Vegetable Research

ARTICLE

Open Access

https://doi.org/10.48130/vegres-0025-0022

Vegetable Research **2025**, 5: e029

Integrated transcriptome and metabolome analysis revealed key genes related to astringency in cucumber fruits

Jiaojiao Zhang[1#], Kunyang Wang[2#], Xiaofei Song[2], Yang Xie[1], Xiaoli Li[1], Shuai Meng[1], Qiushuang Han[1], Jianhua Jia[1], Chen Wang[1] and Liying Yan[1*]

[1] *Hebei Key Laboratory of Horticultural Germplasm Excavation and Innovative Utilization, College of Horticulture Science and Technology, Hebei Normal University of Science and Technology, Qinhuangdao 066004, Hebei, China*

[2] *Analysis and Testing Center, Hebei Normal University of Science & Technology, Qinhuangdao 066000, Hebei, China*

[#] Authors contributed equally: Jiaojiao Zhang, Kunyang Wang

[*] Corresponding author, E-mail: yanliying0665@hevttc.edu.cn

Abstract

Cucumber taste profiles are significantly shaped by astringency, which manifests as a drying and rough sensation in the oral cavity. Nevertheless, the molecular mechanisms underlying cucumber astringency remain largely unexplored. The present study addresses this gap by using an integrated approach encompassing transcriptome and metabolome sequencing. FC (P1) and HC (P2) inbred lines of cucumber were used as experimental materials. P1 fruits at 9 d post-pollination (dpp; P1-9d) and P2 fruits at 3 dpp (P2-3d) exhibited slight astringency, whereas P2 fruits at 9 dpp (P2-9d) showed strong astringency. In the P2-9d vs P2-3d, and P2-9d vs P1-9d comparisons, 5,327 and 1,113 differentially expressed genes (DEGs) were detected, respectively. Among these, 48 differentially expressed transcription factors (TFs) were identified. Phylogenetic analysis, with a focus on the MYB family of TFs, highlighted four candidate TFs potentially involved in flavonoid biosynthesis. Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis highlighted enrichment of phenylpropanoid biosynthesis, phenylalanine metabolism, and flavonoid biosynthesis. Integrated transcriptome and metabolome analysis identified 24 common pathways between P2-9d vs P2-3d, and P2-9d vs P1-9d comparisons, including three flavonoid-related pathways comprising 21 DEGs. Quantitative RT-PCR experiments validated ten candidate genes, shedding light on their putative functions in modulating cucumber astringency. This research advances the understanding of astringency development in cucumbers.

Citation: Zhang J, Wang K, Song X, Xie Y, Li X, et al. 2025. Integrated transcriptome and metabolome analysis revealed key genes related to astringency in cucumber fruits. *Vegetable Research* 5: e029 https://doi.org/10.48130/vegres-0025-0022

Introduction

Cucumber (*Cucumis sativus* L.) is a commonly grown vegetable, usually picked while still unripe and eaten as a juicy, fleshy fruit[1]. Astringency, a key factor affecting cucumber palatability, is a tactile sensation rather than a primary taste, characterized by dryness, roughness, and a tightening effect in the mouth[2]. It results from the formation of insoluble precipitates due to interactions between polyphenols and salivary proteins, which increase oral friction[3]. This property can substantially affect consumer preferences and acceptance. Thus, understanding and managing astringency in cucumber fruits is essential for maintaining and enhancing their marketability and consumer satisfaction.

Currently, the identified astringent compounds include polyphenols, organic acids, inorganic acids, metal ionic salts, and dehydrating agents[4]. The astringency of most horticultural fruits is primarily attributed to polyphenols[5], which are widely distributed secondary metabolites in plants, predominantly as flavonoids and phenolic acids[6]. Flavonoids, an extensive group of polyphenolic metabolites, encompass six major subclasses defined by their distinct molecular architectures: chalcones, flavanediols, flavones, anthocyanidins, flavonols, and proanthocyanidins[7]. The precursor of flavonoids is phenylalanine, which undergoes sequential catalysis by phenylalanine ammonia-lyase (PAL), cinnamic acid 4-hydroxylase (C4H), and 4-coumaroyl CoA ligase (4CL), producing cinnamic acid, coumaric acid, and 4-coumaroyl CoA in the phenylpropanoid metabolic pathway[8,9]. PAL functions as the key regulatory enzyme in this process[10]. Chalcone synthase (CHS) catalyzes the condensation of 4-coumaroyl-CoA and malonyl-CoA to form chalcone, which is then converted into dihydroflavonols through enzymatic actions of chalcone isomerase (CHI) and flavanone 3-hydroxylase (F3H). Dihydroflavonols, precursors for anthocyanins, tannins, and other flavonoid compounds, are enzymatically reduced to leucocyanidin by dihydroflavonol 4-reductase (DFR). In plants, leucocyanidin is converted to proanthocyanidins through two pathways. One pathway involves its conversion into cyanidin by anthocyanin synthase, followed by the catalysis mediated by anthocyanidin reductase (ANR) to form epicatechin. Cyanidin also contributes to proanthocyanidin formation through UDP-flavonoid 3-O-glucosyltransferase (UFGT). The second pathway entails the direct transformation of leucocyanidin into catechin, catalyzed by leucoanthocyanidin reductase (LAR). Finally, epicatechin and catechin from the two pathways polymerize to form proanthocyanidins[11–13].

Transcriptional regulation by transcription factors (TFs) is a key determinant of flavonoid biosynthesis. MYB12, an R2R3-MYB TF in *Arabidopsis*, serves as a specific activator for flavonol biosynthesis[14]. In *Arabidopsis*, the TFs MYB11, MYB12, and MYB111 regulate the expression of key genes in the flavonoid biosynthetic pathway, such as *CHS*, *CHI*, *F3H*, and *FLS1*, during seedling development. Specifically, AtMYB12 controls flavonoid production in the roots, while AtMYB111 regulates it in the cotyledons[15]. Another R3-MYB TF MYBL2 in *Arabidopsis* has been implicated in controlling the production of flavonoids. Deficiency in MYBL2 activity significantly increases anthocyanin accumulation, whereas its overexpression in seeds inhibits the biosynthesis of proanthocyanidins[16,17]. In grapes, VvMYBPA1 upregulates the expression of *LAR* and *ANR*, facilitating

proanthocyanidin accumulation[18]. Moreover, *Arabidopsis TRANSPARENT TESTA8* (*AtTT8*), which encodes a bHLH-type TF, cooperatively regulates the expression level of the flavonoid biosynthesis genes *DFR* and *BANYULS* (*BAN*) alongside *TRANSPARENT TESTA GLABRA1* (*TTG1*) and *TT2*[19].

Combined multi-omics approaches are increasingly employed for genome-wide investigations[20,21]. However, the application of multi-omics technologies for the analysis of cucumber astringency has not been reported yet. In this study, RNA-seq and Metabolome-seq were utilized to analyze the differentially expressed genes (DEGs) and differentially accumulated metabolites (DAMs) in two cucumber inbred lines varying greatly in their astringency levels. Through a comprehensive analysis encompassing DEG analysis and combined transcriptome and metabolome sequencing, this study identified the key regulatory factors and functional genes associated with cucumber astringency. The findings serve as a reference and offer valuable insights for identifying genes linked to cucumber astringency.

Materials and methods

Plant growth and sampling

Cucumber plants utilized in this research were cultivated in the greenhouse of the Hebei Normal University of Science and Technology. Two cucumber inbred lines, namely slightly astringent FC (P1) and the strongly astringent HC (P2), served as the experimental materials. FC fruits at 9 dpp (P1-9d) and HC fruits at 3 dpp (P2-3d) with slight astringency, as well as HC fruits at 9 dpp (P2-9d) with strong astringency, were promptly frozen in liquid nitrogen and kept at −80 °C for later transcriptomic and metabolomic analyses.

Astringency level evaluation

The cucumbers were cut into fan-shaped strips and left to stand for 1 min. Ten evaluators tasted the area near the peel on the cross-section and the juice overflowing from the pulp through the tip of their tongues, providing a score and calculating the astringency index (AI). AI was calculated using the following formula:

$AI = \frac{\sum(si \times ni)}{6N} \times 100$, (si: astringency level; ni: the number of evaluators who give the corresponding astringency level; N: the total number of evaluators).

To ensure accuracy, the evaluators rinsed their mouths with soda water after each evaluation and took a 5-min break before assessing the next sample[22]. The grading criteria for the cucumber astringency index (AI) are defined as follows: AI = 0 represents no astringency, AI < 40 represents slight astringency, 40 ≤ AI < 70 represents moderate astringency, and AI ≥ 70 represents strong astringency.

RNA isolation, library construction and sequencing

Total RNA was isolated from cucumber fruits using the RNAprep Pure Plant Kit (Tiangen, China). A NanoDrop ND2000 spectrophotometer was used to determine the purity of RNA samples. The RNA was sent to Oebiotech (Shanghai, China) for Illumina NovaSeq 6000-based paired-end sequencing. Sequencing data were generated through the platform. To minimize the data error rates, raw data were preprocessed using fastp software (v0.20.1, parameter: --length_required 50) to remove low-quality reads, which yielded clean reads for subsequent statistical analyses[23]. The clean reads were aligned to the cucumber V3 reference genome by using hisat2 software (v2.1.0, parameter: --rna-strandness rf --fr) to obtain information on the positional and sequence characteristics unique to the samples[24]. For gene expression quantification, HTseq-count software (v0.11.2, parameter: -s reverse) was applied to determine read alignments to protein-coding genes in all samples[25].

Transcriptomics data analysis and plotting

Principal component analysis (PCA) with R (v3.2.0) was applied to evaluate the similarity of biological replicates and sample groups based on gene counts. DEGs were identified using the DESeq2 package[26], applying thresholds of $q < 0.05$ and $|\log_2(\text{fold change})| > 1$. Genes meeting these criteria were classified as DEGs. Hypergeometric distribution-based enrichment analysis was performed on DEGs to identify significantly enriched Gene Ontology (GO) terms and KEGG pathways. Visualization of the enriched terms was performed via R (v3.2.0). Heatmaps, bubble charts, Venn diagrams, and bar charts for this analysis were generated using the ggplot2, pheatmap, and readxl packages in R.

Metabolite isolation and liquid chromatography–tandem mass spectrometry

Metabolite extraction was performed on cucumber fruits, with six biological replicates per sample type processed simultaneously. Metabolite extraction and determination were performed as described by Wu et al.[27].

Metabolomics data analysis and plotting

Unsupervised PCA was applied to evaluate the global distribution of samples and assess analytical stability. To ensure accurate differentiation between material groups, orthogonal partial least squares-discriminant analysis (OPLS-DA) was applied, and variable importance in projection (VIP) values was generated for each metabolite. Differential metabolites were identified by analyzing the metabolite data matrix, with the following parameters for each comparison group: p value < 0.05 and VIP > 1. Heatmaps, volcano plots, and bar charts of differential metabolites were created using the ggplot2 (3.5.1), pheatmap (1.0.12), and readxl (1.4.3) packages in R, respectively. KEGG pathway enrichment analysis was conducted using the KEGG IDs of differential metabolites to identify significantly enriched metabolic pathways in each group.

Quantitative real-time PCR

The inbred lines FC, LNTZT, 17S-50, 17S-33, and 17S-23 exhibit less astringency, while HC, 21A127, 21A114, 17S-135, and 17S-139 show higher astringency. Total RNA was isolated from fruits of the ten inbred lines using the aforementioned method, followed by reverse transcription with the FastQuant RT reagent kit (Tiangen, China). Three biological replicates were performed for each sample. qRT-PCR analysis was performed according to the protocol described by Zhang et al.[28]. Primer sequences are provided in Supplementary Table S1.

Statistical analysis

Data were analyzed using one-way ANOVA with Tukey's multiple comparisons test (GraphPad Prism 10.2.2), with statistical significance set at $p \leq 0.05$.

Results

Evaluation of the astringent levels of cucumber fruits

In a previous report, the astringency levels of fruits (9 dpp) were examined from 228 cucumber germplasm resources[22]. Five genetic materials each were selected from those with low astringency and high astringency for subsequent experiments (Supplementary Table S2; Fig. 1a). The cucumber variety with the highest astringency taste out of all the cucumbers was HC, FC with the lowest astringency. The AIs for (FC) P1 fruits at 9 dpp (P1-9d) and (HC) P2 fruits at 3 dpp (P2-3d) were less than 40 and exhibited slight astringency. Meanwhile, the AI for P2 fruits at 9 dpp (P2-9d) exceeded 70 and had strong astringency (Fig. 1b; Supplementary Table S3). Based on

these findings, fruits of P2-3d, P1-9d, and P2-9d were selected for subsequent untargeted metabolomics and transcriptomics analyses.

Transcriptomic profiling during changes in astringency

To investigate the molecular basis of cucumber astringency, RNA-seq-based transcriptome profiling was performed on P1-9d, P2-3d, and P2-9d. Following quality control filtering, high-quality reads were generated, with lengths ranging from 46,315,150 to 47,285,688 bp across the nine libraries. The percentage of Q30 bases exceeded 95%, and the GC was greater than 44%. Of these reads, 98.27%–98.83% were successfully aligned to the cucumber reference genome (Supplementary Table S4). PCA demonstrated high repeatability and stability across the sample groups (Supplementary Fig. S1). Based on the clustered heatmap, the DEGs were grouped into two clusters, indicating distinct gene expression profiles between fruits with different astringency levels (Fig. 2a, b).

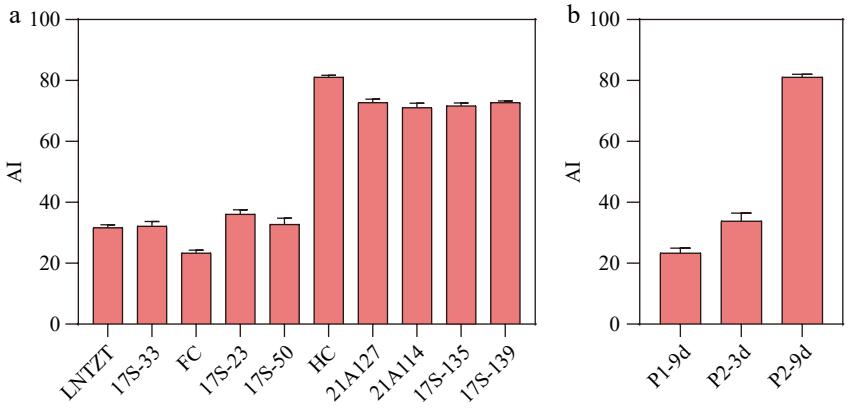

Fig. 1 (a) Evaluation of astringency levels for different cucumber fruits. (b) Evaluation of astringency levels for P1-9d, P2-3d, and P2-9d. AI: Astringency Index. Data are presented as the mean ± SD of three biological replicates.

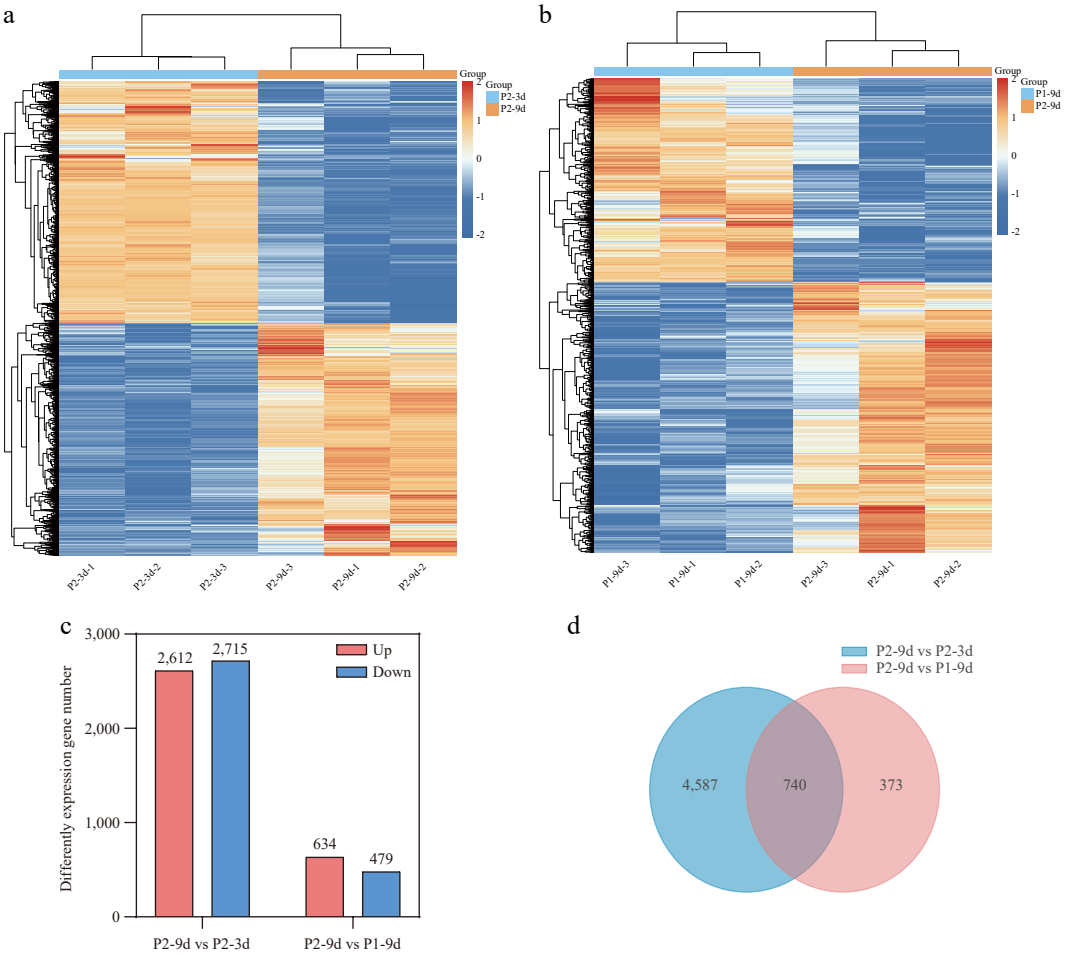

Fig. 2 Transcriptome analysis of cucumber fruits with different astringency levels. (a), (b) Clustered map of DEGs. (c) The number of upregulated and downregulated DEGs in P2-9d vs P2-3d, and P2-9d vs P1-9d comparisons. (d) Venn diagram showing common DEGs in P2-9d vs P2-3d, and P2-9d vs P1-9d comparisons.

In the P2-9d vs P2-3d comparison, a total of 5,327 DEGs (2,612 upregulated, 2,715 downregulated) were identified, whereas the P2-9d vs P1-9d comparison revealed 1,113 DEGs (634 upregulated, 479 downregulated) (Fig. 2c). Volcano plots for each comparison group are presented in Supplementary Fig. S2. Furthermore, 740 DEGs were common to both comparisons (Fig. 2d), effectively excluding DEGs associated with different genotypes and developmental stages.

GO and KEGG enrichment analyses of DEGs

To investigate the biological significance of DEGs, GO enrichment analysis was implemented, which indicated the enrichment of DEGs in both comparisons primarily in molecular functions such as 'sequence-specific DNA binding' and 'DNA-binding TF activity' (Fig. 3a). These results indicate that TFs are critical regulators of cucumber astringency.

To further investigate the functions of DEGs, KEGG pathway enrichment analysis was conducted. DEGs of the two comparisons exhibited marked enrichment in key pathways such as starch and sucrose metabolism (csv00500), plant hormone signal transduction (csv04075), phenylpropanoid biosynthesis (csv00940), phenylalanine metabolism (csv00360), MAPK signaling pathway-plant (csv04016), glyoxylate and dicarboxylate metabolism (csv00630), galactose metabolism (csv00052), flavonoid biosynthesis (csv00941), diterpenoid biosynthesis (csv00904), and alpha-linolenic acid metabolism (csv00592) (Fig. 3b).

Expression of TF genes

TFs, including those of the bHLH and MYB families, are key regulators of astringent substance formation in plants. In this study, 48 TF genes were identified, spanning 24 families (Fig. 4a). Among these, seven and nine TFs belonged to the bHLH and MYB families, respectively. Thirteen TF genes displayed elevated expression levels in P2-9d compared to P2-3d and P1-9d, indicating that they may be positively correlated with astringency. By contrast, two TF genes exhibited lower expression in the P2-9d samples than in the P1-9d and P2-3d samples, suggesting their potential negative regulatory role in astringency (Fig. 4b, c). These expression trends imply that these genes are likely critical regulators of astringent compound biosynthesis in cucumbers.

To better understand the functions of MYB TFs, a phylogenetic analysis of CsMYBs was conducted in comparison to the known anthocyanidin, phenylpropanoid repressor, flavonoid repressor/activator, and proanthocyanidin regulator genes[24]. The results revealed that CsMYB32 (CsaV3_2G016030) and CsMYB6 (CsaV3_5G040410) clustered with the phenylpropanoid biosynthesis repressors AtMYB4 and DkMYB14[29,30]. In addition, CsMYB111 (CsaV3_5G001810) grouped with flavonol regulators, whereas CsMYB4 (CsaV3_3G043510) clustered with proanthocyanidin activators (Fig. 4d). CsMYB4 displayed homology to CsMYB1, which has been linked to tea astringency through its regulation of catechin biosynthesis. This phylogenetic analysis offers valuable insights into the putative functions of these MYB factors in polyphenol biosynthesis[31].

Metabolomics analysis

To further elucidate molecular mechanisms underlying cucumber astringency, untargeted metabolomics analysis was performed on six replicates each of P1-9d, P2-9d, and P2-3d, totaling 18 samples. PCA revealed distinct metabolite profiles among the sample groups, whereas the diversity within each group was relatively low (Fig. 5a). OPLS-DA confirmed the high reproducibility of the results, allowing for a more detailed analysis of DAMs (Supplementary Fig. S3).

A total of 448 DAMs (148 upregulated, 300 downregulated) were detected in the P2-9d vs P1-9d comparison, and 405 DAMs (201 upregulated, 204 downregulated) were obtained in the P2-9d vs P2-3d comparison (Fig. 5b, c). DAMs common to both comparisons included epicatechin 3-glucoside, 2-hydroxycinnamic acid, M-coumaric acid, L-phenylalanine, and vitexin 6-O-glucoside, all of which are closely associated with flavonoid biosynthesis. In addition, metabolites linked to sugar metabolism, such as trehalulose, inulobiose, glucose-6-glutamate, D-glucose, and 6-O-α-D-glucopyranosyl, were detected.

KEGG pathway enrichment analysis of the DAMs highlighted several significant pathways, including purine metabolism

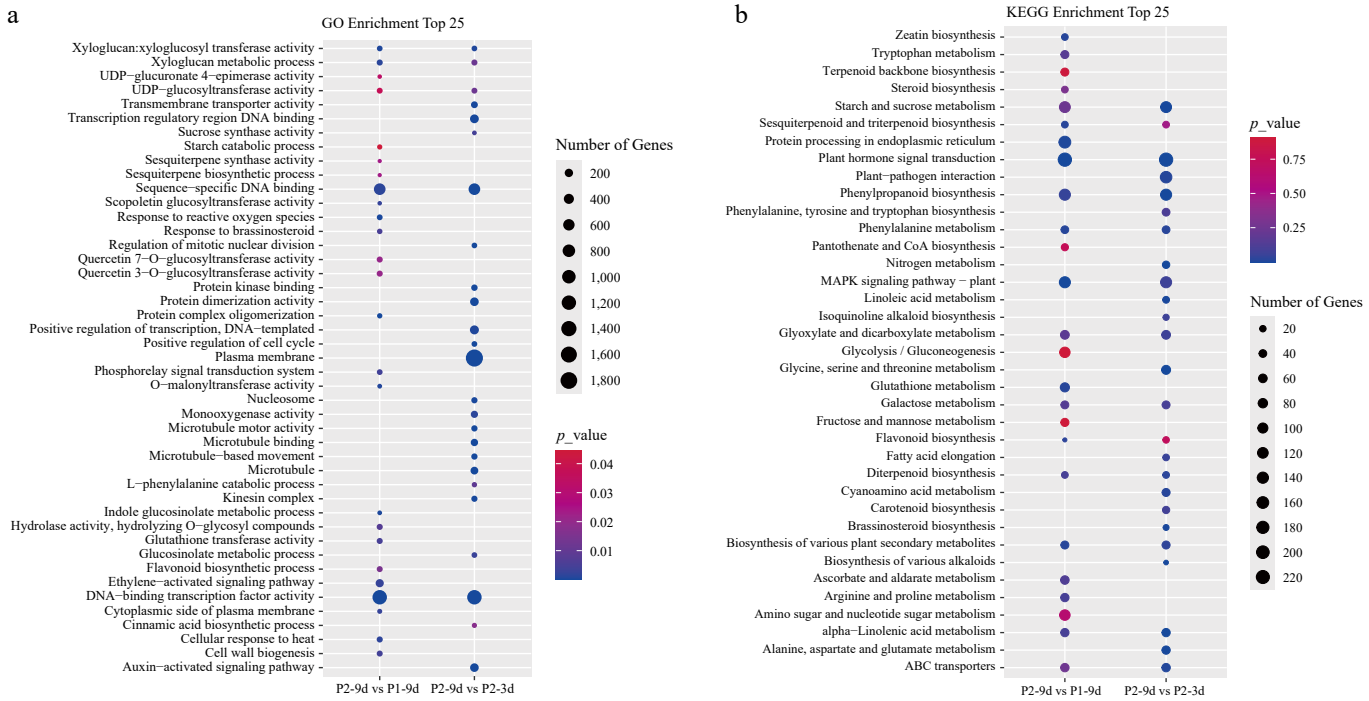

Fig. 3 (a) GO, and (b) KEGG enrichment analyses of DEGs in P2-9d vs P1-9d, and P2-9d vs P2-3d comparisons.

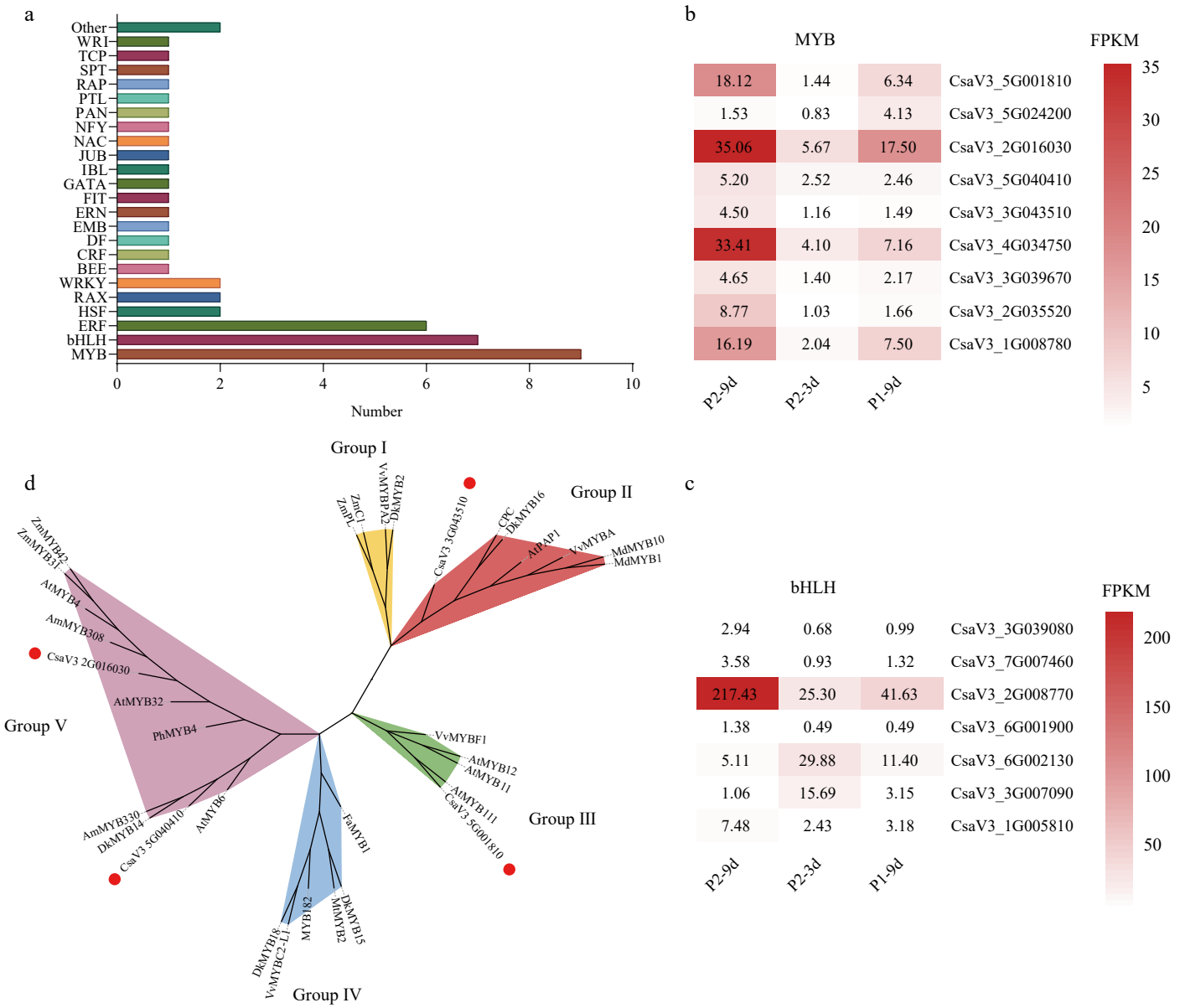

Fig. 4 Differentially expressed transcription factors (TFs) common to both comparisons. (a) Distribution of TF families. Gene expression levels indicated by color bars, showing expression trends of (b) MYB family TFs, and (c) bHLH family TFs. (d) A maximum likelihood phylogenetic tree was inferred for cucumber CsMYBs and orthologous flavonoid biosynthesis regulators from other plant species, incorporating 1,000 bootstrap replicates. Solid red dots represent MYBs isolated in our study; Groups I, II, III, IV, and V indicate proanthocyanidin activators, anthocyanidin, flavonol activators, flavonoid repressors, and phenylpropanoid repressors, respectively. All protein sequence accession numbers from GenBank are tabulated in Supplementary Table S5.

(csv00230), phenylalanine metabolism (csv00360), phenylalanine, tyrosine, and tryptophan biosynthesis (csv00400), phenylpropanoid biosynthesis (csv00940), glycolysis/gluconeogenesis (csv00010), flavonoid biosynthesis (csv00941), starch and sucrose metabolism (csv00500), and glycine, serine, and threonine metabolism (csv00260) (Fig. 5d).

Combined transcriptomic and metabolomic analysis

To elucidate the regulatory network underlying cucumber astringency, an integrated transcriptomic and metabolomic analysis was conducted. KEGG analysis for P2-9d vs P1-9d, and P2-9d vs P2-3d comparisons revealed that DEGs and DAMs showed enrichment in 39 and 55 pathways, respectively (Fig. 6a, b). A total of 24 pathways were common to the two comparisons, including starch and sucrose metabolism, phenylalanine metabolism, flavonoid biosynthesis, ABC transporters, tryptophan metabolism, zeatin biosynthesis, glutathione metabolism, and arginine biosynthesis (Fig. 6c, d).

Joint transcriptome and metabolome profiling of the flavonoid biosynthesis pathway

The transcriptomic and metabolomic analysis suggested that flavonoids are closely associated with cucumber astringency. Functional annotation revealed DEGs and DAMs enriched in flavonoid and phenylpropanoid biosynthesis pathways. Specifically, 21 DEGs, including six *PAL*, four *4CL*, one *CHS*, one *flavonol synthase* (*FLS*), one *F3H*, one *DFR*, three *malonyl-CoA*, one *ANR*, and three *UFGT* genes, were enriched in the two pathways.

A comparative analysis demonstrated that most *PAL* genes exhibited reduced expression in P2-9d compared with both P2-3d and P1-9d. Among the four *4CL* genes, two were downregulated, and two were upregulated. Similarly, one *FLS* gene was downregulated. By contrast, two *malonyl-CoA* genes were upregulated. Among *UFGT* genes, two showed upregulated expression, while one demonstrated downregulation (Fig. 7). The dynamic change of genes

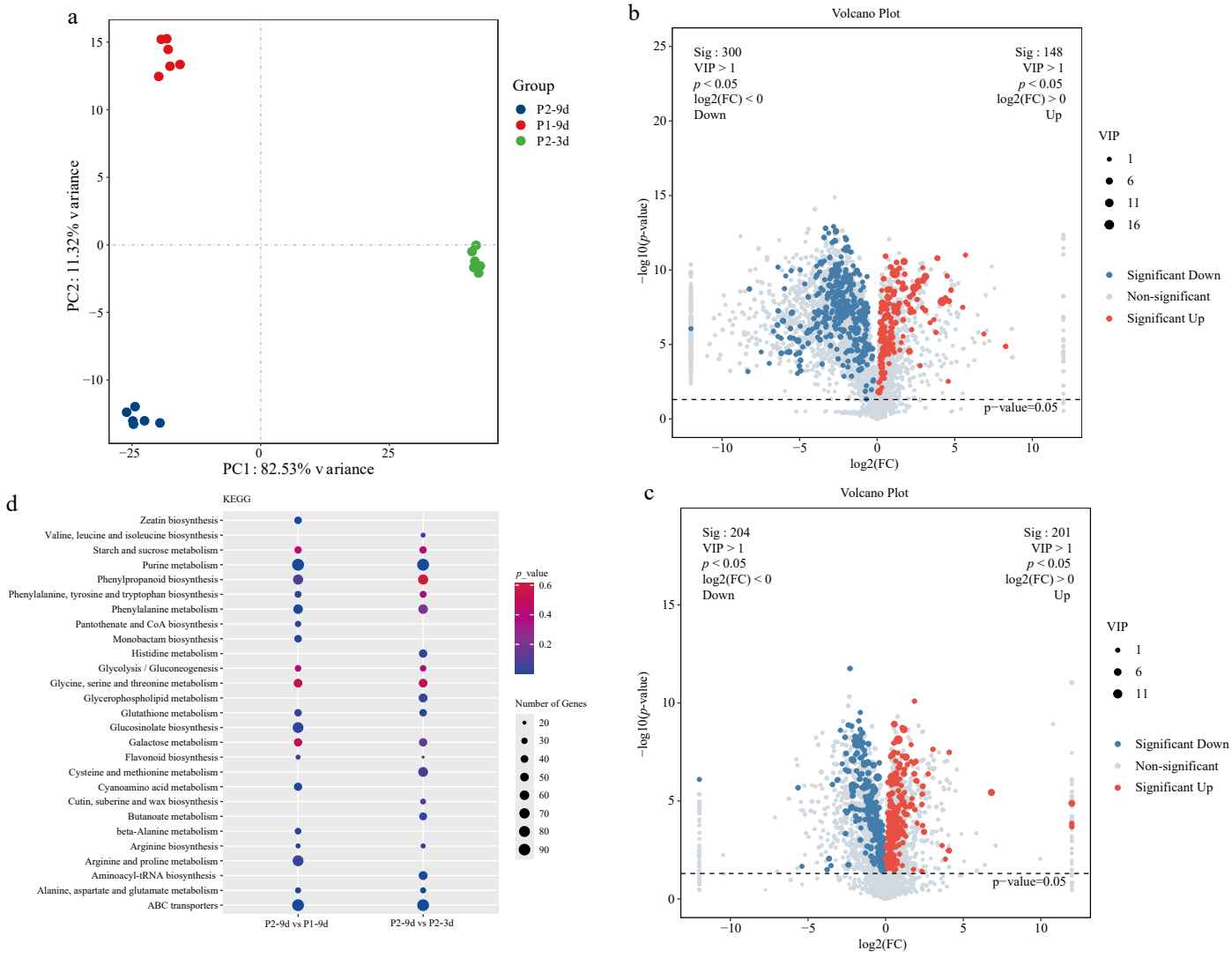

Fig. 5 Metabolomics analysis of different astringency cucumber. (a) PCA score scatter plots of metabolites, with different colors representing different sample groups. (b) Volcano plots of the DAMs for P2-9d vs P1-9d. (c) Volcano plots of the DAMs for P2-9d vs P2-3d. (d) Top 20 KEGG enrichment terms for P2-9d vs P1-9d, and P2-9d vs P2-3d comparisons; dot size in the plot indicates the number of DEGs.

within the flavonoid biosynthesis pathway suggests their potential positive or negative correlation with cucumber astringency.

Identification of candidate genes

Verifying whether the 21 candidate genes and TFs participated in the formation of astringent substances in cucumbers, their expression levels were quantified across different astringency resources (Fig. 1a). Notably, *CsPAL* (CsaV3_4G002290, CsaV3_4G002320) and *Cs4CL* (CsaV3_2G007940, CsaV3_1G031500) exhibited higher expression in fruits with slight astringency than in those with strong astringency, indicating their potential negative regulatory role in cucumber fruit astringency. Conversely, *CsUFGT* (CsaV3_3G018340, CsaV3_4G032410), *CsMYB6* (CsaV3_5G040410), *CsMYB111* (CsaV3_5G001810), *CsbHLH35* (CsaV3_2G008770), and *CsbHLH51* (CsaV3_6G001900) exhibited relatively high expressions in fruits with strong astringency, suggesting their positive regulatory role in cucumber fruit astringency (Fig. 8; Supplementary Fig. S4).

Discussion

Astringency is a key sensory attribute of horticultural plants and a significant factor affecting the taste of vegetables and fruits. This study pioneers the elucidation of molecular mechanisms driving astringency formation in cucumber fruits through an integrated transcriptomic and metabolomic analysis. These findings highlight the critical role of polyphenols, especially flavonoids, in mediating cucumber astringency development.

The phenylpropanoid biosynthesis pathway is a critical route for polyphenol synthesis, involving key enzymes such as PAL, C4H, and 4CL. Among these, PAL and 4CL are considered pivotal in this pathway[32]. Scharbert et al. identified catechins and flavonoid glycosides as primary contributors to the astringency of black tea[33]. During the maturation process of tea leaves, the abundance of catechins exhibits a significant positive correlation with the expression of PAL[34]; however, a negative relationship has also been reported. For instance, *PAL* expression negatively correlates with catechin content in albino tea plants (*AnJiBaiCha*)[35]. Here, it was found that the *CsPALs* (CsaV3_4G002290, CsaV3_4G002320) exhibited higher expressions in slightly astringent cucumber fruit than in strongly astringent ones, mirroring findings in *AnJiBaiCha* (Fig. 8). In *A. thaliana*, four *PAL* members (*AtPAL1-4*) are encoded in the genome, with *AtPAL1* and *AtPAL2* involved in flavonoid synthesis and *AtPAL3* and *AtPAL4* in lignin synthesis[36,37]. Similarly, in this study, six *CsPAL* genes were identified, of which only two correlated with cucumber astringency, suggesting that the other four genes

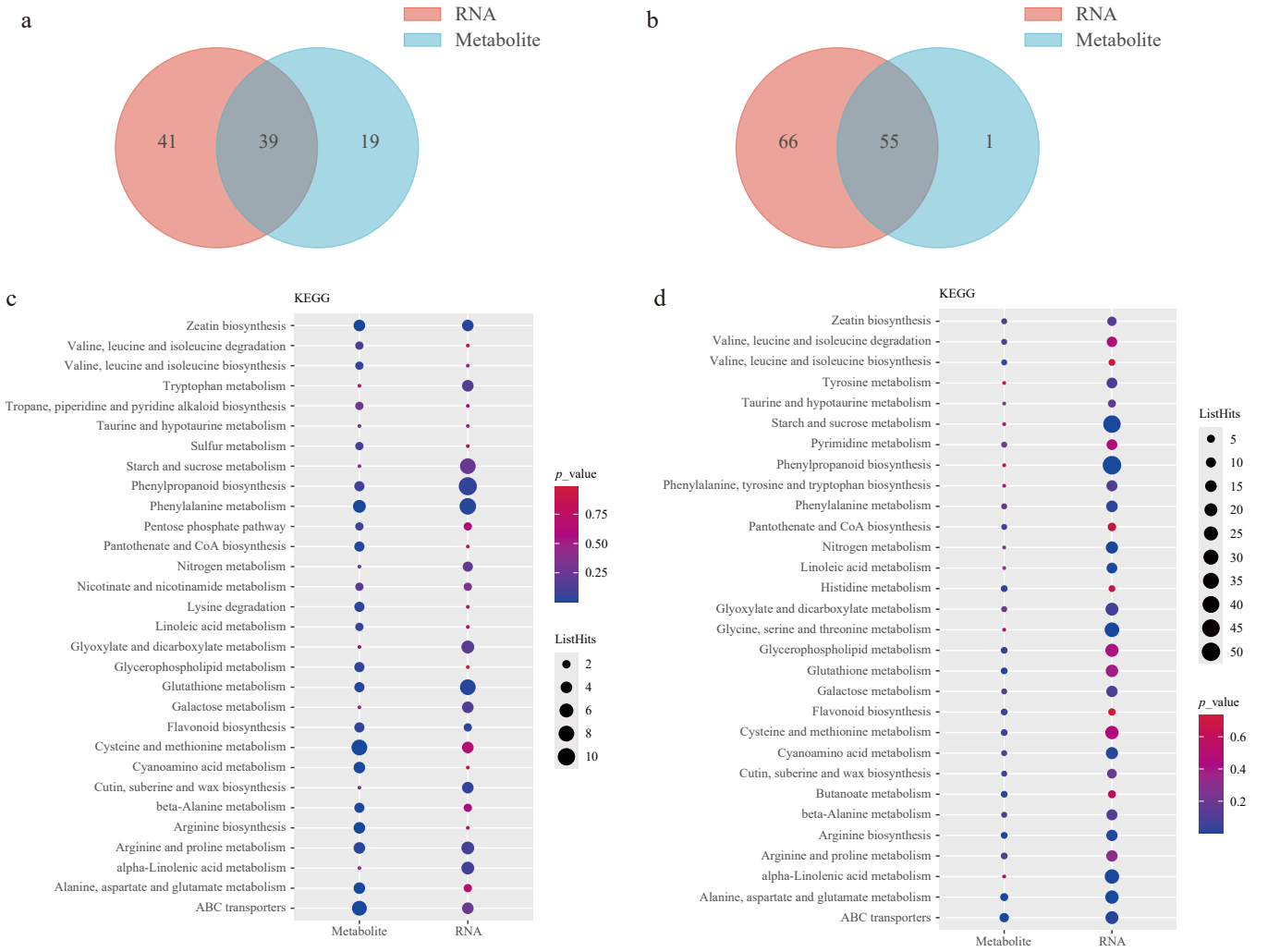

Fig. 6 Integrated analysis of transcriptome and metabolome data. (a), (b) Venn diagram of the mapping pathways of genes and metabolites for P2-9d vs P1-9d, and P2-9d vs P2-3d, respectively. (c), (d) Top 30 KEGG enrichment terms of DEGs and DAMs for P2-9d vs P1-9d, and P2-9d vs P2-3d comparisons, respectively.

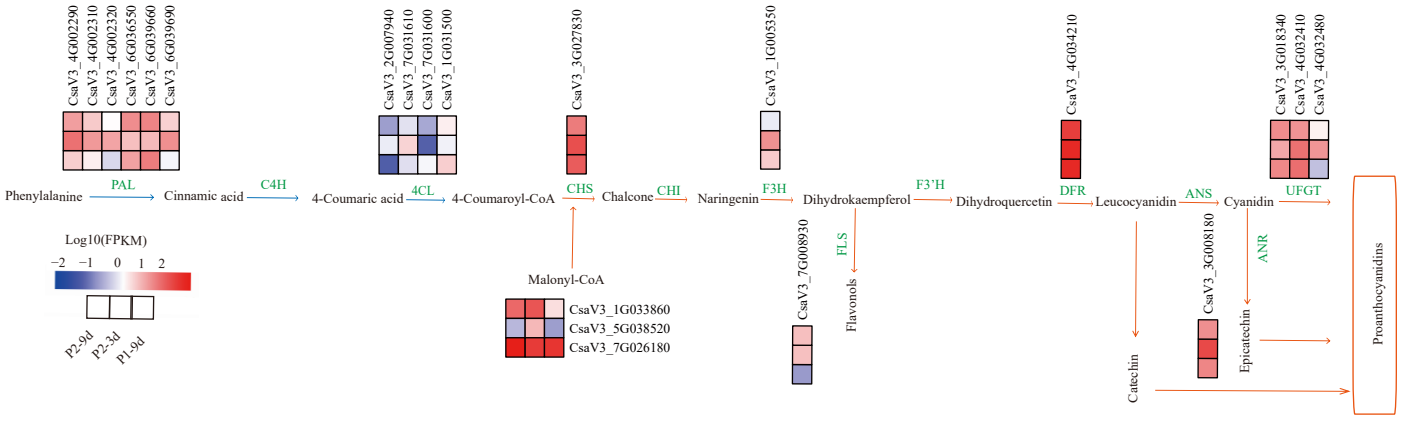

Fig. 7 DEGs involved in flavonoid biosynthetic pathways. The phenylpropanoid metabolic pathway is indicated by blue arrows, the flavonoid metabolic pathway by orange arrows, and green words indicate enzymes in the metabolic process. The red block represents upregulation, and the blue represents downregulation of gene expression. Enzyme annotation: PAL (Phenylalanine ammonia-lyase), C4H (Cinnamate4-hydroxylase), 4CL (4-coumarate CoA ligase), CHS (Chalcone synthase), CHI (Chalcone isomerase), F3H (Flavonoid 3-hydroxylase), FLS (Flavonol synthase), F3'H (Flavonoid 3'-hydroxylase), DFR (Dihydroflavonol 4- reductase), LAR (Leucocyanidin reductase), ANR (Anthocyanidin reductase), ANS (Anthocyanidin synthase), UFGT (UDP-flavonoid 3-O-glucosyltransferase).

may function in lignin synthesis. Future research is necessary to elucidate the regulatory functions of PAL family genes in polyphenol biosynthesis pathways. 4CL enzymes are critical for synthesizing hydroxycinnamate-CoA thioesters, the direct precursors required for flavonoid biosynthesis. The cucumber genome contains six *4CL* genes (*Cs4CL1-6*)[38]. Although the functions of *4CL*

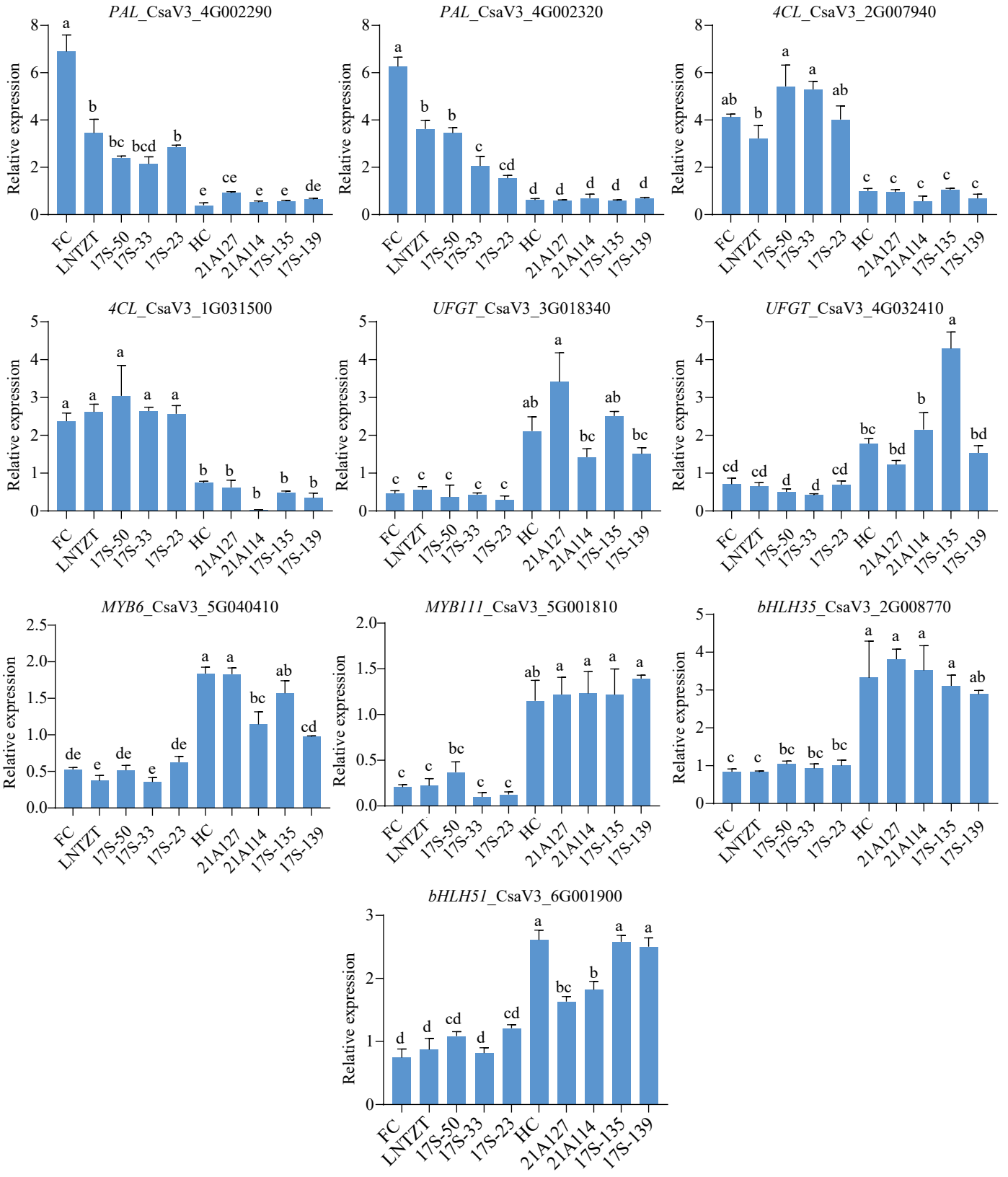

Fig. 8 qRT-PCR analysis of the ten candidate genes. Cucumber *UBIQUITIN* was selected as the housekeeping gene for expression normalization. Data are presented as the mean ± SD of three biological replicates, and statistical significance was calculated by one-way ANOVA followed by Tukey's test ($p \leq 0.05$).

was functionally validated in diverse plants, such as *A. thaliana*[39], rice[40], and liverwort[41], their roles in cucumber remain largely unexplored. The integrated transcriptomic and metabolomic analysis identified four *4CL* genes, with CsaV3_2G007940 and CsaV3_1G031500 exhibiting a negative correlation with cucumber astringency.

The flavonoid biosynthesis pathway is a major route for polyphenol synthesis, culminating in the production of proanthocyanidins (condensed tannins). UFGT is pivotal in mediating flavonoid glycosylation to produce stable proanthocyanidins, which is the primary contributor to astringency in persimmons[13,30]. In this study, two *UFGT* (CsaV3_3G018340 and CsaV3_4G032410) were detected with upregulated expression in strongly astringent cucumber fruits. This suggests that the proanthocyanidin content is relatively high in these fruits, supporting the hypothesis that proanthocyanidins are key compounds responsible for cucumber astringency.

TFs are crucial for polyphenol synthesis. Specifically, flavonoid biosynthesis is regulated by the MYB-bHLH-WD40 transcription complex, a mechanism previously reported in *A. thaliana*[7], strawberries[42], and tea[31]. In this study, nine MYB and seven bHLH TFs were identified (Fig. 4). Further analysis revealed that two MYBs (CsaV3_5G040410 and CsaV3_5G001810) and two bHLHs (CsaV3_2G008770 and CsaV3_6G001900) exhibited higher expression levels in strongly astringent cucumber fruits, indicating their positive correlation with cucumber astringency (Fig. 8). These findings align with the role of *CsMYB60* in promoting flavonoid biosynthesis[43].

KEGG analysis demonstrated that the DEGs and DAMs were functionally linked to the sugar metabolism pathway, implying its role in cucumber astringency. However, the relationship between sugar metabolism and astringency must be investigated in detail.

Xu et al. suggested that cucumber astringency is primarily attributed to catechin levels, which show a decreasing trend as fruit development progresses[38]. However, the present study challenges the findings of Xu et al., as HC fruits at 9 dpp exhibited higher astringency than those at 3 dpp. This discrepancy may be attributed to differences in cucumber ecotypes. Future studies on cucumber astringency may consider the ecotype of cucumber and additional sampling points to verify the findings of this study.

Conclusions

Here, the transcriptome and metabolome data of P1-9d, P2-3d, and P2-9d were analyzed. Combined transcriptomic and metabolomic analysis revealed the molecular mechanisms underlying cucumber fruit astringency as well as the candidate genes and metabolic pathways. *CsPAL* (CsaV3_4G002290, CsaV3_4G002320), *Cs4CL* (CsaV3_2G007940, CsaV3_1G031500), *CsUFGT* (CsaV3_3G018340, CsaV3_4G032410), *CsMYB* (CsaV3_5G040410, CsaV3_5G001810), and *CsbHLH* (CsaV3_2G008770, CsaV3_6G001900) were identified as the key genes related to astringency of cucumber fruits. These findings provide a theoretical foundation for improving cucumber astringency and serve as a reference for studying secondary metabolites in other plants. Future work could experimentally validate gene functions of the candidate genes, dissect regulatory pathways, and leverage multi-omics technologies to improve the understanding of plant secondary metabolites.

Author contributions

The authors confirm their contributions as follows: study design and manuscript revision: Yan L, Zhang J; conducting experiment, data analysis, and manuscript preparation: Zhang J, Wang K; experiments supporting: Song X, Xie Y, Li X, Meng S, Hang Q, Jia J, Wang C. All authors reviewed the results and approved the final version of the manuscript.

Data availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Acknowledgments

This work was supported by a grant from the Scientific Research Fund of Hebei Normal University of Science and Technology (Grant No. 2024YB008) the Key R&D Program of Hebei Province (Grant No. 21326309D).

Conflict of interest

The authors declare that they have no conflict of interest.

Supplementary information accompanies this paper at (https://www.maxapress.com/article/doi/10.48130/vegres-0025-0022)

Dates

Received 6 February 2025; Revised 25 March 2025; Accepted 9 April 2025; Published online 20 August 2025

References

1. Ando K, Carr KM, Grumet R. 2012. Transcriptome analyses of early cucumber fruit growth identifies distinct gene modules associated with phases of development. *BMC Genomics* 13(1):518
2. Green BG. 1993. Oral astringency: a tactile component of flavor. *Acta Psychologica* 84(1):119–25
3. de Freitas V, Mateus N. 2012. Protein/polyphenol interactions: past and present contributions. Mechanisms of astringency perception. *Current Organic Chemistry* 16(6):724–46
4. Liu J, Xie J, Lin J, Xie X, Fan S, et al. 2023. The material basis of astringency and the deastringent effect of polysaccharides: a review. *Food Chemistry* 405:134946
5. Lesschaeve I, Noble AC. 2005. Polyphenols: factors influencing their sensory properties and their effects on food and beverage preferences. *The American Journal of Clinical Nutrition* 81(1):330S–335S
6. Singla RK, Dubey AK, Garg A, Sharma RK, Fiorino M, et al. 2019. Natural polyphenols: chemical classification, definition of classes, subcategories, and structures. *Journal of AOAC International* 102(5):1397–400
7. Li S. 2014. Transcriptional control of flavonoid biosynthesis: fine-tuning of the MYB-bHLH-WD40 (MBW) complex. *Plant Signaling & Behavior* 9(1):e27522
8. Vogt T. 2010. Phenylpropanoid biosynthesis. *Molecular Plant* 3(1):2–20
9. Nabavi SM, Šamec D, Tomczyk M, Milella L, Russo D, et al. 2020. Flavonoid biosynthetic pathways in plants: versatile targets for metabolic engineering. *Biotechnology Advances* 38:107316
10. Winkel-Shirley B. 2001. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiology* 126(2):485–93
11. Dixon RA, Liu C, Jun JH. 2013. Metabolic engineering of anthocyanins and condensed tannins in plants. *Current Opinion in Biotechnology* 24(2):329–35
12. Feng H, Li Y, Wang S, Zhang L, Liu Y, et al. 2014. Molecular analysis of proanthocyanidins related to pigmentation in brown cotton fibre (*Gossypium hirsutum* L.). *Journal of Experimental Botany* 65(20):5759–69
13. Chen W, Xiong Y, Xu L, Zhang Q, Luo Z. 2017. An integrated analysis based on transcriptome and proteome reveals deastringency-related genes in CPCNA persimmon. *Scientific Reports* 7(1):44671
14. Mehrtens F, Kranz H, Bednarek P, Weisshaar B. 2005. The *Arabidopsis* transcription factor MYB12 is a flavonol-specific regulator of phenylpropanoid biosynthesis. *Plant Physiology* 138(2):1083–96
15. Stracke R, Ishihara H, Huep G, Barsch A, Mehrtens F, et al. 2007. Differential regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in different parts of the *Arabidopsis thaliana* seedling. *The Plant Journal* 50(4):660–77
16. Dubos C, Le Gourrierec J, Baudry A, Huep G, Lanet E, et al. 2008. MYBL2 is a new regulator of flavonoid biosynthesis in *Arabidopsis thaliana*. *The Plant Journal* 55(6):940–53

17. Matsui K, Umemura Y, Ohme-Takagi M. 2008. AtMYBL2, a protein with a single MYB domain, acts as a negative regulator of anthocyanin biosynthesis in *Arabidopsis*. *The Plant Journal* 55(6):954–67
18. Bogs J, Jaffé FW, Takos AM, Walker AR, Robinson SP. 2007. The grapevine transcription factor VvMYBPA1 regulates proanthocyanidin synthesis during fruit development. *Plant Physiology* 143(3):1347–61
19. Nesi N, Debeaujon I, Jond C, Pelletier G, Caboche M, et al. 2000. The *TT8* gene encodes a basic helix-loop-helix domain protein required for expression of *DFR* and *BAN* genes in *Arabidopsis siliques*. *The Plant Cell* 12(10):1863–78
20. Szymanski J, Brotman Y, Willmitzer L, Cuadros-Inostroza Á. 2014. Linking gene expression and membrane lipid composition of *Arabidopsis*. *The Plant Cell* 26(3):915–28
21. Chen S, Xu J, Liu C, Zhu Y, Nelson DR, et al. 2012. Genome sequence of the model medicinal mushroom *Ganoderma lucidum*. *Nature Communications* 3(1):913
22. Ren SL, Zhu XY, Yan LY, Li XL. 2023. Establishment of sensory evaluation method for astringency of cucumber and its application in germplasm evaluation. *China Cucurbits and Vegetables* 3:36–41
23. Chen S, Zhou Y, Chen Y, Gu J. 2018. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34(17):i884–i890
24. Kim D, Langmead B, Salzberg SL. 2015. HISAT: a fast spliced aligner with low memory requirements. *Nature Methods* 12(4):357–60
25. Anders S, Pyl PT, Huber W. 2015. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31(2):166–69
26. Love M, Anders S, Huber W. 2014. Differential analysis of count data – the DESeq2 package. *Genome Biology* 15:550
27. Wu Z, Wang Z, Xie Y, Liu G, Shang X, et al. 2023. Transcriptome and metabolome profiling provide insights into flavonoid synthesis in *Acanthus ilicifolius* Linn. *Genes* 14(3):752
28. Zhang J, Wang X, Dong X, Wang F, Cao L, et al. 2022. Expression analysis and functional characterization of tomato Tubby-like protein family. *Plant Science* 324:111454
29. Jin H, Cominelli E, Bailey P, Parr A, Mehrtens F, et al. 2000. Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in *Arabidopsis*. *The EMBO Journal* 19(22):6150–61
30. Chen W, Zheng Q, Li J, Liu Y, Xu L, et al. 2021. DkMYB14 is a bifunctional transcription factor that regulates the accumulation of proanthocyanidin in persimmon fruit. *The Plant Journal* 106(6):1708–27
31. Li P, Fu J, Xu Y, Shen Y, Zhang Y, et al. 2022. CsMYB1 integrates the regulation of trichome development and catechins biosynthesis in tea plant domestication. *New Phytologist* 234(3):902–17
32. Baldi P, Moser M, Brilli M, Vrhovsek U, Pindo M, et al. 2017. Fine-tuning of the flavonoid and monolignol pathways during apple early fruit development. *Planta* 245(5):1021–35
33. Scharbert S, Hofmann T. 2005. Molecular definition of black tea taste by means of quantitative studies, taste reconstitution, and omission experiments. *Journal of Agricultural and Food Chemistry* 53(13):5377–84
34. Singh K, Kumar S, Rani A, Gulati A, Ahuja PS. 2009. Phenylalanine ammonia-lyase (PAL) and cinnamate 4-hydroxylase (C4H) and catechins (flavan-3-ols) accumulation in tea. *Functional & Integrative Genomics* 9(1):125–34
35. Xiong L, Li J, Li Y, Yuan L, Liu S, et al. 2013. Dynamic changes in catechin levels and catechin biosynthesis-related gene expression in albino tea plants (*Camellia sinensis* L.). *Plant Physiology and Biochemistry* 71:132–43
36. Huang J, Gu M, Lai Z, Fan B, Shi K, et al. 2010. Functional analysis of the *Arabidopsis PAL* gene family in plant growth, development, and response to environmental stress. *Plant Physiology* 153(4):1526–38
37. Olsen KM, Lea US, Slimestad R, Verheul M, Lillo C. 2008. Differential expression of four *Arabidopsis PAL* genes; *PAL1* and *PAL2* have functional specialization in abiotic environmental-triggered flavonoid synthesis. *Journal of Plant Physiology* 165(14):1491–99
38. Xu X, Tian H, He M, Gebretsadik K, Qi X, et al. 2019. Changes in catechin contents and expression of catechin biosynthesis-associated genes during early cucumber fruit development. *Acta Physiologiae Plantarum* 41(8):130
39. Gao S, Yu HN, Xu RX, Cheng AX, Lou HX. 2015. Cloning and functional characterization of a 4-coumarate CoA ligase from liverwort *Plagiochasma appendiculatum*. *Phytochemistry* 111:48–58
40. Gui J, Shen J, Li L. 2011. Functional characterization of evolutionarily divergent 4-coumarate: coenzyme a ligases in rice. *Plant Physiology* 157(2):574–86
41. Ehlting J, Büttner D, Wang Q, Douglas CJ, Somssich IE, et al. 1999. Three 4-coumarate: coenzyme A ligases in *Arabidopsis thaliana* represent two evolutionarily divergent classes in angiosperms. *The Plant Journal* 19(1):9–20
42. Xu P, Wu L, Cao M, Ma C, Xiao K, et al. 2021. Identification of MBW complex components implicated in the biosynthesis of flavonoids in woodland strawberry. *Frontiers in Plant Science* 12:774943
43. Liu M, Zhang C, Duan L, Luan Q, Li J, et al. 2019. *CsMYB60* is a key regulator of flavonols and proanthocyanidans that determine the colour of fruit spines in cucumber. *Journal of Experimental Botany* 70(1):69–84