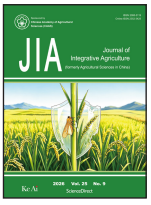
Journal of Integrative Agriculture 2026, 25(9): 3585–3594

Available online at ScienceDirect

Journal of Integrative Agriculture



RESEARCH ARTICLE

Genome- and transcriptome-wide association studies reveal the genetic basis of seed palmitic acid content in *Brassica napus*

Haijiang Liu[1, 2, 3], Yongheng Yuan[1, 2, 3], Yunshan Tang[1, 2, 3], Ruoshui Li[1, 2, 3], Kaijie Ye[1, 2, 3], Mengzhen Zhang[1, 2, 3], Kun Lu[1, 2, 3], Nengwen Yin[1, 2, 3], Huiyan Zhao[1, 2, 3], Yuanyuan Liu[1, 2, 3], Taocui Huang[4], Rui Wang[1, 2, 3], Lei Shi[5], Hai Du[1, 2, 3#], Cunmin Qu[1, 2, 3#]

[1] *Integrative Science Center of Germplasm Creation in Western China (Chongqing) Science City/College of Agronomy and Biotechnology, Southwest University, Chongqing 401335, China*
[2] *Engineering Research Center of South Upland Agriculture, Ministry of Education/Southwest University, Chongqing 400715, China*
[3] *Academy of Agricultural Sciences, Southwest University, Chongqing 400715, China*
[4] *Chongqing Academy of Agricultural Sciences, Chongqing 400216, China*
[5] *National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan 430070, China*

Highlights

- Four candidate genes affecting seed palmitic acid content in *Brassica napus* were identified with combined genome-wide association study and transcriptome-wide association study.
- Four superior haplotypes were identified, and *B. napus* varieties carrying these haplotypes had lower palmitic acid content without affecting the seed oil content, protein content, or seed yield.
- Based on a functional single-nucleotide polymorphism, molecular marker *Bn_A8_SPAC_Marker* was developed to accelerate the breeding of rapeseed with low palmitic acid content.

Abstract

Rapeseed (*Brassica napus* L.) is one of the most important oilseed crops worldwide, and the development of rapeseed varieties with high-quality oil is a long-term breeding goal. Reducing the content of palmitic acid, the main saturated fatty acid in rapeseed oil, can greatly improve oil quality. Here, a genome-wide association study (GWAS) and transcriptome-wide association study (TWAS) of seed palmitic acid content (SPAC) were performed using 393 diverse *B. napus* accessions. Four genes (*BnaA08.DAP*, *BnaA08.PAA1*, *BnaA08.DUF106*, and *BnaC03.DAP*) were identified by both GWAS and TWAS. The transcripts per million (TPM) values of these candidate genes at 20 and 40 days after flowering (DAF) were significantly correlated with SPAC in this association panel. Based on genetic variation in the candidate genes, four low-SPAC haplotypes were identified by combining candidate gene association analysis and haplotype analysis. *Brassica napus* accessions carrying low-SPAC haplotypes had lower SPAC than those carrying high-SPAC haplotypes without affecting seed oil content, seed protein content, or seed yield. Based on the functional single-nucleotide polymorphism (SNP) chrA08_9529850 (C/A) in the promoter of *BnaA08.DUF106*, a molecular marker (*Bn_A8_SPAC_Marker*) was developed that can be used to facilitate breeding for low SPAC in *B. napus*. Our findings provide valuable information for studying the genetic control of SPAC in *B. napus*. Moreover, the candidate genes, favorable haplotypes, and molecular marker identified in this study will be useful for breeding low-SPAC *B. napus* varieties.

Keywords: rapeseed, genome-wide association study, transcriptome-wide association study, seed palmitic acid content, haplotype analysis, molecular marker

Received 28 August 2024; Received in revised form 27 September 2024; Accepted 8 October 2024; Available online 5 November 2024
[#]Correspondence Cunmin Qu, E-mail: drqucunmin@swu.edu.cn; Hai Du, E-mail: haidu81@126.com

doi: 10.1016/j.jia.2024.11.015

1. Introduction

Rapeseed (*Brassica napus*) is the second largest oilseed crop after soybean (*Glycine max*) worldwide (Ge *et al*. 2025). Improving the quality of rapeseed oil will improve human health and significantly increase farmer incomes (Li *et al.* 2022). The fatty acid composition of rapeseed oil includes saturated and unsaturated fatty acids, and the unsaturated fatty acids include oleic, linoleic, linolenic, and erucic acids, while the major saturated fatty acids are palmitic and stearic acids (Huang *et al.* 2021). The long-term consumption of edible oils with high palmitic acid contents can increase low-density lipoprotein cholesterol levels in the blood, thereby increasing the risk of cardiovascular disease (Mancini *et al.* 2015; Guyton *et al.* 2023). Therefore, reducing the saturated fatty acid contents in rapeseed oil, especially palmitic acid, is crucial for improving oil quality.

In recent years, the efficiency of breeding has been greatly improved by mining candidate genes that affect target traits and using genetic engineering techniques (Khlestkina and Shavrukov 2022; Cortés and Du 2023). Overexpressing *GmWRI1c*, a homolog of *Arabidopsis thaliana WRI1*, significantly increased both the seed palmitic acid content (SPAC) in transgenic soybean (*Glycine max*) and the root nodule number (Zheng *et al.* 2022). However, overexpressing *GmWRI1a*, another homolog of *WRI1*, reduced the SPAC from 15 to 8% while increasing the seed oil content (SOC) from 22 to 31% (Wang *et al.* 2022). Knockout of *FATB1A* reduced the SPAC in transgenic soybean seeds from 18 to 11% (Rajcan *et al.* 2022). Silencing the *FATB* gene in transgenic *B. napus* reduced SPAC from 8 to 4.4% compared to the wild type (Belide *et al*. 2012). Overexpressing *BnaFAX1-1* significantly reduced the SPAC and significantly increased the seed oleic acid content of transgenic *B. napus* (Xiao *et al.* 2021). Therefore, exploring the candidate genes affecting SPAC in *B. napus* and identifying favorable alleles conferring low SPAC with no effects on SOC or seed yield could facilitate the development of high-quality *B. napus* through molecular breeding.

Genome-wide association study (GWAS) has been widely used for the genetic analysis and gene mining of seed quality traits in various crops such as rice (*Oryza sativa*) (Mai *et al.* 2023), maize (*Zea mays*) (Li *et al.* 2019), soybean (Qi *et al.* 2023), and rapeseed (Gacek *et al.* 2017; Qu *et al.* 2017). For example, Qi *et al.* (2023) conducted GWAS on the contents of six fatty acids in 547 soybean varieties and identified a multi-effect QTL on chromosome A09 (Qi *et al.* 2023). By combining fatty acid profiling, genome resequencing, and population genomic analysis, the candidate gene *FA9* was identified and knocking out *Gm.FA9* significantly increased the SPAC in soybean (Qi *et al.* 2023). Qu *et al.* (2017) conducted GWAS on the fatty acid contents of 520 *B. napus* varieties, and identified six genetic regions significantly associated with fatty acid content and 24 candidate genes. Facilitated by the continuous development of sequencing technology, the large-scale whole-genome resequencing of *B. napus* has yielded millions of single-nucleotide polymorphism (SNP) markers, leading to the identification of many important candidate genes by GWAS (Liu *et al.* 2022b).

This study examined SPAC in a diverse panel of 393 *B. napus* accessions and identified nine significant QTLs associated with this trait by GWAS. By combining this analysis with a transcriptome-wide association study (TWAS), four candidate genes (*BnaA08.DAP*, *BnaA08.PAA1*, *BnaA08.DUF106*, and *BnaC03.DAP*) associated with SPAC were identified. In addition, several low-SPAC haplotypes were identified and a molecular marker (*Bn_A8_SPAC_Marker*) was developed that can be used for low-SPAC molecular marker-assisted breeding in the future. These findings increase our understanding of the genetic basis of SPAC in *B. napus* and can facilitate the breeding of low-SPAC *B. napus* cultivars.

2. Materials and methods

2.1. Plant materials and field trials

Detailed information on the 393 *B. napus* accessions and genotypes used in this study was published previously (Liu *et al.* 2021). All *B. napus* accessions were planted in three different environments in China: Wuhan, Hubei Province (Trial 1; 30.3°N, 113.6°E), Sansheng, Chongqing (Trial 2; 29.89°N, 106.62°E), and Xiema, Chongqing (Trial 3; 29.72°N, 106.36°E). Each accession was grown in a plot with three rows. Each plot contained 12 plants per row, with 20 cm between plants in each row and 20 cm between rows. After harvesting, the seeds of each accession were scanned with an NIRS DS2500 near-infrared analyzer (FOSS, Hillerød, Denmark) to determine the SPAC. Each accession was scanned three times, and the average of the three scans was used as the final phenotypic value. The mean value, median, maximum, minimum, range, and coefficient of variation of this association panel were calculated using the R package "psych" (https://cran.r-project.org/web/packages/psych/). The R language (4.2.3) was used to calculate the correlation coefficients between different trials through paired analysis using the Pearson method.

2.2. Genome-wide association study

In a previous study, whole-genome resequencing was performed on 505 *B. napus* accessions, and over 10 million SNPs were developed based on the Darmor-*bzh* reference genome (Tang *et al.* 2021). In this study, minor-allele frequency (MAF)>0.05 and missing rate<0.2 were used as the screening criteria to filter those SNPs, and a total of 4.6 million high-quality SNPs were identified for subsequent GWAS. GWAS for SPAC was conducted using factored spectrally transformed linear mixed models (Fast-LMM) (Lippert *et al.* 2011). CMplot was used to construct Manhattan plots and quantile–quantile plots of the data (Yin *et al*. 2021). The linkage disequilibrium (LD) decay value of this association panel (298 kb) was reported previously (Zhang *et al.* 2019; Liu *et al.* 2021). The genes located within the LD decay value (300 kb) upstream and downstream of the lead SNP were considered the candidate genes associated with SPAC.

2.3. Transcriptome-wide association study

Tang *et al*. (2021) previously performed transcriptomic

sequencing of the seeds from 382 *B. napus* accessions at 20 and 40 days after flowering (DAF) in this *B. napus* association panel; and the resulting data were uploaded to the *BnIR* database (Tang *et al.* 2021; Yang *et al.* 2023). In this study, the TWAS of SPAC in this association panel was performed using the *BnIR* database, with 1×10^{-5} used as the threshold for screening candidate genes significantly associated with SPAC (Yang *et al.* 2023). The correlation coefficients between the transcripts per million (TPM) values of the candidate genes at 20 and 40 DAF and SPAC in this accession panel were calculated using the *BnIR* database (Yang *et al.* 2023). A Gene Ontology (GO) enrichment analysis of genes significantly associated with SPAC identified through TWAS was performed using OmicShare (https://www.omicshare.com/).

2.4. Phenotypic analysis of seed quality traits and seed yield between *LSPACHap* and *HSPACHap*

Forty-nine *B. napus* accessions carrying low-SPAC haplotypes at all four candidate loci (hereafter *LSPACHap*; see Results) and 164 accessions carrying high-SPAC haplotypes at these loci (*HSPACHap*) were examined in this study. These varieties were previously planted in Meichuan Town, Wuxue City, Hubei Province, China (29.85°N, 115.55°E) from 2019 to 2020 (Liu *et al.* 2022a, c). After maturity, the seed yield was determined and the seeds of each inbred line were scanned with an NIRS DS2500 near-infrared analyzer to measure the seed oil and protein contents. Student's *t*-test was used to compare the differences in the above traits among lines with the *LSPACHap* and *HSPACHap* haplotypes.

2.5. Association analysis and haplotype analysis of candidate genes

The genotypes of *BnaA08.DAP*, *BnaA08.PAA1*, *BnaA08.DUF106*, and *BnaC03.DAP* in the *B. napus* association panel were obtained using VCFtools software (Danecek *et al.* 2011). An association analysis of the candidate genes *BnaA08.DAP*, *BnaA08.PAA1*, *BnaA08.DUF106*, and *BnaC03.DAP* was performed using the GLM model in Tassel 5.0 software (Bradbury *et al.* 2007). SNP markers in the genomic region from 2 kb upstream of the candidate gene to the termination codon were used to conduct an association analysis with the SPAC of the *B. napus* association panel.

2.6. Development of molecular markers

The sequences of the candidate genes and their homologs were downloaded from the *BnIR* database, and sequence alignment was performed using DNAMAN (https://www.lynnon.com/dnaman.html). SnapGene software (https://www.genoscope.cns.fr/brassicanapus/) was used to analyze the sequences of functional SNPs significantly associated with SPAC in the candidate genes and to identify the SNPs causing changes in restriction endonuclease recognition sites. Specific primers at ~500 bp upstream and downstream of the target SNPs were designed using AmplifX software, and the primers are listed in Appendix A. Five accessions for each of the *LSPACHap* and *HSPACHap* haplotypes were selected for analysis. Following DNA extraction, PCR amplification was performed using the primers listed in Appendix A. The PCR products were digested using restriction enzymes corresponding to the target SNP to verify the accuracy of the newly developed molecular marker.

3. Results

3.1. Phenotypic variation for SPAC in an association panel of *B. napus*

This study investigated SPAC in an association panel of 393 *B. napus* accessions grown in field trials in Wuhan (Trial 1), Sansheng (Trial 2), and Xiema (Trial 3). Extensive phenotypic variation for SPAC was found in the association panel (Table 1), which ranged from 2.98 to 5.6% (1.87-fold) in Trial 1, 3.06 to 5.86% (1.91-fold) in Trial 2, and 3.23 to 5.97% (1.84-fold) in Trial 3 (Table 1). The coefficient of variation values between different trials ranged from 9.51 to 12.89% (Table 1). The SPAC of this association panel was highly correlated between the different trials. The correlation coefficients were 0.70 between Trial 1 and Trial 2, 0.73 between Trial 1 and Trial 3, and 0.82 between Trial 2 and Trial 3 (Appendix B), laying the foundation for the subsequent identification of stable SPAC genetic regions in *B. napus* by GWAS.

3.2. GWAS of SPAC in *B. napus*

The GWAS used factored spectrally transformed linear mixed models (Fast-LMM) to identify genetic region associated with SPAC in *B. napus* (Fig. 1; Appendix C). A total of eight stable QTLs were significantly associated with SPAC in the different trials. These QTLs were distributed on chromosomes A08, A09, C03, C04, and C06, with phenotypic variation explained (PVE) ranging from 18.69 to 47.66% ($P<1\times10^{-10}$) (Fig. 1; Table 2). Among these, five, seven, six, and eight QTLs significantly associated with SPAC were identified in Trial 1, Trial 2, Trial 3, and the mean value (MV), respectively (Fig. 1; Table 2). All the significant QTLs were identified in more than one trial (including MV), which showed high reliability (Fig. 1; Table 2). For example, the significant QTL (*qSPAC.A08.2*) on chromosome A08 could explain 47.86% of the variation in SPAC and was identified simultaneously in all trials (Fig. 1; Table 2). Subsequently, eight low-SPAC alleles were identified based on the lead SNPs among the above stable QTLs (Fig. 2). For example, the T allele in chrA08_229357 (C/T), G allele in chrA08_10194057 (A/G), A allele in chrA09_3050970 (G/A), and C allele in chrA09_31541180 (A/C) were all low-SPAC alleles (Fig. 2-A). A highly significant negative correlation was found between the number of low-SPAC alleles and SPAC (*P*-value, 2.2e–16; *R*, −0.641) (Fig. 2-B).

Table 1 Statistical analysis of seed palmitic acid (C16:0) content in the association panel

Environment[1)]	Mean	Range	CV (%)
WH (Trial 1)	4.42	2.98–5.60	12.89
SS (Trial 2)	4.68	3.06–5.86	10.85
XM (Trial 3)	4.92	3.23–5.97	9.51
MV	4.65	3.24–5.63	10.49

[1)] WH, Wuhan, Hubei Province, China; SS, Sansheng, Chongqing, China; XM, Xiema, Chongqing; MV, mean value.

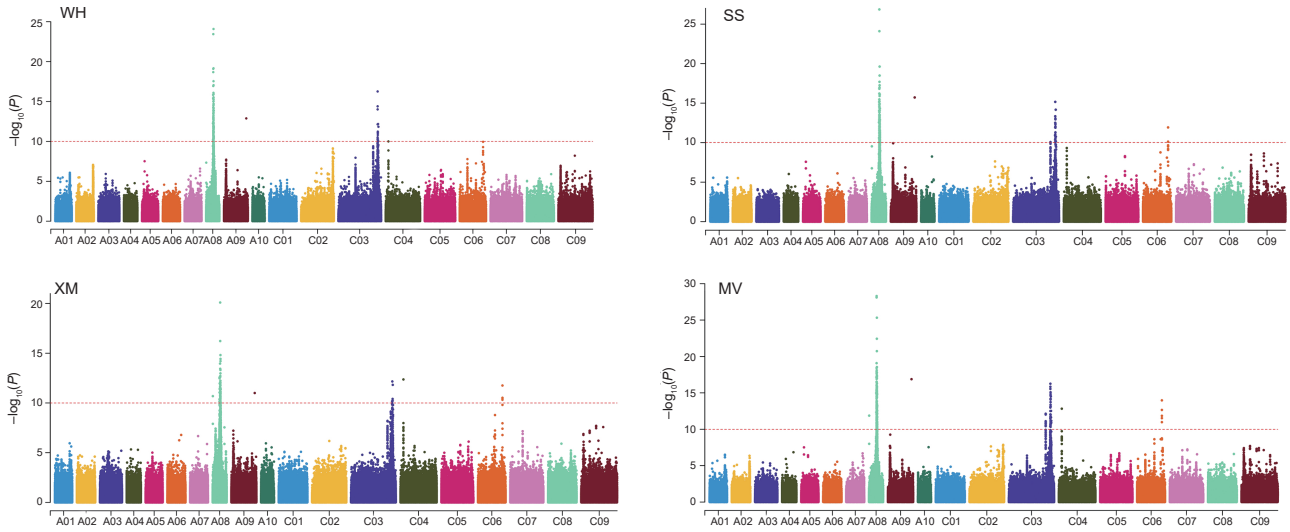


Fig. 1 Genome-wide association study (GWAS) of seed palmitic acid content (SPAC) in the Wuhan (WH, Hubei Province), Sansheng (SS, Chongqing), Xiema (XM, Chongqing), China and mean value (MV) environments. The dashed line represents the threshold for screening significant QTLs.

Table 2 Genome-wide signals significantly associated with seed palmitic acid content in *Brassica napus*

ID	Lead SNP	Position (bp)	PVE (%)	Environment[1)]
qSPAC.A08.1	ChrA08_229357	229,357	18.69	XM/MV
qSPAC.A08.2	ChrA08_10194057	10,194,057	47.66	WH/XM/SS/MV
qSPAC.A09.1	ChrA09_3050970	3,050,970	18.74	SS/MV
qSPAC.A09.2	ChrA09_31541180	31,541,180	23.65	WH/XM/SS/MV
qSPAC.C03.1	ChrC03_49325739	49,325,739	21.62	SS/MV
qSPAC.C03.2	ChrC03_55621804	55,621,804	33.27	WH/XM/SS/MV
qSPAC.C04	ChrC04_3880361	3,880,361	27.53	WH/XM/SS/MV
qSPAC.C06	ChrC06_33011759	33,011,759	29.28	WH/XM/SS/MV

[1)] WH, Wuhan, Hubei Province, China; SS, Sansheng, Chongqing, China; XM, Xiema, Chongqing; MV, mean value.

3.3. Transcriptome-wide association study of SPAC in *B. napus* and candidate gene identification

In previous studies, transcriptome sequencing was performed on the seeds of 382 accessions in an association panel of *B. napus* at 20 and 40 DAF (Tang *et al.* 2021). In the current study, this transcriptome dataset was subjected to TWAS of SPAC in the publicly available *BnIR* database, which identified 16 and 62 candidate genes significantly associated with SPAC at 20 and 40 DAF, respectively (Fig. 3; Appendix D). GO enrichment analysis revealed that these candidate genes were mainly enriched in hydro-lyase activity (GO:0016836), carbonate dehydratase activity (GO:0004089), identical protein binding (GO:0042802), methylglutaconyl-CoA hydratase activity (GO:0004490), and L,L-diaminopimelate aminotransferase activity (GO:0010285), among others (Appendix E).

The candidate genes identified by TWAS were then compared with the genetic regions identified by GWAS. Four genes (*BnaA08.DAP*, *BnaA08.PAA1*, *BnaA08.DUF106*, and *BnaC03.DAP*) showed significant correlations with SPAC based on the TPM values in seeds at 20 and 40 DAF in the association panel (Fig. 3). For example, the TPM values of *BnaA08.DAP* in seeds at 20 DAF (*P*-value, 4.63e–27; *R*, –0.333) and 40 DAF (*P*-value, 2.02e–14; *R*, –0.188) showed significant negative correlations with SPAC in the association panel (Fig. 3-C and D). In addition, the TPM values of *BnaA08.PAA1* (on chromosome A08) in seeds at 20 DAF (*P*-value, 1.96e–20; *R*, −0.259) and 40 DAF (*P*-value, 3.90e–31; *R*, –0.392) showed significant negative correlations with SPAC (Fig. 3-C and D). These genes may play important roles in determining the SPAC of *B. napus*.

3.4. Identification of favorable haplotypes

To identify favorable allelic variation in the above candidate genes and to increase their value for breeding, the SNPs in the 2 kb promoter regions and coding regions of the four candidate genes were extracted, and an association analysis of the candidate genes based on the GLM model was conducted to identify the SNPs significantly associated with SPAC (Fig. 4). For example, the association analysis of *BnaA08.DAP* revealed 10 SNPs significantly associated with SPAC (Fig. 4-A). Haplotype analysis based on the lead SNP chrA08_10438041 (C/T) (*P*-value, 2.74e–19; PVE, 22.16%) in *BnaA08.DAP* revealed that the SPAC of *B. napus* accessions carrying *BnaA08.DAP.Hap1* (T allele) was much lower than that of accessions carrying *BnaA08.DAP.Hap2* (C allele) (Fig. 4-A).

Fourteen SNPs in *BnaA08.PAA1* were significantly associated with SPAC. Among these, chrA08_10480638 (C/A) showed the strongest association (*P*-value, 2.66e–22),

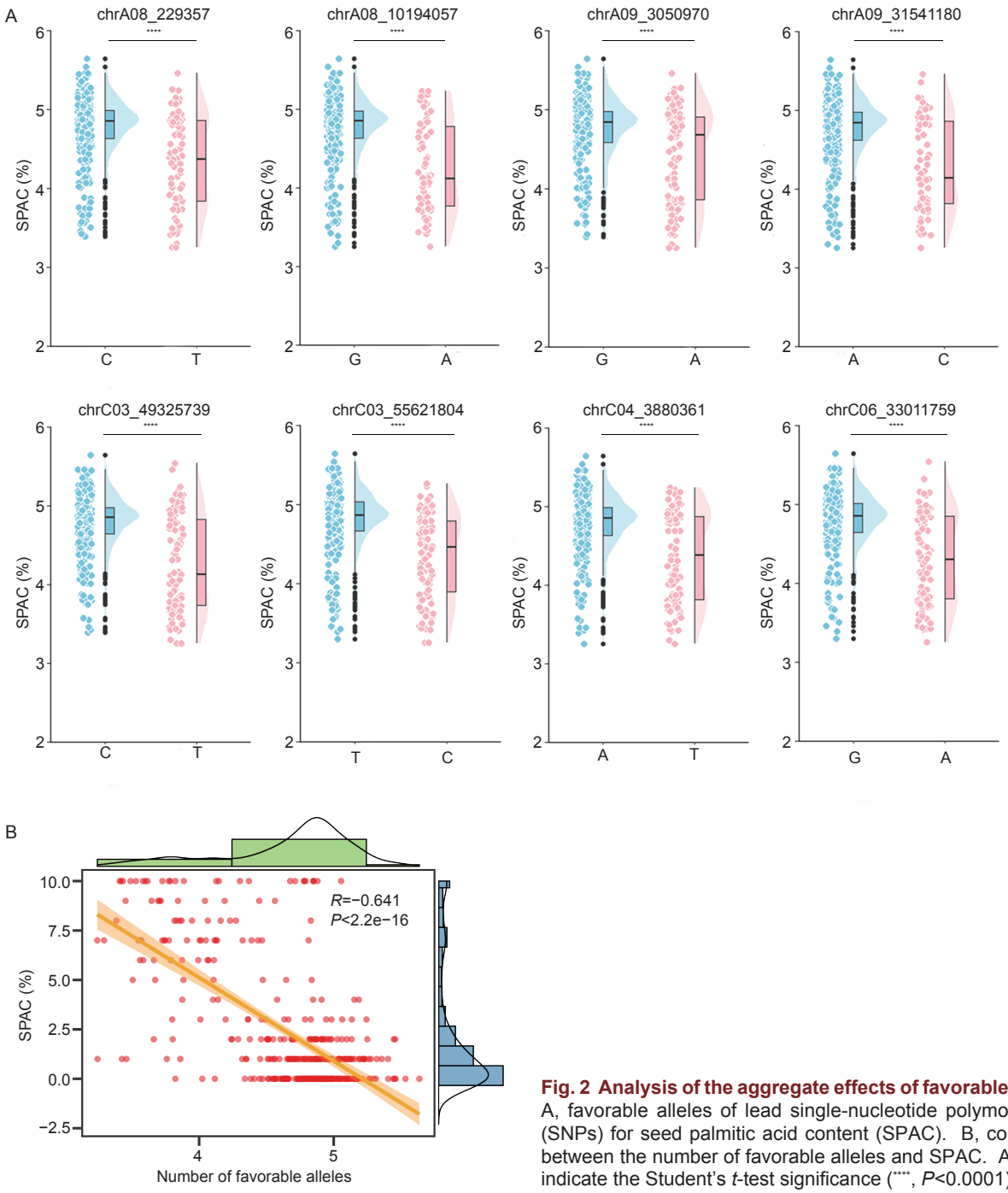


Fig. 2 Analysis of the aggregate effects of favorable alleles. A, favorable alleles of lead single-nucleotide polymorphisms (SNPs) for seed palmitic acid content (SPAC). B, correlation between the number of favorable alleles and SPAC. Asterisks indicate the Student's *t*-test significance (****, $P<0.0001$).

with a PVE of 22.79% (Fig. 4-B). Further analysis demonstrated that *BnaA08.PAA1Hap1* (A allele) was a low-SPAC haplotype (Fig. 4-B). Six SNPs in *BnaA08.DUF106* were significantly associated with SPAC (Fig. 4-C), among which chrA08_10486876 (A/C) showed the most significant correlation (*P*-value, 5.77e–15), with a PVE of 16.85% (Fig. 4-C). Further analysis demonstrated that *BnaA08.DUF106Hap1* (C allele) was another low-SPAC haplotype (Fig. 4-B). On chromosome C03, seven SNPs in *BnaC03.DAP* were significantly associated with SPAC (Fig. 4-D). Haplotype analysis based on the lead SNP chrC03_56246200 (C/T) (*P*-value, 2.81e–11; PVE, 12.65%) in *BnaC03.DAP* revealed that the SPAC of *B. napus* accessions carrying *BnaC03.DAPHap1* (T allele) was much lower than that of accessions carrying *BnaC03.DAPHap2* (C allele) (Fig. 4-D).

To investigate the effects of the above low-SPAC alleles on seed quality traits and seed yield, the low- and high-SPAC alleles of the above four candidate genes (*BnaA08.DAP*, *BnaA08.PAA1*, *BnaA08.DUF106*, and *BnaC03.DAP*) were counted in each accession of the association panel (Fig. 5). Forty-nine accessions carried all low-SPAC alleles (*LSPACHap*) and 164 accessions carried all high-SPAC alleles (*HSPACHap*) (Fig. 5). As shown in Fig. 5 and Table 3, the average SPAC for *LSPACHap* was 3.52%, which was much lower than the 4.66% value for *HSPACHap*. The average SOC and seed protein content (SPC) of *LSPACHap*

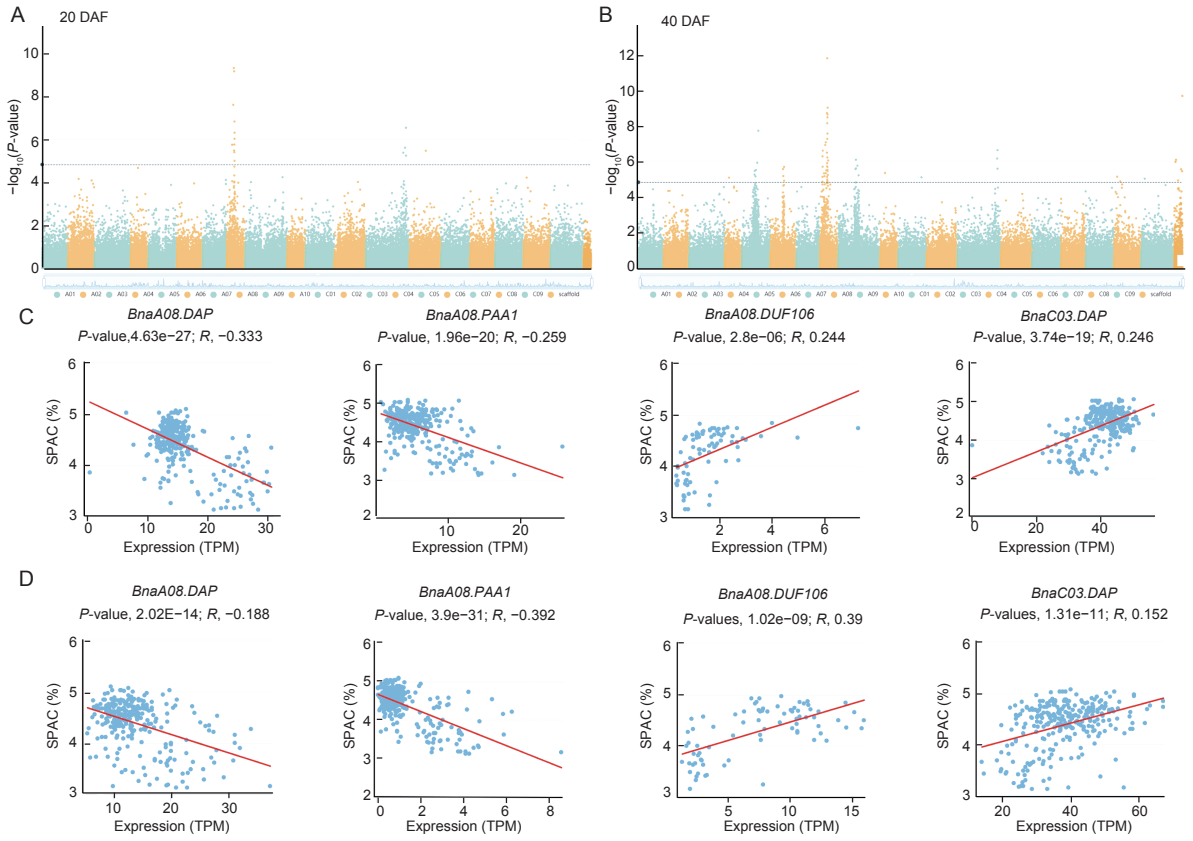


Fig. 3 Transcriptome-wide association studies (TWASs) for seed palmitic acid content (SPAC). A and B, Manhattan plots of the SPAC at 20 days after flowering (DAF) (A) and 40 DAF (B). C and D, correlation analysis between SPAC and the transcripts per million (TPM) expression values of four candidate genes in 20 DAF (C) and 40 DAF (D) seeds of the 382 *Brassica napus* inbred lines.

were 40.21 and 26.41%, respectively, which were not significantly different from the values of *HSPACHap* (40.27 and 25.91%, respectively). However, *LSPACHap* slightly reduced the seed oleic acid content in *B. napus* (53.87 to 48.91%), while the seed linoleic acid content was not affected. Most importantly, *LSPACHap* did not affect the seed yield in *B. napus* (Fig. 5; Table 3). These findings suggest that *LSPACHap* can be used as a high-quality haplotype for the improvement of *B. napus* varieties.

3.5. Development of a cleaved amplified polymorphic sequence (CAPS) molecular marker to facilitate low-SPAC breeding

Among the SNP loci significantly associated with SPAC in the above candidate genes, chrA08_9529850 (C/A) is located in the specific recognition sequence of the restriction endonuclease *Nco*I. Of the 393 accessions examined, 35 carried the A allele in chrA08_9529850 (C/A), and 299 accessions carried the C allele in chrA08_9529850 (C/A). The SPAC of the two types of *B. napus* inbred lines showed significant differences in SS, WH, XM, and MV, but did not affect SOC, SPC or yield traits (Fig. 6-A; Appendix F). Importantly, chrA08_9529850 (C/A) is located in the promoter of gene *BnaA08.DUF106*. The TPM values of *BnaA08.DUF106* were significantly correlated with SPAC in 385 accessions at both 20 and 40 DAF (Figs. 3-B and C and 6-B). Therefore, chrA08_9529850 (C/A) was selected to develop molecular markers.

The accessions L60, L198, L226, L262, and L340 all carried the C allele in chrA08_9529850 (C/A), and the SPAC values of these accessions ranged from 4.62 to 5.04% (Appendix G). The accessions L732, L984, L994, L976, and L1014 all carried the A allele in chrA08_9529850 (C/A), and the SPAC values of these accessions ranged from 3.14 to 3.71% (Appendix G). The genomic fragments of these alleles were amplified by PCR and digested with *Nco*I. Fragments amplified from the five high-SPAC accessions were cut into two fragments of 333 and 157 bp by *Nco*I, whereas the amplified fragments from the five low-SPAC accessions could not be digested by *Nco*I (Fig. 6-C), indicating that the haplotypes could be distinguished by restriction enzyme digestion. In summary, this CAPS molecular marker named *Bn_A8_SPAC_Marker* can be used to identify low-SPAC cultivars based on SNP chrA08_9529850 (C/A).

4. Discussion

4.1. Identification of genes affecting SPAC in *B. napus*

In previous studies, extensive phenotypic variations for SPAC were observed in different populations of *B. napus* (Gacek *et al.* 2017; Guan *et al.* 2019). For example, the SPAC varied from 3.80 to 5.91% among 160 *B. napus* accessions (Gacek *et al.* 2017) and from 1.26 to 5.32% among 435 *B. napus*

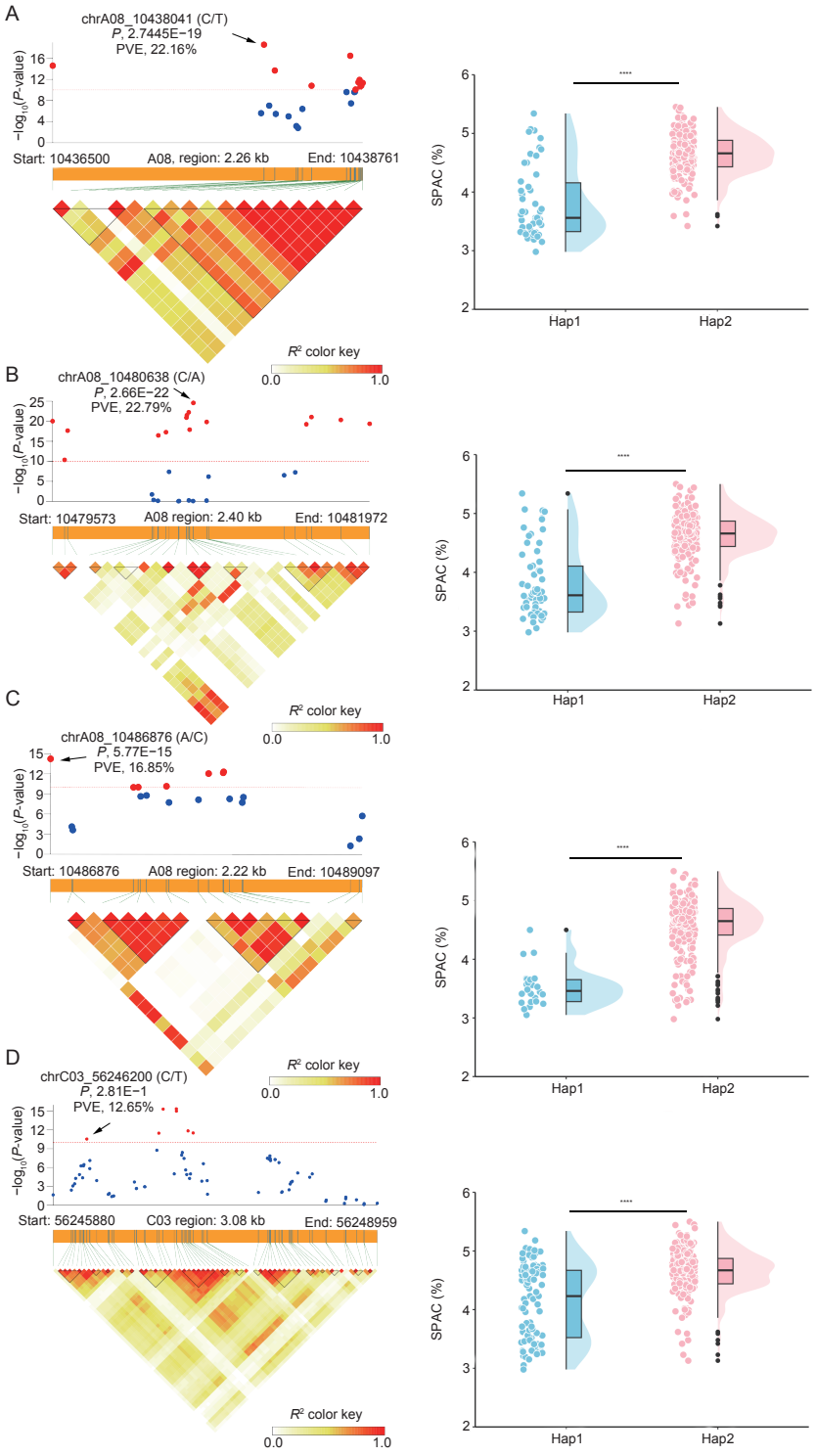


Fig. 4 Association and haplotype analysis of the candidate genes *BnaA08.DAP* (A), *BnaA08.PAA1* (B), *BnaA08.DUF106* (C), and *BnaC03.DAP* (D) with seed palmitic acid content (SPAC). Asterisks indicate the Student's *t*-test significance (****, $P<0.0001$).

accessions across three years (Guan *et al.* 2019). In this study, the SPAC in a panel of 393 *B. napus* accessions grown in three different locations (WH, XM, and SS) was investigated. The SPAC values ranged from 2.98 to 5.97% (Table 1). The SPAC showed relatively high phenotypic variation among the accessions in this association panel grown in different locations, highlighting the complexity of SPAC and the wide genetic diversity within the studied *B. napus* accessions. This extensive phenotypic and genetic variation provides ample opportunities for the genetic improvement of *B. napus* seed quality.

GWAS has been successfully used for the genetic dissection and gene mining of quantitative traits in *B. napus*, such as seed weight (Song *et al.* 2020; Liu *et al.* 2022a),

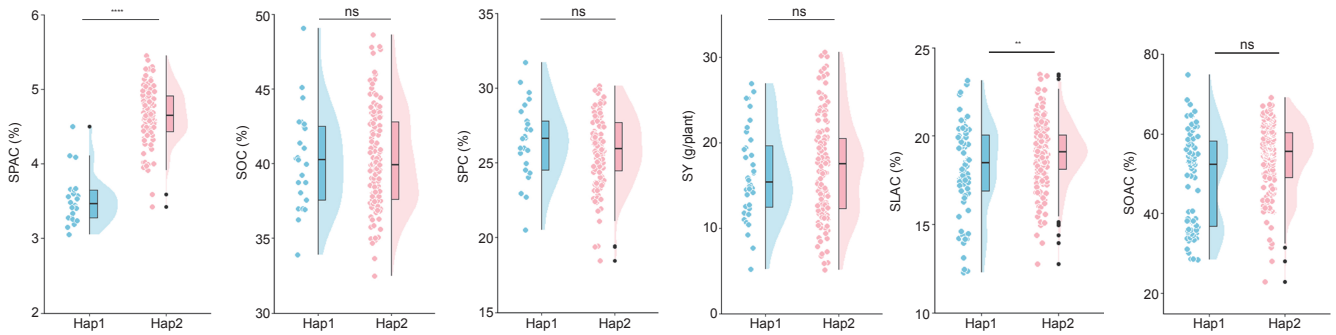


Fig. 5 Differences in seed palmitic acid content (SPAC), seed oil content (SOC), seed protein content (SPC), and seed yield (SY) between *B. napus* varieties carrying *LSPACHap* and *HSPACHap*. Asterisks indicate the Student's *t*-test significance (**, $P<0.01$; ****, $P<0.0001$; ns, non significant).

Table 3 Statistical analysis of seed quality traits and seed yield between *Brassica napus* varieties carrying *LSPACHap* and *HSPACHap*

Haplotype	Trait[1)]	Mean	Median	Min	Max	Range
LSPACHap	SPAC	3.52	3.46	3.05	4.50	1.45
	SOC	40.21	40.25	33.90	49.04	15.14
	Pro	26.41	26.64	20.53	31.69	11.16
	SY	16.34	15.37	5.20	26.92	21.73
	SOAC	48.91	52.36	28.46	23.81	4.65
	SLAC	18.01	18.49	12.27	74.89	62.62
HSPACHap	SPAC	4.66	4.65	3.42	5.45	2.03
	SOC	40.27	39.90	32.49	48.60	16.11
	Pro	25.91	25.95	18.44	30.15	11.71
	SY	17.01	17.49	5.08	30.59	25.51
	SOAC	53.87	55.58	14.67	69.02	54.35
	SLAC	19.04	19.11	12.76	23.84	11.08

[1)] SPAC, seed palmitic acid content (%); SOC, seed oil content (%); Pro, seed protein content; SY, seed yield (g/plant); SOAC, seed oleic acid content; SLAC, seed linoleic acid content.

seed yield (Liu *et al.* 2022c; Zhang *et al.* 2023), Sclerotinia disease resistance (Zuo *et al.* 2022), and SOC (Tang *et al.* 2021; Li *et al.* 2023). Unlike GWAS, TWAS can be used to directly identify candidate genes that are significantly associated with a specific phenotype. Therefore, combining GWAS with TWAS can improve the reliability of candidate gene identification (Wainberg *et al.* 2019). For example, Zhang *et al.* (2022) performed GWAS and TWAS on seed coat color in 382 accessions of *B. napus* and identified the candidate gene *BnaC07.CCR-LIKE* (*CCRL*). The seed coat of the *bnac07ccr* mutant is yellow, in contrast to the black seed coat of the wild type (Zhang *et al.* 2022). However, to date, no studies on the genetic basis and gene mining of SPAC in *B. napus* by GWAS combined with TWAS have been reported.

In this study, the GWAS and TWAS for SPAC in 393 *B. napus* accessions identified four candidate genes (*BnaA08.DAP*, *BnaA08.PAA1*, *BnaA08.DUF106*, and *BnaC03.DAP*) associated with SPAC (Figs. 1 and 3). The TPM values of these genes in seeds at 20 and 40 DAF in this association panel were significantly correlated with SPAC (Fig. 3), suggesting that these genes might be directly involved in the accumulation of palmitic acid in *B. napus* seeds. The discovery of these genes increases our understanding of the oil biosynthetic pathways of rapeseed and provides a new target for the molecular breeding of *B. napus*.

4.2. Breeding of high-quality *B. napus* varieties

By combining candidate gene association analysis and haplotype analysis, favorable alleles that are significantly associated with a specific phenotype can be directly identified. Such alleles are widely used for variety improvement in crops (Sinha *et al.* 2020; Sivabharathi *et al.* 2024). For example, Wang *et al.* (2024) conducted GWAS on salt tolerance in 350 wheat (*Triticum aestivum*) varieties and identified *TaSPL6* as a candidate gene involved in this trait. Indeed, the salt tolerance of the *taspl6* mutant was significantly improved compared to that of the wild type. Based on one InDel in *TaSPL6*, a favorable haplotype (*TaSPL6-DIn*) was identified. Introducing *TaSPL6-DIn* into a modern wheat cultivar lacking this allele increased the seed yield by 5–9% in different saline-alkali soils (Wang *et al.* 2024).

In this study, four low-SPAC haplotypes were identified by the combined association and haplotype analysis of candidate genes (Fig. 4). For example, 10 SNPs in *BnaA08.DAP* were significantly associated with SPAC, and chrA08_10438041 (C/T) had the most significant correlation with SPAC (*P*-value, 2.74e–19; PVE, 22.16%) (Fig. 4-A). The SPAC of *B. napus* accessions carrying the *BnaA08.DAP.Hap1* allele was significantly lower than that of *B. napus* accessions carrying the *BnaA08.DAP.Hap2* allele (Fig. 4-A). The haplotypes identified in this study can facilitate marker-assisted selection in breeding programs. Moreover, identifying low-SPAC haplotypes that do not affect other important traits, such as SOC, SPC, and seed yield, is of great significance for cultivating high-quality *B. napus*. In this study, compared to accessions carrying *HSPACHap*, the accessions carrying *LSPACHap* had lower SPAC while maintaining similar SOC, SPC, and SY, demonstrating the potential of *LSPACHap* for developing high-quality *B. napus* varieties (Fig. 5). Our results indicate the feasibility of selective breeding for reduced SPAC in *B. napus* without sacrificing overall yield or total SOC.

5. Conclusion

This study investigated the SPAC in a panel of 393 *B. napus*

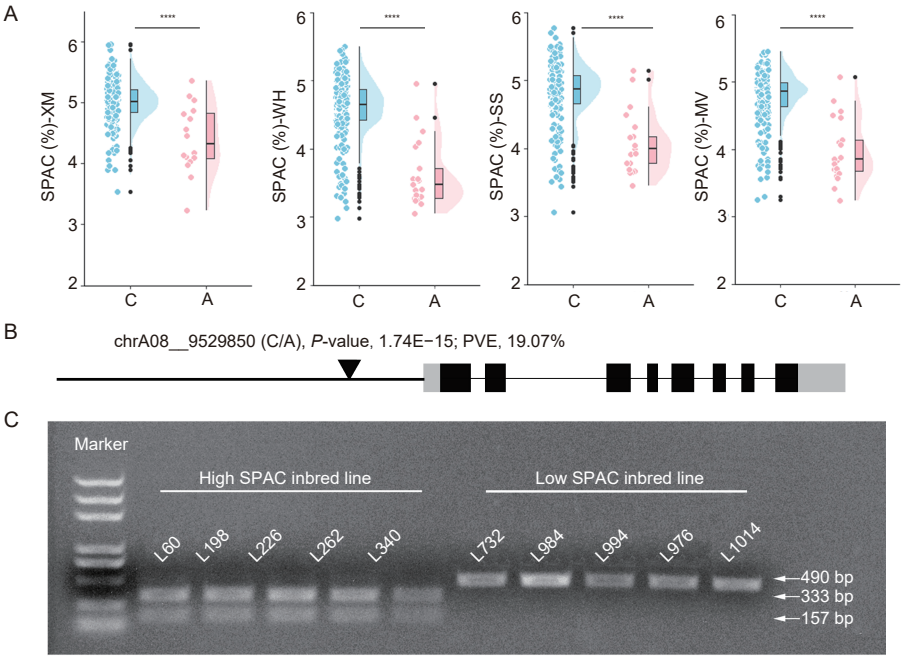


Fig. 6 Molecular markers developed based on chrA08_9529850 (C/A) within the *BnaA08.DUF106* gene. A, differences in seed palmitic acid content (SPAC) among different alleles in chrA08_9529850 (C/A) in the different trials. B, position of the significant SNP chrA08_9529850 (C/A) in *BnaA08.DUF106*. C, PCR products of 10 *Brassica napus* inbred lines subjected to *Nco*I enzyme digestion. Marker, 5,000-bp DNA ladder with band sizes of 5,000, 3,000, 2,000, 1,000, 750, 500, 250, and 100 bp. XM, Xiema, Chongqing; WH, Wuhan, Hubei Province; SS, Sansheng, Chongqing, China; MV, mean value. ****, $P<0.0001$.

accessions collected worldwide in three different field trials. Four candidate genes associated with SPAC were identified *via* GWAS combined with TWAS. In addition, candidate gene association and haplotype analysis of *BnaA08.DAP*, *BnaA08.PAA1*, *BnaA08.DUF106*, and *BnaC03.DAP* revealed several low-SPAC haplotypes. A CAPS molecular marker (*Bn_A8_SPAC_Marker*) that can be used to facilitate *B. napus* breeding for low SPAC was also developed. These significant SNP loci, favorable alleles and haplotypes, and the molecular marker will be useful for breeding high-quality *B. napus* varieties.

Acknowledgements

The authors are grateful for the financial support provided by the National Key Research and Development Program of China (2023YFD1201405), the National Natural Science Foundation of China (32272150 and 32072093), the Natural Science Foundation of Chongqing, China (CSTB2022TIAD-KPX0010, CSTB2023NSCQ-MSX0744 and CSTB2022NSCQ-LZX0034), the earmarked Fund for China Agriculture Research System (CARS-12), and the Innovation and Entrepreneurship Training Program for Undergraduates, China (S202410635206).

Declaration of competing interest

The authors declare that they have no conflict of interest.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this manuscript, the authors used deepseek to assist in improving the clarity and readability of the language. All content generated was carefully reviewed and revised by the authors as necessary, and the authors take full responsibility for the final content of the article.

Appendices associated with this paper are available at https://doi.org/10.1016/j.jia.2024.11.015

References

Belide S, Petrie J R, Shrestha P, Singh S P. 2012. Modification of seed oil composition in *Arabidopsis* by artificial microRNA-mediated gene silencing. *Frontiers in Plant Science*, **3**, 168.

Bradbury P J, Zhang Z, Kroon D E, Casstevens T M, Ramdoss Y, Buckler E S. 2007. TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics*, **23**, 2633–2638.

Cortés A J, Du H. 2023. Molecular genetics enhances plant breeding. *International Journal of Molecular Sciences*, **24**, 9977.

Danecek P, Auton A, Abecasis G, Albers G A, Banks E, DePristo M A, Handsaker R E, Lunter G, Marth G T, Sherry S T, McVean G, Durbin R. 2011. The variant call format and VCFtools. *Bioinformatics*, **27**, 2156–2164.

Gacek K, Bayer P E, Bartkowiak-Broda I, Szala L, Bocianowski J, Edwards D, Batley J. 2017. Genome-wide association study of genetic control of seed fatty acid biosynthesis in *Brassica napus*. *Frontiers in Plant Science*, **7**, 2062.

Ge X Y, Chen J L, Li O Q, Zou M, Tao B L, Zhao L, Wen J, Yi B, Tu J X, Shen J X. 2025. ORF138 causes abnormal lipid metabolism in the tapetum that leads to Ogu cytoplasmic male sterility in *Brassica napus*. *Journal of Integrative Agriculture*, **24**, 2080–2095.

Guan M, Huang X, Xiao Z, Jia L, Wang S, Zhu M, Qiao C, Wei L, Xu X, Liang Y, Wang R, Lu K, Li J, Qu C. 2019. Association mapping analysis of fatty acid content in different ecotypic rapeseed using mrMLM. *Frontiers in Plant Science*, **9**, 1872.

Guyton J R, Ponder M, Kirkpatrick C F. 2023. Advances in understanding palmitic acid metabolism and health risks. *Journal of Clinical Lipidology*, **17**, 571–573.

Huang K L, Tian J, Wang H, Fu Y, Li Y, Zheng Y, Bao X. 2021. Fatty acid export protein *BnFAX6* functions in lipid synthesis and axillary bud growth in *Brassica napus*. *Plant Physiology*, **186**, 2064–2077.
Khlestkina E, Shavrukov Y. 2022. Molecular-genetic basis of plant breeding. *Biomolecules*, **12**, 1392.
Li H, Thrash A, Tang J D, He L, Yan J, Warburton M L. 2019. Leveraging GWAS data to identify metabolic pathways and networks involved in maize lipid biosynthesis. *The Plant Journal*, **98**, 853–863.
Li L, Tian Z, Chen J, Tan Z, Zhang Y, Zhao H, Wu X, Yao X, Wen W, Chen W, Guo L. 2023. Characterization of novel loci controlling seed oil content in *Brassica napus* by marker metabolite-based multi-omics analysis. *Genome Biology*, **24**, 141.
Li Y, Ali U, Cao Z, Cao Z, Zeng C, Xiao M, Wei F, Yao X, Guo L, Lu S. 2022. Fatty acid exporter 1 enhances seed oil content in *Brassica napus*. *Molecular Breeding*, **42**, 75.
Lippert C, Listgarten J, Liu Y, Kadie C M, Davidson R I, Heckerman D. 2011. FaST linear mixed models for genome-wide association studies. *Nature Methods*, **8**, 833–835.
Liu H, Wang J, Zhang B, Yang X, Hammond J, Ding G, Wang S, Cai H, Wang C, Xu F, Shi L. 2021. Genome-wide association study dissects the genetic control of plant height and branch number in response to low-phosphorus stress in *Brassica napus*. *Annals of Botany*, **128**, 919–930.
Liu H J, Wang J C, Zhang B B, Yang X Y, Yuan P, Ding G D, Wang S L, Cai H M, Wang C, Xu F S, Shi L. 2022a. Genome-wide association study identifies new loci for 1000-seed weight in *Brassica napus*. *Euphytica*, **218**, 142.
Liu H J, Wang W, Yang M, Yuan P, Hammond J P, King G J, Ding G D, White P, Wang S L, Cai H M, Wang C, Lu C G, Xu F S, Shi L. 2022b. Genome-wide association studies of important agronomic traits in *Brassica napus*: What we have learned and where we are headed. *Annual Plant Review Online*, **5**, 151–180.
Liu H J, Zou M Y, Zhang B B, Yang X Y, Yuan P, Ding G D, Xu F S, Shi L. 2022c. Genome-wide association study identifies candidate genes and favorable haplotypes for seed yield in *Brassica napus*. *Molecular Breeding*, **42**, 61.
Mai N T P, Nguyen L T T, Tran S G, To H T M. 2023. Genome-wide association study reveals useful QTL and genes controlling the fatty acid composition in rice bran oil using Vietnamese rice landraces. *Functional & Integrative Genomics*, **23**, 150.
Mancini A, Imperlini E, Nigro E, Montagnese C, Daniele A, Orrù S, Buono P. 2015. Biological and nutritional properties of palm oil and palmitic acid: Effects on health. *Molecules*, **20**, 17339–17361.
Qi Z, Guo C, Li H, Qiu H, Li H, Jong C, Yu G, Zhang Y, Hu L, Wu X, Xin D, Yang M, Liu C, Lv J, Wang X, Kong F, Chen Q. 2023. Natural variation in fatty acid 9 is a determinant of fatty acid and protein content. *Plant Biotechnology Journal*, **22**, 759–773.
Qu C, Jia L, Fu F, Zhao H, Lu K, Wei L, Xu X, Liang Y, Li S, Wang R, Li J. 2017. Genome-wide association mapping and Identification of candidate genes for fatty acid composition in *Brassica napus* L. using SNP markers. *BMC Genomics*, **18**, 232.
Rajcan I, Carrero-Colón M, Hudson K. 2022. Reduced palmitic acid content in soybean as a result of mutation in *FATB1a*. *PLoS ONE*, **17**, e0262327.
Sinha P, Singh V K, Saxena R K, Khan A W, Abbai R, Chitikineni A, Desai A, Molla J, Upadhyaya H, Kumar A, Varshney R. 2020. Superior haplotypes for haplotype-based breeding for drought tolerance in pigeonpea (*Cajanus cajan* L.). *Plant Biotechnology Journal*, **18**, 2482–2490.
Sivabharathi R C, Rajagopalan V R, Suresh R, Sudha M, Karthikeyan G, Jayakanthan M, Raveendran M. 2024. Haplotype-based breeding: A new insight in crop improvement. *Plant Science*, **346**, 112129.
Song J, Guan Z, Hu J, Guo C, Yang Z, Wang S, Liu D, Wang B, Lu S, Zhou R, Xie W, Cheng Y, Zhang Y, Liu K, Yang Q, Chen L, Guo L. 2020. Eight high-quality genomes reveal pan-genome architecture and ecotype differentiation of *Brassica napus*. *Nature Plants*, **6**, 34–45.
Tang S, Zhao H, Lu S, Yu L, Zhang G, Zhang Y, Yang Q, Zhou Y, Wang X, Ma W, Xie W, Guo L. 2021. Genome- and transcriptome-wide association studies provide insights into the genetic basis of natural variation of seed oil content in *Brassica napus*. *Molecular Plant*, **14**, 470–487.
Wainberg M, Sinnott-Armstrong N, Mancuso N, Barbeira A, Knowles D, Golan D, Ermel R, Ruusalepp A, Quertermous T, Hao K, Björkegren J, Im H, Pasaniuc B, Rivas M, Kundaje A. 2019. Opportunities and challenges for transcriptome-wide association studies. *Nature Genetics*, **51**, 592–599.
Wang M, Cheng J, Wu J, Chen J, Liu D, Wang C, Ma S, Guo W, Li G, Di D, Zhang Y, Han D, Kronzucker H, Xia G, Shi W. 2024. Variation in *TaSPL6-D* confers salinity tolerance in bread wheat by activating *TaHKT1;5-D* while preserving yield-related traits. *Nature Genetics*, **56**, 1257–1269.
Wang Z, Wang Y, Shang P, Yang C, Yang M, Huang J, Ren B, Zuo Z, Zhang Q, Li W, Song B. 2022. Overexpression of soybean *GmWRI1a* stably increases the seed oil content in soybean. *International Journal of Molecular Sciences*, **23**, 5084.
Xiao Z, Tang F, Zhang L, Li S, Wang S, Huo Q, Yang B, Zhang C, Wang D, Li Q, Wei L, Guo T, Qu C, Lu K, Zhang Y, Guo L, Li J, Li N. 2021. The *Brassica napus* fatty acid exporter *FAX1-1* contributes to biological yield, seed oil content, and oil quality. *Biotechnology for Biofuels*, **14**, 190.
Yang Z, Wang S, Wei L, Huang Y, Liu D, Jia Y, Luo C, Lin Y, Liang C, Hu Y, Dai C, Guo L, Zhou Y, Yang Q. 2023. *BnIR*: A multi-omics database with various tools for *Brassica napus* research and breeding. *Molecular Plant*, **16**, 775–789.
Yin L, Zhang H, Tang Z, Xu J, Yin D, Zhang Z, Yuan X, Zhu M, Zhao S, Li X, Liu X. 2021. rMVP: A memory-efficient, visualization-enhanced, and parallel-accelerated tool for genome-wide association study. *Genomics*, *Proteomics & Bioinformatics*, **19**, 619–628.
Zhang C, Dong S, Xu J, He W, Yang T. 2019. PopLDdecay: A fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics*, **15**, 1786–1788.
Zhang C, Gong R, Zhong H, Dai C, Zhang R, Dong J, Li Y, Liu S, Hu J. 2023. Integrated multi-locus genome-wide association studies and transcriptome analysis for seed yield and yield-related traits in *Brassica napus*. *Frontiers in Plant Science*, **14**, 1153000.
Zhang Y, Zhang H, Zhao H, Xia Y Zheng X, Fan R, Tan Z, Duan C, Fu Y, Li L, Ye J, Tang S, Hu H, Xie W, Yao X, Guo L. 2022. Multi-omics analysis dissects the genetic architecture of seed coat content in *Brassica napus*. *Genome Biology*, **23**, 86.
Zheng H, Zhao D, Shao W, Lu Y, Wang W, Hu Y, Li J, Zhu S, Wang X. 2022. *GmWRI1c* increases palmitic acid content to regulate seed oil content and nodulation in soybean (*Glycine max*). *International Journal of Molecular Sciences*, **23**, 13793.
Zuo R, Xie M, Gao F, Liu J, Tang M, Cheng X, Liu Y, Bai Z, Liu S. 2022. Genome-wide identification and functional exploration of the legume lectin genes in *Brassica napus* and their roles in Sclerotinia disease resistance. *Frontiers in Plant Science*, **13**, 963263.