BMC Genomics

RESEARCH Open Access

Genome-wide investigation of glycoside hydrolase 9 (GH9) gene family unveils implications in orchestrating the mastication trait of *Citrus sinensis* fruits

Chengyan Deng[1†], Yingtian Guo[1†], Jingjuan Zhang[2] and Guizhi Feng[1*]

Abstract

Mastication trait of citrus significantly influences the fruit's overall quality and consumer preference. The accumulation of cellulose in fruits significantly impacts the mastication trait of citrus fruits, and the glycoside hydrolase 9 (GH9) family plays a crucial role in cellulose metabolism. In this study, we successfully identified 32 GH9 genes from the *Citrus sinensis* genome and subsequently conducted detailed bioinformatics analyses of the GH9 family. Additionally, we profiled the spatiotemporal expression patterns of *CsGH9* genes across four distinct fruit tissue types and six crucial developmental stages of citrus fruits, leveraging transcriptome data. Parallel to this, we undertook a comparative analysis of transcriptome profiles and cellulose content among diverse fruit tissues spanning six developmental stages. Furthermore, to identify the pivotal genes involved in cellulose metabolism within the GH9 family during fruit maturity, we employed correlation analysis between cellulose content and gene expression in varying tissues across diverse citrus varieties. This analysis highlighted key genes such as *CsGH9A2/6* and *CsGH9B12/13/14/22*. Collectively, this study provides an in-depth analysis of the GH9 gene family in citrus and offers novel molecular insights into the underlying mechanisms governing the mastication trait formation in citrus fruits.

Keywords Citrus, GH9 family, Bioinformatics analyses, Fruit mastication, Cellulose

Introduction

Cellulose, a ubiquitous polysaccharide found in the cell walls of plants, serves as a fundamental structural component in the plant [1]. It affects the mechanical strength of cell walls and further plays an important role in the formation of fruit quality [2]. In recent years, there has been significant attention on the study of cellulose-degrading enzymes and their roles in processes like fruit hardness, mastication traits, and stone cell formation in the field of the formation and regulation of fruit quality [2–4]. Among these enzymes, the GH9 family of cellulases has emerged as a critical player in cellulose metabolism [5].

†Chengyan Deng and Yingtian Guo contributed equally to the article.

*Correspondence:
Guizhi Feng
fgz@lyu.edu.cn
[1]College of Agriculture and Forestry Science, Linyi University, Linyi 276000, China
[2]National Key Laboratory for Germplasm Innovation & Utilization of Horticultural Crops, Huazhong Agricultural University, Wuhan 430070, PR China

In the initial phases of fruit development, the cellulose content within the primary cell wall assumes a significant role, offering structural reinforcement while facilitating cellular expansion and growth [6, 7]. Specifically, apples exhibit a high concentration of cellulose during their early stages of maturation, which contributes significantly to their structural integrity and brittleness [8]. As fruits progress towards ripening, the cellulose in their cell walls undergoes degradation, a process catalyzed by enzymes like cellulase [9]. In the case of pears, cellulose deposition can lead to the development of stone cells, thereby influencing the overall taste of the fruit [10]. GH9 family comprises endo-β-1,4-glucanases that are pivotal in various biological processes, including cellulose metabolism, cell wall formation, organ abscission, and fruit development [11, 12]. Notably, cellulose degradation is a crucial step in rendering fruits softer and more palatable. In *Arabidopsis*, KOR1, a member of the GH9 family, plays a fundamental role in cellulose biosynthesis for both primary and secondary cell walls [13, 14]. Considerable experimental evidence suggests that KOR1 and its homologs, such as *PttCel9A1*, have a direct impact on plant growth and modulate cellulose content in the stem [15, 16]. Despite their prevalence, the specific functions and regulatory mechanisms of GH9 family genes in fruit development and mastication traits remain relatively understudied.

Citrus is cultivated globally, and its mastication trait plays a crucial role in determining the texture and taste of the fruit, serving as a key indicator of its overall quality. The mastication trait of citrus fruits is closely linked to the cell wall components of both the segment membrane and juice sacs, primarily encompassing pectin, cellulose, and lignin [17–19]. Among these components, cellulose content has been identified as a key factor influencing the mastication trait of citrus fruit; lower cellulose content corresponds to improved mastication characteristics [20]. Comparing citrus varieties with distinct mastication traits reveals significant variations in cellulose content [21]. Efficient cellulose degradation in citrus fruit is pivotal for enhancing fruit quality. Exploring the role of GH9 family genes in cellulose degradation will further our understanding of the regulatory mechanisms governing the mastication trait in citrus fruit.

In this study, we aim to analyze the GH9 family genes in the citrus genome, examine the cellulose content in the fruit, and clarify their roles in cellulose metabolism during fruit development. Overall, the investigation of cellulose-degrading GH9 family genes in citrus fruits will provide insights into GH9 genes contribute to cellulose metabolism and their impact on the mastication traits of citrus fruits.

Result

Identification of glycosyl hydrolase family 9 (GH9) members in *Citrus sinensis*

Through genomic analysis, we identified 32 GH9s from *Citrus sinensis*, in comparison to the previously reported 25 GH9s in *Arabidopsis thaliana* and *Oryza sativa* [22, 23]. *CsGH9* proteins were defined and verified using CDD and SMART, chosen for their established relevance in characterizing GH9 proteins. This classification includes 7 Type-A and 25 Type-B CsGH9 proteins. Basic information on *CsGH9* genes, proteins in *CsGH9* range in length from 110 to 792 amino acids, have an estimated molecular weight range of 11,789.15 to 86,881.3 Da, and have a theoretical pI between 4.57 and 9.84. The results of *CsGH9s* reported in this paper are shown in Table S1.

Phylogenetic analysis, chromosomal localization, intergenomic relationships, and covariate analysis of *CsGH9* genes in *Citrus sinensis*

To investigate the evolutionary relationship among different GH9 gene family members among citrus, rice, and *Arabidopsis*, the full-length protein sequences were used to generate a phylogenetic tree constructed using the Maximum Likelihood Method with MEGA11 (Fig. 1A). 32 genes in the citrus genome were assigned names *CsGH9A1* through *CsGH9A7* and *CsGH9B1* through *CsGH9B25* based on their clustering relationship with GH9 genes in *Arabidopsis* and rice. Further, the GH9s gene exhibited uneven distribution across various chromosomes in the citrus genome, except chr7 (Fig. 1B), suggesting potential genomic dynamics or selective pressures. Among them, the observed collinear relationship between *CsGH9B22* and *CsGH9B17* implies a conserved gene structure and potentially similar biological functions, underscoring the evolutionary link between these two gene groups.

Structural and motif identification of *CsGH9* genes in *Citrus sinensis*

In order to deepen our comprehension of the structural conservation and divergence of GH9 genes across evolutionary epochs, we acquired the UTR-CDS arrangement within the coding sequences of individual GH9 genes for each group's members (Fig. 2A). Moreover, eight distinct motifs were identified within the *CsGH9s*. While the majority of genes exhibited all eight motifs, certain genes displayed a reduced motif count (Fig. 2B, C). The motifs varied in size, with motif 8 measuring 21 amino acids, and motifs 2, 3, 6, and 7 spanning 50 amino acids. Furthermore, members closely related within the GH9 clade demonstrated comparable conserved motifs. When the motifs were sorted by their frequency within the GH9 families, it became evident that Motif 2 and Motif 5 were the most prevalent among genes, appearing in 26 genes,

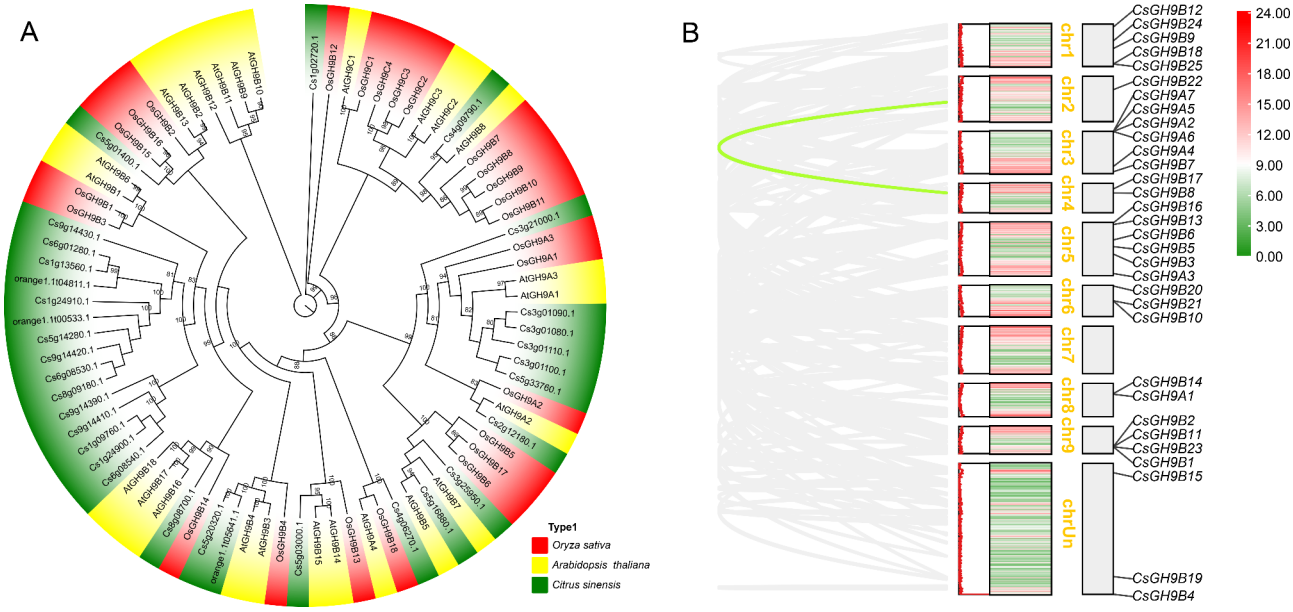

Fig. 1 Phylogenetic relationships among GH9 genes in *Oryza sativa*, *Arabidopsis thaliana*, and *Citrus sinensis* (**A**), and chromosomal localization, gene duplication, and colinearity analyses of *CsPME* genes (**B**). Numbers at each branch node indicate bootstrap probabilities (> 80% based on 1000 bootstrap replicates). Proteins from different species are distinguished by varying colors. The green curved lines depict collinear blocks within the entire citrus genome

whereas Motif 7 was the least common, being found in only 12 genes (Fig. 2B, C).

Promoter characteristics of *CsGH9* genes in *Citrus sinensis*

To elucidate the regulatory functions of the promoters of the *CsGH9* genes in *Citrus sinensis*, we extracted the 2000 bp upstream promoter sequences. As illustrated in Fig. 3A, these sequences contain numerous cis-acting elements associated with hormone responsiveness and environmental stress response. Our analysis of hormone-related cis-acting elements revealed that all *CsGH9* genes include at least one element responsive to phytohormones. Regarding abiotic stress-related elements, we identified several key motifs: MBS and TCCC-motif for drought response, STRE and TC-rich repeats for defense/stress response, WUN-motif and WRE3 for wound response, and LTR for cold response. Among these, MBS is the most widely distributed cis-element within the GH9 gene family, whereas LTR elements are present in 11 members of the GH9 family (Fig. 3A). Furthermore, we characterized the expression response of GH9 family genes to various plant hormones (ABA, IAA, SA, and GA) and different environmental stresses (drought, defense/stress, wound, and cold). The results indicate differential gene expression responses to these factors. Notably, *CsGH9B5* exhibited a significant induction (3-fold) by IAA, while *CsGH9B3* and *CsGH9B19* were significantly induced (5-fold and 2-fold, respectively) by SA. Conversely, *CsGH9A6* was inhibited (2-fold) by cold stress, whereas *CsGH9B22* showed a substantial induction (4-fold) under the same condition (Fig. 3B).

Bioinformatic analysis for the GH9 family members from different citrus types

To understand the variations within the GH9 gene family across diverse citrus varieties, we screened 23, 24, and 26 GH9 genes in *Citrus grandis*, *Citrus clementina*, and *Citrus reticulata*, respectively (Table S1). The distribution of the GH9 gene across chromosomes was found to be uneven among different citrus cultivars (Fig. S1). The subcellular localizations of GH9 proteins, which are closely related to their functions, were predicted, with most GH9 proteins localized to the cytoskeleton and chloroplast (Table S2). A phylogenetic tree was constructed to explore the evolutionary relationships among the different members of the GH9 gene family in these citrus types (Fig. 4A). In this analysis, we identified 11, 12, and 10 genes exhibiting collinear relationships with *Citrus sinensis*, respectively (Fig. 4B; Table S3). We calculated the Ka/Ks values of gene pairs with replication relationships, which ranged from 0 to 0.7 across different citrus varieties, all values being less than 1 (Table S4). This indicates that these gene pairs underwent strong purifying selection. We also compared cis-elements in the promoter regions of GH9 genes from different citrus varieties in response to plant hormones and environmental stresses. The results revealed similar compositions and quantities of cis-elements. Among plant hormones, ABA and JA elicited the most pronounced responses, while

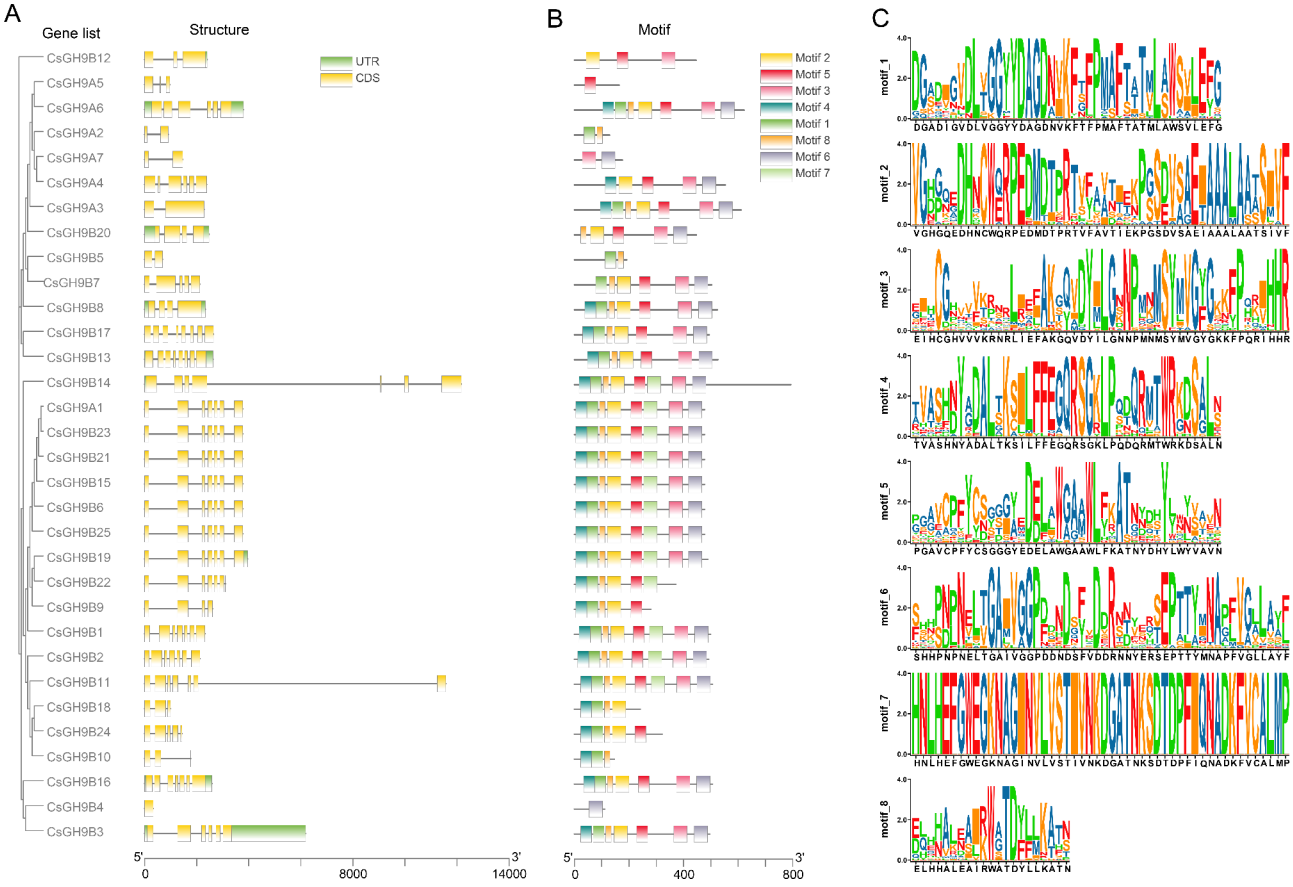

Fig. 2 Structural features **(A)**, motif identification **(B)**, and conserved amino acids **(C)** of *CsGH9* genes. The coding sequence sizes (CDS) and untranslated regions (UTR) are highlighted in yellow and green boxes, respectively. The prediction identified eight conserved motifs within *CsGH9* proteins, with each small box denoting a specific motif. The scale bar represents the length of deduced protein sequences, maintaining proportionality

for environmental stresses, drought and cold stresses triggered the most significant responses (Fig. 4C). Our motif identification and structural analysis showed a high degree of similarity among these genes (Fig. S2).

These findings suggest that while there are distinct differences in the GH9 gene family among various citrus varieties, most characteristics exhibit remarkable similarity.

Spatiotemporal expression patterns of *CsGH9* genes in citrus fruits

The expression patterns of *CsGH9* genes in citrus fruits were investigated in a spatial-temporal context using global gene expression transcriptome data. This data encompassed four distinct fruit tissues (EP, epicarp; AL, albedo; SM, segment membrane; JS, juice sacs) across six developmental stages (50 DAF, 80 DAF, 120 DAF, 155 DAF, 180 DAF, and 220 DAF) as detailed in our previous study (Feng et al., 2021b). Expressions were considered significant when the FPKM (fragments per kilobase of transcript sequence per million base pairs sequenced) value exceeded 0.3 (Kang et al., 2013), leading to the identification of 10 expressed genes from the 32 *CsGH9* gene families. Our focus centered on analyzing the expression characteristics of six highly expressed genes in citrus fruit (Table S5). Figure 5A illustrates that *CsGH9* genes predominantly exhibit low, constitutive, and tissue-specific expression patterns across different citrus fruit tissues during various developmental stages. Notably, *CsGH9B12's* expression increased consistently in all four fruit tissues, while *CsGH9B14* showed a continuous decrease. *CsGH9B13* exhibited specific expression in the later stage of SM, and *CsGH9A2* was specifically expressed in fruit peel tissue (EP and AL), with expression levels steadily increasing. Additionally, *CsGH9A6* and *CsGH9B22* displayed variable expression patterns in different tissues, maintaining high expression in the early stages. Upon examining the overall expression of *CsGH9* family genes in citrus fruits, distinctions emerged among the four tissues. Figure 4B highlights a decreasing expression trend in EP, while AL, SM, and JS demonstrated elevated expression during the initial stage, followed by a continuous increase throughout fruit development and maturation.

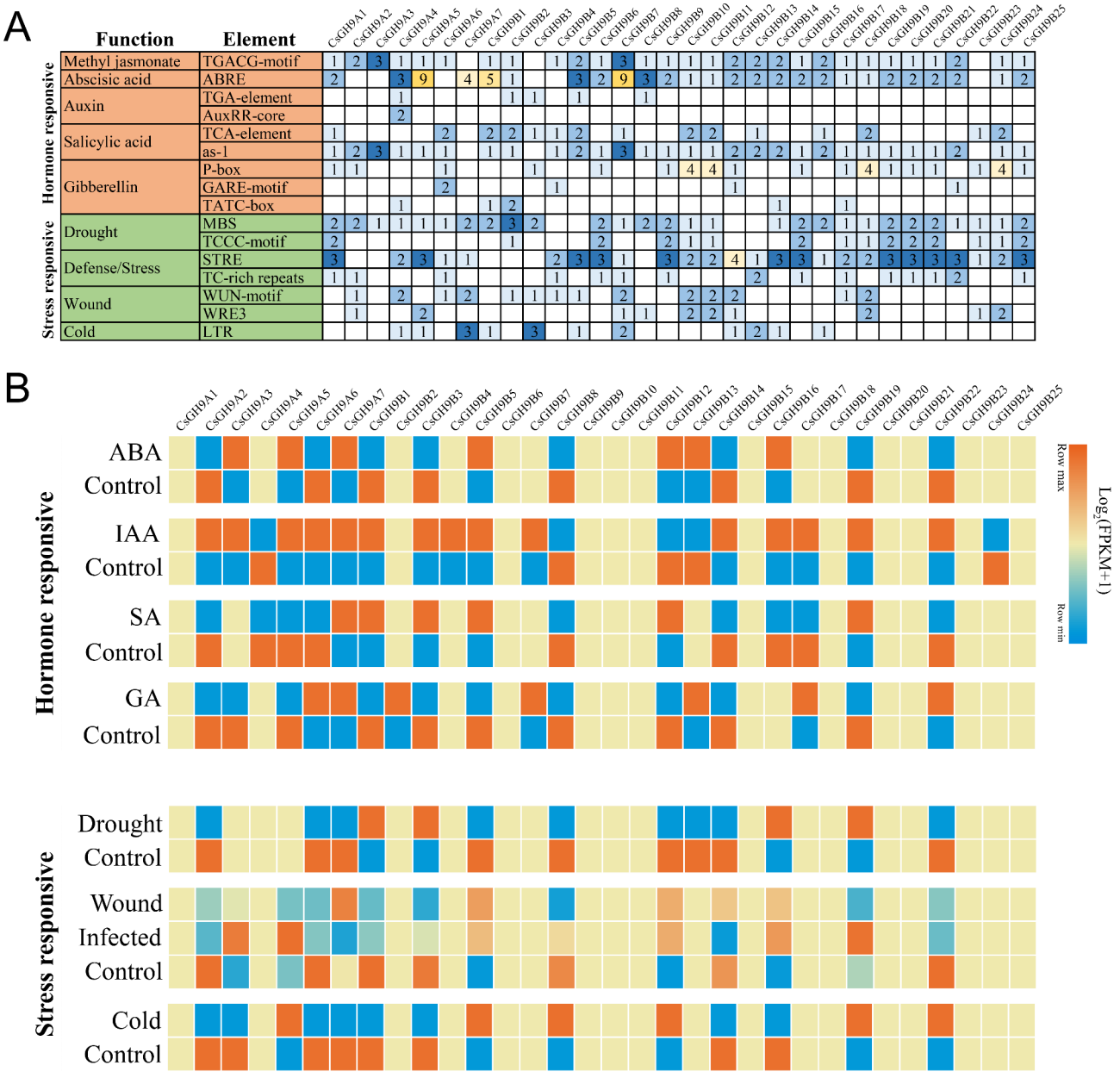

Fig. 3 Analysis of cis-acting regulatory elements **(A)** and their response **(B)** in the promoters of *CsGH9* genes in *Citrus sinensis*. A 2000 bp nucleotide sequence upstream of the initiation codon was considered for each CsGH9 gene to predict cis-regulatory elements

Four distinct fruit tissues were examined: EP (epicarp), AL (albedo), SM (segment membrane), and JS (juice sacs), spanning six developmental stages from young fruits to ripe fruits (50 days after flower (DAF), 80 DAF, 120 DAF, 155 DAF, 180 DAF, and 220 DAF). The hierarchical clustering of gene sets was determined using one minus Pearson correlation as the distance metric and average as the linkage method, based on the $\log_2$ (FPKM+1) values.

Investigation of spatiotemporal cellulose accumulation patterns in citrus fruits

Previous studies have established a correlation between cellulose content and the mastication trait in citrus fruits [20]. To investigate this correlation, we collected samples from three distinct fruit tissues (AL, SM, and JS) at six developmental stages (50 DAF, 80 DAF, 120 DAF, 155 DAF, 180 DAF, and 220 DAF) of the 'Fengjie' navel orange (FJ). Subsequently, we measured the contents of cellulose, and hemicellulose in these citrus fruits (Fig. 6A, Table S6). The cellulose content exhibited a declining trend in all three fruit tissues with the progression of fruit

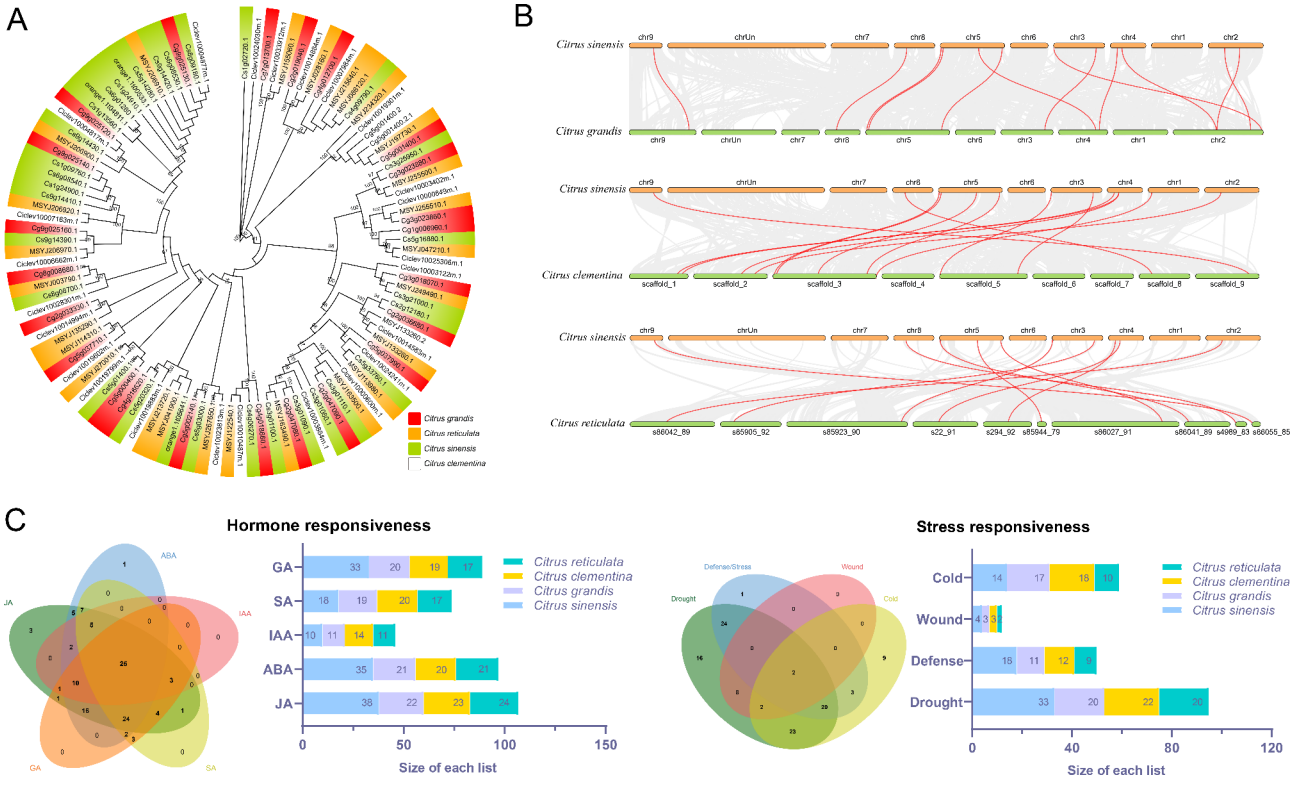

Fig. 4 Differentiation of the GH9 family in diverse citrus varieties. Phylogenetic relationships (**A**), colinearity (**B**), and promoter analyses (**C**). Numbers at each branch node indicate bootstrap probabilities (> 80% based on 1000 bootstrap replicates)

development, and noteworthy variations were observed among these tissues (Fig. 6B). Specifically, cellulose content in the JS tissue increased during the veraison period (155–180 DAF). The hemicellulose content, illustrated by inverse V-shaped curves, displayed distinctive patterns in the AL and SM tissues. Contrarily, in the JS tissue, hemicellulose content exhibited a lower initial level and gradually decreased with fruit development and maturity (Fig. 6B).

Furthermore, for a more comprehensive understanding of the role of cellulose and hemicellulose in the mastication trait of citrus fruits, we collected samples from four major citrus varieties - red tangerine (Clementine Mandarin, CL), pummelo (HB pomelo, HB), ponkan (Huagan No. 2, HG), and nanfeng tangerine (Nanfeng Mandarin, NF) - during the ripening period (Fig. 6C). As shown in Fig. 6D, the hierarchy of cellulose content across various tissues in three citrus varieties is as follows: AL > SM > JS, with the exception of HB. Conversely, the hemicellulose content follows the order: SM > AL > JS, except for HG. Similarly, the JS tissue consistently exhibited the lowest content in both cellulose and hemicellulose among the five varieties.

Identification of key genes involved in cellulose metabolism within the GH9 family of citrus fruits

The significance of the *CsGH9* family of cellulases in cellulose metabolism has been underscored as pivotal [5]. Our investigation delved into the expression patterns across stages and tissue of *CsGH9* genes in citrus fruits, extending our focus to their expression in diverse tissues of different citrus varieties during the ripening stage. As shown in Fig. 7, the expression profiles of the six genes exhibited notable variations among different varieties, with a predominant higher expression observed in AL and SM compared to JS. Among the *CsGH9* family members, the expression levels of *CsGH9A6*, *CsGH9A2*, and *CsGH9B22* genes in AL across different varieties were significantly higher than those in the pulp tissues (SM and JS). Furthermore, the expression of *CsGH9B12* and *CsGH9B13* in SM surpassed that in other tissues, signifying distinct roles within the cellulose metabolism pathway. Finally, we systematically analyzed the correlations between the expression levels of GH9 family genes and cellulose content across various developmental stages, fruit tissues, and citrus varieties. Our findings, detailed in Table S7, reveal a significant correlation between the expression of these six GH9 genes and cellulose accumulation in the fruits. This suggests a potential role for these genes in the regulation of cellulose biosynthesis during fruit development.

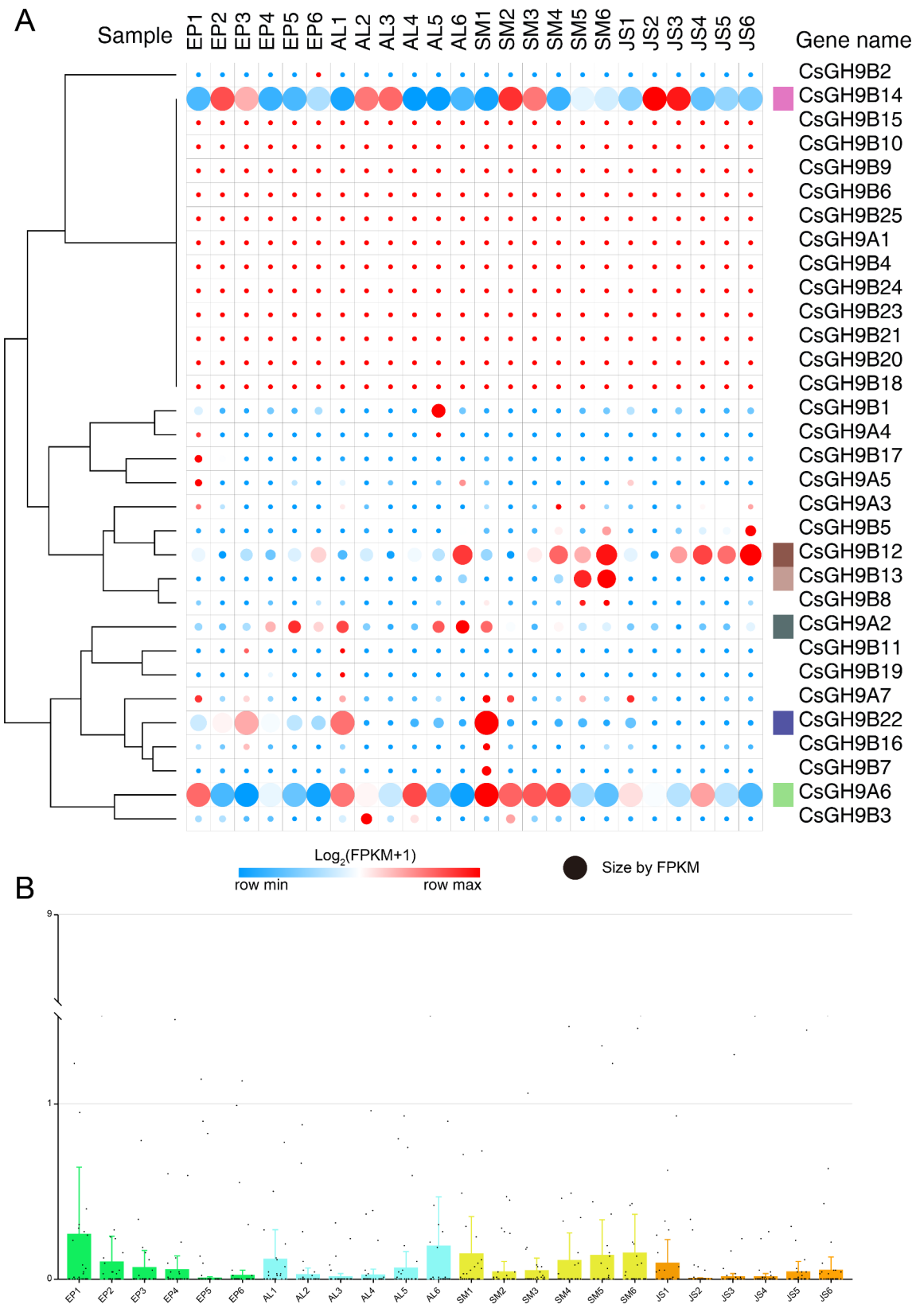

Fig. 5 Spatiotemporal expression of *CsGH9* genes **(A)** and global expression patterns **(B)** in four fruit tissues at six developmental stages in *Citrus sinensis*

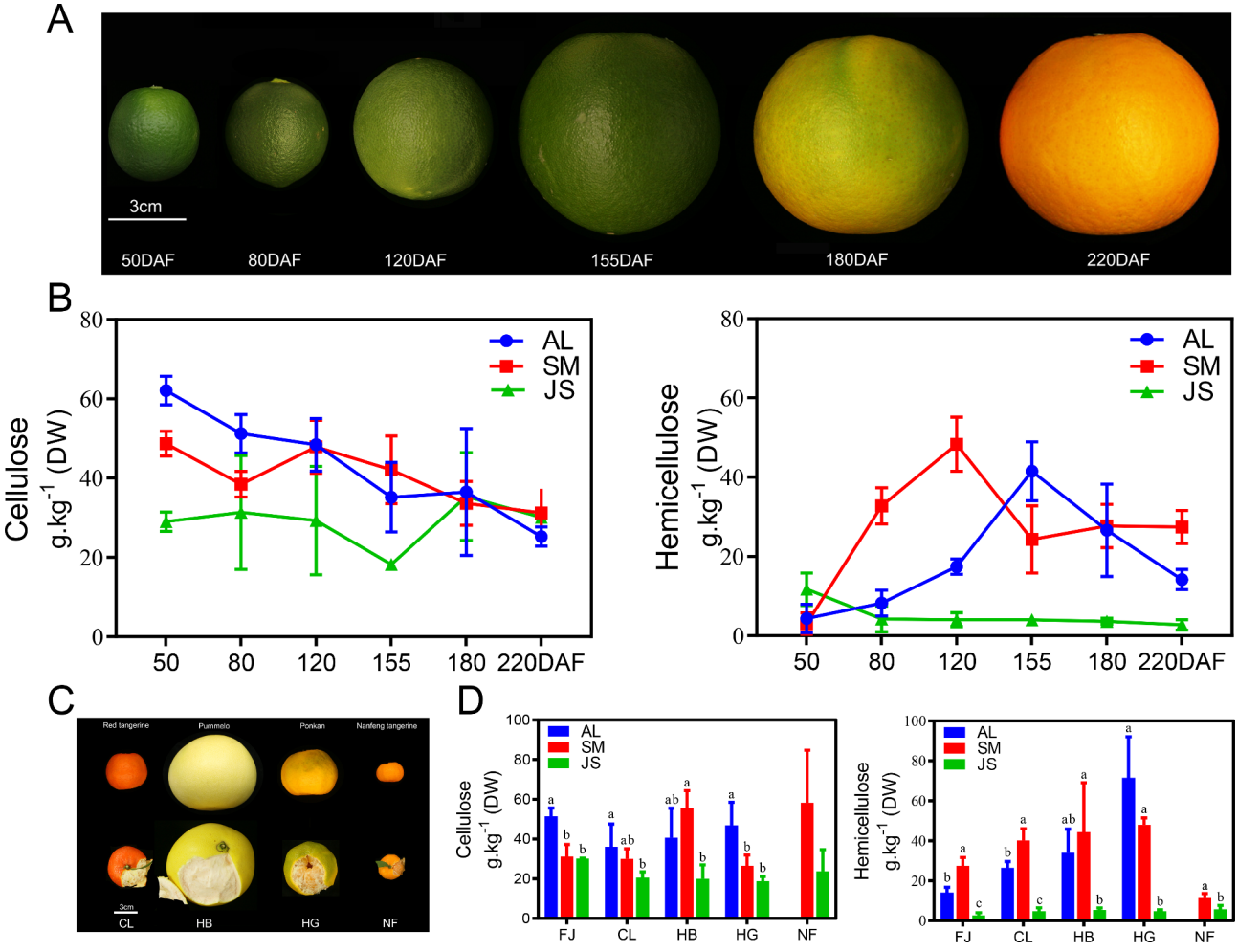

Fig. 6 Cellulose accumulation patterns in citrus fruits. (**A**) and four major citrus varieties including red tangerine (CL), pummelo (HB), ponkan (HG), and nanfeng tangerine (NF) during the ripening period (**C**), collected for cellulose determination. Scale bar = 3 cm. The content of cellulose in *C. sinensis* fruit (**B**) and four major citrus varieties fruit (**D**). Each sample was characterized based on six replicates. Significance was determined by one-way ANOVA in SPSS software, with bars labeled with distinct lowercase letters indicating statistically significant differences ($p < 0.05$)

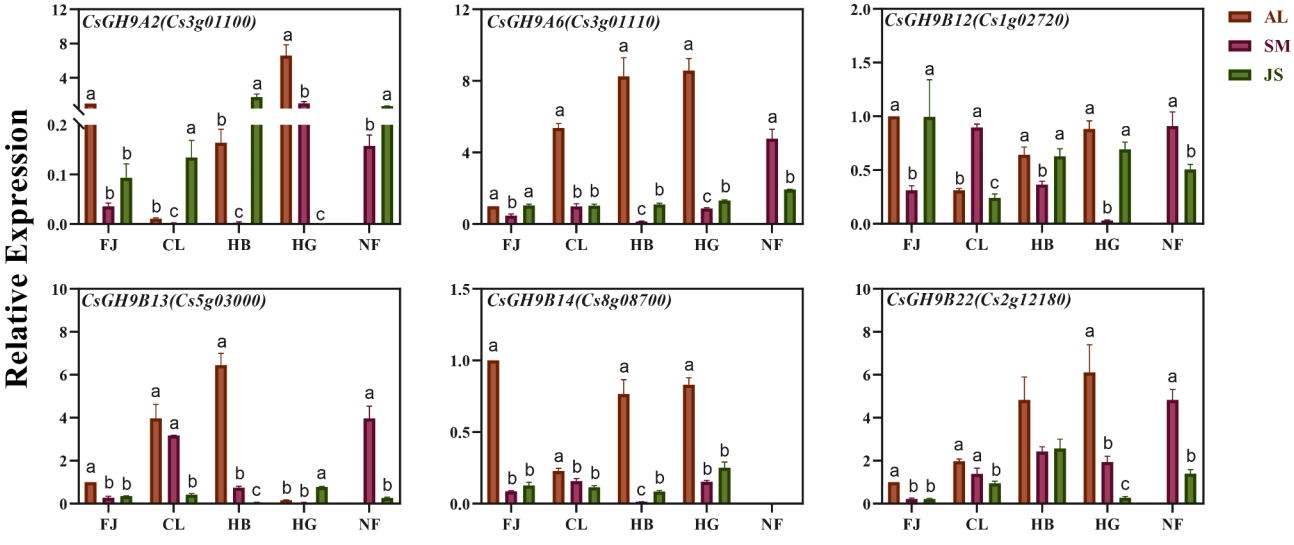

Fig. 7 Expression profiles of *CsGH9* genes analyzed by qRT-PCR

The endogenous reference gene utilized in this study was *CsActin*, and the experimental setup comprised four biological replicates. Significance was determined by one-way ANOVA in SPSS software, with bars labeled with distinct lowercase letters indicating statistically significant differences ($p<0.05$). Citrus varieties examined in the study include navel orange ('Fengjie 72−1' navel orange, FJ), red tangerine (Clementine Mandarin, CL), pummelo (HB pomelo, HB), ponkan (Huagan No. 2, HG), and nanfeng tangerine (Nanfeng Mandarin, NF).

Discussion

Citrus fruits, being globally cultivated, exhibit a distinct mastication trait—a pivotal factor in determining fruit quality. Clarifying the regulatory mechanisms governing the mastication trait in citrus fruits is essential for advancing fruit quality. The current research in this domain is at an early stage, encountering several theoretical and technical challenges demanding thorough exploration. This study contributes valuable insights to the theoretical understanding of superior mastication trait fruit varieties, drawing from the comprehensive analysis of *CsGH9* gene family expression and cellulose content in citrus fruits.

Insights into the role of cellulose in shaping the mastication trait of citrus fruit

The dynamic alterations in cell wall composition throughout fruit ripening play a pivotal role in shaping various horticultural traits [24]. Among the cell wall's primary constituents, including pectin, cellulose, and lignin, cellulose has garnered significant attention due to its potential impact on the mastication properties of citrus fruits [17, 18]. While prior research has primarily concentrated on the overall accumulation of cell wall metabolites in citrus fruits, the specific role of cellulose in relation to mastication is gaining increasing recognition [25, 26]. Our study builds upon previous findings and underscores the importance of cellulose, particularly its spatial and temporal accumulation patterns, in determining the texture and overall quality of citrus fruits. A comparative analysis between the FJ72-1 navel orange and its bud mutant FJWC, which exhibits inferior mastication properties, revealed a higher content of cellulose and hemicellulose in FJWC during fruit ripening [25]. Extending these observations, our current research delves deeper into the spatiotemporal dynamics of cellulose accumulation across different fruit tissues, developmental stages, and citrus varieties. The findings indicate a gradual decrease in cellulose content during fruit development, with notable variations among the albedo (AL), segment membrane (SM), and juice sac (JS) tissues (Fig. 6B). Furthermore, a comparative analysis of cellulose content across different citrus varieties during maturity reveals a consistent pattern, with AL exhibiting the highest cellulose content and JS the lowest (Fig. 6D). Significantly, the HB pomelo, standing as the sole inedible segment membrane (SM) variety, exhibits the highest cellulose content in its SM, a noteworthy finding. This observation underscores the crucial role of cellulose as a determinant of the mastication properties in citrus fruits. The observed patterns of cellulose distribution and accumulation offer profound insights into its influence on fruit texture. To further elucidate the molecular mechanisms and regulatory pathways governing cellulose dynamics and their impact on citrus fruit quality, additional in-depth research is imperative.

Identification of key genes involved in cellulose metabolism of GH9 family of citrus fruit

Previously, research into cellulose and its degrading enzymes in horticultural crops has primarily concentrated on their influence on fruit softening, especially in strawberry [27], tomato [28], and avocado [29]. However, recent studies on *Psidium guajava* have revealed that the endoglucanase of the GH9 family plays a pivotal role in modulating cellulose content during guava fruit development [30]. This discovery suggests that identifying key genes within the GH9 family in citrus fruits may provide crucial insights into the molecular mechanisms underlying the mastication trait. In our investigation of the citrus genome, we successfully identified 32 *CsGH9* family genes (Table S1). Notably, our study highlights specific genes within the GH9 family that may play a vital role in cellulose metabolism in citrus fruits. For instance, *CsGH9A6* exhibited significantly higher expression levels in fruits compared to other GH9 family genes. Specifically, *CsGH9A6* maintained high expression levels in SM and displayed a declining trend throughout fruit development (Fig. 5A). Remarkably, the expression of *CsGH9A6* in the AL of mature citrus fruits surpassed that in other tissues (Fig. 7). These findings indicate that *CsGH9A6* may be a key regulator of cellulose metabolism in citrus fruits, which is essential for maintaining fruit texture and quality.

To elucidate potential regulatory mechanisms, we conducted an in-depth analysis of the cis-acting elements located in the promoter regions of the candidate genes. This comprehensive analysis uncovered several potential regulatory motifs that may contribute to the observed patterns of upregulation and downregulation. Notably, the majority of the GH9 family gene promoters harbor at least one plant hormone responsive element (Fig. 3A), highlighting the pivotal role of plant hormones in regulating the expression of the GH9 family genes. Our previous research on Valencia oranges, which demonstrated that low temperatures can trigger lignin accumulation in the fruit, resulting in compromised mastication qualities

[19]. Thus, the analysis identified the presence of low-temperature responsive elements in the promoter regions of *CsGH9B12/13/14* (Fig. 3A). In contrast, *CsGH9B22* was upregulated 4-fold under low-temperature conditions, despite the absence of low-temperature responsive cis-acting elements in its promoter region (Fig. 3). This discrepancy suggests that *CsGH9B22* may be regulated through indirect mechanisms or influenced by chromatin structure and epigenetic modifications. Further investigation and verification are required to elucidate these regulatory pathways. Although these intriguing findings have piqued our interest, the precise functional roles of these genes in cellulose metabolism remain obscure. A more rigorous exploration of these genes is imperative to gain a comprehensive understanding of their influence on the textural characteristics of citrus fruits, specifically with regards to the mastication trait.

Conclusion

In conclusion, our study has identified a total of 32 *CsGH9* genes in *Citrus sinensis*, and a comprehensive analysis of these genes has been conducted, providing valuable insights for further exploration of *CsGH9* gene functions. Notably, cellulose has emerged as a pivotal factor in shaping the mastication trait of citrus fruits, with specific emphasis on key genes such as *CsGH9A2/6*, *CsGH9B12/13/22*, etc. These findings shed light on the intricate mechanisms underlying the regulation of the citrus fruit mastication trait.

Materials and methods

Plant materials, sample collection, and tissue isolation

Fruit samples of 'Fengjie 72−1' navel orange (*Citrus sinensis* L. Osbeck) were meticulously harvested at distinct developmental stages: 50, 80, 120, 150, 180, and 220 days after flowering (DAF). Additionally, mature stage fruit samples from Huagan No. 2 (*Citrus reticulata* Blanco, PK), HB pomelo (*Citrus grandis* L. Osbeck), Clementine Mandarin (*Citrus clementina* hort. ex Tanaka), and Nanfeng Mandarin (*Citrus reticulata* Blanco cv. Nanfeng tangerine) were collected. For each developmental stage, twelve representative fruits were carefully sampled, ensuring robust biological replicates (three trees per replicate). Employing precise manual dissection techniques, four distinct fruit tissues (Epicarp, Albedo, Segment membrane, and Juice sacs) were isolated and promptly frozen in liquid nitrogen at -80 °C to preserve their biological integrity.

Identification of GH9 gene family members from *Citrus sinensis*

To systematically identify members of the *CsGH9* gene family, two complementary BLAST approaches were employed. Initially, the sweet orange (*Citrus sinensis*, v1.0) genome was queried using the GH9 Hidden Markov Model (HMM) in the HMMER software (version 3.0, http://hmmer.org/, E-value=$1e^{-20}$) [31]. The GH9 domain (PF00759), essential for cellulose hydrolysis, was sourced from the Pfam database (http://pfam.xfam.org/, version 33.1) and employed in the search [32]. Following this, a local BLAST analysis was conducted by comparing 25 *AtGH9* protein sequences from *Arabidopsis thaliana*, obtained from the TAIR database (http://www.Arabidopsis.org/), against the *Citrus sinensis* genome (E-value=$1e^{-5}$).

To confirm the validity of candidate *CsGH9* family genes, each sequence was subjected to verification through the Simple Modular Architecture Research Tool (SMART, http://smart.embl-heidelberg.de/) and the Conserved Domain Database (CCD, http://www.ncbi.nlm.nih.gov/cdd/). This ensured the presence of a GH9 domain, a characteristic feature of GH9 genes [33, 34]. Lastly, the physical and chemical properties of the identified *CsGH9* family proteins were analyzed using ExPASy (https: //web.expasy.org/protparam/) [35]. Subcellular localization predictions were performed using Cell-PLoc (http://www.csbio.sjtu.edu.cn/bioinf/Cell-PLoc-2/) subsequent to gene identification [36].

Phylogenetic analysis, chromosome localization, gene density, and colinear analysis

Phylogenetic analysis aimed to unravel the evolutionary relationships among GH9s utilizing the Maximum Likelihood Method, employing MEGA software. The protein sequences of GH9s from citrus (http://citrus.hzau.edu.cn/index.php), Arabidopsis (https://phytozome-next.jgi.doe.gov/info/Athaliana_TAIR10), and rice (https://phytozome-next.jgi.doe.gov/info/Osativa_v7_0) were utilized for this purpose. Phylogenetic tree was constructed with 1000 bootstraps, and the final visualization and refinement were conducted using tvBOT (Tree Visualization By One Table, https://www.chiplot.online/#) [37]. Chromosome localization, gene density, and colinear analysis were subsequently executed utilizing TBtools software (v1.098769, Advanced Circos) [38].

Gene promoter, motif, domain, and structure analysis

The promoter sequence of *CsGH9s* was obtained by cloning the 2 kb upstream region and subsequently analyzed using the PlantCARE online tool (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) [39]. To identify conservation motifs, the Multiple Em for Motif Elicitation (MEME) online tool (MEME, version 5.1.1, http://meme-suite.org/tools/meme) was employed [40]. The translated amino acid sequences of each *CsGH9s* were subjected to conservation domain identification using the NCBI CD Search tool [41]. Finally, a comprehensive illustration depicting the gene promoter, motifs, domains,

and structures of GH9 family genes was created using the TBtools software [38].

Analysis of gene spatial-temporal expression

Transcriptomic data encompassing 72 samples from four distinct tissues across six developmental stages of sweet orange (*Citrus sinensis* L. Osbeck) were meticulously curated from our prior investigation [42]. The raw transcriptomic data for response analysis were downloaded from the NCBI Sequence Read Archive (SRA) (accession number: SRP325502, SRP155673, SRP345972, SRP404051, SRP483818, and ERP022936), and the analysis was based on previous publication [43]. To unveil the intricate patterns within the gene sets, TBtools, a bioinformatics toolkit, was employed. Hierarchical clustering was executed utilizing the one minus Pearson correlation as the distance metric, coupled with the average linkage method. The gene expression values were processed using $\log_2$ (FPKM+1) [38]. For a comprehensive visualization of the gene expression patterns, a boxplot was meticulously constructed. ChiPlot, a robust tool for visualizing RNA-seq data, was utilized for generating the boxplot, employing $\log_2$ (FPKM+1) as the standardized metric [37].

Determination of cellulose and hemicellulose contents

The quantification of cellulose and hemicellulose concentrations was accomplished through the application of their respective absorbance values, employing the standard curve method. Approximately 5.0 g of citrus flesh tissue underwent dehydration to a constant weight at 80 °C and was subsequently sieved through a 40-mesh sieve. The determination of total cellulose and hemicellulose content adhered to the established procedures outlined in the Plant Cellulose Content Kit (COMIN, CLL-2-Y) and the Plant Hemicellulose Content Kit (COMIN, BXW-2-G).

Quantitative analysis of gene expression

Total RNA extraction was conducted following established protocols [43]. Relative gene expression was assessed in four biological replicates using the QuantStudio™ 6 Flex real-time PCR system (Applied Biosystems), according to the manual of PowerUp™ SYBR™ Green Master Mix (A25742, Thermo Scientific). The housekeeping gene *CsActin* and *CsGAPDH* served as an endogenous reference [44, 45]. Relative abundance of GH9 genes was calculated with the $2^{-\Delta\Delta Ct}$ method. Detailed information on the primers employed for each gene in this study can be found in Supplementary Table S8.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-024-10826-w.

Supplementary Material 1
Supplementary Material 2
Supplementary Material 3
Supplementary Material 4
Supplementary Material 5
Supplementary Material 6
Supplementary Material 7
Supplementary Material 8
Supplementary Material 9
Supplementary Material 10
Supplementary Material 11

Acknowledgements

We sincerely thank professor Hualin Yi and Juxun Wu from Huazhong Agricultural University for the help with plant materials. The authors acknowledge College of Agriculture and Forestry Science, Linyi University, and National Key Laboratory for Germplasm Innovation & Utilization of Horticultural Crops, Huazhong Agricultural University for providing the laboratory and input facility for this experiment.

Author contributions

G.F. conceived the project and designed the research. C.D. and Y.G performed the bioinformatics analysis. G.F. and J.Z. performed most of the experiments. C.D., Y.G. and G.F. wrote the paper. All authors read and approved the manuscript.

Funding

This research was supported by the Natural Science Foundation of China (NSFC, 32302511) and Doctoral Initiating Project of Linyi University, China (No. Z6122057).

Data availability

Data is provided within the manuscript or supplementary information files.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Conflict of interest

The authors declare that they have no conflicts of interest.

Received: 23 May 2024 / Accepted: 23 September 2024
Published online: 30 September 2024

References

1. Pedersen GB, et al. Cellulose synthesis in land plants. Mol Plant. 2023;16(7):1228.
2. Xue Y, et al. The transcription factor PbrMYB24 regulates lignin and cellulose biosynthesis in stone cells of pear fruits. Plant Physiol. 2023;192(3):1997–2014.
3. Tao D. The Relationship between the activity of PG and cx with Dietary Fibre in Sweet Orange Fruit. Acta Horticulturae Sinica; 2007.
4. Li L, et al. Changes in Fruit Firmness, Cell Wall Composition, and Transcriptional Profile in the yellow fruit tomato 1 (yft1) mutant. J Agric Food Chem. 2019;67(1):463–72.

5. Phakeenuya V, et al. A novel multifunctional GH9 enzyme from Paenibacillus curdlanolyticus B-6 exhibiting endo/exo functions of cellulase, mannanase and xylanase activities. Appl Microbiol Biotechnol. 2020;104(5):2079–96.
6. Guerriero G, et al. Callose and cellulose synthase gene expression analysis from the tight cluster to the full bloom stage and during early fruit development in Malus× Domestica. J Plant Res. 2014;127:173–83.
7. Houle A, Conklin-Brittain NL, Wrangham RW. Vertical stratification of the nutritional value of fruit: macronutrients and condensed tannins. Am J Primatol. 2014;76(12):1207–32.
8. Delpino-Rius A, et al. Characterisation of phenolic compounds in processed fibres from the juice industry. Food Chem. 2015;172:575–84.
9. Shen YanHong SY et al. *Isolation of ripening-related genes from ethylene/1-MCP treated papaya through RNA-seq*. 2017.
10. Gong X et al. PbMC1a/1b regulates lignification during stone cell development in pear (Pyrus Bretschneideri) fruit. Hortic Res, 2020. 7.
11. Shani Z, et al. Expression of endo-1,4-β-glucanase (cel1) in Arabidopsis thaliana is associated with plant growth, xylem development and cell wall thickening. Plant Cell Rep. 2006;25(10):1067–74.
12. Wang Y, et al. Genome-wide identification of GH9 gene family and the assessment of its role during fruit abscission zone formation in Vaccinium ashei. Plant Cell Rep; 2023.
13. Shigeru S et al. Role of the putative membrane-bound endo-1,4-beta-glucanase KORRIGAN in cell elongation and cellulose synthesis in Arabidopsis thaliana. Plant Cell Physiol, 2001(3): p. 251.
14. Lane DR, et al. Temperature-sensitive alleles of RSW2 link the KORRIGAN Endo-1,4-beta-Glucanase to Cellulose Synthesis and Cytokinesis in Arabidopsis. Plant Physiol. 2001;126(1):278–88.
15. Junko T et al. KORRIGAN1 and its Aspen Homolog PttCel9A1 decrease cellulose crystallinity in Arabidopsis stems. Plant Cell Physiol, 2009(6): pp. 1099–115.
16. Maloney VJ, Mansfield SD. Characterization and varied expression of a membrane-bound endo-β-1,4-glucanase in hybrid poplar. Plant Biotechnol J, 2010.
17. Waldron KW, Parker M, Smith AC. Plant cell walls and food quality. Compr Rev Food Sci Food Saf. 2003;2(4):128–46.
18. Goulao LF, Oliveira CM. Cell wall modifications during fruit ripening: when a fruit is not the fruit. Trends Food Sci Technol. 2008;19(1):4–25.
19. Feng G et al. Genomic and transcriptomic analyses of Citrus sinensis varieties provide insights into Valencia orange fruit mastication trait formation. Hortic Res, 2021. 8.
20. Dong T, et al. Effect of pre-harvest application of calcium and boron on dietary fibre, hydrolases and ultrastructure in 'Cara Cara' navel orange (Citrus sinensis L. Osbeck) fruit. Sci Hort. 2009;121(3):272–7.
21. Lei Y, et al. Comparison of cell wall metabolism in the pulp of three cultivars of 'Nanfeng'tangerine differing in mastication trait. J Sci Food Agric. 2012;92(3):496–502.
22. Urbanowicz BR, et al. Structural organization and a standardized nomenclature for plant endo-1,4-beta-glucanases (cellulases) of glycosyl hydrolase family 9. Plant Physiol. 2007;144(4):1693–6.
23. Xie G, et al. Global identification of multiple OsGH9 family members and their involvement in cellulose crystallinity modification in rice. PLoS ONE. 2013;8(1):e50171.
24. Brummell DA, et al. Cell wall metabolism during maturation, ripening and senescence of peach fruit. J Exp Bot. 2004;55(405):2029–39.
25. Lei Y, et al. Physicochemical and molecular analysis of cell wall metabolism between two navel oranges (Citrus sinensis) with different mastication traits. J Sci Food Agric. 2010;90(9):1479–84.
26. Wu L-M, et al. Transcriptome analysis unravels metabolic and molecular pathways related to fruit sac granulation in a late-ripening navel orange (Citrus sinensis Osbeck). Plants. 2020;9(1):95.
27. Llop-Tous I, et al. Characterization of two divergent endo-β-1, 4-glucanase cDNA clones highly expressed in the nonclimacteric strawberry fruit. Plant Physiol. 1999;119(4):1415–22.
28. Rose JK, Bennett AB. Cooperative disassembly of the cellulose–xyloglucan network of plant cell walls: parallels between cell expansion and fruit ripening. Trends Plant Sci. 1999;4(5):176–83.
29. Fischer RL, Bennett AB. Role of cell wall hydrolases in fruit ripening. Annu Rev Plant Physiol Plant Mol Biol. 1991;42(1):675–703.
30. Mejía-Mendoza MA, et al. Identification in silico and expression analysis of a β-1-4-endoglucanase and β-galactosidase genes related to ripening in guava fruit. J Genetic Eng Biotechnol. 2022;20:1–11.
31. Xu Q, et al. The draft genome of sweet orange (Citrus sinensis). Nat Genet. 2013;45(1):59–66.
32. El-Gebali S, et al. The pfam protein families database in 2019. Nucleic Acids Res. 2019;47(D1):D427–32.
33. Letunic I, Bork P. 20 years of the SMART protein domain annotation resource. Nucleic Acids Res. 2018;46(D1):D493–6.
34. Lu S, et al. CDD/SPARCLE: the conserved domain database in 2020. Nucleic Acids Res. 2020;48(D1):D265–8.
35. Gasteiger E, et al. Protein identification and analysis tools on the ExPASy server. Springer; 2005.
36. Chou KC, Shen HB. Cell-PLoc: a package of web servers for predicting subcellular localization of proteins in various organisms. Nat Protoc. 2008;3(2):153–62.
37. Xie J et al. Tree visualization by one table (tvBOT): a web application for visualizing, modifying and annotating phylogenetic trees. Nucleic Acids Res, 2023: p. gkad359.
38. Chen C, et al. TBtools: an integrative toolkit developed for interactive analyses of big biological data. Mol Plant. 2020;13(8):1194–202.
39. Lescot M, et al. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Res. 2002;30(1):325–7.
40. Bailey TL, et al. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 2009;37(suppl2):W202–8.
41. Marchler-Bauer A, Bryant SH. CD-Search: protein domain annotations on the fly. Nucleic Acids Res. 2004;32(suppl2):W327–31.
42. Feng G, et al. High-spatiotemporal-resolution transcriptomes provide insights into fruit development and ripening in Citrus sinensis. Plant Biotechnol J. 2021;19(7):1337–53.
43. Feng G, Wu J, Yi H. Global tissue-specific transcriptome analysis of Citrus sinensis fruit across six developmental stages. Sci Data. 2019;6(1):153.
44. Mafra V, et al. Reference genes for accurate transcript normalization in citrus genotypes under different experimental conditions. PLoS ONE. 2012;7(2):e31263.
45. Wu J, et al. Selection of reliable reference genes for gene expression studies using quantitative real-time PCR in navel orange fruit development and pummelo floral organs. Sci Hort. 2014;176:180–8.

Publisher's note