OXFORD

Research Article

Integrating a genome-wide association study with transcriptomic data to predict candidate genes and favourable haplotypes influencing *Brassica napus* seed phytate

Haijiang Liu[1,2†], **Xiaojuan Li**[1,2†], **Qianwen Zhang**[3†], **Pan Yuan**[1,2], **Lei Liu**[4], **Graham J. King** [iD] [4], **Guangda Ding**[1,2], **Sheliang Wang**[1,2], **Hongmei Cai**[2], **Chuang Wang**[2], **Fangsen Xu**[1,2], and **Lei Shi** [iD] [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, Shanghai, 200240, China, and [4]Southern Cross Plant Science, Southern Cross University, Lismore, New South Wales, 2480, Australia

[†]These authors contributed equally to this work.

*To whom correspondence should be addressed. Tel. +8627 87286871. Fax. +8627 87280016. Email: 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 (*Brassica napus*). Here, a panel of 505 diverse *B. napus* accessions was screened in a genome-wide association study (GWAS) using 3.28×10^6 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 (27 kb) 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, favourable haplotypes and the low phytate varieties identified in this study will be useful for low-phytate breeding of *B. napus*.

Key words: oilseed rape, phytate, genome-wide association study, *BnaA9.MRP5*, haplotype analysis

1. Introduction

Phosphorus (P) is an essential macro-element for higher plant growth and development. Phytate is the storage form of P in angiopsperm seeds, accounting for approximately two-thirds of total seed P. Seed phytate provides *myo*-inositol and P used for seed germination and seedling establishment.[1] It also plays vitally important roles in P sensing and homoeostasis during seed development. The function of inositol triphosphate during phytate biosynthesis is as a second messenger mediating intracellular calcium signalling and drought stress responses in plants, along with transcriptional induction of genes associated with inositol metabolism.[2] Additionally, phytate is involved in DNA damage-repair, RNA-editing, mRNA export, and other intracellular regulation in plants.[3] 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.*[4–14] Phytate chelates positive ions of essential minerals, such as Mg, Ca, Fe, and Zn, 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.[1] 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.[15] 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 phytate biosynthesis that may be suitable targets for selective or other advanced breeding. In *Arabidopsis thaliana* (*A. thaliana*), a major quantitative trait locus (QTL) affecting $InsP_6$ and Pi in seeds and leaves has been detected on the top of chromosome 3 by the L*er*/Cvi RIL population, which explained 61.8% of the variation for InsP6 in seeds.[16] 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 (PVE).[17] 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 (GWASs) take advantage of PVE and historical recombination in natural populations without the time and energy required to construct a segregating mapping population.[18] GWAS has become a routine method for analysing complex agronomic traits, such as ionomic variation in rice,[19] arsenic accumulation in maize,[20] and cannabinoid composition in *Cannabis*.[21] Oilseed rape (OSR, rapeseed, canola: *Brassica napus*) provides 18% of global vegetable oil for human consumption, industrial oils, biodiesel, lubricant, and fodder for animal feeds (FAOSTAT, 2014, http://faostat.fao.org/). 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,[22] dozens of agriculturally important traits have been dissected by GWAS,[23,24] although few studies have included *B. napus* seed phytate traits. In this study, a diverse panel of 505 *B. napus* accessions was scored for seed phytate concentration and content and 119 significant associated single nucleotide polymorphism (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.

2. Materials and methods

2.1. 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 centres across China. A total of 443 lines originated in China, 28 from Europe, 10 from Japan, 5 from Canada, 6 from Australia, 3 from Korea, and 10 unknown.[24]

2.2. Field trials of the association panel and trait measurement

The association panel was used to conduct two years of field trails 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}, NH_4OAc-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 $Na_2B_4O_7 \cdot 10H_2O$) applied. Application rates were as follows: N 108 kg ha^{-1}, P_2O_5 90 kg ha^{-1}, K_2O 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 the phytate content was calculated.

2.3. Phytate analysis in seeds

Phytate was analysed using a modification of an existing method.[25] 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 0.4 M of HCl. The mixture was ground for 3 min in a TissueLyser ball mill (Qiagen, Germany). The extracts were centrifuged at 15,000 rpm for 5 min, 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 were 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. A 80 µl of each sample and standard solution (0.016–0.131 mg/ml, phytate, Sigma-Aldrich) from the PCR plate were

transferred into a 96-well plate (flat bottom) for colorimetric determination. A 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.

2.4. 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.[24] Best linear unbiased prediction (BLUP) of seed phytate concentration and content for each line was calculated using the R package 'lme4' (https://cran.r-project.org/web/packages/lme4/index.html). Genome-wide association analysis for seed phytate concentration and content was carried out using general linear models (GLMs) and mixed linear models (MLMs) using the Tassel 5.0 software.[26] 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, the Q + K matrices were both considered in the MLM. 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.[27] The kinship matrix was predicted using the Tassel 5.0 software.[26] The value less than 0 in kinship was defined as 0. The linkage disequilibrium (LD) statistic R^2 was calculated by PopLDdecay software.[28] 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.[26] 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*.

2.5. 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 GLMs and MLMs were identified based on the *B. napus* reference genome.[22] Finally, genes with significantly different levels of transcription in

seeds between high and low phytate varieties were regarded as candidate genes for phytate synthesis.

2.6. Haplotype analysis

HaploView software was used to conduct haplotype analysis. The haplotypes were classified based on all of the SNPs with an MAF >0.05 in a candidate gene. Haplotypes containing at least 15 *B. napus* accessions were used for final comparative analysis, and Student's *t*-test was used to compare the differences in phytate concentration and content among the haplotypes.

2.7. 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 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 × width × height = 8 cm × 6 cm × 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 $(NH_4)_2SO_4$), 1.05 g P (supplied as KH_2PO_4), 1.40 g K (supplied as KCl), 1.75 g Mg (supplied as $MgSO_4.7H_2O$) were applied to the soil, and then 7 ml 46 μM H_3BO_3, 0.32 μM $CuSO_4.5H_2O$, 0.77 μM $ZnSO_4.7H_2O$, 9.14 μM $MnCl_2.4H_2O$, 0.37 μM $Na_2MoO_4 \cdot 2H_2O$ and 50 μM EDTA-Fe (II) were applied, respectively. Each variety was represented by five plants (biological replicates). The plants were grown outdoors with a movable rain shelter. Seeds were sampled at 15, 20, 25, 30, 35, and 40 days after flowering (DAFs), and immediately put in the liquid nitrogen and then stored in the −80°C for subsequent RNA extraction. At the mature stage, root, stem, old leaves, node, silique pericarp, and seed were harvested for determination of biomass and total P.

Total RNA was extracted from sampled seed from the developmental series using a plant RNA purification Kit (Qingke, Beijing, China). cDNA was prepared using cDNA Synthesis Kit (Kangwei, Beijing, China). Primers were listed in Supplementary Table S1. PCR reaction solution contained 5 μl Master Mix ABI Prism™, 0.2 μl forward primer, 0.2 μl reverse primer, 1 μl template DNA and 3.6 μl PCR-grade water. The PCR programme was as follows: 95°C for 5 min , followed by 40 cycles of 95°C for 10 s, 60°C for 20 s and 72°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 $2^{-\triangle\triangle CT}$ method. *BnaTublin* was used as an internal control for normalization.

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

Cultivars	Phytate concentration (mg/g)		Phytate content (mg/5 seeds)		Thousand-seeds weight (g/1000 seeds)		Seed phosphorus concentration (mg/g)	
	2013	2015	2013	2015	2013	2015	2013	2015
HPA972	34.2a	35.43a	0.75a	0.773a	3.71a	3.52a	9.33a	9.04a
LPA20	23.57b	24.48 b	0.48b	0.412b	3.87a	3.79a	8.02a	8.12a

HPA: high phytate content in the seed; LPA: low phytate content in the seed; HPA972: a high phytate variety; LPA20: a low phytate variety.
Values followed by different letters in same column indicate significant differences among cultivars ($P < 0.05$).

2.8. Germination experiment and field trials of the *B. napus* haplotypes

A total of 40 accessions from the association panel, including 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 '*BnaA09g10230 ConcHap1*', 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*', $G_PT_EC_EC_IC_E$ at '*BnaA09g10220ContHap2*', $G_PT_EC_EC_EC_E$ at '*BnaA09g10230ConcHap2*', and $G_PT_EC_EC_ET_EC$ 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. One hundred plump seeds were disinfected with 1% NaClO for 5 min, and then washed six 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 per 100 seeds placed on the germination medium. The 40 accessions were also grown in the field with at a sufficient P 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 four rows and each plot with eight 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.

2.9. Determination of total P concentration in plants

Whole plants were dried at 65°C and then ground to a powder using porcelain mortar. A 0.025 g sample (root, node, stem, silique pericarp, 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 a clarified solution was obtained, which was then diluted to 20 ml, and the P concentration was measured by flow injection analyser (SEAL AA3).

2.10. Determination of Pi concentration in plants

The malachite green method was used to determine inorganic P (Pi). The method involved: adding two steel balls, 0.025 g fresh sample (root, node, stem, silique pericarp, seed, and old leaves) and 25 μl H_2SO_4 into 2 ml centrifugal tube, and then ground three times for 20 s each time. One millilitre ddH$_2$O was added to the treated sample, and then centrifuged at 1000 rpm for 10 min. The supernatant was mixed with malachite green solution of three 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 × 1.5 ml)/(600 μl × 0.001 × M) × C × 0.001. 'M' is the sample weight and 'C' is the dilution.

3. Results

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

Seed phytate concentration and content were determined for the association panel of 505 *B. napus* accessions grown at a sufficient P supply in 2013–14 and 2015–16 field trials. Both phytate traits in this panel showed extensive PVE with an approximately normal distribution (Supplementary Fig. S1 and Table S2). For example, seed phytate concentration ranged from 17.88 to 38.37 mg/g (2.1-fold) in

the 2013–14 field trial and from 12.49 to 57.31 mg/g (4.6-fold) in 2015–16. Variation in phytate content was more consistent between years, ranging from 0.42 to 0.85 mg/5 seeds (2.0-fold) in 2013–14 and ranged from 0.41 to 0.97 mg/5 seeds (2.4-fold) in 2015–16 (Supplementary Fig. S1 and Table S2).

3.2. Population structure, relative kinship, and LD decay

In total, 3.28 million SNP markers were identified for the *B. napus* association panel.[24] SNP number on each chromosome ranged from 86,775 on A08 to 296,088 on C03 (Supplementary Fig. S2a and Table S3). LD decay on each chromosome ranged from 94 kb on C05 to 1000 kb on C01, when r^2 was 0.1 (Supplementary Fig. S2c and d, Table S3). Overall LD decay was 228 kb, and for each genome 135 kb (A genome) and 369 kb (C genome) (Supplementary Fig. S2b and Table S3). To avoid over-adjustment, a total of 412,141 SNPs were selected to assess population structure, relative kinship, and LD decay. The population could be divided into five subgroups based on the cross validation errors (Supplementary Fig. S3). The *r* pairwise relative kinship was close to 0 (Supplementary Figs. S4a and b, Table S4). 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 (Supplementary Fig. S4b and Table S4). 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.

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

We chose the combination of GLMs and MLMs for GWAS analysis. A total of 119 SNPs were identified that were significantly associated with phytate concentration and content ($P < 6.25 \times 10^{-07}$) across two years (Supplementary Figs. S5 and S6, Table S5). Among the 119 SNPs, 31 were identified in 2013–14, 59 in 2015–16, and 40 were identified by using BLUP values (Supplementary Figs. S5 and S6, Table S5). 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 of significant SNPs (10 SNPs; Supplementary Table S5). MLM analysis detected 19 SNPs significantly associated with seed phytate concentration and content on 7 of the 19 *B. napus* chromosomes (Supplementary Table S5). Of these, 19 significant SNPs were identified simultaneously by the GLMs and MLMs (Supplementary Table S5). Additionally, 9.24% (11/119) of the significant SNPs were identified in more than one year (including BLUP), which showed high reliability (Supplementary Table S5).

Of the 119 significant SNPs, 8 were detected simultaneously for phytate concentration and content (Supplementary Table S5). For example, the significant SNP marker (chrA09_5886580) on chromosome A09 could explain 11.76% and 11.85% of the PVE of phytate concentration and phytate content, respectively (Fig. 1b, Supplementary Table S5).

3.4. 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 GLMs and MLMs for seed
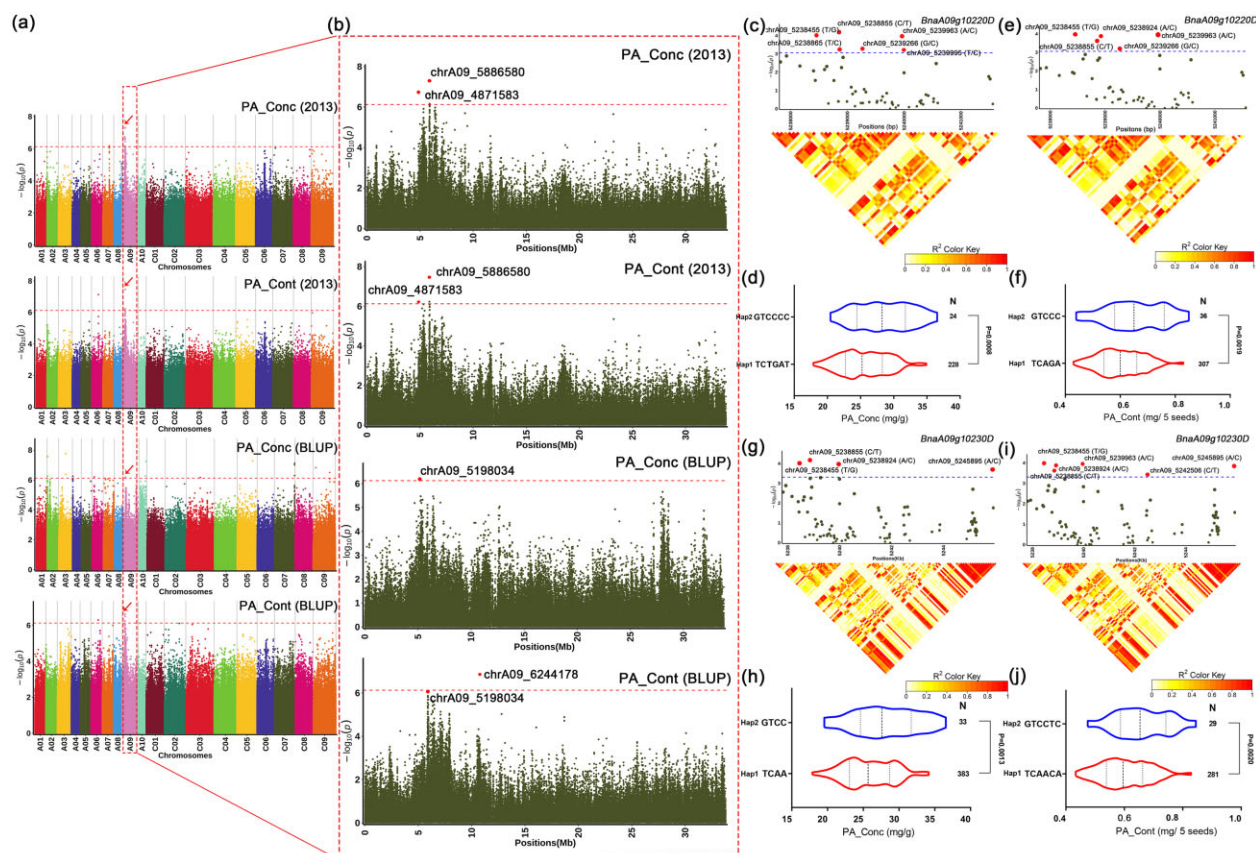
Figure 1.. The co-localized locus and haplotypes on chromosome A09 associated with seed phytate concentration and content of *B. napus*. (a) Manhattan plot of co-localized locus for phytate concentration and content in 2013 and BLUP. (b) Significant SNPs associated with phytate concentration and content (4.87–6.35 Mb) on chromosome A09. The big red dots represented the significant SNPs. (c) Candidate gene association analysis of *BnaA09g10220D* (phytate concentration). (d) Phytate concentration of two haplotypes of *BnaA09g10220D*. (e) Candidate gene association analysis of *BnaA09g10220D* (phytate content). (f) Phytate content of two haplotypes of *BnaA09g10220D*. (g) Candidate gene association analysis of *BnaA09g10320D* (phytate concentration). (h) Phytate concentration of two haplotypes of *BnaA09g10320D*. (i) Candidate gene association analysis of *BnaA09g10230D* (phytate concentration). (j) Phytate content of two haplotypes of *BnaA09g10230D*. Both the mean and BLUP values in the association panel were used to conduct association analysis of phytate concentration and content. The mean values of seed phytate concentration and content each cultivar/inbred line of the association panel in 2013 were used for the haplotype analysis.

phytate concentration and content (Fig. 2a). The transcription of these genes was quantified by qRT-PCR in seed from the developmental series 15–40 DAFs in the high phytate cultivar HPA972 and the low phytate cultivar LPA20 (Fig. 2a). The transcription of *BnaA09g39130D* and *BnaA10g23000D* in the seed of HPA972 at 25 DAFs was relatively high compared with that of LPA20 (Fig. 2b and g). In addition, the transcription of *BnaA09g10220D* and *BnaA09g10230D* in the seed of LPA20 at 30, 35, and 40 DAFs was relatively higher than that of HPA972 (Fig. 2d and e). The relative gene expression level of *BnaC05g33690D* in the seed of LPA20 at 20, 25, and 30 DAFs was higher than that of HPA972 (Fig. 2c), and the relative gene expression level of *BnaA03g31650D* in the seed of LPA20 at 15, 20, 25, 30, 35, and 40 DAFs was higher than that of HPA972 (Fig. 2f). These *B. napus* genes may be associated with functions that play either a direct or indirect role in the seed phytate concentration difference between the low and high phytate varieties.

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

The significant SNP 'chrA09_5198034' on chromosome A09 was detected for seed phytate concentration in the BLUP analysis, and explained 8.33% of the PVEs (Fig. 1a, Supplementary Table S5).

BnaA09g10220D (*BnaA9.MRP5a*) and *BnaA09g10230D* (*BnaA9. MRP5b*) are adjacent tandem genes on A09, where the former was located 27 kb from the peak SNP 'chrA09_5198034'. The significant SNP 'chrA09_4871583' on chromosome A09 was detected for seed phytate concentration and content in the 2013 trial, and explained 10.05% and 9.35% of the PVEs, respectively (Fig. 1a, Supplementary Table S5). Moreover, ChrA09_4871583 was located 300 kb from *BnaA09g10220D* (*BnaA9.MRP5a*). 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 signalling in response to ABA.[29]

Bna.A9.MRP5 was associated with both phytate concentration and phytate content by GLMs and MLMs simultaneously. Candidate gene association analysis and haplotype analysis were therefore performed. *BnaA09g10220D* (*BnaA9.MRP5a*) has 1 intron and 2 exons, and *BnaA09g10230D* (*BnaA9.MRP5b*) has 12 introns and 13 exons (Supplementary Fig. S7). 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 (Supplementary Tables S6 to S9). Six SNPs in *BnaA09g10220D* were detected to be
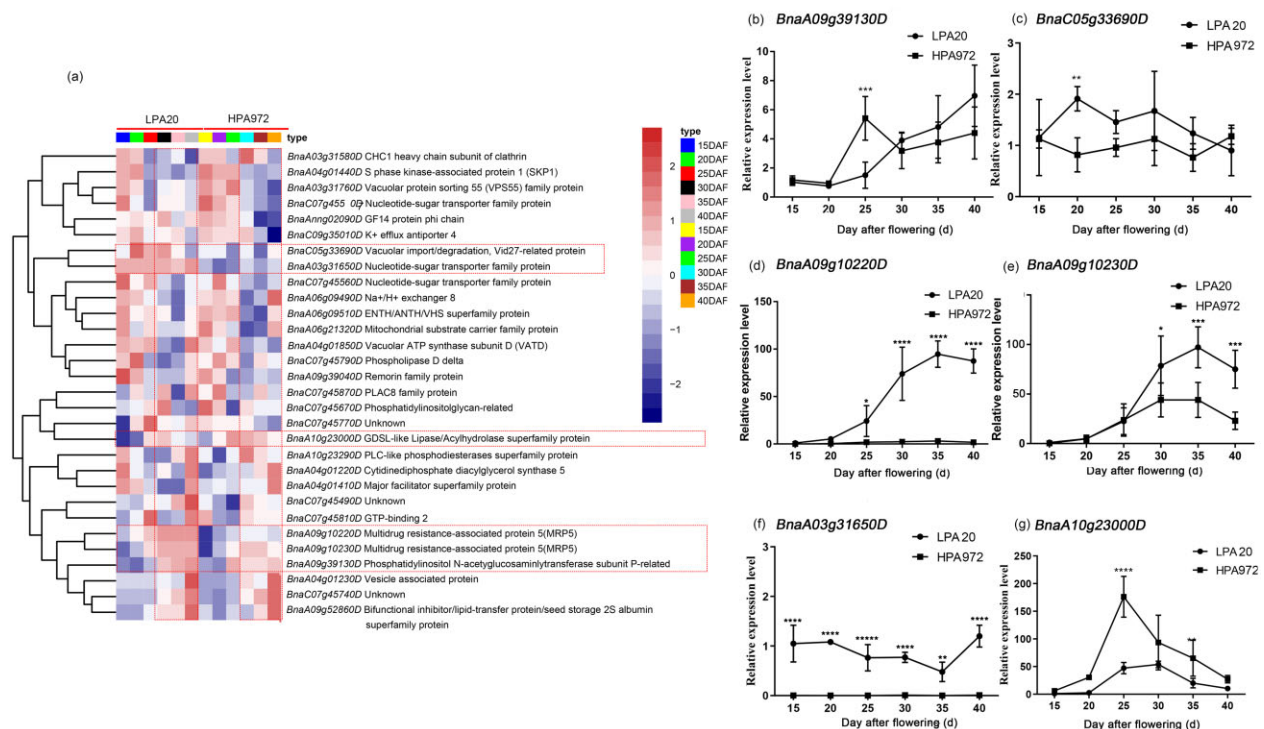
Figure 2.. Difference in the gene expression of candidate genes underlying the significant SNP associated with seed phytate concentration and content in *B. napus*. HPA972, a high phytate variety; LPA20, a low phytate variety. (a) Heatmap of the expression pattern of 30 candidate genes associated with phytate synthesis and transport in the seeds of LPA20 and HPA792 during 15–40 days after flowering (DAFs). (b) to (g) Differences in the gene expression of *BnaA09g39130D* (b), *BnaC05g33690D* (c), *BnaA09g10220D* (d), *BnaA09g10230D* (e), *BnaA03g31650D* (f), and *BnaA10g23000D* (g) in the seed between LPA20 and HPA792 during 15–40 DAFs. Data represent the mean $\pm$ SD of four replicates. Asterisks indicate the significance of Student's *t*-test (*$P < 0.05$, **$P < 0.01$, ***$P < 0.0001$, ****$P < 0.0001$).

significantly associated with phytate concentration (Fig. 1c). 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 (Supplementary Fig. S8a). 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 (Fig. 1d). Five SNPs in *BnaA09g10220D* were detected to be significantly associated with phytate content (Fig. 1e). 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 (Supplementary Fig. S8b).

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 (Fig. 1f). Four SNPs in *BnaA09g10230D* were associated with the phytate concentration in the seed of the association panel (Fig. 1g). 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 (Fig. 1a). Haplotype analysis of the four significant SNPs revealed two major haplotypes (Fig. 1h), 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 (Fig. 1h). Six SNPs in *BnaA09g10230D* were associated with the phytate content in the seed of the association panel (Fig. 1i). 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 (Supplementary Fig. S8). 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 (Fig. 1j). 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 favourable haplotypes for reducing phytate content and concentration in the seed of *B. napus*. The sequence variations in *BnaA9.MRP5* may contribute to most of the phytate variation. Compared with the high phytate *B. napus* variety, the low phytate variety appears to have a weaker ability to synthesize phytate (Table 1; Fig. 5d). One explanation may be that MRP5 is expressed at a higher level during seed development in order to maximize the promotion of phytate accumulation in the seed (Fig. 2d and e) by enhancing the transport of phytic acid from cytoplasm to vacuoles. The reduction in phytic acid within the cytoplasm would also reduce the inhibition of biosynthesis and so facilitate the synthesis of phytate.[30]

3.6. 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 '*BnaA09g10230*
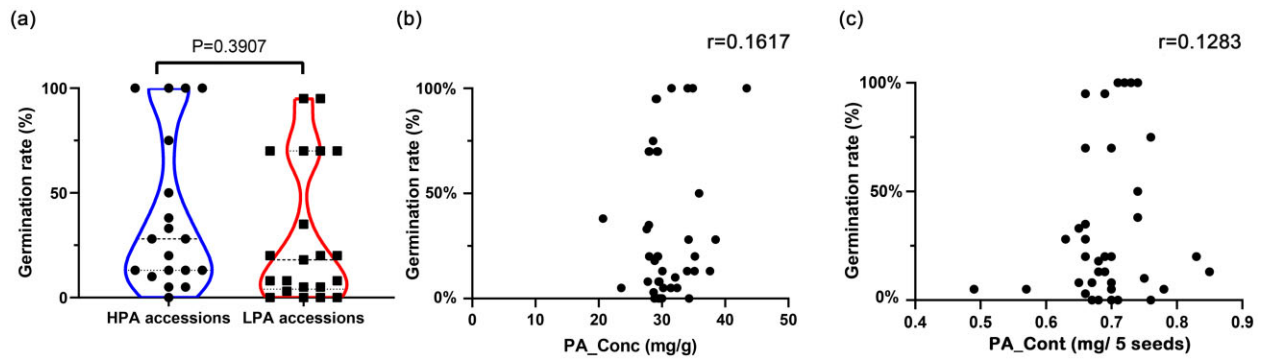
Figure 3.. Germination rate and correlation of seed germination rate with seed phytate concentration and content in *B. napus*. (a) Germination rate (7 d) of high and low phytate haplotypes. (b) Correlation between seed germination rate and phytate concentration in the seed of *B. napus*. (c) Correlation between seed germination rate and phytate content in the seed of *B. napus*.

ConcHap1', 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', $G_PT_EC_EC_IC_E$ at 'BnaA09g10220 ContHap2', $G_PT_EC_EC_E$ at 'BnaA09g10230ConcHap2', and $G_PT_EC_E$ $C_ET_EC_E$ at 'BnaA09g10230ContHap2' on A09 chromosome. In high phytate haplotype cultivars, seed phytate concentration ranged from 27.60 to 43.40 mg/g and seed phytate content from 0.57 to 0.85 mg/5 seeds. This contrasted with low phytate haplotype cultivars, where seed phytate concentration ranged from 23.57 to 30.20 mg/g and content from 0.49 to 0.71 mg/5 seeds (Supplementary Table S1). No significant correlation was detected between germination rate and seed phytate concentration (r=0.1617) or content (r=0.1283) (Fig. 3b and c). Moreover, some LPA cultivars showed a high germination rate of 70–95% (Fig. 3a, Supplementary Table S1). 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, PH, and EBN were detected between cultivars with low phytate and high phytate haplotypes under both P supplies, except for the EBN at a sufficient P supply in 2019–20 (Fig. 4, Supplementary 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 (Supplementary Table S11), although seed yield in 2019 and in 2020 for this cultivar was 15.71 and 16.68 g/plant, respectively at a sufficient P supply (Supplementary 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 $G_PT_EC_EC_IC_EC_E$ in 'BnaA09g10220ConcHap2', $G_PT_EC_EC_IC_E$ in 'BnaA09g10220ContHap2', $G_PT_EC_EC_E$ at 'BnaA09 g10230ConcHap2', and $G_PT_EC_EC_ET_EC_E$ at 'BnaA09g10230Cont Hap2' on chromosome A09 showed much more phytate than LPA20 carrying $T_PC_ET_EG_IA_ET_E$ at 'BnaA09g10220ConcHap1', $T_PC_EA_E$ G_IA_E at 'BnaA09g10220ContHap1', $T_PC_EA_EA_E$ at 'BnaA09g10230 ConcHap1', and $T_PC_EA_EA_EC_EA_E$ at 'BnaA09g10230ContHap1' (Table 1). The biomass and P content of root, lateral branches, and lateral branch silique pericarp of HPA972 were significantly higher than that for LPA20 (Fig. 5a to c). At harvest stage, 59.6% of total P

was distributed in the seed of HPA972 and 75.0% in LPA20 (Fig. 5f and g). No significant difference in seed Pi concentration was detected between HPA972 and LPA20 at 20, 25, 30, and 40 DAFs. At 35 DAFs, the seed Pi concentration of LPA20 was significantly higher than that of HPA972 (Fig. 5e). 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 DAFs (Fig. 5d). Transcriptional level of BnaA09g39130D, BnaA09g10220D, BnaA09g10230D, and BnaA10g23000D in both varieties was higher between 20 and 30 DAFs than during 15 and 25 DAFs (Fig. 2a). Only the transcription of BnaA10g23000D in HPA972 seed was higher than that of LPA20 during seed development (Fig. 2g), which were in agreement with that the difference in seed phytate concentration between HPA972 and LPA20.

4. Discussion

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

In this study, a modified high throughput method using five 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 *B. napus* association panel at a sufficient P supply showed continuous variation and an approximate normal distribution across the 505 cultivars (Supplementary Fig. S1). Seed phytate concentration had 2.31- and 4.6-fold variations in 2013–14 and 2015–16, respectively (Supplementary Fig. S1 and Table S2). Seed phytate content had 2.0- and 2.4-fold variations in 2013–14 and 2015–16, respectively (Supplementary Fig. S1 and Table S2). This compares with 2.8-fold variation (7.91–21.97 mg/g) in phytate concentration among 60 chickpea accessions,[31] and 2.2-fold variation among 69 rice accessions.[32] Since the environment had a large effect on seed phytate concentration and content,[33,34] 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 GWAS.

The decrease of Pi and the increase of phytate during seed development appear to be similar in different crops.[35,36] 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
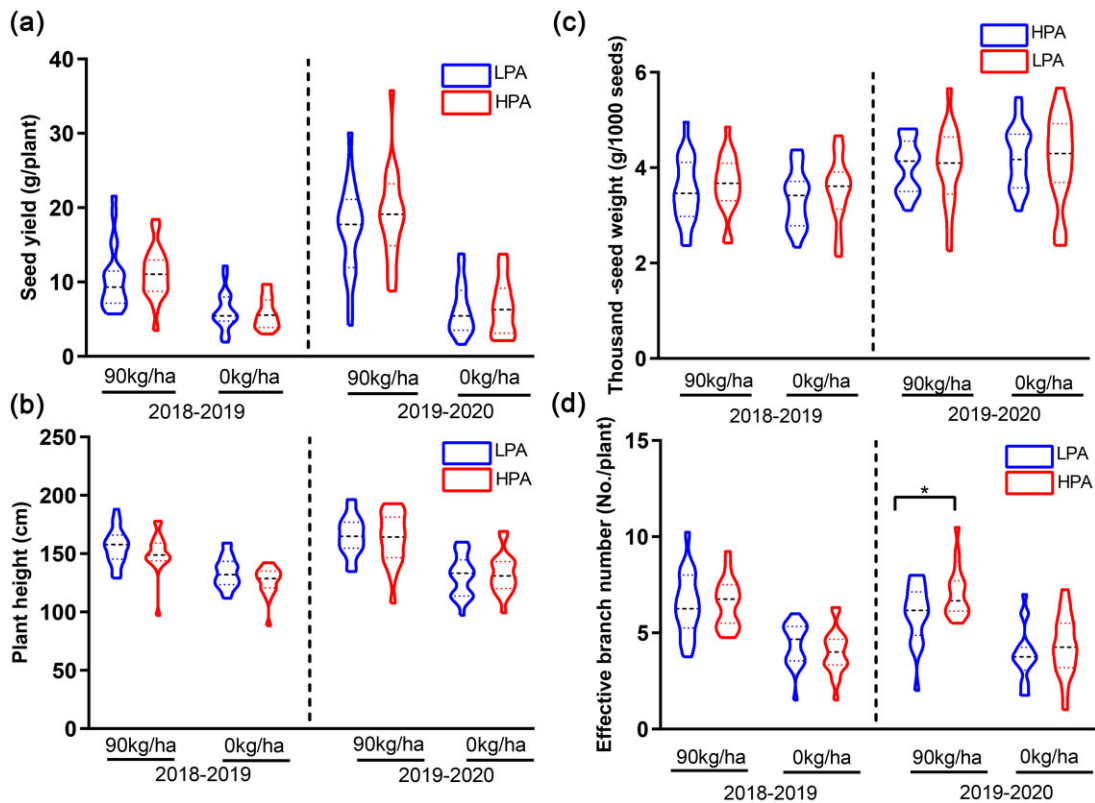
Figure 4.. Difference in the seed yield and yield-related traits between *B. napus* cultivars with high and low phytate haplotypes at a sufficient P supply (90 kg/ha P_2O_5) and a deficient P supply (0 kg/ha P_2O_5). (a) Seed yield (g/plant). (b) Plant height (cm). (c) Thousand-seed weight (g/1000 seeds). (d) Effective branch number (No./plant). Note: *, significant level $P < 0.05$.

phytate variety) decreased slowly between 20 and 30 DAFs, and phytate concentration began to increase at 20 DAFs (Fig. 5d and e). At 35 DAFs, 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 HPA972 was higher than that of LPA20 (Fig. 5e). Seed phytate concentration in HPA972 was always higher than that of LPA20 at any stage of seed development (Fig. 5d). The difference in seed phytate concentration between HPA972 and LPA20 reached a maximum at 30 DAFs (Fig. 5d). Therefore, 25–30 DAFs could be the key time point for the difference of phytate synthesis in the seed between HPA972 and LPA20. Most of the candidate genes associated with the significant SNPs for phytate concentration and content in the seed had increased transcription at 30 DAFs than at 15, 20, and 25 DAFs (Fig. 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 (Fig. 5d and e). *OsSPDT*, a sulphate transporter-like gene in rice, is preferentially located in the vascular bundle and xylem, and is responsible for controlling the distribution of P to grains.[36] Although the mutation of *OsSPDT* decreased seed phytate concentration, it had little effect on the grain germination rate in rice.[37] In low phytate rice mutants Tong et al. (2017) demonstrated an increase in the redistribution of Pi in endosperm led to lysophospholipid biosynthesis.[38] 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.[38] 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 PH 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 (Supplementary Fig. S9).

The P concentration in root and lateral branch silique pericap of HPA972 was significantly higher than in LPA20 (Fig. 5c). In addition to the difference between HPA972 and LPA20 in phytate concentration and content, and the total seed P concentration (Table 1), this may reflect differences in P uptake and translocation ability at maturity (Fig. 2). Total P in a seed is only a small proportion of the total P in the biomass of a mature plant.

4.2. 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,[39] 1.06 cM in 404 inbred lines,[23] 2.0 Mb in 472 inbred lines,[8] and 0.5–1 cM in 192 inbred lines,[40] although larger than that of other species, e.g. 200 kb in Arabidopsis,[41] 187 kb in rice,[19] and 173 kb in maize.[20]

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% PVE for phytate concentration ranging from 6.48% to 11.92% (Supplementary Table S5). A previous study had shown determining the level of transcription of candidate genes in the seed at different time points after
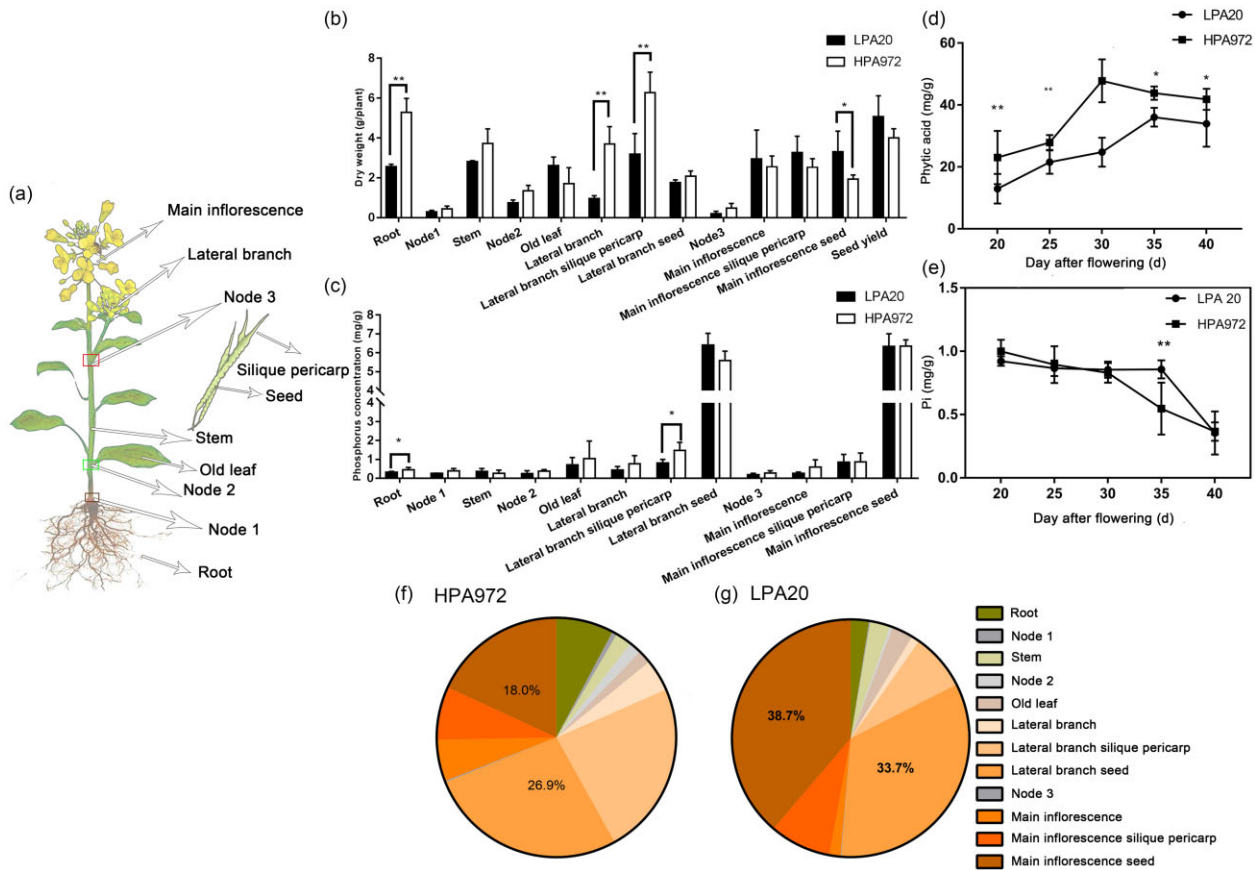
Figure 5.. Differences in dry weight, total P concentration, phytate concentration, Pi concentration, and the percentage of P in different organs between *B. napus* cultivars HPA972 and LPA20. HPA972, a high phytate variety; LPA20, a low phytate variety. (a) Schematic diagram of sampling organs in *B. napus*. Node 1, the transition zone between root and stem; Node 2, the position on the stem from which a true leaf develops; Node 3, the place on the stem from which a branch develops. (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 for each sampling organ in LPA20. Data represent the mean ± SD of eight replicates in (b), (c), (d), and (e). Asterisks indicate the significance of Student's *t*-test (*$P < 0.05$, **$P < 0.01$).

flowering helped to identify candidate genes.[13,14] 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. (Fig. 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 signalling in response to ABA.[29] *BnaC07g45670D* is homologous to *AT4G35530*, which is associated with phosphatidyl inositol glycan. *BnaA10g23000D* in *B. napus* was orthologous to *A. thaliana* *AT5G08460*, which catalyses acyl transfer or hydrolase reactions with lipid and non-lipid substrates.[42] *BnaC05g33690D* was orthologous to *AT3G19240,* which is involved in vacuolar import/degradation.[43] *BnaA03g31650D* is orthologous to *AT3G11320*, which is involved in nucleotide-sugar transport.[44] 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 DAFs (Fig. 2a), which may suggest a close involvement in the seed phytate biosynthesis.

4.3. 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.[1] In previous studies, low phytate (*lpa*) mutants have been generated by random mutagenesis in various crops. In some studies, germination rate, thousand-seed weight, seed yield, and vigour of the *lpa* mutant decreased compared with wild type, which adversely affects breeding value.[33,45–47] However, in other studies *lpa* mutants had little or no effect on field performance. For example in barley, Raboy et al. (2015) measured grain yield of six *lpa* mutations, one of which had no detectable effect on yield, and the other five had yields similar to or within 5% of the wild-type sibling iso-line.[46] In common bean, *lpa* mutants (*lpa*-280-10) had a higher germination rate than the wild type, with seed yield close to that of its parents.[47] 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 phytate concentration or phytate content in the selected varieties (Fig. 3a to c, Supplementary Table S10). Recently, CRISPR-Cas9 mutagenesis has been adopted to knock out three essential *BnITPK* genes to obtain a low phytate line.[14] Compared with wild type, the low phytate mutant had no negative effect on germination rate and seedling vigor.[14] 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 (Fig. 1c to j, Supplementary Tables S6 to S9). There were no significant difference in seed yield, PH, EBN, or thousand-seed weight between cultivars with low and high phytate haplotypes, either at a sufficient P or deficient P supply (Fig. 4, Supplementary Table S11). Marzo et al.[48] 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.[49] Any obvious difference in seed germination, disease resistance, and insect resistance between the two haplotypes of *B. napus* was 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 $T_P C_E T_E G_I A_E T_E$ 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' 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

Conflict of interest

We declare that we have no conflict of interest.

Supplementary data

Supplementary data are available at DNARES online.

References

1. Raboy, V. 2020, Low phytate crops: observations based on four decades of research, *Plants*, **9**, 140.
2. Saxena, S.C., Salvi, P., Kaur, H., et al. 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**, 5623–39.
3. Shears, S.B. 2004, How versatile are inositol phosphate kinases?, *Biochem. J.*, **377**, 265–80.
4. Shi, J., Wang, H., Schellin, K., et al. 2007, Embryo-specific silencing of a transporter reduces phytate content of maize and soybean seeds, *Nat. Biotechnol.*, **25**, 930–7.
5. Xu, X.H., Zhao, H.J., Liu, Q.L., et al. 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**, 75–83.
6. Ali, N., Paul, S., Gayen, D., et al. 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**, e68161.
7. Ali, N., Paul, S., Gayen, D., et al. 2013, RNAi mediated down regulation of myo-inositol-3-phosphate synthase to generate low phytate rice, *Rice*, **6**, 12.
8. 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**, 355–67.
9. Li, W.-X., Huang, J.-Z., Zhao, H.-J., et al. 2014, Production of low phytate rice by hairpin RNA- and artificial microRNA-mediated silencing of OsMIK in seeds, *Plant Cell. Tiss. Organ Cult.*, **119**, 15–25.
10. Khan, M.S.S., Basnet, R., Islam, S.A., et al. 2019, Mutational analysis of OsPLDalpha1 reveals its involvement in phytate biosynthesis in rice grains, *J. Agric. Food Chem.*, **67**, 11436–43.
11. Kishor, D.S., Lee, C., Lee, D., et al. 2019, Novel allelic variant of Lpa1 gene associated with a significant reduction in seed phytate content in rice (*Oryza sativa* L), *PLoS One*, **14**, e0209636.
12. Laha, D., Parvin, N., Hofer, A., et al. 2019, Arabidopsis ITPK1 and ITPK2 have an evolutionarily conserved phytate kinase activity, *ACS Chem. Biol.*, **14**, 2127–33.
13. Sashidhar, N., Harloff, H.J. and Jung, C. 2020, Identification of phytate mutants in oilseed rape (*Brassica napus*) by large-scale screening of mutant populations through amplicon sequencing, *New Phytol.*, **225**, 2022–34.
14. Sashidhar, N., Harloff, H.J., Potgieter, L., et al. 2020, Gene editing of three BnITPK genes in tetraploid oilseed rape leads to significant reduction of phytate in seeds, *Plant Biotechnol. J.*, **18**, 2241–50.
15. Brown, K.H. and Solomons, N.W. 1991, Nutritional problems of developing countries, *Infect. Dis. Clin. North Am.*, **5**, 297–317.
16. Bentsink, L., Yuan, K., Koornneef, M., et al. 2003, The genetics of phytate and phosphate accumulation in seeds and leaves of *Arabidopsis thaliana*, using natural variation, *Theor. Appl. Genet.*, **106**, 1234–43.
17. Zhao, J., Jamar, D.C., Lou, P., et al. 2008, Quantitative trait loci analysis of phytate and phosphate concentrations in seeds and leaves of *Brassica rapa*, *Plant. Cell Environ.*, **31**, 887–900.
18. Nordborg, M. and Weigel, D. 2008, Next-generation genetics in plants, *Nature*, **456**, 720–3.
19. Yang, M., Lu, K., Zhao, F.J., et al. 2018, Genome-wide association studies reveal the genetic basis of Ionomic variation in rice, *Plant Cell*, **30**, 2720–40.
20. Zhao, Z., Zhang, H., Fu, Z., et al. 2018, Genetic-based dissection of arsenic accumulation in maize using a genome-wide association analysis method, *Plant Biotechnol. J.*, **16**, 1085–93.
21. Welling, M.T., Liu, L., Kretzschmar, T., et al. 2020, An extreme-phenotype genome-wide association study identifies candidate cannabinoid pathway genes in Cannabis, *Sci. Rep.*, **10**, 18643.
22. Chalhoub, B., Denoeud, F., Liu, S., et al. 2014, Early allopolyploid evolution in the post-Neolithic *B. napus* oilseed genome, *Science*, **345**, 950–3.
23. Wang, X., Chen, Y., Thomas, C.L., et al. 2017, Genetic variants associated with the root system architecture of oilseed rape (*Brassica napus* L.) under contrasting phosphate supply, *DNA Res.*, **24**, 407–17.
24. Tang, S., Zhao, H., Lu, S., 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**, 470–87.
25. Haung, W. and Lantzsch, H.J. 1983, Sensitive method for the rapid determination of phytate in cereals and cereal products, *J Sci Food Agric.*, **34**, 1423–6.
26. Bradbury, P.J., Zhang, Z., Kroon, D.E., et al. 2007, TASSEL: software for association mapping of complex traits in diverse samples, *Bioinformatics*, **23**, 2633–5.

27. Alexander, D.H., Novembre, J. and Lange, K. 2009, Fast model-based estimation of ancestry in unrelated individuals, *Genome Res.*, **19**, 1655–64.

28. Zhang, C., Dong, S.S., Xu, J.Y., et al. 2019, PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files, *Bioinformatics*, **35**, 1786–8.

29. Nagy, R., Grob, H., Weder, B., et al. 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**, 33614–22.

30. Panzeri, D., Cassani, E., Doria, E., et al. 2011, A defective ABC transporter of the MRP family, responsible for the bean lpa1 mutation, affects the regulation of the phytic acid pathway, reduces seed myo-inositol and alters ABA sensitivity, *New Phytol.*, **191**, 70–83.

31. 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.), *EXBOTJ.*, **66**, 5683–90.

32. Perera, I., Fukushima, A., Arai, M., et al. 2019, Identification of low phytate and high Zn bioavailable rice (*Oryza sativa* L.) from 69 accessions of the world rice core collection, *J Cereal Sci.*, **85**, 206–13.

33. 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**, 1336–9.

34. Oltmans, S.E., Fehr, W.R., Welke, G.A., et al. 2005, Agronomic and seed traits of soybean lines with low–phytate P, *Crop Sci.*, **45**, 593–8.

35. Nakamura, Y., Yuki, K. and Park, S.Y. 2012, Carbohydrate metabolism in the developing endosperm of rice grains, *Plant Cell Physiol.*, **1898**, 833–9.

36. Iwai, T., Takahashi, M., Oda, K., et al. 2012, Dynamic changes in the distribution of minerals in relation to phytate accumulation during rice seed development, *Plant Physiol.*, **160**, 2007–14.

37. Yamaji, N., Takemoto, Y., Miyaji, T., et al. 2017, Reducing phosphorus accumulation in rice grains with an impaired transporter in the node, *Nature*, **541**, 92–5.

38. Tong, C., Chen, Y., Tan, Y., et al. 2017, Analysis of lysophospholipid content in low phytate rice mutants, *J. Agric. Food Chem.*, **65**, 5435–41.

39. Zhou, Q., Zhou, C., Zheng, W., et al. 2017, Genome-wide SNP markers based on SLAF-seq uncover breeding traces in rapeseed, *Front. Plant Sci.*, **8**, 648.

40. Xiao, Y., Cai, D., Yang, W., et al. 2012, Genetic structure and linkage disequilibrium pattern of a rapeseed (*Brassica napus* L.) association mapping panel revealed by microsatellites, *Theor. Appl. Genet.*, **125**, 437–47.

41. Wang, X., Ren, M., Liu, D., et al. 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.*, **62**, 1469–1484.

42. Lai, C.P., Huang, L.M., Chen, L.O., et al. 2017, Genome-wide analysis of GDSL-type esterases/lipases in Arabidopsis, *Plant Mol. Biol.*, **95**, 181–97.

43. Wang, Y., Zhang, W.Z., Song, L.F., et al. 2008, Transcriptome analyses show changes in gene expression to accompany pollen germination and tube growth in Arabidopsis, *Plant Physiol.*, **148**, 1201–11.

44. Parsons, H.T., Christiansen, K., Knierim, B., et al. 2012, Isolation and proteomic characterization of the Arabidopsis Golgi defines functional and novel components involved in plant cell wall biosynthesis, *Plant Physiol.*, **159**, 12–26.

45. Naidoo, R., Tongoona, P., Derera, J., et al. 2012, Combining ability of low phytic acid (lpa1-1) and quality protein maize (QPM) lines for seed germination and vigour under stress and non-stress conditions, *Euphytica*, **185**, 529–41.

46. Raboy, V., Peterson, K., Jackson, C., et al. 2015, A substantial fraction of barley (*Hordeum vulgare* L.) low phytic acid mutations have little or no effect on yield across diverse production environments, *Plants*, **4**, 225–39.

47. Campion, B., Sparvoli, F., Doria, E., et al. 2009, Isolation and characterisation of an lpa (low phytic acid) mutant in common bean (*Phaseolus vulgaris* L.), *Theor. Appl. Genet.*, **118**, 1211–21.

48. Marzo, F., Andres, A., Castiella, M.V., et al. 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**, 1829–33.

49. Dhole, V.J. and Reddy, K.S. 2016, Association of phytate content with biotic stress tolerance in mungbean (*Vigna radiata* L. Wilczek), *Phytoparasitica*, **44**, 261–7.