BMC Plant Biology

RESEARCH

Open Access

Gene mapping and identification of candidate genes controlling carotenoid accumulation of yellow kernels in foxtail millet

Junjie Wang[1], Qi Ma[1], Yuyang Zhang[1], Qian Duan[1], Xiaoxi Zhen[1], Yaoyuan Zhang[1], Hongying Li[1], Yuanhuai Han[1,2*] and Bin Zhang[1,2*]

Abstract

Background Kernel color is an important characteristic of foxtail millet (*Setaria italica*) associated with its market ability, quality, and nutritional value, which is mainly due to the accumulation of carotenoids. Despite its importance, the genetic basis of carotenoid variation in foxtail millet remains largely unexplored. Herein, the molecular mechanisms governing carotenoid accumulation in the kernel of foxtail millet were investigated by an exhaustive methodology encompassing Genome-Wide Association Study (GWAS), Bulk Segregant Analysis sequencing (BSA-seq), and integrated transcriptomic and metabolomic analyses.

Results The total carotenoid content in kernels across 201 foxtail millet germplasms showed a spectrum of variations, which indicated that the kernel color is a quantitative genetic trait controlled by multiple genes. Using GWAS on these germplasms and BSA-seq on an F_6 generation Recombinant Inbred Line (RIL) population derived from the GBS (yellow kernel) and NMB (white kernel) cross, we identified genome regions linked with total carotenoid content in foxtail millet kernels. Integrated transcriptomic and metabolomic profiling during grain filling in both yellow and white varieties pinpointed *SiPSY1* and *SiCCD1* as key genes controlling carotenoid accumulation. Notably, the SNP (G/A) at 364 bp and the Indel (29 bp insertion) at 856 bp within the *SiPSY1* promoter predominantly contributed to the variance in promoter activity. These variations markedly affected *SiPSY1* expression levels, ultimately determining the phenotypic difference between yellow and white kernels.

Conclusions These findings provide crucial genetic insights for the molecular mechanisms involved in carotenoid metabolism and lay a solid foundation for millet color breeding in foxtail millet.

Keywords Foxtail millet, Kernel color, GWAS, BSA-seq, Promoter activity

*Correspondence:
Yuanhuai Han
swgctd@163.com
Bin Zhang
zb0304209@163.com
[1]College of Agriculture, Shanxi Agricultural University, Taigu 030801, Shanxi, China
[2]Houji Laboratory in Shanxi Province, Taiyuan 030031, Shanxi, China

Introduction

Foxtail millet (*Setaria italica*), a traditional and important orphan crop in northern China, demonstrates resilience to drought, adaptability to poor soil conditions, efficient water usage, and consistent high yields [1, 2]. Foxtail millet kernels are rich in nutrition [3]. With increasing public health consciousness, minor crops like foxtail millet are garnering greater attention. Consequently, the advancing foxtail millet industry requires enhanced quality standards. Kernel color, a pivotal parameter for evaluating foxtail millet quality, is a critical aspect of breeding. Kernel color phenotypes exhibit a diverse spectrum, encompassing up to 16 different hues such as yellow, white, gray, and green. Notably, yellow kernel varieties dominate germplasm accessions, followed by light-yellow kernel varieties in foxtail millet [4]. Carotenoids constitute the principal yellow pigments and are important nutrients in kernels. Lutein and zeaxanthin, the predominant carotenoids in kernels, contribute to 90% of their total carotenoid content [5, 6].

Carotenoids, a group of terpenes composed of isoprene units, possess polyene main chains with differing quantities of conjugated double bonds and demonstrate antioxidant characteristics. In plants, carotenoids accumulate in high concentrations in flowers, leaves, fruits, and seeds, producing vibrant yellow, orange, and red pigmentation [7]. These compounds play a crucial role in human health by serving as provitamin A precursors, boosting immunity, delaying aging, and reducing the risk of chronic cardiovascular diseases and cancer [8, 9]. Since mammals cannot synthesize carotenoids endogenously, they must obtain them exclusively through plant-based diets. Therefore, a comprehensive understanding of carotenoid biosynthesis and degradation is essential for enhancing crop nutritional quality via biofortification.

The carotenoid biosynthesis pathway in higher plants is thoroughly understood and involves five critical reactions: condensation, desaturation/isomerization, hydroxylation, oxidation, and epoxidation [7, 10]. Phytoene synthase (PSY) acts as both the initial and most pivotal rate-limiting enzyme in this pathway, catalyzing the conversion of two geranylgeranyl pyrophosphate (GGPP) molecules into one molecule of 15-cis-phytoene. Since PSY activity is essential for directing metabolic flux towards the pathway and significantly impacts carotenoid content, it is a primary focus in carotenoid metabolism research [11, 12]. The sequential desaturation and isomerization of 15-cis-phytoene, facilitated by phytoene desaturase (PDS), ζ-carotene isomerase (ZISO), ζ-carotene desaturase (ZDS), and carotenoid isomerase (CRTISO), yield red-colored all-trans-lycopene, a crucial intermediate in the carotenoid biosynthesis pathway [13, 14]. The cyclization of all-trans-lycopene, catalyzed by either lycopene ε-cyclase (LCYE) or lycopene β-cyclase (LCYB), results in the production of symmetrical orange α- and β-carotene within the respective β-ε and β-β pathways. LCYE and LCYB activities can significantly influence the diversion of metabolic flux toward the β-ε and β-β pathways [15, 16]. Following sequential ring-specific hydroxylation of α-carotene and β-carotene by two cytochrome P450 type hydroxylases (CYP97A and CYP97C) and two non-heme β-ring hydroxylases (BCH1 and BCH2), yellow xanthophylls lutein and zeaxanthin are produced, respectively [17, 18]. Lutein represents the end of the β-ε pathway in the carotenoid biosynthesis. In β-β pathway, Zeaxanthin epoxidase (ZEP) facilitates the conversion of zeaxanthin into antheraxanthin and subsequently violaxanthin. This process can be reversed by violaxanthin de-epoxidase (VDE), known as the xanthophyll cycle [19]. Neoxanthin synthase (NXS) catalyzes the conversion of violaxanthin to neoxanthin, completing the core biosynthetic pathway [20]. Additionally, carotenoid degradation mediated by carotenoid cleavage dioxygenase (CCDs) also impacts the final carotenoid content in plants. The CCD family is divided into two groups: 9-cis-epoxycarotenoid dioxygenases (NCEDs), which specifically cleave 9-cis-violaxanthin and 9-cis-neoxanthin to produce xanthoxin - a pivotal precursor in ABA synthesis, and CCDs (types 1, 2, 4, 7, 8, and 10), capable of catalyzing various carotenoids, leading to the formation of a range of apocarotenoids [21–23]. The enzymes CCD7 and CCD8 sequentially cleave 9-cis-β-carotene to generate carlactone, the precursor of strigolactones (SL) biosynthesis [24]. CCD1 and CCD4 have been shown to negatively regulate carotenoid contents or participate in color formation in various plant species [25–28].

The color of the kernel serves as a direct indicator of quality in foxtail millet. Enhancing carotenoid levels in the kernel to improve color and, consequently, nutritional and commercial value, remains a primary objective in quality-focused foxtail millet breeding efforts. Thus, investigating the molecular mechanisms governing carotenoid accumulation in the kernel and identifying molecular markers for promoting kernel color through molecular breeding are critical tasks. This research systematically identifies and selects key genes influencing kernel carotenoid traits using an integrated approach. This involves conducting a GWAS to explore natural population diversity of carotenoids in foxtail millet grains, analyzing results from BSA-seq of an F_6 RIL population derived from crosse between yellow kernel and white kernel lines. Additionally, it integrates findings from combined transcriptomic and metabolomic analyses conducted at different grain maturation stages on materials representing extreme kernel colors to identify and screen critical genes regulating kernel color formation. Promoter sequences and activities of pivotal genes in different colored kernel varieties were compared and

analyzed to elucidate the genetic and molecular mechanisms underlying kernel pigmentation in foxtail millet. This study will advance understanding of the molecular mechanisms involved in carotenoid metabolism in crops and facilitate the development of molecular breeding strategies to enhance the nutritional quality of foxtail millet.

Materials and methods

Field trials and phenotypic analysis

Field trials and phenotypic analyses were conducted using 201 core Chinese foxtail millet resources from the natural population. The hybrid population was derived from a cross between GBS (♀, yellow kernel) and NMB (♂, white kernel). Table S1 provides detailed information on all varieties. The experiments were conducted in a randomized complete block design at the experimental field of Shanxi Agricultural University. Each variety was planted in 3 rows, with 0.2 m spacing between plants and 0.3 m between rows. Seeds were sown at a depth of 5 cm, and standard field management practices were followed, with three biological replicates. After harvest, well-developed, plump, and uniformly colored grains were chosen for further analysis.

The yellow intensity of the foxtail millet kernels was assessed using a non-contact colorimeter (VS450, X-rite, USA) to measure the b* value (representing yellow and blue intensity, where higher b* values indicate increased yellowness). Total carotenoids were extracted following the AACC method, which involved grinding the kernels into powder under liquid nitrogen, freeze-drying the powder, and then extracting carotenoids with water-saturated n-butanol. The total carotenoid content (mg/kg) was calculated based on the absorbance at 450 nm using the formula [(A - A0) / 0.250] × V / m, where A represents the absorbance of the test liquid, A0 is the absorbance of water-saturated n-butanol, V is the volume of the extraction solution, m is the sample mass, and 0.250 is the carotenoid conversion factor.

Genome-wide association analysis

A GWAS was performed on the total carotenoid content in foxtail millet using the GLM model in GAPIT software. Single-nucleotide polymorphism (SNP) markers were used to account for population structure, with a significant threshold (-log10 ≈ 6) and decay distance (around 50 kb) identified via linkage disequilibrium (LD). The SNP density was visualized through a Manhattan plot using the "CMplot" package in R, highlighting continuous segments in the plot with -log10(p) values of ≥ 6 as potential regions. Genes within these regions were identified, and their protein sequences extracted using TBtools software. The functions of these proteins were then annotated using the Quick Protein Anno plugin. The data from the phytozome (https://phytozome-next.jgi.doe.gov/) and MDSi (https://foxtail-millet.biocloud.net/home) databases were used to identify suitable candidate genes or transcription factors associated with the trait under study.

Analysis of key gene haplotype

The chosen pivotal genes were analyzed using CandiHap software to investigate variations across a region spanning 2000 bp upstream and 500 bp downstream of the target genes in different varieties. These varieties were grouped according to the types of genetic variations, and the top six haplotypes, which were the most prevalent among the primary varieties, were selected for further study. The statistical significance of differences between these haplotypes was assessed with SPSS 25 software, and box-and-whisker plots were generated using Origin 2022b.

BSA-seq analysis

This study was based on the F_5 generation RILs population. For the experimental groups, individuals with grain b* values of ≥ 40 were selected for the F_6 Yellow pool (yellow kernel), while those with grain b* values of ≤ 30 were chosen for the F_6 White pool (white kernel). DNA was extracted from the F_6 generation during the milky stage after the third leaf stage, and fragmented into 350–500 bp segments using a Covaries S220 ultrasonic disruptor (Gene, China). The libraries were prepared with the TruSeq DNA LT Sample Prep kit (Illumina, USA), and paired-end sequencing was carried out by Shanghai OE Biotech Inc. (Shanghai, China).

The sequencing results were processed with fastp software to remove adapter reads, filter out reads with base-calling quality below 95%, discard reads with a base-calling error rate above 1%, and exclude those shorter than 75 bp. The BWA and SAMtools modules in TBtools II software aligned the sequencing data with the foxtail millet reference genome (*Setaria italica* v2.2, https://phytozome-next.jgi.doe.gov/), and Picard software was used to eliminate redundancy. Based on this alignment, the Haplotypecaller module in GATK4 detected SNPs and InDels, keeping variant sites with a Quality/Depth (QD) ratio of ≥ 2. SNP-index values were computed from parental genotypes, and mapping was performed using the ΔSNP-index method. The SNP-index values were calculated in 1 M windows with a 20 kb step, and the average SNP-index within each window, together with the 95% and 99% confidence intervals, were plotted to highlight candidate regions. Candidate genes within these regions were chosen based on their exceeding confidence intervals, and those showing mutations consistent with the mutant parent and significant SNP-index differences were identified as candidate genes.

Detection of carotenoid components and transcriptome sequencing

Four representative samples, exhibiting extreme foxtail millet kernel colors, were selected for transcriptomic sequencing from the parent varieties GBS (yellow kernel), NMB (white kernel), and their RIL population across various stages: mid-grain filling (S3), late stage (S4), and final stage (S5). The yellow kernel samples were labeled as Y-4 and Y-10, and the white kernel samples as W-1 and W-7, each having three biological replicates. After collection, the samples were sent to Metware Biotechnology Co., Ltd., China, for carotenoid component analysis and transcriptome sequencing.

The desiccated samples were ground using a ball mill, and 50 mg of the powdered sample was set aside for extraction with 0.5 mL of an extraction solvent. The mixture was vortexed, centrifuged, and the supernatant collected. The extract was then concentrated, reconstituted, filtered, and stored in amber vials for subsequent LC-MS/MS analysis. An ultra-performance liquid chromatography system (UPLC, ExionLCTM AD) combined with tandem mass spectrometry (MS/MS, QTRAP® 6500+) was employed for the analysis. The liquid phase conditions and the mass spectrometry conditions were also meticulously outlined.

RNA extraction and library construction were executed by Metware Biotechnology Co. Ltd., succeeded by initial quantification and dilution. The insert size was confirmed with an Agilent 2100 bioanalyzer, and accurate quantification was done to assess library quality. After quality control, sequencing was completed on the Illumina HiSeq platform, filtering out adapter reads and excluding low-quality reads. The clean reads were aligned to the *Setaria italica* v 2.2 reference genome, and Fragments Per Kilobase of transcript per Million mapped reads (FPKM) values were calculated. These values were used for sample correlation analysis and Principal Component Analysis (PCA). Differential expression analysis between sample groups was performed with DESeq2 to identify significantly differentially expressed genes, following specific selection criteria.

Cloning and analysis of the *SiPSY1* gene promoter

A 1580 bp sequence upstream of the *SiPSY1* transcription start site was obtained from the foxtail millet genome database. The primers were designed using Primer5 software: (*SiPSY1*-F: TTTAGACCCTGTACTC GGTTTTG, *SiPSY1*-R: GTGTCTGTCGCCTGTTGTTA T). Genomic DNA from selected yellow and white kernel samples of the F_6 generation was extracted using the CTAB method. The extracted DNA was utilized as a template in PCR to amplify the desired fragment. The PCR reaction mixture included 2 × GC Buffer (25 μL), dNTP Mixture (8 μL), each upstream and downstream primer (10 mM, 1 μL), Takara LA Taq enzyme (0.5 μL), template DNA (1 μL), and supplementary ddH_2O up to 50 μL. The PCR protocol included initial denaturation at 94 °C for 5 min, followed by 35 cycles at 94 °C for 30 s (denaturation), 56 °C for 30 s (annealing), and 72 °C for 1 min 30 s (extension), concluding with a final extension at 72 °C for 10 min. The PCR products were identified using 1% agarose gel electrophoresis. The target fragment was cloned into the pMD18-T vector for sequencing. The sequencing outcomes of the *SiPSY1* promoter clone were compared between the representative yellow and white kernel samples using the MultAlin website (http://multalin.toulouse.inra.fr/multalin/multalin.html). The predictions and analyses of cis-acting elements were carried out using the PlantCARE website (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/).

Analysis of *SiPSY1* gene promoter activity

The cloned promoter sequences of *SiPSY1* were amplified from yellow (Y) and white (W) foxtail millet varieties via PCR, employing primers with homologous sequences to pGreenII0800-LUC (*SiPSY1* upstream primer: cccctcgaggtcgacggtatcgataTTTAGACCCTGTACTCG-GTTTTGTGGC; downstream primer: ttttggcgtcttccatg-gtcccccgGTGTCTGTCGCCTGTTGTTATCTGC, the lowercase letters indicating homologous arm sequences, and uppercase letters representing the gene sequence applied as primers). The fragment carrying the SNP mutation (A1, specifically an A to G substitution at position 364 bp) was designed via overlap extension PCR, utilizing the promoter sequence of the yellow variety as the template. Similarly, the fragment harboring the Indel mutation (A2, characterized by a 29 bp insertion at position 856 bp) was generated through overlap extension PCR, employing the promoter sequence of the white variety as the template. All fragments were inserted into the pGreenII0800-LUC vector using the homologous recombination method facilitated by the Seamless Cloning Kit (Biorun). The resulting recombined product was then transformed into *Escherichia coli.* Positive clones were confirmed via colony PCR, and the accurate sequences of the pGreenII0800-LUC-SiPSY1-Ypro, pGreenII0800-LUC-SiPSY1-Wpro, pGreenII0800-LUC-SiPSY1-A1pro and pGreenII0800-LUC-SiPSY1-A2pro vectors were verified through sequencing of the positive clone cultures.

Subsequently, protoplasts were generated from rice seedlings grown in darkness for a week. The previously constructed recombinant plasmids were introduced into these protoplasts via PEG-mediated transformation. The transformed protoplasts were incubated in darkness at 30 °C for 18–24 h before cell lysis. Firefly luciferase activity was measured by adding the luciferase substrate to the cell lysate and immediately detecting the signal with a luminometer. The reaction mixture was then quickly

combined with a freshly prepared Renilla substrate solution, and Renilla luciferase activity was also assessed immediately using the luminometer. Each measurement was conducted in triplicate.

Results

Analysis of the total carotenoid content in kernel

After planting and harvesting 201 varieties of foxtail millet, the grains of each variety were hulled and sorted to obtain uniformly colored and sized kernels. Extraction of total carotenoid contents from each variety aimed to explore genotypic variation in kernel color traits across the 201 foxtail millet accessions. The total carotenoid contents exhibited significant variability within the core germplasm resources of foxtail millet. Among the tested varieties, total carotenoid contents ranged from 4.07 $mg{\cdot}kg^{-1}$ to 27.04 $mg{\cdot}kg^{-1}$, thereby demonstrating extensive diversity (Fig. 1a). Statistical analysis revealed that the total carotenoid content index followed a normal distribution within the 201 natural populations (chi-square value greater than 0.05), suggesting that kernel carotenoid content was a quantitative genetic trait regulated by multiple genes. This observation aligns with the criteria conducive for GWAS analysis.

Genome-wide association study

A GWAS was conducted on the total carotenoid content in kernels using resequencing data from 201 foxtail millet materials. We obtained 3,252,346 high-quality SNP loci, which were analyzed using GWAS, identifying 12 significant regions associated with total carotenoid contents across chromosomes 1, 2, 3, 4, 5, 8, and 9 (Fig. 1b). According to the genome annotation, these regions encompass a total of 342 genes, among which 238 genes have been annotated (Table S2). The quantile-quantile plot depicted an upward shift in the $-\log_{10}$ (P) values, indicating significant variation at these loci beyond random drift. However, since the upward shift was observed in the lower range of $-\log_{10}$ (P) values, it was likely to encompass some false positive loci. Consequently, discontinuous signal points were eliminated to reduce interference from false positives (Fig. 1c). Continuous and stable signal points observed at the termini of chromosome 3 (49808233–49995697 bp, 0.18 M) and chromosome 4 (37851471–38090782 bp, 0.23 M) suggest critical regions involved in regulation of carotenoid biosynthesis and metabolism (Fig. 1b). Within these two regions, there were 27 and 44 annotated genes, respectively. Among all candidate genes, *SiPSY1* and *SiPSY2*, orthologs of the phytoene synthase gene, have been reported to be involved in regulating a major rate-limiting step in carotenoid biosynthesis.

The CandiHap software was utilized for haplotype analysis of the identified key genes to investigate superior haplotypes correlated with the kernel color trait. Within the *SiPSY1* gene region, 10 variations were found with 9 located in the promoter region and 1 in the exon. These variants were classified into 43 distinct haplotypes based

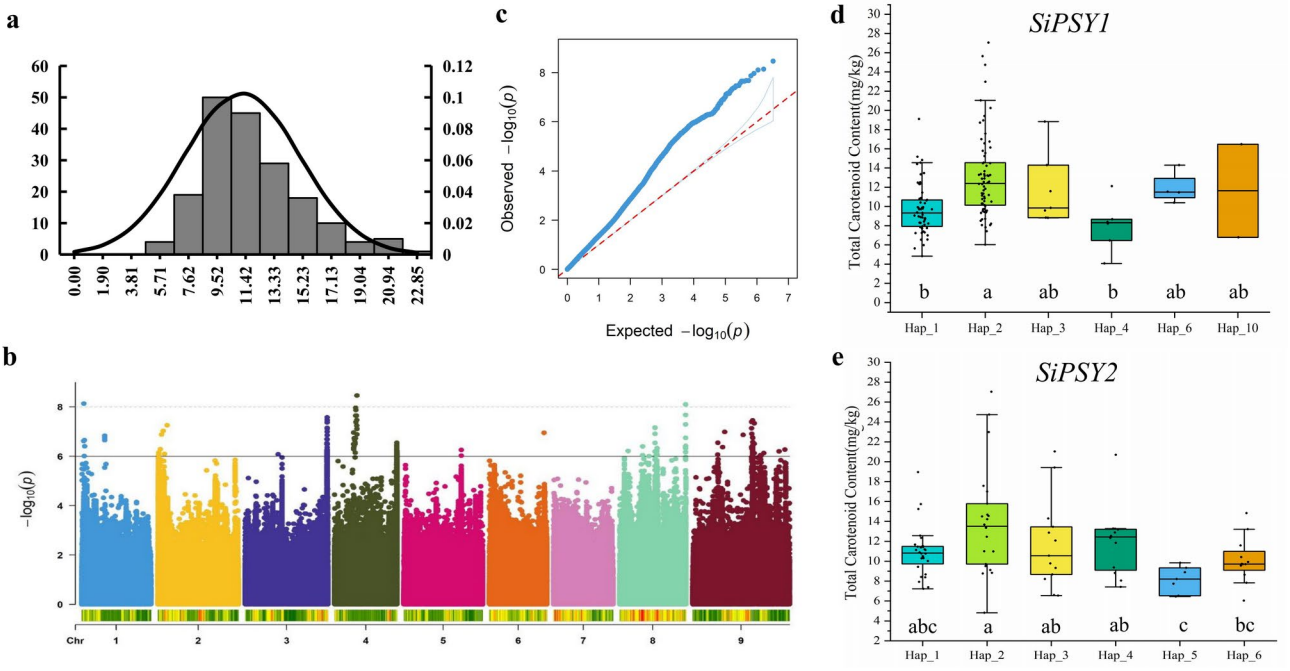

Fig. 1 Identification of key genes regulating carotenoid accumulation in kernels of foxtail millet through GWAS analysis. **(a)** Normal distribution of total carotenoid content in kernels across natural population. **(b)** GWAS results for total carotenoid content in kernels. **(c)** Quantile-quantile plot. **(d)** Haplotype boxplot of the *SiPSY1*. **(e)** Haplotype boxplot of the *SiPSY2*

on mutation types, with the first two exhibiting the highest diversity. Analysis predicted that three variations, SNP4 (A/G), SNP5 (A/C), and SNP6 (C/T), could lead to Hap_1 containing additional cis-elements associated with MYB (CAACAG), abscisic acid (ABRE: ACGTG), and light response (G-box: CACGTT) compared to Hap_2 (Fig. 1d; Table S3). These differences could potentially explain the significantly reduced total carotenoid levels in Hap_1. The synonymous mutation found in the exon region did not alter the protein's structure. In the *SiPSY2* gene region, we identified 16 variations: 12 in the promoter region, 2 downstream, and 2 in the UTR region. Interestingly, relative to other primary haplotypes, Hap_2 possessed unique SNP sites (SNP2: G/A, SNP5: C/T, SNP8: C/T, SNP10: T/C, SNP11: C/G, SNP12: G/C) and demonstrated a markedly higher total carotenoid content, significantly surpassing that of Hap_5 and Hap_6 (Fig. 1e; Table S4). However, these SNPs did not alter the promoter's cis-acting elements. Both *SiPSY1* and *SiPSY2* affected carotenoid concentrations in foxtail millet kernels, though *SiPSY1* appeared to have a more significant effect. We hypothesized that alterations in the promoter region of the *SiPSY1* gene primarily accounted for its diverse roles in carotenoid metabolism among different colored kernel varieties.

Preliminary mapping by BSA-seq

To further identify the key genes that regulate the color traits of foxtail millet kernel, we established a RILs population derived from two parental varieties: GBS, characterized by yellow kernels, and NMB, noted for white kernels (Fig. 2a). Representative samples displaying extreme kernel coloration were chosen based on colorimetric analysis conducted on the F_5 generation, with the yellowness b* value acting as the selection criterion. For the construction of mixed pools, 130 plants with a b* value ≥ 40, indicative of extreme yellow kernels, and another 130 plants with a b* value ≤ 30, representing extreme white kernels, were selected (Fig. 2a). These were designated as yellow kernel (Y) and white kernel (W) pools, respectively (Fig. S1). Simultaneously, the yellow and white kernel parents, GBS and NMB, were utilized to create the corresponding parental mixed pools. The resequencing data yielded a total of 64.47 Gb of data, achieving a sequencing depth exceeding 25X in both parental and progeny pools, and attained a coverage rate above 90% at a sequencing depth of 10X. Following SNP and Indel filtration guidelines, we identified 2,924,668 SNP loci and 425,157 Indel loci. The ΔSNP-index for progeny pools Y and W was calculated, and fitting curves with 95% and 99% confidence intervals were generated (Table S5). The positioning results unveiled two distinct peak signals on the Manhattan plot, situated in the regions chromosome 3: 47,450,376–50,939,361 bp (3.33 M) and chromosome 4: 39,315,152–40,402,503 bp (1.03 M), respectively. Both peaks exceeded the 99% confidence threshold, with peak values of 0.86 and 0.65, underscoring their extreme significance (Fig. 2b). Subsequent analysis unveiled the presence of *SiPSY1*, *SiPSY2*, and *SiCCD1* within these loci, all of which are implicated in the carotenoid biosynthesis pathway (Table S6). This finding strongly suggests that these three structural genes within the carotenoid biosynthesis pathway exert a pivotal influence on determining the color disparities

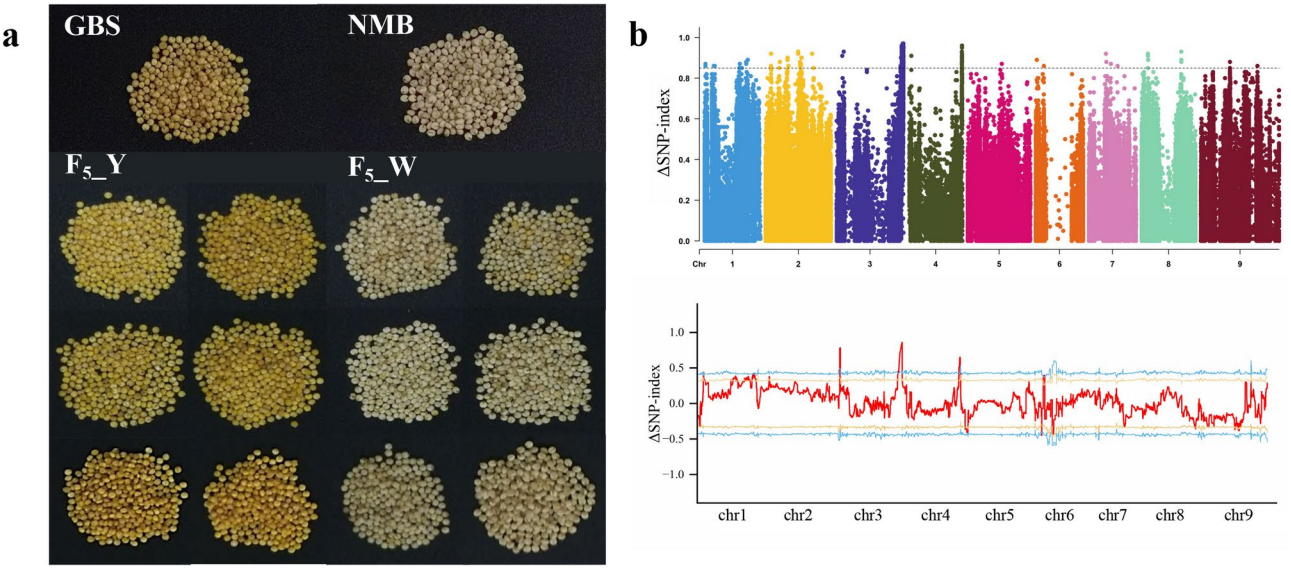

Fig. 2 Identification of key genes responsible for kernel color differences between yellow and white kernels in foxtail millet using BSA-seq. **(a)** Kernel color phenotypes of the two parents, GBS and NMB, along with representative samples from the F_5 population. **(b)** Gene mapping related to kernel color formation in foxtail millet via BSA-seq analysis

between yellow and white kernels pigmentation in foxtail millet.

Targeted metabolomics analysis of carotenoids at different maturation stages in yellow and white kernel

To elucidate the composition and concentration of carotenoids in foxtail millet grains and their role in kernel coloration, we analyzed carotenoid metabolic profiles in representative materials of yellow and white kernels at different grain maturation stages. The analysis revealed a total of 23 carotenoid components, with lutein, zeaxanthin, violaxanthin, violaxanthin myristate, and lutein palmitate identified as the predominant compounds. Lutein emerged as the principal carotenoid in both yellow and white kernel materials, constituting 76-80.8% and 24.5–67.6% of the total carotenoids, respectively, with notably higher abundance in yellow kernels. Zeaxanthin followed as the second most common carotenoid in yellow kernels, representing 7.9–11.8% of total carotenoids, whereas in white kernels, it comprised only 2.7–4.9%. Violaxanthin myristate, the second predominant carotenoid in white kernels, was found to constitute 3.6–58.1% of the total carotenoids, in contrast to 1.1–7.7% in yellow kernels. Interestingly, violaxanthin concentrations were higher in white kernel (2.1–13.5%) than in yellow kernel (0.5–5.4%). The lutein palmitate content ranged from 0.6 to 1.8% in yellow kernel and 0-1.2% in white kernel (Fig. 3). Comparative analysis revealed pronounced disparities in carotenoid profiles, particularly in lutein and zeaxanthin concentrations, between yellow and white kernel materials during grain maturation phases, with yellow kernels exhibiting notably higher levels of both compounds. During grain maturation, levels of lutein and zeaxanthin consistently declined across all the tested materials. Specifically, in the white kernels, these levels significantly decreased to a minimum at the S5 development stage, likely contributing to the white color phenotype. At stage S3, lutein concentration was observed to peak at 33.8 $\mu g \cdot g^{-1}$ in yellow kernel materials Y-4, thereby surpassing the 20.0 $\mu g \cdot g^{-1}$ found in its parent variety, GBS. In white kernel, the maximum lutein level of 12.5 $\mu g \cdot g^{-1}$ was found to occur in the parent variety NMB, but this decreased dramatically to 1 $\mu g \cdot g^{-1}$ by the maturation stage S5.

Transcriptome analysis at different maturation stages in yellow and white kernel

For the parental varieties GBS and NMB, as well as representatives of yellow and white kernel materials from the F_6 population (Y-4, Y-10, W-1, W-7), samples were collected at the mid-grain filling stage (S3), late grain filling stage (S4), and final grain filling stage (S5) for transcriptome sequencing. Following quality control assessment of the sequencing outcomes, we obtained 433.06Gb of high-quality sequencing data, with each sample exceeding 6Gb. The proportion of Q30 bases exceeded 90%, while the GC content ranged from 54.84 to 57.23% (Table S7). A minimum of 78% of reads from each sample were successfully mapped to the reference genome, with 66% uniquely mapped. Notably, up to 12% of reads mapped multiple times (Table S8). The correlation and principal component analyses (PCA) revealed that the biological replicates of identical materials at the same developmental stage exhibited high similarity (Fig. S2). Conversely, distinct differences were apparent among different groups and developmental stages, indicating strong specificity (Fig. S3). Through cross-comparison of the test data from various grain filling stages of yellow and white kernel materials ($|\log_2$Fold change (FC)$| > 1$, False Discovery Rate (FDR) < 0.05), a total of 16,876 differentially expressed genes were identified. To explore the genes contributing to the phenotypic color variation between yellow and white kernels, Venn diagrams were generated to illustrate differentially expressed genes at grain filling stages S3, S4, and S5 for the following comparisons: (1) GBS vs. NMB, GBS vs. W-1, GBS vs. W-7; (2)

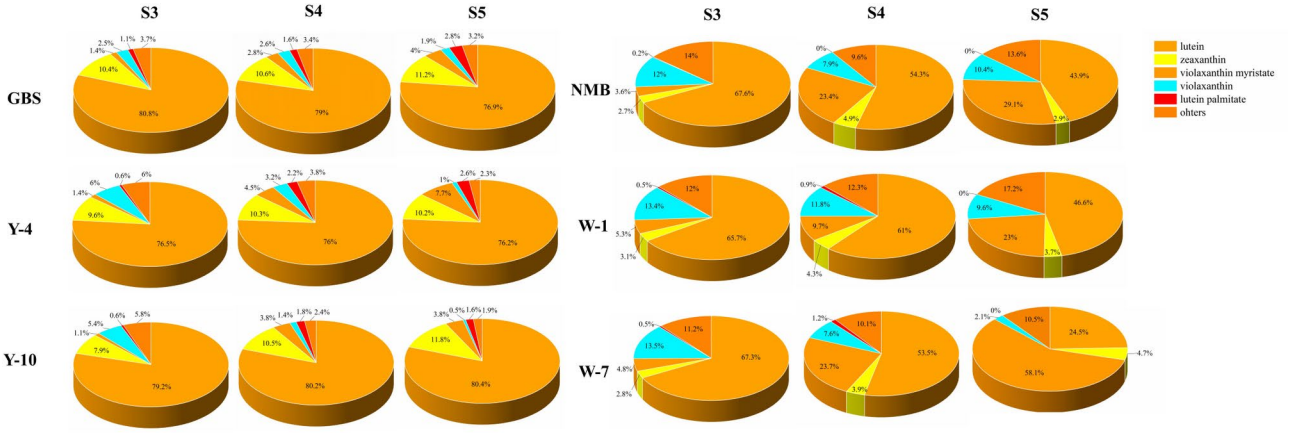

Fig. 3 Analysis of carotenoid composition in yellow and white kernel materials across different grain maturation stages

Y-4 vs. NMB, Y-4 vs. W-1, Y-4 vs. W-7; (3) Y-10 vs. NMB, Y-10 vs. W-1, Y-10 vs. W-7. The results indicated that GBS and three types of white kernel materials exhibited 547, 703, and 696 shared differentially expressed genes at the S3 ~ S5 stages, respectively. Similarly, Y-4 and the same three types of white kernel materials showed 398, 227, and 297 shared differentially expressed genes at the S3 ~ S5 stages, respectively. Likewise, Y-10 alongside these three white kernel materials displayed 583, 650, and 1329 shared differentially expressed genes at the S3 ~ S5 stages, respectively (Fig. S4). By integrating three sets of differentially expressed genes from diverse time intervals and conducting enrichment analysis using the KEGG database, the findings showed significant enrichment of terpenoid compounds, known as precursors to carotenoid synthesis, across all stages. Additionally, pathways such as flavonoid biosynthesis, cytochrome P450, and diterpenoid biosynthesis were notably enriched in two of the periods (Fig. S5). A thorough analysis revealed that gene expression variations between yellow and white kernel during the grain maturation were closely associated with pigment-related pathways, including those related to flavonoids, anthocyanins, and terpenes. This implies that variations in kernel color may be influenced by a variety of pigment-associated metabolites. Furthermore, enrichment frequency analyses indicated a significant six-fold enrichment of terpenoid compounds, which serve as precursors to carotenoid synthesis, during the S3 to S5 stages. Additionally, cytochromes involved in carotenoid metabolism exhibited a two-fold enrichment, suggesting that carotenoids may play a crucial role in determining kernel color.

Transcriptome-metabolome analysis

We performed KEGG pathway enrichment analyses on distinct carotenoid components identified at different stages of grain maturation and differential expression genes from transcriptome analyses. Subsequently, their values were normalized and collectively visualized on the carotenoid biosynthesis pathways (ko00906). The results are shown in Fig. 4, with a total of 32 differentially expressed genes and 9 different metabolites enriched. These include genes of *PSY*, *PDS*, *CRTISO*, *LCYE*, *CYP97C1*, *CHYB*, *CYP97A3*, *DWARF27*, *ZEP*, *VED*, and *CCDs*. *SiPSY1*, as the primary rate-limiting enzyme gene in the carotenoid metabolic pathway, demonstrates a positive correlation with the content of most differential metabolites. Throughout all stages of grain maturation, the expression level of *SiPSY1* is consistently higher in yellow kernel materials compared to white materials. As grains mature, *SiPDS* expression levels decrease in both yellow and white materials, while *SiCRTISO* expression levels gradually increase. Particularly during the S3 and S4 stages, *SiLCYE* expression levels show a significant increase in yellow kernels compared to white kernels, directing a predominant flux of the carotenoid metabolic pathway towards the α-branch in yellow kernels and consequently synthesizing significant amounts of lutein.

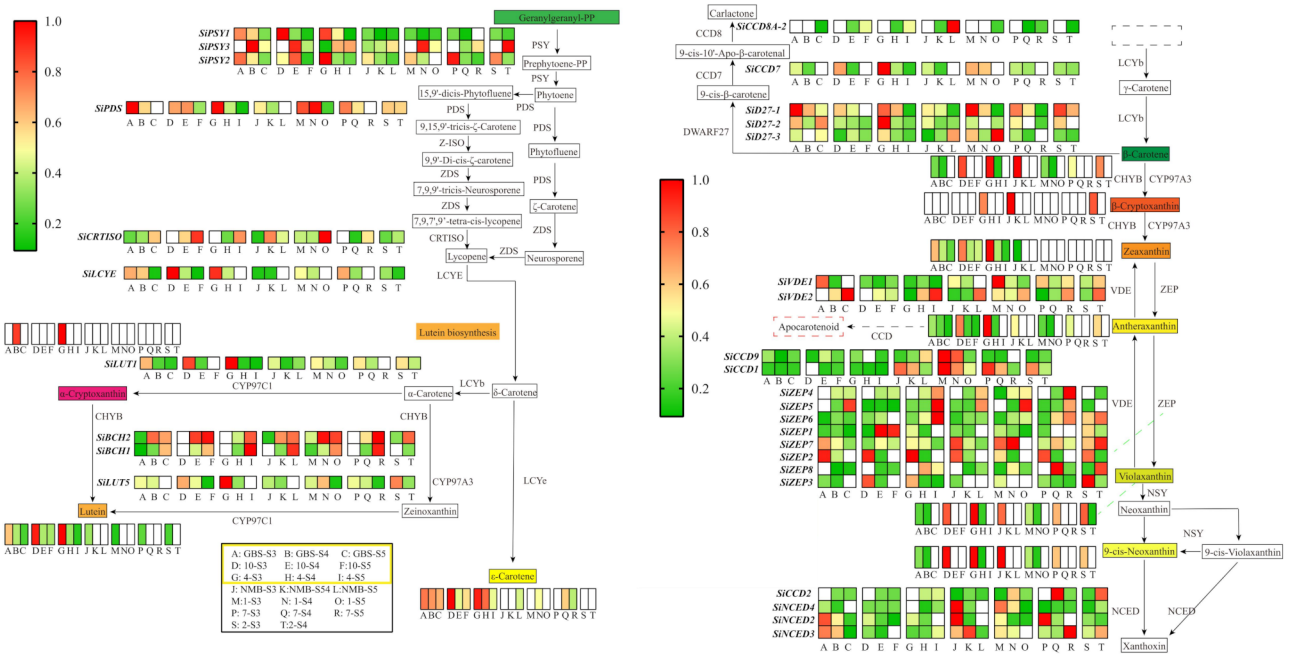

Fig. 4 Heat map of carotenoid biosynthesis pathway

SiCYP97C1 and *SiCYP97A3*, key enzyme genes for lutein synthesis, exhibit a positive correlation with lutein accumulation to some extent. They also exhibit consistent variation trends in relation to differential metabolites, such as β-carotene and zeaxanthin. Among the *ZEP* family, the differential expression genes are most enriched throughout the entire pathway, with *SiZEP2* and *SiZEP3* expressions showing a high degree of positive correlation with most differential metabolites. Meanwhile, the expressions of *SiZEP4*, *SiZEP5*, *SiZEP7*, and *SiZEP8* are higher than those in yellow kernel materials during most stages of white kernel grain maturing. Zeaxanthin is catalyzed by ZEP to sequentially generate antheraxanthin and violaxanthin via two epoxidation reactions. This may explain why white kernels accumulate more violaxanthin while yellow kernels accumulate more zeaxanthin. It is noteworthy that *SiCCD1*, identified as another key carotenoid degradation enzyme gene, shows a significantly higher expression in white kernel compared to yellow kernel during grain maturation. Conversely, the enzyme gene *SiNCED*, also involved in carotenoid degradation, showed a different expression pattern. *SiCCD1* demonstrated an increasing expression trend as the grains gradually mature across all materials, with this upward trend being more pronounced in white kernel materials than in yellow kernel materials. On the other hand, *SiNCED* expression levels generally declined as the grains matured, yet they remained consistently higher in white materials than in yellow ones. In summary, the maturation process of foxtail millet grains involves both the synthesis and degradation of carotenoids. The continuous high expression levels of *SiPSY1* and *SiLCYE* in yellow kernels, along with their consistently low expression levels in white kernels, and the sustained low activity of *SiCCD1* in the carotenoid degradation pathway during all stages of grain maturity in yellow kernels compared to its persistent high activity in white kernels, primarily contribute to the variation in carotenoid content between yellow and white kernels. This variation is responsible for the observable differences in the kernel color phenotype of foxtail millet.

In our study, by integrating GWAS, BSA-seq, and integrated transcriptome-metabolome analysis to identify the key genes responsible for kernel color formation in foxtail millet, *SiPSY1* and *SiCCD1* emerged as crucial genes. As a result, we suggest that variations in lutein and zeaxanthin content primarily contribute to the observed phenotypic differences between yellow and white kernel varieties. Furthermore, *SiPSY1* and *SiCCD1* have been identified as pivotal genes involved in regulating carotenoid accumulation in foxtail millet.

The comparison of *SiPSY1* promoter activity between yellow and while kernel

The outcomes from GWAS, BSA-seq, and integrated transcriptome-metabolome analysis collectively underscore *SiPSY1*'s pivotal role in regulating kernel color formation in foxtail millet. Additionally, haplotype analysis revealed that genetic variations within the promoter region of *SiPSY1* may impact the overall carotenoid content. To delve into these variations in the promoter region and their implications on *SiPSY1* gene expression, we selected extremely yellow and white kernel materials from the F_6 generation of GBS and NMB hybrids, as well as the yellow varieties JG21, CN35, GBS, and the white varieties ZSG as experimental subjects. After extracting DNA from these samples, we successfully cloned a 1580 bp fragment of the promoter upstream of the *SiPSY1* gene's start codon. Comparative analysis coupled with resequencing data indicated that natural populations with the HAP2 genotype, which have higher total carotenoid content, correspond to yellow varieties. Conversely, those with the HAP1 genotype, displaying lower total carotenoid content, aligned with white varieties (Fig. 5a). Among them, the mutation (A/G) at 364 bp position caused significant increase in a cis-element related to abscisic acid ABRE (ACGTG) and a light-responsive element G-box (CACGTT) in white kernel. Moreover, cloning enabled the identification of novel variations at three sites. Specifically, a 29 bp insertion at the 856 bp position leads to the addition of a light-responsive element (G-box, TACGTG) and a gibberellin-responsive element (P-box, CCTTTTG) into in white kernel. Interestingly, distinct variations identified at 1211 bp (A/T) and 1394 bp (C/G) did not affect the cis-elements (Fig. 5b).

We constructed recombinant vectors harboring promoter fragments derived from white and yellow kernels (W and Y, respectively), along with sequences from each mutation site. These constructs were inserted into the pGreenII 0800-LUC binary vector and transiently expressed in rice protoplasts (Fig. 6a). Analysis revealed that the promoter activity of the *SiPSY1* gene, vital in the carotenoid biosynthesis pathway, was 6.5-fold higher in yellow kernels compared to white kernels which demonstrating a statistically significant difference (Fig. 6b). Furthermore, compared to yellow kernels, the SNP mutation (A1 construct with A to G) at position 364 bp and the Indel (A2 construct with a 29 bp insertion) at position 856 bp within the promoter fragment significantly reduced its relative expression activity. Remarkably, the suppression effect exerted by the A1 promoter was more pronounced (Fig. 6b). We propose that the variation in *SiPSY1* promoter sequences between yellow and white kernels significantly influences the differential expression of the *SiPSY1* gene during grain maturation, thereby

Fig. 5 Analysis of *SiPSY1* gene sequence characteristics between yellow and white kernel materials. **(a)** Alignment analysis of *SiPSY1* promoter sequences between yellow and white kernel varieties. **(b)** Prediction of cis-acting elements within the *SiPSY1* gene promoter in yellow and white kernels

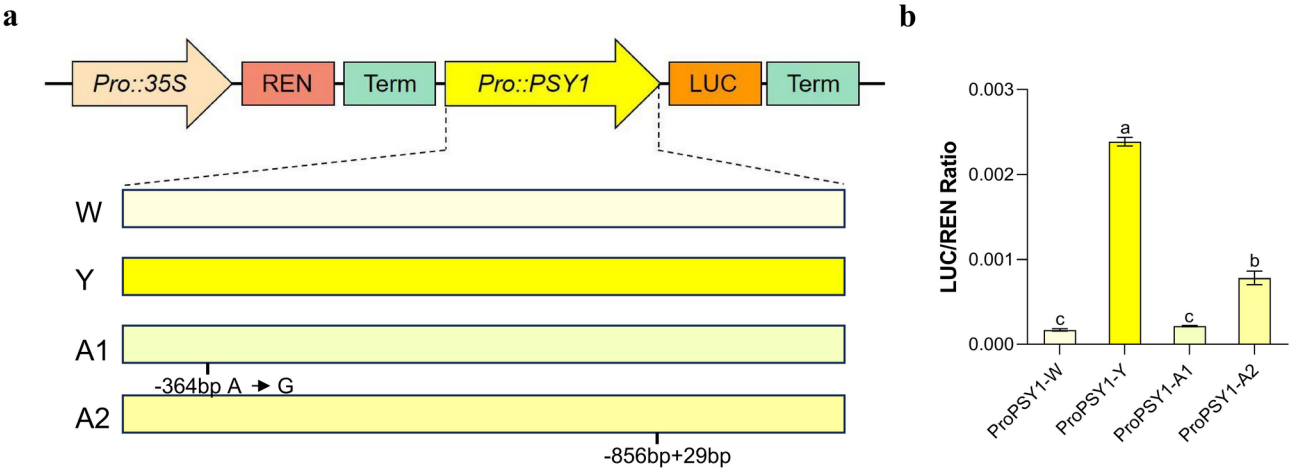

Fig. 6 Comparison of promoter activity of *SiPSY1*. **(a)** Vectors containing promoter fragments from white (W), yellow kernels (Y), the SNP mutation (A1 construct with A to G) at position 364 bp, and the Indel (A2 construct with a 29 bp insertion) at position 856 bp, respectively. **(b)** Comparison of the promoter activities of *SiPSY1*. Significant differences ($P < 0.05$) were indicated by different letters following one-way ANOVA with Duncan's multiple range test

playing a crucial regulatory role in carotenoid accumulation in kernels and the formation of kernel color in foxtail millet.

Discussion

Kernel color characteristic is a quantitative inherited trait

Kernel color is a key quality indicator for foxtail millet, significantly influencing consumer preferences. A more intense yellow color is associated with enhanced aroma, visual appeal, and palatability when cooked. The primary component of this yellow pigment is carotenoids, which are essential for human health as they serve as precursors to Vitamin A, which the body cannot produce on its own and must obtain from external sources. Natural carotenoids offer multiple health benefits, such as protecting vision, boosting the immune system, neutralizing free radicals, preventing various cancers, and aiding in the treatment of oral ulcers and skin diseases [29]. Yang

et al. [30] conducted a comprehensive analysis of yellow pigment levels in 169 foxtail millet varieties, finding significant variation ranging from 1.01 to 19.55 $mg{\cdot}kg^{-1}$, with notable differences across regions. In this study, 201 foxtail millet varieties from China's core germplasm resources were cultivated, and the total carotenoid content in the kernels was measured post-harvest. The results revealed considerable variation among the germplasm resources, ranging from 4.07 to 27.04 $mg{\cdot}kg^{-1}$. This variability is influenced by annual environmental and climatic conditions. Consequently, kernel color is a quantitatively inherited trait governed by multiple genes.

SiPSY1 and *SiCCD1* are the key regulatory factor in determining the color differences between yellow and white kernel in Foxtail millet

Integrating GWAS, BSA-seq, and multi-group analysis constitutes a robust approach for swiftly elucidating the genetic underpinnings of complex quantitative traits. GWAS are widely used to identify genes related to important crop traits. The completion of foxtail millet's whole-genome sequencing provides a foundation for functional genomics studies [31, 32]. Yang et al. [33] employed both second and third-generation sequencing methods to produce a high-quality genome sequence of the ultra-early maturing foxtail millet mutant, *xiaomi*. This progress was found to significantly improve the potential for identifying functional genes in foxtail millet. Jia et al. [34] conducted a haplotype analysis of genome variations in 916 foxtail millet varieties and performed a genome-wide association study on key agronomic traits. They created a haplotype map of foxtail millet using SNPs and identified 512 loci across five latitudinal environments that are linked to 47 agronomic traits, including plant structure, yield, flowering time, and disease resistance. This provides a valuable basis for further genetic research and improvement of foxtail millet. Tian et al. [35] employed an RILs population derived from a cross between yellow and brown hull varieties to identify a significant quantitative trait locus (QTL) on chromosome 1, *QHC.czas1*. They also discovered a gene involved in glume color, *Si1g06530*, which encodes a SANT/Myb protein. Li et al. [36] shed light on the genetic basis of specific metabolite changes during foxtail millet domestication by integrating whole-genome resequencing, transcriptomics, metabolomics, and inflammation response analyses. Their analysis of GWAS signals linked to the b* value, which reflects kernel yellowness, identified *SiPSY1* as a critical gene influencing kernel color domestication. In this study, *SiPSY1* and *SiPSY2* were identified through a GWAS to explore total carotenoid content in kernel, analyzing resequencing data from 201 foxtail millet varieties. We used the F_6 RIL population derived from a cross between the yellow kernel variety GBS and the white kernel variety NMB. To categorize based on kernel color, we created two pools, each containing 130 individuals, for the yellow and white kernel groups. BSA-seq analysis identified two significant peaks, both surpassing the 99% confidence threshold, on chromosomes 3 and 4. In these regions, we specifically targeted the key enzyme gene family involved in carotenoid biosynthesis, selecting *SiPSY1* and *SiPSY2*, along with the carotenoid cleavage dioxygenase gene *SiCCD1*. Our previous research into gene expression patterns in the carotenoid metabolic pathway during grain development revealed that *SiPSY1* consistently had higher expression levels at all stages in yellow kernel varieties compared to white kernel varieties, while *SiCCD1* consistently exhibited lower expression [6]. In this study, joint transcriptomic-metabolomic analysis showed that *SiPSY1* expression levels were significantly higher in yellow kernels than in white kernels, positively correlating with higher concentrations of most differential carotenoid metabolites. On the other hand, *SiCCD1* expression was consistently higher in white kernels across all stages of grain maturation, leading to substantial degradation of lutein and zeaxanthin. In our recent study on the *SiCCD1* gene, we isolated and purified the SiCCD1 protein and co-cultured it in vitro with lutein and zeaxanthin standards to observe enzyme-catalyzed degradation reactions. Using High-Performance Liquid Chromatography (HPLC), we identified and analyzed the enzyme reaction products, revealing a 26.53% reduction in lutein content and an 18.11% reduction in zeaxanthin content. These results demonstrated that the SiCCD1 protein effectively catalyzes the degradation of lutein and zeaxanthin in foxtail millet, with a stronger impact on lutein [37]. The expression of the *ZmPSY1* gene in yellow maize kernels showed a positive correlation with the concentrations of total carotenoids, β-cryptoxanthin, and zeaxanthin in the endosperm, whereas the *PSY1* gene's expression was negligible in white maize and rice kernels [38]. The second generation of 'Golden Rice' was created through genetic engineering techniques by transferring *ZmPSY1* from maize to rice. This process markedly enhanced the concentration of carotenoids, notably β-carotene, within the rice grains, achieving an increase in β-carotene levels by 23 times relative to its precursor [39]. In maize, *CCD1* is the second most influential quantitative trait locus (QTL) affecting total carotenoid content, following *PSY1*. At each stage of kernel development, the copy number of *CCD1* was found to be strongly positively correlated with its expression levels [40]. The variability and expression patterns of pigmentation in yellow and white maize kernels are linked to copy number variations of the *CCD1* gene [27, 41]. In biparental maize populations, the *CCD1* gene regulates the total carotenoid and lutein levels in yellow to orange maize kernels [42]. The findings suggest that

PSY1 and *CCD1* are functionally conserved across cereal crops. The *SiPSY1* and *SiCCD1* genes also play a crucial role in regulating carotenoid accumulation in foxtail millet grains.

The influence of *SiPSY1* promoter activity on gene expression

Promoters are crucial regulatory elements in plant gene expression, affecting both the timing and extent of gene transcription, which in turn influences gene functionality. In this study, extremely colored kernels (yellow and white) from the GBS and NMB F_6 hybrid generation, along with yellow (JG21 CN35 JG21) and white (NMB ZSG) kernel varieties, were used as templates. We successfully cloned approximately 1580 bp of the promoter sequence upstream of the *SiPSY1* gene's start codon. Analysis of the *SiPSY1* promoter sequence revealed two additional light-responsive elements, called G-boxes, a cis-element related to abscisic acid (ABRE, ACGTG), and a single gibberellin-responsive P-box in white kernel samples, which were not present in the yellow kernel samples. Members of the bHLH transcription factor family, particularly Phytochrome Interacting Factors (PIFs), are known to regulate chlorophyll and carotenoid biosynthesis, both key components of chloroplast development. This regulation occurs by adjusting the expression level of the *PSY* gene in response to light signals. For instance, PIFs directly bind to the G-box element in the *PSY* promoter to suppress the expression of specific genes in Arabidopsis. Furthermore, HY5, a transcription factor from the bZIP family, directly interacts with PIFs at the G-box site, creating an activation-inhibition module targeting a single cis-element and working antagonistically. This dynamic system regulates carotenoid synthesis by responding to external signals, ultimately influencing the transcriptional activity of the *PSY* gene [43]. During the maturation of tomato fruits, PIFs can regulate carotenoid biosynthesis through a mechanism similar with that in Arabidopsis. By binding to its promoter, PIF1a inhibits *PSY1* gene expression, reducing *PSY1* activity and limiting carotenoid synthesis. As tomatoes ripen, chlorophyll degradation lessens the shading of tissues, leading to PIF protein breakdown. This process fosters carotenoid accumulation, resulting in the fruit's color transition from green to red [44]. In this study, dual-luciferase reporter gene assays demonstrated that the *SiPSY1* promoter activity is 6.5 times higher in yellow kernels than in white kernels. This significant difference is attributed to an A/G variant at 384 bp and a 29 bp insertion/deletion (indel) at 856 bp within the *SiPSY1* promoter region. These changes lead to the presence of two additional light-responsive G-box elements in white kernel materials. This finding that these variations in the *SiPSY1* promoter sequence are key determinants of differences in *SiPSY1* expression, influencing carotenoid accumulation in foxtail millet and contributing to the variation in coloration between yellow and white kernel.

Conclusions

This study elucidated the genetic and molecular mechanisms governing carotenoid accumulation in foxtail millet kernels, revealing *SiPSY1* and *SiCCD1* as key regulators of color variation. A SNP (G/A) at 364 bp and a 29 bp Indel at 856 bp in the *SiPSY1* promoter reduced its expression, thereby determining the phenotypic distinction between yellow and white kernels. These findings will advance molecular breeding efforts to improve kernel quality in foxtail millet. Further utilizing genomic information to develop molecular breeding strategies aimed at enhancing carotenoid accumulation in foxtail millet kernels, such as marker-assisted selection or genomic selection, will facilitate the development of high-nutrient foxtail millet varieties with improved nutritional quality.

Abbreviations

GWAS	Genome-Wide Association Study
BSA-seq	Bulk Segregant Analysis sequencing
GGPP	Geranylgeranyl pyrophosphate
RIL	Recombinant Inbred Line
GBS	Genotyping-by-Sequencing
SNP	Single-nucleotide polymorphism
LD	Linkage disequilibrium
QD	Quality/Depth
UPLC	Ultra-performance liquid chromatography system
MS/MS	Tandem mass spectrometry
PCA	Principal Component Analysis
FPKM	Fragments Per Kilobase of transcript per Million mapped reads
FDR	False Discovery Rate
FC	Fold change
HPLC	High-Performance Liquid Chromatography
QTL	Quantitative trait locus
PIFs	Phytochrome Interacting Factors

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-025-06585-9.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Supplementary Material 4

Supplementary Material 5

Supplementary Material 6

Acknowledgements

We would like to thank Professor Rupert Fray for discussions and suggestions on related experiments.

Author contributions

Bin Zhang and Yuanhuai Han designed the experiment and constructed the RIL population. Junjie Wang and Qi Ma undertook the GWAS, BSA-seq, transcriptome and metabolome integrated analysis, data analysis and Junjie Wang drafted this manuscript. Yuyang Zhang and Qian Duan coordinated the

carotenoid extraction and gene cloning. Xiaoxi Zhen undertook the promoter activity experiment. Yaoyuan Zhang and Hongying Li provided extensive revision suggestions of the manuscript. All authors read and approved the final manuscript.

Funding
This research was funded by National Natural Science Foundation of China (32370411, 31971906), Key Research Project of Shanxi Province (202302140601003), Science and Technology Special Project in Shanxi Province (202101140601027), Research Project Supported by Shanxi Scholarship Council of China (2022 – 108), High-level Foreign Experts Introduction Project of China (G2022004009L) and CGGL Technology Development Fund (YDZJSX2022B007).

Data availability
All data generated or analyzed during this study are included in this manuscript and its supplementary information files. Illumina sequencing data are available at the Sequence Read Archive (SRA) under accession PRJNA1208398.

Declarations

Ethics approval and consent to participate
Ethical approval was not required for this study. Foxtail millet plants were grown at the experimental station of Shanxi Agriculture University (Shanxi, China) for sample collection. No specific permits were required for the field studies. The research conducted in this study did not require approval from an ethics committee nor did it involve any human or animal subjects.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 13 January 2025 / Accepted: 18 April 2025
Published online: 25 April 2025

References

1. Diao XM, Schnable J, Bennetzen JL, Li JY. Initiation of Setaria as a model plant. Front Agricultural Sci Eng. 2014;1:16–20.
2. Diao XM, Jia GQ. Origin and domestication of Foxtail millet. Springer Int Publishing. 2017. 61–72.
3. Xiang JL, Zhang M, Apea-Bah FB, Beta T. Hydroxycinnamic acid amide (HCAA) derivatives, flavonoid *C*-glycosides, phenolic acids and antioxidant properties of Foxtail millet. Food Chem. 2019;295:214–23.
4. He L, Zhang B, Wang XC, Li HY, Han YH. Foxtail Millet: nutritional and eating quality, and prospects for genetic improvement. Front Agricultural Sci Eng. 2015;2:124–33.
5. Shen R, Yang SP, Zhao GH, Shen Q, Diao XM. Identification of carotenoids in Foxtail millet (*Setaria italica*) and the effects of cooking methods on carotenoid content. J Cereal Sci. 2015;61:86–93.
6. Zhang B, Liu J, Cheng L, Zhang YY, Hou SY, Sun ZX, et al. Carotenoid composition and expression of biosynthetic genes in yellow and white Foxtail millet [*Setaria italica* (L.) Beauv]. J Cereal Sci. 2019;85:84–90.
7. Nisar N, Li L, Lu S, Khin NC, Pogson BJ. Carotenoid metabolism in plants. Mol Plant. 2015;8:68–82.
8. Sanlier N, Yildiz E, Ozler E. An overview on the effects of some carotenoids on health: lutein and Zeaxanthin. Curr Nutr Rep. 2024;13:828–44.
9. Ojobor CC, O'Brien GM, Siervo M, Ogbonnaya C, Brandt K. Carrot intake is consistently negatively associated with cancer incidence: A systematic review and meta-analysis of prospective observational studies. Crit Rev Food Sci Nutr. 2025;65(5):1009–21.
10. Sun TH, Rao S, Zhou XS, Li L. Plant carotenoids: recent advances and future perspectives. Mol Hortic. 2022;2:3.
11. Maass D, Arango J, Wast F, Beyer P, Welsch R. Carotenoid crystal formation in Arabidopsis and Carrot roots caused by increased phytoene synthase protein levels. PLoS ONE. 2009;4:e6373.
12. Giulian G. Provitamin A biofortification of crop plants: a gold rush with many miners. Curr Opin Biotechnol. 2017;44:169–80.
13. Sun T, Tadmor Y, Li L. Pathways for carotenoid biosynthesis, degradation, and storage. Methods Mol Biology. 2020;2083:3–23.
14. Sandmann G. Diversity and origin of carotenoid biosynthesis: its history of Coevolution towards plant photosynthesis. New Phytol. 2021;232:479–93.
15. Bai L, Kim EH, DellaPenna D, Brutnell TP. Novel lycopene epsilon cyclase activities in maize revealed through perturbation of carotenoid biosynthesis. Plant J. 2009;59:588–99.
16. Kim SH, Kim YH, Ahn YO, Ahn MJ, Jeong JC, Lee HS, et al. Downregulation of the lycopene ε-cyclase gene increases carotenoid synthesis via the β-branch-specific pathway and enhances salt-stress tolerance in Sweetpotato Transgenic calli. Physiol Plant. 2013;147:432–42.
17. Kim J, Smith JJ, Tian L, Dellapenna D. The evolution and function of carotenoid hydroxylases in *Arabidopsis*. Plant Cell Physiol. 2009;50:463–79.
18. Arango J, Jourdan M, Geoffriau E, Beyer P, Welsch R. Carotene hydroxylase activity determines the levels of both alpha-carotene and total carotenoids in orange carrots. Plant Cell. 2014;26:2223–33.
19. Jahns P, Holzwarth AR. The role of the xanthophyll cycle and of lutein in photoprotection of photosystem II. Biochim Et Biophyssica Acta Bioenergetics. 2012;1817:182–93.
20. Perreau F, Frey A, Effroy-Cuzzi D, Savane P, Berger A, Gissot L, et al. ABSCISIC-ACID-DEFICIENT4 has an essential function in both cis-violaxanthin and cis Neoxanthin synthesis. Plant Physiol. 2020;184:1303–16.
21. Auldridge ME, Block A, Vogel JT, Dabney-Smith C, Mila I, Bouzayen M, et al. Characterization of three members of the Arabidopsis carotenoid cleavage dioxygenase family demonstrates the divergent roles of this multifunctional enzyme family. Plant J. 2006;45:982–93.
22. Zhong Y, Pan X, Wang R, Xu J, Guo J, Yang T, et al. *ZmCCD10a* encodes a distinct type of carotenoid cleavage dioxygenase and enhances plant tolerance to low phosphate. Plant Physiol. 2020;184:374–92.
23. Beltrán J, Wurtzel ET. Carotenoids: resources, knowledge, and emerging tools to advance apocarotenoid research. Plant Sci. 2025;350:112298.
24. Alder A, Jamil M, Marzorati M, Bruno M, Vermathen M, Bigler P, et al. The path from beta-carotene to Carlactone, a strigolactone-like plant hormone. Science. 2012;335:1348–51.
25. Amaya I, Roldán-Guerra FJ, Ordóñez-Díaz JL, Torreblanca R, Wagner H, Waurich V, Olbricht K, Moreno-Rojas JM, Sánchez-Sevilla JF, Castillejo C. Differential expression of *CCD4(4B)* drives natural variation in fruit carotenoid content in strawberry (*Fragaria* spp). Plant Biotechnol J. 2025;23(3):679–91.
26. García-Limones C, Schnäbele K, Blanco-Portales R, Bellido ML, Caballero JL, Schwab W, et al. Functional characterization of FaCCD1: A carotenoid cleavage dioxygenase from strawberry involved in lutein degradation during fruit ripening. J Agricultural Food Chemmistry. 2008;56:9277–85.
27. Adami M, Franceschi PD, Brandi F, Liverani A, Giovannini D, Rosati C, Dondini L, Tartarini S. Identifying a carotenoid cleavage dioxygenase (*CCD4*) gene controlling yellow/white fruit flesh color of Peach. Plant Mol Biology Report. 2013;31:1166–75.
28. Dutta S, Muthusamy V, Chhabra R, Baveja A, Zunjare RU, Mondal TK, et al. Low expression of carotenoids cleavage *dioxygenase 1* (*ccd1*) gene improves the retention of provitamin-A in maize grains during storage. Mol Genet Genomics. 2021;296:141–53.
29. Asharani VT, Jayadeep A, Malleshi NG. Natural antioxidants in edible flours of selected small millets. Int J Food Prop. 2010;13:41–50.
30. Yang YB, Guan YA, Qin L, Shi H, Wang HL, Zhang HW. The studies on yellow pigment content and appearance quality of millet from different regions. J Chin Cereals Oils Association. 2012;27:14–9.
31. Bennetzen JL, Schmutz J, Wang H, Percifield R, Hawkins J, Pontaroli AC, et al. Reference genome sequence of the model plant Setaria. Nat Biotechnol. 2012;30:555–61.
32. Zhang GY, Liu X, Quan ZW, Cheng SF, Xu X, Pan SK, et al. Genome sequence of Foxtail millet (*Setaria italica*) provides insights into grass evolution and biofuel potential. Nat Biotechnol. 2012;30:549–54.
33. Yang ZR, Zhang HS, Li XK, Shen HM, Gao JH, Hou SY, et al. A mini Foxtail millet with an *Arabidopsis*-like life cycle as a C_4 model system. Nat Plants. 2020;6:1167–78.
34. Jia GQ, Huang XH, Zhi H, Zhao Y, Zhao Q, Li WJ, et al. A haplotype map of genomic variations and genome-wide association studies of agronomic traits in Foxtail millet (*Setaria italica*). Nat Genet. 2013;45:957–61.
35. Tian BH, Zhang LX, Hu JH, Liu YL, Zhou LL, Ping WC, et al. Genetic characterization of hull color using BSR-Seq and genome re-sequencing approaches in Foxtail millet. Front Plant Sci. 2022;13:1019496.

36. Li X, Gao J, Song J, Guo K, Hou S, Wang X, et al. Multi-omics analyses of 398 Foxtail millet accessions reveal genomic regions associated with domestication, metabolite traits and anti-inflammatory effects. Mol Plant. 2022;15:1367–83.
37. He L, Cheng L, Wang JL, Liu J, Cheng JJ, Yang ZR, et al. Carotenoid cleavage dioxygenase 1 catalyzes lutein degradation to influence carotenoid accumulation and color development in Foxtail millet grains. J Agricultural Food Chemmistry. 2022;70:9283–94.
38. da Silva Messias R, Galli V, Dos Anjos e Silva SD, Rombaldi CV. Carotenoid biosynthetic and catabolic pathways: gene expression and carotenoid content in grains of maize landraces. Nutrients. 2014;6:546–63.
39. Paine JA, Shipton CA, Chaggar S, Howells RM, Kennedy MJ, Vernon G, et al. Improving the nutritional value of golden rice through increased pro-vitamin A content. Nat Biotechnol. 2005;23:482–7.
40. Diepenbrock CH, Ilut DC, Magallanes-Lundback M, Kandianis CB, Lipka AE, Bradbury PJ, et al. Eleven biosynthetic genes explain the majority of natural variation in carotenoid levels in maize grain. Plant Cell. 2021;33:882–900.
41. Tan BC, Guan JC, Ding S, Wu S, Saunders JW, Koch KE, et al. Structure and origin of the *white cap* locus and its role in evolution of grain color in maize. Genetics. 2017;206:135–50.
42. LaPorte MF, Vachev M, Fenn M, Diepenbrock C. Simultaneous dissection of grain carotenoid levels and kernel color in biparental maize populations with yellow-to-orange grain. G3 (Bethesda). 2022;12:jkac006.
43. Toledo-Ortiz G, Johansson H, Lee KP, Bou-Torrent J, Stewart K, Steel G, Ret. The HY5-PIF regulatory module coordinates light and temperature control of photosynthetic gene transcription. PLoS Genet. 2014;10:e1004416.
44. Llorente B, D'Andrea L, Ruiz-Sola MA, Botterweg E, Pulido P, Andilla J, et al. Tomato fruit carotenoid biosynthesis is adjusted to actual ripening progression by a light-dependent mechanism. Plant J. 2016;85:107–19.

Publisher's note