



OPEN ACCESS

EDITED BY

Harsh Raman,
NSW Government, Australia

REVIEWED BY

Pilar Soengas,
Spanish National Research Council (CSIC),
Spain
Liu Yingru,
Hebei Agricultural University, China

*CORRESPONDENCE

Bahram Heidari
✉ bheidari@shirazu.ac.ir
Jin Wang
✉ jinwang@ujs.edu.cn
Xiao-Li Tan
✉ xltan@ujs.edu.cn

RECEIVED 28 June 2023

ACCEPTED 03 November 2023

PUBLISHED 11 January 2024

CITATION

Salami M, Heidari B, Batley J,
Wang J, Tan X-L, Richards C
and Tan H (2024) Integration of
genome-wide association studies,
metabolomics, and transcriptomics reveals
phenolic acid- and flavonoid-associated
genes and their regulatory elements under
drought stress in rapeseed flowers.
Front. Plant Sci. 14:1249142.
doi: 10.3389/fpls.2023.1249142

COPYRIGHT

© 2024 Salami, Heidari, Batley, Wang, Tan,
Richards and Tan. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that
the original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Integration of genome-wide association studies, metabolomics, and transcriptomics reveals phenolic acid- and flavonoid-associated genes and their regulatory elements under drought stress in rapeseed flowers

Maryam Salami¹, Bahram Heidari ^{1*}, Jacqueline Batley ²,
Jin Wang^{3*}, Xiao-Li Tan^{3*}, Christopher Richards⁴
and Helin Tan ⁵

¹Department of Plant Production and Genetics, School of Agriculture, Shiraz University, Shiraz, Iran, ²School of Biological Sciences, University of Western Australia, Perth, WA, Australia, ³School of Life Sciences, Jiangsu University, Zhenjiang, China, ⁴United States Department of Agriculture (USDA) Agricultural Research Service (ARS), National Laboratory for Genetic Resources Preservation, Fort Collins, CO, United States, ⁵State Key Laboratory of Crop Genetics and Germplasm Enhancement, Nanjing Agricultural University, Nanjing, China

Introduction: Biochemical and metabolic processes help plants tolerate the adverse effects of drought. In plants accumulating bioactive compounds, understanding the genetic control of the biosynthesis of biochemical pathways helps the discovery of candidate gene (CG)–metabolite relationships.

Methods: The metabolic profile of flowers in 119 rapeseed (*Brassica napus*) accessions was assessed over two irrigation treatments, one a well-watered (WW) condition and the other a drought stress (DS) regime. We integrated information gained from 52,157 single-nucleotide polymorphism (SNP) markers, metabolites, and transcriptomes to identify linked SNPs and CGs responsible for the genetic control of flower phenolic compounds and regulatory elements.

Results: In a genome-wide association study (GWAS), of the SNPs tested, 29,310 SNPs were qualified to assess the population structure and linkage disequilibrium (LD), of which several SNPs for radical scavenging activity (RSA) and total flavanol content (TFLC) were common between the two irrigation conditions and pleiotropic SNPs were found for chlorogenic and coumaric acids content. The principal component analysis (PCA) and stepwise regression showed that chlorogenic acid and epicatechin in WW and myricetin in DS conditions were the most important components for RSA. The hierarchical cluster analysis (HCA) showed that vanillic acid, myricetin, gallic acid, and catechin were closely

associated in both irrigation conditions. Analysis of GWAS showed that 60 CGs were identified, of which 18 were involved in stress-induced pathways, phenylpropanoid pathway, and flavonoid modifications. Of the CGs, *PAL1*, *CHI*, *UGT89B1*, *FLS3*, *CCR1*, and *CYP75B137* contributed to flavonoid biosynthetic pathways. The results of RNA sequencing (RNA-seq) revealed that the transcript levels of *PAL*, *CHI*, and *CYP75B137* known as early flavonoid biosynthesis-related genes and *FLS3*, *CCR1*, and *UGT89B1* related to the later stages were increased during drought conditions. The transcription factors (TFs) *NAC035* and *ERF119* related to flavonoids and phenolic acids were upregulated under drought conditions.

Discussion: These findings expand our knowledge on the response mechanisms to DS, particularly regarding the regulation of key phenolic biosynthetic genes in rapeseed. Our data also provided specific linked SNPs for marker-assisted selection (MAS) programs and CGs as resources toward realizing metabolomics-associated breeding of rapeseed.

KEYWORDS

Brassica napus L., candidate gene, GWAS, haplotype, phytochemicals, SNP, transcriptomic

Introduction

Brassica napus L. ($2n = 4x = 38$, AACC), commonly known as rapeseed or canola, is the most important oil crop of the *Brassica* genus and is widely grown globally (Chalhoub et al., 2014). *B. napus* with high oil (~50%) and protein (~25%) contents is used as fodder and human food. The leaves and flowers of rapeseed are edible and can be used for therapeutic purposes due to their diuretic, anti-scurvy, and anti-inflammatory activities (Raboanatahiry et al., 2021). Furthermore, rapeseed is rich in both nutritional and non-nutritional antioxidants and accumulates phenols, hydroxybenzoic acids, hydroxycinnamic acids, and flavonoids (flavanols, flavones, flavanones, isoflavones, and anthocyanins) at a rate of 10–30 times higher than other oilseeds (Shao et al., 2014). The growing interest in edible flowers is motivated not only by their decorative and nutraceutical properties but also because of the desire for new flavors and new opportunities for food innovation. The antioxidant power of flowers derives from their richness in generic phenolic compounds (Mikolajczak et al., 2020). Phenolic compounds (phenolic acids and flavonoids) are substantial groups of plant secondary metabolites known as markers for biotic and abiotic stress tolerance (Sharma et al., 2019). Phenolic contents and compounds and phenolic oxidizing enzymes are strongly associated with food quality (Tomas-Barberan and Espin, 2001).

The biosynthesis of flavonoids is regulated by various signaling molecules, biosynthetic enzymes, transcription factors (TFs), and phytohormonal regulators (Li et al., 2013). The flavonoid biosynthetic-related genes at early steps including the key enzymes chalcone synthase (CHS), chalcone isomerase (CHI), flavanone 3-hydroxylase (F3H), flavonoid 3'-hydroxylase (F3'H), and flavonoid 3',5'-hydroxylase (F3'5'H) are responsible for

producing common precursors (Zhao, 2015). Then, the later-step biosynthetic enzymes dihydroflavonol-4-reductase (DFR), flavonol synthase (FLS), leucoanthocyanidin reductase (LAR), leucoanthocyanidin dioxygenase (LDOX), anthocyanidin reductase (ANR), and UDP-glucose:flavonoid 3-O-glucosyltransferase (UFGT) catalyze the formation of anthocyanin and other flavonoid end products (flavonols and flavones) and phenolic acid (Han et al., 2012). Besides the effects of genes, environmental factors affect flavonoid accumulation and expression of the flavonoid pathway genes that shows the need for the evaluation of a germplasm under different environmental conditions (Ahmed et al., 2021).

Most parts of Iran's agricultural areas suffer from water scarcity due to low rainfall and poor irrigation systems, which are coincident with flowering and seed filling of rapeseed (Kalantarahmadi and Shoushi Dezfouli, 2021). It has been shown that rapeseed is more sensitive to water stress at the flowering and seed-forming stages (Elferjani and Soolanayakanahally, 2018). Abiotic stress such as drought affects the accumulation of phenolic compounds in plants (Gómez-Caravaca et al., 2012). Drought stress (DS) triggers the accumulation of reactive oxygen species (ROS) that can initiate lipid peroxidation, chlorophyll and betalain bleaching, and protein oxidation (Li et al., 2022). In response to the adverse effects of environmental conditions, plants produce a series of enzymatic and nonenzymatic defense systems for scavenging and detoxifying ROS, resulting in an enhanced antioxidant defense capacity (Huang et al., 2019; Farooq et al., 2021). Nonenzymatic antioxidants comprising of ascorbic acid, carotenoids, phenolic, and flavonoids have a protective role against the production of ROS in plants (Zhang et al., 2019).

Genetic analysis of drought-adaptive traits helps to identify the most important genes involved in networks and pathways for use in

marker-assisted selection (MAS) programs and cultivar development under variable environmental conditions. Drought tolerance is a complex trait regulated by several minor and/or major effects quantitative trait loci (QTL) (Maazou et al., 2016). In rapeseed, a genome-wide association study (GWAS) has been used for genetic analysis of a range of traits such as plant height and primary branch number (Li et al., 2016), time of flowering (Xu et al., 2016), seed coat color (Wang et al., 2017), blackleg disease (Fikere et al., 2020), and oil content (Pal et al., 2021). However, only a few studies have analyzed the use of GWAS to dissect the genetics of drought tolerance-related traits in rapeseed, many of which have been based on low-marker density maps that may not be able to capture the diversity of drought tolerance mechanisms. Zhang et al. (2015) analyzed 66 rapeseed accessions under drought conditions and identified 16 single-nucleotide polymorphisms (SNPs) significantly associated with water stress-related responses. Recently, Shahzad et al. (2021) identified 139 SNPs associated with the water loss ratio using GWAS in 264 rapeseed accessions at the full-bloom stage, of which four were related to putative candidate genes (CGs) involved in drought tolerance in rapeseed.

The genetic basis of the phenolic compounds in response to DS and associated CGs has not yet been characterized in rapeseed germplasm under DS conditions. Rapeseed flowers have nutritional and industrial benefits, but this organ and the flowering stage are vulnerable to drought and, therefore, to seed production. In this study, our aim was to first assess the relation of flower polyphenols and drought tolerance and then identify linked markers/genes to find drought tolerance by selecting high-polyphenol varieties in rapeseed. First, we investigated the effects of irrigation regimes on phytochemical traits and polyphenols and induction of gene regulatory elements. Then, a GWAS analysis was conducted in the well-watered (WW) and DS conditions to assess the genomic architecture of the phytochemical traits and polyphenols. The RNA sequencing (RNA-seq) trial was performed in order to validate genes associated with the linked SNPs and identify key CGs involved in the regulatory pathway of phenolics. Differentially expressed genes (DEGs) between contrasting phenotypes (high- vs. low-phenolic content varieties) were identified and compared with genes underlying the genomic region (GR) of SNPs identified for the tested traits. Our intent was to identify key regulatory genes in phenolic biosynthetic pathways and provide better insight into the genetic basis underlying the establishment of high-phenolic compound varieties showing drought tolerance in rapeseed.

Materials and methods

Field experimental conditions

The plant materials consisted of 119 rapeseed genotypes including inbred lines, hybrids, and commercial cultivars (Supplementary Table S1). The seeds of the genotypes were sown in 11 × 11 lattice experimental design with three replications. Experimental plots were composed of four 1-m-long rows, each with 25 plants in 4-cm spaces. The experiments were performed under two irrigation treatments, one a WW condition and the other

a DS treatment at the Research Farm of Plant Production and Genetics, Shiraz, Iran. The seeds were sown on 17 September 2020 in a silty loam soil. Information on macronutrient and micronutrient elements of soil is shown in Supplementary Table S2. Fertilizers were applied at rates of 250 kg N ha⁻¹, 100 kg P₂O₅ ha⁻¹, and 100 kg K₂SO₄ ha⁻¹. Nitrogen fertilizer was added to the soil prior to sowing, at the time of rosette leaf stage, and at the beginning of flowering. In the WW condition, irrigation was carried out until seed physiological maturity at the time when no further increase in seed weight occurred. In the DS treatment, irrigation was carried out until 50% pod development stage. The weather data during the growing season show that rainfall did not occur during implementation of the DS period (Supplementary Figure S1).

Sampling and preparation of the extracts

In a study, the results showed that the flower accumulates the highest polyphenols and antioxidant activity compared with leaves and seeds (Salami et al., 2023), which supported the results of the study by Khan et al. (2021). Four weeks after ceasing irrigation in the drought treatment and at the end of flowering, 10 flower samples were cut from five main plants in both irrigation regimes, snap frozen in liquid nitrogen, and stored at -80°C. The flower samples were homogenized with 80% methanol of 1:2 (w/v) at 4°C. The homogenization was sonicated for 10 min and mixed at 15 rpm on a tube rotator. Homogenized samples were immediately placed on a stirrer at room temperature for 2 h and centrifuged at 10,000 g for 20 min. Supernatants in triplicate were collected for further analyses (Samec et al., 2011).

Biochemical assays

Total phenolic content (TPC) was quantified as described by Sarker and Oba (2018). Fifty microliters of the extract was mixed with 1 mL of Folin–Ciocalteu phenol reagent. Then, 1 mL of 10% sodium carbonate solution was added to the mixture and shaken thoroughly. The mixture was left for 30 min, and the blue color formed was measured at 760 nm wavelength using a spectrophotometer. A regression equation for the calibration curve ($y = 0.0095x - 0.113$, $R^2 = 1$) of gallic acid was used, and the data were presented as μg gallic acid equivalent (GAE) g⁻¹ dry weight (DW).

Total flavonoid content (TFDC) was determined with an aluminum chloride reagent followed by a colorimetric method (Zou et al., 2004). A total of 0.5 mL of each extract was mixed with 150 μL of 10% aluminum chloride and 2 mL of 1 M sodium hydroxide. The mixtures were incubated for 40 min at room temperature in the dark. The absorbance at 510 nm was measured using a spectrophotometer (Model M550 UV/VIS Spectrophotometer CamSpec). A calibration curve was drawn using a quercetin standard dilution series, and the data were shown as mg quercetin equivalents (QUE) g⁻¹ DW.

The total flavonol content (TFLC) was determined using the *p*-dimethylaminocinnamaldehyde (DMACA) method (Kusznierewicz

et al., 2008). Four hundred microliters of extract was mixed with 600 μL of DMACA solution (0.1% in 1 M HCl in MeOH). The mixture was vortexed, and the absorbance was measured at 640 nm after incubation at room temperature for 10 min. Catechin (CE) was used as a standard, and the TFLC data were expressed as μg catechin equivalents (CE) g^{-1} DW.

Determination of total anthocyanin content and ascorbic acid content

The total anthocyanin content (TAC) was determined using a pH difference method (Egbuna et al., 2018). Two milliliters of the extracts were added to 2 mL of hydrochloric acid (2 N) and 2 mL of ammonia (2N). The extracts were diluted with 0.025 M hydrochloric acid–potassium chloride buffer (pH = 1) and 0.4 M sodium acetate buffer (pH = 4.5). Each sample was diluted with the buffers, and the absorbance of the mixture was measured at 700 nm using a UV-Vis spectrophotometer (UV1601; Shimadzu, Kyoto, Japan). The total anthocyanin content was expressed as cyanidin-3-glucoside equivalents as follows (Zou et al., 2011; Maran et al., 2015):

$$\text{Anthocyanin pigment (mg/L)} = A \times MW \times DF \times V \times 1,000 / a \times l \times m \quad (1)$$

where A is the absorbance, MW is the molecular weight of cyanidin-3-glucoside (449.2 g/mol), DF is the dilution factor, V is the solvent volume (mL), a is the molar absorptivity (26,900 $\text{L mol}^{-1} \text{cm}^{-1}$), l is the cell path length (1 cm), and m is the weight of the sample.

The ascorbic acid content (AAC) of extracts was determined by the 2,6-dichlorophenol-indophenol (DCPIP) method as described by Rao and Deshpande (2006). The freeze-dried samples (0.5 g) were extracted using 20 mL (w/v) metaphosphoric acid (3%) followed by shaking at 300 rpm for 30 min. The extract was centrifuged at 4,000 rpm for 10 min. Here, 1 mL of extract was added to 3 mL of 0.2 mM DCPIP, and the spectrophotometer reading was performed at 515 nm immediately after mixing for 15 s. The results were expressed in mg ascorbic acid per 100 g dry weight (mg/100 g DW).

High-performance liquid chromatography and phenolic compounds

Chromatographic analysis was performed on a Varian ProStar 210 system equipped with a vacuum degasser, binary pump, thermostatted column compartment, and diode array detector (DAD) connected to ChemStation software ver. 4.03.016 (Lopez et al., 2011). The compounds were separated using a reverse-phase Pursuit XRs C18 (250 mm \times 4.6 mm, 5 μm) column and a Pursuit XRs C18 (10 mm \times 4.6 mm, 5 μm) guard column (Varian, Barcelona) at 27°C and based on a gradient system with two

mobile phases. Eluents A and B were water with 0.1% formic acid and methanol, respectively. The flow rate was 1.0 mL/min, and the injection volume was 60 μL of crude extracts. The elution conditions used were as follows: 0.5 min, 20% B isocratic; 5–30 min, linear gradient from 20% to 60% B; 30–35 min, 60% B isocratic; 30–40 min, linear gradient from 60% to 20% B, and finally wash and recondition the column. Simultaneous monitoring for quantification of phenolics was performed at 270 nm (gallic acid, protocatechuic acid, catechin, vanillic acid, epicatechin, and syringic acid), 324 nm (chlorogenic acid, gentisic acid, caffeic acid, coumaric acid, and ferulic acid), and 373 nm (rutin, myricetin, and quercetin). The stock standard solution diluted with methanol was used to obtain an appropriate concentration 1.0–100 $\mu\text{g/mL}$ range for establishing the calibration curves. For quantitative analysis, four different concentrations of 14 analytes were injected in triplicate.

Evaluation of the antioxidant activity by DPPH method

The free radical scavenging activity (RSA) of the extracts was evaluated with 1,1-diphenyl-2-picryl-hydrazil (DPPH) following Molyneux (2004) with some modifications. Briefly, 193 μL methanolic solution of free radical DPPH (60 μM) was added to 7 μL of the sample extracts. After incubation in the dark at room temperature for 30 min, the absorbance of the mixture at 515 nm against methanol was measured as a blank using a spectrophotometer-microplate reader (Epoch, BioTek, Instruments, Inc.). The control solution was 200 μL methanolic solution of free radical DPPH (60 μM). The RSA of the extracts was expressed as a percentage of the reduced DPPH as follows:

$$\text{RSA (\%)} = [(A_0 - A_1) / A_0] \times 100 \quad (2)$$

where A_0 = absorbance of the control and A_1 = absorbance of the test extracts.

Statistical analyses of phytochemical traits

Descriptive statistics including means and standard deviations were obtained using the Statistical Analysis System (SAS, Version 9.3, SAS Institute Inc., Cary, NC, USA). Analysis of variance (ANOVA) was performed based on a generalized linear model (GLM) with genotype as fixed and replication as random effects. The principal component analysis (PCA) and hierarchical cluster analysis (HCA) were performed to visualize the interrelationships between phenolic compounds and similarities of genotypes under two experimental conditions in R version 3.4.0 (<http://www.r-project.org/>). Stepwise regression analysis was used to determine how much antioxidant activity variability could be explained by the independent variables (TPC, TFDC, TFLC, AAC, TAC, and phenolic compounds).

Genotyping and filtering of single-nucleotide polymorphisms

Genomic DNA of 119 rapeseed accessions was extracted from young leaf tissue of greenhouse-grown plants using a cetyltrimethylammonium bromide (CTAB)-based method (Murray and Thompson, 1980). The genomic DNA was quantified using a Nanodrop ND1000 Spectrophotometer (Nanodrop Technologies, Inc., Wilmington, DE, USA). The DNA samples were used for genotyping with the Brassica 60 K Infinium array as described in the manufacturer's protocol (Illumina Inc., San Diego, CA, USA). The SNPs were filtered for site coverage (90%), minimum minor allele frequency (MAF) of 0.05 with only biallelic markers, and low rates of missing data ($\leq 10\%$) using the TASSEL ver. 5 software (Bradbury et al., 2007). The SNPs with low quality and high rates of missing data were removed using a protocol described in Iquira et al. (2015) and Kujur et al. (2015). Analyses of polymorphic information content (PIC) were performed using the software PowerMarker version 3.25 (Liu and Muse, 2005).

Population structure and linkage disequilibrium analysis

Bayesian cluster analysis was used for analysis of the population structure and the fraction of the genetic ancestry of a single accession that belongs to a population using STRUCTURE software (Sehgal et al., 2015). The most appropriate subpopulation number (K) was obtained using the Ln probability of the data [Ln P(D)] and the delta K values were calculated according to Evanno et al. (2005). The maximum of the 1K peak was considered as the true cluster count. The linkage disequilibrium (LD) between pairs of polymorphic loci was analyzed using the TASSEL software package (Bradbury et al., 2007; <http://www.maizegenetics.net/>). LD was estimated using the squared allele frequency correlations (r^2), which is a measure of the correlation between a pair of variables (Hill and Robertson, 1968). To investigate chromosome-wide and genome-specific patterns of LD, the software package TASSEL 4.0 was used to estimate LD (r^2) on each chromosome and across the A- and C-sub genomes, respectively (Bradbury et al., 2007). Pairwise LD explained by r^2 was determined for 29,310 high-quality SNPs with MAF ≥ 0.05 . Locally estimated scatterplot smoothing (LOESS, stands for locally weighted scatterplot smoothing) curves were then fitted into the scatter plot using R software. Specifically, the 95th percentile of the distributions of r^2 of the selected loci was estimated as the threshold r^2 with the assumption that LD was attributable to linkage (Breseghello and Sorrells, 2006).

Genome-wide association study analysis

Marker-trait association (MTA) analysis was performed using the TASSEL program based on the mixed linear model (MLM)

approach with the principal components (PCs) and kinship (K) matrices as cofactor (Bradbury et al., 2007). The MLM-based analysis requires the data for population structure (Q), PCs, K, and a total of 29K high-quality SNPs in 119 rapeseed genotypes (Yu et al., 2006). The K-Matrix was generated with the default parameters by choosing the Centered_IBS method to get a better estimate of the additive genetic variance. Analysis of MTAs was also performed with the general linear model (GLM) and GLM+Q and GLM+PCA models (Yu et al., 2006). The program's default settings were used to filter marker data for minimum genotype number and MAF. The phenotypic variation ($R^2\%$) explained was also estimated for each individual marker. Quantile-quantile (QQ) plots were shown with $-\log_{10}(P)$ of each SNP and expected P -value using the R package qqman (<https://cran.r-project.org/web/packages/qqman/index.html>; Turner, 2014). The results of the MLM analysis were used to obtain the Manhattan plots in TASSEL.

Identification of candidate genes

To determine CGs related to phenolic compounds under WW and drought-stressed conditions, the position of flanking SNPs obtained from the "Darmor-bzh" reference genome (<http://www.genoscope.cns.fr/brassicapapus/data/>) was used to search in the rapeseed genome by EnsemblPlants (<https://plants.ensembl.org/>). Consequently, all of the genes underlying the GR of each SNP were functionally annotated by EnsemblPlants (<https://plants.ensembl.org/>) and online resources (<https://genome.ucsc.edu/> and <https://www.ncbi.nlm.nih.gov/>). The CGs were identified based on their putative function in rapeseed or closely related species.

Allele effect and haplotype analysis

Allele effects for the linked significant SNPs were analyzed as previously described by Alemu et al. (2021). Genotypes were divided into two different groups according to their specific SNP alleles, and the means were compared using Tukey's Honest Significant Difference (HSD) test. Exploring and harnessing haplotype diversity help in the detection of CGs for improvement of target traits in crops (Qian et al., 2017). A haplotype association test was performed to investigate the combined effect of the linked significant SNPs. The SNPs in the same haploblock and the LD of significant SNPs were determined using Haploview 4.2 (Barrett et al., 2005). Standardized disequilibrium coefficient (D') was used to evaluate the LD between markers and generate the LD heatmap. Haplotype blocks were detected based on LD using the confidence interval (CI) method in Haploview 4.2 (Gabriel et al., 2002).

RNA-seq and biological validation of the identified candidate genes

RNA-seq analysis in two contrasting lines, a top high (G67) and a low-phenolics (G59) cultivar, was used to validate the CGs

associated with the SNPs of phenolic compounds. The plants were grown in pots in a greenhouse (26°C–28°C) under short-day (12 h of light and 12 h of dark) conditions until the initial flower stage. When the flower was visible, half of the plants were exposed to DS conditions. During irrigation, drought-treated flower pots maintain a soil relative water content of 10% (irrigated with PEG6000 at a concentration of 20%), while adequately watered flower pots maintain a soil relative water content of 30% (irrigated with sterile water of equal volume) with repeating the experiment three times. The relative water content of soil was calculated according to the following formula (Zhang et al., 2023):

$$\text{Soil moisture content (\%)} = (W2 - W3) / (W3 - W1) \times 100\% \quad (3)$$

where W1 is the weight of the container, W2 is the weight of the container with wet soil, and W3 is the weight of the container with dry soil.

Fifteen days after flowering, the top flowers were harvested from plants in each replicate group and each condition (WW/control or DS). All harvested flowers were immediately frozen using liquid nitrogen and were transferred to a deep freezer (−80°C) for storage.

Total RNA was isolated from each sample using a Plant RNA Mini Kit (Tiangen, Inc., China) according to the manufacturer's protocol. Four cDNA libraries were constructed, and RNA-seq was performed on a DNBSEQ-G400 platform. Low-quality reads were filtered out using NGS QC toolkit v2.2.3 (<https://omictools.com/ngs-qc-toolkit-tool>) (Patel and Jain, 2012), and clean reads were mapped in the reference genome of “Darmor-bzh” (<http://www.genoscope.cns.fr/brassicanaapus/>). The obtained genes from the previous step were quantitatively analyzed using Cluffquant and Cluffnorm of Cufflinks 2.0.0 (<http://cole-trapnell-lab.github.io/cufflinks/releases/v2.0.0/>). Gene expression levels were estimated by the FPKM (fragments per kilobase of exon per million mapped fragments) and an all by all differential expression analysis by combining replications. DEGs in the specific paired sample comparisons were identified with the log₂ fold change between the two samples and the *P*-value given to the comparison along with the FPKMs for each of the two samples in the comparison.

Results

Phenotypic variation for phytochemicals in rapeseed population

The ANOVA revealed a significant effect of drought on TPC, TFDC, TFLC, AAC, TAC, and RSA and variations in 119 genotypes (Supplementary Table S3). The interaction of treatment (WW and DS conditions) × genotype was significant for TPC, TFDC, and TFLC. The means of the traits varied between WW and drought-stressed genotypes as shown in Table 1. The frequency distribution of genotypes and box plots for phytochemicals across two moisture conditions are presented in Supplementary Figure S2 and Figure 1. The largest and smallest variations belonged to the means for TFLC and RSA, respectively (Table 1). The average of TPC, TFDC, TFLC, AAC, TAC, and RSA under DS was 87.09 μg GAE g^{−1} DW, 851.97

mg GAE g^{−1} DW, 1,370.45 μg CE g^{−1} DW, 85.33 mg/100 g DW, 427.67 mg/L, and 92.18%, respectively, which was higher than those in WW conditions. The absolute kurtosis and skewness of most of the phytochemicals were <1, reflecting the normal distribution of the traits (Table 1).

Profiling of phenolic compounds under well-watered and drought stress conditions

Interactions were significant for phenolic compounds except chlorogenic and gentisic acids (Supplementary Table S3). In the analysis of high-performance liquid chromatography (HPLC), 14 phenolic compounds from phenolic acid and flavonoid groups were identified (Figures 2A–F). Gallic, protocatechuic, vanillic, syringic, chlorogenic, gentisic, caffeic, coumaric, and ferulic acids were the major phenolic acids identified. The epicatechin, syringic, caffeic, coumaric, chlorogenic, and protocatechuic acids were more frequent than other compounds in the tested accessions (Table 1; Figures 2D–F).

Results of HCA are shown in Figures 3A, B. In the WW condition, quercetin, vanillic acid, myricetin, gallic acid, and catechin clustered as one group and rutin and coumaric acid in another group with chlorogenic and gentisic acids (Figure 3A). The heat map of genotypes based on their phenolics in the DS condition showed a different spatial distribution of phenolics from WW, where protocatechuic acid, gallic acid, vanillic acid, myricetin, catechin, ferulic acid, rutin, coumaric acid, chlorogenic acid, and gentisic acid were closely associated (Figure 3B).

Phytochemicals and phenolic compounds increased under drought conditions

Under DS, TPC, TFDC, TFLC, AAC, and TAC were increased by 7.04%, 14.64%, 12.82%, 6.86%, and 15.35%, respectively (Table 1; Figure 3C). DS significantly increased the accumulation of phenolic acids and flavonoids such as caffeic acid (115.85%), epicatechin (105.13%), myricetin (72.26%), and syringic acid (14.32%).

Relation of antioxidant activity and phenolics

The interrelationship of the antioxidant activity and phenolics was assessed by both the PCA and stepwise regression. The PCA was performed to identify the relationship between the pattern of phytochemicals and phenolics and the antioxidant activities. Theoretically, 17 PCs can be derived, but the PCs with higher contributions to total variation were selected for further analysis. In WW, seven PCs with eigenvalues greater than 1 explained 70.59% of the variability in the dataset (Supplementary Table S4). The first two PCs (PC1 and PC2) that explained 30.3% of the total variation of traits were used in biplot analysis (Figure 3D). PC1 represented

TABLE 1 Descriptive statistics for the phytochemical traits, phenolic compounds, and antioxidant activity of 119 rapeseed (*Brassica napus*) genotypes grown at two irrigation treatments.

Trait	Unit	Treatment	Range	Mean \pm SD ^a	Skew ^b	Kurt ^c
TPC	$\mu\text{g GAE g}^{-1}$ DW	WW	55.94 - 110.63	81.36 \pm 1.23	-0.12	-1.08
		DS	59.36 - 114.89	87.09 \pm 2.30	-0.2	-0.95
TFDC	mg GAE g^{-1} DW	WW	219.86 - 1388.43	743.15 \pm 9.52	0.35	-0.41
		DS	354.57 - 1362.72	851.97 \pm 8.48	0.15	-0.81
TFLC	$\mu\text{g CE g}^{-1}$ DW	WW	314.5 - 2392	1214.62 \pm 0.78	0.65	-0.81
		DS	553.3 - 2408.3	1370.45 \pm 0.59	0.64	-0.77
RSA	%	WW	87.92 - 93.09	91.48 \pm 1.01	-1.1	2.14
		DS	89.48 - 93.45	92.18 \pm 1.20	-0.84	1.24
AAC	$\text{mg}/100$ g DW	WW	50.07 - 112.565	79.85 \pm 1.87	-0.54	1.44
		DS	63.08 - 119.55	85.33 \pm 1.07	0.44	2.1
TAC	mg/L	WW	212.58 - 563.09	370.75 \pm 0.72	-0.02	-0.65
		DS	288.14 - 570.37	427.67 \pm 0.77	-0.16	-0.56
Gallic acid	$\text{mg}/100$ g DW	WW	2.43 - 6.39	3.21 \pm 1.05	2.48	1.54
		DS	2.83 - 7.17	3.83 \pm 2.64	0.84	1.23
Protocatechuic acid	$\text{mg}/100$ g DW	WW	3.09 - 7.76	4.61 \pm 0.46	0.46	1.04
		DS	2.41 - 16.62	5.7 \pm 0.75	1.09	1.32
Catechin	$\text{mg}/100$ g DW	WW	2.63 - 5.41	3.42 \pm 1.50	0.8	0.77
		DS	3.23 - 8.76	5.34 \pm 1.31	0.7	1.11
Vanillic acid	$\text{mg}/100$ g DW	WW	nd - 7.36	4.11 \pm 1.26	-0.01	-1.74
		DS	nd - 7.50	4 \pm 0.36	-0.17	-0.23
Epicatechin	$\text{mg}/100$ g DW	WW	2.59 - 12.22	4.48 \pm 1.35	2.45	1.32
		DS	2.73 - 55.38	9.19 \pm 1.90	1.21	1.11
Syringic acid	$\text{mg}/100$ g DW	WW	2.60 - 61.67	17.17 \pm 1.14	0.42	-1.13
		DS	2.74 - 44.52	19.63 \pm 2.61	0.27	-1.64
Chlorogenic acid	$\text{mg}/100$ g DW	WW	4.13 - 10.83	7.07 \pm 1.14	-0.52	0.35
		DS	2.87 - 22.65	7.84 \pm 1.31	1.13	0.98
Gentisic acid	$\text{mg}/100$ g DW	WW	3.97 - 9.60	7.32 \pm 1.99	-0.89	0.08
		DS	3.3 - 9.73	7.36 \pm 0.78	-0.89	0.35
Caffeic acid	$\text{mg}/100$ g DW	WW	2.52 - 15.52	4.54 \pm 1.04	1.64	1.36
		DS	3.11 - 45.21	9.8 \pm 1.72	1.06	1.12
Coumaric acid	$\text{mg}/100$ g DW	WW	3.53 - 8.88	6.35 \pm 0.27	-0.25	-0.36
		DS	3.77 - 27.64	7.41 \pm 0.67	1.22	0.78
Ferulic acid	$\text{mg}/100$ g DW	WW	2.86 - 4.81	4.07 \pm 3.37	-1.32	1.73
		DS	2.53 - 8.22	4.71 \pm 3.14	0.89	1.88
Rutin	$\text{mg}/100$ g DW	WW	2.44 - 11.27	5.54 \pm 0.68	0.57	-1.4
		DS	2.34 - 17.98	7.21 \pm 1.52	0.34	-0.64
Myricetin	$\text{mg}/100$ g DW	WW	2.25 - 5.59	2.74 \pm 0.15	1.43	1.24
		DS	nd - 13.97	4.72 \pm 0.23	1.24	1.35

(Continued)

TABLE 1 Continued

Trait	Unit	Treatment	Range	Mean \pm SD ^a	Skew ^b	Kurt ^c
Quercetin	mg/100 g DW	WW	nd - 3.75	2.68 \pm 0.99	1.76	1.32
		DS	nd - 3.17	1.15 \pm 0.36	0.29	-1.34

nd, not detected. ^a Standard deviation. ^b Skewness, which reflects the asymmetry of the probability distribution of a real-valued random variable about its mean. ^c Kurtosis, which reflects the “tailedness” of the probability distribution of a real-valued random variable. TPC, TFDC, TFLC, RSA, AAC, and TAC were the abbreviations of total phenolic content, total flavonoid content, total flavanol content, radical scavenging activity, ascorbic acid content, and total anthocyanin content. WW and DS were the codes of well-watered and drought stress conditions.

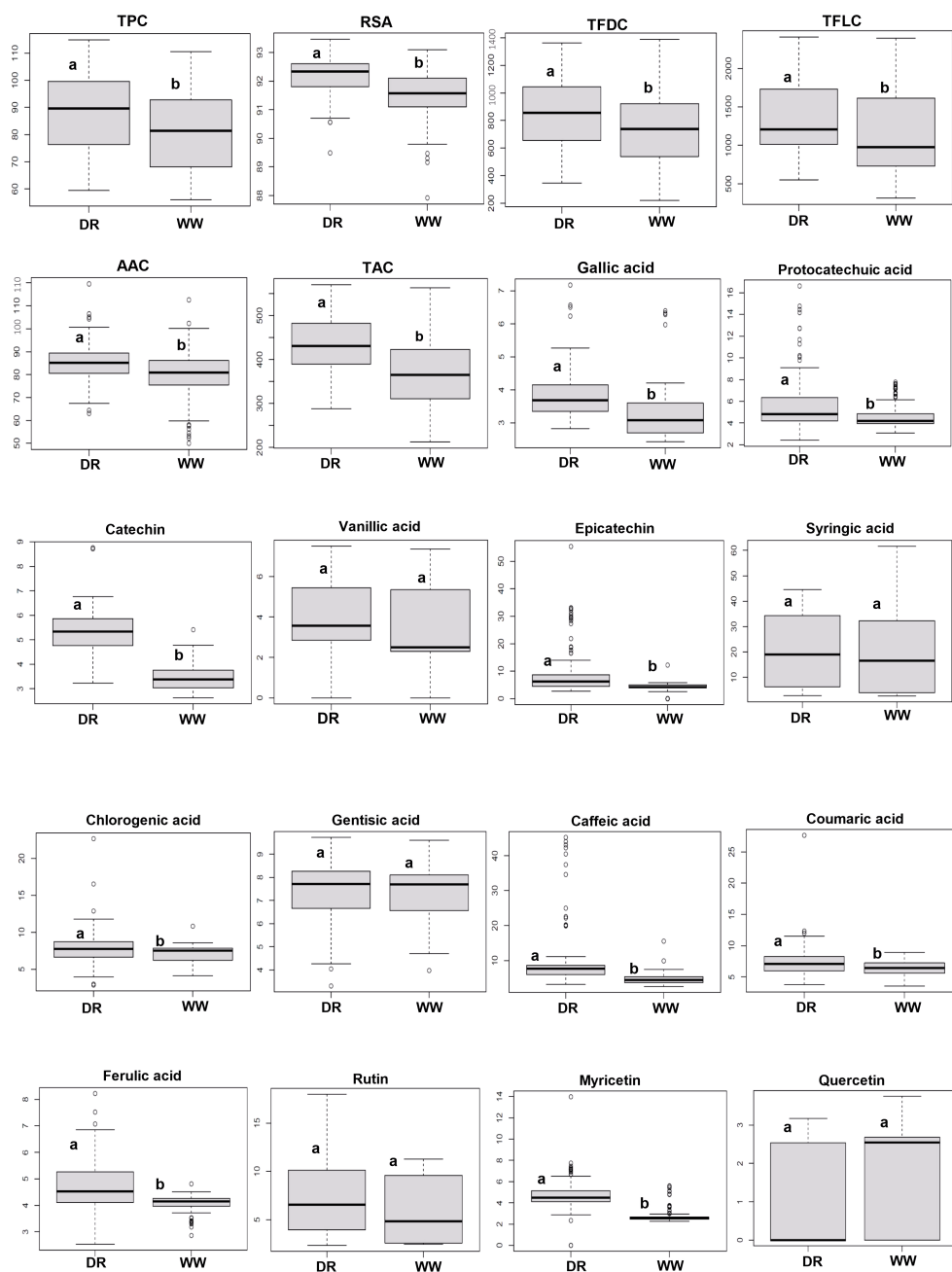
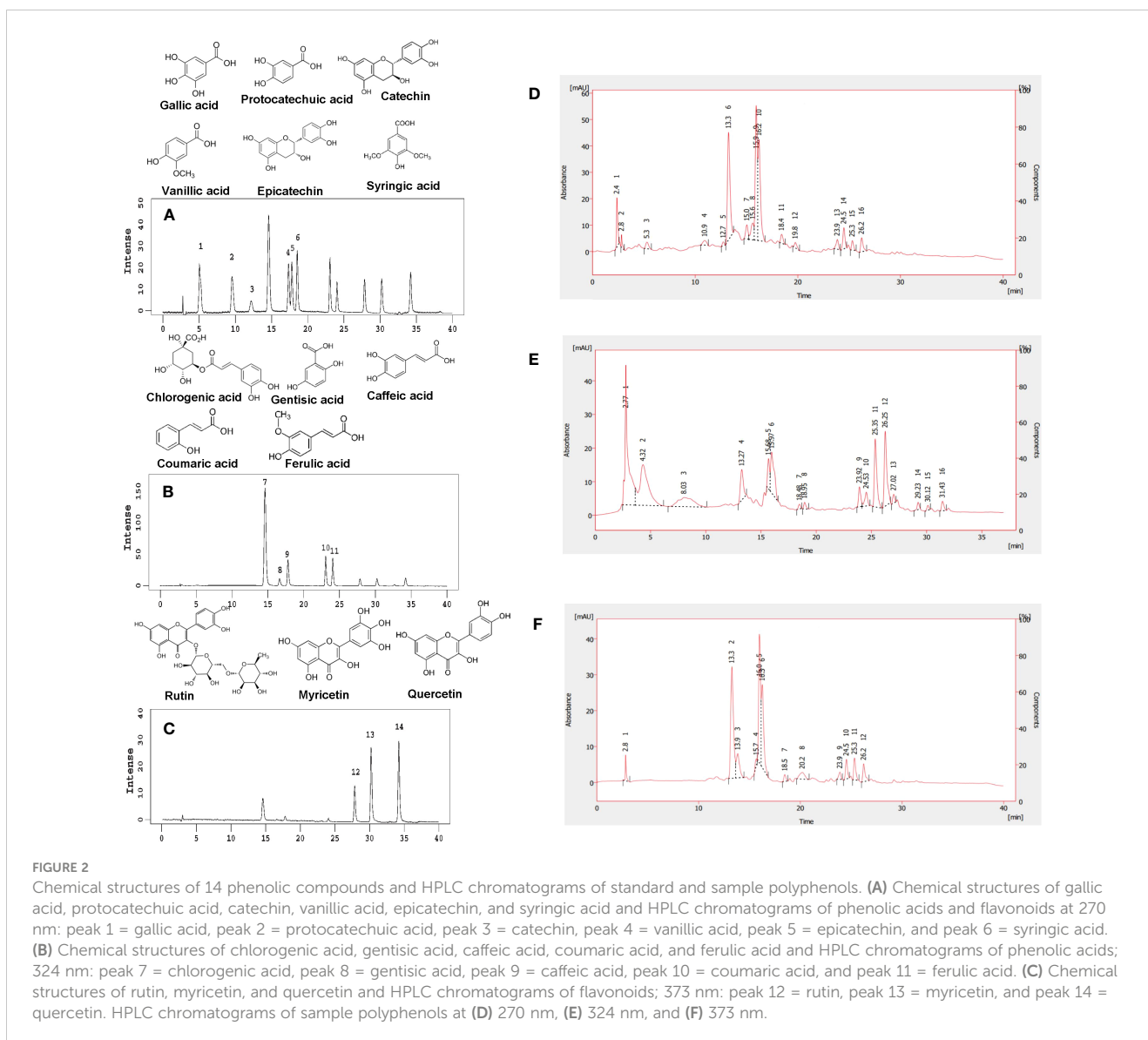


FIGURE 1

Boxplot of trait variations for rapeseed (*Brassica napus*) genotypes in the well-watered condition (WW) and drought stress condition (DR). Different letters on the boxes indicate a significant difference at $P < 0.05$ by Tukey's HSD. The horizontal line within the box represents the median. The lower and upper limit of the box, lower and upper whisker represents Q1 (first quartile/25th percentile), Q3 (third quartile/75th percentile), (Q1–1.5IQR) and (Q3 + 1.5IQR), respectively. IQR, interquartile range.



the variation in gallic acid, catechin, and TPC, while PC2 explained the variation of syringic acid (46%) and rutin (43.4%). Rutin, syringic acid, chlorogenic acid, epicatechin, quercetin, and gentisic acid showed a stronger association with antioxidants, while caffeic, ferulic, and coumaric acids showed a negative correlation with antioxidant activity (Figure 3D).

Under DS, eight PCs with eigenvalues greater than 1 explained 66.13% of the variability in the dataset (Supplementary Table S5). The first two PCs explaining 23.7% of the total variation of the traits were used for the analysis of the relation of traits in a biplot (Figure 3E). Catechin was the most important contributor to PC1. For PC2, vanillic acid, with a positive coefficient, was the most representative variable. Gentisic acid, syringic acid, chlorogenic acid, coumaric acid, myricetin, and rutin showed a strong association with antioxidant activity, while quercetin, catechin, vanillic acid, protocatechuic acid, and gallic acid had negative correlations with antioxidant activity (Figure 3E). Results of the stepwise regression analysis showed that chlorogenic acid in

the WW condition and gentisic acid, coumaric acid, and myricetin in the DS treatment had higher contributions to RSA (Supplementary Table S6).

Distribution of SNPs, population structure, and LD analysis

Of the 52,157 SNPs, 22,847 SNPs with a MAF of <0.05 were removed and the remaining 29,310 SNPs were used to assess the population structure, LD, and association analysis. The filtered SNPs were distributed across the 19 chromosomes with 722 SNPs on chromosome C05 to 2,623 on C04 (Figure 4A).

The results of the population structure analysis are shown in Figures 4B–D. The clustering program estimated the membership probability (Q-matrix) of each rapeseed genotype to combine into a number of hypothetical subpopulations (K3–K10), and the statistic index Δk value was generated for subsequent runs. The most

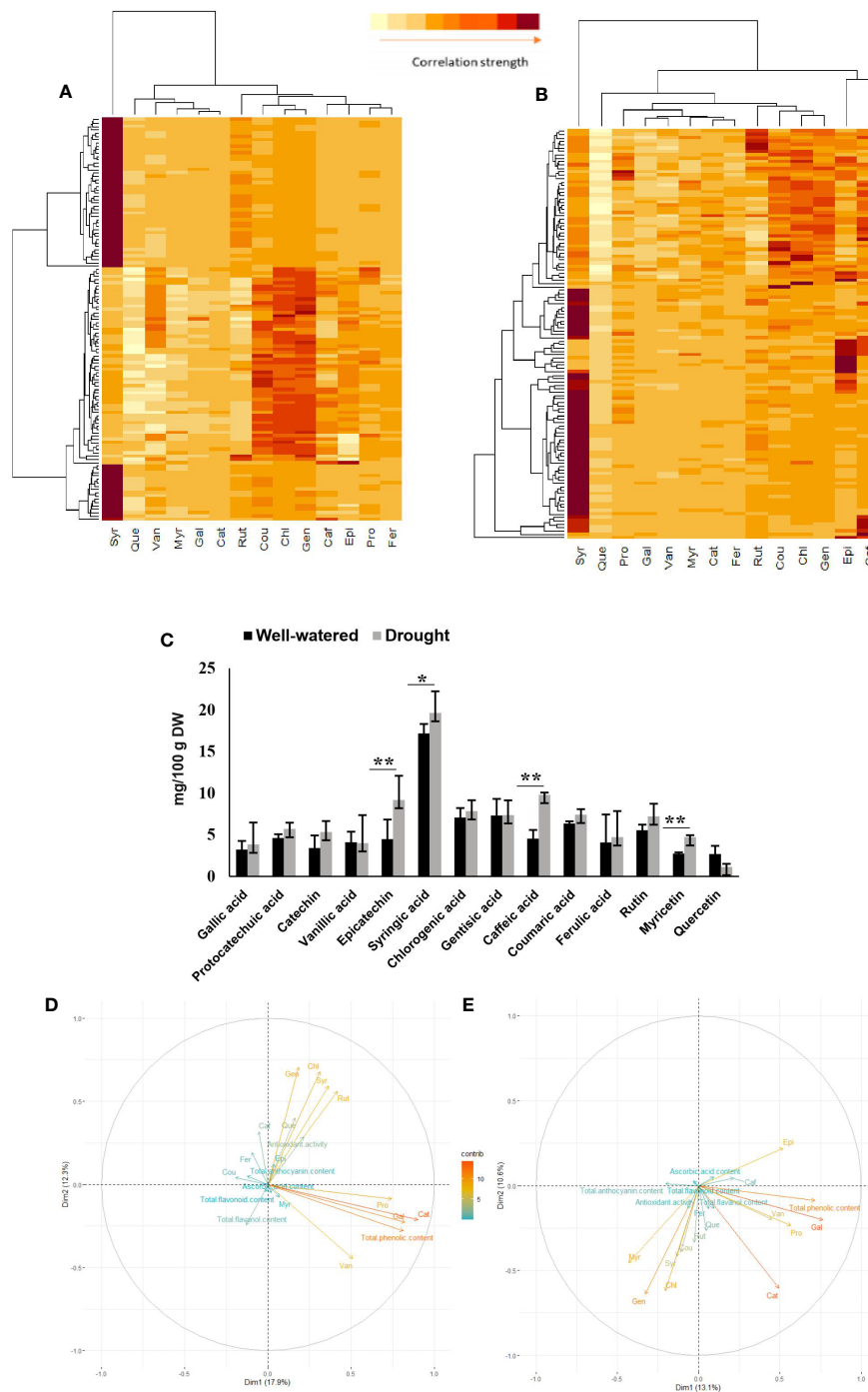


FIGURE 3

Relation of antioxidant activity and phenolics in rapeseed (*Brassica napus*) in the well-watered and drought stress conditions. Heatmap visualization of relative differences of phenolic compounds accumulation among rapeseed accessions in the well-watered condition (A) and drought stress condition (B). Each accession is visualized in a single row and each phenolic compounds is represented by a single column. The dark color indicates high abundance, and light indicates low relative phenolic compounds contents. (C) Average amount of 14 phenolic acids and flavonoids across the rapeseed diversity panel are compared between two conditions (well-watered and drought stress) with Student's t-tests (* $P < 0.05$ and ** $P < 0.01$). Error bars = SE. (D) Principal component analysis (PCA) plot of variables (phytochemical traits, antioxidant activity and phenolic compounds) in the rapeseed accessions in the well-watered condition. (E) Principal component analysis (PCA) plot of variables (phytochemical traits, antioxidant activity and phenolic compounds) in the rapeseed accessions under drought stress condition.

significant change in Δk was observed when K increased from 6 to 7. Overall iterations of the ΔK that had a much higher likelihood for $K = 7$ suggests the presence of seven main subpopulations and one mixed population (Figure 4D).

The critical r^2 value from which the genome-wide LD decayed was estimated at 0.1 (Figure 4D). This revealed large differences in LD decay between different chromosomes, with LD extending from 102 Kbps on chromosome A10 up to 954 Kbps on A08 (Table 2).

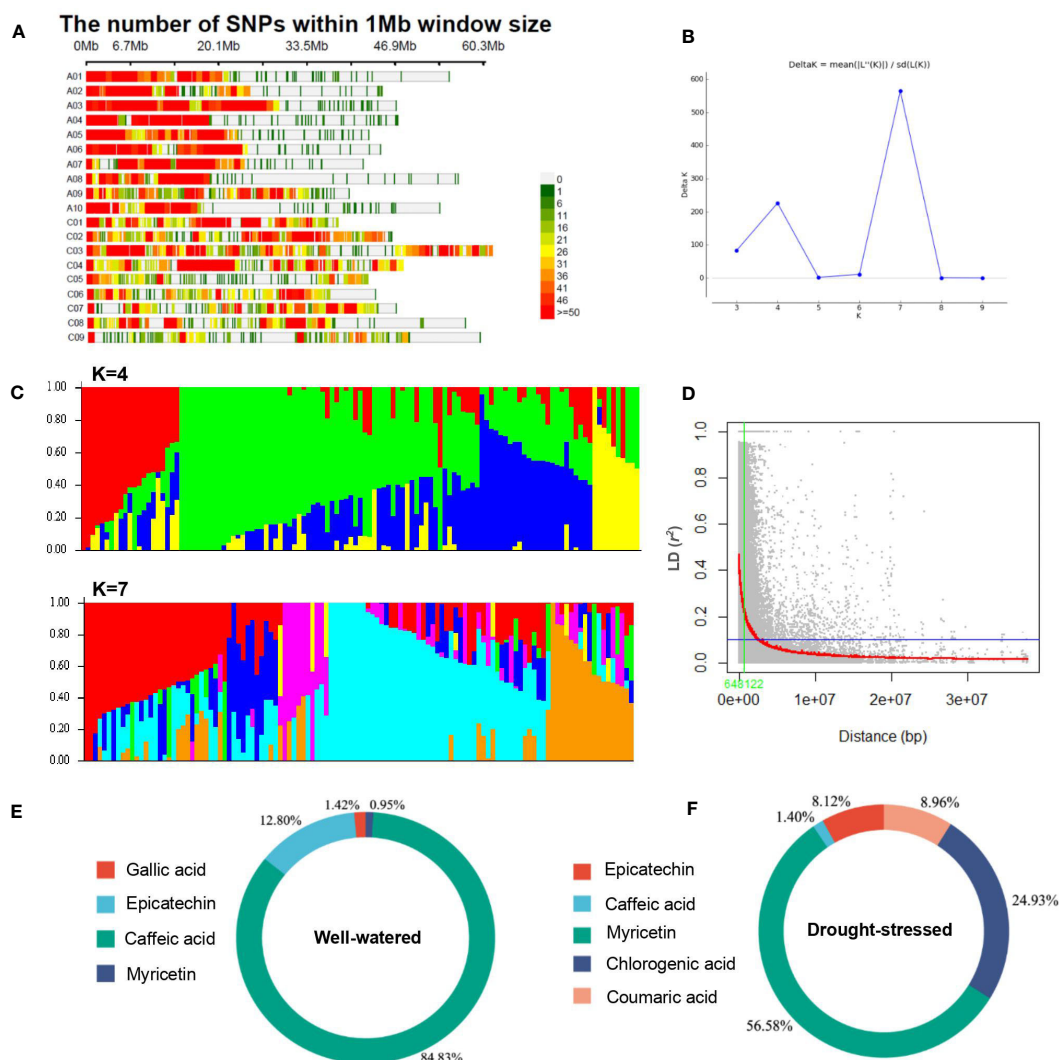


FIGURE 4

Population structure and linkage disequilibrium analysis in the rapeseed (*Brassica napus*) genotypes. (A) SNP density plot chromosome wise representing number of 29,310 SNPs within 1 Mb window size. The horizontal axis shows the chromosome length (Mb); the different color depicts SNP density. (B, C) Delta-K values and population structure analysis and of 119 rapeseed accessions. (B) Delta-K values for different numbers of populations assumed (K) in the STRUCTURE analysis. (C) Population structure analysis, K=4 and K=7. The distribution of the genotypes to different populations is indicated by color code. The number on the y-axis shows the subgroup membership, and the x-axis shows the different genotypes. (D) Decay of linkage disequilibrium (LD) (r^2) as a function of genetic distance (bp) between pairs of loci on all chromosomes in the rapeseed (*Brassica napus*) accessions (n=119). The blue lines indicate the 95th percentile of the distribution of unlinked r^2 , which gives the critical value of r^2 . The LOESS fitting curve (red line) illustrates the LD decay. Summary of significant SNPs detected under well-watered (E) and drought-stressed (F) conditions for phenolic compounds in rapeseed flowers.

The distribution of r^2 with respect to the physical distances over chromosomes is shown in Figure 4D and Table 2. The genome-wide average LD (r^2) was 0.028. The LD values showed variations between chromosomes. The LD varied from 0.022 (chromosomes C03 and C08) to 0.088 (chromosome C05) (Table 2).

Genetic basis of the phenolic acids under drought stress

The MLM using both the PC and K matrix (MLM + K + PC) resulted in effectively controlling type I errors (false positives). Therefore, all subsequent GWAS analyses were performed using

the MLM + K + PC model. A number of 613 significant SNPs ($P < 1.0 \times 10^{-6}$) associated with the phytochemicals and phenolics and RSA in the two irrigation conditions (Supplementary Tables S7, S8), 242 were associated with TPC, RSA, TFDC, TFLC, TAC, AAC, gallic acid, epicatechin, caffeic acid, and myricetin under WW and 371 were associated with RSA, TAC, and TFLC, epicatechin, caffeic acid, myricetin, chlorogenic acid, and coumaric acid under DS conditions (Supplementary Tables S7, S8; Figures 4E, F). The Manhattan and QQ plots for all traits are shown in Figures 5, 6 and Supplementary Figures S3–S5.

GWAS analysis for TPC, TFDC, and TFLC showed that three SNPs, ChrA1:7615643, ChrA1:7669625, and ChrA1:7727474, with P -value = 7.48×10^{-4} were significantly associated with TPC, and 10

TABLE 2 Average distance of linkage disequilibrium (LD) decay ($r^2 < 0.1$) on subgenomes A and C, calculated using 29,311 unique genome-wide SNP markers with minor allele frequency (MAF) ≥ 0.05 , in a collection of 119 rapeseed (*Brassica napus*) genotypes.

	Chromosome	LD decay (kb)	LD (r^2)	SNPs with high LD ($r^2 > 0.8$) (%)
Subgenome A	A01	627.86	0.023	8.76
	A02	708.52	0.025	12.26
	A03	564.58	0.027	5.40
	A04	942.40	0.026	6.78
	A05	653.05	0.027	10.81
	A06	561.67	0.023	9.30
	A07	535.16	0.027	8.00
	A08	953.78	0.027	22.98
	A09	776.33	0.024	9.30
	A10	101.61	0.028	6.61
Subgenome C	C01	103.65	0.027	21.57
	C02	112.52	0.024	1.81
	C03	942.60	0.022	1.04
	C04	102.94	0.023	21.54
	C05	169.42	0.088	1.76
	C06	153.44	0.024	1.22
	C07	175.17	0.024	1.22
	C08	166.58	0.022	7.02
	C09	202.23	0.027	2.99

SNPs on chromosomes A05 (located on the 1838–17079 Kbps) and A06 were significantly associated with TFDC under WW (Supplementary Table S7). Ten SNPs distributed on chromosomes A04, A06, A10, and C04 and two SNPs (ChrA6:21114139 and ChrA10:12986214) on chromosomes A06 and A10 were significantly associated with TFCLC under WW and DS conditions, respectively (Supplementary Tables S7, S8). A total of eight SNPs distributed on chromosomes A02, A05, A09, C04, and C07 were significantly associated with RSA under WW and DS conditions. The Percentage of Variance Explained (PVE) for these SNPs ranged from 13.59% for the SNPs ChrA9:3534018 and ChrC7:3534211 to 17.11% for the SNP ChrA5:21349945 under DS (Supplementary Table S8). GWAS analysis showed that three SNPs distributed on chromosomes A03, C03, and C08 and eight on chromosomes A03, C03, and C07 were significantly associated with the TAC under WW and DS, respectively (Supplementary Tables S7, S8). In the WW condition, two SNPs, ChrA3:9165370 (PVE = 14.24%) and ChrC8:34299868 (PVE = 16.34%), were significantly associated with AAC (Supplementary Table S7).

GWAS for phenolic compounds showed that three SNPs, ChrA10:16316355, ChrC4:1857195, and ChrC4:1891703, with relatively high PVE were significantly associated with gallic acid in the WW condition (Supplementary Table S7; Supplementary Figure S3A). A total of 179 SNPs distributed on chromosomes A01–A10, C02–C07, and C09 were significantly associated with caffeic acid in the WW condition (Supplementary Table S7; Figures 5A, B).

The PVE ranged from 26.83% for the SNP ChrA4:8622635 to 33.50% for the SNP ChrA4:378678. The top 2 hot spots located on the 23–45505 Kbps region of chromosome A02 and 255–43967 Kbps interval of chromosome C02 comprised 35 and 38 SNPs, which covered 50.77% of phenotypic variation on their respective chromosomes (Supplementary Table S7; Figures 5C–F). Among the linked SNPs, ChrA4:10190551 showed the most significant association with caffeic acid under the WW condition (Supplementary Table S7). This linked SNP explained 33.01% of the phenotypic variation of caffeic acid. SNP ChrA4:378678 (PVE = 33.5%) and SNP ChrC4:8636542 (PVE = 32.44%) were also significantly associated with caffeic acid (Supplementary Table S7). Under the DS condition, five SNPs were significantly associated with the caffeic acid with PVE ranging from 24.71% to 30.96% (Supplementary Table S8; Supplementary Figure S3D).

Under the WW condition, the six SNPs including *Bn-A02-p27755843*, *Bn-A02-p25105527*, *Bn-A02-p26154897*, *Bn-scaff_17522_1-p1497074*, *Bn-A09-p28758854*, and *Bn-A09-p35981539* showing significant association with caffeic acid were further used to determine the effects of their individual alleles on caffeic acid (Figures 5G–L). The number of alleles for each of these six SNPs in the whole rapeseed population varied between 2 and 3. The two AG and GG alleles showed a significant difference in the regulation of caffeic acid content. For instance, the allele GG for the SNPs *Bn-A02-p27755843*, *Bn-scaff_17522_1-p1497074*, *Bn-A09-p28758854*, and *Bn-A09-p35981539* was associated with higher

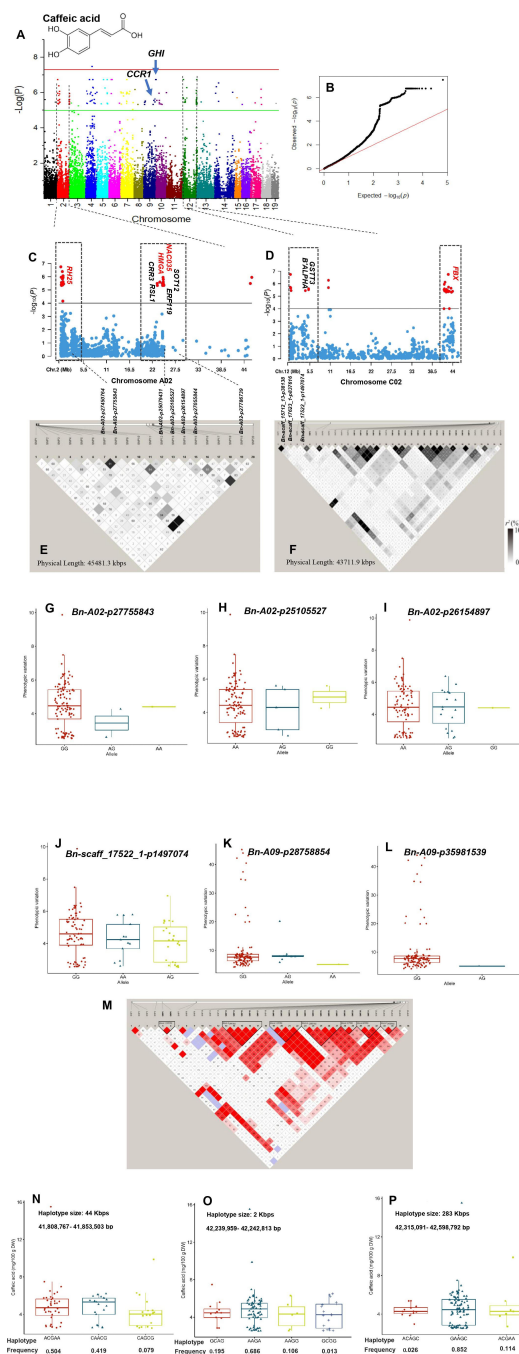


FIGURE 5

GWAS of caffeic acid using the 119 rapeseed (*Brassica napus*) genotypes. (A) Manhattan plot displaying the GWAS result of caffeic acid content in 19 chromosomes (1-10 stand for rapeseed chromosome of A01-A10 and 11-19 stand for rapeseed chromosome of C01-C09 at the horizontal axis) under well-watered condition. SNPs on different chromosomes are denoted by different colors. (B) The corresponding quantile-quantile (QQ) plot for GWAS of caffeic acid content under well-watered condition. (C) Locus zoom plot for caffeic acid associations in the chromosome A02. (D) Locus zoom plot for caffeic acid associations in the chromosome C02. (E) A representation of pairwise r^2 value (displayed as percentages) among polymorphic sites of chromosome 2 (A02) for caffeic acid. (F) A representation of pairwise r^2 value (displayed as percentages) among polymorphic sites of chromosome 12 (C02) for caffeic acid. Allele-effect analysis for six significant SNPs including Bn-A02-p27755843 (G), Bn-A02-p25105527 (H), Bn-A02-p26154897 (I), Bn-scaff_17522_1-p1497074 (J), Bn-A09-p28758854 (K), and Bn-A09-p35981539 (L). The box plot depicts the number of the alleles for each of the six significant SNPs in 119 rapeseed accessions, and the contribution of these alleles to the phenotypic variation observed for caffeic acid level. (M) Haploview plot of the thirty-six significant SNPs on chromosome C02 for caffeic acid under well-watered condition. The coloring of the boxes is based on the scores of two estimates, logarithm of odds (LOD) and LD coefficient (D'). Strong LD (LOD>2 and D'=1) is indicated in red. A decrease in the intensity of red indicates lower LOD and D' values. Regions with weak LD (LOD>2 and 0.21<D'<1) are shown in blue. Regions with no LD (LOD<2 and D'<1) are shown in white. The number within each box indicates the D' statistic value multiplied by 100. Allelic effects of different haplotype blocks associated with caffeic acid in size 44 Kbps (N), 2 Kbps (O), and (P) 283 Kbps. Boxplots indicate the phenotype values corresponding to the different haplotype groups.

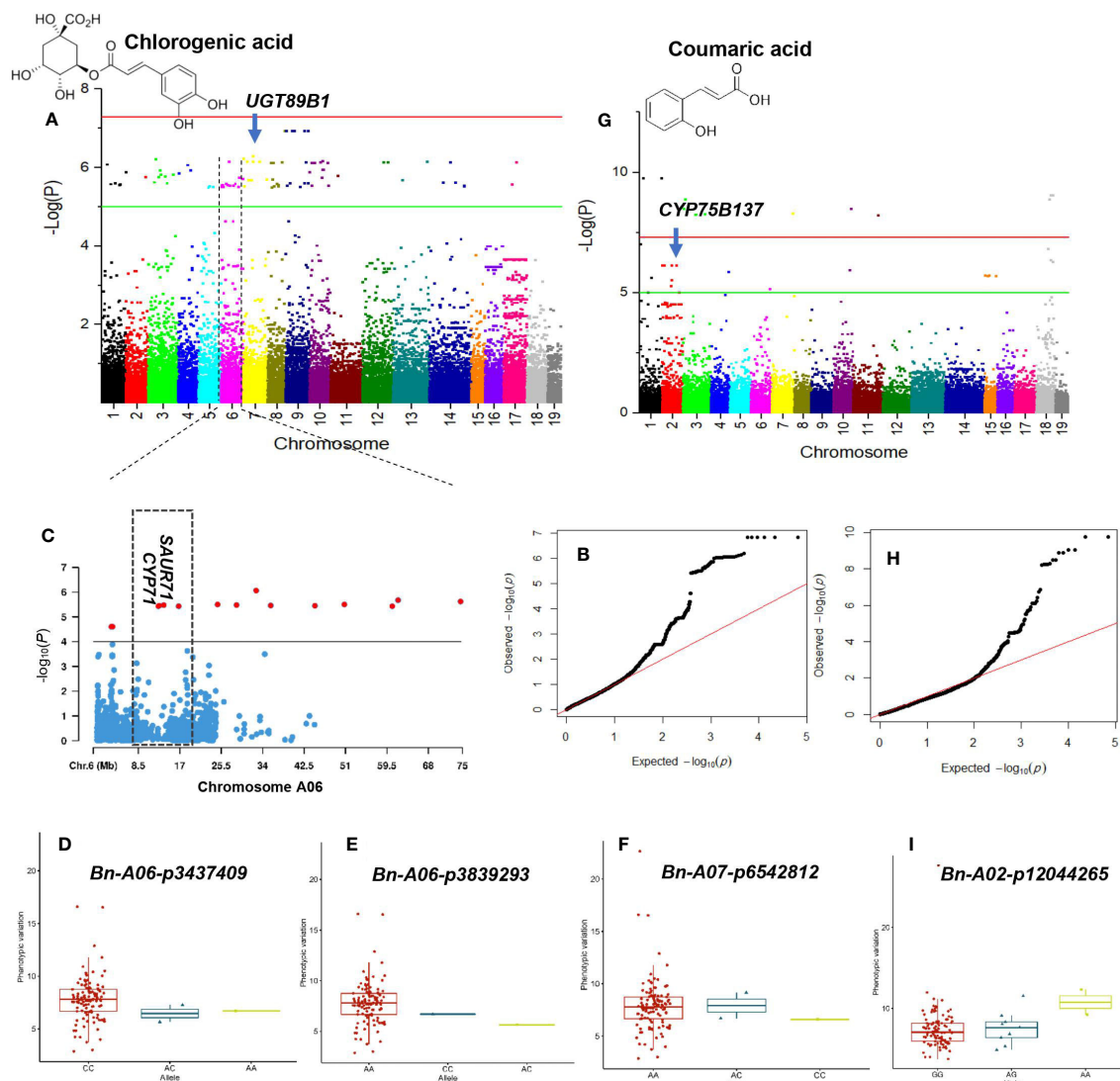


FIGURE 6

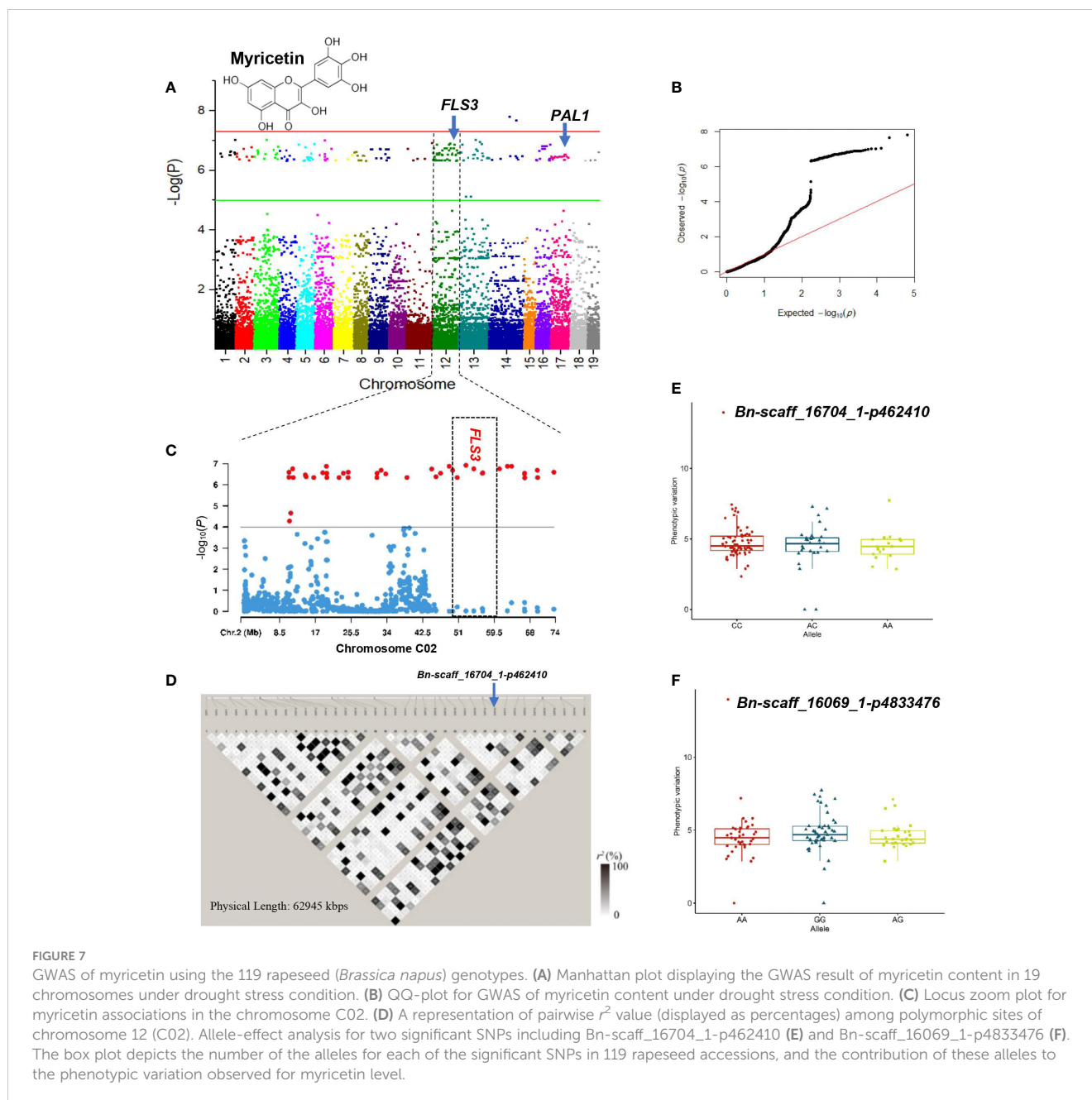
GWAS of chlorogenic acid and coumaric acid using the 119 rapeseed (*Brassica napus*) genotypes. (A) Manhattan plot displaying the GWAS result of chlorogenic acid content in 19 chromosomes (1–10 stand for rapeseed chromosome of A01–A10 and 11–19 stand for rapeseed chromosome of C01–C09 at the horizontal axis) under drought-stressed condition. SNPs on different chromosomes are denoted by different colors. (B) The corresponding quantile–quantile (QQ) plot for GWAS of chlorogenic acid content under drought stress condition. (C) Locus zoom plot for chlorogenic acid associations in the chromosome A06. Allele–effect analysis for significant SNPs including Bn-A06-p3437409 (D), Bn-A06-p3839293 (E), and Bn-A07-p6542812 (F). The box plot depicts the number of the alleles for each of the four significant SNPs in 119 rapeseed accessions, and the contribution of these alleles to the phenotypic variation observed for chlorogenic acid level. (G) Manhattan plot displaying the GWAS result of coumaric acid content in 19 chromosomes under drought stress condition. (H) QQ–plot for GWAS of coumaric acid content under drought stress condition. Allele–effect analysis for significant SNP Bn-A02-p12044265 (I). The box plot depicts the number of the alleles for the significant SNP in 119 rapeseed accessions, and the contribution of these alleles to the phenotypic variation observed for coumaric acid level.

caffeic acid, whereas AG for these SNPs except *Bn-A09-p28758854* reduced this compound (Figures 5G–L).

The 36 significant SNPs associated with caffeic acid on ChrC2 under the WW condition were used for the identification of haplotype blocks on ChrC2 (Figures 5M–P). Three haplotype blocks (blocks 2–4) with strong LD ($D' > 0.9$) were mapped to 44, 2, and 284 Kbps, respectively (Figures 5M–P). In haplotype block 2 (41,808,767–41,853,503 bp), three haplotype alleles were identified that showed significant differences in the phenotypes of caffeic acid. Of these haplotypes, the haplotype alleles ACGAA ($f = 0.504$) and CAACG ($f = 0.419$) were associated with higher caffeic acid

(Figure 5N). In haplotype block 3 (42,239,959–42,242,813 bp), four haplotype alleles showed significant phenotypic variation for caffeic acid. The haplotype allele AAGA with $f = 0.686$ was associated with higher caffeic acid (Figure 5O). Additionally, in haplotype block 4 (42,315,091–42,598,792 bp), the haplotype allele GAAGC regulated higher caffeic acid level ($f = 0.686$) (Figure 5P).

In GWAS analysis, 89 significant SNPs were identified for the chlorogenic acid under the DS condition. The proportion of SNPs in total PVE (R^2) ranged from 30.14% to 34.11%, with a mean of 30.92% (Supplementary Table S8; Figures 7A, B). Of them, 28 SNP–chlorogenic acid associations were identified on chromosome A06



that were located between 126616 and 74617 Kbps (Figure 7C). SNPs ChrA6:615333, ChrA6:12732383, and ChrA6:13700378 were the most significant SNPs associated with chlorogenic acid under the DS condition (Supplementary Table S8). Among the significant SNPs, 12 located on chromosome A09 were correlated with chlorogenic acid (Supplementary Table S8). The three SNPs on chromosome A06 and four on chromosome A07 were associated with higher chlorogenic acid content (Figures 7D–F).

Thirty-two SNPs on chromosomes A01–A04, A06, A07, A10, C01, C05, and C08 were significantly associated with coumaric acid under the DS condition (Supplementary Table S8; Figures 7G, H). The PVE of the SNP ranged from 21.49% to 48.87%. SNPs ChrA1:20708470 and ChrA1:73646490 significantly contributed to 46.86% of the PVE of

coumaric acid. Allele AA for SNP *Bn-A02-p12044265* contributed to higher coumaric acid content (Figure 7I).

Genetic basis of the flavonoid content under drought stress

Twenty-seven SNP–epicatechin associations on chromosomes A01, A05, A07–A10, C03–C05, and C07 (Supplementary Table S7) were detected under the WW condition. The PVE for the linked SNPs ranged from 24.22% to 33.01%. The most significant SNP ChrC4:29181029 had a PVE of 27.02% (Supplementary Table S7). Several SNPs associated with epicatechin were located on

TABLE 3 SNPs and candidate genes significantly associated with phenolic compounds integrating genome-wide association and transcriptome studies.

Phenolic compounds	Lead SNP	Position (bp) ^a	Chr ^b	Candidate Gene ^c	Gene Location	Annotation
Drought stress-related genes						
Caffeic acid	Bn-A02-p27755843	444005	A2	LOC125581627	456830.460131	DEAD-box ATP-dependent RNA helicase 25-like (RH25)
	Bn-A02-p25105527	23180627	A2	LOC125587888	23161297-23165111	E3 ubiquitin-protein ligase (RSL1)
	Bn-A02-p25079431	23154160	A2	LOC125587885	23160045-23161004	Probable NAD(P)H dehydrogenase subunit CRR3, chloroplastic (CRR3)
	Bn-A02-p25105527	23180627	A2	LOC106424243	23185689-23190046	HMG-Y-related protein A (HMGA)
	Bn-A02-p26154897	23778761	A2	LOC106396214	23766736-23769871	NAC domain-containing protein 35 (NAC035)
	Bn-A02-p27055844	24377661	A2	LOC125580157	24382569-24383687	Cytosolic sulfotransferase 12-like (SOT12)
	Bn-A02-p27786739	24744845	A2	LOC111202999	24726755-24728300	Ethylene-responsive transcription factor (ERF119)
	Bn-scaff_15712_13-p38138	372866	C2	LOC106424063	370212-372418	Serine/threonine protein phosphatase 2A 57 kDa regulatory subunit B' (B'ALPHA)
	Bn-scaff_17623_1-p837616	455566	C2	LOC106424074	454162-458783	Glutathione S-transferase T3 (GSTT3)
	Bn-scaff_17522_1-p1497074	5162738	C2	LOC106443306	5088531-5091807	F-box protein At1g30790-like (FBX)
Chlorogenic acid	Bn-A06-p3437409	12615333	A6	LOC106347668	12638041-12641569	Peptidyl-prolyl cis-trans isomerase (CYP71)
	Bn-A06-p3839293	13700378	A6	LOC106433452	13684571-13686129	Auxin-responsive protein (SAUR71)
Metabolite biosynthetic genes						
Caffeic acid	Bn-A09-p28758854	26690641	A9	LOC106367237	26877113-26878867	Cinnamoyl-CoA reductase 1 (CCR1)
	Bn-A09-p35981539	33104863	A9	LOC106412109	33233120-33235006	Chalcone-flavanone isomerase (CHI)
Chlorogenic acid	Bn-A07-p6542812	20717392	A7	LOC106430142	20512963-20517193	Flavonol 3-O-glucosyltransferase (UGT89B1)
Myricetin	Bn-scaff_16704_1-p462410	56681393	C2	LOC106382906	56510517-56512288	Flavonol synthase 3 (FLS3)
	Bn-scaff_16069_1-p4833476	52702475	C7	LOC106388514	53110580-53113485	Phenylalanine ammonia-lyase 1-like (PAL1)
Coumaric acid	Bn-A02-p12044265	16769326	A2	LOC106436650	16093088-16095189	Flavonoid 3'-monooxygenase (CYP75B137)

^a Position in base pairs for the lead SNP according to version 4 of the rapeseed reference sequence, ^b Chromosome, ^c A plausible candidate gene near these loci.

chromosome C05 and covered 43.15% of phenotypic variation. Three significantly linked SNPs, ChrC5:162937, ChrC5:163550, and ChrC5:163557, identified in an interval of 163 Kbps on chromosome C05 showed P -value = 9.34×10^{-6} (Supplementary Table S7). Under the DS condition, 29 SNPs were significantly

associated with the epicatechin on chromosomes A01–A08, A10, C01–C05, C08, and C09 (Supplementary Table S8). The PVE for these SNPs ranged from 27.88% for SNP ChrC1:52661467 to 34.53% for SNP ChrC9:63810360. SNPs ChrC5:43670426, ChrC9:63810360, and ChrA7:28726499 showed stronger GWAS

P-values in relation to epicatechin under the DS condition (Supplementary Table S8).

Association analysis at a threshold of $P < 1.0 \times 10^{-6}$ showed that 202 loci were associated with myricetin under the DS condition (Supplementary Table S8; Figures 6A, B). One hot spot located on the 10688–73679 Kbps region of chromosome C02 comprising 40 loci had a PVE of 34.77% (Supplementary Table S8; Figures 6C, D). Of the significant SNPs, SNP ChrC4:59722505 was the major locus controlling myricetin with a PVE of 40.98%. SNPs ChrC3:45787340 and ChrA6:43739339 explained 37.16% and 37.09% of the variation of myricetin, respectively (Supplementary Table S8). Alleles CC, AC, and AA for SNPs *Bn-scaff_16704_1-p462410* and *Bn-scaff_16069_1-p4833476* showed similar effects on myricetin content (Figures 6E, F).

Characterization of the drought stress-related candidate genes

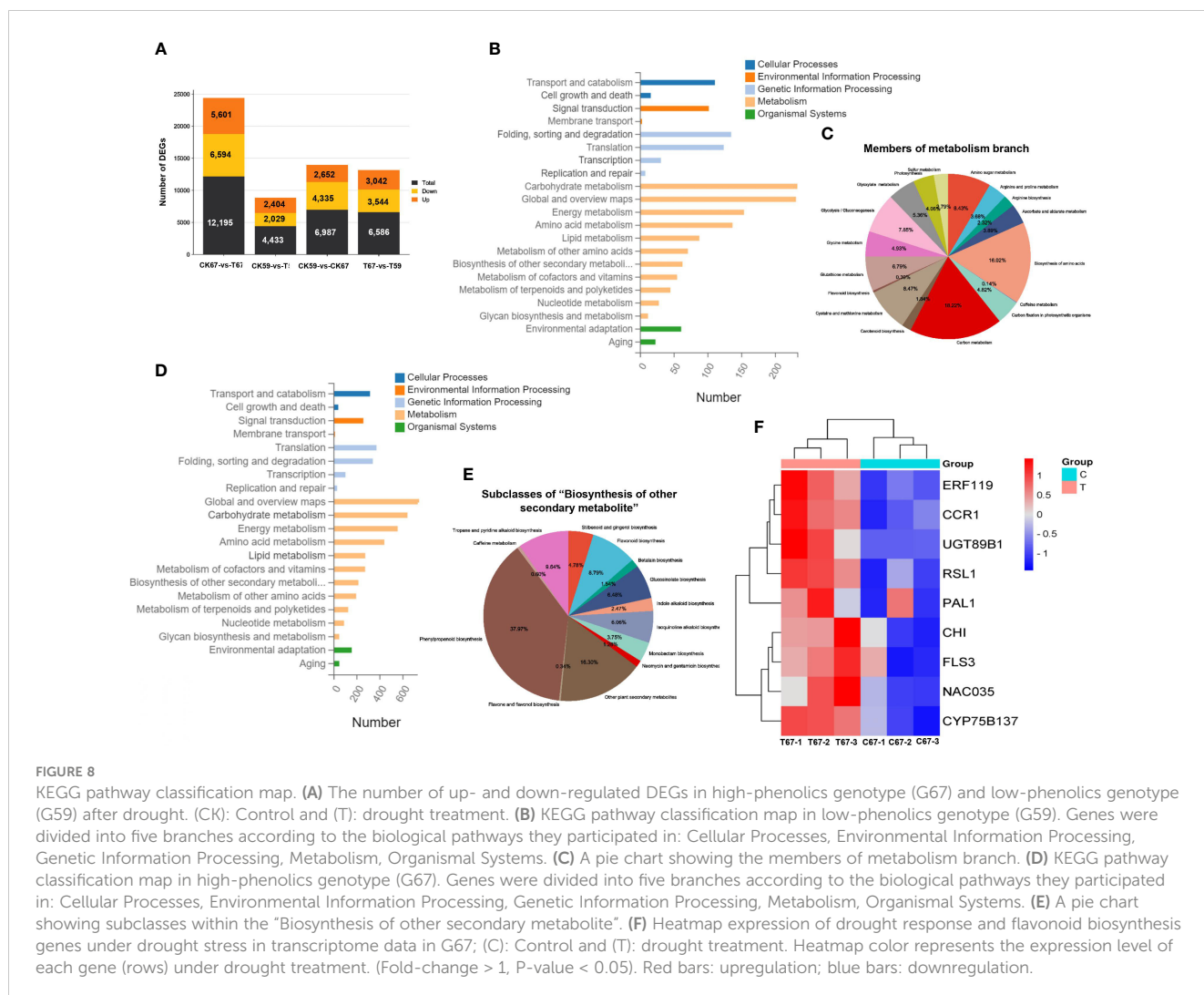
Two LD GRs showed 20 and 36 significant SNP–caffeic acid associations located on chromosomes A02 and C02, respectively, under the WW condition (Figures 5E, F). In the first GR, a highly significant LD ($r^2 = 0.9$) was found between SNPs ChrA2:426966 (tag-SNP #7) and ChrA2:444005 (tag-SNP #8). The position of the CG *DEAD-BOX ATP-DEPENDENT RNA HELICASE 25-LIKE (RH25)* was close to SNP ChrA2: 444005 (Table 3; Figures 5C, E). The *DEAD-box* genes play crucial roles in plant abiotic stress tolerance via regulating some stress-induced pathways. The *RH* genes are negative regulators of ABA signaling via inhibition of *PP2CA* activity. The functions of these RNA helicases influence the plant adaptive response (Baruah et al., 2017). The lead SNP ChrA2:23180627 was close to *E3 UBIQUITIN LIGASE (RSL1)* as a key regulator of hormone signaling (Table 3; Figures 5C, E). E3 ligases that transfer phytohormone abscisic acid (ABA) signals regulate plant responses to different abiotic stresses (Wang et al., 2022). SNP ChrA2:23154160 was close to *HMG-Y-RELATED PROTEIN A (HMGA)*. High-mobility group (HMG) protein is an abundant non-histone chromosomal protein that promotes pollen tube development and regulates the expression of stress response genes (Xu et al., 2021). Several regulatory genes including TFs contribute to abiotic stress responses via regulating downstream stress-responsive genes (Joshi et al., 2016). We identified that the position of SNPs ChrA2:23778761 and ChrA2:24744845 was close to the position of TF CGs *NAC DOMAIN-CONTAINING PROTEIN 35 (NAC035)* and *ETHYLENE-RESPONSIVE TRANSCRIPTION FACTOR-LIKE (ERF119)*, respectively (Table 3; Figures 5C, E). The TF family members involved in ABA-independent [ERF, bHLH, and NAM/ATAF1,2/CUC2 (NAC)] pathways have been upregulated in *B. napus* (Ke et al., 2020). The TFs of this family interact with specific cis-elements, and their overexpression confers stress tolerance (Yao et al., 2021). Moreover, SNP ChrA6:13700378 was close to *AUXIN-RESPONSIVE PROTEIN (SAUR71)*.

In GR2, a significantly complete LD r^2 of 1 was found between 16 SNPs on chromosome C02 under the WW condition (Figures 5D, F). Highly significant LD ($r^2 > 0.6$) was found between SNPs ChrC2:5162738 (tag-SNP #5) and ChrC2:5172871 (tag-SNP #6) and between ChrC2:10513865 (tag-SNP #7) and

ChrC2:10523601 (tag-SNP #8). SNP ChrC2:5162738 was 71 Kbps upstream of *F-BOX PROTEIN (FBX)* (Table 3; Figures 5D, F). F-box protein, as a major subunit of the Skp1, Cullins, F-box proteins (SCF) complex (a multiprotein E3 ubiquitin ligase), plays important roles in abiotic stress responses via the ubiquitin pathway (Song et al., 2023). The lead SNP ChrA6:12615333 was close to the *PEPTIDYL-PROLYL CIS-TRANS ISOMERASE (CYP71)* (Table 3; Figure 7C). Peptidyl-prolyl isomerase regulates protein folding and maturation of newly synthesized proteins under cell stress conditions (Hanes, 2015).

Genes involved in phenolic acid and flavonoid biosynthetic pathways

Our GWAS analysis identified SNP ChrA9:26690641 for caffeic acid in the WW condition that was located at 187 Kbps downstream of *CINNAMOYL-COA REDUCTASE 1 (CCR1)* gene. The *CCR1* gene is a potential control point in the phenylpropanoid pathway (Table 3; Figure 5A). *CCR1* involved in lignification uses feruloyl-CoA as a substrate and balances the abundance of phenolic compounds (Zhou et al., 2010). The linked SNP ChrC7:52702475 identified in the DS treatment was 409 Kbps downstream of *PHENYLALANINE AMMONIA-LYASE 1-LIKE (PAL1)* (Table 3; Figure 6A). *PAL* is the first enzyme in the phenylpropanoid pathway that catalyzes the trans-elimination of ammonia from l-phenylalanine to form trans-cinnamic acid, a precursor of lignins, phenols, flavanoids, and coumarins (Kong, 2015). SNP ChrA9:33104863 identified in the WW condition was located near *CHALCONE-FLAVANONE ISOMERASE (CHI)*, a gene for flavonoid biosynthesis. Chalcone isomerase catalyzes the isomerization of naringenin chalcone into 5,7,4'-trihydroxyflavanone, which is considered to be the earliest intermediate with the flavonoid core (Ni et al., 2020). *CHI* genes co-regulate with other flavonoid-related genes, such as flavonol 3-O-glucosyltransferase, flavonol synthase, and flavonoid 3'-monooxygenase (Xu et al., 2023). Our GWAS results indicated that the lead SNP ChrA7:20717392 identified as a linked SNP in the DS treatment was 200 Kbps upstream of *FLAVONOL 3-O-GLUCOSYLTRANSFERASE (UGT89B1)* (Table 3; Figure 7A). Flavonol 3-O-glucosyltransferase catalyzes the glycosylation of either anthocyanidins or flavonols and is considered as a key enzyme for flavonoid modification (Dai et al., 2017). Glycosylation is usually the last step in the flavonoid biosynthetic pathway, which allows and enables plants to better cope with biotic and abiotic stress including the decrease of toxicity and side effects with the help of glycosylated phenolic compounds (Zhang et al., 2013). The LD GRs were estimated for 39 significant SNP–myricetin associations mapped on chromosome C02 in the DS treatment (Figure 6D). Significant and complete LD ($r^2 = 1$) was found between the two SNPs ChrC2:19721386 (tag-SNP #10) and ChrC2:22620484 (tag-SNP #11), and a highly significant LD ($r^2 > 0.9$) was found between SNP ChrC2:62628403 (tag-SNP #32) and SNP ChrC2:63618405 (tag-SNP #33). The LD plot showed that SNP ChrC2:56681393 was 200 Kbps upstream of *FLAVONOL SYNTHASE 3 (FLS3)* under drought (Table 3; Figures 6C, D).



FLS genes belong to the dioxygenases that catalyze the oxidation of dihydroflavonol to produce flavonol (Wang et al., 2021). The lead SNP ChrA2:16769326 identified as a linked SNP under drought was 675 Kbps upstream of *FLAVONOID 3'-MONOOXYGENASE* (*CYP75B137*).

Transcriptomes of the high-phenolic (G67) and low-phenolic (G59) genotypes

Among the 119 genotypes evaluated in the two conditions, G67 and G59 showed the highest and lowest phenolic contents, respectively. To identify genes potentially involved in drought and flavonoid biosynthesis pathway, we conducted an RNA-seq workflow on flower extracts of both G67 and G59 genotypes. In total, we obtained 90,573,446 clean reads, of which 79.33% of the reads mapped to the *B. napus* reference genome. DEGs were identified with the criteria of $q \leq 0.05$ and \log_2 fold change ≥ 1 . The high-phenolic genotype G67 and low-phenolic G59 showed a

substantial difference in their basal transcriptomes; a comparison of untreated G67 and G59 samples led to the identification of 6,987 DEGs, including 2,652 upregulated DEGs and 4,335 downregulated DEGs (Figure 8A). The high-phenolic genotype G67 had a total of 12,195 DEGs, of which 5,601 were upregulated. By contrast, the low-phenolic genotype had only 4,433 DEGs. Under DS, G67 showed more upregulated DEGs compared to G59 (T67 vs. T59; Figure 8A). We assumed that G67 might have numerous drought tolerance-related genes that are primed for response to DS.

To gain insight into the biological processes that were commonly or uniquely overrepresented between G59 and G67 in response to DS, we analyzed Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment of DEGs. The upregulated common biological processes included cellular processes, environmental information processing, genetic information processing, and metabolism, organismal systems (Figures 8B–E). The upregulated DEGs in G67 were specifically enriched in "phenylpropanoid biosynthesis," "flavone and flavonol biosynthesis," and "other plant secondary metabolites" (Figure 8E).

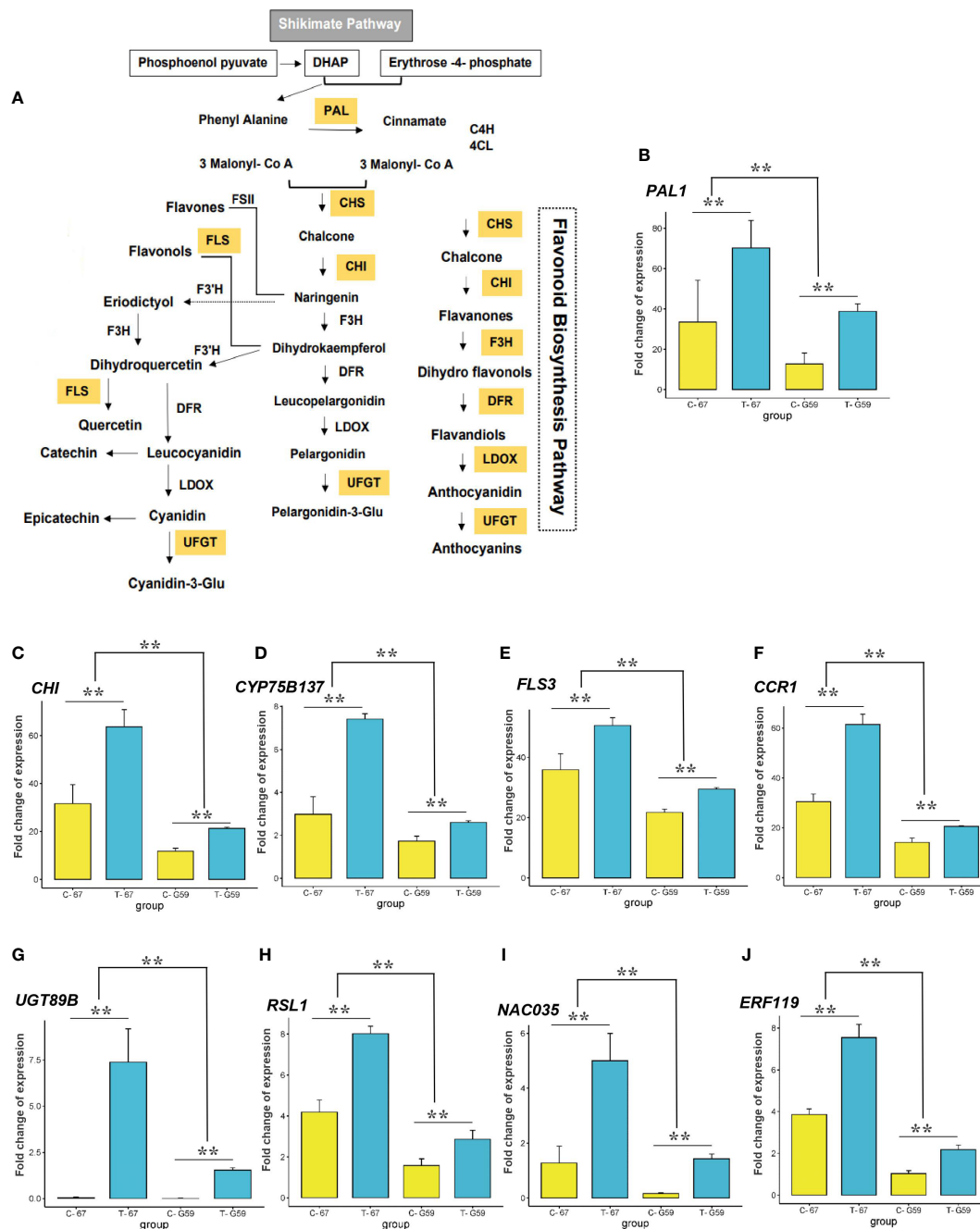


FIGURE 9

Expression variation of candidate genes from GWAS and transcriptome analysis. (A) Schematic diagram of the flavonoid biosynthesis pathway. PAL phenylalanine ammonia lyase, C4H cinnamate 4-hydroxylase, 4CL 4-coumarate-CoA ligase, CHS chalcone synthase, CHI chalcone isomerase, F3H flavanone-3-hydroxylase, F3'H flavonoid 3'-hydroxylase, FLS flavonol synthase, DFR dihydroflavonol reductase, LDOX leucoanthocyanidin dioxygenase, UFGT UDP-glucose flavonoid 3-O-glucosyl transferase are key enzymes involved in the pathway. (B–J) Expression (Fold-change) of PAL1, CHI, CYP75B137, FLS3, CCR1, UGT89B, RSL1, NAC035, and ERF119 in high-phenolics genotype (G67) and low-phenolics genotype (G59) at the flowering stage under drought treatment (T), detected by RNA-seq. Data are average values with standard deviation ($n = 3$ varieties). ** Indicate significant difference relative to control (C) (t-test: $**P < 0.01$).

Integration of the GWAS and transcriptome data

Our aim was to identify CGs differentially expressed after drought that were detected in the GWAS. Among the 60 genes detected in the GWAS, 50 were expressed in rapeseed flower based

on transcriptome data. Of these genes, 18 were involved in stress-induced pathways, phenylpropanoid pathway, and flavonoid modifications (Figure 8F; Table 3). Analysis of the expression pattern of genes showed that the early unbranched part of the flavonoid biosynthesis pathway is encoded by *CHI*, *PAL*, and *CYP75B137* genes (Figure 9A). A significant increase in *PAL* was

observed under drought (2.09-fold). The level of induction of gene expression was genotype-dependent. The relative expression of *PAL* was much higher in the high-phenolic genotype (G67) than in low ones (G59) (Figure 9B). The *CHI* gene also exhibited higher expression (2.01-fold) compared to the control (Figure 9C). The *CYP75B137* gene also showed a higher expression (2.48-fold) compared to the control (Figure 9D). The biosynthesis of flavonoids in the later stages is encoded by *FLS1/FLS3*, *CCR1*, and *UGT89B1* genes. The expression of the *FLS3* gene was higher (1.40-fold) in drought than in the control (Figure 9E). Additionally, under DS, the expression level of *CCR1* was much higher (2.01-fold; Figure 9F). The transcript level of *UGT89B1* was elevated more than 200-fold in the high-flavonoid genotype under water deficit conditions (Figure 9G). The transcript level of flavonoid biosynthesis genes was directly related to the flavonoid content. As shown in Figures 9B–G, the expression level of all flavonoid-related genes was much higher in G67 than that in G59. Transcriptional analysis showed that the genes were differentially expressed between the two genotypes, of which the *RSL1*, *NAC035*, and *ERF119* genes were upregulated under drought (Figures 9H–J).

Discussion

Phenolic compounds in *B. napus* flowers

In recent years, phenolic compounds have been intensively investigated because of their potential health-promoting effects (Cosme et al., 2020; Rahman et al., 2021). Phenolic compounds are considered in *Brassica* as the most abundant group of secondary metabolites known as natural antioxidants (Bennouna et al., 2021). This group of compounds is able to protect cells against the negative effects of ROS, lipid peroxidation, protein denaturation, DNA damage, and consequently reducing the oxidative damage (Jideani et al., 2021; Juan et al., 2021). Moreover, phenolic compounds have an important role in pigmentation, growth, reproduction of plants, and elevated defenses against the adverse effects of environmental stresses (Cheyner et al., 2013; Naik and Al-Khayri, 2016). In the present study, we found that *B. napus* flowers are a rich source of total phenolic, flavonoid, and flavonol content, anthocyanin, ascorbic acid, and natural antioxidants. The major dihydroxybenzoic acid in rapeseed flowers is gentisic acid (2,5-dihydroxybenzoic acid). In addition, caffeic acid (3,4-dihydroxy cinnamic acid) and epicatechin identified as abundant compounds in our study are the most important compounds of hydroxycinnamic acid and flavonoids. Notably, rapeseed flowers offer phenolic and flavonoid in much higher quantities than seed (Jun et al., 2014; Salami et al., 2023). Our results indicated that *B. napus* flowers can serve as natural antioxidants as characterized by a higher inhibition of DPPH (93%), which was higher than 83.17% and 89.25% DPPH activity in *B. juncea* (Lee et al., 2020) and *B. nigra*, respectively (Olgun et al., 2017). Phenolic compounds are often perceived favorably by consumers as healthy natural foods. Therefore, rapeseed flower can be used as a source of phenolic compounds with bioactive potential, which can be

applied in the food sector, as food and as source of natural ingredients that can be an interesting alternative for consumers looking for product innovations.

Drought stress responses in *B. napus* flowers

DS affects flowering and flower metabolism in rapeseed (Tesfamariam et al., 2010; Kirkegaard et al., 2018). There are various reasons for drought-induced damage of the flowering stage, including decreased water potentials in the floral tissues such as pollen, female gametophyte, and stigma and decreased assimilate and nitrogen supplies that ultimately affect the rapeseed yield (Fard et al., 2018; Zhu et al., 2021; Batool et al., 2022). Another phenomenon can occur, which is that drought-induced hormonal changes may cause sterility (Lalay et al., 2022). One of the adverse effects of DS at the molecular level is the accelerated production of ROS that disrupts the physiological and biochemical processes in plants (Saha et al., 2022). Plants have endogenous mechanisms for adapting to ROS production through enhancement of the natural activities of the antioxidant enzymes and nonenzymatic compounds (Fujita and Hasanuzzaman, 2022; Jan et al., 2022). Among the nonenzymatic antioxidant components, ascorbic acid, anthocyanin, and phenolic compounds are effective in resistance and adaptive features in plants against DS through suppression of singlet oxygen and inhibition of ROS-generating enzymes (Ashraf et al., 2022; Celi et al., 2023; Wagay et al., 2023). In this study, we monitored the changes in phenolic compounds between two contrasting irrigation conditions and evaluated their contributions to drought tolerance. The phenolics, flavonoids, and phenolic compounds were increased under drought condition in our rapeseed population. A significant enhancement of epicatechin and myricetin contents was observed in rapeseed flowers under DS conditions. The results of PCA and stepwise regression showed that myricetin had strong associations with antioxidant activity under DS conditions. These results suggest that epicatechin and myricetin play an important role in enhancing drought tolerance in *B. napus*. Similarly, Okafor et al. (2022) reported that flavonoids had a stronger antioxidant activity than phenolic acids. Flavonoids like myricetin help the formation of less reactive phenoxyl radicals by donation of electrons or hydrogen and the ability to chelate transition metals like iron (Semwal et al., 2016). The iron-chelating properties and RSA of flavonoids are closely related; iron is chelated by the flavonoid and ROS are subsequently scavenged by the flavonoids (Kejik et al., 2021; Agraharam et al., 2022). Transcriptomic and metabolomic analyses have shown that flavonoid accumulation is important to improve drought tolerance in *B. napus* (Xiong and Ma, 2022). An increase in phenolic content in rapeseed tested under water scarcity has been shown in several studies (Son et al., 2020; Okafor et al., 2022; Morsi et al., 2023). Information on the response of phenolic compounds and their antioxidant activity to DS will be essential for producers to understand the advantages of this crop for cultivation under water scarcity conditions.

Stable SNPs identified for phytochemicals and phenolics

Information on DNA markers linked with polyphenol components contributing to DS helps screening varieties for drought tolerance improvement. The stability of linked markers over different environmental conditions and pleiotropic effects are two important factors for the use of markers in MAS programs and improvement of crops. In our GWAS analysis, evaluation of the relation between SNPs and traits showed several common and pleiotropic SNPs. Two SNPs for RSA and one for TFLC that were common between the two irrigation treatments significantly contributed to phenotypic variation of these traits. The SNP Bn-A01-p6333910 had pleiotropic effects on both chlorogenic and coumaric acid content under DS conditions. These stable and pleiotropic SNPs with high contribution to phenotypic variation for phenolic content provide valuable genetic information for the improvement of phenolics and indicate promising GR for MAS in rapeseed breeding programs.

Candidate genes involved in phenolic biosynthesis pathway in *B. napus*

All phenolic compounds are synthesized through the shikimic acid/phenylpropanoid pathway (Kumar et al., 2023; Figure 9A). The synthesis of phenylalanine from chorismic acid is considered the branching point that links the shikimate pathway to the phenylpropanoid pathway (Tohge et al., 2013). The general phenylpropanoid pathway from phenylalanine to p-coumaroyl-CoA is considered a major precursor in the synthesis of both hydroxycinnamic acids and flavonoids (Metsamuuronen and Siren, 2019). The early phenylpropanoid pathway involves three key enzymes: PAL, C4H, and 4CL (Zhang and Liu, 2015; Figure 9A). PAL as a key enzyme in primary metabolism catalyzes the non-oxidative deamination of phenylalanine to cinnamic acid and affects the accumulation of lignin, flavonoids, and hydroxycinnamic acid through regulating the rate of phenylalanine entering the phenylpropanoid pathway (Wu et al., 2017). In this study, the drought-related GWAS results consistent with results of a study in *B. napus* (Ali et al., 2023) showed strong associations between *PAL1* and myricetin under drought conditions (Figure 6A, Supplementary Table S8). Moreover, the expression of *PAL1* increased in our highly accumulated myricetin accession and under drought conditions. Our RNA-seq analysis indicated that the stress-inducible expression of the *PAL1* gene might increase the accumulation of flavonoids and enhance abiotic stress tolerance in rapeseed.

CHI as the first enzyme specific for the flavonoid pathway converts naringenin chalcone to naringenin flavanone, which is a branching point for the formation of several groups of flavonoids (Yonekura-Sakakibara et al., 2019). *CHI* is co-regulated with other flavonoid genes, such as *CHS*, *F3H*, and *DFR* (Kayani et al., 2021). The genetic control of *CHI* activity has been assessed in flowers and pollen of several plant species (Sun et al., 2019). In accord with results of Hao et al. (2022), a higher *CHI* expression was reported

under DS in our study. In response to higher *CHI* expression levels under drought, a higher caffeic acid was observed in the high phenolic accession (Figure 9C). Another early flavonoid biosynthetic gene, *FLAVONOID 3'-MONOOXYGENASE* (*CYP75B1*), is mostly involved in monooxygenation/hydroxylation and regulation of antioxidant metabolites (Zandalinas et al., 2017). It has been shown that the *CYP75B1* gene positively correlates with flavonoids and antioxidant flavonol accumulation and induces drought tolerance in *Arabidopsis* (Rao et al., 2020). In our GWAS and RNA-seq analyses, *CYP75B137* was upregulated under drought and associated with coumaric acid, which suggest a protective role of this gene against the adverse effect of drought. In our study, the induction of *FLS3* was quite significant under DS (Figure 9E), which possibly indicates the role of higher *FLS3* expression in drought tolerance in rapeseed. *FLS* catalyzes the oxidation of dihydroflavonol to produce flavonol, which is widely distributed in flowering plants (Yang et al., 2022; Figure 9A). So far, 13 *FLS* genes have been identified for *B. napus* (Schilbert et al., 2021). Higher *FLS* transcripts under DS have been shown in rapeseed (Liu et al., 2015) and *Arabidopsis* (Li et al., 2018).

Glycosylation is one of the most widespread modifications of flavonoids as the last step in the flavonoid biosynthetic pathway in plants (Alseekh et al., 2020). The UDP-glycosyltransferase (UGT) is a key enzyme in flavonoid modification used for transferring UDP-activated sugars to flavonoid compounds and increasing their solubility and stability (Irmisch et al., 2018). Results of GWAS of our study supported that the *UGT89B1* gene was associated with chlorogenic acid content and the expression of *UGT89B1* was strongly induced by DS (Figure 9G). These observations were in agreement with results of a study in *Camellia sinensis* for downregulation of *UGT* enzymes that reduced flavonol glycoside accumulation, ROS detoxification capacity, and lower stress tolerance (Zhao et al., 2019). Our results indicated that the glucosylation catalyzed by *UGT89B1* was effective for drought tolerance in rapeseed through the improvement in the biosynthesis of phenolics.

Key regulators of phenolic biosynthesis pathway in *B. napus*

Biosynthetic enzymes, TFs, phytohormonal regulators, and environmental conditions affect phenylpropanoid biosynthesis and the structural genes (Guo et al., 2020). Different families of TFs including MYB, bHLH, bZIP, NAC, AP2/ERF, and WRKY regulate the expression of the structural genes in the phenylpropanoid pathway (Anwar et al., 2021). The plant-specific NAC proteins constitute one of the largest TF families and are characterized by a highly conserved NAC domain in the N-terminal region (Puranik et al., 2012). Although it has been widely demonstrated that the NAC TFs are involved in the regulation of various biological processes, from plant development to response to stresses, only a few members have been identified as regulators of anthocyanin biosynthesis (Sun et al., 2019). To date, some studies have focused on the regulation of *NAC* genes in anthocyanin biosynthesis (He et al., 2020). For example, *NAC078* protein was

associated with the expression of genes related to flavonoid biosynthesis and leads to the accumulation of anthocyanins in response to high light stress in *Arabidopsis* (Morishita et al., 2009). Results of GWAS in our study showed that the *NAC035* gene was associated with caffeic acid content. Moreover, transcription of *NAC035* that was enhanced by DS was higher in the high-phenolic rapeseed genotype in our study, which was in line with the results of *NAC035* expression in drought in the study by Li et al. (2022). *NAC035* affects the growth and developmental processes from cell division to flower development in response to external signals in plants (Li et al., 2022). We found that the *NAC035* gene could be a positive regulator of flower phenolics in rapeseed. *CHI*, *FLS3*, and *CYP75B137* were upregulated in the higher-phenolic rapeseed under drought conditions, which shows that the expression of these genes under DS is regulated not only by the structural genes but also by TFs.

As one of the largest TF families in plants, AP2/ERF (APETALA2/Ethylene Responsive Factor) plays indispensable roles in plant growth, development, hormone regulation, flavonoid synthesis, and especially in responses to various stresses (Zhang et al., 2020; Qian et al., 2022; Qi et al., 2023). Recently, three AP2/ERF TFs were identified as positive regulators of the chalcone isomerase (*CHI*) gene that enhanced the accumulation of flavanones and flavones in citrus (Zhao et al., 2021). In our study, the *ERF119* gene was associated with caffeic acid content and upregulated under drought and in high-phenolic rapeseed accession. In a study in *Mangifera indica*, *ERF109* promoted the transcription of anthocyanin-related genes and the anthocyanin accumulation (Qian et al., 2022). Under stress, ERFs recognize and bind to cis-regulatory elements in the promoters of target genes to participate in gene expression regulation (Seok et al., 2022). Therefore, our results in line with other studies showed that *ERF119* improves drought tolerance in rapeseed by regulating phenolic-related genes.

Conclusion

In the present study, a conjoint analysis based on metabolome and transcriptome data of rapeseed flowers was performed to identify the key metabolites, genes, and metabolic pathways related to drought response in rapeseed. We identified new sources of variation that contributed to higher polyphenolic compounds in rapeseed under DS conditions. Our work identified genomic locations within *B. napus* that influence the expression of phytochemicals and polyphenolic compounds under DS conditions. This work provides supporting information for two research areas. One is the development of breeding strategies for drought tolerance by disentangling the molecular mechanisms of these traits. The other is to develop genomic resources for developing metabolite-associated breeding in the future. The genomic and metabolomic variability identified in our selected rapeseed led to the detection of SNP-linked markers across the rapeseed genome that could help to facilitate the metabolite-associated breeding of rapeseed in the future. The use of metabolomic, genomic, and transcriptomic approaches increased our confidence in the identification of related genes and markers through GWAS. This workflow enabled the detection of six CGs including *PAL1*, *CHI*, *UGT89B1*,

FLS3, *CCRI1*, and *CYP75B137* associated with major phenolic acids and flavonoids. The analysis of the molecular mechanisms of drought response-related genes will contribute to the cultivation and development of new rapeseed drought-tolerant varieties, providing good materials for cultivating water-saving plants.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: <https://www.ncbi.nlm.nih.gov/sra/?term=PRJNA1052984>, <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE250611>.

Author contributions

MS performed the experiments, prepared samples, analyzed the data, and wrote original draft of the manuscript; BH designed the research and the experiments, prepared samples, wrote and reviewed the first and final draft of the manuscript; JB involved in experiments and reviewed the article; JW involved in the experiments and reviewed the article; XT involved in the experiments and reviewed the article; CR and HT provided scientific comments and reviewed the final draft of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The Handling Editor HR declared a past co-authorship with the author JB.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2023.1249142/full#supplementary-material>

References

- Agraharam, G., Girigoswami, A., and Girigoswami, K. (2022). Myricetin: A multifunctional flavonol in biomedicine. *Curr. Pharmacol. Rep.* 8, 48–61. doi: 10.1007/s40495-021-00269-2
- Ahmed, U., Rao, M. J., Qi, C., Xie, Q., Noushahi, H. A., Yaseen, M., et al. (2021). Expression profiling of flavonoid biosynthesis genes and secondary metabolites accumulation in *populus* under drought stress. *Mol.* 26 (18), 5546. doi: 10.3390/molecules26185546
- Alemu, A., Feyissa, T., Maccaferri, M., Sciara, G., Tuberosa, R., Ammar, K., et al. (2021). Genome-wide association analysis unveils novel QTLs for seminal root system architecture traits in Ethiopian durum wheat. *BMC Genom.* 22 (1), 20–16. doi: 10.1186/s12864-020-07320-4
- Ali, S., Gill, R. A., Ulhassan, Z., Zhang, N., Hussain, S., Zhang, K., et al. (2023). Exogenously applied melatonin enhanced the tolerance of *Brassica napus* against cobalt toxicity by modulating antioxidant defense, osmotic adjustment, and expression of stress response genes. *Ecotoxicol. Environ. Saf.* 252, 114624. doi: 10.1016/j.ecoenv.2023.114624
- Alseikh, S., de Souza, L. P., Benina, M., and Fernie, A. R. (2020). The style and substance of plant flavonoid decoration; towards defining both structure and function. *Phytochem* 174, 112347. doi: 10.1016/j.phytochem.2020.112347
- Anwar, M., Chen, L., Xiao, Y., Wu, J., Zeng, L., Li, H., et al. (2021). Recent advanced metabolic and genetic engineering of phenylpropanoid biosynthetic pathways. *Int. J. Mol. Sci.* 22 (17), 9544. doi: 10.3390/ijms22179544
- Ashraf, M. A., Rasheed, R., Hussain, I., Hafeez, A., Adrees, M., ur Rehman, M. Z., et al. (2022). Effect of different seed priming agents on chromium accumulation, oxidative defense, glyoxalase system and mineral nutrition in canola (*Brassica napus* L.) cultivars. *Environ. pollut.* 309, 119769. doi: 10.1016/j.envpol.2022.119769
- Barrett, J. C., Fry, B., Maller, J. D. M. J., and Daly, M. J. (2005). Haploview: analysis and visualization of LD and haplotype maps. *Bioinform* 21 (2), 263–265. doi: 10.1093/bioinformatics/bth457
- Baruah, I., Debbarma, J., Boruah, H., and Keshavaiah, C. (2017). The DEAD-box RNA helicases and multiple abiotic stresses in plants: A systematic review of recent advances and challenges. *Plant Omics* 10, 252–262. doi: 10.1016/b978-0-12-396546-2.00017-6
- Batool, M., El-Badri, A. M., Hassan, M. U., Haiyun, Y., Chunyun, W., Zhenkun, Y., et al. (2022). Drought stress in *Brassica napus*: effects, tolerance mechanisms, and management strategies. *J. Plant Growth Regul.* 42, (345) 1–25. doi: 10.1007/s00344-021-10542-9
- Bennouna, D., Tourniaire, F., Durand, T., Galano, J. M., Fine, F., Fraser, K., et al. (2021). The *Brassica napus* (oilseed rape) seeds bioactive health effects are modulated by agronomical traits as assessed by a multi-scale omics approach in the metabolically impaired ob-mouse. *Food Chem.* 2, 100011. doi: 10.1016/j.fochms.2021.100011
- Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *J. Bioinform* 23 (19), 2633–2635. doi: 10.1093/bioinformatics/btm308
- Breshegello, F., and Sorrells, M. E. (2006). Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. *Genetics* 172 (2), 1165–1177. doi: 10.1534/genetics.105.044586
- Celi, G. E. A., Grato, P. L., Lanza, M. G. D. B., and Dos Reis, A. R. (2023). Physiological and biochemical roles of ascorbic acid on mitigation of abiotic stresses in plants. *Plant Physiol. Biochem.* 222, 107970. doi: 10.1016/j.plaphy.2023.107970
- Chalhoub, B., Denoeud, F., Liu, S., Parkin, I. A., Tang, H., and Wang, X. (2014). Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science* 345, 950–953. doi: 10.1126/science.1253435
- Cheyrier, V., Comte, G., Davies, K. M., Lattanzio, V., and Martens, S. (2013). Plant phenolics: recent advances on their biosynthesis, genetics, and ecophysiology. *Plant Physiol. Biochem.* 72, 1–20. doi: 10.1016/j.plaphy.2013.05.009
- Cosme, P., Rodríguez, A. B., Espino, J., and Garrido, M. (2020). Plant phenolics: Bioavailability as a key determinant of their potential health-promoting applications. *Antioxidants* 9 (12), 1263. doi: 10.3390/antiox9121263
- Dai, X., Zhuang, J., Wu, Y., Wang, P., Zhao, G., Liu, Y., et al. (2017). Identification of a flavonoid glucosyltransferase involved in 7-OH site glycosylation in tea plants (*Camellia sinensis*). *Sci. Rep.* 7 (1), 5926. doi: 10.1038/s41598-017-06453-z
- Egbuna, C., Ifemeje, J. C., Maduako, M. C., Tijjani, H., Udedi, S. C., Nwaka, A. C., et al. (2018). "Phytochemical test methods: qualitative, quantitative and proximate analysis," in *Phytochemistry* (USA: Taylor & Francis Group), 381–426. doi: 10.1201/9780429426223-15
- Elferjani, R., and Soolanayakanahally, R. (2018). Canola responses to drought, heat, and combined stress: shared and specific effects on carbon assimilation, seed yield, and oil composition. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.01224
- Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14 (8), 2611–2620. doi: 10.1111/j.1365-294x.2005.02553.x
- Fard, N. S., Abad, H. H. S., Rad, A. S., Heravan, E. M., and Daneshian, J. (2018). Effect of drought stress on qualitative characteristics of canola cultivars in winter cultivation. *Ind. Crops Prod.* 114, 87–92. doi: 10.1016/j.indcrop.2018.01.082
- Farooq, M. A., Hong, Z., Islam, F., Noor, Y., Hannan, F., and Zhang, Y. (2021). Comprehensive proteomic analysis of arsenic induced toxicity reveals the mechanism of multilevel coordination of efficient defense and energy metabolism in two *Brassica napus* cultivars. *Ecotoxicol. Environ. Saf.* 208, 111744. doi: 10.1016/j.ecoenv.2020.111744
- Fikere, M., Barbulescu, D. M., Malmberg, M. M., Spangenberg, G. C., Cogan, N. O., and Daetwyler, H. D. (2020). Meta-analysis of GWAS in canola blackleg (*Leptosphaeria maculans*) disease traits demonstrates increased power from imputed whole-genome sequence. *Sci. Rep.* 10 (1), 1–15. doi: 10.1038/s41598-020-71274-6
- Fujita, M., and Hasanuzzaman, M. (2022). Approaches to enhancing antioxidant defense in plants. *Antioxidants* 11 (5), 925. doi: 10.3390/antiox1105925
- Gabriel, S. B., Schaffner, S. F., Nguyen, H., Moore, J. M., Roy, J., Blumenstiel, B., et al. (2002). The structure of haplotype blocks in the human genome. *Science* 296 (5576), 2225–2229. doi: 10.1126/science.1069424
- Gómez-Caravaca, A. M., Iafelice, G., Lavini, A., Pulvento, C., Caboni, M. F., and Marconi, E. (2012). Phenolic compounds and saponins in quinoa samples (*Chenopodium quinoa* Willd.) grown under different saline and nonsaline irrigation regimens. *J. Agric. Food Chem.* 60 (18), 4620–4627. doi: 10.1021/jf3002125
- Guo, Y., Gao, C., Wang, M., Fu, F. F., El-Kassaby, Y. A., Wang, T., et al. (2020). Metabolome and transcriptome analyses reveal flavonoids biosynthesis differences in *Ginkgo biloba* associated with environmental conditions. *Ind. Crops Prod.* 158, 112963. doi: 10.1016/j.indcrop.2020.112963
- Han, Y., Vimolmangkang, S., Soria-Guerra, R. E., and Korban, S. S. (2012). Introduction of apple ANR genes into tobacco inhibits expression of both CHI and DFR genes in flowers, leading to loss of anthocyanin. *J. Exp. Bot.* 63 (7), 2437–2447. doi: 10.1093/jxb/err415
- Hanes, S. D. (2015). Prolyl isomerases in gene transcription. *Biochim. Biophys. Acta Gen. Subj.* 10, 2017–20341850. doi: 10.1016/j.bbagen.2014.10.028
- Hao, Y., Wang, J., Hu, C., Zhou, Q., Mubeen, H. M., and Hou, X. (2022). Regulation of *BcMYB44* on anthocyanin synthesis and drought tolerance in non-heading Chinese cabbage (*Brassica campestris* ssp. *chinensis* Makino). *Hortic* 8 (5), 351. doi: 10.3390/horticulturae8050351
- He, Q., Ren, Y., Zhao, W., Li, R., and Zhang, L. (2020). Low temperature promotes anthocyanin biosynthesis and related gene expression in the seedlings of purple head Chinese cabbage (*Brassica rapa* L.). *Genes* 11 (1), 81. doi: 10.3390/genes11010081
- Hill, W. G., and Robertson, A. (1968). Linkage disequilibrium in finite populations. *Theor. Appl. Genet.* 38 (6), 226–231. doi: 10.1007/bf01245622
- Huang, H., Ullah, F., Zhou, D. X., Yi, M., and Zhao, Y. (2019). Mechanisms of ROS regulation of plant development and stress responses. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.00800
- Iqura, E., Humira, S., and François, B. (2015). Association mapping of QTLs for sclerotinia stem rot resistance in a collection of soybean plant introductions using a genotyping by sequencing (GBS) approach. *BMC Plant Biol.* 15 (1), 1–12. doi: 10.1186/s12870-014-0408-y
- Irmisch, S., Jo, S., Roach, C. R., Jancsik, S., Man Saint Yuen, M., and Madilao, L. L. (2018). Discovery of UDP-glycosyltransferases and BAHD-acyltransferases involved in the biosynthesis of the antidiabetic plant metabolite montbretin A. *Plant Cell* 30 (8), 1864–1886. doi: 10.1105/tpc.18.00406
- Jan, R., Khan, M. A., Asaf, S., Lubna, Waqas, M., Park, J. R., et al. (2022). Drought and UV radiation stress tolerance in rice is improved by overaccumulation of non-enzymatic antioxidant flavonoids. *Antioxidants* 11 (5), 917. doi: 10.3390/antiox11050917
- Jideani, A. I., Silungwe, H., Takalani, T., Omolola, A. O., Udeh, H. O., and Anyasi, T. A. (2021). Antioxidant-rich natural fruit and vegetable products and human health. *Int. J. Food Prop.* 2 (4), 41–67. doi: 10.22270/ajdhs.v2i4.20
- Joshi, R., Wani, S. H., Singh, B., Bohra, A., Dar, Z. A., Lone, A. A., et al. (2016). Transcription factors and plants response to drought stress: current understanding and future directions. *Front. Plant Sci.* 7. doi: 10.3389/fpls.2016.01029
- Juan, C. A., Pérez de la Lastra, J. M., Plou, F. J., and Pérez-Lebeña, E. (2021). The chemistry of reactive oxygen species (ROS) revisited: outlining their role in biological macromolecules (DNA, lipids and proteins) and induced pathologies. *Int. J. Mol. Sci.* 22 (9), 4642. doi: 10.3390/ijms22094642
- Jun, H. I., Wiesenborn, D. P., and Kim, Y. S. (2014). Antioxidant activity of phenolic compounds from canola (*Brassica napus*) seed. *Food Sci. Biotechnol.* 23 (6), 1753–1760. doi: 10.1007/s10068-014-0240-z
- Kalantarhadi, S., and Shoushi Dezfouli, A. A. (2021). Effects of drought stress and foliar application of ascorbic acid, salicylic acid, methanol and post-harvest storage on seed yield and seed vigor of Hyola401 rapeseed cultivar. *J. Crop Ecol. Physiol.* 15, 109–130. doi: 10.22059/jci.2021.299946.2372
- Kayani, S. I., Shen, Q., Rahman, S. U., Fu, X., Li, Y., Wang, C., et al. (2021). Transcriptional regulation of flavonoid biosynthesis in *Artemisia annua* by *AaYABBY5*. *Hortic. Res.* 8, 257. doi: 10.1038/s41438-021-00693-x

- Ke, L., Lei, W., Yang, W., Wang, J., Gao, J., Cheng, J., et al. (2020). Genome-wide identification of cold responsive transcription factors in *Brassica napus* L. *BMC Plant Biol.* 20 (1), 1–13. doi: 10.1186/s12870-020-2253-5
- Kejlik, Z., Kaplanek, R., Masarik, M., Babula, P., Matkowski, A., Filipensky, P., et al. (2021). Iron complexes of flavonoids-antioxidant capacity and beyond. *Int. J. Mol. Sci.* 22 (2), 646. doi: 10.3390/ijms22020646
- Khan, Z., Khan, M. N., Zhang, K., Luo, T., Zhu, K., and Hu, L. (2021). The application of biochar alleviated the adverse effects of drought on the growth, physiology, yield and quality of rapeseed through regulation of soil status and nutrients availability. *Ind. Crops Prod.* 171, 113878. doi: 10.1016/j.indcrop.2021.113878
- Kirkegaard, J. A., Lilley, J. M., Brill, R. D., Ware, A. H., and Walela, C. K. (2018). The critical period for yield and quality determination in canola (*Brassica napus* L.). *Field Crops Res.* 222, 180–188. doi: 10.1016/j.fcr.2018.03.018
- Kong, J. Q. (2015). Phenylalanine ammonia-lyase, a key component used for phenylpropanoids production by metabolic engineering. *RSC Adv.* 5 (77), 62587–62603. doi: 10.1039/c5ra08196c
- Kujur, A., Bajaj, D., Upadhyaya, H. D., Das, S., Ranjan, R., Shree, T., et al. (2015). A genome-wide SNP scan accelerates trait-regulatory genomic loci identification in chickpea. *Sci. Rep.* 5 (1), 1–20. doi: 10.1038/srep11166
- Kumar, K., Debnath, P., Singh, S., and Kumar, N. (2023). An overview of plant phenolics and their involvement in abiotic stress tolerance. *Stresses* 3 (3), 570–585. doi: 10.3390/stresses3030040
- Kusznierewicz, B., Bartoszek, A., Wolska, L., Drzewiecki, J., Gorinstein, S., and Namieśnik, J. (2008). Partial characterization of white cabbages (*Brassica oleracea* var. *capitata* f. *alba*) from different regions by glucosinolates, bioactive compounds, total antioxidant activities and proteins. *LWT - Food Sci. Technol.* 41 (1), 1–9. doi: 10.1016/j.lwt.2007.02.007
- Lalay, G., Ullah, S., and Ahmed, I. (2022). Physiological and biochemical responses of *Brassica napus* L. @ to drought-induced stress by the application of biochar and Plant Growth Promoting Rhizobacteria. *Microsc. Res. Tech.* 85 (4), 1267–1281. doi: 10.1002/jemt.23993
- Lee, J. E., Kim, A. J., Lee, J. E., and Kim, A. J. (2020). Antioxidant activity, whitening and anti-wrinkle effects of leaf and seed extracts of *Brassica juncea* L. Czern. *Asian J. Beauty Cosmetol.* 18 (3), 283–295. doi: 10.20402/ajbc.2020.0038
- Li, F., Chen, B., Xu, K., Gao, G., Yan, G., Qiao, J., et al. (2016). A genome-wide association study of plant height and primary branch number in rapeseed (*Brassica napus*). *Plant Sci.* 242, 169–177. doi: 10.1016/j.plantsci.2015.05.012
- Li, X. W., Li, J. W., Zhai, Y., Zhao, Y., Zhao, X., Zhang, H. J., et al. (2013). A R2R3-MYB transcription factor, *GmMYB12B2*, affects the expression levels of flavonoid biosynthesis genes encoding key enzymes in transgenic *Arabidopsis* plants. *Gene* 532 (1), 72–79. doi: 10.1016/j.gene.2013.09.015
- Li, W., Liu, Z., Feng, H., Yang, J., and Li, C. (2022). Characterization of the gene expression profile response to drought stress in *Populus ussuriensis* using PacBio SMRT and Illumina Sequencing. *Int. J. Mol. Sci.* 23 (7), 3840. doi: 10.3390/ijms23073840
- Li, Q., Yu, H. M., Meng, X. F., Lin, J. S., Li, Y. J., and Hou, B. K. (2018). Ectopic expression of glycosyltransferase UGT7G6E11 increases flavonoid accumulation and enhances abiotic stress tolerance in *Arabidopsis*. *Plant Biol.* 20 (1), 10–19. doi: 10.1111/plb.12627
- Liu, K., and Muse, S. V. (2005). PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinform* 21, 2128–2129. doi: 10.1093/bioinformatics/bti282
- Liu, C., Zhang, X., Zhang, K., An, H., Hu, K., Wen, J., et al. (2015). Comparative analysis of the *Brassica napus* root and leaf transcript profiling in response to drought stress. *Int. J. Mol. Sci.* 16 (8), 18752–18777. doi: 10.3390/ijms160818752
- Lopez, A., Rico, M., Rivero, A., and de Tangil, M. S. (2011). The effects of solvents on the phenolic contents and antioxidant activity of *Styopcaulon scoparium* algae extracts. *Food Chem.* 125 (3), 1104–1109. doi: 10.1016/j.foodchem.2010.09.101
- Maazou, A. R. S., Tu, J., Qiu, J., and Liu, Z. (2016). Breeding for drought tolerance in maize (*Zea mays* L.). *Am. J. Plant Sci.* 7 (14), 1858. doi: 10.4236/ajps.2016.714172
- Maran, J. P., Sivakumar, V., Thirugnanasambandham, K., and Sridhar, R. (2015). Extraction of natural anthocyanin and colors from pulp of jamun fruit. *Int. J. Mol. Sci.* 52 (6), 3617–3626. doi: 10.1007/s13197-014-1429-0
- Metsamuuronen, S., and Siren, H. (2019). Bioactive phenolic compounds, metabolism and properties: A review on valuable chemical compounds in Scots pine and Norway spruce. *Phytochem. Rev.* 18 (3), 623–664. doi: 10.1007/s11101-019-09630-2
- Mikolajczak, N., Sobiechowska, D. A., and Tariska, M. (2020). Edible flowers as a new source of natural antioxidants for oxidative protection of cold-pressed oils rich in omega-3 fatty acids. *Food Res. Int.* 134, 109216. doi: 10.1016/j.foodres.2020.109216
- Molyneux, P. (2004). The use of the stable free radical diphenylpicrylhydrazyl (DPPH) for estimating antioxidant activity. *Songklanakarim. J. Sci. Technol.* 26 (2), 211–219.
- Morishita, T., Kojima, Y., Maruta, T., Nishizawa-Yokoi, A., Yabuta, Y., and Shigeoka, S. (2009). *Arabidopsis* NAC transcription factor, *ANAC078*, regulates flavonoid biosynthesis under highlight. *Plant Cell Physiol.* 50 (12), 2210–2222. doi: 10.1093/pcp/pcp159
- Morsi, N. A., Hashem, O. S., El-Hady, M. A. A., Abd-Elkrem, Y. M., El-temsah, M. E., Galal, E. G., et al. (2023). Assessing drought tolerance of newly developed tissue-cultured canola genotypes under varying irrigation regimes. *Agron* 13 (3), 836. doi: 10.3390/agronomy13030836
- Murray, M. G., and Thompson, W. (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* 8 (19), 4321–4326. doi: 10.1093/nar/8.19.4321
- Naik, P. M., and Al-Khayri, J. M. (2016). “Abiotic and biotic elicitors-role in secondary metabolites production through in vitro culture of medicinal plants,” in *Abiotic and biotic stress in plants—recent advances and future perspectives* (Rijeka: InTech), 247–277. doi: 10.5772/61442
- Ni, R., Zhu, T. T., Zhang, X. S., Wang, P. Y., Sun, C. J., Qiao, Y. N., et al. (2020). Identification and evolutionary analysis of chalcone isomerase-fold proteins in ferns. *J. Exp. Bot* 71 (1), 290–304. doi: 10.1093/jxb/erz425
- Okafor, J. N., Meyer, M., Le Roes-Hill, M., and Jideani, V. A. (2022). Flavonoid and phenolic acid profiles of dehulled and whole *Vigna subterranea* (L.) Verdc seeds commonly consumed in South Africa. *Molecules* 27 (16), 5265. doi: 10.3390/molecules27165265
- Olgun, Ç., Özkan, O., Guney, B., Pattabanoglu, E., Güney, K., and Gür, M. (2017). Chemical composition and antimicrobial activity in cold press oil of fennel, Anise, white and black mustard seeds. *Indian J. Pharm. Educ. Res.* 51 (3), S200–S204. doi: 10.5530/ijper.51.3s.12
- Pal, L., Sandhu, S. K., and Bhatia, D. (2021). Genome-wide association study and identification of candidate genes for seed oil content in *Brassica napus*. *Euphytica* 217 (4), 1–15. doi: 10.1007/s10681-021-02783-2
- Patel, R. K., and Jain, M. (2012). NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. *PLoS One* 7 (2), e30619. doi: 10.1371/journal.pone.0030619
- Puranik, S., Sahu, P. P., Srivastava, P. S., and Prasad, M. (2012). NAC proteins: regulation and role in stress tolerance. *Trends Plant Sci.* 17 (6), 369–381. doi: 10.1016/j.tplants.2012.02.004
- Qi, H., Liang, K., Ke, Y., Wang, J., Yang, P., Yu, F., et al. (2023). Advances of Apetala2/ethylene response factors in regulating development and stress response in maize. *Int. J. Mol. Sci.* 24 (6), 5416. doi: 10.3390/ijms24065416
- Qian, L., Hickey, L. T., Stahl, A., Werner, C. R., Hayes, B., Snowdon, R. J., et al. (2017). Exploring and harnessing haplotype diversity to improve yield stability in crops. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.01534
- Qian, M., Wu, H., Yang, C., Zhu, W., Shi, B., Zheng, B., et al. (2022). RNA-Seq reveals the key pathways and genes involved in the light-regulated flavonoids biosynthesis in mango (*Mangifera indica* L.) peel. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.1119384
- Rabooanahiry, N., Li, H., Yu, L., and Li, M. (2021). Rapeseed (*Brassica napus*): Processing, utilization, and genetic improvement. *Agron* 11 (9), 1776. doi: 10.3390/agronomy11091776
- Rahman, M. M., Rahaman, M. S., Islam, M. R., Rahman, F., Mithi, F. M., Alqahtani, T., et al. (2021). Role of phenolic compounds in human disease: current knowledge and future prospects. *Molecules* 27 (1), 233. doi: 10.5772/66894
- Rao, B., and Deshpande, V. (2006). *Experimental Biochemistry* (Tunbridge Wells, Kent: Anshan).
- Rao, M. J., Xu, Y., Tang, X., Huang, Y., Liu, J., Deng, X., et al. (2020). *CsCYT75B1*, a Citrus CYTOCHROME P450 gene, is involved in accumulation of antioxidant flavonoids and induces drought tolerance in transgenic *Arabidopsis*. *Antioxid* 9 (2), 161. doi: 10.3390/antiox9020161
- Saha, D., Choyal, P., Mishra, U. N., Dey, P., Bose, B., Prathibha, M. D., et al. (2022). Drought stress responses and inducing tolerance by seed priming approach in plants. *Plant Stress* 4, 100066. doi: 10.1016/j.stress.2022.100066
- Salami, M., Heidari, B., and Tan, H. (2023). Comparative profiling of polyphenols and antioxidants and analysis of antiglycation activities in rapeseed (*Brassica napus* L.) under different moisture regimes. *Food Chem.* 399, 133946. doi: 10.1016/j.foodchem.2022.133946
- Samec, D., Piljac-Zegarac, J., Bogovic, M., Habjanic, K., and Gruz, J. (2011). Antioxidant potency of white (*Brassica oleracea* L. var. *capitata*) and Chinese (*Brassica rapa* L. var. *pekinensis* (Lour.)) cabbage: The influence of development stage, cultivar choice and seed selection. *Sci. Hortic.* 128 (2), 78–83. doi: 10.1016/j.scienta.2011.01.009
- Sarker, U., and Oba, S. (2018). Drought stress enhances nutritional and bioactive compounds, phenolic acids and antioxidant capacity of *Amaranthus* leafy vegetable. *BMC Plant Biol.* 18, 1–15. doi: 10.1186/s12870-018-1484-1
- Schilbert, H. M., Schöne, M., Baier, T., Busche, M., Viehöver, P., Weisshaar, B., et al. (2021). Characterization of the *Brassica napus* flavonol synthase gene family reveals bifunctional flavonol synthases. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.733762
- Sehgal, D., Skot, L., Singh, R., Srivastava, R. K., Das, S. P., Taunk, J., et al. (2015). Exploring potential of pearl millet germplasm association panel for association mapping of drought tolerance traits. *PLoS One* 10 (5), e0122165. doi: 10.1371/journal.pone.0122165
- Semwal, D. K., Semwal, R. B., Combrinck, S., and Viljoen, A. (2016). Myricetin: A dietary molecule with diverse biological activities. *Nutrients* 8 (2), 90. doi: 10.3390/nu8020090
- Seok, H. Y., Tran, H. T., Lee, S. Y., and Moon, Y. H. (2022). *AtERF71/HRE2*, an *Arabidopsis* AP2/ERF transcription factor gene, contains both positive and negative cis-regulatory elements in its promoter region involved in hypoxia and salt stress responses. *Int. J. Mol. Sci.* 23 (10), 5310. doi: 10.3390/ijms23105310
- Shahzad, A., Qian, M., Sun, B., Mahmood, U., Li, S., Fan, Y., et al. (2021). Genome-wide association study identifies novel loci and candidate genes for drought stress tolerance in rapeseed. *Oil Crop Sci.* 6 (1), 12–22. doi: 10.1016/j.ocsci.2021.01.002

- Shao, Y., Jiang, J., Ran, L., Lu, C., Wei, C., and Wang, Y. (2014). Analysis of flavonoids and hydroxycinnamic acid derivatives in rapeseeds (*Brassica napus* L. var. *napus*) by HPLC-PDA-ESI (-)-MS n/HRMS. *J. Agric. Food Chem.* 62 (13), 2935–2945. doi: 10.1021/jf404826u
- Sharma, A., Shahzad, B., Rehman, A., Bhardwaj, R., Landi, M., and Zheng, B. (2019). Response of phenylpropanoid pathway and the role of polyphenols in plants under abiotic stress. *Molecules* 24 (13), 2452. doi: 10.3390/molecules24132452
- Son, K. H., Ide, M., and Goto, E. (2020). Growth characteristics and phytochemicals of canola (*Brassica napus*) grown under UV radiation and low root zone temperature in a controlled environment. *Hortic. Environ. Biotechnol.* 61, 267–277. doi: 10.1007/s13580-019-00219-4
- Song, J., Sajad, S., Xia, D., and Jiang, S. (2023). Identification of F-box gene family in *Brassica oleracea* and expression analysis in response to low-temperature stress. *Plant Physiol. Biochem.* 199, 107717. doi: 10.1016/j.plaphy.2023.107717
- Sun, W., Shen, H., Xu, H., Tang, X., Tang, M., Ju, Z., et al. (2019). Chalcone isomerase a key enzyme for anthocyanin biosynthesis in *Ophiorrhiza japonica*. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.00865
- Tesfamariam, E. H., Annandale, J. G., and Steyn, J. M. (2010). Water stress effects on winter canola growth and yield. *Agron* 102 (2), 658–666. doi: 10.2134/agronj2008.0043
- Tohge, T., Watanabe, M., Hoefgen, R., and Fernie, A. R. (2013). Shikimate and phenylalanine biosynthesis in the green lineage. *Front. Plant Sci.* 4. doi: 10.3389/fpls.2013.00062
- Tomas-Barberan, F. A., and Espin, J. C. (2001). Phenolic compounds and related enzymes as determinants of quality in fruits and vegetables. *J. Sci. Food Agric.* 81 (9), 853–876. doi: 10.1002/jsfa.885
- Turner, S. D. (2014). qqman: an R package for visualizing GWAS results using QQ and Manhattan plots. *Biorxiv* 14, 005165. doi: 10.1101/005165
- Wagay, N. A., Rafiq, S., Khan, A., Kaloo, Z. A., Malik, A. R., and Pulate, P. V. (2023). “Impact of phenolics on drought stress and expression of phenylpropanoid pathway genes,” in *Plant Phenolics in Abiotic Stress Management* (Singapore: Springer), 265–285. doi: 10.1007/978-981-19-6426-8_13
- Wang, S., Lv, X., Zhang, J., Chen, D., Chen, S., Fan, G., et al. (2022). Roles of E3 ubiquitin ligases in plant responses to abiotic stresses. *Int. J. Mol. Sci.* 23 (4), 2308. doi: 10.3390/ijms23042308
- Wang, Y., Shi, Y., Li, K., Yang, D., Liu, N., Zhang, L., et al. (2021). Roles of the 2-oxoglutarate-dependent dioxygenase superfamily in the flavonoid pathway: a review of the functional diversity of F3H, FNS I, FLS, and LDOX/ANS. *Mol* 26 (21), 6745. doi: 10.3390/molecules26216745
- Wang, J., Xian, X., Xu, X., Qu, C., Lu, K., Li, J., et al. (2017). Genome-wide association mapping of seed coat color in *Brassica napus*. *J. Agric. Food Chem.* 65 (26), 5229–5237. doi: 10.1021/acs.jafc.7b01226
- Wu, Y., Wang, W., Li, Y., Dai, X., Ma, G., Xing, D., et al. (2017). Six phenylalanine ammonia-lyases from *Camellia sinensis*: Evolution, expression, and kinetics. *Plant Physiol. Biochem.* 118, 413–421. doi: 10.1016/j.plaphy.2017.06.030
- Xiong, J. L., and Ma, N. (2022). Transcriptomic and metabolomic analyses reveal that fullerol improves drought tolerance in *Brassica napus* L. *Int. J. Mol. Sci.* 23 (23), 15304. doi: 10.3390/ijms232315304
- Xu, K., Chen, S., Li, T., Yu, S., Zhao, H., Liu, H., et al. (2021). Overexpression of OsHMGB707, a high mobility group protein, enhances rice drought tolerance by promoting stress-related gene expression. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.711271
- Xu, L., Hu, K., Zhang, Z., Guan, C., Chen, S., Hua, W., et al. (2016). Genome-wide association study reveals the genetic architecture of flowering time in rapeseed (*Brassica napus* L.). *DNA Res.* 23 (1), 43–52. doi: 10.1093/dnares/dsv035
- Xu, W., Xu, X., Han, R., Wang, X., Wang, K., Qi, G., et al. (2023). Integrated transcriptome and metabolome analysis reveals that flavonoids function in wheat resistance to powdery mildew. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1125194
- Yang, H., Li, H., and Li, Q. (2022). Biosynthetic regulatory network of flavonoid metabolites in stems and leaves of *Salvia miltiorrhiza*. *Sci. Rep.* 12 (1), 18212. doi: 10.1038/s41598-022-21517-5
- Yao, T., Zhang, J., Xie, M., Yuan, G., Tschaplinski, T. J., Muchero, W., et al. (2021). Transcriptional regulation of drought response in *Arabidopsis* and woody plants. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.572137
- Yonekura-Sakakibara, K., Higashi, Y., and Nakabayashi, R. (2019). The origin and evolution of plant flavonoid metabolism. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.00943
- Yu, J., Pressoir, G., Briggs, W. H., Vroh Bi, I., Yamasaki, M., Doebley, J. F., et al. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* 38 (2), 203–208. doi: 10.1038/ng1702
- Zandalinas, S. I., Sales, C., Beltrán, J., Gómez-Cadenas, A., and Arbona, V. (2017). Activation of secondary metabolism in citrus plants is associated to sensitivity to combined drought and high temperatures. *Front. Plant Sci.* 7. doi: 10.3389/fpls.2016.01954
- Zhang, Y., Bao, Y., Li, P., Yu, Q., Li, W., Tang, L., et al. (2023). Effects of drought stress on carbon metabolism of Bermudagrass (*Cynodon dactylon* L.). *Plant Soil Environ.* 69 (6). doi: 10.17221/426/2022-PSE
- Zhang, S., Chen, Y., Zhao, L., Li, C., Yu, J., Li, T., et al. (2020). A novel NAC transcription factor, MdNAC42, regulates anthocyanin accumulation in red-fleshed apple by interacting with MdMYB10. *Tree Physiol.* 40 (3), 413–423. doi: 10.1093/treephys/tpaa004
- Zhang, X., and Liu, C. J. (2015). Multifaceted regulations of gateway enzyme phenylalanine ammonia-lyase in the biosynthesis of phenylpropanoids. *Mol. Plant* 8 (1), 17–27. doi: 10.1016/j.molp.2014.11.001
- Zhang, J., Mason, A. S., Wu, J., Liu, S., Zhang, X., Luo, T., et al. (2015). Identification of putative candidate genes for water stress tolerance in canola (*Brassica napus*). *Front. Plant Sci.* 61058. doi: 10.3389/fpls.2015.01058
- Zhang, C., Shi, S., Liu, Z., Yang, F., and Yin, G. (2019). Drought tolerance in alfalfa (*Medicago sativa* L.) varieties is associated with enhanced antioxidative protection and declined lipid peroxidation. *J. Plant Physiol.* 232, 226–240. doi: 10.1016/j.jplph.2018.10.023
- Zhang, P., Zhang, E., Xiao, M., Chen, C., and Xu, W. (2013). Enhanced chemical and biological activities of a newly biosynthesized eugenol glycoconjugate, eugenol α -D-glucopyranoside. *Appl. Microbiol. Biotechnol.* 97 (3), 1043–1050. doi: 10.1007/s00253-012-4351-2
- Zhao, J. (2015). Flavonoid transport mechanisms: how to go, and with whom. *Trends Plant Sci.* 20 (9), 576–585. doi: 10.1016/j.tplants.2015.06.007
- Zhao, M., Jin, J., Gao, T., Zhang, N., Jing, T., Wang, J., et al. (2019). Glucosyltransferase CsUGT78A14 regulates flavonols accumulation and reactive oxygen species scavenging in response to cold stress in *Camellia sinensis*. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.01675
- Zhao, C., Liu, X., Gong, Q., Cao, J., Shen, W., Yin, X., et al. (2021). Three AP2/ERF family members modulate flavonoid synthesis by regulating type IV chalcone isomerase in citrus. *Plant Biotechnol. J.* 19 (4), 671–688. doi: 10.1111/pbi.13494
- Zhou, R., Jackson, L., Shadle, G., Nakashima, J., Temple, S., Chen, F., et al. (2010). Distinct cinnamoyl CoA reductases involved in parallel routes to lignin in *Medicago truncatula*. *Proc. Natl. Acad. Sci.* 107 (41), 17803–17808. doi: 10.1073/pnas.1012900107
- Zhu, J., Cai, D., Wang, J., Cao, J., Wen, Y., He, J., et al. (2021). Physiological and anatomical changes in two rapeseed (*Brassica napus* L.) genotypes under drought stress conditions. *Oil Crop Sci.* 6 (2), 97–104. doi: 10.1016/j.ocsci.2021.04.003
- Zou, Y., Lu, Y., and Wei, D. (2004). Antioxidant activity of a flavonoid-rich extract of *Hypericum perforatum* L. in vitro. *J. Agric. Food Chem.* 52 (16), 5032–5039. doi: 10.1021/jf049571r
- Zou, T. B., Wang, M., Gan, R. Y., and Ling, W. H. (2011). Optimization of ultrasound-assisted extraction of anthocyanins from mulberry, using response surface methodology. *Int. J. Mol. Sci.* 12 (5), 3006–3017. doi: 10.3390/ijms12053006