Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Mapping of transcription factor footprints unveils a transcriptional repression mechanism of capsanthin biosynthesis in pepper (Capsicum annuum)

Xu Wang

xu.wamg@pku-iaas.edu.cn

Peking University Institute of Advanced Agricultural Sciences

Fan Zhang

Peking University Institute of Advanced Agricultural Sciences

Zeyu Zhang

Peking University Institute of Advanced Agricultural Sciences

Xiaotong Sun

Peking University Institute of Advanced Agricultural Sciences

Zhen Zhuang

Peking University Institute of Advanced Agricultural Sciences

Yaping Tang

Xinjiang Academy of Agricultural Sciences

Jing Wang

Peking University Institute of Advanced Agricultural Sciences

Qing Fan

Peking University Institute of Advanced Agricultural Sciences

Xinjing Wang

Peking University Institute of Advanced Agricultural Sciences

Qinyan Li

Yunnan University

Pengfei pan

Peking University Institute of Advanced Agricultural Sciences

Xiaoyan Liang

Peking University Institute of Advanced Agricultural Sciences

Yulei Sui

Peking University Institute of Advanced Agricultural Sciences

Shengbao Yang

Xinjiang Academy of Agricultural Sciences

Liyu Huang
Yunnan University https://orcid.org/0000-0001-9666-6584

Article

Keywords:

Posted Date: September 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10699801/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

Mapping of transcription factor footprints unveils a transcriptional repression mechanism of capsanthin biosynthesis in pepper (*Capsicum annuum*)

Fan Zhang[1,5], Zeyu Zhang[1,5], Xiaotong Sun[1,5], Zhen Zhuang[1,5], Yaping Tang[1,2,3], Jing Wang[1], Qing Fan[1], Xinjing Wang[1], Qinyan Li[4], Pengfei Pan[1], Xiaoyan Liang[1], Yulei Sui[1], Shengbao Yang[3*], Liyu Huang[4*], and Xu Wang[1*]

[1] Peking University Institute of Advanced Agricultural Sciences, Shandong Laboratory of Advanced Agricultural Sciences in Weifang, Shandong 261325, China.
[2] College of Horticulture Science and Engineering, Shandong Agricultural University, Tai'an 271018, China
[3] Key Laboratory of Gene Research and Genetic Improvement of Xinjiang Characteristic Fruits and Vegetables, Xinjiang Academy of Agricultural Sciences, Urumqi 830091, China.
[4] School of Agriculture, Yunnan University, Kunming, Yunnan 650091, China.
[5] These authors contributed equally: Fan Zhang, Zeyu Zhang and Xiaotong Sun
[*] Correspondence: ysb.jack@163.com (S.Y.), lyhuang@ynu.edu.cn (L.H.) and xu.wang@pku-iaas.edu.cn (X.W.)

Abstract

Capsanthin is a rare, oxygenated carotenoid of high nutritional and industrial value, which rapidly accumulates as pepper fruit ripens, resulting in the transition of color from green to red. However, the regulatory mechanisms that control capsanthin biosynthesis at the transcriptional level remain incompletely resolved. Here we comprehensively surveyed the dynamic state and accessibility of chromatin in pepper pericarps across color transition and mapped the upstream transcription factor (TF) footprints of capsanthin biosynthetic pathway genes (CBGs). We thereby unveiled a *trans*-regulatory landscape for the complete set of CBGs. Stable overexpression of selected TFs in peppers nominated two DNA binding with one finger (DOF) TFs as previously unidentified negative regulators of capsanthin accumulation. Lines of evidence further demonstrated that both DOF TFs repressed expression of *Geranylgeranyl Pyrophosphate (GGPP) synthase 1* (*GGPPS1*) gene and reduced GGPP contents *in vivo*. Overall, this study provides an extensive resource for understanding the genetic basis to modulate capsanthin biosynthesis at the transcriptional level and elucidates the first repressive mechanism that restrains capsanthin biosynthesis through limiting carotenoid precursor supply.

Introduction

Red fruit color in pepper (*Capsicum spp.*) is the visible outcome of a developmental shift in carotenoid metabolism. During ripening, flux through the plastidial isoprenoid pathway is redirected towards capsanthin and related oxygenated carotenoids, which accumulate to high levels in the pericarp[1]. Capsanthin and its derivative capsorubin are not only valued commercial pigments but also widely used as nutraceutical or pharmaceutical because of their potent antioxidant properties[2, 3]. Therefore,

understanding how capsanthin biosynthesis is spatiotemporally controlled during ripening is central to both elucidating the molecular mechanisms of capsanthin accumulation and improving fruit color trait in pepper.

The principal capsanthin biosynthetic pathway genes (CBGs) in pepper are well documented, such as *Phytoene synthase* (*PSY*)[4, 5, 6], *β-Carotene hydroxylase* (*CHYB*)[7, 8], *Zeaxanthin epoxidase* (*ZEP*)[9] and *Capsanthin-Capsorubin Synthase* (*CCS*)[10]. Malfunction of these genes in pepper causes dramatic change in color of ripe fruits due to the altered composition and content of carotenoids[7, 8, 9]. However, the regulatory layer that coordinates transcription of these genes has remained less clear. Recent studies have identified several TFs as positive regulators for both CBGs expression and capsanthin accumulation, such as *CaBBX20*[11], *DIVARICATA1*[12], *MADS-box TF RIPENING INHIBITOR* (*MADS-RIN*)[13], *CaNAC2*[14] and *CaLSH10*[15]. These studies suggest that capsanthin biosynthesis is governed by a distributed transcriptional network rather than by structural genes alone. Yet repressors and repressive mechanisms that restrain capsanthin biosynthesis are still not known. Collectively, lack of a genome-wide TF-centric *trans*-regulatory landscape for the complete set of CBGs limits our understanding of the regulatory mechanisms of capsanthin biosynthesis and hindering the efforts to enhance capsanthin production through genetic improvement.

Here, to address this gap, we build a TF-centric regulatory framework for capsanthin biosynthesis during pepper color transition by integrating transcriptomic, epigenomic and chromatin-interaction datasets from ripening pericarps and mapping TF footprints of promoters for all CBGs. Therefrom we identify a DNA binding with one finger (DOF) TF-mediated repressive module that targets *Geranylgeranyl Pyrophosphate (GGPP) synthase 1* (*GGPPS1*) gene, an upstream enzyme controlling precursor entry into carotenoid biosynthesis. Functional validation in transgenic pepper demonstrates that two DOF TFs repress *GGPPS1* expression and limit capsanthin accumulation. Our study therefore provides a multi-omics resource for understanding the genetic basis to modulate capsanthin biosynthesis at the transcriptional level and elucidates the first repressive mechanism that restrains capsanthin biosynthesis through limiting carotenoid precursor supply.

Results and Discussion

A multi-omics atlas links chromatin regulation to capsanthin biosynthesis

To interrogate the links between chromatin dynamics and transcription of CBGs throughout color transition, pepper pericarps of inbred line CA59 developed from green mature stage (G) at 33 days post anthesis (DPA) to red ripe stage (R) at 39 DPA, passing through breaker stages (B1-B3) with progressively increased total carotenoids and capsanthin contents (Fig. 1a, b). We harvested pericarps at three stages (G, B2 and R) to generate 9 RNA-seq libraries (transcriptome sequencing) for profiling global gene expression, 6 ATAC-seq libraries (Assay for transposase accessible chromatin with sequencing) and 48 ChIP-seq libraries (chromatin immunoprecipitation followed by sequencing) targeting six histone modifications (H3K4me1, H3K4me3, H3K9ac,

H3K27ac, H3K9me2, H3K27me3) and two RNA polymerase II states (RNAP II Ser2p and RNAP II Ser5p) for mapping chromatin dynamics (Fig. 1c). To improve annotation quality of pepper CA59 reference genome[16], 6 PacBio ISO-seq datasets were generated as well. The transcript isoforms and the number of both total transcripts and long transcripts (> 3Kb) were substantially increased after re-annotation (Supplementary Fig. 1a-c; Supplementary Table 1). Our high-throughput sequencing datasets were of high quality and reproducibility with strong Pearson correlations ($r \geq 0.96$) and enrichment scores (FRiP $\geq$ 0.3; Supplementary Fig. 1d, e). The genome-wide distributions of ATAC-seq and all ChIP-seq signals were consistent with earlier studies (Supplementary Fig. 2)[17, 18].

Using ChromHMM[19], we condensed chromatin accessibility and histone-mark combinations into ten states (S1-S10) representing seven main classes[17, 18]: open chromatin, enhancer-like (enhancer and weak enhancer), promoter-like (promoter and weak promoter), H3K4me1 island, euchromatin, heterochromatin and other (low signal and no signal) states (Fig. 1d). Open chromatin (S4), enhancer (S5) and promoter (S7) states showed the strongest accessibility and were generally associated with higher RNA abundance, whereas weak promoter and repressive states showed lower transcriptional activity (Supplementary Fig. 3a; Supplementary Table 2). H3K4me1 island (S3) was a separate state also found in other plant species without known biological significance [18, 20]. This state generally displayed low chromatin accessibility but high gene expression levels in pepper (Supplementary Fig. 3a; Supplementary Table 2), yet how this feature reflects the activity of H3K4me1 in pepper is not clear. Further analysis revealed that extensive internal chromatin states switches occurred during color transition (Supplementary Fig. 3b, c). For instance, 29% of the chromatin regions in state S4 at stage G shifted into other states at stage B2, and genes associated with these transitions were enriched for hormone, development and photosynthesis-related processes (Supplementary Fig. 3b-e). Thus, the pericarp epigenome was not static during color transition. Instead, color-transition was accompanied by broad regulatory remodeling that could plausibly shape ripening-associated gene expression.

We further found that 17 CBGs did not follow a single co-expression pattern across G, B2 and R stages (Fig. 1e, f). Genes acting at key entry or branch points, including *Geranylgeranyl diphosphate synthase1* (*GGPPS1*), *PSY1*, *CHYB1*, *Lycopene beta cyclase2* (*LCYB2*) and *CCS*, showed higher expression towards the R stage, consistent with increased carotenoid and capsanthin accumulation (Fig. 1f). Other paralogues such as *PSY2*, *CHYB2*, *ZEP*, *CRTISO* and *LCYB1* showed different temporal preferences (Fig. 1f). This divergent pattern argues against a uniform transcriptional switch across the whole pathway. Rather, ripening-induced expression of selected gene family members at varied pathway nodes, probably allows flux to be adjusted without activating all related enzymes equally. Such partitioning is biologically important because several enzymes act at branch points where carotenoid, chlorophyll-related and xanthophyll pathways intersect.

To explore how gene-specific expression might be achieved, we compared the RNA abundance of CBGs with their local chromatin features and found that, for many ripening-induced CBGs, RNA levels were positively associated with chromatin accessibility, H3K4me3, H3K9ac, H3K27ac and RNA polymerase II Ser5P or Ser2P signals (Fig. 1f). Conversely, H3K27me3 and H3K9me2 displayed weaker or negative relationships for subsets of genes (Fig. 1f). Despite this trend-level association does not represent causal relationship, the direction of the associations was consistent with a model in which active histone marks, open chromatin and polymerase II occupancy accompany transcriptional competence at selected pathway genes. Locus-resolved ChromHMM tracks strengthened this interpretation. Representative genes showed distinct mixtures of accessible, enhancer-like, promoter, euchromatic and repressive states around promoter and gene-body regions (Fig. 1g).

We next asked whether three-dimensional chromatin interaction contributes to this regulatory landscape. High-throughput chromosome conformation capture (Hi-C) maps from G and B2 pericarps identified 11,265 and 8,422 significant loops, respectively (Supplementary Fig. 4a). By integrating loop anchors with chromatin-state annotations[17], these contacts could be classified as Promoter-Promoter loops (P-P; 7,456 for stage G and 6,563 for stage B2), Promoter-Enhancer loops (P-E; 3,316 for stage G and 1,744 for stage B2), Enhancer-Enhancer loops (E-E; 493 for stage G and 115 for stage B2) or other loop types (493 for stage G and 115 for stage B2) (Supplementary Fig. 4b-d) , with a median size of ~60 kb for P-P, P-E and E-E loops (Supplementary Fig. 4e). Promoter-associated loops were detected around the TSS of multiple CBGs, including *ZEP, PDS*, *CHYB1*, *GGPPS1*, *PSY1*, *LCYB2*, *CCS*, *LCYE*, *Z-ISO* and *VDE* (Fig. 1h). The local chromatin contact patterns differed between G and B2 pericarps, indicating that the chromatin neighborhoods of these CBGs are remodeled as the onset of fruit color transition. At *PSY1* and *CHYB1*, B2-associated contacts linked TSS with promoter-proximal or distal regions carrying accessibility and active histone marks such as H3K27ac and H3K9ac (Fig. 1i). These data do not prove that the contacted regions are functional enhancers[17]. They do, however, identify stage-responsive regulatory contacts that could help explain why individual pathway genes acquire transcriptional activity at different points during ripening.

Mapping of transcription factor footprints nominates regulators of CBGs

Having defined the *cis*-regulatory context, we next sought the *trans*-regulators that might act through these promoter landscapes. We inferred TF footprints within the 2-kb promoter regions of the 17 CBGs using PRINT, a multiscale footprinting method that detects local protection patterns within accessible chromatin[21]. The resulting maps showed stage-variable promoter occupancy, with genes such as *CCS* and *GGPPS1*, their promoters contained multiple discrete footprints across G, B2 and R stages, which include DNA motifs recognized by multiple TFs, such as MYB-related, MADS-box, ERF, Trihelix and Dof (Fig. 2a; Supplementary Fig. 5a; Supplementary Table 3). Motif analysis linked the promoter footprints to 324 candidate TFs from 32

families (Fig. 2b). MYB, DOF, HD-ZIP, ERF, MIKC_MADS, bZIP, WRKY, NAC and related families connected to multiple CBGs, whereas other families were more restricted (Fig. 2b). The promoters also differed in predicted regulatory complexity, for example, *GGPPS1*, *CCS* and *Z-ISO* were connected to many candidate TFs, whereas *LCYB2* and *CRTISO* had fewer predicted inputs (Fig. 2b). This promoter-centered bipartite network provided a focused candidate space, while retaining an important caveat that motif-supported footprints nominate potential regulators but do not alone establish direct transcriptional control.

We refined this candidate network by combining motif abundance in accessible chromatin with transcription factor expression. Family-level motif proportions in TF footprints remained broadly similar among G, B2 and R pericarps, but transcription factor family expression changed during color transition (Fig. 2c). This distinction is informative because regulatory potential requires both an accessible binding site and expression of the corresponding factor. Gene Ontology (GO) enrichment of the 328 candidates further linked the network to developmental and hormone-response processes (Fig. 2d). These enrichments suggest that capsanthin biosynthesis is embedded in a broader ripening regulatory program rather than controlled by isolated pathway-specific factors. When computed the Pearson correlation coefficient (PCC) of expression patterns between candidate TFs and their linked CBGs, about 60% of TFs showed positive or negative correlation with their target CBGs (PCC $\geq$0.8 or $\leq$-0.8; Fig. 2e). This result indicates that the stage-associated expressions of CBGs are more likely controlled by a group of TFs with diverse expression patterns. Candidate TFs were selected and most of them were highly expressed in exocarp and mesocarp cells in a ripening-induced manner (Fig. 2f; Supplementary Fig. 5b) [22]. The activity to regulate target gene transcription of these TFs were assessed using a previously established luciferase (LUC) reporter assay [23]. Effectors expressing either GFP or TF-GFP fusion proteins were paired with reporters in which *LUC* expression was independently under control of 13 CBG promoters, to co-infiltrate into tobacco leaves (Fig. 2g; Supplementary Table 4). Among 227 tested effector-reporter combinations, 39% significantly altered reporter activity, whereas most had limited effects (Fig. 2h). Strikingly, 85% of active combinations were repressive rather than activating (Fig. 2h). Previous work in pepper has emphasized positive regulators that promote capsanthin accumulation, but the transient screen indicated that repression is widespread among candidate promoter-TFs interactions. Stable overexpression lines for 12 TFs then nominated two positive regulators, *CaTF308* and *CaTF216*, and two negative regulators, *CaTF16* and *CaTF19*, based on red ripe fruit carotenoid and capsanthin phenotypes (Fig. 2i). *CaTF16* and *CaTF19* both encoded DOF transcription factors and were therefore named *CaDOF1* and *CaDOF2*.

CaDOF1 and CaDOF2 repress *GGPPS1* expression to constrain capsanthin accumulation in pepper

During ripening, *CaDOF1* expression declined while *GGPPS1* increased (Fig. 3a). The reciprocal expression pattern was consistent with a repressive relationship,

particularly because CaDOF1 had reduced *GGPPS1* promoter activity in transient assays. *GGPPS1* expression was positively associated with chromatin accessibility, H3K4me3 and H3K27ac, whereas *CaDOF1* showed a more restricted association with H3K27ac (Fig. 3a). These different chromatin contexts suggest that the repressor and its target are themselves governed by distinct regulatory inputs during fruit development. Both genes were highly expressed in open flowers, but their fruit-stage dynamics diverged (Fig. 3b), indicating that the CaDOF1-*GGPPS1* relationship is developmentally context dependent rather than a simple reflection of constitutive tissue expression. To investigate the *in vivo* function of *CaDOF1*, transgenic pepper lines overexpressing *CaDOF1-GFP* were then generated (Fig. 3c) and validated by immunoblot detection of the fusion protein (Fig. 3d) and whole-genome resequencing (WGR) confirming T-DNA insertions (Fig. 3e). The CaDOF1-GFP proteins were localized to the nucleus, consistent with its expected role as a transcription factor (Fig. 3f). Compared with wild type (WT), two *CaDOF1-GFP* overexpression lines (#231 and #208) developed paler fruits at breaker and red ripe stages (Fig. 3g). This phenotypic alteration was accompanied by significantly reduced geranylgeranyl diphosphate (GGPP) and capsanthin contents in red ripe pericarps, with additional changes in several carotenoid intermediates (Fig. 3h). These results placed *CaDOF1* upstream of capsanthin accumulation and pointed to reduced precursor supply as a major outcome of its overexpression. Spatial transcriptomic profiling further revealed that *CaDOF1* overexpression not only repressed capsanthin biosynthesis but also altered pericarp tissue organization (Fig. 3i-n). UMAP clustering and spatial maps resolved ten pericarp-associated cell types in both WT and *CaDOF1-GFP* fruit, including exocarp, mesocarp, endocarp and vascular tissue domains (Fig. 3i, j). *CaDOF1* overexpression altered the proportions of mesocarp-related cell populations but did not erase major tissue identities (Fig. 3k). *CaDOF1* transcripts increased across multiple cell types in the overexpression line (Fig. 3l), and CBGs were broadly reduced in the corresponding spatial domains (Fig. 3m). Differential expression analysis confirmed downregulation of several pathway genes, including *GGPPS1*, *CCS*, *PSY1*, *CHYB1*, *Zeta-carotene desaturase* (*ZDS*), *Phytoene desaturase* (*PDS*), *Violaxanthin de-epoxidase* (*VDE*), *CHYB2* and *ZEP* (Fig. 3n). Thus, *CaDOF1* acts across pericarp cell populations to repress the expression of multiple CBGs while preserving tissue identity.

Among those CaDOF1 downregulated CBGs (Fig. 3m, n), only *CHYB1*, *GGPPS1* and *PSY1* promoters were transcriptionally repressed by CaDOF1 in LUC reporter assays (Fig. 3o), suggesting that these genes could be direct targets of CaDOF1 in pepper. Amplified DNA Affinity Purification sequencing assay (ampDAP-seq) identified a VCTTTW (V=A/C/G, W=T/A) DOF-binding motif in CaDOF1-bound regions (Fig. 3p and Supplementary Fig. 6a), and CaDOF1 occupancy was detected around *GGPPS1* and additional pathway loci, including *PSY2*, *CHYB1*, *CHYB2* and *GGPPS3* (Fig. 3q and Supplementary Fig. 6b). Fluorescent electrophoretic mobility shift assays (fEMSA) showed that the recombinant CaDOF1 DOF domain bound *GGPPS1* promoter probes (Fig. 3r; Supplementary Fig. 6c-f), and microscale thermophoresis

(MST) revealed that mutations in the core motif substantially abolished CaDOF1-DNA binding (Fig. 3s; Supplementary Fig. 6e, f). ChIP-qPCR in *CaDOF1-GFP* fruits confirmed *in vivo* enrichment at *GGPPS1* promoter regions (Fig. 3t). These results indicate that CaDOF1 binds the *GGPPS1* promoter and represses precursor supply for capsanthin biosynthesis *in vivo*. *CaDOF2* showed fruit-stage expression dynamics similar to *CaDOF1* (Supplementary Fig. 7b). *CaDOF2* overexpression reduced *GGPPS1* expression, GGPP and capsanthin contents (Fig. 3u-w and Supplementary Fig. 7). Recombinant CaDOF2 also bound the same *GGPPS1* promoter motifs in fEMSA and MST assays (Supplementary Fig. 8). Together, these findings argue that *CaDOF2* might play a redundant role with *CaDOF1* to constrain capsanthin biosynthesis through limiting carotenoid precursor supply. More broadly, the study establishes a framework for moving from chromatin state maps and transcription factor footprints to functional regulators of a specialized metabolic trait.

Methods

Plant material and growth conditions

For multi-omics profiling, pericarp tissues were collected from the C. annuum inbred line CA59, with seeds and placentas carefully removed. CA59, wild-type (ST-8 and P2205), as well as transgenic plants were grown in the greenhouse room at 25°C under a 12 h light/12 h dark photoperiod. *Nicotiana benthamiana* seedlings were grown in a climate-controlled chamber at 22°C (60% relative humidity) under 16 h light/8 h dark photoperiod.

Stable transformation of pepper

For functional verification, *35S::CaDOF1-GFP* overexpression plants were generated in "ST-8" background, *35S::CaDOF2*, *35S::CaTF308*, *35S::CaTF216*, *35S::CaTF57*, *35S::CaTF03*, *35S::CaTF83*, *35S::CaTF234*, *35S::CaTF86*, *35S::CaTF20*, *35S::CaTF56* and *35S::CaTF47* overexpression plants were generated in the "P2205" background, respectively. Transgenic pepper plants were generated using the A. tumefaciens-mediated transformation method as described previously (Tang et al., 2025). Briefly, cotyledon segment explants were prepared from 10-day-old seedlings for Agrobacterium-mediated transformation. Thereafter, the explants were directly inoculated with Agrobacterium suspension (OD600=0.6-0.8) under -0.1 Mpa air pressure, followed by a 2-day co-cultivation in the dark. After co-cultivation, the explants were transferred sequentially to a callus-inducing medium and shoot-inducing medium for subsequent regeneration steps, with both media supplemented with 50 mg/L spectinomycin (Spe) for transgenic selection. Once adventitious shoots were elongated to a suitable size, they were excised individually and cultured on a root-inducing medium for approximately 3-4 weeks until a robust root system had formed.

Luciferase (LUC) reporter assay

The LUC reporter assay was performed as previously reported (Zhang et al., 2026).

Briefly, the 1.5~2 kb promoter sequences of *GGPPS1*, *PSY1*, *PSY2*, *PDS*, *CRTISO*, *Z-ISO*, *ZDS*, *LCYB1*, *CHYB1*, *LCYE*, *VDE*, *ZEP* and *CCS* were cloned into *ACT2::LUC-UBQ10:DsRed* vectors as reporter constructs, and the coding sequences of transcription factors were cloned into the *35S::Flag-GFP* vector, respectively. The recombinant plasmids were then co-transformed into *N. benthamiana* leaves via agroinfiltration. At three days post-infiltration, the leaves were imaged by Lumazone PyLoN 2048B (Lumazone) and Amersham ImageQuant 800 Fluor (Cytiva), respectively. The ratio of LUC/DsRed was measured using ImageJ.

Electrophoretic mobility shift assay (EMSA)

The recombinant protein MBP, MBP-CaDOF1dof, and MBP-CaDOF2dof were expressed in Escherichia coli BL21-CodonPlus (DE3)-RIP (Cat. EC1007, WEIDI) and purified with amylose resin (Cat. E8201, NEB). Oligonucleotide probes derived from the promoter regions were 5'-end-labelled with 5-carboxyfluorescein (FAM). The short double-strand DNA probes were synthesized by annealing single-stranded 5'-FAM-labelled oligo. The reaction products were separated using 0.5 × Tris-borate-EDTA (TBE) buffer in 6% native polyacrylamide gel and imaged on the Amersham ImageQuant 800 Fluor (Cytiva).

Microscale thermophoresis (MST)

The MST assay was performed as previously reported (Ning et al., 2025). Briefly, the recombinant protein MBP-CaDOF1dof or MBP-CaDOF2dof were labeled with RED-NHS dye using the Monolith Protein Labeling Kit RED-NHS 2nd Generation (NanoTemper Technologies) following the protocol instructions. The short double-strand DNA probes were synthesized by annealing single-stranded oligo. The MST-based in vitro binding assay was conducted on a Monolith NT.115 Pico instrument (NanoTemper Technologies GmbH, Munich, Germany) with blue/red filters. All binding experiments were conducted with three technical repeats, and analyzed using the program NanoTmper analysis software (NanoTemper Technologies).

Quantification of GGPP and carotenoid contents in pepper fruit pericarps

The pericarp tissues were collected without seeds and placentas, frozen in liquid nitrogen and stored at -80℃. For each biological replicate, pericarp was pooled from at least three fruits. Frozen pericarp tissues were ground into fine powder and freeze-dried for 24-48 h with a freeze-dryer (Cat. CTFD-10S-U, CREATRUST). GGPP and carotenoid contents in pericarp were detected by PKU-IAAS Mass Spectrometry Platform and Biomarker Technologies, respectively. The extraction of pepper GGPP and carotenoids was performed as described previously (Zita et al., 2023; Zhang et al., 2023).

ChIP-seq and data analysis

The ChIP assay performed as previously reported (Yang et al., 2025). Specifically, the pepper chromation was sheared with Bioruptor Plus (Diagenode) for 32 cycles with 30 s ON/30 s OFF and incubated with anti-H3K4me1 (Abcam), anti-H3K4me3

(Abcam), anti-H3K9me2 (Abcam), anti-H3K27me3 (Abcam), anti-H3K27ac (Abcam), Anti-RNA polymerase II CTD repeat YSPTSPS (phospho S5) (Abcam), Anti-RNA Polymerase II CTD mAb (Cosmo Bio) and anti-GFP (Roche), respectively. Two biological replicates were performed for each ChIP, and libraries were sequenced at BerryGenomics using an Illumina NovaSeq X Plus system with 150 bp pair-end reads (PE150). Raw FASTQ files were processed with Trim Galore (v0.6.10) for adaptor removal and quality trimming. Clean reads were aligned to the reference genome using Bowtie2 (v2.5.4, default parameters) to generate SAM files, which were converted to BAM format, sorted, and indexed using SAMtools. Genome-wide ChIP-seq signal tracks were generated using deepTools (v3.5.5) bamCoverage with --normalizeUsing RPKM, producing normalized bigWig (.bw) files for visualization and downstream aggregation analyses (e.g., metaplots/heatmaps). Peak calling was performed using MACS2 (v2.2.9.1) to identify enriched regions, followed by peak annotation and gene association analyses based on genome annotation. Motif enrichment analysis was carried out using HOMER (v5.1) findMotifs.pl. All additional data processing, statistical summaries, signal extraction, visualization, and integrative analyses were implemented in Python (v3.9).

Spatial transcriptomic sequencing and data analysis

Fresh pericarps from WT and *CaDOF1-GFP* (#208) plant were embedded in 100% pre-chilled OCT compound (Sakura, 4583) on ice. Samples were subjected to vacuum infiltration for 5 min at 4℃ followed by centrifugation at 3,000 × g for 5 min to remove air bubbles. The treated samples were cast into 17×17×5 mm molds (ShiTai) and stored at -80°C after OCT solidification. Sections (10-μm thick) were cut using a cryostat microtome (Dakewe, CT520) and used for subsequent spatial transcriptomic sequencing.

The libraries were performed using BMKMANU S3000 gene-expression kits (BMKMANU, ST03002) following the manufacturer's instructions. All libraries were sequenced on the Illumina NovaSeq plotform. Raw sequencing data were aligned to the pepper reference genome (Zhangshugang) using BSTMatrix (v.2.4.f) to generate gene expression count matrices at multiple resolutions. The resulting matrices were processed using the Seurat R package (v.5.1.0) (Hao et al., 2024). To filter out low-quality cells, only cells with >50 expressed genes and genes expressed in >10 cells were retained. Data normalization, scaling and transformation were performed using Seurat (v.5.1.0). Batch effects were corrected using Harmony (v.1.2.0) with default parameters.

Principal component analysis (PCA) was applied for dimensionality reduction, and the top 30 principal components were retained. Clustering was carried out using the Louvain algorithm via the FindNeighbors and FindClusters functions, and visualization was performed using uniform manifold approximation and projection (UMAP) via the RunUMAP function. DEGs between clusters were identified using FindAllMarkers or FindMarkers functions, with parameters set as: only.pos = TRUE, logfc.threshold = 0.5 and min.pct = 0.1.

Data availability

Data supporting the findings of this work are available both in the main text and the additional information files. The raw high throughput sequencing data generated in this study have been deposited in the Genome Sequence Archive in National Genomics Data Center (China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences) with the dataset identifier CRA044885 (https://ngdc.cncb.ac.cn/gsa/browse/CRA044885). The spatial transcriptomic raw datasets have been submitted to the National Center for Biotechnology Information (NCBI) under the BioProjects accession number PRJNA1496594.

Acknowledgements

We thank all the members from XW's lab for productive discussion. We thank Dr. Zhangsheng Zhu (South China Agricultural University, China) for providing seeds of pepper inbred line 'CA59'.

Author contributions

F. Z., L.H., S. Y. and X. W. conceived and designed the experiments. F. Z., X. S., J. W., Q. Z., Z. Y., C. P. and M. L. performed experiments. Z. Z. and X. W. W. performed high throughput sequencing experiments and analyzed the data. K. G. and B. L. analyzed the spatial transcriptomic data. Y. F. and G. Q. performed IP-MS experiments and analyzed the data. P. P. and Y. S. prepared the plant materials. F. Z. and X. W. analyzed the non-omics data. X. W. wrote the manuscript. All authors read, edited, and approved the final manuscript.

Funding

This work was supported by Key R&D Program of Shandong Province, China (2024CXPT034) and Shandong Provincial Natural Science Fund Program (ZR2025MS401 to Fan Zhang).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Figure 1-8.
Supplementary Table 1-4.

References

1. Kim J-S, An CG, Park J-S, Lim YP, Kim S. Carotenoid profiling from 27 types of paprika (Capsicum annuum L.) with different colors, shapes, and cultivation methods. *Food Chemistry* **201**, 64-71 (2016).

2. Maeda H, Nishino A, Maoka T. Biological Activities of Paprika Carotenoids, Capsanthin and Capsorubin. *Adv Exp Med Biol* **1261**, 285-293 (2021).

3. Kennedy LE, Abraham A, Kulkarni G, Shettigar N, Dave T, Kulkarni M. Capsanthin, a Plant-Derived Xanthophyll: a Review of Pharmacology and Delivery Strategies. *AAPS PharmSciTech* **22**, 203 (2021).

4. Thorup TA, Tanyolac B, Livingstone KD, Popovsky S, Paran I, Jahn M. Candidate gene analysis of organ pigmentation loci in the Solanaceae. *Proc Natl Acad Sci U S A* **97**, 11192-11197 (2000).

5. Huh JH*, et al.* A candidate gene approach identified phytoene synthase as the locus for mature fruit color in red pepper (Capsicum spp.). *Theoretical and Applied Genetics* **102**, 524-530 (2001).

6. Berry HM, Zhou N, Rickett DV, Baxter CJ, Enfissi EMA, Fraser PD. The characterisation of phytoene synthase-1 and 2, and 1-D-deoxy-xylulose 5-phosphate synthase genes from red chilli pepper (Capsicum annuum). *The Journal of Horticultural Science and Biotechnology* **98**, 732-744 (2023).

7. Borovsky Y, Tadmor Y, Bar E, Meir A, Lewinsohn E, Paran I. Induced mutation in beta-CAROTENE HYDROXYLASE results in accumulation of beta-carotene and conversion of red to orange color in pepper fruit. *Theor Appl Genet* **126**, 557-565 (2013).

8. Guzman I, Hamby S, Romero J, Bosland PW, O'Connell MA. Variability of Carotenoid Biosynthesis in Orange Colored Capsicum spp. *Plant Sci* **179**, 49-59 (2010).

9. Lee SY*, et al.* A mutation in Zeaxanthin epoxidase contributes to orange coloration and alters carotenoid contents in pepper fruit (Capsicum annuum). *Plant J* **106**, 1692-1707 (2021).

10. Ha SH, Kim JB, Park JS, Lee SW, Cho KJ. A comparison of the carotenoid accumulation in Capsicum varieties that show different ripening colours: deletion of the capsanthin-capsorubin synthase gene is not a prerequisite for the formation of a yellow pepper. *J Exp Bot* **58**, 3135-3144 (2007).

11. Ma J, Dai J-x, Liu X-w, Lin D. The transcription factor CaBBX20 regulates capsanthin accumulation in pepper (Capsicum annuum L.). *Scientia Horticulturae* **314**, 111907 (2023).

12. Song J*, et al.* An R-R-type MYB transcription factor promotes non-climacteric pepper fruit carotenoid pigment biosynthesis. *Plant J* **115**, 724-741 (2023).

13. Wang Y*, et al.* The MADS-RIPENING INHIBITOR-DIVARICATA1 module regulates carotenoid biosynthesis in nonclimacteric Capsicum fruits. *Plant*

Physiol **197**, (2025).

14. Song Y*, et al.* CaNAC2 orchestrates Capsicum annuum fruit ripening by coordinated carotenoid and ABA metabolism under epigenetic control. *Plant Physiology* **199**, (2025).

15. Zhang F*, et al.* An ALOG family transcription factor promotes capsanthin production by activating the expression of Capsanthin-Capsorubin Synthase gene in pepper. *Plant Commun*, 101964 (2026).

16. Liao Y*, et al.* The 3D architecture of the pepper genome and its relationship to function and evolution. *Nat Commun* **13**, 3479 (2022).

17. Yang H*, et al.* Epigenome profiling reveals distinctive regulatory features and cis-regulatory elements in pepper. *Genome Biol* **26**, 121 (2025).

18. Wu LY*, et al.* Dynamic chromatin state profiling reveals regulatory roles of auxin and cytokinin in shoot regeneration. *Dev Cell* **57**, 526-542 e527 (2022).

19. Ernst J, Kellis M. Chromatin-state discovery and genome annotation with ChromHMM. *Nat Protoc* **12**, 2478-2492 (2017).

20. Fu LY*, et al.* ChIP-Hub provides an integrative platform for exploring plant regulome. *Nat Commun* **13**, 3413 (2022).

21. Hu Y*, et al.* Multiscale footprints reveal the organization of cis-regulatory elements. *Nature* **638**, 779-786 (2025).

22. Han J*, et al.* Laminar patterning transcription factors orchestrate spatial metabolite partitioning in Capsicum fruit. *Nat Plants* **12**, 1095-1115 (2026).

23. Zhang F*, et al.* dTAPL, A Method That Expedites the Discovery of Proteins Associated With Specific Genomic Loci in Plants. *Plant Biotechnol J* **23**, 5677-5679 (2025).

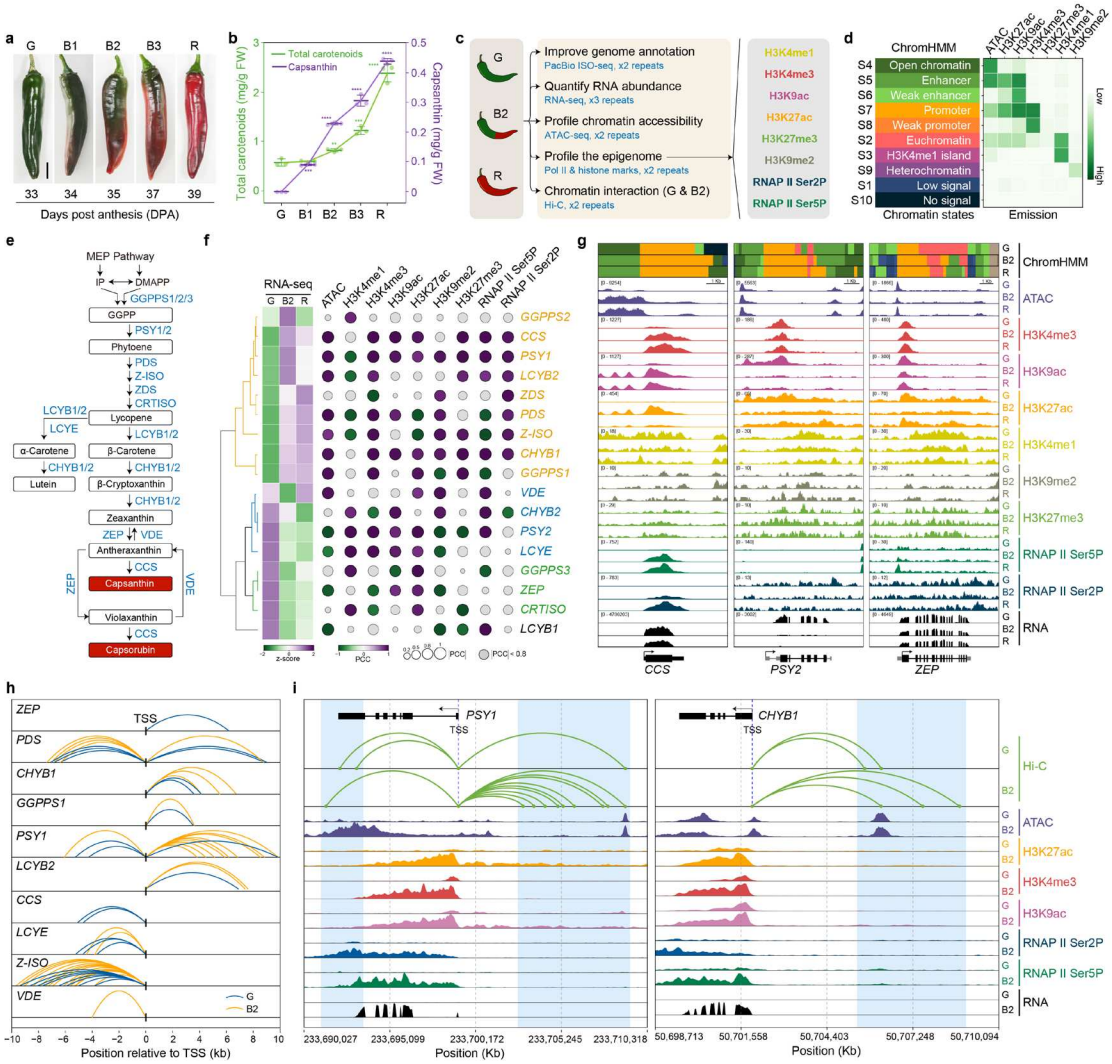

Fig. 1 Chromatin and transcriptional dynamics of CBGs throughout pepper fruit color transition. a Pepper fruit phenotype showing the developmental route from green mature (G) to Breaker (B1-B3) and Red ripe stages (R). Scale bar indicates 2 cm. **b** Total carotenoid and capsanthin contents across developmental stages. Asterisks indicate significant difference between G and other samples (****, $p < 0.0001$; ***, $p < 0.001$; **, $p < 0.005$) based on two-tailed *Student's t test*. **c** Sampling scheme for RNA-seq, ATAC-seq, ChIP-seq and Hi-C profiling at G, B2 and R stages. **d** Ten-state ChromHMM model inferred from chromatin accessibility and six histone modifications. **e** Schematic representation of the carotenoids/capsanthin biosynthetic pathway in pepper. **f** Stage-resolved expression clustering of 17 CBGs and correlations between RNA abundance and chromatin accessibility, histone marks or RNA polymerase II signals. **g** Gene-level chromatin-state annotations for representative pathway genes. **h** Gene-centered Hi-C contact maps and predicted regulatory-element links around pathway loci. **i** Integrated visualization of Hi-C contacts, chromatin accessibility, active histone marks, RNA expression and gene structure around *PSY1* and *CHYB1* genes.

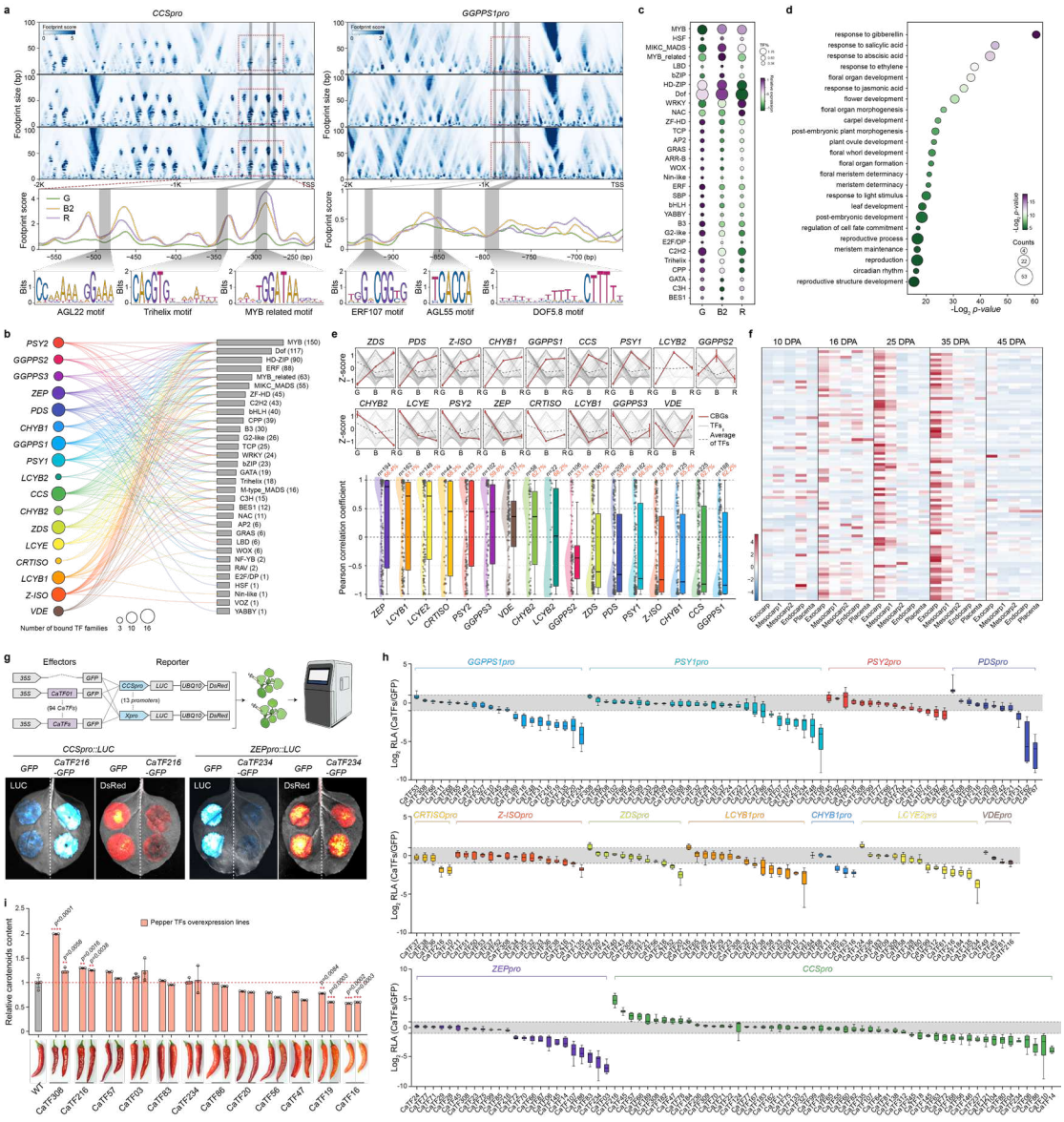

Fig. 2 TF footprints analysis nominates candidate TF regulators of CBGs. a PRINT-derived footprint scores across 2-kb promoter regions of CBGs at G, B2 and R stages. **b** Bipartite network linking CBGs to TF families derived from footprints analysis. **c** Joint analysis of transcription factor-family motif proportions in ATAC peaks and stage-resolved family expression. **d** GO enrichment of 324 candidate TFs. **e** Expression-correlation analysis between candidate TFs and linked CBGs. **f** Stage- and tissue-resolved expression of 94 cloned TFs. **g** Schematic diagram and representative results of LUC reporter assays. **h** LUC reporter assays showing the regulatory activities of candidate TFs on target promoters. Relative LUC activity (RLA) for assay was calculated by normalizing the LUC intensity with DsRed intensity of the same region on the image. The dash lines indicate the FC (RLA^{TF}/RLA^{GFP}) =2 or 0.5. **i** Relative capsanthin contents and red ripe fruit phenotypes of wild-type (WT) and transgenic pepper plants overexpressing indicated TFs. Two independent lines for each TF were shown. *p-value* calculated by two-tailed *Student's t test* is used to determine the

557 statistical significance between genotypes.
558
559

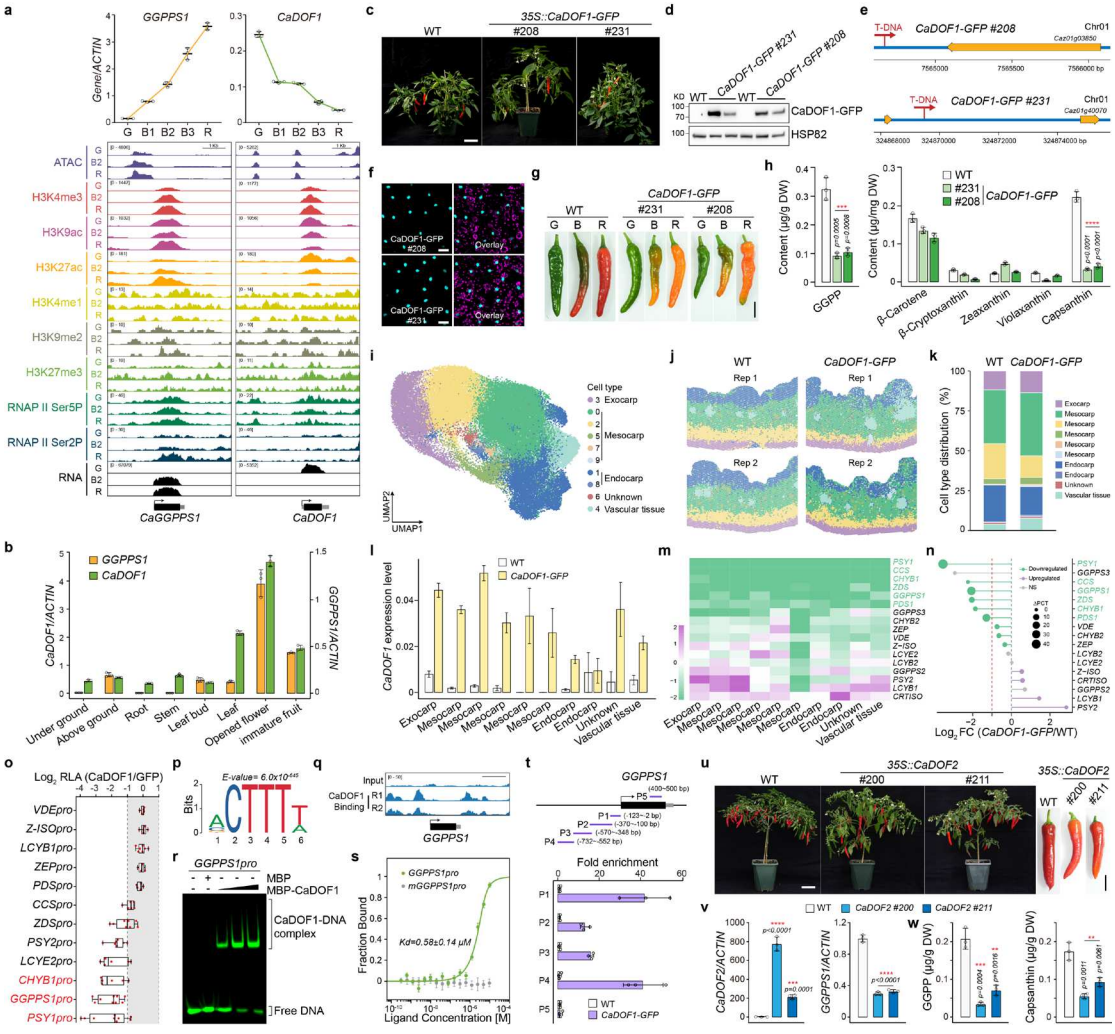

Fig. 3 *CaDOF1* and *CaDOF2* repress capsanthin biosynthesis by targeting *GGPPS1*. **a** Developmental stage-resolved expression and genome-browser views of *GGPPS1* and *CaDOF1* during fruit color transition. **b** Tissue expression patterns of *GGPPS1* and *CaDOF1*. **c-e** Morphology (**c**), CaDOF1-GFP protein level analysis (**d**) and T-DNA insertion validation (**e**) of WT and *CaDOF1-GFP* overexpression pepper lines. Scale bar in (**c**) indicates 10 cm. **f** Nuclear localization of CaDOF1-GFP in transgenic pepper leaves. Scale bar indicates 50 μm. **g** Fruit phenotypes of WT and *CaDOF1-GFP* lines at green mature, breaker and red ripe stages. Scale bar indicates 2 cm. **h** Contents of GGPP and other carotenoid intermediates in red ripe pericarps. *p-value* calculated by two-tailed *Student's t test* is used to determine the statistical significance between genotypes. **i-k** UMAP, spatial maps and cell-type composition of WT and *CaDOF1-GFP* pericarps from spatial transcriptomic analysis. **l-n** Cell-type expression of *CaDOF1* and CBGs, and differential expression analysis of CBGs in WT and *CaDOF1-GFP* overexpression pepper. **o** LUC reporter assays showing the effects of CaDOF1 on promoter activities of CBGs in LUC reporter assays. Relative LUC activity (RLA) for each reporter was calculated by normalizing the LUC intensity with DsRed intensity. The dash lines indicate the FC (RLA^{CaDOF1}/RLA^{GFP}) =2 or 0.5. **p** Enriched CaDOF1-binding motif from ampDAP-seq. **q** CaDOF1 occupancy around

580 *GGPPS1* locus in ampDAP-seq. **r-s** *in vitro* CaDOF1 binding to *GGPPS1* promoter
581 motifs measured by fEMSA (**r**) and MST (**s**). **t** ChIP-qPCR validation of CaDOF1-GFP
582 enrichment at *GGPPS1* promoter regions *in vivo*. **u-w** Morphology (right panel in U),
583 fruit phenotypes (left panel in U), *CaDOF2* and *GGPPS1* expression (V), GGPP and
584 capsanthin contents (W) in WT and *CaDOF2*-overexpression peppers. *p-value*
585 calculated by two-tailed *Student's t test* is used to determine the statistical significance
586 between genotypes. Scale bars for pepper plants and fruits shown in U indicate 10 and
587 2 cm, respectively.
588

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTable4.xlsx
- SupplementaryFigures.pdf
- SupplementaryTable1.xlsx
- SupplementaryTable2.xlsx
- SupplementaryTable3.xlsx