

## ORIGINAL ARTICLE

# Leveraging Phenotypic Plasticity in Seed Oil Content for Climate-Adapted Breeding and Production

Lingju Zeng<sup>1,2</sup> | Xu Han<sup>1,2</sup> | Xiangjian Gou<sup>1,2</sup> | He Pei<sup>1,2</sup> | Yang Shao<sup>1,2</sup> | Yilan Cao<sup>1,2</sup> | Zhenwei Zhang<sup>1,2</sup> | Xianran Li<sup>3</sup>  | Jianming Yu<sup>4</sup>  | Jianbing Yan<sup>1,2</sup> | Liang Guo<sup>1,2,5</sup>  | Tingting Guo<sup>1,2</sup> 

<sup>1</sup>National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan, Hubei, China | <sup>2</sup>Hubei Hongshan Laboratory, Wuhan, Hubei, China | <sup>3</sup>USDA, Agricultural Research Service, Wheat Health, Genetics, and Quality Research Unit, Pullman, Washington, USA | <sup>4</sup>Department of Agronomy, Iowa State University, Ames, Iowa, USA | <sup>5</sup>Yazhouwan National Laboratory, Sanya, Hainan, China

**Correspondence:** Liang Guo ([guoliang@mail.hzau.edu.cn](mailto:guoliang@mail.hzau.edu.cn)) | Tingting Guo ([tguo@mail.hzau.edu.cn](mailto:tguo@mail.hzau.edu.cn))

**Received:** 11 March 2024 | **Accepted:** 13 January 2025

**Funding:** This study was supported by the National Natural Science Foundation of China.

**Keywords:** climate change | haplotype | phenotypic plasticity | plant breeding

## ABSTRACT

Phenotypic plasticity is the property of an organism to change in response to different environments. Understanding and leveraging crop phenotypic plasticity is crucial for mitigating threats caused by climate change. Here, we assessed phenotypic plasticity in multi-environment trials over 4 years, using 505 inbred lines from a *Brassica napus* genetic diversity panel. The variation in seed oil content (SOC) plasticity was primarily associated with three environmental indices: precipitation, diurnal temperature range, and ultraviolet B during the flowering or pod-filling stage, alongside five major plasticity genes. Leveraging this information with climate records, we developed a predictive model to estimate SOC for various planting dates in seven major production regions and validated our predictions in new environments. As climate change necessitates new breeding materials with improved genetics, we examined the genetic potentials of existing lines for enhanced SOC in future climates. Using projected environmental data and the identified major plasticity genes, we predicted SOC of genotypes across production regions. We also identified an optimal haplotype, a specific combination of alleles, for each production region to sustainably produce high SOC for future climates. This study offers insights and selection methods that contribute to mitigating the adverse effects of climate change on agriculture.

## 1 | Introduction

Climate change is increasingly threatening global food security (Lesk et al. 2022). Many regions have already witnessed significant yield losses caused by shifting temperatures and changing weather patterns. These losses are projected to worsen in the future (Tigchelaar et al. 2018; Fu et al. 2023). To ensure the stability of crop production systems, climate-aware strategies should be prioritized in breeding and agronomy (Hickey et al. 2019; Zhao et al. 2022). As a first step, a solid understanding of how climate change affects crop performance and of

the genetic mechanisms underlying crop responses is thus essential before developing efficient and effective strategies in breeding and agronomic practices. When combined with innovative technologies, these strategies are expected to play a central role in addressing the challenge of crop adaptability to future climates.

Phenotypic plasticity, or the property of a genotype to alter performance in response to environmental conditions, is a fundamental characteristic of plants, including crops (Pigliucci, Murren, and Schlichting 2006), that enables them to thrive in

\*Lingju Zeng and Xu Han contributed equally to this work.

diverse environments (Bonamour et al. 2019; Xue and Leibler 2018). To harness this mechanism to enhance crop performance in changing environments, numerous studies have investigated the molecular basis of phenotypic plasticity across various traits, species, and environments (Liu et al. 2021; Kusmec et al. 2017). Genes that underlie phenotypic plasticity have been identified, and their effects have been shown to be dynamic across different environments (Des Marais, Hernandez, and Juenger 2013; Li et al. 2018). Moreover, environmental factors that interact with genes to determine the phenotypes have been examined using algorithms such as Critical Environmental Regressor through Informed Search (CERIS) (Li et al. 2021; Li et al. 2022). Progress has also been made in determining the developmental patterns behind phenotypic plasticity (Mu et al. 2022). While these findings have expanded our understanding of phenotypic plasticity, it is crucial to use this knowledge to streamline strategies for helping crops adjust to current and future climate change.

*Brassica napus* (*B. napus*) is a valuable crop in which phenotypic plasticity can be explored in the context of climate change. Compared to other seed crops, *B. napus* is less tolerant to water deficit and high temperature, highlighting the significant roles of rainfall and temperature as climate change indicators for *B. napus* (Vernon 2006). Despite being predominantly grown in high rainfall areas with long growing seasons, *B. napus* has shown adaptability to low rainfall zones through the release of short season cultivars (Si and Walton 2004). This geographical and seasonal adaptation highlights *B. napus* as an exceptional model crop for studying phenotypic plasticity and applications to adaptation strategies. Seed oil content (SOC) is a crucial trait in *B. napus*, influenced by both genetic and environmental factors (Tan et al. 2022; Tan et al. 2011, Baud and Lepiniec 2010a; Gunasekera et al. 2006; Han et al. 2022). Candidate genes that control SOC in *B. napus* can be involved in transcript regulation, primary metabolism, flux control or material recycling within the cell (Baud and Lepiniec 2010b). Well-understood examples of genes that regulate SOC include *WRINKLED1* (*WRI1*) (Cernac and Benning 2004), *LEAFY COTYLEDON1* (*LEC1*) (Lotan et al. 1998), and *FUSCA3* (*FUS3*) (Luerßen et al. 1998). Biochemical pathway genes such as *ACC* genes (*ACC1* and *ACC2*) (Nerlich et al. 2007), *DIGALACTOSYL DIACYLGLYCEROL DEFICIENT1* (*DGD1*) (Dormann, Balbo, and Benning 1999b), and *PHOSPHATIDYL SERINE DECARBOXYLASE* genes (*PSD1*, *PSD2*, and *PSD3*) (Nerlich et al. 2007) have been identified and validated in *Arabidopsis*. These genes together offer potential to be utilized in breeding for high SOC. However, much of this research has not been carried out across various environmental conditions, leaving a gap in understanding the interactive effects of these genes with different environments and hindering breeding efforts under future climate variability.

Accurate prediction of trait values, such as yield and SOC, is important for making informed decisions in breeding programs and crop production management. With the increasing availability of genomic and phenotypic data in crops, modeling and statistical approaches have been applied to assess which combinations of genetic variants at either the gene or genome level can lead to the optimal trait values (Schulthess et al. 2022; Meuwissen, Hayes, and Goddard 2001; Crossa et al. 2014).

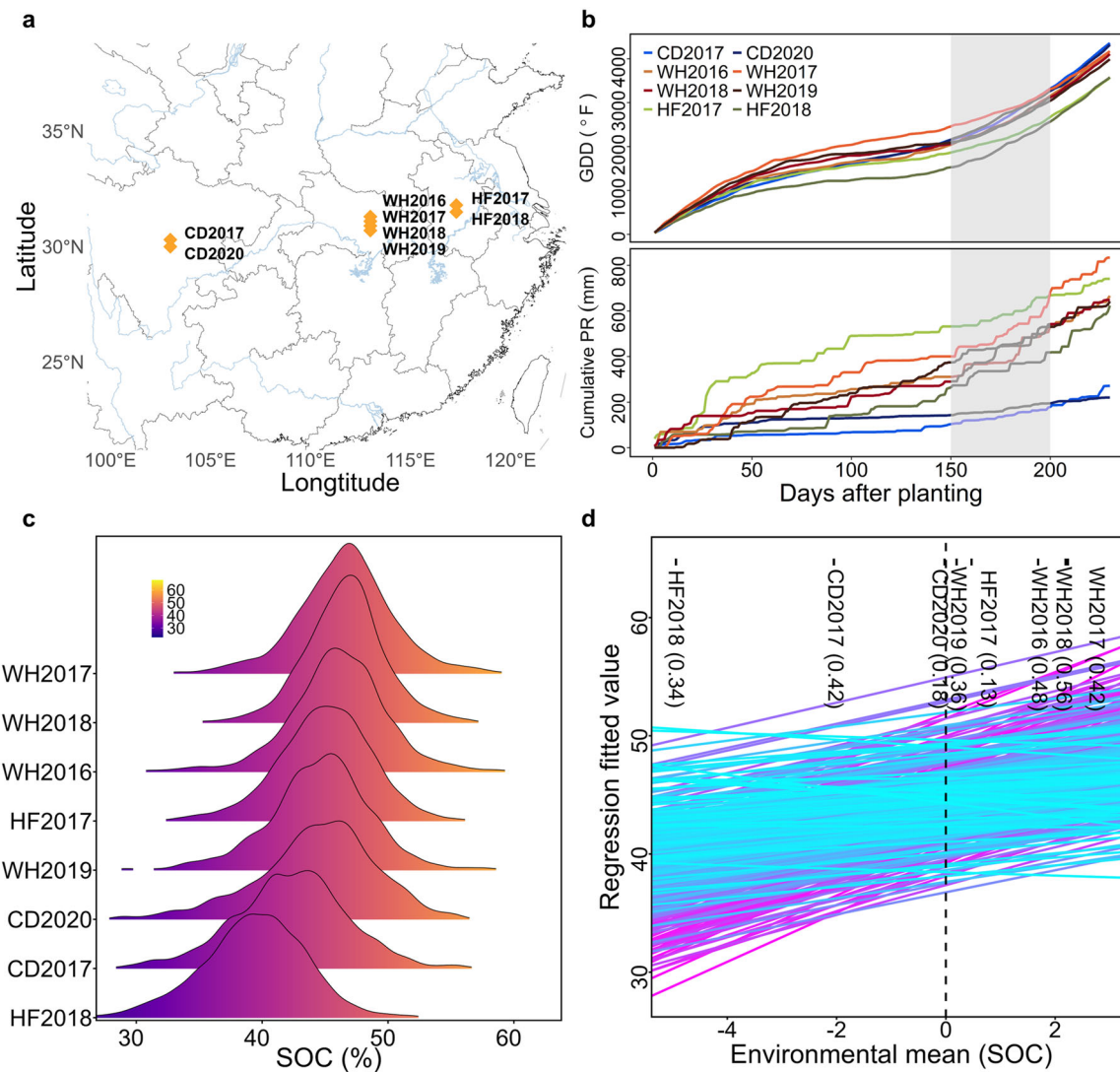
These approaches work well in specific environments or controlled environments with supporting field facilities, but accurately predicting trait values across a range of natural environmental conditions remains a challenge. Two types of research have been suggested to address this challenge: understanding the root causes of phenotypic plasticity of plants in response to environmental variability, and synthesizing this understanding into applied research to create a systematic prediction of traits across environments in crop systems (Messina et al. 2022; Guo et al. 2024).

Here, we aim to investigate how phenotypic plasticity can be leveraged to improve SOC in *B. napus* by addressing key questions: What components influence phenotypic plasticity in SOC in *B. napus*? How can these components be linked up to develop custom *B. napus* varieties for variable local environments affected by climate change? How can agronomic practices, like selecting planting dates, be tailored to each environmental condition to ensure that advanced *B. napus* varieties reach their full genetic potential and meet our needs amidst climate change challenges? By analyzing patterns of phenotypic plasticity in SOC, we hypothesize that identifying genes responsive to specific environmental factors during critical developmental stages can lead to operational optimization in *B. napus* breeding and production processes. Our proposed strategy involves leveraging phenotypic plasticity to determine optimal planting dates and identifying adaptive genotypes to maximize SOC under current and future environmental conditions. This study presents a systematic framework that integrates knowledge discovery in concepts and method designs to address the impact of environmental variability on agricultural production.

## 2 | Results

### 2.1 | Genotype-by-Environment Interaction for SOC

We conducted multi-environment trials with a diversity panel of 505 *B. napus* lines to examine the effect of genotype-by-environment interaction ( $G \times E$ ) on SOC. We carried out these trials at three different sites (Chengdu, CD; Wuhan, WH, and Hefei, HF), spanning a wide geographical range with diverse temperature and precipitation patterns (Figure 1a,b; Supporting Information S1: Tables 1 and 2). To account for potential year-to-year variations in temperature and precipitation, we conducted these trials for up to 4 years at each site. We assessed the contributions of genotype and environment to the observed variation in SOC in the multi-environment trials. It was noted that the average SOC value in this *B. napus* population varies from 39.44% to 46.62% across different environments (Figure 1c). Pearson's correlation coefficients between SOC values in one environment and those in another environment range from 0.27 to 0.68, indicating the moderate influence of the environment on SOC (Supporting Information S2: Figure 1a). Looking at the influence of genotype, correlation coefficients between SOC values in different inbred lines range from  $-0.87$  to  $0.99$  (Supporting Information S2: Figure 1b). An analysis of variance for SOC across all eight environments



**FIGURE 1** | Plasticity for seed oil content in a *Brassica napus* diversity panel across multi-environment trials. (a) Map showing the eight natural field environments at three sites over up to 4 years. (b) Cumulative temperature and precipitation profiles for the eight environments. Gray areas indicate the approximate flowering and pod-filling stages. (c) SOC distribution for each environment. Color key indicates SOC levels. (d) Reaction norms for seed oil content (SOC) in all inbred lines within the diversity panel. Environmental mean is calculated by centering the mean performance of all lines in individual environments. Environments closer to the average of the environmental mean exhibit lower SNP-based heritability (numbers in parentheses). GDD, growing degree days; PR, precipitation. Three tested sites are Chengdu (CD), Wuhan (WH), and Hefei (HF).

indicates that environment, genotype, and  $G \times E$  have similar contributions, explaining 23.75%, 25.26%, and 18.90% of the SOC variation, respectively (Supporting Information S1: Tables 3 and 4). Thus, genotype and environment both influence SOC in this population.

We then analyzed the phenotypic variation in SOC using a joint regression analysis, where observed phenotypes for individuals were modeled as linear responses across the environmental mean (i.e., population mean at each environment) (Finlay and Wilkinson 1963; Eberhart and Russell 1966; Yates and Cochran 1938). Although this model assumes a linear relationship, it was chosen for its high interpretability and practicability. For each genotype, we captured SOC across eight environments by a linear regression model with two parameters: the intercept, which represents the average response of a

given genotype to the eight environments, and the slope, which represents the plasticity of that genotype to different environments (Figure 1c). The resulting models show a noteworthy alignment with the SOC observations, as indicated by the goodness-of-fit ( $R^2 = 0.56$ , on average). We observed substantial variation in both the intercept (39.44% to 46.62%) and the slope ( $-0.59$  to  $2.36$ ) among different genotypes. To investigate whether these variations are influenced by population structure, we examined SOC variation within each subgroup, which has been classified in a previous study (Tang et al. 2021). The subgroups were visualized using principal component analysis, based on the genomic relationship matrix calculated from 8 503 071 single nucleotide polymorphisms (SNPs). The analysis of variance revealed that the subgroup significantly affected both the SOC intercept ( $p = 0.002$ ) and slope ( $p = 0.037$ ) at a threshold of 0.05. However, the relationship between these two

parameters was weak across all subgroups, with little correlation observed in individual subgroups ( $r=0.01-0.24$ ) and in the overall population ( $r=0.01$ ) (Supporting Information S2: Figure 2). This finding suggests that genotypes with higher or lower average responses to all eight environments (intercept) do not necessarily exhibit higher or lower plasticity to different environments (slope).

Change in genetic variation across environments is one of the forms of  $G \times E$  (Huang et al. 2020). To examine genetic variation of SOC across environments, we assessed the SNP-based heritability, which quantifies the proportion of phenotypic variation attributed to additive genetic variation generated by causal variants that are in linkage disequilibrium with the genotyped SNPs (Lee et al. 2011). SNP-based heritability varies among environments, ranging from 0.13 to 0.56, and tends to be lower in environments closer to the average of environmental means (Supporting Information S1: Table 5). This variability underscores the significant  $G \times E$  effects on SOC, reflecting the unique characteristics of individual environments influenced by environmental factors.

## 2.2 | Identification of Environmental Indices That Influence SOC Plasticity

To define the environmental factors influencing SOC plasticity, we examined the general trends of three key climate conditions during the growth season: temperature, precipitation, and solar radiation (Canvin 1965; Pritchard et al. 2000; Rahimi and Bahrani 2011). Among the three experimental sites, Hefei (HF) exhibits the lowest cumulative growing-degree day (GDD) values, while Chengdu (CD) has the least amount of cumulative precipitation (Figure 1b; Supporting Information S2: Figure 3). For solar radiation, Wuhan (WH) has higher levels of ultraviolet B (UV-B) radiation during the leaf stage than the other sites. However, this trend diminishes and eventually reverses as the plants progress to inflorescence emergence and later stages (Supporting Information S2: Figure 4). In addition, variations in environmental conditions between different years at the same site also contributed to SOC plasticity (Figure 1b; Supporting Information S2: Figures 3 and 4). We identified days of cold stress (below  $-3.0^{\circ}\text{C}$ ) as a significant variable, with a noticeable disparity between years, with 33 days in HF2018 and 17 days in HF2017, suggesting that low temperature should be considered when assessing SOC differences (Supporting Information S2: Figure 3). These climate conditions are likely to be pivotal factors influencing SOC plasticity, but further investigation is needed to determine the key phases that affect SOC performance more explicitly.

To identify the explicit environmental indices that contribute to SOC plasticity, we used the CERIS package (Li et al. 2021), a tool designed for analyzing phenotypic plasticity (Supporting Information S2: Figure 5). The environmental indices identified by CERIS characterize the environmental mean by combining critical factors and growth windows. Through the exploration of various combinations of environmental factors and growth windows, CERIS identified three indices:  $\text{DTR}_{183-192}$ ,  $\text{PR}_{166-195}$ , and  $\text{UVB}_{144-186}$ , showing strong correlation with environmental

means, thereby influencing SOC plasticity across eight different environments (Figure 2).  $\text{DTR}_{183-192}$  represents the average diurnal temperature range within the window of 183–192 days after planting. This window falls within the pod-filling stage, highlighting plant sensitivity to temperature changes during this stage.  $\text{PR}_{166-195}$  and  $\text{UVB}_{144-186}$  represent the average precipitation (PR) and average UV-B radiation (UVB) in the windows of 166–195 and 144–186 days after planting, respectively. These two windows coincide with the flowering and pod-filling stages. To quantify the influence of these environmental indices on SOC, we conducted separate regressions of the environmental mean for SOC with each factor, revealing a 0.78% decrease in SOC for every 1-unit increment in  $\text{DTR}_{183-192}$ , a 1.55% increase in  $\text{PR}_{166-195}$ , and a 1.90% decrease in  $\text{UVB}_{144-186}$  (Supporting Information S1: Figure 6). We also developed a multiple regression model to assess the relative importance of each factor while accounting for the effects of the others. The proportionate contributions are 33.17% for  $\text{DTR}_{183-192}$ , 23.76% for  $\text{PR}_{166-195}$ , and 44.88% for  $\text{UVB}_{144-186}$ . This model demonstrates goodness of fit, with an  $R^2$  of 0.96, a statistically significant  $p$ -value of  $8.95 \times 10^{-4}$ , and a standard deviation of the error of 0.477. These findings establish a quantitative link between environmental factors, growth stages, and SOC plasticity, thereby enhancing our understanding of the environmental basis of SOC plasticity.

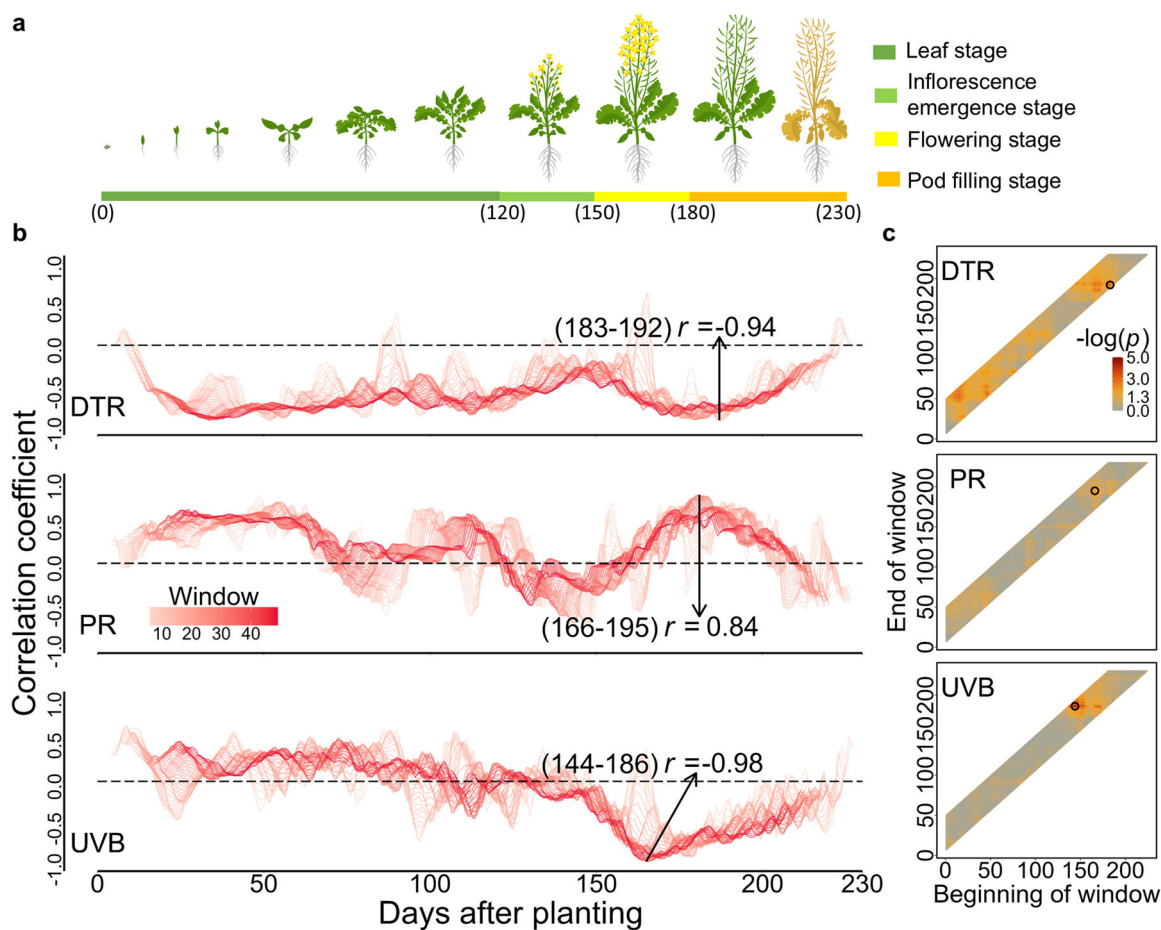
## 2.3 | Identifying Optimal Planting Dates and Empirical Validation

Understanding how *B. napus* lines interact with their environments throughout development drives our search for optimal planting dates that enable the plants to thrive in favorable climates and achieve higher SOC. To this end, we proposed a strategic approach to identify the most advantageous planting dates within practical planting windows. This approach employs a linear regression model that relates SOC to environmental conditions based on the planting date, allowing us to predict SOC for each potential planting date.

We conducted simulation experiments to identify more precise planting dates than the actual dates at the three tested sites (Figure 3a). To ensure a practical planting guide, while minimizing changes to other factors, such as social-economic conditions, we limited the dates to within 10 days before or after the actual planting dates. First, we predicted SOC for these dates using linear regression models of SOC on the identified environmental indices ( $\text{DTR}_{183-192}$ ,  $\text{PR}_{166-195}$  and  $\text{UVB}_{144-186}$ ) and a vernalization covariate (days of low temperature in leaf stage) (Figure 2a; Supporting Information S2: Figure 6). We then compared the measured SOC at actual planting dates to the predicted SOC values for early or late plantings. The results indicate that planting 10 days earlier could increase SOC by as much as 11.33% in HF2018, while planting 10 days later could decrease SOC by up to 12.16% in CD2020 (Figure 3b; Supporting Information S1: Table 6).

To assess the applicability of our approach for selecting optimal planting dates, we predicted SOC at four additional sites not included in the initial testing. We chose these sites from major

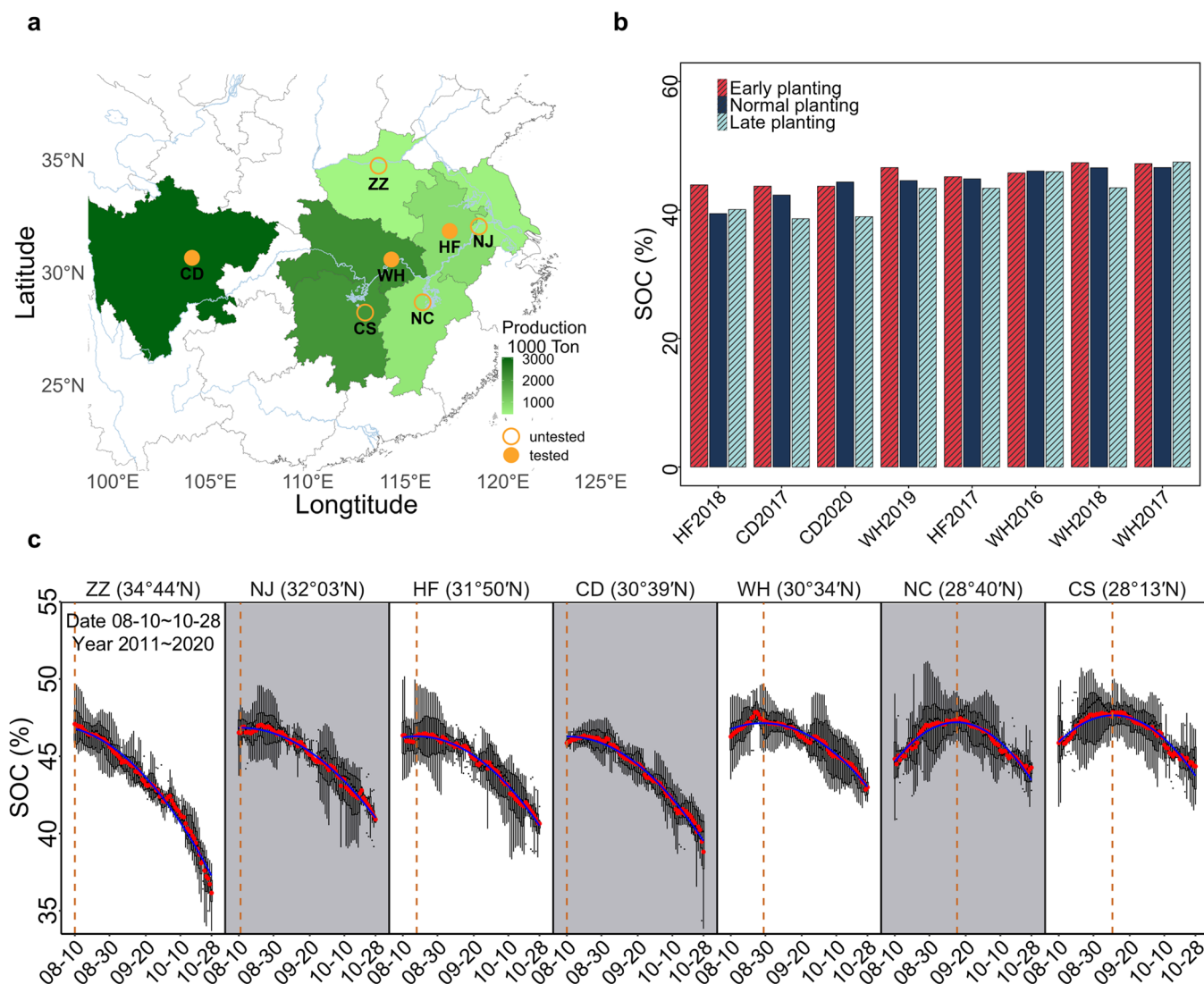




**FIGURE 2** | Three environmental indices determine seed oil content plasticity in *Brassica napus*. (a) Diagram of growth stages in *B. napus*. (b) Correlations between the environmental mean and DTR, PR and UVB values throughout the growing season. DTR<sub>183–192</sub>, PR<sub>166–195</sub> and UVB<sub>144–186</sub> show the most significant correlations with environmental mean. DTR, diurnal temperature range; PR, precipitation; UVB, ultraviolet B. Numbers in parentheses (on the plots) or in subscript for each environmental index indicate days after planting. (c) Significance of correlation between the environmental mean and environmental indices within the search window. [Color figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1111/pce.15408)]

oilseed-producing areas in China (Figure 3a; Supporting Information S1: Table 7). Using historical weather data from 2011 to 2020, we calculated the predicted SOC for each planting date between August 10th and October 28th (80 days) at all seven sites (Figure 3c). This time frame was chosen because these dates represent the majority of planting dates in the farms in the testing areas when fields are typically ready after harvesting the preceding crops. We then fitted a nonlinear function between predicted SOC and planting dates to determine the optimal planting date. The nonlinear function curves display a concave down pattern, with four locations (ZZ, NJ, HF, and CD) showing concave down with negative slopes indicating that early planting is optimal. Three locations show concave down with both positive and negative slopes, suggesting an optimal planting date at the vertex of the curve. For example, the planting dates at vertexes were found to be August 30th, September 12th, and September 17th for WH (30°34'N), CS (28°13'N), and NC (28°40'N), respectively. Besides, SOC varies among the seven planting sites at the optimal planting dates, with the highest SOC observed at CS (47.65%) and the lowest at CD (46.28%). The decrease in SOC is more pronounced at high altitude sites (e.g., ZZ) compared to low altitude sites (e.g., CS) when planting is not conducted at or near the optimal date.

We conducted an empirical validation of the above SOC predictions by planting materials on three dates in two new environments. From the diversity panel, we selected 50 lines, comprising two groups: 25 lines exhibiting greater plasticity and 25 lines demonstrating less plasticity. These lines were planted in Wuhan (114°21'E, 30°28'N) and Ezhou (114°42'E, 30°21'N) in 2022, with planting dates spaced 10 days apart, and harvested in 2023 (Figure 4a). The trials were named WH2023 and EZ2023. We compared the measured SOC with the predicted SOC for these 50 lines across six environments (site-planting date combinations) to assess the accuracy of the prediction model, which was used to identify optimal planting dates. The overall prediction accuracy, indicated by Pearson's correlation coefficient between measured and predicted SOC, is 0.63 for WH2023 and 0.66 for EZ2023 (Figure 4b,c). Both correlation coefficients are statistically significant, with  $p$ -values of  $6.8 \times 10^{-14}$  for WH2023 and  $2.2 \times 10^{-16}$  for EZ2023. For the three planting dates, the prediction accuracies are 0.56, 0.49, and 0.80 at WH2023, and 0.75, 0.59, and 0.56 at EZ2023, all of which are statistically significant. These results consistently show that the prediction model has a reasonable level of accuracy, indicating its usefulness for identifying optimal planting dates.



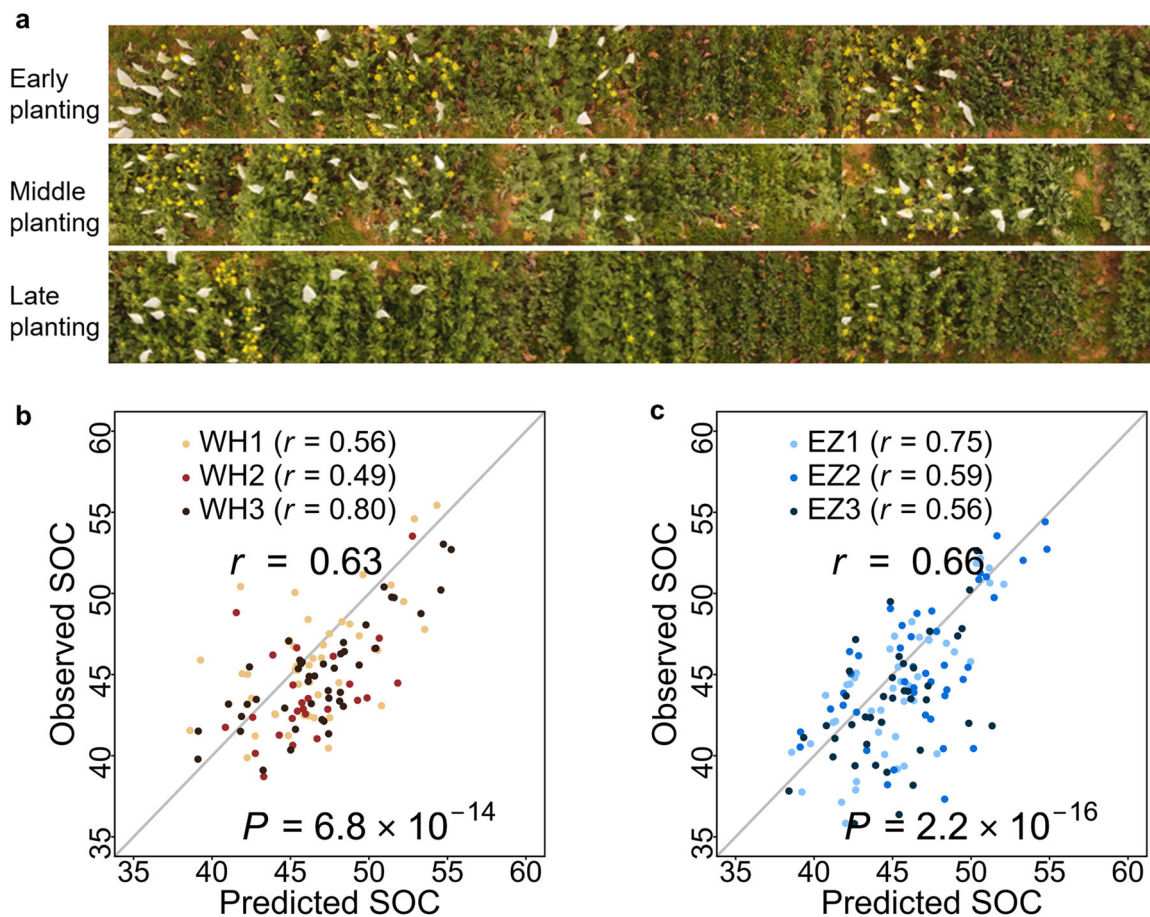
**FIGURE 3** | Identifying optimal planting dates for high seed oil content in predominant *Brassica napus* production areas. (a) Map showing average yields in major *B. napus* production areas, with three tested sites (filled circles) and four sites for prediction (open circles). (b) Observed or predicted SOC at different planting dates (10 days early, normal and 10 days late) in eight tested environments. (c) Predicted SOC with different planting dates in both tested and untested sites. The black dots represent the predicted SOC from 2011 through 2020, while the red dots indicate the average values. The blue lines represent the quadratic function between predicted SOC and planting dates from August 10 to October 28. The orange dashed lines denote the optimal dates from the quadratic fitting. Seven sites are Zhengzhou (ZZ), Nanjing (NJ), Hefei (HF), Chengdu (CD), Wuhan (WH), Nanchang (NC), and Changsha (CS). [Color figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]

## 2.4 | Identifying Genes That Underlie SOC Plasticity

Phenotypic plasticity is often conceptualized and measured in terms of reaction norms: functions that relate individual phenotypes to an environmental index (Martin et al. 2011). To reveal the genetic basis underlying differences in reaction norms, we conducted genome-wide association studies (GWAS) using two reaction norm parameters as phenotypic input, the intercept and the slope, obtained from joint regression analysis based on the environmental mean for SOC across all environments (Supporting Information S2: Figure 7). When using the intercept as the phenotype, we detected genomic regions containing the *PROBABLE METHYLTRANSFERASE PMT6* (*PMT6*) and *DIOXYGENASE FOR AUXIN OXIDATION1* (*DAO1*) genes, with *PMT6* previously reported to influence SOC variation in

*B. napus* (Tang et al. 2021). Using the slope as phenotype, we identified two genomic regions containing genes homologous to *Arabidopsis* (*Arabidopsis thaliana*) *MYB DOMAIN PROTEIN 106* (*MYB106*) and *DIGALACTOSYL DIACYLGLYCEROL DEFICIENT1* (*DGD1*). *MYB106* encodes a MIXTA-like transcription factor involved in cuticle development and lipid transport, while *DGD1* encodes a galactosyltransferase-like protein involved in lipid trafficking (Dormann, Balbo, and Benning 1999a; Oshima et al. 2013).

To uncover the genomic regions that respond to changes in the identified environmental indices (*DTR*<sub>183–192</sub>, *PR*<sub>166–195</sub> and *UVB*<sub>144–186</sub>), we calculated the slope from a regression between SOC and each index, representing the specific response of each genotype to each index. We used the resulting slopes as phenotypes for the GWAS (Figure 5a; Supporting Information S2:



**FIGURE 4** | Empirical validation of seed oil content predictions with three planting dates in two new environments. (a) Aerial photographs of fields from WH2023 when *B. napus* lines were planted at three different dates. (b–c) Scatterplots showing the correlation between the measured and predicted seed oil content (SOC) values in the environments of WH2023 (b) and EZ2023 (c). The numbers 1, 2 and 3 after location names (WH and EZ) indicate early, middle and late planting, respectively. Two sites are Wuhan (WH) and Ezhou (EZ). The  $p$  values of  $6.8 \times 10^{-14}$  for WH2023 and  $2.2 \times 10^{-16}$  for EZ2023 indicate a significant correlation between observed and predicted SOC. [Color figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1111/pe.15408)]

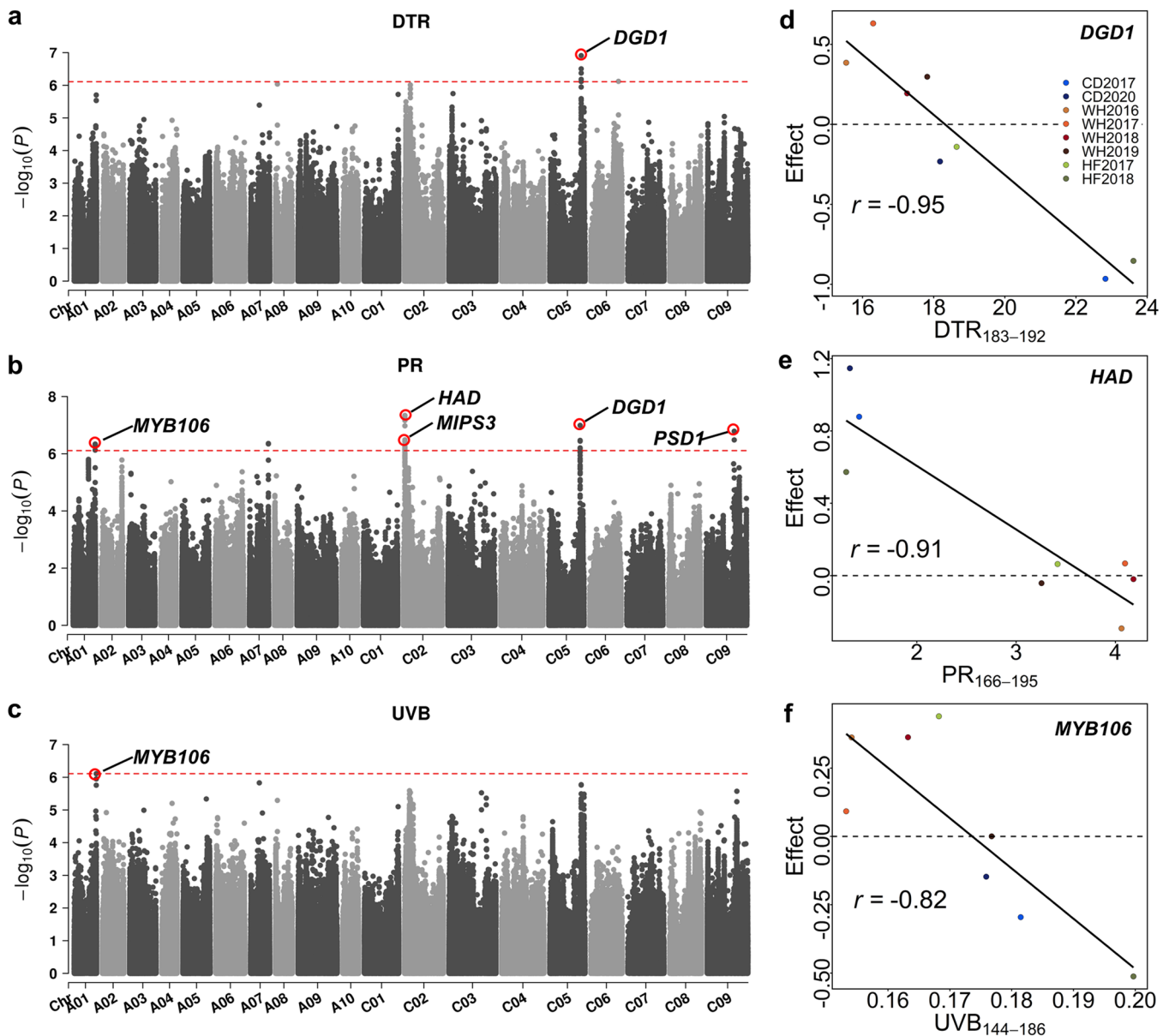
Figure 8). We identified *DGD1* for DTR<sub>183–192</sub> slopes and *MYB106* for UVB<sub>144–186</sub> slopes. We also mapped three additional genomic regions associated with PR<sub>166–195</sub> slopes. We prioritized the candidate genes within these regions with the *Arabidopsis* homologs 3-HYDROXYACYL-[ACYL-CARRIER-PROTEIN] DEHYDRATASE (*HAD*), MYO-INOSITOL-1-PHOSPHATE SYNTHASE3 (*MIPS3*), and PHOSPHATIDYLSERINE DECARBOXYLASE1 (*PSD1*). *HAD* is a 3-hydroxyacyl-ACP dehydratase involved in fatty acid biosynthesis. *MIPS3* catalyzes the rate-limiting step in myo-inositol biosynthesis. *PSD1* is a mitochondrion-localized protein involved in phosphatidylethanolamine biosynthesis (Figure 5a–c). The involvement of these genes in lipid metabolism in *Arabidopsis* is consistent with their putative role in SOC variation in *B. napus*.

To delineate potential polymorphisms in the above candidate genes, we aligned their genomic sequences from eight publicly available whole-genome assemblies of *B. napus* lines (Supporting Information S2: Figure 9; Supporting Information S1: Table 8). We discovered structural variations in *PMT6* and *DGD1*, in the form of transposon insertions in their promoters, defining two haplotypes among these eight *B. napus* inbred lines. *PMT6* in “Gangan,” “QuintaA” and “Shengli” contains an insertion of a *Mutator* retrotransposon; *Helitron* retrotransposon was inserted in the *DGD1* promoter

in “Shengli.” *PMT6* and *DGD1* may therefore affect SOC through changes in their expression levels. We also investigated potential coding variants affecting protein function. The loci most likely to cause amino acid changes are C106G, H32N and S108C in DAO1, MYB106 and HAD, respectively. These variants identified by analyzing sequences of candidate genes of GWAS peaks provide relevant evidence for follow-up studies of validation of functional polymorphisms.

In exploring gene-environment interactions, we hypothesized that the genetic effect would either increase or decrease in response to changes in the environmental index, aligning with the differential allelic effect model (Via 1993). To test this hypothesis, we examined the genetic effects of three significant markers, each linked with a candidate gene (*DGD1*, *MYB106*, and *HAD*), within individual environments. These markers were identified by GWAS as controlling SOC plasticity in response to DTR<sub>183–192</sub>, PR<sub>166–195</sub> or UVB<sub>144–186</sub>. Regression analysis was performed for each marker to illustrate the proposed relationship between genetic effect and environmental index (Figure 5d–f). All three regression models show a linear decrease in genetic effect as the environmental index increases. Additionally, various attributes of these gene-environment interactions were observed, including differential sensitivity





**FIGURE 5** | Identification and characterization of genetic loci associated with phenotypic plasticity in seed oil content in *Brassica napus*. (a–c) Manhattan plots showing the genomic regions associated with SOC plasticity in response to the identified environmental indices: DTR<sub>183–192</sub> (a), PR<sub>166–195</sub> (b), and UVB<sub>144–186</sub> (c). The red dashed lines represent the genome-wide significance threshold. The significant  $p$ -value threshold is set at  $7.77 \times 10^{-7}$ , determined by using the effective numbers of independent tests for multiple-testing adjustment. (d–f) Genetic effect from *DGD1*, *HAD* or *MYB106* on variation for the indicated environmental indices: DTR<sub>183–192</sub> (d), PR<sub>166–195</sub> (e), and UVB<sub>144–186</sub> (f). DTR, diurnal temperature range; PR, precipitation; UVB, ultraviolet B. [Color figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]

(DS; magnitude change), conditional neutrality (CN; effect limiting to specific environment), antagonistic pleiotropy (AP; sign change), and no G×E (no change) (Des Marais, Hernandez, and Juenger 2013). These attributes were determined by comparing genetic effects across two out of eight environments, with the percentage of each attribute calculated for all combinations. For instance, the interactions between *DGD1* and DTR<sub>183–192</sub> show 57.14% AP, 35.71% DS, and 7.14% no G×E (Figure 5d). *MYB106* and UVB<sub>144–186</sub> interaction is predominantly AP (42.86%), followed by 25.00% CN, 21.43% DS, and 10.71% no G×E (Figure 5f). In contrast, the *HAD* and PR<sub>166–195</sub> interaction is primarily CN (39.29%), followed by 32.14% AP, 21.43% DS, and 7.14% no G×E (Figure 5e). Moving forward, we used GWAS-identified genetic variants that are in linkage

disequilibrium with these gene candidates to predict their effects under future climates, as we prioritize the practical application of these genetic variants.

## 2.5 | Identifying Optimal Haplotypes to Enhance SOC in Changing Climates

The identified genetic variants associated with SOC plasticity in these *B. napus* lines guide our search for suitable haplotypes (combinations of genetic variants) that can maintain high SOC levels in future planting site environments. To this end, we trained a model for each genotype (at single and multi-locus levels of candidate genes) by regressing SOC on the identified



environmental index across eight tested environments. We then used this model to predict SOC for each genotype across various environmental index gradients in untested environments. Based on the predicted SOC, we were able to identify the optimal genotypes for each environment (site-year combination).

We used  $PR_{166-195}$  as the primary environmental index for estimating SOC plasticity, opting for it over the other two indices due to data limitation and a lack of noticeable changes in future climate scenarios (Supporting Information S2: Figure 10). At the single locus level for individual candidate genes (*MYB106*, *MIPS3*, *HAD*, *DGD1* and *PSD1*), we examined SOC plasticity in two groups of genotypes, each group containing a distinct pair of homozygous alleles, and observed changes in the ranking of gene-environment interactions (Supporting Information S2: Figure 11). For example, the genotype containing the more plastic alleles (CC) of *MYB106* outperforms the genotype containing the less plastic alleles (AA) in wet environments; however, this pattern reverses in dry environments. This single locus analysis indicates that SOC can be estimated and predicted for individual genotypes based on their genetic variants in candidate genes and the precipitation levels of their environments.

At the multiple locus level (combinations of alleles from five candidate genes), we examined SOC plasticity for 32 haplotypes, focusing on the 10 most abundant among the 505 *B. napus* lines. We observed varied plasticity levels among the 10 major haplotypes (Figure 6a,b). Haplotype 1 (with the allele combination AAGTA for *MYB106*, *MIPS3*, *HAD*, *DGD1* and *PSD1*), consisting of 329 lines and representing 65% of the entire population, displays the lowest SOC plasticity, as illustrated by the lowest slope of its regression with  $PR_{166-195}$ . It performs well in environments with lower precipitation but lags behind in environments with higher precipitation. By contrast, Haplotype 10 (allele combination CGAAA), composed of only five lines, presents the highest SOC plasticity. It significantly outperforms the other haplotypes, particularly in response to extremely high precipitation. This comparison suggests that most of the preserved germplasm has the potential to reach high SOC in relatively dry environments. However, we have a limited number of germplasms with rare haplotypes that are well-suited to wet environments.

To develop high SOC varieties for future climates, we initially examined the precipitation patterns in major oilseed-producing areas. We collected  $PR_{166-195}$  data from seven planting sites, including tested and untested sites, from 1960 to 2080. Based on this data, we identified three distinct categories of precipitation conditions. The first category comprises the sites CD and ZZ, with an average precipitation of 1.0 mm per day with no significant change over time. The second category includes the sites WH, HF and NJ, with average precipitation increasing from 2.0 to 3.9 mm per day throughout the years. The third category involves the sites CS and NC, experiencing an increase in precipitation from 5.0 to 7.7 mm per day on average (Figure 6c). We then determined the optimal haplotype with the highest SOC for the past and future  $PR_{166-195}$  level at each site. In drier regions (CD and ZZ), haplotype 5 consistently demonstrates the highest SOC in both past and future scenarios. For more humid areas (WH, HF and NJ), haplotype 2 was the best

performer in the past but is predicted to be replaced by haplotype 10 in future climates. In high-humidity regions (CS and NC), haplotype 10 is beneficial in both past and future scenarios (Figure 6d). Increasing the prevalence of haplotype 10 in future germplasms could be a viable approach to developing high SOC varieties in future climates.

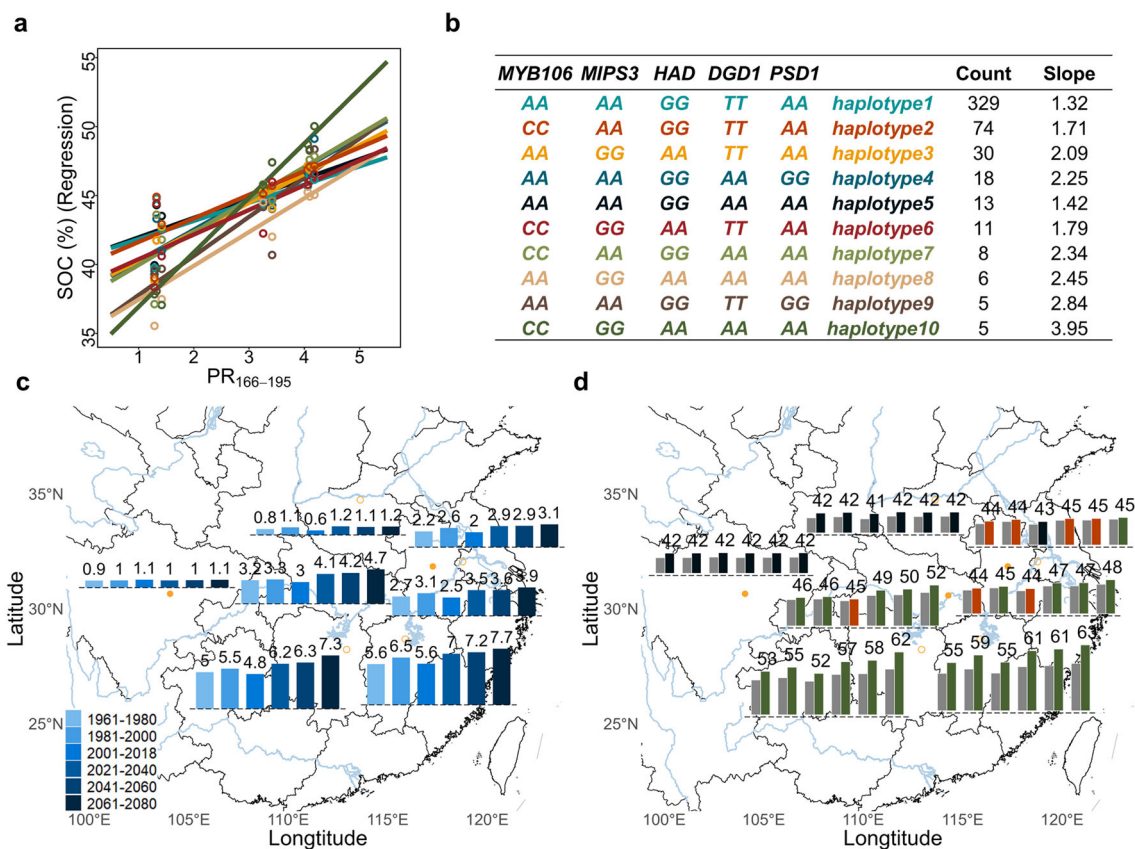
### 3 | Discussion

#### 3.1 | Analyzing Phenotypic Plasticity for Climate-Adapted Practices

Adapting agricultural systems to climate change involves implementing changes in agronomic practices or cultivar selection to ensure successful crop cultivation in altered environmental conditions (Sloat et al. 2020). To achieve these changes, it is important to understand the gene  $\times$  environment interaction and how it affects crop performance. Equally important is translating this mechanistic understanding into practical solutions for agronomic practices or cultivar selection (Bailey-Serres et al. 2019). This study focuses on analyzing phenotypic plasticity to identify environmental indices, genes and their modes of action. We then used this information to predict the most suitable planting dates for various environmental conditions and determine the optimal haplotypes for future climates. In the case of *B. napus*, a key trait is SOC. We identified three environmental indices ( $DTR_{183-192}$ ,  $PR_{166-195}$ , and  $UVB_{144-186}$ ) and five responsive genes (*MYB106*, *HAD*, *MIPS3*, *DGD1*, and *PSD1*) that influence SOC. Using this information, we predicted the SOC for various planting dates and validated these predictions in two new environments. By considering the plasticity of allele combinations from the five identified genes, we determined an optimal haplotype for each production site, which demonstrates the ability to adapt to future climate projections. The identification of production-limiting factors and the development of strategies for agricultural adaptation to climate change are essential for mitigating concerns regarding food security (Mourtzinis, Specht, and Conley 2019).

#### 3.2 | Three Environmental Indices Associated With SOC Plasticity

Unlike previous research, which focused on identifying a single environmental index to represent environmental mean in studies of phenotypic plasticity (Li et al. 2018; Guo et al. 2020), we uncovered three environmental indices as fundamental to plant growth and development. The inclusion of multiple factors is advantageous for capturing the complex and multifaceted nature of the genotype-environment-phenotype relationship, while also enhancing the accuracy of predicting traits in new environments (Millet et al. 2019). Although these factors are closely related, we did not consider the issue of multicollinearity in this study due to our emphasis on prediction, but it should be taken into account when determining the significance of these factors (Kim 2019). We demonstrated that environmental factors have various influences on the final phenotype at different developmental phases, particularly the growth windows of 183–192, 166–195 and 144–186 days after



**FIGURE 6** | Predicting optimal haplotypes across planting sites and time periods. (a) Reaction norms of haplotypes showing seed oil content (SOC) plasticity to the environmental index  $PR_{166-195}$ . (b) List of 10 major haplotypes with their allele combinations at the five plasticity genes. (c) Precipitation levels during the critical growth stage at seven planting sites from 1960 to 2080. (d) Optimal haplotype with the highest SOC from 1960 to 2080, compared to the mean of all haplotypes (gray bar) at seven planting sites during different time periods. Optimal haplotype bars are colored black, red, and dark green. [Color figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]

planting, which align with the flowering or pod-filling stages of *B. napus*. It would be ideal to identify specific environmental indices for each genotype (Sabir et al. 2023), but a more practical approach should use a harmonized index shared by all genotypes for prediction and application purposes.

### 3.3 | Identification of Slope Genes That Differ From Intercept Genes

The genetic basis of SOC variation in *B. napus* has been deciphered through genome- and transcriptome-wide association studies, leading to the identification of numerous genomic regions and/or genes associated with this trait (Tang et al. 2021). In our study, we identified the same set of significant genomic regions when using intercept as the phenotype. Notably, when we used slope as the phenotype and conducted GWAS, we discovered five genes homologous to the *Arabidopsis* genes *MYB106*, *MIPS3*, *HAD*, *DGD1*, and *PSD1* known to be involved in oil metabolism. Moreover, these candidate genes were demonstrated to interact with the identified environmental indices and exhibit dynamic effects in response to changes in these indices. The findings highlight the complex interplay between genes and the environment and offer new insights into the genetic basis of SOC variation from the perspective of phenotypic plasticity. However, further verification

of candidate genes is needed through rigorous follow-up experiments, such as using gene-edited lines to confirm their role in observed phenotypes. Future research will focus on exploring the molecular mechanisms and validating the functions of these genes in response to environmental changes.

### 3.4 | Determining Optimal Planting Dates and Haplotypes in Multiple Environments

Finding the most suitable planting dates can significantly increase growth, development and crop production (Baum, Archontoulis, and Licht 2019). Nevertheless, this undertaking poses difficulties due to various modulating factors (Sacks et al. 2010). While local agricultural facilities often provide recommendations based on historical records, these suggestions are not always reliable. To tackle this issue, our study used the reaction norm to quantify the phenotypic plasticity of each genotype and predicted trait performance based on weather conditions specific to each planting date and location, from which we proposed the optimal planting date.

It is important to note that our selection of suitable planting dates focuses on enhancing seed oil content. Future models should be developed to optimize yield and other economically

significant traits that are of great concern to farmers and consumers. Furthermore, the timing of planting can be adjusted according to factors such as actual field conditions, the maturity of the preceding crops, and nitrogen management. In China, it is customary to delay the planting of *B. napus* in major zones until the preceding crop has reached full maturity and been harvested. Moving forward, our method can be improved by integrating crop models that account for field conditions, preceding crops, and other variables, along with a physiological description of crop growth (Hammer et al. 2005). This integration will enhance our ability to explain and predict crop growth and development, while also simplifying the extensive parameter calibration work required by crop models, especially for large-scale and long-term simulations.

Grounded in an analysis of phenotypic plasticity and the interactions between genes and environmental indices, we generated a predictive model for the relationship between trait performance and these interactions. Using this model, we evaluated the genetic potential of different genotypes to adapt to future climates. In this preliminary study, a precipitation-related environmental index was used to estimate projected future climate conditions. However, other environmental factors, like radiation, soil moisture and nitrogen application, may emerge as critical in other scenarios, and we may be able to model nonlinear relationships among them (Scheres and Van Der Putten 2017; Hammer et al. 2010). While our study used projected future climates for trait performance prediction, it is important to acknowledge that the accuracy of these predictions may require improvement as more precise estimates of climate change become available. With the aid of powerful tools in artificial intelligence and biotechnologies, advanced models are currently being developed to enhance the precision of breeding and agronomy practices in response to a changing climate. By building upon the establishment of mechanistic links and advanced models, we can effectively address fundamental questions related to phenotypic plasticity and develop potential breeding and agronomic strategies.

## 4 | Materials and Methods

### 4.1 | Population and Phenotyping

A diversity panel of 505 *B. napus* accessions was grown in three different sites: Wuhan (WH; 114° 21' E, 30° 28' N) for 4 years (2015–2016, 2016–2017, 2017–2018 and 2018–2019), Hefei (HF; 117° 13' E, 31° 52' N) for 2 years (2016–2017 and 2017–2018) and Chengdu (CD; 104° 12' E, 30° 46' N) for 2 years (2016–2017 and 2019–2020). Additionally, a subset of 50 accessions, consisting of 25 high-plasticity lines and 25 low-plasticity lines, was grown in Wuhan and Ezhou (EZ; 114° 42' E, 30° 21' N) on three different planting dates in 2022. In Wuhan, the planting dates were September 27th, October 7th, and October 17th (WH1, WH2 and WH3); in Ezhou, the dates were October 8th, October 18th, and October 28th (EZ1, EZ2 and EZ3). The phenotypic data for WH2016, WH2017, WH2018, HF2017, HF2018 and CD2017 were obtained from a previous study (Tang et al. 2021). Mature open-pollinated seeds of the natural population were harvested and dried for SOC analysis with a Foss NIRSystems

5000 near-infrared reflectance spectroscope (Li et al. 2003). Six biological replicates per accession were used for analysis. Detailed phenotype and genotype data can be found in supplemental files (Supporting Information S1: Tables 9 and 10) or in the previous study (Tang et al. 2021).

### 4.2 | Analysis of Variance and Estimation of Variance Components

The phenotypic value for genotype  $i$  when tested in replicate  $k$  in the environment  $j$  was modeled as  $y_{ijk} = u + g_i + t_j + b_{k(j)} + (gt)_{ij} + e_{ijk}$ , where  $u$  is the population mean,  $g_i$  is the effect of genotype  $i$ ,  $t_j$  is the effect of environment  $j$ ,  $b_{k(j)}$  is the block effect associated with replicate  $k$  nested in the environment  $j$ ,  $(gt)_{ij}$  is the  $G \times E$  effect associated with genotype  $i$  and environment  $j$  and  $e_{ijk}$  is the error.

The analysis of variance was conducted using the R function “aov” that fits a model with the function “lm” for each stratum. The estimation of variance components was conducted with the R package “VCA” and the function “anovaVCA.”

### 4.3 | Population Structure Analysis and SNP-Based Heritability Estimation

A quality control analysis was conducted on the materials using Plink (Purcell et al. 2007). Initially, SNPs with a minor allele frequency (MAF) < 0.05 were removed. Subsequently, materials with SNP missing rates > 0.1 were excluded. Finally, after filtering the original SNPs (Tang et al. 2021), a final set of 8 503 071 SNPs was obtained, with a median MAF of 0.21. In addition, Plink was used to calculate the genomic relationship matrix (GRM) among the 505 accessions. Principal component analysis (PCA) was subsequently performed based on this matrix using GCTA, with the suggested parameters “--grm --pca 20 --out” (Yang et al. 2011). A prior investigation has examined the population structure, which encompasses five ecological types: semi-winter 1 (SW1), semi-winter 2 (SW2), semi-winter mixed (SWM), spring (SPR) and mixed (Tang et al. 2021). The resulting PCA plot segregated the accessions into subgroups that correspond to their respective ecological types.

The SNP-based heritability was calculated with SOC as the phenotype. This analysis was conducted for each environment using GCTA (Yang et al. 2011), with the recommended parameters “--bfile --make-grm --make-grm-alg 0 --out; --grm --pheno --reml --out.”

### 4.4 | Identifying the Environmental Index

Environmental data were retrieved from the websites of the National Oceanic and Atmospheric Administration (NOAA: <https://www.noaa.gov/weather>) and the Astronomical Applications Department of the US Naval Observatory (<https://www.usno.navy.mil/USNO/astronomical-applications>). Daily temperatures (°F) were converted to growing-degree days (GDDs) for *B. napus* with the formula:  $GDD = [(maximum$



temperature + minimum temperature)/2]–37.4 (Marshall and Squire 1996, Pullens et al. 2019). The daily diurnal temperature range (DTR) was calculated as  $T_{\max} - T_{\min}$ . Phenotypic data were collected from 505 *B. napus* inbred lines planted in eight environments (Supporting Information S1: Table 9), along with environmental data (Supporting Information S1: Table 11) covering the entire growth period in these eight environments. Three categories of environmental parameters were tested: temperature (GDD and DTR), moisture (precipitation [PR] and relative humidity [RH]) and light (clear sky photosynthetically active radiation [CPAR] and ultraviolet B [UVB]) by implementing the CERIS algorithm in R to identify environmental indices ([https://github.com/jmyu/CERIS\\_JGRA](https://github.com/jmyu/CERIS_JGRA)) (Li et al. 2018). The most relevant environmental index in the three categories of environmental parameters for SOC was separately chosen according to the highest correlation between environmental means and environmental index within the corresponding search window (Figure 2b; Supporting Information S2: Figure 5). The importance of the environmental index to SOC was estimated with a single regression model for each index and a multiple regression model, which incorporates multiple indices in one regression model. In the multiple regression model, the relative importance of each variable was calculated by using the R package “relaimpo” (Grömping 2007).

#### 4.5 | Phenotypic Prediction Across Varied Locations

After environmental indices were identified, they were leveraged to predict phenotypic outcomes in different environments (Guo et al. 2020; Li et al. 2022; Li et al. 2021). Here we identified three environmental indices. For each environmental index, we employed the same approach in constructing the model. Firstly, one model was constructed to estimate the magnitude of changes in phenotypic response to each environmental index at the population level by a general linear model (Supporting Information S1: Figure 6):

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

where  $y_i$  is the mean SOC for the population in the  $i$ th environment ( $i = 1, 2 \dots v$ ),  $x_i$  is the mean values for the environmental index in the  $i$ th environment,  $\beta_0$  is the intercept of the linear model,  $\epsilon_i$  is the deviation from regression and  $\beta_1$  is the slope estimate of the linear model, which indicates the expected trait value change per a one-unit change of the corresponding climatic variant (Li et al. 2022). In addition to considering environmental indices in CERIS, vernalization was also taken into account when studying the impact of early sowing on SOC. A nonlinear response model of SOC to the number of low-temperature days during the seedling stage was constructed. This model shows that approximately 8 weeks at 2°–12°C for vernalization result in a significant SOC (Supporting Information S2: Figure 6), aligning with findings from previous research (Filek et al. 2007).

To make prediction, environmental data were collected for each location based on the planting date. Three environmental indices specific to each environment, along with the

low-temperature days during the seedling stage, were extracted. These environmental values were input into individual models, and the averaged output values from the four models were computed to generate the predicted phenotypic value.

#### 4.6 | Phenotypic Prediction Across Varied Planting Dates

To predict the SOC levels when planting either 10 days earlier or later than the actual planting dates in the tested environments, we gathered environmental data for the corresponding period for prediction and followed the approach outlined above (Figure 3b; Supporting Information S1: Table 12). To forecast the trend of SOC variation based on planting dates in seven major *B. napus* producing regions in the middle and lower Yangtze River, environmental data for a span of 10 years (2011–2020) were collected (Supporting Information S1: Table 13). Daily SOC predictions were carried out from August 10th to October 28th, using the environmental data accumulated over the previous decade for prediction on each day (Figure 3c). Finally, a quadratic regression model was constructed to estimate the optimum planting date:

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \epsilon_i$$

where  $y_i$  is the predicted mean SOC for the population on the  $i$ th day ( $i = 1, 2 \dots v$ ) starting from August 10,  $x_i$  is the number of days since August 10,  $\beta_0$  is the intercept,  $\beta_1$  is the linear coefficient,  $\beta_2$  is the quadratic coefficient and  $\epsilon_i$  is the deviation from the regression. To determine the analytical optimum (i.e., the turning point) of the nonlinear regression, we calculated the critical point where the first derivative of the equation with respect to  $x_i$  equals zero. The turning point, corresponding to the day that maximizes the predicted SOC, is given by:

$$x_{opt} = -\frac{\beta_1}{2\beta_2}$$

This formula represents the day at which the rate of change of the SOC begins to reverse, indicating the optimum planting date.

#### 4.7 | Empirical Validation

After the completion of the initial analysis and model building, empirical validation experiments were carried out. Specifically, a subset of 50 inbred lines, consisting of 25 high-plasticity lines and 25 low-plasticity lines, was grown in Wuhan (114° 21' E, 30° 28' N) and Ezhou (114° 42' E, 30° 21' N) on three different planting dates in 2022 (Supporting Information S1: Table 4). These lines were selected based on their phenotypic plasticity values, specifically focusing on the extremes. The 25 high-plasticity lines were characterized by the steepest slopes, and the 25 low-plasticity lines by the least steep slopes (Figure 1d). For any environmental index, we employed the same approach in constructing the model. First, one model was constructed to

estimate the magnitude of changes in phenotypic response to each environmental index at the individual level by joint regression analysis (Guo et al. 2024) (Supporting Information S2: Figure 9; Supporting Information S1: Table 9):

$$Y_{ij} = u_i + \beta_i I_j + \delta_{ij}$$

where  $Y_{ij}$  is the line mean of the  $i$ th line in the  $j$ th environment ( $i = 1, 2 \dots v; j = 1, 2 \dots n$ );  $u_i$  is the mean of the  $i$ th line across all environments;  $\beta_i$  is the regression coefficient that measures the response of the  $i$ th line to environmental input;  $I_j$  is the value of the environmental index in the  $j$ th environment; and  $\delta_{ij}$  is the deviation from regression. After that, the other two models are also constructed using the same theoretical framework. Subsequently, environmental data were collected from NOAA and USNO (Supporting Information S1: Table 14) and separately used as input for these three models to predict the SOC of the 50 lines. The final predicted value for each line was the average output from these three models. Finally, the accuracy of the model was evaluated by examining the correlation between the predicted and observed values of each inbred line, measured as Pearson's correlation coefficients (Figure 4b,c).

#### 4.8 | Genome-Wide Association Study (GWAS) and Linkage Disequilibrium (LD)

Reaction norms connect the environments to which a particular genotype is exposed and the phenotypes produced by that genotype in these environments. From the reaction norm values, two reaction norm parameters (intercept and slope) were obtained for each trait using a joint regression analysis of the observed trait values across environments (Figure 1c). In addition, with the identified environmental indices, the slope was obtained for each index using a joint regression analysis of the observed trait values across environments (Supporting Information S2: Figure 9). Treating the estimates of intercept and slope as a derived trait, an established mixed-model GWAS (Zhang et al. 2010; Yu et al. 2006) implemented in GAPIT (v.3) R package (<http://zzlab.net/GAPIT>) (Wang and Zhang 2021) was used, separately for each trait, to identify genomic regions underlying the observed variation for slope and intercept across different genotypes (Figure 5a–c; Supporting Information S2: Figure 7).

The significance thresholds of association were calculated with Genetic Type I error calculator (GEC) software (Li et al. 2012), and the calculated significance threshold was  $7.77 \times 10^{-7}$ . To identify candidate genes, genes within the 100-kb upstream and downstream regions of significant SNPs were extracted, according to the linkage disequilibrium (LD) decay rate estimated by previous studies (Tang et al. 2021).

#### 4.9 | Identification of Candidate Genes Associated With Phenotypic Plasticity Through GWAS

To identify candidate genes associated with SOC plasticity, three strategies were employed. Firstly, we applied reliable

statistical methods and stringent thresholds to search for candidate genes. Significant SNPs were identified in GWAS using a mixed model (Yu et al. 2006), and genes within linkage disequilibrium block of these markers were considered as candidates, with gene models annotated in “ZS11” reference genome sequence (Sun et al. 2017). Secondly, bioinformatics analysis was conducted to explore potential gene functions. Protein sequences of the candidates from the gene models of ‘ZS11’ were analyzed using the gene database (<http://aralip.plantbiology.msu.edu/>) related to lipid metabolism in Arabidopsis, with many of these genes having been validated. Subsequently, seven candidate genes (Supporting Information S1: Table 15) were found to influence SOC: *PMT6* (BnaA05G0437100ZS), *DAO1* (BnaA09G0636200ZS), *MYB106* (BnaA01G0418000ZS), *HAD* (BnaC02G0039300ZS), *MIPS3* (BnaC02G0039400ZS), *DGD1* (BnaC05G0484900ZS) and *PSD1* (BnaC09G0366500ZS). Thirdly, genes were identified to colocalize with those in a previous study (Tang et al. 2021).

Pan-Genome Sequences Synteny Analysis and Potential Functional Loci Analysis were conducted to unravel potential functional polymorphisms. To investigate the variation in the promoter regions of these genes, the methodology described by Li was followed (Li et al. 2021). In detail, sequences surrounding the genes (*PMT6* and *DGD1*) implicated in SOC phenotypic plasticity were retrieved from the *B. napus* pan-genome information resource (BnPIR; <http://cbi.hzau.edu.cn/bnapus/>) for the inbred lines ZS11, Zheyu7, Gangan, Shengli, No2127, Westar, Quinta and Tapidor. The sequences of these two genes were also extracted from this website. The colinear sequences among these eight inbred lines were used as queries for a BLAST search against each other to identify structural variations. Transposon elements (TEs) present in the sequences were identified with a TE Library constructed by Extensive de-novo TE Annotator (EDTA) (Ou et al. 2019). In addition, sequence variation within the genes (*DAO1*, *MYB106* and *HAD*) was analyzed with polyphen-2 (<http://genetics.bwh.harvard.edu/pph2/>) and provean (<http://provean.jcvi.org/>), noting putative loss-of-function alleles for each gene as predicted by at least one of the software platforms (Supporting Information S1: Table 8).

To model the genetic effect dynamics along the environmental gradient, SNPs significantly associated with slope were tested for association with SOC within each environment using the mixed-model method in GAPIT. These separate genetic effects were then regressed on the environmental gradients to generate the fitted lines as the genetic effect continua, defined as varied genetic effects with different environmental inputs (Figure 5d–f) (Li et al. 2021).

#### 4.10 | Reaction Norms at the Single-Locus and Multi-Locus Combination Level

We constructed reaction norms to reflect the relationship between SOC and PR<sub>166–195</sub> at both single-locus and haplotype levels (Guo et al. 2020). After GWAS and candidate gene examination, the significant SNPs for PR<sub>166–195</sub> were used. Five SNPs were selected, with each SNP showing the most

significant association with SOC plasticity and closely linked to each candidate gene (Figure 5b).

The entire population was divided into two groups based on their homozygous genotype at each selected SNP, resulting in the AA and BB groups. In an individual environment, the genotypic value for AA (or BB) was the average phenotypic value across all individuals carrying the AA (or BB) genotype at the locus. Two lines were drawn in the reaction norm graph, one for each group (Supporting Information S2: Figure 12). Linear regression was applied to show the relationship between the genotypic value and the environment and to show the fitted genotypic values across all environments (Supporting Information S2: Figure 12).

The five SNPs result in  $2^5 = 32$  possible haplotypes, with homozygous genotypes at each SNP. Similar to the calculation of genotypic values for a single locus described above, the genotypic value for each haplotype was the average phenotypic value for all inbred lines harboring each haplotype. There were 10 main haplotypes considered here for each environment (haplotypes represented by fewer than five inbred lines were not considered). Linear regression was applied to show the relationship between genotypic value and the environment and to show the fitted genotypic values across all environments (Guo et al. 2020) (Figure 6a,b).

#### 4.11 | Obtaining Climate Data and Phenotypic Predictions in the Past and Future

To predict the adaptation potential of haplotype materials for past and future planting, meteorological data were acquired with the R package “raster” and the function “getData” (Supporting Information S1: Table 16). The precipitation and temperature data from 1961 to 2018 were retrieved from the historical monthly weather data section in WorldClim (<https://worldclim.org/data/monthlywth.html>). The predicted precipitation and temperature data from 2021 to 2080 were retrieved from the Future climate data section in WorldClim (<https://worldclim.org/data/cmip6/cmip6climate.html>). A spatial resolution of 10 min was selected, with the data aggregated to align with the scales of NOAA and USNO data.

September 20th was set as the simulated planting date for future predictions. The 166–195 interval for the environmental index PR was chosen, corresponding to March 5th to April 3rd. Precipitation data were extracted from this interval. Similarly, the 183–192 interval for the environmental index DTR was chosen, corresponding to March 22nd to March 31st. From this interval, minimum ( $t_n$ ) and maximum ( $t_x$ ) temperature data were extracted, resulting in DTR being defined as:  $DTR = t_x - t_n$ .

However, original models were constructed using data from NOAA and USNO. WorldClim data are on another scale, which was converted to the scales of NOAA and USNO. Subsequently, precipitation data in March from both databases between 1999 and 2018 in the seven locations were collected separately and fitted using a general linear model, revealing a correlation of 0.97 between the two datasets. A model,  $y = 1.107x + 0.052$ ,

was established, where  $y$  represents the precipitation data from NOAA and USNO, while  $x$  represents the precipitation data from WorldClim. Using this equation, WorldClim data were calibrated to match the scales of NOAA and USNO for subsequent predictions.

The means for climate data spanning 20-year intervals in historical records were computed, resulting in three periods: 1961–1980, 1981–2000 and 2001–2018. Incorporating the three future periods 2021–2040, 2041–2060 and 2061–2080 resulted in six data periods for each location (Figure 6c). To assess the performance of each haplotype under varying periods, the environmental data were incorporated into the established multi-locus model. Using a colored bar chart representation on a map, the haplotype with the highest SOC was displayed alongside its corresponding SOC value. Additionally, the gray bars in the bar chart depict the population mean (Figure 6d).

#### Conflicts of Interest

The authors declare no conflicts of interest.

#### Data Availability Statement

All data are included in the manuscript and its supplementary file. Data for phenotype are provided in Supporting Information S1: Tables 9–10. Data for weather information are provided in Supporting Information S1: Tables S11–14 and 16.

#### References

- Bailey-Serres, J., J. E. Parker, E. A. Ainsworth, G. E. D. Oldroyd, and J. I. Schroeder. 2019. “Genetic Strategies for Improving Crop Yields.” *Nature* 575: 109–118.
- Baud, S., and L. Lepiniec. 2010a. “Physiological and Developmental Regulation of Seed Oil Production.” *Progress in Lipid Research* 49: 235–249.
- Baud, S., and L. Lepiniec. 2010b. “Physiological and Developmental Regulation of Seed Oil Production.” *Progress in Lipid Research* 49: 235–249.
- Baum, M. E., S. V. Archontoulis, and M. A. Licht. 2019. “Planting Date, Hybrid Maturity, and Weather Effects on Maize Yield and Crop Stage.” *Agronomy Journal* 111: 303–313.
- Bonamour, S., L. M. Chevin, A. Charmantier, and C. Teplitsky. 2019. “Phenotypic Plasticity in Response to Climate Change: The Importance of Cue Variation.” *Philosophical Transactions of the Royal Society, B: Biological Sciences* 374: 20180178.
- Canvin, D. T. 1965. “The Effect of Temperature on the Oil Content and Fatty Acid Composition of the Oils From Several Oil Seed Crops.” *Canadian Journal of Botany* 43: 63–69.
- Cernac, A., and C. Benning. 2004. “WRINKLED1 Encodes an AP2/EREB Domain Protein Involved in the Control of Storage Compound Biosynthesis in Arabidopsis.” *Plant Journal* 40: 575–585.
- Crossa, J., P. Pérez, J. Hickey, et al. 2014. “Genomic Prediction in CIMMYT Maize and Wheat Breeding Programs.” *Heredity* 112: 48–60.
- Des Marais, D. L., K. M. Hernandez, and T. E. Juenger. 2013. “Genotype-by-Environment Interaction and Plasticity: Exploring Genomic Responses of Plants to the Abiotic Environment.” *Annual Review of Ecology, Evolution, and Systematics* 44: 5–29.
- Dormann, P., I. Balbo, and C. Benning. 1999a. “Arabidopsis Galactolipid Biosynthesis and Lipid Trafficking Mediated by DGD1.” *Science* 284: 2181–2184.



- Dormann, P., I. Balbo, and C. Benning. 1999b. "Arabidopsis Galactolipid Biosynthesis and Lipid Trafficking Mediated by DGD1." *Science* 284: 2181–2184.
- Eberhart, S. A., and W. A. Russell. 1966. "Stability Parameters for Comparing Varieties 1." *Crop Science* 6: 36–40.
- Filek, M., J. Biesaga-Kościelniak, I. Macháčková, and J. Krekule. 2007. "Generative Development of Winter Rape (*Brassica Napus* L.)—The Role of Vernalization." *International Journal of Plant Developmental Biology* 1: 57–63.
- Finlay, K., and G. Wilkinson. 1963. "The Analysis of Adaptation in a Plant-Breeding Programme." *Australian Journal of Agricultural Research* 14: 742–754.
- Fu, J., Y. Jian, X. Wang, et al. 2023. "Extreme Rainfall Reduces One-Twelfth of China's Rice Yield Over the Last Two Decades." *Nature Food* 4: 416–426.
- Grömping, U. 2007. "Relative Importance for Linear Regression in R: The Package Relaimpo." *Journal of Statistical Software* 17: 1–27.
- Gunasekera, C. P., L. D. Martin, K. H. M. Siddique, and G. H. Walton. 2006. "Genotype by Environment Interactions of Indian Mustard (*Brassica juncea* L.) and Canola (*B. napus* L.) in Mediterranean-Type Environments." *European Journal of Agronomy* 25: 1–12.
- Guo, T., Q. Mu, J. Wang, et al. 2020. "Dynamic Effects of Interacting Genes Underlying Rice Flowering-Time Phenotypic Plasticity and Global Adaptation." *Genome Research* 30: 673–683.
- Guo, T., J. Wei, X. Li, and J. Yu. 2024. "Environmental Context of Phenotypic Plasticity in Flowering Time in Sorghum and Rice." *Journal of Experimental Botany* 75: 1004–1015.
- Hammer, G. L., S. Chapman, E. Van Oosterom, and D. W. Podlich. 2005. "Trait Physiology and Crop Modelling as a Framework to Link Phenotypic Complexity to Underlying Genetic Systems." *Australian Journal of Agricultural Research* 56: 947–960.
- Hammer, G. L., E. Van Oosterom, G. Mclean, et al. 2010. "Adapting APSIM to Model the Physiology and Genetics of Complex Adaptive Traits in Field Crops." *Journal of Experimental Botany* 61: 2185–2202.
- Han, X., Q. Tang, L. Xu, et al. 2022. "Genome-Wide Detection of Genotype Environment Interactions for Flowering Time in *Brassica Napus*." *Frontiers in Plant Science* 13: 1065766.
- Hickey, L. T., A. N. Hafeez, H. Robinson, et al. 2019. "Breeding Crops to Feed 10 Billion." *Nature Biotechnology* 37: 744–754.
- Huang, W., M. A. Carbone, R. F. Lyman, R. R. H. Anholt, and T. F. C. Mackay. 2020. "Genotype by Environment Interaction for Gene Expression in *Drosophila Melanogaster*." *Nature Communications* 11: 5451.
- Kim, J. H. 2019. "Multicollinearity and Misleading Statistical Results." *Korean Journal of Anesthesiology* 72: 558–569.
- Kusmec, A., S. Srinivasan, D. Nettleton, and P. S. Schnable. 2017. "Distinct Genetic Architectures for Phenotype Means and Plasticities in *Zea Mays*." *Nature Plants* 3: 715–723.
- Lee, S. H., N. R. Wray, M. E. Goddard, and P. M. Visscher. 2011. "Estimating Missing Heritability for Disease From Genome-Wide Association Studies." *American Journal of Human Genetics* 88: 294–305.
- Lesk, C., W. Anderson, A. Rigden, et al. 2022. "Compound Heat and Moisture Extreme Impacts on Global Crop Yields Under Climate Change." *Nature Reviews Earth & Environment* 3: 872–889.
- Li, G., S. Xiuli, J. Liang, et al. 2003. "Establishment of Math Models of NIRS Analysis for Oil and Protein Contents in Seed of *Brassica Napus*." *Zhongguo Nong ye ke xue= Zhongguo Nongye Kexue* 36: 1609–1613.
- Li, M.-X., J. M. Y. Yeung, S. S. Cherny, and P. C. Sham. 2012. "Evaluating the Effective Numbers of Independent Tests and Significant P-Value Thresholds in Commercial Genotyping Arrays and Public Imputation Reference Datasets." *Human Genetics* 131: 747–756.
- Li, X., T. Guo, G. Bai, et al. 2022. "Genetics-Inspired Data-Driven Approaches Explain and Predict Crop Performance Fluctuations Attributed to Changing Climatic Conditions." *Molecular Plant* 15: 203–206.
- Li, X., T. Guo, Q. Mu, X. Li, and J. Yu. 2018. "Genomic and Environmental Determinants and Their Interplay Underlying Phenotypic Plasticity." *Proceedings of the National Academy of Sciences* 115: 6679–6684.
- Li, X., T. Guo, J. Wang, et al. 2021. "An Integrated Framework Reinstating the Environmental Dimension for GWAS and Genomic Selection in Crops." *Molecular Plant* 14: 874–887.
- Liu, N., Y. Du, M. L. Warburton, Y. Xiao, and J. Yan. 2021. "Phenotypic Plasticity Contributes to Maize Adaptation and Heterosis." *Molecular Biology and Evolution* 38, no. 4: 1262–1275.
- Lotan, T., M. Ohto, K. M. Yee, et al. 1998. "Arabidopsis Leafy COTYLEDON1 Is Sufficient to Induce Embryo Development in Vegetative Cells." *Cell* 93: 1195–1205.
- Luerßen, H., V. Kirik, P. Herrmann, and S. Miséra. 1998. "FUSCA3 Encodes a Protein With a Conserved VP1/ABI3-like B3 Domain Which Is of Functional Importance for the Regulation of Seed Maturation in *Arabidopsis thaliana*." *Plant Journal* 15: 755–764.
- Marshall, B., and G. R. Squire. 1996. "Non-Linearity in Rate-Temperature Relations of Germination in Oilseed Rape." *Journal of Experimental Botany* 47: 1369–1375.
- Martin, J. G. A., D. H. Nussey, A. J. Wilson, and D. Réale. 2011. "Measuring Individual Differences in Reaction Norms in Field and Experimental Studies: A Power Analysis of Random Regression Models." *Methods in Ecology and Evolution* 2: 362–374.
- Messina, C., I. A. Ciampitti, D. Berning, D. Bubeck, G. Hammer, and M. Cooper. 2022. "Sustained Improvement in Tolerance to Water Deficit Accompanies Maize Yield Increase in Temperate Environments." *Crop Science* 62: 2138–2150.
- Meuwissen, T. H. E., B. J. Hayes, and M. E. Goddard. 2001. "Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps." *Genetics* 157: 1819–1829.
- Millet, E. J., W. Kruijjer, A. Coupel-Ledru, et al. 2019. "Genomic Prediction of Maize Yield Across European Environmental Conditions." *Nature Genetics* 51: 952–956.
- Mourtzinis, S., J. E. Specht, and S. P. Conley. 2019. "Defining Optimal Soybean Sowing Dates Across the US." *Scientific Reports* 9: 2800.
- Mu, Q., T. Guo, X. Li, and J. Yu. 2022. "Phenotypic Plasticity in Plant Height Shaped by Interaction Between Genetic Loci and Diurnal Temperature Range." *New Phytologist* 233: 1768–1779.
- Nerlich, A., M. Von Orlow, D. Rontein, A. D. Hanson, and P. Dörmann. 2007. "Deficiency in Phosphatidylserine Decarboxylase Activity in the psd1 psd2 psd3 Triple Mutant of *Arabidopsis* Affects Phosphatidylethanolamine Accumulation in Mitochondria." *Plant Physiology* 144: 904–914.
- Oshima, Y., M. Shikata, T. Koyama, N. Ohtsubo, N. Mitsuda, and M. Ohme-Takagi. 2013. "Mixta-Like Transcription Factors and Wax INDUCER1/SHINE1 Coordinately Regulate Cuticle Development in *Arabidopsis* and *Torenia Fournieri*." *Plant Cell* 25: 1609–1624.
- Ou, S., W. Su, Y. Liao, et al. 2019. "Benchmarking Transposable Element Annotation Methods for Creation of a Streamlined, Comprehensive Pipeline." *Genome Biology* 20: 275.
- Pigliucci, M., C. J. Murren, and C. D. Schlichting. 2006. "Phenotypic Plasticity and Evolution by Genetic Assimilation." *Journal of Experimental Biology* 209: 2362–2367.
- Pritchard, F. M., H. A. Eagles, R. M. Norton, P. A. Salisbury, and M. Nicolas. 2000. "Environmental Effects on Seed Composition of

- Victorian Canola." *Australian Journal of Experimental Agriculture* 40: 679–685.
- Pullens, J. W. M., B. Sharif, M. Trnka, J. Balek, M. A. Semenov, and J. E. Olesen. 2019. "Risk Factors for European Winter Oilseed Rape Production Under Climate Change." *Agricultural and Forest Meteorology* 272–273: 30–39.
- Purcell, S., B. Neale, K. Todd-Brown, et al. 2007. "Plink: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses." *American Journal of Human Genetics* 81: 559–575.
- Rahimi, M., and A. Bahrani. 2011. "Effect of Gamma Irradiation on Qualitative and Quantitative Characteristics of Canola (*Brassica napus* L.)." *Middle-East Journal of Scientific Research* 8: 519–525.
- Sabir, K., T. Rose, B. Wittkop, et al. 2023. "Stage-Specific Genotype-by-Environment Interactions Determine Yield Components in Wheat." *Nature Plants* 9: 1688–1696.
- Sacks, W. J., D. Deryng, J. A. Foley, and N. Ramankutty. 2010. "Crop Planting Dates: An Analysis of Global Patterns." *Global Ecology and Biogeography* 19: 607–620.
- Scheres, B., and W. H. Van Der Putten. 2017. "The Plant Perceptron Connects Environment to Development." *Nature* 543: 337–345.
- Schulthess, A. W., S. M. Kale, F. Liu, et al. 2022. "Genomics-Informed Prebreeding Unlocks the Diversity in Genebanks for Wheat Improvement." *Nature Genetics* 54: 1544–1552.
- Si, P., and G. H. Walton. 2004. "Determinants of Oil Concentration and Seed Yield in Canola and Indian Mustard in the Lower Rainfall Areas of Western Australia." *Australian Journal of Agricultural Research* 55: 367–377.
- Sloat, L. L., S. J. Davis, J. S. Gerber, et al. 2020. "Climate Adaptation by Crop Migration." *Nature Communications* 11: 1243.
- Sun, F., G. Fan, Q. Hu, et al. 2017. "The High-Quality Genome of *Brassica napus* Cultivar 'Zs 11' Reveals the Introgression History in Semi-Winter Morphotype." *Plant Journal* 92: 452–468.
- Tan, H., X. Yang, F. Zhang, et al. 2011. "Enhanced Seed Oil Production in Canola by Conditional Expression of *Brassica napus* Leafy COTYLEDON1 and LEC1-LIKE in Developing Seeds." *Plant Physiology* 156: 1577–1588.
- Tan, Z., Y. Peng, Y. Xiong, et al. 2022. "Comprehensive Transcriptional Variability Analysis Reveals Gene Networks Regulating Seed Oil Content of *Brassica napus*." *Genome Biology* 23: 233.
- Tang, S., H. Zhao, S. Lu, 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*." *Molecular Plant* 14: 470–487.
- Tigchelaar, M., D. S. Battisti, R. L. Naylor, and D. K. Ray. 2018. "Future Warming Increases Probability of Globally Synchronized Maize Production Shocks." *Proceedings of the National Academy of Sciences* 115: 6644–6649.
- Vernon, L. 2006. *Potential Impacts of Climate Change on Agricultural Land Use Suitability: Canola/Luke Vernon and Dennis van Gool*. [South Perth, W.A.]: Dept. of Agriculture.
- Via, S. 1993. "Adaptive Phenotypic Plasticity: Target or by-Product of Selection in a Variable Environment?" *American Naturalist* 142: 352–365.
- Wang, J., and Z. Zhang. 2021. "Gapit Version 3: Boosting Power and Accuracy for Genomic Association and Prediction." *Genomics, Proteomics & Bioinformatics* 19: 629–640.
- Xue, B., and S. Leibler. 2018. "Benefits of Phenotypic Plasticity for Population Growth in Varying Environments." *Proceedings of the National Academy of Sciences* 115: 12745–12750.
- Yang, J., S. H. Lee, M. E. Goddard, and P. M. Visscher. 2011. "GCTA: A Tool for Genome-Wide Complex Trait Analysis." *American Journal of Human Genetics* 88: 76–82.
- Yates, F., and W. G. Cochran. 1938. "The Analysis of Groups of Experiments." *Journal of Agricultural Science* 28: 556–580.
- Yu, J., G. Pressoir, W. H. Briggs, et al. 2006. "A Unified Mixed-Model Method for Association Mapping That Accounts for Multiple Levels of Relatedness." *Nature Genetics* 38: 203–208.
- Zhang, Z., E. Ersoz, C.-Q. Lai, et al. 2010. "Mixed Linear Model Approach Adapted for Genome-Wide Association Studies." *Nature Genetics* 42: 355–360.
- Zhao, Z., E. Wang, J. A. Kirkegaard, and G. J. Rebetzke. 2022. "Novel Wheat Varieties Facilitate Deep Sowing to Beat the Heat of Changing Climates." *Nature Climate Change* 12: 291–296.

### Supporting Information

Additional supporting information can be found online in the Supporting Information section.