Genome-wide association study dissects the genetic control of plant height and branch number in response to low-phosphorus stress in Brassica napus

Article

Accepted Version

Original Article

Genome-wide association study dissects the genetic control of plant height and branch number in response to low-phosphorus stress in *Brassica napus*

Haijiang Liu[1, 2, #], Jingchi Wang[1, 2, #], Bingbing Zhang[1, 2], Xinyu Yang[1, 2], John P Hammond[3,4], Guangda Ding[1,2], Sheliang Wang[1,2], Hongmei Cai[2], Chuang Wang[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] *Southern Cross Plant Science, Southern Cross University, Lismore, NSW 2480, Australia*

Running title: Genetic control of PH and BN of *B. napus* at LP

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

Background and Aims Oilseed rape (*Brassica napus*) is one of the most important oil crops worldwide. Phosphorus (P) deficiency severely decreases the plant height (PH) and branch number (BN) of *B. napus*. However, the genetic bases controlling PH and BN in *B. napus* under P deficiency remain largely unknown. This study aims to mine candidate genes for PH and BN by genome-wide association study (GWAS) and determine low-P tolerance haplotypes.

Methods An association panel of *B. napus* were grown in the field with a low P supply (P, 0 kg/ha) and a sufficient P supply (P, 40 kg/ha) across two years and PH and BN were investigated. More than five million single-nucleotide polymorphisms (SNPs) were used to conduct GWAS of PH and BN at two contrasting P supplies.

Key Results A total of 2127 SNPs were strongly associated ($P < 6.25\times10^{-07}$) with PH and BN at two P supplies. There was significant correlation between phenotypic variation and the number of favorable alleles of associated loci on chromosomes A10 (chrA10_821671) and C08 (chrC08_27999846), which will contribute to breeding improvement by aggregating these SNPs. *BnaA10g09290D* and *BnaC08g26640D* were identified to be associated with the chrA10_821671 and chrC08_27999846, respectively. Candidate gene association analysis and haplotype analysis showed that the inbred lines carrying ATT at '*BnaA10g09290Hap1*' and AAT at '*BnaC08g26640Hap1*' had higher PH than lines carrying other haplotype alleles at low P supply.

Conclusion Our results demonstrate the power of GWAS in identifying genes of interest in *B. napus* and provided insights into the genetic basis of PH and BN at low P

supply in *B. napus*. Candidate genes and favorable haplotypes may facilitate marker-based breeding efforts aimed at improving P use efficiency in *B. napus*.

Keywords: oilseed rape; genome wide association study; plant height; branch number; low phosphorus supply; haplotype analysis

INTRODUCTION

Oilseed rape (*B. napus*) is one of the most important sources of vegetable oil globally. The global production of vegetable oil reached 75 million tons in 2020 (http://www.fao.org/faostat/zh/#data/QC/visualize). Oilseed rape canopy architecture is determined by plant height, the length of the main inflorescence and the number and distribution of ancillary branches (Li et al 2016). Canopy architecture indirectly influences cultivar yield potential by significantly influencing the number of siliques per plant. It has previously been shown that plant height negatively correlates with siliques per plant, owing to the greater lodging risk of taller plants and branch number positively correlates with the number of siliques per plant (Qiu *et al*., 2006; Chen *et al*., 2014).

Phosphorus (P) is an essential macro-element for plant growth and development (Kochian, 2012). Oilseed rape has a high P requirement and P deficiency reduces PH and BN, and the subsequent seed yield of *B. napus* (Shi *et al*., 2013). P in the soil is easily fixed by metal cations to form insoluble compounds or bound in organic forms, which are difficult to be directly absorbed and used by plants (Holford, 1997). Therefore, P fertilizers must be supplied to ensure the normal growth and development of crops. The use of P fertilizer increased from 34.5 million tons in 2001 to 45.4 million tons in 2017 (http://www.fao.org/faostat/zh/#data/QC/visualize). However, long-term excess P fertilizer is problematic because of its limited bioavailability and potential environmental problems such as eutrophication, and the depletion of non-renewable rock phosphate resources (Vance *et al*., 2003). Therefore, uncovering the genetic

mechanisms of *B. napus* tolerance to low P, and breeding high P-efficient varieties is an important way to reduce P use in agricultural systems.

Linkage mapping analysis has been widely used to study the genetic basis of P tolerance in *B. napus* (Yang *et al*., 2010a; Yang *et al*., 2010b; Shi *et al*., 2013; Ding *et al*., 2012). Sixty-two significant quantitative trait locus (QTLs) were associated with plant P uptake, total root surface area, root length, root volume and total dry weight under high and low P conditions in three experiments, explaining 12.7 %~31.9% of the phenotypic variation (Yang *et al*., 2010a). Ding *et al*. (2012) detected 21 significant QTLs associated with PH and BN under high and low P conditions, explaining 8.5%~19.8% of the phenotypic variation. One hundred and fifty-five significant QTLs were associated with seed yield and yield-related traits from three different trials, and 79 QTLs detected at LP (a deficient P supply) and 76 under SP (a sufficient P supply) (Shi *et al*. 2013). Among them, 16 QTLs were associated with PH and 5 QTLs were associated with BN (Shi *et al*., 2013). However, due to the limited number of genetic markers and the frequency of recombination in mapping populations, the physical interval for these QTL are usually large, which makes it difficult to determine the candidate genes (Xiao *et al*., 2017).

Compared with QTL mapping, genome wide association study (GWAS) can achieve a higher resolution of the underlying genetic loci, which greatly improves the efficiency of gene mapping (Nordborg and Weigel, 2008). GWAS has been widely applied to the genetic dissection of complex traits in crops, such as grain size in rice (Si *et al*., 2016), PH in maize (Zhang *et al*., 2019) and silique number in *B. napus* (Li *et al*., 2020).

GWAS combined with QTL analysis has been used successfully to mine candidate genes in *B. napus*, such as root related traits (Wang et al., 2017) and clubroot resistance (Laperche et al., 2017) in *B. napus*, and capsaicinoid content in Capsicum (Han et al., 2018).

GWAS of the PH and BN of *B. napus* under sufficient P conditions has been conducted in several studies (Li et al., 2016; Sun et al., 2016; Zheng et al., 2017). However, the genetic bases controlling PH and BN in *B. napus* under P deficiency remain unknown. In this study, we investigated the PH and BN of a *B. napus* association panel at low and sufficient P supplies across two years. GWAS of PH and BN was performed using high-quality SNPs by whole-genome resequencing (Tang *et al*., 2020). We aimed to identify the (i) genetic diversity of the population, (ii) significant SNPs and candidate genes associated with PH and BN at contrasting P supplies (iii) and reveal the favourable haplotypes for breeding P-efficient *B. napus* cultivars.

MATERIALS AND METHODS

Plant materials and growth conditions

The association panel of *B. napus* comprises 403 cultivars and inbred lines, including 350 semi-winter, 44 spring, 8 winter and 1 unknown type, collected from major *B. napus* breeding centers across China (Supplementary Data Table S1). Of them, 361 lines originated in China, 21 from Europe, 8 from Japan, 5 from Canada, 4 from Australia, 3 from Korea and 1 unknown (Supplementary Data Table S1). The panel were grown in the field with a low P supply (P, 0 kg/ha) and a sufficient P supply (P, 40

kg/ha) with three replications at Meichuan Town, Wuxue city, Hubei province, China (E 115.55, N 29.85°) from 2018 to 2019 (Trial 1) and from 2019 to 2020 (Trial 2). The soil was sandy loam soil. The top soil (0-30 cm) were collected before sowing (before fertilization) for determination of the available nutrients concentrations (Table 1). 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 4 rows and each plot had 8 plants each row. At the mature stage, six plants each plot were selected to measure the PH and BN. PH was the length of the plant from the base of the stem to the tip of the main inflorescence, and BN was calculated by the number of primary branches arising from main shoot. PHr and BNr were defined as the ratio of PH_LP to PH_SP, and that of BN_LP to BN_SP, respectively.

Genome-wide association analysis and candidate gene identification

More than 10 million high-quality SNP markers of the association panel of *B. napus* were derived from previous studie (Tang et al., 2020). After filtering the SNPs with minor-allele frequency (MAF) > 0.05 and missing rate < 0.2, we obtained 5.58 million SNPs for GWAS. GWAS for PH was carried out using general linear models (GLM) and mixed linear models (MLM) by the Tassel 5.0 software (Bradbury *et al*., 2007). In

order to minimize the contribution from regions of extensive strong LD, we scanned the whole genome with a sliding window of 500 kb (in steps of 100 SNPs), and used Plink software to remove SNPs related with other SNPs within the window with correlation coefficient (R^2) > 0.2. Finally, we obtain 497761 SNPs. Admixture software was used to calculate the Q matrix and Tassel 5.0 software was used to calculate the K matrix. The Manhattan plot was drawn by ggplot2 (https://cran.r-project.org/web/packages/ggplot2/index.html) software and the Quantile-Quantile plot was drawn by the CMplot software (https://github.com/YinLiLin/CMplot). The threshold for the significance of associations between SNPs and traits was $P < 6.25\times10^{-07}$. The linkage disequilibrium (LD) statistic was calculated by PopLDdecay software (Zhang *et al*., 2019). According to the LD decay (238kb) of the panel, the genes located within 300 kb upstream and downstream of the peak SNPs were considered as candidate genes. GO enrichment analysis of the candidate genes were performed with the omicshare online platform (https://www.omicshare.com/quote.php). The genotypes of *BnaA02g33340D*, *BnaA10g09290D* and *BnaC08g26640D* in the association panel of *B. napus* were obtained by bcftools software (http://samtools.github.io/bcftools/bcftools.html). Candidate gene association analysis of *BnaA02g33340D*, *BnaA10g09290D* and *BnaC08g26640D* were performed with Tassel 5.0 software (Bradbury *et al*., 2007). The SNP markers from 2 kb up-stream of the gene to 2 kb down-stream of the gene were used to conduct association analysis with the PH of the association panel of *B. napus* at both P supplies.

Haplotype analysis

The haplotype analysis was predicted using the HaploView software (Barrett *et al*., 2005). One-way ANOVA and Student's *t*-test were employed to compare the differences in PH among the haplotypes. Haplotypes containing at least twenty *B. napus* cultivars were used for further comparative analysis. Student's *t*-test was used to compare the differences in PH among the haplotypes.

Statistical Analysis of Phenotypic Data

The mean values of PH and BN of six plants of each replication were calculated using Excel 2007. Best linear unbiased prediction (BLUP) of PH and BN at a deficient P supply (LP) and a sufficient P supply (SP) for each line was calculated using the R package 'lme4' (https://cran.r-project.org/web/packages/lme4/index.html). Trial1_BLUP and Trial2_BLUP were calculated with the phenotypic values of three replications in Trial 1 and Trial 2, respectively (Supplementary Data Table S2). R language was used to calculate the correlation coefficients between phenotypes. The broad-sense heritability was calculated as: $H^2 = V_G/ (V_G/ (V_G+V_E)/nr)$, and among them V_G is genetic variance, V_E is environmental variance, n is the number of environments and r is the number of replicates.

RESULTS

Phenotypic variation for PH and BN of an association panel of B. napus at low and sufficient P supplies

From the seedling stage to the silique stage under LP, the old leaves turned dark and purple, and the inhibition of root and shoot growth was observed in most lines (Supplementary Data Fig. S1A-E). In addition, the PH and BN of *B. napus* decreased under LP compared to SP (Supplementary Data Fig. S1F-I). Extensive phenotypic variations for PH and BN were observed in the association panel of *B. napus* at LP and SP (Supplementary Data Fig. S1G; S2, Table 2, Supplementary Data Table S2). Under LP, PH varied from 57.5 to 179.3 cm (Supplementary Data Fig. S2, Table 2, Supplementary Data Table S2) and BN varied from 0 to 11 per plant across the 2 years (Supplementary Data Fig. S2, Table 2, Supplementary Data Table S2). Under SP, PH varied from 84.0 to 235.0 cm (Supplementary Data Fig. S2, Table 2, Supplementary Data Table S2) and BN varied from 1 to 16 per plant across the 2 years (Supplementary Data Fig. S2, Table 2, Supplementary Data Table S2). PHr varied from 0.40 to 1.00 and BNr varied from 0.10 to 1.00 across the 2 years (Supplementary Data Fig. S2, Table 2, Supplementary Data Table S2). High h^2 values were observed for all traits (Table 2). The correlation coefficients of PH_SP, BN_SP, PH_LP, BN_LP, PHr and BNr between Trial 1 and Trial 2 were 0.57, 0.28, 0.27, 0.09, 0.19 and 0.09, respectively (Supplementary Data Fig. S3). Thus, best linear unbiased prediction (BLUP) analysis was employed to deal with multi- year and field phenotypic values of PH and BN.

Population structure, relative kinship, and LD decay

A total of 5.58 million SNP markers were identified for this *B. napus* association panel (Supplementary Data Table S3). SNP number on each chromosome ranged from

171159 on A08 to 475529 on C03 (Supplementary Data Table S3). LD decay on each chromosome ranged from 1996 kb on C07 to 45 kb on A03, when r^2 was 0.1 (Supplementary Data Table S3). The LD decay on the whole genome was 238 kb (Supplementary Data Fig. S4, Supplementary Data Table S3). A total of 497761 SNPs were selected to assess the population structure, relative kinship and LD. The population could be divided into five subgroups based on the cross validation (CV) errors (Supplementary Data Fig. S5). The r pairwise relative kinship is close to 0 (Supplementary Data Fig. S6, Supplementary Data Table S4). For example, the values of the relative kinships were 0 to 129976 pairs and 0.1 to 148803 pairs, and the ratios to the total value were 80.03% and 91.62%, respectively (Supplementary Data Table S4). These results showed that the genetic distance of the majority of the accessions in the association panel were large enough for the GWAS analysis.

Genome wide association mapping of PH and BN of B. napus at low and sufficient P supplies

We performed GWAS with GLM and MLM approaches to identify SNPs associated with PH and BN at LP and SP in *B. napus*. A total of 1289 SNPs were identified to be significantly associated with PH of *B. napus* at LP and SP across two years ($P < 6.25\times10^{-07}$) (Supplementary Data Fig. S7 A-D, Supplementary Data Table S5-S6). Among them, 133, 685 and 471 SNPs were associated with PH_LP, PH_SP and PHr, respectively. The GLM analysis detected a total of 1275 significant SNPs at both P supplies. Among them, 131, 678 and 466 SNPs were associated with PH_LP, PH_SP and PHr, respectively, and

were distributed on all chromosomes, explaining between 6.03% and 14.01% of the phenotypic variation (Supplementary Data Table S5). Chromosome C03 had the largest number of significant SNPs (244) and chromosome A07 had the least number of significant SNPs (7 SNPs) (Supplementary Data Table S5). MLM analysis detected 106 significant SNPs on 18 of the 19 *B. napus* chromosomes (excluding C02) at both P supplies. Among them, 15, 74 and 17 SNPs were associated with PH_LP, PH_SP and PHr, respectively, and explained between 7.85% and 14.23% of the phenotypic variation (Supplementary Data Table S6). Among them, 93 significant SNPs were simultaneously identified by the two models (Supplementary Data Table S5-S6).

For BN, the GLM analysis detected a total of 837 significant SNPs at both P supplies. Among them, 34, 770 and 33 SNPs were associated with BN_LP, BN_SP and BNr, respectively, which werer distributed on all chromosomes and explained between 6.12% and 15.83% of the phenotypic variation (Supplementary Data Fig. S7 E-H, Supplementary Data Table S5-S6). MLM analysis detected 93 significant SNPs at both P supplies. Among them, three and 90 SNPs were associated with BN_LP and BN_SP, respectively, which explained between 6.92 % and 14.47% of the phenotypic variation. Of them, 90 SNPs were detected by both the GLM and MLM analyses (Supplementary Data Table S5-S6).

Candidate genes for PH and BN at low P supply in B. napus

We identified 1289 significant SNPs significantly associated with PH_LP and PHr. 2, 19, 3, 9, 88, 5 and 5 significant SNPs were associated with PH_LP in Trial 1_R1,

Trial 1_R2, Trial 1_R3, Trial 1_BLUP, Trial 2_R1, Trial 2_R2 and Trial 2_BLUP, respectively (Supplementary Data Fig. S7 A-D, Supplementary Data Table S5-S6). 15, 17, 12, 23, 250, 1 and 158 significant SNPs were associated with PHr in Trial 1_R1, Trial 1_R2, Trial 1_R3, Trial 1_BLUP, Trial 2_R1, Trial 2_R2 and Trial 2_BLUP, respectively (Supplementary Data Fig. S7 A-D, Supplementary Data Table S5-S6). The SNP of chrA10_8216711 on chromosome A10 was associated with PH_LP and PHr in both Trial 2_R2 and Trial 2_BLUP (Fig 1A, B). In this study, the LD decay was 238 kb for this association panel (Supplementary Data Table S3). Based on the LD decay, 300 kb up/downstream of the significant SNPs were selected to identify candidate genes on A10 and 76 candidate genes were detected.

The significant SNP associated with PH_LP and PHr identified at the 8216711 bp position on chromosome A10 was found located in the genetic region of *BnaA10g09290D*, whose function was unknown. Candidate gene association analysis of *BnaA10g09290D* with SNP markers from 2 kb promoter region and the entire coding region showed that three SNPs in *BnaA10g09290D* were significantly associated with the PH_LP and PHr, and had strong LD with each other (Fig. 1C, H). Further analysis demonstrated that A allele of chrA10_8216680, T allele of chrA10_8216711 and T allele of chrA10_8216756 were the P deficiency tolerant alleles (Fig. 1). Two major haplotypes based on the three significant SNPs were detected in this association panel, among which '*BnaA10g09290Hap2*' (TCA) was P deficiency sensitive, while '*BnaA10g09290Hap1*' (ATT) were P deficiency tolerant (Fig. 1G, L). Notably, chrA10_8216680 (T/A) located in the exon region of the gene *BnaA10g09290D* and

resulted in amino acid changes from isoleucine to asparagine and might contribute to the phenotypic difference in PH_LP and PHr (Table 3).

The region significantly associated with PHr in both Trial 1_R2, Trial 1_R3, Trial 1_BLUP and Trial 2_BLUP on chromosome C08 ranged from 27.95 Mb to 28.13 Mb (Fig. 2A, Fig. 2B). Based on the LD decay (238kb), 300 kb upstream and downstream regions of the significant SNP (chrC08_27999846, P = 8.91E-08, PVE = 12.46%) were selected and found to contain 101 genes. The lead SNP of chrC08_27999846 was located within *BnaC08g26640D*. Three SNPs in *BnaC08g26640D* were detected to be significantly associated with the PHr (Fig. 2C), and A allele of chrC08_27999709, A allele of chrC08_27999778 and T allele of chrC08_27999846 were the P deficiency - tolerant alleles (Fig. 2D, E, F). These three significant SNPs revealed two major haplotypes, and '*BnaC08g26640Hap1*' (AAT) had significantly greater PHr value than '*BnaC08g26640Hap2*' (GTC) (P = 1.51E-15) (Fig. 2G). The SNP 'chrC08_27999709' (A/G) was located in the intron region, and did not result in any amino acid change (Table 3). The SNP of chrC08_27999846 (T/C) was in the exon region, and it was a synonymous mutation that did not result in any amino acid change (Table 3). The SNP of chrC08_27999778 (A/T) located in the exon region of the gene *BnaC08g26640D* and resulted in amino acid changes from isoleucine to asparagine and might contribute to the phenotypic difference in PHr (Table 3).

Three SNPs associated with BN_LP, chrA01_13846343, chrA03_20898013 and chrC07_521008, were detected simultaneously by GLM and MLM (Supplementary Data Table S5-S6). On chromosome A01, 40 candidate genes were detected underlying

the region around chrA01_13846343 (Supplementary Data Table S7). The ATP phosphoribosyl transferase 1 (*BnaA01g21560D*) and low temperature and salt responsive protein (*BnaA01g21470D*) were identified in the candidate region. On chromosome A03, the significant SNP chrA03_20898013 was also detected by GLM and MLM model simultaneously, with P = 4.97E-07 and R^2 =9.21% (Supplementary Data Table S5-S6). Two genes encoding heat shock transcription factor A7A (*BnaA03g41540D* and *BnaA03g41550D*) were identified in the upstream region from chrA03_20898013 and their homolog in *Arabidopsis* responds to abiotic stresses (Lin *et al*., 2018). In addition, a phosphatidic acid phosphatase family protein (*BnaA03g41050D*) and a pyruvate orthophosphate dikinase (*BnaA03g41960D*) were identified in the down-stream region of chrA03_20898013. On chromosome C07, a total of 62 genes were detected underlying the candidate region around the SNP 'chrC07_521008' (P= 5.95E-07). Among these candidate genes, an EXP1 gene (*BnaC07g00140D*), which is involved in unidimensional cell growth and plant-type cell wall loosening, and may be involved in BN determination at LP (Supplementary Data Table S8).

Comparison of the significant SNPs with QTLs for P efficiency

Based on the Darmor-*bzh* reference genome, we analyzed the co-localization of the significant SNPs detected in our study and the QTLs detected by previous study (y (Shi *et al*., 2013). Thirty-three significant SNPs detected by GWAS co-localized with the intervals of the QTLs for the same traits in a previous linkage analyses of the

*Bna*TNDH population (Shi *et al*., 2013), including four SNPs on A02, seven SNPs on A09, seven SNPs on C06 and fifteen SNPs on C07 (Supplementary Data Table S9). Among them, four co-located SNPs on A02 chromosome associated with PH_LP (Fig. 3A-3C), including the P-tolerant T allele of chrA02_23692807, C allele of chrA02_23713660, A allele of chrA02_23899688 and T allele of chrA02_23912345 (Fig. 3D). A total of 152 candidate genes were within the QTL (PH_LP1_A02a; Shi *et al*., 2013) confidence interval on A02 (Fig. 3A, Supplementary Data Table S10). Among the four co-localized SNPs, the lead SNP chrA02_23899688 (P = 1.85E-07) located in the CDS region of *BnaA02g33340D* (Supplementary Data Table 3). Since *BnaA02g33340D* was identified by association with PH_LP, we examined the association of sequence variation in *BnaA02g33340D* with the PH_LP of the association panel. Five SNPs in *BnaA02g33340D* were detected to be significantly associated with the PH_LP (Fig. 3E). Further analyses showed that G allele of chrA02_23899623, G allele of chrA02_23899654, T allele of chrA02_23899669, C allele of chrA02_23899686 and A allele of chrA02_23899688 were P-tolerant alleles (Fig. 3F). There were two major haplotypes associated with PH and Hap1 (GGTCA) had higher PH at low P than Hap2 (ATATC) with a P-value of 0.0010, suggesting that the variation in *BnaA02g33340D* was associated with PH_LP (Fig. 3G). Further analysis indicated that chrA02_23899623 (G/A), chrA02_23899654 (G/T) and chrA02_23899669 (T/A) were located in the intron region, and the SNP 'chrA02_23899686 (C/T)' was located in the exon region and was a synonymous mutation, which were unlikely to affect the function of the *BnaA02g33340D* protein

(Table 3). The SNP 'chrA02_23899688 (A/C)' was located in the exon region of the gene *BnaA02g33340D* and resulted in amino acid changes from histidine to proline and may be causative in the difference of PH_LP (Table 3). These observations demonstrated that *BnaA02g33340D* is most likely to be the candidate gene for the locus represented by lead SNP chrA02_23899688.

DISCUSSION

Significant differences in the PH and BN in the association panel of B. napus at low P supply

In this study, the soil available P concentration in three fields ranged from 11.56 to 16.95 mg/kg, which are between P deficient and slightly P deficient (Zou et al. 2009.). Some typical P deficient symptoms were observed in most lines in the field, such as greater anthocyanin visibility in old leaves, restricted shoot growth at the seedling stage (Supplementary Data Fig. S1 A-C), and reduced PH and BN at the mature stage (Supplementary Data Fig. S1 D-F). PH and BN are essential components of plant canopy architecture in *B. napus*, which are closely correlated with its yield (Zheng *et al*., 2020). Compared to SP, PH and BN of 'Eyou Changjia' (a P-efficient variety) decreased by 11% and 25.4%, respectively; and PH and BN of B104-2' (a P-inefficient variety) decreased by 23%, and 46.9%, respectively (Ding *et al*., 2012). In addition, the PH and BN of 'Eyou Changjia' were higher than those of 'B104-2' under LP in two field Trials (Ding *et al*., 2012). Cultivars 'Tapidor' and 'Ningyou 7' are parental cultivars of the *Bna*TNDH population and PH and BN under LP were also less than

those under SP (Shi *et al*., 2013). Compared to the control (no P application), the PH of *B. napus* increased by 11.2%, 16.7% and 19.4%, respectively; and the BN of *B. napus* increased by 25.5%, 36.2% and 40.4% in the treatments of P application rate of 103.0, 206.0 and 309.0 kg/hm^2 (Lu *et al*., 2005).

Extensive phenotypic variations for PH and BN are observed in several natural populations of *B. napus* at SP (Li *et al*., 2016; Sun *et al*., 2016; Luo *et al*., 2017; Zheng *et al*., 2017). For example, the PH varied from 48.3 to 228.3 cm among 496 *B. napus* accessions, with an average ranging from 133.5 cm to 187.3 cm across six environments (Sun *et al*., 2016); and varied from 86.2 cm to 206.0 cm among 333 *B. napus* accessions across 2 years, with 1.6 to 2.4 fold variations across the two years (Zheng *et al*., 2017). In this study, we measured PH and BN in a panel of 403 *B. napus* accessions across two years under both LP and SP conditions. At SP, the PH varied from 84.0 cm to 235.0 cm and BN varied from 1 to 15 per plant (Table 2; Supplementary Data Table S1 and Fig. S2). Compared with the association panel with 496 *B. napus* accessions (Sun *et al*. 2016), the natural population in this study has more extensive genetic variation at SP. At LP conditions, the PH varied from 57.5 cm to 179.2 cm and BN varied from 0 to 11 per plant (Table 2; Supplementary Data Table S1 and Fig. S2). These data show that P deficiency significantly reduced the PH and BN of *B. napus* compared to those grown under SP in our study and others. In addition, extensive phenotypic variations for PH and BN at LP would be useful for the screening the P efficient cultivars that have greater potential to yield higher under LP.

Candidate genes associated with significant SNPs for PH and BN of B. napus under LP

GWAS has been successfully applied to the genetic dissection of low-P tolerance in different crops, such as in maize (Xu *et al*. 2018; Luo *et al*. 2019), rice (Wissuwa *et al*. 2015), wheat (Liu *et al*. 2015), soybean (Zhang *et al*. 2014) and oilseed rape (Wang *et al*. 2017). However, up to now, there is no report on the GWAS of *B. napus* PH and BN under LP, major determinants of seed yield and yield in *B. napus*. A total of 2127 significant SNPs were detected for PH and BN and thirty-three SNPs were identified co-localized with previously-identified QTLs (Supplementary Data Table S5, S6, S9). Closer examination of allelic frequency of the four co-located SNPs on chromosome A02 among this association panel indicated that the inbred lines carrying a minor allele have lower PH than inbred lines carrying major allele under LP (Fig. 3D). Interestingly, the PH of the accessions carrying the minor allele and major allele did not meet significant difference at a sufficient P supply (Supplementary Data Fig. S8). The four co-localized SNPs decrease the PH at low P and might affect the seed yield. ChrA10_8216711 for PH_LP and PHr in both tria2 _R2 and tria2_BLUP was located in the genetic region of a gene with unknown function, *BnaA10g09290D*. ChrC08_27999846 for PHr was located in the exon of *BnaC08g26640D*. We performed candidate gene association analysis and haplotype analysis for *BnaA10g09290D* and *BnaC08g26640D*, and high P efficiency haplotypes of "*BnaA10g09290Hap1*" and "*BnaC08g26640Hap1"* were identified, respectively (Fig. 1G, Fig. 1L, Fig. 2G).

In addition, some genes reported to be associated with P uptake and homeostasis were also detected in our association analysis (Supplementary Data Table S11). The

PHOSPHATE1 (*PHO1*), a gene which encodes a membrane protein consisting of SPX, plays an important role in loading P into xylem in roots (Hamburger *et al*. 2002; Pacak *et al*. 2016; Che *et al*. 2020). The peak SNP 'chrA07_6641098' on A07 was identified to be located within 280 kb of *B. napus BnaA07.PHO1* (Supplementary Data Table S11). Many plant proteins with SPX domains are involved in Pi signaling (Wang *et al*. 2004; Duan *et al*. 2008). In this study, *BnaC03g65110D*, a gene with SPX (SYG1/Pho81/XPR1) domain-containing protein, was located within the interval of the SNP 'chrC03_54249584' for the trait of PHr (Supplementary Data Table S11). *BnaC05g27910D* was a homologous gene of *BnaC03g65110D*, also containing SPX (SYG1/Pho81/XPR1) domain, and was located within the interval of the leading SNP 'chrC05_25541489' for the trait of PHr (Supplementary Data Table S11). In *Arabidopsis*, *AtUBP14* is involved in the response of root systems to P deficiency (Lee *et al*., 2008). In this study, *BnaA08g08470D* (*BnaA08.UBP20*), a gene which contains ubiquitin-specific protease activity, was located within the interval of the SNP 'chrA08_8488725' and linked with the trait for PHr (Supplementary Data Table S11). PHT1 family encode plant Pi transporters to transport P into plants (Huang *et al*., 2019). Recently, *BnaPht1*;*4* was found to participate in the uptake and transportation of P, and promotes seed germination and seedling growth of *B. napus* by regulating the biosynthesis of ABA and GA (Huang *et al*., 2019). In this study, we identified the *BnaA04g22280D* (*BnaA04.PHT11;4*) by SNP 'chrA04_16843321' for BN_LP, which was likely to affect the growth and development of BN under LP. (Supplementary Data Table S11).

In addition, we identified BnaA09g16430D (BnaA09.Pht1;6) by SNP chrA09_9605643 for BN_LP, which is located 170 kb downstream of the SNP chrA09_9605643 (Supplementary Data Table S11). Purple acid phosphatases (PAPs) are a family of binuclear metalloenzymes (Olczak *et al*., 2003). Overexpression of OsPAP10c (purple acid phosphatase 10c) enhances the utilization of phytate-P, root growth and yield at LP in rice (Deng *et al*., 2020). BnaA03g41050D (BnaA03.PAP2) was located 310 kb downstream of the SNP chrA03_20898013 and linked with the trait for BN_LP (Supplementary Data Table S11). These candidate genes might play important roles in the development of PH and BN in *B. napus* at a deficient P supply.

Co-location of the significant SNPs for root system architecture traits and PH and BN of B. napus at a deficient P supply

Previously, 285 SNPs were detected associated with root related traits of *B. napus* at low P supply and a P efficient haplotype '*BnA03Hap*' on A03 chromosome was identified (Wang *et al*., 2017). Fifty-three significant SNPs in this study were adjacent to previously published significant SNPs associated with root system architecture at low P supply, which may control both root system and plant height or branch numbers at low P (Supplementary Data Table S12). For example, SNP chrA08_10449680 was associated with PHr in this study and chrA08_10481443 was associated with Primary Root Length under LP (PRL_LP) in Wang et al. (2017), respectively. *BnaA08g11400D* was detected by chrA08_10449680 and linked with the trait for PHr and located 106 kb upstream of chrA08_10449680 (Supplementary Data Table S12). *BnaA08g11400D*

was homologous to *Arabidopsis* root hair defective 6-like 2 (*RSL2*). *RSL2* encodes a basic Helix- Loop-Helix transcription factors that promotes root hair initiation, growth, and elongation (Han *et al*., 2017; Yi *et al*.,2010; Mangano *et al*., 2018). In addition, mutants lacking RSL2 have disrupted root hair formation under LP (Lan *et al*., 2012; Bhosale *et al*., 2018). *BnaA08g11400D* may effect PH by regulating root hair development at low P or have pleiotropic effects on shoot growth and development in *B. napus.*

CONCLUSIONS

Taken together, the PH and BN of a panel of 403 *B. napus* accessions collected worldwide were investigated under LP and SP, with 2127 significant SNPs associated with these traits identified by GWAS. Thirty-three SNPs co-localized with previous QTLs, including four under LP and 22 under SP. In addition, candidate gene association and haplotype analysis of *BnaA02g33340D*, *BnaA10g09290D* and *BnaC08g26640D* revealed several P tolerant haplotypes. These significant SNP loci, favorable alleles and haplotypes, and the accessions carrying these desired alleles and haplotypes will be useful for breeding low P tolerance *B. napus* cultivars.

FIGURE LEGEDGS

Figure 1. The co-localized locus and haplotypes on chromosome A10 associated with PH_LP and PHr of *B. napus*. (A) Manhattan plot of co-localized locus for PH_LP and PHr in Trial 2_R2 and Trial 2_BLUP. (B) Significant SNPs associated with PH_LP and PHr on chromosome A10. The big red dots represent the significant SNPs.

(C) Candidate gene association analysis of *BnaA10g09290D* with PH_LP. Association of the three alleles in chrA10_8216680 (D), chrA10_8216711 (E) and chrA10_8216756 (F) with PH_LP, respectively. (G) Two haplotypes of *BnaA10g09290D*. (H) Candidate gene association analysis of *BnaA10g09290D* with PHr. Association of the three alleles in chrA10_8216680 (I), chrA10_8216711 (J) and chrA10_8216756 (K) with PHr, respectively. (L) Two haplotypes of *BnaA10g09290D*. The number of inbred lines harboring the corresponding allele are shown in the bracket at the bottom. PH, plant height; LP, low phosphorus supply; SP, sufficient phosphorus supply; PHr, ratio of PH_LP to PH_SP. R2, replication 2; BLUP, best linear unbiased prediction.

Figure 2. The co-localized locus and haplotypes on chromosome C08 associated with PHr of *B. napus*. (A) Manhattan plot of co-localized locus for PHr in Trial 1_R2, Trial 1_R3, Trial 1_BLUP and Trial 2_BLUP. (B) Significant SNP associated with PHr on chromosome C08. The big red dots represent the significant SNPs. (C) Candidate gene association analysis of *BnaC08g26640D* with PHr. Association of the three alleles in chrC08_27999709 (D), chrC08_27999778 (E) and chrC08_27999846 (F) with PHr, respectively. (G) Two haplotypes of *BnaC08g26640D*. The number of inbred lines harboring the corresponding allele are shown in the bracket at the bottom. PH, plant height; LP, low phosphorus supply; SP, sufficient phosphorus supply; PHr, ratio of PH_LP to PH_SP; R2, replication 2; R3, replication 3; BLUP, best linear unbiased prediction.

Figure 3. Co-localized locus on chromosome A02 for plant height of *B. napus* at LP. (A) QTL detected for plant height in the *Bna*TNDH population linkage analysis (Shi et

al., 2013). (B) Manhattan plot of PH_LP (Trial 2_R1). (C) Peak SNP for PH_LP (Trial 2_R1). (D) Association of the co-localized locus with plant height at LP, respectively. (E) Candidate gene association analysis of *BnaA02g33340D*. (F) Association of the five alleles in chrA02_23899623, chrA02_23899654, chrA02_23899669, chrA02_23899686 and chrA02_23899688 with plant height at LP, respectively. (G) Two haplotypes of *BnaA02g33340D*. The number of inbred lines harboring the corresponding allele are shown in the bracket at the bottom. PH, plant height; LP, low phosphorus supply; R1, replication 1.

1 Table 1 Soil physical and chemical properties in the field trials

Trials (Year)	Replications	pH(1:1 H_2O)	Organic matter (g/kg)	Total N concentration (g/kg)	Available P concentration (mg/kg)	Total P concentration (g/kg)	Total K concentration (g/kg)
Trial 1	1	5.61	7.37	0.91	11.56	0.57	13.2
(2018-	2	5.69	8.04	1.12	11.71	0.53	12.1
2019	3	5.65	8.33	1.30	12.19	0.46	14.6
Trial 2	1	5.70	8.42	1.19	12.51	0.63	11.8
(2019-	2	5.58	7.88	1.33	16.95	0.56	10.5
2020)	3	5.56	7.03	1.06	14.10	0.62	12.7

2

3

Table 2. Mean, maximum (max), minimum (min), range, heritability and coefficient of variation (CV, %) of the plant height (PH), branch number (BN), PHr (ratio of PH_LP to PH_SP) and BNr (ratio of BN_LP to BN_SP) in an association panel of *B. napus* in Trial 1 and Trial 2, at low phosphorus (LP) and sufficient phosphorus (SP) supplies.

Trait	Year	Replications	Mean	Min	Max	Range	CV	h^2 (%)
PH_LP	Trial 1	1	121.89	73.75	165.50	91.75	14.80%	63.26
(cm)		2	132.02	74.95	179.25	104.30	14.31%	
		3	132.15	79.00	179.00	100.00	14.31%	
	Trial 2	1	128.66	57.50	176.00	118.50	18.23%	
		2	126.93	60.50	170.50	110.00	16.42%	
BN_LP	Trial 1	1	2.9	0.0	8.0	8.0	53.76%	61.01
(No./		2	4.2	0.0	9.5	9.5	45.80%	
plant)		3	4.3	0.0	10.5	10.5	46.65%	
	Trial 2	1	4.4	0.0	10.0	10.0	45.98%	
		2	4.0	0.0	11.0	11.0	51.15%	
PH_SP	Trial 1	1	137.43	90.50	194.00	103.50	12.72%	88.51
(cm)		2	151.11	84.00	207.50	123.50	12.25%	
		3	159.85	97.00	214.00	117.00	11.02%	
	Trial 2	1	170.31	97.00	216.00	119.00	12.82%	
		2	160.16	107.00	204.00	97.00	11.67%	
		3	171.78	90.00	235.00	145.00	11.76%	
BN_SP	Trial 1	1	4.2	1.0	9.5	8.5	39.48%	70.12
(No./		2	5.9	1.5	13.0	11.5	31.31%	
plant)		3	6.7	3.0	15.0	12.0	28.75%	
	Trial 2	1	7.2	1.0	15.0	14.0	26.85%	
		2	6.7	3.0	14.0	11.0	26.69%	
		3	7.0	2.5	15.5	13.0	26.77%	
PHr	Trial 1	1	0.84	0.45	1.00	0.54	12.22%	69.64
		2	0.83	0.52	1.00	0.47	12.93%	
		3	0.81	0.50	1.00	0.50	13.98%	
	Trial 2	1	0.75	0.40	1.00	0.60	16.17%	
		2	0.73	0.40	0.97	0.57	15.93%	
BNr	Trial 1	1	0.61	0.10	1.00	0.90	40.97%	61.33
		2	0.64	0.10	1.00	0.90	37.03%	
		3	0.63	0.13	1.00	0.88	34.79%	
	Trial 2	1	0.59	0.11	1.00	0.89	35.81%	
		2	0.59	0.11	1.00	0.89	37.74%	

Table 3 List of synonymous and non-synonymous SNP variants identified in the candidate genes of *BnaA02g33090D*, *BnaA10g09290D* and *BnaC08g26640D*

Gene	SNP	Major allele	Minor allele	SNP location	SNP types	Amino acid changes
BnaA02g33090D	chrA02_23899623	G	A	Intron	-	-
	chrA02_23899654	G	T	Intron	-	-
	chrA02_23899669	T	A	Intron	-	-
	chrA02_23899686	C	T	Exon	Synonymous	-
	chrA02_23899688	A	C	Exon	Non-synonymous	histidine to proline
BnaA10g09290D	chrA10_8216680	T	A	Exon	Non-synonymous	isoleucine to asparagine
	chrA10_8216711	C	T	Intron	-	
	chrA10_8216756	A	T	Exon	Synonymous	
BnaC08g26640D	chrC08_27999709	A	T	Intron	-	-
	chrC08_27999778	A	T	Exon	Non-synonymous	isoleucine to asparagine
	chrC08_27999846	T	C	Exon	Synonymous	-

Note: SNP, Single-nucleotide polymorphism.

SUPPLEMENTARY DATA

Figure S1: Shoot and root growth of an association panel of *B. napus* in the field at LP and SP. (A) an association panel of *B. napus* at the seedling stage (left, SP; right, LP); (B) L452 at the seedling stage (left, SP; right, LP); (C) P deficient symptoms of *B. napus* at the seedling stage; (D) P deficient symptoms of *B. napus* at the flowering stage; (E) P deficient symptoms of *B. napus* at the silique stage. (F) Difference in the plant architecture in the selected lines in the association panel of *B. napus* at LP at maturity stage; (G) L528 (left, LP; right, SP); (H) L768 (left, LP; right, SP); (I) L236 (left, LP; right, SP). LP, low phosphorus supply; SP, sufficient phosphorus supply.

Figure S2: Frequency distribution of plant height and branch number of an association panel of *B. napus* in Trial 1and Trial 2 at low and sufficient phosphorus supplies. (A) Plant height in Trial 1. (B) Plant height in Trial 2. (C) Branch number in Trial 1. (D) Branch number in Trial 2. PH, Plant height; BN, Branch number; PHr, ratio of PH_LP to PH_SP; BNr, ratio of BN_LP to BN_SP; LP, low phosphorus supply; SP, sufficient phosphorus supply.

Figure S3: Correlation coefficients of PH and BN of *B. napus* between Trial 1 and Trial 2 at LP and SP. (A) PH. (B) BN. PH, plant height; BN, branch number; LP, low phosphorus supply; SP, sufficient phosphorus supply.

Figure S4: The LD decay of an association panel of *B. napus*

Figure S5: Population structure of an association panel of *B. napus* with K from 2 to 8. (A) The x axis represented the different accessions, the y axis quantified cluster membership, and each accession shown as a vertical line partitioned into K colored

components represented inferred membership in K genetic clusters. (B) The K value estimated for population structure analysis. There was a minimum K-value when K = 5. CV value, cross validation value.

Figure S6: The kinship of an association panel of 403 *B. napus* accessions. (A) A heatmap of the kinship value. (B) The distribution of pairwise relative kinship

Figure S7: Genome-wide association study of PH and BN in an association panel of *B. napus* at low and sufficient phosphorus supplies. (A) Manhattan and QQ plot of the GLM model (Q) for PH in Trial 1. (B) Manhattan and QQ plot of the MLM model (Q+K) for PH in Trial 1. (C) Manhattan and QQ plot of the GLM model (Q) for PH in Trial 2. (D) Manhattan and QQ plot of the MLM (Q+K) model for PH in Trial 2. (E) Manhattan and QQ plot of the GLM model (Q) for BN in Trial 1. (F) Manhattan and QQ plot of the MLM model (Q+K) for BN in Trial 1. (G) Manhattan and QQ plot of the GLM model (Q) for BN inTrial 2. (H) Manhattan and QQ plot of the MLM model (Q+K) for BN in Trial 2. PH, Plant height; BN, Branch number; PHr, ratio of PH_LP to PH_SP; BNr, ratio of BN_LP to BN_SP; LP, low phosphorus supply; SP, sufficient phosphorus supply.

Figure S8: GO enrichment analysis of candidate genes. (A) GO enrichment analysis of candidate genes within 300 kb up and down the lead SNP of chrA10__8216711 on A10 chromosome for PH_LP. (B) GO enrichment analysis of candidate genes within 300 kb up and down the lead SNP of chrC08__27999846 on C08 chromosome for PHr. (C) GO enrichment analysis of candidate genes within 300 kb up and down the lead SNP of chrA01__13846343 on A01 chromosome for BN_LP. (D) GO enrichment

analysis of candidate genes within 300 kb up and down the lead SNP of chrA03__20898013 on A03 chromosome for BN_LP. (E) GO enrichment analysis of candidate genes within 300 kb up and down the lead SNP of chrC07__521008 on C07 chromosome for BN_LP. (F) GO enrichment analysis of candidate genes located in the QTL linkage disequilibrium intervals on chromosome A02 for PH_LP. Blue, green, and red bars indicate molecular function, cellular component and biological process, respectively. PH, Plant height; BN, Branch number; PHr, ratio of PH_LP to PH_SP; LP, low phosphorus supply; SP, sufficient phosphorus supply.

Figure S9: Association of alleles for PH_SP. (A) chrA02_23692807, (B) chrA02_23713660, (C) chrA02_23899688, (D) chrA02_23912345, (E) chrA02_23899623, (F) chrA02_23899654, (G) chrA02_23899669, (H) chrA02_23899686. The number of inbred lines harboring the corresponding allele are shown in the bracket at the bottom. SP, a sufficient phosphorus supply.

FUNDING

This work was supported by the National Nature Science Foundation of China (Grant No. 31972498). We also acknowledge the Applied Basic Research Fronts Program of Wuhan city (Grant No. 2018020401011302), the National Key R & D Program of China (Grant No. 2017YFD0200200) and the Natural and Fundamental Research Funds for the Central Universities of China (Grant No. 2662019PY013).

LITERATURE CITED

Barrett JC, Fry B, Maller J, *et al.* 2005. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**: 263-265.

Bhosale R, Giri J, Pandey BK, *et al.* 2018. A mechanistic framework for auxin dependent *Arabidopsis* root hair elongation to low external phosphate. *Nature Communications* **9**: 1409.

Bradbury PJ, Zhang Z, Kroon DE, *et al*. 2007. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* **23**: 2633-2635.

Chen B, Xu K, Li J, *et al*. 2014. Evaluation of yield and agronomic traits and their genetic variation in 488 global collections of *Brassica napus* L. *Genetic Resources and Crop Evolution* **61**: 979-999.

Che J, Yamaji N, Miyaji T, *et al*. 2020. Node-localized transporters of phosphorus essential for seed development in rice. *Plant cell physiology* **61**: 1387-1398.

Chen ZH, Jenkins I, Nimmo HG, *et al*. 2008. Identification of an F-box protein that negatively regulates Pi starvation responses. *Plant cell physiology* **49**: 1902-1906.

Crenrn JM, Srrupsow GM. 1978. Influence of irrigation and sebding rates on yield and yield components of *Brassica napus*. *Canadian Journal of Plant Science* **58**: 731-737.

Dai X, Wang Y, Zhang WH, *et al*. 2016. OsWRKY74, a WRKY transcription factor, modulates tolerance to phosphate starvation in rice. *Journal of experimental botany* **67**: 947-960.

Deng S, Lu L, Li JL, *et al*. 2020. *Purple acid phosphatase 10c* encodes a major acid phosphatase that regulates plant growth under phosphate-deficient conditions in rice. *Journal of experimental botany* **71**: 4321-4332.

Ding GD, Zhao ZK, Liao Y, *et al*. 2012. Quantitative trait loci for seed yield and yield-related traits, and their responses to reduced phosphorus supply in *Brassica napus*. *Annals of Botany* **109**: 747-759.

Duan K, Yi KK, Dang L, *et al*. 2008. Characterization of a sub-family of Arabidopsis genes with the SPX domain reveals their diverse functions in plant tolerance to phosphorus starvation. *The Plant Journal* **54**: 965-975.

Han HM, Wang HF, Han Y, *et al*. 2017. Altered expression of the TaRSL2 gene contributed to variation in root hair length during allopolyploid wheat evolution. *Planta* **246**: 1019-1028.

Han K, Lee HY, Ro NY, *et al*. 2018. QTL mapping and GWAS reveal candidate genes controlling capsaicinoid content in Capsicum. *Plant biotechnology journal* **16**: 1548-1558.

Hamburger D, Rezzonico E, MacDonald CP, *et al*. 2002. Identification and characterization of the *Arabidopsis* PHO1 gene involved in phosphate loading to the xylem. Plant Cell **14**: 889-902.

He QJ, Lu H, Guo HX, *et al*. 2020. OsbHLH6 interacts with OsSPX4 and regulates the phosphate starvation response in rice. *The Plant Journal* **105**: 649-667.

He YJ, Wu DM, Wei DY, *et al*. 2017. GWAS, QTL mapping and gene expression analyses in *Brassica napus* reveal genetic control of branching morphogenesis.

Scientific Reports **7**: 15971.

Huang KL, Wang H, Wei YL, *et al*. 2019. The high-affinity transporter *BnPHT1;4* is involved in phosphorus acquisition and mobilization for facilitating seed germination and early seedling growth of *Brassica napus*. *BMC Plant Biology* **19**: 156.

Holford ICR. 1997. Soil phosphorus: its measurement and its uptake by plants. *Australian Journal of Soil Research* **35**: 227 -39.

Korte A, Farlow A. 2013. The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods* **9**: 29.

Jain A, Vasconcelos MJ, Raghothama K, *et al*. 2007. Molecular mechanisms of plant adaptation to phosphate deficiency. *Plant breeding reviews* 29: 359 -419.

Kumar SA, Muhlroth A, Jouhet J, *et al*. 2020. The Myb-like transcription factor phosphorus starvation response (PtPSR) controls conditional P acquisition and remodelling in marine microalgae. *New phytologist* **225**: 2380-2395.

Kochian LV. 2012. Plant nutrition: Rooting for more phosphorus. *Nature* **488**: 7412, 466-467.

Lan P, Li W, Wen TN, *et al*. 2012. Quantitative phosphoproteome profiling of iron-deficient Arabidopsis roots. *Plant physiology* **159**: 403-417.

Laperche A, Aigu Y, Jubault M, et al. 2017. Clubroot resistance QTL are modulated by nitrogen input in *Brassica napus*. *Theor Appl Genet* **130**: 669-684.

Lee YS, Huang KX, Florante AQ, *et al*. 2008. Molecular basis of cyclin-CDK-CKI regulation by reversible binding of an inositol pyrophosphate. *Nature Chemical*

Biology **4**: 25 -32.

Li S, Zhu Y, Varshney RK, *et al*. 2020. A systematic dissection of the mechanisms underlying the natural variation of silique number in rapeseed (*Brassica napus* L.) germplasm. *Plant biotechnology journal* **18**: 568-580.

Li TG, Ma XF, Li NY, *et al*. 2017. Genome-wide association study discovered candidate genes of Verticillium wilt resistance in upland cotton (*Gossypium hirsutum* L.). *Plant biotechnology journal* **15**: 1520-1532.

Li C, Liu XY, Ruan H, *et al*. 2019. *GmWRKY45* Enhances Tolerance to Phosphate Starvation and Salt Stress, and Changes Fertility in Transgenic *Arabidopsis*. *Frontiers in Plant Science* **10**: 1714.

Li F, Chen B, Xu K, *et al*. 2016. A genome-wide association study of plant height and primary branch number in rapeseed (*Brassica napus*). *Plant science* **242**: 169-177.

Lu JW, Chen F, Zhang ZQ, *et al*. 2005. Effect of phosphor application rate on rapeseed yield, nutrient absorption and profit. Chinese journal of oil crop sciences 27 (in Chinese with English abstract).

Luo X, Ding Y, Zhang LZ, *et al*. 2017. Genomic prediction of genotypic effects with epistasis and environment interactions for yield-related traits of rapeseed (*Brassica napus* L.). *Front Genet* **8**: 15.

Lin KF, Tsai MY, Lu CA, *et al*. 2018. The roles of *Arabidopsis* HSFA2, HSFA4a, and HSFA7a in the heat shock response and cytosolic protein response. *Botanical studies* **59**: 15.

Liu Y, Wang L, Deng M, *et al*. 2015. Genome-wide association study of phosphorus-deficiency-tolerance traits in *Aegilops tauschii*. *Theoretical and Applied Genetics* **128**: 2203-2212.

Luo B, Ma P, Nie Z, *et al*. 2019. Metabolite profiling and genome-wide association studies reveal response mechanisms of phosphorus deficiency in maize seedling. *The Plant Journal* **97**: 947-969.

Mangano S, Denita-Juarez SP, Marzol E, *et al*. 2018. High auxin and high phosphate impact on RSL2 expression and ROS-homeostasis linked to root hair growth in *Arabidopsis thaliana*. *Frontiers in Plant Science* **9**: 1164.

Murakami H, Kakutani N, Kuroyanagi Y, *et al* 2020. MYB-like transcription factor NoPSR1 is crucial for membrane lipid remodeling under phosphate starvation in the oleaginous microalga *Nannochloropsis oceanica*. *Febs Letters* **594**: 3384-3394.

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

Olczak M, Morawiecka B, Watorek W. 2003. Plant purple acid phosphatases-genes, structures and biological function. *Acta Biochimica Polonica* **50**: 1245-1256.

Pacak A, Barciszewska PM, Swida BA, *et al*. 2016. Heat stress affects Pi-related genes expression and inorganic phosphate deposition/accumulation in barley. *Frontiers in Plant Science* **7**: 926.

Pu Z, Pei Y, Yang J, *et al*. 2018. A QTL located on chromosome 3D enhances the selenium concentration of wheat grain by improving phytoavailability and root

structure. *Plant and Soil* **425**: 287-296.

Qiu D, Morgan C, Shi J, *et al* .2006. A comparative linkage map of oilseed rape and its use for QTL analysis of seed oil and erucic acid content. *Theoretical and Applied Genetics* **114**: 67-80.

Rose TJ, Wissuwa M. 2012. Rethinking internal phosphorus utilization efficiency: a new approach is needed to improve PUE in grain crops. *In Advances in Agronomy Volume* **116**: 185-217.

Shi T, Li R, Zhao Z, *et al*. 2013. QTL for yield traits and their association with functional genes in response to phosphorus deficiency in *Brassica napus*. *Plos One* **8**: e54559.

Si LZ, Chen JY, Huang XH, *et al* .2016. *OsSPL13* controls grain size in cultivated rice. *Nature genetics* **48**: 447-456.

Su JY, Xiao YM, Li M, *et al*. 2006. Mapping QTLs for phosphorus-deficiency tolerance at wheat seedling stage. *Plant and Soil* **281**: 25-36.

Sun CM, Wang BQ, Yan L, *et al*. 2016. Genome-wide association study provides iInsight into the genetic control of plant height in rapeseed (*Brassica napus* L.). *Frontiers in Plant Science* 7:1102.

Tang S, Zhao H, Lu SP, *et al*. 2020. Genome- and transcriptome-wide association studies provide insights into the genetic basis of natural variation of seed oil content in *Brassica napus*. *Molecular plant* **1**: 470-484.

Vance CP, Uhde-Stone C, Allan DL. 2003. Phosphorus acquisition and use: critical adaptations by plants for securing a nonrenewable resource. *New phytologist*

157: 423-447.

Wang Xh, Chen Yl, Thomas CL, *et al*. 2017. Genetic variants associated with the root system architecture of oilseed rape (*Brassica napus* L.) under contrasting phosphate supply. *DNA Research* **24**: 407-417.

Wang Y, Ribot C, Rezzonico E, *et al*. 2004. Structure and expression profile of the *Arabidopsis* PHO1 gene family indicates a broad role in inorganic phosphate homeostasis. *Plant Physiology* **135**: 400-411.

Wissuwa M, Kondo K, Fukuda T, *et al*. 2015. Unmasking novel loci for internal phosphorus utilization efficiency in rice germplasm through genome-wide association analysis. *Plos One* **10**: e0124215.

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. *Theoretical and Applied Genetics* **105**: 890-897.

Wu P, Xu J. 2010. Does OsPHR2, central Pi-signaling regulator, regulate some unknown factors crucial for plant growth? *Plant Signal Behavior* **5**: 712-714.

Xiao YJ, Liu HJ, Wu LJ, *et al*. 2017. Genome-wide association studies in maize: Praise and Stargaze. *Molecular plant* **10**: 359-374.

Xu C, Zhang Hw, Sun JH, *et al*. 2018. Genome-wide association study dissects yield components associated with low-phosphorus stress tolerance in maize. *Theoretical and Applied Genetics* **131**, 1699-1714.

Xu LP, Hu KN, Zhang ZQ, *et al*. 2016. Genome-wide association study reveals the genetic architecture of flowering time in rapeseed (*Brassica napus* L.). *DNA*

Research **23**, 43-52.

Yang M, Ding GD, Shi L, *et al*. 2010a. Quantitative trait loci for root morphology in response to low phosphorus stress in *Brassica napus*. *Theoretical and Applied Genetics* **121**: 181-193.

Yang SY, Huang TK, Kuo HF, *et al*. 2017. Role of vacuoles in phosphorus storage and remobilization. *Journal of experimental botany* **68**: 3045-3055.

Yang M, Ding GD, Shi L, *et al*. 2010b. Detection of QTL for phosphorus efficiency at vegetative stage in *Brassica napus*. *Plant and Soil* **339**: 97-111.

Yi K, Menand B, Bell E, *et al* 2010. A basic helix-loop-helix transcription factor controls cell growth and size in root hairs. *Nature genetics* **42**: 264-267.

Zhang C, Dong SS, Xu JY, *et al*. 2019. PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics* **35**: 1786-1788.

Zhang D, Song H, Cheng H, *et al*. 2014. The acid phosphatase-encoding gene *GmACP1* contributes to soybean tolerance to low-phosphorus stress. *Plos genetics* **10**: e1004061.

Zhang Y, Thomas CL, Xiang JX, *et al*. 2016. QTL meta-analysis of root traits in *Brassica napus* under contrasting phosphorus supply in two growth systems. *Scientific Reports* **6**: 33113.

Zheng M, Peng C, Liu H, et al. 2017. Genome-wide association study reveals candidate genes for control of plant height, branch initiation height and branch number in rapeseed (*Brassica napus* L.). *Frontiers in Plant Science* **8**: 1246.

Zheng M, Zhang L, Tang M, *et al*. 2020. Knockout of two *BnaMAX1* homologs by CRISPR/Cas9-targeted mutagenesis improves plant architecture and increases yield in rapeseed (*Brassica napus* L.). *Plant biotechnology journal* **18**: 644-654.

Zhang Y, Wan J, He L, *et al*. 2019. Genome-wide association analysis of plant height using the maize F1 Population. *Plants (Basel)* **8**: 432.

Zou J, Lu JW, Chen F, *et al*. 2009. Study on abundance and deficiency indices of soil available P, K and B for winter rapeseed in Yangtze River Valley based on ASI method. *Scientia Agricultura Sinica*. **42**: 2028–2033. (in Chinese with English abstract)

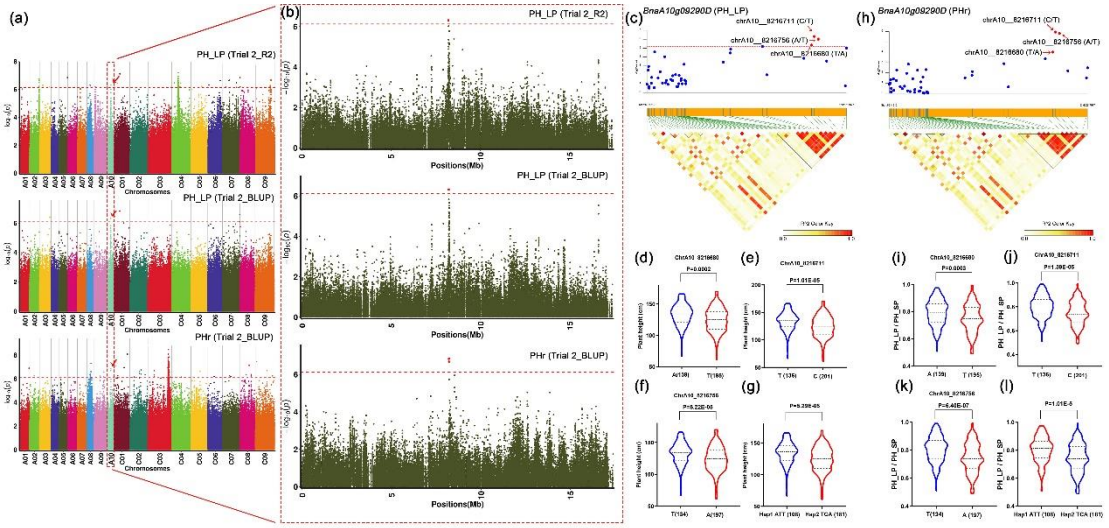

Fig. 1. The co-localized locus and haplotypes on chromosome A10 associated with PH_LP and PHr of *B. napus*. (A) Manhattan plot of co-localized locus for PH_LP and PHr in Trial 2_R2 and Trial 2_BLUP. (B) Significant SNPs associated with PH_LP and PHr on chromosome A10. The big red dots represent the significant SNPs. (C) Candidate gene association analysis of *BnaA10g09290D* with PH_LP. Association of the three alleles in chrA10_8216680 (D), chrA10_8216711 (E) and chrA10_8216756 (F) with PH_LP, respectively. (G) Two haplotypes of *BnaA10g09290D*. (H) Candidate gene association analysis of *BnaA10g09290D* with PHr. Association of the three alleles in chrA10_8216680 (I), chrA10_8216711 (J) and chrA10_8216756 (K) with PHr, respectively. (L) Two haplotypes of *BnaA10g09290D*. The number of inbred lines harboring the corresponding allele are shown in the bracket at the bottom. PH, plant height; LP, low phosphorus supply; SP, sufficient phosphorus supply; PHr, ratio of PH_LP to PH_SP. R2, replication 2; BLUP, best linear unbiased prediction.

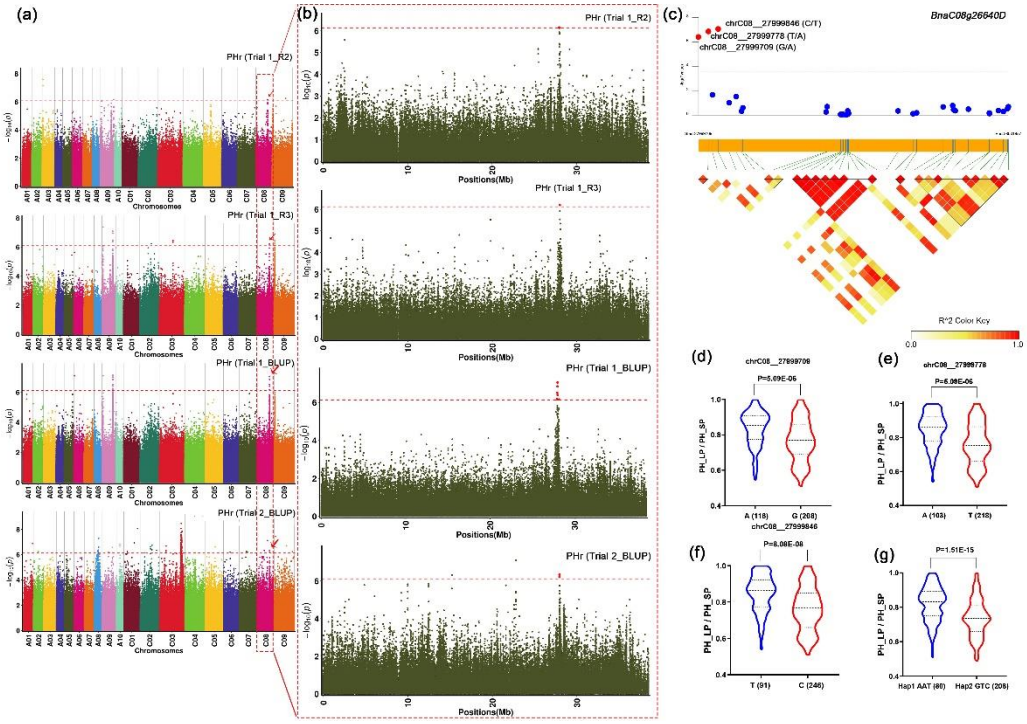

Fig. 2. The co-localized locus and haplotypes on chromosome C08 associated with PHr of *B. napus*. (A) Manhattan plot of co-localized locus for PHr in Trial 1_R2, Trial 1_R3, Trial 1_BLUP and Trial 2_BLUP. (B) Significant SNP associated with PHr on chromosome C08. The big red dots represent the significant SNPs. (C) Candidate gene association analysis of *BnaC08g26640D* with PHr. Association of the three alleles in chrC08_27999709 (D), chrC08_27999778 (E) and chrC08_27999846 (F) with PHr, respectively. (G) Two haplotypes of *BnaC08g26640D*. The number of inbred lines harboring the corresponding allele are shown in the bracket at the bottom. PH, plant height; LP, low phosphorus supply; SP, sufficient phosphorus supply; PHr, ratio of PH_LP to PH_SP; R2, replication 2; R3, replication 3; BLUP, best linear unbiased prediction.

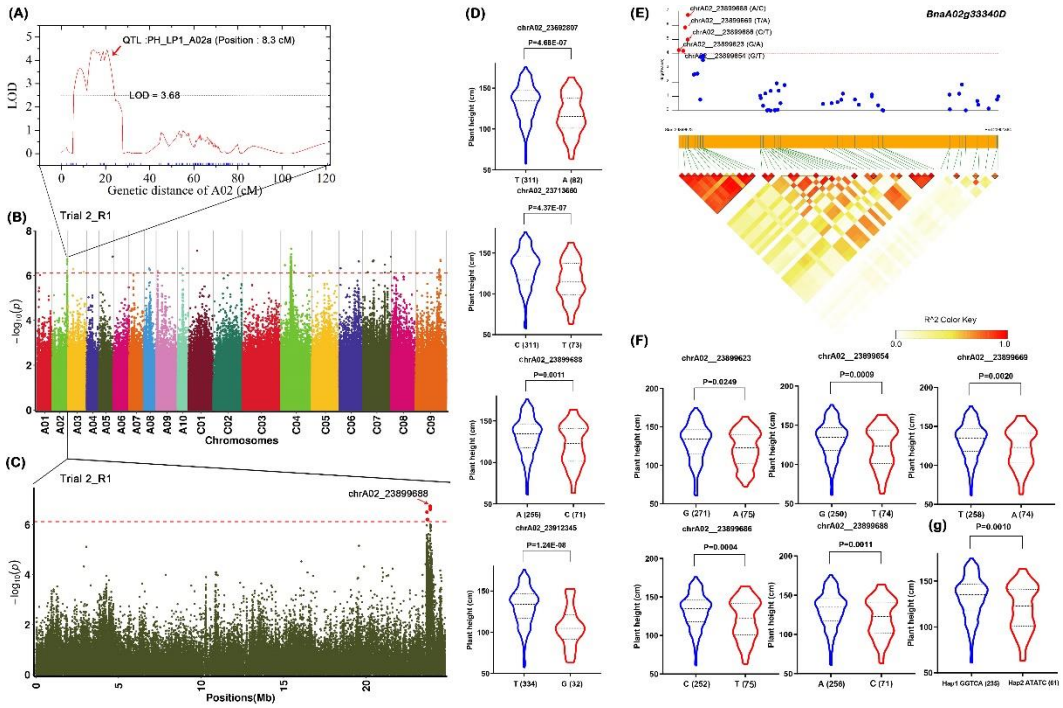

Fig. 3 Co-localized locus on chromosome A02 for plant height of *B. napus* at LP. (A) QTL detected for plant height in the *Bna*TNDH population linkage analysis (Shi *et al*., 2013). (B) Manhattan plot of PH_LP (Trial 2_R1). (C) Peak SNP for PH_LP (Trial 2_R1). (D) Association of the co-localized locus with plant height at LP, respectively. (E) Candidate gene association analysis of *BnaA02g33340D*. (F) Association of the five alleles in chrA02_23899623, chrA02_23899654, chrA02_23899669, chrA02_23899686 and chrA02_23899688 with plant height at LP, respectively. (G) Two haplotypes of *BnaA02g33340D*. The number of inbred lines harboring the corresponding allele are shown in the bracket at the bottom. PH, plant height; LP, low phosphorus supply; R1, replication 1.