Repeat variants for the SbMATE transporter protect sorghum roots from aluminum toxicity by transcriptional interplay in cis and trans

Janaina O. Melo¹,b,¹,² Laura G. C. Martins²,¹,² Beatriz A. Barros³, Maiana R. Pimenta³,⁻,² Ubiraci G. P. Lana⁴, Christiane E. M. Duarte⁵, Maria M. Pastina⁵, Claudia T. Guimaraes⁵, Robert E. Schaffert⁵, Leon V. Kochian⁶, Elizabeth P. B. Fontes⁷,⁴ and Jurandir V. Magalhaes²,²,⁴

¹Embrapa Maize and Sorghum, Brazilian Agricultural Research Corporation, 35701-970 Sete Lagoas, MG, Brazil; ²Departamento de Biologia Geral, Universidade Federal de Minas Gerais, 31270-901 Belo Horizonte, MG, Brazil; ³Department of Biochemistry and Molecular Biology, BIOAGRO, Universidade Federal de Viçosa, 36570-000 Viçosa, MG, Brazil; and ⁴Global Institute for Food Security, University of Saskatchewan, SK S7N 4J8 Saskatoon, Canada

Edited by Julia Bailey-Serres, University of California, Riverside, CA, and approved November 14, 2018 (received for review May 16, 2018)

Acidic soils, where aluminum (Al) toxicity is a major agricultural constraint, are globally widespread and are prevalent in developing countries. In sorghum, the root citrate transporter SbMATE confers Al tolerance by protecting root apices from toxic Al³⁺, but can exhibit reduced expression when introgressed into different lines. We show that allele-specific SbMATE transcription activation occurs and is caused by factors located away from SbMATE. Using expression-QTL mapping and expression genome-wide association mapping, we establish that SbMATE transcription is controlled in a bipartite fashion, primarily in cis but also in trans. Multiallelic promoter transcription-activation and ChiP analyses demonstrated that intermolecular effects on SbMATE expression arise from a WRKY and a zinc finger-DHHC transcription factor (TF) that bind to and trans-activate the SbMATE promoter. A haplotype analysis in sorghum RILs indicates that the Tfs influence SbMATE expression and Al tolerance. Variation in SbMATE expression likely results from changes in tandemly repeated cis sequences flanking a transposable element (a miniature inverted repeat transposable element) insertion in the SbMATE promoter, which are recognized by the Al³⁺-responsive Tfs. According to our model, repeat expansion in Al-tolerant genotypes increases TF recruitment and, hence, SbMATE expression, which is, in turn, lower in Al-sensitive genetic backgrounds as a result of lower TF expression and fewer binding sites. We thus show that even dominant cis regulation of an agronomically important gene can be subjected to precise intermolecular fine-tuning. These concerted cis/trans interactions, which allow the plant to sense and respond to environmental cues, such as Al³⁺ toxicity, can now be used to increase yields and food security on acidic soils.

Significance

Aluminum (Al³⁺) on acidic soils, which represent half of the world’s agricultural lands, damages plant roots. In Africa, where sorghum is a staple food, 20% of the agricultural soils are acidic, significantly reducing yields. SbMATE confers sorghum Al tolerance via root citrate exudation into the soil, where citrate binds and detoxifies Al³⁺, but shows reduced expression in some genetic backgrounds. This phenomenon results from the action of a variable tandem repeat flanking a transposon in the SbMATE promoter. The authors declare no conflict of interest. This article is a PNAS Direct Submission. Published under the PNAS license.


Decisions in plant breeding often reflect the complex interplay between noncoding DNA sequences, acting locally in chromatin, and trans-regulatory elements driving gene expression via intermolecular interactions. Studies in Drosophila have shown that cis elements are evolutionarily important (1), whereas trans factors play pivotal roles in regulating plant stress responses in Arabidopsis (2). However, the manner in which cis and trans factors interact to control phenotypic expression is less clear.

Half the world’s agricultural soils are highly acidic (3), which solubilizes Al³⁺ into the soil solution, damaging plant roots and reducing yields. The Al-activated root citrate transporter SbMATE, which underlies Al tolerance via formation of nontoxic Al–citrate complexes in the rhizosphere (4), increased grain yield by 0.6 ton ha⁻¹ for sorghum grown on acidic soil (5). SbMATE SNPs were associated to sorghum grain yield production in West Africa (6), where sorghum is a staple food. This makes SbMATE important for global food security.

SbMATE expression is up-regulated by Al³⁺ in a time-dependent fashion (4) and is highly correlated with Al tolerance (7). A Tourist-like miniature inverted repeat transposable element (MITE) (8) and its flanking sequences, which are repeated in tandem, were found 2 kb upstream of SbMATE. Variation in the number of these tandem repeats in different sorghum lines was positively correlated with Al tolerance. Nevertheless, introgression of the Al³⁺ locus, where SbMATE resides (9), into Al-sensitive recurrent parents resulted in reduced SbMATE expression and Al tolerance; this suggested involvement of accessory loci acting in trans (7). The current study focuses on the elucidation of the role of the MITE repeats in SbMATE transcriptional regulation, and on the dissection of the genetic background effects that can reduce SbMATE expression.

| Author contributions: J.V.M. designed and supervised research; J.O.M., L.G.C.M., B.A.B., M.R.P., U.G.P.L., C.E.M.D., and R.E.S. performed research; M.M.P. contributed new reagents/analytic tools; J.O.M., L.G.C.M., B.A.B., M.R.P., U.G.P.L., C.E.M.D., M.M.P., C.T.G., L.V.K., and J.V.M. analyzed data; J.V.M. wrote the paper; R.E.S. assisted with genetic stock design and construction; and E.P.B.F. conceived and supervised TF validation in yeast and Arabidopsis. The authors declare no conflict of interest. This article is a PNAS Direct Submission. Published under the PNAS license.
| Data deposition: Data associated with this paper are available to download from the Dryad Digital Repository (doi:10.5061/dryad.18p3h04). The uploaded data (December 10, 2018) include SNP physical positions and association P values with Al tolerance and SbMATE expression. ¹J.O.M. and L.G.C.M. contributed equally to this work.²Present address: Departamento de Ciências Básicas (DCB), Universidade Federal dos Vales do Jequitinhonha e Mucuri, 39100-000 Diamantina, MG, Brazil.³Present address: Núcleo de Graduação em Agronomia, Universidade Federal do Sergipe, 49680-000 Nossa Senhora da Glória, SE, Brazil.⁴To whom correspondence may be addressed. Email: bbfontes@ufv.br or jurandir.magalhaes@embrapa.br. This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1808400115/-/DCSupplemental.
Results

**ShMATE Expression Is Influenced by the Genetic Background.** We studied global (i.e., joint expression of **ShMATE** alleles) and allele-specific expression of **ShMATE** (Fig. 1) to distinguish between *cis* and *trans* regulatory effects. This was done by using stocks derived from the low **ShMATE**-expressing Al-sensitive line BR012 crossed with SC656, a very Al-resistant line with high **ShMATE** expression (7). We generated a homoyzogous stock in which the **ShMATE** allele from SC656 was introgressed into the BR012 genetic background [SC656-near isogenic line (NIL)]. BR012 × SC656 and BR012 × SC656-NIL both have **ShMATE** in heterozygosity, but they differ in genetic backgrounds. Although BR012 × SC656 has a hybrid background, BR012 × SC656-NIL has the homozygous background of BR012. Compared with SC656, there was a consistent reduction of global **ShMATE** expression (Fig. 1A) in the SC656-NIL. **ShMATE** expression was higher in the hybrid background (BR012 × SC656) than in the BR012 background in the BR012 × SC656-NIL. Therefore, **ShMATE** expression is reduced in the Al-sensitive background.

**ShMATE** allele-specific expression was quantified relative to expression in the Al-tolerant and Al-sensitive parents, SC566 and BR012 (Fig. 1 B and C). Here, allele-specific expression was based on a Ta/TA single nucleotide polymorphism (SNP) in the first exon of **ShMATE**, with the A allele present in SC566 and the T allele present in BR012 (7). A marked, 210-fold up-regulation of the Al-sensitive allele (T) was observed in the BR012 × SC656 hybrid (Fig. 1B). Expression of the Al-sensitive allele of **ShMATE** was greatly reduced when present in the BR012 genome (BR012 × SC656-NIL) compared with the hybrid genome (BR012 × SC656; Fig. 1B). Expression changes for the Al-tolerant allele (A) were relatively subtle (Fig. 1C), but expression was reduced in all stocks harboring the BR012 genome compared with SC656.

These findings indicate that **ShMATE** expression is influenced by *trans*-acting factors whose favorable alleles are donated by the Al-tolerant line SC566. These factors are unlinked to **Al tolerant**, with the flanking, 1,749-bp and 2,010-bp sequences harboring the BR012 genome compared with SC656.

**Trans-Acting Loci Influencing **ShMATE** Expression Are Present Within an **ShMATE** Expression/Al Tolerance QTL on Chromosome 9.** Next, we undertook QTL mapping in a BR007 × SC283 recombinant inbred line (RIL) population. SC283 is highly tolerant to Al toxicity, whereas BR007 is highly Al-sensitive (Fig. 2A), and this RIL population was previously used to positionally clone **ShMATE** (4). The availability of a large population size for this highly contrasting cross (9) justified its choice for QTL mapping. Major QTL for both Al tolerance (Fig. 2B) and **ShMATE** expression (Fig. 2C) were colocated with **ShMATE** on chromosome 3. This indicates that higher **ShMATE** expression and Al tolerance in SC283 (7) is achieved predominantly in cis. A comparatively smaller Al tolerance and **ShMATE** expression QTL (eQTL) was detected at ~51 to ~54 Mb [2 < −log10(p) < 8 for the eQTL] on chromosome 9 (zoomed in; Fig. 2 B and C). Although other loci with similar *P* values are found elsewhere, the chromosome 9 QTL was chosen because of its clear joint effect on both Al tolerance and **ShMATE** expression. Multilocus mapping was also undertaken and revealed a possible interaction between the QTL on chromosomes 3 and 9 (SI Appendix, Table S1).

**ShMATE Promoter.** The structure of the **ShMATE** promoter region that contains the MITE insertion is shown in SI Appendix, Fig. S2A and B. The MITE element (unit "b") is flanked by 100-bp (unit "a") and 20-bp (unit "c") sequences. This MITE-containing a-b-c triplet (designated hereafter simply as “MITE repeats”) is followed by a single terminal (unrepeated) 100-bp “a” unit with either an 8-bp deletion (present in SC283 and Tx430) or a 12-bp deletion (present in BR012; SI Appendix, Fig. S2B). Henceforth, the 100-bp “a” sequence within the MITE repeats will be designated as the 100-bp repeat and the terminal, unrepeated units, as the 88- or 92-bp terminal.

Natural, allelic variation at the **ShMATE** promoter arises from tandem variations in the number of identical a-b-c units, which are present either as a singlet or as repeated units in different sorghum lines, with Al-tolerant lines showing in general more repeats compared with Al-sensitive lines (4). For example, the parents of the RIL population, BR007 (Al-sensitive) and SC283 (tolerant), have three and five repeats, respectively. For transactivation assays, we synthesized **ShMATE** promoters containing one MITE repeat (promoter from the Al-sensitive line Tx430, designated as Tx430p), four repeats (BR012p, from BR012, which is Al-sensitive), and five repeats (SC283p, from SC283, Al-tolerant), with the flanking, 1,749-bp and 2,010-bp sequences from the sorghum BAC where **ShMATE** resides (4). Promoter truncations were amplified from SC283.

![Fig. 1.](image-url) Global and allele-specific expression of **ShMATE** assessed with TaqMan probes. The global assay assesses the joint expression of **ShMATE** alleles, and allele-specific expression was based on an SNP (T, present in Al-sensitive BR012; A, present in Al-tolerant SC656) within **ShMATE**. Colored schematics indicate the genetic backgrounds (genome, rectangles) and the **ShMATE** alleles (squares): SC656 (red) and BR012 (green). (A) Global expression, allele-specific expression of the (B) Al-sensitive (T, green) allele and (C) Al-tolerant (A, red) allele, with the probes depicted on Top. The red/green gradient in A shows the joint expression of both the A and T alleles in stocks homozygous for **ShMATE**. Global relative expression values are fold changes relative to the Al-sensitive line. Expression of the T and A alleles are fold changes relative to expression in the parents, BR012 and SC656, which are homozygous for the T and A alleles, respectively. The sorghum genotypes were grown with 0.05 bars (in red) are drawn to scale (Top of the y axis).
Transcription Factors on the Chromosome 9 QTL Transactivate the SbMATE Promoter in Yeast. A qualitative analysis based on the yeast one-hybrid assay indicated that, in the Al tolerance/eQTL region, a WRKY-like transcription factor (TF), Shb09g023500 (SbWRKY1) at 53.14 Mb, and Shb09g021530, a gene encoding a zinc finger DHHC (zf-DHHC) domain-containing protein (SbZNFI) at 50.98 Mb (SI Appendix, Fig. S3 A and B), were both capable of trans-activating SbMATE promoter alleles harboring one, four, and five copies of the MITE repeats (SI Appendix, Fig. S2 C and D). In contrast, ShNFY1, a NFY-like TF (Shb09g022810) located within the same QTL, did not activate the SbMATE promoter, confirming the specificity of ShWRKY1 and SbZNFI transcriptional activation of SbMATE.

Qualitative promoter deletion analysis (SI Appendix, Fig. S2 C and D) showed that a proximal SC283 promoter fragment, extending to position −2102 relative to the SbMATE start codon (−2102pSC283, where “p” stands for promoter), was sufficient for trans-activation by both TFs, but trans-activation was lost when the 92-bp terminal was deleted (−2100pSC283).

SbZNFI and SbWRKY1 Bind Both to the 100-bp Repeat and to the 92-bp Terminal. Leaf protoplasts from transgenic Arabidopsis transformed with constructs containing the first 2,010 bp of the SbMATE promoter including either the 92-bp terminal (−2102p, from the SC283 promoter) or the MITE repeats in BR012 (−5299p; Fig. 3 A and B) were transformed along with constructs encoding 35S-driven YFP::TF cDNA. The immunoprecipitated chromatin (ChIP) fragments obtained with anti-GFP antibody were analyzed by PCR and qPCR. For both SbWRKY1-expressing and SbZNFI-expressing protoplasts, qPCR with IP DNA showed that a fragment within the 92-bp terminal, amplified with primers F1/R4 and F1/R5, was significantly enriched over the input (control) DNA, confirming binding to the 92-bp terminal (Fig. 3 A). In contrast, no enrichment was observed with primers annealing either to the 2,010-bp fragment (F2/R3), which lacks both the 100-bp repeat and the 92-bp terminal, or to the actin gene (endogenous control).

ChIP was also undertaken with the −5299 promoter from BR012, which has four MITE repeats followed by the terminal 88-bp fragment (Fig. 3 B). Amplification of IP DNA with primers F1/R1, which are specific to the 100-bp repeat, was significantly enriched over the input DNA (Fig. 3 B). Collectively, the results in Fig. 3 A and B show that SbWRKY1 and SbZNFI bind both to the 100-bp “a” unit within the MITE repeats and to the 92-bp terminal. The amplification profiles of IP DNA (SI Appendix, Fig. S4) with primers flanking (F3 and F4) and within (R1) the 100-bp repeat (lanes 1 and 3 in −5299p) confirm such binding. Because the R1 primer in Fig. 3 B does not anneal to the 88-bp terminal in −5299p, we cannot rule out that the TFs do not bind to that fragment because of its additional 4-bp deletion compared with the 92-bp terminal fragment in −2102p (see SI Appendix, Fig. S2 B for “a” unit alignments).

The Number of MITE Repeats Correlates With Enhanced SbWRKY1 and SbZNFI Transactivation Activity. Because transactivation assays in yeast are qualitative, we quantified transactivation activity in Arabidopsis. Protoplasts were isolated from Arabidopsis transformed with the truncated −2102 promoter from Al-tolerant SC283 containing the 92-bp terminal (Fig. 3 C). In addition, promoter alleles containing one (Tx430p) and four (BR012p) copies of the MITE repeats followed by the 92-bp and 88-bp terminal (Fig. 3 D), respectively, were tested. Reporter gene-specific activity was higher with protoplasts isolated from Arabidopsis cotransformed with the 92-bp terminal and with SbWRKY1 and SbZNFI compared with the negative controls (Fig. 3 C). In addition, elimination of the 92-bp terminal abolished reporter gene activity (SI Appendix, Fig. S5 A-H), which is consistent with the transactivation results in yeast.

Reporter gene activity was significantly higher with promoter alleles containing four (BR012p) compared with one (Tx430p) copies of the MITE repeats (Fig. 3 D). For both promoter constructs, a cotransactivation assay indicated a synergistic mode of action, with a greater effect on SbMATE promoter activity when both TFs are present (SbWRKY1 + SbZNFI) compared with their individual effects (Fig. 3 D). These two promoters also vary both for the number of MITE repeats and for the presence of an additional 4-bp deletion specifically in the 88-bp unique terminal of the four-repeat promoter. However, we established previously that both TFs bind to the 100-bp “a” repeat and to the unrepeat ed 92-bp “a” terminal (Fig. 3 A and B). In view of that, even if both TFs did not bind to the terminal “a” 88-bp fragment of the four MITE-repeat promoter, this promoter would still harbor four binding sites for the TFs within its MITE region, in contrast to only two binding sites in the one MITE-repeat promoter (within its single 100-bp sequence and in the 92-bp terminal). It is thus unlikely that the additional 4-bp deletion was the cause of the higher transactivation in the 4 MITE-repeat (BR012p) promoter compared with the 1 MITE-repeat promoter (Tx430p). These results strongly suggest that increased binding site abundance in promoters where the number of MITE repeats has expanded leads to enhanced TF recruitment.
**ShbZNF1** and **ShbWRKY1** Alleles from Al-Tolerant and Al-Sensitive Lines Are Differentially Regulated by $\text{Al}^{3+}$.

For clarity, alleles are designated in the text with the gene names (TF is used when referring to both **ShbWRKY1** and **ShbZNF1**), followed by numbers indicating the allele donors. We looked at the expression profiles of **ShbWRKY1** and **ShbZNF1** alleles derived from unrelated Al-tolerant (SC566 and SC283) and Al-sensitive (BR007 and BR012) lines (10) (Fig. 4A). Genetic backgrounds are depicted by the colored rectangles beneath Fig. 4A. The SC566- and SC566-NILs carry the respective **ShbMATE** allele (**ShbMATE**<sub>SC566</sub>, depicted by squares) in the BR012 genetic background, and as such, their TF alleles (green ovals) are the same as the ones in BR012 (**TF<sub>012</sub>**).

In the presence of $\text{Al}^{3+}$, both **ShbWRKY1** and **ShbZNF1** were more highly expressed in Al-tolerant (SC566 and SC283) compared with Al-sensitive (BR012 and BR007; Fig. 4A) lines. In NILs in which tolerant **ShbMATE** alleles (**ShbMATE**<sub>SC283</sub> and **ShbMATE**<sub>SC566</sub>) were introgressed into the Al-sensitive BR012 background (SC283- and SC566-NILs), TF expression was reduced compared with their respective Al-tolerant donors (SC283 and SC566). These responses are similar to their transcriptional target **ShbMATE**, which also showed reduced expression in the SC566- and SC283-NILs compared with the Al-tolerant parents (**SI Appendix**, Fig. S6, 5 d, and ref. 7). Strikingly, the **ShbWRKY1** allele derived from the Al-tolerant line SC566, **ShbWRKY1**<sub>SC566</sub>, was markedly up-regulated by $\text{Al}^{3+}$ (**ShbWRKY1**<sub>SC566</sub> also shows a consistent tendency for $\text{Al}^{3+}$ up-regulation, but slighter). In contrast, the Al-sensitive alleles in BR007 (**ShbWRKY1**<sub>BR007</sub>) and BR012 and NILs (**ShbWRKY1**<sub>BR012</sub>) were strongly down-regulated by $\text{Al}^{3+}$ (Fig. 4A). **ShbWRKY1** and **ShbZNF1** exhibited different transcriptional responses to $\text{Al}^{3+}$ in different Al-tolerant lines, as Al-induced **ShbWRKY1** expression was greater in SC566 compared with SC283, whereas **ShbZNF1** Al$^{3+}$ up-regulation and expression was higher in SC283.

**In Sorghum, ShbWRKY1 and ShbZNF1 Alleles Derived from SC283 (Al-Tolerant) Increase ShbMATE Expression.** A genetic analysis in the BR007 × SC283 RIL population was conducted using **ShbWRKY1** (W) and **ShbZNF1** (Z) gene-specific markers, which were designed based on polymorphisms that differentiate the **TF<sub>283</sub>** and **TF<sub>007</sub>** alleles (Fig. 4B). This was done to select RILs with different combinations between parental alleles of **ShbZNF1** and **ShbWRKY1** (i.e., TF haplotypes). For this analysis, we compared **ShbMATE** expression and Al tolerance of RILs selected to contain both TF alleles from the Al-tolerant parent (**TF<sub>283</sub>** × **Z**/W$^+$), from the Al-sensitive parent (**TF<sub>007</sub>** × Z$^+$W$^-$), and one TF allele from each parent (Z$^-$W$^+$, Z$^+$/W$^-$). A linear regression model fit to all haplotype classes indicated that both TFs enhanced **ShbMATE** expression (Fig. 4C and **SI Appendix**, Fig. S7). **ShbMATE** expression for the double homozygous
haplotypic class containing the Al-tolerant SC283 allele at both loci ($Z^+W^+$) produced the largest increase in $S_bMATE$ expression compared with the other haplotype classes, and was more than 2.3-fold higher than the class containing Al-sensitive alleles from Br007 at both loci ($Z^-W^+$; Fig. 4C).

$S_bZNF1$ exerted a stronger effect on both $S_bMATE$ expression (Fig. 4C) and Al tolerance (Fig. 4D) compared with $S_bWRKY1$, the individual effect of which on Al tolerance was below the statistical power of our haplotype-based approach. This result is likely population-specific, resulting from stronger Al$^{3+}$ up-regulation of $S_bZNF1$ expression compared with $S_bWRKY1$ specifically in the Al-tolerant parent of the RIL population, SC283 (Fig. 4A).

**Time-Dependent Expression in Root Apices Exposed to Al$^{3+}$ for $S_bWRKY1$ and $S_bZNF1$ Is Similar to $S_bMATE$.** Al-induced expression of both $S_bWRKY1$ and $S_bZNF1$ was higher in root apices of Al-tolerant lines compared with the rest of the root system and shoots (SI Appendix, Fig. S8). This response favoring preferential expression in root apices was larger for $S_bWRKY1$ in SC66 compared with $S_bZNF1$ in SC283, which are the genotypes that display the highest expression of each TF gene under Al$^{3+}$ (Fig. 4A). A time-course analysis indicated a general trend for time-dependent increase in TF expression in Al-tolerant lines between 1 and 5 d of Al exposure, which was higher for $S_bWRKY1$ (4.2–4.8-fold) compared with $S_bZNF1$ (1.1–1.6-fold). In general, preferential, time-dependent expression in root apices exposed to Al$^{3+}$ for $S_bZNF1$ and $S_bWRKY1$ parallels the $S_bMATE$ expression measured under the same period in Al. For the Al-sensitive lines, Br007 and Br012, $S_bWRKY1$ and $S_bZNF1$ expression in the presence of Al$^{3+}$ decreased over the same 1-, 3-, and 5-d periods (SI Appendix, Fig. S6).

**Discussion**

We discovered that $S_bMATE$ expression is influenced by a cis-acting tandemly repeated sequence flanking a MITE insertion upstream of $S_bMATE$, which provides sites in which $S_bWRKY1$ and $S_bZNF1$ bind and transcriptionally regulate $S_bMATE$. Possible binding motifs in the binding fragment are the recognition core for Dof1/MNB1a zf-TFs (11, 12) and a motif similar to the WNT-box, where a WRKY TF has been shown to bind (13) (see SI Appendix, Table S3 for cis elements identified in silico).

Our results indicate that $S_bMATE$ and $S_bWRKY1$ are coregulated ($r = 0.3; P = 0.08$; SI Appendix, Fig. S6). This suggests that $S_bWRKY1$ functionality evolved to regulate $S_bMATE$ expression in response to Al$^{3+}$, which is consistent with the active and adaptable nature of Group III C2H-type zs (14). $S_bZNF1$ is a DHHC-like S-acyl transferase zf, and such proteins have been implicated in abiotic stress tolerance (15). $S_bZNF1$ is preferentially expressed in roots of Al-tolerant lines, but its expression is localized to the root tip to a lesser extent than $S_bWRKY1$. This pattern may reflect the more general physiological role of DHHC proteins, stemming from the DHHC cognate function in increasing protein hydrophobicity (16).

Our quantitative analysis of transactivation in *Arabidopsis* protoplasts positively associated $S_bWRKY1$ and $S_bZNF1$ transactivation activity and the number of MITE repeats in the $S_bMATE$ promoter, suggesting a dosage dependency. Hence, we propose that the singular (17), tandemly repeated structure of the MITE repeats has led to differential TF recruitment (SI Appendix, Fig. S9 A and B), resulting in the previously observed positive correlation between the size of the MITE insertion region and the Al tolerance phenotype (4).

Synergistic transactivation in *Arabidopsis* protoplasts, in conjunction with our haplotype analysis of $S_bWRKY1$ and $S_bZNF1$ in RILs derived from parents harboring different TF alleles, suggest that, in sorghum, these TFs cooperate to increase Al tolerance as measured by relative net root growth (%RNRG). Different letters indicate statistical differences (Fisher’s least significant difference, $\alpha = 0.08$).

Fig. 4. Transcription factor expression profile and effect on $S_bMATE$ expression and Al tolerance. (A) $S_bWRKY1$ and $S_bZNF1$ expression in Al-tolerant (SC283 and SC566) and Al-sensitive (Br007 and Br012) lines, and in the SC566-NIL and SC283-NIL (SC566/SC283 $S_bMATE$ in the Br012 background). Colored schematics indicate the genetic backgrounds (genome, rectangles), the $S_bMATE$ alleles (squares), and TF alleles (ovals), along with the Al tolerance phenotype (7). Plants were grown on 100 mM Al$^{3+}$ for 5 d in nutrient solution at pH 4.0 (brackets denote free Al$^{3+}$ activity estimated with GEOCHEM; see SI Appendix, Supplementary Methods), and the root apices (1 cm) were collected. Values are mean ±SD (n = 2). Least significant difference (Fisher’s LSD, $\alpha = 0.10$) bars (in red) are drawn to scale (Top of the y axis). This experiment was repeated with similar results (SI Appendix, Fig. S6, 5 d; n = 3). (B) $S_bWRKY1$ and $S_bZNF1$ polymorphisms in the RIL parents, Br007 (007, Al-sensitive) and SC283 (283, Al-tolerant), which were used to select RILs with all combinations of TF alleles (i.e., TF haplotypes). (C) $S_bWRKY1$ and $S_bZNF1$ effect on $S_bMATE$ expression estimated based on RILs homozygous for the SC283 (Al-tolerant) alleles at both TF loci ($Z^+W^+$), for the Br007 (Al-sensitive) allele ($Z^-W^+$), or showing alternate TF alleles ($Z^+/W^-$ and $Z^-/W^+$). Significant differences based on 5% (***) and 12% (*) confidence intervals. A linear regression model fit to haplotype $S_bMATE$ expression was highly significant (***: $\alpha = 0.01$). Physical (Mb) and genetic (cM) distances between TFs are depicted at the Top. (D) Effect of $S_bWRKY1$ and $S_bZNF1$ on Al tolerance as measured by relative net root growth (%RNRG). Different letters indicate statistical differences (Fisher’s least significant difference, $\alpha = 0.08$).
with RILs fixed for the BR007 alleles (Fig. 4C). The estimated TF effect on Al tolerance (18% increase; Fig. 4D: $Z'/W'$ vs. $Z'/W$) is equal to the decrease in Al tolerance when the SC283 allele of $ShMATE$ was introgressed into the background of the Al-sensitive line, BR012 (~18% in SC283 vs. SC283-NIL; figure 2 in ref. 7), which we show here has low-expressing alleles for both $ShWRKY1$ and $ShZNF1$ (Fig. 4A). This suggests that allelic variation at the TF loci is responsible for our previously observed genetic background effects, which lead to reduced expression of Al-tolerant alleles of $ShMATE$ when introgressed into Al-sensitive backgrounds (7).

Our cis/trans interaction model (SI Appendix, Fig. S9) depicts a possible compensatory mode of action for cis and trans effects in highly Al-tolerant lines. Accordingly, the loss of one MITE repeat in $ShMATE$ was introgressed into the background of the Al-sensitive line, BR012 (~18% in SC283 vs. SC283-NIL; figure 2 in ref. 7), which we show here has low-expressing alleles for both $ShWRKY1$ and $ShZNF1$ (Fig. 4A). This suggests that allelic variation at the TF loci is responsible for our previously observed genetic background effects, which lead to reduced expression of Al-tolerant alleles of $ShMATE$ when introgressed into Al-sensitive backgrounds (7).

Compensatory cis/trans effects (2, 18), which appear to be a rather widespread mechanism that cells use to stabilize gene expression (19), are implicated in coevolution between cis and trans mutations. Although cis variants may provide a more stable control of gene expression under stress, trans regulation is important for environmental responses (2). In the absence of Alloading, both $ShWRKY1$ and $ShZNF1$ are down-regulated in Al-tolerant genotypes (SC283 and SC66), reducing $ShMATE$ expression presumably when root citrate release, which can be costly to the plant, is not needed because of the lack of Al toxicity. Therefore, the interplay between cis-acting elements and TFs that are responsive to Al stress may be advantageous, as a result of a balancing effect on $ShMATE$ expression, which would otherwise be more inflexibly controlled in cis, resulting in genetic load related to the loss of certain cis-regulatory elements involving root citrate release.

In light of the molecular nature of cis and trans variants that modulate $ShMATE$ expression, we can now both predict and circumvent genetic background effects that reduce $ShMATE$ expression to increase grain yield production on acidic, Al toxic soils across the world.

Materials and Methods

Genetic Stocks. Development of NILs, RILs, and hybrid stocks, and the association panel used for GWAS (20), are described in SI Appendix, Supplementary Methods.

Al Tolerance in Hydroponics. Al tolerance was assessed based on root growth inhibition, relative net root growth (RNRG), in nutrient solution with and without (21) $10^{-3}$ M Al3+ at pH 4.0 (20) (brackets denote free Al3+ activity estimated with GEOCHEM; see SI Appendix, Supplementary Methods).

Gene Expression via Quantitative RT-PCR. Sorghum plants were grown in nutrient solution ±(27) $10^{-3}$ M Al3+ for 1 and/or 3 and 5 d, depending on the experiment. Gene expression was assessed either with the TaqMan Gene Expression or SYBR Green assay (Applied Biosystems). Allele-specific expression was assessed (TaqMan) based on an A/T SNP in the first exon of $ShMATE$ (7), with the A allele present in SC566 and the T allele present in all other lines. See SI Appendix, Supplementary Methods.

QTL Mapping in a RIL Population. Al tolerance and $ShMATE$ expression data were obtained in nutrient solution with (27) $10^{-3}$ M Al3+ at pH 4.0 for 5 d and QTL mapping with SNP markers was undertaken with TASSEL (GLM) and by multiple regression.

Genomewide Association Mapping. Genomewide association mapping was undertaken based on a mixed linear model (O + K) with TASSEL. SNP markers were tested for associations with Al tolerance [RRNG; SI Appendix, Table S4 (20)] and $ShMATE$ expression (ΔΔCT) at 5 d of Al exposure.

Transactivation Assays. Full-length promoter fragments and trans-factor cDNA (v.1.4 of the sorghum genome) sequences were commercially synthesized, and transactivation assays were conducted as described in SI Appendix, Supplementary Methods. The experiments were repeated four times with similar results.

Transcription Factor Effects on $ShMATE$ Expression and Al Tolerance via Hyalotype Analysis in an RIL Population. $ShWRKY1$ and $ShZNF1$ genotyping was based on an indel and a SNP polymorphism, respectively, as described in the SI Appendix, Supplementary Methods.

Chromatin Immunoprecipitation Assay. Leaf protoplasts were isolated from transgenic Arabidopsis thaliana plants transformed with different $ShMATE$ promoter fragments and then transformed with the 35S:YFP-$ShWRKY1$ and 35S:YFP-$ShZNF1$ vectors. See SI Appendix, Supplementary Methods.

Acknowledgments. We thank Veridiana Cano for assistance with allele-specific expression assays and William Lucas (University of California, Davis) for critically reading the manuscript. We acknowledge grants from the CGIAR Generation Challenge Program, the Embrapa Macroprogram, the Fundação de Amparo a Pesquisa do Estado de Minas Gerais, and the National Council for Scientific and Technological Development.