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

TESE DE DOUTORADO

Identificação de genes associados com a eficiência na aquisição de fósforo em milho, com foco nos genes *PSTOL1, STR1* e *STR2*

DOUTORANDO: **Gabriel Corradi Azevedo**

ORIENTADORA: **Dra. Cláudia Teixeira Guimarães**

CO-ORIENTADORA: **Dra. Sylvia Morais de Sousa**

BELO HORIZONTE

Fevereiro/2015

GABRIEL CORRADI AZEVEDO

Identificação de genes associados com a eficiência na aquisição de fósforo em milho, com foco nos genes *PSTOL1, STR1* e *STR2*

Tese apresentada ao Programa de Pós-Graduação em Genética do Departamento de Biologia Geral do Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais como requisito parcial para obtenção do título de Doutor em Genética.

Orientadora: Dra. Cláudia Teixeira Guimarães
Co-orientadora: Dra. Sylvia Morais de Sousa

BELO HORIZONTE
Fevereiro/2015

Pós-Graduação em Genética
Departamento de Biologia Geral, ICB
Universidade Federal de Minas Gerais
Av. Antônio Carlos, 6627 - C.P. 486 - Pampulha - 31270-901 - Belo Horizonte - MG
e-mail: pg-gen@icb.ufmg.br FAX: (+31) - 3409-2570

"Identificação de genes associados com a eficiência na aquisição de fósforo em milho, com foco nos genes PSTOL1, STR1 e STR2"

Gabriel Corradi Azevedo

Tese aprovada pela banca examinadora constituída pelos Professores:

Dra. Cláudia Teixeira Guimarães - Orientadora
EMBRAPA

Dra. Sylvia Morais de Sousa
EMBRAPA - MILHO E SORGO

Dra. Fátima Maria de Souza Moreira
UFLA

Dra Eliane Aparecida Gomes
EMBRAPA - MILHO E SORGO

Dr. Luciano da Costa e Silva,
Statistical Analysis System - SAS

Dra. Maria Raquel Santos Carvalho
UFMG

Belo Horizonte, 11 de fevereiro de 2015.

Dedico esta tese aos meus pais Henrique e Zólvia,
À Guilherme e Gabriela,
pelo amor, confiança e compreensão.

AGRADECIMENTOS

Ao programa de Pós-Graduação em Genética e à Universidade Federal de Minas Gerais pela oportunidade de formação profissional.

A CAPES pela concessão da bolsa.

A Embrapa Milho e Sorgo por me dar todas as condições de exercer, com o máximo de excelência, meu trabalho;

Aos meus orientadores, família, amigos e pessoas especiais que estiveram presentes e contribuíram de alguma forma para a realização deste trabalho.

Acima de tudo, agradeço a Deus.

“O todo sem a parte não é todo,
A parte sem o todo não é parte,
Mas se a parte o faz todo, sendo parte,
Não se diga, que é parte, sendo todo.”

Gregório de Matos Guerra (1636-1696)

SUMÁRIO

LISTA DE FIGURAS

CAPÍTULO 1

ADDITIONAL FIGURES

CAPÍTULO 2

MATERIAL SUPLEMENTAR

LISTA DE TABELAS

CAPÍTULO 1

ADDITIONAL TABLES

CAPÍTULO 2

RESUMO

A baixa disponibilidade de fósforo (P) é uma das principais limitações para a produtividade do milho, principalmente em solos tropicais. Considerando a fixação e a baixa mobilidade do P no solo, modificações na morfologia radicular e associação com fungos micorrízicos arbusculares representam importantes estratégias adotadas pelas plantas para maximar a exploração do solo em condições de deficiência desse nutriente. Estudos identificaram o gene *OsPSTOL1* relacionado ao crescimento radicular, absorção de P e produtividade de grãos em arroz . Com objetivo de identificar possíveis homólogos do *OsPSTOL1* em milho foi adotada uma abordagem que combinou genômica comparativa e mapeamento de QTLs para características da morfologia radicular, acúmulo de biomassa e conteúdo de P utilizando uma população de linhagens recombinantes (RILs) derivada do cruzamento entre duas linhagens contrastantes para eficiência no uso do P, L3 e L22. Com base em modelos de características simples e múltiplas, 13 regiões genômicas foram identificadas. Entre os seis genes preditos identificados no genoma do milho com identidade de sequência superior a 55% com *OsPSTOL1*, quatro foram co-localizados com QTLs nos cromossomos 3, 4 e 8. Esses genes foram mais expressos nas raízes das linhagens parentais que contribuiram com os alelos para o aumento dos seus respectivos fenótipos. Conjuntamente, esses resultados indicam que pelo menos quatro genes candidatos podem estar relacionados com a morfologia radicular e ao conteúdo de P em milho. Na linha de associação micorrízica, dois transportadores *half-size* ABCG, *STR1* e *STR2*, foram identificados em arroz e em *Medicago truncatula*, como sendo indispensáveis para desenvolvimento dos arbúsculos e, consequentemente, para o sucesso da associação micorrízica. No presente trabalho, genes similares aos *STR1* e *STR2* foram identificados no genoma do milho, sendo co-localizados com QTLs para morfologia radicular nos cromossomos 10 e 3, respectivamente. As linhagens L3 e L22 apresentaram padrão de expressão similar para ambos os genes *STR*, bem como altos níveis de colonização micorrízica e arbúsculos bem desenvolvidos. Esses genes foram expressos somentes em raízes micorrizadas, seguindo o padrão observado em arroz. Tais resultados sugerem uma possível relação dos genes *STR1* e *STR2* com o desenvolvimento dos arbúsculos em milho e evidenciam que diferenças na eficiência no uso do P entre as linhagens L3 e L22 não são relacionadas com a associação micorrízica. No entanto, estudos adicionais, serão necessários para validar esses genes candidatos como homólogos funcionais ao *PSTOL1*, *STR1* e *STR2* em milho.

Palavras-chave: aquisição de P, genômica comparativa, proteína quinase, *OsPSTOL1*, *STR1* e *STR2*, associação micorrízica, *Zea mays*.

ABSTRACT

Low phosphorus (P) availability is a primary constraint for maize productivity, mainly in tropical soils. Considering the P fixation and low mobility in the soil, modifications on root morphology and association with arbuscular mycorrhizal fungi are important strategies adopted by plants to maximize soil exploitation under phosphorus starvation conditions. *OsPSTOL1* was a gene associated to early root growth, P uptake and grain yield in rice. In order to identify putative homologs to *OsPSTOL1* in maize, we combined comparative genomics and QTL mapping for root traits, seedling dry weight and P content in nutrient solution in a maize RIL population derived from the cross between two maize lines contrasting for P use efficiency, L3 and L22. Based on single and multi-trait models, 13 genomic regions were identified. Among the six predicted maize genes identified sharing more than 55% of amino acid sequence identity with *OsPSTOL1*, four genes were co-localized to QTLs on chromosomes 3, 4 and 8. These genes were more expressed in roots of the parental lines that contributed the alleles enhancing the respective phenotypes. Taken together, these results suggest that at least four of these candidates can be related to root morphology and P content in maize. Considering the mycorrhizal association, two half-size ABCG transporters, *STR1* and *STR2*, were identified in rice and *Medicago truncatula*, as being indispensable to arbuscules development and, consequently, to mycorrhizal association success. Similar genes to *STR1* and *STR2* were identified in the maize genome, being co-localized to QTLs related to root morphology traits identified on chromosomes 10 and 3, respectively. The L3 and L22 lines showed a similar gene expression profile for both *STR* genes, as well high levels of root colonization and arbuscules well developed. The *STR* genes were expressed only in inoculated roots, following the same pattern as observed in rice. These results indicated a possible role of *STR1* and *STR2* to arbuscules development in maize and revealed that the difference in P use efficiency between L3 and L22 is not related to mycorrhizal association. However, additional studies are needed to validate these genes as functional homolog to *OsPSTOL1*, *STR1* and *STR2* in maize.

Keywords: phosphorus acquisition, comparative genomics, protein kinase, *OsPSTOL1,* STR1 and STR2, mycorrhizal association, *Zea mays*.

DELINEAMENTO DA TESE

A tese está constituída por uma introdução geral, seguida por dois capítulos abordando as duas linhas de pesquisa desenvolvidas durante o doutorado, estando subdividida da seguinte maneira:

- Resumo e abstract.
- Introdução geral: contendo uma revisão bibliográfica sobre os principais temas abordados nos dois capítulos da tese.
- Objetivo geral e objetivos específicos.
- Capítulo 1 corresponde a um artigo científico submetido a revista BMC Plant Biology, intitulado: “QTL mapping and expression analyses reveal that *ZmPSTOL* are candidate genes related to root morphology and phosphorus acquisition in maize”.
- Capítulo 2, intitulado: "Análise funcional de homólogos aos genes *STR1* e *STR2*, relacionados à associação micorrízica, em milho", contém o trabalho realizado durante o doutorado sanduíche no laboratório da Dra. Uta Paszkowski, na Universidade de Cambridge, Inglaterra.
- Conclusão geral e perspectivas futuras.
- Anexo, que contém o artigo científico publicado como co-autor na revista *Plant Physiology*, intitulado: "*Duplicate and conquer: multiple homologs of phosphorus-starvation tolerance 1 enhance phosphorus acquisition and sorghum performance on low-P soils"*.

INTRODUÇÃO GERAL

A cultura do milho

O Brasil é um país eminentemente agrícola, exercendo esse setor um forte impacto na sustentabilidade sócio-econômica nacional. Um dos pilares da agricultura do país é o milho (*Zea mays* L.), uma vez que a produção prevista para a safra 2014/15 é superior a 80 milhões de toneladas, o que corresponde a aproximadamente 7% da produção mundial (FAOstat, 2014). Esse cereal é um dos produtos de maior destaque do setor agrícola nacional, devido às suas amplas formas de utilização, que vão desde a alimentação animal, que representa a maior parte de seu consumo, até a indústria de alta tecnologia.

Além da sua importância econômica e social, o milho é uma das espécies cultivadas mais bem estudadas, tendo seu genoma sequenciado e disponibilizado (www.maizesequence.org). O milho é uma espécie diplóide e seu genoma está organizado em 10 cromossomos com tamanho estimado variando de 2,3 a 2,7 Gb, com um número predito de genes superior a 30 mil. Os éxons dos genes de milho são similares em tamanho quando comparados a seus homólogos em sorgo e arroz, no entanto, os íntrons são maiores, devido a inserção de elementos repetitivos (Schnable et al., 2009). Aproximadamente 85% do genoma do milho é composto por milhares de famílias de elementos transponíveis, dispersos de maneira não uniforme ao longo do genoma (Schnable et al., 2009; Llaca et al., 2011).

.

A importância do fósforo na produção agrícola

A baixa disponibilidade de fósforo (P) representa um dos principais entraves à produtividade das culturas agrícolas, incluindo o milho. Mais de 70% da biomassa terrestre ocorre em solos com baixo conteúdo de fósforo, incluindo mais da metade das terras agricultáveis do planeta (Lynch, 2011). O P é considerado um dos mais importantes nutrientes para o desenvolvimento da planta, sendo de fundamental importância em processos metabólicos ligados à energia e à ativação enzimática, participando também na formação de compostos bioquímicos como ácidos nucléicos e fosfolipídeos (Araújo & Machado, 2006; Parentoni et al., 2011).

Comparado a outros nutrientes, o P é considerado o de menor mobilidade e o menos prontamente disponível na rizosfera (Raghotama, 1999; Shen et al., 2011). O P é absorvido do solo pelas plantas na forma de ortofosfato ($H_2PO^-_4$ e HPO^-_4) que ocorrem em concentrações extremamente baixas na solução do solo (0,1 a 10 µM), consideradas subótimas para a produção de grãos (Vance et al., 2003). Além disso, o coeficiente de

difusão do P no solo é muito baixo (10^{-12} a 10^{-15} m^2s^{-1}), fazendo com que a absorção pela planta crie uma zona de depleção desse nutriente ao redor da raiz. Após poucos dias de absorção, a concentração de P na rizosfera pode reduzir de 30 a 50% e a zona de depleção estender-se até cerca de 2 mm da superfície radicular (Jungk, 1987; Araújo & Machado, 2006; Ramaekers et al., 2010). A biodisponibilidade de P é ainda mais reduzida devido ao fato desse elemento poder ser precipitado na forma de fosfatos de ferro, como a estrengita ($FePO_4.2H_2O$), e de Al, como a variscita ($AlPO_4.2H_2O$), em solos ácidos e na forma de forma de hidroxiapatita ($3Ca_3(PO_4)_2.Ca(OH)_2$) e carbonatoapatita ($3Ca_3(PO_4)_2.CaCO_3$) em solos calcáreos, reduzindo a sua solubilidade e a disponibilidade para a planta (Marschner, 1995; Vance et al., 2003).

Nesse sentido, nos diversos sistemas de produção agrícola, principalmente nos solos tropicais, são necessárias aplicações de grandes quantidades de adubos fosfatados no solo para garantir a produtividade, uma vez que a eficiência de absorção do P adicionado pode ser inferior a 20% (McLaughlin & James, 1991; Zhu et al., 2003). Outro grande problema relacionado ao P é que ele constitui um recurso natural não renovável, e, seu uso indiscriminado, pode levar ao esgotamento das jazidas em até 60 anos (Vance et al., 2003; Vance & Chiou, 2011).

Mecanismos para aumento da eficiência de aquisição do fósforo em plantas

Existem várias definições e métodos para calcular a eficiência de uma cultura (Moll et al., 1982; Batten, 1992; Manske et al., 2001; Shenoy & Kalagudi, 2005; Wang et al., 2010). Dentre essas definições, destaca-se a estabelecida por Moll et al. (1982), que definiram eficiência de uso de um nutriente (EFUSO) como gramas de grãos produzidos por grama de nutriente suprido à cultura. Este índice é dividido em dois componentes principais: eficiência de aquisição (EAQ) e eficiência de utilização interna (EUTIL). EAQ é definido como gramas de nutriente absorvido pela planta dividido por gramas de nutriente suprido a cultura, enquanto EUTIL é definido como grama de grãos produzidos por unidade de nutriente absorvido.

Utilizando a metodologia proposta por Moll et al. (1982), Parentoni & Souza Júnior (2008) determinaram a importância relativa da eficiência de aquisição e de utilização interna de fósforo em 28 genótipos tropicais de milho, avaliados em três ambientes com baixa e dois com alta disponibilidade de P. Os autores verificaram que a EAQ foi duas e três vezes mais importante que a EUTIL em solos com baixa e alta disponibilidade de P, respectivamente. Além disso, tais características não foram correlacionadas significativamente entre si, indicando um possível envolvimento de mecanismos e genes distintos no controle de cada uma delas.

Vários são os mecanismos desenvolvidos pelas plantas quando submetidas à limitação de P, indicando que essa é uma característica geneticamente complexa. Tais mecanismos podem ser relacionados ao aumento na eficiência na aquisição do P do solo, que englobam alterações nas características do sistema radicular, interações com microrganismos do solo e modificações das características químicas da rizosfera (Vance et al., 2003; Ramaekers et al., 2010; Lynch, 2011). Já o aumento da eficiência de utilização interna do P pode envolver mecanismos relacionados com o transporte, a remobilização, a distribuição e o metabolismo do P absorvido.

Li et al. (2008) combinaram análises proteômicas, fisiológicas e celulares das raízes de linhagens de milho cultivadas sob condições normais e com deficiência nesse nutriente e verificaram que diferenças na secreção de citrato, no metabolismo de açúcares e na proliferação das células radiculares foram as principais razões para maior tolerancia ao baixo P na linhagem tolerante ao P comparada com a sensível. Os autores concluiram que a linhagem tolerante apresentou um sistema radicular com maior capacidade para mobilização do P externo e um aumento da divisão celular no meristema radicular sob deficiência de P.

Devido à mobilidade limitada do P no solo, sua aquisição é altamente dependente da proximidade com o sistema radicular (Vance et al., 2003; Lambers et al., 2006; Lynch, 2011). Dessa forma, uma das principais estratégias observadas em plantas para aumentar a aquisição de P consiste na maximização da exploração do solo por meio da proliferação e da extensão de suas raízes (Mollier & Pellerin, 1999; Shenoy & Kalagudi, 2005; Ramaekers et al., 2010; Lynch, 2011; Shen et al., 2011).

Dentre as adaptações adotadas pelas plantas para aumentar a superfície de absorção de P, as alterações na morfologia de pelos radiculares são aquelas com menor custo energético (Gahoonia & Nielsen, 2004; Kochian et al, 2004; Araújo & Machado, 2006, Haling et al, 2013). Pelos radiculares são extensões subcelulares de células epidérmicas que, em condições de baixo P, podem ser responsáveis por cerca de 90% do P total absorvido pelas plantas (Raghothama, 1999). Em milho, foi observado no mutante *rth3* (*roothairless 3*), que é afetado na elongação de pelos radiculares, significativas perdas na produtividade de grãos em relação ao controle (Hochholdinger, et al., 2008).

Outra adaptação observada em plantas submetidas a condições limitantes de P é o desenvolvimento de raízes laterais, que exercem importante função para minimizar a deficiência de P, uma vez que possibilitam o aumento da superfície de absorção do sistema radicular (Zhu & Lynch, 2004; Pérez-Torres et al., 2008; Niu et al., 2013). Zhu & Lynch (2004) avaliaram a importância do desenvolvimento de raízes laterais para a aquisição de P em uma população de RILs derivada do cruzamento entre B73 (P-ineficiente) e Mo17 (P-eficiente). Os resultados obtidos demonstraram que os genótipos com maior

desenvolvimento de raízes laterais apresentaram maior aquisição de fósforo e acúmulo de biomassa em comparação aos genótipos com menor desenvolvimento dessas raízes.

Em Arabidopsis, a baixa disponibilidade de P favoreceu o desenvolvimento das raízes laterais, induzindo o aumento da densidade e do comprimento dessas raízes, em detrimento à raiz primária (Williamson et al., 2001). Além disso, esses autores observaram que o número total de ramificações radiculares e o comprimento total do sistema radicular foram maiores nas plantas crescidas em alto P, levando a um consequente aumento da massa seca da parte aérea em relação ao baixo P.

Na maioria dos solos, as camadas superficiais são as que possuem maior biodisponibilidade de P. Portanto, plantas que apresentam o sistema radicular mais superficial e ramificado aumentam significativamente o volume de solo explorado e a probabilidade de adquirir o P (Lynch, 2011). Estudos realizados em feijão indicaram que as características radiculares associadas à maior exploração das camadas superficiais do solo incluem, maior curvatura gravitrópica, aumento de raízes adventícias e maior dispersão de raízes laterais (Lynch & Brown, 2001). Zhu et al. (2005a) verificaram que genótipos de milho com raízes superficiais apresentaram maior desenvolvimento e acúmulo de P em relação aos genótipos que possuíam raízes profundas.

Plantas submetidas à deficiência de P tendem a aumentar a relação de massa seca da raiz em relação à parte aérea, com as raízes tornando-se mais finas, as quais podem explorar um volume maior de solo por unidade de massa radicular. No entanto, o custo de manutenção dessas raízes pode ser maior, uma vez que elas são substituídas mais frequentemente (Gahoonia & Nielsen, 2004; Araújo & Machado, 2006).

Interação com fungos micorrízicos arbusculares

A associação com microrganismos representa outro importante mecanismo para o aumento da eficiência na aquisição de P em plantas. Nesse contexto, destacam-se os fungos micorrízicos arbusculares (FMA) que colonizam as raízes das plantas formando a simbiose mais comum entre plantas e microrganismos (Smith & Read, 2008). Os FMAs geralmente são considerados benéficos por atuarem como extensões do sistema radicular favorecendo o desenvolvimento e a sobrevivência do macrosimbionte (Sieverding, 1991, Moreira & Siqueira, 2006; Berbara et al., 2006). Esses fungos produzem hifas que são estruturas extremamente longas e finas, o que possibilitam a exploração além da zona de depleção, em volumes do solo inatingíveis pela raiz, incluindo os pelos radiculares (Sieverding, 1991; Moreira & Siqueira, 2006, Parniske, 2008). Além disso, o movimento mais rápido do P dentro das hifas, por meio do aumento da afinidade dos íons de P, e a solubilização do P do solo, obtida pela liberação de ácidos orgânicos e enzimas fostatases,

são outros mecanismos relacionados a essa maior eficiência desses microrganismos na absorção do P do solo (Bolan, 1991; Berbara et al., 2006).

Inúmeros estudos apontam para os beneficios da inoculação com FMA em diversas culturas agrícolas, tais como, café (Collozi-Filho et al., 1994), pimenta (Garmendia et al., 2004), banana (Borges et al., 2007), milho (Hu et al., 2009) e oliva (Meddad-Hamza et al., 2010). A inoculação com FMA proporcionou ganhos de matéria seca, produção de grãos e acúmulo de N, P, K e Zn no sorgo e na soja crescidos em casa de vegetação (Bressan et al., 2001) e aumento nos teores e acumulação total de N, P e K na parte aérea do milho e da *Crotalaria juncea* cultivadas a campo (Souza, 2010).

Para que a associação micorrízica tenha início, é necessária uma interação molecular pré-simbiótica entre a planta e o FMA. As raízes das plantas exudam estrigolactonas, que induzem a germinação dos esporos e ramificação das hifas. Em contrapartida, os fungos produzem os fatores *Myc* que ativam os genes da planta relacionados com a simbiose. Após entrar em contato com a superfície radicular, a hifa do FMA se diferencia em uma estrutura de fixação, denominada hifopódio. Como consequência da ligação hifa-raiz, a planta produz o aparato de pré-penetração (PPA). Subsequentemente, a hifa penetra no PPA e cresce em direção ao córtex, onde o fungo deixa a célula vegetal e penetra no eixo radicular, via apoplasto, se ramificando e crescendo lateralmente ao longo do referido eixo. Por fim, a hifa induzirá o desenvolvimento de estruturas *PPA-like* nas células corticais, onde irá penetrar e se ramificar formando estruturas denominadas arbúsculos (Gutjahr et al, 2008; Parniske, 2008).

A troca de nutrientes na associação micorrízica ocorre através dos arbúsculos, que são sempre circundados por uma mebrana derivada da planta, denominada membrana periarbuscular, que é contínua com a membrana plasmática da planta e exclui o fungo do citoplasma vegetal. Transportadores que atuam nessa membrana são fundamentais para a simbiose micorrízica, pois sua atividade é essencial para o desenvolvimento do arbúsculo na simbiose micorrízica (Parniske, 2008).

A partir de um *screen* genético realizado em uma população mutante de *Medicago truncatula*, foi identicado um mutante que apresentou arbúsculos atrofiados em relação ao tipo selvagem quando inoculado com *Glomus versiforme*, sendo denominado *str* (*stunted arbuscules*) (Zhang et al, 2010). O gene responsável por essa mutação, denominado *STR*, foi identificado por meio de clonagem posicional e codifica um transportador *half-size ATP binding cassette* (ABC) da subfamília ABCG (Zhang et al, 2010). A caracterização do *STR* possibilitou que um segundo transportador *half-size* ABC fosse identificado, sendo denominado *STR2*. O heterodímero STR/STR2 é localizado na membrana periarbuscular, onde sua atividade é necessária para o desenvolvimento do arbúsculo e, consequentemente, para uma simbiose micorrízica funcional (Zhang et al., 2010).

Em arroz, Gutjahr et al. (2012) identificaram dois genes induzidos pela micorrização que codificavam transportadores ABC com alta similaridade aos genes *STR1* e *STR2* descritos em *M. truncatula*. Mutações em ambos transportadores de arroz inibiram o desenvolvimento do arbúsculo de diferentes espécies de FMA como, *Glomus intraradices* e *Gigaspora rosea*, gerando mutantes com arbúsculos pequenos e atrofiados (Gutjahr et al., 2012), semelhantes ao fenótipo apresentado pelo mutante *str* de *M. truncatula*.

Mapeamento de QTLs relacionados com morfologia radicular e associação micorrízica em milho

Características relacionadas com a eficiência na aquisição do P são consideradas complexas, sendo, dessa forma, controlada por vários genes (Tuberosa, 2009; Ramaekers et al., 2010). Uma estratégia apropriada para o estudo de características de herança complexa é o mapeamento de QTLs (*Quantitative Trait Loci*), que são definidos como regiões do genoma responsáveis pela expressão de caracteres quantitativos. O mapeamento de QTLs permite mensurar o número de locos envolvidos na herança de uma característica, bem como suas localizações cromossômicas e seus efeitos. Tais informações podem ser úteis aos programas de melhoramento genético possibilitando o uso de marcadores moleculares na seleção assistida para características de interesse (Collard et al., 2005; Schuster & Cruz, 2008).

Em milho, poucos QTLs relacionados com associação micorrízica foram descritos na literatura. Esses QTLs foram relacionados com a responsividade e a colonização micorrízica, sendo responsáveis por 5.6 e 6.5 % da variação fenotípica para essas características, respectivamente (Kaeppler et al., 2000). Em contrapartida, inúmeros QTLs para características da morfologia radicular foram identificados em condições de campo (Chen & Xu, 2011; Cai et al., 2012) e sob condições controladas (Tuberosa et al., 2002; Liu et al., 2008; Ruta et al., 2010; Burton et al., 2014), explicando de 4 a 15% da variação fenotípica para o comprimento, número e plasticidade de raízes laterais, seminais e pelos radiculares sob níveis contrastantes de P (Zhu et al., 2005b, 2005c, 2006). Outros estudos detectaram vários QTLs com efeitos pequenos e *clusters* de QTLs em regiões específicas do genoma para peso seco de raiz, conteúdo de P e índices de eficiência no uso de P (Chen et al., 2008; 2009).

Alguns estudos realizaram uma abordagem de meta-análise de QTLs, visando identificar regiões que concentram vários QTLs controlando diferentes características relacionadas com a eficiência no uso de P (Tuberosa et al., 2003; Hund et al., 2011; Zhang et al., 2014). Uma meta-análise utilizando 15 estudos de QTLs para a morfologia do sistema radicular permitiu a identificação de *clusters* de QTLs nos bins 1.07, 2.04, 2.08, 3.06, 6.05 e

7.04 (Hund et al., 2011). No entanto, mesmo com o grande número de QTLs identificados, poucos genes foram diretamente relacionados a essas características até o momento.

O gene responsável pela tolerância à deficiência de P em arroz (PSTOL1)

Diferenças na eficiência do uso de P foram relacionadas com a aquisição desse nutriente em arroz, em detrimento da utilização interna, que teve um efeito negligente em linhagens derivadas de retrocruzamento (Wissuwa et al., 1998). Além disso, esses autores identificaram QTLs de pequenos efeitos relacionados com a aquisição e a eficiência de utilização interna de P nos cromossomos 2, 6 e 10 e um QTL de grande efeito no cromossomo 12. Esse QTL de efeito maior para a aquisição de P foi também identificado em outra população de RILs (Ni et al., 1998). Posteriormente, tal QTL foi denominado *Phosphorus uptake 1* (*Pup1*) explicando em torno de 80% da variação fenotípica para tolerância a deficiência de P em arroz (Wissuwa et al., 2002). O loco *Pup1* aumentou a absorção de P (Wissuwa & Ae, 2001; Wissuwa et al., 2002) e conferiu aumentos significativos na produtividade de grãos em linhagens semi-isogênicas de arroz (NILs), sob deficiência de P (Chin et al., 2010; 2011). Adicionalmente, *Pup1* foi associado ao aumento da taxa de crescimento radicular sob baixa disponibilidade P (Wissuwa & Ae, 2001).

O mapeamento fino do *Pup1* na variedade doadora do QTL, Kasalath, revelou um locus complexo de 278 kbp com grandes inserções e deleções, e baixa conservação em relação às variedades Nipponbare (genoma de referência *japonica*) e 93-11 (genoma de referência *indica*) (Heuer et al., 2009). Sessenta e oito genes foram identificados nessa região, dentre os quais apenas oito não codificavam elementos transponíveis (Heuer et al., 2009; Chin et al., 2011). A avaliação desses genes revelou que somente o *OsPupk46*, que codificava uma proteína quinase, estava especificamente presente no genoma da Kasalath e apresentou um aumento nos níveis de expressão em baixas condições de P (Chin et al., 2011; Gamuyao et al., 2012). Esse gene foi considerado o mais provável candidato, sendo posteriormente clonado e denominado *Phosphorus-starvation tolerance 1* (*PSTOL1*) (Gamuyao et al., 2012). Quando superexpresso, o *PSTOL1* conferiu um aumento no crescimento radicular, aumentando significativamente a produtividade de grãos sob deficiência de P no solo, provavelmente pela maior aquisição de P e de outros nutrientes (Gamuyao et al., 2012). Através de uma abordagem que combinou genômica comparativa com mapeamento associativo, foram identificados no genoma do sorgo homólogos funcionais ao gene *OsPSTOL1*, que desempenham um papel na absorção de P por meio de alterações na morfologia e na arquitetura do sistema radicular de sorgo (Hufnagel et al., 2014).

Dado o forte efeito fenotípico do *PSTOL1*, e considerando a conservação do genoma entre arroz, milho e sorgo (Hulbert et al., 1990; Ahn & Tanksley, 1993; Schnable et al., 2009; Llaca et al., 2011), a identificação desse gene no arroz e sorgo permite uma estratégia de genômica comparativa para identificar homólogos funcionais no milho e usar essas informações para melhorar a tolerância à deficiência de P nessa cultura.

REFERÊNCIAS BIBLIOGRÁFICAS

Ahn, S., & Tanksley, S. D. Comparative linkage. Proceedings of the National Academy of Sciences, 90:7980-7984, 1993.

Araújo, A. P.; Machado, C. T. T. Fósforo. In: Fernandes, M. S. (ed). Nutrição mineral de plantas. Sociedade Brasileira de Ciência do Solo, Viçosa, pp 253-280, 2006.

Batten, G. D. A review of phosphorus efficiency in wheat. Plant and Soil, 1992, 146:163-168.

Berbara, R. L. L.; de Souza, F. A.; Fonseca, H. M. A. C. Fungos micorrízicos arbusculares: muito além da nutrição. In: Fernandes, M. S. (ed). Nutrição mineral de plantas. Sociedade Brasileira de Ciência do Solo, Viçosa, pp 53-85, 2006.

Bolan, N. S. A critical review on the role of mycorrhizal fungi in the uptake of phosphorus by plants. Plant and Soil, 134:189-207, 1991.

Borges, A. J. S.; Trindade, A. V.; Matos, A. P.; Peixoto, M. F. S. Redução do mal-do-panamá em bananeira-maça pela inoculação de fungo micorrízico arbuscular. Pesquisa Agropecuária Brasileira, 42:35-41, 2007.

Bressan, W.; Vasconcellos, C. A.; Purcino, A. A. C. Fungos micorrízicos e fósforo, no crescimento, nos teores de nutrientes e na produção do sorgo e soja consorciados. Pesquisa Agropecuária Brasileira, 36:315-323, 2001.

Burton, A. L.; Johnson, J. M.; Foerster, J. M.; Hirsch, C. N.; Buell, C. R.; Hanlon, M. T.; Kaeppler, S. M.; Brown, K. M.; Lynch, J. P. QTL mapping and phenotypic variation for root architectural traits in maize (*Zea mays* L.). Theoretical and Applied Genetics, 127:2293-2311, 2014.

Cai, H.; Chen, F.; Mi, G.; Zhang, F.; Maurer, H. P.; Liu, W.; Reif, J. C.; Yuan, L. Mapping QTLs for root system architecture of maize (*Zea mays* L.) in the field at different developmental stages. Theoretical and Applied Genetics, 125:1313-1324, 2012.

Chen, J.; Xu, L. The candidate QTLs affecting phosphorus absorption efficiency. Frontiers Agriculture China, 5:456-462, 2011.

Chen, J.; Xu, L.; Cai, Y.; Xu, J. Identification of QTLs for phosphorus utilization efficiency. Euphytica, 167:245-252, 2009.

Chen, J.; Xu, L.; Cai, Y.; Xu, J. QTL mapping of phosphorus efficiency and relative biologic characteristics in maize (*Zea mays* L.) at two sites. Plant and Soil, 313:251-266, 2008.

Chin, J.; Gamuyao, R.; Bustamam, M.; Wissuwa, M.; Heuer, S. Developing Rice with high yield under phosphorus deficiency: Pup1 sequence to application. Plant Physiology, 156:1202-1216, 2011.

Chin, J.; Gamuyao, R.; Heuer, S.; Lu, X.; Ismail, A.; Haefele, S.; Wissuwa, M. Development and application of gene-based markers for the major rice QTL Phosphorus uptake. Theoretical Applied Genetics, 120:1073-1086, 2010.

Collard, B. C. Y.; Jahufer, M. Z. Z.; Brouwer, J. B.; Pang, E. C. K. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. Euphytica, 142:169-196, 2005.

Colozzi-Filho, A.C.; Siqueira, J. O.; Saggin Junior, O. J.; Guimaraes, P. T. G.; Oliveira, E. Efetividade de diferentes fungos micorrízicos arbusculares na formação de mudas, crescimento pós-trasplante e produção do cafeeiro. Pesquisa Agropecuária Brasileira, 29:1397-1406, 1994.

Gahoonia, T. S.; Nielsen, N. E. Root traits as tools for creating phosphorus efficient crop varieties. Plant and Soil, 260:47-57, 2004.

Gamuyao, R.; Pariaska-Tanaka, J.; Pesaresi, P.; Catausan, S.; Dalid, C.; Slamet-Loedin I.; Tecson-Mendoza, E. M.; Wissuwa, M.; Heuer, S. The protein kinase Pstol1 from traditional rice confers tolerance of phosphorus deficiency. Nature, 488:535-539, 2012.

Garmendia, I., Goicoechea, N., Aguirreolea, J. Effectiveness of three Glomus species in protecting pepper (*Capsicum annuum* L.) against verticillium wilt. Biological Control, 31:296-305, 2004.

Gutjahr, C.; Banba, M.; Croset, V.; An, K.; Miyao, A.; An, G.; Hirochika, H.; Imaizumi-Anraku, H.; Paszkowski, U. Arbuscular mycorrhiza-specific signaling in rice transcends the common symbiosis signaling pathway. The Plant Cell, 20:2989-3005, 2008.

Gutjahr, C.; Radovanovic, D.; Geoffroy, J.; Zhang, Q.; Siegler, H.; Chiapello, M.; Casieri, L.; An, K.; An, G.; Guiderdoni, E.; Kumar,C. S.; Sundaresan, V.; Harrison, M. J.; Paszkowski, U. The half-size ABC transporters STR1 and STR2 are indispensable for mycorrhizal arbuscule formation in rice. The Plant Journal, 69:906-920, 2012.

Haling, R. E.; Brown, L. K.; Bengough, A. G.; Young, I. M.; Hallett, P. D.; White, P. J.; George, T. S. Root hairs improve root penetration, root-soil contact, and phosphorus acquisition in soils of different strength. Journal of Experimental Botany, 64: 3711-3721, 2013.

Heuer, S.; Lu, X.; Chin, J. H.; Tanaka, J. P.; Kanamori, H.; Matsumoto, T.; De Leon, T.; Ulat, V. J.; Ismail, A. M.; Yano, M.; Wissuwa, M. Comparative sequence analyses of the major quantitative trait locus phosphorus uptake 1 (Pup1) reveal a complex genetic structure. Plant Biotechnology Journal, 7:456-457, 2009.

Hochholdinger, F.; Wen, T. J.; Zimmermann, R.; Chimot-Marolle, P.; da Costa e Silva, O.; Bruce, W.; Lamkey, K. R.; Wienand, U.; Schnable, P. S: The maize (*Zea mays* L.) ROOTHAIRLESS3 gene encodes a putative GPI-anchored, monocot-specific, COBRAlike protein that significantly affects grain yield. Plant Journal, 54:888-898, 2008.

Hu, J.; Lin, X.; Wang, J.; Dai, J.; Cui, X.; Chen, R.; Zhang, J. Arbuscular mycorrhizal fungus enhances crop yield and P-uptake of maize (*Zea mays* L.): A field case study on a sandy loam soil as affected by long-term P-deficiency fertilization. Soil Biology and Biochemistry, 41:2460-2465, 2009.

Hulbert, S. H.; Richter, T. E.; Axtello, J. D.; Bennetzen, J. L. Genetic mapping and characterization of sorghum and related crops by means of maize DNA probes. Proceedings of the National Academy of Sciences, 87:4251-4255, 1990.

Hund, A.; Reimer, R.; Messmer, R. A consensus map of QTLs controlling the root length of maize. Plant and Soil, 344:143-158, 2011.

Jungk, A. Soil root interactions in the rhizosphere affecting plant availability of phosphorus. Journal of Plant Nutrition, 10:1197-1224, 1987.

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

Lambers H, Shane M. W.; Cramer, M. D.; Pearse, S.J.; Veneklaas, E.J. Root structure and functioning for efficient acquisition of phosphorus: matching morphological and physiological traits. Annals of Botany, 98:693–713, 2006.

Li, K.; Xu, C.; Li, Z.; Zhang, K.; Yang, A.; Zhang, J. Comparative proteome analyses of phosphorus responses in maize (*Zea mays* L.) roots of wild-type and a low-P-tolerant mutant reveal root characteristics associated with phosphorus efficiency. The Plant Journal, 55:927–39, 2008.

Llaca, V.; Campbell, M. A.; Deschamps, S. Genome diversity in maize. Journal of Botany, 2011:1-10, 2011.

Lynch, J. P. Root phenes for enhanced soil exploration and phosphorus acquisition: tools for future crops. Plant Physiology, 156:1041-1049, 2011.

Lynch, J.P.; Brown, K. M. Topsoil foraging – an architectural adaptation of plants to low phosphorus availability. Plant and Soil, 237:225-237, 2001.

Manske, G. G. B.; Ortiz-monasterio, J. I.; Ginkel, M. Van, Gonza, R. M. Importance of P uptake efficiency versus P utilization for wheat yield in acid and calcareous soils in Mexico. European Journal of Agronomy, 14:261–274, 2001.

Marschner, H. Mineral nutritional of higher plants. 2. ed. Boston: Academic Press, 1995.

McLaughlin, M. L.; James, T. R. Effects of surface-applied phosphorus and superphosphate on the solution chemistry and phytotoxicity of subsurface aluminium: sand/solution and soil experiments. Australian Journal of Agricultural Research, 42:859-874, 1991.

Meddad-Hamza, A.; Beddiar, A.; Gollotte, A.; Lemoine, M. C.; Kuszala, C., Gianinazzi, S.; Biologie, D. De.; Mokhtar, U. B. Arbuscular mycorrhizal fungi improve the growth of olive trees and their resistance to transplantation stress. African Journal of Biotechnology, 9:1159-1167, 2010.

Moll, R.H.; Kamprath, E.J.; Jackson, W.A. Analysis and interpretation of factors, which contribute to efficiency of nitrogen utilization. Agronomy Journal, 74: 562-564, 1982.

Mollier, A.; Pellerin, S. Maize root system growth and development as influenced by phosphorus deficiency. Journal of Experimental Botany 50:487-497, 1999.

Moreira, F.S.; Siqueira, J.O. Microbiologia e bioquímica do solo. 2.ed. Lavras: UFLA, 726p, 2006.

Ni, J. J.; Wu, P.; Senadhira, D.; Huang, N. Mapping QTLs for phosphorus deficiency tolerance in rice (*Oryza sativa* L.). Theoretical and Applied Genetics, 97:1361-1369, 1998.

Niu, Y. F.; Chai, R. S.; Jin, G. L.; Wang, H.; Tang, C. X.; Zhang, Y. S. Responses of root architecture development to low phosphorus availability: a review. Annals of Botany, 112:391-408, 2013.

Parentoni, S. N.; Mendes, F. F.; Guimarães, L. J. M. Melhoramento para eficiência no uso de fósforo. In: Neto, R. F.; Borém, A. Melhoramento de plantas para condições de estresses abióticos. Viçosa: UFV, p.101-126, 2011.

Parentoni, S. N.; Souza Júnior, C. L. S. Phosphorus acquisition and internal utilization efficiency in tropical maize genotypes. Pesquisa Agropecuária Brasileira, 43:893-901, 2008.

Parniske, M. Arbuscular mycorrhiza: the mother of plant root endosymbioses. Nature Reviews Microbiology, 6:763-675, 2008.

Pérez-Torres, C.A.; López-Bucio, J.; Cruz-Ramírez, A.; Ibarra-Laclette, E.; Dharmasiri, S.; Estelle, M.; Herrera-Estrella, L. Phosphate availability alters lateral root development in Arabidopsis by modulating auxin sensitivity via a mechanism involving the TIR1 auxin receptor. Plant Cell, 20:3258-3272, 2008.

Raghothama, K. G. Phosphate acquisition. Annual Review of Plant Physiology and Plant Molecular Biology, 50:665-693, 1999.

Ramaekers, L.; Remans, R.; Rao, I. M.; Blair, M. W.; Vanderleyden, J. Strategies for improving phosphorus acquisition efficiency of crop plats. Field Crops Research,117:169-176, 2010.

Ruta, N.; Liedgens, M.; Fracheboud, Y.; Stamp, P.; Hund. A. QTLs for the elongation of axile and lateral roots of maize in response to low water potential. Theoretical and Applied Genetics, 120:621-631, 2010.

Schnable, P.S.; Ware, D.; Fulton, R.S.; Stein, J. C.; Wei, F.; Pasternak, S.; Liang, C.; Zhang, J.; Fulton, L.; Graves, T. A.; Minx, P.; Reily, A. D.; Courtney, L.; Kruchowski, S. S.; Tomlinson, C.; Strong, C.; Delehaunty, K.; Fronick, C.; Courtney B, Rock SM, Belter E, Du F, Kim K, Abbott RM, Cotton M, Levy A, Marchetto P, Ochoa, K.; Jackson, S. M.; Gillam, B.; et al: The B73 maize genome: complexity, diversity, and dynamics. Science, 326:1112-1115, 2009.

Schuster, I.; Cruz C. D. Análise de QTLs - Caráter quantitativo. In: Schuster, I.; Cruz C. D. (eds). Estatístiva genômica aplicada a populações derivadas de cruzamentos controlados. Viçosa, pp 279-322, 2008.

Shen, J.; L, Yuan.; J. Zhang.; H. Li.; Z. Bai.; X. Chen.; W. Zhang.; F, Zhang. Phosphorus dynamics: from soil to plant. Plant Physiology 156:997-1005, 2011.

Shenoy, V. V.; Kalagudi, G. M. Enhancing plant phosphorus use efficiency for sustainable cropping. Biotechnology Advances, 23:501-513, 2005.

Sieverding, E.; Friedrichsen, J.; Suden, W. Vesicular-arbuscular mycorrhiza management in tropical agrosystems. 371p, 1991.

Smith, S. E.; Read, D. J. Mycorrhizal Symbiosis (Third Edition), New York, 787p, 2008.

Souza, J. R. Influência dos cultivos de milho e Crotalaria juncea inoculados com fungo micorrízico arbuscular (FMA) no desempenho agronômico de batata-doce em sucessão. 52p. Dissertação (Mestrado em Fitotecnia). Instituto de Agronomia, Departamento de Fitotecnia, Universidade Federal Rural do Rio de Janeiro, Seropédica, RJ, 2010.

Tuberosa, R.; Salvi, S.; Sanguineti, M. C.; Maccaferri, M.; Giuliani, S.; Landi, P. Searching for quantitative trait loci controlling root traits in maize : a critical appraisal. Plant and Soil, 255:35-54, 2003.

Tuberosa, R.; Sanguineti, M.C.; Landi, P.; Giuliani, M.M.; Salvi, S.; Conti, S. Identification of QTLs for root characteristics in maize grown in hydroponics. Plant Molecular Biology, 48:697-712, 2002.

Vance, C. P.; Chiou, T. J. Phosphorus focus editorial. Plant Physiology, 155:987-988, 2011.

Vance, C. P.; Uhde-Stone, C.; Allan, D. L. Phosphorus acquisition and use: critical adaptations by plants for securing a nonrenewable resource. New Phytologist, 157:423-447, 2003.

Wang, X.; Shen, J.; Liao, H. Acquisition or utilization, which is more critical for enhancing phosphorus efficiency in modern crops? Plant Science, 179:302-306, 2010.

Williamson, L. C.; Ribrioux, S. P.; Fitter, A. H.; Leyser, H. M. Phosphate availability regulates root system architecture in Arabidopsis. Plant Physiology, 126:875-882, 2001.

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

Wissuwa, M.; Wegner, J.; Ae, N.; Yano, M. Substitution mapping of Pup1: a major QTL increasing phosphorus uptake of rice from a phosphorus-deficient soil. Theoretical and Applied Genetics, 105:890-897, 2002.

Wissuwa, M.; Yano, M.; Ae, N. Mapping of QTLs for phosphorus-deficiency tolerance in rice (*Oryza sativa* L.). Theoretical and Applied Genetics, 97:777-783, 1998.

Zhang, H.; Uddin, M. S.; Zou, C.; Xie, C.; Xu, Y.; Li, W-X. Meta-analysis and candidate gene mining of low-phosphorus tolerance in maize. Journal of integrative plant biology, 56:262-270, 2014.

Zhang, Q.; Blaylock, L.; Harrison, M. Two half-ABC transporters are essential for arbuscule development in arbuscular mycorrhizal symbiosis. Plant Cell, 22:1483–1497, 2010.

Zhu, J.; Kaeppler, S. M.; Lynch, J. Mapping of QTLs for lateral root branching and length in maize (*Zea mays* L.) under differential phosphorus supply. Theoretical and Applied Genetics, 111: 688-695, 2005b.

Zhu, J.; Kaeppler, S. M.; Lynch, J. P. Mapping of QTL controlling root hair length in maize (*Zea mays* L.) under phosphorus deficiency. Plant and Soil, 270:299-310, 2005a.

Zhu, J.; Kaeppler, S. M.; Lynch, J. P. Topsoil foraging and phosphorus acquisition efficiency in maize (*Zea mays*). Functional Plant Biology, 32:749-762, 2005a.

Zhu, J.; Lynch, J. The contribution of lateral rooting to phosphorus acquisition efficiency in maize (*Zea mays*) seedlings. Functional Plant Biology, 31:949-958, 2004.

Zhu, J.; Mickelson. S. M.; Kaeppler, S.M.; Lynch, J.P. Detection of quantitative trait loci for seminal root traits in maize (*Zea mays* L.) seedlings grown under differential phosphorus levels. Theoretical and Applied Genetics, 113:1-10, 2006.

Zhu, Y-G.; Smith, F.A.; Smith, S. E. Phosphorous efficiency and responses of barley (*Hordeum vulgare* L.) to arbuscular mycorrhizal fungi grown in highly calcareous soil. Mycorrhiza, 13:93-100, 2003.

OBJETIVOS

Objetivo geral

O objetivo desse estudo foi identificar genes e regiões genômicas relacionadas com a eficiência na aquisição de fósforo em milho, com foco nos homólogos aos genes *PSTOL1*, *STR1* e *STR2* de arroz.

Objetivos específicos

- Mapear QTLs e genes candidatos associados com morfologia radicular e parâmetros de eficiência no uso de P em milho;

- Avaliar os genes *STR1* e *STR2*, relacionados com associação micorrízica, em linhagens de milho contrastantes para aquisição de P.

CAPÍTULO 1

QTL mapping and gene expression analyses reveal that *PSTOL* candidate genes are associated with root morphology and phosphorus acquisition in maize

Gabriel C Azevedo[1], Adriana Cheavegatti-Gianotto[2], Bárbara F Negri[3], Bárbara Hufnagel[1], Luciano da Costa e Silva[2], Jurandir V Magalhaes[4], Antonio Augusto F Garcia[2], Ubiraci GP Lana[4], Sylvia M de Sousa[4], Claudia T Guimaraes[4*]

Author Affiliations:

[1]Departamento de Biologia Geral, Universidade Federal de Minas Gerais, Avenida Presidente Antônio Carlos, 6627, Belo Horizonte, MG 31270-901, Brazil

[2]Departamento de Genética, Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Caixa Postal 83, Piracicaba, SP 13400-970, Brazil

[3]Departamento de Bioengenharia, Universidade Federal de São João del-Rei, Praça Dom Helvécio, 74, São João del-Rei, MG 36301-160, Brazil

[4]Embrapa Maize and Sorghum, Road MG 424, km 65, Caixa Postal 151, Sete Lagoas, MG 35701-970, Brazil.

*Corresponding author: Claudia T Guimaraes, Nucleus of Applied Biology, Embrapa Maize and Sorghum, Road MG424, km 65, Sete Lagoas, MG, 35701-970, Brazil

Email: claudia.guimaraes@embrapa.br

phone: +55 31 3027 1300, fax: +55 31 3027 1279

Email addresses:

GCA: gabrielcorradiazevedo@yahoo.com.br

AC-G: acheavegatti@gmail.com

BFN: babinegri@hotmail.com

BH: barbarahufnagel@gmail.com

LCS: ldcesilva@gmail.com

JVM: jurandir.magalhaes@embrapa.br

AAFG: augusto.garcia@usp.br

UGPL: ubiraci.lana@embrapa.br

SMS: sylvia.sousa@embrapa.br

Abstract

Background: Phosphorus (P) is an essential macronutrient for crop production. Most agricultural soils contain low P availability due to P fixation and low mobility in the soil. Thus, the modification of root morphology is an important mechanism to improve P acquisition in plants. The gene *PSTOL1* was shown to increase early root growth, P acquisition and grain yield in rice and sorghum.

Results: Two complementary models for quantitative trait loci (QTL) mapping were applied to dissect the genetic complexity of root morphology and P acquisition in maize RILs during the seedling stage under low P. A total of 13 QTL regions were associated with single and/or multiple traits. Six predicted maize genes sharing more than 55% amino acid sequence identity with *OsPSTOL1* and three genes previously associated with root traits were mapped in our current linkage map. Of these candidate genes, four *ZmPSTOL* genes were co-localized with QTLs for root morphology, biomass accumulation and/or P content on chromosomes 3 (*ZmPSTOL3.06*), 4 (*ZmPSTOL4.05*) and 8 (*ZmPSTOL8.02* and *ZmPSTOL8.05_1*). Expression analysis revealed that these genes were more expressed in roots of the parental lines that contributed the alleles enhancing the respective phenotypes.

Conclusions: The combination of QTL and expression data indicates that *ZmPSTOL* genes may have a functional relationship with root morphology leading to enhanced P acquisition in maize, similar to the function of their homologs in rice and sorghum. Additionally, some QTLs identified in this study were also associated with yield performance in other studies, suggesting a possible application in marker-assisted selection to improve P-use efficiency in maize.

Keywords: phosphorus acquisition, protein kinase, SNP marker, *OsPSTOL1, Zea mays*

Background

The increasing demand for agricultural production poses a global challenge to improve the efficiency of phosphorus (P) use in plants due to its low availability in a large proportion of arable lands [1,2]. A large fraction of soil P is tightly fixed to the clay´s surface, which requires high amounts of P fertilizers for high-yielding farming systems, increasing production costs and hampering soil fertility management [3-5]. However, low-input farmers have limited access to P fertilizer, which is the second most used fertilizer for plant growth [6]. Maize is the most common grain produced worldwide and a major staple food in Africa and Latin America [7], where soils often show low P availability. Thus, improving maize P-use efficiency is expected to increase yield stability and, consequently, food security [1,4,8].

Due to limited soil mobility, P acquisition is greatly dependent on the proximity of this nutrient to the root system [3,9]. Additionally, P acquisition efficiency was the most important component of P-use efficiency in tropical maize germplasms evaluated in low- and high-P soils [10]. Thus, a well-developed root system is considered to be an important adaptation to maximize soil exploitation, enabling plants to improve P acquisition efficiency [11-13]. Studies have shown that plants that are more efficient in P acquisition presented higher root:shoot dry weight ratios [14,15], reduced root diameters [16], longer and denser root hairs [17], increased lateral roots [18], greater lateral branching and shallower basal roots [17,19]. These changes in root morphology are key strategies used by plants to improve soil exploitation at a minimal metabolic cost [5,20].

Root morphology is controlled by multiple genes in maize [13,21], but only a few of them such as *roothairless* (*Rth1*) [22], *brittle stalk-2-like protein 3* (*Bk2l3*) [23], and *rootless concerning crown and seminal roots* (*Rtcs*) [24] have been cloned and characterized. Nevertheless, several quantitative trait loci (QTLs) were mapped for root traits under contrasting conditions of P availability in nutrient solution [25-27], in glasshouse [28] and in the field [29-32]. These QTLs individually explained from 1 to 14% of the phenotypic variation, confirming the genetic complexity of root traits. Recently, a meta-analysis integrating ten QTL studies of traits associated with low-P tolerance in maize highlighted candidate genes in the consensus QTL regions that were putative homologs to previously characterized genes from other plants [33]. However, no candidate genes were directly associated with root morphology and P acquisition in maize.

In rice, a major QTL controlling phosphorus uptake (*Pup1*) was mapped to chromosome 12, explaining approximately 80% of the phenotypic variance of this trait [34]. Rice near isogenic lines (NILs) carrying the *Pup1* QTL showed a three-fold increase in P uptake and enhanced root surface area when grown in P-deficient soil [34,35]. Additionally, irrigated and upland rice varieties introgressed with *Pup1* showed a significant enhancement in grain yield in different low-P soils compared to their parents [36,37]. The gene underlying the *Pup1* locus was identified and named *Phosphorus-starvation tolerance 1* (*PSTOL1*), which encodes a serine/threonine kinase of the LRK10L-2 subfamily [38]. The overexpression of *PSTOL1* in two transgenic rice varieties enhanced the grain yield over 60% under low-P conditions due to larger root system (i.e., root length, and total root surface area), which also improved the uptake of P and other nutrients [38]. Furthermore, [39] showed that sorghum homologs to *OsPSTOL1* were associated with enhanced P uptake and grain yield in sorghum grown in a low-P soil due effects on root morphology and root system architecture. A remarkable conservation of protein-encoding genes among maize, sorghum and rice has been confirmed *in silico* using a genome sequencing comparison, which found that approximately 89% of the 11,892 maize gene families predicted in the B73

genome were shared with rice and sorghum [40]. Of these, genes encoding important adaptive traits are expected to be shared within these grass species.

Our major goal was to identify maize homologs to rice *PSTOL1* based on evidences of co-localization with QTLs for root morphology, biomass accumulation and P content in seedlings combined with analysis of gene expression.

Results

Transgressive segregation of phenotypic traits in RILs

Significant genetic variation in root morphology traits, biomass and phosphorus content in the seedlings were observed for the RIL population with high broad sense heritability estimates, which ranged from 0.65 for root:shoot dry weight ratio to 0.82 for root length (Additional File 1: Table S1). The P-efficient line, L3, tended to present superior phenotypic measurements for all traits compared with the P-inefficient line L22, with the exception of root diameter and root:shoot ratio (Figure 1). RILs showed larger phenotypic variations than both parents, suggesting transgressive segregation for all phenotypic traits (Figure 1). The parental lines belonged to distinct heterotic groups (L3, flint and L22, dent) [41], were genetically divergent based on SNP markers [42] and contrasted for grain yield under low P [43].

Surface area is an important root trait contributing to seedling dry weight and P content

Root length, root surface area and the surface area of fine roots were high and positively correlated with each other (Table 1). These traits exhibited comparable magnitudes to those observed in a sorghum diversity panel composed of 287 accessions [39] and 30 maize lines [44]. These root traits also showed strong correlation coefficients with seedling biomass accumulation (0.77 to 0.86), moderate correlations with total P content in the seedling (0.31 to 0.48), and negative correlations with root:shoot ratio (-0.29 to -0.38) (Table 1). In contrast, root diameter was negatively correlated with root length, root surface area and total seedling dry weight, but no significant correlation was found with total P content. The negative correlation between root length and root diameter (-0.62) was similar to the coefficients obtained for root diameter with lateral (-0.65) and non-lateral (-0.68) roots in temperate maize RILs [18].

To further investigate the relative importance of root traits on seedling dry weight and P content, we performed a path analysis, using the root traits as explanatory variables and the total seedling dry weight and P content as dependent variables. As P content was calculated as the product of total seedling dry weight and P concentration as previously

proposed [30,31], both dependent variables were significantly correlated (0.58) and were evaluated separately. The partitioning of the correlation coefficients revealed that root length had the lowest direct effect on total seedling dry weight (-4.831) and P content (-0.997), but contributed with a strong indirect effect through root surface area (5.816 and 1.565, respectively) (Table 2). Thus, the negative direct effect of root length was counterbalanced by the indirect effect via root surface area, probably due to the high positive correlation between these traits (r = 0.98). A similar pattern was observed for surface area of fine roots, which was also positively correlated with root surface area (r = 0.8). This trait exerted a minor negative direct influence on total seedling dry weight (-0.681) and P content (-0.117), which were mitigated by the positive indirect effect via root surface area (4.748 and 1.277, respectively). Root diameter also played a more important effect indirectly via root surface area on both variables, masking its direct contribution, which corroborates with the negative correlation between these root traits (r = -0.48). Therefore, root surface area contributed the highest direct effect on total seedling dry weight (5.935) and P content (1.597) and mediated an important proportion of the indirect effects of the other root traits on the dependent variables.

Thus, the path analysis clarified the direct and indirect importance of a greater root surface area, which is a combination of longer roots with smaller diameters, to improve biomass accumulation and P content in the seedlings under P deficiency. This root morphology also promoted an additional advantage for shoot over root development, confirming that the investment in root growth was beneficial to P acquisition as discussed by [18].

Two distinct QTL mapping strategies reveal complementary results

A linkage map was constructed using 292 markers that covered 1787.5 cM of the maize genome, with an average interval of 6.1 cM between adjacent markers. In addition to SSR and SNP markers, six *ZmPSTOL* candidate genes and three genes previously associated with root morphology (*Rth1*, *Bk2l3,* and *Rtcs*) were mapped to their predicted physical positions. Multiple interval mapping models for single (MIM) and multiple traits (MT-MIM) provided statistical evidence for 13 genomic regions harboring QTLs on all maize chromosomes, with the exception of chromosome 5 (Tables 3 and 4). The QTL regions were named using the trait initials if they were detected through single trait analysis or as “multi” if they were detected by multiple trait analysis, followed by their genomic position in bin (Figure 2).

MIM models for individual traits detected QTLs controlling root length, root diameter, surface area of fine roots, and root:shoot ratio. The proportion of the phenotypic variance

explained by each QTL (R^2) ranged from 6.84% (*qRD4.05*) to 15.12% (*qSA2_10.03*). The magnitude of QTL effects ranged from 0.270 standard deviations from the progeny mean (sd) to -0.393 sd (Table 3).

The MT-MIM analysis revealed the presence of ten QTLs with R^2 ranging from 2.04% (*qMulti8.02* for R:S) up to 15.17% (*qMulti10.03* for SA2). The highest additive effect was also observed for *qMulti10.03* for SA2 (-0.403 sd) (Table 4).

LOD estimates for the MT-MIM model were higher and the confidence intervals were narrower than those for the individual MIM models, suggesting superior statistical power of MT-MIM compared with the MIM models applied to each trait individually (Figure 2). Despite these differences, the MT-MIM and MIM models were coincident in revealing QTLs at bins 1.03, 1.07, 3.06, and 10.03, whereas QTLs at bins 1.06, 2.08, 3.04, 6.06, 8.02, and 9.04 were only revealed using the MT-MIM model. Conversely, QTLs at regions 4.05, 7.02, and 8.05 were only detected by MIM models. Therefore, combining the results of MIM and MT-MIM analyses conveyed the most accurate information regarding the genetic architecture of the traits under investigation in this particular study.

Using a simulation, [45] showed that when a QTL affects only a small subset of the traits, the MT-MIM model might have a lower power than MIM models to identify this QTL due to a greater genome-wide threshold for the MT-MIM model. This may be the reason why MT-MIM failed to identify QTLs at regions 4.05, 7.02, and 8.05. Although the MT-MIM LOD profile revealed peaks at these regions, the values were not statistically significant according to the score threshold employed.

The additive main effect of QTLs detected by MIM and MT-MIM had both positive and negative signs, confirming the contribution of favorable alleles coming from both parental lines for most of the traits analyzed (Tables 3 and 4). Additionally, five epistatic interactions were identified using the MT-MIM model, including some with magnitudes comparable to the main additive effects (Table 5). No epistatic effect was detected based on single trait analysis. Taken together, the additive and epistatic effects on the MT-MIM model explained between 23.41% and 35.54% of the phenotypic variance for each trait (Tables 4 and 5).

ZmPSTOL predicted proteins share a conserved serine/threonine kinase domain with OsPSTOL1

Using OsPSTOL1 [GenBank: BAK26566] as a query, six predicted proteins were selected on the maize genome, sharing more than 55% amino acid sequence identity. The genes encoding these proteins were predicted to be located on chromosomes 3, 4 and 8, and named according to their genetic position in bin (Table 6).

A phylogenetic analysis revealed that the six predicted ZmPSTOL proteins clustered together with rice PSTOL1 and two receptor-like kinases (SNC4 and PR5) from *Arabidopsis* (circled in Figure 3). This clade included serine/threonine receptor-like kinases belonging to the LRK10L-2 subfamily, as originally observed by [38]. In a detailed alignment of structural predictions, all maize PSTOL-like proteins shared conserved ATP-binding and serine/threonine protein kinase domains with OsPSTOL1 (Additional File 2. Figure S1). A distinct glycosyl hydrolase domain was predicted for ZmPSTOL4.05. The maize proteins ZmPSTOL4.05, ZmPSTOL8.02, ZmPSTOL8.05_1 and ZmPSTOL8.05_2 were classified as receptor-like kinases (RLKs), which are characterized by the presence of a transmembrane domain for signal perception and an intracellular kinase domain [46,47]. In contrast, the proteins ZmPSTOL3.04 and ZmPSTOL3.06 contained the intracellular kinase domain but lacked the transmembrane domain in a manner similar to OsPSTOL1 [38], and thus were classified as receptor-like cytoplasmic kinases (RLCKs) [47].

ZmPSTOL candidate genes have distinct expression patterns

The expression analyses revealed that *ZmPSTOL4.05*, *ZmPSTOL8.02* and *ZmPSTOL8.05_1* exhibited consistently higher expression in the roots of the P-inefficient genotype (L22) under low (2.5 µM) and high (250 µM) P conditions (Figure 4). In contrast, *ZmPSTOL3.06* was preferentially expressed in the roots of the P-efficient line L3, when exposed to both levels of P (Figure 4). *ZmPSTOL8.05-2* was primarily expressed in L22 independent of the tissue and P availability, whereas *ZmPSTOL3.04* showed a random expression pattern (Figure 4).

Discussion

Complex inheritance of root traits, seedling dry weight and P content in maize

Two distinct and powerful statistical models for QTL mapping (MIM and MT-MIM) were applied to dissect root morphology traits, total seedling dry weight and P content in tropical maize RILs. The QTLs identified using the MIM model explained from 6 to 15% of the total phenotypic variance for each trait, which was similar to the QTLs previously mapped for root traits and P efficiency indices in maize [25-27,29,31]. However, the proportion of phenotypic variance explained by all QTLs and their epistatic interactions using MT-MIM ranged from 23.41 to 35.54%, which was higher than in previous studies. To the best of our knowledge, the present study demonstrates the first QTL mapping of root morphology traits, seedling biomass and P content in maize using the MT-MIM model, confirming that complementary

information can be generated when this strategy is combined with single trait MIM analyses, as previously suggested by [45].

Indeed, the quantitative inheritance of these traits seems to be even more complex as indicated by the presence of epistatic interactions among QTLs, which showed effects of magnitudes comparable to those of main effects. These results are consistent with previous studies, demonstrating that epistatic interactions significantly contributed to root traits in nutrient solution [25-27,48] and for P-use efficiency indices under field conditions [30,49].

QTL mapping revealed that both parents contributed favorable alleles for most of the traits evaluated, which possibly leads to transgressive segregation. The occurrence of segregating progenies with extreme phenotypes, out of the parental range, has been detected in plants subjected to different abiotic stresses under field or nutrient solution [48,50,51]. Transgressive segregation was also observed in maize RILs derived from a cross between Mo17 and B73 for the length and number of lateral and seminal roots [25,27] and for root hair length [26]. More recently, L3 and L22 were shown to donate favorable alleles for P acquisition efficiency and P-use efficiency based on grain yield in a design III experiment under field conditions [49].

The QTL mapping results strongly reflected the phenotypic correlations among the target traits. The high correlation between root length and root surface area (r = 0.98) reflected the coincidence in position and sign of the QTLs detected based on MT-MIM (*qMulti1.07*, *3.04*, *8.02* and *10.03*; Table 4). Additionally, some significant QTLs for root diameter were either mapped to unique regions or with opposite signs compared with other root traits, corroborating the negative correlations of these traits. Moreover, the importance of the root surface area to seedling dry weight and P content was supported by the presence of multi-trait QTLs that significantly affected these traits (*qMulti1.07*, *8.02* and *10.03*; Table 4), which could be a result of determinant genes with pleiotropic effects or the presence of linked genes.

Coincidence of QTLs for root morphology in the seedling stage with QTLs for grain yield

Even though early root growth enhancement has not always led to superior yield performance in the field [52], a large number of QTL studies have indicated that some genomic regions consistently affect root morphology traits during the seedling stage and agronomic performance under different environments. These regions will be further discussed below.

Three genes previously associated with root morphology in maize (*Rth1*, *Bk2l3* and *Rtcs*) were mapped to chromosome 1, but did not overlap with currently mapped QTLs

(Figure 2). However, the three QTLs mapped for multiple traits on chromosome 1 were coincident with QTLs previously reported for early root traits and for yield components in the field. *qMulti1.03* associated with root diameter, total seedling dry weight and root:shoot ratio overlapped with the QTL influencing seminal root length and weight [48], primary axile root diameter [53], daily elongation rate of axile roots [54] and total length of second-order lateral roots [55] in nutrient solution. This region was coincident with QTLs for drought tolerance index [48], grain yield, kernel number and weight in low P soil [56]. This genomic region was also detected in a meta-analysis for low-P tolerance in maize as the consensus cQTL1-2 [33].

qMulti1.06 was detected based on the multi-trait MIM model and was significantly associated only with the total P content in the seedling (Table 4 and Figure 2). A major QTL at bin 1.06 was associated with root traits in nutrient solution, grain yield under well-watered and water-stressed conditions [48], and root-pulling resistance in adult plants [57]. Due to the consistency in the effects of this genomic region on root traits and grain yields in different studies, this QTL was named *root-yield-1.06*, and validated as constitutively affecting roots, agronomic features and grain yield under different water regimens [58]. A QTL for P utilization efficiency (*qPUTIL1*) based on grain yield under low P soil also overlapped in this region [49].

The third QTL region on chromosome 1 was mapped at bin 1.07 spanning from 214 to 223 Mbp (*qMulti1.07*), which was associated with root length, root surface area, surface area of fine roots, total seedling dry weight and root:shoot ratio. This region was coincident with a cluster of QTLs named Ax-2 that controlled the root numbers and lengths identified in a meta-analysis combining 15 QTL studies [52]. QTLs for grain yield and drought tolerance index were also mapped to this genomic region [48].

The association of root morphology QTLs in early stages of plant development with yield performance, including the validation of the *root-yield-1.06*, suggested that at least some of these genomic regions can be further used in marker-assisted selection to improve yield stability under drought and other mineral stresses in maize.

ZmPSTOL genes may be involved with root morphology and/or P content in maize

Another QTL mapped using MT-MIM (*qMulti8.02*) that was associated with root length, root surface area, root:shoot ratio and P content co-localized with *ZmPSTOL8.02* (Figure 2). This *ZmPSTOL* candidate gene was highly expressed in the roots of L22 (Figure 4), the donor line of the favorable QTL alleles for all traits mentioned above. In rice, overexpression of *OsPSTOL1* enhanced total root length and root surface area in transgenic seedlings in nutrient solution as well as grain yield of transgenic varieties cultivated in P-deficient soils

[38]. According to these authors, the larger root system contributed to a significant increase in the uptake of nutrients such as phosphorus, nitrogen and potassium in transgenic rice lines overexpressing *OsPSTOL1*. The sequence similarity and conserved domains of these protein kinases from rice and maize combined with the evidences shown here suggest that *ZmPSTOL8.02* could be one of the genes underlying *qMulti8.02*, sharing similar functions in root development and P acquisition efficiency in maize to *OsPSTOL1* in rice. Additionally, QTLs associated with the seminal root number in high P [27], shoot dry weight [32] and primary root length [54] overlapped with *qMulti8.02*, confirming that genes controlling root and shoot development in this genomic region are also expressed in other genetic backgrounds.

The other three *ZmPSTOL* genes co-localized with single trait QTLs for root length (*ZmPSTOL8.05_1*), root diameter (*ZmPSTOL4.05*), and root:shoot ratio (*ZmPSTOL3.06*). These genes were preferentially expressed in the roots of the donor line that contributes positive alleles for the respective QTLs under low- and high-P conditions.

ZmPSTOL8.05_1 was mapped to 117 cM on chromosome 8, flanking *qRL8.05*. Additionally, a weak LOD peak for P content was coincident with this candidate gene (Figure 2). The MIM model based on score as a significance threshold was not able to detect this QTL, but a minor effect QTL was detected using the Bayesian Information Content threshold with LOD 1.8 and explaining 6.5% of the total variance for the total P content in the seedling (data not shown). The LOD profile of this QTL was similar to the one as shown for P content in Figure 2. This genomic region also harbored QTLs explaining 5 to 6% of the phenotypic variance for P acquisition efficiency based on grain yield [49], indicating that this genomic region consistently contributed to P acquisition during the seedling and adult plant stages. Additionally, QTLs in this genomic region were mapped for root length and grain yield under field conditions [32], root length and root dry weight in nutrient solution [26,48,59]. The coincidence of QTLs for root traits and for grain yield from different studies associated with the superior expression of *ZmPSTOL8.05_1* in roots under low-P are highly compatible with the role of its putative homolog (*OsPSTOL1*) in rice.

ZmPSTOL3.06 was mapped to 169.6 cM on chromosome 3 within the confidence interval of *qRS3.06* for root:shoot ratio (Figure 2). This genomic region spans bin 3.06 that harbored QTLs for root traits in a meta-analysis using 15 QTL studies in nine maize mapping populations [52]. This candidate gene was highly expressed in roots cultivated with both P levels of the P-efficient line L3 (Figure 4), which contributed with positive alleles for the root:shoot ratio QTL. *ZmPSTOL3.06* had the lowest e-value with *OsPSTOL1* (Table 6) and its predicted protein lacks the transmembrane domain, similarly to OsPSTOL1 (Additional File 2: Figure S1). This combined information makes this predicted gene a strong candidate to *OsPSTOL1* homolog.

ZmPSTOL3.04 was the flanking marker of *qMulti3.04*, but mapped out of its confidence interval. However, a consensus QTL including six QTLs for traits associated with low-P tolerance was coincident with this region, and *ZmPSTOL3.04* was one of the candidate genes identified based on bioinformatics mining [33]. This gene was expressed in roots and shoots independent of the maize genotype and P availability, thereby exhibiting a different expression pattern compared to the other *ZmPSTOL* candidate genes (Figure 4).

The finding that *ZmPSTOL* genes were preferentially expressed in the roots of the lines that contributed the allelic enhancing root traits, seedling dry weight and P content suggested that at least *ZmPSTOL3.06*, *ZmPSTOL4.05*, *ZmPSTOL8.02* and *ZmPSTOL8.05_1* have a functional relationship with root morphology and/or with P acquisition in maize. These results are strongly supported by the role of *OsPSTOL1* in enhancing early root length, root surface area, and P uptake in rice [38]. More recently, multiple homologs of *OsPSTOL1* were identified in sorghum and shown to improve root morphology and grain yield in low-P soil [39]. Thus, it could be expected that genes encoding important adaptive traits would be shared among rice, sorghum and maize, such as the case for the major Al tolerance gene in sorghum (*SbMATE*) [60] that was found to be functionally conserved in maize (*ZmMATE1*) [61] and rice (*OsFRDL4*) [62].

Conclusions

Comprehensive QTL analyses revealed important regions associated with root traits, seedling dry weight and P content in maize under low-P availability. Using the multiple trait-multiple interval mapping model, these QTLs explained a larger extent of the phenotypic variance for the target traits compared with previous studies. Additionally, this study provides evidence that some *ZmPSTOL* genes are preferentially expressed in the roots and co-localize with QTLs for root traits and P content under low-P availability. Considering that *OsPSTOL1* improves grain yield under P starvation in rice, it is plausible that the *ZmPSTOL* genes are putative functional homologs of this gene in maize. However, further functional studies are required to validate the role of *ZmPSTOL*. Finally, these genomic regions may be target for molecular breeding aiming to improve P efficiency in maize.

Material and methods

Mapping population

A total of 145 maize recombinant inbred lines (RILs) derived from a cross between L3 (P-efficient) and L22 (P-inefficient) were analyzed in this study. L3 and L22 were developed

in the Embrapa Maize and Sorghum breeding program and contrast for P acquisition efficiency [43] and root morphology traits [44].

Quantitative analysis of root traits, seedling biomass accumulation and P content using a paper pouch system

Maize seeds were surface sterilized with 0.5% (v/v) sodium hypochlorite for 5 min and germinated in moistened germination paper rolls. After four days, uniform seedlings were transferred to moist blots in paper pouches after removing the endosperm to eliminate seed reserves [44]. A modified Magnavaca nutrient solution [63] containing 2.5 µM P was replaced every three days and the pH was maintained at 5.6. The containers were maintained in a growth chamber with a 12 hr photoperiod at 27/20°C day/night temperatures and 330 µmol m^{-2} s^{-1} of light intensity. After 13 days, root images were captured using a digital photography setup and analyzed using RootReader2D (http://www.plantmineralnutrition.net/rootreader.htm) and WinRHIZO (http://www.regent.qc.ca/assets/winrhizo_about.html) software. Four different root traits including total length (RL) (cm), average diameter (RD) (cm^3), total surface area (SA) (cm^2) and surface area of fine roots (SA2) (1.0 < d ≤ 2.0 mm) (cm^2) were measured according to [44].

Root and shoot tissues were dried separately at 65°C in a forced-air oven until a constant weight was obtained to determine the root:shoot dry weight ratio (R:S) and total seedling dry weight (TDW). For P analysis, root and shoot tissues were subjected to nitric perchloric acid digestion [64]. The total P content in the seedling (Pcont) was calculated as the sum of the P content in each seedling component, which was the product of the dry weight and the P concentration in the root and shoot.

The mapping population and parents were evaluated in randomized complete block design with four biological replicates, which were composed for three plants in bulk. Analysis of variance (ANOVA), correlations between pairs of traits and path analysis were performed using the GENES software [65]. The phenotypic correlations were calculated based on the mean values. Broad sense heritability was estimated as $h^2 = \hat{\sigma}_G^2 / \left(\hat{\sigma}_G^2 + \hat{\sigma}_E^2\right)$ with $\hat{\sigma}_G^2 = (MS_G - MS_E)/r$ and $\hat{\sigma}_E^2 = MS_E$, where $\hat{\sigma}_G^2$ and $\hat{\sigma}_E^2$ are the estimates of genetic and error variance, respectively; MS_G and MS_E are the genetic and error mean squares, respectively, and r is the number of replications.

Five variables were included in the path-coefficient analysis [66]. The independent variables x_i (i = 1, 2, …, 5) were the total root length, root average diameter, surface area and surface area of fine roots. The total seedling dry weight and phosphorus content were considered as dependent variables y_j (j = 1, 2) in two distinct path analysis studies. The

estimated path coefficient (P_{ij}) was considered as the direct effect of variable x_i on y_j. Indirect effects of x_i on y_j mediated by variable $x_{i'}$ were calculated by multiplying the correlation between x_i and $x_{i'}$ ($r_{ii'}$) by $P_{i'j}$. Root:shoot ratio was excluded from the analysis due to its contribution to both dependent variables.

Linkage map

DNA was isolated from young leaves using the CTAB method [67]. Initially, 60 polymorphic SSR markers were genotyped in the RIL population according to [68]. The primer sequences and genomic locations of the SSRs were obtained from the Maize Genetics and Genomics Database (http://www.maizegdb.org/locus.php). A total of 332 SNPs (Single Nucleotide Polymorphisms) were genotyped in the population using Kompetitive Allele-Specific PCR or the KASP^{TM} assay (LGC Genomics, Teddington, UK). Information for the sequences and the physical positions of the SNPs is available at Panzea (http://www.panzea.org/db/searches/webform/marker_search).

The markers were tested for the expected segregation ratio of 1:1 using chi-square statistics ($p < 0.05$) corrected for multiple tests based on Bonferroni's method. The linkage map was constructed using MapMaker/EXP 3.0 [69] considering a minimum LOD of 3.0 and a maximum frequency of recombination (r) of 0.4. The mapping function Kosambi [70] was used to convert recombination fractions into map distances.

QTL mapping

Phenotypic fitted values were obtained from the following statistical model adjusted for each single trait:

$$y_{ij} = \mu + Bi + G_j + \varepsilon_{ij}$$

where y_{ij} is the phenotypic observation from the i^{th} (i= 1, …, 4) replication on the j^{th} genotype (j = 1, …, 145); μ is the phenotypic average; G_j is the j^{th} genotype; and ε_{ij} is the residual associated with the y_{ij} observation. We tested the fit of two models by first assuming that the residuals were normally distributed with constant variance, $\varepsilon_{ij} \sim N(0, \sigma^2)$ and by second considering the heteroscedasticity, $\varepsilon_{ij} \sim N(0, \sigma_j^2)$. Both models were fitted using the *gls* function from the *nlme* R package [71] and compared using the *ANOVA* function from the same package. For each trait, the fitted phenotypic values were extracted from the model with the best fit.

Due to the large variability in the absolute values, all phenotypic fitted values were standardized to achieve unity as standard deviations and zero means as follows:

$$z_{it} = \frac{y_{it} - \mu_t}{\sigma_t},$$

where z_{it} is the standardized observation of trait t (t = 1, 2, …, 7) on subject i; y_{it} is the observation of trait t on subject i; μ_t is the average of trait t; and σ_t is the standard deviation of trait t.

For the joint QTL analysis, a multiple trait-multiple interval mapping (MT-MIM) model was evaluated as previously described [45]. The complete model was fitted using the following equation:

$$z_{ti} = \mu_t + \sum_{r=1}^{m} \beta_{tr} x_{ir} + \sum_{r<l}^{p} w_{trl} x_{ir} x_{il} + \varepsilon_{ti},$$

where z_{ti} is the standardized observation of trait t on subject i; μ_t is the intercept for trait t; the parameter β_{tr} has the genetic interpretation of the additive effect of QTL r on trait t (r = 1, 2, …, m QTLs included in the model); and the regressive variables x_{ir} represent the contrast coefficients codified according to the Cockerham model [72,73] (i.e., x_{ir} is 1 for the dominant and -1 for the recessive homozygous). The third component on the right side of the model refers to a subset of the p pairwise interactions among QTLs previously included in the model, where w_{trl} is the epistatic effect between QTL r and QTL l on trait t and the random error ε_{ti} was assumed to be independent and identically distributed according to a multivariate normal distribution, with a mean vector of zero and a positive definite symmetric variance-covariance matrix Σ_ε, i.e., $\varepsilon_{ti} = MVN(0, \Sigma_\varepsilon)$.

Multi-trait QTL mapping was initiated with a forward search for the main effect QTL using a grid of 1 cM in the genome and a 15% genome-wide significance level. After three rounds of QTL search, the positions of all QTLs in the model were re-estimated as along with all other parameters in the model. After the inclusion of the main effects in the model, the forward search for epistasis was performed by testing all pairwise interactions among QTLs already included in the model, employing a 5% genome-wide significance level. Only the epistatic effects that displayed at least one significant marginal effect were kept in the final model. The Haley-Knott regression [74] was used to estimate the model parameters, and the resampled score statistics [45,75] were employed to obtain the empirical genome-wide threshold for the QTL mapping analysis. Using this approach, a final model was selected to calculate the proportion of phenotypic variance explained by all QTLs as the ratio between the genotypic variance of the QTL effect to the phenotypic variance times 100 (coefficient of determination R^2), and the LOD profile along the chromosomes. The R^2 values were estimated using the fitted full model, including non-significant QTL effects. The QTL confidence intervals were obtained using the drop 1.5-LOD support interval method with approximately 0.95 confidence levels [76].

Multiple interval mapping (MIM) analysis was performed for each single trait [72,77] in a similar procedure to that performed for the joint analysis, considering $t = 1$. All QTL analyses were performed using R software (version 2.15.2) and a QTL mapping package named OneQTL that is under development by L. da Costa e Silva.

Identification of maize *PSTOL1* homologs

Using the OsPSTOL1 amino acid sequence [GenBank: BAK26566] we performed searches against the maize genome database (www.maizesequence.org) using BLASTp. Six predicted maize proteins with more than 55% sequence identity to rice PSTOL1 were selected and aligned using ClustalX software version 1.83 [78]; the alignment included OsPSTOL1 and the Arabidopsis protein kinases SNC4 [79] and PRK5 [80]. The phylogenetic tree was constructed based on the maximum likelihood method with 1000 bootstraps [81] using MEGA software [82]. The protein domains were identified using the CDART (Conservative Domain Architecture Retrieval Tool (http://www.ncbi.nlm.nih.gov/Structure/lexington/lexington.cgi).

Mapping candidate genes

Specific primers for the maize candidate genes were designed using Primer Blast (ncbi.nlm.nih.gov/tools/primer-blast/index.cgi). PCR reactions were performed using 30 ng of DNA, 0.2 mM of each dNTP, 2 mM of $MgCl_2$, 10 pmols of each primer, 5% (v/v) dimethyl sulfoxide (DMSO) and 1 U of Taq DNA polymerase (Invitrogen, Carlsbad, CA). The amplification profile included an initial step of 95ºC for 1 min, followed by 35 cycles of denaturing at 94ºC for 1 min, annealing at 58 to 60ºC for 30 sec, depending on the primers, and extension at 72ºC for 1 min. The amplification products were treated with ExoSAP-IT reagent (USB Corporation, Cleveland, OH) and sequenced using the BigDye Terminator v3.1 cycle sequencing kit on an ABI PRISM 3100 genetic analyser (Applied Biosystems, Foster City, CA) to identify polymorphisms between the parental lines.

Sequence-tagged site (STS) markers were developed to map genes previously associated with root morphology in maize. For the *roothairless* gene (*Rth1*) [22], a 22-bp indel was amplified using the primers 5'-TTGCCCACGGCTGGCAAGAG-3' and 5'-GGCTCTGTAGCACGCCCCTC - 3' and resolved on a silver-stained polyacrylamide gel [83]. The same strategy was used for the *brittle stalk-2-like protein 3* gene (*Bk2l3*) [23], which was revealed after the amplification of a 15-bp indel using the primer pair: 5'-GCTGGTTAGATCCCCCGCCCA-3' and 5'-GCACTGGAGCCACCGACACTG-3'. The *rootless concerning crown and seminal roots* gene (*Rtcs*) [24] was genotyped as a CAPS marker obtained after digestion of the amplification product of genomic DNA with the primers

5’-CGCGCCATAGCCCGCAGTAA-3' and 5’-GATTGGCACGGGCCGGTCAG-3' with the restriction enzyme *Aci*I and was visualized on a silver-stained polyacrylamide gel [83].

Cleaved amplified polymorphic sequence (CAPS) markers were developed for the other candidate genes. For *ZmPSTOL3.04* the PCR product amplified using the primers 5’-ACGGGGCTTGGAGGCACATG-3' and 5’-TGAGACCGCGTGGGGAAGGG-3' was digested with the restriction enzyme *Stu*I. The polymorphism for *ZmPSTOL8.02* was obtained after digestion with *Rsa*I of the genomic fragment amplified with the primers 5’-TGACTGGTGCCAGAGGTACGC-3' and 5’-TGCATACAAGGGACTGCTTCGGA-3'. CAPS markers were resolved on silver stained polyacrylamide gels [83]. The images were digitally captured using a Nikon digital camera.

ZmPSTOL3.06 was mapped based on the presence and absence of the amplification product using the primers 5’-AAGGGCGTCCAACCGCCTTG-3’ and 5’-TTGTTGGCCGGTCCGTTGGG-3’ on a 1% (w/v) agarose gel stained with ethidium bromide.

A G/A SNP was revealed after sequencing the amplified product obtained using the primer pair 5’-CCGCTACGCCTTGGTTGCCA-3' and 5’-CGCCGTAGTTAGCGGAGCCG-3' to map *ZmPSTOL4.05,* the primer pair 5’-AGCCTCCACGATGGCCGACA-3' and 5’-TGCATTTGTGTGACCTGGAA-3' to map *ZmPSTOL8.05_1*, and the primers 5’-TCCACGGCCGACAGGTAGCA-3' and 5’-GCTCAAGAGAACTCAGGGTGGC-3' to map *ZmPSTOL8.05_2*.

Gene expression analysis

The expression profiles of the candidate genes were assessed in the roots and shoots of the L3 and L22 genotypes harvested after 13 days in modified Magnavaca’s nutrient solution containing low (2.5 µM) and high (250 µM) P. Total RNA was extracted from a bulk of three plants using the RNeasy Plant Mini kit (Qiagen, Valencia, CA), and 1 µg of total RNA pretreated with DNase I was used for cDNA synthesis using the High Capacity cDNA Reverse Transcription kit (Applied Biosystems, Foster City, CA) according to the manufacturer's instructions. Gene expression was determined by quantitative PCR (qPCR-RT) using SYBR Green I and TaqMan assays in the ABI Prism 7500 Fast System (Applied Biosystems, Foster City, CA). Primers were designed for all target genes using Primer Express Software (Applied Biosystems, Foster City, CA), and 18S rRNA was used as an endogenous constitutive control (Additional File 3: Table S2). With the exception of *ZmPSTOL4.05*, two primers pair were designed for each gene to confirm their expression pattern. However, only the expression profile obtained with the primer pairs highlighted in bold in the Additional File 2 are shown. The relative gene expression was calculated using

the $2^{-\Delta\Delta CT}$ method [84], with three technical replicates and L3 roots under low-P conditions as a calibrator.

List of abbreviations

BIC: Bayesian Information Content, bp: base pair, CAPs: Cleaved Amplified Polymorphic sequence, CDART: Conservative Domain Architecture Retrieval Tool, CI: confidence interval, cM: centiMorgan, CTAB: cetyl trimethyl ammonium bromide, ICP inductively-coupled argon, indel: insertion/deletion, KASP: Kompetitive Allele-Specific PCR, LOD: likelihood of odds, Mbp: mega base pairs, MIM: single trait-multiple interval mapping, MT-MIM: multiple traits-multiple interval mapping, µM: micro Molar, mM: mili Molar, P cont: total phosphorus content in the seedling, PCR: polymerase chain reaction, P: phosphorus, *PSTOL1: Phosphorus-starvation tolerance 1*, qPCR: quantitative real-time PCR, QTL: quantitative trait loci, RD: average of root diameter, RIL recombinant inbred line, RL: total root length, R:S: root:shoot dry weight ratio, SA: total root surface area, SA2: surface area of fine roots, sd: standard deviations from the progeny mean, SNP: single nucleotide polymorphism, SSR: simple sequence repeats, STS: sequence-tagged site, TDW: total seedling dry weight, v/v: volume per volume, w/v: weight per volume.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

Conceived and designed experiments: CTG, SMS, JVM. Performed experiments: GCA, BFN, BH, UGPL. Analyzed the data: GCA, ACG, LCS, AAFG. Wrote and revised the paper: GCA, ACG, LCS, JVM, AAFG, SMS, CTG.

Acknowledgments

The authors would like to acknowledge funding from the Generation Challenge Programme (project G7010.03.01), the Embrapa Macroprogram, FAPEMIG and CNPq/Embrapa REPENSA. PhD fellowship for GCA was granted by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES/Embrapa) and BH was supported by Fundação de Amparo a Pesquisa do Estado de Minas Gerais (FAPEMIG). JVM and AAFG received fellowships from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). We are thankful to Dr. Sigrid Heuer (ACPFG, Australia) for sharing information about the rice *PSTOL1*.

References

1. Cordell D, Drangert J-O, White S: **The story of phosphorus: Global food security and food for thought.** *Global Environ Chang* 2009, **19**:292-305.

2. Dawson CJ, Hilton J: **Fertilizer availability in a resource-limited world: Production and recycling of nitrogen and phosphorus.** *Food Policy* 2011, **36**:S14-S22.

3. Vance C, Uhde-Stone C, Allan DL: **Phosphorus acquisition and use: critical adaptations by plants for securing a nonrenewable resource.** *New Phytol* 2003, **157**:423-447.
4. Shen J, Yuan L, Zhang J, Li H, Bai Z, Chen X, Zhang W, Zhang F: **Phosphorus dynamics: from soil to plant.** *Plant Physiol* 2011, **156**:997-1005.
5. Lynch JP, Brown KM: **New roots for agriculture: exploiting the root phenome.** *Phil Trans R Soc B* 2012, **367**:1598-1604.

6. Batten GD: **A review of phosphorus efficiency in wheat.** *Plant Soil* 1992, **146**:163-168.

7. Awika JM: **Major cereal grains production and use around the world**. In: *Advances in Cereal Science: Implications to Food Processing and Health Promotion*. Edited by Awika JM, Piironen V, Bean S. Washington: American Chemical Society; 2011:1-13.

8. Calderón-Vázquez C, Sawers RJH, Herrera-Estrella L: **Phosphate deprivation in maize: genetics and genomics.** *Plant Physiol* 2011, **156**:1067-1077.

9. Lambers H, Raven JA, Shaver GR, Smith SE: **Plant nutrient-acquisition strategies change with soil age.** *Trends Ecol Evol* 2008, **23**:95-103.
10. Parentoni SN, Souza Jr CL: **Phosphorus acquisition and internal utilization efficiency in tropical maize genotypes.** *Pesqui Agropecu Bras* 2008, **43**:893-901.
11. Raghothama KG: **Phosphate acquisition.** *Annu Rev Plant Phys* 1999, **50**:665-693.

12. Horst WJ, Kamh M, Jibrin JM, Chude VO: **Agronomic measures for increasing P availability to crops.** *Plant Soil* 2001, **237**:211-223.

13. Ramaekers L, Remans R, Rao IM, Blair MW, Vanderleyden J: **Strategies for improving phosphorus acquisition efficiency of crop plants.** *Field Crop Res* 2010, **117**:169-176.
14. Mollier A, Pellerin S: **Maize root system growth and development as influenced by phosphorus deficiency.** *J Exp Bot* 1999, **50**:487-497.

15. Hermans C, Hammond JP, White PJ, Verbruggen N: **How do plants respond to nutrient shortage by biomass allocation?** *Trends Plant Sci* 2006, **11**:610-617.
16. Silberbush M, Barber SA: **Sensitivity of simulated phosphorus uptake to parameters used by a mechanistic-mathematical model.** *Plant Soil* 1983, **100**: 93-100.
17. Lynch JP: **Roots of the second green revolution.** *Aust J Bot* 2007, **55**:493-512.
18. Zhu J, Lynch JP: **The contribution of lateral rooting to phosphorus acquisition efficiency in maize (*Zea mays*) seedlings.** *Func Plant Biol 2004, **31***:949-958.
19. Lynch JP, Brown KM: **Topsoil foraging – an architectural adaptation of plants to low phosphorus availability.** *Plant Soil* 2001, **237**:225-237.
20. York LM, Nord EA, Lynch JP: **Integration of root phenes for soil resource acquisition.** *Fron Plan Sci* 2013, **4**:1-15.
21. Hochholdinger F, Tuberosa R: **Genetic and genomic dissection of maize root development and architecture.** *Curr Opin Plant Biol* 2009, **12**:1-6.
22. Wen T-S, Hochholdinger F, Sauer M, Bruce W, Schnable PS: **The *roothairless1* gene of maize encodes a homolog of *sec3*, which is involved in polar exocytosis.** *Plant Physiol* 2005, **138**:1637-1643.
23. Brady S, Song S, Dhugga KS, Rafalski JA, Benfey PN: **Combining expression and comparative evolutionary analysis. The COBRA gene family.** *Plant Physiol* 2007, **143**:172-187.
24. Taramino G, Sauer M, Stauffer Jr JL, Multani D, Niu X, Sakai H, Hochholdinger F: **The maize (*Zea mays* L.) *RTCS* gene encodes a LOB domain protein that is a key regulator of embryonic seminal and post-embryonic shoot-borne root initiation.** *Plant J* 2007, **50**:649-659.
25. Zhu J, Kaeppler SM, Lynch JP: **Mapping of QTL controlling root hair length in maize (*Zea mays* L.) under phosphorus deficiency.** *Plant Soil* 2005a, **270**:299-310.
26. Zhu J, Kaeppler SM, Lynch JP: **Mapping of QTLs for lateral root branching and length in maize (*Zea mays* L.) under differential phosphorus supply.** *Theor Appl Genet* 2005b, **111**:688-695.
27. Zhu J, Mickelson SM, Kaeppler SM, Lynch JP: **Detection of quantitative trait loci for seminal root traits in maize (*Zea mays* L.) seedlings grown under differential phosphorus levels.** *Theor Appl Genet* 2006, **113**:1-10
28. Kaeppler SM, Parke JL, Mueller SM, Senior L, Stuber C, Tracy WF: **Variation among maize inbred lines and detection of quantitative trait loci for growth at low phosphorus and responsiveness to arbuscular mycorrhizal fungi.** *Crop Sci* 2000, **40**:358-364.

29. Chen J, Xu L, Cai Y, Xu J: **QTL mapping of phosphorus efficiency and relative biologic characteristics in maize (*Zea mays* L.) at two sites.** *Plant Soil* 2008, **313**:251-266.
30. Chen J, Xu Li, Cai Y, Xu J: **Identification of QTLs for phosphorus utilization efficiency.** *Euphytica* 2009, **167**:245-252.
31. Chen J, Xu L: **The candidate QTLs affecting phosphorus absorption efficiency.** *Front Agric China* 2011, **5**:456-462.
32. Cai H, Chen F, Mi G, Zhang F, Maurer HP, Liu W, Reif JC, Yuan L: **Mapping QTLs for root system architecture of maize (*Zea mays* L.) in the field at different developmental stages.** *Theor Appl Genet* 2012a, **125**:1313-1324.
33. Zhang H, Uddin MS, Zou C, Xie C, Xu Y, Li W-X: **Meta-analysis and candidate gene mining of low-phosphorus tolerance in maize.** *J Int Plant Biol* 2014, **56**:262-270.
34. Wissuwa M, Wegner J, Ae N, Yano M: **Substitution mapping of Pup1: a major QTL increasing phosphorus uptake of rice from a phosphorus-deficient soil.** *Theor Appl Genet* 2002, **105**:890-897.
35. Wissuwa M, Ae N: **Further characterization of two QTLs that increase phosphorus uptake of rice (*Oryza sativa* L.) under phosphorus deficiency.** *Plant Soil* 2001, **237**:275-286.
36. Chin JH, Lu X, Haefele SM, Gamuyao R, Ismail A, Wissuwa M, Heuer S: **Development and application of gene-based markers for the major rice QTL *Phosphorus uptake 1*.** *Theor Appl Genet* 2010, **120**:1073-1086.
37. Chin JH, Gamuyao R, Dalid C, Bustaman M, Prasetiyono J, Moeljopawiro S, Wissuwa M, Heuer S: **Developing rice with high yield under phosphorus deficiency: *Pup1* sequence to application.** *Plant Physiol* 2011, **156**:1202-1216.
38. Gamuyao R, Chin JH, Pariasca-Tanaka J, Pesaresi P, Catausan S, Dalid C, Slamet-Loedin I, Tecson-Mendoza EM, Wissuwa M, Heuer S: **The protein kinase Pstol1 from traditional rice confers tolerance of phosphorus deficiency.** *Nature* 2012, **488**:535-539.
39. Hufnagel B, de Sousa SM, Assis L, Guimaraes CT, Leiser W, Azevedo GC, Negri BF, Larson BG, Shaff JE, Pastina MM, Barros BA, Weltzien E, Rattunde HFW, Viana JH, Clark RT, Falcão A, Gazaffi R, Garcia AAF, Schaffert RE, Kochian LV, Magalhaes JV: **Duplicate and conquer: Multiple homologs of PHOSPHORUS-STARVATION TOLERANCE1 enhance phosphorus acquisition and sorghum performance on low-phosphorus soils.** *Plant Physiol* 2014, **166**:659-677.
40. Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, Minx P, Reily AD, Courtney L, Kruchowski SS, Tomlinson C, Strong C, Delehaunty K, Fronick C, Courtney B, Rock SM, Belter E, Du F, Kim K, Abbott

RM, Cotton M, Levy A, Marchetto P, Ochoa K, Jackson SM, Gillam B, et al: **The B73 maize genome: complexity, diversity, and dynamics.** *Science* 2009, **326**:1112-1115.

41. Parentoni SN, Magalhaes JV, Pacheco CAP, Santos MX, Abadie T, Gama EEG, Guimarães PEO, Meirelles WF, Lopes MA, Vasconcelos MJV, Paiva E: **Heterotic groups based on yield-specific combining ability data and phylogenetic relationship determined by RAPD markers for 28 tropical maize open pollinated varieties.** *Euphytica* 2001, **121**:197-208.
42. Lu Y, Yan J, Guimaraes CT, Taba S, Hao Z, Gao S, Chen S, Li J, Zhang S, Vivek BS, Magorokosho C, Mugo S, Makumbi D, Parentoni SN, Shah T, Rong T, Crouch JH, Xu Y: **Molecular characterization of global maize breeding germplasm based on genome-wide single nucleotide polymorphisms.** *Theor Appl Genet* 2009, **120**:93-115.
43. Parentoni SN, Souza Jr CL, Alves VMC, Gama EEG, Coelho AM, Oliveira AC, Guimarães PEO, Guimaraes CT, Vasconcelos MJV, Pacheco CAP, Meirelles WF, Magalhaes JV, Guimaraes LJM, Silva AR, Mendes FF, Schaffert RE: **Inheritance and breeding strategies for phosphorus efficiency in tropical maize (*Zea mays* L.).** *Maydica* 2010, **55**:1-15.
44. de Sousa SM, Clark RT, Mendes FF, de Oliveira AC, Vasconcelos MJV, Parentoni SN, Kochian LV, Guimaraes CT, Magalhaes JV: **A role for root morphology and related candidate genes in P acquisition efficiency in maize.** *Func Plant Biol* 2012, **39**:925-935.
45. da Costa e Silva L, Wang S, Zeng Z-B: **Multiple trait multiple interval mapping of quantitative trait loci from inbred line crosses.** *BMC Genet* 2012, **13**:67.
46. Shiu SH, Bleecker AB: **Receptor-like kinases from *Arabidopsis* form a monophyletic gene family related to animal receptor kinases.** *P Natl Acad Sci USA* 2001, **98**:10763-10768.
47. Vij S, Giri J, Dansana PK, Kapoor S, Tyagi AK: **The receptor-like cytoplasmic kinase (OsRLCK) gene family in rice : organization , phylogenetic relationship, and expression during development and stress.** *Mol Plant* 2008, **1**:732-750.
48. Tuberosa R, Sanguineti MC, Landi P, Giuliani MM, Salvi S, Conti S: **Identification of QTLs for root characteristics in maize grown in hydroponics.** *Plant Mol Biol* 2002, **48**:697-712.
49. Mendes FF, Guimarães LJM, Souza JC, Guimarães PEO, Magalhaes JV, Garcia AAF, Parentoni SN, Guimaraes CT: **Genetic architecture of phosphorus use efficiency in tropical maize cultivated in a low-P soil.** *Crop Sci* 2014, **54**:1-9.

50. Li M, Guo X, Zhang M, Wang X, Zhang G, Tian Y, Wang Z: **Mapping QTLs for grain yield and yield components under high and low phosphorus treatments in maize (*Zea mays* L.).** *Plant Sci* 2010, **178**:454-462.
51. Ruta N, Liedgens M, Fracheboud Y, Stamp P, Hund A: **QTLs for the elongation of axile and lateral roots of maize in response to low water potential.** *Theor Appl Genet* 2010, **120**:621-631.
52. Hund A, Fracheboud Y, Soldati A, Frascaroli E, Salvi S, Stamp P: **QTL controlling root and shoot traits of maize seedlings under cold stress.** *Theor Appl Genet* 2004, **109**:618-629.
53. Hund A, Reimer R, Messmer R: **A consensus map of QTLs controlling the root length of maize.** *Plant Soil* 2011, **344**:143-158.
54. Trachsel S, Messmer R, Stamp P, Hund A: **Mapping of QTLs for lateral and axile root growth of tropical maize.** *Theor Appl Genet* 2009, **119**:1413-1424.
55. Burton AL, Johnson JM, Foerster JM, Hirsch CN, Buell CR., Hanlon MT, Kaeppler SM, Brown KM, Lynch JP**: QTL mapping and phenotypic variation for root architectural traits in maize (*Zea mays* L.).** *Theor Appl Genet* 2014, **127**:2293-2311.
56. Cai H, Chu Q, Gu R, Yuan L, Liu J, Zhang X, Chen F, Mi G, Zhang, F: **Identification of QTLs for plant height, ear height and grain yield in maize (*Zea mays* L.) in response to nitrogen and phosphorus supply.** *Plant Breeding* 2012b, **131**:502-510.
57. Landi P, Sanguineti MC, Darrah LL, Giuhani MM, Salvi S, Conti S, Tuberosa R: **Detection of QTLs for vertical root pulling resistance in maize and overlap with QTLs for root traits in hydroponics and for grain yield under different water regimes.** *Maydica* 2002, **47**:233-243.
58. Landi P, Giuliani S, Salvi S, Ferri M, Tuberosa R, Sanguineti MC: **Characterization of root-yield-1.06, a major constitutive QTL for root and agronomic traits in maize across water regimes.** *J Exp Bot* 2010, **61**:3553-3562.
59. Liu J, Li J, Chen F, Zhang F, Ren T, Zhuang Z, Mi G: **Mapping QTLs for root traits under different nitrate levels at the seedling stage in maize (*Zea mays* L.).** *Plant Soil* 2008, **305**:253-265.
60. Magalhaes JV, Liu J, Guimaraes CT, Lana UGP, Alves VMC, Wang YH, Schaffert R, Hoekenga OA, Pineros MA, Shaff JE, Klein PE, Carneiro NP, Coelho CM, Trick HN, Kochian, LV: **A gene in the multidrug and toxic compound extrusion (MATE) family confers aluminum tolerance in sorghum.** *Nat Genet* 2007, **39**:1156-1161.

61. Maron LG, Piñeros MA, Guimaraes CT, Magalhaes JV, Pleiman JK, Mao C, Shaff J, Belicuas SNJ, Kochian LV: **Two functionally distinct members of the MATE (multi-drug and toxic compound extrusion) family of transporters potentially underlie two major aluminum tolerance QTLs in maize.** *Plant J* 2010, **61**:728-740.

62. Yokosho K, Yamaji N, Ma JF: **An Al-inducible MATE gene is involved in external detoxification of Al in rice.** *Plant J* 2011, **68**:1061-1069.

63. Magnavaca R, Gardner CO, Clark RB: **Evaluation of maize inbred lines for aluminum tolerance in nutrient solution.** In *Genetic Aspects of Plant Mineral Nutrition*. Edited by Gabelman HW, Longhman BC. Dordrecht: Martinus Nijhoff Publishers; 1987:255-265.
64. da Silva FC: *Manual de Análises Químicas de Solos, Plantas e Fertilizantes*. Brasília: Embrapa Informação Tecnológica; 2009.
65. Cruz CD: *Genes Software: biometrics*. Viçosa: UFV; 2006.
66. Wright S: **Correlation and causation.** J Agric Res 1921, **20**:557-585.
67. Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW: **Ribosomal DNA spacer length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics.** *P Natl Acad Sci USA* 1984, **81**:8014-8018.
68. Ninamango-Cárdenas FE, Guimaraes CT, Martins PR, Parentoni SN, Carneiro NP, Lopes MA, Moro JR, Paiva E: **Mapping QTLs for aluminum tolerance in maize.** *Euphytica* 2003, **130**:223-232.
69. Lander ES, Green P, Abrahamon J, Barlow A, Daly M, Lincoln SE, Newburg L: **MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations.** *Genomics* 1987, **1**:174-181.
70. Kosambi DD: **The estimation of map distances from recombination values.** *Ann Eugenic* 1944, **12**:172-175.
71. Pinheiro J, Bates D, Debroy S, R-CORE-TEAM: **nlme: Linear and nonlinear mixed effects models.** 2014. [http://cran.r-project.org/package=nlme].
72. Kao CH, Zeng ZB, Teasdale RD: **Multiple interval mapping for quantitative trait loci.** *Genetics* 1999, **152**:1203-1216.
73. Kao CH, Zeng ZB: **Modelling epistasis of quantitative trait loci using Cockerham's model.** *Genetics* 2002, **160**:1243-1261.
74. Haley CS, Knott SA: **A simple regression method for mapping quantitative trait loci in line crosses using flanking markers.** *Heredity* 1992, **69**:315-324.

75. Zou F, Fine JP, Hu J, Lin DY: **An efficient resampling method for assessing genome-wide statistical significance in mapping quantitative trait loci.** *Genetics* 2004, **168**: 2307-2316.
76. Lynch M, Walsh B: *Genetics and Analysis of Quantitative Traits*. Sunderland: Sinauer Assoc; 1998.
77. Zeng ZB, Kao CH, Basten C: **Estimating the genetic architecture of quantitative traits.** *Genet Res* 1999, **74**:279-289.
78. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG: **The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools.** *Nucleic Acids Res* 1997, **25**:4876–4882.
79. Bi D, Cheng YT, Li X, Zhang Y: **Activation of plant immune responses by a gain-of-function mutation in an atypical receptor-like kinase.** *Plant Physiol* 2010, **153**:1771-1779.
80. Wang X, Zafian P, Choudhary M, Lawton M: **The PR5K receptor protein kinase from Arabidopsis thaliana is structurally related to a family of plant defense proteins.** *P Natl Acad Sci USA* 1996, **93**:2598-2602.
81. Jones DT, Taylor WR, Thornton JM: **The rapid generation of mutation data matrices from protein sequences.** *Comput Appl Biosci* 1992, **8**:275-282.
82. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S: **MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods.** *Mol Biol Evol* 2011, **28**:2731-2739.
83. Creste S, Neto AT, Figueira A: **Detection of single sequence repeat polymorphisms in denaturing polyacrylamide sequencing gels by silver staining.** *Plant Mol Biol Rep* 2001, **19**:299-306.
84. Schmittgen TD, Livak KJ: **Analyzing real-time PCR data by the comparative C_T method.** *Nat Protoc* 2008, **3**:1101-1108

FIGURES

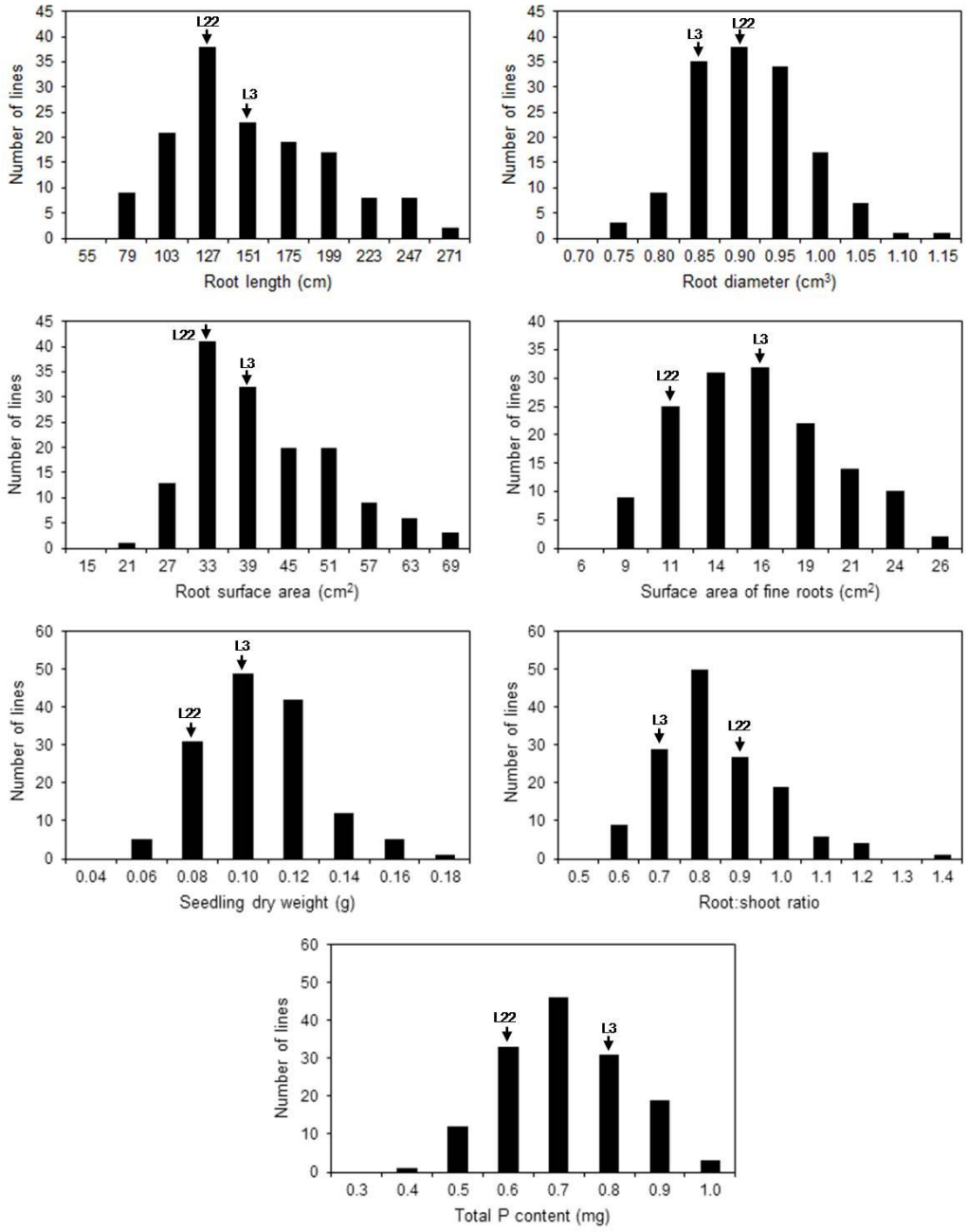


Figure 1 Frequency distribution of traits evaluated in 145 maize RILs grown in low-P (2.5 µM). The P-efficient (L3) and the P-inefficient (L22) parental lines are indicated by arrows.

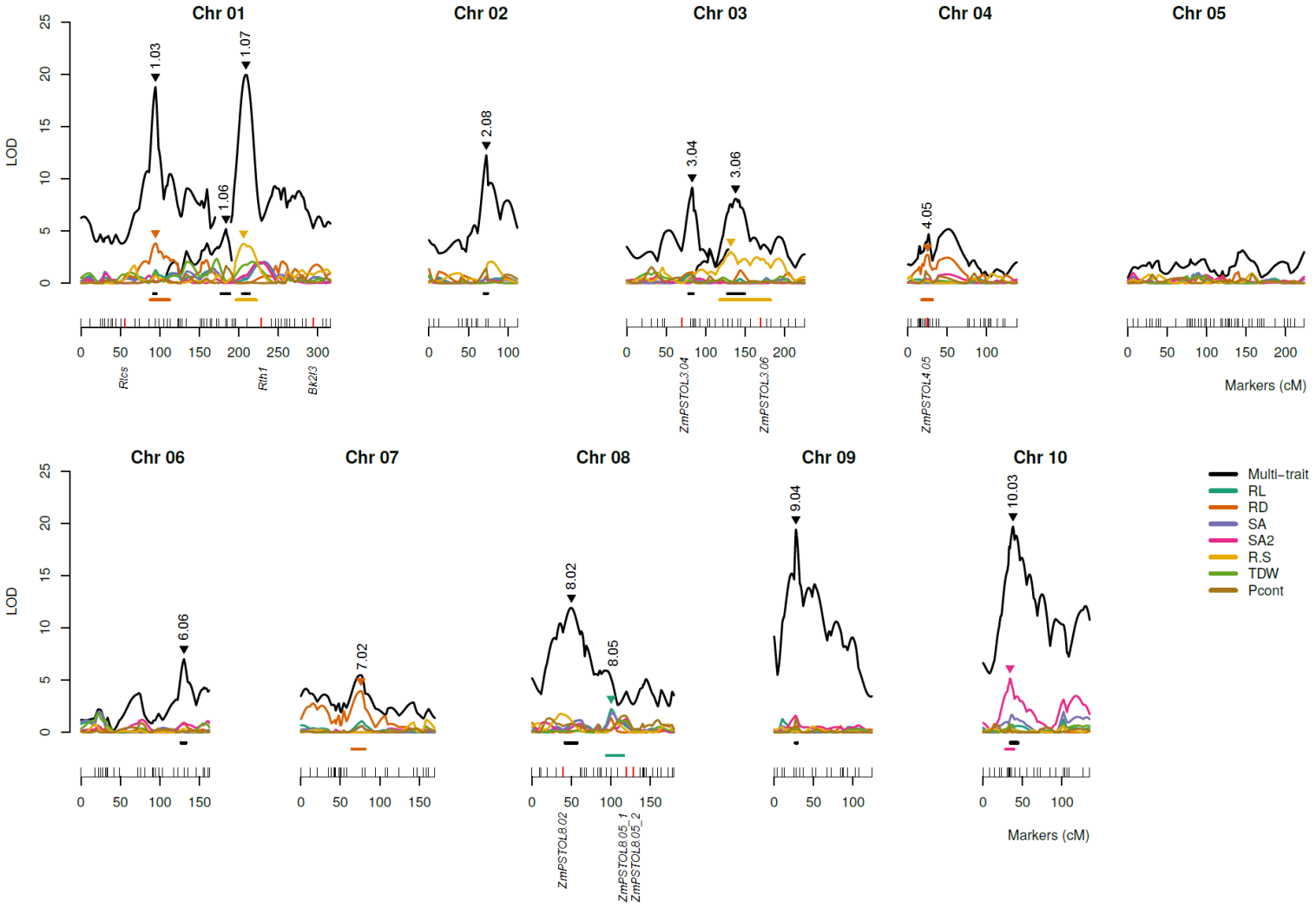

Chr 01
Chr 02
Chr 03
Chr 04
Chr 05
Chr 06
Chr 07
Chr 08
Chr 09
Chr 10
LOD
1.03
1.06
1.07
2.08
3.04
3.06
4.05
6.06
7.02
8.02
8.05
9.04
10.03
Rtcs
Rth1
Bk2i3
ZmPSTOL3.04
ZmPSTOL3.06
ZmPSTOL4.05
ZmPSTOL8.02
ZmPSTOL8.05_1
ZmPSTOL8.05_2
Markers (cM)
Multi−trait
RL
RD
SA
SA2
R.S
TDW
Pcont

Figure 2 QTLs identified for root traits, seedling dry weight and P content using single and multiple traits MIM analyses. The markers are represented as vertical traces along the horizontal lines, which represent the chromosomes and are numbered in centiMorgans. The candidate genes are depicted below the red vertical traces. QTL profiles for single trait MIM are shown as colored lines according to the legend for root length (RL), root average diameter (RD), root surface area (SA), surface area of fine roots (SA2), root:shoot ratio (R:S), total seedling dry weight (TDW) and total P content (Pcont). Multi trait QTL profiles are shown as black line. The QTL peaks are depicted with an inverted triangle colored according to the legend followed by the bin. The confidence interval (95%) for each QTL is represented by a horizontal line above the chromosomes colored according to the legend.

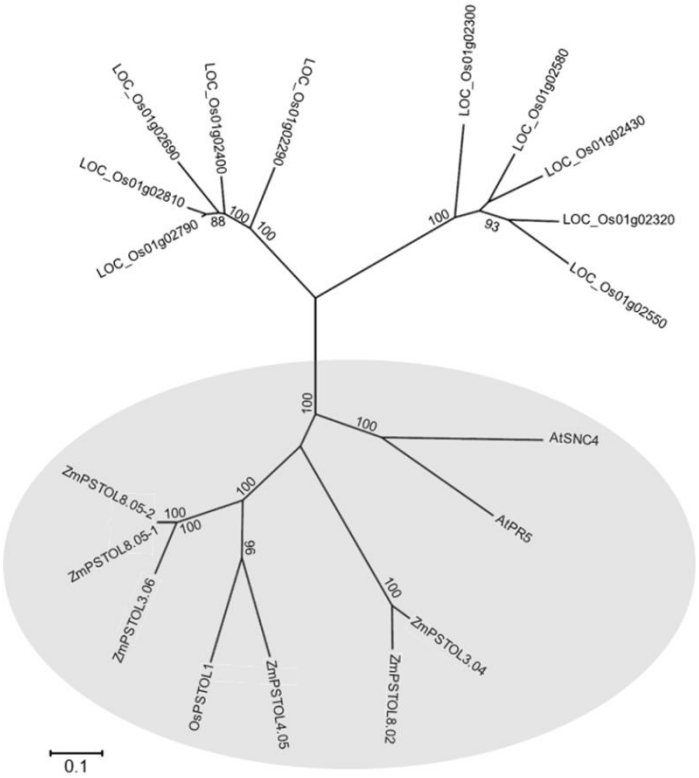


Figure 3 Phylogenetic tree of predicted serine/theronine receptor-like kinases from maize, rice and Arabidopsis thaliana. The rice PSTOL1, the six maize proteins sharing more than 55% sequence identity to OsPSTOL1, PR5K and SNC4 from Arabidopsis thaliana were grouped separately from other rice kinases, and are highlighted. Numbers on branches are bootstrap values for the percentage of coincidence (%) inferred from 1,000 replicates. Only percentage values higher than 80% are shown.

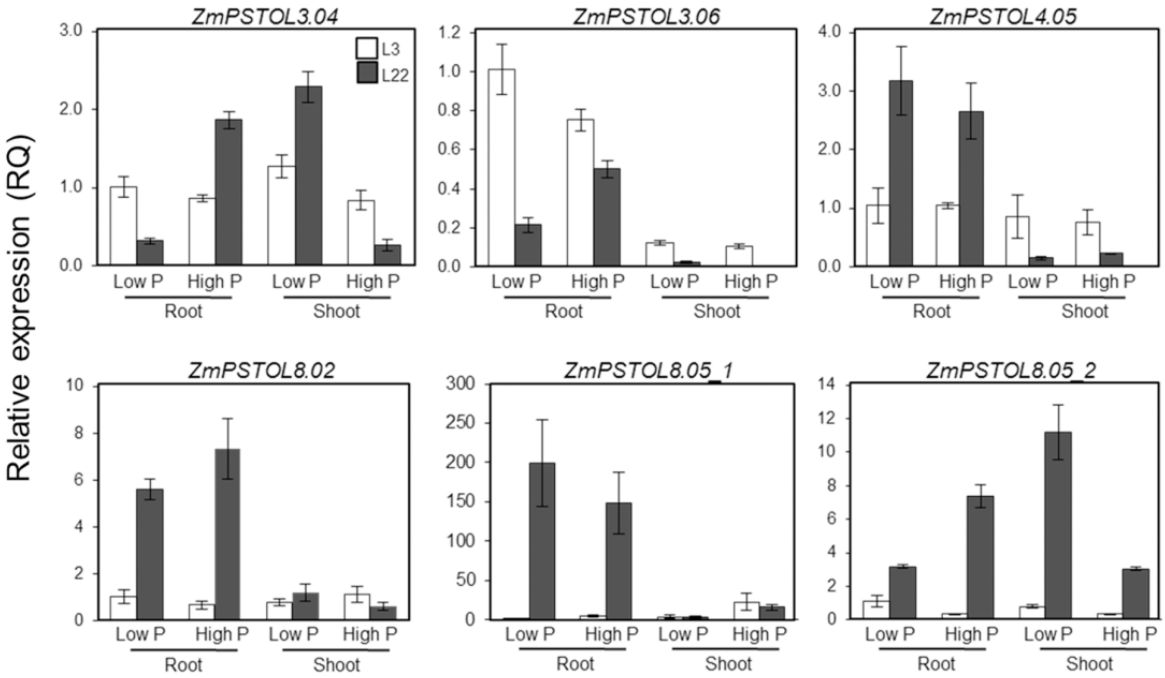


Figure 4 Expression profiles of the ZmPSTOL genes. The expression of the maize candidate genes are presented as relative gene expression (RQ) evaluated in roots and shoots of maize seedlings of the two parental lines L3 (P-efficient) and L22 (P-inefficient) grown grown under two levels of P (2.5 and 250 µM) after 13 days of treatment. Error bars indicate the standard errors of three technical replicates composed of three seedlings each.

TABLES

Table 1 Phenotypic correlation coefficients (r) among traits evaluated in the RILs under low-P condition in nutrient solution

Traits	RL	SA2	RD	R:S	TDW	Pcont
SA	0.98**	0.80**	-0.48**	-0.38**	0.86**	0.39**
RL		0.69**	-0.62**	-0.38**	0.79**	0.31**
SA2			0.03	-0.29**	0.77**	0.48**
RD				0.26**	-0.25**	0.14
R:S					-0.43**	-0.02
TDW						0.58**

Traits: root surface area (SA), root length (RL), surface area of fine roots (SA2), root diameter (RD), root:shoot dry weight ratio (R:S), total seedling dry weight (TDW), and total P content (Pcont).
Correlation coefficients followed by ** are significant at $p < 0.01$.

Table 2 Path analysis showing the partitioning of the phenotypic correlations into direct and indirect effects of root traits on total seedling dry weight and P content

Independent Variables		Dependent Variables	
		Total Dry Weight	**Total P Content**
Root surface area	**Direct effect**	**5.935**	**1.597**
	Indirect effect via root length	-4.734	-0.977
	Indirect effect via surface area of fine roots	-0.544	-0.094
	Indirect effect via root diameter	0.203	-0.135
	Phenotypic correlation (*r*)	0.86	0.39
Root length	**Direct effect**	**-4.831**	**-0.997**
	Indirect effect via surface area	5.816	1.565
	Indirect effect via surface area of fine roots	-0.462	-0.081
	Indirect effect via root diameter	0.267	-0.177
	Phenotypic correlation (*r*)	0.79	0.31
Surface area of fine roots	**Direct effect**	**-0.681**	**-0.117**
	Indirect effect via root length	-3.284	-0.678
	Indirect effect via surface area	4.748	1.277
	Indirect effect via root diameter	-0.012	0.008
	Phenotypic correlation (*r*)	0.77	0.48
Root diameter	**Direct effect**	**-0.424**	**0.281**
	Indirect effect via root length	3.043	0.628
	Indirect effect via surface area	-2.848	-0.766
	Indirect effect via surface area of fine roots	-0.020	-0.003
	Phenotypic correlation (*r*)	-0.25	0.14
	Coefficient of determination	0.869	0.295

Table 3 Quantitative trait loci (QTLs) identified using single trait-multiple interval mapping analysis for root length (RL), root diameter (RD), surface area of fine roots (SA2) and root:shoot ratio (R:S) under low-P

Trait	QTL	Bin	cM	Marker Position (Mbp)	LOD	Flanking Markers Position (Mbp)		R^2 (%)	Effect	R^2_T (%)
RL	*qRL8.05*	8.05	100.4	PZA00766_1 133.8	2.24	PHM934_19 116.8	*ZmPSTOL8.05_1* 152.0	6.87	-0.271**	6.87
RD	*qRD1.03*	1.03	94.4	umc1073 32.9	3.80	bnlg1083 27.5	PZA03742_1 44.5	9.60	0.307***	25.64
	qRD4.05	4.05	25.0	*ZmPSTOL4.05* 39.8	2.74	PHM15427_11 33.9	PHM14618_11 180.3	6.84	0.270***	
	qRD7.02	7.02	76.0	PZA01690_7 123.1	3.95	PZA01933_3 98.1	PZA01946_7 123.6	10.01	-0.331***	
SA2	*qSA2_10.03*	10.03	34.2	PHM2770_19 72.6	5.16	PHM1155_14 62.1	PZA01877_2 77.5	15.12	-0.393***	15.12
R:S	*qRS1.07*	1.07	206.0	PHM12693_8 223.5	3.76	PZA01963_15 203.7	PZA03301_2 240.6	10.85	0.377***	16.47
	qRS3.06	3.06	132.0	PZA02212_1 174.5	3.00	PZA00186_4 165.8	PZA01154_1 216.0	8.53	0.310***	

QTLs are named using the trait initials followed by their genomic position in bin
cM and Mbp indicate the marker position in centiMorgans and in mega base pairs at maximum LOD value
Flanking markers are based on -1.5 LOD support interval
R^2: Ratio of the genotypic variance of the QTL effect to the phenotypic variance, times 100
Effects measured as standard deviation from the progeny mean: Positive values indicate that L3 carries the allele for an increase in the trait, and negative values indicate that L22 contributes the allele for an increase in the trait. Effect significances based on *p-values* estimated via score statistics resampling ($^{+}p< 0.1$, $*p< 0.05$, $**p< 0.01$, $***p< 0.001$)
R^2_T: genotypic variance of the full model

Table 4 Quantitative trait loci (QTLs) identified using multiple traits-multiple interval mapping analysis for root length (RL), root diameter (RD), root surface area (SA), surface area of fine roots (SA2), root:shoot ratio (R:S), total seedling dry weight (TDW) and total P content (Pcont)

QTL	Bin	cM	Marker Position (Mbp)	LOD	Flanking Markers Position (Mbp)		Main Effect R^2 (%)						
							RL	RD	SA	SA2	R:S	TDW	Pcont
qMulti1.03	1.03	94.4	umc1073 32.9	18.49	bnlg1083 27.5	PZA03742_1 44.5	-0.109 1.18	0.163* 2.64	-0.075 0.56	0.110 1.19	-0.168* 2.82	-0.147* 2.14	0.026 0.06
qMulti1.06	1.06	183.8	PZA00619_3 195.4	5.07	bnlg1598 187.8	umc1335 196.9	0.093 0.85	-0.028 0.08	-0.084 0.70	0.078 0.60	-0.054 0.27	-0.004 0.00	0.210** 4.33
qMulti1.07	1.07	209.0	PHM114614_22 205.6	20.01	PZA01963_15 203.7	PHM12693_8 223.5	-0.185* 3.03	0.010 0.01	-0.231** 4.73	-0.224** 4.41	0.373*** 12.28	-0.299*** 7.89	-0.138^{+} 1.67
qMulti2.08	2.08	72.7	PZA01885_2 206.9	16.10	PZA02077_1 206.5	PZA01885_2 206.9	0.116 1.28	-0.109 1.13	0.083 0.65	-0.041 0.02	-0.111 1.16	-0.049 0.23	-0.262*** 6.50
qMulti3.04	3.04	83.0	PZA00297_2 39.9	7.90	*ZmPSTOL3.04* 20.2	PHM5502_31 67.2	0.217** 4.28	-0.216** 4.25	0.192* 3.34	0.119^{+} 1.29	0.124^{+} 1.40	-0.023 0.05	-0.059 0.31
qMulti3.06	3.06	138.0	PZA01962 178.2	8.25	PZA02212_1 174.5	PZA03735_1 180.5	-0.018 0.03	-0.097 0.81	-0.049 0.20	-0.047 0.19	0.306*** 8.07	-0.077 0.51	-0.025 0.05
qMulti6.06	6.06	130.5	PHM16607_11 160.2	6.88	PHM597_18 157.9	PZB01569_7 160.7	-0.027 0.07	0.108 1.11	0.010 0.01	0.178* 3.01	-0.064 0.39	0.111 1.17	0.045 0.02
qMulti8.02	8.02	48.0	*ZmPSTOL8.02* 13.3	15.76	*ZmPSTOL8.02* 13.3	PHM1978_111 21.8	-0.252** 4.14	0.164 1.76	-0.243** 3.83	-0.184* 2.20	-0.177* 2.04	-0.095 0.59	-0.239** 3.71
qMulti9.04	9.04	27.5	PHM13183_12 104.7	20.30	PZA0225_8 104.5	PZB01358_1 106.8	-0.022 0.04	-0.135^{+} 1.71	-0.079 0.58	-0.286*** 7.67	0.083 0.65	-0.164^{+} 2.53	-0.148^{+} 2.05
qMulti10.03	10.03	38.0	PHM1155_14 62.1	19.96	PHM1812_32 47.7	PZA01877_2 77.5	-0.180* 3.03	-0.106 1.06	-0.240** 5.38	-0.403*** 15.17	-0.047 0.21	-0.171* 2.74	-0.013 0.02
						R^2_T	**26.17**	**35.54**	**24.51**	**33.53**	**34.04**	**23.41**	**27.28**

QTLs are named using the “multi”, indicating that were detected using MT-MIM, followed by their genomic position in bin
cM and Mbp indicate the marker position in centiMorgans and in mega base pairs at maximum LOD value

Flanking markers are based on -1.5 LOD support interval

Effect measured as standard deviation from the progeny mean: Positive values indicate that L3 carries the allele for an increase in the trait, and negative values indicate that L22 contributes the allele for an increase in the trait. Effect significances based on *p-values* estimated via score statistics resampling ($^{+}p< 0.1$, $*p< 0.05$, $**p< 0.01$, $***p< 0.001$)

R^2: Ratio of the genotypic variance of the QTL effect to the phenotypic variance, times 100

R^2_T: genotypic variance of the full model (including epistasis shown in Table 4).

Table 5 Epistatic interactions for root morphology traits, total seedling dry weight and P content evaluated in low-P conditions

Interactions	Interaction Effect R^2 (%)						
	RL	**RD**	**SA**	**SA2**	**R:S**	**TDW**	**Pcont**
qMulti1.03* X *qMulti1.07	-0.060 0.33	0.067 0.41	-0.078 0.56	-0.183* 3.07	0.054 0.27	-0.053 0.25	-0.005 0.03
qMulti1.03* X *qMulti10.03	-0.009 0.01	-0.029 0.08	-0.029 0.08	-0.141^{+} 1.85	0.148* 2.04	-0.123 1.41	-0.235** 5.12
qMulti1.07* X *qMulti9.04	0.324*** 9.02	-0.358*** 11.00	0.280** 6.76	0.127 1.38	-0.108 1.00	0.196* 3.31	0.067 0.39
qMulti2.08* X *qMulti10.03	-0.071 0.45	0.179** 2.86	-0.013 0.02	0.115 1.18	0.166* 2.47	0.028 0.07	0.077 0.53
qMulti8.02* X *qMulti9.04	0.231** 3.24	-0.150 1.37	0.254* 3.93	0.188^{+} 2.16	-0.221* 2.99	0.307** 5.74	0.265** 4.29

Effects measured as standard deviation from the progeny mean; Positive values indicate that L3 carries the allele for an increase in the trait, and negative values indicate that L22 contributes the allele for an increase in the trait. Interaction effect significances based on *p*-values were estimated via score statistics resampling (^{+}p< 0.1, *p< 0.05, ** p< 0.01, ***p< 0.001)

R^2: Ratio of the genotypic variance of the QTL effect to the phenotypic variance, times 100.

Traits: root length (RL), root average diameter (RD), root surface area (SA), surface area of fine roots (SA2), total seedling dry weight (TDW), root:shoot ratio (R:S) and total P content (Pcont).

Table 6 Maize candidate genes sharing more than 55% amino acid sequence identity to OsPSTOL1

Predicted Gene	Gene ID	Physical position (bp)	Identity (%)	Coverage (%)	E-value
GRMZM2G412760	*ZmPSTOL3.04*	Chr3: 20,172,140	55	99	5.1e-104
GRMZM2G448672	*ZmPSTOL3.06*	Chr3: 206,918,421	66	97	4.7e-186
AC193632.2_FG002	*ZmPSTOL4.05*	Chr4: 39,792,602	69	95	2.0e-105
GRMZM2G172396	*ZmPSTOL8.02*	Chr8: 13,267,001	55	99	9.6e-123
GRMZM2G451147	*ZmPSTOL8.05_1*	Chr8: 152,043,859	70	97	3.4e-131
GRMZM2G164612	*ZmPSTOL8.05_2*	Chr8: 152,100,275	70	97	2.3e-127

ADDITIONAL FILES

Additional File 1: Table S1 Trait means, coefficients of variation (CV), genetic variances (σ_G^2), environmental variances (σ_E^2) and heritability estimates (h^2) in 145 RILs derived from a cross between maize lines L3 and L22

Trait	Mean	CV(%)	σ_G^2	σ_E^2	h^2
Root length (cm)	142.06	26.2	1683.2	353.4	0.82
Root diameter (cm^3)	0.88	7.4	0.004	0.001	0.79
Root surface area (cm^2)	38.48	23.4	91.6	20.6	0.81
Surface area of fine roots (cm^2)	14.50	24.3	13.5	3.2	0.81
Total seedling dry weight (g)	0.09	20.2	0.004	0.0001	0.81
Root:shoot ratio	0.79	20.5	0.012	0.006	0.65
Total P content (mg)	0.65	17.4	0.012	0.003	0.75

```
OsPSTOL1            M-----------------------------------------------------------
ZmPSTOL3.04         M---------------ARLEPPLTRVV---------------------------------
ZmPSTOL3.06         M-----------------------------------------------------------
ZmPSTOL4.05         MACNRRRKLLCACLLLACAAPAAAAAVNISVYWGQNSNEGSLGQTCSSGRYALVAMAFLS
ZmPSTOL8.02         M---------------DMIVQPLFCTVDCIVH--------------------------IV
ZmPSTOL8.05_1       MAMG--------------------------------------------------------
ZmPSTOL8.05_2       MYVK------------------LEEVVHTEV---------------------------E
                    *

OsPSTOL1            -----------------------------------------------------------
ZmPSTOL3.04         -----------------------------------------------------------
ZmPSTOL3.06         -----------------------------------------------------------
ZmPSTOL4.05         TFGSGQTPVLNLAGHCDPASGGCTALAADIAACQARGVRVLLSIGGGAGSYNLSSASDAE
ZmPSTOL8.02         NFEKGN------------------------------------------------------
ZmPSTOL8.05_1       -----------------------------------------------------------
ZmPSTOL8.05_2       HMQEGHTEVL--------------------------------------------------

OsPSTOL1            -----------------------------------------------------------
ZmPSTOL3.04         -----------------------------------------------------------
ZmPSTOL3.06         -----------------------------------------------------------
ZmPSTOL4.05         SCPYPDASLGAALATGLFDHVWVQFYNNPGCEYQQKDGDGVANLAASWKAWTQSLPSSAS
ZmPSTOL8.02         -----------------------------------------------------------
ZmPSTOL8.05_1       ----------------------LDF------------------------------------
ZmPSTOL8.05_2       -------------------RKKFDFH-----------------------------------

OsPSTOL1            -----------------------------------------------------------
ZmPSTOL3.04         -------------------------------------------------TGLGG----
ZmPSTOL3.06         -----------------------------------------------------------
ZmPSTOL4.05         VFLGLPASPAAAGSGYVPPDDLVSRVLPAVSGSANYGGLMLWNRY--YDNSTGYSARILS
ZmPSTOL8.02         ---------------------------------------------EIGDMMGTKGNGRK
ZmPSTOL8.05_1       ----------------------------SHGNSSIKGWGMWKFLPQGKNSPPTGSKRKE
ZmPSTOL8.05_2       ------------------PQQV------EGHTDVPRKKLEFWLGKLEVDHMKAARSKRNE

OsPSTOL1            ------------------------------------------LLCQRA-----SKNA
ZmPSTOL3.04         --------------------------TWCQRERQEDNTNRQKLRFFLCKKT----SSSIQ
ZmPSTOL3.06         -----------------------------------------------LPFKSKDE
ZmPSTOL4.05         -SNIIAGTAGVSGL---CAIIALAALMW----------------WYKRRYGMVIPWRRGV
ZmPSTOL8.02         -IPLIVSISVAASLLLPCIYV----LVW----------HRQKLEFFLCKKT----SSAIE
ZmPSTOL8.05_1       AAPIVGA-VAVAFL---CLVI----LTS----------------FLACRYGLLPFKSKNK
ZmPSTOL8.05_2       -GPIVGAVVAVAFL---CLVI----LTC----------------FLACRHGSLPFKSKNK

OsPSTOL1            P--RIESFLQKQE-TSNPKRYTLSEVKRMTKSFAHKLGRGGFGTVYKGSLPDGREIAVKM
ZmPSTOL3.04         E--NIEALISSYG-SLAPKRYKYSEVTKITSCINNKLGEGGYGVVFKGMLDDSRLVAVKF
ZmPSTOL3.06         P--RIESFLQKNG-NLHPKRYTYADVKRMTRSFTEKLGQGGFGAVYRGSLHDGRQVAVKM
ZmPSTOL4.05         S--GVESFLQKQGALLHPKGYTYSEVKRMTRSFAHKLGQGGYGAVYRGSMPDGREVAVKM
ZmPSTOL8.02         E--NIEALILAHG-SLAPKRYRYSEVTKITSSLNIKLGEGGYGMVFKGRLDDGRLVAVKF
ZmPSTOL8.05_1       PGTRIESFLQKNE-SIHPKRYTYADVKRMTKSFAVKLGQGGFGAVYKGSLHDGRQVAVKM
ZmPSTOL8.05_2       PGTRIESFLQKNESSIHPKRYTYADVKRMTKSFAVKLGQGGFGAVYKGSLH-GRQVAVKM
                        :*:::        ** *  ::*.::* .:  ***.**:* *::* :  .* :***:

OsPSTOL1            LKDT-KGDGEEFINEVAGISKTSHINVVNLLGFSLQGSKRALIYEYMPNGSLDRYSFGDS
ZmPSTOL3.04         LHDS-KGDGEEFVNEVMSIGRTSHINIVSLFGFCLEGSKRALIYEYMPNGSLDKYIY---
ZmPSTOL3.06         LKDT-KGDGEEFMNEVASISRTCHVNIVTLTGFCLQGSKRALVYEYMPNGSLERYAF---
ZmPSTOL4.05         LTGMLEGDGEEFMNEVASISRTSHVNIVTLVGYCLQGPKRALLYEYMPNGSLERYTF--G
ZmPSTOL8.02         LHDS-KGDGEEFVNEVMSIGRTSHINIVSLFGFCLEGSKRALIYEYMSNGSLDKYIY---
ZmPSTOL8.05_1       LKDT-QGDGEEFMNEVASISRTSHVNVVTLLGFCLQGSKRALIYEYMPNGSLERYAF-TG
ZmPSTOL8.05_2       LKDT-QGDGEEFMNEVASISRTSHVNVVTLLGFCLQGSKRALIYEYMPNGSLERYAF-TG
                    * .  :******:*** .*.:*.*:*:*.* *:.*:*.****:****.****::* :

OsPSTOL1            SVQGDNTLSWDRLFNIIVGIARGLEYLHCHCNIRIVHFDIKPQNILLAQDFCPKISDFGL
ZmPSTOL3.04         SDNPKEVLGWDRLYTIAIGIARGLEYLHHSCNTRIVHFDIKPQNILLDQNFQPKIADFGL
ZmPSTOL3.06         --RAENTLSWEKLFDVATGTARGLEYLHRGCNTPIVHFDIKPHNILLDQDFCPKISDFGL
ZmPSTOL4.05         SSSGEDALSWDRLFGIVVGVARGLEYLHTGCNTRIVHFDIKPHNILLDQDMCPKISDFGL
ZmPSTOL8.02         TENPKAVLGWDKLYTIAIGIARGLEYLHHSCNTRIVHFDIKPQNILLDQNFHPKIADFGL
```

```
ZmPSTOL8.05_1        DMNSENLLTWERLFDIAIGTARGLEYLHRGCNTRIVHFDIKPHNILLDQDFCPKISDFGL
ZmPSTOL8.05_2        DMNSENLLSWERLFDIAIGTARGLEYLHRGCNTRIVHFDIKPHNILLDQDFCPKISDFGL
                         .  * *::*: :  * ********  **  ********:**** *:: ***:****

OsPSTOL1             SKLCHLK-ESRISINGLRGTPGYIAPEVFSRQYGSASSKSDVYSYGMVVLEMAGAK-KNI
ZmPSTOL3.04          AKLCHTK-ESKLSMTGARGTPGFIAPEVHSRTFGVVSTKSDVYSYGMMLLEMVGGRKNVK
ZmPSTOL3.06          AKLCPNKASSAVSIVGARGTVGYIAPEVYSKQFGVVSSKSDVYSYGMMVLEMVGARDKST
ZmPSTOL4.05          AKLCGQK-ASRVSIAGARGTVGYIAPEVFSRSYEAVGSKADVYSYGMVVLEMVGAR-KNV
ZmPSTOL8.02          AKLCHMK-ESKLSMTGARGTPGFIAPEVHSRTFGVVSTKADVYSYGMMLLEMVGGRKNVN
ZmPSTOL8.05_1        AKLCLNK-ESAISIVGARGTIGYIAPEVYSKQFGTISSKSDVYSYGMMVLEMVGARDRNT
ZmPSTOL8.05_2        AKLCLNK-ESAISIVGARGTIGYIAPEVYSKQFGTISSKSDVYSYGMMVLEMVGARDRNT
                     :***  *  * :*: * *** *:*****.*: :   .:*:*******::***.*.: .

OsPSTOL1             NVS------TGSSSKYFPQWLYDNL-DQFCCPTGEISSQ-------TTD--LVRKMVVVG
ZmPSTOL3.04          SFA------QVSSEKYFPHWIYDHFGQNDGLLACEGTPE-------KEE--IAKKMALVG
ZmPSTOL3.06          SAD------SERSSQYFPQWIYEHL-DDYCVSASEVDGG-------TTE--LVRKMIVVG
ZmPSTOL4.05          HVSATDDGGNSSSSRYFPQWLYENL-DQFCRPTTTSNGEIRGDDDDATEVLLVRKMVVVG
ZmPSTOL8.02          SAA------QESSEKYFPHWIYDHFGQEDGLQACEVTRE-------NEG--IAKKLSVIG
ZmPSTOL8.05_1        SAD------SDHSSQYFPQWLYEHL-DDYCVGASEINGE-------TTE--LVRKMIVVG
ZmPSTOL8.05_2        SAD------SDHSSQYFPQWLYEHL-DDYCVGASEINGE-------TTE--LVRKMIVVG
                             *.:***:*:*::: ::    :                 :.:*: ::*

OsPSTOL1             LWCIQLVPTDRPSMREVLEMLESNGRDLPLPPKGL-------------------------
ZmPSTOL3.04          MWCIQILPLHRPTITKVLEMFDRGLDELDMPPRQNFSQTFEDPTYNFNAENMSTSSGTKT
ZmPSTOL3.06          MWCIQLIPTDRPTMTRVVEMLEGSTSNLELPPKVLLSWQASIDVFLLRPQLI--------
ZmPSTOL4.05          LWCIQSKPDSRPSMGQVLEMLESNAADLQLPPKALCT-----------------------
ZmPSTOL8.02          LSCIQILPMHRPTIGKVMEMFERGSDELDMPPRQNFSQIFEDPVHSLNAETMSMISGTKA
ZmPSTOL8.05_1        LWCIQVIPTDRPTMTRVVEMLEGSTSNLELPPRVLLS-----------------------
ZmPSTOL8.05_2        LWCIQVIPTDRPTMTRVVEMLEGSTSNLELPPRVLLS-----------------------
                     : ***  *  **:: .*:**:: .  :* :**:

OsPSTOL1             --------------------------
ZmPSTOL3.04          QR------------------------
ZmPSTOL3.06          KGY-----------IN----------
ZmPSTOL4.05          -AY-----------------------
ZmPSTOL8.02          KAYSEVLKMKEISVVNSKTIQRLPTL
ZmPSTOL8.05_1        --------------------------
ZmPSTOL8.05_2        --------------------------
```

* = identical amino acids residues in all sequences
: = highly conserved amino acids
. = different but somewhat similar amino acid
blank = dissimilar amino acids or gaps

Additional File 2: Figure S1 Alignment of the OsPSTOL1 and six maize serine/theronine receptor-like kinases highlighting the predicted domains. Letters in green indicate the glycosil hydrolase domain and in red, the transmembrane domain. The kinase domain is represented by a gray background, whereas the ATP-binding site is highlighted with black background with white letters and the Serine/Threonine protein kinase active site with yellow background.

Additional File 3: Table S2 Primers used for gene expression analyses

Gene ID	Primer	Chemical detection	Primer sequence (5' - 3')	Probe sequence	Product length (bp)
ZmPSTOL3.04	**Zm3.04_1F**	TaqMan	ACCATAGCAATTGGGATCGC	TCGAGGGTTGGAATACTTGCACCAT	75
	Zm3.04_1R		ATCGAAATGGACTATGCGTGT		
	Zm3.04_2F	TaqMan	CCTGGGTTTATAGCTCCTGAAG	TGTTACTAGAGATGGTTGGAGGCAGGA	141
	Zm3.04_2R		ACTTGAAACTTGAGCAAATGATTTAAC		
ZmPSTOL3.06	**Zm3.06_1F**	TaqMan	TCAACATCGTGACACTGACG	CATGTACTCGTAAACAAGCGCCCTCT	92
	Zm3.06_1R		GTACCTTTCAAGCGAACCATTC		
	Zm3.06_2F	TaqMan	AGTATCAGCAGGACTTGTCATG	CAAGCAGAACCCCGTCAGTGTCA	97
	Zm3.06_2R		CGCCCTCTTGGATCCTTG		
ZmPSTOL4.05	**Zm4.05_1F**	SYBR Green	GTTACAGTGCCCGGATTCTGA		110
	Zm4.05_1R		ACGTCTCTTGTACCACCACATCA		
ZmPSTOL8.02	**Zm8.02_1F**	TaqMan	TGGTTTTCAAGGGAAGGCTAG	CAGGAATTTCACTGCAACTAGACGACCA	73
	Zm8.02_1R		CCGTCACCTTTGGAGTCATG		
	Zm8.02_2F	SYBR Green	CCAGTACACAGTTTGAATGCG		133
	Zm8.02_2R		GTTTTACAGAGTTGGCAACCG		
ZmPSTOL8.05-1	**Zm8.05.1_1F**	TaqMan	ATCAAAAAGAAAAGAAGCAGCA	AACGGCAACAGCACCAACAATAGG	78
	Zm8.05.1_1R		AAGGATGTGAGAATGACTAGACAC		
	Zm8.05.1_2F	SYBR Green	AGAAGCAGCACCTATTGTTGG		73
	Zm8.05.1_2R		AGCCAAGAAGGATGTGAGAATG		
ZmPSTOL8.05-2	**Zm8.05.2_1F**	TaqMan	GGGCATACGGACGTACC	CTTGAATTCCGGCTTGGGAAGTTGG	74
	Zm8.05.2_1R		ACCGTCTTCATATGGTCAACC		
	Zm8.05.2_2F	TaqMan	TGGAGGTATCTGGAAAGAAGC	CGTATGCCCCTCCTCCACCTG	84
	Zm8.05.2_2R		ATTCAAGCCTCATTCCCGG		
18S	18S_Zm_F	SYBR Green	CGTCCTAGTCTCAACCATAAACG		82
	18S_Zm_R		CCCCGGAACCCAAAGACT		

Results with primers highlighted in bold are presented on Figure 4.

18S is a primer pair re-designed from the TaqMan® Eukaryotic 18S rRNA (Applied Biosystems, Foster City, CA) and used as endogenous control.

CAPÍTULO 2

Análise funcional de homólogos aos genes *STR1* e *STR2*, relacionados à associação micorrízica, em milho

INTRODUÇÃO

As plantas desenvolveram vários mecanismos para aumentar a superfície de exploração e absorção das raízes, visando maximizar a aquisição de fósforo sob condições limitantes desse nutriente (Williamson et al., 2001; Zhu & Lynch, 2004; Haling et al., 2013). Dentre esses mecanismos, destaca-se a associação com os fungos micorrízicos arbusculares (FMA), que atuam como extensões do sistema radicular, possibilitando uma exploração mais eficiente dos recursos do solo, notadamente o fósforo (Jakobsen, 1999; Moreira & Siqueira, 2006). Assim, a associação micorrízica resulta em ganhos no desenvolvimento, na taxa fotossintética e na taxa de transferência de carboidratos para as raízes em várias espécies vegetais (Sieverding et al., 1991; Rai et al., 2013). A relação planta-FMA é a expressão de um mutualismo nutricional, onde ambos os simbiontes são beneficiados (Berbara et al., 2006), uma vez que as plantas suprem o fungo com o carbono (fixado via processos fotossintéticos pelo simbionte autotrófico), enquanto os fungos provêm as plantas de nutrientes (Berbara et al., 2006; Parniske, 2008; Smith & Read, 2008).

A maioria dessa troca ocorre nas células do córtex radicular, em uma estrutura de hifas altamente ramificadas, denominada arbúsculo. Os arbúsculos representam o ponto central na associação entre o fungo e a planta, uma vez que é o local onde existe maior superfície de contato entre os simbiontes envolvidos (Siddiqui & Pichtel, 2008; Yang & Paszkowski, 2011). Dois genes relacionados ao desenvolvimento dos arbúsculos foram identificados em *Medicago truncatula* (Zhang et al., 2010) e arroz (Gutjahr et al., 2012). Esses genes codificam transportadores *half-size* ATP *binding cassette* (ABC) da subfamília G, denominados *STR1* e *STR2*. Mutações em ambos os genes inibem o desenvolvimento arbuscular, produzindo arbúsculos pequenos e atrofiados, o que impede que a simbiose micorrízica ocorra (Zhang et al., 2010; Gutjahr et al., 2012).

Uma análise filogenética sugeriu que os genes *STR* são conservados nas angiospermas, sendo identificados ortólogos putativos para ambos os genes em milho, nos cromossomos 3 e 10 (Gutjahr et al, 2012). De forma interessante, esses genes candidatos foram co-localizados com QTLs para morfologia radicular, acúmulo de biomassa e conteúdo de P, detectados no capítulo 1. Além das linhagens L3 e L22 serem os progenitores da população segregante avaliada no capítulo 1, essas linhagens são contrastantes quanto ao uso do P em condições de campo (Parentoni et al., 2010) e quanto a morfologia radicular

em solução nutritiva (de Sousa et al., 2012). Assim, o objetivo desse estudo foi a caracterização funcional dos homólogos aos genes *STR1* e *STR2* nas linhagens de milho L3 e L22.

MATERIAL E MÉTODOS

Avaliação dos domínios protéicos dos STR1 e STR2 de milho, sorgo, arroz e *M. truncatula*

Com o auxílio do programa CDART (*Conservad Domain Architecture Retrieval Tool*, http://www.ncbi.nlm.nih.gov/Structure/lexington/lexington.cgi), os domínios das proteínas codificadas pelos genes *STR1* e *STR2* de arroz e *M. truncatula* e de seus putativos homólogos em milho e sorgo, foram identificados.

Experimento de inoculação em casa de vegetação

Sementes de milho das linhagens L3 e L22 foram germinadas em areia durante uma semana em casa de vegetação, e as plântulas foram transferidas para vasos de 1L contendo solo autoclavado e areia na proporção 1:1. O delineamento experimental adotado foi o inteiramente casualizado, com cinco repetições, cinco épocas de coleta (10, 20, 30, 40 e 50 dias após a transplantio, dat), em plantas controle (sem inoculação) e inoculadas. A espécie de FMA utilizada foi *Rhizophagus irregulares* (Biorize), na concentração final de 500 esporos/planta. Durante as duas primeiras semanas, as plantas foram regadas com 100 mL de água deionizada a cada dois dias. A partir da terceira semana, o uso da água foi intercalado com solução nutritiva de Hoagland com concentração de H_2PO_4 a 100 µM.

Nas épocas das coletas, as plantas foram cuidadosamente retiradas dos vasos e o excesso de substrato preso às raízes foi removido com água. A raiz e a parte aérea foram separadas e o peso fresco registrado. Em seguida, amostras de ambos os tecidos foram coletadas e imediatamente transferidas para nitrogênio líquido para extração de RNA. Amostras da raiz e da parte aérea também foram coletadas para a quantificação de P e para avaliar a colonização quantitativa e qualitativamente nas raízes. Por fim, o tecido restante foi seco a 70°C por uma semana para se estabelecer o peso seco. É importante destacar que o peso fresco dos tecidos foi avaliado antes e depois das amostragens, de modo a permitir uma normalização do peso seco final após as amostragens.

Os dados obtidos foram submetidos à análise de variância e interação genótipo (L3 e L22) ambiente (controle e inoculado) para cada tempo de coleta e característica avaliada, considerando 5 % de probabilidade, com o auxílio do programa SISVAR (Ferreira, 2000).

Quando não detectada interação entre genótipo e ambiente foram considerados quatro tratamentos: L3 inoculado, L3 controle, L22 inoculado e L22 controle. As médias comparadas pelo teste *t*, a 5% de probabilidade.

Avaliação quantitativa e qualitativa da colonização micorrízica

As raízes de milho foram lavadas em água corrente e imersas em KOH 10% por 12 horas, sendo então lavadas por três vezes com água deionizada e imersas em HCl 0,3 M por duas horas. O HCl foi retirado e as raízes aquecidas a 95ºC por 1 hora em solução de glicerol acidificada (1% HCl) com *Trypan blue* 0,1%. A colonização micorrízica foi quantificada seguindo o método *grid-line intersect* modificado descrito por Paszkowski et al. (2006), com modificações. Resumidamente, a colonização total foi determinada pela quantidade total do FMA e pela presença de estruturas fúngicas específicas em microscópio ótico (Olympus CX31) com aumento de 10 vezes em 100 pontos aleatórios por amostra radicular. Para uma representação confiável da colonização radicular por estruturas micorrízicas específicas, todas as estruturas presentes em um mesmo ponto aleatório foram consideradas separadamente.

O desenvolvimento dos arbúsculos nas diferentes épocas de coleta foi avaliado por meio da coloração com aglutinina de germen de trigo (*wheat germ aglutinin* - WGA) conjugado com Alexa Fluor 488. Primeiramente, as raízes foram lavadas com água e transferidas para tubos contendo etanol 50% por 4 horas. Após esse período, as raízes foram enxaguadas com água deionizada e incubadas em KOH 20%, onde permaneceram por 3 dias. Posteriormente, as raízes foram novamente enxaguadas com água e adicionou-se HCl 0,1 M em quantidade suficiente para cobrir as raízes. Após duas horas, as amostras foram enxaguadas duas vezes com água deionizada e uma vez com tampão fosfato salino (PBS) NaCl 0,8 %, KCl 0,02 %, Na_2HPO_4 0,14 % e KH_2PO_4 0,02 %, pH 7,4. Por fim, foi adicionada a solução de coloração WGA conjugada com AlexaFluor488 na concentração final 0.2 µg/ml no tampão PBS e as amostras mantidas à temperatura ambiente por 12 horas. Após a coloração, as raízes foram transferidas para uma solução de iodeto de propídio (PI) por 5 minutos para corar a parede das células vegetais. Após esse procedimento, fragmentos aleatórios das raízes coradas foram transferidos para lâminas para serem analisados no microscópio confocal, seguindo o procedimento descrito por Zhang et al. (2010).

Um total de 200 arbúsculos foram medidos para cada genótipo e em cada época de coleta com o auxílio do programa Image J (REF). A distribuição do tamanho desses arbúsculos foi avaliada com a função *box plot* utilizando no programa R versão 3.1.0.

Determinação do conteúdo de fósforo nos tecidos da planta

Aproximadamente 20 mg de cada tecido foram transferidos para microtubos de 2.5 mL contendo 500 µL de água deionizada. As amostras foram aquecidas por 45 minutos a 85°C. Posteriormente, 100 µL e 50 µL das amostras aquecidas de raiz e folha, respectivamente, foram transferidos para novos tubos e adicionou-se água até completar 300 µL. A essa mistura, foram adicionados 600 µl de uma solução de moblidato de amônio (0.44%) e 100 µl de ácido ascórbico (10%). As amostras foram incubadas a 42°C por uma hora e a absorbância medida a 820 nm. O conteúdo de P foi calculado multiplicando o peso seco pela concentração de P de cada tecido.

Extração de RNA e síntese de cDNA

O RNA da parte aérea e das raízes foi extraído utilizando o reagente TRIZOL. Inicialmente, os tecidos das plantas foram macerados e aproxidamente 100 mg de tecido foi transferido para tubos de 2 mL, que receberam a adição de 1 mL de TRIZOL. Após 5 minutos à temperatura ambiente, a extração do RNA ocorreu seguindo as etapas: (i) adição de 0,2 mL de clorofórmio e agitação em vortex por 3 minutos; (ii) centrifugação a 12000 x g a 4°C por 15 minutos; (iii) transferência da fase aquosa para um novo tubo seguido da adição de 0,5 mL de isopropanol por 5 minutos à temperatura ambiente; (iv) centrifugação a 12000 x g a 4°C por 15 minutos e eliminação do sobrenadante; (v) o *pellet* foi lavado com 1 mL de etanol 75% gelado (- 20°C) e centrifugação a 10000 g a 4°C por 10 minutos; (vi) após a retirada do etanol, 1 mL de LiCl 4M foi adicionado e as amostras foram incubadas por 1 hora a -20°C; (vii) após esse período, as amostras foram centrifugadas a 12000 x g a 4°C por 20 minutos e o sobrenadante descartado; (viii) o *pellet* foi novamente lavado com 200 µL de etanol 75% gelado (-20°C) e centrifugado a 12000 x g a 4°C por 5 minutos; (vi) por fim, o etanol foi retirado completamente e as amostras ressuspendidas em 20 µL de água MiliQ.

O RNA foi quantificado no Nano Drop ND-1000 Spectrophotometer (Saveen 1 Werner, Malmo, Suécia) e 1 µL utilizado para avaliar sua qualidade em gel de agarose 1% submetido a 150 V por 15 minutos. Para cada amostra, um volume correspondente a 1,1 µg de RNA foi tratado por 15 minutos a 25°C com 1 µL de DNase I (Invitrogen, Carlsbad, USA). A digestão da DNase foi terminada utilizando 1 µL de EDTA (25 mM) por 10 minutos a 65°C. A qualidade do RNA foi novamente avaliada em gel de agarose 1%. A síntese de cDNA foi realizada utilizando o kit SuperScript II reverse transcriptase (Invitrogen, Carlsbad, USA) seguindo as orientações do fabricante.

Análise da expressão gênica

A expressão dos genes foi avaliada nas raízes e na parte aérea das duas linhagens de milho, L3 e L22, coletadas a 20 e 50 dias após o transplantio. As reações foram realizadas utilizando o kit SYBR green I master mix (Roche, Mannheim, Alemanha) no equipamento Roche 480 Light Cycler, seguindo as recomendações do fabricante. Foram utilizadas três repetições biológicas e três técnicas, além de dois controles negativos contendo água para cada gene testado. O gene gliceraldeído-3-fosfato desidrogenase (*GAPDH*) foi utilizado como normalizador. Um total de seis genes alvo foram testados, incluindo *STR1* e *STR2* de milho, dois transportadores de fosfato *ZmPT3* e *ZmPT6* (Nagy et al., 2006), o gene marcador de micorriza arbuscular *ZmAM3*, homólogo ao *OsAM3* (Gutjahr et al., 2008) e o fator de elongação de *Rhizophagus irregulares, RiEF* (Sokolski et al., 2010). A expressão gênica relativa e as curvas de *melting* foram avaliadas de acordo com as condições descritas por Quan (2014). Os *primers* utilizados estão apresentados na Tabela 1.

Tabela 1. *Primers* utilizados para as análises de expressão gênica.

Identificador GRMZM	Gene	Sequência (5'>3')
GRMZM2G357034	*STR1*	For - ACACCAAGAACACCATACCAG Rev - AGGACTCGAAGTTTGACATGG
GRMZM2G035276	*STR2*	For - TGTTTCAGCCCTCGATTGG Rev - CAAATCAACCCGGCTTTCC
GRMZM2G112377	*ZmPT3*	For - GCCTTCCGTTACGTCATTGT Rev - AGCACGTCTCTGATCCCATC
GRMZM2G881088	*ZmPT6*	For - GATCCAGCTCATCGGTTTCT Rev - GAGCGTGGTGTGTTTGTTCT
GRMZM2G135244	*ZmAM3*	For - ATCTGTCGTTGCGTTCCTCT Rev - GCATCTATCACTGCGGGAAT
-	*RiEF*	For - TGTTGCTTTCGTCCCAATATC Rev - GGTTTATCGGTAGGTCGAG
GRMZM2G046804	*GAPDH*	For - CTTCGGCATTGTTGAGGGTTTG Rev - TCCTTGGCTGAGGGTCCGTC

RESULTADOS E DISCUSSÃO

Conservação de domínios entre STR1 e STR2 de milho, sorgo, arroz e *M. truncatula*

As sequências aminoacídicas dos STR1 e STR2 de arroz (Gutjahr et al, 2012) e de *M. truncatula* (Zhang et al, 2010) foram alinhadas com as sequências dos seus putativos homólogos em milho e sorgo mostrando que o domínio ABC é conservado entre essas espécies (Figuras suplementares 1 e 2). A conservação desses domínios indica que esses genes são relacionados, conforme observado por Gutjahr et al. (2012). Com base na conservação dos genes *STR1* e *STR2* em diferentes espécies, essem autores sugerem que o repertório de genes requerido para o desenvolvimento da micorriza arbuscular pode ter emergido em um ancestral comum das plantas terrestres.

Produção de matéria seca e acúmulo de P nos tecidos das linhagens de milho contrastantes para o uso de fósforo

A linhagem L3 apresentou maiores produção de matéria seca de raiz em relação a L22 (Figura 1). A diferença entre as linhagens chegou a 45 e 38% para os tratamentos inoculado e não inoculado, respectivamente, aos 50 dias após o transplantio. Além da diferença na biomassa radicular entre as linhagens, a inoculação com *R. irregularis* favoreceu a produção de raízes em ambas as linhagens a partir de 20 dias após o transplantio (Figura 1). A inoculação micorrízica teve maior influência sobre L3, uma vez que a diferença entre a biomassa das raízes de plantas inoculadas e não inoculadas foi superior a 30% para L3, enquanto essa diferença ficou em torno de 20% para L22 (Figura 1).

Esses resultados indicaram que a inoculação com *R. irregularis* provocou alterações no sistema radicular das plantas micorrizadas, levando a uma maior produção de massa seca, principalmente na linhagem eficiente no uso do P (Figura 1; Figura suplementar 3). Os efeitos da inoculação de *Glomus etunicatum* e *Glomus clarum* e da adição de níveis de P (0, 50, 100 e 200 mg/kg de solo) sobre a morfologia radicular e concentração de P foram avaliados em plantas de milho em diferentes épocas de coleta (Bressan & Vasconcellos, 2002). Os autores observaram que a inoculação com *Glomus clarum* levou ao aumento do peso das raízes seminais e adventícias, do número de raízes laterais e maior produção de biomassa radicular em milho (Bressan & Vasconcellos, 2002). De acordo com Hetrick et al. (1988), os FMA alteram significativamente o sistema radicular reduzindo a quantidade relativa da ramificação da raiz. Esses autores defendem que as plantas micorrizadas desenvolvem um padrão de crescimento mais longo e exploratório, permitindo que as hifas do FMA explorem nutrientes em um volume maior de solo. Em contrapartida, as plantas não

micorrizadas mantiveram um padrão de crescimento radicular mais ramificado, uma vez que as próprias raízes vão exercer a função de extrair diretamente os nutrientes do solo.

No entanto, o aumento na matéria seca de raiz da linhagem L3 não refletiu em maior acúmulo de massa seca da parte aérea, uma vez que a L22 apresentou parte aérea mais desenvolvida com e sem inoculação em comparação com a L3 (Figura 1). Para a L22, somente aos 50 dias a inoculação apresentou aumento significativo no desenvolvimento da parte aérea em relação ao tratamento não inoculado. Já para L3, o tratamento inoculado apresentou diferença significativa em relação ao não inoculado aos 20 e 30 dias após o transplantio no desenvolvimento da parte aérea, que se tornou equivalente após 40 e 50 dias do transplantio (Figura 1).

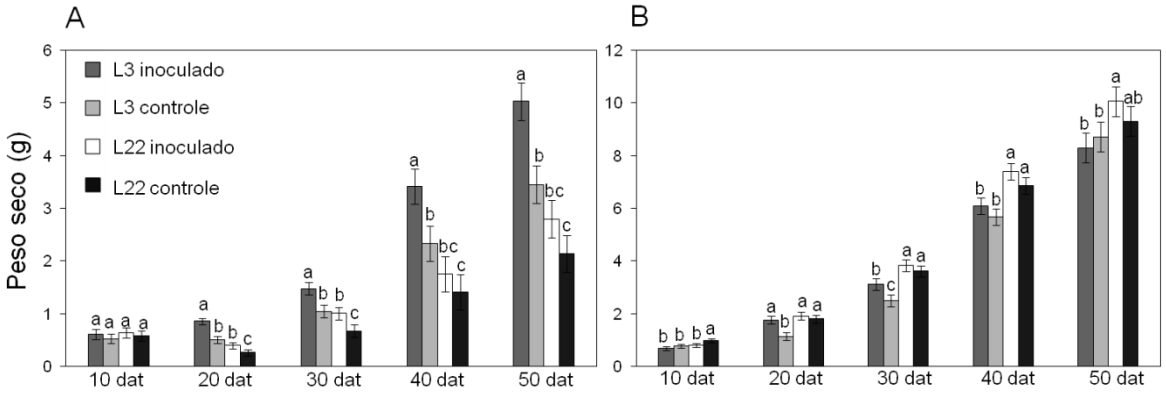


Figura 1. Produção de matéria seca da (A) raiz e da (B) parte aérea em L3 e L22 com e sem inoculação de *Rhizophagus irregularis*. A biomassa foi medida 10, 20, 30, 40 e 50 dias após o transplantio. As barras de erro indicam a DMS (diferença mínima significativa) considerando cinco repetições biológicas. L3 - linhagem eficiente no uso do P; L22 - linhagem ineficiente no uso do P.

Apesar de ser considerada ineficiente para o uso do P, a linhagem L22 apresenta maior produção de biomassa de parte aérea que L3 de forma constitutiva, inclusive em condições de campo, como foi apresentando por Mendes et al. (2014). Esses autores mostraram que a população de RILs retrocruzada com a L22 apresentou médias de altura de plantas e de massa seca de palhada em solo com baixo P superiores em comparação com a média das progênies retrocruzadas com a L3. Como os índices de eficiência no uso de P são calculados com base na produção de grãos (Parentoni et al., 2010), o maior desenvolvimento da parte aérea intrínseco à linhagem L22 não é considerado como um componente no índice de eficiência no uso de nutrientes, como o P. Adicionalmente, a L22

possui menor produção de grãos que a L3, sendo portanto considerada ineficiente no uso de P.

Em relação à razão do peso seco entre raiz:parte aérea, diferenças significativas entre as linhagens foram observadas a partir de 20 dias de transplantio, enquanto os resultados entre os tratamentos variando de acordo com a linhagem e época de coleta (Figura 2). Com exceção da primeira época de coleta, a razão raiz:parte aérea foi superior na L3 comparada com a L22. Esses resultados são justificados uma vez que a L3 apresenta maior crescimento radicular em comparação a L22, e em contrapartida a L22 apresentou maior biomassa da parte aérea (Figura 1). Assim, foi observado um maior crescimento radicular em relação à parte aérea na L3, principalmente na presença de FMA. Considerando os tratamentos, a inoculação aumentou a relação raiz:parte aérea na L3 apenas aos 40 dias após o transplantio, enquanto o efeito significativo da inoculação na L22 ocorreu apenas aos 30 dias do transplantio (Figura 2). Aos 50 dias após o transplantio, foi observada uma interação significativa entre genótipos e ambientes para a relação raiz:parte aérea, com o efeito da inoculação observado somente para a linhagem L3, que apresentou uma relação raíz:parte aérea superior no ambiente inoculado (Tabela 2).

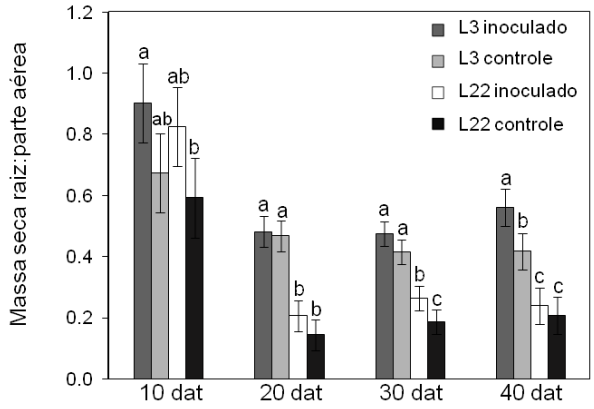


Figura 2. Razão da massa seca da raiz em relação à massa seca da parte aérea para L3 e L22 com e sem inoculação por *Rhizophagus irregularis*. A biomassa foi medida 10, 20, 30 e 40 dias após o transplantio. As barras de erro indicam a DMS (diferença mínima significativa) para cinco repetições biológicas. L3 - linhagem eficiente no uso do P; L22 - linhagem ineficiente no uso do P.

Tabela 2. Razão da massa seca da raiz em relação a massa seca da parte aérea (raíz:parte aérea) para linhagense L3 e L22, com e sem inoculação de *Rhizophagus irregularis* aos 50 dias após o transplantio.

Linhagem	Inoculado	Não inoculado
L3	0,606 a[1] A[2]	0,398 a B
L22	0,282 b A	0,222 b A

[1] Médias seguidas pela mesma letra minúscula na coluna não diferem entre si a 5% pelo teste t. [2]Médias seguidas pela mesma letra maiúscula na linha não diferem entre si a 5% pelo teste t.

O aumento da razão raiz:parte aérea é uma adaptação observada em plantas sob condições limitantes de P (Mollier & Pellerin, 1999; Liu et al, 2004; Niu et al, 2013). Gill et al. (1992) verificaram que a relação raiz:parte aérea foi maior nas plântulas de milho crescidas sob baixo nível de P em solução nutritiva. Em condições limitantes desse nutriente, as raízes atuam como os drenos preferenciais dos fotoassimilados, visando o maior desenvolvimento do sistema radicular e, consequentemente, o aumento da área de absorção do nutriente.

O conteúdo de P na parte aérea foi superior no tratamento inoculado aos 20 dias para ambos os genótipos, já nas demais coletas não houve diferença significativa (Figura 3). Na raiz, pode-se verificar que as linhagens micorrizadas apresentaram maior conteúdo de P aos 40 e 50 dias após o transplantio, tendo um maior efeito na L3 em relação a L22. Esses resultados indicam que a inoculação micorrízica favoreceu uma maior absorção de P pelas raízes, que foi direcionado para o desenvolvimento radicular, sem efeito significativo na parte aérea. É importante salientar que o peso seco dos tecidos foi utilizado para calcular o conteúdo de P, influenciando fortemente essa característica. Como destacado anteriormente, mesmo sendo considerada eficiente a linhagem L3 apresenta parte aérea menos desenvolvida em relação a L22.

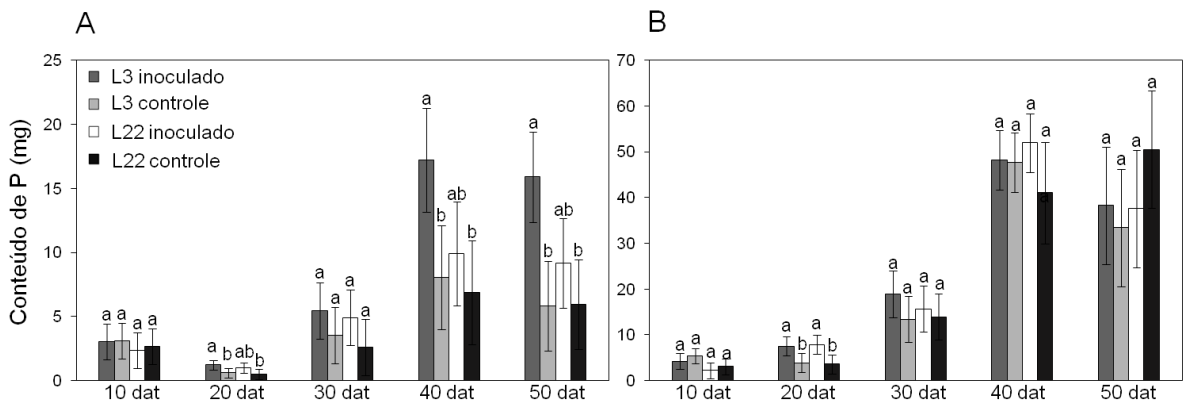


Figura 3. Conteúdo de P na (A) raiz e na (B) parte aérea de L3 e L22 com e sem inoculação de *Rhizophagus irregularis*. O conteúdo de P foi medido aos 10, 20, 30, 40 e 50 dias após o transplantio, multiplicando a massa seca pela concentração de P de cada tecido da planta. Barras de erro indicam a DMS (diferença mínima significativa) para cinco repetições biológicas. L3 - linhagem eficiente no uso do P; L22 - linhagem ineficiente no uso do P.

Avaliação da colonização radicular e do desenvolvimento dos arbúsculos de *R. irregularis*

Com o intuito de verificar se o contraste na eficiência do uso do P em L3 e L22, observada em campo por Parentoni et al. (2010), poderia ser relacionada com a diferenças na associação micorrízica entre essas linhagens, uma avaliação quantitativa e qualitativa da colonização foi realizada. A extensão da colonização micorrízica foi avaliada em detalhes aos 10, 20, 30, 40 e 50 dat, considerado a colonização total, que englobava as principais estruturas fúngicas (vesículas, arbúsculos, esporos e hifa intraradicular) e a colonização baseada apenas nos arbúsculos e vesículas. As estruturas fúngicas estão apresentadas nas Figuras 4 A-D e as quantificações na Figura 4E. Os níveis de colonização foram muito elevados em ambas as linhagens, ultrapassando 90% aos 50 dias, fase do desenvolvimento da planta onde a L22 apresentou maior porcentagem de colonização considerando todas as estruturas fúngicas (Figura 4E). Nas quatro primeiras épocas de coleta, independente da forma de avaliação adotada, não foram detectadas diferenças significativas entre L3 e L22, exceto para as vesículas aos 30 dias após o transplantio, detectadas em maior quantidade na L22.

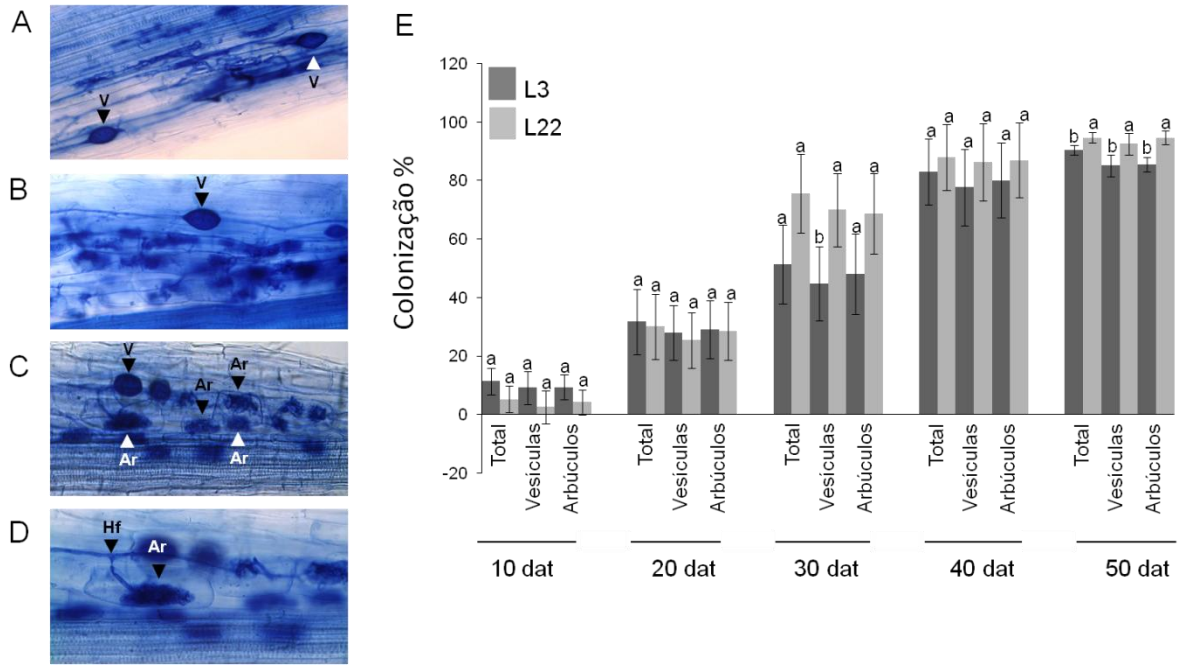


Figura 4. Colonização micorrízica nas raízes das linhagens L3 e L22 coradas com *Trypan blue*. (A) e (B) Vesículas (v) nas raízes da L22; (C) Arbúsculos (Ar) e vesículas (v) na L3; (D) Arbúsculos (Ar) e hifa (Hf) na L3; (E) colonização total, de arbúsculos e vesículas em diferentes épocas de coleta. Barras de erro indicam a DMS (diferença mínima significativa) para cinco repetições biológicas. L3 - linhagem eficiente no uso do P; L22 - linhagem ineficiente no uso do P.

A análise da colonização micorrízica por meio da coloração por *Trypan blue* permite uma abordagem quantitativa da colonização. No entanto, para uma avaliação precisa do desenvolvimento e da morfologia dos arbúsculos, que são as principais estruturas responsáveis pela troca de nutrientes entre a planta e o fungo, a coloração com WGA-AlexaFluor488 visualizado em microscópio confocal é uma ferramenta mais adequada (Javot et al., 2007; Zhang et al., 2010; Gutjahr et al., 2012). Utilizando essa metodologia, os arbúsculos já se encontram bem desenvolvidos, altamente ramificados e ocupando completamente as células a partir dos 20 dias após o transplantio, sendo que alguns dele já iniciam o processo de senescência mais acentuada aos 40 dias (Figura 5). De maneira geral, não foram observadas diferenças quanto à formação dos arbúsculos entre os genótipos, que apresentaram bom desenvolvimento de todas as estruturas fúngicas necessárias para a colonização por *R. irregularis* (Figuras 4 e 5).

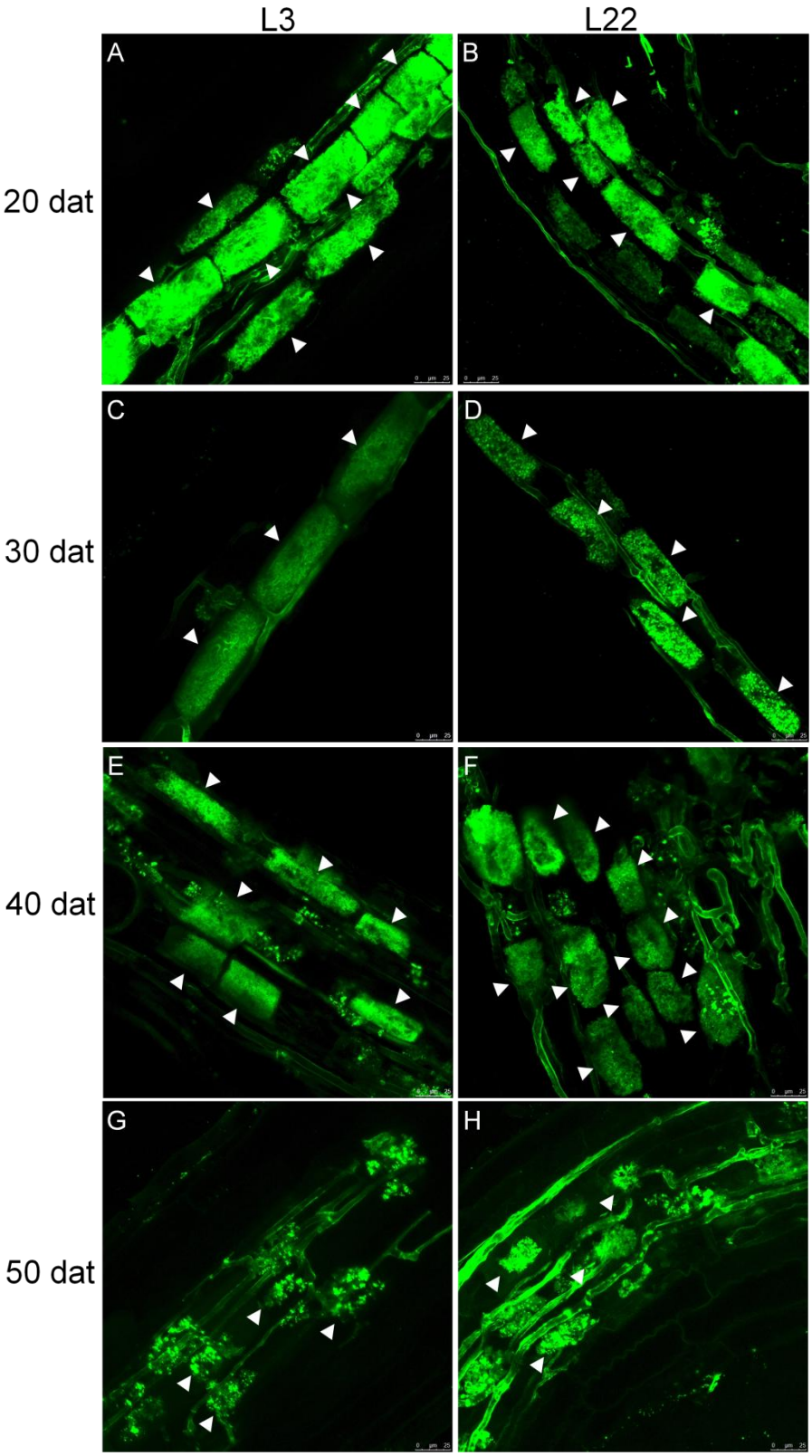

L3
L22
20 dat
30 dat
40 dat
50 dat
A
B
C
D
E
F
G
H

Figura 5. Imagens do microscócopio confocal de arbúsculos de *R. irregularis* corados com WGA-AlexaFluor488 em células corticais de L3 e L22. (A), (C), (E) e (G) arbúsculos visualizados em raízes de L3 coletadas aos 20, 30, 40 e 50 dias após o transplantio (dat), respectivamente; (B), (D), (F) e (H) arbúsculos visualizados em raízes de L22 coletadas aos 20, 30, 40 e 50 dat, respectivamente. Setas brancas indicam os arbúsculos. L3 - linhagem eficiente no uso do P; L22 - linhagem ineficiente no uso do P.

Adicionalmente, o tamanho dos arbúsculos foi avaliado, indicando que o tamanho dos arbúsculos foi reduzindo ao longo do tempo de colonização (Figura 6). Entre 20 e 40 dias após o transplantio, a L3 apresentou em média, arbúsculos maiores que a L22 (Figura 6). Já aos 50 dias, os arbúsculos foram menores na L3, sugerindo um processo de senescência mais acelerado nessa linhagem em comparação com a L22 (Figura 6). É importante salientar que o tamanho das células não foi considerado nessa abordagem, uma vez que a coloração com iodeto de propídio (PI), que permitiria delimitar a parede da célula não foi eficiente nas células que possuiam arbúsculos.

O tamanho e a quantidade dos arbúsculos estão relacionados com a área disponível para transferência de P (Mosse, 1973; Alexander et al., 1989). Aos 50 dias após o transplantio, a linhagem L22 apresentou arbúsculos maiores e mais numerosos em comparação a L3. No entanto, tal fato não refletiu em um maior conteúdo de P nas raízes de L22. Isso evidencia, uma maior capacidade da linhagem L3 em absorver P, independente do nível de colonização e do estágio de desenvolvimento dos arbúsculos dos FMA. Combinando todos os resultados apresentados, podemos sugerir que as diferenças na eficiência de uso do P entre L3 e L22 observada em campo por Parentoni & Souza-Júnior (2008), não estão relacionadas com diferenças na associação micorrízica entre essas linhagens.

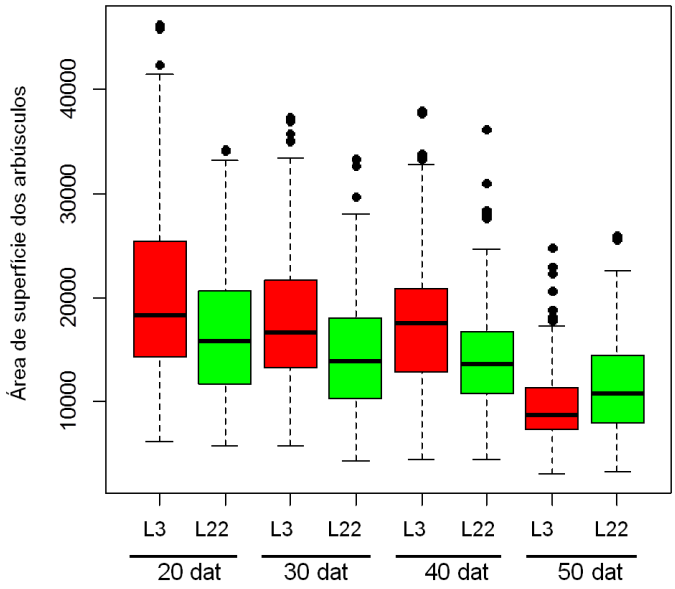


Figura 6. Análise de box plot representando a distribuição do tamanho dos arbúsculos nas linhagens L3 e L22 aos 20, 30, 40 e 50 dias após o transplantio. L3 - linhagem eficiente no uso do P; L22 - linhagem ineficiente no uso do P.

Análise temporal e espacial da expressão dos genes *STR1* e *STR2* em milho e de outros genes relacionados com a associação micorrízica

As épocas para análises da expressão dos genes candidatos foram definidos em função dos maiores contrastes para o tamanho dos arbúsculos, que foram aos 20 e 50 dias após o transplantio (Figura 6). Os genes *STR1* e *STR2* apresentaram padrões de expressão muito similares, sendo expressos especificamente em raízes micorrizadas e em maiores níveis aos 50 dias após o transplantio (Figura 7). Apesar das diferenças no tamanho dos arbúculos entre L3 e L22, a partir dos 20 dias os arbúsculos já estavam bem desenvolvidos em ambas as linhagens (Figura 5). Como Gutjahr et al. (2012) observaram que a expressão dos genes *STR1* e *STR2* era restrita às células das raízes que possuiam arbúsculos em arroz, poderíamos justificar o mesmo nível de expressão desses genes nos dois genótipos de milho. Já o aumento da expressão dos genes *STR* ao longo do tempo de inoculação foi altamente associado com o aumento na porcentagem de micorrização (Figura 4). Adicionalmente, os genes *STR* não foram expressos nas raízes não inoculadas com *R. irregularis*, nem na parte aérea (dados não mostrados).

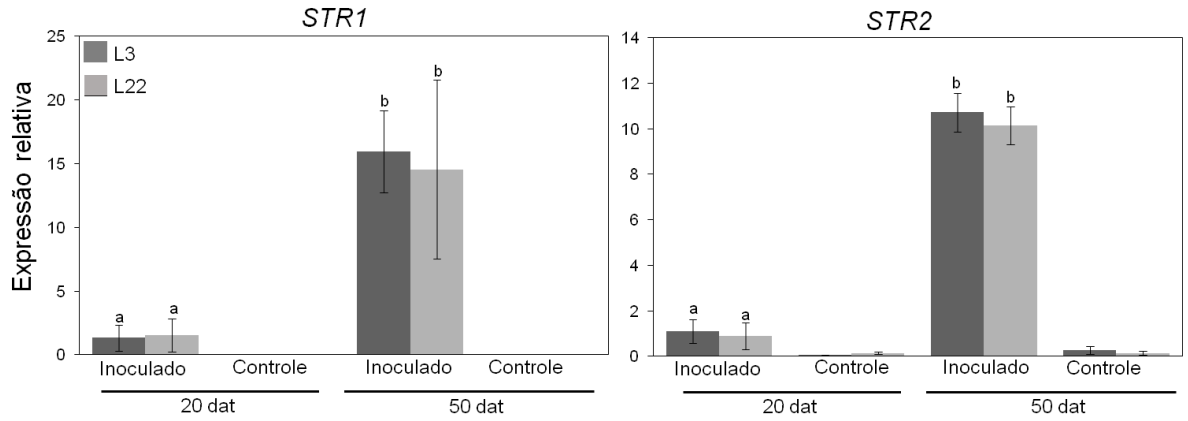


Figura 7. Análise da expressão relativa dos genes *STR1* e *STR2* em milho. A expressão foi analisada por PCR em tempo real nas raízes de duas linhagens de milho contrastantes para a eficiência no uso do P, L3 (eficiente) e L22 (ineficiente), aos 20 e 50 dias após o transplantio. A inoculação foi realizada com 500 esporos de *R. irregulares* e o tratamento controle não recebeu inoculação. As barras de erro indicam o desvio padrão para três repetições biológicas.

Além dos genes *STR*, outros genes relacionados com associação micorrízica foram avaliados (Figura 8). Os genes *ZmAM3 e RiEF* foram utilizados para quantificar o FMA inoculado. Conforme esperado, ambos os genes foram expressos somente nas raízes micorrizadas (Figura 8). A expressão do *ZmAM3* foi aumentada ao longo do tempo de inoculação em ambas as linhagens, sendo mais expresso na L22 aos 50 dias (Figura 8). Um padrão semelhante de expressão foi obtido para o gene *RiEF* (Figura 8). As diferenças na expressão desses genes ao longo do tempo podem estar relacionadas ao aumento da porcentagem de colonização observados entre 20 e 50 dat. Tais resultados seriam esperados considerando que o gene *ZmAM3* é especificamente induzido com o desenvolvimento da simbiose micorrízica (Sawers & Paszkowski, dados não publicados) e o *RiEF* é um gene expresso pelo FMA inoculado *R. irregularis* (Sokolski et al., 2010; Quan, 2014).

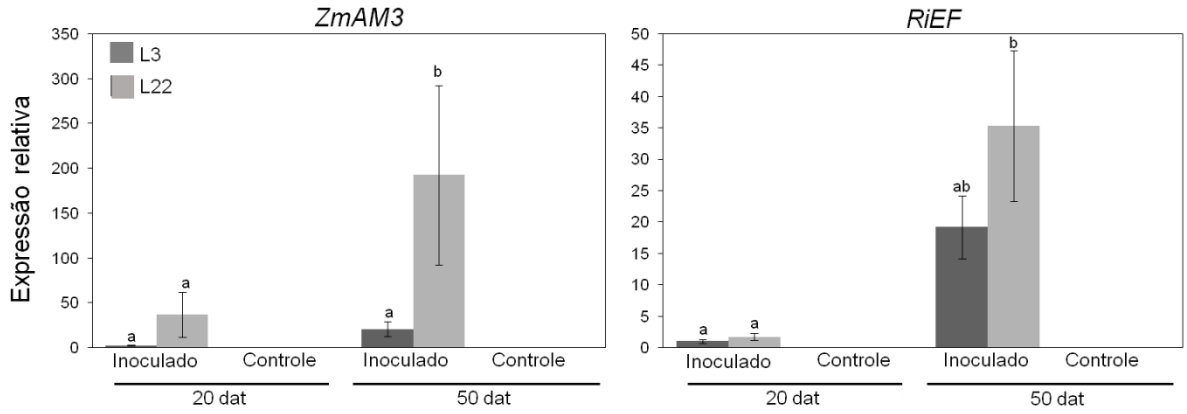


Figura 8. Análise da expressão relativa dos genes *ZmAM3* e *RiEF* em milho. A expressão foi analisada por PCR em tempo real nas raízes de duas linhagens de milho contrastantes para a eficiência no uso do P, L3 (eficiente) e L22 (ineficiente), aos 20 e 50 dias após o transplantio. A inoculação foi realizada com 500 esporos de *R. irregulares* e o tratamento controle não recebeu inoculação. As barras de erro indicam o desvio padrão para três repetições biológicas.

Quanto ao padrão de expressão dos transportadores de P em L3 e L22, o *ZmPT3* foi expresso preferencialmente nas raízes em relação à parte aérea, sendo expresso em raízes micorrizadas e não micorrizadas independentemente do genótipo e do tempo (Figura 9a). Adicionalmente, esse gene foi altamente reprimido nas raízes micorrizadas aos 50 dias em ambas as linhagens comparado aos 20 dias, e apresentou uma alta expressão na L3 não micorrizada 20 dias após o transplantio. Esses resultados corroboram com os obtidos por Quan (2014) que verificou uma indução na expressão do *ZmPT3* em raízes de milho não micorrizadas cultivadas a 100 µM, a mesma concentração de P utilizada no nosso experimento.

Já o *ZmPT6* foi preferencialmente expresso nas raízes micorrizadas aos 20 e 50 dias após o transplantio. Aos 20 dias, a expressão do *ZmPT6* foi superior na L22 em relação à L3, atingindo níveis similares nas duas linhagens aos 50 dias após o transplantio (Figura 9b). O *ZmPT6* não foi expresso na parte aérea das linhagens avaliadas, diferente do observado por Nagy et al. (2006) e Quan (2014).

Baseado em análises filogenéticas e no padrão de expressão, os transportadores de P de milho *ZmPT3* e *ZmPT6* foram associados com a aquisição de P de forma direta e via simbiose micorrízica, respectivamente. Apesar de serem relacionados com a aquisição de P do solo, esses genes também são expressos na parte aérea das plantas de milho (Nagy et al., 2006; Quan, 2014). Os perfis de expressão observados nesse estudo para *ZmPT3* e *ZmPT6* respeitam a classificação proposta por Quan (2014), uma vez que *ZmPT3* foi

reprimido na presença de micorrização, sendo altamente expresso em raízes não micorrizadas, enquanto *ZmPT6* foi induzido na micorrização.

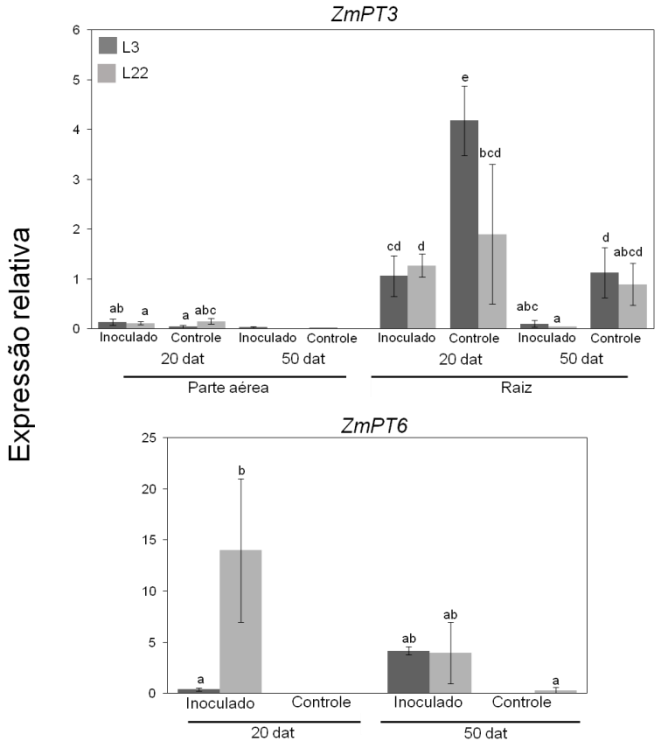


Figura 9. Análise da expressão relativa dos genes *ZmPT3* e *ZmPT6* em milho. A expressão foi analisada por PCR em tempo real na parte aérea (somente para *ZmPT3*) e raízes de duas linhagens de milho contrastantes para a eficiência no uso do P, L3 (eficiente) e L22 (ineficiente), aos 20 e 50 dias após o transplantio. A inoculação foi realizada com 500 esporos de *R. irregulares* e o tratamento controle não recebeu inoculação. As barras de erro indicam o desvio padrão para três repetições biológicas.

CONCLUSÃO

Os resultados obtidos no presente trabalho demonstraram que os genes *STR1* e *STR2* possuem expressão apenas em raízes micorrizadas e induzida ao longo do tempo de micorrização em duas linhagens de milho. Apesar das linhagens serem contrastantes quanto à eficiência no uso de P em campo, elas apresentaram altos níveis de colonização micorrízica, com arbúsculos bem desenvolvidos a partir de 20 dias após o transplantio. Adicionalmente, a micorrização conferiu aumento significativo apenas no desenvolvimento do sistema radicular, principalmente na linhagem L3 entre 40 e 50 dias após o transplantio. Assim, podemos sugerir que os genes *STR1* e *STR2* são induzidos pela colonização micorrízica e que tal associação não está relacionada com diferença na eficiência no uso de P entre essas linhagens. É necessário destacar, que o perfil de expressão de *STR1* e *STR2* demonstrado para milho é compatível com a expressão dos seus homólogos em arroz (Gutjahr et al. 2012). Além disso, STR1 e STR2 de milho, de arroz e de *M. truncatula* compartilham o domínio ABC altamente conservado, sugerindo que esses transportadores ABC podem estar associados ao desenvolvimento dos arbúsculos em milho. No entanto, estudos adicionais são necessários para a validação da função desses genes como homólogos do *STR1* e *STR2* em milho.

REFERÊNCIAS BIBLIOGRÁFICAS

Alexander, T. O. M.; Toth, R.; Meier, R.; Weber, H. C. Dynamics of arbuscule development and degeneration in onion, bean , and tomato with reference to vesicular - arbuscular mycorrhizae in grasses. Canadian Journal of Botany, 67:2505–2513, 1989.

Berbara, R. L. L.; de Souza, F. A.; Fonseca, H. M. A. C. Fungos micorrízicos arbusculares: muito além da nutrição. In: Fernandes, M. S. (ed). Nutrição mineral de plantas. Sociedade Brasileira de Ciência do Solo, Viçosa, pp 53-85, 2006.

Bolan, N. S. A critical review on the role of mycorrhizal fungi in the uptake of phosphorus by plants. Plant and Soil, 134:189-207, 1991.

Bressan, W.; Vasconcellos, C. A.; Purcino, A. A. C. Fungos micorrízicos e fósforo, no crescimento, nos teores de nutrientes e na produção do sorgo e soja consorciados. Pesquisa Agropecuária Brasileira, 36:315-323, 2001.

de Sousa S, Clark R, Mendes F, Oliveira A, Vasconcelos M, Parentoni S, Kochian L, Guimarães C, Magalhães J: A role for root morphology and related candidate genes in P acquisition efficiency in maize. Functional Plant Biology, 39:925-935, 2012.

Ferreira, D. F. "Análises estatísticas por meio do Sisvar para Windows versão 4.0."Reunião anual da região brasileira da sociedade internacional de biometria, 45:235, 2000.

Gill, M. A., Salim, M., Zia, M. S. Maize growth and uptake of phosphate and copper at different ambient phosphate concentrations. Soil Science and Plant Nutrition, 38:631-636, 1992.

Gutjahr, C.; Banba, M.; Croset, V.; An, K.; Miyao, A.; An, G.; Hirochika, H.; Imaizumi-Anraku, H.; Paszkowski, U. Arbuscular mycorrhiza-specific signaling in rice transcends the common symbiosis signaling pathway. The Plant Cell, 20:2989–3005, 2008.

Gutjahr, C.; Radovanovic, D.; Geoffroy, J.; Zhang, Q.; Siegler, H.; Chiapello, M.; Casieri, L.; An, K.; An, G.; Guiderdoni, E.; Kumar,C. S.; Sundaresan, V.; Harrison, M. J.; Paszkowski, U. The half-size ABC transporters STR1 and STR2 are indispensable for mycorrhizal arbuscule formation in rice. The Plant Journal, 69:906-920, 2012.

Haling, R. E.; Brown, L. K.; Bengough, A. G.; Young, I. M.; Hallett, P. D.; White, P. J.; George, T. S. Root hairs improve root penetration, root-soil contact, and phosphorus acquisition in soils of different strength. Journal of Experimental Botany, 64:3711-3721, 2013.

Hodge, A.; Robinson, D.; Fitter, H. An arbuscular mycorrhizal inoculum enhances root proliferation in , but not nitrogen capture from , nutrient-rich patches in soil. New Phytologist, 145:575-584, 2000.

Javot, H., Penmetsa, R. V., Terzaghi, N., Cook, D. R., & Harrison, M. J. A Medicago truncatula phosphate transporter indispensable for the arbuscular mycorrhizal symbiosis. Proceedings of National Academy of Sciences, 104:1720-1725, 2007.

Liu, Y., Mi, G., Chen, F., Zhang, J., Zhang, F. Rhizosphere effect and root growth of two maize (*Zea mays* L.) genotypes with contrasting P efficiency at low P availability. Plant Science, 167: 217–223, 2004.

McLaughlin, M. L.; James, T. R. Effects of surface-applied phosphorus and superphosphate on the solution chemistry and phytotoxicity of subsurface aluminium: sand/solution and soil experiments. Australian Journal of Agricultural Research, 42:859-874, 1991.

Mendes, F.F; Guimarães, L. J. M.; Souza, J. C.; Guimarães, P. E. O.; Magalhaes, J. V.; Garcia, A. A. F.; Parentoni, S. N.; Guimaraes, C. T. Genetic architecture of phosphorus use efficiency in tropical maize cultivated in a low-P soil. Crop Science, 54:1-9, 2014.

Miransari, M.; Bahrami, H. A.; Rejali, F.; Malakouti, M. J. Effects of soil compaction and arbuscular mycorrhiza on corn (*Zea mays* L.) nutrient uptake. Soil and Tillage Research, 103:282-290, 2009.

Moreira, F. S.; Siqueira, J. O. Microbiologia e bioquímica do solo. 2.ed. Lavras: UFLA, 726p, 2006.

Mollier, A.; Pellerin, S. Maize root system growth and development as influenced by phosphorus deficiency. Journal of Experimental Botany 50:487-497, 1999.

Mosse, B. Advances in the study of vesicular-arbuscular mycorrhizal. Annual Review of Phytopathology, 11:171-196, 1973.

Nagy, R.; Vasconcelos, M. J. V.; Zhao, S.; McElver, J.; Bruce, W.; Amrhein, N.; Raghothama, K. G.; Bucher, M. Differential regulation of five Pht1 phosphate transporters from maize (*Zea mays* L.). Plant Biology, 8:186-197, 2006.

Niu, Y. F.; Chai, R. S.; Jin, G. L.; Wang, H.; Tang, C. X.; Zhang, Y. S. Responses of root architecture development to low phosphorus availability: a review. Annals of Botany, 112:391-408, 2013.

Parentoni, S. N. ; Souza Jr, C. L.; Alves, V. M. C.; Gama, E. E. G.; Coelho, A. M.; Oliveira, A. C.; Guimarães, P. E. O.; Guimaraes, C. T.; Vasconcelos, M. J. V.; Pacheco, C. A.P.; Meirelles, W.F.; Magalhaes, J. V.; Guimaraes, L. J. M.; Silva, A. R.; Mendes, F.F.; Schaffert, R. E. Inheritance and breeding strategies for phosphorus efficiency in tropical maize (*Zea mays* L.). Maydica, 55:1-15, 2010.

Parniske, M. Arbuscular mycorrhiza: the mother of plant root endosymbioses. Nature Reviews. Microbiology, 6:763-675, 2008.

Paszkowski, U.; Jakovleva, L.; Boller, T. Maize mutants affected at distinct stages of the arbuscular mycorrhizal symbiosis. Plant Journal, 47:165-173, 2006.

Quan, C. Genetic variation of phosphate transporter (PHT1) genes in tolerance to low Pi-fertilization and AM responsiveness in Maize (*Zea mays*). 47p. Universite de Lausanne, Faculte de Biologie et de Medecine, 2014.

Rai, A.; Rai, S.; Rakshit, A. Mycorrhiza-mediated phosphorus use efficiency in plants, Environmental and Experimental Biology, 11:107-117, 2013.

Redecker, D.; Kodner, R.; Graham, L. E. Glomalean fungi from the ordovician. Science, 289:1920-1930, 2000.

Remy, W.; Taylor, T. N.; Hass, H.; Kerp, H. Four hundred-million-yearold vesicular arbuscular mycorrhizae. Proceedings of the National Academy of Sciences USA, 91:11841-11843, 1994.

Richardson, A. E.; Barea J-M.; McNeill A. M.; Prigent-Combaret C. Acquisition of phosphorus and nitrogen in the rhizosphere and plant growth promotion by microorganisms. Plant and Soil, 321:305–339, 2009.

Siddiqui, Z.; Pichtel, J. Mycorrhizae: sustainable agriculture and forestry, 385p, 2008.

Sieverding, E.; Friedrichsen, J.; Suden, W. Vesicular-arbuscular mycorrhiza management in tropical agrosystems. 371p, 1991.

Smith, S. E.; Read, D. J. Mycorrhizal Symbiosis (Third Edition), New York, 787p, 2008.

Sokolski, S.; Dalpé, Y.; Séguin, S.; Khasa, D.; Lévesque, C. A.; Piché, Y. Conspecificity of DAOM 197198, the model arbuscular mycorrhizal fungus, with Glomus irregulare : molecular evidence with three protein-encoding genes. Botany, 88:829-838, 2010.

Verrier, P. J.; Bird, D.; Burla B.; Dassa, E.; Forestier, C.; Geisler, M.; Klein, M.; Kolukisaoglu, U.; Lee, Y.; Martinoia, E.; Murphy, A.; Rea, P. A.; Samuels, L.; Schulz, B.; Spalding, E. P.; Yazaki, K.; Thedoulou, F. L. Plant ABC proteins – a unified nomenclature and updated inventory. Trends in Plant Science, 13:151-159, 2008.

Williamson, L. C.; Ribrioux, S. P.; Fitter, A. H.; Leyser, H. M. Phosphate availability regulates root system architecture in Arabidopsis. Plant Physiology, 126:875-882, 2001.

Wu, Q.-S.; Zou, Y.-N.; He, X.-H. Contributions of arbuscular mycorrhizal fungi to growth, photosynthesis, root morphology and ionic balance of citrus seedlings under salt stress. Acta Physiologiae Plantarum, 32:297-304, 2010.

Yang, S.-Y.; Paszkowski, U. Phosphate import at the arbuscule: just a nutrient? Molecular Plant-Microbe Interactions : Molecular Plant-Microbe Interactions, 24:1296-1299, 2011.

Zhang, Q.; Blaylock, L.; Harrison, M. Two half-ABC transporters are essential for arbuscule development in arbuscular mycorrhizal symbiosis. Plant Cell, 22:1483–1497, 2010.

Zhu, J.; Lynch, J. The contribution of lateral rooting to phosphorus acquisition efficiency in maize (*Zea mays*) seedlings. Functional Plant Biology, 31:949-958, 2004.

MATERIAL SUPLEMENTAR

```
STR1_Maize      MHRQQQQQHEVDLATGARRPSRGHANGGHRAAAEMPAGGH--------------------
STR1_Sorghum    MHR--QQQHEV--TTGARRPGREHANGAHRAAAEMPAGGHRAERVERPAGHRTERAEMPG
STR1_Rice       MQ---QQQRSREMAR-AGEPAAAHRHRAERAAAGEPA----------------------
STR1_Medicago   MA------------------------------------------------------------
                *

STR1_Maize      --------RPERPAGHRMDRAETRAT--HH----------------------HHHHATER
STR1_Sorghum    AGGHHRAERAERPAGHRVERAEMPVAGGHHRAERAERPAGHRVERAEMQHATHHHHATER
STR1_Rice       ------------------------------------------------------TRTHRTER
STR1_Medicago   ------------------------------------------------------------RLER
                                                                            **

STR1_Maize      A-GE-GRRTT---AAAEMPAWQ-TERKKSLESLLDAAGDARGRPQHQ--HQHHHRGSGG-
STR1_Sorghum    APGD-ARRTTTAAAAAEMPAWQTTERKKSLESLLGAAGDARGWQQQQQHHHHHHRAGGGD
STR1_Rice       A-GEPAAATTT--TTTRRPPPTTTERKESLESLLDATDAARG----------GRRGGG--
STR1_Medicago   D-GT----------------------NKSLESLMDS-----------------HKPGGTT
                  *                       ::*****:.:                 :: .*

STR1_Maize      -----HVPVRPVPVPGVKVINFPGQGLEFKELSYSVIKKQKKDGVKVKKEVYLLNDISGQ
STR1_Sorghum    GGGGHHVPVRPVPVPGEKVINFPGQGLEFKELSYSVIKKQKKDGVKVKKEVYLLNDISGQ
STR1_Rice       ---------------GVKAAVASRQGLEFKNLSYSVVKKQKKEGVKVKKEVYLLNDISGE
STR1_Medicago   TNLNQLRTQKSI----------PGYGLEFTNLSYSIIKKQKKDGVWINKETYLLHDISGQ
                                  .  ****.:****::*****:** ::**.***:****:

STR1_Maize      ALRGQVTAILGPSGAGKSTFLDALAGRIAKGSLEGSVSIDGRSVTTSYMKQISSYVMQDD
STR1_Sorghum    ALRGQVTAILGPSGAGKSTFLDALAGRIAKGSLEGSVSIDGRSVTTSYMKQISSYVMQDD
STR1_Rice       APRGQVTAILGPSGAGKSTFLDALAGRIAKGSLEGSVRIDGRAVTTSYMKQISSYVMQDD
STR1_Medicago   AIKGEIMAIMGPSGAGKSTFLDALAGRIAKGSLQGSVRIDGKPVTTSYMKMVSSYVMQDD
                * :*:: **:**********************:*** ***:.******* :********

STR1_Maize      QLFPMLTVLETLRFAAEVRLPPSLSRAEKLNRVWELIEQLGLQTTAHTYIGDEGIRGVSG
STR1_Sorghum    QLFPMLTVLETLRFAAEVRLPPSLSRAEKLKRVWELIEQLGLQTTAHTYIGDEGIRGVSG
STR1_Rice       QLFPMLTVLETLTFAAEVRLPPSLSRAEKLKRVWELIDQLGLQTTAHTYIGDEGTRGVSG
STR1_Medicago   QLFPMLTVFETFMFAAEVRLPPSISRDEKKKRVHELLNKLGLQSATHTYIGDEGRRGVSG
                ********:**: **********:** ** :** **:::****:::******** *****

STR1_Maize      GERRRVSIGTDIIHKPSLLFLDEPTSGLDSTSAYSVVEKVKEIAKGGSIVLMTIHQPSFR
STR1_Sorghum    GERRRVSIGTDIIHKPSLLFLDEPTSGLDSTSAYSVVEKVKEIAKGGSIVLMTIHQPSFR
STR1_Rice       GERRRVSIGIDIIHKPSLLFLDEPTSGLDSTSAHSVVEKVKDIARGGSIVLMTIHQPSFR
STR1_Medicago   GERRRVSIGIEIIHKPSLLFLDEPTSGLDSTSAYSVVEKIKDIAQGGSIVLMTIHQPSFR
                ********* :**********************:*****:*:**:***************

STR1_Maize      IQMLLDRIVILARGRLIYLGTPLTLPTHLAGFGRPVPDGENSIEYLLDVIKEYDESTLGL
STR1_Sorghum    IQMLLDRIVILARGRLIYLGSPITLPAHLAGFGRPVPDGENSIEYLLDVIKEYDESTLGL
STR1_Rice       IQMLLDRLVILARGRLIYLGSPSTLQTHLAGFGRPVPDGENSIEYLLDVIKEYDESTSGL
STR1_Medicago   IQMLLDKITILARGRLIYMGRPDALHTHLSGFGRPVPDGENNIEYLLDVITEYDQATVGL
                ******::.*********:* * :* :**:***********.********.***::* **

STR1_Maize      EPLVAYQRDGSKPDEAAKTPIPKTPRTPYQKSVQFRQ--IQLKSNQFSVTTGTPHA-NPM
STR1_Sorghum    DPLVAYQRDGSKPDEAAKTPIPKTPRTPHQKSVQFRQ--MQLKSNQFSAATGTPHATNTF
STR1_Rice       EPLVAYQRDGTKPDGAAKTPVPRTPRTPHQKSVQFRQ--IQLKSNQFSLNSGAANG-NTF
STR1_Medicago   DPLVQYQHDGHKPDPAAMTPVPKPPRTPYRRNTPASKHMISLRSQGFTAGTPQPDS----
                :*** **:** *** ** **:*:.****::::..   :  :.*:*: *:  :  ...

STR1_Maize      SNFE-SYNI----DDEDEEFDNSLERKS-HTPLHTG--TSTYHPRLASQFY-----KDFS
STR1_Sorghum    SNFE-SYNI----DDEEEEFDNSLERKS-HTPLHTG--TSTYHPRLASQFY-----KDFS
STR1_Rice       SNFESSYNVDGGGDDDDEDFDNSLERKL-QTPMHAGGPASGYQPRLASQFY-----KDFS
STR1_Medicago   ----SQFGLDDDDNDDDENFDNSLERRSVQTSRNIV--TSGVYPRLASQFYQDFSAKDFS
                     .:.:    :*::*:*******:  :*. :    :*   ********     ****

STR1_Maize      VWVYHGVTG-TPHRKPTW-TPARTPARTPVSSYQQRSR-VA---TPQH-----APPPQSP
STR1_Sorghum    VWVYHGVTG-TPHRKPTWSTPARTPARTPMSSYHQRSR-FA---TPQH-----APPPQSP
STR1_Rice       VWVYHGVTGSTPHRRPTW-----TPARTPVSSFQRGRA-VTMTPTPQNNPQRRPPPPPSP
STR1_Medicago   VWLYNGVVG-TPRRPPSW-TPARTPGWTPGKTPLSGPRSFV---SNQH-----SASYQDP
                **:*:**.* **:* *:*     **. ** .:       ..   : *:     ...  .*
```

```
STR1_Maize      R-------------EPVFKPEEPSYHEYQLELE-PLDAPEDG------PKFANPWFREVV
STR1_Sorghum    H-------------EPVFKPEEPSYHEYQLELE-PLDAPEDG------PKFANPWLREVA
STR1_Rice       H-------------VPVFKPEEPTYHEYELDLEPPLDAPEEDYNGGHRPKFANPWPREVA
STR1_Medicago   YYIQKTNTVVGQSMDYSATSYAPSYEEFEIE-E-VLDEPDLG------PKYANPWLREVA
                                ..  *:*.*:::: *  ** *: .      **:**** ***.

STR1_Maize      VLSWRTALNVVRTPELFLSREIVLAVMALILSTLFHRLSDANFTTINRLLNFYIFAVCLV
STR1_Sorghum    VLSWRTALNVVRTPELFLSREIVLTVMALILSTLFHRLSDSNFITINRLLNFYIFAVCLV
STR1_Rice       VLSWRTVLNVVRTPELFLSREVVLAAMAVILSTMFRRLGAGDVPTVNRLLNFYIFAVCLV
STR1_Medicago   VLSWRTVLNVIRTPELFASREIVLTVMALVLSTIFKNLGDTTFIDINRLLNFYIFAVCLV
                ******.***:****** ***:**:.**::***:*:.*.   .  :**************

STR1_Maize      FFSSNDAVPTFIQERFIFIRERSHNAYRASSYVVSSLIVYLPFFAIQGFTFAVITKFMLH
STR1_Sorghum    FFSSNDAVPTFIQERFIFIRERSHNAYRASSYVISSLIVYLPFFAIQGFTFAVITKYMLH
STR1_Rice       FFSSNDAVPTFIQERFIFIRERSHNAYRASSYVVASLVVYLPFFAVQGLTFAVITKLMLR
STR1_Medicago   FFSSNDAVPSFIMERFIFIRETSHNAYRASSYVISSLIVYLPFFAVQGLTFAVITKLMLH
                *********:** ******** ***********::**:*******:**:******* **:

STR1_Maize      LQSSLVNFWIILFASLITTNAYVMLVSALVPSYITGYAVVIATTALFFITCGFFLKRTMI
STR1_Sorghum    LHSNLVNFWIILFASLITTNAYVMLVSALVPSYITGYAVVIATTALFFITCGFFLKRTKI
STR1_Rice       MESSLLHFWVILFASLITTNAYVMLVSALVPSYITGYAVVIATTALFFLTCGFFLKRTLI
STR1_Medicago   LKSNLFNFWMILFASLITTNAYVMLVSALVPSYITGYAVVIATTALFFLTCGFFLKRTQI
                :.*.*.:**:**************************************:********* *

STR1_Maize      PMAWRWLHYISAIKYPFEALLVNEFKGSE-CYVGTQNQLSPGPLGQVS------NLNSTS
STR1_Sorghum    PMAWRWLHYISAIKYPFEALLVNEFKGDH-CYVGTQNQLSPGPLGQVS------NLNATS
STR1_Rice       PVGWRWLHYASAIKYPFEALLVSEFKGGR-CYAGDRADLSPGPLGGFKPSSLRRELNASD
STR1_Medicago   PAYWKWLHYISAIKYPFEGLLINEFKNNRGCYSGNKADLSPGPLGDVKPS---KHHNASL
                *  *:**** ********.**:.***... ** * : :******* ..      . *::

STR1_Maize      TTCPLVGQDVLSTMDIAIDNIWIDVAILLAWGVLYRLVFYVVLRFYSKNERK
STR1_Sorghum    TTCPLVGQDVLDTMDISIDNIWIDVAILLAWGVLYRLIFYVVLRFYSKNERK
STR1_Rice       AACPLMGQDVLSTLDITIDSIWVDVAILLAWGVLYRLLFYVVLRFYSKNERK
STR1_Medicago   PLNCLLGEDVLSTMDITMESLWYDILILLAWGVLYRFFFYLVLRFYSKNERK
               .   *:*:***.*:**::::.:* *: **********:.**:***********
```

* = resíduos de aminoácidos idênticos em todas as sequências
: = resíduos altamente conservados entre as sequências
. = resíduos moderadamente conservados entre as sequências
branco = aminoácidos distintos entre as sequências ou *gaps*.

Figura suplementar 1. Alinhamento do STR1 de milho, sorgo, arroz e *Medicago truncatula*. O domínio ABC é destacado em amarelo.

```
STR2_Maize      MLVRSIQAACLQIKLIGDHTGAYVHSMAHQHHHQHNGRHGGGHRLETVIDMAQQQPEKPA
STR2_Sorghum    ---------------------------MAHHNHHQHNGRHGGGHRLETVINMAPPPPEKPA
STR2_Rice       ---------------------------MAHRHVGDLEAAAYGG------------------
STR2_Medicago   ---------------------------MKTQ-----------GLELETVIDIKHKP-----
                                            *  :           *

STR2_Maize      GVGRFGFTGGLEFTGLTYTVVKKQRGAGGEWEKKDVDLLHEITGYAPKGCVTAVMGPSGA
STR2_Sorghum    GVGRFGFTGGLEFTSLTYTVVKKQRGAGGEWEKKDVDLLHEITGYAPKGCVTAVMGPSGA
STR2_Rice       --GRFGFTGGLEFTGLTYTVTKKQRGAGGEWEKKEVDLLHEVTGYAPKGCVTAVMGPSGA
STR2_Medicago   ----VSFTGGLEFESLTYTVTKKKK-VDGKWSNEDVDLLHDITGYAPKGCITAVMGPSGA
                    ..******* .*****.**:: ..*:*.:::*****::********:*********

STR2_Maize      GKSTFLDALAGRIAS--LDGRVALDGVEMSPSLIKRSSAYVMQDDRLFPMLTVYETLMFA
STR2_Sorghum    GKSTFLDALAGRIAS--LDGRVALDGVEMSPSLIKRSSAYVMQDDRLFPMLTVYETLMFA
STR2_Rice       GKSTLLDALAGRIAA-RLGGRVALDGVEVSPGLVKRCSAYVMQDDRLFPMLTVRETLMFA
STR2_Medicago   GKSTLLDGLAGRIASGSLKGKVSLDGNSVNASLIKRTSAYIMQEDRLFPMLTVYETLMFA
                ****:**.******:  * *:*:*** .:...*:** ***:**:********* ******

STR2_Maize      ADFRLGSSVSASDKKLRVENLIEQLGLTSSRNTYIGDEGTRGVSGGERRRVSIGVDIIHG
STR2_Sorghum    ADFRLGSSVSASDKKLRVENLIEQLGLTSSRNTYIGDEGTRGVSGGERRRVSIGVDIIHG
STR2_Rice       ADLRLGASVPAADKRRRVDALIDQLGLAASGNTYIGDEGTRGVSGGERRRVSIGVDIIHG
STR2_Medicago   ADFRLG-PLSAVDKRQRVEKLIEQLGLSSSRNTYIGDEGTRGVSGGERRRVSIGVDIIHG
                **:*** .:.* **: **: **:****::* ****************************

STR2_Maize      PALLFLDEPTSGLDSTSAHSVIEKVHDIACAGSTVVLTIHQPSSRILLLLDHLIILARGQ
STR2_Sorghum    PALLFLDEPTSGLDSTSAHSVIEKVHDIACAGSTVVLTIHQPSSRILLLLDHLIILARGQ
STR2_Rice       PALLFLDEPTSGLDSTSAHSVVEKVRDIACAGSTVVLTIHQPSSRILQLLDHLVILARGQ
STR2_Medicago   PSLLFLDEPTSGLDSTSALSVIEKLHDIARNGSTVILTIHQPSSRIQLLLDHLIILARGQ
                *:**************** **:**::***  ****:**********  *****:******

STR2_Maize      LMYSGGPKEVTAHLGRMGRKVPKGESSIEHLLDVIQEYEQSEF-GVKALAEFCLTGLKPP
STR2_Sorghum    LMYSGGPKEVTAHLGRMGRKVPKGENSIENLLDVIQEYEQSEF-GVKALAEFCLTGLKPP
STR2_Rice       LMYSGAPREVAAHL----------------------------------------------
STR2_Medicago   LMFQGSLKDVGHHLNRMGRKIPKGENPIENLIDVIQEYDQCDFVGVEVLAEFARTGMKPP
                **:.*. ::*  **

STR2_Maize      KLTATYGAEGLSTVSSIVQTPISGPGGGD-------------------DDDFDRSLRSQH
STR2_Sorghum    KLTATYGPEGLSTVSSIVQTPISGPGGGD-------------------DDDFDRSLRSQH
STR2_Rice       ------------------------------------------------------------
STR2_Medicago   LLSDM--EEIISYTNSIAPSPSPLHRGSKYEEKSQDFSYSSQISRRSLNDEFDHSIRSPY

STR2_Maize      SKSP--WSG------VRLTPSRRPRHKDQQHGRPGPGPGPGNHRYTPDIVMGTPTPLSSA
STR2_Sorghum    SKSP--WSG------VRLTPSRRPKHK-------GPGAPHDNHRYTPEIVMGTPTPMSSA
STR2_Rice       ------------------------------------------------------------
STR2_Medicago   NNTPMSWSASNSAAFLKFTPSRLKNENKVQ--KPPSHASPGIYTYSSEILPATPTPHSSD

STR2_Maize      YTVNEDDYLTPTHRAA-PNATGAPGVGVNALGHRG-KFANSYVGEVWVLMRRNFTNIWRT
STR2_Sorghum    YTVNEDDYLTPTHRAAGPNATGAPGMGVNALGHRG-KFANSYVVEVWVLMRRNFTNIWRT
STR2_Rice       ------------------------------------------------------------
STR2_Medicago   YVVDENDYLTPTNSSQ---------------EHLGPKFANSYIGETWILMRRNFTNIRRT

STR2_Maize      PELFLSRLMVLTVMGFLMATMFTKPKDNAQGITNRLSFFIFTVCVFFFSSNDAVPAFIQE
STR2_Sorghum    PELFLSRLMVLTVMGFLMATMFTKPKDNTQGITNRLSFFIFTVCVFFFSSNDAVPAFIQE
STR2_Rice       ------------------------------------------------------------
STR2_Medicago   PELFLSRLMVLTFMGVMMATMFHNPKNTLQGITNRLSFFIFTVCLFFFSSNDAVPAFIQE

STR2_Maize      RFIFIRETSHNAYRASAYVVAGLITYLPFLLLQSATYAAIVWFALKLHGQFLYFLVMLYA
STR2_Sorghum    RFIFIRETSHNAYRASAYVVAGIITYLPFLLLQSATYAAIVWFALRLHGQFLYFLVMLYA
STR2_Rice       ------------------------------------------------------------
STR2_Medicago   RFIFIRETSHNAYRASCYTIASLITHMPFLALQALAYAAIVWFALELRGPFIYFFLVLFI

STR2_Maize      SLLSTNSFVVFISSVVPNFILGYAAVIAFTALFFLFCGYFLDSHSIPVAWKWMNTVSTMK
STR2_Sorghum    ALLSTNSFVVFISSVVPNFILGYAAVIAFTALFFLFCGYFLDSHSIPVAWKWMNTISTMK
```

```
STR2_Rice       ------------------------------------------------------------
STR2_Medicago   SLLSTNSFVVFVSSIVPNYILGYAAVIAFTALFFLFCGYFLSSEDIPLYWRWMNKVSTMT


STR2_Maize      YPYEGLLMNEFDGDRVFASDPAVGLTLTGNNILQQLGISVEEDRKWWMVLYLLGWAVFYR
STR2_Sorghum    YPYEGLLMNEFNGGHVFASERAIGLTLTGNDILKQLGISTEEDRKWWMVLYLLGWAVFYR
STR2_Rice       ------------------------------------------------------------
STR2_Medicago   YPYEGLLMNEYQTNETFGSND--GVSITGFDILKSLHIGTEEIKKRNNVLIMLGWAVLYR


STR2_Maize      VLFYLVLRFASKNKRK
STR2_Sorghum    VLFYLVLRFASKNKRK
STR2_Rice       ----------------
STR2_Medicago   ILFYIILRFASKNQRS
```

* = resíduos de aminoácidos idênticos em todas as sequências
: = resíduos altamente conservados entre as sequências
. = resíduos moderadamente conservados entre as sequências
branco = aminoácidos distintos entre as sequências ou *gaps*.

Figura suplementar 2. Alinhamento do STR2 de milho, sorgo, arroz e *Medicago truncatula*. O domínio ABC é destacado em amarelo.

Figura suplementar 3. Raízes de milho coletadas a 40 dias após o transplantio. (A) L3 inoculada com *R. irregularis*; (B) L3 sem inoculação; (C) L22 inoculada com *R. irregularis* e (D) L22 sem inoculação.

CONCLUSÃO GERAL E PERSPECTIVAS FUTURAS

Utilizando uma abordagem de genômica comparativa que combinou mapeamento de QTLs de características relacionadas com a eficiência no uso do P, análises de bioinformática e de expressão gênica, foram identificados no genoma do milho seis genes candidatos a homólogos ao gene *PSTOL1* de arroz. Os resultados apresentados no capítulo 1 indicam que pelo menos quatro desses genes podem estar relacionados com a morfologia radicular e com a aquisição de P em milho. Dentre eles, os genes candidatos *ZmPSTOL3.06*, *ZmPSTOL8.02* and *ZmPSTOL8.05_1* foram selecionados para estudos detalhados visando a validação funcional dos mesmos. A expressão temporal e espacial desses genes está sendo analizada nas linhagens L3 e L22. Adicionalmente, esses genes estão sendo transformados em milho e tabaco, cujas plantas transgênicas serão avaliadas quanto a parâmetros morfológicos do sistema radicular e conteúdo de P.

No capítulo 2, abordamos o efeito da colonização micorrízica no desenvolvimento das plantas, no conteúdo de P e na expressão de genes candidatos em duas linhagens de milho contrastantes para a eficiência na aquisição de P. Apesar da colonização micorrízica não justificar diferenças na eficiência no uso de P entre as linhagens, os genes *STR1* e *STR2* estão associados com o desenvolvimento dos arbúsculos, além de apresentarem evidências estruturais para serem candidato a homólogos aos genes de arroz e de *M. truncatula*. A validação desses genes em milho envolverá estudos funcionais em quatro linhagens mutantes de milho que possuem inserções na região do gene *STR2*. Outra estratégica será a busca por variações naturais para a colonização micorrízica em um painel de linhagens de milho da Embrapa.

ANEXO

Este anexo contém o artigo científico intitulado: "Duplicate and conquer: multiple homologs of PHOSPHORUS-STARVATION TOLERANCE1 enhance phosphorus acquisition and sorghum performance on low-phosphorus soils", publicado na Plant Physiology (2014), com co-autoria de Gabriel Corradi Azevedo.

Hufnagel B, de Sousa S M, Assis L, Guimarães C T, Leiser W, Azevedo G C, Negri B F, Larson B G, Shaff J E, Pastina M M, Barros B A, Weltzien E, Rattunde H F W, Viana J H, Clark R T, Falcão, Gazaffi R, Garcia A A F, Schaffert R E, Kochian L V, Magalhães, J V: **Duplicate and Conquer: Multiple Homologs of PHOSPHORUS-STARVATION TOLERANCE1 Enhance Phosphorus Acquisition and Sorghum Performance on Low-Phosphorus Soils**. *Plant Physiology*, 166:659-677, 2014.

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.

[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

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

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

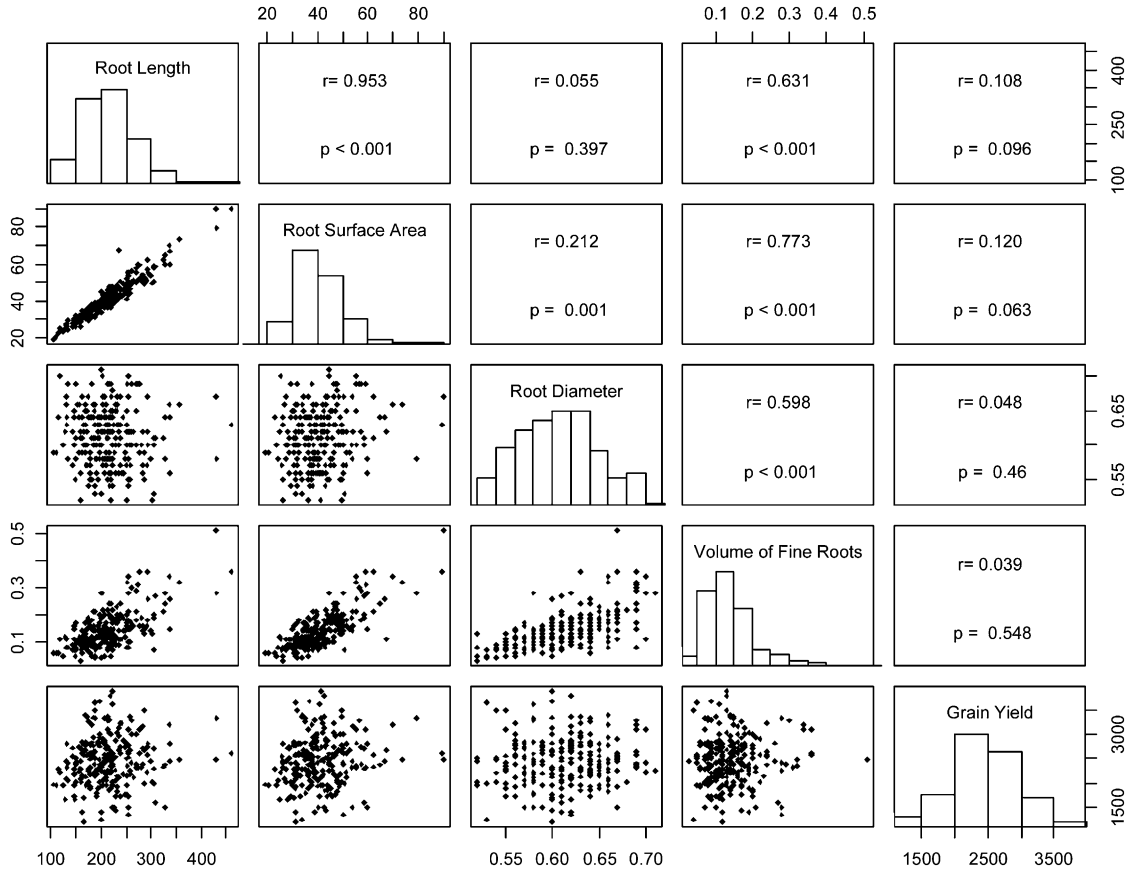


Figure 1. Correlation matrix of root length (mm), root surface area (cm^2), average root diameter (mm), and volume of fine roots (mm^3), all assessed in hydroponics-grown plants, as well as grain yield (kg ha^{-1}) for plants grown on a low-P soil for the SAPst. Correlation coefficients (*r*) and *P*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

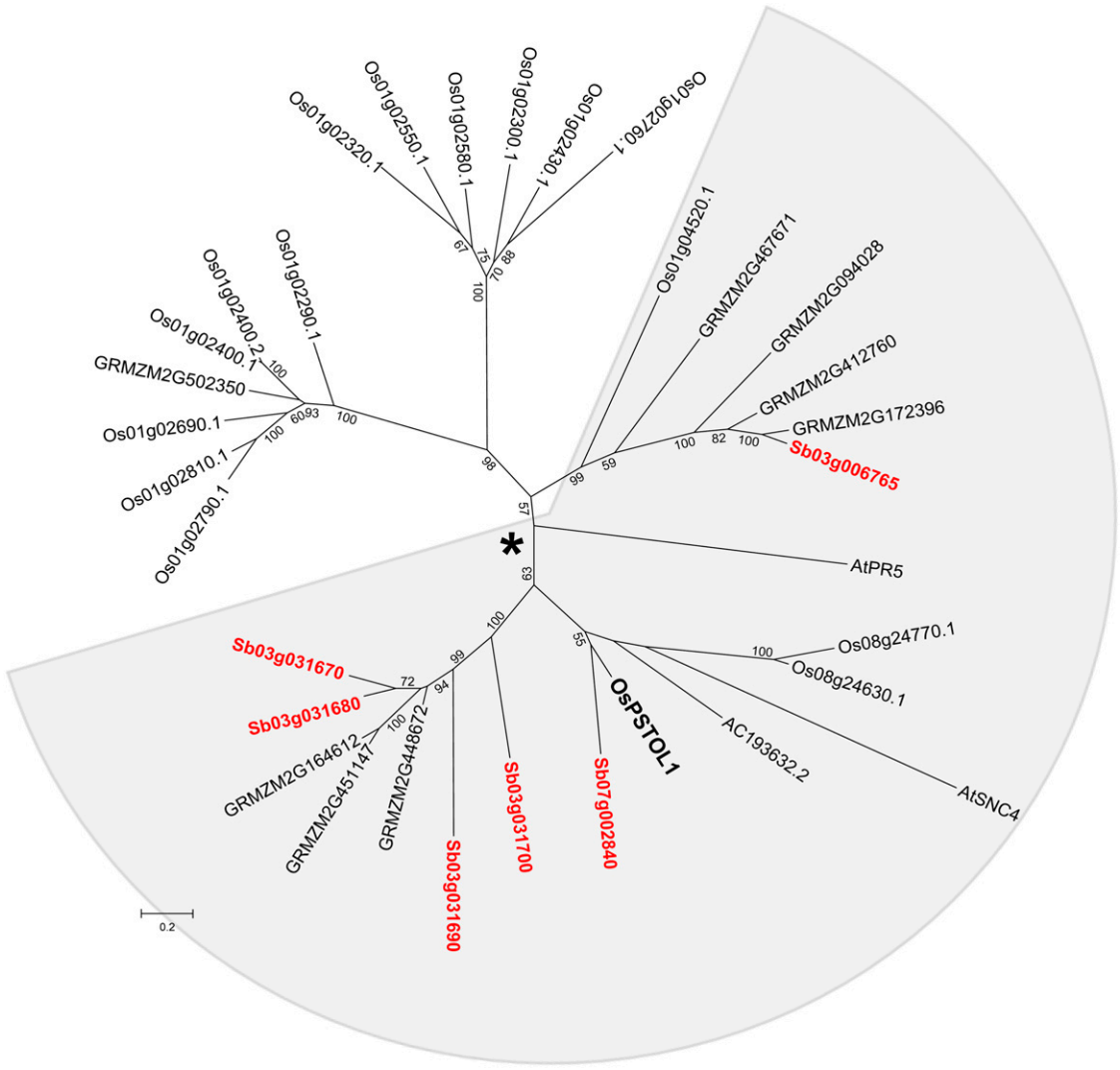


Figure 2. Relationship among *PSTOL1* homologs. The phylogenetic tree comprises PSTOL1-like proteins in sorghum (Sb), rice (LOC_Os), maize (GRMZM and AC), and Arabidopsis (At) in addition to OsPSTOL1. The highlighted clade comprises selected sorghum proteins with high identity with OsPSTOL1, and the 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

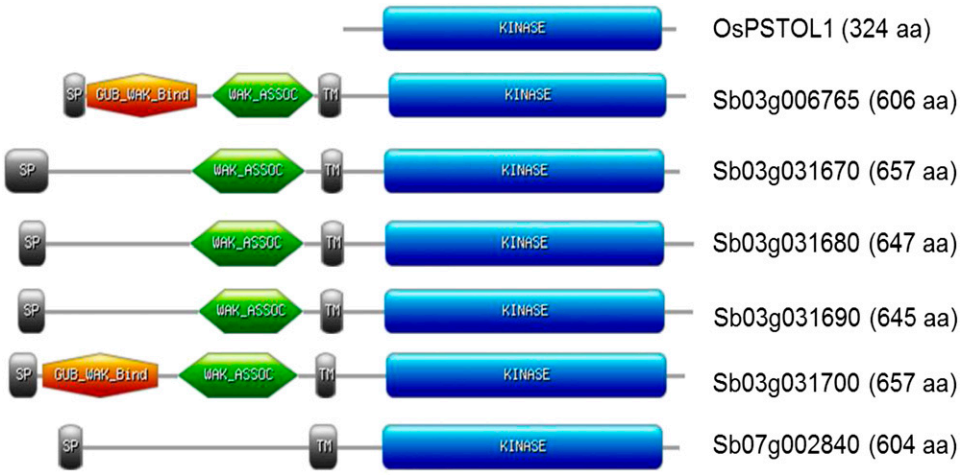


Figure 3. Schematic representation of protein domains of OsPSTOL1 and SbPSTOL1. Depicted are the predicted signal peptide (SP), the GUB_WAK_Bind domain, the wall-associated receptor kinase C-terminal domain (WAK_ASSOC), the transmembrane domain (TM), and the Ser/Thr kinase domain (KINASE). Full complementary 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**

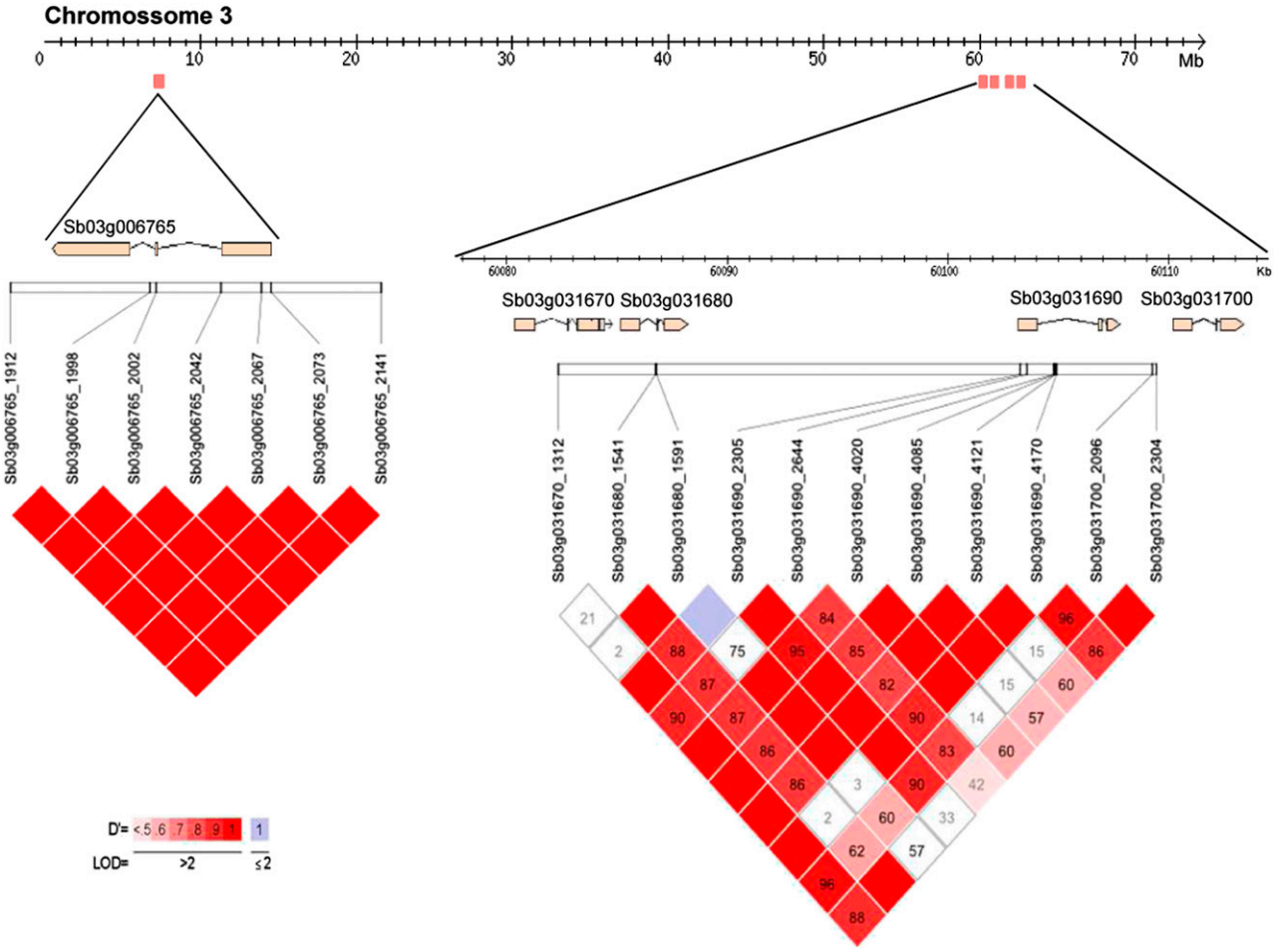


Figure 4. LD across *SbPSTOL1* homologs on chromosome 3. The standardized LD coefficient (D′; Lewontin, 1964) was estimated among all pairwise combinations between associated SNP loci within *Sb03g006765* (position 7 Mb) and the *SbPSTOL1* cluster at position 60 Mb. Squares shows D′ between two loci 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

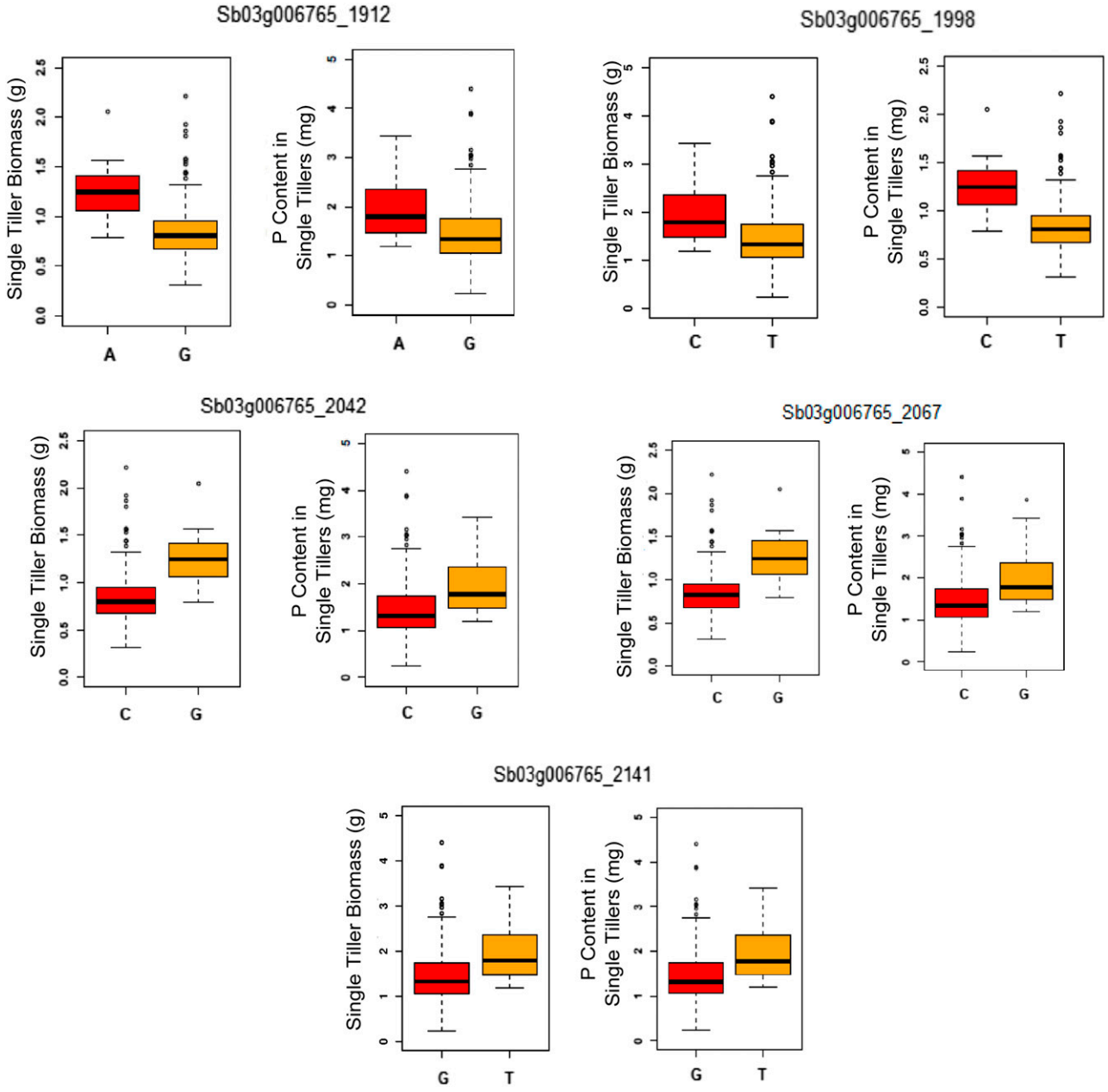


Figure 5. Effects for SNPs within *Sb03g006765* estimated in the WAP in Africa. Effects for single tiller biomass and P content in single tillers are shown. The *P* values from association analysis are shown in Table IV. [See online article for color version of this figure.]

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

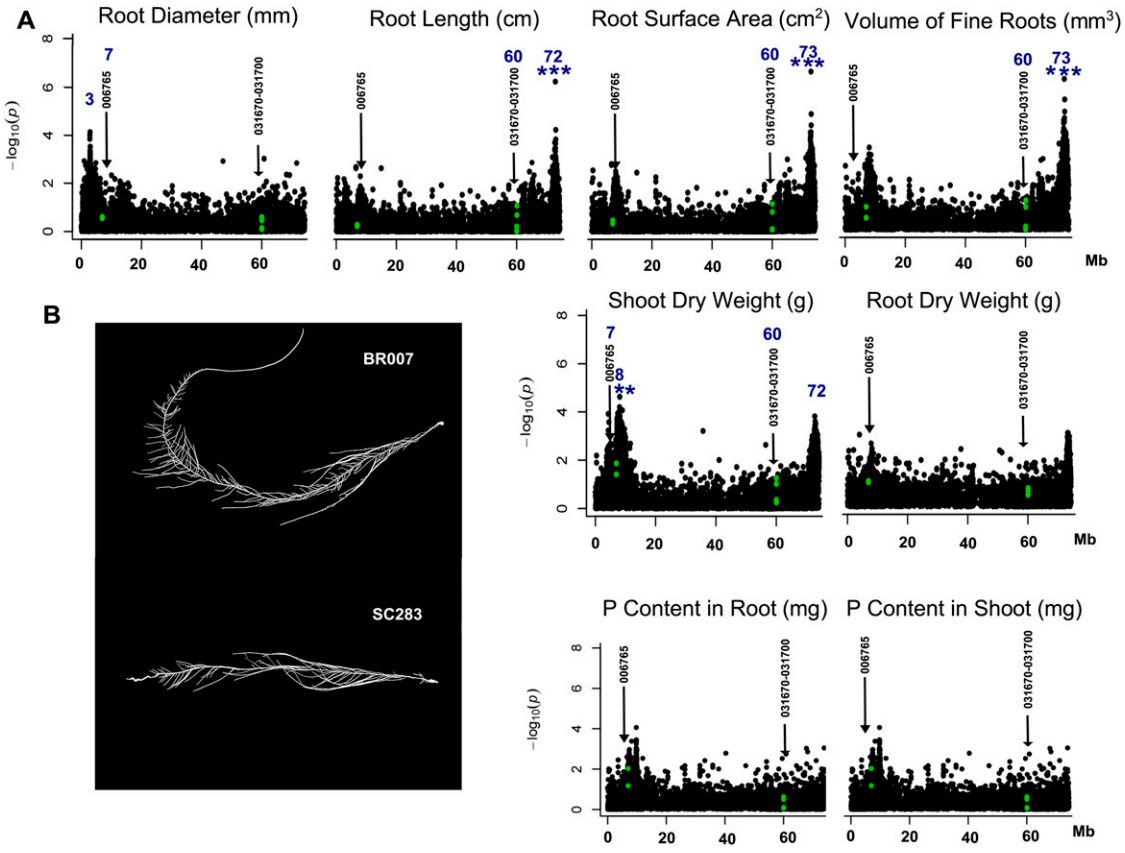


Figure 6. Validation of *PSTOL1* homologs on chromosome 3 by biparental mapping. A, QTL analysis of root morphology traits, shoot and root dry weight accumulation, and P content in the BR007 × SC283 RIL population. Physical positions on chromosome 3 are shown on the *x* axis (Mb). The positions for *Sb03g006765* (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 **	

**** = $p < 0.001$ *** = $0.001 \leq p < 0.01$ ** = $0.01 \leq p < 0.05$

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:

$$h^2 = \frac{\sigma^2_g}{\sigma^2_g + \frac{\sigma^2_\varepsilon}{r}}$$

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:

$$h^2 = \frac{\sigma^2_g}{\sigma^2_g + \frac{\sigma^2_{ge}}{e} + \frac{\sigma^2_\varepsilon}{er}}$$

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:

$$h^2 = \frac{\sigma^2_g}{\sigma^2_g + \frac{\sigma^2_\varepsilon}{r}}$$

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 < 1*e*-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 < 1*e*-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:

$$h^2 = \frac{\sigma^2_g}{\sigma^2 + \frac{\sigma^2_\varepsilon}{r}}$$

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

SUPPLEMENTAL FIGURE S1

```
PSTOL1          KGSLPDGREIAVKMLKDTKGDGEEFINEVAGISKTSHINVVNLLGFSLQGSKRALIYEYM 114
Sb07g002840     RGNLADGREIAVKTLKDTEGDGEDFMNEVASISRTSHVNIVTLLGFCLQGRKRALIYEYM 390
Sb03g031670     RGNLYDGRQVAVKMLKDTKGDGEEFMNEVSSISRTSHVNVVTLLGFCLQGSKRALIYEYM 444
Sb03g031680     KGKLYDGRQVAVKMLKDTNGDGEEFMNEVASISRTSHVNVVTLLGFCLQGSKRALIYEYM 432
Sb03g031690     RGSLSDGGQVAVKMLKDPKGDGEEFMNEVSSISRTSHVNVVMLIGFCLEGSKRALIYEYM 433
Sb03g031700     KGSLPDGRLVAVKMLKELKGDDEQFMNEVASISTTSHVNIVTLLGFCVQGSKRALVYDYM 441
Sb03g006765     RGRLDDGRLVAVKFLHDSKGDGEEFVNEVMSIGRTSHINIVSLFGFCLEGSKRALIYEYM 392
                :* * **  :*** *:: :**.*:*:*** .*. ***:*:* *:**.::* ****:*:**

PSTOL1          PNGSLDRYSFGDSSVQGDNTLSWDRLFNIIVGIARGLEYLHCHCNIRIVHFDIKPQNILL 174
Sb07g002840     PNGSLERYTFGSMSAEGDNSLCWDKLFEIVIGIARGLEYLHNGCNTRIVHFDIKPQNILL 450
Sb03g031670     PNGSLERYAFNS-NMNSENSLSWEKLFDIAIGTARGLEYLHRGCNTRIVHFDIKPHNILL 503
Sb03g031680     PNGSLERYAFNS-NMNCENSLSWEKLFDIAIGTARGLEYLHRGCNTRIVHFDIKPHNILL 491
Sb03g031690     PNGSLERYAFNS-NMNNQNSLGWEKLFDIAIGIARGLEYLHRGCNTRIVHFDIKPHNILL 492
Sb03g031700     TNGSLERFIFSR-HLEDKNSLSWGKLFEIVVGIARGLEYLHRGCKTRIVHFDIKPHNILL 500
Sb03g006765     PNGSLDKYIYSE---NPKAILGWDKLYTIAIGIARGLEYLHHSCNTRIVHFDIKPPNILL 449
                .****::: :.    : .  * * :*: * :* ********  *: ********* ****

PSTOL1          AQDFCPKISDFGLSKLCHLKESRISINGLRGTPGYIAPEVFSRQYGSASSKSDVYSYGMV 234
Sb07g002840     DQNFCPKISDFGLAKLCQQKQSKISMVGMRGTIGYIAPEVFNRSYGAVSNKSDVYSYGMM 510
Sb03g031670     DQDFCPKISDFGLAKLCLNKESAISIAGARGTIGYIAPEVYSKQFGLVSSKSDVYSYGMM 563
Sb03g031680     DLDFCPKISDFGLAKLCLNKESAISIVGARGTIGYIAPEVYSKQFGTVSSKSDVYSYGMM 551
Sb03g031690     DHDFCPKISDFGLAKLCPNKESAISIAGARGTIGYIAPEVFSKQFGTVSSKSDVYSYGMM 552
Sb03g031700     DENFCPKISDFGLAKLSVQKESAISIGVARGTIGYIAPEVFSRQFGVVTSKSDVYSYGMM 560
Sb03g006765     DQNFHPKIADFGLAKLCRSKESKLSMTGARGTPGFIAPEVHSRTFGVVSTKSDVYSYGMM 509
                  :* ***:****:**.  *:* :*:   *** *:*****..: :* .:.*********:

PSTOL1          VLEMAGAKKNI--NVSTGSSSK-YFPQWLYDNLDQFCCPTG-EISSQTTDLVRKMVVVGL 290
Sb07g002840     VLEMVGARKQI--DVGIDTSSN-YFPRWLYDKLDQFCGATISEIGSDTTELVRKMIIVGL 567
Sb03g031670     VLEMVGARDKNNTSASVESSSQ-YFPQWLYEHLDEYCISAS-EINGETTELVRKMIVVGL 621
Sb03g031680     VLEMVGARDKN-TSGNSESSSQ-YFPQWIYEHLDDYCISAS-EIDGETTELVRKMIVVGL 608
Sb03g031690     VLEMVGARGKN-MNADSESSSQ-YFPQWIYEHLDDYCISAS-EVNSETTELVRKMIVVGL 609
Sb03g031700     ILEMVGSTRTTNNNTNCESTDELYFPLWIYDNLDQYCLDAS-ETSKDDGEVVRKMIVVGL 619
Sb03g006765     LLEMVGGRKNV-KSLAQESSEK-YFPHWIYDHFGQDDGLPACEVTSENEGIAKKMSVIGL 567
                :***.*.      .    ::.: *** *:*:::.:    .  *   :   :.:** ::**

PSTOL1          WCIQLVPTDRPSMREVLEMLESNGRDLPLPPKGL----- 324
Sb07g002840     WCIQLRPIDRPSMSKVLEMLESNTSDLQLPPKAFWTS-- 604
Sb03g031670     WCIQVIPTDRPTMTRVVEMLEGGTSNLELPPKVLLS--- 657
Sb03g031680     WCIQLIPTDRPTMTRVVEMLEGSTSNLELPPKVLLSYQA 647
Sb03g031690     WCIQVIPTDRPTMTKVVEMLEGSTSNLELPPKVIQI--- 645
Sb03g031700     WCVQVMPIDRPSMSQVVEMLESDLKDLQLPSKPLTLSS- 657
Sb03g006765     WCIQILPMHRPTIGKVLEMFDRGLDELDMPPRQNFSEIL 606
                **:*: * .**:: .*:**:: .  :* :*.:
```

```
* = identical amino acid residues in all sequences
: = highly conserved amino acids
. = different but somewhat similar amino acids
blank = dissimilar amino acids or gaps
```

Signal Peptide

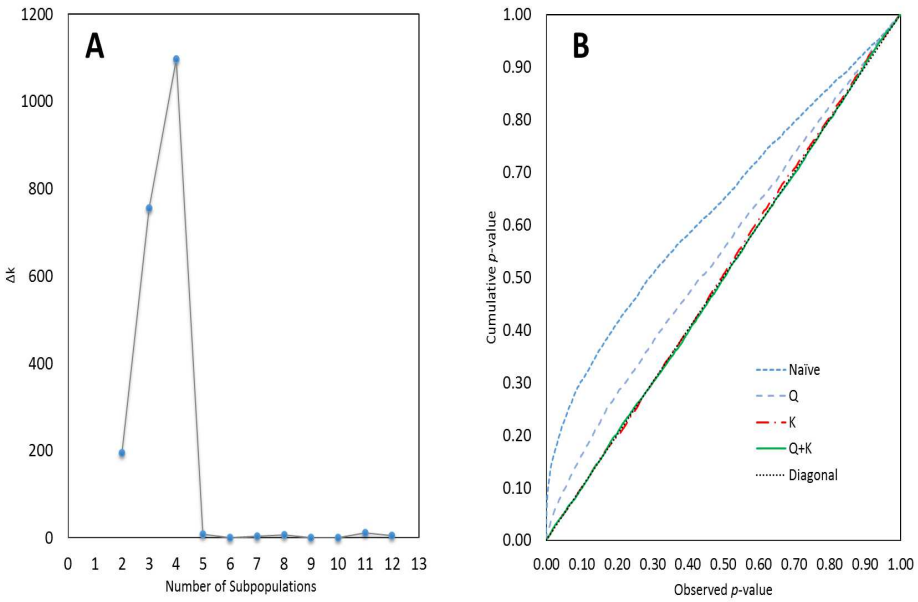
GUB-WAK-bind Domain – Wall-associated receptor kinase galacturonan-binding (GUB-WAK-Bind)
WAK ASSOC Domain – Wall-associated kinase C-Terminal (WAK_ASSOC)

Transmembrane Domain

Kinase Domain

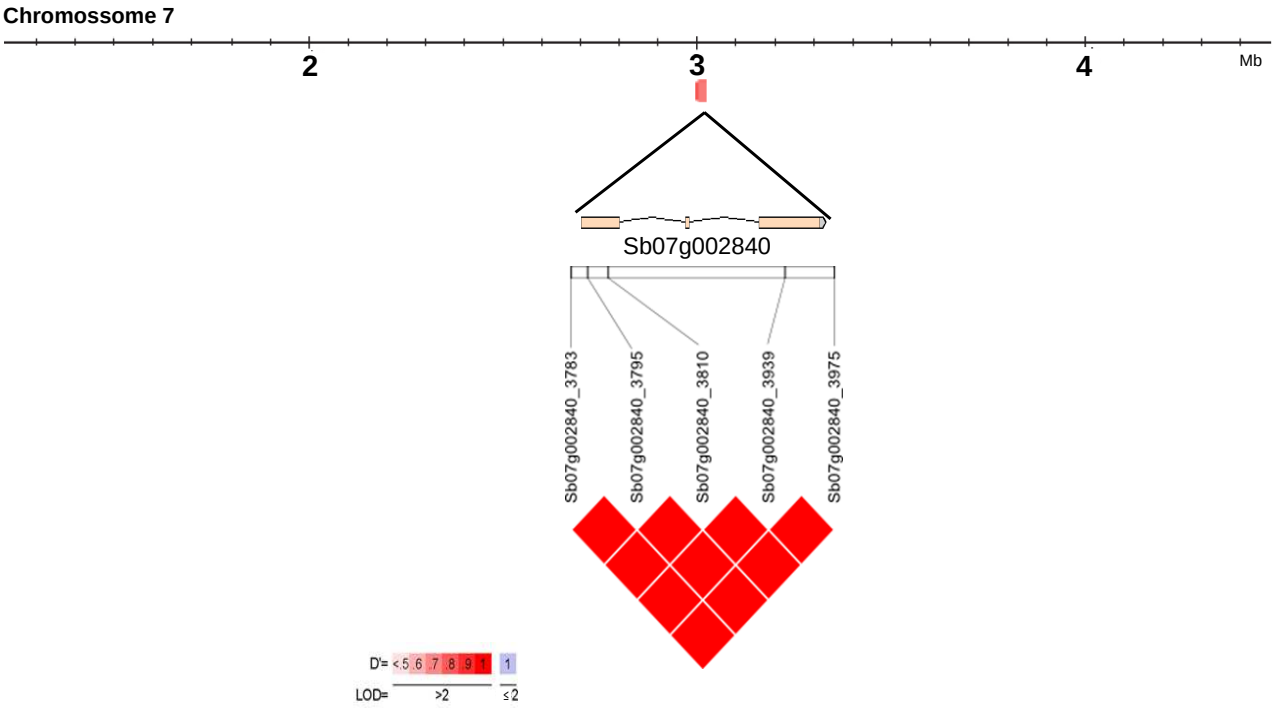
Supplemental Figure S1: Multiple sequence alignment of OsPSTOL1 and SbPSTOL1 proteins depicting structural predictions.

SUPPLEMENTAL FIGURE S2



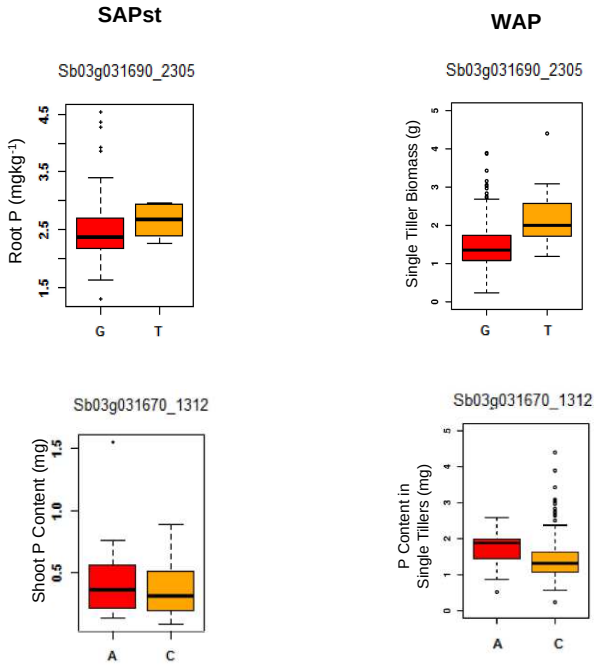
Supplemental Figure S2: Association model selection. (**A**) Second order change in the log probability of data (Δk) across different numbers of subpopulations (Evanno et al., 2005). The log likelihood of data was plotted as a function of successive number of subpopulations. (**B**) Type I error simulation with different models based on 310 SNP markers randomly distributed across the sorghum genome. Cumulative distribution of *p-values* is shown for the naïve, Q (population structure), K (kinship) and Q+K models. Methods that appropriately control for Type I error should show a uniform distribution of *p*-values near the diagonal line.

SUPPLEMENTAL FIGURE S3



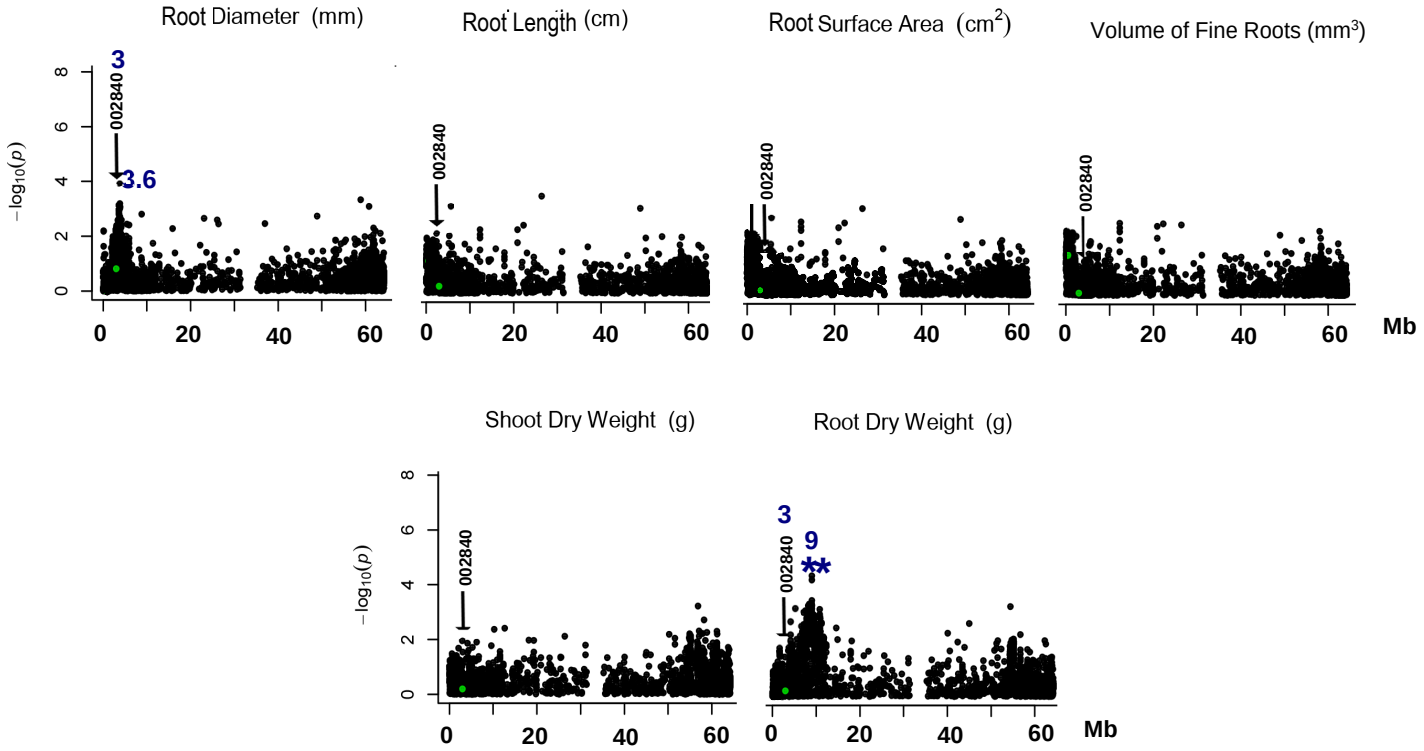
Supplemental Figure S3: Linkage disequilibrium across the *SbPSTOL1* gene on chromosome 7. The standardized linkage disequilibrium coefficient (D', Lewontin, 1964) was estimated among all pairwise combinations between associated SNP loci within *Sb07g002840*. Squares shows D' between two loci and the 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 1.0)**.** LD plots were generated using Haploview v4.2 (Barrett et al., 2005).

SUPPLEMENTAL FIGURE S4



Supplemental Figure S4: Allele effects of SNPs *Sb03g031690*_2305 and *Sb03g031670*_1312. Effects on root P concentration and shoot P content (SAPst), and single tiller biomass and P content in single tillers (WAP) are shown. The *p*-values from association analysis are shown in Tables II and IV.

SUPPLEMENTAL FIGURE S5



Supplemental Figure S5: Validation of *PSTOL1* homologs on chromosome 7 by bi-parental mapping. QTL analysis of root morphology traits, shoot and root dry weight accumulation in the BR007 x SC283 RIL population. Physical positions on chromosome 7 are shown on the x-axis (Mb). The positions for *Sb07g002840* (depicted as 002840 in the graphs) in the beginning of chromosome 7 are shown. Blue asterisks denote significant QTL declared based on permutation tests at $0.01 \leq p < 0.05$ (**). Numbers in blue indicate the physical positions (Mb) of SNPs at the $-\log_{10}(p)$ peak within the different QTL and the closest *SbPSTOL1* genes. The SNPs that are closest to *SbPSTOL1* homologs are highlighted in green.

Supplemental Table SI: Correlation matrix for traits assessed in hydropinically-grown plants. The Pearson Correlations among the 10 traits measured in hydroponics were estimated using Hmisc package (http://biostat.mc.vanderbilt.edu/wiki/Main/Hmisc) on R. The correlation coefficents (top value) and *p*-values (bottom value) are shown for each pair of traits.

Root Length	(cm)	**Root Length**									
Root Diameter	(mm)	-0.05 *0.397*	**Root Diameter**								
Volume of Fine Roots	(mm^3)	0.63 *0.000*	0.60 *0.000*	**Volume of Fine Roots**							
Root Surface Area	(cm^2)	0.95 *0.000*	0.21 *0.001*	0.77 *0.000*	**Root Surface Area**						
Root P	(g kg^{-1})	0.06 *0.370*	-0.02 *0.719*	0.01 *0.890*	0.03 *0.631*	**Root P**					
Root Dry Weight	(g)	0.16 *0.012*	-0.35 *0.000*	0.05 *0.425*	0.06 *0.384*	0.12 *0.070*	**Root Dry Weight**				
Root P Content	(mg)	0.20 *0.002*	-0.23 *0.000*	0.13 *0.053*	0.12 *0.070*	0.58 *0.000*	0.83 *0.000*	**Root P Content**			
Shoot P	(g kg^{-1})	0.10 *0.107*	-0.26 *0.000*	0.03 *0.637*	0.03 *0.641*	0.37 *0.000*	0.27 *0.000*	0.37 *0.000*	**Shoot P**		
Shoot Dry Weight	(g)	0.14 *0.036*	-0.39 *0.000*	0.01 *0.835*	0.02 *0.722*	0.14 *0.032*	0.70 *0.000*	0.60 *0.000*	0.33 *0.000*	**Shoot Dry Weight**	
Shoot P Content	(mg)	0.28 *0.000*	-0.24 *0.000*	0.23 *0.000*	0.21 *0.001*	0.28 *0.000*	0.50 *0.000*	0.53 *0.000*	0.78 *0.000*	0.74 *0.000*	**Shoot P Content**

Supplemental Table SIII: SNP effects on RSA traits for *SbPSTOL1* homologs. Genotypic effects were estimated as deviations from the alleles with effects set as zero.

Gene	Polymorphism Position (bp)	Allele	Bushiness (n n^{-1})	Centroid (cm)	Median Roots (n)	SRL (cm cm^{-3})
Sb03g006765	2002	G		0		
		C		0.44		
Sb03g031670	1312	C			2.19	
		A			0	
Sb03g031680	1541	A				0
		G				24.66
	1591	T				0
		C				28.94
Sb03g031690	2305	T		0.49		
		G		0		
	4020	C	0	0.46	2.78	
		A	0.51	0	0	
	4085	A	0	0.48	2.86	
		G	0.55	0	0	
	4121	A	0	0.47	2.48	
		G	0.50	0	0	
	4170	G			1.98	
		A			0	

Supplemental Table SIV: Sorghum *PSTOL1* primers sequences. Sequence of each primer amplifying *SbPSTOL1* genes. D: Direction of the primer. Ta: annealing temperature. F: forward. R: reverse

Sequence (5'->3')	D	Ta (°C)
Sb03g006765		
TGGCTTGGAGCAAAATACGTGGC	F	55
GGGCCCGCTTTGAGCCTTCC	R	
SNPs: 1912 (G/A); 1998 (C/T); 2002 (G/C); 2042 (G/C); 2042 (G/C); 2067 (G/C); 2073 (T/C); 2141(T/C)		
Sb03g031670		
CGTTGGTTTTGCATCGGCCACC	F	55
TCGTCTCGTCGCGTTGAACACA	R	
SNPs: 1312 (A/C)		
Sb03g031680		
ATGGCGCTGCGAAAATCCGT	F	55
TGTTGCCCAAAGGGAGGATCAAATC	R	
SNPs: 1541 (A/G); 1591(T/C)		
Sb03g031690		
GGCAAAAACGGTCCGAAGGGC	F	55
CGCTGCATCAATGGCGTTGCA	R	
SNPs: 305 (T/G); 2330 (G/A); 2433 (A/T); 2644 (A/C)		
ACGGCGGTCAGGTAGCAGTCA	F	55
GCACCGGCAATGGAAATGGCA	R	
SNPs: 4020 (A/C); 4085 (G/A); 4121 (G/A); 4170 (G/A)		
Sb03g031700		
TCGACTGCGTGGACGGCAAC	F	55
TCGTAAGGGCTGAGTGTGCGT	R	
SNPs: 381 (C/T); 463 (A/G)		
AGGGCATTAAACAGAAGCAGCAGT	F	55
ACACCACATCCACCTCGCCCT	R	
SNPs: 2006 (C/T); 2096 (G/A); 2304 (G/A)		
Sb07g002840		
GCAACCTGGCTGATGGCCGT	F	58
AGGCCTTCGGTGGCAACTGC	R	
SNPs: 3783 (A/G); 3795 (A/G); 3810 (C/T); 3939 (G/A); 3975 (T/C); 3987 (G/A); 4019 (G/C)		