Integrating genome-wide association studies with selective sweep reveals genetic loci associated with tolerance to low phosphate availability in Brassica napus

Article

Accepted Version

Original Article

Integrating genome wide association studies with selective sweep reveals genetic loci associated with tolerance to low phosphate availability in *Brassica napus*

Haijiang Liu[1,2], Yuan Pan[1,2], Rui Cui[1,2], John P. Hammond[3], Philip J. White[1,4], Yuting Zhang[1], Maoyan Zou[1,2], Guangda Ding [1,2], Sheliang Wang [1,2], Hongmei Cai[2], Fangsen Xu[1,2], Lei Shi[1,2]*

1 National Key Lab of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan 430070, China

2 Key Lab of Cultivated Land Conservation, Ministry of Agriculture and Rural Affairs/ Microelement Research Centre, Huazhong Agricultural University, Wuhan 430070, China

3 School of Agriculture, Policy and Development, University of Reading, Reading RG6 6AR, UK

4 The James Hutton Institute, Dundee, UK

Running title: Genetic bases of LP tolerance in *B. napus*

*For correspondence. E-mail: leish@mail.hzau.edu.cn

Abstract

Oilseed rape (*Brassica napus* L.; *B. napus*) is an important oil crop around the world. However, the genetic mechanisms of *B. napus* adaptations to low phosphate (P) stress are largely unknown. In this study, a genome-wide association study (GWAS) identified 68 SNPs significantly associated with seed yield (SY) under low P (LP) availability, and 7 SNPs significantly associated with phosphorus efficiency coefficient (PEC) in two trials. Among these SNPs, two, chrC07__39807169 and chrC09__14194798, were co-detected in two trials, and *BnaC07.ARF9* and *BnaC09.PHT1;2* were identified as candidate genes of them, respectively, by combine GWAS with quantitative reverse-transcription PCR (qRT-PCR). There were significant differences in the gene expression level of *BnaC07.ARF9* and *BnaC09.PHT1;2* between P -efficient and -inefficiency varieties at LP. SY_LP had a significant positive correlation with the gene expression level of both *BnaC07.ARF9* and *BnaC09.PHT1;2*. *BnaC07.ARF9* and *BnaA01.PHR1* could directly bind the promoters of *BnaA01.PHR1* and *BnaC09.PHT1;2*, respectively. Selective sweep analysis was conducted between ancient and derived *B. napus*, and detected 1280 putative selective signals. Within the selected region, a large number of genes related to P uptake, transport and utilization were detected, such as purple acid phosphatase (PAP) family genes and phosphate transporter (PHT) family genes. These findings provide novel insights into the molecular targets for breeding P efficiency varieties in *B. napus*.

Key words: *Brassica napus*, genome wide association study, selective sweep, phosphorus efficiency, *BnaC07.ARF9*, *BnaC09.PHT1;2*.

Abbreviations: GWAS, genome-wide association study; SNP, single-nucleotide polymorphisms; LP, low phosphorus; QTL, quantitative trait loci; PVE, phenotypic variation

INTRODUCTION

Oilseed rape (*Brassica napus* L.) is an important oil crop, which is widely cultivated around the world (Tang et al. 2021). *B. napus* is sensitive to phosphate (P) availability, resulting in reduced plant height and branch number, and subsequently a reduced yield (Liu et al. 2021a). Maintaining adequate supply of P fertilizer is important for maintaining *B. napus* yields. However, P rock used in the manufacture of P fertilizers, is a non-renewable resource and will eventually be exhausted in 300-400 years (Tiessen 2008; Cordell et al. 2009; Stutter et al. 2012). Excessive application of P fertilizers has also resulted in environmental problems such as eutrophication of aquatic ecosystems (Syers et al. 2008). In addition, the rising price of P fertilizer in recent years is placing economic pressures on farmers. Therefore, identifying genetic mechanisms for tolerance to low P availability in crops, and the alleles of genes associated with these tolerance mechanisms can be used to breed P-efficient crops that could contribute to addressing the "P crisis".

Genes associated with P uptake, transport and signal transduction have been identified previously through reverse genetics, which provides important gene targets for breeding P-efficient crops (Chiou and Lin. 2011; Lopez-Arredondo et al. 2014). For example, phosphate starvation response 1 (PHR1) and PHR1-like1 (PHL1) are central transcription factors involved in P signaling (Chiou and Lin. 2011). They regulate several P starvation-induced (PSI) genes by binding to the cis-element PHR1-binding sequence (P1BS) (Chiou and Lin. 2011). Transgenic plants overexpressing the *PHR1* gene in *Arabidopsis*, maize, rice, wheat and oilseed rape all show increased root hair growth, up-regulation of high-affinity P transporter genes and improved P uptake (Nilsson et al. 2007; Zhou et al. 2008; Wang et al. 2013a; Wang et al. 2013b; Lopez-Arredondo et al. 2014). In addition to *PHR1*,

65 the transcription factors AUXIN RESPONSE FACTOR7 (ARF7) and ARF19 have also been
66 reported to be involved in responses to low P availability in *Arabidopsis* (Huang et al. 2018).
67 Compared with the wild type, *arf7*, *arf19*, and *arf7 arf19* double mutants exhibit fewer lateral roots
68 and lower shoot P content (Huang et al. 2018). Some P uptake and utilization genes also play
69 important roles in the efficient utilization of P in plants (Lambers et al 2022). Overexpression of
70 *Arabidopsis AtPAP15* gene in soybean increased root P utilization efficiency, dry weight and seed
71 yield under low P availability (Wang et al. 2009). In rice, the transgenic plants overexpressing
72 *OsPAP10c*, a purple acid phosphatase, showed higher seed yield than WT in both hydroponic and
73 field experiments under different levels of P availability (Deng et al. 2020). In addition,
74 overexpression *OsPht1;1* in rice increased P content in shoots, and seed yield of transgenic plants
75 (Seo et al. 2008). In *B. napus*, overexpression *BnPht1;4* in *Arabidopsis* increased root P uptake,
76 primary root length and root biomass of transgenic plants under low P availability (Ren et al. 2014).
77 GWAS has become a common method for identifying candidate genes in *B. napus* (Liu et al.
78 2022a). For example, *BnaTT8*, a gene that controls seed coat content has been identified by GWAS
79 combined with transcriptome sequencing, and the seed coat content of *BnaTT8* knockout mutants
80 was significantly lower than that of the wild type (Zhang et al. 2022). Recently, three candidate
81 genes (*BnaA02g33340D*, *BnaA10g09290D* and *BnaC08g26640D*) are identified for plant height
82 and branch number under LP using a GWAS approach (Liu et al. 2021a). In addition, selective
83 sweep analysis has also been used to map and reveal the candidate genes controlling important
84 agronomic traits in *B. napus*, such as flowering time-related genes (Wu et al. 2019; Lu et al. 2019),
85 seed oil content (Tang et al. 2021) and seed glucosinolate content (Tan et al. 2021).
86 High seed yield is one of the desired targets in *B. napus* breeding. P deficiency decreased the

seed yield of *B. napus* significantly (Ding et al. 2012; Shi et al. 2013; Yuan et al. 2016). Additionally, the seed yield of crops under abiotic stresses have been widely used in the GWAS to explore candidate genes, such as, seed yield of maize under low phosphorus stress (Xu et al. 2018), seed yield of soybean under drought stress (Liyanage et al 2023), seed yield of rice under water-deficit stress (Kadam et al. 2018). In this study, we identified 68 reliable SNPs significantly associated with SY_LP, and *BnaC07.ARF9* and *BnaC09.PHT1;2* were identified to be candidate genes associated with P efficiency. In addition, we detected several selected regions between ancient and derived *B. napus* varieties by selective sweep analysis and 128 genes involved in P uptake, transport and utilization were identified.

MATERIALS AND METHODS

Plant materials and phenotypic investigation

In this study, 403 diverse rapeseed accessions were used to investigate the seed yield under LP availability. Among these *B. napus* varieties, 83 are ancient varieties and 184 are derived varieties from the ancient varieties (Tan et al. 2021; Tang et al. 2021; Supplementary Table S1). The panel was grown under field conditions with a low P supply (P, 0 kg ha^{-1}) with three replications at Meichuan Town, Wuxue city, Hubei province, China (115.55° E, 29.85° N) from 2018 to 2019 (Trial 1) and from 2019 to 2020 (Trial 2). The soil was sandy loam soil, and the available soil P concentration in three fields ranged from 11.56 to 16.95 mg kg^{-1}, which are between P deficient and slightly P deficient (Liu et al. 2021a). All the plots received basal fertilizer, and the application rate was as follows (per hectare): 108 kg of N (supplied as urea), 0 or 40 kg of P (supplied as calcium superphosphate), 87 kg of K (supplied as potassium chloride) and 6 kg of B (supplied as borax).

These fertilizers were thoroughly mixed and applied in bands near the crop rows. The remaining N (72 kg/ha) was top dressed as urea in equal amounts at the four to five-leaf stage. Each accession had four rows and each plot had eight plants in each row. Six plants in each plot were selected to investigate the seed yield at maturity stage. The mean values, maximum, minimum and coefficient of variation of SY_LP were calculated using R software. Additionally, the phosphorus efficiency coefficient (the ratio of seed yield under LP availability to that under normal condition (Liu et al. 2022b) of the association panel of *B. napus* were calculated.

Selective sweep analyses

The nucleotide diversity (π), fixation index (F_{ST}), and reduction of diversity (ROD) between ancient and derived varieties were analyzed using the VCFtools package (Danecek et al. 2011) with a step size of 10-kb and a 100-kb sliding window. The cross-population composite likelihood ratio test (XP-CLR) was analyzed using the XP-CLR software (Chen et al. 2010). The top 5% values of F_{ST}, ROD, and XP-CLR were applied to confirm highly differentiated regions.

GO enrichment analysis for candidate genes

To identify biological functions of candidate genes associated with the selection, gene ontology (GO) enrichment analysis of genes in the selected region was performed using a publicly available online website (http://www.pantherdb.org/). GO terms with P value < 0.05 were considered as those in which candidate genes were significantly enriched.

Genome-wide association study and candidate gene identification

In a previous study, more than 10 million SNPs across this association panel of *B. napus* were

identified (Tang et al. 2021). After filtering the SNPs with missing rate <0.2 and minor-allele frequency (MAF) >0.05, more than 1.60 million high-quality SNPs were obtained for GWAS of SY_LP. GWAS was carried out using factored spectrally transformed linear mixed models (Fast-LMM) (Lippert et al. 2011). The population are divided into five subgroups based on the cross-validation errors and the r pairwise relative kinship is close to 0 (Liu et al. 2021a). The Quantile-Quantile plot was drawn by the CMplot software (Yin et al. 2021) and the manhattan plot was drawn by ggplot2 software (https://cran.r-project.org/web/packages/ggplot2). The threshold for the significance of associations between SNPs and SY_LP is $P < 6.25\times10^{-07}$ (Liu et al. 2021a). The genes located within LD decay value (298 kb; Liu et al. 2021a) upstream and downstream of the significant SNPs were considered as candidate genes. The genotypes of *BnaC07.ARF9*, and *BnaC09.PHT1;2* in the P efficient varieties and P inefficient varieties were obtained by Vcftools software (Danecek et al. 2011).

RNA purification and gene expression analysis

B. napus varieties with high seed yield at LP were defined as P-efficient varieties and *B. napus* varieties with low yield at LP were defined as P-inefficient varieties. Sixteen P -efficient varieties and fourteen P -inefficient varieties were selected for gene expression analysis of putative candidate genes (*BnaA01.PHR1*, *BnaC07.ARF9*, and *BnaC09.PHT1;2*). Plants were grown with Hoagland's solution for 7 days, and then transplanted to zero P for a further 7 days. After 14 days of growth, leaves and roots were sampled and immediately snap-frozen in liquid nitrogen and stored at -80℃ for RNA extraction. Total RNA was extracted from tissues using a plant RNA purification Kit (Shanghai Promega Biological, China). cDNA was prepared using cDNA Synthesis Kit (Kangwei,

153 Beijing, China). Gene specific primers are listed in Supplementary Table S2. Quantitative
154 polymerase chain reaction (qPCR) reaction solution contained 0.2 μL reverse primer, 0.2 μL forward
155 primer, 1 μL template cDNA, 3.6 μL PCR-grade water and 5 μL Master Mix ABI Prism TM. The
156 PCR program was as follows: 95°C for 5 min, followed by 40 cycles of 95 °C for 15 s, 60 °C for 30
157 s and 72°C for 20 s. All reactions were performed in four biological repetitions and the average
158 expression value was calculated. $2^{-\triangle\triangle CT}$ method was used to evaluate relative expression levels.
159 *BnaTubulin* was used as an internal control for normalization.
160
161 **Yeast one-hybrid assays**
162 Matchmaker one-hybrid system (Clontech, Mountain View, CA) was used to perform the yeast one-
163 hybrid assays. The *BnaA01.PHR1* and *BnaC07.ARF9* coding region was amplified from *B. napus*
164 variety 'Y127' and cloned into the pGADT7-Rec2 prey vector by the ABclonal multiF seamless
165 assembly mix to create a translational fusion between the GAL4 activation domain and the
166 transcription factor. Additionally, the *BnaC09.PHT1.2* and *BnaA01.PHR1* promoter fragment
167 upstream of the transcription starting sites (spanned from −1 to −3000 bp) were amplified and cloned
168 into the pHIS2 reporter vector by *EcoRI* and *SacI*. The prey vector and reporter vector were co-
169 transformed into the yeast strain Y187. The pGADT7-53 and pHIS2-53 were used as a positive
170 control. The pGADT7-53 and pHIS2 empty vectors were used as a negative control. Cells were
171 grown in SD/-Trp-Leu liquid media to an OD 600 of 0.1 and then diluted with 10- and 100-fold
172 sterile water. For each dilution, 6 μl was plated on solid SD/-Trp-Leu-His supplemented with 0, 50
173 and 80 mM3-AT to test the strength of the interaction. 3-amino-1,2,4-triazole was a competitive
174 inhibitor of the yeast HIS3 protein and used to suppress background growth. The plates were then

incubated for 3 to 4 days at 30 °C.

Subcellular localization analysis

The coding sequences of *BnaC07.ARF9* and *BnaC09.PHT1;2* without stop codons were amplified by PCR and cloned into the binary vector pCAMBIA1300-YFP vectors. After confirmation by sequencing, the vectors were transferred into *A. tumefaciens* GV3101 and agroinfiltrated into 4-week-old *N. benthamiana* leaves. The cell plasma membrane marker construct AtPIP2a::RFP was used as a cell plasma membrane marker protein (Cutler et al. 2000). Nuclei of tobacco cells were stained with DAPI solution at 10 μg mL^{-1}(w/v) for 10 min, and then washed three times with ddH_2O. YFP and DAPI signals were examined using a confocal laser microscope (LSM 510 Meta, Carl Zeiss lnc.) with excitation wave lengths 488 nm for YFP, 514 nm for RFP, and 405 nm for DAPI.

Statistical analysis

The data were analyzed using Student's t-test, and significance was defined as $P<0.05$. r and p values of the correlation analysis were determined by Pearson correlation analysis.

RESULTS

Significant SNPs associated with the seed yield and PEC **of *B. napus* under LP supply**

The seed yield and PEC of the 403 diverse *B. napus* accessions showed an approximate normal distribution with extensive phenotypic variations under LP availability (Supplementary Fig. 1). For example, the seed yield at LP (SY_LP) ranged from 0.49 to 12.36 g in Trial 1 and from 1.16 to 9.83 g in Trial 2, and the PEC ranged from 0.11 to 0.87 in Trial 1 and from 0.10 to 0.80 in Trial 2

(Supplementary Fig. 1). A total of 68 SNPs and seven SNPs were identified significantly associated with SY_LP and PEC by GWAS (FAST-LMM), respectively. For the 68 SNPs significantly associated with SY_LP, 20 and 48 were identified in Trial 1 and Trial 2, respectively, explaining between 7.80 % to 10.55 % of the phenotypic variation (PVE) in Trial 1 and between 6.17% to 9.92% in Trial 2 (Supplementary Table 3). Chromosome C07 had the largest number of significant SNPs (14) and chromosome C09 had the second largest number of significant SNPs (9 SNPs) (Supplementary Table 3). In Trial 1, five PEC -associated significant SNPs, namely chrA03__7876388, chrA05__14295709, chrA07__15986664, chrC03__22658012 and chrC03__22657988 were detected by Fast-LMM model with the P values of $2.22\times^{-07}$, $4.54\times^{-07}$, $7.43\times^{-07}$, $4.30\times^{-07}$ and $1.96\times^{-07}$, respectively (Fig. 1; Supplementary Table 3). In Trial 2, chrA05__19953598 and chrC05__1385078 were significant associated with PEC ($P = 4.95\times^{-07}$, PVE = 7.04%; $P = 1.34\times^{-07}$, PVE = 11.82%) (Fig. 1; Supplementary Table 3).

Several reported genes related to P uptake, transport and utilization were identified for these SNPs (Supplementary Table 4). For example, the significant SNP chrA05__16581233 (P=4.92E-07, PVE=9.62%) on A05 was identified to be located within 280 kb of PAP17 ortholog (*BnaA05g22460D*; chrA05: 17114844-17116381) (Supplementary Table 4). In addition, *BnaC09.PHT1;2* (*BnaC09g17520D*; chrC09:14227709-14229865) was located within the interval of the significant SNP chrC09__14194798 (P=4.93E-8, PVE=9.92%) (Supplementary Table 4). *BnaA03g18240D* (chrA03:8545197-8547691), a gene with an SPX (SYG1/Pho81/XPR1) domain-containing protein, was located within the interval of the significant SNP chrA03__8376196 (P=1.22E-07, PVE=8.10%) for SY_LP (Supplementary Table 4).

BnaC07.ARF9 may be a key gene associated with P -efficiency in _B. napus_

The most significant SNPs on chromosome C07, chrC07__39807169 (P= 6.41 ×10^{-8}, PVE= 10.11%) and chrC07__39841117 (P= 2.90 × 10^{-8}, PVE=8.34%), were associated with SY_LP in Trial 1 and in Trial 2, respectively (Fig. 1; Supplementary Table 3). The distance between chrC07__39807169 and chrC07__39841117 is only 33 kb (Fig. 1; Supplementary Table 3). In this study, the LD decay was 298 kb for this association panel. Based on the LD decay, 300 kb up/downstream of the peak SNP (chrC07__39807169) were selected to identify candidate genes on C07 and 116 genes were detected (Supplementary Table 5). Among them, 19 genes were induced by P deficiency, 17 genes were inhibited by P deficiency, and 88 genes did not respond to P deficiency (Supplementary Table 5). The peak SNP chrC07__39807169 was identified to be located within 98 kb of *B. napus* *BnaC07.ARF9* (*BnaC07g38640D*; chrC07:39905968-39909666) (Fig. 2). A total of 52 SNPs were located within the 2 kb promoter region and the entire coding region of *BnaC07.ARF9* in the P efficient varieties and P inefficient varieties. Among them, four were in the promoter, 23 in the intron, and 25 in the exon (Supplementary Table 6). Previous studies have shown that *ARF7* and *ARF19* are important transcription factors regulating *Arabidopsis* tolerance to LP availability (Huang et al. 2018). Our previous transcriptome data showed that the relative expression of *BnaC07.ARF9* in the roots of cultivar 'Eyou changjia' were significantly increased under low P supply (Supplementary Fig. 2A, B; Du et al. 2017), which was confirmed by subsequent qRT-PCR experiments (Supplementary Fig 2). The overlap of the YFP and DAPI signals of 35S:BnaC07.ARF9:YFP indicated the BnaC07.ARF9 protein is a nucleus-localized auxin response factor (Fig. 2E). In addition, we analyzed the π value 1000 kb before and after *BnaC07.ARF9***,** and results showed that the intervals of the ancient and derived varieties had low and similar π values (Fig. 2B). In the two

field trials of P -efficient (16) and P -inefficient (14) *B. napus* varieties, seed yield of P -efficient varieties was significantly higher than that of P inefficient varieties at low P availability (Supplementary Fig. 3). Furthermore, the P -efficient varieties at LP exhibited significantly higher relative expression of *BnaC07.ARF9* than P -inefficient varieties (Fig. 2C-D). Additionally, the relative expression level of *BnaC07.ARF9* of these varieties was positively correlated with SY_LP (Supplementary Fig. 4A), suggesting that an increased mRNA abundance of *BnaC07.ARF9* at LP is associated with the low P tolerance of *B. napus*.

A putative signaling cascade *BnaC07.ARF9/ BnaA01.PHR1/ BnaC09.PHT1;2* for P efficiency of *B. napus*

The significant SNPs, 'chrC09__14194940 (Trial 1, P = 1.56×10^{-7}, PVE=9.83%), and 'chrC09__14194798' (Trial 2, P = 4.93×10^{-8}, PVE=9.92%), on chromosome C09, were associated with SY_LP (Fig. 1; Supplementary Table 3). Based on the LD decay, 300 kb up/downstream of the peak SNP (chrC09__14194798) were selected to identify candidate genes on C09 and 83 genes were detected (Supplementary Table 7). Among them, 10 genes were induced by P deficiency, 6 genes were inhibited by P deficiency, and 72 genes did not respond to P deficiency (Supplementary Table 7). *BnaC09g17520D* (chrC09:14227709-14229865), the ortholog of *PHT1;2* in *B. napus* was identified by lead SNP chrC09__14194798, which was located 32 kb downstream of the lead SNP chrC09__14194798 (Fig. 1; Supplementary Table 3). There were 53 SNPs located within the 2 kb promoter region and the entire coding region of *BnaC09.PHT1;2* in the P efficient varieties and P inefficient varieties, of which 38 were located in the promoter, two were located in the intron, and 13 were located in the exon (Supplementary Table 6). The *PHT1* family encode plant P transporters which transport P across plasma membranes, including at the soil-root interface (Lopez-Arredondo

et al. 2014). Transcriptome analysis of cultivar 'Eyou changjia' showed that the gene expression of *BnaC09.PHT1;2* increased significantly in roots under LP supply, which was confirmed by subsequent qRT-PCR experiments in ZS11 (Supplementary Fig. 2C, D). The merging images obtained from the YFP and RFP channels showed that the BnaC09.PHT1;2-YFP and AtPIP2A-RFP fluorescence proteins co-localized to the plasma membrane (Fig. 3E), which indicated that BnaC09.PHT1;2 is localized at the cell membrane (Supplementary Fig. 2C, D; Fig. 3E). The relative expression level of the *BnaC09.PHT1;2* in root among 30 inbred lines (16 P efficient varieties and 14 P inefficient varieties) under LP supply were significantly positively correlated with the SY_LP (Supplementary Fig. 4B). In addition, the relative expression level of *BnaC09.PHT1;2* in roots of P -efficient varieties was significantly higher than that in P -inefficient varieties under LP supply (Fig. 3C, D). The π value 1000 kb before and after *BnaC09.PHT1;2* was analyzed and showed that the intervals of the ancient and derived varieties had low π values (Fig. 3B).

In *Arabidopsis*, *AtARF7* and *AtARF19* have been shown to regulate response to low P availability via the MYB-CC transcription factor *PHR1* (Huang et al 2018). There were significant differences in the relative expression level of *BnaA01.PHR1* in roots between P -efficient and -inefficient varieties (Fig. 4A, B). Moreover, the relative expression level of *BnaA01.PHR1* was significantly positively correlated with the SY_LP in these varieties (Supplementary Fig. 4C). In addition, there was significant positive correlations in the relative expression level between *BnaC07.ARF9* and *BnaA01.PHR1* under LP supply in P -efficient and -inefficient varieties (Fig. 4C). Two AuxRE elements (AuxRE-1, TGTCTC −2666 to −2672; and AuxRE-2, TGTCTC −2299 to −2305) and one TGA elements (TGA, AACGAC, −449 to −455) were identified in the 3000 bp fragment upstream of the *BnaA01.PHR1* start codon isolated from *B. napus* variety 'Y127'. This

fragment was cloned into pHIS2 vector for Y1H assay. The transformants of the pHIS2-*BnaA01.PHR1* plus pGADT7- *BnaC07.ARF9* and the positive control grew well, but the negative control was completely inhibited on selective media with both 50 and 80 mM 3-AT (Fig. 4E). These data suggested that the *BnaC07.ARF9* could directly bind to the promoter regions of *BnaA01.PHR1* and the latter might be the target gene downstream of the former.

There was a significant positive correlation in the relative expression level between *BnaA01.PHR1* and *BnaC09.PHT1;2* in P -efficient and -inefficient varieties under LP supply (Fig. 4D). Three PHR1-binding sequences (P1BS) in the *BnaC09.PHT1;2* promoter (P1BS-1 GTATATGC, -627 to 635; P1BS-2 GTATATCC, -464 to 472; P1BS-2 GTATATCC, -426 to 434) were identified using ensembl (http://plants.ensembl.org/Brassica_napus/). Yeast one hybrid assays were performed to explore the relationship between *BnaA01.PHR1* and *BnaC09.PHT1;2* (Fig. 4F). All transformed yeasts grew well on selective medium without 3-AT. When the 3-AT concentration was increased up to 50 mM or 80 mM, the transformants of the pHIS2- *BnaC09.PHT1;2* plus pGADT7- *BnaA01.PHR1* and the positive control grew normally, but almost all the negative controls did not grow (Fig. 4F). These findings suggest that *BnaA01.PHR1* could interact with the promoters of *BnaC09.PHT1;2* and that *BnaC09.PHT1;2* might be the target gene downstream of *BnaA01.PHR1*.

Combined use of F_{ST}, ROD and XP-CLR to detect P uptake and utilization related selective signals

Root system architecture, plant height, branch number, and seed yield of ancient and derived varieties under LP have been investigated previously (Wang et al. 2017; Liu et al. 2021a). Compared with derived *B. napus*, ancient *B. napus* had more developed root system architecture at the seedling

307 stage (Fig. 5A-G), higher branch numbers, higher plant height and greater seed yield at maturity
308 under LP (Fig. 5H-J). These indicated that tolerance to P deficiency in *B. napus* has been
309 significantly reduced through *B. napus* breeding.

310 Selective sweeps between ancient and derived *B. napus* were analyzed to find the genes related
311 to P uptake, transport and utilization in *B. napus* (Fig. 6). All genomic regions in 100-kb sliding
312 windows (a step of 10 kb) were scanned, and the regions with the top 5% of population fixation
313 statistics (F_{ST}) values were defined as significantly different windows. A total of 195 selective
314 sweeps (covering 52.32 Mb) were identified between ancient and derived *B. napus* varieties (Fig.
315 6A; Supplementary Table S8). There were 43 P uptake, transport and utilization, and root
316 development related genes in these selective-sweep regions. (Fig. 6A; Table 1). These genes (with
317 the top 5% of F_{ST} values) were mainly enriched in several significant P uptake related pathways,
318 such as, acid phosphatase activity (GO:0003993), root cap development (GO:0048829) and primary
319 root development (GO:0080022) (Supplementary Fig. 5A and Supplementary Table 9).

320 In addition to F_{ST}, 285 selective sweep signals (with the top 5% of ROD value) were identified
321 by comparison between ancient and derived varieties, which covered 61.32 Mb of the Darmor-*bzh*
322 reference genome (Fig. 6B; Supplementary Table 10). A total of 57 known P uptake and homeostasis
323 related genes were found in these selective-sweep regions, (Fig. 6b; Table 1). GO enrichment
324 analysis showed that these genes (with the top 5% of ROD value) were significantly enriched in the
325 pathways of acid phosphatase activity (GO:0003993), nutrient reservoir activity (GO:0045735),
326 transporter activity (GO:0005215) and sucrose transport (GO:0015770) (Supplementary Fig. 5B
327 and Supplementary Table 11).

328 A total of 804 selective sweeps signals (with the top 5% of XP-CLR value) were identified

between the ancient and derived varieties by XP-CLR analysis, which covered 117.38 Mb of the Darmor-*bzh* reference genome (Fig. 6C; Supplementary Table 12). The strong signal of a selective sweep was found on A08 chromosome between the ancient and derived varieties (Fig. 6C). There were two known P uptake related genes in this region, including an *ARF19* ortholog (*BnaA08g22150D*) and a *PHT1;8* ortholog (*BnaA08g21590D*) (Fig. 6C; Table 1). Another strong selective signal was found on A04 and A06 chromosome, respectively, including six orthologs of PHT family genes, *PHT1;3* (*BnaA06g36740D* and *BnaA06g36750D*), PHT1;4 (*BnaA04g21000D* and *BnaA04g21010D*), *PHT1;7* (*BnaA06g36760D*), and PHT4;2 (*BnaA04g21790D*) (Fig. 6C; Table 1). In addition, some regions on C02, C03, C07 and C09 chromosome that encoded several SPX (SYG1/Pho81/XPR1) domain-containing protein genes displayed strong selective-sweep signals (Fig. 6C; Table 1). GO enrichment analysis showed that these genes were significantly enriched three pathways, namely, lateral root formation (GO:0010311), phosphate ion homeostasis (GO:0055062), and cellular response to auxin stimulus (GO:0071365) (Supplementary Fig. 5C and Supplementary Table 13).

DISCUSSION

Genetic architecture of LP tolerance traits

Linkage mapping analysis has been widely used to dissect the genetic bases of P tolerance in *B. napus* (Ding et al. 2012; Shi et al. 2013; Zhang et al. 2016). Thirty-two significant quantitative trait loci (QTLs) are associated with plant height, branch number, seed weight, P efficiency coefficient, and seed yield under LP supply, and explain between 7.8 to 21% of the phenotypic variation (Ding et al. 2012). Moreover, 131 QTLs were detected related to primary root length, total root length, and

lateral root density under low and sufficient P conditions by the Tapidor/Ningyou7 population in *B. napus* (Zhang et al. 2016). Compared with linkage analysis, association analysis has the advantages of no need to construct mapping groups and high analysis accuracy (Liu et al. 2022a). In this study, through GWAS with high-density SNP markers, 68 SNPs were significantly associated with SY_LP, which could explain 6.17%-10.55% of the phenotypic variation (Supplementary Table 3). In a previous study, 1773 SNPs were detected associated with seed yield of *B. napus* at s sufficient P supply and three candidate genes (*BnaA01g17200D*, *BnaA02g08680D*, and *BnaA09g10430D*) were identified (Liu et al. 2022b). Eleven significant SNPs in this study were adjacent to previously published significant SNPs associated with seed yield at a sufficient P supply, which may control seed yield at both P levels (Supplementary Table 14).

Compared with other crops, the LD value of *B. napus* is relatively large, which leads to the linkage disequilibrium region of significant SNPs containing hundreds of genes, making it extremely difficult to identify the real candidate genes (Zhang et al. 2022). Therefore, GWAS combined with transcriptome sequencing to mine new candidate genes has become a common method (Liu et al. 2022a). In this study, two reliable SNP cluster (chrC07: 17114844-17116381; chrC09: 17114844-17116381) were identified on C07 and C09 chromosome, and two genes (*BnaC07.ARF9* and *BnaC09.PHT1;2*) significantly induced by P deficiency were identified by GWAS combined with previous transcriptome data, respectively (Fig. 2-3). The expression levels of *BnaC07.ARF9* and *BnaC09.PHT1;2* at LP supply were significantly positively correlated with SY_LP (Supplementary Fig 2). In addition, the expression levels of *BnaC07.ARF9* and *BnaC09.PHT1;2* in P -efficient varieties was significantly higher than that in P -inefficient varieties (Fig. 2C-D; Fig. 3C-C). These results suggest that *BnaC07.ARF9* and *BnaC09.PHT1;2* may be P -efficient candidate genes.

373 *PHR1* is a component of the auxin regulatory pathway involved in low-Pi responses in different
374 crops (Lopez-Arredondo et al. 2014). The *Arabidopsis* transgenic plants overexpressed of *B. napus*
375 gene *BnaA01.PHR1* have more lateral roots, more root fresh weight, and higher P concentration
376 than wild type (Ren et al. 2012). The results of rice and *Arabidopsis* showed that *PHR1* could
377 achieve P -efficiency by regulating PHT1 family genes (Lopez-Arredondo et al. 2014). In this study,
378 we demonstrated that *BnaC07.ARF9* could directly bind the promoter of *BnaA01.PHR1*, and
379 *BnaA01.PHR1* directly bind the promoter of *BnaC09.PHT1;2* by yeast one-hybrid assays and qRT-
380 PCR analysis (Fig .4).

381 **P uptake and utilization related selective signals**

382 *B. napus* (AACC) originated 7,500 years ago by crossing *Brassica rapa* (AA) (n = 10) and *Brassica*
383 *oleracea* (CC) (n =9) (Wu et al. 2019). Recently, the breeding history of *B. napus* flowering time
384 (Wu et al. 2019; Lu et al. 2019), seed glucosinolate content (Tan et al. 2021) and seed oil content
385 (Tang et al. 2021) have been analyzed, and a series of candidate genes related to them are mined. P
386 is an essential macronutrients for plant growth and development (Hawkesford et al et al. 2012). In
387 this study, the ancient varieties had more developed roots, root exudates and higher yield than the
388 derived varieties under LP supply and showed stronger resistance to LP stress (Fig. 5). This indicated
389 that during the improvement process of *B. napus* from ancient varieties to derived varieties, the
390 selection leads to the loss of diversity in P -efficiency genes. The possible reason is that almost all
391 the derived varieties were selected under sufficient fertilizer supply, and the agronomic traits, such
392 as seed yield, seed quality, flowering time and disease traits of *B. napus* were focused, while the
393 low P stress tolerance was ignored. The ancient varieties have stronger resistance to stress than
394 derived varieties, which are widely existed in the process of cultivar improvement in other crops

395 (Liu et al. 2021b; Zhang et al. 2021). For example, in the process of soybean improvement, although
396 the seed weight and oil content of derived varieties were significantly improved, but compared with
397 the ancient varieties, their salt and drought stress tolerance were significantly decreased (Zhang et
398 al. 2021). Additionally, ancient rice varieties have a stronger low nitrogen stress tolerance than
399 derived rice varieties, and one nitrogen efficient gene "*OsTCP19*" is lost during rice improvement
400 (Liu et al. 2021b).

401 In this study, the selective sweep analysis between ancient and derived varieties were performed
402 by F_{ST}, ROD, and XP-CLR method, and excavated a large number of PAP, PHT family genes, and
403 root development related genes (Table 1). GO analysis showed that the most significant GO terms
404 enriched with genes in the selective sweep regions included “acid phosphatase activity
405 (GO:0003993)”, “primary root development (GO:0080022)”, “root cap development
406 (GO:0048829)”, “transporter activity (GO:0005215)”, “nutrient reservoir activity (GO:0045735)”,
407 “sucrose transport (GO:0015770)”, “lateral root formation (GO:0010311)”, “phosphate ion
408 homeostasis (GO:0055062)”, and “cellular response to auxin stimulus (GO:0071365)”
409 (Supplementary Fig 5). This confirmed that *B. napus* was likely lost its adaptability to LP tolerance
410 genes during the domestication process from ancient and derived varieties (Supplementary Fig 5).
411 The π value of 1000 kb before and after *BnaC07.ARF9* and *BnaC09.PHT1;2* were analyzed, and
412 results showed that the intervals of the ancient and derived varieties had similar and low π values,
413 so we did not detect these two genes through selective sweep analysis (Fig. 3B). This suggests that
414 *BnaC07.ARF9* and *BnaC09.PHT1;2* is highly conserved in *B. napus* and have indispensable
415 functions in response to P deficient tolerance (Fig. 3B).

416 In addition, a total of twelve significant SNPs loci were located in the selected interval by

417 comparing GWAS with selective sweep results. For example, on chromosome A05, three SNPs
418 (chrA05__16831572, chrA05__16881222, and chrA05__16881233) significantly associated with
419 SY_LP were located in the selected interval (Position: 16730000- 16890000) (Supplementary Table
420 15). Root system architecture, plant height, and branch number of the *Brassica napus* panel under
421 low P supply have been conducted genome-wide association study, and 188 SNPs are located in the
422 selected interval (Wang et al. 2017; Liu et al. 2021; Duan et al. 2021; Supplementary Table 15). For
423 example, SNP Bn-A01-p21527187, which was significantly associated with primary root length,
424 total root length and lateral root length, co-localized with the selected interval on chromosome A01
425 identified in this study (Position: 17900000- 18320000) (Supplementary Table 15). These SNPs and
426 nearby genes may cause difference in the tolerance ability to LP stress of ancient *B. napus* and
427 derived *B. napus.* Unfortunately, no co-located SNP loci were found by comparing the GWAS
428 results of PEC and SY_LP. In total, ancient *B. napus* that exhibited more resistant to LP stress may
429 therefore serve as a resource for germplasm improvement in the development of P -efficient varieties
430 in further.

431 ## CONCLUSIONS

432 Taken together, our study analyzes the genetic differences related to LP stress tolerance
433 between ancient and derived varieties of *B. napus* for the first time. In addition, our study identified
434 two new potential *B. napus* P -efficient candidate genes, which could be valuable gene resources for
435 the subsequent cultivation of P -efficient *B. napus* varieties.

436

437 ## FUNDING

438 This work was supported by the National Nature Science Foundation of China (Grant Nos.

31972498 and 32172662). The computations in this paper were run on the bioinformatics computing platform of the National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University.

Author contribution statement

Haijiang Liu and Lei Shi designed research, reviewed writing and drafted the manuscript. Haijiang Liu, Pan Yuan, Rui Cui, Maoyan Zou, and Yuting Zhang: participated in the experiments. Lei Shi, Pan Yuan, Rui Cui, Guangda Ding, Sheliang Wang, Hongmei Cai, John Hammond and Fangsen Xu: Participated in manuscript revision and data interpretation.

Data Availability

All the raw sequencing data generated during this study are available in the Genome Sequence Archive (https://bigd.big.ac.cn/gsa/) with Bioproject IDs PRJCA002835 and PRJCA002836 (Tang et al. 2021).

Declaration

Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Human and animal rights

This study does not include human or animal subjects.

461
462

463 **LITERATURE CITED**

464 Chen H, Patterson N, Reich D (2010) Population differentiation as a test for selective sweeps.
465 Genome Res 20:393-402

466 Chiou TJ, Lin SI (2011) Signaling network in sensing phosphate availability in plants. Annu Rev
467 Plant Biol 62:185-206

468 Cordell D, Drangert JO, White S (2009) The story of phosphorus: global food security and food for
469 thought. Glob Environ Chang 19:292–305

470 Cutler SR, Ehrhardt DW, Griffitts JS, Somerville CR (2000) Random GFP::cDNA fusions enable
471 visualization of subcellular structures in cells of *Arabidopsis* at a high frequency. Proc Natl Acad
472 Sci USA 97: 3718-3723

473 Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G,
474 Marth GT, Sherry ST, McVean G, Durbin R, Genomes Project Analysis G (2011) The variant call
475 format and VCFtools. Bioinformatics 27:2156-2158

476 Deng S, Lu L, Li J, Du Z, Liu T, Li W, Xu F, Shi L, Shou H, Wang C (2020) Purple acid phosphatase
477 10c encodes a major acid phosphatase that regulates plant growth under phosphate-deficient
478 conditions in rice. J Exp Bot 71:4321-4332

479 Ding G, Zhao Z, Liao Y, Hu Y, Shi L, Long Y, Xu F (2012) Quantitative trait loci for seed yield and
480 yield-related traits, and their responses to reduced phosphorus supply in *Brassica napus*. Ann Bot
481 109:747-759

482 Du HY, Yang C, Ding GD, Shi L, Xu FS (2017) Genome-wide identification and characterization
483 of SPX domain-containing members and their responses to phosphate deficiency in *Brassica napus*.

Front Plant Sci. 8:35

Hawkesford M, Horst W, Kichey T, Lambers H, Schjoerring J, Skrumsager Møller I, White P (2012) Chapter 6: Functions of macronutrients. In: Marschner P (ed) Marschner's mineral nutrition of higher plants, 3rd edn. Academic Press, London, pp 135–189

Hu J, Chen B, Zhao J, Zhang F, Xie T, Xu K, Gao G, Yan G, Li H, Li L, Ji G, An H, Li H, Huang Q, Zhang M, Wu J, Song W, Zhang X, Luo Y, Chris Pires J, Batley J, Tian S, Wu X (2022) Genomic selection and genetic architecture of agronomic traits during modern rapeseed breeding. Nat Genet 54:694-704

Huang KL, Ma GJ, Zhang ML, Xiong H, Wu H, Zhao CZ, Liu CS, Jia HX, Chen L, Kjorven JO, Li XB, Ren F (2018) The ARF7 and ARF19 transcription factors positively regulate PHOSPHATE STARVATION RESPONSE1 in *Arabidopsis* roots. Plant Physiol 178:413-427

Kadam N N, Struik P C, Rebolledo M C (2018) Genome-wide association reveals novel genomic loci controlling rice grain yield and its component traits under water-deficit stress during the reproductive stage. J Exp Bot 69: 4017-4032.

Kochian LV (2012) Plant nutrition: rooting for more phosphorus. Nature 488: 466–467

Lambers H (2022) Phosphorus acquisition and utilization in plants. Annu Rev Plant Biol 73: 17–42

Li Y, Wang X, Zhang H, Wang S, Ye X, Shi L, Xu F, Ding G (2019) Molecular identification of the phosphate transporter family 1 (*PHT1*) genes and their expression profiles in response to phosphorus deprivation and other abiotic stresses in *Brassica napus*. PLoS One 14:e0220374

Lippert C, Listgarten J, Liu Y, Kadie CM, Davidson RI, Heckerman D (2011) FaST linear mixed models for genome-wide association studies. Nat Methods 8:833-835

Liu H, Wang J, Zhang B, Yang X, Hammond JP, Ding G, Wang S, Cai H, Wang C, Xu F, Shi L

(2021a) Genome wide association study dissects the genetic control of plant height and branch number in response to lowphosphorus stress in *Brassica napus*. Ann Bot 128: 919-930

Liu H, Wang W, Yang M, Yuan P, Hammond JP, King GJ, Ding G, White PJ, Wang S, Cai H, Wang C, Lu C, Xu F, Shi L (2022a) Genome‐wide association studies of important agronomic traits in *Brassica napus* : what we have learned and where we are headed. Annual Plant Reviews online, pp 151-180

Liu H, Zou M, Zhang B, Yang X, Yuan P, Ding G, Xu F, Shi L (2022b) Genome-wide association study identifes candidate genes and favorable haplotypes for seed yield in *Brassica napus*. Mol Breeding 42: 61

Liu Y, Wang H, Jiang Z, Wang W, Xu R, Wang Q, Zhang Z, Li A, Liang Y, Ou S, Liu X, Cao S, Tong H, Wang Y, Zhou F, Liao H, Hu B, Chu C (2021b) Genomic basis of geographical adaptation to soil nitrogen in rice. Nature 590:600-605

Liyanage D K, Torkamaneh D, Belzile F (2023) The Genotypic variability among short-season soybean cultivars for nitrogen fixation under drought Stress. Plants 12: 1004.

Lopez-Arredondo DL, Leyva-Gonzalez MA, Gonzalez-Morales SI, Lopez-Bucio J, Herrera-Estrella L (2014) Phosphate nutrition: improving low-phosphate tolerance in crops. Annu Rev Plant Biol 65:95-123

Lu K, Wei L, Li X, Wang Y, Wu J, Liu M, Zhang C, Chen Z, Xiao Z, Jian H, Cheng F, Zhang K, Du H, Cheng X, Qu C, Qian W, Liu L, Wang R, Zou Q, Ying J, Xu X, Mei J, Liang Y, Chai YR, Tang Z, Wan H, Ni Y, He Y, Lin N, Fan Y, Sun W, Li NN, Zhou G, Zheng H, Wang X, Paterson AH, Li J (2019) Whole-genome resequencing reveals *Brassica napus* origin and genetic loci involved in its improvement. Nat Commun 10:1154

528 Nilsson L, Muller R, Nielsen TH (2007) Increased expression of the MYB-related transcription
529 factor, *PHR1*, leads to enhanced phosphate uptake in *Arabidopsis* thaliana. Plant Cell Environ
530 30:1499-1512
531 Ren F, Guo QQ, Chang LL, Chen L, Zhao CZ, Zhong H, Li XB (2012) *Brassica napus PHR1* gene
532 encoding a MYB-like protein functions in response to phosphate starvation. PLoS One 7:e44005
533 Ren F, Zhao CZ, Liu CS, Huang KL, Guo QQ, Chang LL, Xiong H, Li XB (2014) A *Brassica napus*
534 *PHT1* phosphate transporter, *BnPht1;4*, promotes phosphate uptake and affects roots architecture of
535 transgenic Arabidopsis. Plant Mol Biol 86:595-607
536 Seo HM, Jung Y, Song S, Kim Y, Kwon T (2008) Increased expression of *OsPT1*, a high-affinity
537 phosphate transporter, enhances phosphate acquisition in rice. Biotechnol Lett. 30:1833–38
538 Shi T, Li R, Zhao Z, Ding G, Long Y, Meng J, Xu F, Shi L (2013) QTL for yield traits and their
539 association with functional genes in response to phosphorus deficiency in *Brassica napus*. PLoS
540 One 8:e54559
541 Stutter MI, Shand CA, George TS, Blackwell MSA, Bol R, MacKay RL, Richardson AE, Condron
542 LM, Turner BL, Haygarth PM (2012) Recovering phosphorus from soil: a root solution? Environ
543 Sci Technol 46:1977–1978
544 Syers JK, Johnston AE, Curtin D (2008) Efficiency of soil and fertilizer phosphorus use:
545 Reconciling changing concepts of soil phosphorus behavior with agronomic information. FAO and
546 Fertilizer and Plant Nutrition Bulletin 18. Rome: Food and Agricultural Organization
547 Tan Z, Xie Z, Dai L, Zhang Y, Hu Z, Tang S, Wan L, Yao X, Guo L, Hong D (2021) Genome- and
548 transcriptome-wide association studies reveal the genetic basis and the breeding history of seed
549 glucosinolate content in *Brassica napus*. Plant Biotechnol J 20: 211-225

550 Tang S, Zhao H, Lu S, Yu L, Zhang G, Zhang Y, Yang QY, Zhou Y, Wang X, Ma W, Xie W, Guo L
551 (2021) Genome- and transcriptome-wide association studies provide insights into the genetic basis
552 of natural variation of seed oil content in *Brassica napus*. Mol Plant 14:470-484
553 TiessenH (2008) Phosphorus in the global environment. In: White PJ, Hammond JP (eds) The
554 ecophysiology of plant-phosphorus interactions. Springer, Dordrecht, pp 1–7
555 Vance C, Uhde-Stone C, Allan D (2003) Phosphorus acquisition and use: critical adaptations by
556 plants for securing a nonrenewable resource. New Phytologist 157: 423–447
557 Wang J, Sun J, Miao J, Guo J, Shi Z, He M, Chen Y, Zhao X, Li B, Han F, Tong Y, Li Z (2013a) A
558 phosphate starvation response regulator *Ta-PHR1* is involved in phosphate signalling and increases
559 grain yield in wheat. Ann Bot 111:1139-1153
560 Wang W, Ding G-D, White PJ, Wang X-H, Jin K-M, Xu F-S, Shi L (2018) Mapping and cloning of
561 quantitative trait loci for phosphorus efficiency in crops: opportunities and challenges. Plant and
562 Soil 439:91-112
563 Wang X, Bai J, Lui H, Sun Y, Shi X, Ren Z (2013b) Overexpression of a maize transcription factor
564 ZmPHR1 improves shoot inorganic phosphate content and growth of *Arabidopsis* under low-phosphate
565 conditions. Plant Mol. Biol. Rep 31:665–77
566 Wang X, Chen Y, Thomas CL, Ding G, Xu P, Shi D, Grandke F, Jin K, Cai H, Xu F, Yi B, Broadley
567 MR, Shi L (2017) Genetic variants associated with the root system architecture of oilseed rape
568 (*Brassica napus* L.) under contrasting phosphate supply. DNA Res 24:407-417
569 Wang X, Wang Y, Tian J, Lim BL, Yan X, Liao H (2009) Overexpressing AtPAP15 enhances
570 phosphorus efficiency in soybean. Plant Physiol 151:233–40
571 Wu D, Liang Z, Yan T, Xu Y, Xuan L, Tang J, Zhou G, Lohwasser U, Hua S, Wang H, Chen X,

572 Wang Q, Zhu L, Maodzeka A, Hussain N, Li Z, Li X, Shamsi IH, Jilani G, Wu L, Zheng H, Zhang
573 G, Chalhoub B, Shen L, Yu H, Jiang L (2019) Whole-genome resequencing of a worldwide
574 collection of rapeseed accessions reveals the genetic basis of ecotype divergence. Mol Plant 12:30-
575 43
576 Xu C, Zhang H, Sun J (2018) Genome-wide association study dissects yield components associated
577 with low-phosphorus stress tolerance in maize. Theor Appl Genet 131:1-16.
578 Yin L, Zhang H, Tang Z, Xu J, Yin D, Zhang Z, Yuan X, Zhu M, Zhao S, Li X, Liu X (2021) rMVP:
579 A memory-efficient, visualization-enhanced, and parallel-accelerated tool for genome-wide
580 association study. Genomics Proteomics Bioinformatics 19:619-628
581 Yuan P, Ding G, Cai H, Jin K, Broadley MR, Xu F, Shi L (2016) A novel Brassica-rhizotron system
582 to unravel the dynamic changes in root system architecture of oilseed rape under phosphorus
583 deficiency. Annals of Botany 118:173-184
584 Zhang W, Xu W, Zhang H, Liu X, Cui X, Li S, Song L, Zhu Y, Chen X, Chen H (2021) Comparative
585 selective signature analysis and high-resolution GWAS reveal a new candidate gene controlling seed
586 weight in soybean. Theor Appl Genet 134:1329-1341
587 Zhang Y, Thomas CL, Xiang J, Long Y, Wang X, Zou J, Luo Z, Ding G, Cai H, Graham NS,
588 Hammond JP, King GJ, White PJ, Xu F, Broadley MR, Shi L, Meng J (2016) QTL meta-analysis of
589 root traits in *Brassica napus* under contrasting phosphorus supply in two growth systems. Sci Rep
590 6:33113
591 Zhang Y, Zhang H, Zhao H, Xia Y, Zheng X, Fan R, Tan Z, Duan C, Fu Y, Li L, Ye J, Tang S, Hu
592 H, Xie W, Yao X, Guo L (2022) Multi-omics analysis dissects the genetic architecture of seed coat
593 content in *Brassica napus*. Genome Biol 23:86

Zhou J, Jiao F, Wu Z, Li Y, Wang X, He X, Zhong W, Wu P (2008) *OsPHR2* is involved in phosphate-starvation signaling and excessive phosphate accumulation in shoots of plants. Plant Physiol 146:1673-1686

Figure legends

Fig. 1 Genome wide association study for SY and PEC at LP in an association panel of *B. napus* in Trial 1 and Trial 2 by Fast-LMM model. (a) SY in Trial 1, (b) SY in Trial 2, (c) PEC in Trial 1, (d) PEC in Trial 2. SY, seed yield; LP, low phosphorus; PEC, phosphorus efficiency coefficient.

Fig. 2 Nucleotide diversity (π), gene expression, and subcellular localization analysis of *BnaC07.ARF9* gene. (A) Manhattan plot for SY_LP in Trial 1; (B) π analysis of 1000 kb up and downstream of candidate gene *BnaC07.ARF9*; (C-D) The relative expression level of *BnaC07.ARF9* in root of P -efficient and -inefficient varieties; (E) Subcellular localization of *BnaC07.ARF9*. P, phosphorus.

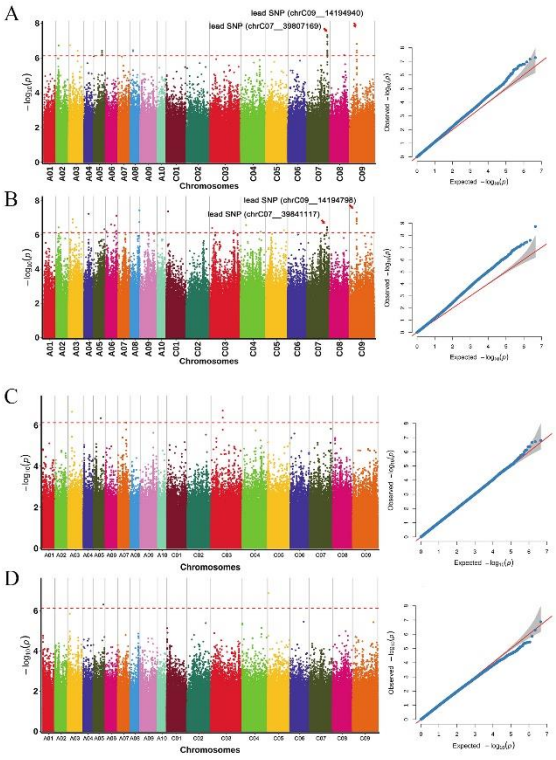
Fig. 3 Nucleotide diversity (π), gene expression, and subcellular localization analysis of *BnaC09.PHT1;2* gene. (A) Manhattan plot for SY_LP in Trial 1; (B) π analysis of 1000 kb up and downstream of candidate gene *BnaC09.PHT1;2*; (C-D) The relative expression level of *BnaC09.PHT1;2* in root of P-efficient and -inefficient varieties; (E) Subcellular localization of *BnaC09.PHT1;2*. P, phosphorus.

Fig. 4 BnaC07.ARF9 regulates BnaC07.PHT1;2 by regulating BnaA01.PHR1 to achieve high phosphorus efficiency. (A-B) The relative expression level of BnaA01.PHR1 in root of P-efficient and -inefficient varieties; (C) The correlation between the relative expression level of BnaC07.ARF9 at LP and the relative expression level of BnaA01.PHR1 at LP in P-efficient and -inefficient varieties; (D) The correlation between the relative expression level of BnaA01.PHR1 at LP and the relative expression level of BnaC09.PHT1;2 at LP in P-efficient and -inefficient varieties; (E) Yeast one-hybrid assays of binding activity of BnaC07.ARF9 with BnaA01.PHR1;

(F) Yeast one-hybrid assays of binding activity of BnaA01.PHR1 with BnaC09.PHT1;2. P, phosphorus; LP, low phosphorus.

Fig. 5 Differences in the root system architecture (A-F), SY and SY related traits (G-I) between ancient and derived *B. napus* varieties at LP. LRL, lateral root length; LRN, lateral root number; TRL, total root length; SDW, shoot dry weight; RDW, root dry weight; PRL, primary root length; BN, branch number; PH, plant height; SY, seed yield; LP, low phosphorus.

Fig. 6 Genome-wide distribution of selective-sweep signals identified through comparisons between ancient and derived varieties by F_{ST} (A), ROD (B), and XP-CLR (C) method. Genes related to phosphorus uptake, transport and utilization in selection outlier regions between ancient and derived varieties were labeled on the top of the corresponding chromosomes. The red dashed lines represent the thresholds (top 5% of F_{ST}, ROD, and XP-CLR values).



631

632 Fig. 1 Genome wide association study for SY and PEC at LP in an association panel of *B. napus* in

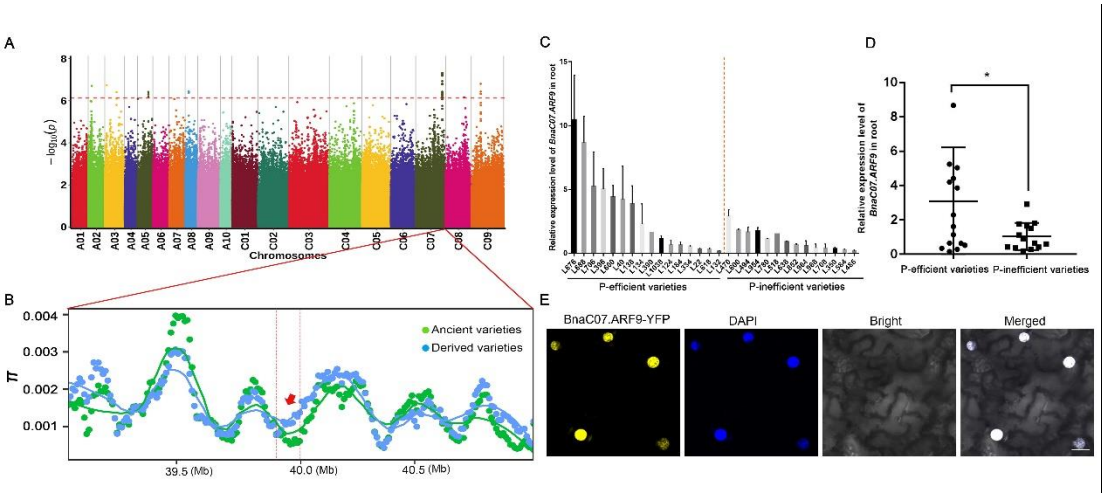
633 Trial 1 and Trial 2 by Fast-LMM model. (a) SY in Trial 1, (b) SY in Trial 2, (c) PEC in Trial 1, (d)

634 PEC in Trial 2. SY, seed yield; LP, low phosphorus; PEC, phosphorus efficiency coefficient.

635

636

637



638

639 Fig. 2 Nucleotide diversity (π), gene expression, and subcellular localization analysis of

640 *BnaC07.ARF9* gene. (a) Manhattan plot for SY_LP in Trial 1; (b) π analysis of 1000 kb up and

641 downstream of candidate gene *BnaC07.ARF9*; (c-d) The relative expression level of *BnaC07.ARF9*

642 in root of P-efficient and inefficient varieties; (e) Subcellular localization of *BnaC07.ARF9*. P,

643 phosphorus.

644

645

646

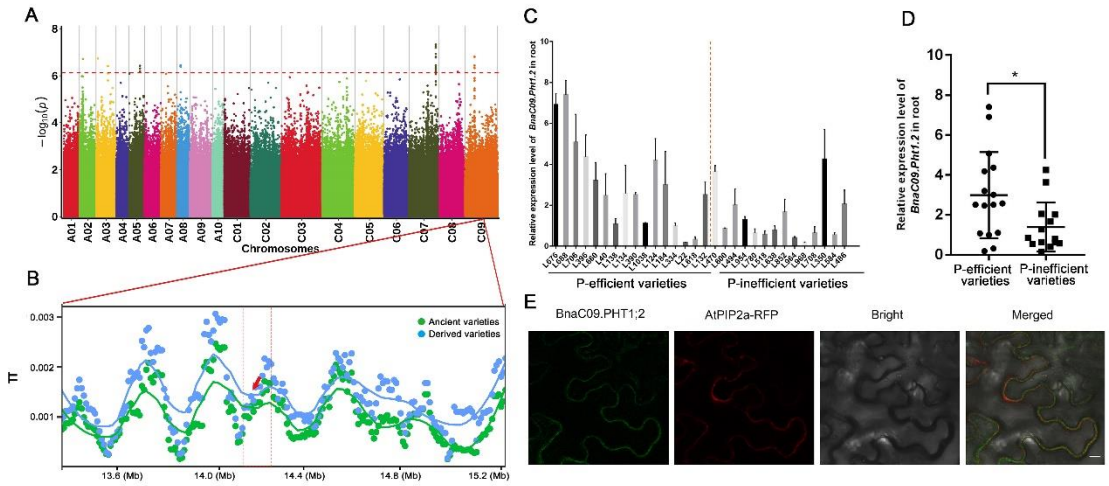


Fig. 3 Nucleotide diversity (π), gene expression, and subcellular localization analysis of *BnaC09.PHT1;2* gene. (a) Manhattan plot for SY_LP in Trial 1; (b) π analysis of 1000 kb up and downstream of candidate gene *BnaC09.PHT1;2*; (c-d) The relative expression level of *BnaC09.PHT1;2* in root of P-efficient and inefficient varieties; (e) Subcellular localization of *BnaC09.PHT1;2*. P, phosphorus.

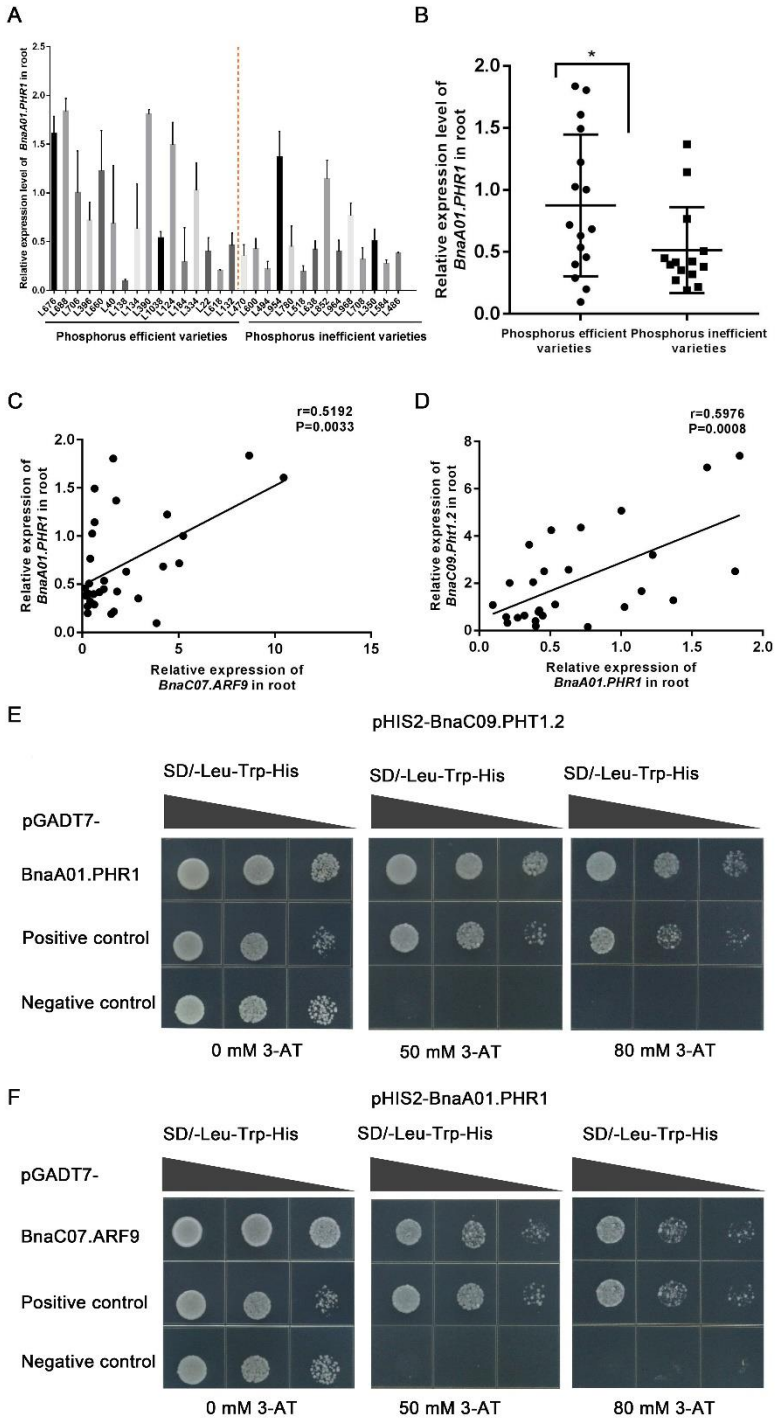
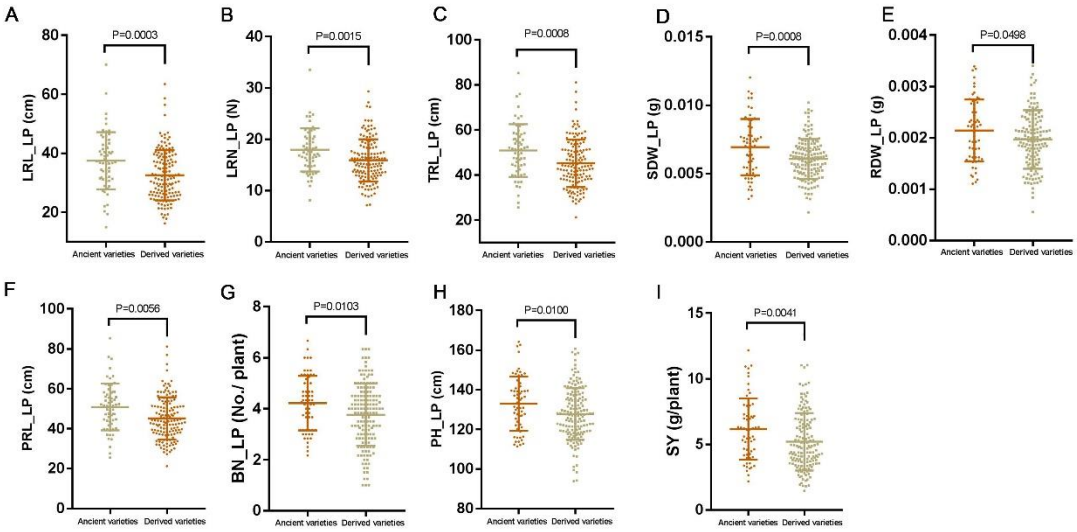


Fig. 4 *BnaC07.ARF9* regulates *BnaC07.PHT1;2* by regulating *BnaA01.PHR1* to achieve high phosphorus efficiency. (a-b) The relative expression level of *BnaA01.PHR1* in root of P-efficient and -inefficient varieties; (c) The correlation between the relative expression level of *BnaC07.ARF9* at LP and the relative expression level of *BnaA01.PHR1* at LP in P-efficient and -inefficient varieties; (d) The correlation between the relative expression level of *BnaA01.PHR1* at

LP and the relative expression level of *BnaC09.PHT1;2* at LP in P-efficient and -inefficient varieties; (e) Yeast one-hybrid assays of binding activity of *BnaC07.ARF9* with *BnaA01.PHR1*; (f) Yeast one-hybrid assays of binding activity of *BnaA01.PHR1* with *BnaC09.PHT1;2*. P, phosphorus; LP, low phosphorus.



667

668 Fig. 5 Differences in the root system architecture (a-f), SY and SY related traits (g-i) between ancient

669 and derived *B. napus* varieties at LP. LRL, lateral root length; LRN, lateral root number; TRL, total

670 root length; SDW, shoot dry weight; RDW, root dry weight; PRL, primary root length; BN, branch

671 number; PH, plant height; SY, seed yield; LP, low phosphorus.

672

673

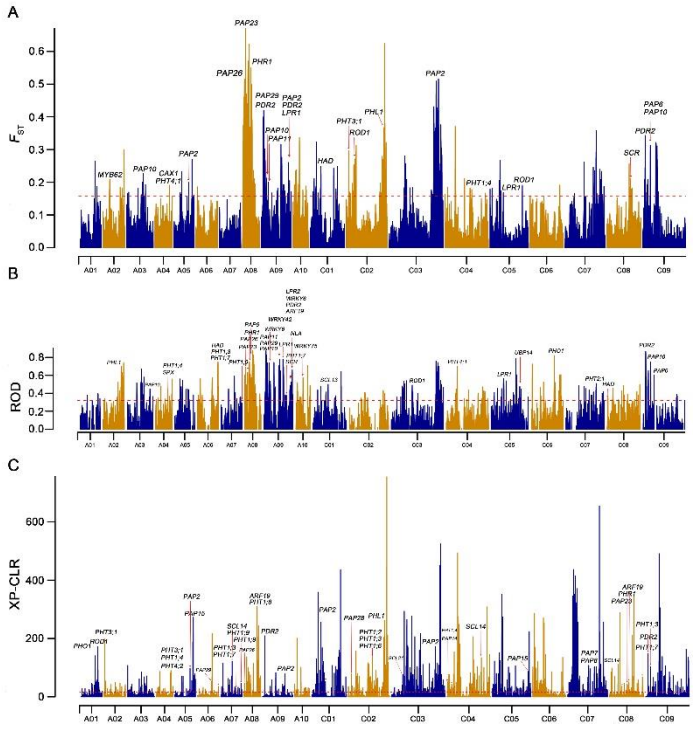


Fig. 6 Genome-wide distribution of selective-sweep signals identified through comparisons between ancient and derived varieties by F_{ST} (a), ROD (b), and XP-CLR (c) method. Genes related to phosphorus uptake, transport and utilization in selection outlier regions between ancient and derived varieties were labeled on the top of the corresponding chromosomes. The red dashed lines represent the thresholds (top 5% of F_{ST}, ROD, and XP-CLR values). F_{ST}, the fixation index; ROD, reduction of diversity; XP-CLR, the cross-population composite likelihood ratio test.

686 Table 1 Candidate genes associated with phosphorus uptake, transport and utilization in the selected sweep regions between the ancient and derived *B. napus*

Gene_ID	Arabidopsis homologous gene	Method	Description
BnaA01g23840D	AT3G23430.1	XP-CLR	Phosphate 1 (PHO1)
BnaA01g28480D	AT3G15820.1	XP-CLR	Reduced oleated desaturation 1(ROD1)
BnaA02g01980D	AT5G14040.1	XP-CLR	Phosphate transporter 3;1 (PHT3;1)
BnaA02g13730D	AT1G68320.1	F_{ST}	MYB domain protein 62 (MYB62)
BnaA02g30980D	AT5G29000.2	ROD	Homeodomain-like superfamily protein
BnaA03g36360D	AT3G21610.1	F_{ST}, ROD	Acid phosphatase/vanadium-dependent haloperoxidase-related protein
BnaA03g39080D	AT2G16430.2	F_{ST}, ROD	Purple acid phosphatase 10 (PAP10)
BnaA03g39090D	AT2G16430.2	F_{ST}, ROD	Purple acid phosphatase 10 (PAP10)
BnaA03g39100D	AT2G16430.2	F_{ST}, ROD	Purple acid phosphatase 10 (PAP10)
BnaA04g07160D	AT5G15070.1	ROD	Phosphoglycerate mutase-like family protein
BnaA04g21000D	AT2G38940.1	XP-CLR	Phosphate transporter 1;4 (PHT1;4)
BnaA04g21010D	AT2G38940.1	XP-CLR	Phosphate transporter 1;4 (PHT1;4)
BnaA04g21790D	AT2G38060.1	F_{ST}, XP-CLR	Phosphate transporter 1;4 (PHT1;4)
BnaA04g22260D	AT2G38920.1	ROD	SPX (SYG1/Pho81/XPR1) domain-containing protein
BnaA04g22280D	AT2G38940.1	ROD	phosphate transporter 1;4 (PHT1;4)
BnaA05g19610D	AT3G21610.1	ROD	Acid phosphatase/vanadium-dependent haloperoxidase-related protein
BnaA05g21920D	AT3G18220.1	F_{ST}, XP-CLR	Phosphatidic acid phosphatase (PAP2) family protein
BnaA05g30450D	AT3G07130.1	XP-CLR	Purple acid phosphatase 15 (PAP15)
BnaA06g05800D	AT1G09870.1	ROD	Histidine acid phosphatase family protein
BnaA06g22550D	AT5G63140.1	XP-CLR	Purple acid phosphatase 29 (PAP29)
BnaA06g22560D	AT5G63140.1	XP-CLR	Purple acid phosphatase 29 (PAP29)
BnaA06g24240D	AT5G15070.2	ROD	Phosphoglycerate mutase-like family protein
BnaA06g36380D	AT5G44020.1	ROD	HAD superfamily, subfamily IIIB acid phosphatase

BnaA06g36740D	AT5G43360.1	ROD, XP-CLR	Phosphate transporter 1;3 (PHT1;3)
BnaA06g36750D	AT5G43360.1	ROD, XP-CLR	Phosphate transporter 1;3 (PHT1;3)
BnaA06g36760D	AT3G54700.1	ROD, XP-CLR	Phosphate transporter 1;7 (PHT1;7)
BnaA07g19030D	AT3G61770.1	ROD	Acid phosphatase/vanadium-dependent haloperoxidase-related protein
BnaA07g21370D	AT1G76430.1	XP-CLR	Phosphate transporter 1;9 (PHT1;9)
BnaA07g26520D	AT1G67600.1	XP-CLR	Acid phosphatase/vanadium-dependent haloperoxidase-related protein;
BnaA07g32730D	AT1G76430.1	ROD, XP-CLR	Phosphate transporter 1;9 (PHT1;9)
BnaA07g32740D	AT1G76430.1	ROD, XP-CLR	Phosphate transporter 1;9 (PHT1;9)
BnaA07g32750D	AT1G76430.1	ROD, XP-CLR	Phosphate transporter 1;9 (PHT1;9)
BnaA08g01500D	AT1G52340.1	F_{ST}	ABA DEFICIENT 2 (ABA2)
BnaA08g04670D	AT3G01310.2	F_{ST}, XP-CLR	Phosphoglycerate mutase-like family protein
BnaA08g05220D	AT4G13700.1	F_{ST}, ROD	Purple acid phosphatase 23 (PAP23)
BnaA08g05520D	AT4G14930.1	F_{ST}, ROD, XP-CLR	Survival protein SurE-like phosphatase/nucleotidase
BnaA08g06550D	AT3G01310.2	F_{ST}, ROD	Phosphoglycerate mutase-like family protein;
BnaA08g07270D	AT5G34850.1	F_{ST}, ROD, XP-CLR	Purple acid phosphatase 26 (PAP26)
BnaA08g13620D	AT4G28610.1	F_{ST}, ROD	Phosphate starvation response 1 (PHR1)
BnaA08g14490D	AT4G26080.1	F_{ST}	ABA INSENSITIVE 1 (ABI1)
BnaA08g15210D	AT1G56360.1	ROD	Purple acid phosphatase 6 (PAP6)
BnaA08g19060D	AT1G24350.1	XP-CLR	Acid phosphatase/vanadium-dependent haloperoxidase-related protein
BnaA08g21590D	AT1G20860.1	XP-CLR	Phosphate transporter 1;8 (PHT1;8)
BnaA08g22150D	AT1G19220.1	XP-CLR	Auxin response factor 19 (ARF19)
BnaA09g03540D	AT5G29000.2	ROD	Homeodomain-like superfamily protein
BnaA09g05200D	AT5G23630.1	F_{ST}, XP-CLR	Phosphate deficiency response 2 (PDR2)
BnaA09g06490D	AT5G63140.1	F_{ST}, ROD	Purple acid phosphatase 29 (PAP29)

BnaA09g08700D	AT2G16430.2	F_{ST}, ROD	Purple acid phosphatase 10 (PAP10)
BnaA09g09490D	AT2G18130.1	F_{ST}, ROD	Purple acid phosphatase 11 (PAP11)
BnaA09g13370D	AT1G62300.1	ROD	WRKY6
BnaA09g20470D	AT4G04450.1	ROD	WRKY42
BnaA09g30170D	AT1G23010.1	F_{ST}, ROD	Low phosphate root1 (LPR1)
BnaA09g31580D	AT5G66450.2	F_{ST}, XP-CLR	Phosphatidic acid phosphatase (PAP2) family protein
BnaA09g34140D	AT3G54220.1	ROD	SCARECROW (SCR)
BnaA09g34510D	AT3G54700.1	ROD, XP-CLR	Phosphate transporter 1;7 (PHT1;7)
BnaA09g43980D	AT5G23630.1	F_{ST}, ROD	Phosphate deficiency response 2 (PDR2)
BnaA09g44280D	AT1G19220.1	ROD	Auxin response factor 19 (ARF19)
BnaA09g45250D	AT1G15080.1	F_{ST}, ROD	Lipid phosphate phosphatase 2 (LPP2)
BnaA09g45970D	AT1G13900.1	ROD	Purple acid phosphatases superfamily protein
BnaA09g46120D	AT1G13750.1	ROD	Purple acid phosphatases superfamily protein
BnaA09g51120D	AT1G02860.2	ROD	Nitrogen limitation adaptation (NLA)
BnaA09g51130D	AT1G02860.1	ROD	Nitrogen limitation adaptation (NLA)
BnaA10g20210D	AT5G13080.1	ROD	WRKY DNA-binding protein 75 (WRKY75)
BnaC01g17060D	AT4G25150.1	F_{ST}, XP-CLR	HAD superfamily, subfamily IIIB acid phosphatase
BnaC01g21110D	AT4G17230.1	ROD	SCARECROW-like 13 (SCL13);
BnaC01g22010D	AT3G50920.2	XP-CLR	Phosphatidic acid phosphatase (PAP2) family protein
BnaC01g28660D	AT1G60600.2	F_{ST}	ABERRANT CHLOROPLAST DEVELOPMENT 4 (ABC4)
BnaC02g05120D	AT5G14040.1	F_{ST}	Phosphate transporter 3;1 (PHT3;1);
BnaC02g11790D	AT5G57140.1	XP-CLR	Purple acid phosphatase 28 (PAP28);
BnaC02g13150D	AT3G15820.1	F_{ST}	REDUCED OLEATE DESATURATION 1 (ROD1);
BnaC02g13220D	AT3G01310.2	F_{ST}	Phosphoglycerate mutase-like family protein
BnaC02g15050D	AT5G52510.1	XP-CLR	SCARECROW-like 8 (SCL8)
BnaC02g28490D	AT4G25150.1	XP-CLR	HAD superfamily, subfamily IIIB acid phosphatase

BnaC02g30200D	AT5G43340.1	XP-CLR	Phosphate transporter 1;6 (PHT1;6)
BnaC02g30210D	AT5G43360.1	XP-CLR	Phosphate transporter 1;3 (PHT1;3)
BnaC02g30220D	AT5G43370.2	XP-CLR	Phosphate transporter 2 (PHT1;2)
BnaC02g30230D	AT5G43360.1	XP-CLR	Phosphate transporter 1;3 (PHT1;3)
BnaC02g30240D	AT5G43360.1	XP-CLR	Phosphate transporter 1;3 (PHT1;3)
BnaC02g30270D	AT5G43360.1	XP-CLR	Phosphate transporter 1;3 (PHT1;3)
BnaC02g34860D	AT2G03240.1	XP-CLR	EXS (ERD1/XPR1/SYG1) family protein
BnaC02g37670D	AT3G29060.1	F_{ST}, XP-CLR	EXS (ERD1/XPR1/SYG1) family protein
BnaC02g38370D	AT5G48150.2	XP-CLR	SCARECROW-like 21
BnaC02g39280D	AT5G29000.2	F_{ST}, XP-CLR	Homeodomain-like superfamily protein
BnaC02g39290D	AT5G29000.3	F_{ST}, XP-CLR	Homeodomain-like superfamily protein
BnaC03g18120D	AT2G32770.3	XP-CLR	Purple acid phosphatase 13 (PAP13)
BnaC03g28680D	AT2G04890.1	F_{ST}, XP-CLR	SCARECROW-like 21 (SCL21)
BnaC03g39290D	AT3G15820.1	ROD	Reduced oleated desaturation 1(ROD1)
BnaC03g65000D	AT4G22550.1	F_{ST}, XP-CLR	Phosphatidic acid phosphatase (PAP2)
BnaC03g65110D	AT4G22990.1	F_{ST}, XP-CLR	Major Facilitator Superfamily with SPX (SYG1/Pho81/XPR1) domain-containing protein
BnaC03g68600D	AT1G50420.1	F_{ST}, ROD	SCARECROW-like 3
BnaC04g00570D	AT2G46880.1	XP-CLR	Purple acid phosphatase 14 (PAP14)
BnaC04g06470D	AT2G38940.1	XP-CLR	Phosphate transporter 1;4 (PHT1;4)
BnaC04g06480D	AT2G38940.1	XP-CLR	Phosphate transporter 1;4 (PHT1;4)
BnaC04g15030D	AT2G29650.1	ROD	Phosphate transporter 4;1 (PHT4;1)
BnaC04g40230D	AT2G29060.1	XP-CLR	SCARECROW-like 14
BnaC04g46040D	AT2G38920.1	F_{ST}	SPX (SYG1/Pho81/XPR1) domain-containing protein
BnaC04g46050D	AT2G38940.1	F_{ST}	Phosphate transporter 1;4 (PHT1;4)
BnaC05g07350D	AT1G09870.1	XP-CLR	Histidine acid phosphatase family protein
BnaC05g18410D	AT1G23010.1	F_{ST}, ROD	Low Phosphate Root1 (LPR1)

BnaC05g26710D	AT1G50600.1	ROD	SCARECROW-like 5
BnaC05g26720D	AT1G50600.1	ROD	SCARECROW-like 5
BnaC05g32240D	AT3G20630.1	ROD	Ubiquitin-specific protease 14 (UBP14)
BnaC05g37480D	AT3G15820.1	F_{ST}	Reduced oleated desaturation 1(ROD1)
BnaC05g44840D	AT3G07130.1	XP-CLR	Purple acid phosphatase 15 (PAP15)
BnaC06g25630D	AT1G68740.1	ROD	Phosphate 1 (PHO1)
BnaC06g25640D	AT1G68740.1	ROD	Phosphate 1 (PHO1)
BnaC07g05120D	AT1G63010.4	XP-CLR	Major Facilitator Superfamily with SPX (SYG1/Pho81/XPR1) domain-containing protein
BnaC07g10420D	AT1G15080.1	XP-CLR	Lipid phosphate phosphatase 2 (LPP2)
BnaC07g11220D	AT1G24350.1	XP-CLR	Acid phosphatase/vanadium-dependent haloperoxidase-related protein
BnaC07g21290D	AT2G01880.1	XP-CLR	Purple acid phosphatase 7 (PAP7)
BnaC07g21320D	AT2G01880.1	XP-CLR	Purple acid phosphatase 7 (PAP7)
BnaC07g21330D	AT2G01880.1	XP-CLR	Purple acid phosphatase 7 (PAP7)
BnaC07g21340D	AT2G01890.1	XP-CLR	Purple acid phosphatase 8 (PAP8)
BnaC07g23650D	AT3G26570.1	ROD	Phosphate transporter 2;1 (PHT2;1)
BnaC08g00490D	AT1G04040.1	ROD	HAD superfamily, subfamily IIIB acid phosphatase
BnaC08g02280D	AT1G07530.1	XP-CLR	SCARECROW-like 14 (SCL14)
BnaC08g08870D	AT4G13700.1	XP-CLR	Purple acid phosphatase 23 (PAP23)
BnaC08g13160D	AT4G28610.1	XP-CLR	Phosphate starvation response 1 (PHR1)
BnaC08g18670D	AT1G19220.1	XP-CLR	Auxin response factor 19 (ARF19)
BnaC08g25070D	AT3G54220.1	F_{ST}	SCARECROW (SCR)
BnaC08g25460D	AT3G54700.1	XP-CLR	Phosphate transporter 1;7 (PHT1;7)
BnaC08g42510D	AT1G09870.1	XP-CLR	Histidine acid phosphatase family protein
BnaC09g04770D	AT5G23630.1	F_{ST}, ROD, XP-CLR	Phosphate deficiency response 2 (PDR2)
BnaC09g08970D	AT2G16430.2	F_{ST}, ROD	Purple acid phosphatase 10 (PAP10)

BnaC09g09650D	AT1G56360.1	F_{ST}, ROD	Purple acid phosphatase 6 (PAP6)
BnaC09g12880D	AT1G63010.4	XP-CLR	Major Facilitator Superfamily with SPX (SYG1/Pho81/XPR1) domain-containing protein
BnaC09g17240D	AT5G43360.1	XP-CLR	Phosphate transporter 1;3 (PHT1;3)
BnaC09g23750D	AT4G11810.1	XP-CLR	Major Facilitator Superfamily with SPX (SYG1/Pho81/XPR1) domain-containing protein

687

688