1 **Phosphate transporter PHT1;1 as a key determinant of phosphorus** 2 **acquisition in *Arabidopsis* natural accessions**

3

4 Pei-Shan Chien^a, Ya-Ting Chao^{a,1}, Chia-Hui Chou^{b,1}, Yu-Ying Hsu^a, Su-Fen Chiang^a,
5 Chih-Wei Tung^{b,2} and Tzyy-Jen Chiou^{a,2}

6

7 ^aAgricultural Biotechnology Research Center, Academia Sinica, Taipei, Taiwan

8 ^bDepartment of Agronomy, National Taiwan University, Taipei, Taiwan

9

10 ¹These authors contributed equally to this work.

11

12 ²Corresponding authors

13 Chih-Wei Tung

14 Email: chihweitung@ntu.edu.tw

15 ORCID ID: 0000-0001-5337-0177

16 Tzyy-Jen Chiou

17 Email: tjchiou@gate.sinica.edu.tw

18 ORCID ID: 0000-0001-5953-4144

19

20

21 **Running title**

22 Genetic variation governs phosphorus acquisition

23

24 **One sentence summary**

25 Stepwise GWAS analyses reveal insights into the genetic basis in regulating

26 phosphorus acquisition and associations between the phosphate transporter *PHT1;1*

27 haplotype and *Arabidopsis* habitats.

28 ABSTRACT

29 To understand the genetic basis in governing phosphorus (P) acquisition, we
 30 performed genome-wide association studies (GWAS) on a diversity panel of
 31 *Arabidopsis thaliana* by two primary determinants of P acquisition, phosphate
 32 (Pi)-uptake activity and PHOSPHATE TRANSPORTER 1 (PHT1) protein abundance.
 33 Association mapping revealed a shared significant peak on chromosome 5 (Chr5)
 34 where the *PHT1;1/2/3* genes reside, suggesting a strong correlation between the
 35 regulation of Pi-uptake activity and PHT1 protein abundance. Genes encoding
 36 transcription factors, kinases, and a metalloprotease associated with both traits were
 37 also identified. Conditional GWAS followed by statistical analysis of
 38 genotype-dependent expression of *PHT1;1* and transcription activity assays revealed
 39 an epistatic interaction between *PHT1;1* and *MYB DOMAIN PROTEIN 52* (*MYB52*) on
 40 Chr1. Analyses of F1 hybrids generated by crossing two subgroups of natural
 41 accessions carrying specific SNPs associated with *PHT1;1* and *MYB52* further
 42 revealed the strong effects of potential variants on *PHT1;1* expression and Pi
 43 uptake activity. Notably, the soil P contents in *A. thaliana* habitats were found to
 44 coincide with *PHT1;1* haplotype, underscoring how fine-tuning of the activity of P
 45 acquisition by natural variants allows plants to adapt to their environments. This
 46 study sheds light on the complex regulation of P acquisition and offers a framework
 47 to systematically assess the effectiveness of GWAS approaches in the study of
 48 quantitative traits.

49 INTRODUCTION

50 Plants have evolved an array of strategies to sustain growth across diverse habitats
 51 where soil nutrients differ widely. Phosphate (Pi), the primary form of phosphorus (P)
 52 acquired by plants, is a limited resource because of its low solubility and mobility in
 53 soil (Theodorou and Plaxton, 1993; Raghothama, 1999). Modulation of P transport
 54 systems (Misson et al., 2004; Shin et al., 2004; Ai et al., 2009), root system
 55 architecture (Lynch, 2011; Pandey et al., 2013), secretion of root exudates (Wang et
 56 al., 2011), and reallocation and recycling of internal P (Raghothama, 1999) are known
 57 to be the major adaptive responses plants used to cope with low availability of P in
 58 the environments. Remarkable progress has been made in understanding the genetic
 59 and molecular mechanisms that control these responses in *Arabidopsis thaliana* (*A.*
 60 *thaliana*), particularly in the Columbia (Col-0) accession. For example, PHOSPHATE
 61 STARVATION RESPONSE 1 (PHR1) and PHR1-LIKE 1 (PHL1), which are MYB-coiled coil
 62 (CC) type transcriptional activators, are known to play major roles in the activation of
 63 P starvation-inducible genes (Rubio et al., 2001; Bustos et al., 2010; Sun et al., 2016).
 64 Members of the PHOSPHATE TRANSPORTER 1 (PHT1), PHT5, and PHOSPHATE 1
 65 (PHO1) transporter families were found to mediate P acquisition, storage, and
 66 root-to-shoot translocation, respectively (Hamburger et al., 2002; Misson et al., 2004;
 67 Shin et al., 2004; Liu et al., 2016). The activity of PHT1 is increased by its trafficking to
 68 the plasma membrane through *PHOSPHATE TRANSPORTER TRAFFIC FACILITATOR1*
 69 (PHF1), but phosphorylation suppresses its trafficking (González et al., 2005; Bayle et
 70 al., 2011; Chen et al., 2015; Yang et al., 2020). Moreover, several regulatory modules
 71 of miRNAs and their target genes, such as miR399-PHO2 (a ubiquitin-conjugating E2
 72 enzyme), miR827-Nitrogen Limitation Adaptation (NLA; a RING type ubiquitin E3
 73 ligase), and miR156-SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 3 (transcription

74 factor) are intertwined to regulate Pi transport (Huang et al., 2013; Lin et al., 2013;
75 Lei et al., 2016).

76 Many important agronomic traits are controlled by quantitative trait loci (QTLs).
77 Identification of QTLs associated with a trait of interest and subsequent analyses of
78 candidate genes often offer new strategies for crop improvement, including
79 improvement of the use efficiency of nutrient ions. For example, using a QTL
80 mapping strategy, (Reymond et al., 2006) identified *LOW PHOSPHATE ROOT1*, which
81 explained 52% of the variation in primary root growth response to P deficiency in an
82 *A. thaliana* Bay-0 × Shahdara recombinant inbred line population. In rice,
83 *Phosphorus-Starvation Tolerance 1 (PSTOL1)*/Phosphorus uptake1, which encodes a
84 protein kinase, is an important gene/QTL that enhances grain yield by promoting root
85 growth (Wissuwa et al., 2002; Gamuyao et al., 2012). The absence of *PSTOL1* from
86 modern rice cultivars highlights the importance of exploring the diversity of natural
87 variants for breeding crops with improved P acquisition ability (Gamuyao et al., 2012).
88 Tremendous efforts have also been made to dissect the genetic basis of various
89 P-related traits in other crops (Wang et al., 2010; Wang et al., 2010; Hufnagel et al.,
90 2014; Gu et al., 2016). Despite the identification of several potential P use efficiency
91 QTLs, only few have been functionally validated (Gamuyao et al., 2012; Zhang et al.,
92 2014; Luo et al., 2019).

93 *A. thaliana*, which is grown as an inbred line and for which there is a wealth of
94 natural accessions with publicly available whole genome sequences, is an attractive
95 model for QTL identification using the genome-wide association (GWAS) approach
96 (Alonso-Blanco and Koornneef, 2000; The 1001 Genomes Consortium, 2016). There
97 have been several recent studies related to P deficiency in *A. thaliana* accessions,
98 such as identification of candidate loci regulating responses to a combination of

99 salinity and P starvation (Kawa et al., 2016), root growth under P and iron limitation
 100 (Bouain et al., 2019), and leaf anion homeostasis under different P supplies (El-Soda
 101 et al., 2019). These analyses highlight the power of using GWAS in *A. thaliana* to
 102 decipher new players in P stress responses. However, this approach may have some
 103 potential drawbacks. These include false positive associations resulting from cryptic
 104 population structure and kinships and false negatives caused by overcorrection of
 105 confounding factors or insufficient statistical power to detect rare alleles (Nordborg
 106 and Weigel, 2008). Thus, to validate the biological roles of detected loci, systematic
 107 follow-up analyses are necessary (Verslues et al., 2014; Bouain et al., 2019).

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

122

123 RESULTS

124 ***GWAS of Pi-uptake activity and PHT1;1/2/3 protein abundance under*** 125 ***moderately low Pi***

126 Two panels of *A. thaliana* 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 the medium containing different concentrations of Pi (KH₂PO₄) ranging from 10 to 132 250 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 [³³P]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.

145 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

148 GWA-Portal database (Seren et al., 2012; The 1001 Genomes Consortium, 2016;
149 Seren, 2018), 264 and 178 accessions were used for GWAS analysis of Pi-uptake
150 activity and PHT1;1/2/3 protein level, respectively. These traits were normally
151 distributed (Figure 1, A and B; Supplemental Table S1), suggesting that both are
152 quantitative traits determined by multiple genes. Principal component analysis (PCA)
153 of genotypic data revealed that PC1 only explained 3.15% and 2.9% of the total
154 variation in the two populations analyzed for PHT1;1/2/3 protein level and Pi-uptake
155 activity, respectively (Supplemental Figure S2), suggesting that the influence of
156 population structure in detecting genotype-trait association is negligible. Therefore,
157 we employed the accelerated mixed model (AMM) implemented in GWA-Portal
158 (Seren et al., 2012) for our association analysis, in which kinship is considered to
159 effectively control confounding signals due to relatedness, and the variance
160 components are estimated to obtain the overall pseudo-heritability. Based on the
161 phenotype and kinship matrix, we calculated the heritability and estimated how
162 much variation was contributed by genotypic variation. Heritabilities of 0.388 and
163 0.735 were estimated for Pi-uptake activity and PHT1;1/2/3 protein level,
164 respectively. As shown in the Manhattan plots, several major SNP clusters were
165 associated with Pi-uptake activity, but only one cluster with $-\log_{10}(P) > 6$ was
166 identified for PHT1;1/2/3 protein level (Figure 1, C and D). Compared with PHT1
167 protein abundance, it is conceivable that the degree of complexity in controlling
168 Pi-uptake activity is much higher, which involves other factors in addition to Pi
169 transporters, such as root architecture and root exudate. Remarkably, both traits
170 shared a significant peak on chromosome 5 (Chr5), suggesting that these two traits
171 are controlled by the same locus. Intriguingly, *PHT1;1/2/3*
172 (*AT5G43350/AT5G43370/AT5G43360*) are located within the SNP peak on Chr5. This

173 observation shows the effectiveness of our experimental design and the importance
174 of PHT1 family members for natural variation in Pi-uptake activity.

175

176 Identification of candidate genes associated with P-acquisition traits

177 To identify the candidate genes located in the regions adjacent to the significant
178 peaks, we first defined significant SNPs and then searched for trait-associated
179 candidate genes. The SNPs with a minor allele count (MAC) > 10 and a $-\log_{10}(P) > 4$,
180 which was the cutoff based on the deviation of observed $-\log_{10}(P)$ from expected $-\log_{10}(P)$ in the quantile-quantile (Q-Q) plots (Figure 1, C and D, right panel), were
181 selected. Using these criteria, 478 and 508 significant SNPs were identified for
182 Pi-uptake activity and PHT1;1/2/3 protein abundance, respectively (Figure 2;
183 Supplemental Table S2, A and B). As the decay of linkage disequilibrium (LD) in *A.*
184 *thaliana* accessions is within 10 kb on average (Kim et al., 2007), we selected 462 and
185 282 protein-coding genes located within the 10-kb regions flanking the significant
186 SNPs as potential candidates (Figure 2; Supplemental Table S3). Among them, 450
187 and 270 genes were unique for Pi-uptake activity and PHT1;1/2/3 protein level,
188 respectively, and 12 genes were associated with both traits, including 4 Pi
189 transporters, PHT1;1, PHT1;2, PHT1;3, and PHT1;6 (Figure 2; Supplemental Table S3).

190
191 To prioritize candidate genes for further functional assays, we assigned a value,
192 denoted as “hit frequency”, based on the number of times a gene overlapped with
193 the 10-kb region surrounding a significant SNP (Supplemental Figure S3). For instance,
194 the hit frequency of *PHT1;1* was calculated to be 61 for uptake activity (Supplemental
195 Table S3), meaning that any part of the *PHT1;1* gene body is within the 20-kb
196 windows centered on 61 significant SNPs (Supplemental Figure S3). Genes were
197 selected using the following criteria: (1) genes were correlated with either Pi-uptake

activity or PHT1;1/2/3 protein level and had a hit frequency >10; or (2) genes were correlated 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 correlated with both were identified (Figure 2; Supplemental Table S3).

In line with the previous studies (Shin et al., 2004; Huang et al., 2013), *pht1;1* single mutant showed a significant and striking reduction in PHT1;1/2/3 protein level, Pi-uptake activity (Supplemental Figure S4), and Pi concentration (Figure 3) under a moderately low Pi condition (50 μ M). In addition, we analyzed 97 available T-DNA insertional mutants covering 43 out of 91 candidate genes to evaluate their P-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 gene mutants (Supplemental Figure S4, 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-DEPENDENT MALATE DEHYDROGENASE 2*; *C-NAD-MDH2*), *AT5G43320* (*CASEIN KINASE I-LIKE 8*; *CKL8*), *AT1G17950* (*MYB DOMAIN PROTEIN 52*; *MYB52*), and *AT2G27610* (tetratrico peptide 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, A and B). 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 *A. thaliana* 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 KH₂PO₄) and low Pi (50 μM KH₂PO₄) conditions (Supplemental Figure S5). Under sufficient Pi conditions, the RNA levels of most candidate genes were higher in the shoots than in the roots except that the *MYB52* transcription factor was mainly expressed in the roots and *GER2*, which encodes an enzyme involved in the biosynthesis of GDP-fucose, 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.

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), to identify additional loci, we performed conditional GWAS implemented in GWAPP (GWAS-Web-App, <http://gwas.gmi.oeaw.ac.at>; Seren et al., 2012), a predecessor of GWA-Portal. With that, the effect from a specific causal SNP could be included in the model to increase the power to detect other causal markers in a structured dataset when using AMM (Seren et al., 2012). This approach has been successfully applied in flowering phenotype and other GWAS studies (Atwell et al., 2010; Yang et al.,

2012). After masking the effect of the most significant SNP on Chr5 (17,400,077 bp; the allele is C or T) (Figure 4A), a novel SNP peak on Chr1 became evident (Figure 4B). This result implies a biological interaction between the genes where the significant SNPs on Chr1 and Chr5 are located. The most significant SNP is located on *HOMEODOMAIN GLABROUS 12* (*HDG12*; *AT1G17920*), which belonged to an HD-ZIP IV family. However, the T-DNA mutants of *HDG12* did not show differences in Pi concentrations from WT controls. We thus designated the second significant SNP on Chr1 located at 186 bp downstream of the 3'-UTR of *MYB52* (*AT1g17950*), whose association with Pi accumulation in roots was identified in our analysis (Figure 3A), as SNP locus 1 (6,179,475 bp; the allele is A or T; Supplemental Figure S6) and the one on Chr5 located in the third exon of *PHT1;1* as SNP locus 2 (Supplemental Figure S7A). There are four possible allele combinations, {A,C}, {A,T}, {T,C} and {T,T}, of these two significant SNP loci.

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 6 annotated candidate genes surrounding the SNP locus 1 (Supplemental Figure S6A; the lower panel) and *PHT1;1/2/3* genes (SNP locus 2) in each accession. 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 S8B). As revealed by two-way ANOVA, *MYB52* expression was significantly higher in the accessions with allele “A” than in those with allele “T” at SNP locus 1 ($P < 1.00E-04$), 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* ($P = 1.60E-02$) (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* ($P = 4.16E-01$) or *PHT1;3* ($P = 6.99E-01$) (Supplemental Figure S8, 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 supporting 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 S9). 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.

298

299 ***{A,C} genotype is correlated to the higher *PHT1;1* expression and the superior***
 300 ***Pi uptake activity***

301 Given that the accessions with the {A,C} allele combination exhibited the
 302 highest *PHT1;1* gene expression among the four combinations (Figure 4E), we
 303 were curious how P acquisition is affected when one genomic copy of the {A,T}
 304 allele is replaced by one copy of the allele from the {A,C} genetic background.
 305 Thus, we crossed {A,T} accessions (Cnt-1, Shigu-2, An-1, Alst-1, CIBC2, Ba-1, and
 306 Ws) or {A,C} accessions (Paw-3, Ciste-1, Voeran-1, Sei-0, Ta-0, and Ga-2) to Col-0,
 307 which is also {A,C} (Table 1). Although we have yet to dissect potential factors at
 308 other genomic loci that may also contribute to the phenotype, we observed
 309 significantly increased *PHT1;1* expression and Pi-uptake activity relative to the
 310 maternal {A,T} accessions when one copy of the {A,C} genome was introduced
 311 into the {A,T} background. For instance, F1 plants from Cnt-1 X Col-0, which are
 312 {A,T}X{A,C} hybrids, showed a 7.19-fold increase in *PHT1;1* gene expression and a
 313 1.57-fold increase in Pi uptake activity when compared with the maternal line
 314 Cnt-1. This phenomenon was not observed in the {A,C}X{A,C} F1 hybrid (Table 1).
 315 We surmise that the varied P acquisition ability in the {A,C}X{A,C} hybrids may
 316 cause by heterosis or hybrid weakness. These results showed that introduction of
 317 one copy of the {A,C} genome can improve the P acquisition of the {A,T}
 318 accessions, probably through the introduction of functional *PHT1;1*.

319 We also site-by-site generated another F1 population by crossing a homozygous
 320 *pht1;1* knockout mutant (Col-0 background) with various {A,T}- and {A,C}-accessions,
 321 the accessions selected above (Table 1). We found that the *PHT1;1* expression was
 322 rescued partially in both {A,T}X*pht1;1* and the {A,C}X*pht1;1* hybrids (Table 1),

323 indicating that *PHT1;1* introduced from either {A,T}- or {A,C}-genotype is functional.
 324 The partial complementation may result from one copy of functional *PHT1;1* (i.e.
 325 *pht1;1*^{Col-0}/*PHT1;1*^{A,T} or *pht1;1*^{Col-0}/*PHT1;1*^{A,C}). It is interesting to note that the
 326 *PHT1;1* expression in the {A,C}X*pht1;1* hybrid was higher than that in the
 327 {A,T}X*pht1;1* hybrid. However, the difference in the Pi-uptake activity between these
 328 two hybrids was not evident, implying other effects from unidentified loci complex
 329 with *PHT1;1* expression.

330 The results of F1 analysis support the genetic interaction between SNP locus 1 and
 331 locus 2: (1) *PHT1;1* gene expression increased significantly in the {A,C}X*pht1;1*
 332 hybrids than in the {A,T}X*pht1;1* hybrids, and (2) the performance for P acquisition is
 333 significantly increased in the {A,T}X{A,C} hybrids but not the {A,C}X{A,C} hybrids.
 334 Conclusively, {A,C} genotype is correlated to the higher *PHT1;1* expression and the
 335 superior Pi uptake activity.

336 Because *PHT1;1* is a major determinant of Pi-uptake activity and its expression
 337 is highly dependent on the allelic variation at SNP locus 2, i.e., the accessions
 338 with allele “C” had significantly higher transcript levels than those with allele “T”
 339 (Figure 4E), we divided the *A. thaliana* diversity panel into two populations
 340 based on the allelic variation at SNP locus 2 (either “C” or “T”) and searched for
 341 additional loci controlling Pi-uptake activity independent of *PHT1;1*. Using these
 342 two artificial populations composed of 153 accessions with allele “C” and 111
 343 accessions with allele “T”, we conducted GWAS for Pi-uptake activity. Although a
 344 few SNP peaks were observed in the Manhattan plots, none were significant
 345 ($-\log_{10}(P) > 6$, based on the threshold used in Figure 1, C and D) (Supplemental Figure
 346 S10). Therefore, Pi-uptake activity is mainly controlled by SNP locus 2 on Chr5.

347 The effect of other QTLs may not be substantial enough to be detected by this
348 approach.

349

350 ***The PHT1 haplotype as a hallmark linking P-acquisition traits and the soil P***
351 ***contents in A. thaliana habitats***

352 Because SNP locus 2 is a synonymous SNP, we searched for non-synonymous SNPs
353 and found four nearly consecutive significant SNPs in the second exon of *PHT1;1*
354 (Chr5: 17,401,478, 17,401,472, 17,401,471, and 17,401,466 bp) denoted as SNP3,
355 SNP4, SNP5, and SNP6, respectively (Supplemental Figure S7B). These SNPs are in
356 high LD ($r^2 = 0.99$ in the group of accessions assayed for PHT1;1/2/3 protein level,
357 Supplemental Table S5A) and explain 9.11% and 10.25% of the genotypic variance in
358 Pi uptake activity and PHT1;1/2/3 protein level, respectively (Supplemental Table S2,
359 A and B). These SNPs code for two forms of the PHT1;1 protein, the Col-0 type
360 Asn-Pro-Glu-Ser-Ala (NPESA) and the non-Col-0 type Asp-Pro-Thr-Ser-Pro (DPTSP)
361 (Supplemental Figure S7B). To investigate whether these non-synonymous SNPs
362 could alter P-acquisition activity, we first inspected the subcellular localization of the
363 proteins encoded by these SNP alleles and found no difference in their plasma
364 membrane localization (Supplemental Figure S11A). Then, we examined the protein
365 level and Pi-uptake activity in *pht1;1* mutants expressing PHT1;1^{NPESA} or PHT1;1^{DPTSP}
366 driven by the native *PHT1;1* promoter (Supplementary Figure S11, B and C). Despite
367 expression variation among independent transgenic lines, in general, both
368 transgenes PHT1;1^{NPESA} and PHT1;1^{DPTSP} could restore the phenotypes (PHT1;1/2/3
369 protein level and Pi-uptake activity) of the *pht1;1* mutant to the WT level
370 (Supplemental Figure S11, B and C). However, no consistent significant differences

371 were observed between plants expressing these transgenes, suggesting that the
372 corresponding nucleotides may serve as markers rather than causative SNPs.

373 SNP locus 2 is in high LD with SNP3/4/5/6 ($r^2 = 0.67$ pairwise in the group of
374 accessions assayed for PHT1 protein level, Supplemental Table S5A). The
375 combination of these five SNPs results in four haplotypes, GACCT, GACCC, AGAGT,
376 and AGAGC (Supplemental Figure S7C), in which GACCT (105 and 68 accessions
377 assayed for Pi-uptake activity and PHT1;1/2/3 protein level, respectively) and AGAGC
378 (139 and 94 accessions assayed for Pi-uptake activity and PHT1;1/2/3 protein level,
379 respectively) are the two major haplotypes. It is of interest to note that the
380 accessions with the “AGAGC” haplotype showed significantly higher Pi-uptake
381 activities ($P = 3.39E-08$) and PHT1;1/2/3 protein levels ($P = 3.64E-08$) than those with
382 the “GACCT” haplotype (Figure 6A and 6B). Notably, when we located the habitats of
383 these natural accessions in Europe, we found a significant association ($P = 1.47E-02$)
384 between these two haplotypes and the edaphic topsoil conditions (Tóth et al., 2014)
385 (Figure 6C). The high Pi-uptake accessions (with the “AGAGC” haplotype) tended to
386 be collected from growing sites containing low soil P while the low Pi-uptake
387 accessions (with the “GACCT” haplotype) were collected from high soil P sites (Figure
388 6D). However, the relatively weak correlation with Pi levels suggests that soil P is not
389 the only factor to impose selection pressure on *A. thaliana* populations. In summary,
390 our analyses revealed the importance of the *PHT1;1* haplotype in connecting the
391 P-acquisition activities of natural *A. thaliana* accessions to the soil P contents in their
392 habitats.

393

DISCUSSION

PHT1;1 is a key determinant of P-acquisition activity as evidenced by analysis of A. thaliana natural accessions

In this study, we employed GWAS to analyze the genetic architecture of two major traits representing the capability to acquire P. Figure 7 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), the *PHT1* was also identified as one of the crucial loci conferring natural variations in low Pi tolerance using five Pi starvation response traits. These findings reinforce the importance of the PHT1 family for P acquisition from the perspective of genetic variations in a natural population. Nevertheless, (Sakuraba et al., 2018) employed ³³Pi radioactive imaging to analyze 200 natural accessions of *A. thaliana* but did not identify any QTL markers 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.

Among PHT1;1, PHT1;2 and PHT1;3, PHT1;1 plays the dominant role, possibly because of its relatively high expression level (Supplemental Figure S8, 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 *A. thaliana* natural accessions, revealing that *PHT1;1* is controlled by MYB52, but not

419 *PHT1;2* or *PHT1;3* (Figure 4; Supplemental Figure S8). Importantly, we provide
420 evidence to connect the *PHT1;1* haplotype of natural accessions to the soil P
421 contents in their habitats.

422

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

424 Expression of *A. thaliana PHT1;1* was previously reported to be positively
425 regulated by PHR1 (Rubio et al., 2001), WRKY DNA-BINDING PROTEIN 75 (WRKY75)
426 (Devaiah et al., 2007), WRKY45 (Wang et al., 2014), WRKY42 (Su et al., 2015),
427 phytochrome interacting factors PIF4/PIF5 and ELONGATED HYPOCOTYL 5 (Sakuraba
428 et al., 2018), and NITRATE-INDUCIBLE, GARP-TYPE TRANSCRIPTIONAL REPRESSOR
429 1/HYPERSENSITIVE TO LOW PI-ELICITED PRIMARY ROOT SHORTENING 1.2 (Wang et
430 al., 2020). Our study employed conditional GWAS followed by epistasis analysis and
431 transcription activity assays and identified MYB52 as an additional positive regulator
432 of *PHT1;1* expression (Figures 4 and 5). The observation of reduced shoot Pi
433 concentration in *myb52* mutants (17.5% reduction compared with WT, Figure 3A)
434 further supports this notion. However, we did not observe significant changes in
435 *PHT1;1/2/3* protein level or short-term Pi-uptake activity in the *myb52* mutants
436 (Supplemental Figure S4). We think the reason for this is that the T-DNA insertional
437 mutants generated in a single Col-0 background may not fully represent the
438 complexity of genetic variations found in the diversity panel of *A. thaliana* natural
439 accessions analyzed in this work. In addition, in the *myb52* single mutant, other
440 positive transcriptional activators may compensate *PHT1;1* expression for the loss of
441 MYB52.

442 So far, only a limited number of studies have explored the function of MYB52, and
443 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), 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, *A. thaliana* 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 *MYB52* does not respond to changes of external Pi supply (Supplemental Figure S5), underlining the value of GWAS, in which analysis is independent of gene expression.

457

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

Based on the epistatic analysis of SNP locus 1 and SNP locus 2, the strength of the regulation of *PHT1;1* gene expression could be differentiated by allele combination, 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 (Table 1) revealed the potential for increasing P acquisition ability by introduction of the {A,C} allele. These observations highlighted the importance of SNPs within (or near) the *PHT1;1* gene body in regulating PHT1 protein abundance and P acquisition. Notably, we found that two groups of accessions, discriminated by *PHT1;1* haplotype (SNP3/4/5/6/locus 2), showed different P-acquisition phenotypes and different soil P

469 contents in their habitats. As shown in Figure 6, the “AGAGC” haplotype is found in
 470 high-activity accessions preferentially collected from low-P soil. In contrast, the
 471 “GACCT” haplotype corresponds to low-activity accessions preferentially collected
 472 from high-P soil. This result suggests that soil P level imposes selection pressure on *A.*
 473 *thaliana* populations. The accessions with the “AGAGC” haplotype may adapt better
 474 to poor P soil than those with the “GACCT” haplotype because of improved ability to
 475 acquire P. To our knowledge, this is the first report to show the association between
 476 P-related traits and soil P content.

477

478 ***Discovery of additional regulators of P acquisition***

479 Apart from *PHT1;1/2/3*, we did not identify other well-known genes that regulate
 480 P-acquisition activity, such as *PHR1*, *PHO2*, *NLA*, and *PHF1* (Rubio et al., 2001;
 481 González et al., 2005; Huang et al., 2013; Lin et al., 2013). It is not surprising that
 482 GWAS fails to identify known genes. In a similar attempt, *PHT1* loci is the only
 483 well-known Pi regulatory cluster identified in the GWAS analysis, even though five
 484 P-acquisition-related traits were analyzed (Yi et al., 2021). The detection power is
 485 affected by many factors, such as allele frequency of variants, population size, and
 486 demographic history of the analyzed trait (Korte and Farlow, 2013). Nonetheless, we
 487 identified several new loci potentially involved in P acquisition (Figure 3). For
 488 example, C-NAD-MDH2 can convert oxaloacetate to malate, secreting into the
 489 rhizosphere to liberate Pi from the soil (Ryan et al., 2001). CKL8, a Ser/Thr kinase,
 490 was reported to modify the activity of ACC synthase 5 (Tan and Xue, 2014) and may
 491 regulate root elongation via ethylene under Pi-deprivation (Ma et al., 2003). EGY3
 492 was classified as a site-2 protease-like putative metalloprotease implicated in
 493 endoplasmic reticulum stress (Iwata et al., 2009; Iwata et al., 2017), which could

494 trigger Pi-dependent activation of autophagy (Naumann et al., 2019). The roles of
495 CKL8 and EGY3 in P acquisition require further investigation. It is of interest to note
496 that the expression of these candidate genes is not responsive to changes of Pi
497 supply (Supplemental Figure S5), making it impossible to identify them by
498 transcriptomic analysis (Misson et al., 2005; Liu et al., 2016). Thus, GWAS analysis
499 complements transcriptomic analysis.

500 It is always challenging to validate the result of GWAS. Even with sufficient
501 statistical power to detect genetic variants, the effect contributed by a single
502 SNP-associated candidate gene may be too minor to validate. Furthermore, given the
503 strong correlation of PHT1 to the P-acquisition phenotypes, the contribution by other
504 single associated genes may not be easy to reveal. Therefore, the ineffectiveness of
505 recovering the original phenotypes in the loss-of-function mutants of a single
506 candidate gene (other than PHT1) strengthens the predominant role of PHT1 in
507 determining the P-acquisition phenotypes.

508 Our findings extend current knowledge by uncovering additional genetic regulators
509 of P acquisition and linking haplotypes to nutrient ion contents in the habitats of *A.*
510 *thaliana* accessions. Furthermore, our phenotyping and analytical approaches
511 provide a framework to systematically assess the effectiveness of the GWAS
512 approach for studies of quantitative traits.

513 MATERIALS AND METHODS

514 *Plant materials and growth conditions*

515 Two panels of *A. thaliana* natural accessions (CS76309 and CS76427) used in this
516 study were purchased from the Arabidopsis Biological Resource Center (ABRC).
517 CS76309, which was the largest population of natural accessions available when this
518 project was initiated, is a core set of 360 natural accessions genotyped with a
519 250k-SNP/ high-density tiling array by Justin Borevitz's laboratory. (Horton et al.,
520 2012). CS76427 is a set of 80 natural accessions sequenced using the Illumina GA
521 platform by Detlef Weigel's laboratory as part of the 1001 Genomes Project (Cao et
522 al., 2011). Ninety-seven T-DNA mutants covering 43 candidate genes associated
523 with GWAS traits were obtained from the ABRC and Nottingham Arabidopsis
524 Stock Centre (Supplemental Table S3). Homozygous lines were isolated and
525 analyzed under moderately low Pi conditions (50 μ M KH_2PO_4).

526 Seeds were surface sterilized and germinated on agar plates with one-half
527 strength modified Hoagland nutrient solution containing 250 μ M KH_2PO_4 , 1%
528 sucrose, and 0.8% bactoagar (Chiou et al., 2006). Six-day-old seedlings were
529 transferred to Pi-sufficient (+Pi) and Pi-deficient (−Pi) media supplemented with
530 1 mM and 50 μ M KH_2PO_4 , respectively, and grown for another 6 days. Plants
531 were grown at 22°C under a 16-h photoperiod with cool fluorescent white light
532 at 100–150 $\mu\text{E m}^{-2} \text{sec}^{-1}$.

533

534 *Plasmid construction and plant transformation*

535 For complementation of the *pht1;1* mutant, the open reading frame of *PHT1;1*
536 was cloned into pCR8/GW/TOPO (Invitrogen) and then recombined into the
537 Gateway destination vector pMDC32, in which the constitutive 35S promoter was

538 replaced by the *PHT1;1* promoter (3.3 kb upstream from the start codon).
 539 Modification of the sequence of *PHT1;1* was performed by site-directed
 540 mutagenesis using the Col-0 sequence as the template. The resulting constructs
 541 were designated PHT1;1^{DPTSP} and PHT1;1^{NPESA}. Transgenic plants were generated
 542 using an *Agrobacterium tumefaciens* dipping procedure (Clough and Bent, 1998)
 543 with strain GV3101.

544

545 ***RNA extraction and quantitative RT-PCR analysis***

546 Total RNA from 14-day-old seedlings was isolated using TRIzol reagent
 547 (Invitrogen) followed by DNase I (Ambion) treatment. cDNA was synthesized
 548 from 1 µg total RNA using Moloney murine leukemia virus reverse transcriptase
 549 (Invitrogen) with oligo(dT) primer. Quantitative RT-PCR was performed using the
 550 Power SYBR Green PCR Master Mix Kit (Applied Biosystems) on a 7500 Fast
 551 Real-Time PCR System (Applied Biosystems) according to the manufacturer's
 552 instructions. Relative expression levels were normalized to an internal control,
 553 *Actin8* (Supplemental Table S6). The quantitative PCR primers designed to
 554 examine transcript levels are located at the same position in the selected genes
 555 with a perfect match in each accession.

556

557 ***Western blot analysis***

558 Roots from ten 12-day-old seedlings were collected, ground in liquid nitrogen,
 559 and dissolved in protein lysis buffer containing 2% SDS, 60 mM Tris-HCl, pH 8.5,
 560 2.5% glycerol, 0.13 mM EDTA, and 1× complete protease inhibitor (Roche). Total
 561 proteins (50 µg) were loaded onto 4% to 12% Bis-Tris sodium dodecyl
 562 sulfate-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/>).

570

571 ***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 [^{33}P] KH_2PO_4 for 2 h. The seedlings were grown under the same conditions as samples used for western blot analysis.

580

581 ***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 (SD) 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 *A. thaliana* 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 accelerated mixed model (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. 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

613 calculated from the phenotype and kinship matrix in the PyGWAS package (Provided
614 by Ümit Seren, software engineer of GWA-portal). The LDheatmap R package was
615 used to produce pairwise linkage disequilibrium (LD) measurements for SNPs (Shin et
616 al., 2006).

617

618 ***Dual luciferase assay***

619 For the dual luciferase assay, the promoter of Col-0-*PHT1;1* (2 kb upstream
620 from its start codon) was amplified and inserted into the pGreenII0800-LUC
621 vector, which carries the firefly luciferase gene as a reporter and the renilla
622 luciferase gene driven by the mosaic virus 35S promoter for normalization
623 (Hellens et al., 2005). The coding sequence of Col-0-*MYB52* fused with 10 × Myc
624 was cloned into the pGWB520 vector. *A. thaliana* (Col-0) protoplasts were
625 isolated and transformed with the two types of plasmids according to the
626 method described previously (Sheen, 2001). After 12–16 h of incubation in the
627 medium containing 50 μM KH₂PO₄, the protoplasts were harvested and the
628 *PHT1;1* promoter activity was quantified with a dual-luciferase assay kit
629 (Promega).

630 **Supplemental Data**

631 **Supplemental Figure S1. Phenotypes of Col-0 seedlings grown under different** 632 **concentrations of Pi.**

633 **(A)** Images of 12-day-old seedlings grown on media containing 250 μ M Pi for 6 days,
634 then transferred to media containing different Pi concentrations and grown for
635 another 6 days. Bar = 1.3 cm. Fresh weights of roots **(B)** and shoots **(C)** from 10
636 seedlings. Pi concentration in roots **(D)** and shoots **(E)**. Error bar indicates mean $\pm$ SD
637 ($n = 3$ biologically independent samples). Different letters above bars indicate
638 statistically different groups according to one-way ANOVA followed by the least
639 significant difference test at a probability of $P < 0.05$.

640

641 **Supplemental Figure S2. Principal component analysis (PCA) estimation of** 642 **population structure in a genetic dataset.**

643 An imputed full sequence dataset (The 1001 Genomes Consortium, 2016) was used
644 for PCA. PCA of PHT1;1/2/3 protein abundance **(A)** and Pi-uptake activity **(B)** (see
645 details in “MATERIALS AND METHODS”).

646

647 **Supplemental Figure S3. A schematic map showing how a candidate gene was** 648 **selected within the 10 kb region flanking a significant SNP from the association** 649 **mapping of Pi-uptake activity.**

650 **(A)** Manhattan plot of Pi-uptake activity (also shown in Figure 1C). **(B)** Taking a
651 significant SNP (17,400,077; $-\log_{10}(P) = 7.11$) on Chr 5 as an example, four genes,
652 namely *C-NAD-MDH2* (AT5G43330), *PHT1;6* (AT5G43340), *PHT1;1* (AT5G43350), and
653 *PHT1;3* (AT5G43360) were located within the 20-kb window centered on the SNP;
654 each gene was scored as having one “hit”. Taking *PHT1;1* as an example, the gene

body overlaps with the 20-kb windows of 61 significant SNP hits; thus, the “hit frequency” of *PHT1;1* is 61 (Supplemental Table S3). The threshold for Bonferroni correction provided by GWA-Portal is indicated on the top of the Manhattan plot **(B)** as a red dashed line.

Supplemental Figure S4. Pi-related phenotypes of selected T-DNA mutants.

Seedlings of 12-day-old mutant and wild-type (Col-0) plants grown under 50 μ M Pi (–Pi) conditions were analyzed. **(A)** Relative protein levels of PHT1;1/2/3 in roots, $n = 7$ to 21 biological repeats. **(B)** Relative Pi uptake activity of whole seedlings, $n = 7$ to 10. The results are shown as a box and whisker diagram with dots. The bars within the box plots represent the 25th and 75th percentiles, the center line indicates the median, and the whiskers indicate the minimum and maximum. Each dot represents an individual sample. Significant differences between mutants and WT were determined by a two-sided Student’s *t*-test, and *P*-values were corrected by Benjamini Hochberg multiple testing (* Adjusted *P* < 0.05; ** Adjusted *P* < 0.01).

Supplemental Figure S5. The transcript levels of candidate genes in Col-0 under different Pi concentrations.

Relative mRNA levels of *PHT1;1*, *PHT1;4* **(A)** and selected candidate genes **(B)** in roots and shoots grown under sufficient Pi (+Pi; 1 mM KH₂PO₄, black bars) and low Pi (–Pi; 50 μ M KH₂PO₄, white bars) conditions. Data are presented as mean $\pm$ SE ($n = 4$ biologically independent samples). Significance of differences was determined by two-sided Student’s *t*-test (–Pi versus +Pi treatment, ** *P* < 0.01; * *P* < 0.05).

Supplemental Figure S6. A snapshot of the genomic region associated with the significant peak on Chr1.

(A) The upper panel shows the results of a conditional GWAS of PHT1;1/2/3 protein levels for Chr1 using GWAPP. The genomic map (lower panel) shows the candidate genes surrounding the significant peak on Chr1. **(B)** All identified SNPs within the *MYB52* (*AT1G17950*) (Chr1:6,178,123, 6,178,793, and 6,178,993) gene body are in high LD ($0.8 < r^2 < 1$; Supplemental Table S5B) with SNP locus 1.

Supplemental Figure S7. LD relationships for polymorphisms within and surrounding the *PHT1;1* locus.

(A) The upper panel shows the results from association mapping of PHT1;1/2/3 protein level for Chr5. The lower panel is a genomic map of the significant SNPs, including SNP locus 2, within the *PHT1;1* gene body. **(B)** SNP3 to SNP6, which are located in the second exon of the *PHT1;1* gene, result in two types of PHT1;1 protein sequence: Col-0 (NPESA) and non-Col-0 (DPTSP). **(C)** Four non-synonymous SNPs together with SNP locus 2 form four combinations of haplotypes, among which “AGAGC” (marked in blue) and “GACCT” (marked in light green) are dominant.

Supplemental Figure S8. Expression levels of candidate genes co-localized with the significant peak identified by conditional GWAS.

The mRNA levels of *PHT1;1*, *PHT1;2*, and *PHT1;3* **(A)** and candidate genes **(B)** in roots of different accessions. Two-way ANOVA was performed to test the epistatic relationship between two loci based on averaged mRNA levels ($2^{-\Delta\Delta CT} \times 100$) of *PHT1;2* **(C)** and *PHT1;3* **(D)**. The letters in blue and red shown in **(C)** and **(D)** indicate SNP locus 1 and SNP locus 2, respectively. Relative expression levels of all selected genes

are shown. Data are presented as mean $\pm$ SD (n = 3 biologically independent samples).

Supplemental Figure S9. No direct binding of MYB52 to the selected region of *PHT1;1* promoter was detected.

(A) Schematic illustration of the binding sites of PHR1 (P1BS, green bars) (Rubio et al., 2001) and MYB52 (MBSI and MBSII) (Franco-Zorrilla et al., 2014) in the *PHT1;1* promoter, as analyzed by PlantPAN 2.0 (Chow et al., 2016). **(B)** Results of yeast-one-hybrid analysis conducted using the Gold Yeast One-Hybrid Library Screening System (Matchmaker). PHR1 and MYB52 were cloned into the prey vector pGADT7, and three tandem repeats of P1BS (GAATATCC) and MBSII (ATTGGC) were cloned into the pAbAi vector (Supplementary Table S4). Colony growth on leucine and uracil drop-out medium (-L-U) represents the success of co-transformation. Positive interactions were detected by selection with Aureobasidin A (AbA100 and AbA200 indicate concentrations of 100 and 200 ng/mL, respectively). Interaction between Rec-p53 and P53 served as a control.

Supplemental Figure S10. GWAS of Pi-uptake activity among two subgroups of *A. thaliana* natural accessions based on SNP locus 2 “T” and “C” on Chr5.

Manhattan plots from GWAS of Pi-uptake activity among 111 accessions carrying allele “T” at SNP locus 2 **(A)** or among 153 accessions carrying allele “C” at SNP locus 2 **(B)**. Thresholds for Bonferroni (red dashed line) and Benjamini Hochberg (blue dashed line) correction provided by GWA-Portal are indicated on the top of each plot. The two lines in **(B)** overlap.

Supplemental Figure S11. Functional characterization of two forms of *PHT1;1* protein.

(A) Plasma membrane localization of *PHT1;1* proteins encoded by Col-0 type and non-Col-0 type *PHT1;1* transgenes. Bar = 20 μ m. Relative protein levels of *PHT1;1/2/3* **(B)** and Pi-uptake activities **(C)** under moderately low Pi conditions (50 μ M Pi) among different plants are shown. For *PHT1;1/2/3* protein level, n= 4 to 14 biologically independent samples; for uptake activity, n= 4 to 17. Numbers indicate different independent lines. Vector, pMDC32 vector in which the 35S promoter is substituted with the *PHT1;1* promoter. The results are shown by a box and whisker diagram with dots. The bars within the box plots represent the 25th and 75th percentiles, the central line indicates the median, and the whiskers indicate the maximum and minimum. Each dot represents an individual sample. Different letters above the bars indicate statistical differences from one-way ANOVA followed by the least significant difference test at a probability of $P < 0.05$.

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 $-\log_{10}(P) > 4$.

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.

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

756

757 **Supplemental Table S5. LD matrix of SNPs located inside *PHT1;1* (A) and near the**
 758 ***MYB52* (B) gene body.**

759

760 **Supplemental Table S6. List of the primer pairs used in this work.**

761 **ACKNOWLEDGMENTS**

762 We acknowledge the ABRC for providing the *A. thaliana* natural accessions. We
 763 thank the Plant Tech Core Facility in the Agricultural Biotechnology Research Center
 764 of Academia Sinica for protoplast transformation and Ms. Shu-Chen Shen in the
 765 Advanced Optical Microscope Core Facility of Academia Sinica for operation of the
 766 confocal microscope. We are grateful to Ümit Seren who is the webpage manager of
 767 the GWA-Portal for helpful suggestions on GWAS analysis and technical
 768 troubleshooting. We also thank Dr. Bernhard Schäfke, a Research Assistant Professor
 769 at the Academy for Advanced Interdisciplinary Studies & Department of Biology of
 770 Southern University of Science and Technology, for constructive discussion, and Dr.
 771 Zhengrui Wang for constructive discussion and technical advice. This work was
 772 supported by the Investigator Award (AS-IA-106-L02), Academia Sinica, Taiwan.

773

774

775

776

777 **AUTHOR CONTRIBUTIONS**

778 P.-S. C., C.-W. T., and T.-J. C. conceived and designed the study. P.-S. C., Y.-Y. H., and
 779 S.-F. C. performed the experiments. P.-S. C., C.-H. C., Y.-T. C., C.-W. T., and T.-J. C.
 780 analyzed the data. P.-S. C., T.-J. C., C.-H. C., Y.-T. C., and C.-W. T. wrote the manuscript.

781

782 **Table 1. P-acquisition phenotypes of F1 hybrids.**

Accession ID	Genotype	<i>PHT1;1</i> expression ^a			Uptake activity ^b		
		Maternal line	F1 ^d (X Col-0)	F1 ^d (X <i>pht1;1</i>)	Maternal line	F1 ^d (X Col-0)	F1 ^d (X <i>pht1;1</i>)
An-1	{A,T}	12.3 ± 3.7	18.4 ± 3.8	5.4 ± 0.3	1.4 ± 0.1	1.7 ± 0.1*	1.4 ± 0.1
Ba-1	{A,T}	29.0 ± 6.3	52.4 ± 1.9*	11.2 ± 1.0	1.6 ± 0.0	1.9 ± 0.1*	1.3 ± 0.1*
CIBC2	{A,T}	8.1 ± 1.2	22.8 ± 1.6**	8.8 ± 1.7	1.6 ± 0.0	1.8 ± 0.1*	1.7 ± 0.2
Cnt-1	{A,T}	4.8 ± 1.6	34.5 ± 5.3**	4.6 ± 1.1	1.0 ± 0.0	1.6 ± 0.0**	1.3 ± 0.1**
Shigu-2	{A,T}	21.5 ± 2.2	38.5 ± 3.3**	12.4 ± 1.9*	1.0 ± 0.1	1.7 ± 0.0**	1.4 ± 0.1*
Ws	{A,T}	20.0 ± 1.8	25.3 ± 5.4	15.6 ± 2.9	1.7 ± 0.1	1.7 ± 0.1	1.4 ± 0.1*
Col-0	{A,C}	43.9 ± 6.0	—	23.5 ± 4.3*	1.4 ± 0.1	—	1.2 ± 0.1*
Alst-1	{A,C}	9.3 ± 1.1	47.6 ± 6.5**	16.1 ± 2.9	1.5 ± 0.1	2.0 ± 0.1**	1.9 ± 0.1*
Ciste-1	{A,C}	68.5 ± 9.8	72.8 ± 9.6	29.1 ± 10.6*	1.9 ± 0.1	1.7 ± 0.1	1.6 ± 0.1
Ga-2	{A,C}	104.4 ± 22.8	87.1 ± 8.9	41.9 ± 4.7*	2.3 ± 0.1	2.1 ± 0.2	1.4 ± 0.1**
Paw-3	{A,C}	62.6 ± 5.0	47.3 ± 1.7*	26.7 ± 7.3*	1.5 ± 0.1	1.8 ± 0.1	1.4 ± 0.1
Sei-0	{A,C}	83.9 ± 6.7	47.9 ± 6.6*	33.4 ± 3.3**	2.1 ± 0.1	2.0 ± 0.1	1.7 ± 0.1*
Ta-0	{A,C}	31.8 ± 3.2	50.4 ± 6.0*	23.7 ± 2.9	2.3 ± 0.1	1.9 ± 0.1	1.6 ± 0.1**
Voeran-1	{A,C}	58.8 ± 6.5	31.8 ± 3.5*	25.6 ± 4.5*	1.9 ± 0.0	1.9 ± 0.1	1.7 ± 0.1
<i>pht1;1</i>	{A,C}	0.0 ± 0.0	—	—	1.0 ± 0.1	—	—

783 The indicated trait value was averaged from three to four biological repeats and the value is shown as
784 average ± SE.

785 Asterisks mark values significantly different between the maternal line and F1 hybrids (* $P < 0.05$; ** P
786 < 0.01 , Student's t -test).

787 ^a indicates the relative *PHT1;1* gene expression value ($2^{-\Delta CT} \times 10$).

788 ^b indicates the relative Pi-uptake activity from each accession normalized to the level from the *pht1;1*
789 mutant.

790 ^c Ratio indicates the average Pi trait value of F1 hybrids divided by that of inbred lines.

791 ^d F1 indicates maternal line x Col-0.

792 —, not applicable.

793 **FIGURE LEGENDS**

794 **Figure 1. GWA analysis of two P-acquisition traits.**

795 Phenotypic distribution of raw values for Pi-uptake activity **(A)** and square
796 root-transformed (SQRT) values for PHT1;1/2/3 protein abundance **(B)**. Manhattan
797 plots of Pi-uptake activity **(C)** and PHT1;1/2/3 protein level **(D)**; the respective Q-Q
798 plots are shown on the right. The analyses were conducted using GWA-Portal.
799 Thresholds for Bonferroni (red dashed line) and Benjamini Hochberg (blue dashed
800 line) correction provided by GWA-Portal are indicated. The two lines overlap in **(C)**.

801

802 **Figure 2. GWAS of two Pi-related traits and the step-by-step scoring system used** 803 **for the selection of candidate genes.**

804 In addition to the well-known Pi transporters *PHT1;1* and *PHT1;2*, our approach
805 identified seven candidate genes (e.g., transcription factors, a kinase, an
806 oxidoreductase, and a protease) involved in P acquisition under low Pi.

807

808 **Figure 3. Analyses of Pi-related traits in the seedlings of T-DNA insertional mutants.**

809 Pi concentrations in mutants relative to Col-0 wild type in the roots **(A)** and shoots
810 **(B)**. Twelve-day-old seedlings grown under moderately low Pi conditions (see
811 “MATERIALS AND METHODS”) were sampled. The results are shown in a box and
812 whisker diagram with individual data points superimposed. n = 9 biological repeats.
813 The bars within the box plots represent the 25th and 75th percentiles, the center line
814 represents the median, and the whiskers show the minimum and maximum values.
815 Significant differences between mutants and Col-0 were determined by two-sided
816 Student’s *t*-test, and *P*-values were adjusted for multiple testing using Benjamini–
817 Hochberg correction (* Adjusted *P* < 0.05; ** Adjusted *P* < 0.01). Insertion mutants

for the following genes were tested: *GER2* (*AT1G17890*); *EGY3* (*AT1G17870*); Rhodanese (*AT1G17850*); *C-NAD-MDH2* (*AT5G43330*); *CKL8* (*AT5G43320*); *MYB52* (*AT1G17950*), and tetratrico peptide repeat (TPR)-like superfamily protein (*AT2G27610*).

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 arrowhead in **(A)**) was included as a cofactor (**B**, conditioning) in GWAPP. 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.

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 $\pm$ SE (n = 4 biologically independent samples). The significance of differences was determined by two-sided Student's *t*-test (** *P* < 0.01).

843

844 **Figure 6. *PHT1;1* haplotype connects the P-acquisition phenotypes of *A. thaliana***
845 **accessions to the soil P contents in their habitats.**

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

857

858 **Figure 7. Illustration of our GWAS approach.**

859 **(A)** GWAS of two Pi-related traits. Our approach identified the well-known PHT1 Pi
860 transporters and an additional seven genes potentially involved in regulating P
861 acquisition under low Pi. **(B)** Conditional GWAS of PHT1;1/2/3 protein level revealed
862 a biological interaction between SNP locus 1 and locus 2 and an MYB52-PHT1;1
863 regulatory module. **(C)** The results of heterozygosity analysis revealed improvement
864 of Pi phenotypes by introduction of the {A,C} genome into {A,T} lines. **(D)** Different
865 *PHT1;1* haplotypes discriminate soil P contents in the *A. thaliana* habitats.

Figure 1. GWA analysis 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**).

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.

Figure 3. Analyses of Pi-related traits in the seedlings of T-DNA insertional mutants.

Pi concentrations in mutants relative to Col-0 wild type in the roots (**A**) and shoots (**B**). Twelve-day-old seedlings grown under moderately low Pi conditions (see “MATERIALS AND METHODS”) were sampled. The results are shown in a box and whisker diagram with individual data points superimposed. $n = 9$ biological repeats. The bars within the box plots represent the 25th and 75th percentiles, the center line represents the median, and the whiskers show the minimum and maximum values. Significant differences between mutants and Col-0 were determined by two-sided Student’s t -test, and P -values were adjusted for multiple testing using Benjamini–Hochberg correction (* Adjusted $P < 0.05$; ** Adjusted $P < 0.01$). Insertion mutants for the following genes were tested: *GER2* (AT1G17890); *EGY3* (AT1G17870); *Rhodanese* (AT1G17850); *C-NAD-MDH2* (AT5G43330); *CKL8* (AT5G43320); *MYB52* (AT1G17950), and tetratricopeptide repeat (TPR)-like superfamily protein (AT2G27610).

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 arrowhead in **A**) was included as a cofactor (**B**, conditioning) in GWAPP. 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.

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).

Figure 6. *PHT1;1* haplotype connects the P-acquisition phenotypes of *A. thaliana* 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, while 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 gray lines within violin plots represent the 25th percentile, median, 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.

Figure 7. Illustration of our GWAS approach.

(A) GWAS of two Pi-related traits. Our approach identified the well-known PHT1 Pi transporters and an additional seven genes potentially involved in regulating P acquisition under low Pi. **(B)** Conditional GWAS of PHT1;1/2/3 protein level revealed a biological interaction between SNP locus 1 and locus 2 and an MYB52-PHT1;1 regulatory module. **(C)** The results of heterozygosity analysis revealed improvement of Pi phenotypes by introduction of the {A,C} genome into {A,T} lines. **(D)** Different *PHT1;1* haplotypes discriminate soil P contents in the *A. thaliana* habitats.

Parsed Citations

Ai P, Sun S, Zhao J, Fan X, Xin W, Guo Q, Yu L, Shen Q, Wu P, Miller AJ, Xu G (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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Atwell S, Huang YS, Vilhjalmsdottir BJ, Willems G, Horton M, Li Y, Meng D, Platt A, Tarone AM, Hu TT, Jiang R, Muliyil NW, Zhang X, Amer MA, Baxter I, Brachi B, Chory J, Dean C, Debieu M, de Meaux J, Ecker JR, Faure N, Kniskern JM, Jones JD, Michael T, Nemri A, Roux F, Salt DE, Tang C, Todesco M, Traw MB, Weigel D, Marjoram P, Borevitz JO, Bergelson J, Nordborg M (2010) Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines. Nature 465: 627-631

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Baek D, Kim MC, Chun HJ, Kang S, Park HC, Shin G, Park J, Shen M, Hong H, Kim WY, Kim DH, Lee SY, Bressan RA, Bohnert HJ, Yun DJ (2013) Regulation of miR399f transcription by AtMYB2 affects phosphate starvation responses in Arabidopsis. Plant Physiol. 161: 362-373

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Bayle V, Arrighi JF, Creff A, Nespoulous C, Vialaret J, Rossignol M, Gonzalez E, Paz-Ares J, Nussaume L (2011) Arabidopsis thaliana high-affinity phosphate transporters exhibit multiple levels of posttranslational regulation. Plant Cell 23: 1523-1535

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Bouain N, Korte A, Satbhai SB, Nam HI, Rhee SY, Busch W, Rouached H (2019) Systems genomics approaches provide new insights into Arabidopsis thaliana root growth regulation under combinatorial mineral nutrient limitation. PLoS Genet. 15: e1008392

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Bustos R, Castrillo G, Linhares F, Puga MI, Rubio V, Perez-Perez J, Solano R, Leyva A, Paz-Ares J (2010) A central regulatory system largely controls transcriptional activation and repression responses to phosphate starvation in Arabidopsis. PLoS Genet. 6: e1001102

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Cao J, Schneeberger K, Ossowski S, Gunther T, Bender S, Fitz J, Koenig D, Lanz C, Stegle O, Lippert C, Wang X, Ott F, Muller J, Alonso-Blanco C, Borgwardt K, Schmid KJ, Weigel D (2011) Whole-genome sequencing of multiple Arabidopsis thaliana populations. Nat. Genet. 43: 956-963

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Chen J, Wang Y, Wang F, Yang J, Gao M, Li C, Liu Y, Liu Y, Yamaji N, Ma JF, Paz-Ares J, Nussaume L, Zhang S, Yi K, Wu Z, Wu P (2015) The rice CK2 kinase regulates trafficking of phosphate transporters in response to phosphate levels. Plant Cell 27: 711-723

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Chow CN, Zheng HQ, Wu NY, Chien CH, Huang HD, Lee TY, Chiang-Hsieh YF, Hou PF, Yang TY, Chang WC (2016) PlantPAN 2.0: an update of plant promoter analysis navigator for reconstructing transcriptional regulatory networks in plants. Nucleic Acids Res. 44: D1154-1160

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

15: 573-581

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Franco-Zorrilla JM, Lopez-Vidriero I, Carrasco JL, Godoy M, Vera P, Solano R (2014) DNA-binding specificities of plant transcription factors and their potential to define target genes. Proc. Natl Acad. Sci. USA 111: 2367-2372

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

González E, Solano R, Rubio V, Leyva A, Paz-Ares J (2005) PHOSPHATE TRANSPORTER TRAFFIC FACILITATOR1 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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Hamburger D, Rezzonico E, Petétot 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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Hellens RP, Allan AC, Friel EN, 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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Horton MW, Hancock AM, Huang YS, Toomajian C, Atwell S, Auton A, Mulyati NW, Platt A, Sperone FG, Vilhjálmsson BJ, Nordborg M, Borevitz JO, Bergelson J (2012) Genome-wide patterns of genetic variation in worldwide *Arabidopsis thaliana* accessions from the RegMap panel. Nat. Genet. 44: 212-216

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Hufnagel B, de Sousa SM, Assis L, Guimaraes CT, Leiser W, Azevedo GC, Negri B, Larson BG, Shaff JE, Pastina MM, Barros BA, Weltzien E, Rattunde HF, Viana JH, Clark RT, Falcao A, Gazaffi R, Garcia AA, Schaffert RE, Kochian LV, Magalhaes JV (2014) Duplicate and conquer: multiple homologs of PHOSPHORUS-STARVATION TOLERANCE1 enhance phosphorus acquisition and sorghum performance on low-phosphorus soils. Plant Physiol. 166: 659-677

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Kang HM, Zaitlen 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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Kawa D, JULKOWSKA 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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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 SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 3 in *Arabidopsis*. Plant Cell Physiol. 57: 192-203

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Luo B, Ma P, Nie Z, Zhang X, He X, Ding X, Feng X, Lu Q, Ren Z, Lin H, Wu Y, Shen Y, Zhang S, Wu L, Liu D, Pan G, Rong T, Gao S (2019) Metabolite profiling and genome-wide association studies reveal response mechanisms of phosphorus deficiency in maize seedling. Plant J. 97: 947-969

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Misson J, Raghothama KG, Jain A, Jouhet J, Block MA, Bligny R, Ortet P, Creff A, Somerville S, Rolland N, Doumas P, Nacry P, Herrera-Estrella L, Nussaume L, Thibaud MC (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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Misson J, Thibaud MC, Bechtold N, Raghothama K, Nussaume L (2004) Transcriptional regulation and functional properties of Arabidopsis Pht1;4, a high affinity transporter contributing greatly to phosphate uptake in phosphate deprived plants. Plant Mol. Biol. 55: 727-741

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Naumann C, Muller J, Sakonwasee S, Wiegand 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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Pandey BK, Mehra P, Giri J (2013) Phosphorus starvation response in plants and opportunities for crop improvement. Wiley-VCH Verlag GmbH & Co. KGaA Ch37: 991-1012

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Reymond 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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Rubio V, Linhares F, Solano R, Martín 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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Seren Ü (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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Seren Ü, Vilhjalmsdóttir 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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Shin H, Shin HS, Dewbre GR, Harrison MJ (2004) Phosphate transport in Arabidopsis: Pht1;1 and Pht1;4 play a major role in phosphate acquisition from both low- and high-phosphate environments. Plant J. 39: 629-642

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Shin JH, Blay S, McNeney 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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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 phosphate starvation. Plant Physiol. 170: 499-514

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Tan ST, Xue HW (2014) Casein Kinase 1 regulates ethylene synthesis by phosphorylating and promoting the turnover of ACS5. Cell Rep. 9: 1692-1702

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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 accumulation in Arabidopsis. Plant Physiol. 164: 144-159

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Wang H, Xu Q, Kong YH, Chen Y, Duan JY, Wu WH, Chen YF (2014) Arabidopsis WRKY45 transcription factor activates PHOSPHATE TRANSPORTER1;1 expression in response to phosphate starvation. Plant Physiol. 164: 2020-2029

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Wang L, Li Z, Qian W, Guo W, Gao X, Huang L, Wang H, Zhu H, Wu JW, Wang D, Liu D (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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Wang S, Yin L, Tanaka H, Tanaka K, Tsujimoto H (2010) Identification of wheat alien chromosome addition lines for breeding wheat with high phosphorus efficiency. Breed. Sci. 60: 371-379

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Wang X, Yan X, Liao H (2010) Genetic improvement for phosphorus efficiency in soybean: a radical approach. Ann. Bot. 106: 215-222

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Yang J, Ferreira T, Morris AP, Medland SE, Genetic Investigation of ATC, Replication DIG, Meta-analysis C, Madden PA, Heath AC, Martin NG, Montgomery GW, Weedon MN, Loos RJ, Frayling TM, McCarthy MI, Hirschhorn JN, Goddard ME, Visscher PM (2012) Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nat. Genet. 44: 369-375, S361-363

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Yang WT, Baek D, Yun DJ, Lee KS, Hong SY, Bae KD, Chung YS, Kwon YS, Kim DH, Jung KH, Kim DH (2018) Rice OsMYB5P improves plant phosphate acquisition by regulation of phosphate transporter. PLoS ONE 13: e0194628

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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 Online ahead of print <https://doi.org/10.1093/plphys/kiab441>

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)