

Haplotypes within tandemly duplicated candidate genes at BnaA9.MRP5 modulate phytate concentration in canola (*Brassica napus* L.)

Haijiang Liu¹, xiaojuan Li², Qianwen Zhang³, pan yuan¹, Lei Liu⁴, Graham King⁴, Guangda Ding¹, Sheliang Wang¹, Hongmei Cai¹, Chuang Wang¹, Fangsen XU¹, and Lei Shi¹

¹Huazhong Agricultural University

²Affiliation not available

³Shanghai Jiao Tong University School of Agriculture and Biology

⁴Southern Cross University

January 26, 2021

Abstract

Phytate is the storage form of phosphorus in angiosperm seeds and plays vitally important roles during seed development. However, in crop plants phytate decreases bioavailability of seed-sourced mineral elements for humans, livestock and poultry, and contributes to phosphate-related water pollution. However, there is little knowledge about this trait in oilseed rape *B. napus* (oilseed rape). Here, a panel of 505 diverse *B. napus* accessions was screened in a genome-wide association study (GWAS) using 3.28 x 10⁶ single nucleotide polymorphisms (SNPs). This identified 119 SNPs significantly associated with phytate concentration (PA_Conc) and phytate content (PA_Cont) and six candidate genes were identified. Of these, BnaA9.MRP5 represented the candidate gene for the significant SNP chrA09_5198034 (27kb) for both PA_Cont and PA_Conc. Transcription of BnaA9.MRP5 in a low -phytate variety (LPA20) was significantly elevated compared with a high -phytate variety (HPA972). Association and haplotype analysis indicated that inbred lines carrying specific SNP haplotypes within BnaA9.MRP5 were associated with high- and low-phytate phenotypes. No significant differences in seed germination and seed yield were detected between low and high phytate cultivars examined. Candidate genes, favorable haplotypes and the low phytate varieties identified in this study will be useful for low-phytate breeding of *B. napus*.

Haplotypes within tandemly duplicated candidate genes at *BnaA9.MRP5* modulate phytate concentration in canola (*Brassica napus* L.)

Haijiang Liu^{1,2, #}, Xiaojuan Li^{1,2, #}, Qianwen Zhang^{3#}, Pan Yuan^{1,2}, Lei Liu⁴, Graham John King⁴, Guangda Ding^{1,2}, Sheliang Wang^{1,2}, Hongmei Cai², Chuang Wang², 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 and Biology, Shanghai JiaoTong University, 800 Dongchuan Road, Minhang District, Shanghai, 40, China

4Southern Cross Plant Science, Southern Cross University, Lismore, NSW 2480, Australia

*Corresponding Author: Tel: (+8627) 87286871; E-mail: leish@mail.hzau.edu.cn

Abstract : Phytate is the storage form of phosphorus in angiosperm seeds and plays vitally important roles during seed development. However, in crop plants phytate decreases bioavailability of seed-sourced mineral elements for humans, livestock and poultry, and contributes to phosphate-related water pollution. However, there is little knowledge about this trait in oilseed rape *B. napus*(oilseed rape). Here, a panel of 505 diverse *B. napus* accessions was screened in a genome-wide association study (GWAS) using 3.28 x 10⁶ single nucleotide polymorphisms (SNPs). This identified 119 SNPs significantly associated with phytate concentration (PA_Conc) and phytate content (PA_Cont) and six candidate genes were identified. Of these, *BnaA9.MRP5* represented the candidate gene for the significant SNP chrA09_5198034 (27kb) for both PA_Cont and PA_Conc. Transcription of *BnaA9.MRP5* in a low -phytate variety (LPA20) was significantly elevated compared with a high -phytate variety (HPA972). Association and haplotype analysis indicated that inbred lines carrying specific SNP haplotypes within *BnaA9.MRP5* were associated with high- and low-phytate phenotypes. No significant differences in seed germination and seed yield were detected between low and high phytate cultivars examined. Candidate genes, favorable haplotypes and the low phytate varieties identified in this study will be useful for low-phytate breeding of *B. napus* .

KEYWORDS : Oilseed rape, phytate, genome-wide association study, *BnaA9.MRP5* , haplotype analysis.

INTRODUCTION

Phosphorus (P) is an essential macro-element for higher plant growth and development. Phytate is the storage form of P in angiosperm seeds, accounting for approximately two-thirds of total seed P. Seed phytate provides *myo* -inositol and P used for seed germination and seedling establishment (Raboy et al., 2000). It also plays vitally important roles in P sensing and homeostasis during seed development. The function of inositol triphosphate during phytate biosynthesis is as a second messenger mediating intracellular calcium signaling and drought stress responses in plants, along with transcriptional induction of genes associated with inositol metabolism (Berridge and Irvine, 1989; Boominathan et al., 2004; Kaur et al., 2008; Saxena et al., 2013). Additionally, phytate is involved in DNA damage-repair, RNA-editing, mRNA export and other intracellular regulation in plants (Shears et al., 2004). Several enzymes involved in the biosynthesis of phytate have been identified in Arabidopsis and several crop species including rice, maize, and rapeseed, including genes that could be engineered for seed- or tissue-specific regulation of phytic-acid synthesis, such as *AtITPK1* , *AtITPK2*, *OsMIPS* , *OsIPK*, *OsMIK* , *OsMRP5* , *OsPLD1* , *OSLPA1* , *ZmMRP4* , *BnPGK2* , *BnITPK1* and *BnITPK2* (Shi et al., 2007; Xu et al., 2009; Ali et al., 2013a ; Ali et al., 2013b; Li et al., 2014a; Li et al., 2014b; Khan et al ., 2019; Kishor et al ., 2019; Laha et al., 2019; Sashidhar et al ., 2020a; Sashidhar et al., 2020b). Phytate chelates positive ions of essential minerals, such as Mg²⁺, Ca²⁺, Fe³⁺ and Zn²⁺, forming phytate that is unable to be absorbed and used by human and other monogastric animals without phytase in the digestive system or exogenous phytase pre-treatment (Raboy, 2020). This leads to a decrease in the bioavailability of seed-sourced mineral elements for humans, livestock and poultry, as well as phosphate-related water pollution (Brown and Solomons, 1991). Thus, an optimal phytate content in the seed of crops should not only satisfy requirements for seedling establishment but have minimal negative effects on animal nutrition and the environment. Hence, it is crucial to resolve genetic loci associated with the phytate concentration in the seed and identify key genes in the pathway of phytic acid biosynthesis that may be suitable targets for selective or other advanced breeding.

Genetic linkage mapping and identification of quantitative trait loci has been widely used to study the genetic basis of phytate related traits in plants (Bentsink et al., 2003; Stangoulis et al., 2006; Blair et al., 2009; Sompong et al., 2012). In *Arabidopsis thaliana*(*A. thaliana*), a major QTL affecting InsP₆ and Pi in seeds and leaves has been detected on the top of chromosome 3 by the *Ler* /*Cvi* RIL population, which explained 61.8% of the variation for InsP₆ in seeds (Bentsink et al., 2003). In *Brassica rapa* , a total of 25 QTLs for phytate and phosphate concentrations in seeds and leaves were detected on seven linkage groups with five genetic populations, explaining 12.7–31.9% of the phenotypic variation (Zhao et al., 2008). However, due to a limited number of genetic markers and low frequency of recombination in the mapping population, few candidate genes were map-based cloning.

Genome-wide association studies (GWAS) take advantage of phenotypic variation and historical recombina-

nation in natural populations without the time and energy required to construct a segregating mapping population (Nordborg and Weigel., 2008). GWAS has become a routine method for analyzing complex agronomic traits, such as ionic variation in rice (Yang et al., 2018), arsenic accumulation in maize (Zhao et al., 2018), lateral root development in soybean (Prince et al., 2019), seed size in *Arabidopsis* (Ren et al., 2019) and cannabinoid composition in *Cannabis* (Welling et al., 2020).

Oilseed rape (OSR, rapeseed, canola: *B.napus*) provides 18% of global vegetable oil for human consumption, industrial oils, biodiesel, lubricant, and fodder for animal feeds (FAOSTAT, 2014). After extracting oil, OSR cakes and OSR meals function as an important poultry feed and fertilizer. Since the release of the *B. napus* reference genome (Chalhoub et al., 2014), dozens of agriculturally important traits have been dissected by GWAS (Wang et al., 2017; Wu et al., 2019; Lu et al., 2019; Tang et al., 2020) although few studies have included *B. napus* seed phytate traits. In this study, a diverse panel of 505 *B. napus* accessions were scored for seed phytate concentration and content and 119 significant associated SNP loci were identified by GWAS. A multidrug resistance-associated protein 5 (*BnaA9.MRP5*) gene was identified as a candidate gene, with eight distinct haplotypes associated with seed phytate concentration and content. These results will contribute to improving our understanding of the genetic mechanism of seed phytate accumulation in the seed of *B. napus* and breeding of low phytate *B.napus* cultivars.

METHODS

Plant materials

An association panel of *B. napus* 505 diverse accessions collected worldwide was used in this study, including 420 semi-winter, 59 spring, 16 winter and 10 unknown types, collected from major breeding centers across China. 443 lines originated in China, 28 from Europe, 10 from Japan, 5 from Canada, 6 from Australia, 3 from Korea, and 10 unknown (Tang et al., 2020).

Field trials of the association panel and Trait Measurement

The association panel was used to conduct two years of field trials at the experimental site of Huazhong Agricultural University in Wuhan (114.32° E, 30.52° N) from October 2013 to May 2014, and from October 2015 to May 2016. The soil was a yellow-brown soil (Alfisol), with the following properties: pH 6.84 (1:5 soil solution ratio), organic matter 12.99 g kg⁻¹, NH₄OAc-extracted potassium (K) 141.14 mg kg⁻¹, total nitrogen (N) 0.70 g kg⁻¹, available N 51.30 mg kg⁻¹, and Olsen-P 9.79 mg kg⁻¹. All plots received basal fertilizer, including 60% of the total N applied (supplied as urea), and 100% of the total P (supplied as calcium superphosphate), K (supplied as potassium chloride) and boron (supplied as Na₂B₄O₇·10H₂O) applied. Application rates were as follows: N 108 kg ha⁻¹, P₂O₅ 90 kg ha⁻¹, K₂O 120 kg ha⁻¹ and Borax 15 kg ha⁻¹. These fertilizers were thoroughly mixed and applied in bands near the crop rows. The remaining N was top dressed as urea in equal amounts at the four to five-leaf stage and at the stem elongation stage, respectively. Each accession was planted in three replicates blocks with 30 plants (3 rows of 10 plants) per replicate in a randomized complete block design. In order to avoid intercrossing, a voile bag was used to cover each individual inflorescence prior to flowering. Three groups of five well-filled, self-pollinated seeds randomly selected from each accession were used to measure the phytate concentration, from which was the phytate content was calculated.

Phytate analysis in seeds

Phytate was analyzed using a modification of an existing method (Shi et al., 2003). Five seeds were weighed, placed into a 1.5 mL Microfuge tube together with a diameter 3 mm tungsten ball bearing (Qiagen, Germany) and 1 mL of HCl. The mixture was ground for 3 minutes in a TissueLyser ball mill (Qiagen, Germany). The extracts were centrifuged at 15,000 rpm for 5 minutes, and 0.5 mL of supernatant was transferred to a 1.5 mL Microfuge tube. Because the seed oil layer was heavy, this step was repeated twice and finally 0.2 ml of supernatant removed into a fresh 1.5 ml Microfuge tube. Subsequently, 15 µl water, 30 µl 0.2 M HCl and 15 µl of each extracted sample was added to a 96 well PCR plate, which was centrifuged for 20 s at 300 rpm, following which 120 µl of 0.02% (w/v) ammonium iron (III) sulphate-0.2 M HCl was added to each well.

Wells were capped and the plate heated at 99 °C for 30 min. after which the PCR plate was cooled in an ice bath for 10 min and then centrifuged at 3,000 *g* for 30 min. 80 µl of each sample and standard solution (0.016 mg/ml - 0.131 mg/ml, phytic acid, Sigma-Aldrich) from the PCR plate were transferred into a 96 well plate (flat bottom) for colorimetric determination. 120 µl of 1% (w/v) 2, 2'-bipyridine-1% (v/v) thioglycolic acid was then added and mixed for 10 min on a plate shaker. Finally, absorbance at 519 nm was determined using a KC4 multi-detection microplate reader (Bio-Tek Instruments, USA) with three biological replicates.

Genome-wide association analysis

A total of 3,280,623 SNPs with a minor-allele frequency (MAF) > 0.05 and missing rate of <0.2, were called in the association panel used in this study (Tang et al., 2020). Genome-wide association analysis for seed phytate concentration and content was carried out using general linear models (GLM) and mixed linear models (MLM) using the Tassel 5.0 software (Bradbury et al., 2007). Fixed effects were calculated with a Q (population structure) matrix, and random effects were calculated with a K (Kinship) matrix. While only the Q matrix was taken into account in the GLM model, the Q+K matrices were both considered in the MLM model. 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 any SNPs related with other SNPs within the window with correlation coefficient (R^2) > 0.1. Finally, a total of 412,141 SNPs were used to infer the population structure within the association panel using the Admixture software (Alexander et al., 2009). The kinship matrix was predicted using the Tassel 5.0 software (Bradbury et al., 2007). The value less than 0 in kinship was defined as 0. The linkage disequilibrium (LD) statistic R^2 was calculated by PopLDdecay software (Zhang et al., 2019). The significant P-value thresholds for the association panel was 6.25×10^{-07} . The genotypes of *BnaA9.MRP5* (*BnaA09g10220D* and *BnaA09g10230D*) in the association panel were obtained by vcftools software(<https://vcftools.github.io/index.html>). Candidate gene association analysis of *BnaA09g10220D* and *BnaA09g10230D* was performed using Tassel 5.0 software (Bradbury et al., 2007). The SNP markers from 2 kb up- the gene to 2 kb down-stream of the gene were used to conduct association analysis with the phytate concentration and content in the seed of the association panel of *B. napus*.

Identification of candidate genes for phytate concentration and content in the seed

B. napus genes orthologous to *A. thaliana* phytate synthesis related genes located within 300 kb of the SNPs co-located with GLM and MLM models were identified based on the *B. napus* reference genome (Chalhoub et al., 2014). Finally, genes with significantly different levels of transcription in seeds between high and low phytic acid varieties were regarded as candidate genes for phytate synthesis.

Haplotype analysis

HaploView software was used to conduct haplotype analysis (Barrett et al., 2005). The haplotypes were classified based on all of the SNPs with an MAF > 0.05 in a candidate gene. Haplotypes containing at least fifteen *B.napus* accessions were used for final comparative analysis, and Student's *t*-test were used to compare the differences in phytate concentration and content among the haplotypes.

Pot culture experiment and gene expression analysis of HPA972 and LPA20

Among the association panel, HPA972 (a high phytate variety) and LPA20 (a low phytate variety) had similar thousand seed weight, but seed phytate concentration of the former was far higher than that the latter (Table 1). They were used for comparing the difference in the the gene expression of the candidate genes identified in this study in the developing seeds between the high phytate variety and the low phytate variety by a pot culture experiment. Seeds were sown in plastic round pots (length x width x height =8 cm x 6 cm x 7 cm) with holes in the bottom. A plastic bag was placed in the pot to avoid nutrient leaching. Each pot was filled with 7 kg of soil, and 1.4 g N (supplied as $(\text{NH}_4)_2\text{SO}_4$), 1.05 g P (supplied as KH_2PO_4), 1.40 g K (supplied as KCl), 1.75 g Mg (supplied as $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$) were applied to the soil, and then 7 ml 46 µM H_3BO_3 , 0.32 µM $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 0.77 µM $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$, 9.14 µM $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$, 0.37µM $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$ and 50 µM EDTA-Fe (II) were applied, respectively. Each variety was represented by five plants (biological

replicates). The plants were grown in a nature condition with a movable rain shelter. Seeds were sampled at 15, 20, 25, 30, 35 and 40 days after flowering (DAF), and immediately put in the liquid nitrogen and then stored in the -80 for subsequent RNA extraction. At the mature stage, root, stem, old leaves, node, pod hull and seed were harvested for determination of biomass and total P.

Total RNA was extracted from the developmental series seeds after flowering using a plant RNA purification Kit (Qingke, Beijing, China). cDNA was prepared using cDNA Synthesis Kit (Kangwei, Beijing, China). Primers were listed in Table S5. PCR reaction solution contained 5 μ l Master Mix ABI Prism, 0.2 μ l forward Primer, 0.2 μ l reverse Primer, 1 μ l Template DNA and 3.6 μ l PCR-grade water. The PCR program was as follows: 95 °C for 3 s and 40 cycles of 95 °C for 3 s, followed by 60 °C for 20 s. All of the reactions were performed in four technical repetitions and then the average expression value was calculated. Relative expression levels were evaluated using the 2^{-T} method. *BnaTublin* was used as an internal control for normalization.

Germination experiment and field trials of the *B. napushaplotypes*

A total of 40 accessions from the association panel, including 21 carrying $T_P C_E T_E G_I A_E T_E$ at '*BnaA09g10220ConcHap1*', $T_P C_E A_E G_I A_E$ at '*BnaA09g10220ContHap1*', $T_P C_E A_E A_E$ at '*BnaA09g10230ConcHap1*' and $T_P C_E A_E A_E C_E A_E$ at '*BnaA09g10230ContHap1*' on A09 chromosome and 19 accessions carrying $G_P T_E C_E C_I C_E C$ at '*BnaA09g10220ConcHap2*', $G_P T_E C_E C_I C_E$ at '*BnaA09g10220ContHap2*', $G_P T_E C_E C_E$ at '*BnaA09g10230ConcHap2*' and $G_P T_E C_E C_E T_E C$ at '*BnaA09g10230ContHap2*' on A09 chromosome were used for the germination experiment. "P", "E" and "I" each SNP in a haplotype indicated the SNP located in the promoter region, exon region, and the intron region, respectively. 100 plump seeds were disinfected with 1% NaClO for 5 min, and then washed 6 times with pure water and soaked in pure water overnight at 4 °C. The next day, in a dark environment, the seeds were sown on the Petri dish. Seed germination rate was scored on the seventh day after sowing as number of germinated seeds / 100. The 40 accessions were also grown in the field with a P sufficient supply (90 kg/ha) and a P deficient supply (0 kg/ha) with three replications at Meichuan Town, Wuxue city, Hubei province, China (E 115.55 °, N 29.85 °) from 2018 to 2020. The two years of field trails involved a randomized block design. Each accession had 4 rows and each plot with 8 plants each row. Four plants each accession were selected to measure plant height (PH) and effective branch number (EBN) at the mature stage. After harvest, the seed yield and thousand seed weight were investigated.

Determination of total P concentration in plants

Whole plants were dried at 65 °C and then ground to a powder using porcelain mortar. 0.025 g sample (root, node, stem, pod hull, seed and old leaves) was put into a digestive tube, and then 1 ml $H_2SO_4-H_2O_2$ was added to digest the sample until was a clarified solution, which was then diluted to 20 ml, and the P concentration measured by flow injection analyzer (SEAL AA3).

Determination of Pi concentration in plants

The malachite green method was used to determine inorganic P (Pi). The method involved: adding 2 steel balls, 0.025 g fresh sample (root, node, stem, pod hull, seed and old leaves) and 25 μ l H_2SO_4 into 2 ml centrifugal tube, and then ground three times for 20 s each time. 1 ml ddH₂O was added to the treated sample, and then centrifuged at 1000 rpm for 10 min. The supernatant was mixed with malachite green solution of 3 times of the supernatant, shaken well and left to stand for 30 min. The OD value at 650 nm was determined by microplate reader. Pi concentration (mg / g) = (OD value \times 1.5 ml) / (600 μ L \times 0.001 \times M) \times C \times 0.001. 'M' is the sample weight and 'C' is the dilution.

RESULTS

Phenotypic variation in the phytate concentration and content in the seed of a *B. napus* association panel

Seed phytate concentration and content was determined for the association panel of 505 *B. napus* accessions

grown in 2013-2014 and 2015-2016 field trials at P sufficient supply. Both phytate traits in this panel showed extensive phenotypic variation with an approximately normal distribution (Figure 1). For example, seed phytate concentration ranged from 17.88 to 38.37 mg/g (2.1 fold) in the 2013-2014 field trial and from 12.49 to 57.31 mg/g (4.6 fold) in 2015-2016. Variation in phytate content was more consistent between years, ranging from 0.42 to 0.85 mg/ 5 seeds (2.0 fold) in 2013-2014 and ranged from 0.41 to 0.97 mg/ 5 seeds (2.4 fold) in 2015-2016 (Figure 1).

Population structure, relative kinship and LD decay

3.28 million SNP markers were identified for the *B. napus* association panel (Tang et al., 2020). SNP number on each chromosome ranged from 86,775 on A08 to 296,088 on C03 (Figure S3a, Table S3). LD decay on each chromosome ranged from 94 kb on C05 to 1000 kb on C01, when r^2 was 0.1 (Figure S3c, Figure S3d, Table S3). Overall LD decay was 228 kb, and for each genome 135 kb (A genome) and 369 kb (C genome) (Figure S3b, Table S3). To avoid over-adjustment, a total of 412,141 SNPs were selected to assess population structure, relative kinship and LD. The population could be divided into five subgroups based on the cross validation (CV) errors (Figure S1). The r pairwise relative kinship was close to 0 (Figure S2a, Figure S2b, Table S2). For example, the values of the relative kinships were 0 to 142,342 pairs and 0.1 to 204,811 pairs, where the ratios to the total value were 55.8% and 80.31%, respectively (Figure S2b, Table S2). These results indicated that the majority of the accessions had a wide distribution of the genetic relationships and the association panel was suitable for the GWAS.

Genome wide association mapping of phytate concentration and content in *B. napus* seed

We chose the combination of GLM and MLM models for GWAS analysis. A total of 119 SNPs was identified to be significantly associated with phytate concentration and content ($P < 6.25 \times 10^{-07}$) across two years (Figure S4, Figure S5, Table S4). Among the 119 SNPs, 31 were identified in 2013-2014, 59 in 2015-2016 and 40 were identified by using BLUP values (Figure S4, Figure S5, Table S4). The GLM analysis detected a total of 119 SNPs significantly associated with the two phytate-related traits, distributed on 16 of the 19 *B. napus* chromosomes (excluding A01, A05 and C04). Chromosome A09 had the largest number of significant SNPs (35) and C03 the second largest number significant SNPs (10 SNPs) (Table S4). MLM analysis detected 19 SNPs significantly associated with seed phytate concentration and content on seven of the 19 *B. napus* chromosomes (Table S4). Of these, 19 significant SNPs were identified simultaneously by the GLM and MLM models (Table S4). Additionally, 9.24% (11/119) of the significant SNPs were identified in more than one year (including BLUP), which showed high reliability (Table S4).

Of the 119 significant SNPs, eight were detected simultaneously for phytate concentration and content (Table S4). For example, the significant SNP marker (chrA09.5886580) on chromosome A09 could explain 11.76% and 11.85% of the phenotypic variation (PVE) of phytate concentration and phytate content, respectively. (Figure 3b, Table S4).

Candidate genes regulating phytate concentration and content in the seed of *B. napus* by GWAS and gene expression analyses

Thirty candidate genes were within the confidence interval of the SNPs simultaneously identified by GLM and MLM models for seed phytate concentration and content (Figure 2a). The transcription of these genes was quantified by qRT-PCR in the developmental series seed after flowering in the high phytate cultivar HPA972 and the low phytate cultivar LPA20 (Figure 2a). The transcription of *BnaA09g39130D* and *BnaA10g23000D* in the seed of HPA972 at 25 days after flowering (DAF) was relatively higher than that of LPA20 (Figure 2b, Figure 2g). In addition, the transcription of *BnaA09g10220D* and *BnaA09g10230D* in the seed of LPA20 at 30, 35, and 40 DAF was relatively higher than that of HPA972 (Figure 2d, Figure 2e). The relative gene expression level of *BnaC05g33690D* in the seed of LPA20 at 20, 25 and 30 DAF were higher than that of HPA972 (Figure 2c), and the relative gene expression level of *BnaA03g31650D* in the seed of LPA20 at 15, 20, 25, 30, 35, and 40 DAF were higher than that of HPA972 (Figure 2f). These genes could be closely associated with phytate synthesis in the seed of *B. napus*.

Candidate gene association analysis and haplotype analysis of *BnaA9.MRP5*

For seed phytate concentration, the significant SNP ‘chrA09_5198034’ on chromosome A09 was detected in BLUP and explained 8.33% of the phenotypic variations (Figure 3a, Table S4). chrA09_5198034 was located 27 kb from gene the Darmor-*bzh* reference gene models *BnaA09g10220D* and *BnaA09g10230D* (*BnaA9.MRP5*). For seed phytate concentration the significant SNP ‘chrA09_4871583’ on chromosome A09 was detected in the 2013 trial and explain 10.05% of the phenotypic variations, and for phytate content 9.35% (Figure 3a, Table S4). chrA09_4871583 was located 300 kb from gene models *BnaA09g10220D* and *BnaA09g10230D* (*BnaA9.MRP5*). In Arabidopsis, *AtMRP5* is a high affinity transporter of phytate orthologous to *BnaA9.MRP5*. *AtMRP5* is not only involved in the storage of phytate, but also associated with guard cell signaling in response to ABA (Nagy et al., 2009).

Fifty seven SNPs located within the 2 kb promoter region and the entire coding region of *BnaA09g10220D* and 102 SNPs within the corresponding region of *BnaA09g10230D* were identified and confirmed as being associated with seed phytate concentration and content in the association panel (Table S6-S9). Six SNPs in *BnaA09g10220D* were detected to be significantly associated with phytate concentration (Figure 3c). Further analyses showed that the T allele of chrA09_5238455 and C allele of chrA09_5238855 were associated with low phytate, and the G allele of chrA09_5239266 and A allele of chrA09_5239963 were associated with low phytate (Figure 4a). Two major haplotypes were detected, with Hap1 (T_PC_ET_EG_IA_ET_E) cultivars having much lower seed phytate concentration than cultivars with Hap2 (G_PT_EC_EC_IC_EC_E), at a P-value of 0.0008 (Figure 3d). Five SNPs in *BnaA09g10220D* were detected to be significantly associated with phytate content (Figure 3e). We observed that the T allele of chrA09_5238455, C allele of chrA09_5238855, A allele of chrA09_5238924, G allele of chrA09_5239266 and A allele of chrA09_5239963 were low phytate alleles (Figure 4b).

Haplotype analysis of the five associated SNPs also revealed two major haplotypes, and cultivars with Hap1 (T_PC_EA_EG_IA_E) had a lower phytate content in the seed than cultivars with Hap2 (G_PT_EC_EC_IC_E) with a P-value of 0.0019 (Figure 3f). Four SNPs in *BnaA09g10230D* were associated with the phytate concentration in the seed of the association panel (Figure 3g). Further analyses showed that the T allele of chrA09_5238455, C allele of chrA09_5238855, A allele of chrA09_5238924 and A allele of chrA09_5245895 could be designated as low phytate alleles (Figure 4a). Haplotype analysis of the four significant SNPs revealed two major haplotypes (Figure 3h), with Hap1 (T_PC_EA_EA_E) cultivars having significantly lower phytate concentration in the seed than cultivars with Hap2 (G_PT_EC_EC_E) with a P-value of 0.0013 (Figure 3h). Six SNPs in *BnaA09g10230D* were associated with the phytate content in the seed of the association panel (Figure 3i). Four of the six alleles (T allele of chrA09_5238455, C allele of chrA09_5238855, A allele of chrA09_5238924 and A allele of chrA09_5239963) were the low phytate allele (Figure 4b). In the association panel, two Haps were detected in *BnaA09g10230D*, and cultivars with Hap1 (T_PC_EA_EA_EC_EA_E) had much lower phytate content in the seed than cultivars with Hap2 (G_PT_EC_EC_ET_EC_E) with a P-value of 0.0020 (Figure 3j). Haplotypes of T_PC_ET_EG_IA_ET_E, T_PC_EA_EG_IA_E, T_PC_EA_EA_E and T_PC_EA_EA_EC_EA_E could be considered as favorable haplotypes for reducing phytate content and concentration in the seed of *B. napus*. Taking into account the difference in the expression level of *BnaA9.MRP5* observed in the HPA 972 and LPA 20 (Figure 2d, Figure 2e), the sequence variations in *BnaA9.MRP5* might mainly contribute to the phenotypic difference in phytate concentration and content.

Differences in seed germination, seed yield and yield related traits between cultivars with high and low phytate haplotypes

A seed germination experiment was conducted to establish whether there were significant differences in germination rate between low and high phytate haplotypes. Forty accessions were assessed, 21 carrying T_PC_ET_EG_IA_ET_E at ‘*BnaA09g10220ConcHap1*’, T_PC_EA_EG_IA_E at ‘*BnaA09g10220ContHap1*’, T_PC_EA_EA_E at ‘*BnaA09g10230ConcHap1*’ and T_PC_EA_EA_EC_EA_E at ‘*BnaA09g10230ContHap1*’ on A09 chromosome and 19 accessions carrying G_PT_EC_EC_IC_EC_E at ‘*BnaA09g10220ConcHap2*’, G_PT_EC_EC_E at ‘*BnaA09g10230ConcHap2*’ and G_PT_EC_EC_ET_EC_E at ‘*BnaA09g10230ContHap2*’ on A09 chromosome. In high phytate haplotype cultivars,

seed phytate concentration ranged from 27.60 mg/g to 43.40 mg/g and seed phytate content from 0.57 mg/5 seeds to 0.85 mg/5 seeds. This contrasted with low phytate haplotype cultivars, where seed phytate concentration ranged from 23.57 mg/g to 30.20 mg/g and content from 0.49 mg/5 seeds to 0.71 mg/5 seeds (Table S10). No significant correlation was detected between germination rate and seed phytate concentration ($r=0.1617$) or content ($r=0.1283$) (Figure 5b, Figure 5c). Moreover, some LPA cultivars showed a high germination rate of 70% - 95% (Figure 5a, Table S10). Together, this indicated that the low phytate in the seed did not affect the germination in this study.

Field trials were conducted to establish whether there were significant differences in seed yield and yield related traits between low and high phytate haplotypes at a sufficient P supply and a deficient P supply. No significant difference in seed yield, thousand -seed weight, plant height and effective branch number were detected between cultivars with low phytate and high phytate haplotypes under both P supplies, except for the effective branch number at a sufficient P supply in 2019-2020 (Figure 6, Table S11). In addition, some low phytate varieties had high seed yield, which could be useful for low phytate *B. napus* breeding. For example, the seed phytate concentration of the low phytate cultivar LPA20 was 0.48 mg/g and phytate content 23.57 mg/5 seeds (Table S1), although seed yield in 2019 and in 2020 for this cultivar was 15.71 g/plant and 16.68 g/plant, respectively at a sufficient P supply (Table S11).

There was no significant difference in thousand -seed weight between the high phytate cultivar of HPA972 and the low phytate cultivar LPA20. However, the phytate concentration and content in HPA972 carrying $G_P T_E C_E C_1 C_E$ in '*BnaA09g10220ConcHap2*', $G_P T_E C_E C_1 C_E$ in '*BnaA09g10220ContHap2*', $G_P T_E C_E C_E$ at '*BnaA09g10230ConcHap2*' and $G_P T_E C_E C_E T_E C_E$ at '*BnaA09g10230ContHap2*' on A09 chromosome showed much more phytate than LPA20 carrying $T_P C_E T_E G_1 A_E T_E$ at '*BnaA09g10220ConcHap1*', $T_P C_E A_E G_1 A_E$ at '*BnaA09g10220ContHap1*', $T_P C_E A_E A_E$ at '*BnaA09g10230ConcHap1*' and $T_P C_E A_E A_E C_E A_E$ at '*BnaA09g10230ContHap1*' (Table 1). The biomass and P concentration of root, lateral branches and lateral branches pod hull of HPA972 were significantly higher than that for LPA20 (Figure 7a, Figure 7b, Figure 7c). At harvest stage 59.6% of total P was distributed in the seed of HPA972 and 75.0% in LPA20. (Figure 7 f, g). No significant difference in seed Pi concentration was detected between HPA972 and LPA20 at 20, 25, 30 and 40 DAF. At 35 DAF, the seed Pi concentration of LPA20 was significantly higher than that HPA972 (Figure 7e). Seed phytate concentration of HPA972 was higher than that of LPA20 during the whole period of seed development, with the most significant difference between the cultivars at 30 DAF (Figure 7d). Transcriptional level of *BnaA09g39130D*, *BnaA09g10220D*, *BnaA09g10230D* and *BnaA10g23000D* in both varieties was higher between 20 -30 DAF than during 15 -25 DAF (Figure 2a). Only the transcription of *BnaA10g23000D* in HPA972 seed was higher than that of LPA20 during seed development (Figure 2g), which were in agreement with that the difference in seed phytate concentration between HPA972 and LPA20.

DISCUSSION

Variation in the phytate concentration and content in *B. napus*

In this study, a modified high throughput method using 5 seeds per sample was used to determine the phytate concentration in the seed of a large association panel of *B. napus* cultivars. Phytate concentration and content in the association panel of *B. napus* at a sufficient P supply showed continuous variation and an approximate normal distribution across the panel of 505 cultivars (Figure 1). Seed phytate concentration had 2.31 and 4.6 fold variations in 2013-2014 and in 2015-2016, respectively (Figure 1). Seed phytate content had 2.0 and 2.4 fold variations in 2013-2014 and in 2015-2016, respectively (Figure 1). This compares with 2.8 fold variation (7.91 to 21.97 mg/g) in phytate concentration in a 60 chickpea accessions (Joshi-Saha and Reddy, 2015), and 2.2 fold (8.18 to 18.04 mg/) among 69 rice accessions (Perera et al., 2019). Since the environment had a large effect on seed phytate concentration and content (Shane et al., 2003; Oltmans et al., 2005), BLUP analysis was adopted to reduce these effects on the two traits. The mean phenotypic values each year and the BLUP values were all used for genome wide association study.

The decrease of Pi and the increase of phytate during seed development appears to be similar in different

crops (Nakamura et al., 1989; Matthew et al 2005., Iwai et al., 2012). In the initial stage of maize seed formation, seed Pi maintains a high level. When the rate of phytate synthesis in the seed increases, the concentration of Pi decreases significantly. The decreased Pi is mainly used for the synthesis and storage of phytate. In this study, the concentration of Pi in the seed of HPA972 (a high phytate variety) and LPA20 (a low phytate variety) decreased slowly between 20-30 DAF, and phytate concentration began to increase at 20 DAF (Figure 7d, Figure 7e). At 35 DAF, the Pi concentration in the seed of LPA20 was significantly higher than that of HPA972, which indicated that the Pi consumption in the seed of LPA972 was higher than that of LPA20 (Figure 7e). Seed phytate concentration in HPA972 was always higher than that of LPA20 at any stage of seed development (Figure 7d). The difference in seed phytate concentration between HPA972 and LPA20 reached a maximum at 30 DAF (Figure 7d). Therefore, 25-30 DAF could be the key time point for the difference of phytate synthesis in the seed between HPA972 and LPA20. Most of the candidate genes underlying the significant SNPs for phytate concentration and content in the seed showed higher gene expression level at 30 DAF than at 15, 20 and 25 DAF (Figure 2a). This is consistent with the increase in phytate concentration and the decrease in Pi concentration in the seed of high and low phytate varieties (Figure 7d, Figure 7e). The seed P content accounted for 75% of the total P in LPA20, compared with only 59% of the total P in HPA972 (Figure 7f, Figure 7g). HPA972 therefore appeared to have a stronger P uptake ability and higher total P content than LPA20. However, the former had less ability to distribute P to the seed than the latter. There were no significant differences in total P content in the seed between LPA20 and HPA972 (Table 1), although the phytate content of LPA20 was significantly lower than that of HPA972 (Table 1). This indicated that the difference in phytate content between HPA972 and LPA20 was caused by the difference of phytate synthesis process. In low phytate rice mutants Tong (2017) demonstrated an increase in the redistribution of Pi in endosperm led to lysophospholipid biosynthesis. Increased lysophospholipid in the endosperm of low phytate mutants may have practical applications in rice breeding to produce “healthier” rice, which increased the digestive absorption of mineral elements. Similarly, if a greater proportion on Pi is incorporated into phospholipids and less Pi is converted into phytate in *B. napus* seed, then such cultivars would have low phytate and high oil content in the seed. Moreover, although the plant height of LPA20 was reduced compared with HPA972, the seed yield of the primary inflorescence of LPA20 was significantly longer than that of HPA972, which would be more suitable for mechanized harvesting (Figure S6).

Genetic control of phytate related traits and exploration of candidate genes

The LD decay observed in this study (228 kb) was much smaller than that in the previous studies in *B. napus*, such as 298 kb in 300 inbred lines (Zhou et al., 2017), 1.06 cM in 404 inbred lines (Wang et al., 2017), 2.0 Mb in 472 inbred lines (Li et al., 2014), and 0.5–1 cM in 192 inbred lines (Xiao et al., 2012), although larger than that of other species e.g. 200 kb in *Arabidopsis* (Wang et al., 2020), 187 kb in rice (Yang et al., 2018) and 173 kb in maize (Luo et al., 2019).

A total of 119 significant SNPs associated with seed phytate concentration and content were detected on 16 of the 19 *B. napus* chromosomes, accounting for an average of 8.81% phenotypic variation for phytate concentration ranging from 6.48% to 11.92% (Table S4). A previous study had shown determining the level of transcription of candidate genes in the seed at different time points after flowering helped to identify candidate genes (Sashidhar et al., 2020a; Sashidhar et al., 2020b). In this study, the genes related to phytate concentration and content within the LD decay region (~300 kb) of each significant SNP in *B. napus* were screened. The transcriptional profile of these genes was employed to help narrow down the list of candidate genes, and finally 30 candidates were identified. These genes included the ABC transporter *MRP5*, phospholipid cytidylate transferase CDs, cation transport protein, vacuolar protein sorting, etc (Figure 2a). Of these, both *BnaA09g10220D* and *BnaA09g10230D* showed homology to *AT1G04120* (*AtMRP5*), which in *A. thaliana* has been related to the storage of phytate and associated with guard cell signaling in response to ABA (Nagy et al., 2009). *BnaC07g45670D* is homologous to *AT4G35530*, which is associated with phosphatidyl inositol glycan. *BnaA10g23000D* in *B. napus* was orthologous to *A. thaliana* *AT5G08460*, which catalyzes acyl transfer or hydrolase reactions with lipid and non-lipid substrates (Lai et al., 2017). *BnaC05g33690D* was orthologous to *AT3G19240*, which is involved in vacuolar import/degradation (Wang et al., 2008). *BnaA03g31650D* is orthologous to *AT3G11320*, which is involved in nucleotide-sugar transport

(Parsons et al., 2012). In addition, these genes all showed significant differential expressed in the seed between high phytate and low phytate varieties at 15, 20, 25, 30, 35 and 40 DAF (Figure 2a), which may suggest a close involvement in the seed phytate biosynthesis.

Breeding of low phytate cultivars in *B. napus*

P in crop seed is the main driving force of the global P cycle, and waste of finite P resources can be reduced by reducing the concentration of phytate in seeds (Raboy et al., 2020). In previous studies, low phytate (*lpa*) mutants have been generated by random mutagenesis in various crops. Such mutants are most often associated with lower thousand seed weight and viability than the wild type, reducing breeding value (Meis et al., 2003; Oltmans et al., 2005; Raboy, 2007; Su et al., 2018). In our study, some low phytate cultivars had a higher germination rate, and some high phytate cultivars showed low germination rate. In general, we found no significant correlation between seed germination rate and phytic acid concentration or phytic acid content in the selected varieties (Figure 5a-c, Table S10). Recently, CRISPR-Cas9 mutagenesis has been adopted to knock out three essential *BnITPK* genes to obtain a low phytate line (Sashidhar et al., 2020). Compared with wild type, the low phytate mutant had no negative effect on germination rate and seedling vigor (Sashidhar et al., 2020b). This result is consistent with our findings and suggests there is a valuable pool of natural alleles that confer similar function. Haplotype analysis of *BnaA9.MRP5*, a candidate gene underlying the significant SNP for phytate concentration and content, identified the high and low phytate haplotypes (Figure 3c-j, Table S6-S9). There were no significant difference in seed yield, plant height, effective branch number, or thousand seed weight between cultivars with low and high phytate haplotypes, either under a sufficient or deficient P supply (Figure 6). Marzoz et al (1997) reported that lower phytate and protein contents increase the risk of *Bruchus* weevil infestation in pea seeds. In Mung beans, if the phytate concentration in the seed is reduced below 18 mg/g the tolerance to biotic stresses decreased, even with resistance genes present in the plant (Dhole et al., 2016). Any obvious difference in seed germination, disease resistance and insect resistance between the two haplotypes of *B. napus* were not observed in the field in this study. Therefore, the low phytate haplotypes represent an important genetic resource for breeding low phytate varieties without negative effects on seed germination and seed yield.

In conclusion, phytate concentration and content in the seed of *B. napus* is a complex quantitative trait controlled by multiple genes. A total of 119 SNPs were identified to be significantly associated with phytate concentration and content in the seed across two years, and eight of them were detected simultaneously for seed phytate concentration and content. No significant differences in the seed germination and seed yield and yield-related traits were found among low and high phytate crops examined. The discovery of the low phytate haplotype of T_PC_ET_EG_IA_ET_E at '*BnaA09g10220ConcHap1*', T_PC_EA_EG_IA_E at '*BnaA09g10220ContHap1*', T_PC_EA_EA_E at '*BnaA09g10230ConcHap1*' and T_PC_EA_EA_EC_EA_E at '*BnaA09g10230ContHap1*' could enable the accurate selection of *B. napus* oilseed cultivars with lower phytate content and lacking any negative effect on seed germination and seed yield.

ACKNOWLEDGEMENTS

This work was supported by the National Key R&D Program of China (Grant No. 2017YFD0200200) and 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) and the Natural and Fundamental Research Funds for the Central Universities of China (Grant No. 2662019PY013).

Conflict interest

The authors declare there is no conflict interest.

Author contributions

L.S. and G.J.K. conceived and designed the research study. P.Y., G.D.D., S.L.W., H.M.C., W.C., F.S.X and L.S supervised the experiments. H.J.L., X.J.L and Q.W.Z. conducted fieldwork and analyzed data. H.J.L., L.S., L.L. and G.J.K. wrote the manuscript.

References

- Alexander, D.H., Novembre, J. and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* . 19(9), 1655-1664.
- Ali, N., Paul, S., Gayen, D., Sarkar, S.N., Datta, K. and Datta, S.K. (2013). Development of low phytate rice by RNAi mediated seed-specific silencing of inositol 1,3,4,5,6-pentakisphosphate 2-kinase gene (IPK1). *PLoS One* . 8(7), e68161.
- Ali, N., Paul, S., Gayen, D., Sarkar, S.N., Datta, S.K. and Datta, K. (2013). RNAi mediated down regulation of myo-inositol-3-phosphate synthase to generate low phytate rice. *Rice* . 6(1), 1-12.
- Barrett, J.C., Fry, B., Maller, J. and Daly, M.J. (2005). Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* . 21(2), 263-265.
- Bentsink, L., Yuan, K., Koornneef, M. and Vreugdenhil, D. (2003). The genetics of phytate and phosphate accumulation in seeds and leaves of *Arabidopsis thaliana*, using natural variation. *Theor Appl Genet* . 106(7), 1234-1243.
- Berridge, M.J. and Irvine, R.F. (1989). Inositol phosphates and cell signalling. *Nature* . 341(6239), 197-205.
- Blair, M.W., Sandoval, T.A., Caldas, G.V., Beebe, S.E. and Paez, M.I. (2009). Quantitative Trait locus analysis of seed P and seed phytate content in a recombinant inbred line population of common Bean. *Crop Science* . 49(1), 237-246.
- Boominathan, P., Shukla, R., Kumar, A., Manna, D., Negi, D., Verma, P.K. and Chattopadhyay, D. (2004). Long term transcript accumulation during the development of dehydration adaptation in *Cicer arietinum*. *Plant Physiol* . 135(3), 1608-1620.
- Bradbury, P.J., Zhang, Z., Kroon, D.E., Casstevens, T.M., Ramdoss, Y. and Buckler, E.S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* . 23(19), 2633-2635.
- Brown, K. H. and Solomons, N.W. (1991). Nutritional problems of developing countries. *Infect Dis Clin North Am* . 5, 297.
- Chalhoub, B., Denoeud, F., Liu, S., Parkin, I.A.P., Tang, H., Wang, X., Chiquet, J. et al. (2014). Early allopolyploid evolution in the post-Neolithic *B.napus* oilseed genome. *Science* . 345(6199), 950-953.
- Dhole, V.J. and Reddy, K.S. (2016). Association of phytate content with biotic stress tolerance in mungbean (*Vigna radiata* L. Wilczek). *Phytoparasitica* . 44(2), 261-267.
- Emilsson, V., Thorleifsson, G., Zhang, B., Leonardson, A.S., Zink, F., Zhu, J., Carlson, S., Helgason, A., Walters, G.B., Gunnarsdottir, S., Mouy, M., Steinthorsdottir, V., Eiriksdottir, G.H., Bjornsdottir, G., Reynisdottir, I., Gudbjartsson, D., Helgadottir, A., Jonasdottir, A., Jonasdottir, A., Styrkarsdottir, U., Gretarsdottir, S., Magnusson, K.P., Stefansson, H., Fossdal, R., Kristjansson, K., Gislason, H.G., Stefansson, T., Leifsson, B.G., Thorsteinsdottir, U., Lamb, J.R., Gulcher, J.R., Reitman, M.L., Kong, A., Schadt, E.E. and Stefansson, K. (2008). Genetics of gene expression and its effect on disease. *Nature* . 452(7186), 423-428.
- Iwai, T., Takahashi, M., Oda, K., Terada, Y. and Yoshida, K.T. (2012). Dynamic changes in the distribution of minerals in relation to phytate accumulation during rice seed development. *Plant Physiol* . 160(4), 2007-2014.
- Joshi-Saha, A. and Reddy, K.S. (2015). Repeat length variation in the 5'UTR of myo-inositol monophosphatase gene is related to phytate content and contributes to drought tolerance in chickpea (*Cicer arietinum* L.). *J Exp Bot* . 66 (19), 5683-5690.
- Kaur, H., Shukla, R.K., Yadav, G., Chattopadhyay, D. and Majee, M. (2008). Two divergent genes encoding L-myo-inositol 1-phosphate synthase1 (CaMIPS1) and 2 (CaMIPS2) are differentially expressed in chickpea. *Plant Cell Environ* . 31(11), 1701-1716.

- Khan, M.S.S., Basnet, R., Islam, S.A. and Shu, Q. (2019). Mutational Analysis of OsPLDalpha1 Reveals Its Involvement in Phytic Acid Biosynthesis in Rice Grains. *J Agric Food Chem* . 67(41), 11436-11443.
- Kishor, D.S., Lee, C., Lee, D., Venkatesh, J., Seo, J., Chin, J.H., Jin, Z., Hong, S.K., Ham, J.K. and Koh, H.J. (2019). Novel allelic variant of Lpa1 gene associated with a significant reduction in seed phytic acid content in rice (*Oryza sativa* L.). *PLoS One* . 14(3), e0209636.
- Kochian, L.V. (2012) Plant nutrition: Rooting for more P. *Nature*488. 466-467.
- Laha, D., Parvin, N., Hofer, A., Giehl, R.F.H., Fernandez-Rebollo, N., von Wiren, N., Saiardi, A., Jessen, H.J. and Schaaf, G. (2019). Arabidopsis ITPK1 and ITPK2 have an evolutionarily conserved phytic acid kinase activity. *ACS Chem Biol* . 14(10), 2127-2133.
- Lai, C.P., Huang, L.M., Chen, L.O., Chan, M.T. and Shaw, J.F. (2017). Genome-wide analysis of GDSL-type esterases/lipases in Arabidopsis. *Plant Mol Biol* . 95(1-2), 181-197.
- Lee, H.S., Lee, D.H., Cho, H.K., Kim, S.H., Auh, J.H. and Pai, H.S. (2015). InsP 6 -sensitive variants of the gle1 mRNA export factor rescue growth and fertility defects of the ipk1 low-phytic-acid mutation in Arabidopsis. *Plant Cell Online* . 27(2), 417-431.
- Li, F., Chen, B., Xu, K., Wu, J., Song, W., Bancroft, I., Harper, A.L., Trick, M., Liu, S., Gao, G., Wang, N., Yan, G., Qiao, J., Li, J., Li, H., Xiao, X., Zhang, T. and Wu, X. (2014a). Genome-wide association study dissects the genetic architecture of seed weight and seed quality in rapeseed (*B.napus* L.). *DNA Res* . 21(4), 355-367.
- Li, W.X., Huang, J.Z., Zhao, H.J., Tan, Y.Y., Cui, H.R., Poirier, Y. and Shu, Q.Y. (2014b). Production of low phytate rice by hairpin RNA- and artificial microRNA-mediated silencing of OsMIK in seeds. *Plant Cell, Tissue and Organ Culture* . 119(1), 15-25.
- Li, W.X., Zhao, H.J., Pang, W.Q., Cui, H.R., Poirier, Y. and Shu, Q.Y. (2014c). Seed-specific silencing of OsMRP5 reduces seed phytate and weight in rice. *Transgenic Res* . 23(4), 585-599.
- Lu, K., Wei, L., Li, X., Wang, Y., Wu, J., Liu, M., Zhang, C., Chen, Z., Xiao, Z., Jian, H., Cheng, F., Zhang, K., Du, H., Cheng, X., Qu, C., Qian, W., Liu, L., Wang, R., Zou, Q., Ying, J., Xu, X., Mei, J., Liang, Y., Chai, Y.R., Tang, Z., Wan, H., Ni, Y., He, Y., Lin, N., Fan, Y., Sun, W., Li, N.N., Zhou, G., Zheng, H., Wang, X., Paterson, A.H. and Li, J. (2019). Whole-genome resequencing reveals *B.napus* origin and genetic loci involved in its improvement. *Nat Commun* . 10(1), 1154.
- Luo, X., Wang, B., Gao, S., Zhang, F., Terzaghi, W. and Dai, M. (2019). Genome-wide association study dissects the genetic bases of salt tolerance in maize seedlings. *J Integr Plant Biol* . 61(6), 658-674.
- Marisa, S.O., Roberta, C.L. and Andrew, S. (2002). Developing seeds of Arabidopsis store different minerals in two types of vacuoles and in the endoplasmic reticulum. *Plant Cell* . 14(6), 1311-1327.
- Marzo, F., Andres, A., Castiella, M.V. and Alonso, R. (1997). Fertilization effects of phosphorus and sulfur on chemical composition of seeds of *Pisum sativum* L. and relative infestation by *Bruchus pisorum* L. *J Agric Food Chem* . 45 , 1829-1833.
- Meis, S.J., Fehr, W.R. and Schnebly, S.R. (2003). Seed source effect on field emergence of soybean lines with reduced phytate and raffinose saccharides. *Crop Sci* . 43 , 1336-1339.
- Morell, M.K. and Myers, A.M. (2005). Towards the rational design of cereal starches. *Curr Opin Plant Biol* . 8(2), 204-210.
- Nagy, R., Grob, H., Weder, B., Green, P., Klein, M., Frelet, A., Schjoerring, J.K., Brearley, C. and Martinoia, E. (2009). The Arabidopsis ATP-binding cassette protein AtMRP5/AtABCC5 is a high affinity inositol hexakisphosphate transporter involved in guard cell signaling and phytate storage. *J Biol Chem* . 284(48), 33614-33622.

- Nakamura, Y., Yuki, K. and Park, S.Y. (1989). Carbohydrate metabolism in the developing endosperm of rice grains. *Plant Cell Physiol* . 1898, 833-839.
- Nordborg, M. and Weigel, D. (2008). Next-generation genetics in plants. *Nature* . 456(7223), 720-723.
- Oltmans, S.E., Fehr, W.R., Welke, G.A., Raboy, V. and Peterson, K.L. (2005). Agronomic and seed traits of soybean lines with low-phytate P. *Crop Sci* . 45, 593-598.
- Paran, I. and Zamir, D. (2003). Quantitative traits in plants: beyond the QTL. *Trends in Genet* . 19(6), 303-306.
- Parsons, H.T., Christiansen, K., Knierim, B., Carroll, A., Ito, J., Batth, T.S., Smith-Moritz, A.M., Morrison, S., McInerney, P., Hadi, M.Z., Auer, M., Mukhopadhyay, A., Petzold, C.J., Scheller, H.V., Loque, D. and Heazlewood, J.L. (2012). Isolation and proteomic characterization of the Arabidopsis Golgi defines functional and novel components involved in plant cell wall biosynthesis. *Plant Physiol* . 159(1), 12-26.
- Perera, I., Fukushima, A., Arai, M., Yamada, K., Nagasaka, S., Seneweera, S. and Hirotsu, N. (2019). Identification of low phytate and high Zn bioavailable rice (*Oryza sativa* L.) from 69 accessions of the world rice core collection. *Journal of Cereal Science* .85 , 206-213.
- Prince, S.J., Valliyodan, B., Ye, H., Yang, M., Tai, S., Hu, W., Murphy, M., Durnell, L.A., Song, L., Joshi, T., Liu, Y., Vandepoele, K., Grover Shannon, J. and Nguyen, H.T. (2019). Understanding genetic control of root system architecture in soybean: Insights into the genetic basis of lateral root number. *Plant Cell Environ* . 42(1), 212-229.
- Raboy, V. (2020). Low phytate crops: Observations based on four decades of research. *Plants* . 9(2) 1-27.
- Raboy, V., Gerbasi, P.F., Young, K.A., Stoneberg, S.D., Pickett, S.G., Bauman, A.T., Murthy, P.P., Sheridan, W.F. and Ertl, D.S. (2000). Origin and Seed Phenotype of Maize low phytate 1-1 and low phytate 2-1. *Plant Physiol* . 124(1), 255-368.
- Raboy, V. and Dickinson D.B. (1984). Effect of P and zinc Nutrition on soybean seed phytic acid and zinc. *Plant Physiol* . 75(4), 1094-1098.
- Ren, D., Wang, X., Yang, M., Yang, L., He, G. and Deng, X.W. (2019). A new regulator of seed size control in Arabidopsis identified by a genome-wide association study. *New Phytol*, 222(2), 895-906.
- Rose, T.J., Rose, M.T., Pariasca-Tanaka, J., Heuer, S. and Wissuwa, M. (2011). The Frustration with Utilization: Why Have Improvements in Internal P Utilization Efficiency in Crops Remained so Elusive? *Front Plant Sci* . 2, 73.
- Sashidhar, N., Harloff, H.J. and Jung, C. (2020a). Identification of phytate mutants in oilseed rape (*B.napus*) by large-scale screening of mutant populations through amplicon sequencing. *New Phytol* . 225(5), 2022-2034.
- Sashidhar, N., Harloff, H.J., Potgieter, L. and Jung, C. (2020b). Gene editing of three BnITPK genes in tetraploid oilseed rape leads to significant reduction of phytate in seeds. *Plant Biotechnol J* . 1, 1-10.
- Saxena, S.C., Salvi, P., Kaur, H., Verma, P., Petla, B.P., Rao, V., Kamble, N. and Majee, M. (2013). Differentially expressed myo-inositol monophosphatase gene (CaIMP) in chickpea (*Cicer arietinum* L.) encodes a lithium-sensitive phosphatase enzyme with broad substrate specificity and improves seed germination and seedling growth under abiotic stresses. *J Exp Bot* . 64(18), 5623-5639.
- Shane, J., Meis, W.R. and Fehr, S.R. (2003). Seed source effect on field emergence of soybean lines with reduced phytate and raffinose saccharides. *Crop Sci* . 43, 1336-1339.
- Shears, S.B. (2004). How versatile are inositol phosphate kinases?. *Biochem J* . 377, 265-280.

- Shi, J., Wang, H., Schellin, K., Li, B., Faller, M., Stoop, J.M., Meeley, R.B., Ertl, D.S., Ranch, J.P. and Glassman, K. (2007). Embryo-specific silencing of a transporter reduces phytate content of maize and soybean seeds. *Nat Biotechnol.* 25(8), 930-937.
- Shi, J., Wang, H., Wu, Y., Hazebroek, J., Meeley, R.B. and Ertl, D.S. (2003). The maize low-phytate mutant lpa2 is caused by mutation in an inositol phosphate kinase gene. *Plant Physiol.* 131(2), 507-515.
- Sompong, U., Somta, P., Raboy, V. and Srinives, P. (2012). Mapping of quantitative trait loci for phytate and P contents in seed and seedling of mungbean (*Vigna radiata* (L.) Wilczek). *Breed Sci.* 62(1), 87-92.
- Stangoulis, J.C.R., Huynh, B.L., Welch, R.M., Choi, E.Y. and Graham, R.D. (2006). Quantitative trait loci for phytate in rice grain and their relationship with grain micronutrient content. *Euphytica.* 154(3), 289-294.
- Stewart, W.M., Dibb, D.W., Johnston, A.E. and Smyth, T.J. (2005). The contribution of commercial fertilizer nutrients to food production. *Agronomy J.* 97, 1-6.
- Su, D., Zhou, L., Zhao, Q., Pan, G. and Cheng, F. (2018). Different P supplies altered the accumulations and quantitative distributions of phytic acid, zinc, and iron in rice (*Oryza sativa* L.) Grains. *J Agric Food Chem.* 66, 1601-1611.
- Tang, S., Zhao, H., Lu, S., Yu, L., Zhang, G., Zhang, Y., Yang, Q.Y., Zhou, Y., Wang, X., Ma, W., Xie, W. and Guo, L. (2020). Genome- and transcriptome-wide association studies provide insights into the genetic basis of natural variation of seed oil content in *Brassica napus*. *Mol Plant.*
- Tong, C., Chen, Y., Tan, Y., Liu, L., Waters, D.L.E., Rose, T.J., Shu, Q. and Bao, J. (2017). Analysis of lysophospholipid content in low phytate rice mutants. *J Agric Food Chem.* 65(26), 5435-5441.
- Wang, B., Wu, Z., Li, Z., Zhang, Q., Hu, J., Xiao, Y., Cai, D., Wu, J., King, G.J., Li, H. and Liu, K. (2018). Dissection of the genetic architecture of three seed-quality traits and consequences for breeding in *B.napus*. *Plant Biotechnol J.* 16(7), 1336-1348.
- Wang, X., Chen, Y., Thomas, C.L., Ding, G., Xu, P., Shi, D., Grandke, F., Jin, K., Cai, H., Xu, F., Yi, B., Broadley, M.R. and Shi, L. (2017). Genetic variants associated with the root system architecture of oilseed rape (*B.napus* L.) under contrasting phosphate supply. *DNA Res.*24(4), 407-417.
- Wang, X., Ren, M., Liu, D., Zhang, D., Zhang, C., Lang, Z., Macho, A.P., Zhang, M. and Zhu, J.K. (2020). Large-scale identification of expression quantitative trait loci in *Arabidopsis* reveals novel candidate regulators of immune responses and other processes. *J Integr Plant Biol.* doi: 10.1111/jipb.12930.
- Wang, Y., Zhang, W.Z., Song, L.F., Zou, J.J., Su, Z. and Wu, W.H. (2008). Transcriptome analyses show changes in gene expression to accompany pollen germination and tube growth in *Arabidopsis*. *Plant Physiol.* 148(3), 1201-1211.
- Welling, M.T., Liu, L., Kretzschmar, T., Mauleon, R., Ansari, O. and King, G.J. (2020). An extreme-phenotype genome-wide association study identifies candidate cannabinoid pathway genes in *Cannabis*. *Sci Rep* 10(1), 18643.
- White, P.J. and Veneklaas, E.J. (2012). Nature and nurture: the importance of seed P content. *Plant and Soil.* 357(1-2), 1-8.
- Wu, D., Liang, Z., Yan, T., Xu, Y., Xuan, L., Tang, J., Zhou, G., Lohwasser, U., Hua, S., Wang, H., Chen, X., Wang, Q., Zhu, L., Maodzeka, A., Hussain, N., Li, Z., Li, X., Shamsi, I.H., Jilani, G., Wu, L., Zheng, H., Zhang, G., Chalhoub, B., Shen, L., Yu, H. and Jiang, L. (2019). Whole-genome resequencing of a worldwide collection of rapeseed accessions reveals the genetic basis of ecotype divergence. *Mol Plant.* 12(1), 30-43.
- Xiao, Y., Cai, D., Yang, W., Ye, W., Younas, M., Wu, J. and Liu, K. (2012). Genetic structure and linkage disequilibrium pattern of a rapeseed (*B.napus* L.) association mapping panel revealed by microsatellites. *Theor Appl Genet.* 125(3), 437-447.

Xu, X.H., Zhao, H.J., Liu, Q.L., Frank, T., Engel, K.H., An, G. and Shu, Q.Y. (2009). Mutations of the multi-drug resistance-associated protein ABC transporter gene 5 result in reduction of phytate in rice seeds. *Theor Appl Genet* . 119(1), 75-83.

Xuan, L., Yan, T., Lu, L., Zhao, X., Wu, D., Hua, S. and Jiang, L. (2020). Genome-wide association study reveals new genes involved in leaf trichome formation in polyploid oilseed rape (*B.napus* L.). *Plant Cell Environ* . 43(3), 675-691.

Yamaji, N., Takemoto, Y., Miyaji, T., Mitani-Ueno, N., Yoshida, K.T. and Ma, J.F. (2017). Reducing P accumulation in rice grains with an impaired transporter in the node. *Nature* . 541(7635), 92-95.

Yang, M., Lu, K., Zhao, F.J., Xie, W., Ramakrishna, P., Wang, G., Du, Q., Liang, L., Sun, C., Zhao, H., Zhang, Z., Liu, Z., Tian, J., Huang, X.Y., Wang, W., Dong, H., Hu, J., Ming, L., Xing, Y., Wang, G., Xiao, J., Salt, D.E. and Lian, X. (2018). Genome-Wide Association Studies Reveal the Genetic Basis of Ionic Variation in Rice. *Plant Cell* . 30(11), 2720-2740.

Zhang, C., Dong, S.S., Xu, J.Y., He, W.M. and Yang, T.L. (2019). PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics* . 35(10), 1786-1788.

Zhao, J., Jamar, D.C., Lou, P., Wang, Y., Wu, J., Wang, X., Bonnema, G., Koornneef, M. and Vreugdenhil, D. (2008). Quantitative trait loci analysis of phytate and phosphate concentrations in seeds and leaves of *Brassica rapa*. *Plant Cell Environ* . 31(7), 887-900.

Zhao, Z., Zhang, H., Fu, Z., Chen, H., Lin, Y., Yan, P., Li, W., Xie, H., Guo, Z., Zhang, X. and Tang, J. (2018). Genetic-based dissection of arsenic accumulation in maize using a genome-wide association analysis method. *Plant Biotechnol J* . 16(5), 1085-1093.

Zhou, Q., Zhou, C., Zheng, W., Mason, A.S., Fan, S., Wu, C., Fu, D. and Huang, Y. (2017). Genome-Wide SNP Markers Based on SLAF-Seq Uncover Breeding Traces in Rapeseed (*B.napus* L.). *Front Plant Sci* . 8, 648.

Table 1. Phytate concentration and content in the seed of HPA and LPA *B. napus* cultivars

Cultivars	Phytate concentration (mg/g)	Phytate concentration (mg/g)	Phytate content (mg/5 seeds)	Phytate content (mg/5 seeds)	Thousand ;seeds weight (g/1000 seeds)	Thousand ;seeds weight (g/1000 seeds)	Total phosphorus content (mg)	Total phosphorus content (mg)
	2013	2015	2013	2015	2013	2015	2013	2015
HPA972	34.29	35.43	0.757	0.773	3.71	3.52	0.933	0.904
LPA20	23.57	24.48	0.487	0.412	3.87	3.79	0.802	0.812

Note, HPA, high phytate content in the seed; LPA, low phytate content in the seed. HPA972, a high phytate variety; LPA20, a low phytate variety.

Figure 1. Frequency distribution of phytate concentration and content in the seed of an association panel of *B. napus* in 2013 and 2015. (a) Phytate concentration. (b) Phytate content. PA_Conc, phytate concentration; PA_Cont, phytate content.

Figure 2. Difference in the gene expression of candidate genes underlying the significant SNP associated with phytate concentration and content in *B. napus*. HPA972, a high phytate variety; LPA20, a low phytate variety. (a) Heatmap of the expression pattern of 30 candidate genes associated with phytate synthesis and transport. (b) - (g) Differences in the gene expression of *BnaA09g39130D* (b), *BnaC05g33690D* (c), *BnaA09g10220D* (d), *BnaA09g10230D* (e), *BnaA03g31650D* (f), *BnaA10g23000D* (g) in the seed development between LPA20 and HPA792. DAF, day after flowering. Data represent the mean \pm SD of four

replicates in (b), (c), (d), (e), (f) and (g). Asterisks indicate the significance of Student's t test (* $P < 0.1$, ** $P < 0.01$, *** $P < 0.0001$, **** $P < 0.0001$).

Figure 3. The co-localized locus and haplotypes on chromosome A09 associated with phytate concentration and content of *B.napus* . (a) Manhattan plot of co-localized locus for phytate concentration and content in 2013 and BLUP. (b) Significant SNP associated with phytate concentration and content (4.87 Mb–6.35 Mb) on chromosome A09, the big red dots represented the significant SNPs. (c) Candidate gene association analysis of *BnaA09g10220D* (phytate concentration). (d) Two haplotypes of *BnaA09g10220D* . (e) Candidate gene association analysis of *BnaA09g10220D* (phytate content). (f) Two haplotypes of *BnaA09g10220D* . (g) Candidate gene association analysis of *BnaA09g10230D* (phytate concentration). (h) Two haplotypes of *BnaA09g10230D* . (i) Candidate gene association analysis of *BnaA09g10230D* (phytate content). (j) Two haplotypes of *BnaA09g10230D* .

Figure 4. Association of the six alleles in chrA09_5238455, chrA09_5238855, chrA09_5238924, chrA09_5239266, chrA09_5239963 and chrA09_5245895 with phytate concentration and content in *B. napus* . (a) Association of the six alleles with phytate concentration. (b) Association of the six alleles with phytate content. The numbers of inbred lines harboring the corresponding allele were shown in the bracket at the bottom. PA_Conc, phytate concentration; PA_Cont, phytate content.

Figure 5. Germination rate and correlation of seed germination rate with seed phytate concentration and content in *B. napus* . (a) Germination rate (7 d) of high and low phytate haplotypes. (b) Correlation between seed germination rate and phytate concentration in the seed of *B. napus* . (c) Correlation between seed germination rate and phytate content in the seed of *B. napus* .

Figure 6. Difference in the seed yield and yield-related traits between *B. napus* cultivars with high and low phytate haplotypes under a sufficient P supply (90 kg/ha P_2O_5) and a deficient P supply ((0 kg/ha P_2O_5). (a) Seed yield (g/plant). (b) Plant height (cm). (c) Thousand -seed weight (g/1000 seeds). (d) Effective branch number (No. /plant).

Note: *, significant level $p < 0.05$.

Figure 7. Differences in dry weight, total P concentration, Pi concentration, phytate concentration and the percentage of P in different organs between *B. napus* cultivars HPA972 and LPA20. HPA972, a high phytate variety; LPA20, a low phytate variety. (a) Schematic diagram of sampling organs in *B. napus*. Node 1, the transition zone between root and stem; Node 2, the place on the stem from which a leaf growth; Node 3, the place on the stem from which a branch growth. (b) Dry weight. (c) Total P concentration. (d) Phytate concentration. (e) Pi concentration. (f) Percentage of P each sampling organ in HPA972. (g) Percentage of P each sampling organ in LPA20. Data represented the mean \pm SD of eight replicates in (b), (c), (d) and (e). Asterisks indicate the significance of Student's t test (* $P < 0.1$, ** $P < 0.01$).

Figure S1. Population structure of an association panel of *B. napus* with K from 2 to 8. (a) The ΔK estimated for population structure. The x axis represented the different accessions, the y axis quantified cluster membership, and each accession shown as a vertical line that is partitioned into K colored segments 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.

Figure S2. The kinship of an association panel of 505 *B. napus* accessions. (a) A heatmap of the kinship value among association panel of 505 *B. napus* accessions. (b) The distribution of pairwise relative kinship.

Figure S3. The SNP density and LD decay of an association panel of *B. napus* . (a) Distribution of SNP across the whole genome of *B. napus* . (b) The LD decay of the whole genome. (c) LD decay of each chromosome in the A genome and whole genome. (d) LD decay of each chromosome in the C genome and whole genome.

Figure S4. Genome wide association analysis (GLM (Q) model) for phytate concentration and content in an association panel of *B. napus* in 2013, 2015 and BLUP. (a) Phytate concentration in 2013. (b) phytate content

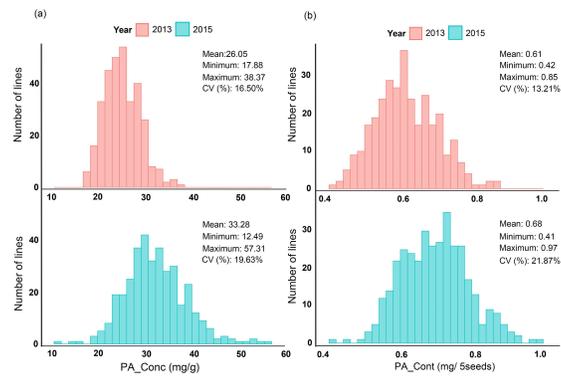
in 2013. (c) phytate concentration in 2015. (d) phytate content in 2015. (e) BLUP phytate concentration in 2013 and 2015. (f) BLUP phytate content in 2013 and 2015. BLUP, the best linear unbiased prediction.

Figure S5. Genome wide association analysis (MLM (Q+K) model) for phytate concentration and content in an association panel of *B. napus* in 2013, 2015 and BLUP. (a) Phytate concentration in 2013. (b) phytate content in 2013. (c) phytate concentration in 2015. (d) phytate content in 2015. (e) BLUP phytate concentration in 2013 and 2015. (f) BLUP phytate content in 2013 and 2015. BLUP, the best linear unbiased prediction.

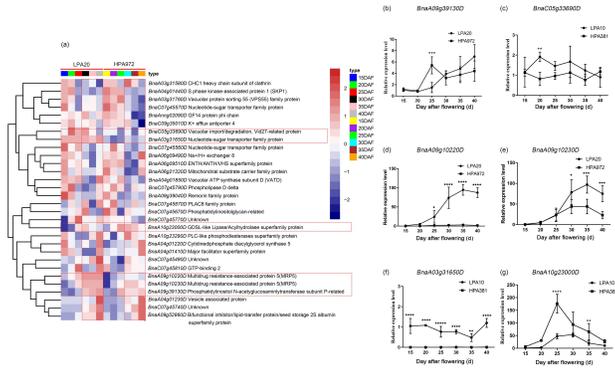
Figure S6. Difference in shoot and root between *B. napus* cultivars HPA972 and LPA20. HPA972, a high phytate variety; LPA20, a low phytate variety. (a) Shoot and root. (b) Branches.

Hosted file

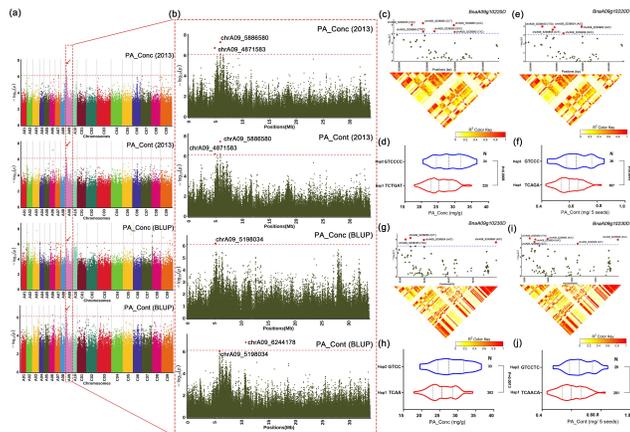
main document-2021-01.15-PCE.pdf available at <https://authorea.com/users/391783/articles/505796-haplotypes-within-tandemly-duplicated-candidate-genes-at-bnaa9-mrp5-modulate-phytate-concentration-in-canola-brassica-napus-l>



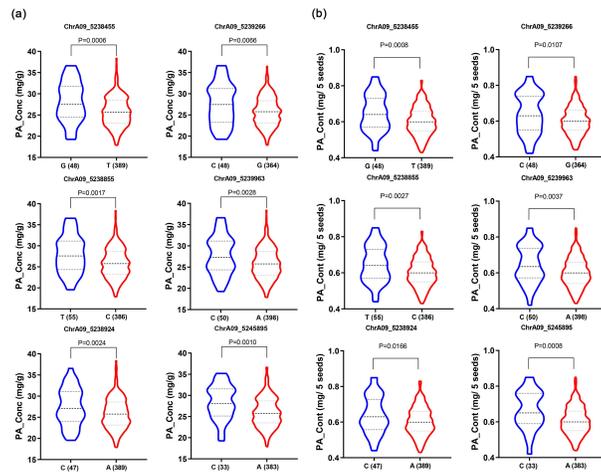
1
2 Figure 1. Frequency distribution of phytate concentration and content in the seed of an association
3 panel of *B. napus* in 2013 and 2015. (a) Phytate concentration. (b) Phytate content. PA_Conc,
4 phytate concentration; PA_Cont, phytate content.



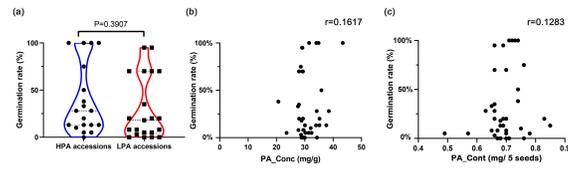
1
 2 Figure 2. Difference in the gene expression of candidate genes underlying the significant SNP
 3 associated with phytate concentration and content in *B. napus*. HPA972, a high phytate variety;
 4 LPA20, a low phytate variety. (a) Heatmap of the expression pattern of 30 candidate genes
 5 associated with phytate synthesis and transport. (b) - (g) Differences in the gene expression of
 6 *BnaA09g39130D* (b), *BnaC05g33690D* (c), *BnaA09g10220D* (d), *BnaA09g10230D* (e),
 7 *BnaA03g31650D* (f), *BnaA10g23000D* (g) in the seed development between LPA20 and HPA972.
 8 DAF, day after flowering. Data represent the mean \pm SD of four replicates in (b), (c), (d), (e), (f)
 9 and (g). Asterisks indicate the significance of Student's t test (* $P < 0.1$, ** $P < 0.01$, *** $P <$
 10 0.0001 , **** $P < 0.0001$).



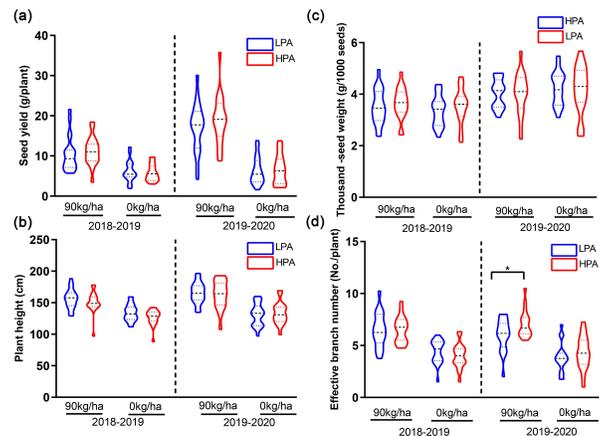
1
 2 Figure 3. The co-localized locus and haplotypes on chromosome A09 associated with phytate
 3 concentration and content of *B.napus*. (a) Manhattan plot of co-localized locus for phytate
 4 concentration and content in 2013 and BLUP. (b) Significant SNP associated with phytate
 5 concentration and content (4.87 Mb–6.35 Mb) on chromosome A09, the big red dots represented
 6 the significant SNPs. (c) Candidate gene association analysis of *BnaA09g10220D* (phytate
 7 concentration). (d) Two haplotypes of *BnaA09g10220D*. (e) Candidate gene association analysis
 8 of *BnaA09g10220D* (phytate content). (f) Two haplotypes of *BnaA09g10220D*. (g) Candidate gene
 9 association analysis of *BnaA09g10230D* (phytate concentration). (h) Two haplotypes of
 10 *BnaA09g10230D*. (i) Candidate gene association analysis of *BnaA09g10230D* (phytate content). (j)
 11 Two haplotypes of *BnaA09g10230D*.



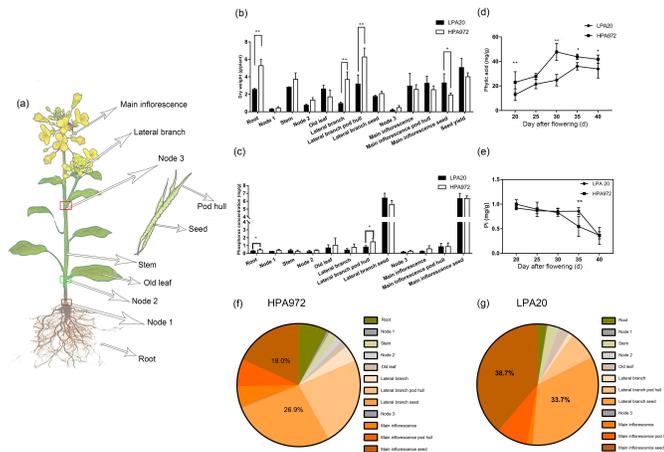
1
2 Figure 4. Association of the six alleles in chrA09_5238455, chrA09_5238855, chrA09_5238924,
3 chrA09_5239266, chrA09_5239963 and chrA09_5245895 with phytate concentration and content
4 in *B. napus*. (a) Association of the six alleles with phytate concentration. (b) Association of the six
5 alleles with phytate content. The numbers of inbred lines harboring the corresponding allele were
6 shown in the bracket at the bottom. PA_Conc, phytate concentration; PA_Cont, phytate content.



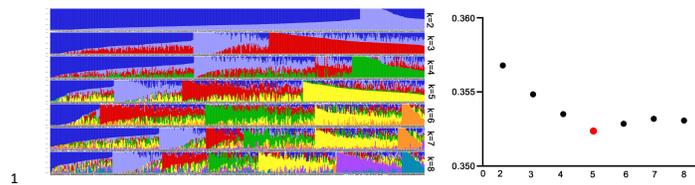
1
2 Figure 5. Germination rate and correlation of seed germination rate with seed phytate
3 concentration and content in *B. napus*. (a) Germination rate (7 d) of high and low phytate
4 haplotypes. (b) Correlation between seed germination rate and phytate concentration in the seed of
5 *B. napus*. (c) Correlation between seed germination rate and phytate content in the seed of *B.*
6 *napus*.



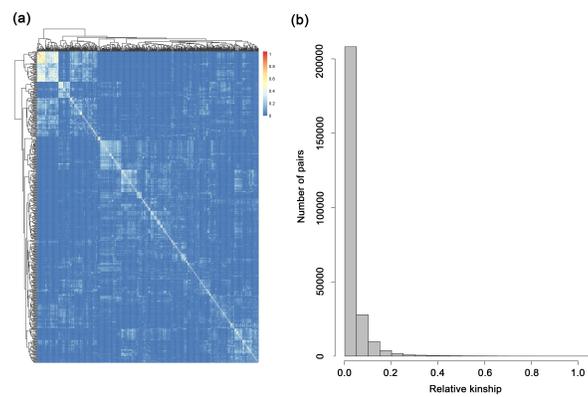
1
2 Figure 6. Difference in the seed yield and yield-related traits between *B. napus* cultivars with high
3 and low phytate haplotypes under a sufficient P supply (90 kg/ha P_2O_5) and a deficient P supply
4 ((0 kg/ha P_2O_5). (a) Seed yield (g/plant). (b) Plant height (cm). (c) Thousand-seed weight (g/1000
5 seeds). (d) Effective branch number (No./plant).
6 Note: *, significant level $p < 0.05$.



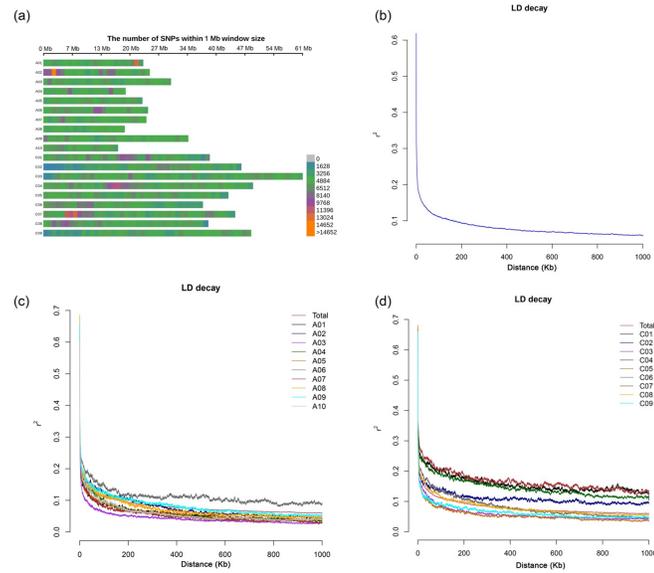
1
 2 Figure 7. Differences in dry weight, total P concentration, Pi concentration, phytate concentration
 3 and the percentage of P in different organs between *B. napus* cultivars HPA972 and LPA20.
 4 HPA972, a high phytate variety; LPA20, a low phytate variety. (a) Schematic diagram of sampling
 5 organs in *B. napus*. Node 1, the transition zone between root and stem; Node 2, the place on the
 6 stem from which a leaf growth; Node 3, the place on the stem from which a branch growth. (b)
 7 Dry weight. (c) Total P concentration. (d) Phytate concentration. (e) Pi concentration. (f)
 8 Percentage of P each sampling organ in HPA972. (g) Percentage of P each sampling organ in
 9 LPA20. Data represented the mean ± SD of eight replicates in (b), (c), (d) and (e). Asterisks
 10 indicate the significance of Student's t test (*P < 0.1, **P < 0.01).



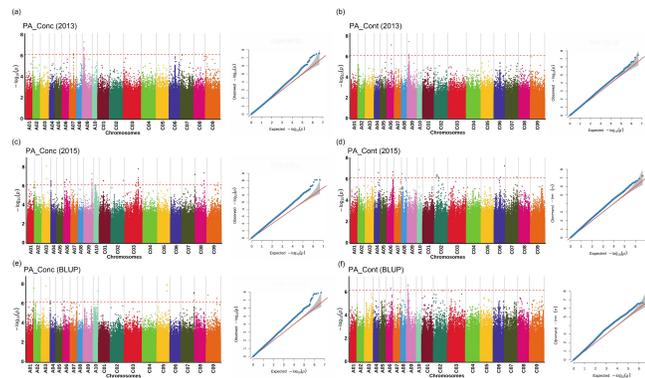
1
2 Figure S1. Population structure of an association panel of *B. napus* with K from 2 to 8. (a) The ΔK
3 estimated for population structure. The x axis represented the different accessions, the y axis
4 quantified cluster membership, and each accession shown as a vertical line that is partitioned into
5 K colored segments represented inferred membership in K genetic clusters. (b) The K value
6 estimated for population structure analysis. There was a minimum K-value when K = 5.



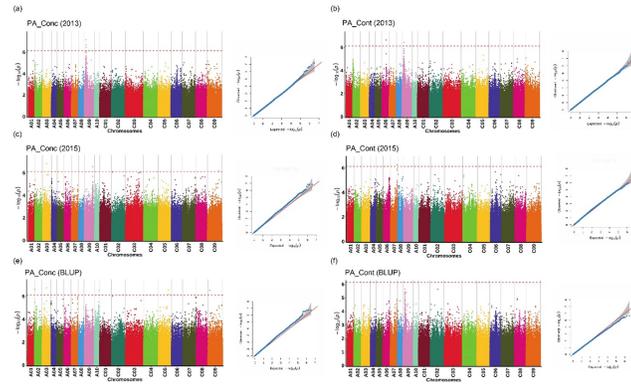
1
2 Figure S2. The kinship of an association panel of 505 *B. napus* accessions. (a) A heatmap of the
3 kinship value among association panel of 505 *B. napus* accessions. (b) The distribution of pairwise
4 relative kinship.



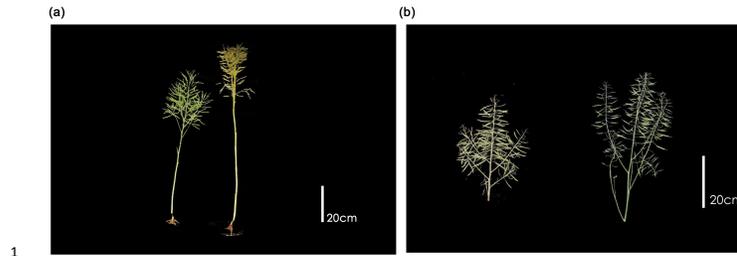
1
2 Figure S3. The SNP density and LD decay of an association panel of *B. napus*. (a) Distribution of
3 SNP across the whole genome of *B. napus*. (b) The LD decay of the whole genome. (c) LD decay
4 of each chromosome in the A genome and whole genome. (d) LD decay of each chromosome in
5 the C genome and whole genome.



1
2 Figure S4. Genome wide association analysis (GLM (Q) model) for phytate concentration and
3 content in an association panel of *B. napus* in 2013, 2015 and BLUP. (a) Phytate concentration in
4 2013. (b) phytate content in 2013. (c) phytate concentration in 2015. (d) phytate content in 2015.
5 (e) BLUP phytate concentration in 2013 and 2015. (f) BLUP phytate content in 2013 and 2015.
6 BLUP, the best linear unbiased prediction.



1
2 Figure S5. Genome wide association analysis (MLM (Q+K) model) for phytate concentration and
3 content in an association panel of *B. napus* in 2013, 2015 and BLUP. (a) Phytate concentration in
4 2013. (b) phytate content in 2013. (c) phytate concentration in 2015. (d) phytate content in 2015.
5 (e) BLUP phytate concentration in 2013 and 2015. (f) BLUP phytate content in 2013 and 2015.
6 BLUP, the best linear unbiased prediction.



1
2 Figure S6. Difference in shoot and root between *B. napus* cultivars HPA972 and LPA20. HPA972, a
3 high phytate variety; LPA20, a low phytate variety. (a) Shoot and root. (b) Branches.

Hosted file

Figure 1.pdf available at <https://authorea.com/users/391783/articles/505796-haplotypes-within-tandemly-duplicated-candidate-genes-at-bnaa9-mrp5-modulate-phytate-concentration-in-canola-brassica-napus-1>

Hosted file

Figure 2.pdf available at <https://authorea.com/users/391783/articles/505796-haplotypes-within-tandemly-duplicated-candidate-genes-at-bnaa9-mrp5-modulate-phytate-concentration-in-canola-brassica-napus-1>

Hosted file

Figure 3.pdf available at <https://authorea.com/users/391783/articles/505796-haplotypes-within-tandemly-duplicated-candidate-genes-at-bnaa9-mrp5-modulate-phytate-concentration-in-canola-brassica-napus-1>

Hosted file

Figure 4.pdf available at <https://authorea.com/users/391783/articles/505796-haplotypes-within-tandemly-duplicated-candidate-genes-at-bnaa9-mrp5-modulate-phytate-concentration-in-canola-brassica-napus-1>

Hosted file

Figure 5.pdf available at <https://authorea.com/users/391783/articles/505796-haplotypes-within-tandemly-duplicated-candidate-genes-at-bnaa9-mrp5-modulate-phytate-concentration-in-canola-brassica-napus-1>

Hosted file

Figure 6.pdf available at <https://authorea.com/users/391783/articles/505796-haplotypes-within-tandemly-duplicated-candidate-genes-at-bnaa9-mrp5-modulate-phytate-concentration-in-canola-brassica-napus-1>

Hosted file

Figure 7.pdf available at <https://authorea.com/users/391783/articles/505796-haplotypes-within-tandemly-duplicated-candidate-genes-at-bnaa9-mrp5-modulate-phytate-concentration-in-canola-brassica-napus-1>

Hosted file

Figure S1.pdf available at <https://authorea.com/users/391783/articles/505796-haplotypes-within-tandemly-duplicated-candidate-genes-at-bnaa9-mrp5-modulate-phytate-concentration-in-canola-brassica-napus-1>

Hosted file

Figure S2.pdf available at <https://authorea.com/users/391783/articles/505796-haplotypes-within-tandemly-duplicated-candidate-genes-at-bnaa9-mrp5-modulate-phytate-concentration-in-canola-brassica-napus-1>

Hosted file

Figure S3.pdf available at <https://authorea.com/users/391783/articles/505796-haplotypes-within-tandemly-duplicated-candidate-genes-at-bnaa9-mrp5-modulate-phytate-concentration-in-canola-brassica-napus-1>

Hosted file

Figure S4.pdf available at <https://authorea.com/users/391783/articles/505796-haplotypes-within-tandemly-duplicated-candidate-genes-at-bnaa9-mrp5-modulate-phytate-concentration-in-canola-brassica-napus-1>

Hosted file

Figure S5.pdf available at <https://authorea.com/users/391783/articles/505796-haplotypes-within-tandemly-duplicated-candidate-genes-at-bnaa9-mrp5-modulate-phytate-concentration-in-canola-brassica-napus-1>

Hosted file

Figure S6.pdf available at <https://authorea.com/users/391783/articles/505796-haplotypes-within-tandemly-duplicated-candidate-genes-at-bnaa9-mrp5-modulate-phytate-concentration-in-canola-brassica-napus-l>