Genome-wide association study reveals candidate genes controlling root system architecture under low phosphorus supply at seedling stage in Brassica napus

Article

Accepted Version

Genome-wide association study reveals candidate genes controlling root system architecture under low phosphorus supply at seedling stage in *Brassica napus*

Running title: RSA of *B. napus* at LP

Pan Yuan[1,2], Haijiang Liu[1,2], Xiaohua Wang[3], John P. Hammond[4], 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] *College of Agriculture and Forestry Science, Linyi University, Linyi 276000, China*

[4] *School of Agriculture, Policy and Development, University of Reading, Reading RG6 6AR, UK*

* **Correspondence:** Lei shi (leish@mail.hzau.edu.cn), Tel: +86 027-87286871.

Keywords: Oilseed rape (*B. napus* L.), root system architecture (RSA), low phosphorus supply, genome-wide association study (GWAS), haplotype analysis.

Abstract

Optimal root system architecture (RSA) is essential for vigorous growth and yield in crops. Plants have evolved adaptive mechanisms in response to low phosphorus (LP) stress, and one of those is changes in RSA. Here, more than five million single-nucleotide polymorphisms (SNPs) obtained from whole genome re-sequencing data (WGR) of an association panel of 370 oilseed rape (*Brassica napus* L.) were used to conduct genome-wide association study (GWAS) of RSA traits of the panel at LP in 'pouch and wick' system. Fifty-two SNPs were forcefully associated with lateral root length (LRL), total root length (TRL), lateral root density (LRD), root number (RN), mean lateral root length (MLRL) and root dry weight (RDW) at LP. There were

significant correlations between phenotypic variation and the number of favorable alleles of the associated loci on chromosomes A06 (chrA06_20030601), C03 (chrC03_3535483) and C07 (chrC07_42348561), respectively. Three candidate genes (*BnaA06g29270D*, *BnaC03g07130D* and *BnaC07g43230D*) were detected by combining transcriptome, candidate gene association analysis and haplotype analysis. Cultivar carrying "CCGC" at *BnaA06g29270DHap1*, "CAAT" at *BnaC03g07130DHap1* and "ATC" at *BnaC07g43230DHap1* had greater LRL, LRN and RDW than lines carrying other haplotypes at LP supply. The RSA of cultivar harboring the three favorable haplotypes was further confirmed by solution culture experiments. These findings define exquisite insights into genetic architectures underlying *B. napus* RSA at LP, and provide valuable gene resources for root breeding.

Introduction

Oilseed rape (*Brassica napus* L.; *B. napus*) is a globally important oil crop (Chalhoub et al. 2014). *B. napus* has high phosphorus (P) demand and P deficiency affects root development and reduces seed yield (Wang et al. 2017; Liu et al. 2021a). As populations expand around the world, the demand for edible oil is also growing. However, P fertilizers are derived from non-renewable P rock resources and over application of P fertilizers are associated with environmental issues, such as eutrophication (Vance et al. 2003; Liu et al. 2015). Therefore, detecting the genetic mechanisms for tolerance to low P (LP) availability and breeding P-efficient *B. napus* cultivars to reduce P use in agricultural systems is crucial.

Root system architecture (RSA) is the spatial configuration of different types and ages of roots emerging from a single plant and is important for plant survival (Lynch et al. 1995). Traditional linkage analysis has identified abundant quantitative trait loci (QTLs) for root traits in *B. napus* in response to low P availability. A *B. napus* DH population (*Bna*TNDH) has been constructed from a cross between a P-efficient cultivar, 'Ningyou7', and a P-inefficient cultivar, 'Tapidor' (Shi et al. 2013). Thirty QTLs were detected for primary root length (PRL), total root length (TRL), root number (RN) and lateral root density (LRD) in an agar-based growth system with P-deficient

and P-sufficient supply in *Bna*TNDH population of *B. napus* (Shi et al. 2013). Among them, QTLs clustered on A03 chromosome associated with LRN and RDW at a low P supply. Altogether, 131 QTLs were detected for RSA traits using a high-density single-nucleotide polymorphism (SNP)-based genetic linkage map (Zhang et al. 2016). The intervals of QTLs for root and shoot biomass overlapped on chromosome A03, and QTLs for root biomass and lateral root emergence on chromosome A04 and C04, C08 and C09 were co-located. A major QTL explained 18.0% of the phenotypic variation for lateral root density on chromosome C09 (Zhang et al. 2016).

Compared with traditional linkage analysis, GWAS (genome-wide association study) does not require construction of mapping populations, which can save time (Xiao et al. 2017). For *B. napus*, a 60K SNP chip has been developed for genotyping line in GWAS studies and was used GWAS analysis of RSA at LP. Two hundred and eighty-five SNPs were significantly associated with RSA in response to LP by GWAS study involving 405 accessions (Wang et al. 2017). Nine SNPs linked with RSA were consistent with the published QTL mapping of RSA in the *Bna*TNDH population (Wang et al. 2017; Zhang et al. 2016). Eleven SNPs were identified by a combined GWAS and QTL mapping approach, with the SNPs located on A06, A08 and C01 co-located with QTLs associated with root angle at LP (Duan et al. 2021).

GWAS of RSA of *B. napus* under nutrient conditions have also been reported. Thirty-one significant SNPs were detected for root traits by shovelomics approach in the field using GWAS analysis with 216 *B. napus* accessions (Arifuzzaman et al. 2019). Two association mapping panels were used to perform GWAS on root related traits in *B. napus*, and 27 significant SNPs were identified (He et al. 2019). Thirty-two SNPs are identified as significantly associated with root related traits at five vegetative stages by a GWAS analysis with 280 *B. napus* accessions, with *BnaA03g52990D*, *BnaA06g37280D,* and *BnaA09g07580D* identified as candidate genes (Li et al. 2021).

Compared with SNP chip genotyping, whole genome re-sequencing (WGR) provides higher accuracy in mapping the location of recombination events and is capable of detecting more genetic variants. Recently, WGR technology has been used to analyze the genetic structure of agronomic traits of *B. napus* including seed oil content (Tang et

al. 2021), glucosinolate content (Tan et al. 2022), phytate (Liu et al. 2021a) and low-temperature tolerance (Luo et al. 2021).

In this study, the genetic architecture of RSA at LP in the seeding stage was investigated by GWAS using a panel of 370 *B. napus* accessions and more than 5 million SNPs. Here we report the (1) significant SNPs and candidate genes linked with RSA traits at LP and (2) identify the favourable haplotypes for breeding cultivars with optimal RSA in *B. napus*.

Materials and methods

Plant materials

The *B. napus* association panel contained 370 cultivars and inbred lines, comprising 327 semi-winter, 37 spring, 4 winter and 2 vegetable types, converged from major *B. napus* breeding centers across China (Table S1).

Phenotypic investigation

The phenotype data for RSA of the association panel from the 'pouch and wick' system of *B. napus* at low P (0 mM Pi) published by Wang et al. (2007) were used in this study. For individual genotypes, at least 16 replicates were sown. Fourteen days after sowing on free-P paper, seedling's root system was detected using a digital single-lens reflex camera (Canon EOS 1100D, Canon Inc., Tokyo, Japan) with a marker at the base. Root images were analyzed using RootReader2D (RR2D) and automatically calculates PRL (primary root length), LRL (lateral root length) and LRN (lateral root number) (Clark et al. 2013). TRL (total root length) = PRL + LRL, LRD (lateral root density) = LRN / PRL and MLRL (mean lateral root length) = LRL/LRN as described by Wang et al. (2017). Root samples were oven-dried at 80°C for two days and dry biomass were collected. The mean, maximum, range, skewness and coefficient of variation were calculated using the psych package in R software (https://cran.r-project.org/web/packages/psych/index.html). The correlation coefficients between phenotypes were calculated by R language.

Cultivar L133 with three favourable haplotypes and cultivar L154 without these

haplotypes were used to investigate the difference in RSA between them. Seedlings were grown at sufficient P (SP: 250 μmol L^{-1} P) and low P (LP: 5 μmol L^{-1} P) in hydroponics in growth chamber (16 h light (7:00 am-23:00 pm)/ 8 h dark regime at 22°C) (Wang et al. 2017). Twenty-four-day-old seedlings were harvested after the measurement of net photosynthetic rate (Pn). Pn of the youngest fully expanded leaf was measured between 10 am to 12 am using a portable photosynthesis system (Li6400; LI-COR, Lincoln, NE, USA). The root was scanned with a modified flatbed scanner (Epson V700, Nagano-ken, Japan) at 400 dpi. Total root length (TRL, cm), root volume (RV, cm^3), root area (RA, cm^2) and lateral root number (LRN, N) were determined with the WinRHIZO program (Regent Instruments Inc., Quebec, Canada). Leaves were scanned with the scanner and the whole area of leaves was measured using the Image J software (http://rsbweb.nih.gov/ij/download.html). Shoot and root samples were dried at 80°C for two days for dry biomass measurements.

Genome-wide association study and candidate gene identification

High-quality SNP markers (more than 10 million) across the *B. napus* association panel, comprising 403 distinct genotypes, cultivars and inbred lines were obtained previously (Tang et al. 2021). The SNPs filtered with minor-allele frequency (MAF) (>0.05) and missing rate (<0.2), 1.60 million SNPs were obtained for GWAS in this study. We used the factored spectrally transformed linear mixed models (Fast-LMM) (https://www.softpedia.com/get/Science-CAD/FaST-LMM-Set.shtml). The threshold for screening significant SNPs in this association panel is set to 6.25×10^{-07} (Liu et al. 2021a; Liu et al. 2021b). Use GGplot2 package (https://cran.r-project.org/web/packages/ggplot2/index.html) in R software to draw Manhattan plot and CMplot package (https://cran.r-project.org/web/packages/CMplot/) to draw Quantile-Quantile plot. Population structure and kinship was calculated by Admixture software and Tassel 5.0 software, respectively (Bradbury et al. 2007; Alexander et al. 2009). Analysis of molecular variance (AMOVA) was performed with the arlequin software (3.5.2.2) with 1000 permutations to statistically examine differences between populations (Excoffier et al. 2010). PopLDdecay software was used to calculated LD

decay value (Zhang et al. 2019). Genes located within 173 kb upstream and downstream of the peak SNPs were viewed as candidate genes according to the LD decay of the panel (173 kb).

Candidate gene association analysis and haplotype analysis

The coding region and the upstream 2 kb promoter region of candidate genes were extracted using vcftools software (https://vcftools.github.io/index.html, Danecek et al. 2011). Candidate gene association analysis of *BnaA06g29270D*, *BnaC03g07130D* and *BnaC07g43230D* were performed with Tassel 5.0 and LDBlockShow software (Bradbury et al. 2007; Dong et al. 2021). Haploview.4.2 software was used for haplotype analysis (Barrett et al. 2005). Further comparative analysis was conducted based on haplotypes containing at least 10 *B. napus* accessions. To compared the differences in the haplotypes of RSA, student's t-test was applied.

Expression profile of putative candidate genes in P efficient *B. napus* cultivar

RNA was extracted using the EastepR super total RNA extraction kit (Promega, Madison, WI). RNA concentration was quantified using the NanoDropND-1000 spectrophotometer (Thermo Scientific). One μg RNA was convert into cDNA with the ReverTra Ace qPCR RT master mix with gDNA remover (TOYOBO, Osaka, Japan). cDNA samples were diluted at 1:10 in ddH_2O. RT-qPCR was performed in 10 μL reaction volumes with gene specific primers (Table S2), and conducted with a QuantStudio^{TM} 6 Flex System (Applied Biosystems, Foster City, CA) using the SYBR Green Real-Time PCR Master Mix Kit (TOYOBO). The RT-qPCR was performed as following: 95°C for 5 min followed by 40 cycles of 95°C for 10 s, 55°C for 20 s, and 72°C for 20 s, and a subsequent standard dissociation protocol to validate the presence of a unique PCR product. Each sample was analyzed in three technical replicates, and the $2^{-\Delta\Delta CT}$ method was applied to calculate the relative expression from the normalizing of two housekeeping genes (*Tubulin* and *Actin2*). Primer sequences are listed in Table S2.

Transcriptome sequencing

The transcriptome data of P deficient *B. napus* cultivars 'Eyou changjia' at LP and SP conditions have been published (Du et al. 2017). Fifteen-day old seedlings were grown at 250 μM Pi (SP) and P-free nutrient solution (LP, 0 μM P) for 10 days. The roots and leaves were harvested with three biological replicates. Twelve RNA samples were subjected to the Illumina HiSeq 2000 platform (Illumina, USA). Data analysis was performed as described in Du et al. (2017).

Statistical analysis

Data were presented as means ± SEM. Student's *t* test using Excel software (Microsoft, Redmond, WA) and SigmaPlot (SPSS Science Inc., IL) were applied for comparisons between samples. The asterisks indicate significant differences between samples at the *P < 0.05, **P < 0.01, ***P < 0.0001, ****P < 0.0001 level using the *t* test. The phenotypic data of root traits were analyzed by restricted maximum likelihood (REML) programs to evaluate the line means and sources of variation (Shi et al. 2013). Means were approximated using the [Genotype] term as a fixed factor, retaining [(Replicate/Side of paper/Tray/Frame/Room)] as random factors. The sources of variation were estimated using the Random terms and no defined fixed factor. Statistical analysis was performed by GenStat15th (VSN International, Oxford, UK).

Results

Phenotypic variation for root system architecture (RSA) of an association panel of *B. napus* at LP supply

The phenotype data for RSA of 405 genotypes in the association panel of *B. napus* at low P published by Wang et al. (2017) and the re-sequencing of 370 cultivars (Tang et al. 2021) from the 405 were used in this study. Extensive phenotypic variations for TRL, LRD, LRN, MLRL and RDW were detected in the association panel at LP in 'pouch and wick' system (Fig. 1; Table S3). At LP, the LRL varied from 15.49 to 53.41 cm with a coefficient of variation of 24.26% and a mean value of 33.72 cm (Table 1). The LRN varied from 7.5 to 24.8 per plant with a coefficient of variation of 23.27%

and a mean value of 16.2 per plant (Table 1). In addition, the skewness values of traits were close to 0, which indicates that these traits conform to a normal distribution and were suitable for GWAS analysis (Table 1). The correlation coefficient (r) between LRL and TRL, TRL and LRN, TRL and RDW, LRN and RDW were 0.97, 0.85, 0.73 and 0.77, respectively (Fig. S1).

Population structure, relative kinship and LD decay

To identify genes involved in RSA in response to low P availability, we used published genotype data consisting of more than 1.6 million high-quality SNPs for the 370 accessions (Tang et al. 2021) to conduct a GWAS based on the six RSA traits at low P in the 'pouch and wick' system. When r^2 was 0.1, the LD decay in this 370 *B. napus* association panel was 173 kb (Fig. S2). Quality control for the imputed data was conducted using Plink software, where SNPs were removed if they had a minor allele frequency lower than 0.2. The number of markers retained for subsequent analyses was 417,787. Population was divided into four subgroups (Fig. S3). There were significant differences between the K = 4 assigned populations ($P < 0.01$, Table S4). The kinship heatmap shows that most cultivars are distantly related, which showed that this association panel was suitable for GWAS analysis (Fig. S4). A total of 433,432 LD blocks were identified in the whole genome (Table S5), and the majority of LD blocks were less than 10 kb, and only 15 LD blocks were larger than 100 kb (Fig. S5). These data indicated that the genetic distance of the accessions in the association panel was appropriate for GWAS analysis.

Genome-wide association study of RSA related traits of *B. napus* at LP supply

GWAS for all six RSA traits was performed using the Fast-LMM. The quantile-quantile plots indicated that the model of the six RSA could be used to detect association signals (Fig. 2). Fifty-two SNPs were confirmed to be significantly associated with these RSA traits at LP with a $P < 6.25 \times 10^{-07}$ in *B. napus* (Fig. 2; Table S6). Of the 52 SNPs, 18 were located in the LD block interval (Table S7). Among them, 6, 4, 12, 6, 3 and 21 SNPs were significantly associated with LRL, TRL, LRD, LRN, MLRL and

RDW, respectively, and were distributed on 12 chromosomes, explaining between 7.35 and 12.37% of the phenotypic variance explained (PVE) (Fig. 2; Table S6). Chromosome C07 had the largest number of significant SNPs (11) and chromosome A01 had the least number of significant SNPs (1) (Fig. 2; Table S6). Four SNPs on chromosomes C03 and C07 (chrC03_3535476, chrC03_3535483, chrC07_42348561 and chrC07_42348533) were significantly associated with RSA and were consistently associated with two RSA traits (Table S6). SNP chrC03_3535476 was associated with LRL and TRL, and explained 9.82% and 9.09% of the PVE, respectively (Table S6). SNP chrC07_42348561 was associated with LRN and RDW, and explained 8.99% and 9.54% of the PVE, respectively (Table S6). In addition, 42 of the 52 significant SNPs were located in the 18 LD blocks in this study. These results indicated that RSA traits such as LRL were highly correlated with TRL, and LRN was highly correlated with RDW (Fig. S1).

Prediction of candidate genes for RSA at LP supply in *B. napus*

The lead SNP chrC03_3535483 (P=2.37E-07, PVE=10.19%) on C03 was significantly associated with LRL and TRL, and may be a major genetic locus responsible for RSA in *B. napus* at LP (Fig. 2; Table S6). LD decay was 173 kb for this 370 *B. napus* association panel (Fig. S2). Based on the LD decay, 93 candidate genes were explored in the 173 kb up/down-stream of the significant SNP (chrC03_3535483) (Table S8). The transcriptome data of P-efficient *B. napus* cultivars 'Eyou changjia' under LP and SP conditions (Li et al. 2019) were used to determine differential expression of candidate genes identified by GWAS analysis. Transcript abundance of *BnaC03g07130D* was significantly increased at LP in both shoots and roots, and was also evidenced by qRT-PCR (Table 2; Fig. 3A). Lead SNP (chrC07_42348561, P = 1.82E-07, PVE = 9.54 %) on C07 was associated with both LRN and RDW at LP (Fig. 2; Table S6). Based on the LD decay, 76 candidate genes were detected (Table S8). Transcriptome data and qRT-PCR analysis showed that the transcript of *BnaC07g43230D* was induced by P deficiency in shoots and roots (Table 2; Fig. 3B).

In addition, one region significantly associated with LRD on chromosome A06

ranged from 19.93 to 20.03 Mb (Fig. 2; Table S6). Seventy genes were detected underlying the region around lead SNP chrA06_20030601, and two of the candidate genes were significantly expressed by LP (Table 2; Table S8). The expression levels of *BnaA06g29270D* in shoots and roots were significantly decreased at LP supply by qRT-PCR analysis (Table 2; Fig. 3C). Twenty-four SNPs were located within the 2 kb promoter region and the entire coding region of *BnaA06g29270D*. However, no SNPs were identified within the corresponding region of *BnaA06g29530D* (Table 2). Therefore, *BnaA06g29270D* was predicted to be the candidate gene for further study.

Candidate gene association study and haplotype analysis of BnaA06g29270D, BnaC03g07130D and BnaC07g43230D

To further understand the intragenic variation affecting the phenotypic values and identify the favorable alleles and haplotypes, we extracted the SNPs within the entire coding regions and promoters (upstream, 2 kb) of three candidate genes *BnaC03g07130D*, *BnaA06g29270D* and *BnaC07g43230D*, respectively. Thirty-seven SNPs were identified in *BnaC03g07130D*, and 26 in *BnaC07g43230D* and 24 *BnaA06g29270D* (Fig. 4-6).

Ten and four SNPs in *BnaC03g07130D* were significantly associated with LRL and TRL, respectively (Fig. 4A; 3G). Among the ten significant SNPs, four were located in the promoter, five in the exon and one in the intron region (Table 3). The five SNPs identified in the exons resulted in synonymous mutations, which were unlikely to influence the function of *BnaC03g07130D* protein (Table 3). Further, the 'C' allele of chrC03_3396486, 'A' allele of chrC03_3396712, 'A' allele of chrC03_3396812 and 'T' allele of chrC03_3396932 located in the promoter of *BnaC03g07130D* were all identified as strong alleles associated with LRL and TRL (Fig. 4B-E, H-K). Two major haplotypes were detected, among which *BnaC03g07130Hap1* (CAAT) had higher LRL and TRL, while *BnaC03g07130Hap2* ("TGCC") had lower LRL and TRL ($P = 0.0002$) (Fig. 4F, L).

For candidate gene *BnaC07g43230D* three SNPs were associated with the LRN and RDW (Fig. 5A, F), and the 'A' allele of chrC07_42368601, 'T' allele of

chrC07_42368638 and 'C' allele of chrC07_42368650 were the positive alleles for root development at LP (Fig. 5B-D, G-I). Two major haplotypes were identified, with *BnaC07g43230Hap1* ("ATC") having significantly greater LRN and RDW than *BnaC07g43230Hap2* ("TCT") (Fig. 5E, J).

For candidate gene *BnaA06g29270D*, 11 SNPs were significantly associated with the LRD, and had strong LDs with each other (Fig. 6A; Table 3). Among the 11 significant SNPs, 3, 6 and 2 were located in the promoter, exon and intron region, respectively (Table 3). Five SNPs (chrA06_19976766, chrA06_19976785, chrA06_19976887, chrA06_19977025 and chrA06_19977610) were in the exon region, but were synonymous mutations which resulted in no amino acid changes (Table 3). The SNP of chrA06_19977004 (C/A) was located in the exon region of the gene *BnaA06g29270D* and resulted in amino acid change from cysteine to a stop codon, and might contribute to the phenotypic difference in LRD (Table 3). The 'C' allele of chrA06_19975674, 'C' allele of chrA06_19975707, 'G' allele of chrA06_19975749 and 'C' allele of chrA06_19977004 were associated with higher LRD (Fig. 6B-E). Two major haplotypes were identified, with *BnaA06g29270Hap1* ("CCGC") having significantly higher LRD than *BnaA06g29270Hap2* ("TACA") ($P < 0.0001$) (Fig. 6F).

A total of 73 *B. napus* cultivars in this association panel contained these three favorable haplotypes. As expected, RSA (such as, LRL, TRL, LRN and RDW) were significantly higher in the 73 cultivars than in the other 297 cultivars at low P (Fig. 7). To better understand the relative influence of the three favorable haplotypes on the root and shoot growth of *B. napus*, a cultivar harboring the three favorable haplotypes (L133) and a cultivar with none of these haplotypes (L154) with extremely different in root system architecture traits at low P (Figure 7) were chose to dissect root system architecture, biomass, whole leaf area and net photosynthesis at both sufficient P and low P (Fig. 8). TRL and WLA (whole leaf area) were higher in L133 cultivar than in L154 cultivar at SP (Fig. 8A, B, H). P deficiency decreased the RSA (TRL, RV, RA, LRN and RDW), shoot dry weight (SDW) and WLA significantly and resulted in decreased net photosynthesis rate. Moreover, L154 cultivar had lower RSA, SDW, WLA and NP than L133 cultivar at LP (Fig. 8). L154 was more sensitive to P deficiency

than L133.

Discussion

Advantages of WGR for genotyping in GWAS

GWAS is a powerful method for connecting identified phenotypic differences and the underlying causative loci (Korte et al. 2013). Currently, SNP chips and WGR are widely employed genotyping methods to identify genetic variants within genotypes combined with GWAS. By far, more than half of GWAS in *B. napus* were performed based on 60K SNP chip (Liu et al. 2022). As the continuous development of sequencing technology and the continuous reduction in costs, WGR has become a routine method to identify genetic variants. Compared with SNP chip, WGR data have more genetic loci and cover the whole genome (Liu et al. 2022). In *B. napus*, a total of 24,403 SNPs are selected to perform association mapping of flowering time (Jun et al. 2019), and 11,804 SNPs were used (Raman et al. 2019), which all the two panel were genotyped with a 60 K SNP array. However, 5.56 million SNPs were confirmed used for association analyses by WGR, and with identified 22 SNPs linked with 37 genes were associated with flowering time in *B. napus* (Wu et al. 2019). A total of 26,841 SNPs were selected by SNP array to perform association mapping of seed weight in *B. napus*, and detected a major QTL on chromosome A09 (Li et al. 2014). While, 690,953 SNPs high quality were used by WGR and 20 SNPs were detected to associate with seed weight in *B. napus* (Dong et al. 2018).

Previously, a total of 19,397 SNPs with minor allele frequency (MAF) > 0.05 were selected to conduct association analyses of RSA in *B. napus* at LP with the panel of *B. napus* lines genotyped with a 60 K *Brassica* Infinium SNP array(Wang et al. 2017). However, in this study, a total of 1.6 million higher quality SNPs (MAF > 0.2) identified by WGR were used for association analyses. As the 60 K SNP chip contains fewer SNPs than WGR, some SNPs that are significantly associated with the target traits could not be detected in the previous association analysis. A total of 52 significant SNPs associated with RSA at LP were identified on 12 of the 19 *B. napus* chromosomes, explaining a higher average of PVE (9.14%; Table S6) in this study than observed by

60 K SNP chip (an average of 6.09% PVE) (Wang et al. 2017). In this study, 12 significant SNPs detected by WGR for RSA were adjacent to previously reported significant SNPs detected by 60 K SNP chip, proving the accuracy of the results of GWAS in this study (Table S9). For example, significant SNP chrA03_26604503 located at 26604503 bp on chromosome A03 was associated with RDW in this study, and SNP Bn-A03-p28127216 located at 26598176 bp on chromosome A03 was also associated with RDW by Wang et al. (2017) (Table S9).

Favorable haplotype of candidate genes for RSA of *B. napus* at LP

GWAS of six RSA traits (LRL, TRL, LRD, LRN, MLRL and RDW) identified a total of 52 significant SNPs in *B. napus* at LP (Fig. 2; Table S6), and a total of 184 candidate genes were assessed based on the lead SNPs (Table S8). We did not compare the SNPs for RSA identified at LP and SP as the RSA of the association panel of *B. napus* at high P was not investigated in Wang et al. (2017). Some significant SNPs, candidate genes and haplotypes identified at low P might also have a function under sufficient P. Using GWAS data in combination with transcriptome data to mine candidate genes has become a conventional method in *B. napus*, such as harvest index (Lu et al. 2017), seed oil content (Tang et al. 2021), flowering time (Huang et al. 2021) and glucosinolate content (Kittipol et al. 2019). Based on our previously published transcriptome sequencing of the root and shoot of P-efficient cultivar (‘Eyouchangjia’) at SP and LP, four significantly differentially expressed genes were detected at LP in close proximity to the significant SNPs associated with RSA traits (Table 2). Gene transcripts of *BnaC03g07130D* and *BnaC07g43230D* were significantly increased, and the transcript of *BnaA06g29270D* was significantly decreased by Pi deficiency (Table 2; Fig. 3). These candidate genes might play an important part in root development of *B. napus* at a deficiency P supply. In this study, *BnaA06g29270D* was located within the interval of the SNP chrA06_20030601 for LRD (Fig. 2). *BnaA06g29270D* is homologous to *Arabidopsis At5g28770*, basic leucine zipper 63 (*bZIP63*), which was phosphorylated by Snf1-related-kinase1 (SnRK1) and then activates the promoter of AUXIN RESPONSE FACTOR 19 (ARF19), promoting LR emergence and subsequently LRD (Muralidhara et al. 2021). In *Arabidopsis*, *bzip63* mutants have increased PRL and

391 emerged LRD (defined as the number of LRs per primary root length) than wild type
392 plants under control conditions (Muralidhara et al. 2021). These data support the role
393 *BnaA06g29270D* in controlling LR density in *B. napus*.
394 *BnaC03g07130D*, a gene that encodes a reversibly glycosylated polypeptide 2
395 (RGP2), was located within the interval of the SNP chrC03_3535483 and linked with
396 the trait for LRL and TRL (Fig. 2; Table S6). In castor oilseeds (*Ricinus communis*),
397 sucrose synthase (SUS) interacts with RGP1 in roots, which suggests that these proteins
398 interact to directly channel UDPG derived toward RGP glycosyl chain initiation and
399 extension (Fedosejevs et al. 2017). Further study of the protein interaction between SUS
400 and RGP in root development will be conducive to a better understanding of RSA and
401 root carbon balance.
402 *BnaC07g43230D* was identified by significant SNP chrC07_42348561, which was
403 likely to influence the LRN and RDW at LP (Fig. 2; Table S6). *BnaC07g43230D* is an
404 AIG2-like (avirulence induced gene) family protein and its function is unknown. In
405 *Arabidopsis*, FIT (FER-LIKE IRON DEFICIENCY-INDUCED TRANSCRIPTION
406 FACTOR) is a key regulator of iron uptake in root. *AtAIG2*-like was detected by iron
407 deficiency in the roots of proteomic and transcriptomic study of WT, *fit* knock-out
408 mutant and FIT overexpression lines (Mai et al. 2015). Primary root extension is
409 inhibited and lateral root formation is stimulated at Pi deprivation in members of the
410 Brassicaceae family in agar system. Studies have confirmed external Pi sensing at the
411 root tips depends on Fe availability by hormone and peptide signaling pathways (Ward
412 et al., 2008; Ticconi et al., 2009; Müller et al., 2015; Singh et al., 2018). The function
413 of this BnaAIG2-like protein may is likely to be involved in RSA at LP and should be
414 examined in further studies.
415 Candidate gene association analysis is widely used to detect favorable haplotypes of
416 candidate genes. Favorable haplotypes of candidate gene *BnaC02.GTR2* associated
417 with seed glucosinolate content in *B. napus* (Tan et al. 2022) and the favorable
418 haplotype of candidate gene *BnaA09.MRP5,* which influenced rapeseed phytate content
419 (Liu et al. 2021b) have been reported previously. Additionally, overexpressing
420 favorable haplotype of *BnaA03.NIP5* increased low boron tolerance in boron inefficient

cultivar of *B. napus* (He et al. 2021).

The favorable haplotypes of *BnaA06g29270D*, *BnaC03g07130D* and *BnaC07g43230D* were determined as "CCGC", "CAAT" and "ATC" by candidate gene association analysis, respectively (Fig. 4-6). Furthermore, 73 *B. napus* cultivars containing three favorable haplotypes have higher LRL, LRN, RDW, and TRL than other *B. napus* cultivars (Fig. 7). L133, the cultivar containing three favorable haplotypes showed more tolerance to Pi starvation than L154, the cultivar which did not have these haplotypes confirm the root breeding value of these favorable haplotypes.

Conclusions

A total of 52 significant SNPs were significantly associated with RSA at LP by GWAS based on WGR of an association panel of *B. napus*. "CCGC", "CAAT" and "ATC" were identified to be favorable haplotypes in candidate genes, which could be used for molecular marker-assisted breeding of optimal RSA in response to low P availability in *B. napus*. In addition, gene editing and modification can be applied to reveal the function and the underlying mechanism of these candidate genes.

Data availability statement

The original contributions presented in this study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contribution

Pan Yuan, Haijiang Liu and Lei Shi designed research, reviewed the writing and drafted the manuscript. **Pan Yuan, Haijiang Liu and Xiaohua Wang** participated the experiments. **John P. Hammond** participated in manuscript revision.

Funding

This work was supported by the National Nature Science Foundation of China (Grant No. 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.

Data availability

Raw sequencing data of genome re-sequencing are available in the Genome Sequence Archive (https://bigd.big.ac.cn/gsa/) with Bio-project IDs PRJCA002835 and PRJCA002836. All the materials in this study are available upon request.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

Figure legend

Fig. 1 Distribution of RSA (root system architecture) phenotypes in 370 *B. napus* accessions. Histograms of the distribution of LRL (A), TRL (B), LRD (C), RN (D), MLRL (E) and RDW (F) measured at low phosphorus supplies. LRL, lateral root length; TRL, total root length; LRD, lateral root density; RN, lateral root number; MLRL; mean lateral root length; RDW, root dry weight.

Fig. 2 Quantile-quantile (QQ) and Manhattan plots for the RSA (root system architecture) by genome-wide association study (GWAS). GWAS of LRL (A), MLRL (B), TRL (C), RN (D), LRD (E) and RDW (F) in 370 *B. napus* at low phosphorus supplies by Fast-LMM model. LRL, lateral root length; MLRL, mean lateral root length; TRL, total root length; RN, lateral root number; LRD, lateral root density; RDW, root dry weight.

Fig. 3 Relative expression level of three candidate genes (A, *BnaA06g29270D*; B, *BnaC03g07130D* and C, *BnaC07g43230D*) under low phosphorus and sufficient phosphorus conditions. LP, low phosphorus; SP, sufficient phosphorus. Data are means ± SEM (n=4). Asterisks indicate the significance of Student's t-test (*P < 0.05, ***P < 0.0001, ****P < 0.0001).

Fig. 4 Candidate gene association and haplotypes analysis of *BnaC03g07130D* with LRL (lateral root length) and TRL (total root length). (A) Candidate gene association analysis of *BnaC03g07130D* with LRL. (B-E) Association of the two alleles in chrC03_3396486 (B), chrC03_3396712 (C), chrC03_3396812 (D) and chrC03_3396932 (E) with LRL, respectively. (F) Two haplotypes of *BnaC03g07130D*. (G) Candidate gene association analysis of *BnaC03g07130D* with TRL. (H-K) Association of the two alleles in chrC03_3396486 (H), chrC03_3396712 (I), chrC03_3396812 (J) and chrC03_3396932 (K) with TRL, respectively. (L) Two haplotypes of *BnaC03g07130D*. The number of inbred lines harbouring the corresponding allele is shown in brackets at the bottom.

Fig. 5 Candidate gene association and haplotypes analysis of *BnaC07g43230D* with LRN (lateral root number) and RDW (root dry weight). (A) Candidate gene association analysis of *BnaC07g43230D* with LRN. (B-D) Association of the three alleles in chrC07_42368601 (B), chrC07_42368638 (C) and chrC07_42368650 (D) with LRN, respectively. (E) Two haplotypes of *BnaC07g43230D*. (F) Candidate gene association analysis of *BnaC07g43230D* with LRN. (G-I) Association of the three alleles in chrC07_42368601 (G), chrC07_42368638 (H) and chrC07_42368650 (I) with RDW, respectively. (J) Two haplotypes of *BnaC07g43230D*. The number of inbred lines harbouring the corresponding allele is shown in brackets at the bottom.

Fig. 6 Candidate gene association and haplotypes analysis of *BnaA06g29270D* with LRD (lateral root density). (A) Candidate gene association analysis of *BnaA06g29270D* with LRD. (B-E) Association of the four alleles in chrA06_19975674 (B), chrA06_19975707 (C), chrA06_19975749 (D) and chrA06_19977004 (E) with LRD, respectively. (F) Two haplotypes of *BnaA06g29270D*. The number of inbred lines harbouring the corresponding allele is shown in brackets at the bottom.

Fig. 7 Differences in root system architecture at LP between *B. napus* cultivars with three favorable haplotypes (CCGC, CAAT and ATC) and *B. napus* cultivars without the three favorable haplotypes in the association panel. Favorable haplotypes of LRL (A), TRL (B), LRD (C), RN (D), MLRL (E) and RDW (F) were analysised. (73) represents 73 cultivars with the three favorable haplotypes; (297) represents the rest of the

association panel after removing the above 73 cultivars. LRL, lateral root length; TRL, total root length; LRD, lateral root density; RN, lateral root number; MLRL, mean lateral root length; RDW, root dry weight.

Fig. 8 Differences in root system architecture, biomass, whole leaf area and net photosynthesis between *B. napus* cultivar (L133) with three favorable haplotypes (CCGC, CAAT and ATC) and *B. napus* cultivar (L154) without the three favorable haplotypes (TGCC, TACA and TCT). (A) Root growth of L133 and L154 at LP for 14 d in the 'pouch and wick' system. (B) TRL. (C) RV. (D) RA. (E) LRN. (F) RDW. (G) SDW. (H) WLA. (I) NP. LP, low phosphorus; TRL, total root length; RV, root volume; RA, root area; LRN, lateral root number; RDW, root dry weight; SDW, shoot dry weight; WLA, whole leaf area; NP, net photosynthesis. Data are means ± SEM (n=6). Asterisks indicate the significance of Student's t-test (*P < 0.05, **P < 0.01, ***P < 0.0001, ****P < 0.0001). Scale bar is 2 cm in (A).

Supplementary material

The following supplemental materials are available.

Fig. S1. Correlation of eight root related traits at low phosphorus supplies.

Fig. S2. The LD decay of an association panel of *B. napus*.

Fig. S3. Population structure of an association panel of *B. napus* with K from 2 to 8.

Fig. S4. The kinship of an association panel of 370 *B. napus* accessions.

Fig. S5. Distribution of linkage disequilibrium block sizes across all chromosomes.

Table S1. List of 370 accessions of *B. napus* used in the study.

Table S2. Primers used for qRT-PCR.

Table S3. Root related traits at low phosphorus supplies in an association panel of *B. napus*.

Table S4. AMOVA analysis between the K = 4 assigned populations in *Brassica napus*.

Table S5. Linkage disequilibrium block in this study.

Table S6. Significant SNP loci for root related traits of *B. napus* by genome wide association study at low phosphorus supplies.

Table S7. LD blocks harboring significant SNPs associated with RSA.

Table S8. Candidate genes within LD decay value up and down the lead SNPs (chrA06_19934701, chrC03_3535476 and chrC07_42348526) for root related traits.

Table S9. Comparison of SNPs detected by WGR for RSA in this study with previously identified SNPs by 60 K SNP chip for RSA at a low phosphorus supply.

References

Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. Genome Res 19 (9):1655-1664. https://doi.org/10.1101/gr.094052.109

Arifuzzaman M, Oladzadabbasabadi A, McClean P et al (2019) Shovelomics for phenotyping root architectural traits of rapeseed/canola (*Brassica napus* L.) and genome-wide association mapping. Mol Genet Genomics 294 (4):985-1000. https://doi.org/10.1007/s00438-019-01563-x

Barrett JC, Fry B, Maller J et al (2005) Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21 (2):263-265. https://doi.org/10.1093/bioinformatics/bth457

Bradbury PJ, Zhang Z, Kroon DE et al (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23 (19):2633-2635. https://doi.org/10.1093/bioinformatics/btm308

Chalhoub B, Denoeud F, Liu SY et al (2014) Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. Science 345 (6199):950-953. https://doi.org/10.1126/science.1253435

Clark RT, Famoso AN, Zhao K et al (2013) High-throughput two-dimensional root system phenotyping platform facilitates genetic analysis of root growth and development. Plant Cell Environ 36 (2):454-466. https://doi.org/10.1111/j.1365-3040.2012.02587.x

Danecek P, Auton A, Abecasis G et al (2011) The variant call format and VCFtools. Bioinformatics 27 (15):2156-2158. https://doi.org/10.1093/bioinformatics/btr330

Dong HL, Tan CD, Li YZ et al (2018) Genome-wide association study reveals both

overlapping and independent genetic loci to control seed weight and silique length in *Brassica napus*. Front Plant Sci 9:921. https://doi.org/10.3389/fpls.2018.00921

Dong SS, He WM, Ji JJ et al (2021) LDBlockShow: a fast and convenient tool for visualizing linkage disequilibrium and haplotype blocks based on variant call format files. Brief Bioinform 22 (4). https://doi.org/10.1093/bib/bbaa227

Du HY, Yang C, Ding GD (2017) Genome-wide identification and characterization of SPX domain-containing members and their responses to phosphate deficiency in *Brassica napus*. Front Plant Sci 8. https://doi.org/10.3389/fpls.2017.00035

Duan XJ, Wang XH, Jin KM et al (2021) Genetic dissection of root angle of *Brassica napus* in response to low phosphorus. Front Plant Sci 12. https://doi.org/10.3389/Fpls.2021.697872

Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10 (3):564-567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Fedosejevs ET, Liu LNC, Abergel M et al (2017) Coimmunoprecipitation of reversibly glycosylated polypeptide with sucrose synthase from developing castor oilseeds. Febs Lett 591 (23):3872-3880. https://doi.org/10.1002/1873-3468.12893

He ML, Wang SL, Zhang C et al (2021) Genetic variation of *BnaA3.NIP5;1* expressing in the lateral root cap contributes to boron deficiency tolerance in *Brassica napus*. Plos Genet 17 (7). https://doi.org/10.1371/journal.pgen.1009661

He YJ, Hu DX, You JC et al (2019) Genome-wide association study and protein network analysis for understanding candidate genes involved in root development at the rapeseed seedling stage. Plant Physiol 137:42-52. https://doi.org/10.1016/j.plaphy.2019.01.028

Huang LY, Min Y, Schiessl S et al (2021) Integrative analysis of GWAS and transcriptome to reveal novel loci regulation flowering time in semi-winter rapeseed. Plant Sci 310. https://doi.org/10.1016/j.plantsci.2021.110980

Jan HU, Guan M, Yao M et al (2019) Genome-wide haplotype analysis improves trait predictions in *Brassica napus* hybrids. Plant Sci 283:157-164. https://doi.org/10.1016/j.plantsci.2019.02.007

Kittipol V, He ZS, Wang LH et al (2019) Genetic architecture of glucosinolate variation in *Brassica napus*. J Plant Physiol 240. https://doi.org/10.1016/J.Jplph.2019.06.001

Korte A, Farlow A (2013) The advantages and limitations of trait analysis with GWAS: a review. Plant Methods 9. https://doi.org/10.1186/1746-4811-9-29

Li F, Chen B, Xu K et al (2014) Genome-wide association study dissects the genetic architecture of seed weight and seed quality in rapeseed (*Brassica napus* L.). DNA Res 21 (4):355-367. https://doi.org/10.1093/dnares/dsu002

Li KQ, Wang J, Kuang LQ et al (2021) Genome-wide association study and transcriptome analysis reveal key genes affecting root growth dynamics in rapeseed. Biotechnol Biofuels 14 (1). https://doi.org/10.1186/s13068-021-02032-7

Li Y, Wang X, Zhang H et al (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 (7). https://doi.org/10.1371/journal.pone.0220374

Liu HJ, Wang JC, Zhang BB et al (2021a) Genome-wide association study dissects the genetic control of plant height and branch number in response to low-phosphorus stress in *Brassica napus*. Ann Bot 128 (7):919-929. https://doi.org/10.1093/aob/mcab115

Liu HJ, Li XJ, Zhang QW et al (2021b) Integrating a genome-wide association study with transcriptomic data to predict candidate genes and favourable haplotypes influencing *Brassica napus* seed phytate. DNA Res 28 (5). https://doi.org/10.1093/dnares/dsab011

Liu HJ, Wang W, Yang M et al (2022) Genome-wide association studies of important agronomic traits in *Brassica napus*: what we have learned and where we are headed. Annual Plant Reviews 5:1-30.

Liu YX, Wang L, Deng M et al (2015) Genome-wide association study of phosphorus-deficiency-tolerance traits in *Aegilops tauschii*. Theor Appl Genet 128 (11):2203-2212. https://doi.org/10.1007/s00122-015-2578-x

631 Lu K, Peng L, Zhang C et al (2017) Genome-wide association and transcriptome
632 analyses reveal candidate genes underlying yield-determining traits in *Brassica*
633 *napus*. Front Plant Sci 8. https://doi.org/10.3389/Fpls.2017.00206
634 Luo T, Zhang YT, Zhang CN et al (2021) Genome-wide association mapping unravels
635 the genetic control of seed vigor under low-temperature conditions in rapeseed
636 (*Brassica napus* L.). Plants (Basel) 10 (3).
637 https://doi.org/10.3390/Plants10030426
638 Lynch J (1995) Root Architecture and Plant Productivity. Plant Physiol 109 (1):7-13.
639 https://doi.org/10.1104/Pp.109.1.7
640 Mai HJ, Lindermayr C, von Toerne C et al (2015) Iron and FER-LIKE IRON
641 DEFICIENCY-INDUCED TRANSCRIPTION FACTOR-dependent regulation
642 of proteins and genes in *Arabidopsis thaliana* roots. Proteomics 15 (17):3030-
643 3047. https://doi.org/10.1002/pmic.201400351
644 Müller J, Toev T, Heisters M. et al (2015). Iron-dependent callose deposition adjusts
645 root meristem maintenance to phosphate availability. Dev Cell 33 (2): 216-230.
646 https://doi.org/10.1016/j.devcel.2015.02.007
647 Muralidhara P, Weiste C, Collani S et al (2021) Perturbations in plant energy
648 homeostasis prime lateral root initiation via SnRK1-bZIP63-ARF19 signaling.
649 Proc Natl Acad Sci USA 118 (37). https://doi.org/10.1073/pnas.2106961118
650 Raman H, Raman R, Qiu Y et al (2019) GWAS hints at pleiotropic roles for
651 FLOWERING LOCUS T in flowering time and yield-related traits in canola. BMC
652 Genomics 20 (1):636. https://doi.org/10.1186/s12864-019-5964-y
653 Shi L, Shi TX, Broadley MR et al (2013) High-throughput root phenotyping screens
654 identify genetic loci associated with root architectural traits in *Brassica napus*
655 under contrasting phosphate availabilities. Ann Bot 112 (2):381-389.
656 https://doi.org/10.1093/aob/mcs245
657 Singh AP, Fridman Y, Holland N et al (2018). Interdependent nutrient availability and
658 steroid hormone signals facilitate root growth plasticity. Dev Cell 46 (1):59-
659 72.e4. https://doi.org/10.1016/j.devcel.2018.06.002
660 Tan ZD, Xie ZQ, Dai LH et al (2022) Genome- and transcriptome-wide association

studies reveal the genetic basis and the breeding history of seed glucosinolate content in *Brassica napus*. Plant Biotechnol J 20 (1):211-225. https://doi.org/10.1111/pbi.13707

Tang S, Zhao H, Lu SP et al (2021) Genome- and transcriptome-wide association studies provide insights into the genetic basis of natural variation of seed oil content in *Brassica napus*. Mol Plant 14 (3):470-487. https://doi.org/10.1016/j.molp.2020.12.003

Ticconi CA, Lucero RD, Sakhonwasee S et al (2009) ER-resident proteins PDR2 and LPR1 mediate the developmental response of root meristems to phosphate availability. Proc Natl Acad Sci U S A. 106 (33):14174-14179. https://doi.org/10.1073/pnas.0901778106

Vance CP, Uhde-Stone C, Allan DL (2003) Phosphorus acquisition and use: critical adaptations by plants for securing a nonrenewable resource. New Phytol 157 (3):423-447. https://doi.org/10.1046/j.1469-8137.2003.00695.x

Ward JT, Lahner B, Yakubova E, Salt DE, Raghothama KG (2008) The effect of iron on the primary root elongation of *Arabidopsis* during phosphate deficiency. Plant Physiol 147:1181-1191. https://doi.org/10.1104/pp.108.118562

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 Res 24 (4):407-417. https://doi.org/10.1093/dnares/dsx013

Wu D, Liang Z, Yan T et al (2019) Whole-Genome Resequencing of a Worldwide Collection of Rapeseed Accessions Reveals the Genetic Basis of Ecotype Divergence. Mol Plant 12 (1):30-43. https://doi.org/10.1016/j.molp.2018.11.007

Xiao YJ, Liu HJ, Wu LJ et al (2017) Genome-wide Association Studies in Maize: Praise and Stargaze. Mol Plant 10 (3):359-374. https://doi.org/10.1016/j.molp.2016.12.008

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

691 35 (10):1786-1788. https://doi.org/10.1093/bioinformatics/bty875
692 Zhang Y, Thomas CL, Xiang JX et al (2016) QTL meta-analysis of root traits in
693 *Brassica napus* under contrasting phosphorus supply in two growth systems.
694 Sci Rep 6. https://doi.org/10.1038/Srep33113
695
696

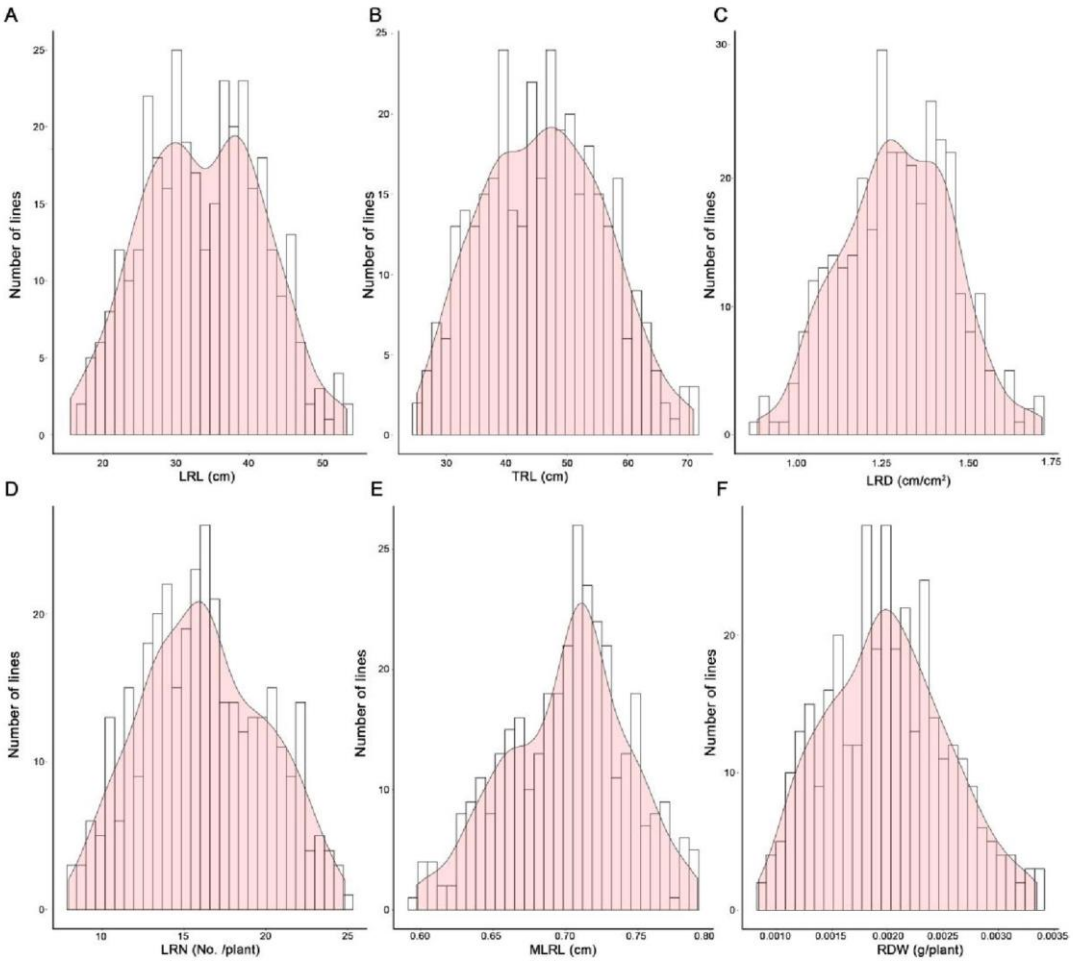


Fig. 1 Distribution of RSA (root system architecture) phenotypes in 370 *B. napus* accessions. Histograms of the distribution of LRL (A), TRL (B), LRD (C), RN (D), MLRL (E) and RDW (F) measured at low phosphorus supplies. LRL, lateral root length; TRL, total root length; LRD, lateral root density; RN, lateral root number; MLRL; mean lateral root length; RDW, root dry weight.

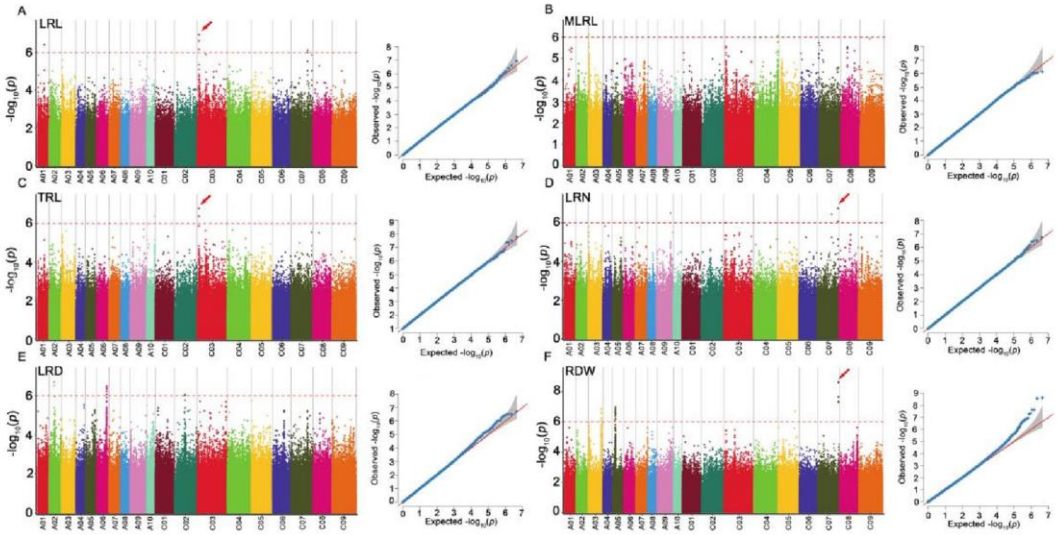


Fig. 2 Quantile-quantile (QQ) and Manhattan plots for the RSA (root system architecture) by genome-wide association study (GWAS). GWAS of LRL (A), MLRL (B), TRL (C), RN (D), LRD (E) and RDW (F) in 370 *B. napus* at low phosphorus supplies by Fast-LMM model. LRL, lateral root length; MLRL, mean lateral root length; TRL, total root length; RN, lateral root number; LRD, lateral root density; RDW, root dry weight.

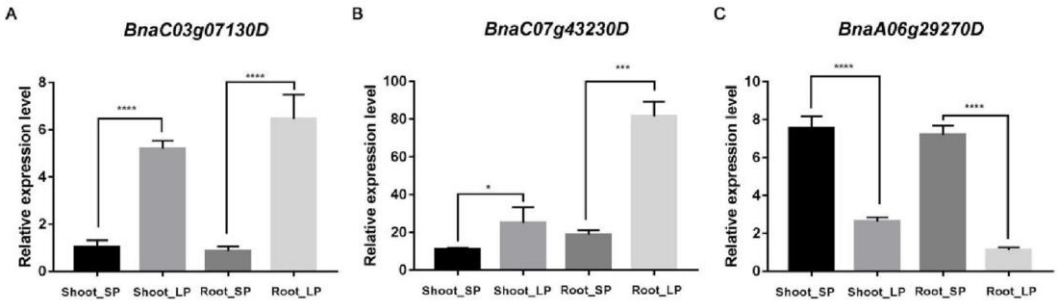
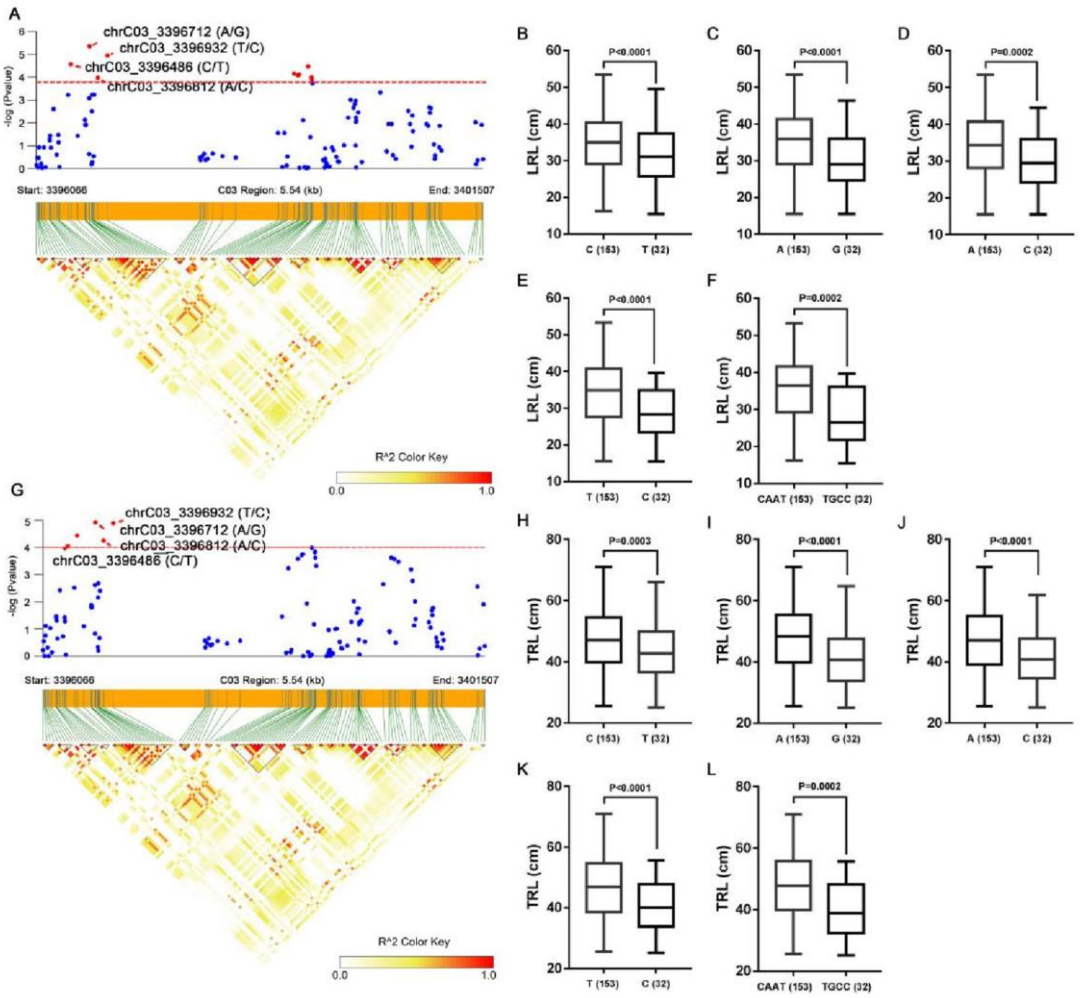


Fig. 3 Relative expression level of three candidate genes (A, *BnaA06g29270D*; B, *BnaC03g07130D* and C, *BnaC07g43230D*) under low phosphorus and sufficient phosphorus conditions. LP, low phosphorus; SP, sufficient phosphorus. Data are means ± SEM (n=4). Asterisks indicate the significance of Student's t-test (*P < 0.05, ***P < 0.0001, ****P < 0.0001).



723

724 **Fig. 4** Candidate gene association and haplotypes analysis of *BnaC03g07130D* with
725 LRL (lateral root length) and TRL (total root length). (A) Candidate gene association
726 analysis of *BnaC03g07130D* with LRL. (B-E) Association of the two alleles in
727 chrC03_3396486 (B), chrC03_3396712 (C), chrC03_3396812 (D) and
728 chrC03_3396932 (E) with LRL, respectively. (F) Two haplotypes of *BnaC03g07130D*.
729 (G) Candidate gene association analysis of *BnaC03g07130D* with TRL. (H-K)
730 Association of the two alleles in chrC03_3396486 (H), chrC03_3396712 (I),
731 chrC03_3396812 (J) and chrC03_3396932 (K) with TRL, respectively. (L) Two
732 haplotypes of *BnaC03g07130D*. The number of inbred lines harbouring the
733 corresponding allele is shown in brackets at the bottom.

734

735

736

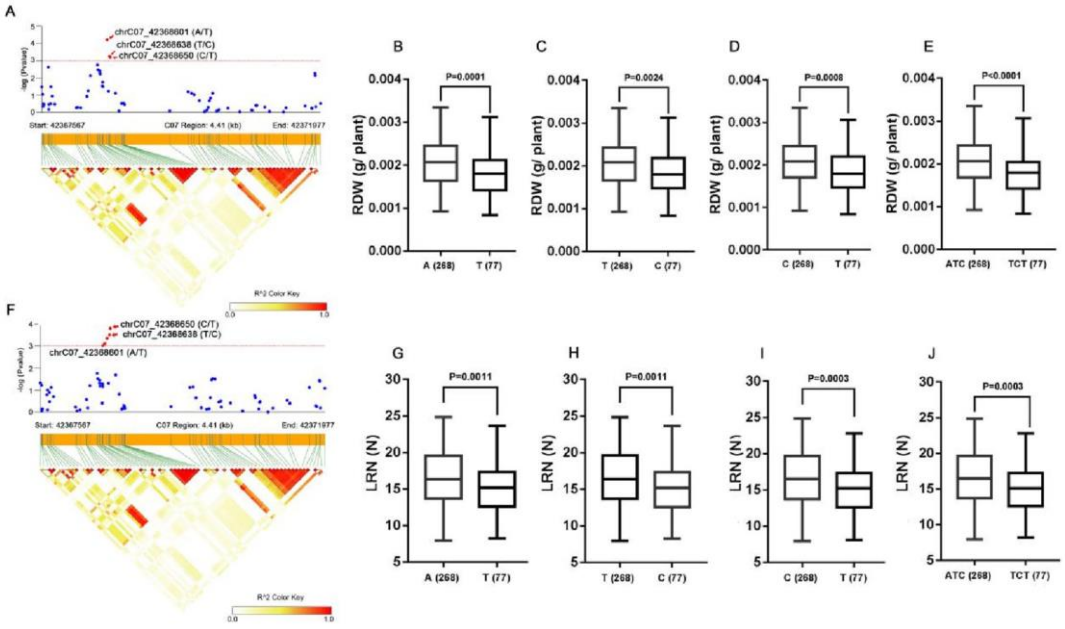


Fig. 5 Candidate gene association and haplotypes analysis of *BnaC07g43230D* with LRN (lateral root number) and RDW (root dry weight). (A) Candidate gene association analysis of *BnaC07g43230D* with LRN. (B-D) Association of the three alleles in chrC07_42368601 (B), chrC07_42368638 (C) and chrC07_42368650 (D) with LRN, respectively. (E) Two haplotypes of *BnaC07g43230D*. (F) Candidate gene association analysis of *BnaC07g43230D* with LRN. (G-I) Association of the three alleles in chrC07_42368601 (G), chrC07_42368638 (H) and chrC07_42368650 (I) with RDW, respectively. (J) Two haplotypes of *BnaC07g43230D*. The number of inbred lines harbouring the corresponding allele is shown in brackets at the bottom.

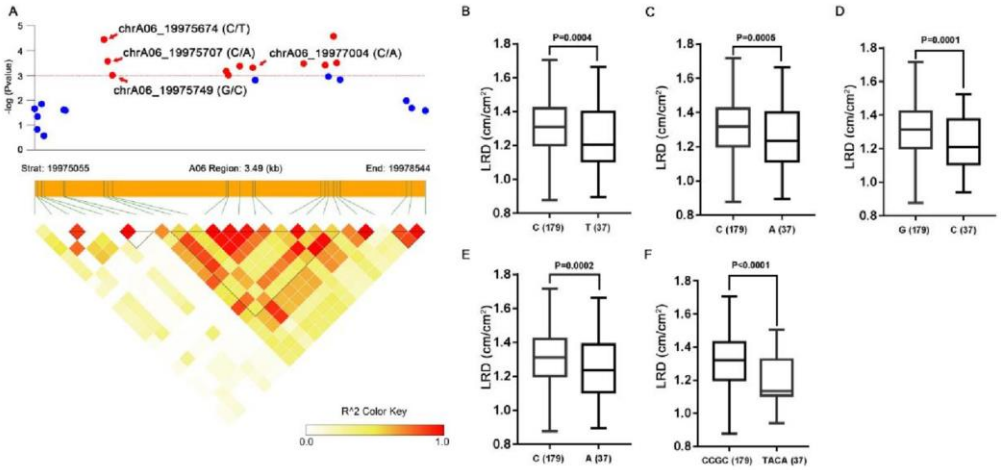
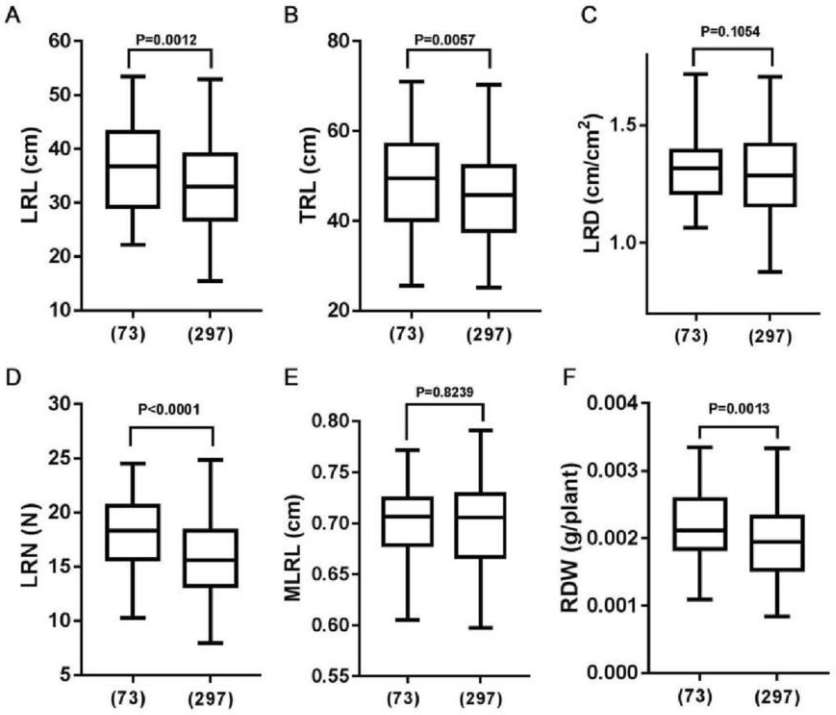


Fig. 6 Candidate gene association and haplotypes analysis of *BnaA06g29270D* with LRD (lateral root density). (A) Candidate gene association analysis of *BnaA06g29270D* with LRD. (B-E) Association of the four alleles in chrA06_19975674 (B), chrA06_19975707 (C), chrA06_19975749 (D) and chrA06_19977004 (E) with LRD, respectively. (F) Two haplotypes of *BnaA06g29270D*. The number of inbred lines harbouring the corresponding allele is shown in brackets at the bottom.

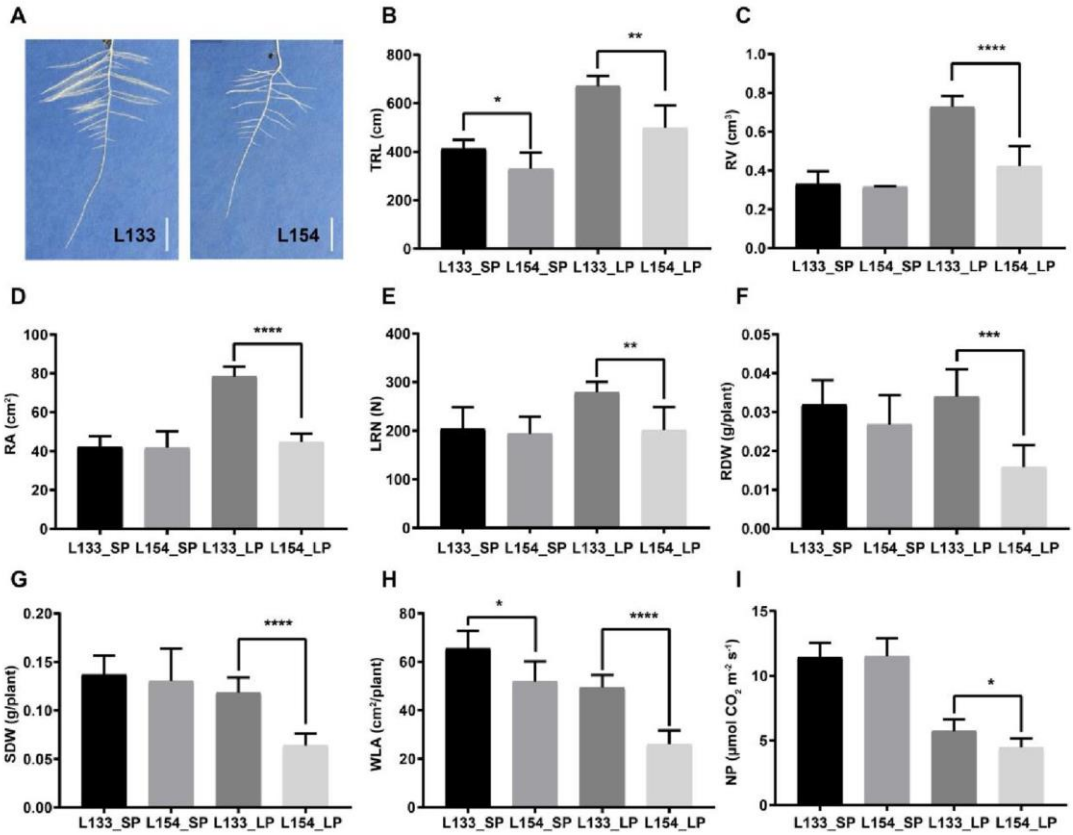


758

759 **Fig. 7** Differences in root system architecture at LP between *B. napus* cultivars with
760 three favorable haplotypes (CCGC, CAAT and ATC) and *B. napus* cultivars without the
761 three favorable haplotypes in the association panel. Favorable haplotypes of LRL (A),
762 TRL (B), LRD (C), RN (D), MLRL (E) and RDW (F) were analysised. (73) represents
763 73 cultivars with the three favorable haplotypes; (297) represents the rest of the
764 association panel after removing the above 73 cultivars. LRL, lateral root length; TRL,
765 total root length; LRD, lateral root density; RN, lateral root number; MLRL, mean
766 lateral root length; RDW, root dry weight.

767

768



769

770 **Fig. 8** Differences in root system architecture, biomass, whole leaf area and net
771 photosynthesis between *B. napus* cultivar (L133) with three favorable haplotypes
772 (CCGC, CAAT and ATC) and *B. napus* cultivar (L154) without the three favorable
773 haplotypes (TGCC, TACA and TCT). (A) Root growth of L133 and L154 at LP for 14
774 d in the 'pouch and wick' system. (B) TRL. (C) RV. (D) RA. (E) LRN. (F) RDW. (G)
775 SDW. (H) WLA. (I) NP. LP, low phosphorus; TRL, total root length; RV, root volume;
776 RA, root area; LRN, lateral root number; RDW, root dry weight; SDW, shoot dry weight;
777 WLA, whole leaf area; NP, net photosynthesis. Data are means ± SEM (n=6).
778 Asterisks indicate the significance of Student's t-test (*P < 0.05, **P < 0.01, ***P <
779 0.0001, ****P < 0.0001). Scale bar is 2 cm in (A).

780