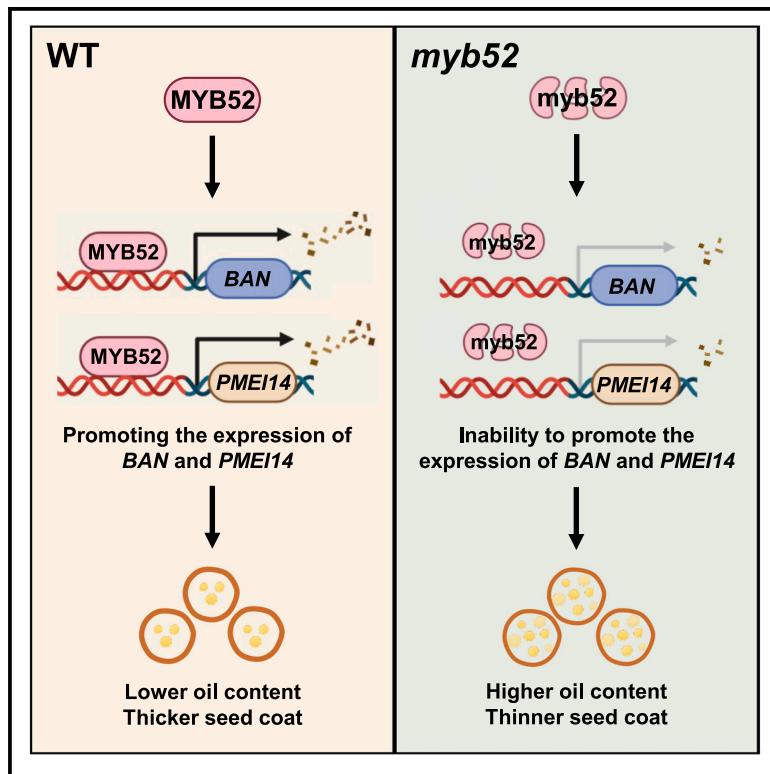


### Manipulation of seed coat content for increasing oil content via modulating *BnaMYB52* in *Brassica napus*

#### Graphical abstract



#### Authors

Jiang Ye, Xiaowei Wu, Xiang Li, ...,  
Zengdong Tan, Liang Guo, Xuan Yao

#### Correspondence

tanzengdong@yzwlab.cn (Z.T.),  
guoliang@mail.hzau.edu.cn (L.G.),  
xuanyao@mail.hzau.edu.cn (X.Y.)

#### In brief

Ye et al. report a major QTL gene in *B. napus* and find that *BnaA09.MYB52* positively regulates seed coat content and negatively affects seed oil content. *BnaA09.MYB52* and its downstream genes are breeding targets for genetic manipulation to improve oil content and meal quality in *B. napus*.

#### Highlights

- A pleiotropic QTL gene is identified by GWAS for seed coat content and seed oil content
- The expression of *BnaA09.MYB52* is predominant in the seed coat during seed development
- *BnaA09.MYB52* directly targets *PMEI14* and *BAN* to affect SOC and SCC
- *BnaMYB52* mutation results in increased SOC, decreased SCC, and reduced lignin content



## Article

# Manipulation of seed coat content for increasing oil content via modulating *BnaMYB52* in *Brassica napus*

Jiang Ye,<sup>1,3</sup> Xiaowei Wu,<sup>1,2,3</sup> Xiang Li,<sup>1</sup> Yuting Zhang,<sup>2</sup> Yuqing Li,<sup>1</sup> Hui Zhang,<sup>1</sup> Jie Chen,<sup>2</sup> Yuyan Xiang,<sup>1</sup> Yefan Xia,<sup>1</sup> Hu Zhao,<sup>1</sup> Zengdong Tan,<sup>2,\*</sup> Liang Guo,<sup>1,2,4,\*</sup> and Xuan Yao<sup>1,2,\*</sup>

<sup>1</sup>National Key Laboratory of Crop Genetic Improvement, Hubei Hongshan Laboratory, Huazhong Agricultural University, Wuhan, Hubei 430070, China

<sup>2</sup>Yazhouwan National Laboratory, Sanya, Hainan 572025, China

<sup>3</sup>These authors contributed equally

<sup>4</sup>Lead contact

\*Correspondence: [tanzengdong@yzwlab.cn](mailto:tanzengdong@yzwlab.cn) (Z.T.), [guoliang@mail.hzau.edu.cn](mailto:guoliang@mail.hzau.edu.cn) (L.G.), [xuanyao@mail.hzau.edu.cn](mailto:xuanyao@mail.hzau.edu.cn) (X.Y.)  
<https://doi.org/10.1016/j.celrep.2025.115280>

## SUMMARY

Seed oil is synthesized in the embryo, which is surrounded by the seed coat. However, the genetic and molecular mechanisms of how seed coat development affects seed oil content (SOC) remains less studied. Through genome-wide association studies in *Brassica napus* accessions, we identify *BnaA09.MYB52* transcription factor as a candidate gene underlying *ZS11\_qSCC.A09*. Knocking out *BnaMYB52* results in a 7.7%–8.2% reduction in seed coat content (SCC) but a 12.3%–13.4% increase in SOC, whereas overexpression of *BnaA09.MYB52* leads to the opposite effects. Intriguingly, *BnaA09.MYB52* has predominant expression in seed coat during seed development. *BnaA09.MYB52* activates the expression of *PMEI14* and *BAN*. Mutant analyses indicate that the SOC is increased by 4.3%–7.7%, while seed coat thickness shows a 12.9%–22.8% reduction in the *ban* and *pmei14* mutants. Thus, our findings identify the molecular mechanism of MYB52-BAN/PMEI14 modules regulating SCC and SOC and provide a potential target for oil content improvement in *B. napus*.

## INTRODUCTION

Seed is composed of embryo, endosperm, and seed coat in angiosperms.<sup>1</sup> In *Brassica napus*, one of the critical resources of edible vegetable oil and biofuel worldwide,<sup>2,3</sup> oil (triacylglycerol [TAG]) is mainly stored in the embryo, which develops from the fertilized zygote. The seed coat is derived from two layers of the outer integument and three layers of the inner integument, which protects the embryo and endosperm from abiotic/biotic stresses and influences seed dormancy, germination, and dispersal.<sup>4,5</sup> Owing to synthesizing flavonoid compounds, mucilage, lignin, and other metabolites,<sup>6,7</sup> the seed coat competes with the embryo for photosynthetic assimilates or common precursors, suggesting an association between seed coat and TAG accumulation.

To date, more than 10 quantitative trait loci (QTLs) have been identified for seed coat content (SCC, seed coat mass divided by seed weight) in *B. napus*.<sup>8–11</sup> However, none of these QTLs have been cloned. In addition, through searching for homologs of *Arabidopsis* genes, several SCC-regulating genes have been identified, and they also affect seed oil content (SOC) in *B. napus*. For example, knockout or knockdown *TRANSPARENT TESTA* (*TT*) genes (such as *BnaTT1*, *BnaTT2*, *BnaTT7*, *BnaTT10*, *BnaTT12*, and *BnaTT18*), involved in the flavonoid biosynthesis pathway,

results in seeds with decreased pigment, increased SOC, and thinner seed coat.<sup>4,12–14</sup> Targeted mutation of *TRANSPARENT TESTA GLABRA 1* (*BnaTTG1*, encoding a WD40-repeat protein) also produces seeds with altered pigment, SOC, and seed coat thickness.<sup>15</sup> It is worth mentioning that *BnaC07.CCRL* (cinnamoyl-coenzyme A [CoA] reductase-like) and *BnaTT8s* are identified as regulators involved in SCC and SOC by transcriptome-wide association studies (TWASs) in *B. napus*. *BnaC07.CCRL* encodes an NAD(P)-binding Rossmann-fold superfamily protein that is involved in lignin biosynthesis, and *BnaTT8s* encode a basic-helix-loop-helix (bHLH) transcription factor 42 that participates in proanthocyanidin-specific accumulation in seed coat. Both genes regulate SCC and SOC by affecting the deposition of different metabolites.<sup>8,12</sup> Overall, progress in SCC regulation remains slow and the underlying molecular mechanisms are largely unknown. Further research will help improve understanding of the regulatory mechanisms of SCC in *B. napus* and other crucifers.

*MYB52* is a member of the R2R3 MYB transcription factors that contain an MYB DNA-binding domain (with R2 and R3 repeats of about 52 amino acids) at the N terminus and an activation/repression domain at the C terminus and evolve to have functional diversity in plants such as metabolism, development, and responses to biotic/abiotic stresses.<sup>16,17</sup> In *Arabidopsis*,



*AtMYB52* functions as an important regulator of seed coat mucilage. Loss-of-function mutants of *AtMYB52* result in a defect of mucilage extrusion and unaltered seed coat morphology. Molecular experiments demonstrate that *AtMYB52* directly binds to the promoters of the *PECTIN METHYLESTERASE INHIBITOR 6 (PMEI6)*, *PMEI14*, and *SUBTILISIN-LIKE SERPROTEASE 1.7 (SBT1.7)* involved in seed coat mucilage demethylesterification and activates their expression.<sup>18</sup> Furthermore, *AtMYB52* physically interacts with *AtERF4* protein (an AP2/ERF transcription factor) to repress each other's transcriptional activity. *AtMYB52* and *AtERF4* function in the same pathway to regulate seed coat mucilage demethylesterification.<sup>19</sup> Mutant analysis shows that *AtMYB52* also affects lignin deposition to control secondary cell wall formation in stems and fiber cells.<sup>20,21</sup> Overexpression of *AtMYB52* in *Arabidopsis* confers ABA hypersensitivity, salt sensitivity, and drought tolerance.<sup>22</sup> In soybean, the transcription levels of *GmMYB52* are increased under ABA, salt, drought, and cold stresses.<sup>23</sup> *BnaA09.MYB52* positively controls seed coat mucilage accumulation and tolerance to osmotic stress during seed germination in *Arabidopsis*.<sup>24</sup> Although *MYB52* is involved in the development of seed coat mucilage, lignin, and response to stresses, whether it impacts SCC and SOC remains unknown.

In this study, we performed genome-wide association studies (GWASs) for SCC and characterized a major QTL *ZS11\_qSCC.A09* in detail, which is co-localized with *ZS11\_qSOC.A09* regulating SOC. Through a gene prioritization framework (POCKET), haplotype analysis, and seed trait analysis of transgenic plants, we identify *BnaA09.MYB52* as a pleiotropic regulator conferring SCC and SOC variation. During seed development, *BnaA09.MYB52* transcript specifically accumulates in the seed coat, but not in the seed embryo. *BnaA09.MYB52* regulates seed coat content/thickness and SOC by targeting *BnaPMEI14* and *BnaBAN (BANYULS)*, which are also highly expressed in seed. Our study provides important insights into the role of *BnaA09.MYB52* in the transcriptional regulation of SCC and SOC and likely facilitates breeding of high-SOC varieties in *B. napus*.

## RESULTS

### *ZS11\_qSCC.A09* affects both SOC and SCC in *B. napus*

To dissect the genetic basis of SCC and SOC in *B. napus*, we previously measured SCC and SOC of 505 *B. napus* accessions.<sup>8,25</sup> A total of 11,700,689 single-nucleotide polymorphisms (SNPs) were identified based on the Zhongshuang11 (*ZS11*) v0 genome (<https://yanglab.hzau.edu.cn/>) in the previous studies.<sup>8,25,26</sup> GWAS was carried out for the two traits, and five and 16 QTLs are found to be significantly associated with SCC and SOC, respectively (Figures 1A and S1A; Tables S1 and S2). Among them, three QTLs affecting both SCC as well as SOC are localized on chromosome A05 and A09 (Figure 1B). Further analysis of the co-localized QTLs revealed that the variants corresponding to the lead SNP of each QTL resulted in opposite changes in SCC and SOC. For example, in *ZS11\_qSCC.A09*, the variation in lead SNP (A to T) leads to a decrease in SCC and an increase in SOC ( $p_{\text{SCC}} = 2.38 \times 10^{-5}$ ,  $p_{\text{SOC}} = 1.49 \times 10^{-4}$ ) (Figures 1C and 1D). Additionally, we collected 1,956 SOC- or SCC-related genes based on functional annotation (Table S3). By further analyzing

the eGWAS result in seeds identified in the previous study,<sup>26</sup> it was found that 19.89% of the related genes are regulated by hotspot 87 (Hot87). Interestingly, Hot87 is co-located with *ZS11\_qSCC.A09* (Figures 1E and 1F). The above results indicate that *ZS11\_qSCC.A09* is an important SCC- and SOC-related QTL, and this QTL may affect SCC and SOC by regulating the expression of a number of downstream genes.

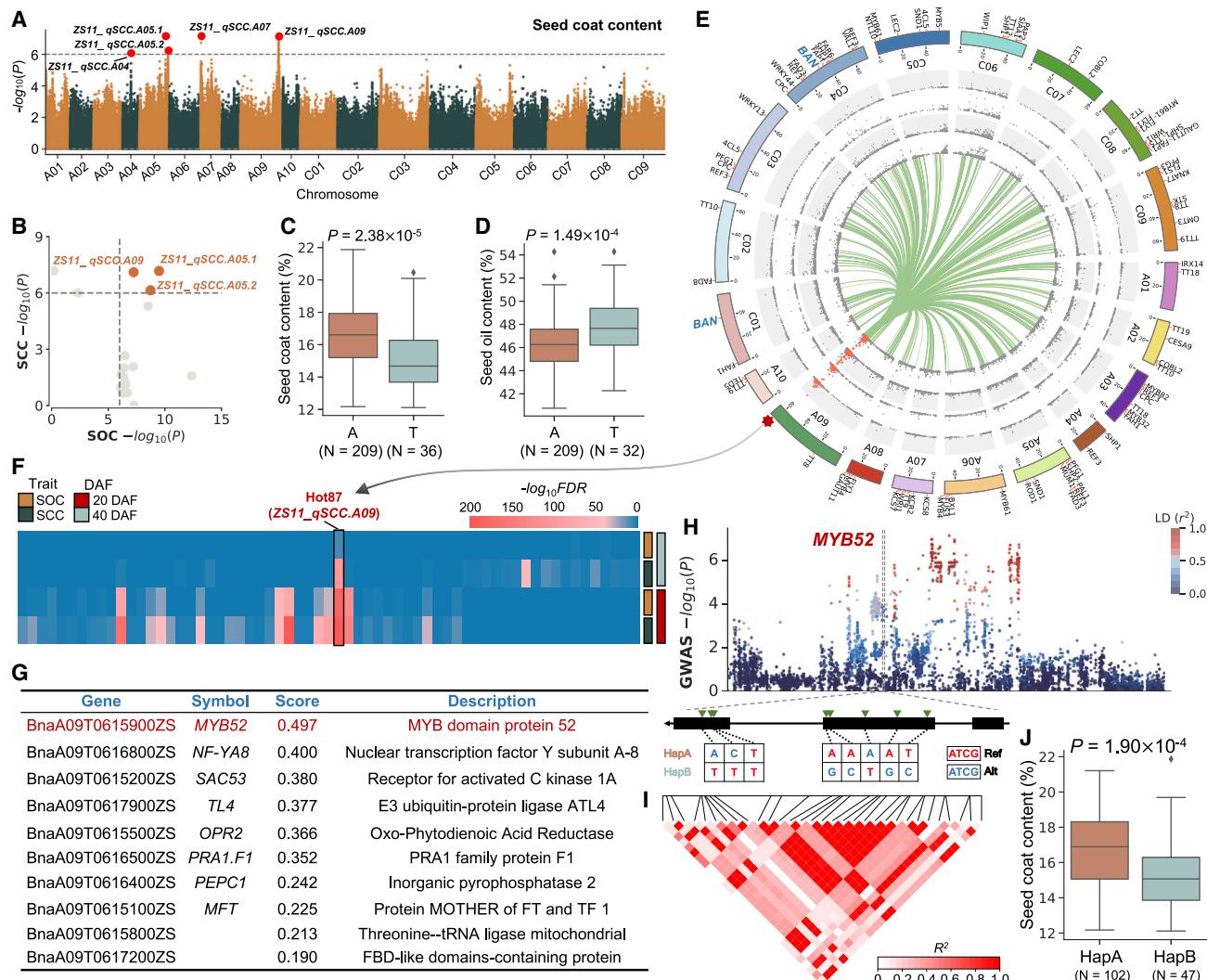
### *BnaA09.MYB52* is the candidate gene of *ZS11\_qSCC.A09*

Since *ZS11\_qSCC.A09* co-localizes with the eQTL hotspot regulatory region, it is hypothesized that there are important transcription factors in *ZS11\_qSCC.A09* that affect SCC and SOC. To explore the candidate gene in *ZS11\_qSCC.A09*, POCKET<sup>25</sup> was used to predict the candidate gene in *ZS11\_qSCC.A09*. The results reveal that *BnaA09.MYB52* is the top-ranked regulator (Figure 1G). There are multiple missense variants in the coding region of *BnaA09.MYB52* (Figures 1H and 1I), and these genetic variants are sorted into two major haplotypes of this gene. Compared to accessions harboring the haplotype A, there is a lower SCC and higher SOC for accessions with the haplotype B ( $p_{\text{SCC}} = 1.90 \times 10^{-4}$ ,  $p_{\text{SOC}} = 0.01$ ) (Figures 1J and S1B). Therefore, we focused on identifying the potentially pleiotropic functions of *BnaA09.MYB52* underlying *ZS11\_qSCC.A09* and elucidating its regulatory mechanism on seed coat development and oil accumulation.

Through BLAST analysis in the *B. napus* genome, there are four homologous copies of *BnaMYB52*, namely *BnaA09.MYB52* (*BnaA09G0615900ZS*), *BnaA08.MYB52* (*BnaA08G0259100ZS*), *BnaC08.MYB52-1* (*BnaC08G0471400ZS*), and *BnaC08.MYB52-2* (*BnaC08G0242900ZS*). Sequence analysis indicates that these four genes have a conserved MYB domain and belong to the R2R3-MYB transcription factor family (Figure S1C). Structural analysis shows that *BnaA09.MYB52* protein contains one R2 repeat with 54 amino acids and an R3 repeat with 51 amino acids in the N terminus, and several motifs are found in the C terminus (Figure S1C). These data demonstrate that *BnaA09.MYB52* is a typical R2R3-MYB transcription factor.

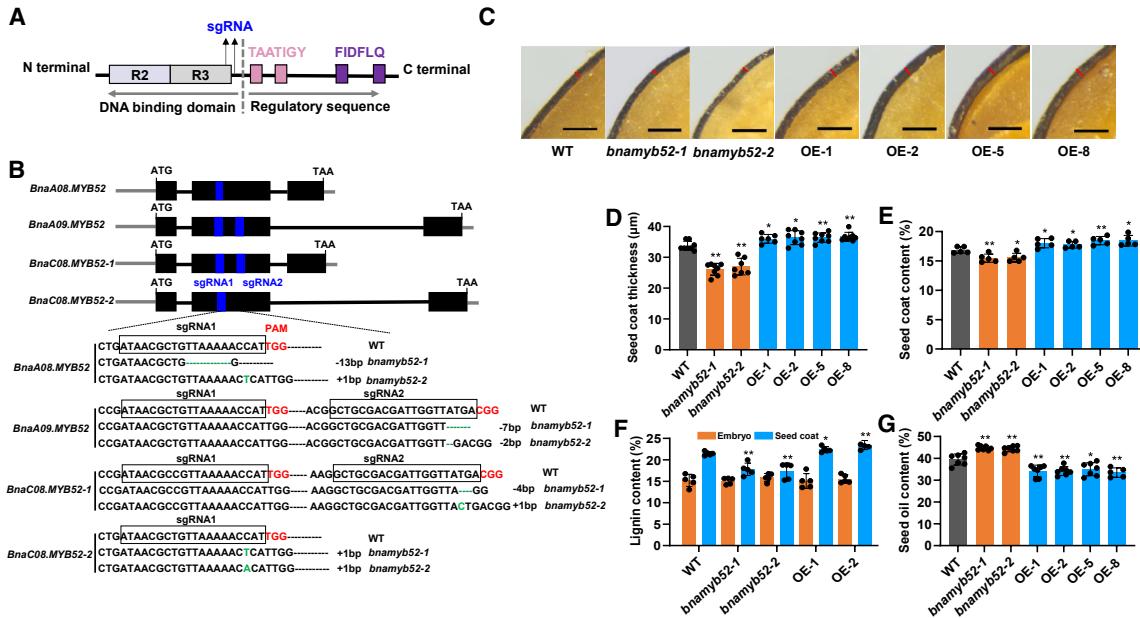
### *BnaA09.MYB52* positively regulates SCC

To explore the function of *BnaA09.MYB52* in seed coat development, we generated *BnaMYB52* knockout plants (*bnamyb52-1* and *bnamyb52-2* with four homologous genes of *BnaMYB52* knocked out) using the CRISPR-Cas9 method. We also generated overexpression lines of *BnaA09.MYB52* driven by the constitutive 35S promoter in *B. napus* cultivar Westar. The *bnamyb52-1* and *bnamyb52-2* display premature stop codons leading to truncated proteins of four copies of *BnaMYB52* (Figures 2A and 2B). The expression levels of OE lines are significantly higher than that of wild type (WT) in the embryo and seed coat (Figures S2A and S2B). Histological observations indicate that seed coat thickness is decreased in *bnamyb52* mutants, whereas it is increased in OE lines as compared with WT (Figure 2C). Statistical analysis further shows that, compared to WT (33.8 μm), the average seed coat thickness of *bnamyb52* mutants is only 26.1 and 27.0 μm, respectively, exhibiting a 20.0%–21.8% reduction, whereas OE plants' seed coat thickness is 36.0 μm (OE-1), 36.4 μm (OE-2), 36.4 μm (OE-5), and 36.7 μm (OE-8),



**Figure 1.** *BnaA09.MYB52* is a key gene for SCC and SOC in *B. napus*

- (A) GWAS analysis of SCC. Red points annotate the QTLs. The ZS11 indicates *B. napus* cultivar Zhongshuang 11 and is used as the reference genome in this study.
- (B) The co-localization of SOC and SCC QTLs. Each point represents a QTL related to SOC and SCC. The horizontal coordinate indicates  $-\log_{10}$  of  $p$  value for each QTL in the SOC GWAS. The vertical coordinate indicates the  $-\log_{10}$  of  $p$  value for each QTL in the SCC GWAS.
- (C) The lead SNP haplotypes of ZS11\_qSCC.A09 in the SCC GWAS. The diamond indicates the phenotype outlier. Differences were analyzed by two-sided Student's t test.
- (D) The lead SNP haplotypes of ZS11\_qSCC.A09 in the SOC GWAS. The diamonds indicate phenotype outliers. Differences were analyzed by two-sided Student's t test.
- (E) The eQTL for SOC- and SCC-related genes co-localized with ZS11\_qSCC.A09. Circles represent, from the outside to the inside, the chromosomes, the eGWAS results of SOC-related genes at 20 DAF, the eGWAS results of SCC-related genes at 20 DAF, the eGWAS results of SOC-related genes at 40 DAF, the eGWAS results of SCC-related genes at 40 DAF, and the connectivity of the individual genes localized in Hot87, respectively.
- (F) Enrichment analysis of each hotspot-regulated gene related to SOC and SCC. The color represents the range of false discovery rate (FDR). ZS11\_qSCC.A09 is the most significantly enriched.
- (G) Summary of the top 10 predicted genes in ZS11\_qSCC.A09. *BnaA09.MYB52* is identified as the most likely candidate gene responsible for ZS11\_qSCC.A09.
- (H) Local Manhattan plot (top) surrounding the peak on chromosome A09 and two haplotypes (bottom) of *BnaA09.MYB52*.
- (I) The linkage disequilibrium (LD) heatmap for the variant in *BnaA09.MYB52*.
- (J) Boxplot for SCC based on the *BnaA09.MYB52* haplotypes. The diamond indicates the phenotype outlier. The differences were measured by two-sided Student's t test.



**Figure 2.** *BnaA09.MYB52* positively regulates SCC and negatively affects SOC in *B. napus*

(A) Domain structure of *BnaA09.MYB52* transcription factor. “R” means repeats including about 52 amino acids. “TAATIGY” and “FIDFLQ” indicate two predicted motifs in the C-terminal region.

(B) Mutation types of two homozygous *BnaMYB52* mutants edited by the CRISPR-Cas9 system in *B. napus*. The two sgRNAs (blue boxes above the sequence) were designed in the second exon of *BnaMYB52* protein-coding region. Black boxes and lines indicate the exons and introns. Gray lines indicate the promoter and UTR sequences. ATG and TAA represent the start and stop codon, respectively. The four copies of *BnaMYB52* are all knocked out in the two mutants (*bnamyb52-1* and *bnamyb52-2*). sgRNA sequences and protospacer-adjacent motif (PAM) are denoted in the black boxes and red font, respectively. Green dashed/nucleotides indicate deletions or insertions, respectively. Deletion/insertion nucleotides are represented as a minus or plus on the right side of the sequence. bp, base pair.

(C) Representative images of seed coat at the mature stage in the T<sub>2</sub> generation. Scale bars, 200 μm

(D–G) SCC (D), seed coat thickness (E), lignin content (F), and SOC (G) of WT (Westar), *BnaMYB52* knockout lines, and *BnaA09.MYB52* overexpression lines. Data are presented as mean ± SD. Student’s t test: \*p < 0.05; \*\*p < 0.01.

exhibiting a 6.7%–8.7% increase (Figure 2D). A previous study shows that seed coat thickness directly affects SCC.<sup>8</sup> Compared to WT, both *bnamyb52* mutants show a significant reduction in SCC (an 8.2% reduction in *bnamyb52-1* plants and a 7.7% reduction in *bnamyb52-2* plants), while SCC of OE lines is increased by 5.3%–9.5% (Figure 2E). These results demonstrate that *BnaA09.MYB52* functions as a positive regulator of SCC by controlling seed coat thickness in *B. napus*.

A previous study shows that seed coat content/thickness is positively correlated with lignin content.<sup>12</sup> Thus, lignin content was measured in embryo and seed coat, respectively. There is no difference in the lignin content of embryo among WT, *bnamyb52* mutants, and OE lines (OE-1 and OE-2). However, lignin content is significantly decreased in the seed coat of *bnamyb52* mutants (reduced by 17.2%–18.9%) and, conversely, increased by 4.4%–9.3% in OE plants as compared to WT (Figure 2F). Additionally, metabolite analysis indicates that the seeds of *bnamyb52* mutants display a significant reduction in L-phenylalanine, p-coumaric acid, sinapaldehyde, and pinoresinol 4-O-β-D-glucopyranoside related to lignin biosynthetic pathway, whereas the contents of these three metabolites (except sinapaldehyde) are markedly increased in OE lines compared with WT, implying that lignin content may be altered in the *bnamyb52* mutant seeds (Figure S2C). Together, these results show that *BnaA09.MYB52*

positively regulates seed coat content/thickness in *B. napus* and it is the causal gene of the ZS11\_qSCC.A09 for SCC.

#### ***BnaMYB52* mutation improves SOC without compromising yield**

As *BnaA09.MYB52* is also located in ZS11\_qSOC.A09 regulating oil content (Figure S1A) and SCC is negatively correlated with SOC,<sup>8,25</sup> we next tested whether *BnaA09.MYB52* affected SOC in *B. napus*. As expected, compared with WT, SOC measured by GC-FID in the *bnamyb52* mutants is elevated by 12.3%–13.4% (p < 0.01), while the OE plants display a reduction of 11.3%–15.2% (p < 0.05) in the T<sub>3</sub> generation (Figure 2G). The result is further validated using a near-infrared reflectance spectroscope in the T<sub>2</sub> and T<sub>3</sub> generations (Figures S2D and S2E). The fatty acid (FA) composition in seeds of WT, *bnamyb52* mutants, and OE plants is changed slightly as shown in Figure S2F. *WRI1* functions as a master regulator of TAG biosynthesis.<sup>27</sup> *MCMT*, *ENR*, *FATA*, *GPDH*, *LPAAT1*, *PDAT*, and *DGAT1* are considered critical genes involved in FA synthesis or TAG assembly.<sup>4</sup> We conducted RT-qPCR analysis to determine the expression level of these genes in the seeds, and the results show significantly increased expression of these genes in the *bnamyb52* mutants compared to WT (Figures S3A–S3H), which is in line with the trend in oil-content change. To further validate the function of *MYB52* in the

determination of SOC, two independent homozygous transfer DNA insertional mutants, *atmyb52-1* and *atmyb52-2*, were obtained in *Arabidopsis* (Figures S3I and S3J). The RT-qPCR results show that the *AtMYB52* transcripts are basically undetectable, indicating that they are null mutants (Figure S3K). In the T<sub>3</sub> generation, the average SOC of *atmyb52-1* and *atmyb52-2* is significantly higher than that of WT (Figure S3L), which is further validated in the T<sub>4</sub> generation (Figure S3M). These results suggest that *MYB52* is a negative regulator of oil accumulation in *B. napus* seeds, and its function is conserved in *Arabidopsis* and *B. napus*.

In *B. napus*, SOC generally has a negative relationship with the protein content in seeds.<sup>28</sup> Compared to WT, *bnamyb52* mutants display a lower protein content of mature seeds, while OE lines exhibit a higher protein content (Figure S4A). No significant difference is observed in soluble sugar content between WT and the transgenic plants (Figure S4B). Next, we investigated whether *BnaMYB52* influenced yield-related traits under field conditions in *B. napus*. At the mature stage, the yield components (thousand-seed weight, seed number per pod, and pod number) and yield-related traits (branch number, plant height, main inflorescence length, and pod length) of *bnamyb52* mutants and OE-1 are comparable to those of WT (Figures S4C–S4I). Additionally, *bnamyb52* mutants and overexpression lines do not show significantly altered growth and flowering time (Figures S4J and S4K). To test whether *MYB52* affects germination rate, we conducted germination assays for WT and *bnamyb52* mutant seeds harvested in May 2024 on filter paper. From 18 to 24 h, the germination rate of *bnamyb52* mutants is significantly lower than that of WT ( $p < 0.05$ ). However, from 24 to 30 h, the germination rate between WT and *bnamyb52* mutants has no significant difference. The germination rates of all materials are ultimately over 95% (Figure S4L). Those data indicate that genetic manipulation of *BnaMYB52* does not result in a significant impact on yield in *B. napus*.

#### ***BnaA09.MYB52* is predominantly expressed in the seed coat during seed development**

To determine the expression pattern of *BnaA09.MYB52*, RT-qPCR was performed in multiple tissues of the *B. napus* variety Westar. Transcripts of *BnaA09.MYB52* are highly enriched in pod walls, seeds, stems, and roots, but with relatively low expression levels in leaves, petals, and buds. During seed development, *BnaA09.MYB52* expression increases rapidly after 25 days after flowering (DAF), peaks at 40 DAF, and then decreases at 45 DAF (Figure 3A). We performed a histochemical staining analysis of transgenic *Arabidopsis* plants expressing the *pBnaA09.MYB52:GUS* ( $\beta$ -glucuronidase) reporter gene to precisely investigate the expression pattern of *BnaA09.MYB52*. GUS staining shows that *BnaA09.MYB52* is strongly expressed in siliques, seeds (at the early developmental stages), seedlings, and stems, whereas it is weakly expressed in inflorescences and rosette leaves (Figures 3B–3G), which is consistent with RT-qPCR results. Of the tissues observed, we mainly focused on the *BnaA09.MYB52* expression in seeds that consist of the seed coat and the embryo. Intriguingly, among the two parts of seeds, *BnaA09.MYB52* is only transcribed in the seed coat, whereas it is barely expressed in the embryo (Figures 3H and 3I). In *Arabidopsis*, *MYB52* is also shown to be mainly expressed

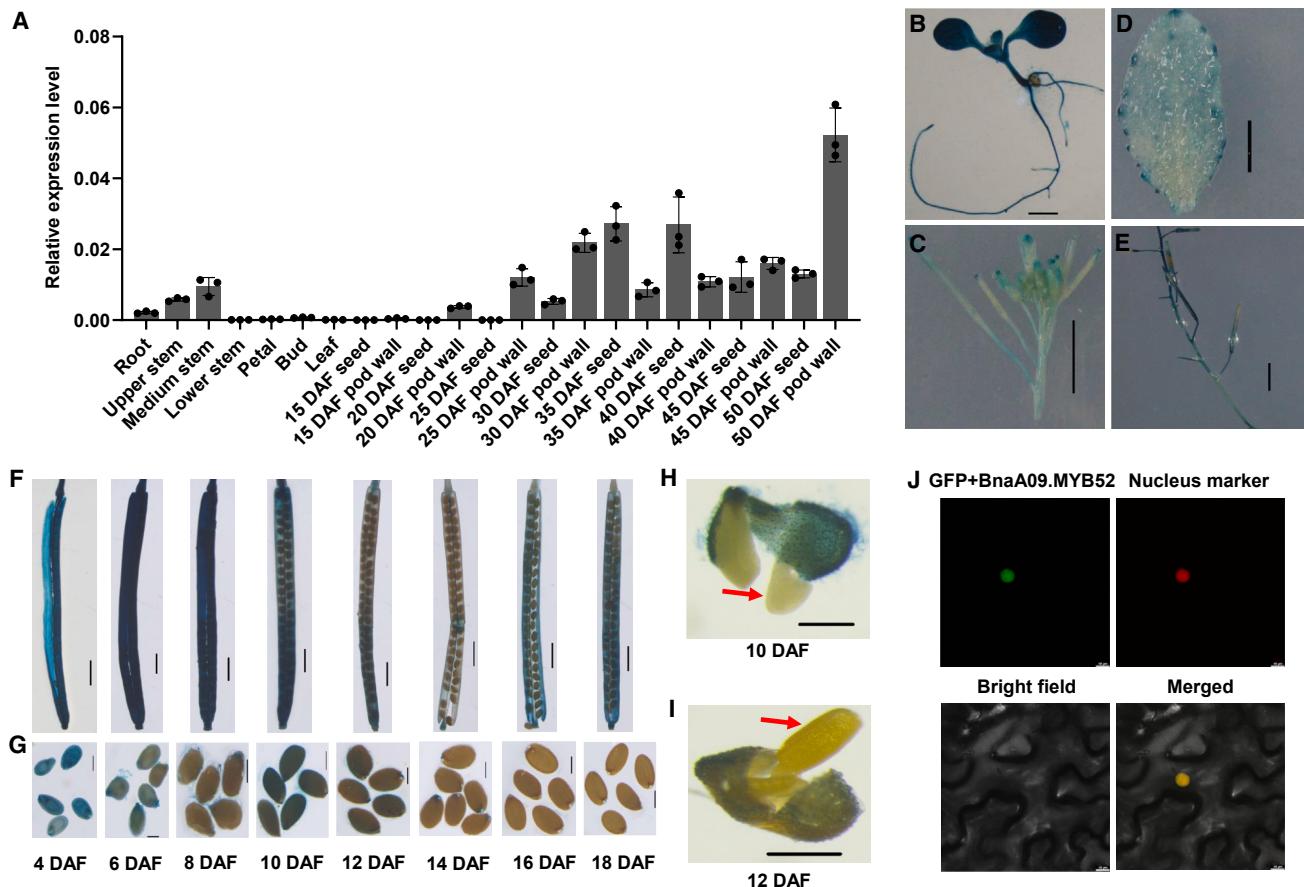
in the seed coat during seed development.<sup>18</sup> In addition, confocal microscopy analysis using *Nicotiana benthamiana* leaves expressing *BnaA09.MYB52-GFP* reveals that *BnaA09.MYB52* is localized in the nucleus (Figure 3J), supporting its function as a transcription factor. These results together suggest that *BnaA09.MYB52* has a potential function in seed coat development in *B. napus*.

#### ***BnaA09.MYB52* activates *BnaPMEI6*, *BnaPMEI14*, and *BnaSBT1.7* gene expression**

In *Arabidopsis*, *MYB52* binds to the promoters of *PMEI6*, *PMEI14*, and *SBT1.7* containing an MYB-binding site (AA(A/C)AAAC motif).<sup>18</sup> *PMEI6*, *PMEI14*, and *SBT1.7* are shown to regulate seed coat mucilage biosynthesis, yet whether these genes affect SCC and SOC is unknown. Therefore, we were interested in novel functions of the three genes and they were chosen for functional characterization. Here, we questioned whether *BnaA09.MYB52* could regulate *BnaPMEI6*, *BnaPMEI14*, and *BnaSBT1.7* in *B. napus*. *BnaPMEI6*, *BnaPMEI14*, and *BnaSBT1.7* contain four, two, and five homologous genes, respectively, in *B. napus*. Of these, *BnaA05.PMEI6*, *BnaC06.PMEI14*, and *BnaA09.SBT1.7* have the highest expression levels or MBS motifs and were selected for further analysis. We searched in 2-kb promoter regions of the three genes and found that they all contain AACAAAC fragments. Electrophoretic mobility shift assay (EMSA) and yeast one-hybrid (Y1H) assay together support that *BnaA09.MYB52* directly binds to the promoters of *BnaA05.PMEI6*, *BnaC06.PMEI14*, and *BnaA09.SBT1.7* (Figures 4A–4F).

To test the capacity of *BnaA09.MYB52* protein on driving the *BnaA05.PMEI6*, *BnaC06.PMEI14*, and *BnaA09.SBT1.7* expression, we performed a luciferase (LUC) assay in *Arabidopsis* leaf protoplasts. The reporter constructs containing the LUC gene driven by the *BnaA05.PMEI6*, *BnaC06.PMEI14*, and *BnaA09.SBT1.7* promoters, respectively, and the effector construct driven by the *BnaA09.MYB52* expression were co-transformed into *Arabidopsis* protoplasts. LUC activities driven by the promoters of *BnaA05.PMEI6*, *BnaC06.PMEI14*, and *BnaA09.SBT1.7* are markedly increased in the presence of *BnaA09.MYB52* relative to the control (Figures 4G–4I). To further validate that *BnaA09.MYB52* activates *BnaA05.PMEI6*, *BnaC06.PMEI14*, and *BnaA09.SBT1.7*, we detected the transcripts of the three genes in WT and the *BnaMYB52* knockout plants in *B. napus*. The expression levels of *BnaPMEI6*, *BnaPMEI14*, and *BnaSBT1.7* are downregulated in the *bnamyb52* mutants compared to WT, implying that *BnaA09.MYB52* activates the expression of *BnaPMEI6*, *BnaPMEI14*, and *BnaSBT1.7* *in vivo* (Figures 4J–4L).

In addition, we analyzed the expression patterns of *BnaPMEI6*, *BnaPMEI14*, and *BnaSBT1.7* genes in different tissues using the RNA sequencing (RNA-seq) data (<https://yanglab.hzau.edu.cn/BnIR>). The results show that *BnaPMEI6* is expressed in all tissues and has the highest expression level in seeds. The expression of *BnaSBT1.7* is high in leaves and siliques; medium in stems, buds, petals, sepals, and seeds; and low in roots. We also find that five homologous genes of *BnaSBT1.7* have expression differentiation and only *BnaC02.SBT1.7* is highly transcribed in seeds. It is noteworthy that two homologous genes of *BnaPMEI14* are both specifically expressed in seeds and barely transcribed in other tissues. During seed development, *BnaPMEI14* has the highest



**Figure 3. Expression patterns and subcellular localization of BnaA09.MYB52**

(A) Expression analysis of *BnaA09.MYB52* in different *B. napus* tissues. Samples were collected from the *B. napus* variety Westar. Error bars represent  $\pm$ SD. (B–I) GUS staining analysis of seedlings (B), inflorescences (C), rosette leaves (D), stems (E), siliques (F), and seeds (G) at different developmental stages, and embryo and seed coat at 10 (H) and 12 DAF (I) in *Arabidopsis*. DAF, day after flowering. Scale bar in (B), 0.1 mm. Scale bars in (C)–(E), 5 mm. Scale bar in (F), 1 mm. Scale bars in (G)–(I), 0.2 mm.

(J) Subcellular localization of BnaA09.MYB52-GFP fusion protein in *N. benthamiana* leaves. The PIF4-mCherry plasmid is used as a nucleus marker. Scale bars, 10  $\mu$ m.

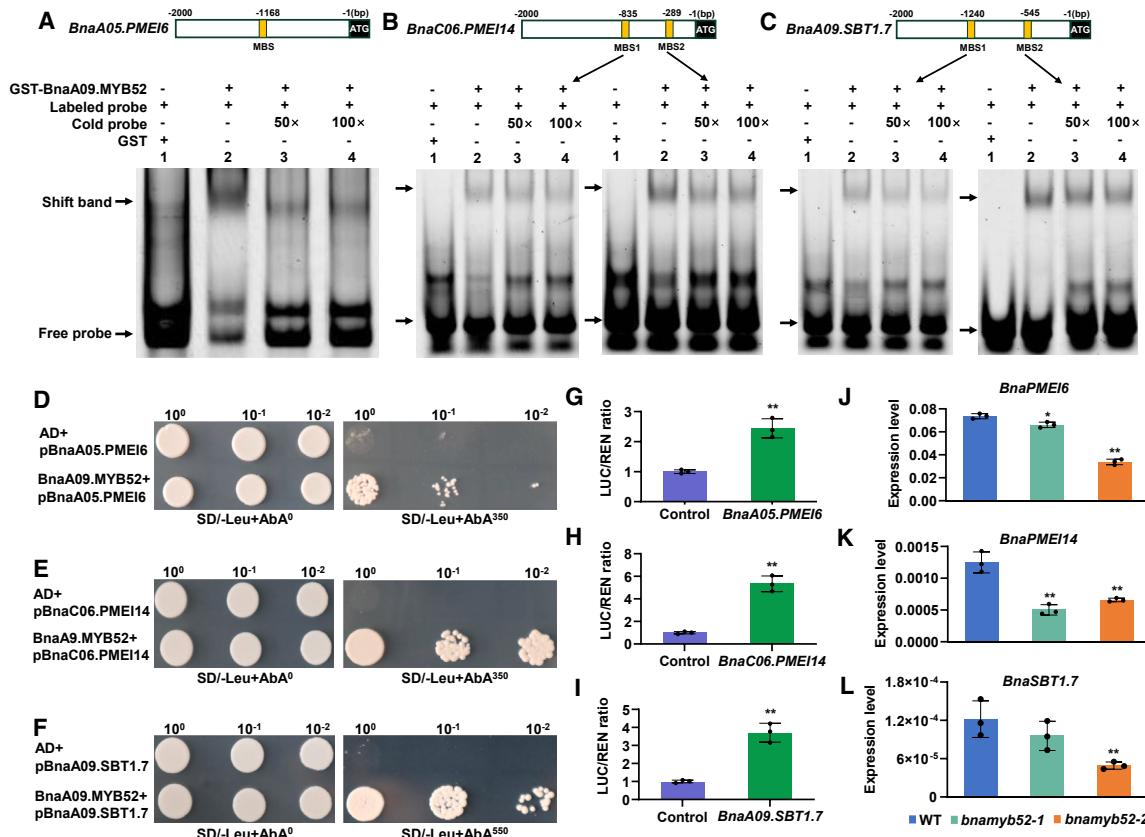
expression level at 24 DAF, and its expression disappears from 38 DAF to the mature stage, indicating that *BnaPME14* functions at the early stages of seed development.

#### ***BnaBAN* is a newly identified target gene activated by BnaMYB52**

Intriguingly, more than 100 candidates co-expressed with *MYB52* are characterized in *Arabidopsis*.<sup>18</sup> Therefore, we raised a possibility that, in addition to binding to the promoters of *BnaPME6*, *BnaPME14*, and *BnaSBT1.7*, *BnaA09.MYB52* likely targets other downstream genes. A previous study shows that the key genes in the phenylpropane metabolic pathway influence oil accumulation in *B. napus*.<sup>29</sup> Here, we find that three regulators (*TT8*, *F3'H*, and *BAN*), involved in the phenylpropane metabolic pathway, are co-expressed with *MYB52* (Figure S5A). The eGWAS indicates that the gene expression of *BnaBAN* is influenced by the eQTL localized to *BnaA09.MYB52* (Figures 1E and 5A). The variation at the eQTL affects *BnaBAN* expression at 20 DAF and SCC (Figures 5B and 5C). By performing the Y1H assay, the results

show that *BnaA09.MYB52* binds to the promoter of *BnaA01.BAN* but does not target *BnaA09.TT8*, *BnaC09.TT8*, and *BnaA10.F3'H* in *B. napus* (Figure S5B). *BnaA09.MYB52* can bind to the MBS motif and the promoters of four homologous genes of *BnaBAN* all contain one MBS fragment (Figures S5C and S5D). These data support that *BnaA09.MYB52* targets *BnaBAN* in *B. napus*.

To further study the relationship of *BnaA09.MYB52* and *BnaBAN*, RT-qPCR analysis was conducted, and the results show that the expression of *BnaBAN* is significantly decreased in the *bnaMYB52* mutants (Figure 5D). *BnaBAN* contains four copies in *B. napus*. Of these, *BnaA01.BAN* is highest expressed in seeds and was selected for the next analysis. EMSA and Y1H assays further demonstrate that *BnaA09.MYB52* directly binds to the promoter of *BnaA01.BAN* (Figures 5E and 5F). In addition, the dual-LUC assay shows that LUC expression levels driven by promoters of four homologous genes of *BnaBAN*, respectively, are significantly upregulated with co-transformation of the *BnaA09.MYB52* effector and the reporter construct in *Arabidopsis* protoplasts (Figures 5G and 5H). Taken together, these



**Figure 4.** BnaA09.MYB52 activates the expression of *BnaA05.PMEI6*, *BnaC06.PMEI14*, and *BnaA09.SBT1.7*

(A–C) EMSA assays showing that GST-BnaA09.MYB52 fusion protein directly binds to the MBS-containing regions within the promoters of *BnaA05.PMEI6* (A), *BnaC06.PMEI14* (B), and *BnaA09.SBT1.7* (C). The start codon is designated as –1 and the positions of the putative MYB-binding sequences are shown above the black boxes. The DNA probes are 5'-FAM-labeled sequences containing the MBS motif. Unlabeled probes are used as competitors (50- and 100-fold that of the hot DNA probe). The GST-tagged BnaA09.MYB52 recombinant protein was purified. The GST tag alone is used as a negative control. The MBS-containing sequences are highlighted in yellow color.

(D–F) Y1H assays showing the interaction between transcription factor BnaA09.MYB52 and the promoters of *BnaA05.PMEI6* (D), *BnaC06.PMEI14* (E), and *BnaA09.SBT1.7* (F). The empty vector pGAD7 is used as a negative control. SD/-Leu, synthetic dropout medium without leucine. AbA<sup>0</sup>, AbA<sup>350</sup>, and AbA<sup>550</sup> indicate SD/-Leu containing 0, 350, and 550 ng/mL of AbA. 10<sup>0</sup>, 10<sup>-1</sup>, and 10<sup>-2</sup> represent yeast cell concentration.

(G–I) Relative LUC activity shows that BnaA09.MYB52 activates *BnaA05.PMEI6* (G), *BnaC06.PMEI14* (H), and *BnaA09.SBT1.7* (I) expression in *Arabidopsis* protoplasts. The *REN* gene expression level is used as an internal control. The LUC/REN ratio denotes the relative activity of the promoters. Each analysis includes three technical replicates. Values are mean ± SD. Student's t test: \*\*p < 0.01.

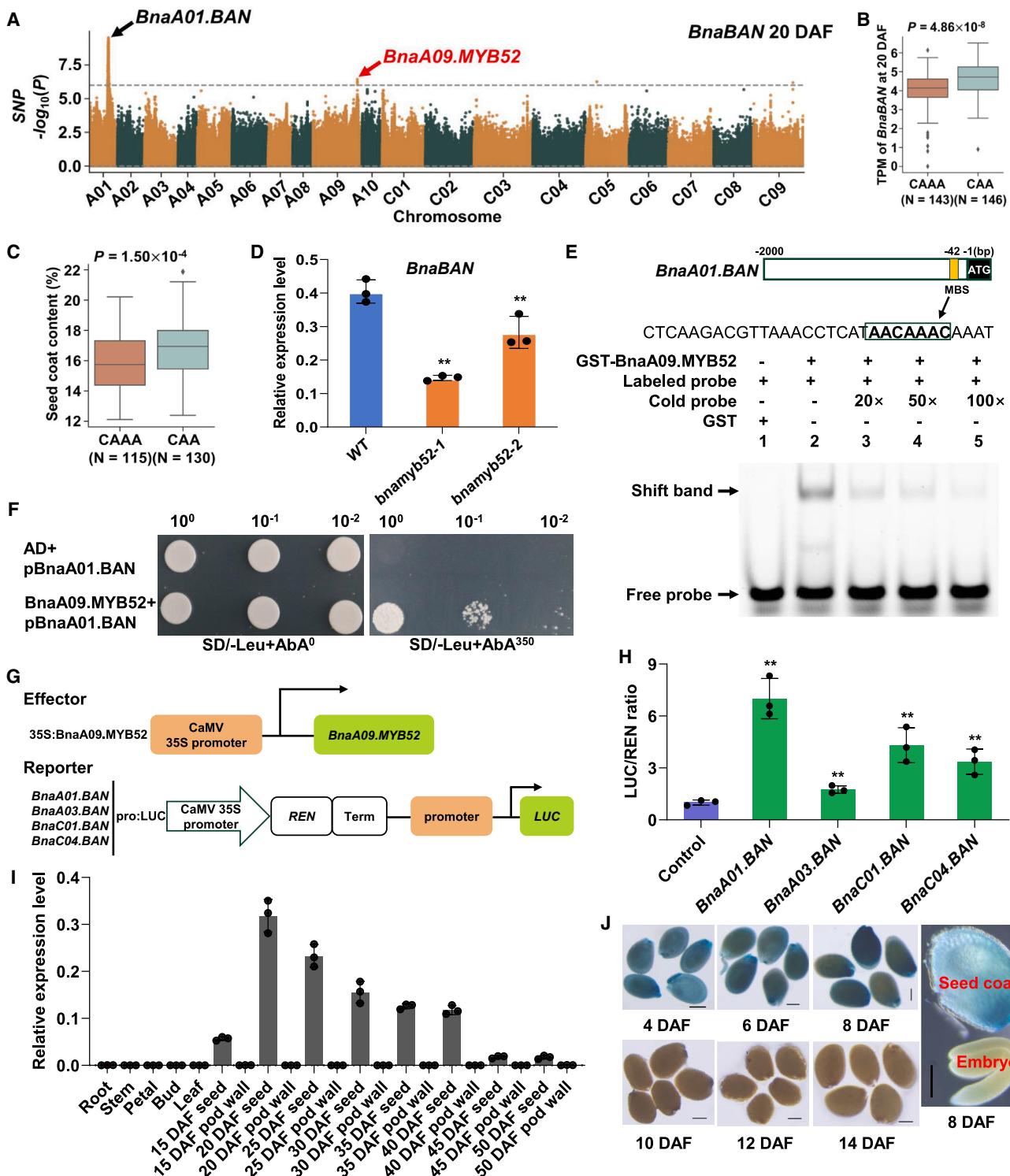
(J–L) The relative expression levels of *BnaA05.PMEI6* (J), *BnaC06.PMEI14* (K), and *BnaA09.SBT1.7* (L) in WT and BnaMYB52 knockout lines in *B. napus*. Total RNA was extracted from the seeds at 40 DAF. Data represent the mean ± SD. \*\*p < 0.01, Student's t test.

data further support that BnaA09.MYB52 directly targets the *BnaBAN* promoters and promotes *BnaBAN* expression in *B. napus*.

*BAN*, encoding anthocyanidin reductase, which converts anthocyanidins to 2,3-cis-flavan-3-ols compounds (proanthocyanidins starter units), is involved in the flavonoid biosynthesis pathway.<sup>30–33</sup> Metabolic analysis demonstrates that the content of several metabolites (such as L-phenylalanine, p-coumaric acid, gossypetin, and astragalin) in the flavonoid pathway is reduced in *bnamyb52* mutants, whereas overexpression lines result in the opposite trend compared with WT (Table S4), further implying that BnaMYB52 interacts with *BnaBAN*.

Given that *BnaBAN* acts as a downstream target of BnaA09.MYB52, we hypothesized that *BnaBAN* also functions

in seed coat development. To test this possibility, we examined the expression pattern of *BnaBAN* using the RNA-seq data, RT-qPCR, and GUS staining. The expression profile in the different tissues of the *B. napus* cultivar ZS11 shows that *BnaBAN* is specifically expressed in the seeds, whereas almost no transcript is detected in other tissues (<https://yanglab.hzau.edu.cn/BnIR>). Consistently, RT-qPCR analysis shows that *BnaBAN* has a specific expression in the seeds, with the peaked expression in the seeds at 20 DAF (Figure 5I). Histochemical staining analysis on various tissues of transgenic *Arabidopsis* lines harboring the *pBnaA01.BAN:GUS* construct also reveals that the expression is detected predominantly in the seeds, especially in the seed coat, and gradually disappears in the seeds at 10 DAF (Figures 5J, S5E–S5H). These observations together



**Figure 5. BnaA09.MYB52 directly targets BnaBAN**

(A) eGWAS Manhattan plot of *BnaBAN* at 20 DAF.

(B) Effect of variation in lead SNP on the expression of *BnaBAN* at 20 DAF. The diamonds indicate phenotype outliers.

(C) Effect of variation in lead SNP on SCC. The diamond indicates the phenotype outlier.

(D) RT-qPCR showing the relative expression levels of *BnaBAN* in the seeds at 40 DAF of WT and *BnaMYB52* knockout lines in *B. napus*. Data represent the mean ± SD. Student's t test: \*\* $p < 0.01$ .

(legend continued on next page)

imply that the highly specific expression of *BnaBAN* in the seed coat may contribute to seed coat content/thickness in *B. napus*.

### **BAN and PMEI14 regulate seed coat thickness and SOC**

Because *BnaPMEI6*, *BnaPMEI14*, *BnaSBT1.7*, and *BnaBAN* genes function as the downstream target of *BnaA09.MYB52*, we are interested in exploring whether these four genes affect seed coat content/thickness and SOC. To this end, we obtained homozygous *pmei6*, *pmei14*, *sbt1.7*, and *ban* mutants in *Arabidopsis*. All mutants were identified by PCR amplification using the left primer (LBb1.3) of the transfer DNA insertion and gene-specific primers (Figure S6A). The transfer DNA insertion sites were validated through Sanger sequencing (Figure S6B). RT-qPCR analysis shows that *ban-1* and *pmei14-1* have no transcript of *AtBAN* and *AtPMEI14*, respectively, and *ban-2*, *pmei14-3*, and *sbt1.7* show significantly reduced transcription levels of *AtBAN*, *AtPMEI14*, and *AtSBT1.7*, respectively, relative to that of WT (Col-0 background) in developing siliques (Figures S6C–S6H). However, the *AtPMEI6* transcript level in the *pmei6* mutant is approximately 1.9-fold that of WT (Figure S6G).

Subsequently, we determined SOC in WT, *ban-1*, *ban-2*, *pmei14-1*, *pmei14-3*, and *sbt1.7* (except *pmei6*) in the T<sub>2</sub> generation, and found that, in contrast to WT, *ban-1*, *ban-2*, *pmei14-1*, and *pmei14-3* show significantly elevated SOC, whereas *sbt1.7* displays an unchanged SOC at the mature stage (Figure 6A). The SOC is increased by 4.3%–4.8% and 6.3%–7.7% in the *ban* (*ban-1* and *ban-2*) and *pmei14* (*pmei14-1* and *pmei14-3*) mutants, respectively. Furthermore, the statistical analysis shows that SOC of *ban-1*, *ban-2*, *pmei14-1*, and *pmei14-3* is also elevated in the T<sub>3</sub> generation compared to WT (Figure 6B).

To evaluate whether *BAN* and *PMEI14* also affected seed coat content/thickness, we measured seed coat thickness, and statistical analysis shows that the average seed coat thickness in WT is 19.16 μm, whereas it is only 14.8, 16.2, 16.7, and 15.0 μm in *ban-1*, *ban-2*, *pmei14-1*, and *pmei14-3* plants, respectively, in the T<sub>2</sub> generation (Figure 6C). The seed coat thickness of *ban* and *pmei14* mutant plants exhibits a 15.7%–22.8% and 12.9%–21.6% reduction, respectively. Additionally, histological observation reveals that *ban* and *pmei14* mutants all display an obviously decreased seed coat thickness compared to WT (Figure 6D). The SOC exhibited a significantly negative correlation with the seed coat thickness (correlation coefficient is 0.82,  $p < 0.0001$ ) calculated by ANOVA analysis, which is in line with the previous study in *B. napus*.<sup>8,34</sup> These data demonstrate that *BAN* and *PMEI14* positively regulate seed coat thickness and negatively affect SOC.

Taken together, our results show that MYB52-BAN/PMEI14 modules play a key role during seed coat development and can be manipulated to increase oil content (Figure 6E).

### **DISCUSSION**

Seed coat is a key component of seed and its development is closely associated with oil accumulation in *B. napus*. Previously, several *TT* family genes such as *BnaTT1*, *BnaTT2*, *BnaTT7*, *BnaTT8*, *BnaTT10*, *BnaTT12*, and *BnaTT18* in the flavonoid biosynthesis pathway have been shown to decrease seed coat content/thickness and increase SOC through CRISPR-Cas9-mediated mutation in *B. napus*.<sup>4</sup> *BnaCCRL*, encoding cinnamoyl-CoA reductase-like protein involved in lignin synthesis, and double-knockout mutants displayed reduced seed coat thickness and elevated SOC.<sup>8</sup> In addition, Yu et al.<sup>29</sup> performed high-resolution time-course transcriptomes and co-expression network analysis to identify candidate genes controlling SCC and SOC in *B. napus*. Several co-located QTLs responsible for SCC and SOC were found using a RIL population,<sup>35</sup> yet how these two traits are coordinated is poorly understood. In this study, we characterize a major QTL (ZS11\_qSCC.A09) controlling SCC and SOC through GWAS, which is also characterized in our previous study and co-localized with *uqA9-12* identified by SOC GWAS in the other study.<sup>6,8</sup> *BnaA09.MYB52* is identified as a candidate gene by machine learning and haplotype analysis. Functional disruption and overexpression analysis demonstrate that *BnaA09.MYB52* positively regulates seed coat content/thickness and negatively affects SOC in *B. napus*, which is conserved in *Arabidopsis*. *MYB52* is reported to regulate seed coat mucilage formation, secondary wall thickening, as well as drought tolerance in *Arabidopsis*.<sup>18,21,22</sup> Thus, these results extend our understanding of *MYB52* function in plants. *BnaA09.-MYB52* is predominantly expressed in the seed coat during seed development, not in the embryo where a majority of the oil is accumulated. Thus, we speculate that *BnaA09.MYB52* directly regulates seed coat development and indirectly affects oil accumulation. In addition, *BnaMYB52* knockout and overexpression lines influence seed-quality-related traits (oil content, protein content, SCC, and lignin content), and no difference is observed in thousand-seed weight compared to WT. Therefore, *BnaMYB52* may not affect total photo-assimilated products into the seeds, whereas it functions in carbon re-allocation between the embryo and the seed coat in *B. napus*.

Lignin is considered an anti-nutrient that is an undesirable factor in human and animal diets.<sup>36,37</sup> In our study, *BnaMYB52* knockout lines show significantly decreased lignin content relative to WT. Moreover, our results show significantly elevated oil content in the *BnaMYB52* knockout lines compared to WT, while no significant difference is observed in yield-related traits between WT and *BnaMYB52* knockout lines. Thus, *BnaA09.MYB52* is a high-value functional gene that can be utilized not only to increase SOC but also to improve oil and meal quality in *B. napus* breeding.

(E) EMSA analysis showing that GST-BnaA09.MYB52 recombinant protein directly binds to the AACAAAC motif-containing region of the *BnaA01.BAN* promoters. GST is used as a negative control. Unlabeled probes are 20-, 50-, and 100-fold that of the hot DNA probe.

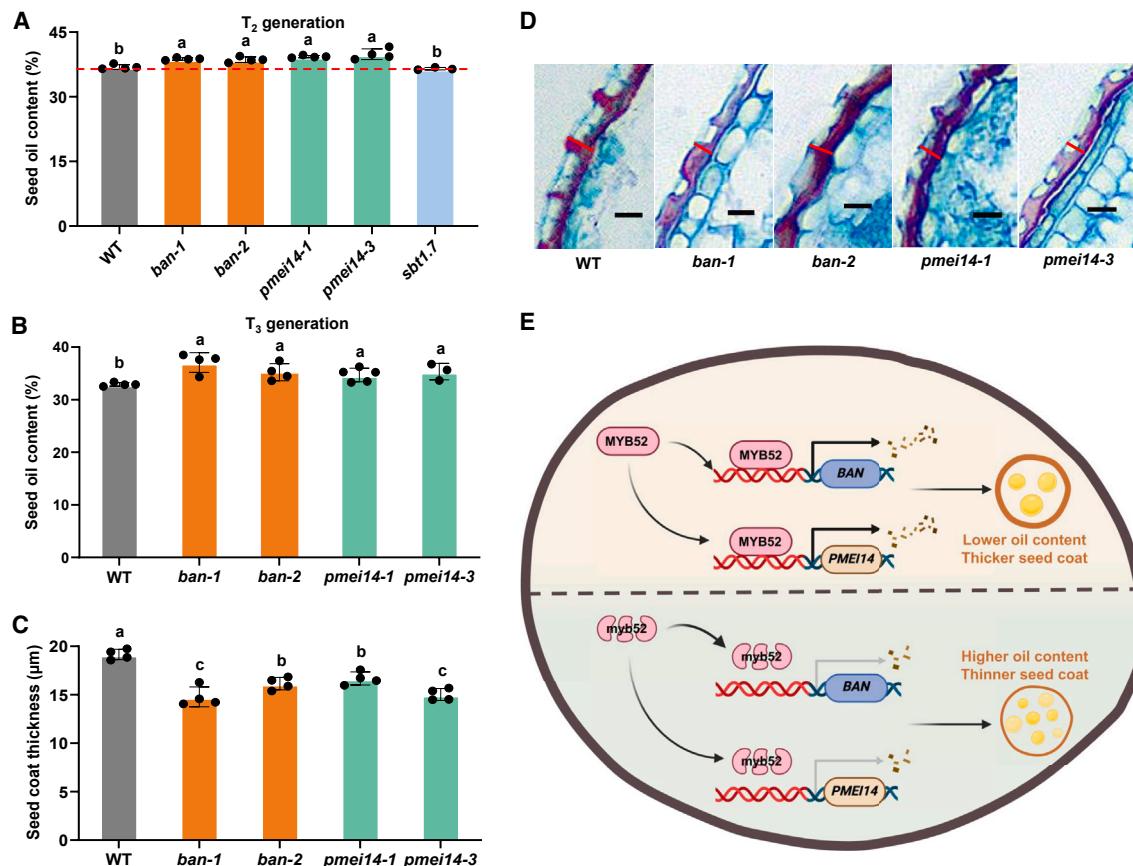
(F) Y1H assay to explore the interaction between transcription factor *BnaA09.MYB52* and *BnaA01.BAN*.

(G) Schematic diagram of vectors used in the transcriptional activity assay. The effector construct includes the *BnaA09.MYB52* CDS driven by the 35S promoter. The reporter construct contains the *REN* gene driven by the 35S promoter and the *LUC* reporter gene driven by the *BnaBAN* promoter.

(H) LUC assays in *Arabidopsis* protoplasts demonstrate that *BnaA09.MYB52* induces the transcription of *BnaBAN*. Student's t test: \*\* $p < 0.01$ .

(I) *BnaBAN* transcript levels in different *B. napus* tissues. Values represent the mean ± SD.

(J) *pBnaA01.BAN:GUS* activity levels in seeds from 4 to 14 DAF in *Arabidopsis*. Scale bars, 0.2 mm.



**Figure 6. Two downstream targets of BnaA09.MYB52, BAN and PMEI14, control seed coat thickness and SOC**

(A) The SOC of WT, *ban-1*, *ban-2*, *pmei14-1*, *pmei14-3*, and *sbt1.7* in the T<sub>2</sub> generation in *Arabidopsis*. *ban-1* and *ban-2* are two types of *BAN* transfer DNA insertional mutants. *pmei14-1* and *pmei14-3* are two types of *PMEI14* transfer DNA insertional mutants. The different lowercase letters represent significant differences ( $p < 0.05$ ) calculated by one-way ANOVA analysis with Tukey's multiple comparison test.

(B) The SOC of WT, *ban-1*, *ban-2*, *pmei14-1*, and *pmei14-3* in the T<sub>3</sub> generation in *Arabidopsis*. The different lowercase letters represent significant differences ( $p < 0.05$ ) calculated by one-way ANOVA analysis with Tukey's multiple comparison test.

(C) The seed coat thickness of WT, *ban-1*, *ban-2*, *pmei14-1*, and *pmei14-3* in the T<sub>2</sub> generation in *Arabidopsis*. The different lowercase letters represent significant differences ( $p < 0.05$ ) calculated by one-way ANOVA analysis with Tukey's multiple comparison test.

(D) The representative images of seed coat in WT, *ban-1*, *ban-2*, *pmei14-1*, and *pmei14-3*. Scale bars, 20  $\mu$ m.

(E) A proposed working model depicting the role of MYB52 in coordinating seed coat development and seed oil accumulation. MYB52 directly binds to the promoters of *BAN* and *PMEI14* and increases their transcript levels, resulting in increased SCC and decreased SOC (top), while MYB52 mutation decreases *BAN* and *PMEI14* expression, which reduces seed coat content/thickness and elevates the content of seed oil (bottom).

*PMEI14*, encoding a pectin methylesterase inhibitor, interacts with the pectin methylesterase (PME) and inhibits the protein activity of PME that catalyzes the demethylesterification of cell wall pectin to release acidic pectin and methanol.<sup>18,38</sup> It was previously reported that a *pmei14* mutant showed increased seed coat mucilage demethylesterification and decreased seed coat radial cell wall thickness in *Arabidopsis*.<sup>18,39</sup> However, whether *PMEI14* regulates SCC and SOC remains unknown. In this study, we demonstrate that BnaA09.MYB52 directly binds to the promoter of *BnaC06.PMEI14* and activates its expression. The *pmei14* mutants display reduced seed coat thickness and elevated SOC, indicating an important role of *PMEI14* in regulating two traits. Because *PMEI14* mutation results in stronger attachment of seed coat mucilage and reduced radial cell wall thickness,<sup>39</sup> which may explain why seed coat

thickness of the *pmei14* mutants is decreased. Besides, PME activity is increased in the *pmei14* mutants and more methanol is released in the seed coat.<sup>39</sup> Methanol can act as the carbon source into metabolites through the Calvin-Benson cycle.<sup>40</sup> Thus, we speculate that the content of seed oil is increased due to more carbon source provided by the seed coat in the *pmei14* mutants. BnaA09.MYB52 protein also targets the promoters of *BnaA05.PMEI6* and *BnaA09.SBT1.7* involved in seed coat mucilage biosynthesis.<sup>41,42</sup> However, compared to WT, no difference was observed for SOC in two mutants in the T<sub>2</sub> generation. Whether *PMEI6* and *SBT1.7* genes regulate those two traits needs further investigation using more mutants.

Moreover, our results show that BnaA09.MYB52 directly regulates the expression of *BnaBAN* to control seed coat development and oil accumulation. *BAN* functions as a negative regulator of

flavonoid biosynthesis and affects pigment accumulation in the seed coat.<sup>43</sup> GUS staining and RT-qPCR both indicated that *BAN* is specifically expressed in the seeds, especially in the seed coat, which is in line with its functions in the seed coat. Thus, the promoter of *BAN* can be used to observe seed coat development in cell biology and synthesize metabolites specific for the seed coat in synthetic biology in the future. *BAN* mutation results in significantly reduced seed coat thickness and increased SOC, which enhances our understanding of *BAN* function in plant science. Different from *PME14*, *BAN* influences the accumulation of flavonoids in the seed coat that compete with the embryo and endosperm for carbon source, such as malonyl-CoA and phosphoenol pyruvate, resulting in changed seed coat thickness and oil content.<sup>6,29</sup> Flavonoid and lignin are derived from two branches of the phenylpropanoid pathway. They are both synthesized from phenylalanine and catalyzed by phenylalanine ammonia lyase (PAL), cinnamate 4-hydroxylase (C4H), and 4-coumarate coenzyme A ligase (4CL).<sup>44,45</sup> The content of lignin is reduced in *bnamyb52* mutants, possibly because of affected flavonoid synthesis. Taken together, our results establish a molecular link between seed coat development and oil accumulation.

*B. napus* is the third-largest source of vegetable oil worldwide and increasing oil content is a long-term breeding target.<sup>46</sup> However, elevating oil content in *B. napus* is slow due to a lack of more elite alleles/genes. Our study identified *BnaMYB52-BnaPME14/BnaBAN* modules regulating SOC through re-allocating carbon source between the seed coat and the embryo. SCC and SOC show a trade-off effect that can be utilized in breeding cultivars with high SOC. It is possible to introduce the superior *BnaA09.MYB52* allele into *B. napus* cultivars with thicker seed coat and lower oil content using marker-assisted breeding. Additionally, we explore knocking out *BnaMYB52* and downstream target genes via the genome-editing method to improve oil content and meal quality in more *B. napus* cultivars. In this regard, we provide valuable functional genes for genetic and molecular manipulation toward breeding high-oil varieties in *B. napus*.

### Limitations of the study

We demonstrated that *BnaMYB52* mutation results in increased SOC in the spring-type variety Westar. However, whether *BnaMYB52* can be utilized to improve oil content in the different genetic background varieties needs to be further explored. Although we identified the downstream target genes of *BnaA09.MYB52* and revealed the role of *PME14* and *BAN* in regulating seed coat development and oil accumulation in *B. napus*, we still need more investigations to identify other target genes regulated by *BnaMYB52*. In addition, identifying the interacting proteins with *BnaMYB52* and investigating how the interaction affects the expression of downstream target genes are important issues that will facilitate our understanding of the regulatory network mediated by *BnaMYB52*.

### RESOURCE AVAILABILITY

#### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Liang Guo ([guoliang@mail.hzau.edu.cn](mailto:guoliang@mail.hzau.edu.cn)).

### Materials availability

Plasmids and transgenic plants generated in this study are available from the lead contact upon request. This study did not generate new unique reagents.

### Data and code availability

- The metabolomics data have been deposited at the National Metabolomics Data Repository (NMDR) website, the Metabolomics Workbench, and are publicly available as of the date of publication. Accession numbers and links to the repositories are listed in the [key resources table](#). This paper also analyzes existing, publicly available data from ZS11.v0 genome and RNA-seq, accessible at the URLs listed in the [key resources table](#).
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

### ACKNOWLEDGMENTS

We thank Dr. Anming Ding (Tobacco Research Institute, Chinese Academy of Agricultural Sciences) for providing seeds of *Arabidopsis* mutants *pmei14-1*, *pmei14-3*, *pmei6*, and *sbt1.7*. This work is supported by the Biological Breeding-National Science and Technology Major Project (2023ZD04069), National Science Fund for Distinguished Young Scholars (32225037), and Hubei Provincial Natural Science Foundation (2024AFB176).

### AUTHOR CONTRIBUTIONS

L.G., X.Y., and Z.T. conceived and supervised the project. J.Y., X.W., and Y.L. performed most of the experiments. X.W., X.L., and Y.Xiang performed the *B. napus* transformations. H. Zhang performed the phenotyping assay of the *B. napus* germplasms. J.C. conducted the liquid chromatography-mass spectrometry (LC-MS) analysis. Z.T., Y.Z., Y.Xia and H. Zhao performed the bioinformatics analysis. J.Y. and L.G. wrote the manuscript. L.G., X.Y., Z.T., X.W., Y.Z., and J.C. revised the manuscript. All authors read and approved the final manuscript.

### DECLARATION OF INTERESTS

The authors declare no competing interests.

### STAR METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Plant materials and growth conditions
- METHOD DETAILS
  - Trait investigation
  - GWAS and candidate gene prioritization
  - Plasmid construction and plant transformation
  - Fatty acid extraction and GC-FID analysis
  - Metabolite analysis by LC-MS/MS
  - Determination of lignin content
  - GUS staining and subcellular localization
  - Cytological analysis
  - Yeast one-hybrid (Y1H) assay
  - Electrophoretic mobility shift assay (EMSA)
  - Dual-luciferase assay
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Real-time PCR
  - Statistical analysis

### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.celrep.2025.115280>.

Received: September 16, 2024  
 Revised: November 25, 2024  
 Accepted: January 16, 2025  
 Published: February 11, 2025

## REFERENCES

1. Figueiredo, D.D., and Köhler, C. (2014). Signalling events regulating seed coat development. *Biochem. Soc. Trans.* 42, 358–363. <https://doi.org/10.1042/BST20130221>.
2. Zheng, M., Terzaghi, W., Wang, H., and Hua, W. (2022). Integrated strategies for increasing rapeseed yield. *Trends Plant Sci.* 27, 742–745. <https://doi.org/10.1016/j.tplants.2022.03.008>.
3. Qu, C., Zhu, M., Hu, R., Niu, Y., Chen, S., Zhao, H., Li, C., Wang, Z., Yin, N., Sun, F., et al. (2023). Comparative genomic analyses reveal the genetic basis of the yellow-seed trait in *Brassica napus*. *Nat. Commun.* 14, 5194. <https://doi.org/10.1038/s41467-023-40838-1>.
4. Li, H., Yu, K., Zhang, Z., Yu, Y., Wan, J., He, H., and Fan, C. (2024). Targeted mutagenesis of flavonoid biosynthesis pathway genes reveals functional divergence in seed coat colour, oil content and fatty acid composition in *Brassica napus* L. *Plant Biotechnol. J.* 22, 445–459. <https://doi.org/10.1111/pbi.14197>.
5. Haughn, G., and Chaudhury, A. (2005). Genetic analysis of seed coat development in *Arabidopsis*. *Trends Plant Sci.* 10, 472–477. <https://doi.org/10.1016/j.tplants.2005.08.005>.
6. Miao, L., Chao, H., Chen, L., Wang, H., Zhao, W., Li, B., Zhang, L., Li, H., Wang, B., and Li, M. (2019). Stable and novel QTL identification and new insights into the genetic networks affecting seed fiber traits in *Brassica napus*. *Theor. Appl. Genet.* 132, 1761–1775. <https://doi.org/10.1007/s00122-019-03313-4>.
7. Vijayan, S.S., Nagarajappa, N., and Ranjitha, H.P. (2023). Seed coat composition in black and white soybean seeds with differential water permeability. *Plant Biol.* 25, 935–943. <https://doi.org/10.1111/plb.13551>.
8. Zhang, Y., Zhang, H., Zhao, H., Xia, Y., Zheng, X., Fan, R., Tan, Z., Duan, C., Fu, Y., Li, L., et al. (2022). Multi-omics analysis dissects the genetic architecture of seed coat content in *Brassica napus*. *Genome Biol.* 23, 86. <https://doi.org/10.1186/s13059-022-02647-5>.
9. Wang, J., Jian, H., Wei, L., Qu, C., Xu, X., Lu, K., Qian, W., Li, J., Li, M., and Liu, L. (2015). Genome-wide analysis of seed acid detergent lignin (ADL) and hull content in rapeseed (*Brassica napus* L.). *PLoS One* 10, e0145045. <https://doi.org/10.1371/journal.pone.0145045>.
10. Jin, M.Y., Li, J.N., Fu, F.Y., Zhang, Z.S., Zhang, X.K., and Liu, L.Z. (2007). QTL analysis of the oil content and the hull content in *Brassica napus* L. *Agric. Sci. China* 6, 414–421. [https://doi.org/10.1016/S1671-2927\(07\)60064-9](https://doi.org/10.1016/S1671-2927(07)60064-9).
11. Liu, L.Z., Lin, N., Chen, L., Tang, Z.L., Zhang, X.K., and Li, J.N. (2006). QTL analysis of 5 important traits in *Brassica napus*. *Journal of Agricultural Biotechnology* 14, 747–751. <https://doi.org/10.3969/j.issn.1674-7968.2006.05.021>.
12. Zhai, Y., Yu, K., Cai, S., Hu, L., Amoo, O., Xu, L., Yang, Y., Ma, B., Jiao, Y., Zhang, C., et al. (2020). Targeted mutagenesis of *BnTT8* homologs controls yellow seed coat development for effective oil production in *Brassica napus* L. *Plant Biotechnol. J.* 18, 1153–1168. <https://doi.org/10.1111/pbi.13281>.
13. Xie, T., Chen, X., Guo, T., Rong, H., Chen, Z., Sun, Q., Batley, J., Jiang, J., and Wang, Y. (2020). Targeted knockout of *BnTT2* homologues for yellow-seeded *Brassica napus* with reduced flavonoids and improved fatty acid composition. *J. Agric. Food Chem.* 68, 5676–5690. <https://doi.org/10.1021/acs.jafc.0c01126>.
14. Lian, J., Lu, X., Yin, N., Ma, L., Lu, J., Liu, X., Li, J., Lu, J., Lei, B., Wang, R., and Chai, Y. (2017). Silencing of *BnTT1* family genes affects seed flavonoid biosynthesis and alters seed fatty acid composition in *Brassica napus*. *Plant Sci.* 254, 32–47. <https://doi.org/10.1016/j.plantsci.2016.10.012>.
15. Cheng, H., Cai, S., Hao, M., Cai, Y., Wen, Y., Huang, W., Mei, D., and Hu, Q. (2024). Targeted mutagenesis of *BnTTG1* homologues generated yel-
- low-seeded rapeseed with increased oil content and seed germination under abiotic stress. *Plant Physiol. Biochem.* 206, 108302. <https://doi.org/10.1016/j.plaphy.2023.108302>.
16. Millard, P.S., Kragelund, B.B., and Burow, M. (2019). R2R3 MYB transcription factors - functions outside the DNA-binding domain. *Trends Plant Sci.* 24, 934–946. <https://doi.org/10.1016/j.tplants.2019.07.003>.
17. Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C., and Lepiniec, L. (2010). MYB transcription factors in *Arabidopsis*. *Trends Plant Sci.* 15, 573–581. <https://doi.org/10.1016/j.tplants.2010.06.005>.
18. Shi, D., Ren, A., Tang, X., Qi, G., Xu, Z., Chai, G., Hu, R., Zhou, G., and Kong, Y. (2018). *MYB52* negatively regulates pectin demethylesterification in seed coat mucilage. *Plant Physiol.* 176, 2737–2749. <https://doi.org/10.1104/pp.17.01771>.
19. Ding, A., Tang, X., Yang, D., Wang, M., Ren, A., Xu, Z., Hu, R., Zhou, G., O'Neill, M., and Kong, Y. (2021). ERF4 and MYB52 transcription factors play antagonistic roles in regulating homogalacturonan de-methylesterification in *Arabidopsis* seed coat mucilage. *Plant Cell* 33, 381–403. <https://doi.org/10.1093/plcell/kocab031>.
20. Cassan-Wang, H., Goué, N., Saidi, M.N., Legay, S., Sivadon, P., Goffner, D., and Grima-Pettenati, J. (2013). Identification of novel transcription factors regulating secondary cell wall formation in *Arabidopsis*. *Front. Plant Sci.* 4, 189. <https://doi.org/10.3389/fpls.2013.00189>.
21. Zhong, R., Lee, C., Zhou, J., McCarthy, R.L., and Ye, Z.H. (2008). A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in *Arabidopsis*. *Plant Cell* 20, 2763–2782. <https://doi.org/10.1105/tpc.108.061325>.
22. Park, M.Y., Kang, J.Y., and Kim, S.Y. (2011). Overexpression of *At-MYB52* confers ABA hypersensitivity and drought tolerance. *Mol. Cell.* 31, 447–454. <https://doi.org/10.1007/s10059-011-0300-7>.
23. Ling, X.U., Wang, Y.C., Xiao-Lan, H.E., Huang, Y.H., and Zhang, D.Y. (2017). Isolation, expression and binding function analysis of the transcription factor GmMYB52 in soybean. *Acta Agron. Sin.* 43, 1458. <https://doi.org/10.3724/SP.J.1006.2017.01458>.
24. Ma, J., He, T., Yu, R., Zhao, Y., Hu, H., Zhang, W., Zhang, Y., Liu, Z., and Chen, M. (2024). *Brassica napus BnaA09.MYB52* enhances seed coat mucilage accumulation and tolerance to osmotic stress during seed germination in *Arabidopsis thaliana*. *Plant Biol.* 26, 602–611. <https://doi.org/10.1111/plb.13641>.
25. Tang, S., Zhao, H., Lu, S., Yu, L., Zhang, G., Zhang, Y., Yang, Q.Y., Zhou, Y., Wang, X., Ma, W., 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, 470–487. <https://doi.org/10.1016/j.molp.2020.12.003>.
26. Tan, Z., Peng, Y., Xiong, Y., Xiong, F., Zhang, Y., Guo, N., Tu, Z., Zong, Z., Wu, X., Ye, J., et al. (2022). Comprehensive transcriptional variability analysis reveals gene networks regulating seed oil content of *Brassica napus*. *Genome Biol.* 23, 233. <https://doi.org/10.1186/s13059-022-02801-z>.
27. Chapman, K.D., and Ohlrogge, J.B. (2012). Compartmentation of triacylglycerol accumulation in plants. *J. Biol. Chem.* 287, 2288–2294. <https://doi.org/10.1074/jbc.R111.290072>.
28. Hu, Z.Y., Hua, W., Zhang, L., Deng, L.B., Wang, X.F., Liu, G.H., Hao, W.J., and Wang, H.Z. (2013). Seed structure characteristics to form ultrahigh oil content in rapeseed. *PLoS One* 8, e62099. <https://doi.org/10.1371/journal.pone.0062099>.
29. Yu, L., Liu, D., Yin, F., Yu, P., Lu, S., Zhang, Y., Zhao, H., Lu, C., Yao, X., Dai, C., et al. (2023). Interaction between phenylpropane metabolism and oil accumulation in the developing seed of *Brassica napus* revealed by high temporal-resolution transcriptomes. *BMC Biol.* 21, 202. <https://doi.org/10.1186/s12915-023-01705-z>.
30. Xie, D.Y., Sharma, S.B., Paiva, N.L., Ferreira, D., and Dixon, R.A. (2003). Role of anthocyanidin reductase, encoded by *BANYULS* in plant flavonoid biosynthesis. *Science* 299, 396–399. <https://doi.org/10.1126/science.1078540>.

31. Yu, K., Song, Y., Lin, J., and Dixon, R.A. (2023). The complexities of proanthocyanidin biosynthesis and its regulation in plants. *Plant Commun.* 4, 100498. <https://doi.org/10.1016/j.xplc.2022.100498>.
32. Jun, J.H., Lu, N., Docampo-Palacios, M., Wang, X., and Dixon, R.A. (2021). Dual activity of anthocyanidin reductase supports the dominant plant proanthocyanidin extension unit pathway. *Sci. Adv.* 7, eabg4682. <https://doi.org/10.1126/sciadv.abg4682>.
33. Baudry, A., Heim, M.A., Dubreucq, B., Caboche, M., Weisshaar, B., and Lepiniec, L. (2004). TT2, TT8, and TTG1 synergistically specify the expression of BANYULS and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. *Plant J.* 39, 366–380. <https://doi.org/10.1111/j.1365-313X.2004.02138.x>.
34. Tan, Z., Han, X., Dai, C., Lu, S., He, H., Yao, X., Chen, P., Yang, C., Zhao, L., Yang, Q.Y., et al. (2024). Functional genomics of *Brassica napus*: Progress, challenges, and perspectives. *J. Integr. Plant Biol.* 66, 484–509. <https://doi.org/10.1111/jipb.13635>.
35. Yan, X.Y., Li, J.N., Fu, F.Y., Jin, M.Y., Chen, L., and Liu, L.Z. (2009). Co-location of seed oil content, seed hull content and seed coat color QTL in three different environments in *Brassica napus* L. *Euphytica* 170, 355–364. <https://doi.org/10.1007/s10681-009-0006-5>.
36. Hannoufa, A., Pillai, B.V.S., and Chellamma, S. (2014). Genetic enhancement of *Brassica napus* seed quality. *Transgenic Res.* 23, 39–52. <https://doi.org/10.1007/s11248-013-9742-3>.
37. Chen, Y.Y., Lu, H.Q., Jiang, K.X., Wang, Y.R., Wang, Y.P., and Jiang, J.J. (2022). The flavonoid biosynthesis and regulation in *Brassica napus*: A review. *Int. J. Mol. Sci.* 24, 357. <https://doi.org/10.3390/ijms24010357>.
38. Wang, M., Yuan, D., Gao, W., Li, Y., Tan, J., and Zhang, X. (2013). A comparative genome analysis of PME and PMEI families reveals the evolution of pectin metabolism in plant cell walls. *PLoS One* 8, e72082. <https://doi.org/10.1371/journal.pone.0072082>.
39. Allen, P.J., Napoli, R.S., Parish, R.W., and Li, S.F. (2023). MYB-bHLH-TTG1 in a multi-tiered pathway regulates *Arabidopsis* seed coat mucilage biosynthesis genes including *PECTIN METHYLESTERASE INHIBITOR14* required for homogalacturonan demethylesterification. *Plant Cell Physiol.* 64, 906–919. <https://doi.org/10.1093/pcp/pcad065>.
40. Michel, F. (2001). Pectin methylesterases: cell wall enzymes with important roles in plant physiology. *Trends Plant Sci.* 6, 414–419. [https://doi.org/10.1016/s1360-1385\(01\)02045-3](https://doi.org/10.1016/s1360-1385(01)02045-3).
41. Rautengarten, C., Usadel, B., Neumetzler, L., Hartmann, J., Büsis, D., and Altmann, T. (2008). A subtilisin-like serine protease essential for mucilage release from *Arabidopsis* seed coats. *Plant J.* 54, 466–480. <https://doi.org/10.1111/j.1365-313X.2008.03437.x>.
42. Saez-Aguayo, S., Ralet, M.C., Berger, A., Botran, L., Ropartz, D., Marion-Poll, A., and North, H.M. (2013). *PECTIN METHYLESTERASE INHIBITOR 6* promotes *Arabidopsis* mucilage release by limiting methylesterification of homogalacturonan in seed coat epidermal cells. *Plant Cell* 25, 308–323. <https://doi.org/10.1105/tpc.112.106575>.
43. Albert, S., Delseny, M., and Devic, M. (1997). BANYULS, a novel negative regulator of flavonoid biosynthesis in the *Arabidopsis* seed coat. *Plant J.* 11, 289–299. <https://doi.org/10.1046/j.1365-313x.1997.11020289.x>.
44. Lan, H.N., Liu, R.Y., Liu, Z.H., Li, X., Li, B.Z., and Yuan, Y.J. (2023). Biological valorization of lignin to flavonoids. *Biotechnol. Adv.* 64, 108107. <https://doi.org/10.1016/j.biotechadv.2023.108107>.
45. An, J.P., Li, H.L., Liu, Z.Y., Wang, D.R., You, C.X., and Han, Y. (2023). The E3 ubiquitin ligase SINA1 and the protein kinase BIN2 cooperatively regulate *PHR1* in apple anthocyanin biosynthesis. *J. Integr. Plant Biol.* 65, 2175–2193. <https://doi.org/10.1111/jipb.13538>.
46. Liu, J., Li, M., Zhang, Q., Wei, X., and Huang, X. (2020). Exploring the molecular basis of heterosis for plant breeding. *J. Integr. Plant Biol.* 62, 287–298. <https://doi.org/10.1111/jipb.12804>.
47. Li, N., Song, D., Peng, W., Zhan, J., Shi, J., Wang, X., Liu, G., and Wang, H. (2019). Maternal control of seed weight in rapeseed (*Brassica napus* L.): the causal link between the size of pod (mother, source) and seed (offspring, sink). *Plant Biotechnol. J.* 17, 736–749. <https://doi.org/10.1111/pbi.13011>.
48. Ye, J., Liang, H., Zhao, X., Li, N., Song, D., Zhan, J., Liu, J., Wang, X., Tu, J., Varshney, R.K., et al. (2023). A systematic dissection in oilseed rape provides insights into the genetic architecture and molecular mechanism of yield heterosis. *Plant Biotechnol. J.* 21, 1479–1495. <https://doi.org/10.1111/pbi.14054>.
49. Shi, J., Zhan, J., Yang, Y., Ye, J., Huang, S., Li, R., Wang, X., Liu, G., and Wang, H. (2015). Linkage and regional association analysis reveal two new tightly-linked major-QTLs for pod number and seed number per pod in rapeseed (*Brassica napus* L.). *Sci. Rep.* 5, 14481. <https://doi.org/10.1038/srep14481>.
50. Hu, J., Chen, B., Zhao, J., Zhang, F., Xie, T., Xu, K., Gao, G., Yan, G., Li, H., Li, L., et al. (2022). Genomic selection and genetic architecture of agronomic traits during modern rapeseed breeding. *Nat. Genet.* 54, 694–704. <https://doi.org/10.1038/s41588-022-01055-6>.
51. Li, H. (2014). Toward better understanding of artifacts in variant calling from high-coverage samples. *Bioinformatics* 30, 2843–2851. <https://doi.org/10.1093/bioinformatics/btu356>.
52. Widmer, C., Lippert, C., Weissbrod, O., Fusi, N., Kadie, C., Davidson, R., Listgarten, J., and Heckerman, D. (2014). Further improvements to linear mixed models for genome-wide association studies. *Sci. Rep.* 4, 6874. <https://doi.org/10.1038/srep06874>.
53. Li, M.X., Yeung, J.M.Y., Cherny, S.S., and Sham, P.C. (2012). Evaluating the effective numbers of independent tests and significant P-value thresholds in commercial genotyping arrays and public imputation reference datasets. *Hum. Genet.* 131, 747–756. <https://doi.org/10.1007/s00439-011-1118-2>.
54. Xing, H.L., Dong, L., Wang, Z.P., Zhang, H.Y., Han, C.Y., Liu, B., Wang, X.C., and Chen, Q.J. (2014). A CRISPR/Cas9 toolkit for multiplex genome editing in plants. *BMC Plant Biol.* 14, 327. <https://doi.org/10.1186/s12870-014-0327-y>.
55. Zheng, M., Zhang, L., Tang, M., Liu, J., Liu, H., Yang, H., Fan, S., Terzaghi, W., Wang, H., and Hua, W. (2020). Knockout of two *BnaMAX1* homologs by CRISPR/Cas9-targeted mutagenesis improves plant architecture and increases yield in rapeseed (*Brassica napus* L.). *Plant Biotechnol. J.* 18, 644–654. <https://doi.org/10.1111/pbi.13228>.
56. Hong, Y., Xia, H., Li, X., Fan, R., Li, Q., Ouyang, Z., Tang, S., and Guo, L. (2022). *Brassica napus BnaNTT1* modulates ATP homeostasis in plastids to sustain metabolism and growth. *Cell Rep.* 40, 111060. <https://doi.org/10.1016/j.celrep.2022.111060>.
57. Liu, Y., Du, Z., Li, Y., Lu, S., Tang, S., and Guo, L. (2024). Improving linolenic acid content in rapeseed oil by overexpression of *CsFAD2* and *CsFAD3* genes. *Mol. Breed.* 44, 9. <https://doi.org/10.1007/s11032-024-01445-0>.
58. Shi, T., Zhu, A., Jia, J., Hu, X., Chen, J., Liu, W., Ren, X., Sun, D., Fernie, A.R., Cui, F., and Chen, W. (2020). Metabolomics analysis and metabolite-agronomic trait associations using kernels of wheat (*Triticum aestivum*) recombinant inbred lines. *Plant J.* 103, 279–292. <https://doi.org/10.1111/tpj.14727>.
59. Li, L., Tian, Z., Chen, J., Tan, Z., Zhang, Y., Zhao, H., Wu, X., Yao, X., Wen, W., Chen, W., and Guo, L. (2023). Characterization of novel loci controlling seed oil content in *Brassica napus* by marker metabolite-based multi-omics analysis. *Genome Biol.* 24, 141. <https://doi.org/10.1186/s13059-023-02984-z>.
60. Chen, J., Hu, X., Shi, T., Yin, H., Sun, D., Hao, Y., Xia, X., Luo, J., Fernie, A.R., He, Z., and Chen, W. (2020). Metabolite-based genome-wide association study enables dissection of the flavonoid decoration pathway of wheat kernels. *Plant Biotechnol. J.* 18, 1722–1735. <https://doi.org/10.1111/pbi.13335>.
61. Xu, W., Bao, W., Liu, H., Chen, C., Bai, H., Huang, M., Zhu, G., Zhao, H., Gou, N., Chen, Y., et al. (2021). Insights into the molecular mechanisms of late flowering in *Prunus sibirica* by whole-genome and transcriptome analyses. *Front. Plant Sci.* 12, 802827.

## STAR★METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Antibodies</b>		
Rabbit monoclonal anti-GST	ABclonal	Cat# AE077; RRID: AB_3674395
<b>Bacterial and virus strains</b>		
Trelief 5a Chemically Competent Cell	Qingke	Cat# TSC-C01-100
GV3101 Chemically Competent Cell	Qingke	Cat# TSC-A01
BL21 (DE3) Competent Cell	TransGen Biotech	Cat# CD601-02
<i>S.cerevisiae</i> strain (Y1H Gold)	Lab stock	N/A
<b>Chemicals, peptides, and recombinant proteins</b>		
SD/-Ura with Agar	Coolaber	Cat# PM2272
SD/-Leu with Agar	Coolaber	Cat# PM2202
Aureobasidin A	Coolaber	Cat# CA2332G
Yeast Nitrogen Base	Coolaber	Cat# PM2070
LiAC	Sigma	Cat# L4158
PEG3350	Sigma	Cat# 88276
EDTA	Sigma	Cat# EDS
Tris	Sigma	Cat# T1503
EMSA/Gel-Shift Loading Buffer	Beyotime	Cat# GS006
EMSA/Gel-Shift Binding Buffer	Beyotime	Cat# GS005
T4 DNA Ligase	New England Biolabs	Cat# M0202S
Protease Inhibitor Cocktail	Sigma	Cat# P9599
GST.Bind Resin	Merck	Cat# 70541
KOD DNA Polymerase	TOYOBO	Cat# KFX-101
Murashige & Skoog Medium	Duchefa Biochemie	Cat# M0221.0050
Kanamycin sulfate	Sangon Biotech	Cat# A600286
<b>Critical commercial assays</b>		
β-Glucuronidase Reporter Gene Staining Kit	Coolaber	Cat# SL7160
Dual-luciferase Reporter Assay Kit	Promega	Cat# E1910
RNAprep Pure Plant Plus Kit	TIANGEN	Cat# 4992239
HiScript IV 1st Strand cDNA Synthesis Kit	Vazyme	Cat# R412
SupRealQ Ultra Hunter SYBR qPCR Master Mix	Vazyme	Cat# Q713
ClonExpress II One Step Cloning Kit	Vazyme	Cat# C112
FastPure Gel DNA Extraction Mini Kit	Vazyme	Cat# DC301
Lignin Content Assay Kit	Boxbio	Cat# AKSU010M
TIANprep Mini Plasmid Kit	TIANGEN	Cat# DP103
HighPure Maxi Plasmid Kit	TIANGEN	Cat# DP116
<b>Deposited data</b>		
Raw data of metabolome	This paper	( <a href="https://www.metabolomicsworkbench.org">https://www.metabolomicsworkbench.org</a> ) Metabolomics Workbench:ST003643
ZS11.v0 genome data (public)	BnTIR	<a href="https://yanglab.hzau.edu.cn/BnIR">https://yanglab.hzau.edu.cn/BnIR</a>
RNA-seq data (public)	BnTIR	<a href="https://yanglab.hzau.edu.cn/BnIR">https://yanglab.hzau.edu.cn/BnIR</a>
<b>Experimental models: Organisms/strains</b>		
<i>Nicotiana benthamiana</i>	This paper	N/A
Arabidopsis: Col-0	This paper	N/A
Arabidopsis: <i>atban-1</i>	AraShare	N660393
Arabidopsis: <i>atban-2</i>	AraShare	N680216

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Arabidopsis: <i>atpmei14-1</i>	Ding et al. <sup>19</sup>	SALK_206157C
Arabidopsis: <i>atpmei14-3</i>	Ding et al. <sup>19</sup>	SALK_020742
Arabidopsis: <i>atpemi6</i>	Ding et al. <sup>19</sup>	GABI_790B12
Arabidopsis: <i>atsbt1.7</i>	Ding et al. <sup>19</sup>	GABI_140B02
Arabidopsis: <i>atmyb52-1</i>	AraShare	N686878
Arabidopsis: <i>atmyb52-2</i>	AraShare	N681644
Arabidopsis: <i>pMYB52-GUS</i>	This paper	N/A
Arabidopsis: <i>pBAN:GUS</i>	This paper	N/A
<i>Brassica napus</i> : Westar	This paper	N/A
<i>Brassica napus</i> : ZS11	This paper	N/A
<i>Brassica napus</i> : BnaA09.MYB52-OE-1	This paper	N/A
<i>Brassica napus</i> : BnaA09.MYB52-OE-2	This paper	N/A
<i>Brassica napus</i> : BnaA09.MYB52-OE-5	This paper	N/A
<i>Brassica napus</i> : BnaA09.MYB52-OE-8	This paper	N/A
<i>Brassica napus</i> : <i>bnamyb52-1</i>	This paper	N/A
<i>Brassica napus</i> : <i>bnamyb52-2</i>	This paper	N/A
<b>Oligonucleotides</b>		
Primers used in this study	This paper	Table S5
<b>Recombinant DNA</b>		
pCAMBIA1305-BnaA09.MYB52-GFP	This paper	N/A
pCAMBIA2300-pBnaA09.MYB52-GUS	This paper	N/A
pCAMBIA2300-pBnaA01.BAN-GUS	This paper	N/A
pGADT7-BnaA09.MYB52	This paper	N/A
pAbAi-pBnaA09.TT8	This paper	N/A
pAbAi-pBnaC09.TT8	This paper	N/A
pAbAi-pBnaA10.F3'H	This paper	N/A
pAbAi-pBnaA01.BAN	This paper	N/A
pAbAi-pBnaC06.PMEI14	This paper	N/A
pAbAi-pBnaA05.PMEI6	This paper	N/A
pAbAi-pBnaA09.SBT1.7	This paper	N/A
pGreenII 62-SK-BnaA09.MYB52	This paper	N/A
pGreen II-0800-pBnaA01.BAN	This paper	N/A
pGreen II-0800-pBnaA03.BAN	This paper	N/A
pGreen II-0800-pBnaC01.BAN	This paper	N/A
pGreen II-0800-pBnaC04.BAN	This paper	N/A
pGreen II-0800-pBnaC06.PMEI14	This paper	N/A
pGreen II-0800-pBnaA05.PMEI6	This paper	N/A
pGreen II-0800-pBnaA09.SBT1.7	This paper	N/A
pCAMBIA6p-1-BnaA09.MYB52-GST	This paper	N/A
pCAMBIA2306-BnaA09.MYB52-FLAG	This paper	N/A
pKSE401-BnaMYB52	This paper	N/A
<b>Software and algorithms</b>		
BWA	N/A	<a href="https://github.com/lh3/bwa">https://github.com/lh3/bwa</a>
FaST-LMM	N/A	<a href="https://github.com/fastlmm/FaST-LMM/">https://github.com/fastlmm/FaST-LMM/</a>
ImageJ	N/A	<a href="https://imagej.net/">https://imagej.net/</a>
Adobe Photoshop	N/A	<a href="https://www.adobe.com/products/photoshop.html">https://www.adobe.com/products/photoshop.html</a>
GraphPad Prism 10	N/A	<a href="https://www.graphpad.com/scientific-software/prism">https://www.graphpad.com/scientific-software/prism</a>

## EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

### Plant materials and growth conditions

The *B. napus* germplasm resources were supplied by the National Engineering Research Center of Rapeseed (China) and cultivated in Wuhan in October 2016 and 2017 at the experimental station of Huazhong Agricultural University. The field planting followed a randomized complete block design with three biological replications, and the field management followed local standard agriculture practice. *B. napus* cultivar Westar was used as the WT for *BnaMYB52*-CRISPR and *BnaA09.MYB52*-OE. To obtain homozygous lines, the T<sub>0</sub> and T<sub>1</sub> transgenic plants were grown under 22 h of light at 22°C and 2 h of dark at 18°C from 2021 to 2022. Homozygous mutants and WT plants were grown in the genetically modified field of Huazhong Agricultural University in Wuhan from October 2022 to May 2023 and from October 2023 to May 2024. The Arabidopsis transfer DNA insertion mutants *myb52* (SALK\_118938C and SALK\_138624C) and *ban* (SALK\_040250C and SALK\_122474C) were purchased from the AraShare (<https://www.arashare.cn/index/Product/index>). The Arabidopsis transfer DNA insertion lines *pmei14* (SALK\_206157C and SALK\_020742), *pmei6* (GABI\_790B12), and *sbt1.7* (GABI\_140B02) were kindly provided by A. Ding. Seeds of Arabidopsis Col-0 and *N. benthamiana* were stored in our lab. Arabidopsis seeds were surface sterilized with 75% ethanol for 3 min and 10% NaClO (v/v) for 10 min and then washed three times using sterile deionized water. The seeds were germinated for 7 days on 1/2 Murashige and Skoog solid medium (pH 5.8), transplanted to soil, and grown under long-day conditions (16 h light/8 h dark) at 22°C. The light was provided by white fluorescent tubes, with the intensity of approximately 120 μmol m<sup>-2</sup>.s<sup>-1</sup>. The tobacco seedlings were grown under the same conditions as those of the Arabidopsis seedlings.

## METHOD DETAILS

### Trait investigation

Seed oil and protein content were measured by a near-infrared reflectance spectroscopy (FOSS NIRSystems 5000). Total sugar content was determined according to Li et al. (2019).<sup>47</sup> Thousand-seed weight was defined as the average weight of 1000 well-developed dry seeds. The representative seeds (close to average value) were first dissected to separate seed coat from embryo under a light microscope, and then weighed with an electronic balance, respectively. SCC was measured as seed coat weight divided by seed weight. Additionally, plant height, main inflorescence length, branch number, seed number per pod, pod number, and pod length were measured, as previously described.<sup>48–50</sup>

### GWAS and candidate gene prioritization

Genome-wide association study (GWAS) for SCC was performed using our previously re-sequenced *B. napus* accessions.<sup>8,25</sup> The sequencing data was aligned to the Zhongshuang 11 v0 reference genome (<https://yanglab.hzau.edu.cn/>) with BWA software.<sup>51</sup> An association analysis was conducted by a mixed linear model (MLM) in the FaST-LMM Pathom package.<sup>52</sup> Genetic Type I error calculator (GEC) software was used to calculate the significance threshold.<sup>53</sup>

To prioritize the candidate genes in QTL regions, the POCKET algorithm based on multi-omics data and machine learning was performed as described by Tang et al. (2021).<sup>25</sup> Briefly, we first evaluated variation effects in intragenic regions, such as frameshift, and premature stop codon. Secondly, we evaluated the effects of gene expression levels on the target trait by transcriptome-GWAS (TWAS) and *cis*-eQTL. Thirdly, we identified haplotypes and evaluated the effects of haplotypes on the target trait based on genetic variants in the natural population. Fourthly, we searched related databases (Pfam, InterPro, Tair, Gene Ontology, and KEGG) and literature to predict the association of candidates and the phenotype. Lastly, we calculated the scores from the above four steps and identified candidate target genes for functional verification through mutants and overexpression lines.

### Plasmid construction and plant transformation

To knock out *BnaMYB52* using the CRISPR/Cas9 system,<sup>54</sup> two single-guide RNAs (sgRNAs) targeting four copies of *BnaMYB52* were designed by the CRISPR-P web (<http://crispr.hzau.edu.cn/CRISPR2/>) and they were integrated into a single binary vector (pKSE401), which contained a codon-optimized Cas9 genes driven by the double CaMV 35S promoter. The resulting vector was introduced into *Agrobacterium tumefaciens* strain GV3101, and then the positive strain was transformed into the *B. napus* variety Westar by *Agrobacterium tumefaciens*-mediated hypocotyl transformation.<sup>55</sup> PCR amplification and DNA sequencing were performed to identify mutations in *BnaMYB52*.

To generate *BnaA09.MYB52* overexpression plants, the full-length coding sequence without stop codon of *BnaA09.MYB52* was amplified with the primers *BnaA09.MYB52-pCAMBIA2306-F* and *BnaA09.MYB52-pCAMBIA2306-R* and cloned into the pCAMBIA2306 vector containing the 35S promoter. The resulting construct was introduced into *Agrobacterium tumefaciens* strain GV3101 and the *B. napus* transformation was performed as described by Zheng et al. (2020).<sup>55</sup> The overexpression transgenic lines were detected by PCR and RT-qPCR. All primers are listed in Table S5.

### Fatty acid extraction and GC-FID analysis

Fatty acid (FA) was extracted and measured as described with minor modifications.<sup>56,57</sup> Briefly, about 15 mg seeds were weighed and placed in the glass tubes. Four-milliliter extraction solution (95% methanol, 5% H<sub>2</sub>SO<sub>4</sub>, and 0.01% butylated hydroxyl toluene)

and 100 µL heptadecanoic acid (16.2 µmol/mL) were added to each sample. Heptadecanoic acid was used as an internal standard. The samples were incubated at 85°C for 2 h, and then cooled to room temperature. Finally, 3 mL hexane and deionized water were added to each sample, respectively. Fatty acid species of seeds were quantified using QP2010 Ultra GC-FID (Gas Chromatography-Flame Ionization Detector).<sup>58–60</sup>

#### Metabolite analysis by LC-MS/MS

Mature seeds were ground into powder in liquid nitrogen, and 0.1 g of dry powder was weighed. The metabolites were extracted overnight under 4°C with 1.0 mL 70% methanol (V/V) containing 0.1 mg/L acyclovir (an internal standard). After centrifugation at 10,000 g for 5 min, the supernatant was kept and filtered with a membrane (0.22 µm pore size). The metabolites of samples were measured by an LC-MS/MS system (SCIEX QTRAP 6500) according to previous studies.<sup>58–60</sup>

#### Determination of lignin content

Mature seeds were ground into powder, and lignin content was measured using the Lignin Content Assay Kits (AKSU010U, Beijing Boxbio Science& Technology), as described by Li et al.<sup>4</sup>

#### GUS staining and subcellular localization

A 2502-bp fragment of the *BnaA09.MYB52* promoter and a 1941-bp fragment of the *BnaA01.BAN* promoter were amplified from Westar's gDNA and cloned into the pCAMBIA2300-GUS vector, respectively, with ClonExpress II One Step Cloning Kit (C112, Vazyme). The positive plasmid was transferred into *Agrobacterium tumefaciens* GV3101 strain and then transformed into *Arabidopsis* (Col-0) by floral dipping method. Samples of the transgenic plants were stained overnight at 37°C according to the manufacturer's protocol (SL7160, Coolaber), and chlorophyll was removed in 75% ethanol prior to analysis. The GUS signals were observed under a SZX2 microscope (Olympus).

The full-length coding sequence (without stop codon) of *BnaA09.MYB52* was cloned into the pCAMBIA1305-GFP vector to generate CaMV35S::*BnaA09.MYB52*-GFP. The positive plasmid pCAMBIA1305-*BnaA09.MYB52*-GFP was transferred into *Agrobacterium tumefaciens* GV3101 strain and then injected into *N. benthamiana* leaves. The GFP signals were imaged under confocal microscopy (Leica SP8). Primers used for these constructs are listed in Table S5.

#### Cytological analysis

The seeds at the mature stage were collected and fixed in FAA solution (70% ethanol: formaldehyde: glacial acetic acid, 18: 1: 1). Embedding, slicing, and staining (Fast Green and Safranin) were performed as described by Xu et al. (2021).<sup>61</sup> Images were taken using a Nikon ECLIPSE Ci microscope. Seed coat thickness was calculated by ImageJ software. Four biological replicates were performed for each sample.

#### Yeast one-hybrid (Y1H) assay

The procedures were conducted according to the manufacturer's instructions (Clontech). Approximately 2 kb fragments of *BnaA09.TT8*, *BnaC09.TT8*, *BnaA10.F3'H*, *BnaA01.BAN*, *BnaA05.PMEI6*, *BnaC06.PMEI14*, and *BnaA09.SBT1.7* promoters were separately cloned into the pAbAi vector. The recombinant plasmid was linearized by *Bst*BI (R0519, New England Biolabs) and introduced into the yeast Y1H Gold strain to generate a reporter strain. The reporter strain was screened for optimal AbA (Aureobasidin A) concentration on SD-/Ura medium. *BnaA09.MYB52* full-length coding sequence was cloned into the pGADT7 to generate a prey vector *BnaA09.MYB52*-pGADT7. The *BnaA09.MYB52*-pGADT7 and empty vector pGADT7 were separately transformed into the reporter strain. Interactions were examined on selection plates (SD-/Leu+AbA). All primers are listed in Table S5.

#### Electrophoretic mobility shift assay (EMSA)

The full-length coding sequence of *BnaA09.MYB52* was amplified from the cDNA of Westar and cloned into the pGEX-6p-1 vector. The GST-*BnaA09.MYB52* fusion protein was expressed in *Escherichia coli* BL21 at 16°C for 16–20 h in the presence of 0.5 mM isopropyl β-D-thiogalactoside (IPTG). GST-*BnaA09.MYB52* protein was purified using GST.Bind Resin (70541, Merck) according to the manufacturer's protocol. *BnaA1.BAN*, *BnaA05.PMEI6*, *BnaC06.PMEI14*, and *BnaA09.SBT1.7* promoter fragments containing MBS motifs were synthesized with 5'-end labeled with FAM. Forward and reverse probes were heated at 98°C for 10 min and then cooled to room temperature to generate double-strand probes. The FAM-labeled and unlabeled probes were individually incubated with GST-*BnaA09.MYB52* protein at 23°C for 30 min in the EMSA/Gel -shift Binding Buffer (GS005, Beyotime). The reaction mixtures were loaded on a 6% (w/v) native polyacrylamide gel, and electrophoresis was performed at 80 V for 1 h in 0.5×TBE buffer (45 mM Tris-base, 45 mM boric acid, and 0.5 mM EDTA, pH 8.3) at 4°C in the dark. FAM-labeled DNA was detected using Amersham Typhoon (Cytiva). All primers used for EMSA are listed in Table S5.

#### Dual-luciferase assay

The full-length coding sequence of *BnaA09.MYB52* was cloned into pGreenII 62-SK to generate the effector. The promoter sequences of *BnaA01.BAN* (1941 bp), *BnaA03.BAN* (1,995 bp), *BnaC01.BAN* (1,968 bp), *BnaC04.BAN* (2,182 bp), *BnaA05.PMEI6* (2,052 bp), *BnaC06.PMEI14* (2,078 bp), and *BnaA09.SBT1.7* (1,991 bp) were separately inserted into pGreenII0800-LUC to generate

seven reporters. The effector and reporters were co-transformed into *Arabidopsis* protoplasts. The CaMV35S-driven *Renilla* luciferase (*REN*) gene was used as an internal control. Co-transformation of the reporters and empty vector pGreenII 62-SK was used as negative controls. LUC and REN activities were measured according to the manufacturer's instructions (E1910, Promega). All primers used for vector construction are listed in [Table S5](#).

## QUANTIFICATION AND STATISTICAL ANALYSIS

### Real-time PCR

Total RNA was extracted using Eastep Super Total RNA Extraction Kit (LS1040, Promega), and cDNA was synthesized using EasyScript First-Strand cDNA Synthesis SuperMix (AE301, TransGen Biotech). RT-qPCR was conducted to detect the relative expression levels of genes using ChamQ Universal SYBR qPCR Master Mix (Q711, Vazyme) on the CFX Connect Real-Time System (Bio-Rad). The *BnaActin7* and *Actin2* were used as the internal reference in *B. napus* and *A. thaliana*, respectively. Primers for RT-qPCR are included in [Table S5](#).

### Statistical analysis

The Excel statistical function T.TEST was used to calculate the significance level ( $P_{t\text{-test}}$ ) of differences between two samples. GraphPad Prism 10 was used for plotting and One-way ANOVA analysis. All data are presented as mean  $\pm$  SD of multiple independent experiments (at least three independent experiments). The statistical test for each individual experiment is indicated in the figure legends. Statistical significance was as follows: \* $p < 0.05$ , \*\* $p < 0.01$  and indicated in the figure legends.