Duplicate and Conquer: Multiple Homologs of *PHOSPHORUS-STARVATION TOLERANCE1* Enhance Phosphorus Acquisition and Sorghum Performance on Low-Phosphorus Soils[1[C][W][OPEN]]

Barbara Hufnagel, Sylvia M. de Sousa, Lidianne Assis[2], Claudia T. Guimaraes, Willmar Leiser, Gabriel C. Azevedo, Barbara Negri, Brandon G. Larson, Jon E. Shaff, Maria Marta Pastina, Beatriz A. Barros, Eva Weltzien, Henry Frederick W. Rattunde, Joao H. Viana, Randy T. Clark[3], Alexandre Falcão, Rodrigo Gazaffi, Antonio Augusto F. Garcia, Robert E. Schaffert, Leon V. Kochian, and Jurandir V. Magalhaes*

Departamento de Biologia Geral, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, 31270–901, Brazil (B.H., C.T.G., G.C.A., J.V.M.); Embrapa Maize and Sorghum, Sete Lagoas, Minas Gerais, 35701–970, Brazil (B.H., S.M.d.S., L.A., C.T.G., G.C.A., B.N., M.M.P., B.A.B., J.H.V., R.E.S., J.V.M.); International Crops Research Institute for the Semi-Arid Tropics, BP 320 Bamako, Mali (W.L., E.W., H.F.W.R.); Institute of Plant Breeding, Seed Science, and Population Genetics, University of Hohenheim, 70593 Stuttgart, Germany (W.L.); Departamento de Bioengenharia, Universidade Federal de São João del-Rei, Praça Sao Joao del-Rei, Minas Gerais, 36301–160, Brazil (B.N.); Robert W. Holley Center for Agriculture and Health, United States Department of Agriculture-Agricultural Research Service, Cornell University, Ithaca, New York 14850 (B.G.L., J.E.S., R.T.C., L.V.K.); University of Campinas, Campinas, Sao Paulo, 13083–852, Brazil (A.F.); and Departamento de Genética, Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, Sao Paulo, 13400–970, Brazil (R.G., A.A.F.G.)

Low soil phosphorus (P) availability is a major constraint for crop production in tropical regions. The rice (*Oryza sativa*) protein kinase, PHOSPHORUS-STARVATION TOLERANCE1 (OsPSTOL1), was previously shown to enhance P acquisition and grain yield in rice under P deficiency. We investigated the role of homologs of *OsPSTOL1* in sorghum (*Sorghum bicolor*) performance under low P. Association mapping was undertaken in two sorghum association panels phenotyped for P uptake, root system morphology and architecture in hydroponics and grain yield and biomass accumulation under low-P conditions, in Brazil and/or in Mali. Root length and root surface area were positively correlated with grain yield under low P in the soil, emphasizing the importance of P acquisition efficiency in sorghum adaptation to low-P availability. *SbPSTOL1* alleles reducing root diameter were associated with enhanced P uptake under low P in hydroponics, whereas *Sb03g006765* and *Sb03g0031680* alleles increasing root surface area also increased grain yield in a low-P soil. *SbPSTOL1* genes colocalized with quantitative trait loci for traits underlying root morphology and dry weight accumulation under low P via linkage mapping. Consistent allelic effects for enhanced sorghum performance under low P between association panels, including enhanced grain yield under low P in the soil in Brazil, point toward a relatively stable role for *Sb03g006765* across genetic backgrounds and environmental conditions. This study indicates that multiple *SbPSTOL1* genes have a more general role in the root system, not only enhancing root morphology traits but also changing root system architecture, which leads to grain yield gain under low-P availability in the soil.

Increasing food production is one of the major global challenges in dealing with continuously growing population and food consumption (Godfray et al., 2010). One of the major obstacles to improve crop production in tropical regions is phosphorus (P) deficiency caused by P fixation in the soil clays. P is one of the most important plant nutrients, contributing approximately 0.2% of a plant's dry weight, and is a component of key organic molecules such as nucleic acids, phospholipids, and ATP (Schachtman et al., 1998). On tropical soils, even when the total amount of soil P is high, its bioavailability is low due to P fixation by aluminum and iron oxides in clay minerals (Marschner, 1995) and immobilization into organic forms (Schachtman et al., 1998). Approximately half of the world's agricultural lands occurs on low-P soils (Lynch, 2011); hence, crop adaptation to P insufficiency should be a major breeding target to enable sustainable agricultural production worldwide. In addition, because phosphate rock fertilizer

[1] This work was supported by the Generation Challenge Programme (grant no. G7010.03.06), the Embrapa Macroprogram, the German Federal Ministry for Economic Cooperation and Development, the U.S. National Science Foundation (grant no. DBI–0820624), the Fundação de Amparo a Pesquisa do Estado de Minas Gerais (Ph.D. fellowship to B.H.), the Fundação de Amparo a Pesquisa do Estado de São Paulo (to A.F.), the National Council for Scientific and Technological Development (to J.V.M.), and the Dryland Cereals Research Program.

[2] Present address: Universidade Federal do Acre, Rio Branco, Acre 69920-900, Brazil.

[3] Present address: DuPont Pioneer, Johnston, IA 50131.

* Address correspondence to jurandir.magalhaes@embrapa.br.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Jurandir V. Magalhaes (jurandir.magalhaes@embrapa.br).

[C] Some figures in this article are displayed in color online but in black and white in the print edition.

[W] The online version of this article contains Web-only data.

[OPEN] Articles can be viewed online without a subscription.

www.plantphysiol.org/cgi/doi/10.1104/pp.114.243949

is a nonrenewable resource that is being depleted by agricultural demands, increasing fertilizer prices negatively impact agriculture, particularly for small-holder farmers in developing countries in the tropics and subtropics (Cordell et al., 2009; Sattari et al., 2012). In sorghum (*Sorghum bicolor*), breeding strategies for low-P adaptation have been developed based on multienvironment trials in West Africa, indicating the importance of undertaking selection in low-P soil conditions (Leiser et al., 2012a, 2012b). Therefore, developing crops with greater ability to grow and maintain satisfactory yields on soils with reduced P availability is expected to substantially improve food security worldwide.

Tolerance to P deficiency in plants can be achieved by mechanisms underlying both P acquisition and P internal utilization efficiency (Parentoni and Souza Junior, 2008). One of the major mechanisms that plants have evolved to overcome low-P availability is to maximize the ability of the roots to acquire and absorb P from the soil. Plants can mobilize P through the exudation of organic acids, acid phosphatases, and ribonucleases, resulting in enhanced P availability and uptake (Hinsinger, 2001; Ryan et al., 2001; Dakora and Phillips, 2002; Hammond and White, 2008; Ma et al., 2009; Pang et al., 2009). Another strategy to cope with low-P availability is to increase the soil volume accessed by root systems by forming mycorrhizal symbioses (Li et al., 2012; Smith and Smith, 2012; Rai et al., 2013). Due to low-P mobility on tropical soils, changes in root architecture and morphology enhance P uptake by facilitating soil exploration (Williamson et al., 2001; Ho et al., 2005; Walk et al., 2006; Svistoonoff et al., 2007; Lynch, 2011; Ingram et al., 2012; Niu et al., 2013). Root structural changes leading to higher P uptake include increased root hair growth (Yan et al., 2004; Haling et al., 2013; Lan et al., 2013) and length and enhancing lateral root over primary root growth (Williamson et al., 2001; Wang et al., 2013). In addition, increased root surface area is achieved by a combination of reduced root diameter and enhanced elongation of relatively thinner roots (Fitter et al., 2002). There is both intraspecific and interspecific genetic variation for P deficiency tolerance in crop species (Lynch and Brown, 2001, 2012; Mudge et al., 2002; Paszkowski et al., 2002; Rausch and Bucher, 2002; Huang et al., 2011; Zhang et al., 2011; Leiser et al., 2014a) that can be explored to develop P-efficient cultivars.

In rice (*Oryza sativa*), *Phosphorus uptake1* (*Pup1*), a major quantitative trait locus (QTL) for P deficiency tolerance donated by an *aus*-type Indian variety, Kasalath, was mapped to the long arm of chromosome 12 (Ni et al., 1998; Wissuwa et al., 1998, 2002; Heuer et al., 2009). Near-isogenic lines bearing the Kasalath allele at *Pup1* showed 3-fold higher P uptake and grain yield in low-P trials compared with the recurrent parent, cv Nipponbare, which is intolerant to P starvation (Wissuwa and Ae, 2001). Following high-resolution mapping of *Pup1*, comparative sequence analyses of homologous bacterial artificial chromosomes showed that a Kasalath genomic fragment contained several genes not present in cv Nipponbare, highlighting an approximately 90-kb deletion in the cv Nipponbare reference genome that encompassed the *Pup1* locus (Heuer et al., 2009). Within this insertion/deletion, *OsPupK46-2*, a gene encoding a Ser/Thr kinase of the Receptor-like Protein Kinase LRK10L-2 subfamily, was found to enhance grain yield and P uptake in rice lines overexpressing this gene, indicating that this protein kinase underlies the *Pup1* locus (Gamuyao et al., 2012). *OsPupK46-2*, which is now designated *PHOSPHORUS-STARVATION TOLERANCE1* (*OsPSTOL1*), was found to be up-regulated in the root tissues of tolerant near-isogenic lines under P-deficient conditions and was shown to increase P uptake by a physiological mechanism based on the enhancement of early root growth and development. Furthermore, lines overexpressing *OsPupK46-2* showed an approximately 30% grain yield increase in comparison with the null lines, suggesting that *PSTOL1* has potential for molecular breeding applications to improve crop performance under low-P conditions. Consistent with the proposed physiological mechanism underlying *OsPSTOL1*, the superior performance of the transgenic lines was related to enhanced root dry weight, root length, and root surface area (Gamuyao et al., 2012).

Sorghum is the world's fifth most important cereal crop and is a staple food for more than half a billion people. It is widely adapted to harsh environmental conditions, and more specifically, to arid and semiarid regions of the world (Doumbia et al., 1993, 1998). It has been estimated that rice diverged from its most recent common ancestor with sorghum and maize (*Zea mays*) approximately 50 million years ago (Kellogg, 1998; Paterson et al., 2000, 2004; Paterson, 2008). About 60% of the genes in the sorghum genome are located in syntenic regions to rice (Paterson et al., 2009), emphasizing the potential for using comparative genomics for cross-species identification of genes underlying abiotic stress tolerance in the grass family. Here, we applied association analysis to specifically study the role of sorghum homologs of rice *OsPSTOL1* in tolerance to P starvation in sorghum. Single-nucleotide polymorphisms (SNPs) within *PSTOL1* homologs in sorghum, collectively designated *SbPSTOL1*, were significantly associated with grain yield under low-P conditions and also root morphology and root system architecture (RSA) traits phenotyped from hydroponically grown plants. Under low P, *SbPSTOL1* genes increased biomass accumulation and P content in the African landrace panel and grain yield in the sorghum association panel phenotyped in a low-P Brazilian soil. This suggests a stable effect across environments and sorghum genotypes that potentially can be used for molecular breeding applications. QTL mapping with a large sorghum recombinant inbred line population was used to validate the association results, indicating that *SbPSTOL1* homologs colocalize with QTLs related to root morphology and performance under low P. Our results indicate that *SbPSTOL1* homologs have the ability to enhance P uptake and sorghum performance in low-P soils by a mechanism related not only to early root growth enhancement, as was the case for rice *OsPSTOL1*, but also by modulating RSA.

RESULTS

Phenotypic Characterization of the Sorghum Association Panels

Two association panels were used in this study: the Sorghum Association Panel Subset (SAPst), which is composed of both tropical converted and breeding accessions, and the West African Association Panel (WAP), consisting of landraces and breeding lines primarily from Mali, Niger, Senegal, and Burkina Faso. The SAPst was phenotyped for grain yield performance in a low-P soil in Brazil, root morphology, P uptake, and biomass accumulation in a paper pouch system where nutrients are supplied to the root systems hydroponically, and RSA in nutrient solution using a mesh system as detailed in "Materials and Methods." The WAP was phenotyped in pots for biomass and P uptake under low P in the soil.

Extensive genotypic variation for agronomic performance under low-P conditions was observed for both association panels. Broad sense heritabilities were high, with values of 0.63 for grain yield under low P in the SAPst and 0.42 and 0.59 for P content in single tillers and single tiller biomass, respectively, in the WAP. For assessing root morphology traits in a paper pouch system, our RootReader2D system (Clark et al., 2013) was used to process the root systems images and the WhinRhizo software was used to automatically calculate a number of root morphology traits including root length, root diameter, volume of fine roots and root surface area under low-P conditions (de Sousa et al. 2012). Heritability estimates for the root morphology traits were high ($h^2 > 0.72$), indicating good experimental precision for these measurements.

Root surface area was highly correlated both with root length ($r = 0.95$, $P < 0.001$) and the volume of fine roots ($r = 0.77$, $P < 0.001$; Fig. 1) in the SAPst. The volume of fine roots showed strong but comparatively lower correlation coefficients of approximately 0.6 with root length and diameter, whereas root diameter tended to be independent from root length and weakly associated with root surface area. Root length and root surface area were positively correlated with grain yield ($r = 0.11$–0.12, $P < 0.10$; Fig. 1).

Supplemental Table S1 shows the results of a correlation analysis among the same root morphological traits shown in Figure 1 and traits related to P uptake and biomass accumulation in hydroponics. Significant positive correlations were found between root length and P content as well as between root length and biomass accumulation for roots and shoots, whereas root diameter was negatively correlated with these same traits. These results indicate that greater root length combined with smaller root diameter contributed to enhanced P uptake and biomass accumulation for sorghum plants subjected to low P in hydroponics.

Identification of *PSTOL1* Homologs in Sorghum and SNP Discovery

A sequence similarity search using the amino acid sequence of rice PSTOL1 (OsPSTOL1; GenBank accession no. BAK26566) identified a large gene family containing approximately 100 members in the sorghum genome. Based on BLASTP and BLASTN, six likely *PSTOL1* homologs in sorghum (*SbPSTOL1*; E < *e*-35 and identity > 50%) were selected (Table I). The sorghum homolog with the closest sequence similarity to *OsPSTOL1* was *Sb07g002840*, which is located on sorghum chromosome 7, and another five highly similar homologs are located on sorghum chromosome 3. Similar to OsPSTOL1, all six sorghum PSTOL1 proteins (SbPSTOL1) are predicted to have a Ser/Thr kinase domain (Pfam identifier PF00069.18) and are members of the Corn root protein kinase1 family (CRPK1) (Takezawa et al., 1996).

A phylogenetic analysis including OsPSTOL1, the six SbPSTOL1 proteins in Table I, and related kinases from rice, Arabidopsis (*Arabidopsis thaliana*), and maize revealed the presence of three distinct clades (Fig. 2). All sorghum PSTOL1-like proteins clustered together in the largest clade (highlighted in Fig. 2), which also included maize, Arabidopsis, and rice proteins. In this clade, a branch including five sorghum homologs and OsPSTOL1 was identified (split depicted by the asterisk in Fig. 2), also depicting a close relationship between OsPSTOL1, Sb07g002840, and the maize protein AC193632.2. The two remaining clades were composed entirely of other rice *PSTOL1* homologs except for a single sequence from maize.

Protein structural predictions indicate that the kinase domain is commonly present in OsPSTOL1 and in the six selected SbPSTOL1 proteins (Fig. 3; Supplemental Fig. S1). However, distinctly different features were predicted for the SbPSTOL1 proteins, namely a signal peptide suggesting a secretory pathway, a transmembrane domain, and cell wall interaction domains. These are a Cys-rich wall-associated receptor kinase galacturonan-binding (GUB-WAK-bind) domain, which is the extracellular part of the Ser/Thr kinase that is expected to bind to the cell wall pectins, and a conserved Cys-rich wall-associated receptor kinase C-terminal (WAK_association) domain, located C terminal to the GUB_WAK_bind domain. These cell wall association domains were neither found in OsPSTOL1 nor in the closely related proteins, Sb07g002840 and AC193632.2, and differed in the remaining SbPSTOL1 proteins. While the WAK_association domain was found in all SbPSTOL1 proteins except for Sb07g002840, the GUB_WAK_bind domain was specifically found in Sb03g006765 and Sb03g031700. In agreement with the structural predictions, the sequence alignment of the sorghum and rice proteins indicates that the sequences were conserved mostly at the kinase domain (Supplemental Fig. S1), which is typical from receptor-like kinases (RLKs; Zhang et al., 2005).

Despite the high similarity among *SbPSTOL1* homologs, particularly with regard to the kinase domain, visual inspection of the multiple alignments of polypeptide

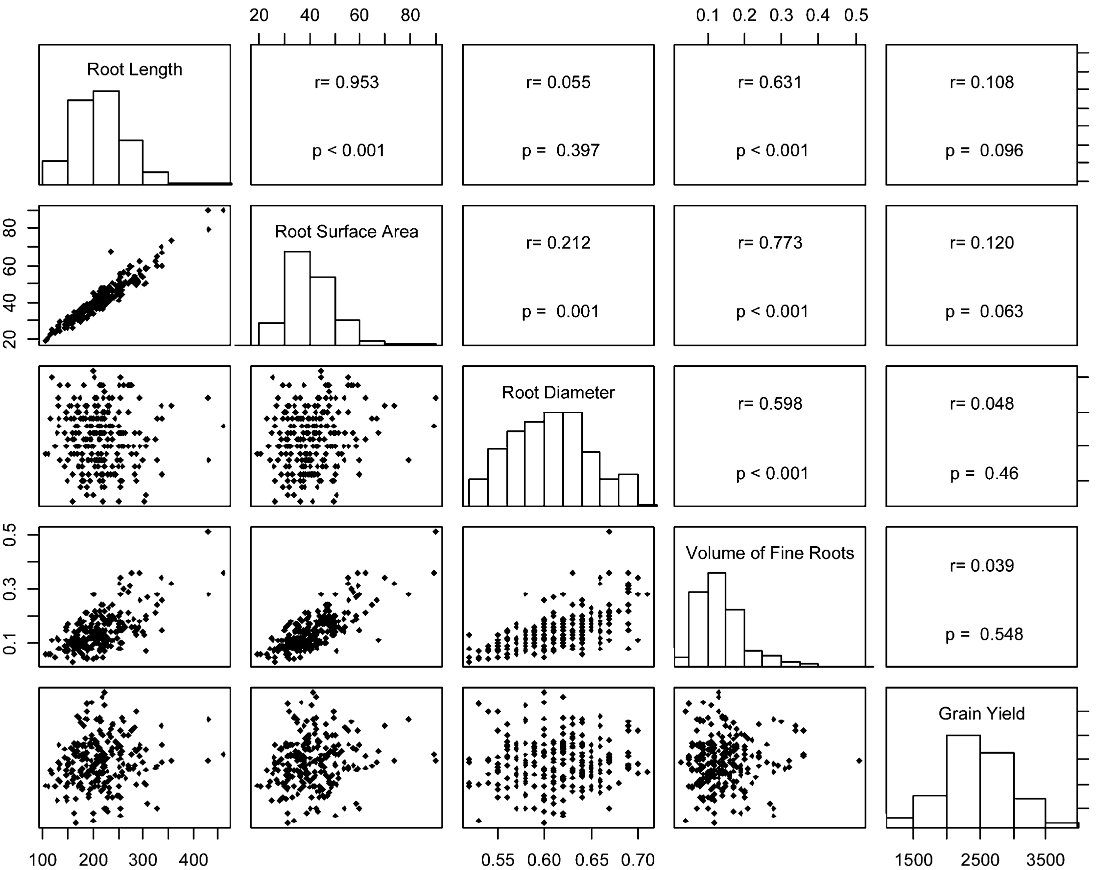

Figure 1. Correlation matrix of root length (mm), root surface area (cm^2), average root diameter (mm), and volume of fine roots (mm^3), all assessed in hydroponics-grown plants, as well as grain yield (kg ha^{-1}) for plants grown on a low-P soil for the SAPst. Correlation coefficients (*r*) and *P* values are shown.

sequences confirmed that PrimerBlast was effective in identifying paralog-specific primers. We were able to identify 31 paralog-specific SNPs across the six sorghum homologs, with the number of SNPs per gene ranging from five to eight for *Sb03g006765*, *Sb03g031690*, *Sb03g031700*, and *Sb07g002840*, whereas only one and two SNPs were discovered within *Sb03g031670* and *Sb03g031680*, respectively.

Population Structure and Association Model Selection

The presence of population structure in the SAPst was assessed by estimating ΔK, the second-order rate of change of the log-likelihood of the data obtained with STRUCTURE (Pritchard et al., 2000), Ln(k), divided by its SD (Evanno et al., 2005). Upon plotting ΔK as a function of the number of subpopulations, a break in

Table I. *Selected SbPSTOL1 homologs*

Physical locations in bp on sorghum chromosomes 3 and 7 are shown. Sequence similarity searches in sorghum were performed as described in "Materials and Methods" using the NCBI (www.ncbi.nlm.nih.gov) and Phytozome (version 1.4; www.phytozome.net) databases using the OsPSTOL1 sequence as query. E, Expect value of the alignment. Chromosomal positions are denoted by the first two numbers after *Sb* in the gene designations.

Gene	Location	Size	BLASTN		BLASTP	
			E	Identity	E	Identity
	bp			%		%
Sb03g006765	7,009,497–7,012,497	3,001	2.00e-37	83	2.00e-108	55
Sb03g031670	60,080,358–60,084,458	4,100	2.00e-123	71	1.00e-129	70
Sb03g031680	60,085,127–60,088,181	3,054	3.00e-121	71	2.00e-130	70
Sb03g031690	60,103,142–60,107,812	4,670	7.00e-73	73	6.00e-76	69
Sb03g031700	60,110,148–60,113,362	3,214	9.00e-78	68	1.0e-115	63
Sb07g002840	3,011,700–3,016,004	4,304	0	76	6.0e-143	73

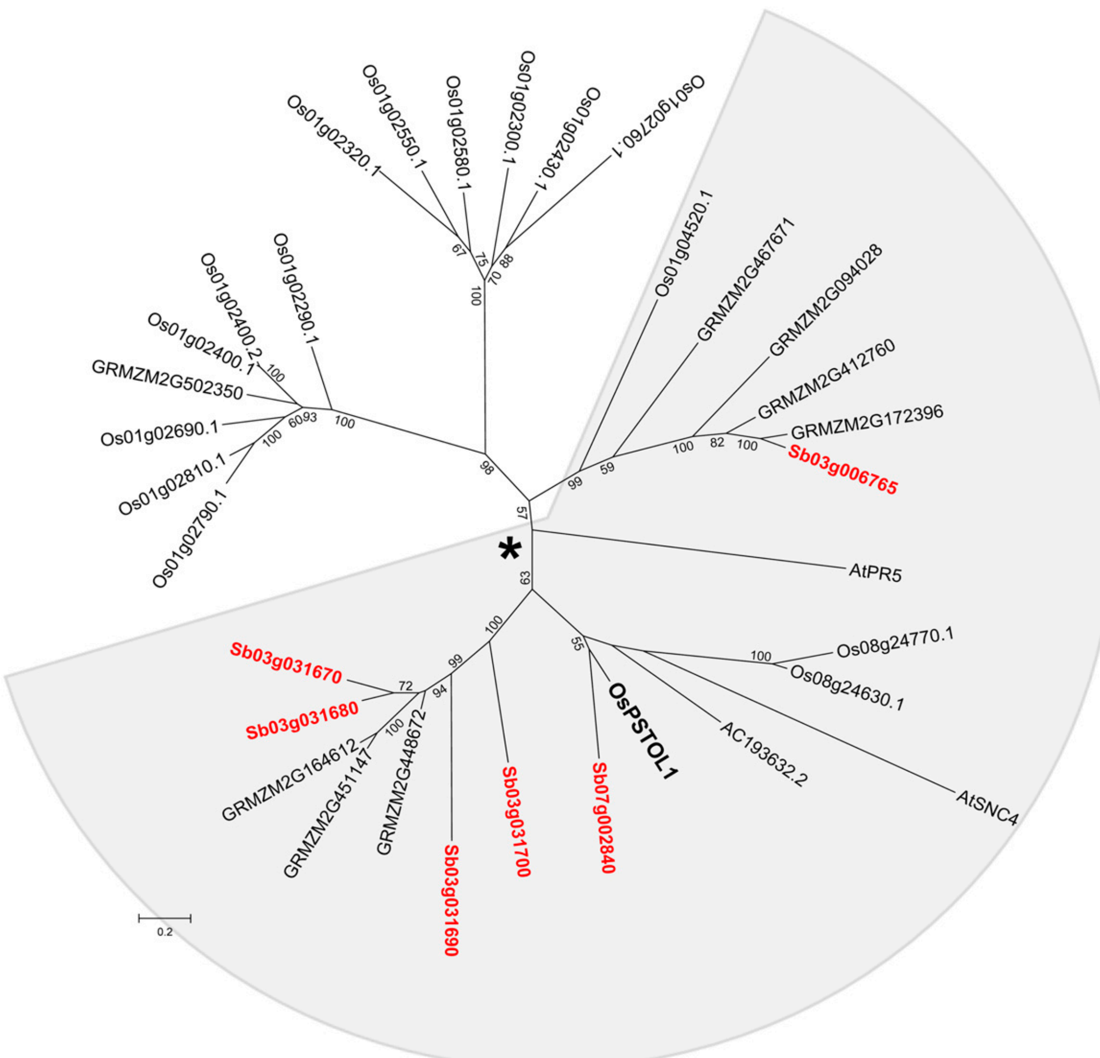

Figure 2. Relationship among *PSTOL1* homologs. The phylogenetic tree comprises PSTOL1-like proteins in sorghum (Sb), rice (LOC_Os), maize (GRMZM and AC), and Arabidopsis (At) in addition to OsPSTOL1. The highlighted clade comprises selected sorghum proteins with high identity with OsPSTOL1, and the asterisk depicts a branch split including five of the six selected SbPSTOL1 proteins. Amino acid sequences of 33 proteins were used. The unrooted phylogenetic tree was constructed using the maximum likelihood method and displayed using the MEGA6 software (http://www.megasoftware.net/index.html). [See online article for color version of this figure.]

the slope is expected at the most likely subpopulation number. Supplemental Figure S2A reveals an evident change in ΔK with four subpopulations, indicating that this is a reasonable level of differentiation for the SAPst. As described previously, the pattern of genetic diversity in sorghum largely reflects racial and geographical origins (Casa et al., 2008; Caniato et al., 2011; Morris et al., 2013). The largest subpopulation, with 112 accessions, comprised mostly caudatum sorghums, followed by subpopulations with prevalence of durra, guinea/kafir, and breeding lines, with 93, 50, and 32 accessions, respectively. The probability distribution under the null hypothesis was then obtained by plotting the *P* values resulting from association analysis against the cumulative *P* values (Supplemental Fig. S2B). The naïve model, without correction for population structure (Q) or kinship (K), produced a markedly nonuniform distribution of *P* values resulting in the highest type I error rate. Considering the Q model, 9.5% of all *P* values are under a $P < 0.05$ threshold, indicating improvement in type I error control over the naïve model, but still there was considerable detection of false-positive associations. The model incorporating familial relatedness (K) resulted in efficient type I error control, with 5.7% of *P* values under $P < 0.05$. Because the Q+K model led to a negligible improvement in false-positive control, the K model was selected for association analysis with *SbPSTOL1* homologs. Similar results for model selection were found by Casa et al. (2008) when evaluating the complete Sorghum Association Panel for

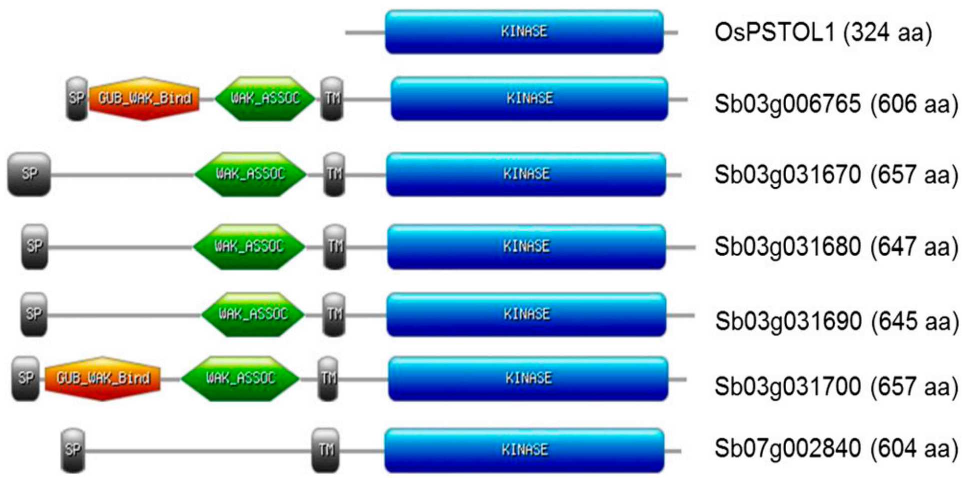

Figure 3. Schematic representation of protein domains of OsPSTOL1 and SbPSTOL1. Depicted are the predicted signal peptide (SP), the GUB_WAK_Bind domain, the wall-associated receptor kinase C-terminal domain (WAK_ASSOC), the transmembrane domain (TM), and the Ser/Thr kinase domain (KINASE). Full complementary DNA sequences were obtained for Sb07g002840, Sb03g031670, Sb03g031690, and Sb07g006765. Reverse transcription-PCR of *SbPSTOL1* genes was undertaken with root tissues from the sorghum line, BR007, after 13 d in nutrient solution (Magnavaca et al., 1987) with low P (2.5 μM P) as described for root morphology analysis. Roots were collected, and total RNA was isolated using the RNeasy Plant Mini Kit (Qiagen; http://www.qiagen.com/). For Sb03g006765, a start codon located 72 bp upstream of the first codon (TCA) in the coding sequence reported in Phytozome (http://www.phytozome.net/sorghum) was considered for structural predictions. aa, Amino acids. [See online article for color version of this figure.]

different agronomic traits, with the K model resulting in a good approximation to a uniform distribution of *P* values.

Association Analysis of Root Morphology, P Uptake, and Traits Related to Performance under Low P

Association analysis was undertaken following the same procedure used above for type I error simulations but using SNPs within *SbPSTOL1* genes and the models that provided the appropriate error control, as detailed in "Materials and Methods." Based on the SAPst, a total of 23 SNPs from all six *SbPSTOL1* genes were significantly associated with different traits related to root morphology, P uptake, and biomass accumulation in hydroponics in addition to grain yield under low P in the field (Table II). SNPs within the *SbPSTOL1* homologs *Sb03g006765* and *Sb03g031690* were associated with root diameter and root surface area, whereas two *Sb03g031690* SNPs were weakly associated with the volume of fine roots.

For the SNP loci associated with root diameter changes, the alleles underlying smaller root diameter for *Sb03g006765* (allele G at SNP 2002), *Sb03g31690* (allele G at SNP 4085), and *Sb03g002840* (allele C at SNP 3975) were associated with increased biomass accumulation and/or P uptake traits (Supplemental Table S2). This is consistent with the negative correlation observed between root diameter and biomass accumulation and P uptake traits (Supplemental Table S1), indicating that genotypes with smaller root diameter performed better under low-P conditions in hydroponics. The only exception to this trend was SNP 2644 within *Sb03g0031690*, where the A allele increased root diameter and also root P concentration. However, because the A allele was also associated with higher root surface area, changes in root surface area likely more strongly influence the higher root P concentration for SNP 2644 within *Sb03g0031690*.

For *Sb03g031690*, SNP 2305 was consistently associated with root surface area ($0.01 < P < 0.05$) and P concentration in the root, with the same allele, T, positively contributing to both traits (Supplemental Table S2). The results for both *Sb03g031670* and *Sb03g031700* in principle do not point toward an effect in root morphology for these genes. However, because we were able to identify only one and two paralog-specific SNPs, respectively, for these genes, we cannot rule out the possibility that the lack of association with root traits was due to the few SNPs that were tested.

Six SNPs within *Sb03g006765* and one SNP within *Sb03g031680* were associated with grain yield in the low-P soil in the SAPst (Tables II and III). Considering the SNP effects, the resulting grain yield increase varied from 6.4% to 8.1% relative to the grain yield under low-P conditions. Within *Sb03g031680*, SNP 1541 was significantly associated with grain yield (Table II), resulting in the highest increase in yield among all the SNPs associated with this trait, with the A allele increasing grain yield by 200 kg ha^{-1} (Table III). A possible positive effect in P acquisition via changes in root morphology was observed for *Sb03g031680*, where SNP 1541 was jointly associated with root surface area and grain yield, with the A allele positively contributing to both traits (Table III; Supplemental Table S2). Another case of consistent allelic effects between root surface area and grain yield was found for *Sb03g006765*, where

Table II. *Association results for SNP loci within SbPSTOL1 homologs in the SAPst phenotyped for root morphology traits (Hydroponics) and for grain yield performance in a low-P soil (Field) using the K model*

Numbers represent the proportion of the phenotypic variation explained by each SNP in percentage. Asterisks are as follows: ****$P < 0.001$, ***$0.001 \leq P < 0.01$, **$0.01 \leq P < 0.05$, and *$0.05 \leq P < 0.1$.

Gene	Position	Hydroponics									Field
		Root Morphology			P Uptake and Biomass Accumulation						
		Root Diameter	Volume of Fine Roots	Root Surface Area	Root P	Root P Content	Shoot P	Shoot P Content	Root Dry Weight	Shoot Dry Weight	Grain Yield
	bp	*mm*	*mm*^{3}	*cm*^{2}	*g kg*$^{-1}$	*mg*	*g kg*$^{-1}$	*mg*	*g*		*kg ha*$^{-1}$
Sb03g006765	1,912										2.00**
	1,998			2.46**							1.95**
	2,002	1.80**			2.28*		2.14*	2.23*		2.65**	
	2,042										1.53*
	2,067										2.46**
	2,073										1.65*
	2,141										2.27**
Sb03g031670	1,312						1.70**	1.78*			
Sb03g031680	1,541			1.83*							2.26*
	1,591					2.25*		2.17*	2.51**	2.42**	
Sb03g031690	2,305			2.74**	7.56****						
	2,644	4.09***		2.24**	3.24**						
	4,020		1.22*								
	4,085	1.10*	1.38*				1.09*				
	4,121										
	4,170						1.41**				
Sb03g031700	2,096				3.71***						
	2,304								1.42*		
Sb07g002840	3,783					1.80*		2.05*	2.12*	3.06**	
	3,795							1.92*	2.02*	2.69**	
	3,810								1.96*	2.09*	
	3,939					1.68*			2.03*	2.32**	
	3,975	1.00*						2.03*	2.13*	2.98**	

the C allele at SNP 1998 increased both traits (Table III; Supplemental Table S2). Overall, the fraction of the total phenotypic variation explained by each SNP (r^2) ranged from 1 (*Sb07g002840*_3975; with root diameter) to 7.5% (*Sb03g031690*_2305; P concentration in the roots).

Association analysis was also conducted with the WAP with the same SNP loci used in the SAPst (Table IV). Multiple associations were detected for SNPs within *Sb03g006765* with P concentration and content in the shoots in addition to single tiller biomass accumulation. Strikingly, in the SAPst, the same SNPs were associated with grain yield under low-P conditions (Table II). The single SNP within *Sb03g031670*, SNP 1312, was associated with P content in single tillers in the WAP and shoot P content in the SAPst grown on hydroponics (Tables II and IV). For *Sb03g031690* and *Sb03g031700*, associations were detected for P content and traits related to biomass accumulation (Table IV). The phenotypic variance explained in the WAP ranged from 2.8% (*Sb03g006765_2042*; P concentration in the shoot) to 6.8% (*Sb03g006765_2042*; P content of single tiller biomass).

Genome Structure and Linkage Disequilibrium

The genome structure of the *SbPSTOL1* homologs selected for association analysis revealed a possible double tandem duplication event encompassing a rather narrow physical region of approximately 33 kb at position approximately 60 Mb on chromosome 3 (Fig. 4). This duplication includes *Sb03g031670* and *Sb03g031680* at positions 60,080 to 60,088 kb and *Sb03g031690* and *Sb03g031700* at 60,103 to 60,113 kb. The remaining *SbPSTOL1* gene on chromosome 3, *Sb03g006765*, is located at position 7 Mb, thus toward the beginning of chromosome 3 and physically distant from the other homologs at position 60 Mb. As expected, all SNPs within *Sb03g006765* were under strong linkage disequilibrium (LD), which was also in general the case for the cluster at position 60 Mb. In that region, occasional low LD was observed mostly for SNPs within different genes, occasionally within genes belonging to possibly different tandem duplicates, and frequently involving *Sb03g031700* at the outer edge of the 60-Mb region. As expected, SNPs within *Sb03g006765* were not in LD with SNPs within the genes in the 60-Mb cluster, indicating that association results between the 7- and 60-Mb region were independent. The *SbPSTOL1* homolog on chromosome 7, *Sb07g002840*, was located at position 3 Mb, and complete LD was observed for SNPs in this gene (Supplemental Fig. S3).

Consistency of SNP Effects across Association Panels Phenotyped in Brazil and Africa

Next, we compared the SNP effects estimated in the SAPst and WAP, looking for alleles consistently

Table III. *SNP effects on grain yield for SbPSTOL1 homologs*

The alleles marked in boldface contribute positively to grain yield. Genotypic effects were estimated as deviations from the alleles with effects set as 0.

Gene	Position	Genotype	Effect
	bp		*kg ha^{-1}*
Sb03g006765	1,912	**A**	173.3
		G	0
	1,998	**C**	169.1
		T	0
	2,042	**G**	158.4
		C	0
	2,067	**G**	181.7
		C	0
	2,073	**C**	154.5
		T	0
	2,141	**T**	180.9
		G	0
Sb03g031680	1,541	**A**	200.1
		G	0

contributing to better performance under low-P conditions. In general, the effects in grain yield in the SAPst ranged from 154 to 200 kg ha^{-1} (Table III). Five SNPs within *Sb03g006765* were strongly associated with P content in single tillers and biomass accumulation in the WAP grown on low-P conditions (Table IV). All of these five SNPs, which were under complete LD within *Sb03g006765* in the SAPst (Fig. 4), had the same alleles positively contributing to biomass accumulation and shoot P content in the WAP (Fig. 5) and grain yield in the SAPst grown on a low-P soil (Table III). For example, the A and C alleles at SNP loci 1912 and 1998, respectively, enhanced single tiller biomass and total P content in single tillers in the WAP (Fig. 5) and grain yield in the SAPst (Table III), with similar results being observed for SNP loci 2042, 2067, and 2141.

SbPSTOL1 genes belonging to each one of the tandem duplications at position 60 Mb on chromosome 3, *Sb03g031670* and *Sb03g031690*, also showed consistent allelic effects across the two association panels for certain traits. This was the case for *Sb03g031690*_2305, where the T allele had a positive effect on root P concentration in the SAPst and single tiller biomass accumulation in the WAP, although no association was found for P content in the WAP (Supplemental Fig. S4). The A allele of *Sb03g031670* at SNP 1312 increased P content in both the SAPst and the WAP.

Validation of *PSTOL1* Homologs via Biparental Mapping

Due to the reduced power of association analysis to detect associations with quantitative traits such as P efficiency, validation of the association results was undertaken within a higher LD context, in a large sorghum recombinant inbred line (RIL) population whose parents contrasted for root morphology traits. This population was genotyped by genotyping-by-sequencing (GBS; Elshire et al., 2011) and phenotyped for root morphology, biomass accumulation and P uptake traits in hydroponics, following the same procedures used for the SAPst. A fixed-effects linear model was applied to test for associations between the different SNP loci and the phenotypes, as detailed in "Materials and Methods."

Sb03g006765, within which SNP 2002 was associated with root diameter in the SAPst (Table II), closely colocalized with a QTL for shoot dry weight accumulation in the RIL population ($P < 0.05$), being separated from the SNP with the highest $-\log_{10}(P)$ by 1 Mb (Fig. 6). Interestingly, our data suggest the presence of a possible QTL for root diameter at position 3 Mb, thus separated from *Sb03g006765* by 4 Mb, although the significance for this QTL ($P = 0.146$) was below the $P < 0.05$ threshold. In addition, a highly significant QTL controlling root length, root surface area, and volume of fine roots was found 12 to 13 Mb from the 60-Mb physical position on chromosome 3, suggesting that the *SbPSTOL1* cluster at position 60 Mb underlies these QTLs. With regard to the *SbPSTOL1* gene on chromosome 7 and using a $P < 0.05$ threshold, *Sb07g002840* was near a QTL for root dry weight, located 6 Mb from the SNP with the highest $-\log_{10}(P)$ (Supplemental Fig. S5). In addition, this *SbPSTOL1* gene was also 0.6 Mb from a possible QTL for root diameter at position 3.6 Mb, although the P value for this QTL ($P = 0.114$) was below the $P < 0.05$ threshold.

Table IV. *Association results for SNP loci within SbPSTOL1 homologs in the WAP phenotyped in pots*

Numbers represent values of explained phenotypic variation. Asterisks are as follows: ****$P < 0.001$, ***$0.001 \leq P < 0.01$, **$0.01 \leq P < 0.05$, and *$0.05 \leq P < 0.1$.

Gene	SNP Position	P Shoot	P Content of Single Tiller Biomass	Single Tiller Biomass	Biomass per Plant
	bp	*mg kg^{-1}*	*mg*	*g*	*g*
Sb03g006765	1,912	2.72**	6.66****	5.19***	
	1,998	2.72**	6.66****	5.20***	
	2,042	2.77**	6.80****	5.41***	
	2,067		3.59**	3.94**	
	2,141	2.71**	6.66****	5.19***	
Sb03g031670	1,312		4.16**		
Sb03g031690	2,305			1.91*	
Sb03g031700	2,096		1.61*		2.50**
	2,304		1.77*		2.86**

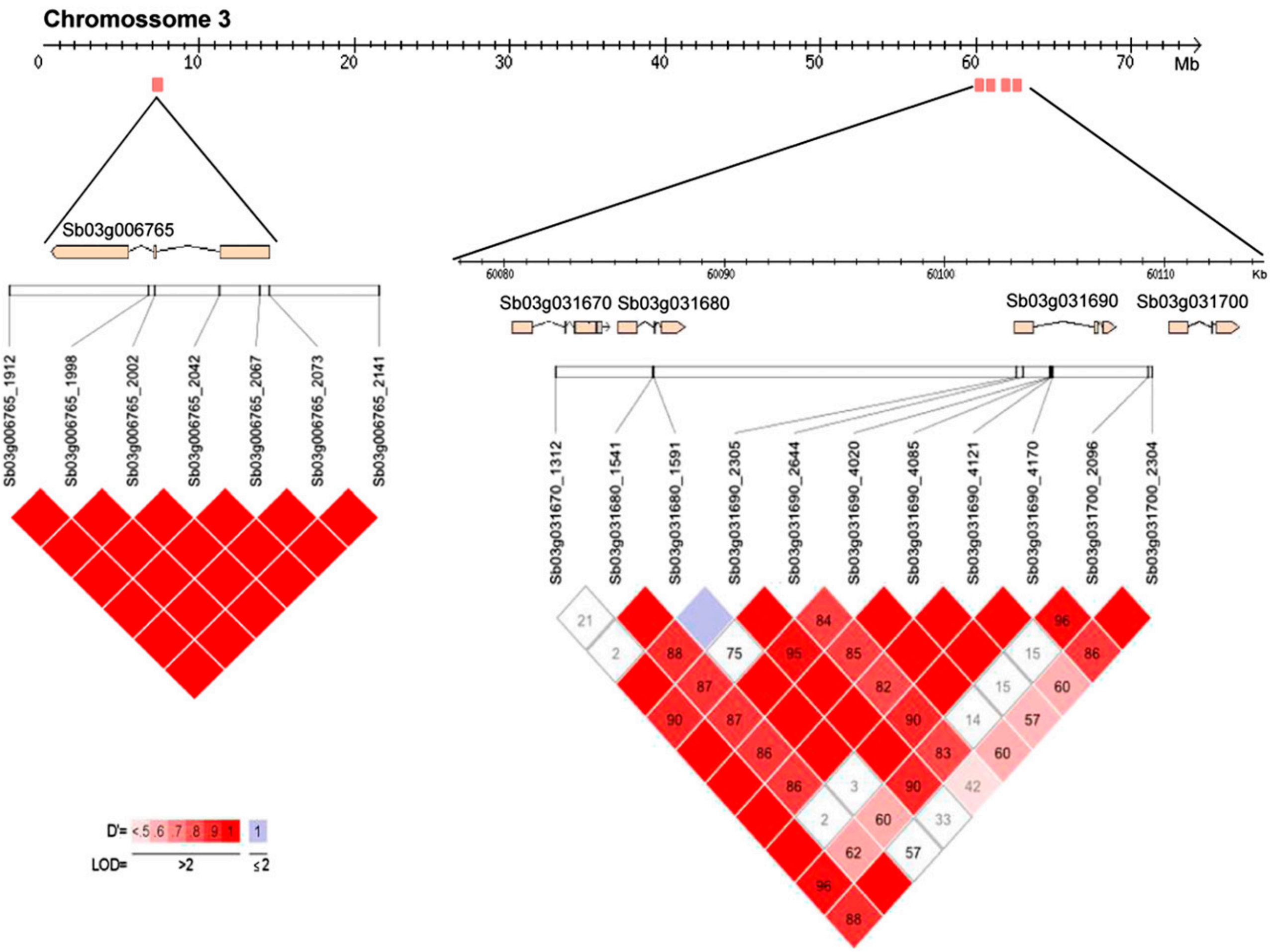

Figure 4. LD across *SbPSTOL1* homologs on chromosome 3. The standardized LD coefficient (D′; Lewontin, 1964) was estimated among all pairwise combinations between associated SNP loci within *Sb03g006765* (position 7 Mb) and the *SbPSTOL1* cluster at position 60 Mb. Squares shows D′ between two loci and the log of the odds (LOD) of two loci being in LD (Barrett et al., 2005). The red color indicates strong (D′ > 0.8) and statistically significant (LOD > 2) LD. Numbers inside squares represent D′ values expressed in percentages (squares without numbers represent D′ values of 1). LD plots were generated using Haploview version 4.2 (Barrett et al., 2005).

Association Analysis of RSA Traits in the SAPst

RSA traits were assessed in the SAPst using our RootReader3D platform as described by Clark et al. (2011) to image and generate three-dimensional (3D) reconstructions for the entire root systems of the sorghum accessions in the SAPst. Then, the RootReader3D software was used to compute 19 different RSA traits, which were then employed as different phenotypes in the association analysis. Large genotypic variation in RSA traits was observed among sorghum accessions in this panel. Heritability estimates for the four RSA traits shown in Figure 7A, which are related to root branching and depth of the root system, varied from 0.16 (for centroid) to 0.39 (median number of roots). Bushiness, which is the ratio of the maximum number of roots divided by the median number of roots (Iyer-Pascuzzi et al., 2010), was negatively correlated with centroid, which is the vertical position of the center of mass of the entire root system ($r = -0.24$, $P < 0.001$). Bushiness was also negatively correlated with the median number of roots in a root system ($r = -0.46$, $P < 0.001$). Specific root length, which is the ratio of total root length to the volume of the entire root system and a measure of the benefit-to-cost ratio for root systems, was negatively correlated with the median number of roots ($r = -0.27$, $P < 0.001$), whereas centroid was positively correlated with the median number of roots ($r = 0.42$, $P < 0.001$).

Association analysis performed on the SAPst indicated that *SbPSTOL1* homologs on chromosome 3 were associated with four RSA traits, bushiness, centroid, median number of roots, and specific root length (Fig. 7A; Supplemental Table S3). Three SNPs within *Sb03g031690* (4020, 4085, and 4121) were associated with variations in bushiness and centroid, which were negatively correlated. In agreement with that, alternative alleles at these SNPs increased centroid and bushiness (Supplemental Table S3). These same SNPs, in addition

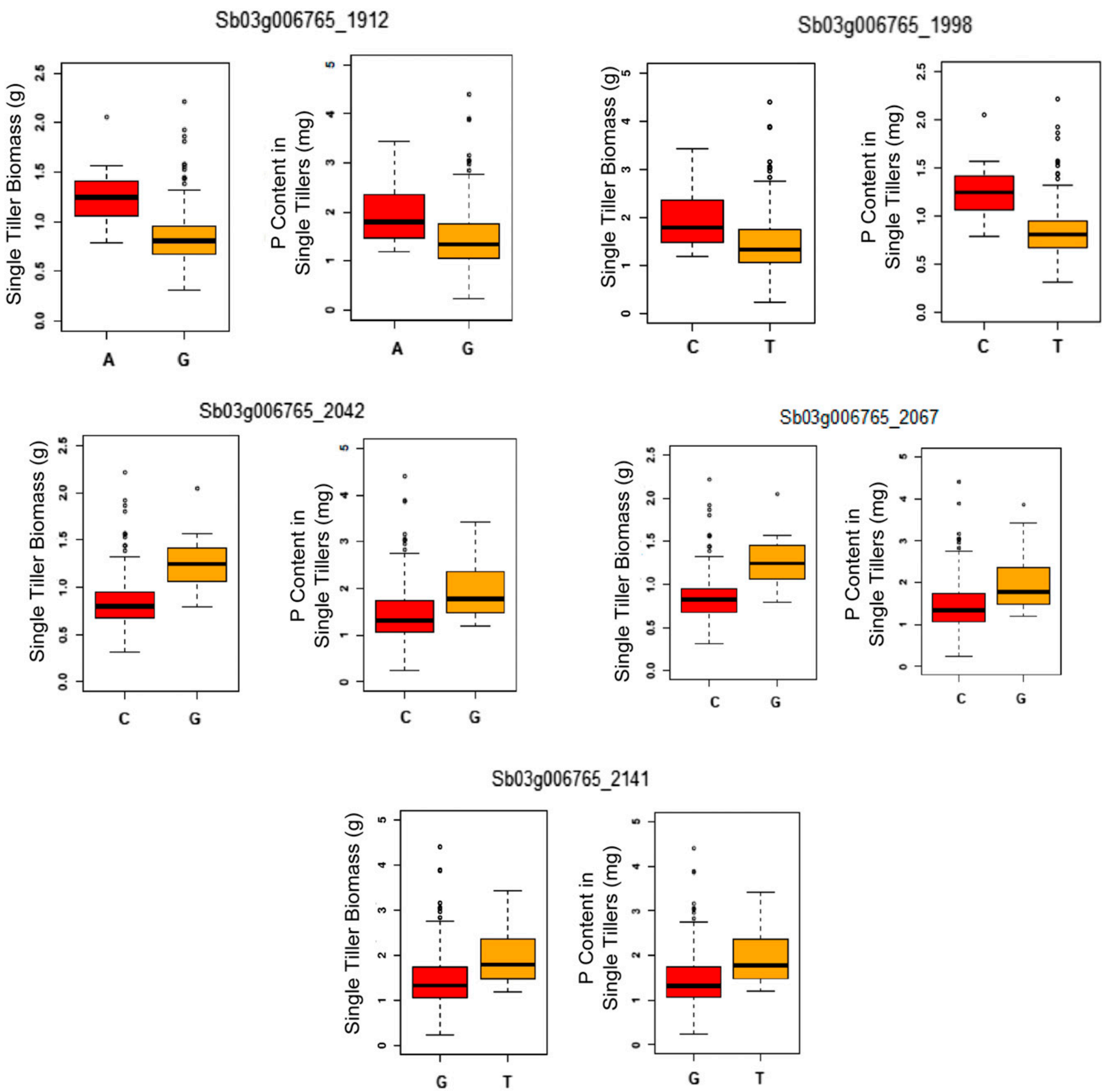

Figure 5. Effects for SNPs within *Sb03g006765* estimated in the WAP in Africa. Effects for single tiller biomass and P content in single tillers are shown. The *P* values from association analysis are shown in Table IV. [See online article for color version of this figure.]

to SNP 4170 in *Sb03g031690*, were also associated with the median number of roots, whereas the two SNPs within *Sb03g031680* were associated with specific root length. The same alleles at commonly associated SNP loci within *Sb03g031690* increased the value for centroid and the median number of roots, which again is consistent with the positive correlation between these two traits. Associations for *Sb03g006765* and *Sb03g031670* were found for centroid and median number of roots, respectively, with higher *P* values compared with those for *Sb03g031690* and *Sb03g031680*. Using a $P < 0.05$ threshold, *SbPSTOL* homologs on chromosome 7 were not associated with RSA traits.

DISCUSSION

Exploiting the conservation of gene content and order across genomes (Bonierbale et al., 1988) has held great promise in the identification of important genes in the grass family. Within the comparative genomics framework, information regarding functional variation in one

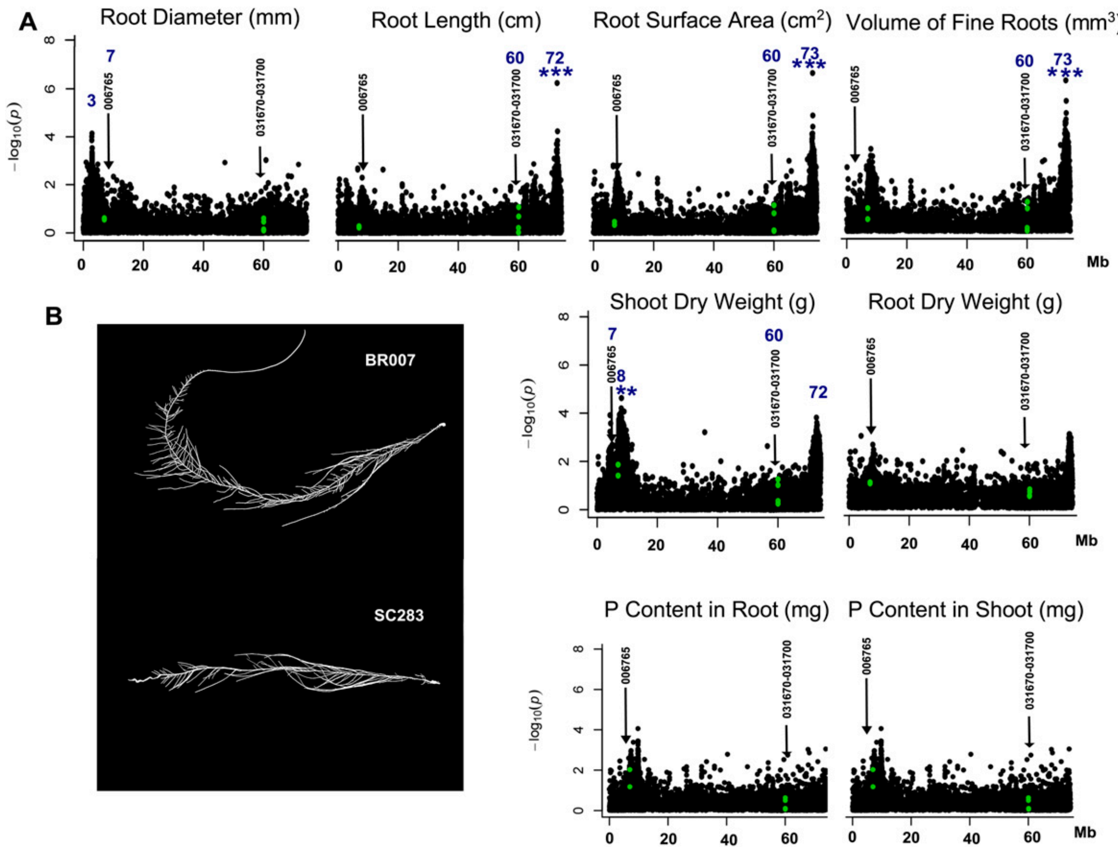

Figure 6. Validation of *PSTOL1* homologs on chromosome 3 by biparental mapping. A, QTL analysis of root morphology traits, shoot and root dry weight accumulation, and P content in the BR007 × SC283 RIL population. Physical positions on chromosome 3 are shown on the *x* axis (Mb). The positions for *Sb03g006765* (depicted as 006765 in the graphs) at the beginning of chromosome 3 and for the cluster encompassing *Sb03g031670*, *Sb03g031680*, *Sb03g031690*, and *Sb03g031700* (031670–031700) at the end are shown. Blue asterisks denote significant QTLs declared based on permutation tests at $0.01 \leq P < 0.05$ (**) and $0.001 \leq P < 0.01$ (***). Numbers in blue indicate the physical positions (Mb) of SNPs at the $-\log_{10}(P)$ peak within the different QTLs in addition to the closest *SbPSTOL1* genes. The SNPs that are closest to *SbPSTOL1* homologs are highlighted in green. B, Root morphology for the parents of the mapping population.

species can be transferred to related species, and in that scenario, molecular breeding applications may be greatly expedited. However, examples of the successful application of comparative genomics targeting agronomically relevant genes are still scarce, which can be attributed at least in part to the dynamic nature of the fine-scale structure among plant genomes (Bancroft, 2001). The availability of complete sequences for several plant genomes now provides for a much more complete framework where functional variants can be identified via comparative genomics, even in potentially complex situations arising from ancient gene duplication events. This is likely the case of RLKs in plants (Lehti-Shiu et al., 2009), a protein class to which PSTOL1 belongs. In view of that, based on the isolation of *OsPSTOL1* in rice, we undertook a cross-species effort to identify homologs in the sorghum genome that are functional with regard to enhancing agronomic performance under low-P conditions, a common constraint for sustainable crop yield worldwide (Ismail et al., 2007).

According to Parentoni and Souza Junior (2008), P acquisition efficiency in maize was more important than P internal utilization efficiency to explain the total variation for P use efficiency in a tropical soil. This was also the case for the SAPst phenotyped in a tropical soil in Brazil, where P availability to roots is reduced due to the high P fixation to clay minerals in the soil. Because *OsPSTOL1* in rice has been shown to enhance early root growth (Gamuyao et al., 2012), which is an important mechanism leading to enhanced P acquisition efficiency under low-P availability, we assessed root morphology traits in sorghum plants from the SAPst grown hydroponically using low-P nutrient solution in a paper pouch system. We observed that total root length and root surface area were positively correlated with grain yield under low-P conditions, which is in agreement with reports indicating that results of root morphology analysis in paper pouches correlate with those of plants grown in soil (Liao et al., 2001; Hund et al., 2009; Miguel et al., 2013). This makes

A

Gene	Position (bp)	Bushiness (n n^{-1})	Centroid (cm)	Median Number of Roots (n)	SRL (cm cm^{-3})
Sb03g006765	2002		1.75 **		
Sb03g031670	1312			2.12 **	
Sb03g031680	1541				3.99 ***
	1591				2.09 **
Sb03g031690	4020	2.62 **	2.50 **	4.89 ****	
	4085	2.98 ***	2.75 ***	5.09 ****	
	4121	2.56 **	2.75 **	3.75 ***	
	4170			1.69 **	

**** = p < 0.001 *** =0.001 ≤ p < 0.01 ** = 0.01 ≤ p < 0.05

B

Figure 7. Association analysis between *SbPSTOL1* homologs and RSA traits in 266 accessions of the SAPst. A, Association results. Numbers indicate the fraction of the phenotypic variation explained by the SNP loci. B, Lines contrasting for centroid. Line 071 has a large centroid (9.27 cm), whereas line 291 has a small centroid (4.86 cm).

root morphology assessed in younger plants a useful proxy to identify genes related to agronomic performance under low-P conditions. The correlation coefficients were nevertheless low, which is expected given the complex genetic architecture governing both grain yield and different P efficiency mechanisms that are likely to take place in such a diverse association panel.

SNP markers within sorghum homologs to *OsPSTOL1* were associated with both root morphology traits and P acquisition and biomass accumulation in hydroponically grown plants and also with grain yield assessed in a low-P soil in Brazil. The SNPs used for association mapping within *Sb03g006765* at position 7 Mb on chromosome 3 are in complete LD. Indeed, multiple SNPs within *Sb03g006765* were associated with grain yield under low P availability in the SAPst. Jointly considering the WAP and SAPst, the same alleles based on five SNPs within *Sb03g006765* enhanced P acquisition and biomass accumulation (assessed in the WAP) and grain yield (assessed in the SAPst), suggesting that this gene has a broad role in contributing to sorghum performance under low-P conditions across diverse genetic backgrounds and environmental conditions. This may be achieved by enhancing root surface area, since the C allele at SNP 1998 was associated with both increased root surface area and grain yield in a low-P soil in the SAPst.

The functionality for *Sb03g006765* in sorghum P efficiency is consistent with the reported role of *OsPSTOL1* in rice, which was shown to enhance P acquisition and performance under low P (Wissuwa et al., 1998; Gamuyao et al., 2012). One SNP within *Sb03g006765* was associated with root traits, namely root diameter. Using biparental mapping, we found indications of a QTL for root diameter approximately 4 Mb from *Sb03g006765*, although the significance of this QTL was above the $P < 0.05$ threshold. However, we detected a significant QTL for shoot dry weight with this *PSTOL1* homolog, supporting a role for *Sb03g006765* in enhancing P acquisition in sorghum, possibly by reducing root diameter, thus increasing the root surface area and reducing the costs to the plant of exploring the soil for P (Bates and Lynch, 2001).

Four additional *PSTOL1* homologs, *Sb03g031670*, *Sb03g031680*, *Sb03g031690*, and *Sb03g031700*, are close together at position 60 Mb on sorghum chromosome 3 and are possibly the result of tandem duplication events. SNP 2305 within *Sb03g031690* jointly increased both root P concentration in the SAPst and biomass accumulation in the WAP under low P. The same allele was also associated with increased root surface area in hydroponics, suggesting that altered root morphology is the physiological mechanism underlying improved P acquisition via *Sb03g031690*. A connection between enhanced root surface area and grain yield under low-P conditions in the soil was also found for *Sb03g031680*, where the A allele at SNP 1541 jointly increased both traits in the SAPst. Indeed, P deficiency has been shown to enhance root growth, leading to significant increases in root length and root surface area in *Cucumis melo* (Fita et al., 2011) and *Brassica napus* (Zhang et al., 2011). The hypothesis that *Sb03g031680* and *Sb03g031690* control root morphology, as was the case for *OsPSTOL1* in rice (Gamuyao et al., 2012), is reinforced by the presence of highly significant QTLs related to root morphology traits at approximately 12 Mb from the PSTOL1 homologs at position 60 Mb on chromosome 3. Because the *SbPSTOL1* genes at position 60 Mb are in close physical proximity on chromosome 3, multiple associations detected for these genes may be the result of LD in the region. However, SNPs within *Sb03g006765* at position 7 Mb are not in LD with those within the *SbPSTOL1* cluster at position 60 Mb, suggesting that multiple *SbPSTOL1* homologs enhance sorghum performance under low-P conditions, likely by the same overall physiological mechanism based on root morphology modulation. Lehti-Shiu et al. (2009) found that RLKs present in tandem repeats were more likely than nontandem RLKs to be up-regulated by biotic stress factors and UV-B light in Arabidopsis, suggesting that the *SbPSTOL1* homologs close together at position 60 Mb on chromosome 3 may constitute a functional module controlled by common regulatory elements. If so, similar to the case for maize aluminum tolerance (Maron et al., 2013), tandem duplication events may provide an additional level of regulation, leading to enhanced gene expression and increased P acquisition controlled by *PSTOL1* homologs in sorghum.

In addition to root morphology, changes in the spatial configuration of the root system, known as RSA, is another mechanism implicated in P acquisition efficiency. Because on acidic tropical soils phosphate ions tend to bind fairly tightly to aluminum and iron oxides on the surface of clay minerals, the soil phosphate tends to be fixed in the surface soil horizons, and it is logical that root systems that participate in topsoil P foraging would be an adaptive advantage on these soils. In general, compared with the associations detected with two-dimensional (2D) root morphology traits, we observed stronger associations with RSA traits, particularly for bushiness, centroid, and median number of roots for *Sb03g031690*. Centroid measures the tendency of a root to grow at different depths (i.e. a larger centroid value is indicative of a more deeply growing root system), whereas bushiness is a measure of the global branching of the root system (Clark et al., 2011). The associations detected for centroid and bushiness for *Sb03g031690* suggest that *SbPSTOL1* enhances the capacity of the sorghum root system to forage the P that is more available in superficial soil layers. This is consistent with studies in maize, bean (*Phaseolus vulgaris*), and soybean (*Glycine max*), indicating that the proliferation of shallow lateral roots increases topsoil foraging of P (Liao et al., 2001; Lynch, 2011).

Although the mechanism(s) by which *PSTOL1* increases root growth and presumably changes RSA has not yet been determined, Gamuyao et al. (2012) reported on a dirigent gene (*OsPupK20-2*) that is located close to *OsPSTOL1* in the rice genome, and its expression was induced both in 35S::*OsPSTOL1* plants and in a near-isogenic line harboring the *Pup1* QTL region including *OsPSTOL1*. A dirigent domain-containing protein, ENHANCED SUBERIN1, was recently shown to play a critical role in the formation of the Casparian strip in Arabidopsis, as it is required for the correct patterning of lignin depositions into this endodermis structure (Hosmani et al., 2013). In view of observations that environmental stresses such as salinity may change the fine structure of plant roots and the development of the Casparian strip (Degenhardt and Gimmler, 2000), it is tempting to speculate that dirigent genes link root system architectural changes to *PSTOL1*.

RLKs constitute the largest family of receptors in plants, and the diverse structures in the receptor domains suggest that there are likely to be several biological functions for these proteins (Lehti-Shiu et al., 2009; Tör et al., 2009). Although RLKs have been found to be important regulators of root hair growth in Arabidopsis under low-P conditions (Lan et al., 2013), their role seems to be much more general, as they also function in defense and stress responses, including drought, salinity, and cold (Lehti-Shiu et al., 2009; Ye et al., 2009; Das and Pandey, 2010; Marshall et al., 2012; Schulz et al., 2013; Vivek et al., 2013), possibly acting as hubs in stress signaling and development (Das and Pandey, 2010; Asano et al., 2012; Schulz et al., 2013). Several RLKs such as CRPK1 have been shown to play key roles in abiotic stress tolerance, positively or negatively regulating stress tolerance by modulating abscisic acid signaling and reducing the accumulation of reactive oxygen species (Lan et al., 2013).

In addition to a conserved Ser/Thr kinase domain, the selected SbPSTOL1 proteins are predicted to share distinctly different structural features from rice OsPSTOL1, including a transmembrane domain and extracellular domains that are likely in contact with the pectin fraction of the cell wall. These features are typical of cytoplasmic Ser/Thr wall-associated kinase (WAK) proteins, which span the plasma membrane and extend to the extracellular region to bind tightly to the cell wall (He et al., 1999). According to our sequence data, all selected sorghum *SbPSTOL1* genes possess three exons and two introns, which is the typical intron-exon organization of Arabidopsis (He et al., 1999; Kanneganti and Gupta, 2008), barley (*Hordeum vulgare*; Kaur et al., 2013), and rice (Zhang et al., 2005) WAKs. In addition, five highly similar WAK genes were found to lie in a 30-kb cluster in Arabidopsis (He et al., 1999), which resembles the genome organization of the *SbPSTOL1* genes at position 60 Mb on chromosome 3.

By physically linking the cell wall to the cytoplasm, these unique WAK proteins have been hypothesized to function as cell wall sensors (Steinwand and Kieber, 2010), possibly mediating cross talk whereby environmental cues may alter pectin content, structure, and biological activity (Kanneganti and Gupta, 2008). Indeed, expression of WAK1 in Arabidopsis was found to be induced by aluminum treatment (Sivaguru et al., 2003), and WAKL4 is related to mineral nutrition responses (Hou et al., 2005). Wall-associated kinases are necessary for normal cell enlargement (Cosgrove, 2001; Wagner and Kohorn, 2001), and in barley and rice, *WAK* genes were found to affect root growth (Kanneganti and Gupta, 2008, 2011; Kaur et al., 2013). Our results indicate that multiple SNPs within Sb03g006765 are consistently associated with sorghum performance under low P in the two association panels. The Sb03g006765 protein was predicted to possess a GUB-WAK-Bind domain in addition to a WAK-association domain. Thus, it is possible that a singular and perhaps tighter interaction between this protein and the cell wall results in increased root surface area, P uptake, and finally sorghum performance under low P in the soil via an as yet unidentified mechanism. As hypothesized by Gamuyao et al. (2012), it is conceivable that *PSTOL1* plays a broader role in abiotic stress tolerance than just P efficiency. As such, our results in sorghum reveal a possibly larger potential for *SbPSTOL1* genes to contribute to overall adaptation to a range of abiotic stresses via changes in root development, such as drought tolerance and the uptake of other mineral nutrients besides P.

In a genome-wide association study targeting adaptation to low P in sorghum (Leiser et al., 2014b), the most significant association with grain yield was found for an SNP at position 71.1 Mb on chromosome 3 (71178053), only 35.7 kb from the aluminum tolerance gene *Sorghum bicolor multidrug and toxic compound extrusion* (*SbMATE*), which underlies the aluminum tolerance locus *Aluminum tolerance Sorghum bicolor* (Alt_{SB}; Magalhaes et al., 2007). As SNPs associated with grain yield were also found within Alt_{SB} and *SbMATE*, this suggests a possible pleiotropic effect of *SbMATE* in providing tolerance to aluminum toxicity and to low-P conditions (Leiser et al., 2014b). The 71178053 SNP and one Alt_{SB}-specific SNP showed grain yield effects of 240 kg ha^{-1} (W. Leiser, personal communication) and 250 kg ha^{-1}, respectively, under low P (Leiser et al., 2014b). In turn, the grain yield effect for SNPs within Sb03g006765 and Sb03g031680 were 154 to 200 kg ha^{-1}, which was comparable to the effects reported by Leiser et al. (2014b). The grain yield effect for SNPs within *SbPSTOL1* genes described in this study are now being validated in different sorghum breeding contexts, including random mating populations, which should provide for a more realistic view of the breeding value of *SbPSTOL1* genes when targeting adaptation to acid soil conditions.

Together, aluminum toxicity and low P availability are two of the major abiotic stress factors limiting sorghum production on tropical soils. We recently reported on the development of tag markers for the major aluminum tolerance locus in sorghum, Alt_{SB}, which confers aluminum tolerance by a physiological mechanism based on aluminum-activated root citrate release to the rhizosphere (Caniato et al., 2014). Aluminum toxicity and low P availability in general coexist in tropical soils. As such, markers developed for sorghum *SbPSTOL1* genes as well as for Alt_{SB}, which are available through the Integrated Breeding Platform of the Generation Challenge Program (https://www.integratedbreeding.net/), now have the potential to benefit sorghum production and thus food security in the large areas of acid soils that exist worldwide.

MATERIALS AND METHODS

Plant Material

SAPst

A subset of the association panel described by Casa et al. (2008) with 287 accessions, comprising both tropical converted (185) and breeding (102) accessions, was used in this study. Tropical converted accessions were produced by introgressing genomic regions conferring early maturity and dwarfing genes from an elite donor into exotic accessions (Stephens et al., 1967) and were selected to broadly represent the genetic diversity of cultivated sorghum (*Sorghum bicolor*). The breeding accessions included photoperiod-insensitive elite inbreds, improved cultivars, and landraces that either have been or are still used extensively in the United States and Brazil for sorghum breeding (Casa et al., 2008).

WAP

The West African sorghum diversity panel consists of 187 sorghum genotypes from six West and Central African countries, consisting of seven racial groups (three pure races and four intermediate groups) including breeding accessions and landraces with various degrees of photoperiod sensitivity and internode lengths (Leiser et al., 2014b).

Root Morphology Analysis in the SAPst

Root morphology data were obtained in a paper pouch system supplied with nutrient solution as described by de Sousa et al. (2012). The experiments were organized in a completely randomized design with three replicates. Seeds were surface sterilized with sodium hypochlorite (0.5% [w/v] for 5 min) and germinated in moistened paper rolls. After 4 d, uniform seedlings were transferred to moistened blotting papers placed into paper pouches (24 × 33 × 0.02 cm) as described by Hund et al. (2009). Each experimental unit consisted of one pouch (three seedlings per paper pouch), whose bottom 3 cm was immersed in containers filled with 5 L of the nutrient solution described by Magnavaca et al. (1987) with a P concentration of 2.5 μM (as phosphate; 10 pouches per container). The nutrient solution was changed every 3 d, and the pH was maintained at 5.6. The containers were kept under continuous aeration in a growth chamber with 27°C/20°C day/night temperatures and a 12-h photoperiod.

After 13 d, root images were captured with a digital photography setup and analyzed using both the RootReader2D program (Clark et al., 2013; http://www.plantmineralnutrition.net/rr2d.php) as well as WinRHIZO (http://www.regent.qc.ca/assets/winrhizo_about.html) software. The following root traits were measured as described by de Sousa et al. (2012): total root length (designated root length; in cm), average root diameter (designated root diameter; in mm), volume of fine roots (diameter between 1 and 2 mm; in mm^3), and total root surface area (designated root surface area; in cm^2).

ANOVA was performed using GenStat version 16 software (Payne, 2009). Broad sense heritability was calculated as:

$$h^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_\varepsilon^2}{r}}$$

where $\sigma^2_g = (MS_g - MS_\varepsilon)/r$ and $\sigma^2_\varepsilon = MS_\varepsilon$, where σ^2_g is the genetic variance, σ^2_ε is the error variance, MS_g and MS_ε are the genetic and error mean squares, respectively, and r is the number of replications.

Phenotyping for Low-P Adaptation

SAPst

A set of 243 accessions of the SAPst consisting of three dwarf types were selected for the field trials. Soil P (Mehlich 1) was 5 ppm at 0- to 20-cm soil depth and varied between 0 and 5 ppm at the subsuperficial soil layer (20–40 cm). The field trial was located on a clayed and highly weathered tropical soil, representative of the types found in the agriculture areas of Central Brazil, dominated by kaolinite and iron and aluminum oxides in the clay fraction. The soil fertility in natural conditions is low, presenting low pH and high toxic aluminum levels and low P. One additional area, with approximately 10 ppm P, was used as the P sufficiency control. Experiments were set up as three α-lattice designs with nine incomplete blocks, nine accessions from the association panel and two checks per incomplete block, and three replications in 2 years of evaluation (2010 and 2011). Statistical analyses were performed jointly for both years based on mixed models using the GenStat version 16 software, considering genotypes and blocks within replications as random effects and years, experiments, replications, and checks as fixed effects. First, a series of model selection steps were performed, starting from a model including both spatial (rows and columns; taken as random) and phenology-related (i.e. plant height, flowering time, and number of harvested panicles; fixed effects) covariates, which were kept in the model based on the likelihood ratio test and the Wald test, respectively, considering a significance level of 5%. Then, different structures of the variance-covariance matrices for the genetic (G matrix) and residual (R matrix) effects across years were compared via the Akaike information criterion (Akaike, 1974). An unstructured and a diagonal variance-covariance model were selected for the G and R matrices, respectively. Best linear unbiased predictions for grain yield were finally obtained for genotypes, combining the information for both years of evaluation. Broad sense heritability was calculated based on variance components as:

$$h^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_{ge}^2}{e} + \frac{\sigma_\varepsilon^2}{er}}$$

where σ^2_g, σ^2_{ge}, and σ^2_ε are the variance components for the genetic, genotype-by-environment interaction, and error terms, respectively, in which environment corresponds to the different years of evaluation; and e and r are the number of years and replications, respectively.

P uptake was assessed in 13-d sorghum plants grown in hydroponics (i.e. shoot and root P) at the end of the root morphology experiments. Plant tissues were dried at 65°C until constant weight. For P analysis, plant tissues were subjected to a nitric/perchloric acid digestion (da Silva, 2009), and P quantification was done using inductively-coupled argon plasma emission spectrometry. Total P content, designated as P content, was calculated by multiplying plant total dry weight by tissue P concentration.

WAP

Experiments were conducted at the International Crops Research Institute for the Semi-Arid Tropics experimental station in Mali. A pot trial with four replications laid out in an α-design was carried out for 38 d. Ten-liter pots were filled with soil from a low-P field, having a Bray-1 P value of 5.2 ppm, a pH of 6.9, and an exchangeable acidity of 12%. In each pot, 10 seeds of each genotype were sown at 1.5-cm depth into five holes (two seeds per hole) and covered with sand. Fourteen days after sowing, plantlets were thinned to two plants per pot. Plants were treated three times during the experiment primarily against shoot fly (*Atherigona soccata*) and other insects using insecticides commonly used in the region. Some genotypes exhibited prolific tillering (more than four tillers per plant), observed in some materials grown in the off-season (night temperatures of less than 15°C and short days), while others did not tiller at all. Total shoot biomass was harvested 38 d after sowing, with plants having approximately six to eight leaves. Number of plants and number of tillers were recorded to calculate shoot biomass accumulated per plant and per tiller. Fresh plant material was stored in paper bags and dried at 45°C for 7 d. Total shoot dry matter was recorded, and P concentration was analyzed using inductively coupled plasma emission spectrometry as described by VDLUFA (2011).

Biomass, P concentration, and P content data were separately analyzed using mixed models, considering the effect of genotypes as fixed and replications and incomplete blocks nested within replications as random. Broad sense heritability was estimated, considering genotype as random:

$$h^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_\varepsilon^2}{r}}$$

where σ^2_g is the genotypic variance component, σ^2_ε is the error variance, and r is the number of replications.

In Silico Identification of Sorghum *PSTOL1* Homologs

A sequence similarity search based on the OsPSTOL1 amino acid and nucleotide sequences (GenBank accession no. BAK26566) was undertaken with the National Center for Biotechnology Information (NCBI; www.ncbi.nlm.nih.gov) and Phytozome (www.phytozome.org) databases. At the NCBI, the sorghum bicolor taxid ID 4558 database was used for BLASTP and BLASTN searches. At Phytozome, the sorghum genome sequence version 1.4 (Paterson et al., 2009) was used for BLASTP and BLASTN analyses. We selected the seven top hits that were coincident across procedures considering $E < 1e\text{-}100$ as the inclusion threshold. These amino acid sequences were aligned with those of proteins highly similar to *PSTOL1* in rice (*Oryza sativa*), Arabidopsis (*Arabidopsis thaliana*), and maize (*Zea mays*). For that, we selected proteins with $E < 1e\text{-}100$ when aligned with rice *PSTOL1*, and a total of 33 sequences were selected. The phylogenetic tree was inferred using the maximum likelihood method based on the JTT matrix-based model developed by D.T. Jones, W.R. Taylor, and J.M. Thorton (Jones et al., 1994; Rzhetsky and Nei, 1992). Initial trees from the heuristic search were obtained automatically by applying the neighbor-joining and BioNJ (Gascuel, 1997) algorithms to a matrix of pairwise distances estimated using the JTT model and then selecting the topology with the superior log-likelihood value. Five hundred bootstrap resampling steps were applied to the data. The tree was drawn to scale, with branch lengths measured as the number of substitutions per site. Evolutionary analyses were conducted in MEGA6 (Tamura et al., 2013).

Protein sequences were predicted by the Translate tool (http://web.expasy.org/translate/), and multiple sequence alignment analysis was performed using ClustalW (http://www.ebi.ac.uk/clustalw/index.html). Structural predictions were carried out using Pfam (http://pfam.xfam.org/), the TargetP 1.1 server (http://www.cbs.dtu.dk/services/TargetP/), Phobius (http://www.ebi.ac.uk/Tools/pfa/phobius/), and iTAK (Plant Transcription Factor & Protein Kinase Identifier and Classifier; http://bioinfo.bti.cornell.edu/cgi-bin/itak/index.cgi). MyDomains (http://prosite.expasy.org/mydomains/) was used as a tool for diagramming domain positions for predicted proteins. All tools were used with default parameters.

SNP Discovery

Primers amplifying amplicons specific to each of the six selected candidates were designed using Primer-Blast (ncbi.nlm.nih.gov/tools/primer-blast/index.cgi) using the sorghum genome as reference. This tool uses local and global alignment algorithms to identify specific primers based on a user-selected database (Ye et al., 2012).

A set of 22 diverse accessions was used for SNP discovery. Genomic DNA was isolated from approximately 500 mg of leaf tissue (five plants per accession) as described by Saghai-Maroof et al. (1984). PCR was done in a 20-μL reaction volume with 30 ng of genomic DNA, 1 unit of Taq DNA polymerase (Invitrogen), 10× Taq reaction buffer, 2 mM $MgCl_2$, 2 mM deoxyribonucleotide triphosphates, 5% (v/v) dimethyl sulfoxide, and 1 pmol of each primer. Thermocycling consisted of 4 min at 94°C followed by 35 cycles of 30 s at 94°C, 30 s at the specific annealing temperature for each primer (described in Supplemental Table S4), and 45 s at 72°C, followed by 7 min at 72°C. PCR products (5 μL) were treated with 2 μL of ExoSAP-IT reagent (Invitrogen) in a volume of 12 μL according to the manufacturer's specifications. A sample of 3 μL of treated PCR products was sequenced with the same forward and reverse primers used for PCR amplification with the BigDye Terminator version 3.1 cycle sequencing kit on an ABI PRISM 3100 sequencer (Applied Biosystems). The resulting sequences were subjected to quality-control procedures, trimmed with the Lasergene 7 software, and aligned by using the MegAlign program (DNAStar).

SNPs with a minor allele frequency higher than 0.1 were converted to the Kompetitive Allele Specific PCR genotyping system (KASP, LGC Genomics) following genotyping of the sorghum association panels. SNP loci were named based on the respective gene within which each SNP is located followed by the position relative to the beginning of the gene sequence (http://www.phytozome.net). For example, *Sb03g031690*_2305 refers to an SNP locus at position 2,305 bp within gene *Sb03g031690*.

Association Analysis

SAPst

The Bayesian cluster analysis implemented in the STRUCTURE software (Pritchard et al., 2000) was used to estimate population structure in the association panel based on 310 SNPs distributed across the sorghum genome (Murray et al., 2009). The number of subpopulations, k, was allowed to vary from 1 to 12, with three independent runs for each value of k. The admixture model with correlated allele frequencies was adopted, with a burn-in length of 50,000 and a 50,000 run length, and a correction factor for Hardy-Weinberg equilibrium departure, λ, was estimated with an independent run with $k = 1$. The optimal k value was determined based on the estimated log-likelihood of the data (Yu et al., 2006) and on the second-order rate of change in the log-likelihood of the data (ΔK; Evanno et al., 2005). Based on this procedures, we chose $k = 4$ as the optimal number of subpopulations.

The familial relatedness or kinship (K) matrix depicting similarity between individuals (Yu et al., 2006) was estimated in TASSEL version 4.0 based on the same 310 SNPs used to assess population structure. In this matrix, each element d_{ij} was equal to the proportion of the SNPs that are different between taxon i and taxon j. The distance matrix was converted into a similarity matrix by subtracting all values from 2 and then scaling so that the minimum value in the matrix is 0 and the maximum value is 2.

Model selection for association analysis was based on type I error simulations to select the model yielding the fewest false-positive associations, based on root traits and grain yield. The 310 randomly distributed SNP markers used here do not provide genome saturation within the low LD context in the association panel (Bouchet et al., 2012); consequently, the chances of association with the phenotypic traits can be considered negligible. Thus, it provides a null distribution to test the efficiency of different models to control for false-positive associations. The different models were (1) $y = A\alpha + e$, a naïve model that does not control for population structure or relatedness; (2) an association model incorporating population structure (Q), $y = A\alpha + Qv + e$; (3) a model controlling for familial relatedness or kinship (K), $y = A\alpha + Zu + e$; and (4) a mixed model that jointly accounts for both population structure and kinship (Q+K; Yu et al., 2006), $y = A\alpha + Qv + Zu + e$. In these models, y is a vector of phenotypic observations, α and v are vectors of fixed effects related to SNPs and population structure, respectively, whereas u and e are vectors of polygenic background effects related to familial relatedness and residuals, respectively. Q is a population membership assignment matrix relating y to v, and A and Z are incidence matrices of 1 s and 0 s relating y to α and u, respectively. The naïve and Q models were fitted using the General Linear Model procedure, while Mixed Linear Model was used for the K and Q+K models in TASSEL version 4.0 (Bradbury et al., 2007). SNPs with minor allele frequency > 0.05 were considered for simulating type I error. The quantile-quantile plots of estimated $-\log_{10}(P)$ were constructed using the observed *P* values from marker-trait associations and the expected *P* values based on the assumption that no association exists between markers and all traits. Finally, association analysis was undertaken with SNPs within *SbPSTOL1* genes and the different phenotypic traits using the kinship model, which was found to provide appropriate control of type I error, following the same model definition and procedures used for type I error simulations as described above.

WAP

Association analysis was carried out with 31 SNPs within sorghum *PSTOL1* homologs using TASSEL version 4.0. A compressed mixed linear model correcting for population structure and kinship was fitted, using two principal components estimated with SNP relate (Zheng et al., 2012) and a kinship matrix calculated using the EMMA algorithm within GAPIT (Lipka et al., 2012). The principal components and the kinship matrix were calculated using 220,934 SNPs genotyped by sequence (Elshire et al., 2011) for this population, excluding the SNPs within *SbPSTOL1* genes. Adjusted means of single tiller biomass, biomass per plant, shoot P, and P content in single tillers were used as response variables.

Biparental Mapping in Sorghum RILs

A total of 400 RILs derived from the cross between inbred lines SC283 and BR007, which show contrasting root morphology phenotypes, was used for QTL detection. This RIL population was phenotyped for root traits according to the same conditions described for the converted association panel. Genotyping was done by GBS (Elshire et al., 2011) using GBS pipeline 3.0 in the TASSEL software package (Bradbury et al., 2007; Glaubitz et al., 2014). Missing genotype calls were imputed using the NPUTE version 4.0 software (Roberts et al., 2007). SNPs with minor allele frequency > 0.4 were considered for the analysis. Genome scans were undertaken using a general linear model with TASSEL version 3.0, with a total of 10,810 SNPs on chromosome 3 and 4,218 SNPs on chromosome 7, looking for QTLs related to root morphology, dry matter accumulation, and P content. The high marker density provided by GBS precluded the need for estimating conditional probabilities of marker genotypes in the high-LD context of the RIL population. Marker significance was determined based on a significance level of 5% and 1,000 permutations (Anderson and Braak, 2003). Accordingly, this method calculates the predicted and residual values of the reduced model (excluding markers) and then permutes the residuals, which showed greater power compared with permutation of the raw data.

RSA Association on the SAPst

RSA data were assessed in the 266 accessions of the SAPst in a hydroponics-based 3D RSA system. The experiments consisted of a randomized block designs with two replicates. Seeds were surface sterilized with sodium hypochlorite (0.5% [w/v] for 5 min) and germinated in moistened paper rolls. After 4 d, seedlings were planted between the two top mesh layers using polyethylene foam in a mesh system, created from Acrylonitrile Butadiene Styrene plastic circles of 20-cm diameter made with a 3D printer. Each mesh system consists of eight layers of the plastic circles spaced 10 cm from each other by four anodized aluminum rods. This mesh system serves to constrain the roots but not impede their growth. The mesh systems were placed into clear glass cylinders or in large polyethylene tank containers filled with nutrient solution as described by Magnavaca et al. (1987) with 2.5 μM P and maintained at pH 5.6. The containers were kept under continuous aeration in a growth chamber with 27°C/20°C day/night temperatures and a 12-h photoperiod.

Root images for 3D reconstruction of the RSA and computation of RSA traits were taken after 8 and 10 d with a digital camera. In general, the methods described in detail by Clark et al. (2011) were used except that the plants were grown as described above hydroponically instead of in gel cylinders. For each plant's root system, 100 2D digital images as the plant was rotated through 360° with 2D images were taken every 3.6°. Our imaging and analysis software first digitally eliminates the image of the colored mesh and any other extraneous elements from the photographs of the root systems. Then, the RootReader3D software reconstructs the 3D image of the specific root system from the 100 2D digital images and automatically calculates 19 different RSA traits as described by Clark et al. (2011). The 19 RSA traits calculated from each sorghum root system are listed here: total root length (cm), maximum width of the root system (cm), minimum width of the root system (cm), maximum depth of the root system (cm), centroid (cm), exploitation volume (cm^3), exploitation index (cm^3 cm^{-1}), median number of roots, maximum number of roots, bushiness (i.e. maximum number of roots/median number of roots), surface area (cm^2), surface area-volume ratio (cm^3 cm^{-2}), surface area-length ratio (cm^3 cm^{-1}), one-third/two-third volume distribution (cm^3 cm^{-3}), convex hull volume (cm^3), solidity (i.e. volume/convex hull volume in cm^3 cm^{-3}), narrowness index (cm cm^{-1}), volume (cm^3), and specific root length (i.e. root length/root volume in cm cm^{-3}). Broad sense heritability was estimated via mixed models, considering genotype as random:

$$h^2 = \frac{\sigma_g^2}{\sigma^2 + \frac{\sigma_\varepsilon^2}{r}}$$

where σ^2_g is the genotypic variance component, σ^2_ε is the error variance, and *r* is the number of replications. Replications and blocks within replications were considered as fixed and random effects, respectively.

Sequence data from this article can be found in the GenBank/EMBL data libraries under accession numbers EES02527.1, EES01294.1, EES01295.1, EES01297.1, EES01298.1, EES13320.1, and BAK26566.

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure S1. Multiple sequence alignment of OsPSTOL1 and SbPSTOL1 proteins depicting structural predictions.

Supplemental Figure S2. Association model selection.

Supplemental Figure S3. Linkage disequilibrium across the *SbPSTOL1* gene on chromosome 7.

Supplemental Figure S4. Effects of SNPs *Sb03g031690*_2305 and *Sb03g031670*_1312.

Supplemental Figure S5. Validation of *PSTOL1* homologs on chromosome 7 by biparental mapping.

Supplemental Table S1. Correlation matrix for traits assessed in hydroponically grown plants.

Supplemental Table S2. Effects for associated SNPs within *SbPSTOL1* genes.

Supplemental Table S3. SNP effects on RSA traits for *SbPSTOL1* homologs.

Supplemental Table S4. Sorghum *PSTOL1* primer sequences.

ACKNOWLEDGMENTS

We thank Drs. Matthias Wissuwa (Japan International Research Center for Agricultural Sciences) and Sigrid Heuer (Australian Centre for Plant Functional Genomics) for useful discussions about *PSTOL1* genes in rice and sorghum.

Received May 28, 2014; accepted September 1, 2014; published September 2, 2014.

LITERATURE CITED

Akaike HAI (1974) A new look at the statistical model identification. IEEE Trans Automat Contr **19:** 716–723

Anderson M, Braak C (2003) Permutation tests for multi-factorial analysis of variance. J Stat Comput Simul **73:** 85–113

Asano T, Hayashi N, Kikuchi S, Ohsugi R (2012) CDPK-mediated abiotic stress signaling. Plant Signal Behav **7:** 817–821

Bancroft I (2001) Duplicate and diverge: the evolution of plant genome microstructure. Trends Genet **17:** 89–93

Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics **21:** 263–265

Bates TR, Lynch JP (2001) Root hairs confer a competitive advantage under low phosphorus availability. Plant Soil **236:** 243–250
Bonierbale MW, Plaisted RL, Tanksley SD (1988) RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. Genetics **120:** 1095–1103
Bouchet S, Pot D, Deu M, Rami JF, Billot C, Perrier X, Rivallan R, Gardes L, Xia L, Wenzl P, et al (2012) Genetic structure, linkage disequilibrium and signature of selection in Sorghum: lessons from physically anchored DArT markers. PLoS ONE **7:** e33470
Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics **23:** 2633–2635
Caniato FF, Guimarães CT, Hamblin M, Billot C, Rami JF, Hufnagel B, Kochian LV, Liu J, Garcia AAF, Hash CT, et al (2011) The relationship between population structure and aluminum tolerance in cultivated sorghum. PLoS ONE **6:** e20830
Caniato FF, Hamblin MT, Guimaraes CT, Zhang Z, Schaffert RE, Kochian LV, Magalhaes JV (2014) Association mapping provides insights into the origin and the fine structure of the sorghum aluminum tolerance locus, Alt_{SB}. PLoS ONE **9:** e87438
Casa M, 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 Sci **48:** 30–40
Clark RT, Famoso AN, Zhao K, Shaff JE, Craft EJ, Bustamante CD, McCouch SR, Aneshansley DJ, Kochian LV (2013) High-throughput two-dimensional root system phenotyping platform facilitates genetic analysis of root growth and development. Plant Cell Environ **36:** 454–466
Clark RT, MacCurdy RB, Jung JK, Shaff JE, McCouch SR, Aneshansley DJ, Kochian LV (2011) Three-dimensional root phenotyping with a novel imaging and software platform. Plant Physiol **156:** 455–465
Cordell D, Drangert JO, White S (2009) The story of phosphorus: global food security and food for thought. Glob Environ Change **19:** 292–305
Cosgrove DJ (2001) Plant cell walls: wall-associated kinases and cell expansion. Curr Biol **11:** R558–R559
Dakora F, Phillips D (2002) Root exudates as mediators of mineral acquisition in low-nutrient environments. Plant Soil **245:** 35–47
Das R, Pandey GK (2010) Expressional analysis and role of calcium regulated kinases in abiotic stress signaling. Curr Genomics **11:** 2–13
da Silva FC (2009) Manual de Análises Químicas de Solos, Plantas e Fertilizantes, Ed 2. Embrapa Informação Tecnológica, Brasilia, Brazil
Degenhardt B, Gimmler H (2000) Cell wall adaptations to multiple environmental stresses in maize roots. J Exp Bot **51:** 595–603
de Sousa S, Clark R, Mendes F, de Oliveira AC, de Vasconcelos MJV, Parentoni SN, Kochian LV, Guimarães CT, Magalhaes JV (2012) A role for root morphology and related candidate genes in P acquisition efficiency in maize. Funct Plant Biol **39:** 925–935
Doumbia MD, Hossner LR, Onken AB (1998) Sorghum growth in acid soils of West Africa: variations in soil chemical properties. Arid Soil Res Rehabil **12:** 179–190
Doumbia MD, Hossner LR, Onken B (1993) Variable sorghum growth in acid soils of subhumid West Africa. Arid Soil Res Rehabil **7:** 335–346
Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS ONE **6:** e19379
Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol **14:** 2611–2620
Fita A, Nuez F, Picó B (2011) Diversity in root architecture and response to P deficiency in seedlings of Cucumis melo L. Euphytica **181:** 323–339
Fitter A, Williamson L, Linkohr B, Leyser O (2002) Root system architecture determines fitness in an Arabidopsis mutant in competition for immobile phosphate ions but not for nitrate ions. Proc Biol Sci **269:** 2017–2022
Gamuyao R, Chin JH, Pariasca-Tanaka J, Pesaresi P, Catausan S, Dalid C, Slamet-Loedin I, Tecson-Mendoza EM, Wissuwa M, Heuer S (2012) The protein kinase Pstol1 from traditional rice confers tolerance of phosphorus deficiency. Nature **488:** 535–539
Gascuel O (1997) BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. Mol Biol Evol **14:** 685–695
Glaubitz JC, Casstevens TM, Lu F, Harriman J, Elshire RJ, Sun Q, Buckler ES (2014) TASSEL-GBS: a high capacity genotyping by sequencing analysis pipeline. PLoS ONE **9:** e90346
Godfray HCJ, Beddington JR, Crute IR, Haddad L, Lawrence D, Muir JF, Pretty J, Robinson S, Thomas SM, Toulmin C (2010) Food security: the challenge of feeding 9 billion people. Science **327:** 812–818
Haling RE, Brown LK, Bengough AG, Young IM, Hallett PD, White PJ, George TS (2013) Root hairs improve root penetration, root-soil contact, and phosphorus acquisition in soils of different strength. J Exp Bot **64:** 3711–3721
Hammond JP, White PJ (2008) Sucrose transport in the phloem: integrating root responses to phosphorus starvation. J Exp Bot **59:** 93–109
He ZH, Cheeseman I, He D, Kohorn BD (1999) A cluster of five cell wall-associated receptor kinase genes, Wak1-5, are expressed in specific organs of Arabidopsis. Plant Mol Biol **39:** 1189–1196
Heuer S, Lu X, Chin JH, Tanaka JP, Kanamori H, Matsumoto T, De Leon T, Ulat VJ, Ismail AM, Yano M, et al (2009) Comparative sequence analyses of the major quantitative trait locus *phosphorus uptake 1* (*Pup1*) reveal a complex genetic structure. Plant Biotechnol J **7:** 456–457
Hinsinger P (2001) Bioavailability of soil inorganic P in the rhizosphere as affected by root-induced chemical changes: a review. Plant Soil **273:** 173–195
Ho MD, Rosas JC, Brown KM, Lynch JP (2005) Root architectural tradeoffs for water and phosphorus acquisition. Funct Plant Biol **32:** 737–748
Hosmani PS, Kamiya T, Danku J, Naseer S, Geldner N, Guerinot ML, Salt DE (2013) Dirigent domain-containing protein is part of the machinery required for formation of the lignin-based Casparian strip in the root. Proc Natl Acad Sci USA **110:** 14498–14503
Hou X, Tong H, Selby J, Dewitt J, Peng X, He Z (2005) Involvement of a cell wall-associated kinase, WAKL4, in Arabidopsis mineral responses. Plant Physiol **139:** 1704–1716
Huang CY, Shirley N, Genc Y, Shi B, Langridge P (2011) Phosphate utilization efficiency correlates with expression of low-affinity phosphate transporters and noncoding RNA, *IPS1*, in barley. Plant Physiol **156:** 1217–1229
Hund A, Trachsel S, Stamp P (2009) Growth of axile and lateral roots of maize. I. Development of a phenotying platform. Plant Soil **325:** 335–349
Ingram PA, Zhu J, Shariff A, Davis IW, Benfey PN, Elich T (2012) High-throughput imaging and analysis of root system architecture in Brachypodium distachyon under differential nutrient availability. Philos Trans R Soc Lond B Biol Sci **367:** 1559–1569
Ismail AM, Heuer S, Thomson MJ, Wissuwa M (2007) Genetic and genomic approaches to develop rice germplasm for problem soils. Plant Mol Biol **65:** 547–570
Iyer-Pascuzzi AS, Symonova O, Mileyko Y, Hao Y, Belcher H, Harer J, Weitz JS, Benfey PN (2010) Imaging and analysis platform for automatic phenotyping and trait ranking of plant root systems. Plant Physiol **152:** 1148–1157
Jones DT, Taylor WR, Thornton JM (1994) A mutation data matrix for transmembrane proteins. FEBS Lett **339:** 269–275
Kanneganti V, Gupta AK (2008) Wall associated kinases from plants: an overview. Physiol Mol Biol Plants **14:** 109–118
Kanneganti V, Gupta AK (2011) RNAi mediated silencing of a wall associated kinase, OsiWAK1 in Oryza sativa results in impaired root development and sterility due to anther indehiscence: wall associated kinases from Oryza sativa. Physiol Mol Biol Plants **17:** 65–77
Kaur R, Singh K, Singh J (2013) A root-specific wall-associated kinase gene, HvWAK1, regulates root growth and is highly divergent in barley and other cereals. Funct Integr Genomics **13:** 167–177
Kellogg EALA (1998) Relationships of cereal crops and other grasses. Proc Natl Acad Sci USA **95:** 2005–2010
Lan P, Li W, Schmidt W (2013) Genome-wide co-expression analysis predicts protein kinases as important regulators of phosphate deficiency-induced root hair remodeling in Arabidopsis. BMC Genomics **14:** 210
Lehti-Shiu MD, Zou C, Hanada K, Shiu SH (2009) Evolutionary history and stress regulation of plant receptor-like kinase/pelle genes. Plant Physiol **150:** 12–26
Leiser WL, Rattunde H, Weltzien E, Cisse N, Abdou M, Diallo A, Tourè AO, Magalhaes JV, Haussmann B (2014b) Two in one sweep: aluminum tolerance and grain yield in P-limited soils are associated to the same genomic region in West African sorghum. BMC Plant Biol **14:** 206
Leiser WL, Rattunde HF, Piepho HP, Parzies HK (2012a) Getting the most out of sorghum low-input field trials in West Africa using spatial adjustment. J Agron Crop Sci **198:** 349–359
Leiser WL, Rattunde HFW, Piepho HP, Weltzien E, Diallo A, Melchinger AE, Parzies HK, Haussmann BIG (2012b) Selection strategy for sorghum

targeting phosphorus-limited environments in West Africa: analysis of multi-environment experiments. Crop Sci **52:** 2517

Leiser WL, Rattunde HFW, Weltzien E, Haussmann BIG (2014a) Phosphorus uptake and use efficiency of diverse West and Central African sorghum genotypes under field conditions in Mali. Plant Soil **377:** 383–394

Lewontin RC (1964) The interaction of selection and linkage. I. General considerations; heterotic models. Genetics **49:** 49–67

Li C, Gui S, Yang T, Walk T, Wang X, Liao H (2012) Identification of soybean purple acid phosphatase genes and their expression responses to phosphorus availability and symbiosis. Ann Bot (Lond) **109:** 275–285

Liao H, Rubio G, Yan X, Cao A, Brown KM, Lynch JP (2001) Effect of phosphorus availability on basal root shallowness in common bean. Plant Soil **232:** 69–79

Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, Gore MA, Buckler ES, Zhang Z (2012) GAPIT: genome association and prediction integrated tool. Bioinformatics **28:** 2397–2399

Lynch J, Brown K (2001) Topsoil foraging: an architectural adaptation of plants to low phosphorus availability. Plant Soil **237:** 225–237

Lynch JP (2011) Root phenes for enhanced soil exploration and phosphorus acquisition: tools for future crops. Plant Physiol **156:** 1041–1049

Lynch JP, Brown KM (2012) New roots for agriculture: exploiting the root phenome. Philos Trans R Soc Lond B Biol Sci **367:** 1598–1604

Ma Y, Szostkiewicz I, Korte A, Moes D, Yang Y, Christmann A, Grill E (2009) Regulators of PP2C phosphatase activity function as abscisic acid sensors. Science **324:** 1064–1068

Magalhaes JV, Liu J, Guimarães CT, Lana UGP, Alves VMC, Wang Y-H, Schaffert RE, Hoekenga OA, Piñeros MA, Shaff JE, et al (2007) A gene in the multidrug and toxic compound extrusion (MATE) family confers aluminum tolerance in sorghum. Nat Genet **39:** 1156–1161

Magnavaca R, Gardner CO, Clark RB (1987) Inheritance of aluminum tolerance in maize. *In* WH Gabelman, BC Loughman, eds, Genetic Aspects of Plant Mineral Nutrition. Martinus Nijhoff Publishers, Dordrecht, The Netherlands, pp 201–212

Maron LG, Guimarães CT, Kirst M, Albert PS, Birchler JA, Bradbury PJ, Buckler ES, Coluccio AE, Danilova TV, Kudrna D, et al (2013) Aluminum tolerance in maize is associated with higher MATE1 gene copy number. Proc Natl Acad Sci USA **110:** 5241–5246

Marschner H (1995) Mineral Nutrition of Higher Plants, Ed 2. Academic Press, London

Marshall A, Aalen RB, Audenaert D, Beeckman T, Broadley MR, Butenko MA, Caño-Delgado AI, de Vries S, Dresselhaus T, Felix G, et al (2012) Tackling drought stress: receptor-like kinases present new approaches. Plant Cell **24:** 2262–2278

Miguel MA, Widrig A, Vieira RF, Brown KM, Lynch JP (2013) Basal root whorl number: a modulator of phosphorus acquisition in common bean (*Phaseolus vulgaris*). Ann Bot (Lond) **112:** 973–982

Morris GP, Ramu P, Deshpande SP, Hash CT, Shah T, Upadhyaya HD, Riera-Lizarazu O, Brown PJ, Acharya CB, Mitchell SE, et al (2013) Population genomic and genome-wide association studies of agroclimatic traits in sorghum. Proc Natl Acad Sci USA **110:** 453–458

Mudge SR, Rae AL, Diatloff E, Smith FW (2002) Expression analysis suggests novel roles for members of the Pht1 family of phosphate transporters in Arabidopsis. Plant J **31:** 341–353

Murray SC, Rooney WL, Hamblin MT, Mitchell SE, Kresovich S (2009) Sweet sorghum genetic diversity and association mapping for brix and height. Plant Genome J **2:** 48–62

Ni JJ, Wu P, Senadhira D, Huang N (1998) Mapping QTLs for phosphorus deficiency tolerance in rice (*Oryza sativa* L.). Theor Appl Genet **97:** 1361–1369

Niu YF, Chai RS, Jin GL, Wang H, Tang CX, Zhang YS (2013) Responses of root architecture development to low phosphorus availability: a review. Ann Bot (Lond) **112:** 391–408

Pang J, Tibbett M, Denton MD, Lambers H, Siddique KHM, Bolland MDA, Revell CK, Ryan MH (2009) Variation in seedling growth of 11 perennial legumes in response to phosphorus supply. Plant Soil **328:** 133–143

Parentoni S, Souza Junior CL (2008) Phosphorus acquisition and internal utilization efficiency in tropical maize genotypes. Pesquisa Agropec Bras **43:** 893–901

Paszkowski U, Kroken S, Roux C, Briggs SP (2002) Rice phosphate transporters include an evolutionarily divergent gene specifically activated in arbuscular mycorrhizal symbiosis. Proc Natl Acad Sci USA **99:** 13324–13329

Paterson AH (2008) Genomics of sorghum. Int J Plant Genomics **2008:** 362451

Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros T, Poliakov A, et al (2009) The Sorghum bicolor genome and the diversification of grasses. Nature **457:** 551–556

Paterson AH, Bowers JE, Burow MD, Draye X, Elsik CG, Jiang CX, Katsar CS, Lan TH, Lin YR, Ming R, et al (2000) Comparative genomics of plant chromosomes. Plant Cell **12:** 1523–1540

Paterson AH, Bowers JE, Chapman BA (2004) Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. Proc Natl Acad Sci USA **101:** 9903–9908

Payne RW (2009) GenStat. Wiley Interdiscip Rev Comput Stat **1:** 255–258

Pritchard JK, Stephens M, Rosenberg NA, Donnelly P (2000) Association mapping in structured populations. Am J Hum Genet **67:** 170–181

Rai A, Rai S, Rakshit A (2013) Mycorrhiza-mediated phosphorus use efficiency in plants. Environ Exp Biol **11:** 107–117

Rausch C, Bucher M (2002) Molecular mechanisms of phosphate transport in plants. Planta **216:** 23–37

Roberts A, McMillan L, Wang W, Parker J, Rusyn I, Threadgill D (2007) Inferring missing genotypes in large SNP panels using fast nearest-neighbor searches over sliding windows. Bioinformatics **23:** i401–i407

Ryan P, Delhaize E, Jones D (2001) Function and mechanism of organic anion exudation from plant roots. Annu Rev Plant Physiol Plant Mol Biol **52:** 527–560

Rzhetsky A, Nei M (1992) A simple method for estimating and testing minimum-evolution trees. Mol Biol Evol **9:** 945–967

Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. Proc Natl Acad Sci USA **81:** 8014–8018

Sattari SZ, Bouwman AF, Giller KE, van Ittersum MK (2012) Residual soil phosphorus as the missing piece in the global phosphorus crisis puzzle. Proc Natl Acad Sci USA **109:** 6348–6353

Schachtman DP, Reid RJ, Ayling SM (1998) Phosphorus uptake by plants: from soil to cell. Plant Physiol **116:** 447–453

Schulz P, Herde M, Romeis T (2013) Calcium-dependent protein kinases: hubs in plant stress signaling and development. Plant Physiol **163:** 523–530

Sivaguru M, Ezaki B, He ZH, Tong H, Osawa H, Baluska F, Volkmann D, Matsumoto H (2003) Aluminum-induced gene expression and protein localization of a cell wall-associated receptor kinase in Arabidopsis. Plant Physiol **132:** 2256–2266

Smith SE, Smith FA (2012) Fresh perspectives on the roles of arbuscular mycorrhizal fungi in plant nutrition and growth. Mycologia **104:** 1–13

Steinwand BJ, Kieber JJ (2010) The role of receptor-like kinases in regulating cell wall function. Plant Physiol **153:** 479–484

Stephens J, Miller F, Rosenow D (1967) Conversion of alien sorghums to early combine genotypes. Crop Sci **7:** 1967

Svistoonoff S, Creff A, Reymond M, Sigoillot-Claude C, Ricaud L, Blanchet A, Nussaume L, Desnos T (2007) Root tip contact with low-phosphate media reprograms plant root architecture. Nat Genet **39:** 792–796

Takezawa D, Patil S, Bhatia A, Poovaiah BW (1996) Calcium-dependent protein kinase genes in corn roots. J Plant Physiol **149:** 329–335

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol **30:** 2725–2729

Tör M, Lotze MT, Holton N (2009) Receptor-mediated signalling in plants: molecular patterns and programmes. J Exp Bot **60:** 3645–3654

VDLUFA (2011) Umweltanalytik. VDLUFA-Verl, Darmstadt, Germany

Vivek PJ, Tuteja N, Soniya EV (2013) CDPK1 from ginger promotes salinity and drought stress tolerance without yield penalty by improving growth and photosynthesis in Nicotiana tabacum. PLoS ONE **8:** e76392

Wagner TA, Kohorn BD (2001) Wall-associated kinases are expressed throughout plant development and are required for cell expansion. Plant Cell **13:** 303–318

Walk TC, Jaramillo R, Lynch JP (2006) Architectural tradeoffs between adventitious and basal roots for phosphorus acquisition. Plant Soil **279:** 347–366

Wang J, Sun J, Miao J, Guo J, Shi Z, He M, Chen Y, Zhao X, Li B, Han F, et al (2013) A phosphate starvation response regulator Ta-PHR1 is involved in phosphate signalling and increases grain yield in wheat. Ann Bot (Lond) **111:** 1139–1153

Williamson LC, Ribrioux SPCP, Fitter AH, Leyser HMO (2001) Phosphate availability regulates root system architecture in Arabidopsis. Plant Physiol **126:** 875–882

Wissuwa M, Ae N (2001) Further characterization of two QTLs that increase phosphorus uptake of rice (*Oryza sativa* L.) under phosphorus deficiency. Plant Soil **237:** 275–286

Wissuwa M, Wegner J, Ae N, Yano M (2002) Substitution mapping of Pup1: a major QTL increasing phosphorus uptake of rice from a phosphorus-deficient soil. Theor Appl Genet **105:** 890–897

Wissuwa M, Yano M, Ae N (1998) Mapping of QTLs for phosphorus-deficiency tolerance in rice (*Oryza sativa* L.). Theor Appl Genet **97:** 777–783

Yan X, Liao H, Beebe SE, Blair MW, Lynch JP (2004) QTL mapping of root hair and acid exudation traits and their relationship to phosphorus uptake in common bean. Plant Soil **265:** 17–29

Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL (2012) Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics **13:** 134

Ye S, Wang L, Xie W, Wan B, Li X, Lin Y (2009) Expression profile of calcium-dependent protein kinase (CDPKs) genes during the whole lifespan and under phytohormone treatment conditions in rice (*Oryza sativa* L. ssp. *indica*). Plant Mol Biol **70:** 311–325

Yu J, Pressoir G, Briggs WH, Vroh Bi I, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, et al (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet **38:** 203–208

Zhang H, Huang Y, Ye X, Xu F (2011) Genotypic variation in phosphorus acquisition from sparingly soluble P sources is related to root morphology and root exudates in Brassica napus. Sci China Life Sci **54:** 1134–1142

Zhang S, Chen C, Li L, Meng L, Singh J, Jiang N, Deng XW, He ZH, Lemaux PG (2005) Evolutionary expansion, gene structure, and expression of the rice wall-associated kinase gene family. Plant Physiol **139:** 1107–1124

Zheng X, Levine D, Shen J, Gogarten SM, Laurie C, Weir BS (2012) A high-performance computing toolset for relatedness and principal component analysis of SNP data. Bioinformatics **28:** 3326–3328