatro (Ye et al., 2001) *primers* para amplificar um fragmento maior, contendo o SNP, e um ou dois fragmentos menores representando o(s) alelo(s) do loco SNP. Os *primers* são desenhados de tal forma que os fragmentos alelo-específicos e o fragmento comum possuam diferentes tamanhos e possam ser resolvidos por eletroforese em gel de agarose. Harjes et al. (2008) desenvolveram marcadores ARMS-PCR para quatro polimorfismos do gene *lycopene epsilon cyclase* (*lcyE*), que foram responsáveis por 58% da variação presente na via biossintética da vitamina A em milho. Essa técnica foi utilizada também em trigo, para detectar SNPs em cinco locos distintos entre diferentes cultivares (Chiaparino et al., 2004).

No presente trabalho, foram gerados marcadores moleculares de baixo-custo, utilizando a técnica ARMS-PCR, para polimorfismos altamente associados à tolerância ao alumínio em sorgo, com base no gene *Alt_{SB}*. Estes marcadores foram utilizados na mineração alélica em painéis de sorgo para a identificação de haplótipos do gene *Alt_{SB}* no germoplasma alvo. Finalmente, os acessos de sorgo contendo esses haplótipos foram caracterizados ao nível fisiológico e molecular para verificar a possível ocorrência de diversidade haplotípica funcional em sorgo.

2. Objetivos

2.1. Objetivo Principal

Esse trabalho foi desenvolvido com o objetivo de gerar marcadores haplótipo-específicos de baixo custo para polimorfismos associados ao fenótipo de tolerância ao alumínio em sorgo, utilizá-los na mineração de alelos em painéis de ampla diversidade e caracterizar os haplótipos identificados.

2.2. Objetivos Específicos

- Construir marcadores haplótipo-específicos que sejam capazes de identificar acessos de sorgo altamente tolerantes ao Al;
- Validar os marcadores haplótipo-específicos em três painéis de sorgo de origem diversa, verificando-se, entre outras coisas, a capacidade de recuperação de haplótipos superiores por meio da combinações de marcadores;
- Caracterizar o mecanismo fisiológico e molecular da tolerância ao Al conferido pelos haplótipos identificados.

3. Material e Métodos

3.1. Desenho e Teste dos *Primers*

Foram gerados marcadores para seis polimorfismos SNPs e um indel, localizados na região do loco *Alt_{SB}*, que foram significativamente associados com a tolerância ao AI em sorgo segundo Caniato et al. (2009). Para os polimorfismos do tipo indel foram desenhados sistemas de marcadores dominantes, compostos por um par de *primers*, no qual um dos *primers* corresponde ao fragmento de inserção (Figura 1), sendo que a presença de banda representou a inserção e a sua ausência, a deleção. As condições para o desenhos dos *primers* foram: tamanho do fragmento gerado entre 80 e 1000 pb; tamanho dos *primers* entre 18 e 25 bases; Tm (*melting temperature*, temperatura de fusão) de 50°C a 63°C e diferença máxima de 5°C entre as Tm dos *primers*; conteúdo de CG entre 20-80%. Todos os *primers* foram gerados utilizando o programa BatchPrimer3 (You et al., 2008), disponível em <http://wheat.pw.usda.gov/demos/BatchPrimer3/>.

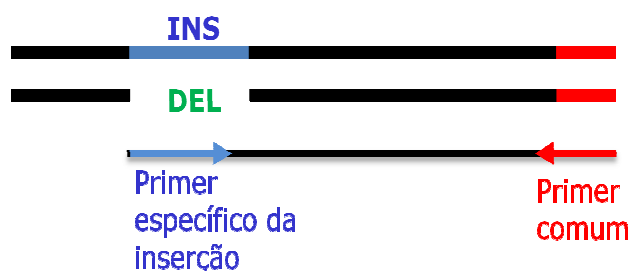


Figura 2 - Esquema de amplificação dos *primers* para os polimorfismos do tipo indel. As setas representam os *primers* e as cores, azul e vermelho, os sítios de anelamento. INS= Inserção. DEL=deleção. O *primer forward* é específico para inserção.

Para o desenho dos marcadores do tipo SNP foi adotada a metodologia ARMS-PCR (Ye et al., 2001). Nessa técnica, dois *primers* genéricos são desenhados nas orientações *forward* e *reverse*, de forma a amplificar a região que contém o SNP e, no sistema com quatro *primers*, outros dois *primers* internos, alelo-específicos, são desenhados. Os fragmentos alelo-específicos são produzidos pela combinação de um *primer* genérico e um primer alelo-específico. Os *primers* foram desenhados de forma que os fragmentos alelo-específicos possuíssem tamanhos diferentes, para que pudessem ser facilmente resolvidos por eletroforese em gel de agarose 1,5%. O SNP alvo fica localizado na última base 3' do sítio de anelamento dos *primers* internos e, para aumentar a especificidade da reação, foi adicionado um segundo pareamento imperfeito (*mismatch*) na antepenúltima base da extremidade 3' dos mesmos (Figura 3). Quando não foi possível a obtenção de *primers*

internos em orientações opostas que viabilizasse a técnica com quatro *primers*, esses foram desenhados seguindo a mesma orientação. Neste caso, uma reação de PCR foi feita para cada alelo, separadamente, com um sistema de três *primers* (Figura 4).

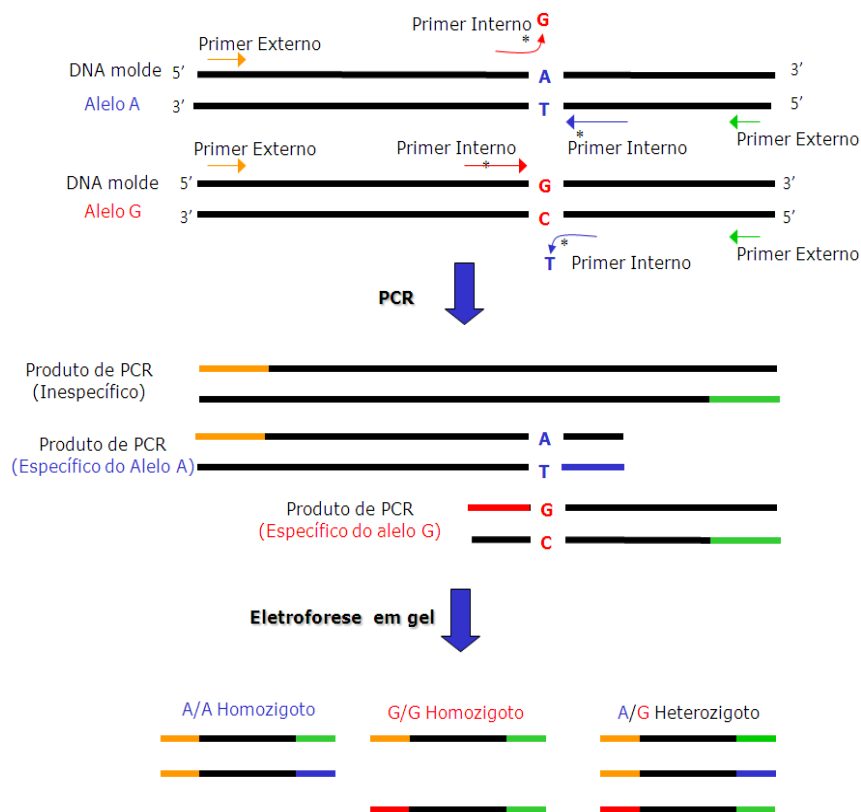


Figura 3 - Ilustração esquemática do desenho de *primers* na técnica ARMS-PCR (adaptado de Ye et al., 2001). Nesse exemplo foi utilizado, a título de exemplificação, um SNP R (G/A). Quatro *primers*, dois alelos-específicos, internos, e dois *primers* externos são utilizados em uma única reação de PCR. Dois produtos alelos específicos, um para o alelo G e outro para o alelo A, são amplificados por dois pares de *primers*. O primeiro corresponde à amplificação do *primer* A alelo-específico com o *primer* externo reverso e o outro com o *primer* G alelo-específico com o *primer* externo *forward*. Um *mismatch*, representado pelo asterisco *, é adicionado na antepenúltima base do final 3' do *primer*. Os dois *primers* externos são desenhados de tal forma que os fragmentos alelo-específicos amplificados possuam tamanhos diferentes e possam ser resolvidos em gel de agarose.

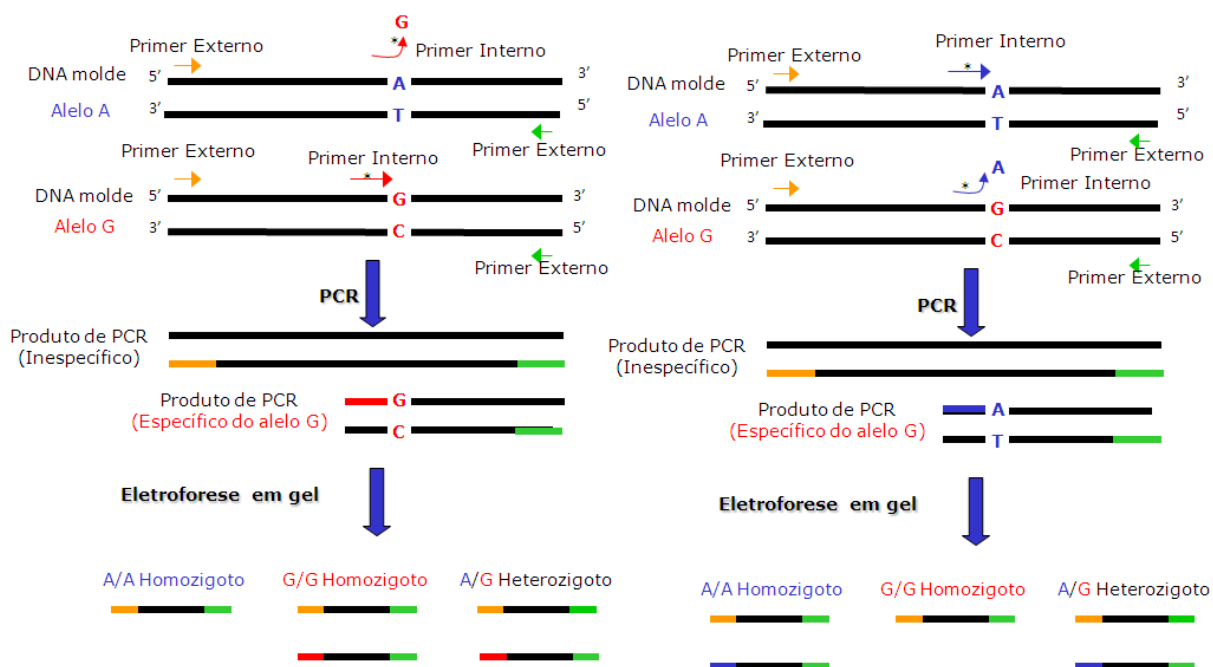


Figura 4 - Sistema ARMS-PCR utilizando três *primers*. Nesse caso, um *primer* alelo-específico, interno, e dois *primers* externos são utilizados em cada reação de PCR, sendo necessárias, portanto, duas reações de PCR separadas para a amplificação dos alelos. Um *mismatch*, representado pelo asterísco *, é adicionado na antepenúltima base do final 3' do *primer* interno, para garantir a especificidade da reação. As setas representam os *primers* e as suas cores o fragmento ao qual eles correspondem.

Os oligonucleotídeos foram desenhados limitando a diferença de tamanho entre os fragmentos internos de 1,1 a 1,5 vezes, mantendo-se o conteúdo GC dos *primers* entre 20 e 80%, com T_m entre 50 e 80°C e diferença entre T_m dos *primers* de no máximo 5°C. Todos os *primers* foram gerados utilizando o programa BatchPrimer3 (You et al., 2008), disponível em <http://wheat.pw.usda.gov/demos/BatchPrimer3/>.

Para validar os sistemas de *primers*, foram utilizadas duas linhagens com os alelos alternativos nos locos marcadores, BR007 e SC283 ou SC566, sensível e tolerantes ao AI, respectivamente. Para o polimorfismo do tipo indel, as reações de PCR foram realizadas em 20 μ L de volume final contendo 30 ng de DNA genômico, 1X tampão Taq, 0,2 mM dNTP, 0,2 μ M de cada *primer* e 1U da enzima *Taq polymerase* (Phoneutria, Belo Horizonte, MG). Para os polimorfismos SNP, as reações de PCR foram realizadas em 13 μ L de volume final contendo 30 ng de DNA genômico, 0,2 μ M de cada um dos *primers* externos, 1 μ M dos *primers* internos e 1X Go Taq® Green Master Mix (Promega, Foster City, CA), com exceção do SNP ORF8_8365pb, onde a concentração dos *primers* internos foi de 1,2 μ M. A amplificação consistiu de um ciclo de desnaturação inicial de 94°C por 4 minutos, seguido de 35 ciclos de 94°C por 1 minuto, 1 minuto para anelamento dos *primers* em temperaturas específicas de para cada conjunto de *primers*, e 72°C por 1 minuto, finalizado por um ciclo de extensão de 72°C por 7 minutos. As temperaturas de anelamento para cada conjunto de

primers estão mostradas na Tabela 3 (Seção Resultados e Discussão) Os produtos de amplificação foram separados por eletroforese em gel de agarose 1,5%, por uma hora a 110V. A seguir, o gel foi incubado em solução de brometo de etídio (1 µg/mL) por 15 minutos, sendo visualizado sob luz ultravioleta e fotografado em equipamento Kodak Gel Logic 200 Imaging System (Carestream Health, New Haven, CT).

3.2. Estratégia de *Pooling*

A estratégia de *pooling* é uma maneira eficiente de se genotipar uma grande quantidade de indivíduos em busca de alelos de baixa frequência, como é o caso dos alelos-alvo do gene *Alt_{SB}*, minimizando-se o número de reações de PCR. Para se definir uma estratégia de *pooling* que seja eficiente para o uso com os marcadores do loco *Alt_{SB}*, seguiu-se a metodologia de Amos et al. (2000), revisada por Gastwirth (2000), descrita a seguir, de forma a determinar o tamanho ideal do *pool* a ser utilizado na genotipagem de populações.

Inicialmente, foi determinada a sensibilidade dos marcadores, ou seja, a capacidade desses em identificar, em um *pool* de DNA com diferentes proporções de indivíduos com o alelo alvo e o alelo alternativo, a presença de pelo menos um indivíduo com o alelo-alvo. A sensibilidade foi utilizada na determinação do tamanho ideal do *pool*. A eficiência do método de *pooling* (1-F) foi calculada a partir da frequência do alelo favorável. Supondo que a frequência do alelo de interesse é π e que n indivíduos são testados, para *pools* de tamanho s , a probabilidade γ , de que pelo menos um membro do *pool* possua o alelo é:

$$\gamma = 1 - (1 - \pi)^s$$

O modelo assume que o teste possui 100% de acurácia na classificação de um *pool* como tendo ou não o alelo de interesse.

O número esperado de reações necessárias para se identificar completamente todas as mutações em determinada população é calculado por:

$$E(Y) = \frac{n}{s} + n\gamma$$

Sendo Y a quantidade de reações necessárias para se genotipar toda uma população para um determinado polimorfismo, essas fórmulas calculam a $E(Y)$, que é um número estimado de reações. Se a razão $Y:n$ for menor que 1, a estratégia de *pooling* é eficiente. Esta razão é denominada de F (fração de testes requeridos relativos aos testes individuais) e a eficiência de um método de *pooling* é dada por 1-F, que representa a fração de testes que é economizada.

Essas fórmulas foram utilizadas para se calcular o tamanho ideal dos *pools* em populações compostas por 100, 150, 200 e 400 indivíduos, levando em consideração as frequências dos alelos de interesse encontradas por Caniato (2009).

3.3. Mineração de alelos do loco *Alt_{SB}*

3.3.1. Material Genético

- Germoplasma ICRISAT/CIRAD

Esse germoplasma consiste de 209 acessos mantidos pelo *Centre de Cooperation Internationale en Recherche Agronomique pour le Développement* (CIRAD, França) e pelo *International Crops Research Institute for the Semi-Arid Tropics* (ICRISAT, Patancheru - Índia). Esse germoplasma é comumente utilizado por agricultores na África e na Ásia e por melhoristas do ICRISAT e do CIRAD (Deu et al., 2006). Adicionalmente, 45 linhagens utilizadas pelos programas de melhoramento da Embrapa Milho e Sorgo e dos Estados Unidos foram incorporadas a esse grupo. Essa coleção constitui uma amostra representativa da diversidade genética da espécie e foi elaborada com base na classificação racial e na origem geográfica, em função da resposta ao fotoperiodismo e também por padrões de diversidade avaliados com marcadores microssatélite (Deu et al. 2006).

- Germoplasma IGD

Essa coleção, mantida pelo *Institute for Genomic Diversity* (IGD, EUA), é constituída de 377 acessos, compostos por 228 linhagens convertidas para insensibilidade ao fotoperiodismo e porte baixo (*Sorghum Converted*, SC) e 149 linhagens de melhoramento. Esse germoplasma foi elaborado com base na diversidade geográfica e genotípica, tentando garantir a inclusão de materiais provenientes da maioria das regiões ecogeográficas da África, raças cultivadas e de importância histórica em programas de melhoramento da África e dos Estados Unidos, incluindo linhagens que contribuíram significativamente para o melhoramento genético de sorgo nos EUA como fontes de resistência e tolerância a estresses bióticos e abióticos (Casa et al., 2008).

- Germoplasma INRAN

Esse germoplasma é proveniente do *Institut National de la Recherche Agronomique du Niger* (INRAN, Niger) e contém 164 acessos, sendo 142 *landraces* e 22 linhagens de melhoramento. Esse material foi coletado em 79 vilas distribuídas por toda a extensão do Niger, compreendendo todas as regiões agro-ecológicas desse país, e é composto pelos mais diversos materiais.

3.3.2. Isolamento do DNA genômico

Sementes de sorgo de cada acesso foram germinadas em câmara de crescimento e transplantadas para bandejas de isopor contendo substrato em casa de vegetação. Para cada acesso foram coletados discos foliares de cinco plantas de cada acesso com 2 a 3 semanas de idade, que foram utilizados para o isolamento de DNA genômico segundo o método descrito por Saghai-Marooft et al. (1984), com modificações. Aproximadamente 500 mg do material vegetal foram triturados no equipamento Geno/Grinder 2000 (BT&C, Bridgewater, NJ), onde foram adicionados 600 µL de tampão CTAB [2 % (m/v) CTAB; 0,2 M Tris- HCl (pH 7,5); 1,4 M NaCl; 0,02 M EDTA (pH 8,0); 2 % (v/v) 2-mercaptoetanol]. A mistura foi mantida em banho-maria a 65°C durante uma hora, com homogeneização a cada 15 minutos. Em seguida, promoveu-se a lavagem com igual volume de solução de clorofórmio-octanol (24:1, v/v), com homogeneização constante por 20 minutos. O material foi centrifugado a 16.000 x g por 10 minutos e o sobrenadante transferido para novo microtubo, onde foram adicionados 500 µL de isopropanol mantido à 20°C. Os microtubos foram centrifugados a 16.000 x g por 10 minutos, o sobrenadante descartado e o precipitado lavado com 200 µL de etanol 70 % (v/v) gelado. Os tubos foram novamente centrifugados por 5 minutos a 16.000 x g, o sobrenadante foi descartado e o etanol residual retirado em centrífuga a vácuo por cinco minutos. Os precipitados foram ressuspensos em 100 µL de tampão TE contendo RNase A (10 mM Tris-HCl; 1 mM EDTA, pH 8,0; 0,1 µg/µL RNase A). As amostras foram quantificadas no NanoDrop 3000 (Thermo Fisher Scientific, Worcester, MA) e em gel de agarose 0,8 % (m/v) em tampão TAE (40 mM Tris-acetato; 1 mM EDTA, pH 8,0), comparando-se com um padrão de DNA de concentração conhecida. Após a eletroforese realizada a 100 V durante uma hora, o gel foi incubado em solução de brometo de etídio (1 µg/mL) por 15 minutos, visualizado sob luz ultravioleta e fotografado em equipamento Kodak Gel Logic 1500 Imaging System (Carestream Health, New Haven, CT). O DNA estoque foi diluído em água ultrapura para a concentração de 10 ng/µL e armazenado a -20°C.

3.3.3. Análise da Tolerância ao Alumínio em Solução Nutritiva

- Germoplasma IGD

As 35 linhagens selecionadas do painel IGD, sendo 30 que possuíam um dos seis haplótipos considerados candidatos à tolerante, ou seja, haplótipos semelhantes ao das linhagens tolerantes, e 5 controles, foram avaliadas quanto ao crescimento da raiz seminal em solução nutritiva contendo Al nas concentrações de 0 e 148 µM que correspondem às

atividades livres de Al de {0} e {27} μM de Al, a pH 4,0 (estimativas de atividade livre obtidas com o software de especiação GEOCHEM-PC, Parker et al., 1995). As sementes de cada uma das linhagens foram esterilizadas com hipoclorito de sódio 0,525% por 10 minutos sob agitação constante e enxaguadas oito vezes com água destilada. A seguir, procedeu-se a germinação em rolos de papel de germinação umedecidos em água deionizada por um período de três dias em câmara de germinação com temperatura diurna média de $27\pm 3^\circ\text{C}$, noturna de $20\pm 3^\circ\text{C}$ e fotoperíodo de 12 horas. As plântulas foram então transferidas para copos plásticos perfurados, acomodados em placas de PVC dentro de bandejas plásticas com capacidade para 8 litros de solução nutritiva completa (Maganavaca et al., 1987) sem alumínio por um período de 24 horas de aclimação, quando o comprimento radicular inicial (C_{0d}) foi obtido. A solução nutritiva foi então trocada por solução contendo Al nas atividades de {0} e {27} μM Al. O comprimento radicular foi novamente obtido após 3 e 5 dias de crescimento em solução nutritiva contendo {0} e {27} μM Al^{3+} , (C_{3d} e C_{5d}). O crescimento radicular líquido (CL) após 3 e 5 dias foi calculado pela diferença de valores entre comprimento radicular obtidos após 3 e 5 dias e comprimento radicular inicial, $\text{CL}_{3d} = C_{3d} - C_{0d}$ e $\text{CL}_{5d} = C_{5d} - C_{0d}$, respectivamente. Para cada tratamento foram avaliadas duas repetições de sete plantas, distribuídas de acordo com o delineamento inteiramente casualizado, em câmara de crescimento, em condições similares às aquelas utilizadas na germinação das sementes. Para cada tratamento com Al as médias de crescimento líquido radicular, $\text{CL}_{3d+\text{Al}}$ e $\text{CL}_{5d+\text{Al}}$, de cada uma das duas repetições foram divididas pela média global do crescimento líquido radicular sem Al, $\text{CL}_{3d-\text{Al}}$ e $\text{CL}_{5d-\text{Al}}$, sendo assim gerados dois valores percentuais de crescimento líquido relativo, CLR_{3d} e CLR_{5d} .

- Germoplasma INRAN

Dos 16 acessos selecionados do painel INRAN, 10 possuíam um dos dois haplótipos considerados candidatos a fontes de tolerância, por possuírem haplótipos semelhantes aos de linhagens tolerantes, e 6 destes eram acessos do programa de melhoramento do Niger, além de três controles da Embrapa. Os genótipos foram avaliados quanto ao crescimento da raiz seminal em solução nutritiva contendo Al na concentração de atividades livres de Al de {27} μM , a pH 4,0. Devido à natureza heterogênea desses materiais, o controle sem Al não foi utilizado. As condições de germinação das sementes e aclimação foram semelhantes às utilizadas para as linhagens do painel IGD. Após o período de 24 horas de aclimação foi obtido o comprimento radicular inicial (C_{0d}). A solução nutritiva foi então trocada por solução contendo Al na atividade {27} μM Al. O comprimento radicular foi novamente obtido após 3 e 5 dias de crescimento em solução nutritiva contendo Al (C_{3d} e C_{5d}). O crescimento radicular líquido (CL) após 3 e 5 dias foi calculado pela diferença de valores entre

comprimento radicular obtidos após 3 e 5 dias e comprimento radicular inicial, $CL_{3d} = C_{3d} - C_{0d}$ e $CL_{5d} = C_{5d} - C_{0d}$, respectivamente. Para cada tratamento o crescimento líquido radicular de cada plântula, CL_{3d+Al} e CL_{5d+Al} , foi dividido pelo C_{0d} , sendo assim gerados dois valores percentuais de crescimento radicular relativo, CRR_{3d} e CRR_{5d} , $CRR_{3d} = CL_{3d}/C_{0d}$ e $CRR_{5d} = CL_{5d}/C_{0d}$. Novamente, para cada tratamento foram avaliadas duas repetições de sete plantas, distribuídas de acordo com o delineamento inteiramente casualizado, em câmara de crescimento, em condições similares às da germinação das sementes.

Notas para Injúria Radicular Visual (IRV) foram atribuídas após 5 dias de exposição ao Al de acordo com uma escala variando de 1 a 5, sendo 1) ausência de raízes secundárias, raiz principal escurecida, ápice radicular translúcido, ligeiramente curvada e muito danificada. 2) ausência de raízes secundárias, presença de pequenos pontos pretos na raiz principal, ápice radicular inchado, danificado e ligeiramente amarronzado. 3) ausência de raízes secundárias, raiz principal com pontos pretos ocasionais, ápice radicular branco, ligeiramente inchado e danificado. 4) ausência de raízes secundárias, raiz principal branca, ápice radicular não danificado. 5) presença abundante de raízes secundárias, raiz principal branca, ápice radicular não danificado.

3.3.4. Genotipagem dos Painéis com Marcadores do Loco *Alt_{SB}*

A genotipagem dos painéis foi realizada utilizando os marcadores desenvolvidos para os polimorfismos do loco *Alt_{SB}* nas mesmas condições descritas na seção anterior. Para o painel INRAN foram montados *pools* de quatro indivíduos, contendo 100 ng de DNA genômico de cada indivíduo, que foram adicionados a um microtubo e misturados para homogeneização do material. Os *pools* foram submetidos a reações de PCR com os marcadores, sendo que para os SNPs genotipados com o sistema ARMS-PCR de três *primers* foram feitas reações somente para os alelos favoráveis. Os *pools* que apresentaram o alelo favorável foram, então, desmembrados e seus indivíduos genotipados individualmente para cada um dos marcadores.

3.4. Análise dos Haplótipos do Loco *Alt_{SB}*

3.4.1. Análise da expressão do gene *SbMATE*

O centímetro inicial do ápice radicular de 28 plantas de cada linhagem selecionada, cultivadas em solução nutritiva contendo Al na concentração de {27} μM de Al, a pH 4,0, foram coletados cinco dias após a exposição ao Al.

As amostras de RNA foram isoladas com o “RNeasy Plant Mini Kit” (Qiagen, Valencia, CA), segundo as recomendações do fabricante. Ao final do procedimento, as amostras foram eluídas em 35 µL de água tratada com DEPC. Uma alíquota de 5 µL do RNA foi diluída com 4 µL de água DEPC e tratada com 10 U da enzima DNase I (Amersham Pharmacia Biotech, Piscataway, NJ) por 15 minutos à temperatura ambiente. Para a quantificação do RNA total foram utilizados 2 µL de cada amostra acrescidos de 398 µL de água DEPC. A pureza e a concentração do RNA foram determinadas pela leitura das amostras em 260 e 280 nm, utilizando-se o NanoDrop 3000 (Thermo Fisher Scientific, Worcester, MA). A integridade das amostras foi verificada por meio de eletroforese em gel de agarose 1 % (m/v) contendo 0,25 µg/mL de brometo de etídio. Após a eletroforese realizada a 200 V durante 30 minutos, o gel foi visualizado sob luz ultravioleta e fotografado em equipamento Kodak Gel Logic 200 Imaging System (Carestream Health, New Haven, CT). Para a síntese do DNA complementar foi utilizado o kit *High Capacity RNA-to-cDNA* (Applied Biosystem, Foster City, CA), utilizando os ciclos de amplificação de 25°C por 10 minutos, 27°C por 120 minutos e 85°C por 5 minutos.

O ensaio de expressão foi realizado no equipamento 7500 Faster Real-Time PCR System (Applied Biosystem, Foster City, CA) utilizando o kit TaqMan Gene Expression (Applied Biosystem, Foster City, CA) desenhado para avaliar a expressão do gene *SbMATE*, utilizando o gene 18S como controle endógeno como descrito por Magalhães et al. (2007).

3.4.2. Análise de acúmulo de alumínio nos ápices radiculares

Vinte e uma plantas de cada linhagem selecionada foram cultivadas em solução nutritiva contendo Al na concentração de {27} µM, em pH 4,0, por um período de cinco dias. As raízes das plantas intactas foram lavadas em oito litros de água ultra-pura (18 mΩ) por 15 minutos, sob aeração e, depois de trocada a água, permaneceram por mais 25 minutos sob as mesmas condições. As lavagens foram realizadas para remover o Al presente na solução externa às raízes. O primeiro centímetro dos ápices radiculares foi cortado com um bisturi, seguindo-se de secagem a 55°C, *overnight*, em estufa. O peso seco dos ápices radiculares foi determinado utilizando uma microbalança XP2U (Mettler, Greifensee, Suíça). As amostras secas foram digeridas com 100 µL de ácido perclórico 70%, ressuspensas em 2 mL de ácido nítrico e analisadas quanto ao conteúdo de Al na raiz por espectrometria de plasma (*inductively-coupled argon plasma*, ICP) na universidade de Cornell (Ithaca, NY, EUA).

3.4.3. Análise da exsudação de citrato

Sementes dos acessos selecionados foram esterilizadas com hipoclorito de sódio 0,525% por 10 minutos sob agitação constantes e enxaguadas oito vezes com água destilada. A seguir, foram germinadas em rolos de papel de germinação umedecidos em água deionizada por um período de três dias em câmara de crescimento com temperatura diurna média de $27\pm 3^{\circ}\text{C}$, noturna de $20\pm^{\circ}\text{C}$ e fotoperíodo de 12 horas. As plântulas foram então transferidas para copos plásticos perfurados, acomodados em placas de PVC dentro de bandejas plásticas com capacidade para 8 litros de solução nutritiva completa (Maganavaca et al., 1987) sem alumínio por um período de 24 horas de aclimação. A solução nutritiva foi então trocada por solução contendo Al nas atividades de {0} e {27} μM Al e os exsudados foram coletados depois de 5 dias sob tratamento de {0} e {27} μM Al. As parcelas experimentais foram compostas de sete plântulas, utilizando-se três repetições para cada tratamento. A coleta de exsudatos radiculares começou pela transferência de sete plântulas para tubos Falcon de 50 mL contendo solução de $\text{CaCl}_2 \cdot 6\text{H}_2\text{O}$ (pH 4,5), com ou sem $\text{AlCl}_3 \cdot 6\text{H}_2\text{O}$ e, no último caso, com a mesma atividade de {27} μM de Al utilizada na solução anterior. Os exsudados foram coletados por 6 horas e a solução de exsudação resultante foi passada por uma coluna cromatográfica aniônica de prata OnGuard® II Ag (Dionex, Sunnyvale, CA) e tratada com a resina catiônica Dowex® 50WX8 (Sigma Aldrich St. Louis, MO) para remoção do alumínio e outros ânions que pudessem interferir na determinação dos ácidos orgânicos. Em seguida, 1 mL dessas amostras foram liofilizados e ressuspendidos em 0,175 mL de água ultra-pura. A análise de ácidos orgânicos foi realizada por eletroforese capilar, num sistema como descrito por Piñeros et al. (2002).

4. Resultados e Discussão

4.1. Geração de Marcadores Haplótipos-específicos

Dentre os polimorfismos associados com a tolerância ao AI no loco Alt_{SB} encontrados por Caniato (2009), sete foram utilizados para o desenho de marcadores haplótipo-específicos, sendo quatro SNPs localizados no gene *SbMATE* (um no primeiro exon, E1, e três no segundo intron, I2), dois SNPs localizados na ORF8 e um indel de 19 pb na ORF9 (Figura 5, Tabela 1).

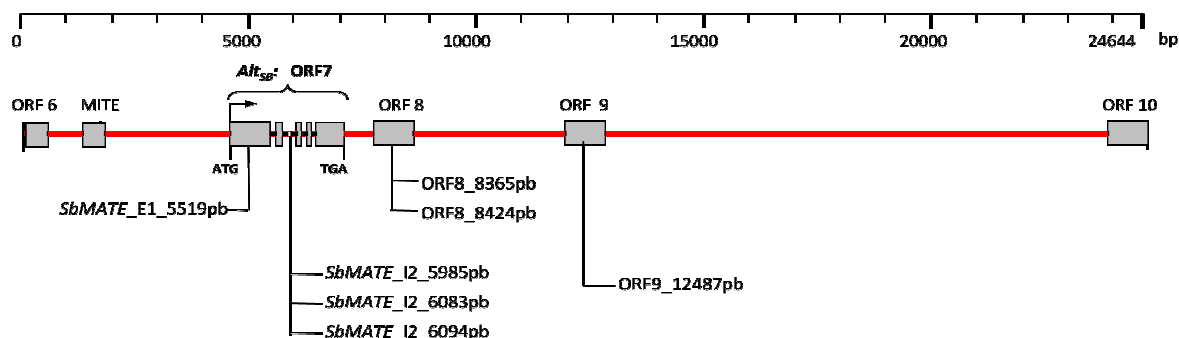


Figura 5 - Representação do loco Alt_{SB} com as seis regiões amostradas para o mapeamento associativo com as respectivas posições em pares de bases (pb). Estão representados o gene *SbMATE* (ORF7, exons representados por caixas cinza e introns por linhas pretas), ORF8 e ORF9.

Tabela 1 – Polimorfismos do loco Alt_{SB} . Para cada polimorfismo estão representados o tipo de polimorfismo, indel ou SNP, com o alelo favorável indicado em **negrito**, a região em que o polimorfismo está localizado (vide Figura 5), o valor de significância de associação do polimorfismo à característica (p-valor) e a porcentagem da variância fenotípica explicada para CLR_{5d} (r^2). Fonte: Caniato et al. (2009).

Polimorfismo	Tipo do Polimorfismo	Região	p-valor	r^2
<i>SbMATE_EX1_5519</i>	SNP W (A/T)	<i>SbMATE</i> Primeiro exon	-	-
<i>SbMATE_I2_5985pb</i>	SNP R (A/G)	<i>SbMATE</i> Segundo intron	5.57E-08	7,01
<i>SbMATE_I2_6083pb</i>	SNP M (A/C)	<i>SbMATE</i> Segundo intron	4.69E-16	14,34
<i>SbMATE_I2_6094pb</i>	SNP S (C/G)	<i>SbMATE</i> Segundo intron	1.75E-11	10,34
ORF8_8365pb	SNP M (A/C)	ORF8	4.45E-10	9,97
ORF8_8424pb	SNP K (G/T)	ORF8	3.09E-10	9,45
ORF9_12487pb	Indel 19 pb (19/0)	ORF9	4.40E-12	10,94

Para os seis polimorfismos do tipo SNP buscou-se desenvolver sistemas ARMS-PCR com quatro *primers*, com o objetivo de obter marcadores co-dominantes cujos alelos pudessem ser resolvidos simultaneamente em uma única reação. No entanto, isso só foi

possível para os polimorfismos *SbMATE_I2_6083pb*, *ORF8_8365pb* e *ORF8_8424pb*, sendo que para os demais SNPs foi utilizado o sistema ARMS-PCR com três *primers* para cada um dos alelos. Nestes casos, não foi possível obter *primers* internos, específicos para cada um dos alelos nas duas orientações, devido a fatores como conteúdo de CG nas regiões que flanqueiam esses SNPs, falta de estabilidade dos *primers* gerados, formação de dímeros entre os *primers*, dentre outros. Sempre que possível, foram desenhados mais de um sistema de *primers* para cada polimorfismo, para que fosse possível avaliar mais de uma combinação entre os *primers* e garantir o sucesso na geração dos marcadores. Assim, para *ORF9_12487pb* foram sintetizados dois conjuntos de *primers*, para *SbMATE_I2_5985pb* foram sintetizados três e para *SbMATE_I2_6083pb* dois, sendo avaliadas até oito combinações entre os *primers* dos conjuntos. Para os outros polimorfismos, foi possível obter apenas um conjunto. Pelo menos um sistema de marcadores (Tabela 2) funcionou satisfatoriamente para cada polimorfismo.

Tabela 2 – Descrição dos marcadores haplótipo-específicos para o loco *Alt_{SB}*. Para cada polimorfismo está indicado o tipo do polimorfismo de acordo com a nomenclatura IUPAC, o tipo de sistema ARMS-PCR que foi desenvolvido para cada loco e a combinação de *primers* para amplificar cada fragmento específico com os tamanhos de fragmentos em pares de bases (pb).

Polimorfismo	Tipo do Polimorfismo	Sistema ARMS-PCR	Marcadores Haplótipo-Específicos	
			<i>Primers</i>	Fragmentos
<i>SbMATE_EX1_5519</i>	SNP W (A/T)	três <i>primers</i>		Fragmento inespecífico 709 pb
			HB48+HB49+HB50	Fragmento específico alelo A- 230 pb
			HB48+HB49+HB51	Fragmento específico alelo T - 230 pb (alelo favorável)
<i>SBMATE_I2_5985pb</i>	SNP R (A/G)	três <i>primers</i>		Fragmento inespecífico 700 pb
			HB23+HB24+HB26	Fragmento específico alelo G - 361 pb
			HB23+HB24+HB27	Fragmento específico alelo A- 361 pb (alelo favorável)
<i>SbMATE_I2_6083pb</i>	SNP M (A/C)	quatro <i>primers</i>		Fragmento inespecífico 700 pb
			HB23+HB24+HB30+HB37	Fragmento específico alelo A - 240 pb (alelo favorável)
				Fragmento específico alelo C - 484 pb
<i>SbMATE_I2_6094pb</i>	SNP S (C/G)	três <i>primers</i>		Fragmento inespecífico 700 pb
			HB23+HB24+HB32	Fragmento específico alelo G - 254 pb
			HB23+HB24+HB33	Fragmento específico alelo C - 254 pb (alelo favorável)
<i>ORF8_8365pb</i>	SNP M (A/C)	quatro <i>primers</i>		Fragmento inespecífico 1666 pb
			HB56+HB57+HB58+HB59	Fragmento específico alelo A - 496 pb
				Fragmento específico alelo C - 1116 pb (alelo favorável)
<i>ORF8_8424pb</i>	SNP K (G/T)	quatro <i>primers</i>		Fragmento inespecífico 615 pb
			HB52+HB53+HB54+HB55	Fragmento específico alelo T - 496 pb
				Fragmento específico alelo G - 166 pb (alelo favorável)
<i>ORF9_12487pb</i>	Indel 19 pb	-	HB1+HB2	Fragmento específico inserção – 685 pb (alelo desfavorável)

Para o polimorfismo *SbMATE_E1_5519pb*, um SNP W (T/A), foi gerado um sistema ARMS-PCR com três *primers*. Um alelo extremamente raro está presente especificamente na linhagem SC566, que é uma das mais tolerantes ao AI entre as linhagens avaliadas por Caniato et al. (2007) e, portanto, uma fonte importante para programas de melhoramento. Assim, tais marcadores serão muito úteis para a seleção assistida por marcadores (SAM). O esquema de amplificação desse marcador, assim como o resultado do teste de amplificação para esse polimorfismo estão representados na Figura 6.

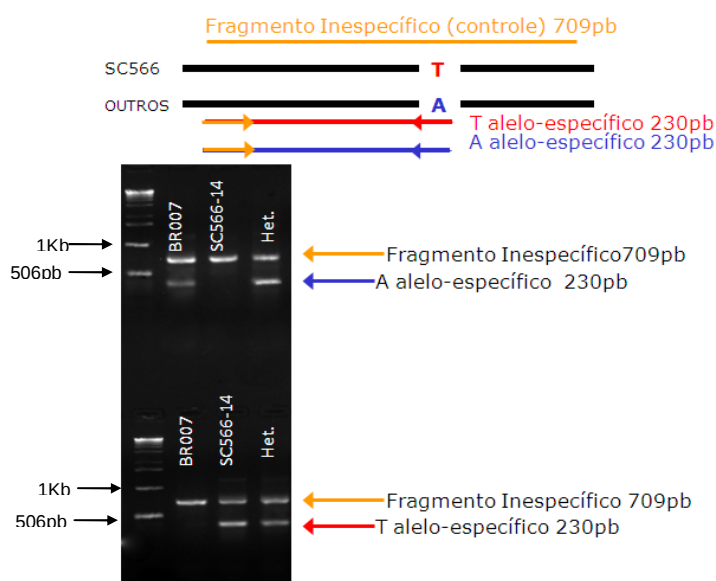


Figura 6 - Esquema de amplificação do marcador *SbMATE_E1_5519pb*. Os padrões de amplificação são apresentados para dois indivíduos contendo os haplótipos contrastantes, SC566 (tolerante ao AI) e BR007 (sensível), e para o indivíduo heterozigoto. Foram feitas duas reações diferentes, sendo uma para cada alelo. Na primeira canaleta está o marcador de tamanho molecular.

Para o polimorfismo *SbMATE_I2_5985pb* do tipo R (G/A) foram gerados três conjuntos de marcadores, que possibilitaram até oito diferentes combinações de *primers*, sendo que apenas uma combinação no sistema de três *primers* apresentou resultado satisfatório (Figura 7). Para esse marcador, duas reações são feitas separadamente, uma para cada alelo, obtendo-se um sistema de marcadores co-dominante.

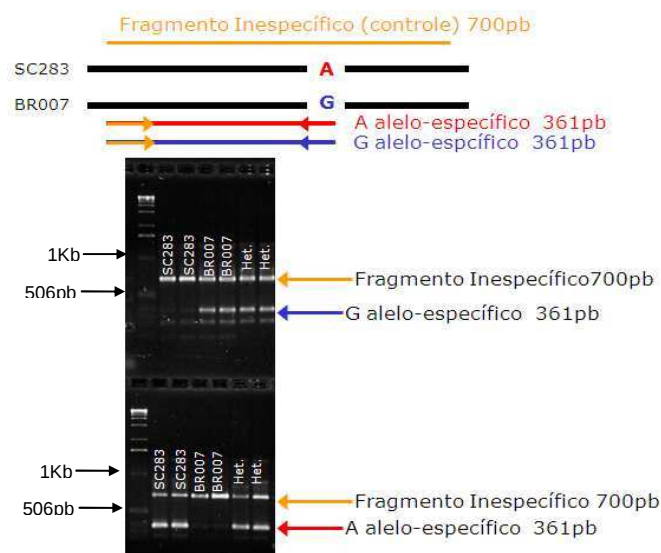
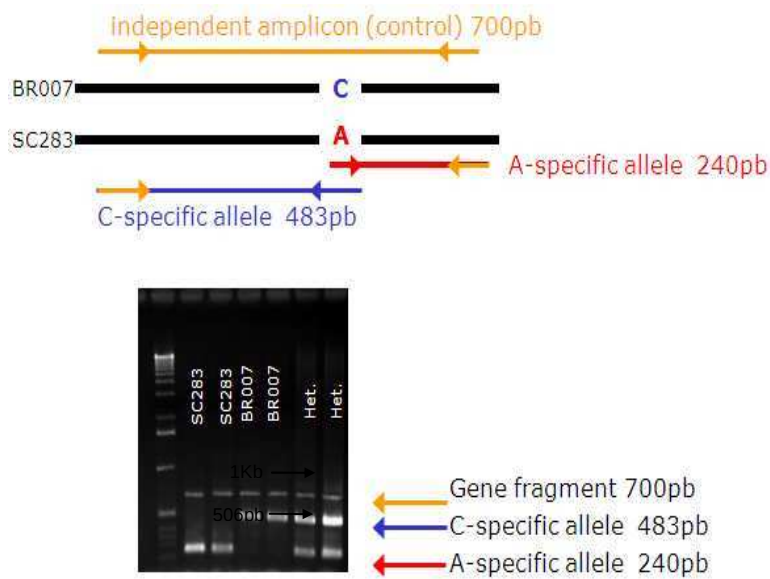


Figura 7 – Esquema de amplificação do marcador *SbMATE_I2_5985pb*. Os padrões de amplificação são apresentados para dois indivíduos contendo os haplótipos contrastantes, SC283 (tolerante ao AI) e BR007 (sensível), e para o indivíduo heterozigoto. Foram feitas duas reações diferentes, sendo uma para cada alelo. Na primeira canaleta está o marcador de tamanho molecular.

O loco *SbMATE_I2_6083pb* possui o polimorfismo que explica a maior porção da variação fenotípica para a característica tolerância ao AI, sugerindo a possibilidade de um efeito funcional para esse polimorfismo. Além disso, esse polimorfismo associado ao alelo favorável possui frequência tão baixa quanto a da tolerância ao AI, o que pode ser um diferencial na mineração dos alelos de interesse, uma vez que poderá resultar na seleção mais precisa de acessos altamente tolerantes. Entretanto, embora o loco *SbMATE_I2_6083pb* possua um efeito diferenciado na característica de tolerância, não se pode afirmar que seja este o polimorfismo causativo. Isto ocorre devido à extensão do desequilíbrio de ligação existente entre esse e os outros polimorfismos da região, que impossibilita a distinção entre um SNP em forte desequilíbrio de ligação com o QTN, mascarando o resultado da associação. Foram desenhados dois conjuntos ARMS-PCR para essa mutação, sendo que somente um sistema ARMS-PCR de quatro *primers* co-dominante funcionou adequadamente (Figura 8).



Legenda em inglês-alterar!

Figura 8 - Esquema de amplificação do marcador *SbMATE_I2_6083pb*. Os padrões de amplificação são apresentados para dois indivíduos com haplótipos contrastantes, SC283 (tolerante ao Al) e BR007 (sensível) e para o indivíduo heterozigoto. Na primeira canaleta está o marcador de tamanho molecular.

O polimorfismo *SbMATE_I2_6094pb* é um SNP S (G/C) localizado a apenas 11 pb do SNP anterior. A região 3' subsequente a esse SNP é pobre em CG e possui uma sequência poli-A, que não permitiu desenho de *primers* alelo-específicos na direção reversa. Assim, foi desenhado um sistema ARMS-PCR com três *primers*, sendo que em função da proximidade com o loco *SbMATE_I2_6083pb*, os *primers* internos do sistema tiveram que incluir esse segundo polimorfismo (Figura 9). Deste modo, foi necessário uma 'correção de fase' nesses *primers* para que não houvesse *mismatch* em nenhum dos dois SNPs compreendidos no sítio de anelamento dos mesmos.

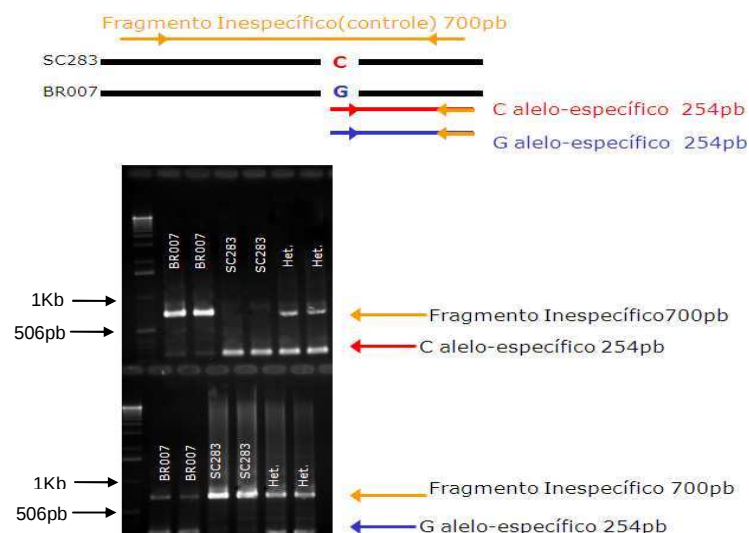


Figura 9 - Esquema de amplificação do marcador *SbMATE_I2_6094pb*. Os padrões de amplificação são apresentados para dois indivíduos contendo os haplótipos contrastantes, SC283 (tolerante ao Al) e BR007 (sensível), e para o indivíduo heterozigoto. Foram feitas duas reações diferentes, sendo uma para cada alelo. Na primeira canaleta está o marcador de tamanho molecular.

Para os polimorfismos da ORF8 foi sintetizado apenas um conjunto ARMS-PCR para cada SNP no esquema de quatro *primers*, cujos perfis de amplificação estão apresentados nas Figuras 10 e 11. O SNP M do loco ORF8_8365pb apresentou instabilidade na amplificação do alelo favorável C. Isso pode ter ocorrido devido ao tamanho do fragmento gerado para esse alelo, que é muito grande (1165 pb), gerando uma preferência de amplificação para o outro alelo. Para corrigir esse problema, a concentração do *primer* que amplifica esse alelo foi aumentada em 20% na reação de PCR, a fim de se favorecer a amplificação do mesmo.

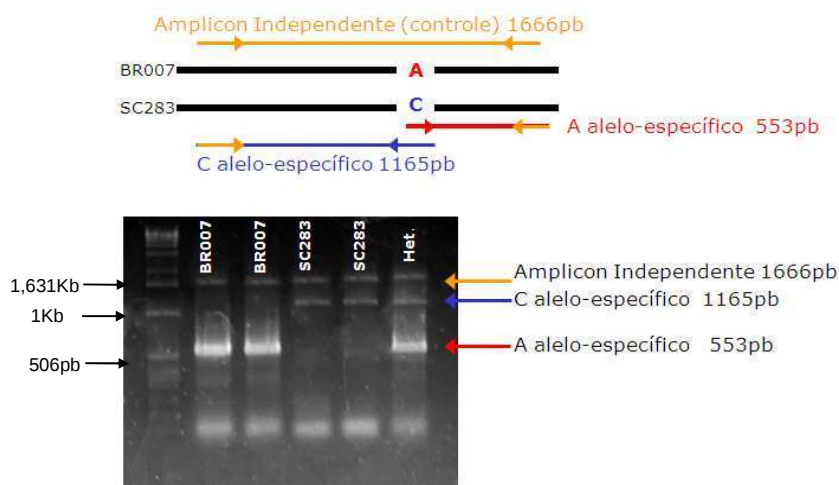


Figura 10 - Esquema de amplificação do marcador ORF8_8365pb. Os padrões de amplificação são apresentados para dois indivíduos com haplótipos contrastantes, SC283 (tolerante ao Al) e BR007 (sensível) e para o indivíduo heterozigoto. Na primeira canaleta está o marcador de tamanho molecular.

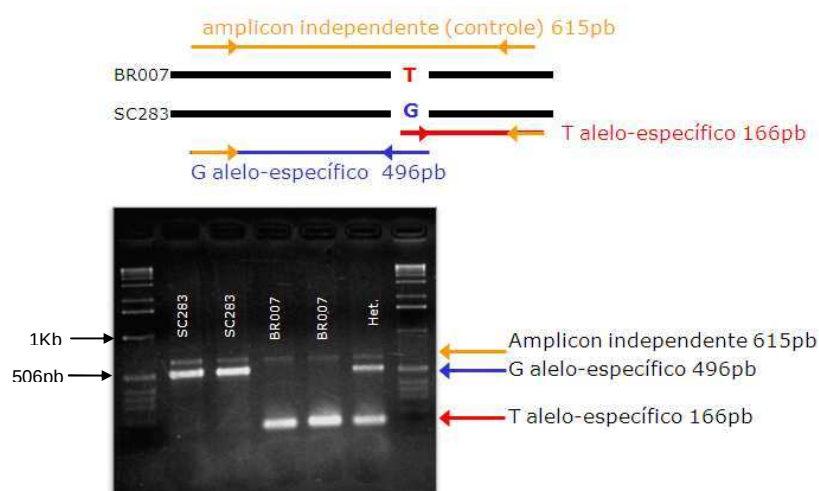


Figura 11 - Esquema de amplificação do marcador ORF8_8424pb. Os padrões de amplificação são apresentados para dois indivíduos com haplótipos contrastantes, SC283 (tolerante ao Al) e BR007 (sensível) e para o indivíduo heterozigoto. Na primeira canaleta está o marcador de tamanho molecular.

O polimorfismo ORF9_12487pb do tipo Indel apresenta uma inserção de 19 pb no alelo não favorável. Para esse polimorfismo foram desenhados dois pares de *primers* contendo a inserção como um dos componentes, um na orientação *forward* e outro na reversa. O par no qual a inserção foi representada pelo *primer forward* apresentou ótimo perfil de amplificação (Figura 12). Já o outro par, não apresentou uma grande especificidade. Como há a presença de apenas uma banda, esse é um marcador dominante.

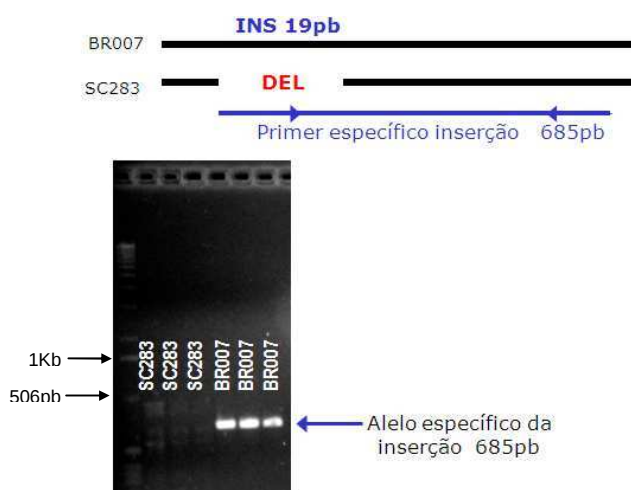


Figura 12 - Esquema de amplificação do marcador ORF9_12487pb. Os padrões de amplificação são apresentados para dois indivíduos com haplótipos contrastantes, SC283 (tolerante ao Al) e BR007 (sensível). Na primeira canaleta está o marcador de tamanho molecular.

Os marcadores ARMS-PCR possuem uma banda inespecífica, que pode servir como um controle da reação. No entanto, essa banda não se mostrou muito estável, principalmente nos sistemas de quatro *primers*. Isso ocorre devido a uma competição por amplificação entre os *primers* internos e externos. Como a estabilidade dos *primers* internos assim como a concentração destes é maior, a amplificação dos fragmentos internos ocorre em detrimento dos fragmentos externos. Para tentar evitar esse problema, foi feita uma tentativa de se diminuir a razão entre as concentrações dos *primers* internos e externos entretanto, essa alteração afetou a estabilidade da reação como um todo e as concentrações iniciais dos *primers* foram mantidas. Contudo, a ausência dessas bandas não prejudicou a visualização dos polimorfismos de interesse. Casos como esse foram descritos também por Chiaparino et al. (2004), no desenvolvimento de marcadores ARMS-PCR para trigo e por Ye et al. (2001), em humanos. Além disso, Bui e Liu (2009) relataram que os *primers* internos que terminam na extremidade 3' com C ou G apresentam, muitas vezes, bandas fantasmas nos indivíduos que portam o alelo alternativo e outras bandas de diferentes tamanhos. Nos marcadores aqui desenvolvidos, algumas dessas bandas fantasmas persistiram mesmo após a tentativa de aumento da temperatura de anelamento. Contudo, essas bandas foram muito fracas e não interferiram na identificação das bandas principais. A descrição completa dos marcadores está descrita na Tabela 3.

Tabela 3 – Descrição completa dos sistemas de marcadores, indicando o tipo do polimorfismo, o sistema ARMS-PCR em que foi baseado o marcador, a temperatura de anelamento para cada conjunto o identificador dos *primers* (Id), o tipo do *primer*, a temperatura de dissociação (T_m , *melting temperature*) assim como a sequência de cada *primer*. Os alelos favoráveis estão destacados em negrito.

Polimorfismo	Tipo do Polimorfismo	Sistema SAP	Temperatura de anelamento	Primers			
				Id	Tipo	T_m (°C)	Sequência dos <i>Primers</i> (5'→3')
SbMATE_EX1_5519	SNP W (A/T)	três primers	53°C	HB48	Externo <i>forward</i>	56.88	GCTTTGAGATGTGTGTTTGC
				HB49	Externo <i>reverse</i>	57.03	GAAGAATTAGAAGCGGGACA
				HB50	alelo A	59.64	TGGAACAGCCCCAGGA
				HB51	alelo T	59.60	CTGGAACAGCCCCAGGT
SbMATE_I2_5985pb	SNP R (A/G)	três primers	51°C	HB23	Externo <i>forward</i>	59.87	GCATGCAACGGTGAAGTAGA
				HB24	Externo <i>reverse</i>	60.70	TTCGCCATTGTCAGTTGTTG
				HB26	alelo G	58.13	GATGTGTGTTTGCACATTTCTC
				HB27	alelo A	57.66	GATGTGTGTTTGCACATTTCTT
SbMATE_I2_6083pb	SNP M (A/C)	quatro primers	50°C	HB23	Externo <i>forward</i>	59.87	GCATGCAACGGTGAAGTAGA
				HB24	Externo <i>reverse</i>	60.70	TTCGCCATTGTCAGTTGTTG
				HB30	alelo A	59.16	GGGTGGGAGAGTGAGAGTGTA
				HB37	alelo C	58.65	GTGTTTTTTTCCGTTAAGGGTAG
SbMATE_I2_6094pb	SNP S (C/G)	três primers	51°C	HB23	Externo <i>forward</i>	59.87	GCATGCAACGGTGAAGTAGA
				HB24	Externo <i>reverse</i>	60.70	TTCGCCATTGTCAGTTGTTG
				HB32	alelo G	59.95	GGGAGAGTGAGAGTATCTGCCCTTTAG
				HB33	alelo C	59.84	GGGAGAGTGAGAGTATATGCCCTTTAC

Polimorfismo	Tipo do Polimorfismo	Sistema SAP	Temperatura de anelamento	Primers			
				Id	Tipo	Tm (°C)	Sequência dos <i>Primers</i> (5'→3')
ORF8_8365pb	SNP M (A/C)	quatro primers	51°C	HB56	Externo <i>forward</i>	59.03	AGAAACAAATAGTCATGATGAAAACAAT
				HB57	Externo <i>reverse</i>	59.02	CATGTACTTGGAATTCATTTTATAGATC
				HB58	alelo A	59.65	GCTAGTTTAAACCGCCCATACA
				HB59	alelo C	59.85	GAAAAAGAAAACTAACTATCCATACAACCTG
ORF8_8424pb	SNP K (G/T)	quatro primers	51°C	HB52	Externo <i>forward</i>	65.00	CGTTGAAGCTGAAGCTAGTCTTTTTCAC
				HB53	Externo <i>reverse</i>	64.00	TTCAAGTTTATGAACCTATGCGCATTTA
				HB54	alelo G	64.19	TATTTGCATTGCATTGCGGG
				HB55	alelo T	65.12	GTAGTATTTGCGTGGTCAACCGGTAATA
ORF9_12487pb	Indel 19 pb	-	53°C	HB1	<i>forward</i>	58.16	AAGCACCAGGCAAATCATT
				HB2	<i>reverse</i>	55.63	GATGAGAAGCTGATCACCAC

4.2. Identificação de haplótipos com os marcadores do loco *Alt_{SB}*

Os polimorfismos utilizados para o desenvolvimento de marcadores estão em desequilíbrio de ligação, como descrito por Caniato (2009). De fato, desequilíbrio de ligação extenso tem sido observado em sorgo (Hamblin et al., 2005), como para o gene *Opaco-2*, no qual desequilíbrio de ligação completo ocorre em dentro de uma região de 4250pb (de Alencar Figueiredo et al., 2008). A ocorrência de blocos com desequilíbrio de ligação permite que poucos SNPs, chamados de *tag* SNPs, sejam necessários para representar uma região. Os *tag* SNPs podem servir como representantes dos outros SNPs em desequilíbrio, identificando os haplótipos mais comuns observados em um painel (Johnson et al., 2001). Em geral, o número de *tag* SNPs necessários para se amostrar uma região é proporcional ao número de SNPs e à taxa de recombinação entre eles. Assim, a partir de valores dos quadrados dos coeficientes de correlação entre frequências alélicas dos mesmos (r^2 , Hill e Robertson, 1968) é possível se definir uma estratégia para a seleção de *tag* SNPs (Hamblin et al., 2007). Estudos baseados nas sequências de genes da via biossintética de amido em sorgo mostrou que entre 2 e 5 *tag* SNPs são necessários para diferenciar seis ou menos haplótipos comuns para esses genes (Hamblin et al., 2007).

Os valores de r^2 entre os sete polimorfismos do loco *Alt_{SB}* variaram de 0,3 a 1, sendo que os dois polimorfismos da ORF8 estão em desequilíbrio de ligação completo entre si ($r^2=1$) e em forte desequilíbrio com os polimorfismos *SbMATE_I2_6094pb* e ORF9_12487pb ($r^2=0,9$) (Caniato, 2009). Já o polimorfismo *SbMATE_I2_6083pb* possui um desequilíbrio moderado com todos os outros ($0,5 < r^2 < 0,7$) e o polimorfismo *SbMATE_I2_5985pb* possui um desequilíbrio mais baixo com os outros polimorfismos ($0,3 < r^2 < 0,5$). Assim, uma estratégia para a representação da região do loco *Alt_{SB}* seria utilizar uma combinação entre os polimorfismos *SbMATE_I2_6083pb*, *SbMATE_I2_5985pb* e um dos polimorfismos da ORF8, ou *SbMATE_I2_6094pb* ou ORF9_12487pb, de forma a capturar uma gama maior de haplótipos na região.

Como esses polimorfismos foram associados com a tolerância ao alumínio em um painel de ampla variabilidade genética em sorgo, eles são excelentes candidatos para explorar a diversidade genética na espécie, buscando-se os haplótipos mais favoráveis para a característica em questão. A outra grande vantagem desses polimorfismos é que os marcadores gerados a partir deles podem ser utilizados em diferentes germoplasmas de sorgo e não somente em populações estruturadas.

4.3. Estratégia de *pooling* para os marcadores do loco *Alt_{SB}*

A mineração de alelos demanda um processo de genotipagem em larga escala, a fim de se analisar grandes coleções de germosplasma em busca de novas fontes de alelos superiores. O DNA *pooling* é uma estratégia para se reduzir o número de genotipagens e, consequentemente os custos e o volume de trabalho, principalmente considerando que os alelos favoráveis são de baixa frequência (Amos et al., 2000; Sham et al., 2002), como é o caso dos polimorfismos do loco *Alt_{SB}*. Para a utilização dessa abordagem, é fundamental que o alelo de interesse seja identificado em cada *pool* de DNA, mesmo contendo proporções mais elevadas do alelo alternativo para cada loco marcador. Assim, um fator limitante no uso do *pooling* é a sensibilidade do ensaio, ou seja, a probabilidade de se detectar um determinado alelo dado que ele esteja presente em pelo um dos acessos que compõem o *pool*. Assim, o primeiro passo foi verificar a sensibilidade dos marcadores desenvolvidos. O teste foi feito construindo-se amostras com diferentes proporções de DNA genômico dos indivíduos BR007 e SC283, contrastantes para os polimorfismos, nas proporções de 1:0 à 1:20 (SC283:BR007 e BR007:SC283), inicialmente para o loco *SbMATE_I2_6083pb* (Figura 12). A concentração de DNA genômico dos estoques utilizados para esse experimento foi de 10 ng/μL. O marcador para esse loco apresentou uma alta sensibilidade na detecção do alelo de interesse em todas as proporções testadas, ficando clara a maior eficiência de amplificação do alelo do acesso tolerante, mesmo quando este estava presente em proporção 20 vezes menor. Já o alelo sensível só foi detectado claramente até a proporção de 1:8 BR007:SC283. Entretanto, como o alelo tolerante é o alelo de interesse e possui baixa frequência na espécie, é importante que ele seja detectado em qualquer proporção avaliada, consistente com o padrão representado na Figura 13.

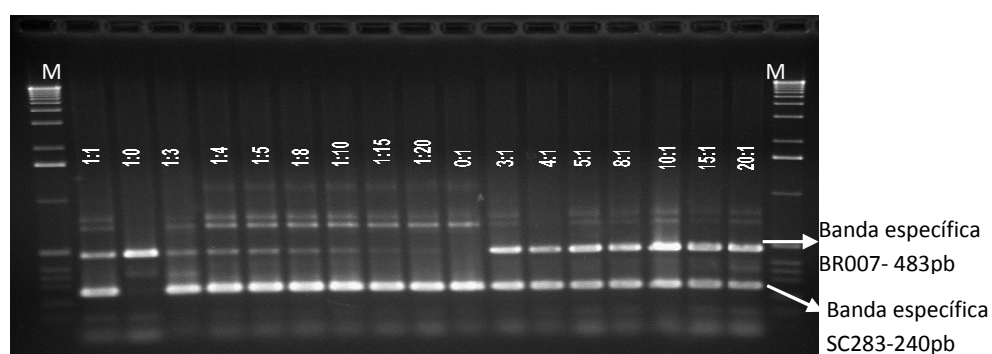


Figura 13 - DNA *pooling* para o marcador *SbMATE_I2_6083pb*. Cada caneleta representa uma mistura de DNA genômico entre os indivíduos BR007 e SC283 nas proporções que estão devidamente identificadas na figura (proporção BR007:SC283). M= marcador de tamanho molecular 1Kb.

Em função da alta sensibilidade de detecção do alelo de interesse em diversas proporções, os cálculos para se determinar o tamanho ideal dos *pools* foram feitos tendo como base as frequências dos alelos de interesse encontradas por Caniato (2009). O valor ótimo para o tamanho de *pool* mais frequente nas situações calculadas foi de quatro indivíduos, com a eficiência do processo de *pooling*, ou seja, a quantidade de reações economizadas, variando entre 43,18% em populações de 100 indivíduos até 85,80% para populações de 400 indivíduos para o polimorfismo *SbMATE_I2_6083pb*, por exemplo. Os resultados dos cálculos de eficiência da estratégia de *pooling* calculados para diferentes situações para três dos polimorfismos estão mostrados na Tabela 4.

Tabela 4 - Resultados dos cálculos para determinação do tamanho ideal do *pool* de DNA. Os valores correspondem à eficiência da estratégia (1-F) para diferentes tamanhos de população (n) e diferentes números de indivíduos no pool (n), considerando a frequência do alelo de interesse previamente determinada. Os valores máximos para cada condição estão destacados em negrito.

Polimorfismo	Frequência	n <i>pool</i>	n população			
			100	150	200	400
<i>SbMATE_I2_6083pb</i>	0,0913	3	0,4170	0,4170	0,4170	0,4170
		4	0,4318	0,6212	0,7159	0,8580
		5	0,4196	0,6131	0,3869	0,6131
		6	0,3964	0,5976	0,4024	0,5976
		7	0,3688	0,5792	0,4208	0,5792
		⋮	⋮	⋮	⋮	⋮
		20	0,0974	0,3982	0,5487	0,7743
<i>SbMATE_I2_5985pb</i>	0,1345		100	150	200	400
		3	0,3150	0,3150	0,3150	0,3150
		4	0,3111	0,5408	0,6556	0,8278
		5	0,2857	0,5238	0,6428	0,8214
		6	0,2537	0,5024	0,6268	0,8134
		7	0,2209	0,4806	0,6105	0,8052
		⋮	⋮	⋮	⋮	⋮
<i>SbMATE_I2_6094pb</i>	0,0613		100	150	200	400
		3	0,4938	0,4938	0,4938	0,4938
		4	0,5264	0,6843	0,7632	0,8816
		5	0,5288	0,6859	0,7644	0,8822
		6	0,5175	0,6783	0,7587	0,8794
		7	0,4994	0,6662	0,7497	0,8748
		⋮	⋮	⋮	⋮	⋮
		20	0,2322	0,4881	0,6161	0,8080

Em seguida, um teste com os marcadores gerados para os outros polimorfismos permitiu verificar que a estratégia de *pools* de quatro indivíduos é eficiente para todos os sistemas ARMS-PCR, com exceção do polimorfismo ORF8_8365pb. O marcador desenvolvido para esse polimorfismo apresentou bandas mais fracas para o alelo de interesse. Assim, ao se adotar a estratégia de *pooling*, não foi possível visualizar a banda referente a esse alelo. A estratégia de *pooling* foi efetiva para cinco dos seis marcadores desenvolvidos no presente trabalho, excluindo-se o indel do loco ORF9_12487bp cuja fase invertida não permite a utilização dessa estratégia. Essa estratégia pode representar uma grande economia de reações de PCR para a genotipagem de populações, de até 76% para uma população de 200 indivíduos, por exemplo. Assim, a estratégia de *pooling* é factível para os polimorfismos do loco *Alt_{SB}* e deve ser adotada na mineração de alelos para a tolerância ao alumínio.

Os marcadores desenvolvidos possuem um custo estimado de R\$3,70 por amostra, enquanto que o sequenciamento de uma amostra, alternativa existente para a genotipagem desse tipo de polimorfismo, possui um custo estimado de R\$6,00. Deste modo, o custo para a genotipagem de uma população de 200 indivíduos, por exemplo, ficaria em torno de R\$740,00, para cada marcador, podendo ser reduzido para R\$177,60 quando utilizada a estratégia de *pooling*, dependendo do marcador. Este valor contrasta com os R\$1200,00 que seriam gastos para genotipar essa mesma população por meio das técnicas de sequenciamento clássico de DNA. Portanto, os marcadores gerados nesse trabalho, além de serem eficientes na detecção dos alelos de interesse possuem fácil manipulação e baixo-custo, sendo fermentas viáveis para programas de melhoramento.

4.4. Mineração de alelos do Loco *Alt_{SB}*

Os marcadores haplótipo-específicos foram subsequentemente testados quanto à capacidade de identificar acessos contendo haplótipos-alvo em duas diferentes coleções de germoplasma de sorgo. Inicialmente, os haplótipos do loco *Alt_{SB}* foram identificados por meio das informações de sequenciamento de DNA gerados por Caniato (2009) e associados com os dados fenotípicos de tolerância ao alumínio no painel de associação de ICRISAT/CIRAD, também obtidos por Caniato (2009). Nesse germoplasma, dentre os 254 acessos de sorgo, somente 14 acessos (6%) apresentaram $CLR_{sd} > 80\%$, sendo considerados altamente tolerantes ao Al. Os haplótipos desses 14 acessos para os sete marcadores utilizados nesse estudo estão representados na Tabela 5.

Os seis polimorfismos associados à tolerância ao Al conferida pelo loco *Alt_{SB}* recuperaram de 15 a 50 acessos, dependendo do loco analisado, sendo que 9 (64,3%) dos 14 acessos altamente tolerantes foram recuperados. Utilizando-se o marcador

SbMATE_I2_6083pb foram recuperados 26 dos 254 acessos, dos quais 9 acessos são altamente tolerantes, apresentando 46% de eficiência de recuperação de acessos altamente tolerantes ao alumínio. Os marcadores ORF9_12487pb e *SbMATE_I2_6094pb* recuperaram 15 e 18 acessos, respectivamente, dos quais apenas 5 e 6 acessos foram altamente tolerantes. No total, cinco dos 14 acessos altamente tolerantes identificados por Caniato (2009) não foram recuperados por esses marcadores, sugerindo que a seleção baseada somente nesses marcadores acarretaria uma perda de tais acessos altamente tolerantes.

Tabela 5 - Haplótipos dos 14 acessos tolerantes do painel ICRISAT/CIRAD para o loco *Alt_{SB}*. Estão apresentados os polimorfismos que foram utilizados para o desenvolvimento dos marcadores ARMS-PCR. Os alelos favoráveis para cada polimorfismo estão destacados em negrito.

Acesso	CLR _{SD} (%)	<i>SbMATE_E1_5519pb</i>	<i>SbMATE_I2_5985pb</i>	<i>SbMATE_I2_6083pb</i>	<i>SbMATE_I2_6094pb</i>	ORF8_8365pb	ORF8_8424pb	ORF9_12487pb
		SNP W (A/T)	SNP R (A/G)	SNP M (A/C)	SNP S (C/G)	SNP M (A/C)	SNP K (G/T)	INDEL (19/0)
IS 14351	262	A	A	A	C	C	G	0
IS 23142	161	A	G	C	G	A	T	19
SC283	157	A	A	A	C	C	G	0
IS 25077	120	A	G	C	G	A	T	19
IS 21519	106	A	A	A	C	C	G	0
SC566	99	T	A	A	G	A	T	19
IS 26457	98	A	A	A	G	A	T	19
IS 29691	97	A	G	C	G	A	T	19
IS 26554	95	A	A	A	G	A	T	19
5DX61-6-2	89	A	G	C	G	A	T	19
SC175-14	88	A	G	C	G	A	T	19
BR501	82	A	A	A	C	A	T	19
CMS225	82	A	A	A	C	C	G	0
IS 10801	80	A	A	A	C	C	G	0

O painel pertencente ao IGD é composto por plantas de pequeno porte e florescimento independente do comprimento do dia, que foram desenvolvidas para permitir o cultivo de sorgo em regiões temperadas. Um total de 377 acessos dessa coleção também

foi genotipado com três marcadores do loco *SbMATE_I2* e um da ORF9, sendo identificados oito diferentes haplótipos do loco *Alt_{SB}* (Tabela 6).

Tabela 6 - Haplótipos do loco *Alt_{SB}* presentes no painel do IGD. Foram considerados os polimorfismos utilizados na genotipagem do painel com quatro marcadores ARMS-PCR. A localização dos polimorfismos está representada na Figura 4. As células em negrito indicam a presença do alelo favorável. Nas duas primeiras linhas estão descritos os haplótipos das linhagens BR007 e SC283, padrões de sensibilidade e tolerância ao AI, respectivamente.

Número de acessos		<i>SbMATE_I2_5985pb</i>	<i>SbMATE_I2_6083pb</i>	<i>SbMATE_I2_6094pb</i>	ORF9_12487pb
		SNP R (A/G)	SNP M (A/C)	SNP S (C/G)	INDEL (19/0)
BR007	-	G	C	G	19
SC283	-	A	A	C	0
Hap_1	304	G	C	G	19
Hap_2	39	G	C	G	0
Hap_3	2	A	C	G	19
Hap_4	2	A	A	C	19
Hap_5	11	A	A	G	19
Hap_6	6	A	A	C	0
Hap_7	8	G	A	C	0
Hap_8	1	A	C	G	0

Para cada haplótipo, os acessos que possuíam pelo menos um dos alelos favoráveis (i.e. ligado em fase de acoplamento com o alelo de tolerância) nos locos *SbMATE_I2* foram considerados como candidatos a fontes de tolerância e foram fenotipados para a tolerância ao AI em solução nutritiva, utilizando-se a característica Crescimento Líquido Relativo da raiz aos 5 dias (CLR_{5D}). Assim, 30 acessos portadores dos haplótipos Hap_3, Hap_4, Hap_5, Hap_6, Hap_7 e Hap_8 foram fenotipados, além de quatro outros genótipos sensíveis como controle negativo e um tolerante como controle positivo. Como controles negativos foram selecionados acessos que possuíam os haplótipos Hap_1 e Hap_2, e a linhagem BR007. Já o controle positivo foi a linhagem SC283. Dentre os 30 acessos, 15 (50%) apresentaram CLR_{5D}<30% e foram classificados como sensíveis, 8 (27%) tiveram

CLR_{5D} entre 30 e 80% sendo considerados tolerantes e os outros 7 (23%) restantes possuíam CLR_{5D} superior a 80% e foram considerados altamente tolerantes (Figura 14).

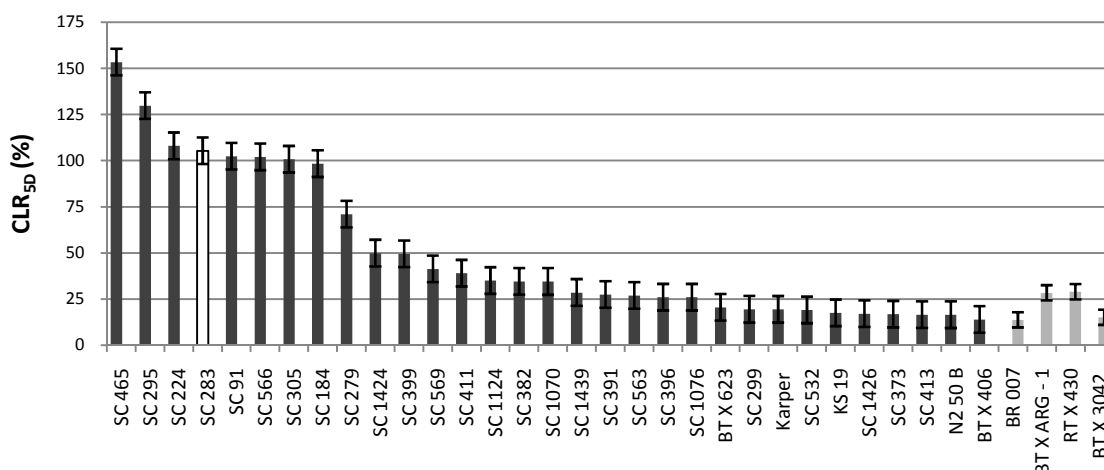


Figura 14 – Crescimento Líquido Relativo para os 30 acessos do painel do IGD que foram selecionados com base nos marcadores desenhados para os polimorfismos do loco *Alt_{SB}*. Esses acessos foram expostos à {27} μ M de Al por cinco dias em solução nutritiva pH 4.0. As barras na cor cinza representam os controles negativos e com a cor branca o controle positivo, SC283.

Os sete acessos classificados como altamente tolerantes possuem o alelo favorável dos SNPs *SbMATE_I2_6083pb* (alelo A) e do *SbMATE_I2_5985pb* (alelo A). Já para os outros dois marcadores, apenas quatro linhagens possuem o alelo favorável. Todos os acessos tolerantes pertencem ou ao haplótipo Hap_6, que possui todos os alelos favoráveis, ou ao Hap_5, que possui os alelos favoráveis para os marcadores *SbMATE_I2_6083pb* e *SbMATE_I2_5985pb*. Os marcadores recuperaram números de acesso variáveis conforme apresentado na Tabela 7. O marcador *SbMATE_I2_6083pb*, por exemplo, recuperou 27 acessos, o que corresponde à 90% das linhagens recuperadas por todos os marcadores. Dentre esses, foram recuperados todos os acessos altamente tolerantes e 87,5% dos tolerantes no conjunto dos 30 acessos. Embora os alelos favoráveis dos locos *SbMATE_I2_6094pb* e *OR9_12487pb* tenham uma frequência mais baixa, eles não foram capazes de identificar todos os acessos altamente tolerantes presentes no grupo de 30 acessos recuperados com todos os marcadores.

Tabela 7 - Acessos de sorgo do painel do IGD recuperados em três classes de tolerância, altamente tolerante (AT, $CLR_{5d} > 80\%$), tolerante (T, $30\% < CLR_{5d} < 80\%$) e sensível ($CLR_{5d} < 30\%$) pela genotipagem com quatro marcadores no loco Alt_{SB} . A posição dos marcadores no contexto no loco Alt_{SB} está mostrada na Figura 4. %Total denota a porcentagem de indivíduos recuperados por cada marcador em relação ao total de indivíduos recuperados pela genotipagem. %Marcador denota a porcentagem de indivíduos de cada classe em relação ao número total de indivíduos recuperados por cada marcador.

	TOTAL	AT	T	S
<i>SbMATE_I2_5985pb</i>	20	7	4	9
% Total	66,67	100,00	50,00	60,00
% Marcador		35,00	20,00	45,00
<i>SbMATE_I2_6083pb</i>	27	7	7	13
% Total	90,00	100,00	87,50	86,67
% Marcador		25,93	25,93	48,15
<i>SbMATE_I2_6094pb</i>	17	4	3	9
% Total	56,67	57,14	37,50	60,00
% Marcador		23,53	17,65	52,94
<i>ORF9_12487pb</i>	14	4	4	6
% Total	46,67	57,14	50,00	40,00
% Marcador		28,57	28,57	42,86

O terceiro painel utilizado nesse trabalho é mantido pelo INRAN, sendo uma coleção muito heterogênea. O painel é composto principalmente por *landraces* cultivadas pela população local do Niger em pequenas propriedades onde é limitada a adoção de variedades melhoradas. As *landraces* são localmente adaptadas ao ambiente severo, heterogêneo e imprevisível da África. Esse painel é composto por 164 materiais representando 14 *landraces* distintas e 23 acessos de melhoramento. A descrição da distribuição das *landraces* dos acessos desse painel está disposta no Anexo 1.

Esse germoplasma foi genotipado com os seis marcadores desenvolvidos para os SNPs *SbMATE_I2_5985pb*, *SbMATE_I2_6083pb*, *SbMATE_I2_6094pb*, *ORF8_8424pb* utilizando-se a estratégia de DNA *pooling*, além dos polimorfismos *ORF8_8365pb* e *ORF9_12487pb*. Dos 164 acessos do painel, 29 foram descartados devido à baixa germinação, número escasso de sementes ou rendimento insuficiente de DNA. Para o processo de genotipagem foram montados 41 *pools* compostos por quatro indivíduos em cada *pool*. Esse procedimento apresentou uma economia de 54% do total de reações (Tabela 8), superior aos resultados encontrados com os dados simulados no delineamento da estratégia de DNA *pooling*.

Tabela 8 - Eficiência da estratégia de *pooling* quanto à redução do número de genotipagens no painel INRAN. Nº de *pools* positivos denota o número de *pools* que apresentaram o alelo de interesse. Nº de reações necessárias para abertura dos *pools* denota o número de reações que são necessárias ao se genotipar todos os *pools* que apresentaram o alelo de interesse. Total indica o total de reações para genotipar o painel com cada marcador (os 41 *pools* iniciais mais os *pools* positivos). Eficiência (1-F) indica a eficiência da metodologia de *pooling*, sendo F a proporção entre o número de reações realizadas e o número de reações necessárias sem a metodologia *pooling*.

Polimorfismo	Nº de <i>Pools</i> Positivos	Nº de Reações necessárias para abertura dos <i>pools</i>	Total	Eficiência (1-F)
<i>SbMATE_I2_5985pb</i>	10	80	121	0,55
<i>SbMATE_I2_6083pb</i>	12	48	89	0,34
<i>SbMATE_I2_6094pb</i>	12	96	137	0,49
ORF8_8365pb	12	48	89	0,34
ORF8_8424pb	-	135	135	-
ORF9_12487pb	-	135	135	-
Total			706	0,54

¹Total: refere-se ao total de reações necessárias para se genotipar todos os *pools* (41) mais o total de reações necessárias para se genotipar os *pools* positivos.

Com a genotipagem desse painel foi possível a identificação de seis haplótipos para os polimorfismos, como apresentados na Tabela 9. Os acessos que possuem os haplótipos Hap_5 e Hap_6 foram considerados candidatos a fontes de tolerância ao AI e foram fenotipados em solução nutritiva. Dos acessos recuperados pelos marcadores, três eram *landrace* Bagoba, três *landrace* JJ, duas *landrace* Malle, uma *landrace* MDK, uma Sokomba, uma MG e duas linhagens de melhoramento. Dos 14 acessos candidatos, apenas 10 foram avaliados devido à disponibilidade de sementes, sendo incluídos ainda seis acessos desse mesmo painel que fazem parte do programa de melhoramento do Niger e três linhagens controles, BR007 (sensível), SC283 e SC566 (tolerantes ao AI).

Tabela 9 - Haplótipos do loco *Alt_{SB}* presentes no painel do INRAN. Foram considerados os polimorfismos utilizados na genotipagem do painel com seis marcadores ARMS-PCR. A localização dos polimorfismos está representada na Figura 4. As células em negrito indicam a presença do alelo favorável. Nas duas primeiras linhas estão descritos os haplótipos das linhagens BR007 e SC283, padrões de sensibilidade e tolerância ao AI, respectivamente

Número de acessos		<i>SbMATE</i> _12_5985pb	<i>SbMATE</i> _12_6083pb	<i>SbMATE</i> _12_6094pb	ORF8_8365pb	ORF8_8424pb	ORF9_12487pb
		SNP R (A/G)	SNP M (A/C)	SNP S (C/G)	SNP M (A/C)	SNP K (T/G)	INDEL (19/0)
BR007	-	G	C	G	A	T	19
SC283	-	A	A	C	C	G	0
Hap_1	63	G	C	G	A	T	19
Hap_2	44	G	C	G	A	T	0
Hap_3	8	A	C	G	A	T	19
Hap_4	6	A	C	G	A	T	0
Hap_5	3	A	A	C	C	G	19
Hap_6	11	A	A	C	C	G	0

Devido às características já citadas do material, foi observada uma variabilidade acentuada para tolerância ao AI dentro dos acessos, que já seria esperado. Em função dessa variabilidade, a avaliação do crescimento radicular de cada acesso foi realizada com base em cada planta individualmente, sendo consideradas tolerantes as plantas que apresentaram um índice de injúria radicular visual (IRV) maior que 3,0 (Tabela 10).

Tabela 10 – Avaliação dos acessos do painel do INRAN para tolerância em {27} μM de Al solução nutritiva. $\text{CRR}_{3\text{D}}$ e $\text{CRR}_{5\text{D}}$ denotam o crescimento radicular relativo aos 3 e 5 dias de estresse de Al, respectivamente; $\text{CRR}_{5\text{D}}\text{MIN}$ denota o valor mínimo do $\text{CRR}_{5\text{D}}$ encontrado para aquele acesso; $\text{CRR}_{5\text{D}}\text{MAX}$ denota o valor máximo do $\text{CRR}_{5\text{D}}$ encontrado para aquele acesso; IRV Média denota a injúria radicular visual média das raízes dos acessos e N° Plantas Tolerantes denota o número de plantas com $\text{IRV} > 3$ que foram classificadas como tolerantes para cada acesso.

ACESSO	NOMEMCLATURA INRAN	CRR _{3D}	CRR _{5D}	CRR _{5D} MIN	CRR _{5D} MAX	IRV MÉDIA	Nº Plantas Tolerantes
Com Alelo de Tolerância Favorável							
200923001	BAGOBAB	20,11	23,75	4,00	39,00	1,00	0
200923010	BAGOBAB	20,60	31,52	17,53	52,17	3,81	11
200923015	BAGOBAB	22,56	28,36	12,50	43,18	2,35	4
200923037	JJ	28,59	37,14	11,76	137,50	2,33	4
200923040	JJ	28,36	39,95	10,00	117,14	1,46	1
200923041	JJ	34,59	40,17	13,43	82,86	1,67	2
200923075	KORBIU	34,59	40,17	15,52	65,52	2,11	4
200923077	MALLE	31,40	40,19	14,58	65,12	3,46	12
200923144	P967083xSEPON68	18,21	22,53	13,33	38,98	1,61	3
200923162	MR732xSEPON82	15,13	20,62	8,74	38,67	2,81	7
Controles – Linhagens de Melhoramento do INRAN							
200923149	90SN-7	22,23	28,00	17,14	45,95	2,39	5
200923157	SEPON 82	19,77	23,95	11,83	31,33	2,32	0
200923150	IRAT-204	18,16	20,54	11,43	26,87	1,96	0
200923021	BABADIA FARA	24,19	26,71	10,53	63,79	2,11	2
200923060	KAOURA	15,89	16,87	5,41	27,27	0,36	0
200923089	MDK	18,86	20,99	12,50	28,57	1,39	0
Controles – Linhagens de Melhoramento da Embrapa							
BR007		16,29	16,22	13,95	20,27		-
SC283		34,55	71,12	62,00	91,43		-
SC566		35,13	62,86	50,62	73,02		-

Dos 10 acessos selecionados por meio dos marcadores, nove foram classificados como tolerantes, uma vez que possuem plantas com um bom crescimento radicular sob níveis tóxicos de alumínio em solução nutritiva, ou seja, $\text{IRV} \geq 3,0$. A *landrace* JJ é a mais representada no painel, tendo 27 acessos. Destes, três apresentaram os haplótipos de interesse e foram tolerantes. Dos três acessos com haplótipos tolerantes e caracterizados como *landrace* Bagoba, um não apresentou tolerância ao Al. Essa *landrace* é representada por 19 acessos no painel. Tais resultados reforçam a heterogeneidade do material e demonstram que uma seleção baseada apenas por tipo de *landraces* não é eficiente, sendo necessário também considerar o loco *Alt_{SB}* para a recuperação de acessos tolerantes ao Al.

Algumas plantas apresentaram tolerância superior em comparação com os controles positivos, SC283 e SC566. Essas plantas foram transplantadas para vasos e estão sendo mantidas em casa de vegetação para serem autofecundadas e possivelmente utilizadas no programa de melhoramento do INRAN. Esses acessos serão de grande valia para os programas de melhoramento do INRAN, uma vez que já possuem outras características desejáveis para aquele local, como a sensibilidade ao fotoperiodismo (Weltzien et al., 2006). A importância da mineração de alelos no germoplasma alvo é que as fontes dos alelos de interesse já são adaptadas aos ambientes-alvo, sem a necessidade de introgressão de fonte exóticas pouco adaptadas às condições locais e a consequente necessidade de retrocruzamentos para recuperação das características adaptativas.

Na análise dos três painéis foi confirmada a baixa frequência da tolerância ao alumínio em sorgo, já reportada por Caniato (2009). Apesar da baixa frequência, os acessos tolerantes foram recuperados utilizando os marcadores haplótipo-específicos do loco *Alt_{SB}*, indicando que tais marcadores serão úteis para o *screening* de germoplasmas buscando novas fontes de tolerância ao Al, como mostrado nesse trabalho, além da seleção assistida visando à introgressão desta característica em linhagens elites.

Com o rápido avanço das técnicas moleculares associado a uma redução dos custos na geração desses dados, os marcadores de DNA vêm tornando-se de uso rotineiro nos programas de melhoramento genético de plantas, aumentando a eficiência e proporcionando maiores ganhos genéticos nas principais culturas de interesse econômico. No entanto, um dos maiores desafios dessa área de pesquisa ainda é a transferência de novas tecnologias para os programas de melhoramento. A genotipagem com os marcadores gerados nesse trabalho pode ser realizada em laboratórios com infra-estrutura mínima, podendo ser utilizados efetivamente em programas de melhoramento para identificar novos alelos do gene *Alt_{SB}* em coleções de germoplasma diversas. Adicionalmente, a estratégia de DNA *pooling* também se mostrou eficiente para a genotipagem de grandes coleções, uma vez que minimiza o número de reações de PCR necessárias para a identificação de alelos de baixa frequência, economizando tempo e recursos. Contudo, não se deve esquecer que outras fontes de tolerância, não associadas ao loco *Alt_{SB}*, podem ser perdidas nesse processo de seleção.

4.5. Caracterização fisiológica e molecular dos acessos recuperados com os marcadores haplótipo-específicos do loco *Alt_{SB}*

Uma caracterização dos diferentes acessos que compõem o loco *Alt_{SB}* é necessária para se elucidar os mecanismos que modulam a variabilidade fenotípica da tolerância ao alumínio. Caniato (2009), trabalhando com o painel de associação ICRISAT/CIRRAD, identificou 30 haplótipos diferentes, dos quais treze apresentaram baixa frequência (0,5%). Nove dos haplótipos descritos estão presentes em acessos altamente tolerantes, entretanto, a maioria desses haplótipos está presente também em indivíduos sensíveis. Esses nove haplótipos foram selecionados para um estudo fisiológico e molecular (Tabela 11).

Tabela 11 - Descrição dos nove haplótipos do loco *Alt_{SB}* que apresentaram acessos tolerantes. Cada coluna apresenta um loco polimórfico associado com a tolerância ao Al e o número de linhagens que apresentam o haplótipo. A linhagem BR007, padrão de sensibilidades, possui o haplótipo *Alt_{SB}_2*, o segundo mais frequente do painel e presente principalmente em linhagens sensíveis. Já a linhagem SC283, padrão de tolerância, possui o haplótipo *Alt_{SB}_22*, presente em linhagens altamente tolerantes e tolerantes. Esses dois haplótipos estão destacados em negrito.

Haplótipo	Número de linhagens	ORF6_160pb	ORF6_200pb	mite	SbMATE_I2_5947pb	SbMATE_I2_5985pb	SbMATE_I2_6083pb	SbMATE_I2_6094pb	ORF8_8365pb	ORF8_8424pb	ORF9_12487pb	ORF10_24804pb	ORF10_25094pb
<i>Alt_{SB}_2</i>	62	7	G	1184	G	G	C	G	A	T	19	T	C
<i>Alt_{SB}_8</i>	1	0	A	1184	A	A	A	C	C	G	19	C	C
<i>Alt_{SB}_12</i>	1	0	A	1184	A	G	C	G	A	T	0	C	C
<i>Alt_{SB}_19</i>	2	0	A	1544	A	A	A	C	C	G	19	C	T
<i>Alt_{SB}_20</i>	15	7	G	1912	G	G	C	G	A	T	0	T	C
<i>Alt_{SB}_21</i>	7	7	G	1544	G	G	C	G	A	T	0	C	C
<i>Alt_{SB}_22</i>	5	0	A	1912	A	A	A	C	C	G	19	C	T
<i>Alt_{SB}_23</i>	5	0	A	1544	A	A	A	G	A	T	0	C	C
<i>Alt_{SB}_27</i>	2	7	G	1544	A	A	A	G	A	T	0	C	C

O haplótipo *Alt_{SB}_2* é o segundo mais frequente no painel, estando presente em 61 indivíduos sensíveis e um indivíduo tolerante ao Al. Dos indivíduos que possuem o haplótipo *Alt_{SB}_21*, apenas um é tolerante ao Al. Os haplótipos *Alt_{SB}_8* e *Alt_{SB}_12* estão presentes em apenas um indivíduo tolerante cada. Dois indivíduos possuem o *Alt_{SB}_19*, um tolerante e outro moderadamente tolerante, assim como o *Alt_{SB}_27*, que está presente em um indivíduo tolerante e em outro sensível. Os haplótipos *Alt_{SB}_22* e *Alt_{SB}_23* estão presentes em três indivíduos tolerantes e em indivíduos com tolerância moderada ao alumínio.

Para o estudo desses haplótipos foram selecionadas 16 linhagens, procurando abrangendo a diversidade fenotípica e genotípica dos haplótipos identificados (Tabela 12), mas incluindo 12 linhagens altamente tolerantes ao Al no painel ICRISAT/CIRAD. Essas linhagens já haviam sido previamente caracterizadas por Caniato (2009) para a tolerância ao Al em solução nutritiva na presença de $\{27\} \mu\text{M}$ de Al em pH 4,0, quando foram avaliadas as variáveis de Crescimento Líquido Relativo em três e cinco dias de estresse de Al ($\text{CLR}_{5\text{d}}$ e $\text{CLR}_{3\text{d}}$, respectivamente) e injúria radicular visual (IRV), em um experimento semelhante ao descrito para avaliação de tolerância das linhagens do painel do IGD.

Tabela 12 - Acessos de sorgo selecionados para a caracterização fisiológica e molecular dos haplótipos. Nome dos acessos, haplótipo do loco Alt_{SB} , número de acessos no painel de associação ICRISAT/CIRAD que apresentam esse haplótipo, classe de tolerância ao Al, país de origem do acesso e as variáveis fenotípicas $\text{CLR}_{3\text{d}}$, $\text{CLR}_{5\text{d}}$ e IRV. Os valores representam médias de duas repetições e sete plantas por repetição obtidos por Caniato (2009).

Acesso	Haplótipo	Número de Acessos do Haplótipo	Classe de Tolerância ao AL	País de Origem	$\text{CLR}_{3\text{d}}$ (%)	$\text{CLR}_{5\text{d}}$ (%)	IRV
BR007B	Alt_{SB}_2	62	S	USA	$19 \pm 1,2$	$12 \pm 0,8$	1 ± 0
IS 29691	Alt_{SB}_2	62	AT	ZWE	$117 \pm 5,6$	$99 \pm 6,5$	$1 \pm 0,3$
IS 21519	Alt_{SB}_8	1	AT	MWI	$95 \pm 5,8$	$113 \pm 3,9$	5 ± 0
IS 25077	Alt_{SB}_12	1	AT	GHA	$118 \pm 5,9$	$120 \pm 8,9$	$3 \pm 0,4$
IS 10801	Alt_{SB}_19	2	AT	TCD	$65 \pm 4,9$	$78 \pm 6,5$	$4 \pm 0,5$
SC283	Alt_{SB}_20	5	AT	TZA	$105 \pm 5,4$	157 ± 9	$4 \pm 0,3$
IS 14351	Alt_{SB}_20	5	AT	MWI	$233 \pm 7,6$	$262 \pm 4,6$	$5 \pm 0,2$
IS 23142	Alt_{SB}_22	7	AT	TZA	$86 \pm 5,9$	$163 \pm 9,9$	$5 \pm 0,3$
BR012R	Alt_{SB}_22	7	S	BRA	$24 \pm 1,8$	$16 \pm 1,2$	$1 \pm 0,3$
IS 26457	Alt_{SB}_21	5	AT	BEM	$96 \pm 6,1$	$97 \pm 9,3$	$4 \pm 0,2$
IS 26554	Alt_{SB}_21	5	AT	BEM	$74 \pm 3,1$	$95 \pm 4,3$	$5 \pm 0,2$
SC549	Alt_{SB}_23	5	T	NGA	$70 \pm 3,4$	$61 \pm 4,9$	$3 \pm 0,3$
SC566	Alt_{SB}_23	2	AT	NGA	$97 \pm 5,8$	$99 \pm 3,9$	5 ± 0
9929030	Alt_{SB}_23	2	S	BRA	$28 \pm 2,1$	$25 \pm 2,3$	$2 \pm 0,5$
SC175-14	Alt_{SB}_27	15	AT	ETH	$72 \pm 2,4$	87 ± 3	$4 \pm 0,3$
SC112-14	Alt_{SB}_27	15	S	ETH	$38 \pm 4,3$	$24 \pm 2,9$	1 ± 0

Os códigos dos países estão descritos a seguir: BEN, Benin; BRA, Brasil; ETH, Etiópia; GHA, Ghana; MWI, Malawi; NGA, Nigéria; TCD, Chad; TZA, Tanzânia; USA, Estados Unidos da América; ZWE, Zimbabwe.

Foram realizados dois experimentos para análise da expressão do gene *SbMATE*, que apresentaram uma alta correlação entre si ($r=0,94$, $p<0,001$) e puderam ser analisados conjuntamente. Os indivíduos BR007, SC112, IS29691 e BR012 apresentaram uma baixa expressão do gene *SbMATE* e, com exceção da IS29691, foram sensíveis ao Al tóxico (Figura 15, Tabela 12).

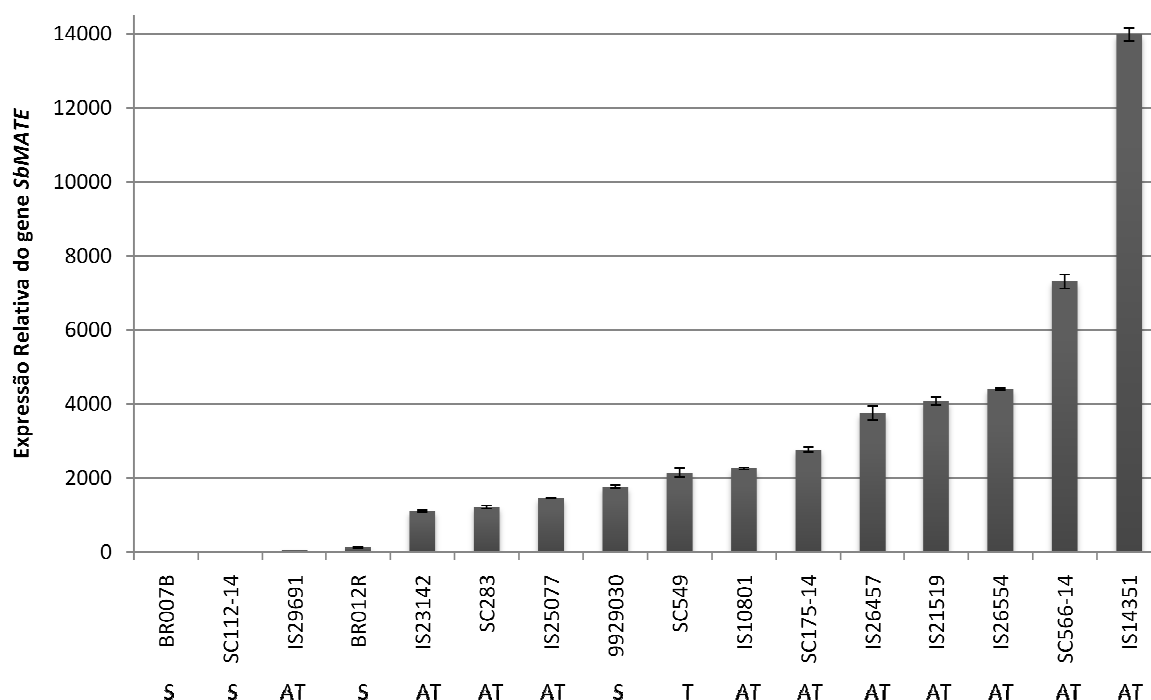


Figura 15 - Expressão relativa do gene *SbMATE* em linhagens selecionadas de sorgo, tomando-se como calibrador a linhagem BR007. As linhagens de sorgo foram expostas a, {27} μ M de Al por cinco dias em pH 4,0. Os valores representam médias de três repetições e as barras verticais indicam o desvio padrão. As letras abaixo do genótipo referem-se às classes de tolerância ao Al previamente caracterizados: (S) sensível, (T) tolerante e (AT) altamente tolerante.

A expressão relativa do gene *SbMATE* apresentou uma correlação positiva de 0,76 ($p=0,0006$) e 0,68 ($p=0,0035$) com CLR_{3d} e CLR_{5d} , respectivamente (Tabela 13). Magalhães et al. (2007) encontraram uma correlação de 0,98 ($p<0.001$) entre expressão do gene e CLR_{5d} .

Tabela 13 - Correlação entre as variáveis estudadas para as 16 linhagens representantes dos haplótipos do loco *Alt_{SB}*. Para cada correlação estão representados os coeficientes de correlação de Pearson, *r*, na primeira linha, seguidos dos p-valores. Exsudação denota a exsudação de citrato pelas raízes por 6 horas em 5 dias na presença de {27} μ M de Al. CLR_{3d} e CLR_{5d} denotam o crescimento líquido relativos das raízes 3 e 5 dias na presença de {27} μ M de Al, respectivamente. Expressão denota a expressão relativa do gene *SbMATE* após 5 dias na presença de {27} μ M de Al. 'Ac Al' denota a acumulação de Al nos ápices radiculares após um período de 5 dias na presença de {27} μ M de Al.

	Exsudação	CLR _{3d}	CLR _{5d}	Expressão	Ac Al
Exsudação	1,00	0,60	0,55	0,85	-0,67
		0,0146	0,0269	<0.0001	0,0049
CLR _{3d}		1,00	0,93	0,76	-0,50
			<0,0001	0,0006	0,0049
CLR _{5d}			1,00	0,68	-0,56
				0,0035	0,0252
Expressão				1,00	-0,54
					0,0252
AcAl					1,00

Nesse experimento foi evidente a grande expressão do gene *SbMATE* na linhagem IS14351, que apresentou um comportamento diferenciado em relação às demais. Esta linhagem é altamente tolerante ao Al como evidenciado pelo seu alto valor de CLR_{5d} (262%) e, aparentemente, a maior parte da sua tolerância deve estar associada ao gene *Alt_{SB}*. Como é uma linhagem proveniente do Malawi, na África, espera-se que essa linhagem seja adaptada às condições ambientais locais, podendo ser utilizada diretamente em programas de seleção assistida com os marcadores do gene *Alt_{SB}*, evitando a necessidade de recuperação de características adaptativas.

Embora a linhagem IS29691 tenha apresentado uma baixa expressão do gene *Alt_{SB}* quando comparada aos demais, mostrou-se altamente tolerante ao Al, indicando que tal característica não está envolvida com o loco *Alt_{SB}* nesse acesso. De fato, marcadores flanqueando o loco *Alt_{SB}* não apresentaram ligação significativa com a tolerância ao Al na população de retrocruzamento dessa linhagem com a linhagem sensível BR007 (Caniato et al., 2007). Uma possível explicação é que a tolerância desse acesso seja proveniente de outros locos de tolerância distintos do *Alt_{SB}*.

O acúmulo de alumínio nos ápices radiculares das linhagens apresentou uma correlação negativa de 0,72 (*p* = 0,0017) com o CLR_{5d} (Figura 16, Tabela 13), com a expressão do gene *Alt_{SB}* (*r*=0,54; *p*=0,03) e com a IRV (*r*=0,86; *p*<0,0001). A exsudação de

citrato pelos ápices radiculares (Figura 17) apresentou também uma correlação de 0,60 ($p=0,146$) com CLR_{3d} e 0,55 ($p=0,0269$) com CLR_{5d} e com a expressão do gene *SbMATE* ($r=0,85$, $p<0.0001$). Esses resultados são condizentes com o mecanismo de tolerância controlado pelo *Alt_{SB}*, de imobilização do Al externamente à célula. A alta correlação encontrada entre a exsudação de citrato e a expressão do *Alt_{SB}* indica, ainda, que o esse gene é o principal responsável pela exsudação de citrato no painel.

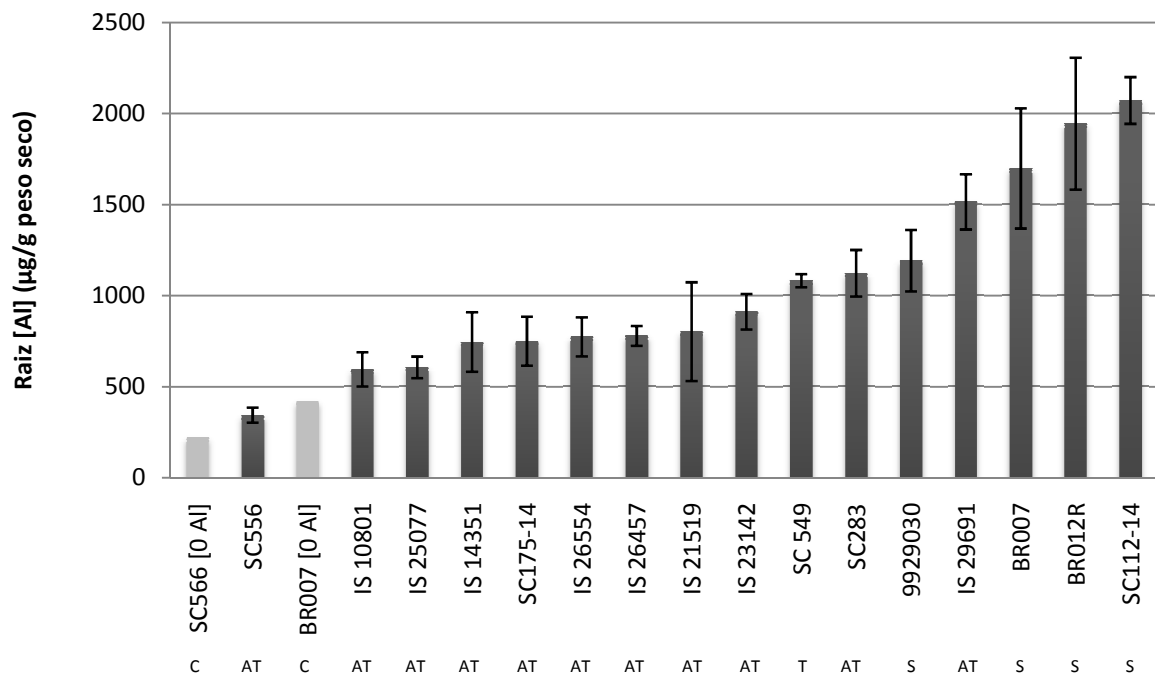


Figura 16 - Acumulação de Al nos ápices radiculares de linhagens de sorgo submetidas à {27} µM de Al por 5 dias em solução nutritiva pH 4,0. SC566 [0 Al] e BR007 [0 Al] são os controles negativos do experimento e estão destacados em cinza claro. Os valores representam a média de três repetições e as barras verticais indicam o desvio padrão. As letras abaixo do genótipo referem-se às classes de tolerância ao Al, previamente caracterizados por Caniato (2009): (S) sensível, (T) tolerante, (AT) altamente tolerante e (C) controles.

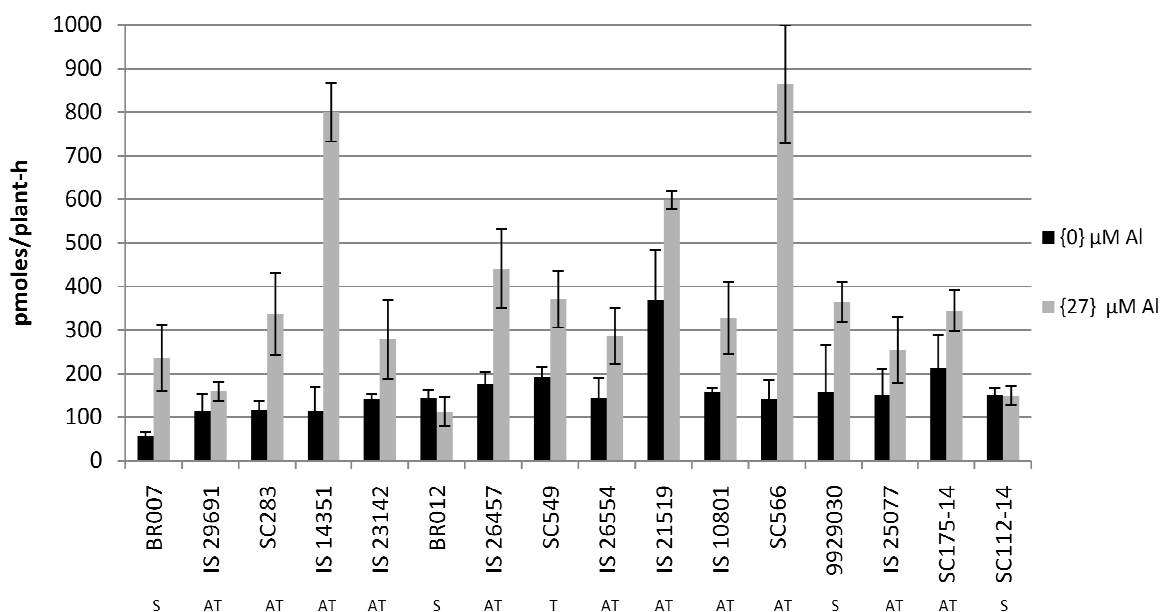


Figura 17 – Exsudação de citrato pelos ápices radiculares de linhagens de sorgo submetidas a {0} μM (barras pretas) e {27} μM (barras cinza) de Al por 5 dias em solução nutritiva pH 4,0. Os valores representam a média de três repetições e as barras verticais indicam o erro padrão. As letras abaixo do genótipo referem-se às classes de tolerância ao Al, previamente caracterizados: (S) sensível, (T) tolerante e (AT) altamente tolerante.

A linhagem IS29691, apesar de ser tolerante ao Al, mostrou um acúmulo de alumínio no ápice radicular similar ao das linhagens sensíveis ao Al e não apresentou diferença entre as taxas de exsudação de citrato quando submetido ou não ao estresse de Al. Esse acesso possui o haplótipo *Alt_{SB}_2* e é o único acesso altamente tolerante dentro dos 62 que possuem esse haplótipo no painel de associação, que inclui a linhagem BR007. Assim, fica evidente que a tolerância desse acesso não é devida ao loco *Alt_{SB}*. É possível que esse acesso possua um mecanismo de detoxificação interna de Al baseado em transportadores ABC, como em arroz (Huang et al., 2009) e arabidopsis (Larsen et al., 2005). O mecanismo de IS29691, porém, parece assemelhar-se com aquele descrito para arroz, uma vez que grande acúmulo de alumínio foi encontrado no ápice radicular desse acesso, enquanto que o mecanismo apresentado por arabidopsis baseia-se na redistribuição do Al em outros tecidos, mantendo-o longe das partes sensíveis da planta (Larsen et al., 2005). Entretanto, mecanismos internos de acúmulo de Al não foram descritos ainda para o sorgo, fazendo-se necessários estudos mais aprofundados para elucidar essa questão. Assim, essa linhagem se apresenta promissora como uma nova fonte de tolerância ao Al, podendo ser utilizada na piramidação de genes, buscando níveis superiores de tolerância.

Com os estudos dos diferentes haplótipos do loco *Alt_{SB}* observou-se que nem toda tolerância apresentada pelos acessos de sorgo é proveniente desse loco, assim como

descrito por Caniato et al. (2007). Além disso, outros mecanismos de tolerância podem ser os responsáveis pelo desenvolvimento da planta sob o estresse de Al. Uma estratégia de seleção baseada somente nesse loco pode, portanto, levar a uma perda de fontes diferentes de tolerância ao Al.

5. Conclusão

No presente trabalho foram gerados sete marcadores moleculares para os polimorfismos associados com a tolerância ao Al que possuem um grande potencial para a mineração de novos alelos do gene *Alt_{SB}* e para a introgressão em linhagens elite, uma vez que são de fácil manipulação e de baixo-custo. Com esses marcadores pode-se, eficientemente, buscar os haplótipos superiores para a característica de tolerância ao Al na diversidade genética da espécie, além da aplicação na seleção e piramidação de novas fontes de tolerância ao alumínio em sorgo. A estratégia de *pooling* desenhada para esses marcadores também se mostrou útil para acelerar e para reduzir os custos da genotipagem no processo de seleção assistida por marcadores.

A mineração de alelos nos painéis do IGD e do INRAN utilizando esses marcadores identificou acessos que poderão ser utilizados como fontes de tolerância nos programas de melhoramento, evitando a necessidade de recuperação de características adaptativas, como fotoperiodismo, por exemplo.

O estudo molecular e fisiológico dos haplótipos do loco *Alt_{SB}* demonstrou que a tolerância ao Al (CLR_{5d}) está correlacionada com a expressão do gene *SbMATE* ($r=0,68$ $p=0,0035$), a exsudação de Al pelos ápices radiculares ($r=0,55$ $p=0,0269$) e o acúmulo de alumínio na raiz ($r=-0,56$, $p=0,0252$). Além disso, com essa caracterização dos haplótipos foi determinado que a elevada tolerância ao Al da linhagem IS29691 pode estar associada com genes distintos ao *Alt_{SB}* e que possivelmente possui um mecanismo de detoxificação interna de Al, totalmente distinto daquele já descrito na espécie. Estudos futuros devem ser realizados para caracterizar tais mecanismos de tolerância ao Al.

6. Referencias Bibliográficas

- Aldrich PR, Doebley J, Schertz KF, Stec A. 1992. Patterns of allozyme variation in cultivated and wild sorghum bicolor. *Theoretical Applied Genetics* 85:451-460.
- Amaral SR, Lira MA, Tabosa JN, Santos MVF, Mello ACL, Santos VF. 2003. Comportamento de linhagens de sorgo forrageiro submetidas a déficit hídrico sob condição controlada. *Pesquisa Agropecuária Brasileira* 38: 973-979.
- Aniol A, Gustafson JP. 1984. Chromosome location of genes controlling aluminum tolerance in wheat, rye and triticale. *Canadian Journal of Genetics and Cytology* 26: 701-705.
- Aniol A. 1990. Genetics of tolerance to aluminum in wheat (*Triticum aestivum* L. Thell). *Plant and Soil* 123: 223-227.
- Amos CI, Frazier ML, Wang W. 2000. DNA pooling mutation detection with reference to sequence analysis. *American Journal of Human Genetics* 66: 1689-1692.
- Barnaud A, Deu M, Garine E, Mckey D, Joly HI. 2007. Local genetic diversity of sorghum in a village in northern Cameroon: structure and dynamics of landraces. *Theoretical Applied Genetics* 114:237-248
- Batley J, Barker G, O'Sullivan H, Edwards KJ, Edwards D. 2003. Mining for single nucleotide polymorphisms and insertions/deletions in maize expressed sequence tag data. *Plant Physiology* 132: 84-91.
- Bui M, Liu Z. 2009. Simple allele-discriminating PCR for cost-effective and rapid genotyping and mapping. *Plant Methods* 5: 1-8.
- Bundock PC, Cross MJ, Shapter FM and Henry RJ. 2006. Robust allele-specific polymerase chain reaction markers developed for single nucleotide polymorphisms in expressed barley sequences. *Theoretical Applied Genetics* 112:358-365.
- Caniato FF, Guimarães CT, Schaffert RE, Alves VMC, Kochian LV, Borém A, Klein PE, Magalhães JV. 2007. Genetic diversity for aluminum tolerance in sorghum. *Theoretical and Applied Genetics* 114: 863-876.

- Caniato FF. 2009. Estrutura de população e mapeamento associativo para tolerância ao alumínio em regiões candidatas do loco *Alt_{SB}* de sorgo. 115 p. Tese (Doutorado em Genética e Melhoramento) – Universidade Federal de Viçosa, Viçosa, Minas Gerais.
- Casa AM, Pressoir G, Brown PJ, Mitchell SE, Rooney WL, Tuinstra MR, Franks CD, Kresovich S. 2008. Community resources and strategies for association mapping in sorghum. *Crop Science* 48: 30-40.
- Chagnè D, Gasic K, Crowhurst RN, Han Y Bassett CH, Bowatte DR, Lawrence TJ, Rikkerink EHA, Gardiner SE, Korban SS. 2008. Development of a set of SNP markers present in expressed genes of the apple. *Genomics* 92: 353-358.
- Chiapparino E, Lee D, Donini P. 2004. Genotyping single nucleotide polymorphisms in barley by tetra-primer ARMS-PCR. *Genome* 47:414-420.
- Ciamprova M. 2002. Morphological and structural responses of plant roots to aluminum at organ, tissue, and cellular levees. *Biological Plants Pragues* 45: 161-171.
- Collins NC, Shirley NJ, Saeed M, Pallotta M, Gustafson JP. 2008. An *ALMT1* gene cluster controlling aluminum tolerance at the *Alt4* locus of rye (*Secale cereale* L.). *Genetics* 179: 669-682.
- Cui YX, Xu GW, Magill CW, Schertz KF. 1995. RFLP-based assay of *Sorghum bicolor* (L) Moench genetic diversity. *Theoretical Applied Genetics* 90:787-796.
- De Alencar Figueiredo LF, Calatayud C, Dupuits C, Billot C, Rami JF, Brunel D, Perrier X, Courtois M, Deu M, Glaszmann JC. 2008. Phylogenetic evidence of crop neodiversity in sorghum. *Genetics* 179: 997-1008.
- Delhaize E, Ryan PR, Randall PJ. 1993. Aluminum tolerance in wheat (*Triticum aestivum* L.). II. Aluminum-stimulated excretion of malic acid from root apices. *Plant Physiology* 103: 695-702.
- Delhaize E, Ryan PR, Hebb DM, Yamamoto Y, Sasaki T, Matsumoto H. 2004. Engineering high-level aluminum tolerance in barley with the *ALMT1* gene. *Proceedings of the National Academy of Sciences of the USA* 101: 15249-15254.
- Delhaize E, Gruber BD, Ryan PR. 2007. The roles of organic anion permeases in aluminum resistance and mineral nutrition. *FEBS Letters* 581: 2255-2262.

- Deu M, Rattunde F, Chantereau J. 2006. A global view of genetic diversity in cultivated sorghums using a core collection. *Genome* 49:168-180.
- Deu M, Sagnard F, Chantereau J, Calatayud C, Hérault D, Mariac C, Pham J-L, Vigouroux Y, Kapran I, Traore TS, Mamadou A, Gerard B, Ndjeunga J, Bezançon G. 2008. Niger-wide assessment of in situ sorghum genetic diversity with microsatellite markers. *Theoretical and Applied Genetics* 116:903–913.
- Djè Y, Heuertz M, Lefèbvre C, Vekemans X. 2000. Assessment of genetic diversity within and among germoplasm accessions in cultivated sorghum using microsatellite markers. *Theoretical Applied Genetics* 100: 918-925.
- Dricoll CT, Schecher WD. 1990. The chemistry of aluminum in the environment. *Environmental Geochemistry and Health* 12: 28-49.
- Durrett TP, Gassmann W, Rogers EE. 2007. The FRD3-mediated efflux of citrate into the root vasculature is necessary for efficient iron translocation. *Plant Physiology* 144: 197-205.
- FAO Statistical Database. 2009. Disponível em <<http://faostat.fao.org/site/408/default.aspx>> Acesso em: 11 outubro 2009.
- Foy CD, Carter TE, JR, Duke JA, Devine TE. 1993. Correlation of shoot and root growth and its role in selecting for aluminum tolerance in soybean. *Journal of Plant Nutrition* 16: 305-325.
- Galego FJ, Calles B, Beninto C. 1998. Molecular markers linked to the aluminum tolerance gene *Alt1* in rye (*Secale cereal* L.). *Theoretical Applied Genetics* 97:1104-1109.
- Gastwirth JL. 2000. The efficiency of pooling in the detection of rare mutations. *American Journal of Human Genetics* 67: 1036-1039.
- Gebhardt C, Ballvora A, Walkemeier B, Oberhagemann P, Schuler K. 2004. Assessing genetic potential in germplasm collections of crop plants by marker-trait association: a case study for potatoes with quantitative variation of resistance to late blight and maturity type. *Molecular Breeding* 13:93-102.
- Hamblin MT, Fernandez MGS, Casa AM, Mitchell SE, Paterson AH, Kresovich S. 2005. Equilibrium processes cannot explain high levels of short- and medium-range linkage disequilibrium in the domesticated grass *Sorghum bicolor*. *Genetics* 171: 1247-1256.

- Hamblin MT, Fernandez MGS, Tuinstra MR, Rooney WR, Kresovich S. 2007. Sequence variation at candidate loci in the starch metabolism pathway in sorghum: prospects for linkage disequilibrium mapping. *Crop Science* 47: 125-134.
- Harlan JR e De Wet JM. 1972. A simplified classification of cultivated sorghum. *Crop Science* 12: 172-176.
- Harjes CE, Rochefort TR, Bai L, Brutnell TP, Kandianis CB, Sowinski SG, Stapleton AE, Vallabhaneni R, Williams M, Wurtzel ET, Yan J, Buckler ES. 2008. Natural genetic variation in Lycopene epsilon cyclase tapped for maize biofortification. *Science* 319: 330-333.
- Hill WG, Robertson A. 1968. Effects of inbreeding at loci with heterozygote advantage. *Genetics* 60: 615–628.
- Hoekenga OA, Vision TJ, Shaff JE, Monforte AJ, Lee GP, Howell SH, Kochian LV. 2003. Identification and characterization of aluminum tolerance loci in *Arabidopsis* (*Arabidopsis thaliana* x *Columbia*) by quantitative trait locus mapping. A physiologically simple but genetically complex trait. *Plant Physiology* 132: 936-948.
- Hoekenga OA, Maron LG, Piñeros MA, Cançado GMA, Shaff J, Kobayashi Y, Ryan PR, Dong B, Delhaize E, Sasaki T, Matsumoto H, Yamamoto Y, Koyama H, Kochian LV. 2006. *AtALMT1*, which encodes a malate transporter, is identified as one of several genes critical for aluminum tolerance in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the USA* 103: 9738-9743.
- Huang L, Brooks SA, Li W, Fellers JP, Trick HN, Gill BS. 2003. Map-based cloning of leaf rust resistance gene Lr21 from the large and polyploidy genome of bread wheat. *Genetics* 164: 655-664.
- Huang CF, Yamaji N, Mitani N, Yano M, Nagamura Y, Ma JF. 2009. A bacterial-type ABC transporter is involved in aluminum tolerance in rice. *The Plant Cell* 21: 655-667.
- Hue NV, Craddock GR, Adams F. 1986. Effect of organic anions on aluminum toxicity in subsoil. *Soil Science Society American Journal* 50: 28–34.
- Ishikawa S, Wagatsuma T, Sasaki R, Ofei-Manu P. 2000. Comparison of the amount of citric and malic acids in Al media of seven plant species and two cultivars each in five plant species. *Soil Science Plant Nutrition* 46: 751-758.

- Iuchi S, Koyama H, Iuchi A, Kobayashi Y, Kitabayashi S, Kobayashi Y, Ikka T, Hiratama 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 of the USA* 104: 9900-9905.
- Johnson GCL, Esposito L, Barratt BJ, Smith AN, Heward J, Di Genova G, Ueda H, Cordell HJ, Eaves IA, Dudbridge F, Twells RCJ, Payne F, Hughes W, Nutland S, Stevens H, Carr P, Tuomilehto-Wolf E, Tuomilehto J, Gough SCL, Clayton DG, Todd JA. 2001. Haplotype tagging for the identification of common disease genes. *Nature Genet.* 29:233-237.
- Jones D, Gilroy S, Larsen P, Howell S, Kochian L. 1998. Effect of aluminum on cytoplasmic Ca^{2+} homeostasis in root hairs of *Arabidopsis thaliana* (L.). *Planta* 206: 378-387.
- Kaur N, Street K, Mackay M, Yahiaoui N, Keller B. 2008. Molecular approaches for characterization and use of natural disease resistance in wheat. *European Journal of Plant Pathology* 121: 387-397
- Kobayashi Y, Hoekenga OA, Itoh H, Nakashima M, Saito S, Shaff JE, Maron LM, Piñeros MA, Kochian LV, Koyama, H. 2007. Characterization of *AtALMT1* expression in aluminum-inducible malate release and its role for rhizotoxic stress tolerance in *Arabidopsis*. *Plant Physiology* 145: 843-852.
- Kochian, LV. 1995. Cellular mechanisms of aluminum toxicity and resistance in plants. *Annual Review of Plant Physiology and Plant Molecular Biology* 46: 237-260.
- Kochian LV, Piñeros MA, Hoekenga OA. 2005. The physiology, genetics and molecular biology of plant aluminum resistance and toxicity. *Plant Soil* 274:175-195.
- Konieczny A, Ausubel FM. 1993. A procedure for mapping *Arabidopsis* mutation using co-dominant ecotype-specific PCR-based markers. *Plant Journal* 3: 403-410.
- Larkin PJ. 1987. Calmodulin levels are not responsible for aluminum tolerance in wheat. *Australian Journal of Plant Physiology* 106: 1107-1114.
- Larsen PB, Geisler MJB, Jones CA, Williams KM, Cancel JD. 2005. *ALS3* encodes a phloem-localized ABC transporter-like protein that is required for aluminum tolerance in *Arabidopsis*. *The Plant Journal* 41: 353-363.
- Latha R, Rubia L, Bennett J, Swaminathan MS. 2004. Allele mining for stress tolerance genes in *Oryza* species and related germplasm. *Molecular Biotechnology* 27: 101-108.

- Li XF, Ma JF, Matsumoto H. 2000. Pattern of aluminum-induced secretion of organic acids differs between rye and wheat. *Plant Physiology* 123: 1537-1544.
- Ligaba A, Katsuhara M, Ryan PR, Shibasaka M, Matsumoto H. 2006. The *BnALMT1* and *BnALMT2* genes from rape encode aluminum-activated malate transporter that enhanced the aluminum resistance of plant cells. *Plant Physiology* 142: 1294-1303.
- Liu J, Magalhães JV, Shaff J, Kochian LV. 2009. Aluminum-activated citrate and malate transporters from the MATE and ALMT families function independently to confer Arabidopsis aluminum tolerance. *Plant Journal* 57: 389-399.
- Luo MC, Dvorak J. 1993. Molecular mapping of an aluminum tolerance locus on chromosome 4D of Chinese Spring wheat. *Euphytica* 91: 31-35.
- Ma JF, Hiridate S, Nomoto K, Iwashita T, Matsumoto H. 1997. Internal detoxification mechanism of Al form in the leaves. *Plant Physiology* 113: 1033-1039.
- Ma JF, Taketa S, Yang ZM. 2000. Aluminum tolerance genes on the short arm of chromosome 3R are linked to organic acid release in triticale. *Plant Physiology* 122: 687-694.
- Ma QF, Rengel Z, Kuo J. 2002. Aluminum toxicity in rye (*Secale cereal*): Root growth dynamics of cytoplasmatic Ca^{2+} in intact root tips. *Annals of Botany* 89: 241-244.
- Magalhães JV, Garvin DF, Wang Y, Sorrells ME, Klein PE, Schaffert RE, Li L and Kochian LV. 2004. Comparative mapping of a major aluminum tolerance gene in sorghum and other species in the Poaceae. *Genetics* 167:1905-1914.
- Magalhães JV, Liu J, Guimarães CT, Lana UGP, Alves VMC, Wang Y-H, Schaffert RE, Hoekenga OA, Piñeros MA, Shaff JE, Klein PE, Carneiro NP, Coelho CM, Trick HN, Kochian LV. 2007. A gene in the multidrug and toxic compound extrusion (MATE) family confers aluminum tolerance in sorghum. *Nature Genetics* 39:1156-1161.
- Magnavaca R, Gardner COE, Clark RB. 1987. Inheritance of aluminum tolerance in maize. In: Gabelman HW, Loughman BC (Ed.). *Genetic aspects of plant mineral nutrition*. Dordrecht: 201-212.
- Martin RB. 1992. Aluminum speciation in biology. In: Chadwick DJ, Whelan L. (Ed.) *Aluminum in Biology and Medicine*. John Wiley and Sons: p. 5-25.

- Matos M, Camacho MV, Pérez-Flores V, Pernaute B, Pinto-Carnide O, Benito C. 2005. A new aluminum tolerance gene located on rye chromosome 7RS. *Theoretical Applied Genetics* 111: 360-369.
- Miyasaka S, Bute J, Howell R, Foy C. 1991. Mechanism of aluminum tolerance in snapbeans. Root exudation of citric acid. *Plant Physiology* 96: 737-743.
- Nguyen VT, Burow MD, Nguyen HT, Le BT, Le TD, Paterson AH. 2001. Molecular mapping of genes conferring aluminum tolerance in rice (*Oryza sativa* L.). *Theoretical and Applied Genetics* 102:1002–1010.
- Nguyen VT, Nguyen BD, Sarkarung S, Martinez C, Paterson, AH, Nguyen HT. 2002. Mapping of genes controlling aluminum tolerance in rice: comparison of different genetic backgrounds. *Molecular Genetics and Genomics* 267:772–780.
- Nguyen BD, Brar DS, Bui BC, Nguyen TV, Pham LN, Nguyen HT. 2003. Identification and mapping of the QTL for aluminum tolerance introgressed from new source, *Oryza rufipogon* Griff., into indica rice (*Oryza sativa* L.). *Theoretical and Applied Genetics* 106:583–593.
- Omote H, Hiasa M, Matsumoto V, Otsuka V, Moriyama V. 2006. The MATE proteins as fundamental transporters of metabolic and xenobiotic organic cations. *Trends in Pharmacological Sciences* 27: 587–593.
- Osawa H, Matsumo H. 2001. Possible involvement of protein phosphorylation in aluminum-responsive malate efflux from wheat root apex. *Plant Physiology* 126: 411-420.
- Parker DR, Norvell WA, Chaney RL. 1995. GEOCHEM-PC: a chemical speciation program for IBM and compatible computers. In: Loeppert RH et al. (Ed.) *Chemical equilibrium and reaction models*. Soil Science Society of America: 253-269.
- Prada D. 2009. Molecular population genetics and agronomic alleles in seed banks: searching for a needle in a haystack? *Journal of Experimental Botany* 60: 2541-2552.
- Piñeros MA, Magalhães JV, Carvalho Alves VM, Kochian LV. 2002. The physiology and biophysics of an aluminum tolerance mechanism based on root citrate exudation in maize. *Plant Physiology* 129: 1194-1206.
- Raman H, Zhang K, Cakir M, Appels R, Garvin DF, Maron LG, Kochian LV, Moroni JS, Raman R, Imtiaz M, Drake-Brockman F, Waters I, Martin P, Sasaki T, Yamamoto

- Y, Matsumoto H, Hebb DM, Delhaize E, Ryan PR. 2005. Molecular characterization and mapping of ALMT1, the aluminum-tolerance gene of bread wheat (*Triticum aestivum* L.). *Genome* 48: 781-791.
- Riede CR, Anderson JA. 1996. Linkage of RFLP markers to an aluminum tolerance gene in wheat. *Crop Science* 36: 905-1028.
- Rocha, CMC. 1997. A região dos cerrados e as pesquisas desenvolvidas pela EMBRAPA Cerrados. In: Simpósio sobre os cerrados do meio norte. Anais. EMBRAPA-CPAMN, p. 57-80.
- Ryan PR, Shaff JE, Kochian LV. 1993. Aluminum toxicity in roots. Correlation among ionic currents, ion fluxes, and root elongation in aluminum-sensitive and aluminum-tolerant wheat cultivars. *Plant Physiology* 99:1193-1200.
- Ryan PR, Raman H, Gupta S, Horst WJ, Delhaize E. 2009. A second mechanism for aluminum resistance in wheat relies on the constitutive efflux of citrate from roots. *Plant Physiology* 149: 340-351.
- Sagghai-Marroof MA, Soliman KA, Jorgensen RA, Allard RW. 1984. Ribosomal DNA spacer length polymorphism in barley: Mendelian inheritance, chromosomal location and population dynamics. *Proceedings of the National Academy of Science of USA* 81: 8014-8018.
- Sasaki T, Yamamoto Y, Ezaki B, Katsuhara M, Ahn SJ, Ryan PR, Delhaize E, Matsumo H. 2004. A wheat gene encoding an aluminum-activated malate transporter. *Plant Journal* 37: 645-653.
- Sawaki Y, Iuchi S, Kobayashi Y, Kobayashi Y, Ikka T, Sakurai N, Fujita M, Shinozaki, Shibata D, Kobayashi M, Koyama H. 2009. *STOP1* regulates multiple genes that protect arabidopsis from proton and aluminum toxicities. *Plant Physiology* 150: 281-294.
- Schertz KF, Dalton LG. 1988. Sorghum. In: Fehr WR, Hadley HH. *Hybridization of crop plants*. Crop Science Society of America: 577-588.
- Schulz B e Kolukisaoglu HU. 2006. Genomics of plant ABC transporter: the alphabet of photosynthetic life forms or just holes in membranes? *FEBS Letters* 580: 1010-1016.
- Sham P, Bader JS, Craig I, O'Donovan M, Owen M. 2002. DNA Pooling: A tool for large scale association studies. *Nature Reviews Genetic* 3: 862–871.

- Shewayrga H, Jordan DR, Godwin ID. 2006. Genetic erosion and changes in distribution of sorghum (*Sorghum bicolor* L. Moench) landraces in north-eastern Ethiopia. *Plant Genetic Resources: Characterization and Utilization* 6: 1-10.
- Sivaguru M, Baluska F, Volkman D, Felle H, Horst W. 1999. Impacts of aluminum on the cytoskeleton of the maize root apex, short-term effects on the distal part of the transition zone. *Plant Physiology* 119: 1073-1082.
- Takeda K, Kariuda M, Itoi H. 1985. Blueing of sepal color of *Hydrangea macrophylla*. *Phytochemistry* 24: 2251-2254.
- von Uexküll HR, Mutert E. 1995. Global extent, development and economic impact of acid soils. In *Plant-Soil Interactions at Low pH: Principles and Management*, eds. RA Date, NJ Grundon, GE Raymet, ME Probert. Kluwer Academic Publishers: 5-19.
- Weltzein E, Rattunde HFW, Clerget B, Siart S, Toure A, Signard F. 2006. Sorghum diversity and adaptation to drought in West Africa. In: Jarvis D, Mar I, Sears L (eds) *Enhancing the use of crop genetic diversity to manage abiotic stress in agricultural production systems*. Soil Science Society of America: 253–269.
- Wight CP, Kibite S, Tinker NA, Molnar SJ. 2006. Identification of molecular markers for aluminum tolerance in diploid oat through comparative mapping and QTL analysis. *Theoretical and Applied Genetics* 112: 222-231.
- Wu P, Liao CY, Hu B, Yi KK, Jin WZ, Ni JJ, He C. 2000. QTLs and epistasis for aluminum tolerance in rice (*Oryza sativa* L.) at different seedling stages. *Theoretical and Applied Genetics* 100:1295-1303.
- Ye S, Dhillon S, Ke X, Collins AR, Day INM. 2001. An efficient procedure for genotyping single nucleotide polymorphisms. *Nucleic Acid Research* 29: 88-96.
- You FM, Huo N, Gu YQ, Luo M, Ma Y, Hane D, Lazo GR, Dvorak J, Anderson OD. 2008. BatchPrimer3: a high throughput web application for PCR and sequencing primer design. *BMC Bioinformatics*, 9:253.
- Yu J, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, Kresovich S, Buckler ES. 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics* 38: 203-208.

Zhang W, Ryan P, Tyerman S. 2001. Malate-permeable channels and cation channel activated by aluminum in the apical cells of wheat roots. *Plant Physiology* 103: 209-214.

Anexo 1

Tabela 1 – Distribuição dos acessos do Painei INRAN por *landraces*. Nomenclatura Local refere-se aos nomes dados a cada acesso de acordo com o local ao qual foi coletado. Esses acessos foram caracterizados morfologicamente pelo INRAN e agrupados em *landraces*.

Nome da <i>Landrace</i>	Número de acessos	Nomenclatura Local
BAGOBÁ	19	Bogoba/Bagoba
BDF	5	Babadia/ Babadia Fara/ Babba Da Fara / Babba Dia Fara
BDJ	4	Babadia Já/ Baba Da Ja/ Babba Dia Já
HK	1	Hakorin Karoua
JJ	27	Jenjari / Janjaré / Janjari
KAOURA	4	Kaoura
KIERMA	6	Kierma / Kieram
KORBIU	9	Korbi / Hamo Tawaye / Korbiyu / Zibi fo moizé hinka
MALLE	11	Karama Male / Male / Male Mezai / Mallé
MDK	18	Matche Dakomia / Matche Doukoumia / Balonkoma / El Koumniya / El Mandi / Kouroumbassaou/ Matché Da Kouminiya / Matché da Koumniya Fara / Matché da Koumniya Já
MDW	11	Makadoayo / Makaho Da Wayo / Makaho Dawayo / Makamniya / Makamniya Dawa
MG	7	El Galmi / El Kellin Kellin / El Mota / El Zahi / Enmatan Bana / Tagazani / Tallabani
MM	1	Mota
SOKOMBA	19	Sikombe / Sokomba / Sokombe Kire / Sokombe Koire / Socomba / Socomba Locale / Socomba Précoce / Socomba Tardif
Linhagens de Melhoramento	22	-