Phosphate transporter PHT1;1 is a key determinant of phosphorus acquisition in Arabidopsis natural accessions

Pei-Shan Chien 1, Ya-Ting Chao 1, Chia-Hui Chou 2, Yu-Ying Hsu 1, Su-Fen Chiang 1, Chih-Wei Tung 2,* and Tzyy-Jen Chiou 1,* †

1 Agricultural Biotechnology Research Center, Academia Sinica, Taipei, Taiwan
2 Department of Agronomy, National Taiwan University, Taipei, Taiwan

*Author for correspondence: cihweitung@ntu.edu.tw (C.-W.T.), tjchiou@gate.sinica.edu.tw (T.-J.C.).
These authors contributed equally (Y.-T.C., C.-H.C.).
†Senior authors
P.-S. C., C.-W. T., and T.-J. C. conceived and designed the study. P.-S. C., Y.-Y. H., and S.-F. C. performed the experiments. P.-S. C., C.-H. C., Y.-T. C., C.-W. T., and T.-J. C. analyzed the data. P.-S. C., T.-J. C., C.-H. C., Y.-T. C., and C.-W. T. wrote the manuscript.
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 (https://academic.oup.com/plphys/pages/general-instructions) is: Tzyy-Jen Chiou (tjchiou@gate.sinica.edu.tw).

Abstract

Phosphorus (P) is a mineral nutrient essential for plant growth and development, but most P in the soil is unavailable for plants. To understand the genetic basis of P acquisition regulation, we performed genome-wide association studies (GWASs) on a diversity panel of Arabidopsis (Arabidopsis thaliana). Two primary determinants of P acquisition were considered, namely, phosphate (Pi)-uptake activity and PHOSPHATE TRANSPORTER 1 (PHT1) protein abundance. Association mapping revealed a shared significant peak on chromosome 5 (Chr5) where the PHT1;1/2/3 genes reside, suggesting a connection between the regulation of Pi-uptake activity and PHT1 protein abundance. Genes encoding transcription factors, kinases, and a metalloprotease associated with both traits were also identified. Conditional GWAS followed by statistical analysis of genotype-dependent PHT1;1 expression and transcriptional activity assays revealed an epistatic interaction between PHT1;1 and MYB DOMAIN PROTEIN 52 (MYB52) on Chr1. Further, analyses of F1 hybrids generated by crossing two subgroups of natural accessions carrying specific PHT1;1- and MYB52-associated single nucleotide polymorphisms (SNPs) revealed strong effects of these variants on PHT1;1 expression and Pi uptake activity. Notably, the soil P contents in Arabidopsis habitats coincided with PHT1;1 haplotype, emphasizing how fine-tuned P acquisition activity through natural variants allows environmental adaptation. This study sheds light on the complex regulation of P acquisition and offers a framework to systematically assess the effectiveness of GWAS approaches in the study of quantitative traits.

Introduction

Plants have evolved an array of strategies to sustain growth across diverse habitats where soil nutrients differ widely. Phosphate (Pi), the primary form of phosphorus (P) acquired by plants, is a limited resource because of its low solubility and mobility in soil (Theodorou and Plaxton, 1993; Raghothama, 1999). Modulation of P transport systems (Misson et al., 2004; Shin et al., 2004; Ai et al., 2009), root system architecture (Lynch, 2011; Pandey et al., 2013), secretion of root exudates (Wang et al., 2011), and reallocation...
and recycling of internal P (Raghothama, 1999) are known to be major adaptive responses plants use to cope with low availability of P in their environment. Remarkable progress has been made in understanding the genetic and molecular mechanisms that control these responses in Arabidopsis (Arabidopsis thaliana), particularly in the Columbia (Col-0) accession. For example, PHOSPHATE STARVATION RESPONSE 1 (PHR1) and PHR1-LIKE 1 (PHL1), which are MYB-coiled coil (CC)-type transcriptional activators, are known to play major roles in the activation of P starvation-inducible genes (Rubio et al., 2001; Bustos et al., 2010; Sun et al., 2016). Members of the PHOSPHATE TRANSPORTER 1 (PHT1), PHT5, and PHOSPHATE 1 (PHO1) transporter families were found to mediate P acquisition, storage, and root-to-shoot translocation, respectively (Hamburger et al., 2002; Misson et al., 2004; Shin et al., 2004; Liu et al., 2016). The activity of PHT1 is increased by its trafficking to the plasma membrane through PHOSPHATE TRANSPORTER TRAFFIC FACILITATOR1 (PHF1), but phosphorylation suppresses its trafficking (González et al., 2005; Bayle et al., 2011; Chen et al., 2015; Yang et al., 2020). Moreover, several regulatory modules of miRNAs and their target genes, such as miR399-PHO2 (a ubiquitin-conjugating enzyme), miR827-Nitrogen Limitation Adaptation (NLA; a RING type ubiquitin E3 ligase), and miR156-SQUAMOSA PROMOTER-BINDING PROTEIN-LIKE 3 (transcription factor) are intertwined to regulate P transport (Huang et al., 2013; Lin et al., 2013; Lei et al., 2016).

Many important agronomic traits are controlled by quantitative trait loci (QTLs). Identification of QTLs associated with a trait of interest and subsequent analyses of candidate genes often offer new strategies for crop improvement, including improvement of nutrient use efficiency. For example, using a QTL mapping strategy, Reymond et al. (2006) identified LOW PHOSPHATE ROOT1, which explained 52% of the variation in primary root growth response to P deficiency in an Arabidopsis Bay-0 × Shahdara recombinant inbred line population. In rice (Oryza sativa), Phosphorus-Starvation Tolerance 1 (PSTOL1)/Phosphorus uptake1, which encodes a protein kinase, is an important gene/QTL that enhances grain yield by promoting root growth (Wissuwa et al., 2002; Gamuyao et al., 2012). The absence of PSTOL1 from modern rice cultivars highlights the importance of exploring the diversity of natural variants for breeding crops with improved P acquisition ability (Gamuyao et al., 2012). Tremendous efforts have also been made to dissect the genetic basis of various P-related traits in other crops (Wang et al., 2010a, 2010b; Hufnagel et al., 2014; Gu et al., 2016). Despite the identification of several potential P use efficiency QTLs, few have been functionally validated (Gamuyao et al., 2012; Zhang et al., 2014; Luo et al., 2019).

Arabidopsis, which is grown as an inbred line and for which there is a wealth of natural accessions with publicly available whole genome sequences, is an attractive model for QTL identification using the genome-wide association (GWAS) approach (Aloano-Blanco and Koornneef, 2000; The 1001 Genomes Consortium, 2016). There have been several recent studies related to P deficiency in Arabidopsis accessions, such as identification of candidate loci regulating responses to P starvation (Yi et al., 2021), a combination of salinity and P starvation (Kawa et al., 2016), root growth under P and iron limitation (Bouain et al., 2019), and leaf anion homeostasis under different P supplies (El-Soda et al., 2019). These analyses highlight the power of using GWAS in Arabidopsis to decipher further players in P stress responses. However, this approach may have some potential drawbacks. These include false-positive associations resulting from cryptic population structure and kinships, and false negatives caused by overcorrection of confounding factors or insufficient statistical power to detect rare alleles (Nordborg and Weigel, 2008). Thus, to validate the biological roles of detected loci, systematic follow-up analyses are necessary (Verslues et al., 2014; Bouain et al., 2019).

P acquisition is a key determinant of P use efficiency (PUE) in plants; however, little is known about how genotypic variations govern P acquisition in natural environments. In this study, we analyzed a set of approximately 200 Arabidopsis accessions to dissect the genetic architecture controlling Pi-uptake activity and the protein levels of PHT1 Pi transporters under moderately low Pi conditions (50 μM). We established a platform for analyses including phenotyping followed by conventional and conditional GWAS mapping and the association of phenotypes with selected F1 heterozygotes. We identified several genes involved in regulating P acquisition-related traits. Importantly, we found a connection between the PHT1;1 haplotypes and P acquisition ability and the edaphic conditions in the growth sites of natural accessions. Our findings draw attention to genotypic variants of Arabidopsis accessions that are adapted to the soil P content in their habitats and present information for potentially improving PUE in crops along with a strategic approach for GWAS to identify causal genes.

Results

GWAS of Pi-uptake activity and PHT1;1/2/3 protein abundance under moderately low Pi

Two panels of Arabidopsis accessions (see details in “Materials and methods”), representing worldwide diversity and used to develop a high-resolution haplotype map and whole-genome sequences (Cao et al., 2011; Horton et al., 2012), were investigated in this GWAS. To determine the conditions for phenotyping P-acquisition ability, we first examined the growth and Pi concentrations of Col-0 seedlings grown in medium containing different concentrations of P (KH₂PO₄) ranging from 10 to 250 μM (Supplemental Figure S1). Because we intended to focus on the responses to Pi limitation while circumventing potential adverse effects caused by impaired growth, a moderately low concentration of Pi (50 μM) was chosen (Supplemental Figure S1A). The Pi concentrations of seedlings grown under this condition were significantly lower than those grown under Pi-sufficient conditions (250 μM), but there was no significant difference in fresh weight (Supplemental Figure S1, B–E). Using this
condition, we measured two important traits contributing to P-acquisition ability in 12-day-old seedlings. The first trait, root \([^{32}P]\)Pi-uptake activity, was measured by performing liquid scintillation counting of whole seedlings of 306 natural accessions. The second trait, protein abundance of PHT1 Pi transporters in the root, was quantified in 180 accessions by immunoblot analysis using an antibody that recognizes three Pi transporters (PHT1;1, PHT1;2, and PHT1;3) because of their high similarity.

After removing the accessions that showed inconsistent phenotypes across three to four biological replicates (15 accessions for uptake activity and two for protein abundance) and 27 accessions for uptake activity that could not be identified in the GWA-Portal database (Seren et al., 2012; The 1001 Genomes Consortium, 2016; Seren, 2018), 264 and 178 accessions were used for GWAS analysis of Pi-uptake activity and PHT1;1/2/3 protein level, respectively. These traits were normally distributed (Figure 1, A and B; Supplemental Display Activity and PHT1;1/2/3 protein level, respectively. These traits were normally distributed (Figure 1, A and B; Supplemental Table S1), suggesting that both are quantitative traits determined by multiple genes. Principal component analysis (PCA) of genotypic data revealed that PC1 only explained 3.15% and 2.9% of the total variation in the two populations analyzed for PHT1;1/2/3 protein level and Pi-uptake activity, respectively (Supplemental Figure S2). Based on the phenotype and kinship matrix, we calculated the pseudo-heritability and estimated how much variation was contributed by genotypic variation. A heritability of 0.388 and 0.735 was estimated for Pi-uptake activity and PHT1;1/2/3 protein level, respectively. Next, we employed the accelerated mixed model (AMM), in which kinship is considered to control confounding signals due to relatedness, implemented in model (AMM), in which kinship is considered to control confounding signals due to relatedness, implemented in GWA-Portal (Seren et al., 2012) as suggested (Atwell et al., 2010) for our association analysis. As shown in the Manhattan plots, several major SNP clusters were associated with Pi-uptake activity, but only one cluster with \(-log_{10}(P) > 6\) was identified for PHT1;1/2/3 protein level (Figure 1, C and D). Compared with PHT1 protein abundance, it is conceivable that the degree of complexity in controlling Pi-uptake activity is much higher, which involves other factors in addition to Pi transporters, such as root architecture and root exudates. The Pi uptake activity and PHT1;1/2/3 protein level showed a positive correlation (Pearson's \(r = 0.36, P < 1.00E-05;\) Supplemental Figure S3). Remarkably, both traits shared a significant peak on chromosome 5 (Chr5), suggesting that these two traits are controlled by the same locus. Intriguingly, PHT1;1/2/3 (AT5G43350/AT5G43370/AT5G43360) are located within the SNP peak on Chr5. This observation shows the effectiveness of our experimental design and the importance of PHT1 family members for natural variation in Pi-uptake activity.

Identification of candidate genes associated with P-acquisition traits

To identify the candidate genes located in the regions adjacent to the significant peaks, we first defined significant SNPs and then searched for trait-associated candidate genes. The SNPs with a minor allele count \(> 10\) and a \(-log_{10}(P) > 4\) based on the deviation of observed \(-log_{10}(P)\) from expected \(-log_{10}(P)\) in the Q–Q plots were selected (Figure 1, C and D, right). This threshold \(-log_{10}(P) > 4\) has been considered reasonable to determine significant SNP-trait association in Arabidopsis GWAS studies (Togninalli et al., 2018, 2020). Using these criteria, 478 and 508 significant SNPs were identified for Pi-uptake activity and PHT1;1/2/3 protein abundance, respectively (Figure 2; Supplemental Table S2, A and B). As the decay of linkage disequilibrium (LD) in Arabidopsis accessions is within 10 kb on average (Kim et al., 2007), we selected 462 and 282 protein-coding genes located within the 10-kb regions flanking the significant SNPs as potential candidates (Figure 2; Supplemental Table S3). Among them, 450 and 270 genes were unique for Pi-uptake activity and PHT1;1/2/3 protein level, respectively, and 12 genes were associated with both traits, including 4 Pi transporters, PHT1;1, PHT1;2, PHT1;3, and PHT1;6 (Figure 2; Supplemental Table S3).

To prioritize candidate genes for further functional assays, we assigned a value, denoted as “hit frequency”, based on the number of times a gene overlapped with the 10-kb region surrounding a significant SNP (Supplemental Figure S4). For instance, the hit frequency of PHT1;1 was calculated to be 61 for uptake activity (Supplemental Table S3), meaning that any part of the PHT1;1 gene body is within the 20-kb windows centered on 61 significant SNPs (Supplemental Figure S4). Genes were selected using the following criteria: (1) genes were associated with either Pi-uptake activity or PHT1;1/2/3 protein level and had a hit frequency \(> 10\) or (2) genes were associated with both traits and had a hit frequency \(> 5\). Accordingly, 81 genes associated with either Pi uptake activity (48 genes) or PHT1;1/2/3 protein level (33 genes) and 10 genes associated with both were identified (Figure 2; Supplemental Table S3).

In line with the previous studies (Shin et al., 2004; Huang et al., 2013), the pht1;1 single mutant showed a significant and striking reduction in PHT1;1/2/3 protein level, Pi-uptake activity (Supplemental Figure S5), and Pi concentration (Figure 3) under a moderately low Pi condition (50 \(\mu\)M). In addition, we analyzed 97 available T-DNA insertion mutants, covering 43 out of 91 candidate genes, to evaluate their Pi-acquisition ability (Supplemental Table S4). We considered the genes promising candidates only when two independent allelic mutants showed consistent and significant changes. Using this criterion, we did not observe significant changes in the PHT1;1/2/3 protein level or short-term Pi-uptake activity for any of the candidate mutant lines (Supplemental Figure S5, A and B). However, when the Pi concentration of 12-day-old seedlings was analyzed, we found that mutations in seven of the genes, namely AT1G17890 (GDP-keto-6-deoxymannose 3,5-epimerase/4-reductase; GER2), AT1G17870 (ETHYLENE-DEPENDENT GRAVITROPISM-DEFICIENT AND YELLOW-GREEN-LIKE 3; EGY3), AT1G17850 (Rhodanese/Cell cycle control phosphatase superfamily protein), AT5G43330 (CYTOSOLIC-NAD- 

### Reference

Shin, J. et al. (2004). [Reference](http://example.com). Huang, Y. et al. (2013). [Reference](http://example.com).
DEPENDENT MALATE DEHYDROGENASE 2; C-NAD-MDH2), AT5G43320 (CASEIN KINASE I-LIKE 8; CKL8), AT1G17950 (MYB DOMAIN PROTEIN 52; MYB52), and AT2G27610 (tetratricopeptide repeat-like superfamily protein), showed a significant reduction (7%–23%) either in roots or shoots compared with the Col-0 wild-type (WT) control (Figure 3, Aa n dB). Since Pi concentration represents a final consequence of various activities, including P acquisition, metabolism, and storage, we reasoned that these genes might participate in any of these activities despite a minor role in directly regulating Pi uptake. Nevertheless, it is also necessary to keep in mind that all these T-DNA insertional mutants are generated in a single accession, Col-0, and the complexity of naturally occurring variants from the Arabidopsis genomic backgrounds in the diversity panel is not fully represented.

We then examined the RNA levels of selected genes in the roots and shoots of Col-0 under sufficient Pi (1-mM KH$_2$PO$_4$) and low Pi (50-l KH$_2$PO$_4$) conditions (Supplemental Figure S6). Under sufficient Pi conditions, the RNA levels of most candidate genes were higher in the shoots than in the roots except for the MYB52 transcription factor, which was mainly expressed in the roots, and GER2, which encodes an enzyme involved in the biosynthesis of GDP-fucose that was expressed equally in both tissues. In response to low Pi conditions, only the expression level of GER2 in the shoots was reduced slightly (0.8-fold). None of the other candidate genes were responsive to low Pi in the roots.

Figure 1 GWAS of two P-acquisition traits. Phenotypic distribution of raw values for Pi-uptake activity (A) and square root-transformed (SQRT) values for PHT1;1/2/3 protein abundance (B). Manhattan plots of Pi-uptake activity (C) and PHT1;1/2/3 protein level (D); the respective Q–Q plots are shown on the right. The analyses were conducted using GWA-Portal. Thresholds for Bonferroni (red dashed line) and Benjamini Hochberg (blue dashed line) correction provided by GWA-Portal are indicated. The two lines overlap in (C). Red arrows in (C) and (D) indicate the shared significant peaks located at Chr5.
Figure 2 GWAS of two Pi-related traits and the step-by-step scoring system used for the selection of candidate genes. In addition to the well-known Pi transporters PHT1;1 and PHT1;2, our approach identified seven candidate genes (e.g. transcription factors, a kinase, an oxidoreductase, and a protease) involved in P acquisition under low Pi.

Epistatic interaction between PHT1;1 and MYB52
Because there was only one prominent SNP peak on Chr5 associated with PHT1;1/2/3 protein level (Figure 1D), our approach identified seven candidate genes (e.g. transcription factors, a kinase, an oxidoreductase, and a protease) involved in P acquisition under low Pi.

Next, we selected accessions based on the four allele combinations to test the potential interaction between these two SNP loci. Because the population in four different genotype combinations among the 178 accessions analyzed is varied, based on the availability of genome sequences in Arabidopsis 1001 Genome (https://tools.1001genomes.org/vcfsubset/#select_strains), we randomly selected 5–9 accessions from each allele combination as representatives (6 out of 8 {T,T} accessions, 9 out of 63 {A,T} accessions, 7 out of 92 {A,C} accessions, and 5 out of 15 {T,C} accessions; Figure 4C). We then evaluated the expressions of six annotated candidate genes surrounding the SNP locus 1 (Supplemental Figure S7A; the lower) and PHT1;1/2/3 genes (SNP locus 2) in each accession. Since PHT1 antibody cannot discriminate three Pi transporters (PHT1;1, PHT1;2, and PHT1;3) because of their high similarity, the current step...
focuses on dissecting how genetic regulation affects the gene expression of individual PHT1s. Here, we thus detected the transcript level of PHT1;1, PHT1;2 and PHT1;3 instead of an overall protein abundance of PHT1;1/2/3. Among the genes surrounding SNP locus 1, only MYB52 showed a clear difference in expression between the two allelic variants at SNP locus 1 (Supplemental Figure S9B). As revealed by two-way analysis of variance (ANOVA), MYB52 expression was significantly higher in the accessions with allele “A” than in those with allele “T” at SNP locus 1 (\( P = 1.00 \times 10^{-4} \)), regardless of the allele at SNP locus 2 (Figure 4D). Intriguingly, a significant allelic interaction between SNP locus 1 and SNP locus 2 was observed for the expression of PHT1;1 (\( P = 1.60 \times 10^{-2} \); Figure 4E). These observations highlight the existence of an interaction between MYB52 and PHT1;1. On the other hand, there was no significant allelic interaction between SNP locus 1 and 2 for PHT1;2 (\( P = 4.16 \times 10^{-1} \)) or PHT1;3 (\( P = 6.99 \times 10^{-1} \); Supplemental Figure S9, C and D).

Because MYB52 was annotated as an R2R3-MYB transcription factor (Dubos et al., 2010), we hypothesized that there is an epistatic interaction between MYB52 and PHT1;1 resulting from MYB52-dependent transcriptional regulation of PHT1;1. To examine this possibility, we co-expressed Col-0-MYB52 as an effector with a luciferase reporter driven by the promoter of Col-0-PHT1;1 in protoplasts isolated from the leaves of Col-0 (Figure 5A). In support of our hypothesis, MYB52 positively regulated PHT1;1 promoter activity (1.8-fold of the vector control; Figure 5B). However, the activation was not as high as the major Pi regulator PHR1 (5.4-fold of the vector control; Figure 5B). On the other hand, we were unable to detect a physical interaction between MYB52 and its potential binding sites (MBSII) in the promoter of PHT1;1 by yeast one-hybrid analysis despite the occurrence of positive interaction between PHR1 and the PHR1-binding sequences (P1BS) in the PHT1;1 promoter (Supplemental Figure S10). Because MYB is one of the transcription factor families that can act as homo- or heterodimers to regulate distinct cellular processes (Pireyre and Burow, 2015), MYB52-mediated transcriptional activation of PHT1;1 may require other factors.

---

**Figure 4** Conditional GWAS and statistical analyses reveal an association between MYB52 and PHT1;1. Manhattan plots of GWAS analyses performed using AMM (A, no conditioning) or modified AMM in which the most significant SNP located at Chr5 (the red arrowhead in A) was included as a cofactor (B, conditioning) in GWAPP; a further SNP peak on Chr1 became evident (the red arrowhead in B). The 5% false discovery rate threshold is plotted as a dashed horizontal line. C, List of candidate accessions used for examination of the epistatic relationship between Chr1 and Chr5. Asterisks indicate the SNPs that were corrected after re-sequencing. Analyses of the relationship between SNP locus 1 and SNP locus 2 by two-way ANOVA based on average gene expression levels of MYB52 (2\(^{-\Delta CT} \times 1000\)) (D) and PHT1;1 (2\(^{-\Delta CT} \times 100\)) (E). The letters in blue and red shown in (C), (D), and (E) indicate SNP locus 1 and SNP locus 2, respectively.
{A,C} genotype is associated with higher PHT1;1 expression and superior Pi uptake activity

Given that the accessions with the {A,C} allele combination exhibited the highest PHT1;1 gene expression among the four combinations (Figure 4E), we were curious how P acquisition is affected when one genomic copy of the {A,T} allele is replaced by one copy of the allele from the {A,C} genetic background. Thus, we crossed {A,T} accessions (Cnt-1, Shigu-2, An-1, A1t-1, C1BC2, Ba-1, and Wa) or {A,C} accessions (Paw-3, Ciste-1, Voeran-1, Sei-0, Ta-0, and Ga-2) to Col-0, which is also {A,C} (Figure 6; Supplemental Table S5). Although we have yet to dissect potential factors at other genomic loci that may also contribute to the phenotype, we observed significantly increased PHT1;1 expression (P = 2.91E-02) and Pi-uptake activity (P = 1.33E-02) relative to the maternal {A,T} accessions when one copy of the {A,C} genome was introduced into the {A,T} background (Figure 6, A and B). For instance, F1 plants from Cnt-1 X Col-0, which are {A,T} × {A,C} hybrids, showed a 7.19-fold increase in PHT1;1 gene expression and a 1.57-fold increase in Pi uptake activity when compared with the maternal line Cnt-1 (Supplemental Table S5). However, this phenomenon was not observed in the {A,C} × {A,C} F1 hybrid. We surmise that the varied P acquisition ability in the {A,C} × {A,C} hybrids may be caused by heterosis or hybrid weakness. These results showed that introduction of one copy of the {A,C} genome can improve the P acquisition of the {A,T} accessions, probably through the introduction of functional PHT1;1.

We also site-by-site generated another F1 population by crossing a homozgyous pht1;1 knockout mutant (Col-0 background) with various {A,T}- and {A,C}-accessions, the same accessions selected above. We found that the PHT1;1 expression was rescued partially in both {A,T} X pht1;1 and the {A,C} X pht1;1 hybrids relative to null expression in pht1;1 mutant (Figure 6A), indicating that PHT1;1 introduced from either {A,T}- or {A,C} genotype is functional. The partial complementation may result from one copy of functional PHT1;1 (i.e. pht1;1-Col-0/PHT1;1(A,T) or pht1;1-Col-0/PHT1;1(A,C)). It is interesting to note that the PHT1;1 expression in the {A,C} X pht1;1 hybrid was higher than that in the {A,T} X pht1;1 hybrid (P = 2.47E-04, Figure 6A). However, the difference in the Pi-uptake activity between these two hybrids was not evident (P = 1.84E-01, Figure 6B), implying other effects from unidentified loci complex with PHT1;1 expression to regulate the overall activity of Pi uptake.

The results of F1 analysis support the genetic interaction between SNP locus 1 and locus 2: (1) PHT1;1 gene expression increased significantly in the {A,C} X pht1;1 hybrids than in the {A,T} X pht1;1 hybrids (P = 2.47E-04, Figure 6A) and (2) the performance for P acquisition is significantly increased in the {A,T} X Col-0 hybrids (P = 1.33E-02) but not the {A,C} X Col-0 hybrids (P = 9.93E-01, Figure 6B). Conclusively, {A,C} genotype is associated with higher PHT1;1 expression and superior Pi uptake activity.

Because PHT1;1 is a major determinant of Pi-uptake activity and its expression is highly dependent on the allelic variation at SNP locus 2, that is the accessions with allele “C” had significantly higher transcript levels than those with allele “T” (Figure 4E), we divided the Arabidopsis diversity panel into two populations based on the allelic variation at SNP locus 2 (either “C” or “T”) and searched for additional loci controlling Pi-uptake activity independent of PHT1;1. Using these two artificial populations composed of 153 accessions with allele “C” and 111 accessions with allele “T”, we conducted GWAS for Pi-uptake activity. Although a few SNP peaks were observed in the Manhattan plots, none were significant (−log10(P) > 6), based on the threshold used in Figure 1, C and D; Supplemental Figure S11). Therefore, Pi-uptake activity is mainly controlled by SNP locus 2 on Chr5. The effect of other QTLs may not be substantial enough to be detected by this approach.

The PHT1 haplotype as a hallmark linking P-acquisition traits and the soil P contents in Arabidopsis habitats

Because SNP locus 2 is a synonymous SNP, we searched for non-synonymous SNPs and found four nearly consecutive significant SNPs in the second exon of PHT1;1 (Chr5: 17,401,478, 17,401,472, 17,401,471, and 17,401,466 bp)

Figure 5 MYB52 regulates the promoter activity of PHT1;1. A: A schematic illustration of constructs used for the analysis of PHT1;1 promoter activity in a dual-luciferase assay. B: PHT1;1 promoter activity was defined as the ratio of LUC:REN (firefly:renilla). Transformation with the empty vector served as a negative control. Data are presented as mean ± se (n = 4 biologically independent samples). The significance of differences was determined by two-sided Student’s t test (**P < 0.01).
denoted SNP3, SNP4, SNP5, and SNP6, respectively (Supplemental Figure S8B). These SNPs are in high LD (\(r^2 = 0.99\) in the group of accessions assayed for PHT1;1/2/3 protein level, Supplemental Table S6A) and explain 9.11% and 10.25% of the genotypic variance in Pi uptake activity and PHT1;1/2/3 protein level, respectively (Supplemental Table S2, A and B). These SNPs code for two forms of the PHT1;1 protein, the Col-0 type Asn–Pro–Glu–Ser–Ala (NPESA) and the non-Col-0 type Asp–Pro–Thr–Ser–Pro (DPTSP; Supplemental Figure S8B). To investigate whether these nonsynonymous SNPs could alter P-acquisition activity, we first inspected the subcellular localization of the proteins encoded by these SNP alleles and found no difference in their plasma membrane localization (Supplemental Figure S12A). Then, we examined the protein level and Pi-uptake activity in \(pht1;1\) mutants expressing PHT1;1 NPESA or PHT1;1 DPTSP driven by the native PHT1;1 promoter (Supplementary Figure S12, B and C). Despite expression variation among independent transgenic lines, in general, both transgenes PHT1;1 NPESA and PHT1;1 DPTSP could restore the phenotypes (PHT1;1/2/3 protein level and Pi-uptake activity) of the \(pht1;1\) mutant to the WT level (Supplemental Figure S12, B and C). However, no consistent significant differences were observed between plants expressing these transgenes, suggesting that the corresponding nucleotides may serve as markers rather than causative SNPs.

SNP locus 2 is in high LD with SNP3/4/5/6 (\(r^2 = 0.67\) pairwise in the group of accessions assayed for PHT1 protein level, Supplemental Table S6A). The combination of these five SNPs results in four haplotypes, GACCT, GACC, AGAGT, and AGAGC (Supplemental Figure S8C), in which GACCT (105 and 68 accessions assayed for Pi-uptake activity and PHT1;1/2/3 protein level, respectively) and AGAGC (139 and 94 accessions assayed for Pi-uptake activity and PHT1;1/2/3 protein level, respectively) are the two major haplotypes. It is of interest to note that the accessions with the “AGAGC” haplotype showed significantly higher Pi-uptake activities (\(P = 3.39E-08\)) and PHT1;1/2/3 protein levels (\(P = 3.64E-08\)) than those with the “GACCT” haplotype (Figure 7, A and B). Notably, when we located the habitats of these natural accessions in Europe, we found a significant association (\(P = 1.47E-02\)) between these two haplotypes and the edaphic topsoil conditions (Tóth et al., 2014; Figure 7C). The high Pi-uptake accessions (with the “AGAGC” haplotype) tended to be collected from growing sites containing low soil P whereas the low Pi-uptake accessions (with the “GACCT” haplotype) were collected from high soil P sites (Figure 7D). However, the relatively weak association with Pi levels suggests that soil P is not the only factor to impose selection pressure on Arabidopsis populations. In summary, our analyses revealed the importance of the PHT1;1 haplotype in connecting the P-acquisition activities of natural Arabidopsis accessions to the soil P contents in their habitats.

Discussion

**PHT1;1** is a key determinant of P-acquisition activity as evidenced by analysis of Arabidopsis natural accessions

In this study, we employed GWAS to analyze the genetic architecture of two major traits representing the capability to acquire P. Figure 8 illustrates the critical steps of our systematic approach. We identified a significant SNP peak for both traits on Chr5 where the Pi transporters PHT1;1, PHT1;2, and PHT1;3 are located (Figure 1), highlighting the importance of these Pi transporter genes in regulating their own protein abundance as well as the Pi-uptake activity of whole plants. In agreement with a very recent report (Yi et al., 2021), PHT1 was also identified as one of the crucial loci conferring natural variations in low Pi tolerance using five Pi starvation response traits. Nevertheless, Sakuraba et al.
(2018) employed $^{33}$Pi radioactive imaging to analyze 200 natural accessions of Arabidopsis but did not identify any QTL for P-acquisition ability. This is probably because the sensitivity of the imaging system was lower than that of the scintillation counter employed in this study. These findings strengthened the importance of the PHT1 family for P acquisition from the perspective of genetic variations in a natural population, and also demonstrated the power and limitation of GWAS.

Figure 7 PHT1;1 haplotype connects the P-acquisition phenotypes of Arabidopsis accessions to the soil P contents in their habitats. Comparison of Pi-uptake activity (A) and PHT1;1/2/3 protein level (B) between accessions with different PHT1;1 haplotypes. Each gray dot represents an individual accession. C, Distribution of two different genotypic accessions and soil P-contents (in purple) on a map of the European Union. Symbols in green are accessions carrying the “GACCT” haplotype, whereas symbols in blue are those with “AGAGC”. D, Comparison of soil P content for two types of accessions: those with the “GACCT” haplotype and those with the “AGAGC” haplotype. The dark purple lines within violin plots represent the 25th percentile, median (dashed), and 75th percentile. The red lines indicate the means of the quantified phenotypes. The $P$-value determined by two-sided Student’s $t$ test is indicated on the top of each plot. Values in parentheses shown below the SNPs indicate the numbers of accessions assayed.

Among PHT1;1, PHT1;2, and PHT1;3, PHT1;1 plays the dominant role, possibly because of its relatively high expression level (Supplemental Figure S9, C and D). Indeed, unlike the pht1;1 mutant, the pht1;2 and pht1;3 mutants showed no reduction in the Pi concentration compared with the WT Col-0 (Huang et al., 2013). Yi et al. (2021) hypothesized this might result from gene duplication followed by a strong purifying selection to adapt to the environment. Our analyses clarified the difference in regulating the transcription
level of these PHT1 homologs among Arabidopsis natural accessions, revealing that PHT1;1 is controlled by MYB52, but not PHT1;2 or PHT1;3 (Figure 4; Supplemental Figure S9). Importantly, we provide evidence to connect the PHT1;1 haplotype of natural accessions to the soil P contents in their habitats.

**MYB52 regulates PHT1;1 expression and Pi accumulation**

Expression of Arabidopsis PHT1;1 was previously reported to be positively regulated by PHR1 (Rubio et al., 2001), WRKY DNA-BINDING PROTEIN 75 (WRKY75, Devaiah et al., 2007), WRKY45 (Wang et al., 2014), WRKY42 (Su et al., 2015), phytochrome-interacting factors PIF4/PIF5 and ELONGATED HYPOCOTYL 5 (Sakuraba et al., 2018), and NITRATE-INDUCIBLE, GARP-TYPE TRANSCRIPTONAL REPRESSOR 1/ HYPERSENSITIVE TO LOW PI-ELICITED PRIMARY ROOT SHORTENING 1.2 (Wang et al., 2020). Our study employed conditional GWAS followed by epistasis analysis and transcriptional activity assays and identified MYB52 as an additional positive regulator of PHT1;1 expression (Figures 4 and 5).

Although MYB52 gene expression is not Pi-starvation responsive, both PHT1;1 and MYB52 were highly expressed in the root relative to the shoot (Supplemental Figure S6). Moreover, results of the promoter–reporter analysis revealed a ubiquitous expression pattern of PHT1;1 in the root, including epidermis, cortical cells, central cylinder, lateral roots, and root tip (Mudge et al., 2002; Nussaume et al., 2011), and the expression of MYB52 in the basal part and the maturation zone of lateral roots (Park et al., 2011). Together, these observations support the coincided PHT1;1 and MYB52 expression in the same cells and the potential role of MYB52 in regulating root-expressed PHT1;1 and then contributing to P acquisition. The observation of reduced shoot Pi concentration in myb52 mutants (17.5% reduction compared with the WT, Figure 3B) further supports this notion. However, we did not observe significant changes in PHT1;1/2/3 protein level or short-term Pi-uptake activity in the myb52 mutants (Supplemental Figure S5). We reasoned that the T-DNA insertional mutants generated in a single Col-0 background do not represent the complexity of genetic variations in diverse Arabidopsis natural accessions analyzed in this work. In addition, in the myb52 single mutant, other positive transcriptional activators may compensate PHT1;1 expression for the loss of MYB52.

So far, only a limited number of studies have explored the function of MYB52, and these studies have focused on its potential roles in regulating ABA-mediated drought tolerance and cell wall thickening (Zhong et al., 2008; Park et al., 2011). Although cell wall modification is associated with the change of root architecture and P re-utilization (Lin et al., 2011; Zhu et al., 2015; Naumann et al., 2019), the involvement of MYB52 in these processes requires further investigation. Several other R2R3 MYB transcription factors were identified to be upregulated by P deficiency and involved in P acquisition or P-starvation responses. For example, Arabidopsis MYB2 and MYB62 regulate P-starvation responses via an miR399-mediated pathway and gibberellic acid biosynthesis, respectively (Devaiah et al., 2009; Baek et al., 2013). In rice, overexpression of OsMYB5P and OsMYB2P-1 could improve P acquisition by enhancing the expression of Pi transporters (Dai et al., 2012; Yang et al., 2018). Unlike these MYB transcription factors, the
expression of MYBS2 does not respond to changes of external Pi supply (Supplemental Figure S6), underlining the value of GWAS, in which analysis is independent of gene expression.

**PHT1;1 haplotype connects the P-acquisition phenotypes of Arabidopsis accessions to the soil P contents in their habitats**

Based on the epistatic analysis of SNP locus 1 and SNP locus 2, the regulation of P acquisition in different allele combinations only indicates a genetic interaction with the expression level of PHT1;1, with the highest PHT1;1 expression occurring in accessions carrying \{A,C\} and the lowest in those having \{A,T\} (Figure 4E). Moreover, analysis of heterozygotes (Figure 6) revealed the potential for increasing P acquisition ability by the introduction of the \{A,C\} allele. Nonetheless, SNP locus 1 "A" is located downstream of MYBS2 and SNP locus 2 "C" is a synonymous SNP within the PHT1;1 gene body. Thus, modification of MYBS2 or PHT1;1 protein sequence should not be the reason; these two SNPs may serve as markers rather than causal alleles. These observations highlighted the potential of the SNPs that linked either to the SNP locus 1 or the SNP locus 2 involved in regulating PHT1;1 expression and Pi-uptake activity, assuming that other factors might contribute to the P acquisition as well.

Notably, we found that two groups of accesses, discriminated by PHT1;1 haplotype (SNP3/4/5/6/locus 2), showed different P-acquisition phenotypes and different soil P contents in their habitats. As shown in Figure 7, the "AGAGC" haplotype is found in high-activity accessions preferentially collected from low-P soil. By contrast, the "GACCT" haplotype corresponds to low-activity accessions preferentially collected from high-P soil. This result suggests that soil P level imposes selection pressure on Arabidopsis populations. The accessions with the "AGAGC" haplotype may adapt better to poor P soil than those with the "GACCT" haplotype because of improved ability to acquire P. This report shows the association between P-related traits and soil P content.

**Discovery of additional regulators of P acquisition**

Our GWAS study also identified several loci with minor effects potentially involved in P acquisition (Figure 3). For example, C-NAD-MDH2 can convert oxaloacetate to malate, secreting into the rhizosphere to liberate Pi from the soil (Ryan et al., 2001). CKL8, a Ser/Thr kinase, was reported to modify the activity of ACC synthase 5 (Tan and Xue, 2014) and may regulate root elongation via ethylene under Pi-deprivation (Ma et al., 2003). EGY3 was classified as a site-2 protease-like putative metalloprotease implicated in endoplasmic reticulum stress (Iwata et al., 2009; Iwata et al., 2017), which could trigger Pi-dependent activation of autophagy (Naumann et al., 2019). The roles of CKL8 and EGY3 in P acquisition require further investigation. It is of interest to note that the expression of these candidate genes is not responsive to changes in Pi supply (Supplemental Figure S6), making it impossible to identify them by transcriptomic analysis (Misson et al., 2005; Liu et al., 2016).

Apart from PHT1;1/2/3, we did not identify other well-known genes that regulate P-acquisition activity, such as PHR1, PHO2, NLA, and PHF1 (Rubio et al., 2001; González et al., 2005; Huang et al., 2013; Lin et al., 2013). These known Pi-starvation regulators were initially identified based on the loss-of-function mutant phenotypes in a single genetic background. By contrast, the GWAS analyzed the statistical association between phenotypic variation and a single SNP variant in diverse accessions. Although a common assumption is to search for a single causal variant, it is often overlooked that the phenotype might be conferred by several variants acting together (Burgess, 2022). In a recent similar study, PHT1 loci was also the only Pi regulatory cluster being identified by GWAS, even though five P-acquisition-related traits were analyzed (Yi et al., 2021). GWAS detection power is affected by many factors concerning the population studied, such as allele frequency of variants, population size, and demographic history of the analyzed trait, which could contribute to the “missing heritability” (Korte and Farlow, 2013). Furthermore, most detected genetic variants reside outside protein-coding regions, making it challenging to investigate the biological mechanism controlling the traits without identifying the causal genes.

It is always challenging to validate the result of GWAS. Even with sufficient statistical power to detect genetic variants, the effect contributed by a single candidate gene could be mild. In our case, the SNPs located within PHT1;1 explain 9%–10% Pi-uptake activity and 10%–14% PHT1;1/2/3 protein level (Supplemental Table S2, A and B), the remaining 90% of the phenotype contributed by individual factors could be mild, as no substantial phenotypic changes were observed in the single gene mutants analyzed (Supplemental Figure S5). Nevertheless, MYBS2 is relatively promising because it was identified by the first round GWAS and the subsequent conditional GWAS (Figure 4), followed by validation of PHT1;1 regulation by promoter-reporter assay (Figure 5).

Our findings extend current knowledge by uncovering additional genetic regulators of P acquisition and by linking haplotypes to nutrient ion contents in the habitats of Arabidopsis accessions. Furthermore, our phenotyping and analytical approaches provide a framework to systematically assess the effectiveness of the GWAS approach for studies of quantitative traits, shedding light on revealing a haplogroup of natural accession with superior Pi uptake activity for the trait prediction and crop breeding.

**Materials and methods**

**Plant materials and growth conditions**

Two panels of Arabidopsis (Arabidopsis thaliana) natural accessions (CS76309 and CS76427) used in this study were purchased from the Arabidopsis Biological Resource Center (ABRC). CS76309, which was the largest population of natural accessions available when this project was initiated, and...
is a core set of 360 natural accessions genotyped with a 250 k-SNP/high-density tiling array by Justin Borevitz’s laboratory. (Horton et al., 2012). CS76427 is a set of 80 natural accessions sequenced using the Illumina GA platform by Detlef Weigel’s laboratory as part of the 1001 Genomes Project (Cao et al., 2011). Ninety-seven T-DNA mutants covering 43 candidate genes associated with GWAS traits were obtained from the ABRC and Nottingham Arabidopsis Stock Centre (Supplemental Table S3). Homozygous lines were isolated and analyzed under moderately low Pi conditions (50–100 μM KH2PO4).

Seeds were surface sterilized and germinated on agar plates with half-strength modified Hoagland nutrient solution containing 250-μM KH2PO4, 1% (w/v) sucrose, and 0.8% (w/v) bactoagar (Chiou et al., 2006). Six-day-old seedlings were transferred to Pi-sufficient (+Pi) and Pi-deficient (−Pi) media supplemented with 1-mM and 50-μM KH2PO4, respectively, and grown for another 6 days. Plants were grown at 22°C under a 16-h photoperiod with cool fluorescent white light at 100–150 μE m⁻² s⁻¹.

**Plasmid construction and plant transformation**

For complementation of the pht1;1 mutant, the open reading frame of PHT1;1 was cloned into pCR8/GW/TOPO (Invitrogen) and then recombined into the Gateway destination vector pMDC32, in which the constitutive 3SS promoter was replaced by the PHT1;1 promoter (3.3-kb upstream from the start codon). Modification of the sequence of PHT1;1 was performed by site-directed mutagenesis using the Col-0 sequence as the template. The resulting constructs were designated PHT1;1<sup>DPTSP</sup> and PHT1;1<sup>NPESA</sup>. Transgenic plants were generated using an *Agrobacterium tumefaciens* dipping procedure (Clough and Bent, 1998) with strain GV3101.

**RNA extraction and RT–qPCR analysis**

Total RNA from 14-day-old seedlings was isolated using TRIzol reagent (Invitrogen) followed by DNase I (Ambion) treatment. cDNA was synthesized from 1-μg total RNA using Moloney murine leukemia virus reverse transcriptase (Invitrogen) with oligo(dT) primer. Reverse transcription quantitative–polymerase chain reaction (RT–qPCR) was performed using the Power SYBR Green PCR Master Mix Kit (Applied Biosystems) on a 7500 Fast Real-Time PCR System (Applied Biosystems) according to the manufacturer’s instructions. Relative expression levels were normalized by an internal control Actin8. The RT–qPCR primers designed to examine transcript levels are located at the same position in the selected genes with a perfect match in each accession. The primers used for the gene expression are listed in Supplemental Table S7.

**Immunoblot analysis**

Roots from ten 12-day-old seedlings were collected, ground in liquid nitrogen, and dissolved in protein lysis buffer containing 2% (w/v) sodium dodecyl sulfate (SDS), 60-mM Tris–HCl, pH 8.5, 2.5% (v/v) glycerol, 0.13-mM EDTA, and 1× complete protease inhibitor (Roche). Total proteins (50 μg) were loaded onto 4%–12% Bis–Tris SDS–polyacrylamide gel electrophoresis (SDS–PAGE) gels (NuPAGE System) and transferred to polyvinylidene difluoride membranes. PHT1 proteins were then detected using PHT1;1/2/3 primary rabbit polyclonal antibody (Liu et al., 2011). ADP-ribosylation factor 1 (Arf-1), a Golgi marker, recognized by anti-Arf-1 antibody (Agrisera, AS08 325), was used as a loading control. The chemiluminescence signal was acquired by an UVP BioSpectrum 815 Image System (Level Biotechnology) under a linear dynamic range without saturation and quantified using ImageJ (https://imagej.nih.gov/ij/).

**Pi concentration and Pi uptake activity**

Pi concentration and Pi uptake activity were measured as described previously (Chiou et al., 2006). The shoots and roots of 12-day-old seedlings were harvested separately for the Pi concentration analysis. To measure Pi uptake activity, five 11-day-old seedlings were first incubated in sugar-free–Pi liquid medium overnight. The next day, the Pi uptake activity was measured using whole seedlings after transferring to a medium containing 50-μM [33P] KH2PO4 for 2 h. The seedlings were grown under the same conditions as samples used for western blot analysis.

**GWAS**

Among the 440 (360 + 80) natural accessions, we excluded those with the same ID but a different stock number (e.g. Vezzano-2 [CS76349 and CS76350]), a similar ID (e.g. Ciste-1 and Ciste-2), or seeds unavailable after propagation. We phenotyped 306 lines for Pi-uptake activity. Because some accessions could not be identified in GWA-Portal or GWAPP, we reduced the number to 180 for phenotyping PHT1;1/2/3 protein abundance.

To detect PHT1;1/2/3 protein level, we randomized the order of loading for a set of samples in one SDS–PAGE gel to avoid variation caused by transfer efficiency among replicates. In addition, an identical protein sample extracted from Col-0 accession was loaded on every SDS–PAGE gel. Each blot was also probed with anti-Arf-1 antibody to acquire the relative value of PHT1 (PHT1/Arf-1) among different accessions. This relative value was then normalized to the identical Col-0 sample between different blots to obtain the final relative PHT1;1/2/3 protein abundance in each accession from three to four biological replicates. We applied the coefficient of variation (CV) as a threshold to take the variability of biological replicates into account (Luo et al., 2019). Since CV was defined as the ratio of standard deviation (σ) to the mean, setting the filtering threshold using CV instead of SD could avoid the problem of sensitivity of SD to outliers. When CV was set to 0.3, the number of accessions was sufficient for GWAS analysis (264 and 178 Arabidopsis accessions for Pi uptake activity and the protein levels of PHT1;1/2/3, respectively).

The quantified data were analyzed by GWA-Portal against an imputed full sequence dataset using an AMM, which takes kinship into account to reduce the confounding
effects of population structure (Kang et al., 2008; Seren et al., 2012; Consortium, 2016). The raw values of Pi uptake activity (Supplemental Table S1A) were subjected to analysis directly; however, square root-transformed (SQRT) values for PHT1;1/2/3 protein abundance (Supplemental Table S1B), an option implemented in GWA-Portal, were used to improve normality. We performed conditional GWAS, whose function is available in GWAPP, in which a 250 k SNP dataset was used for the association study. Although the different datasets used resulted in some variations in the association mapping (GWAPP in Figure 4A versus GWA-Portal in Figure 1C), the same most significant SNP (Chr5: 17,400,077 bp) was identified by both analyses. For measuring the population structure of PHT1;1/2/3 protein abundance, 3,319,069 SNPs (minor allele frequency [MAF] > 0.01) were included in the datasets. For measuring Pi uptake activity, 3,417,097 SNPs (MAF > 0.01) were included. Population structure was analyzed using TASSEL (Bradbury et al., 2007). The trait heritability was calculated from the phenotype and kinship matrix in the PyGWAS package (Provided by Ümit Seren, software engineer of GWA-portal). The LDheatmap R package was used to produce pairwise linkage disequilibrium measurements for SNPs (Shin et al., 2006).

**Dual-luciferase assay**

For the dual-luciferase assay, the promoter of Col-0-PHT1;1 (2-kb upstream from its start codon) was amplified and cloned into the XhoI–BamH1 site of pGreenII0800-LUC vector, which carries the firefly luciferase gene as a reporter and the renilla luciferase gene driven by the mosaic virus 35S promoter for normalization (Hellens et al., 2005). The coding sequence of Col-0-MYB52 fused with 10×Myc was cloned into the pGWBS20 vector. Arabidopsis (Col-0) protoplasts were isolated and transformed with the two types of plasmids according to the method described previously (Sheen, 2001). After 12–16 h of incubation in the medium containing 50-μM KH2PO4, the protoplasts were harvested and the PHT1;1 promoter activity was quantified with a dual-luciferase assay kit (Promega).

**Accession numbers**

Sequence data from this article can be found in the GenBank/EMBL data libraries under the accession numbers listed in Supplemental Table S3.

**Supplemental data**

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

**Supplemental Figure S1.** Phenotypes of Col-0 seedlings grown under different concentrations of Pi.

**Supplemental Figure S2.** PCA estimation of population structure in a genetic dataset.

**Supplemental Figure S3.** Phenotypic correlation between Pi-uptake activity and PHT1;1/2/3 protein abundance.

**Supplemental Figure S4.** A schematic map showing how a candidate gene was selected within the 10-kb region flanking a significant SNP from the association mapping of Pi-uptake activity.

**Supplemental Figure S5.** Pi-related phenotypes of selected T-DNA mutants.

**Supplemental Figure S6.** The transcript levels of candidate genes in Col-0 under different Pi concentrations.

**Supplemental Figure S7.** A snapshot of the genomic region associated with the significant peak on Chr1.

**Supplemental Figure S8.** LD relationships for polymorphisms within and surrounding the PHT1;1 locus.

**Supplemental Figure S9.** Expression levels of candidate genes co-localized with the significant peak identified by conditional GWAS.

**Supplemental Figure S10.** No direct binding of MYB52 to the selected region of PHT1;1 promoter was detected.

**Supplemental Figure S11.** GWAS of Pi-uptake activity among two subgroups of Arabidopsis natural accessions based on SNP locus 2 “T” and “C” on Chr5.

**Supplemental Figure S12.** Functional characterization of two forms of PHT1;1 protein.

**Supplemental Table S1.** Phenotypic values of Pi-uptake activity (A) and square root-transformed (SQRT) PHT1 protein levels (B) for GWAS analysis.

**Supplemental Table S2.** SNPs associated with Pi-uptake activity (A) and PHT1;1/2/3 protein level (B) filtered by a minor allele count (MAC) > 10 and −log10(P) > 4.

**Supplemental Table S3.** A list of the candidate protein-coding genes associated with Pi-uptake activity and PHT1;1/2/3 protein level located within the 10-kb regions flanking the significant SNPs.

**Supplemental Table S4.** List of the T-DNA insertional mutants examined in this work.

**Supplemental Table S5.** P-acquisition phenotypes of F1 hybrids.

**Supplemental Table S6.** LD matrix of SNPs located inside PHT1;1 (A) and near the MYB52 (B) gene body.

**Supplemental Table S7.** List of the primer pairs used in this work.

**Acknowledgments**

We acknowledge the ABRC for providing the Arabidopsis natural accessions. We thank the Plant-Tech Core Facility in the Agricultural Biotechnology Research Center of Academia Sinica for protoplast transformation and Ms. Shu-Chen Shen in the Advanced Optical Microscope Core Facility of Academia Sinica for the operation of the confocal microscope. We are grateful to Ümit Seren, who is the webpage manager of the GWA-Portal, for helpful suggestions on GWAS analysis and technical troubleshooting. We also thank Dr Bernhard Schäfke, a Research Assistant Professor at the Shenzhen Institute of Advanced Technology of the Chinese Academy of Sciences, for constructive discussion, and Dr Zhengrui Wang for technical advice and helpful discussion.
Funding
This work was supported by the Investigator Award (AS-IA-106-L02), Academia Sinica, Taiwan.

Conflict of interest statement. None declared.

References

Ai P, Sun S, Zhao J, Fan X, Xin W, Guo Q, Yu L, Shen Q, Wu P, Miller AJ, et al. (2009) Two rice phosphate transporters, OsPht1:2 and OsPht1:6, have different functions and kinetic properties in uptake and translocation. Plant J 57: 798–809

Alonso-Blanco C, Koornneef M (2000) Naturally occurring variation in Arabidopsis: an underexploited resource for plant genetics. Trends Plant Sci 5: 22–29

Atwell S, Huang YS, Vilhjalmsson BJ, Willems G, Horton M, Li Y, Devaiah BN, Madhuvanthi R, Karthikeyan AS, Raghothama KG (1998) Floral dip: a simplified method for transformation of Arabidopsis. Trends Plant Sci 3: 307–311

Baek D, Kim MC, Chun HJ, Kang S, Park HC, Shin G, Park J, Shen M, Hong H, Kim WY, et al. (2013) Regulation of miR399 transcript accumulation by AtMYB2 affects phosphate starvation responses in Arabidopsis. Plant Physiol 161: 362–373

Bayle V, Arrighi JF, Creff A, Nespoulous C, Vialaret J, Rouached H (2006) Identification and characterization of the plant-specific SEC12-related protein that enables the endoplasmic reticulum exit of a high-affinity phosphate transporter in Arabidopsis. Plant Physiol 141: 956–963

Buckler ES, Selle C, Sikkema K, Zhang Z, Buckler-Edwards V,追踪 K, et al. (2009) A central regulatory system associated with leaf anion concentrations in response to phosphate supply in Arabidopsis thaliana. BMC Plant Biol 9: 410

Chen J, Wang Y, Wang F, Yang J, Gao M, Li C, Liu Y, Liu Y, Yamaji N, Ma JF, et al. (2015) The rice CK2 kinase regulates trafficking of phosphate transporters in response to phosphate levels. Plant Cell 27: 711–723

Chiou TJ, Aung K, Lin SI, Wu CC, Chiang SF, Su CL (2006) Regulation of phosphate homeostasis by microRNA in Arabidopsis. Plant Cell 18: 412–421

Clough SJ, Bent AF (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J 16: 735–743

Dai X, Wang Y, Yang A, Zhang WH (2012) OsMYB2P-1, an R2R3 MYB transcription factor, is involved in the regulation of phosphate-starvation responses and root architecture in rice. Plant Physiol 159: 169–183

Devaiah BN, Karthikeyan AS, Raghothama KG (2007) WRKY75 transcription factor is a modulator of phosphate acquisition and root development in Arabidopsis. Plant Physiol 143: 1789–1801

Devaiah BN, Madhuvanthi R, Karthikeyan AS, Raghothama KG (2009) Phosphate starvation responses and gibberellic acid biosynthesis are regulated by the MYB62 transcription factor in Arabidopsis. Mol Plant 2: 43–58

Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L (2010) MYB transcription factors in Arabidopsis. Trends Plant Sci 15: 573–581

El-Soda M, Neris Moreira C, Goredema-Matongera N, Jamar D, Koornneef M, Aarts MGM (2019) QTTL and candidate genes associated with leaf anion concentrations in response to phosphate supply in Arabidopsis thaliana. BMC Plant Biol 19: 410

Gamuyao R, Chin JH, Pariasca-Tanaka J, Peseares P, Catasus M, Dalid C, Slomet-Loedin I, Tecson-Mendoza EM, Wissuwa M, Heuer S (2012) The protein kinase Psot1 from traditional rice confers tolerance of phosphate deficiency. Nature 488: 535–539

González E, Solano R, Rubio V, Leyva A, Paz-Ares J (2005) PHOSPHATE TRANSPORTER TRAFFIC FACTOR1 is a plant-specific SEC12-related protein that enables the endoplasmic reticulum exit of a high-affinity phosphate transporter in Arabidopsis. Plant Cell 17: 3500–3512

Gu R, Chen F, Long L, Cai H, Liu Z, Yang J, Wang L, Li H, Li J, Liu W, et al. (2016) Enhancing phosphate uptake efficiency through QTL-based selection for root system architecture in maize. J Genet Genom 43: 663–672

Hamburger D, Rezzonico E, Petetot JM-C, Somerville C, Poirier Y (2002) Identification and characterization of the Arabidopsis PHO1 gene involved in phosphate loading to the xylem. Plant Cell 14: 889–902

Hellens RP, Allan AC, Frien EL, Bolitho K, Grafton K, Templeton MD, Karunairetnam S, Gleave AP, Laing WA (2005) Transient expression vectors for functional genomics, quantification of promoter activity and RNA silencing in plants. Plant Methods 1: 13

Horton MW, Hancock AM, Huang YS, Toomajian C, Atwell S, Auton A, Mulliyati NW, Platt A, Sperone FG, Vilhjalmsson BJ, et al. (2012) Genome-wide patterns of genetic variation in worldwide Arabidopsis thaliana accessions from the RegMap panel. Nat Genet 44: 212–216

Huang TK, Han CL, Lin SI, Chen YJ, Tsai YC, Chen YR, Chen JW, Lin WY, Chen PM, Liu TY, et al. (2013) Identification of downstream components of ubiquitin-conjugating enzyme PHOSPHATE2 by quantitative membrane proteomics in Arabidopsis roots. Plant Cell 25: 4044–4060

Hufnagel B, de Sousa SM, Assis L, Guimaraes CT, Leiser W, Azevedo GC, Negri B, Larson BG, Shaff JE, Pastina MM, et al. (2014) Duplicate and conquer: multiple homologs of PHOSPHORUS-STAVRATION TOLERANCE1 enhance phosphorus acquisition and sorghum performance on low-phosphorus soils. Plant Physiol 165: 659–677

Iwata Y, Ashida M, Hasegawa C, Tabara K, Mishiba KI, Koizumi N (2017) Activation of the Arabidopsis membrane-bound transcription factor bZIP28 is mediated by site-2 protease, but not site-1 protease. Plant J 91: 408–415

Iwata Y, Fedoroff NV, Koizumi N (2009) The Arabidopsis membrane-bound transcription factor AtbZIP60 is a novel plant-specific endoplasmic reticulum stress transducer. Plant Signal Behav 4: 514–516

Kang HM, Zahtlen NA, Wade CM, Kirby A, Heckerman D, Daly MJ, Eskin E (2008) Efficient control of population structure in model organism association mapping. Genetics 178: 1709–1723

Kawa D, Julkowski MM, Sommerfeld HM, Ter Horst A, Haring MA, Testerink C (2016) Phosphate-dependent root system architecture responses to salt stress. Plant Physiol 172: 690–706

Kim S, Plagnol V, Hu TT, Toomajian C, Clark RM, Ossowski S, Ecker JR, Weigel D, Nordborg M (2007) Recombination and linkage disequilibrium in Arabidopsis thaliana. Nat Genet 39: 1151–1155

Korte A, Farlow A (2013) The advantages and limitations of trait analysis with GWAS: a review. Plant Methods 9: 1–9

Lei KJ, Lin YM, Ren J, Bai L, Miao YC, An GY, Song CP (2016) Modulation of the phosphate-deficient responses by microRNA156 and its targeted QUAMOSA PROMOTER BINDING PROTEIN-LIKE 3 in Arabidopsis. Plant Cell Physiol 57: 192–203

PLANT PHYSIOLOGY 2022: 190; 682–697 | 695
Lin WD, Liao YY, Yang TJ, Pan CY, Buckhout TJ, Schmidt W (2011) Coexpression-based clustering of Arabidopsis root genes predicts functional modules in early phosphate deficiency signaling. Plant Physiol 155: 1383–1402

Lin WY, Huang TK, Chiou TJ (2013) Nitrogen limitation adaptation, a target of microRNA827, mediates degradation of plasma membrane-localized phosphate transporters to maintain phosphate homeostasis in Arabidopsis. Plant Cell 25: 4061–4074

Liu TY, Aung K, Tseng CY, Chang TY, Chen YS, Chiou TJ (2011) Vacular Ca2+/H+ transport activity is required for systemic phosphate homeostasis involving shoot-to-root signaling in Arabidopsis. Plant Physiol 156: 1176–1189

Liu TY, Huang TK, Yang SY, Hong YT, Huang SM, Wang FN, Chiang SF, Tsai SY, Lu WC, Chiou TJ (2016) Identification of plant vacuolar transporters mediating phosphate storage. Nat Commun 7: 11095

Luo B, Ma P, Nie Z, Zhang X, He X, Ding X, Feng X, Lu Q, Ren Z, Lin H, et al. (2019) Metabolite profiling and genome-wide association studies reveal response mechanisms of phosphorus deficiency in maize seeding. Plant J 97: 947–969

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

Ma Z, Baskin TI, Brown KM, Lynch JP (2003) Regulation of root elongation under phosphorus stress involves changes in ethylene responsiveness. Plant Physiol 131: 1381–1390

Misson J, Raghothama KG, Jain A, Jouhet J, Block MA, Bligny R, Ortet P, Creff A, Somerville S, Rolland N, et al. (2005) A genome-wide transcriptional analysis using Arabidopsis thaliana Affymetrix gene chips determined plant responses to phosphate deprivation. Proc Natl Acad Sci USA 102: 11934–11939

Misson J, Thibaud MC, Bechtold N, Raghothama K, Nussaume L (2004) Transcriptional regulation and functional properties of Arabidopsis Pho1, a high affinity transporter contributing greatly to phosphate uptake in phosphate deprived plants. Plant Mol Biol 55: 727–741

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

Naumann C, Muller J, Sakhonsawee S, Wieghaus A, Hause G, Heisters M, Burstenbinder K, Abel S (2019) The local phosphate deficiency response activates endoplasmic reticulum stress-dependent autophagy. Plant Physiol 179: 460–476

Nordborg M, Weigel D (2008) Next-generation genetics in plants. Nature 456: 720–723

Nussaume L, Kanno S, Javot H, Marin E, Pochon N, Ayadi A, Nakashima TM, Thibaud MC (2011) Phosphate import in plants: focus on the PHt1 transporters. Front Plant Sci 2: 83

Pandey BK, Mehra P, Giri J (2013) Phosphorus starvation response in plants and opportunities for crop improvement. In N Tuteja, SS Gill, eds, Climate Change and Abiotic Stress Tolerance. Wiley-VCH: 69–99

Park MY, Kang JY, Kim SY (2011) Overexpression of AtMYB52 confers ABA hypersensitivity and drought tolerance. Mol Cells 31: 447–454

Pireyre M, Burow M (2015) Regulation of MYB and bHLH transcription factors: a glance at the protein level. Mol Plant 8: 378–388

Raghothama KG (1999) Phosphate acquisition. Annu Rev Plant Physiol Plant Mol Biol 50: 665–693

Raymond M, Svistoonoff S, Loudet O, Nussaume L, Desnos T (2006) Identification of QTL controlling root growth response to phosphate starvation in Arabidopsis thaliana. Plant Cell Environ 29: 115–125

Rubio V, Linhares F, Solano R, Martin AC, Iglesias J, Leyva A, Paz-Ares J (2001) A conserved MYB transcription factor involved in phosphate starvation signaling both in vascular plants and in unicellular algae. Genes Dev 15: 2122–2133

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

Sakuraba Y, Kanno S, Mabuchi A, Monda K, Iba K, Yanagisawa S (2018) A phytochrome-B-mediated regulatory mechanism of phos- phorus acquisition. Nat Plants 4: 1089–1101

Segura V, Vijnhalmsso Bj, Platt A, Korte A, Seren U, Long Q, Nordborg M (2012) An efficient multi-locus mixed-model approach for genome-wide association studies in structured populat- ions. Nat Genet 44: 825–830

Seren U (2018) GWA-portal: genome-wide association studies made easy. In D Ristova, E Barbez, eds, Root Development: Methods and Protocols. Springer New York, New York, NY, pp 303–319

Seren U, Vijnhalmsso Bj, Horton MW, Meng D, Forai P, Huang YS, Long Q, Segura V, Nordborg M (2012) GWAPP: a web application for genome-wide association mapping in Arabidopsis. Plant Cell 24: 4793–4805

Sheen J (2001) Signal transduction in maize and Arabidopsis mesophyll protoplasts. Plant Physiol 127: 1466–1475

Shin H, Shin HS, Dewbre GR, Harrison MJ (2004) Phosphate transport in Arabidopsis: Pho1 and Pho14 play a major role in phosphate acquisition from both low- and high-phosphate environments. Plant J 39: 629–642

Shin JH, Bays S, Mcnney B, Graham J (2006) LDheatmap: an R function for graphical display of pairwise linkage disequilibria between single nucleotide polymorphisms. J Stat Softw 16: 1–9

Su T, Xu Q, Zhang FC, Chen Y, Li LQ, Wu WH, Chen YF (2015) WRKY42 modulates phosphate homeostasis through regulating phosphate translocation and acquisition in Arabidopsis. Plant Physiol 167: 1579–1591

Sun L, Song L, Zhang Y, Zheng Z, Liu D (2016) Arabidopsis PHL2 and PHR1 act redundantly as the key components of the central regulatory system controlling transcriptional responses to phos- phate starvation. Plant Physiol 170: 499–514

Töth G, Guicharnaud R-A, Töth B, Herrmann T (2014) Phosphorus levels in croplands of the European Union with implications for P fertilizer use. Eur J Agron 55: 42–52

Tan ST, Xue HW (2014) Casein Kinase 1 regulates ethylene synthesis by phosphorylating and promoting the turnover of ACCS. Cell Rep 9: 1692–1702

The 1001 Genomes Consortium (2016) 1,135 genomes reveal the global pattern of polymorphism in Arabidopsis thaliana. Cell 166: 481–491

Theodorou ME, Plaxton WC (1993) Metabolic adaptations of plant respiration to nutritional phosphate deprivation. Plant Physiol 101: 339–344

Togninalli M, Seren U, Meng D, Fitz J, Nordborg M, Weigel D, Borgwardt K, Korte A, Grimmel D (2018) The AraGWAS Catalog: a curated and standardized Arabidopsis thaliana GWAS catalog. Nucleic Acids Res 46: D1150–D1156

Togninalli M, Seren U, Freudenthal JA, Monroe JG, Meng D, Nordborg M, Weigel D, Borgwardt K, Korte A, Grimmel D (2020) AraPheno and the AraGWAS Catalog 2020: a major data- base update including RNA-Seq and knockout mutation data for Arabidopsis thaliana. Nucleic Acids Res 48: D1063–D1068

Verslues PE, Lasky JR, Juenger TE, Liu TW, Kumar MN (2014) Genome-wide association mapping combined with reverse genetics identifies new effectors of low water potential-induced proline ac- cumulation in Arabidopsis. Plant Physiol 164: 144–159

Wang H, Xu Q, Kong YH, Chen Y, Duan JY, Wu WH, Chen YF (2014) Arabidopsis WRKY45 transcription factor activates PHOSPHATE TRANSPORTER1 expression in response to phosphate starvation. Plant Physiol 164: 2020–2029

Wang L, Li Z, Qian W, Guo W, Gao X, Huang L, Wang H, Zhu H, Wu JW, Wang D, et al. (2011) The Arabidopsis purple acid phosphatase AtPAP10 is predominantly associated with the root surface and plays an important role in plant tolerance to phosphate limitation. Plant Physiol 157: 1283–1299

Wang S, Yin L, Tanaka H, Tanaka K, Tsujimoto H (2010a) Identification of wheat alien chromosome addition lines for breeding wheat with high phosphorus efficiency. Breed Sci 60: 371–379
Wang X, Wang HF, Chen Y, Sun MM, Wang Y, Chen YF (2020) The transcription factor NIGT1.2 modulates both phosphate uptake and nitrate influx during phosphate starvation in Arabidopsis and maize. Plant Cell 32: 3519–3534
Wang X, Yan X, Liao H (2010b) Genetic improvement for phosphorus efficiency in soybean: a radical approach. Ann Bot 106: 215–222
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
Yang J, Ferreira T, Morris AP, Medland SE, Genetic Investigation of ATC, Replication DIG, Meta-analysis C, Madden PA, Heath AC, Martin NG, (2012) Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nat Genet 44: 369–375, S361–363
Yang WT, Baek D, Yun DJ, Lee KS, Hong SY, Bae KD, Chung YS, Kwon YS, Kim DH, Jung KH, et al. (2018) Rice OsMYBSP improves plant phosphate acquisition by regulation of phosphate transporter. PLoS One 13: e0194628
Yang Z, Yang J, Wang Y, Wang F, Mao W, He Q, Xu J, Wu Z, Mao C (2020) PROTEIN PHOSPHATASE95 regulates phosphate homeostasis by affecting phosphate transporter trafficking in rice. Plant Cell 32: 740–757
Yi C, Wang X, Chen Q, Callahan DL, Fournier-Level A, Whelan J, Jost R (2021) Diverse phosphate and auxin transport loci distinguish phosphate tolerant from sensitive Arabidopsis accessions. Plant Physiol 187: 2656–2673
Zhang D, Song H, Cheng H, Hao D, Wang H, Kan G, Jin H, Yu D (2014) The acid phosphatase-encoding gene GmACP1 contributes to soybean tolerance to low-phosphorus stress. PLoS Genet 10: e1004061
Zhong R, Lee C, Zhou J, McCarthy RL, Ye ZH (2008) A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in Arabidopsis. Plant Cell 20: 2763–2782
Zhu XF, Wang ZW, Wan JX, Sun Y, Wu YR, Li GX, Shen RF, Zheng SJ (2015) Pectin enhances rice (Oryza sativa) root phosphorus remobilization. J Exp Bot 66: 1017–1024