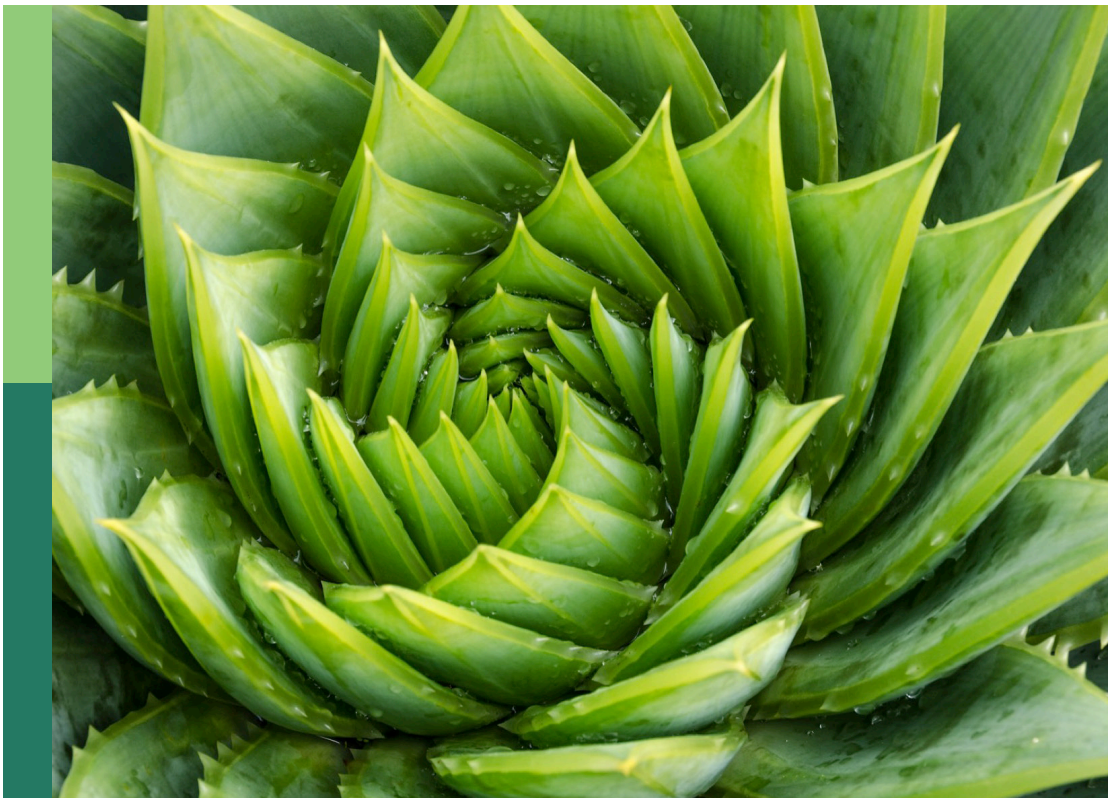
frontiers | Research Topics

Transcriptomics of fruit growth, development and ripening

Edited by
Neftali Ochoa-Alejo, Maria Carmen Gomez-Jimenez and Octavio Martínez

Published in
Frontiers in Plant Science



ISSN 1664-8714
ISBN 978-2-8325-4787-8
DOI 10.3389/978-2-8325-4787-8

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Transcriptomics of fruit growth, development and ripening

Topic editors

Neftali Ochoa-Alejo — Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Mexico

Maria Carmen Gomez-Jimenez — University of Extremadura, Spain

Octavio Martínez — Centro de Investigación y Estudios Avanzados del IPN (CINVESTAV), Mexico

Citation

Ochoa-Alejo, N., Gomez-Jimenez, M. C., Martínez, O., eds. (2024). *Transcriptomics of fruit growth, development and ripening*. Lausanne: Frontiers Media SA.

doi: 10.3389/978-2-8325-4787-8

Table of contents

frontiers | Frontiers in Plant Science

TYPE Editorial
PUBLISHED 05 April 2024
DOI 10.3389/fpls.2024.1399376

OPEN ACCESS

EDITED AND REVIEWED BY
Neelima Roy Sinha,
University of California, Davis, United States

*CORRESPONDENCE
Neftali Ochoa-Alejo
neftali.ochoa@cinvestav.mx

RECEIVED 11 March 2024
ACCEPTED 27 March 2024
PUBLISHED 05 April 2024

CITATION
Ochoa-Alejo N, Gómez-Jiménez MC and Martínez O (2024) Editorial: Transcriptomics of fruit growth, development and ripening. *Front. Plant Sci.* 15:1399376. doi: 10.3389/fpls.2024.1399376

Editorial: Transcriptomics of fruit growth, development and ripening

Neftali Ochoa-Alejo[1]*, Maria Carmen Gómez-Jiménez[2] and Octavio Martínez[3]

[1]Departamento de Ingeniería Genética, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Unidad Irapuato, Irapuato, Guanajuato, Mexico, [2]Laboratory of Plant Physiology, Universidad de Extremadura, Badajoz, Spain, [3]Unidad de Genómica Avanzada, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Irapuato, Guanajuato, Mexico

KEYWORDS

carbamoyltransferases, carotenoids, core genes, flavor, fruit, histone acetyltransferases, metabolomics, regulatory network

Editorial on the Research Topic
Transcriptomics of fruit growth, development and ripening

Introduction

Fruits are organs that hold seeds in plant species. From a botanical point of view, fruit is defined as a mature ovary (Roth, 1977), or as a structure developing from the gynoecium of one flower as the result of pollination or parthenocarpy (Bobrov and Romanov, 2019), or, additionally, as the flower in the state of seed germination (Knoll, 1939). An enormous variety of fruits with different forms, sizes, textures, colors, flavors, and aromas exists in nature (Klee, 2010; Rodríguez et al., 2012; Stournaras et al., 2012; Wu et al., 2018; Kapoor et al., 2022). Fruits are classified as fleshy or dry. Fleshy fruits are distributed in nature primarily by animals, whereas dry fruits may be dispersed by animals, wind, or water (Carey et al., 2019). Dry fruits are classified as dehiscent when they release seeds into environment (the seeds are discarded before or after consuming), or indehiscent those that release seeds in protected fruit wall propagules. Fleshy fruits are classified as climacteric (bananas, tomatoes, apricot, pears, mangoes, apricots, peaches, apples, papayas, guava, nectarines, blueberry, plum, passion fruit, cantaloupe, and avocados) or non-climacteric (grapefruit and lemon, berries such as raspberry, strawberry, cherry, grapes, pineapple, melon, watermelon, and pomegranate). In climacteric fruits, a burst of ethylene biosynthesis and an increase in respiration is observed at the onset of ripening. On the other hand, non-climacteric fruits lack the autocatalytic ethylene burst. Fruits are plant organs of nutritional value for animals and humans since they are sources of food, fiber, vitamins, minerals, carbohydrates, organic acids, amino acids, proteins, polyphenols (flavonoids and stilbenes), sterols, fatty acids, lipids, and pigments with antioxidant properties, among others (McKee and Latner, 2000; Zamora et al., 2001; Avila-Sosa et al., 2019; Wu et al., 2019; Golovinskaia and Wang, 2021; Alvi et al., 2022; Kasampalis

et al., 2022; Yun et al., 2022; Bures et al., 2023; Cavalcante de Oliveira et al., 2023). Fruits are formed from single ovaries and may or may not involve inclusion of accessory floral tissues like the floral receptacle. After pollination, flowers undergo complex processes involving cell division, cell enlargement, and cell differentiation mediated by the expression of hundreds or even thousands of genes under a fine, harmonic, and sequentially regulated program to promote growth, development, ripening and, finally, senescence (Gillaspy et al., 1993; Karlova et al., 2014). Fruit initiation, growth, development, ripening, and senescence are influenced by genetic, epigenetic, hormonal, and environmental factors (Seymour et al., 2008, 2013; Chen et al., 2022). A powerful approach to study those genes expressed during the growth, development, ripening, and senescence is transcriptomics through RNA-Seq analysis (Wang et al., 2009; Tarazona et al., 2011; Trapnell et al., 2013; Li D, et al., 2022). The set of all RNA molecules transcribed in an organ or tissue at a particular point of time under a given set of environmental conditions constitute the transcriptome (Velculescu et al., 1997; Martínez-López et al., 2014). *Arabidopsis thaliana* and tomato (*Solanum lycopersicum*) have been used as model plants to investigate dry and fleshy fruit biology, respectively (Seymour et al., 2013; Gómez et al., 2014). This Research Topic focuses on transcriptomic research in different plant species revealing changes in gene expression and key regulatory gene networks involved in fruit growth, development, ripening, and senescence.

Fruit transcriptomics

Transcriptional changes in fleshy fruits during growth, development, and ripening were previously reviewed (Karlova et al., 2014); transcriptomic analysis of the dry fruit (silique) of the model plant *Arabidopsis thaliana* was published by Mizzotti et al. (2018). In general, fruit formation, development, ripening, and senescence involve metabolic changes regulated by hormones and environmental factors; these changes include variations in metabolites, color, flavor, and aroma, and the softening process in the case of fleshy fruits. All these changes are usually regulated at the transcriptional level by diverse transcription factors (TFs). Transcriptional studies on fruit growth, development, ripening, and senescence have been published by different authors for several tropical plant species, including mango (*Mangifera indica* L.) (Pandit et al., 2010), pineapple (*Ananas comosus*) (Koia et al., 2012), melon (*Cucumis melo* L.) (Saladié et al., 2015), watermelon [*Citrullus lanatus* (Thunb.) Matsun. & Nakai)] (Zhu et al., 2017), litchi (*Litchi chinensis* Sonn.) (Liu et al., 2017), citrus (*Citrus sinensis*) (Feng et al., 2019), and goji berry (*Lycium barbarum*) (Zhao et al., 2020), or on temperate species such as almond (*Prunus dulcis*) (Guo et al., 2021), apple (*Malus domestica*) (Xu et al., 2020; Li M, et al., 2022), peach (*Prunus persica*) (Li M, et al., 2022), pear (*Pyrus* spp.) (Xie et al., 2013; Zhang et al., 2016), blueberry (*Vaccinia* spp.) (Gao et al., 2021; Yang L, et al., 2021), grape berry (*Vitis vinifera*) (Minio et al., 2019), and raspberry (*Rubus idaeus*) and strawberry (*Fragaria vesca*) (Zhou et al., 2023), have been reported. Transcriptomics of fruits used as vegetables, such as cucumber (*Cucumber sativus*), and chili pepper (*Capsicum* spp.), have been also published (Ando et al., 2012; Martínez et al., 2021, respectively). As an example of transcriptomic analysis, Galla et al. (2009) described a computational annotation of differentially expressed genes in fruits of olive (*Olea europea* L. cv. Leccino), a non-climacteric species, at three developmental stages (initial fruit set, completed pit hardening, and veraison) using subtractive hybridization libraries; 1,132 clones were sequenced, 60% of which presented similarity with known proteins, and were annotated by Gene Ontology (GO). Bioinformatic analysis revealed a significantly different distribution of the annotated GO category. The olive fruit-specific transcriptome dataset was used to query all known KEGG (Kyoto Encyclopedia of Genes and Genomes) metabolic pathways for characterizing and positioning retrieved EST (Expressed Sequence Tags) records, finding a predominance of KEGS maps associated to carbohydrate (enzymes involved in starch and sucrose metabolism, glycolysis and gluconeogenesis), fatty acid (fatty acid biosynthesis and lipid degradation), and secondary metabolism (phenylpropanoids, terpenoids, flavonoids, alkaloids, and caffeine biosynthesis, and limonene and pinene degradation). Moreover, genes involved in amino acid biosynthesis and metabolism were also differentially expressed. Genes related to hormone biosynthesis and action were also found to be differentially expressed. Auxin responsive transcription factor genes, such as *ARF1* and *ARF7*, were up- and down-regulated, respectively. Biosynthesis of abscisic acid (ABA) was stimulated as well as the expression of the ABA-biosynthesis related enzyme genes *ABA2* (*ABA DEFICIENT 2*), and *AAO3* (*ABSCISIC ALDEHYDE OXIDASE*). Genes related to the biosynthesis and action of indoleacetic acid (IAA), ABA, gibberellic acid (GA), ethylene, cytokinins, jasmonate, and salicylic acid were differently regulated according to the type of hormone. Genes encoding transcription factors were mainly down-regulated throughout fruit development, while some others were related to protein modification and degradation. In a further study (Parra et al., 2013), a comparative transcriptional profiling analysis of pericarp and the abscission zone (AZ) tissues in olive ripe-fruit from cv. Picual was conducted, and it was found that 4,391 genes were differentially expressed (DEG); in general, AZ tissue exhibited higher response to external stimuli than did ripe fruit, with higher expression of auxin-signaling genes, lignin catabolic and biosynthetic pathway, and of biosynthetic pathways of aromatic amino acid, isoprenoids, protein amino acid dephosphorylation, photosynthesis, and amino acid transport. Ripe fruit showed an enrichment in transcripts involved in ATP synthesis coupled proton transport, glycolysis, and cell-wall organization. Additionally, approximately 150 transcripts encoding putative TFs of diverse families were identified (37 fruit TFs and 113 AZ TFs); the most abundant TFs in ripe fruit were MADS-box proteins (TAGL2, AGL9, AG1), homeobox domain proteins, zinc finger proteins (ZF), basic helix-loop-helix proteins (bHLH), and basic leucine zipper proteins (bZIP). Among the 37 TF genes, 25 were exclusively expressed in fruit [6 *ZF*, 5 homeobox proteins, 5 *bHLH* domain class, 3 *bZIP*, 1 *MADS*-box (*AG1*), 1 *MYB* (*MYBA22*), 1 *NAC*, 1 *Aux/IAA* (*IAA1*), 1 *CAMTA*, and 1 *C2H21*. In the AZ tissue *ZF*, *bHLH*, and *bZIP* genes were highly expressed, but at different proportions,

and one member of the *E2F* family and 9 *WRKY* TFs genes were exclusively expressed.

In some cases, transcriptional studies have been focused on just early or late stages of fruit development and ripening (Ando et al., 2012; Gómez et al., 2014; Liu et al., 2017), or on a specific tissue (Parra et al., 2013; Tafolla-Arellano et al., 2017; Luo et al., 2018; Zhang et al., 2021; Zhou et al., 2022), or even on cell transcriptomics (Martin et al., 2016; Shinozaki et al., 2018). Moreover, transcriptomics of hormones-related gene expression (Zhu et al., 2011; Huang et al., 2014; Van de Poel et al., 2014; Tang et al., 2015; Briegas et al., 2020; Kou et al., 2021; Qiao et al., 2021; Camarero et al., 2023), or the effects of environmental/stress factors on fruit gene expression (Li et al., 2019; Cramer et al., 2020; Waite et al., 2023), biosynthesis/metabolism genes expressed during fruit development (Yu et al., 2021; Zhang et al., 2021; Diao et al., 2023), transcriptomic changes due to the evolution/domestication processes (Martínez et al., 2021; Borredá et al., 2022; Gramzou et al., 2022), expression of transcription factors genes and gene networks (Ye et al., 2015; Villa-Rivera et al., 2022), transcriptomics of fruit quality (Yang H, et al., 2021; Lei et al., 2022), fruit shape (Tsaballa et al., 2011; Shi et al., 2023) and size-related genes (Huang et al., 2023; Liu et al., 2023), or postharvest transcriptomic changes (Wang et al., 2018; Romero et al., 2022), have been documented as an approach to reveal the mechanisms and factors involved in the growth, development, ripening, and senescence of fruits.

Articles and insights

This Research Topic is composed of nine articles, and among them, two deal with different aspects of fruit transcriptomics, including growth, development, and ripening (Gaete-Eastman et al.; Rajewski et al.), and in one case the transcriptomic analysis was combined with metabolomics to study fruit flavor (Lu et al.). Basically, Rajewski et al. looked for differentially expressed genes in pericarp tissues of dry fruits from *Nicotiana obtusifolia* and *Solanum pimpinellifolium*, two wild species, and of the fleshy fruits from *Solanum lycopersicum* and *Cucumis melo*, two climacteric species, and, interestingly, they found core genes (121) during fruit development and ripening when *Arabidopsis thaliana* fruits (dry) were also included in the analysis; on the other hand, in the comparative gene expression profiles between the wild tomato (*S. pimpinellifolium*) and the domesticated species (*S. lycopersicum*), 1,472 genes showed divergent expression patterns and there were Gene Ontology enrichments for plant-type cell wall organization and lipid biosynthetic processes. Furthermore, expression analysis of ethylene, pigment, and flavor biosynthesis-related genes exhibited statistically significant differences between cultivated and wild tomato. Additionally, fruit size-, firmness-, and lignification-related transcription factors differed in expression between wild and domesticated tomato. This study revealed insights on the effects of domestication on gene expression profiles and on the evolutionary process in dry and fleshy fruits. In another article, Gaete-Eastman et al. reported an RNA-Seq transcriptomic analysis across different developmental stages and ripening of the non-climacteric fruits of *Fragaria chiloensis*, a Chilean strawberry species, and the main findings refer to the differential expression of ABA biosynthesis-related genes involved in softening, color, and aroma production, which are regulated by transcription factors such as FcMYB1. In an integrated metabolomic and transcriptomic study of jujube (*Ziziphus jujuba*) fruits during development and maturation, Lu et al., described differentially expressed transcription factor genes highly correlated with sugars and organic acids accumulation, important compounds involved in the fruit flavor.

Two articles were focused on the expression analysis of specific gene families during fruit development and ripening (Cai et al.; Dhar et al.). A genome-wide analysis of histone acetyltransferase (HAT) and histone deacetylase (HDAC) gene families and their expression in chili pepper (*Capsicum annuum* L.) fruits during development and ripening revealed a total of 30 HAT and 15 HDAC, which were differentially expressed and may be involved in the regulation of fruit development- and ripening-related phytohormone metabolism and signaling through changes in chromatin acetylation/deacetylation activities (Cai et al.). In the other case, Dhar et al. reported a genome and transcriptome-wide analysis of carbamoyltransferase genes (play roles by regulating the urea cycle, *de novo* pyrimidine biosynthesis, and arginine biosynthesis in prokaryotes and eucaryotes) in major fleshy fruits (30) from an evolutionary point of view, and they found 393 carbamoyltransferase genes conserved in the plant kingdom, indicating a fundamental biological relevance.

Additionally, the overexpression of the carotenoid biosynthesis-related structural gene *PSY1* as an approach to increase the carotenoid content in the skin and flesh of apple (*Malus domestica*) fruits was investigated, and an increase in carotenoid content (β-carotene being the most accumulated) was observed (Ampomah-Dwamena et al.).

A postharvest molecular study on the climacteric apple (*Malus domestica* cv. Golden Delicious) fruit quality after wax coating treatment was reported here, and the main findings were the inhibition of the expression of ethylene biosynthesis, chlorophyl degradation, and carotenoid biosynthesis-related genes (Si et al.). In one article, key genes associated with seed germination dormancy as affected by cold stratification in *Fritillaria taipaiensis* P.Y.Li (a traditional Chinese medicinal plant) were analyzed and they found that stratification at 4 °C induced an up-regulation of genes involved in gibberellic acid and auxin biosynthesis (Yang et al.). Finally, a comparative transcriptomic analysis between plants bearing prickles and non-prickled plants revealed the possible developmental mechanism of prickle formation in the important forest species *Zanthoxylum bungeanum* (Su et al.).

In conclusion, this Research Topic included not only interesting articles on transcriptomics covering different aspects of growth, development, and ripening of fruits from different plant species, but involving even transcriptomic changes occurring during postharvest conditions and germination of seeds, which certainly will be of motivation for those researchers working on this important biological process.

Author contributions

NO-A: Conceptualization, Funding acquisition, Investigation, Visualization, Writing – original draft, Writing – review & editing. MG-J: Writing – review & editing. OM: Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The author (s) thank Consejo Nacional de Humanidades, Ciencias y Tecnologías (Conahcyt, Mexico) for the financial support through the project FC1570.

Acknowledgments

We would like to express our gratitude to Frontiers in Plant Science for giving us the opportunity to be Guest Editors for the Research Topic "*Transcriptomics of fruit growth, development and ripening*". Our plenty gratitude also goes to all authors, contributors and experts who have played a significant role in making this Research Topic a valuable and engaging resource.

Conflict of interest

The authors declare that this Editorial was elaborated in the absence of any commercial or financial relationships that could be constructed as a potential conflict of interest.

Publisher's note

References

Alvi, T., Khan, M. K. I., Maan, A. I., and Razzaq, Z. U. (2022). Date fruit as a promising source of functional carbohydrates and bioactive compounds: A review on its nutraceutical potential. *J. Food Biochem.* 46, e14325. doi: 10.1111/jfbc.14325

Ando, K., Carr, K. M., and Grumet, R. (2012). Transcriptome analyses of early cucumber fruit growth identifies distinct gene modules associated with phases of development. *BMC Genomics* 13, 518. doi: 10.1186/1471-2164-13-518

Avila-Sosa, R., Montero-Rodríguez, A. F., Aguilar-Alonso, P., Vera-López, O., Lazcano-Hernández, M., Morales-Medina, J. C., et al. (2019). Antioxidant properties of Amazonian fruits: A mini review of in *vivo* and in *vitro* studies. *Oxid. Med. Cell. Longev.* 2019, 8204129. doi: 10.1155/2019/8204129

Bobrov, A. V. F. C., and Romanov, M. S. (2019). Morphogenesis of fruits and types of fruit of angiosperms. *Bot. Lett.* 166, 366–399. doi: 10.1080/23818107.2019.1663448

Borredá, C., Perez-Roma, E., Talon, M., and Terol, J. (2022). Comparative transcriptomics of wild and commercial Citrus during early ripening reveals how domestication shaped fruit gene expression. *BMC Plant Biol.* 22, 123. doi: 10.1186/s12870-022-03509-9

Briegas, B., Corbacho, J., Parra-Lobato, M. C., Paredes, M. A., Labrador, J., Gallardo, M., et al. (2020). Transcriptome and hormone analyses revealed insights into hormonal and vesicle trafficking regulation among *Olea europaea* fruit tissues in late development. *Int. J. Mol. Sci.* 21, 4819. doi: 10.3390/ijms21144819

Bures, M. S., Bandic, L. M., and Vlahovicek-Kahlina, K. (2023). Determination of bioactive components in mandarin fruits: A review. *Crit. Rev. Anal. Chem.* 53, 1489–1514. doi: 10.1080/10408347.2022.2035209

Camarero, M. C., Briegas, B., Labrador, J., Gallardo, M., and Gomez-Jimenez, M. C. (2023). Characterization of transcriptome dynamics during early fruit development in olive (*Olea europaea* L.). *Int. J. Mol. Sci.* 24, 961. doi: 10.3390/ijms24020961

Carey, S., Mendler, K., and Hall, J. C. (2019). How to build a fruit: Transcriptomics of a novel fruit type in the Brassicaceae. *PloS One* 14, e0209535. doi: 10.1371/journal.pone.0209535

Cavalcante de Oliveira, A., Moreira Mar, J., Frota Corrêa, R., Sanches, E. A., Campelo, P. H., da Silva Ramos, A., et al. (2023). *Pouteria* spp. fruits: Health benefits of bioactive compounds and their potential for the food industry. *Food Res. Int.* 173, 113310. doi: 10.1016/j.foodres.2023.113310

Chen, K., Huang, G., Li, Y., Zhang, X., Lei, Y., Li, Y., et al. (2022). Illumina MiSeq sequencing reveals correlations among fruit ingredients, environmental factors, and AMF communities in three *Lycium barbarum* producing regions of China. *Microbiol. Spectr.* 10, e02293-21. doi: 10.1128/spectrum.02293-21

Cramer, G. R., Cochetel, N., Ghan, R., Destrac-Irvine, A., and Delrot, S. (2020). A sense of place: Transcriptomics identifies environmental signatures in Cabernet Sauvignon berry skins in the late stages of ripening. *BMC Plant Biol.* 20, 41. doi: 10.1186/s12870-020-2251-7

Diao, Q., Tian, S., Cao, Y., Yao, D., Fan, H., and Zhang, Y. (2023). Transcriptome analysis reveals association of carotenoid metabolism pathway with fruit color in melon. *Sci. Rep.* 13, 5004. doi: 10.1038/s41598-023-31432-y

Feng, G., Wu, J., and Yi, H. (2019). Global tissue-specific transcriptome analysis of Citrus sinensis fruit across six developmental stages. *Sci. Rep.* 6, 153. doi: 10.1038/s41597-019-0162-y

Galla, G., Barcaccia, G., Ramina, A., Collani, S., Alagna, F., Baldoni, L., et al. (2009). Computational annotation of genes differentially expressed along olive fruit development. *BMC Plant Biol.* 9, 128. doi: 10.1186/1471-2229-9-128

Gao, X., Wang, L., Zhang, H., Zhu, B., Lv, G., and Xiao, J. (2021). Transcriptome analysis and identification of genes associated with floral transition and fruit development in rabbiteye blueberry (*Vaccinium ashei*). *PloS One* 16, e0259119. doi: 10.1371/journal.pone.0259119

Gillaspy, G., Ben-David, H., and Gruissem, W. (1993). Fruits: A developmental perspective. *Plant Cell* 5, 1439–1451. doi: 10.2307/3869794

Golovinskaia, O., and Wang, C.-K. (2021). Review of functional and pharmacological activities of berries. *Molecules* 26, 3904. doi: 10.3390/molecules26133904

Gómez, M. D., Vera-Sirera, F., and Pérez-Amador, M. A. (2014). Molecular programme of senescence in dry and fleshy fruits. *J. Exp. Bot.* 65, 4515–4526. doi: 10.1093/jxb/eru093

Gramzou, L., Klupsch, K., Fernández-Pozo, N., Höltzer, M., Marz, M., Rensing, S. A., et al. (2022). Comparative transcriptomics identifies candidate genes involved in the evolutionary transition from dehiscent to indehiscent fruits in *Lepidium* (Brasicaceae). *BMC Plant Biol.* 22, 340. doi: 10.1186/s12870-022-03631-8

Guo, C., Wei, Y., Yang, B., Ayup, M., Li, N., Liao, K., et al. (2021). Developmental transcriptome profiling uncovered carbon signaling genes associated with almond fruit drop. *Sci. Rep.* 11, 3401. doi: 10.1038/s41598-020-69395-z

Huang, G., Li, T., Li, X., Tan, D., Jiang, Z., Wei, Y., et al. (2014). Comparative transcriptome analysis of climacteric fruits of Chines pear (*Pyrus ussuriensis*) reveals new insights into fruit ripening. *PloS One* 9, e107562. doi: 10.1371/journal.pone.0107562

Huang, M., Zhu, X., Bai, H., Wang, C., Gou, N., Zhang, Y., et al. (2023). Comparative anatomical and transcriptomics reveal the larger cell size as a major contributor to larger fruit size in apricot. *Int. J. Mol. Sci.* 24, 8748. doi: 10.3390/ijms24108748

Kapoor, L., Simkin, A. J., Doss, G. P., and Siva, R. (2022). Fruit ripening: dynamics and integrated analysis of carotenoids and anthocyanins. *BMC Plant Biol.* 22, 27. doi: 10.1186/s12870-021-03411-w

Karlova, R., Chapman, N., David, K., Angenent, G. C., Seymour, G. B., and de Maagd, R. A. (2014). Transcriptional control of fleshy fruit development and ripening. *J. Exp. Bot.* 65, 4527–4541. doi: 10.1093/jxb/eru316

Kasampalis, D. S., Tsouvaltzis, P., Ntouros, K., Gertsis, A., Gitas, I., Moshou, D., et al. (2022). Nutritional composition changes in bell pepper as affected by the ripening stage of fruits at harvest or postharvest storage and assessed non-destructively. *J. Sci. Food Agric.* 102, 445–454. doi: 10.1002/jsfa.11375

Klee, H. J. (2010). Improving the flavor of fresh fruits: genomics, biochemistry, and biotechnology. *New Phytol.* 187, 44–56. doi: 10.1111/j.1469-8137.2010.03281.x

Knoll, F. (1939). "Über den begriff "Frucht".". *Der Biol.* 8, 154–160.

Koia, J. H., Moyle, R. L., and Botella, J. R. (2012). Microarray analysis of gene expression profiles in ripening pineapple fruits. *BMC Plant Biol.* 12, 240. doi: 10.1186/1471-2229-12-240

Kou, X., Feng, Y., Yuan, S., Zao, X., Wu, C., Wang, C., et al. (2021). Different regulatory mechanisms of plant hormones in the ripening of climacteric and non-climacteric fruits: A review. *Plant Mol. Biol.* 107, 477–497. doi: 10.1007/s11103-021-01199-9

Lei, D., Lin, Y., Chen, Q., Zhao, B., Tang, H., Zhang, Y., et al. (2022). Transcriptomic analysis and the effect of maturity stage on fruit quality reveal the importance of the L-galactose pathway in the ascorbate biosynthesis of hardy kiwifruit (*Actinidia arguta*). *Int. J. Mol. Sci.* 23, 6816. doi: 10.3390/ijms23126816

Li, M., Galimba, K., Xiao, Y., Dardick, C., Mount, S. M., Callahan, A., et al. (2022). Comparative transcriptomic analysis of apple and peach fruits: Insights into fruit type specification. *Plant J.* 109, 1614–1629. doi: 10.1111/tpj.15633

Li, T., Wang, Y.-H., Liu, J.-X., Xu, Z.-S., and Xion, A.-S. (2019). Advances in genomic, transcriptomic, proteomic, and metabolomic approaches to study biotic stress in fruit crops. *Crit. Rev. Biotechnol.* 39, 680–692. doi: 10.1080/07388551.2019.1608153

Li, D., Zand, M. S., Dye, T. D., Goniewicz, M. L., Rahman, I., and Xie, Z. (2022). An evaluation of RNA-seq differential analysis methods. *PloS One* 17, e0264246. doi: 10.1371/journal.pone.0264246

Liu, W., Chen, M., Bai, L., Zhuang, Z., Fan, C., Jiang, N., et al. (2017). Comprehensive transcriptomics and proteomics analyses of pollinated and parthenocarpic litchi (*Litchi chinensis* Sonn.) fruits during early development. *Sci. Rep.* 7, 5401. doi: 10.1038/s41598-017-05724-z

Liu, H., Zhang, X., Li, J., Zhang, G., Fang, H., and Li, Y. (2023). Transcriptome analysis reveals the mechanisms of different fruit appearance between apricot (*Armeniaca vulgaris* Lam.) and its seedlings. *Mol. Biol. Rep.* 50, 7995–8003. doi: 10.1007/s11033-023-08631-x

Luo, X., Cao, D., Li, H., Zhao, D., Xue, H., Niu, J., et al. (2018). Complementary iTRAQ-based proteomic and RNA sequencing-based transcriptomic analyses reveal a complex network regulating pomegranate (*Punica granatum* L.) fruit colour. *Sci. Rep.* 8, 12362. doi: 10.1038/s41598-018-30088-3

Martin, L. B. B., Nicolas, P., Matas, A. J., Shinozaki, Y., Catalá, C., and Rose, J. K. (2016). Laser microdissection of tomato fruit cell and tissue types for transcriptome profiling. *Nat. Protoc.* 11, 2376–2388. doi: 10.1038/nprot.2016.146

Martínez, O., Arce-Rodríguez, M. L., Hernández-Godínez, F., Escoto-Sandoval, C., Cervantes-Hernández, F., Hayano-Kanashiro, C., et al. (2021). Transcriptome analyses throughout chili pepper fruits development reveal novel insights into the domestication process. *Plants* 10, 585. doi: 10.3390/plants10030585

Martínez-López, L. A., Ochoa-Alejo, N., and Martínez, O. (2014). Dynamics of the chili pepper transcriptome during fruit development. *BMC Genomics* 15, 143. doi: 10.1186/1471-2164-15-143

McKee, L. H., and Latner, T. A. (2000). Underutilized sources of dietary fiber: A review. *Plant Foods Hum. Nutr.* 55, 285–304. doi: 10.1023/A:1008144310986

Minio, A., Massonnet, M., Figueroa-Balderas, R., Vondras, A. M., Blanco-Ulate, B., and Cantu, D. (2019). Iso-seq allows genome-independent transcriptome profiling of grape berry development. *G3 Genes Genomes Genet.* 9, 755–767. doi: 10.1534/g3.118.201008

Mizzotti, C., Rotasperti, L., Moretto, M., Tadini, L., Resentini, F., Galliani, B., et al. (2018). Time-course transcriptome analysis of *Arabidopsis thaliana* siliques discloses genes essential for fruit development and maturation. *Plant Physiol.* 178, 1249–1268. doi: 10.1104/pp.18.00727

Pandit, S. S., Kulkarni, R. S., Giri, A. P., Kölner, T. G., Degenhardt, J., Gershenzon, J., et al. (2010). Expression profiling of various genes during development and ripening of mango. *Plant Physiol. Biochem.* 48, 426–433. doi: 10.1016/j.plaphy.2010.02.012

Parra, R., Paredes, M. A., Sanchez-Calle, I. M., and Gomez-Jimenez, M. C. (2013). Comparative transcriptional profiling analysis of olive ripe-fruit pericarp and abscission zone tissues shows expression differences and distinct patterns of transcriptional regulation. *BMC Genomics* 14, 866. doi: 10.1186/1471-2164-14-866

Qiao, H., Zhang, H., Wang, Z., and Shen, Y. (2021). Fig fruit ripening is regulated by the interaction between ethylene and abscisic acid. *J. Integr. Plant Biol.* 63, 553–569. doi: 10.1111/jipb.13065

Rodríguez, A., Alquézar, B., and Peña, L. (2012). Fruit aromas in mature fleshy fruits as signals of readiness for predation and seed dispersal. *New Phytol.* 197, 36–48. doi: 10.1111/j.1469-8137.2012.04382.x

Romero, I., Escribano, M. I., Merodio, C., and Sanchez-Ballesta, M. T. (2022). Postharvest high CO_2 treatments on the quality of soft fruit berries: An integrated transcriptomic, proteomic, and metabolomic approach. *J. Agric. Food Chem.* 70, 8593–8597. doi: 10.1021/acs.jafc.2c01305

Roth, I. (1977). "Fruits of angiosperms," in *Encyclopedia of plant anatomy* (Gebrüder Borntraeger, Berlin, Stuttgart), 675.

Saladié, M., Cañizares, J., Phillips, M. A., Rodriguez-Concepcion, M., Larrigaudière, C., Gibon, Y., et al. (2015). Comparative transcriptional profiling analysis of developing melon (Cucumis melo L.) fruit from climacteric and non-climacteric varieties. *BMC Genomics* 16, 440. doi: 10.1186/s12864-015-1649-3

Seymour, G. B., Østergaard, L., Chapman, N. H., Knapp, S., and Martin, C. (2013). Fruit development and ripening. *Annu. Rev. Plant Biol.* 64, 219–241. doi: 10.1146/annurev-arplant-050312-120057

Seymour, G., Poole, M., Manning, K., and King, G. J. (2008). Genetics and epigenetics of fruit development and ripening. *Curr. Opin. Plant Biol.* 11, 58–63. doi: 10.1016/j.pbi.2007.09.003

Shi, S., Li, D., Li, S., Wang, Y., Tang, X., Liu, Y., et al. (2023). Comparative transcriptomic analysis of early development in eggplant (*Solanum melongena* L.) and functional characterization of SmOvate5. *Plant Cell Rep.* 42, 321–336. doi: 10.1007/s00299-022-02959-7

Shinozaki, Y., Nicolas, P., Fernandez-Pozo, N., Ma, Q., Evanich, D. J., Shi, Y., et al. (2018). High-resolution spatiotemporal transcriptome mapping of tomato fruit development and ripening. *Nat. Commun.* 9, 363. doi: 10.1038/s41467-017-02782-9

Stournaras, K. E., Lo, E., Böhning-Gaese, K., Cazzeta, E., Dehling, D. M., Schleuning, M., et al. (2012). How colorful are fruits? Limited color diversity in fleshy fruits on local and global scales. *New Phytol.* 198, 617–629. doi: 10.1111/nph.12157

Tafolla-Arellano, J. C., Zheng, Y., Sun, H., Jiao, C., Ruiz-May, C., Hernandez-Oñate, A., et al. (2017). Transcriptome analysis of mango (*Mangifera indica* L.) fruit epidermal peel to identify putative cuticle-associated genes. *Sci. Rep.* 7, 46163. doi: 10.1038/srep46163

Tang, N., Deng, W., Hu, G., Hu, N., and Li, Z. (2015). Transcriptome profiling reveals the regulatory mechanism underlying pollination dependent and parthenocarpic fruit set mainly mediated by auxin and gibberellin. *PloS One* 10, e0125355. doi: 10.1371/journal.pone.0125355

Tarazona, S., García-Alcalde, F., Dopazo, J., Ferrer, A., and Conesa, A. (2011). Differential expression in RNA-seq: A matter of depth. *Genome Res.* 21, 2213–2223. doi: 10.1101/gr.124321.111

Trapnell, C., Hendrickson, D. G., Sauvageau, M., Goff, L., Rinn, J. L., and Pachter, L. (2013). Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat. Biotechnol.* 31, 46–53. doi: 10.1038/nbt.2450

Tsaballa, A., Pasentsis, K., Darzentas, N., and Tsaftaris, A. S. (2011). Multiple evidence for the role of an *Ovate*-like gene in determining fruit shape in pepper. *BMC Plant Biol.* 11, 46. doi: 10.1186/1471-2229-11-46

Van de Poel, B., Bulens, I., Hertog, M. L. A. T. M., Nicolai, B. M., and Geeraerd, A. H. (2014). A transcriptomics-based kinetic model for ethylene biosynthesis in tomato (*Solanum lycopersicum*) fruit: Development, validation and exploration of novel regulatory mechanisms. *New Phytol.* 202, 952–963. doi: 10.1111/nph.12685

Velculescu, V. E., Zhang, L., Zhou, W., Vogelstein, J., Basrai, M. A., Bassett, D. E., et al. (1997). Characterization of the yeast transcriptome. *Cell* 88, 243–251. doi: 10.1016/S0092-8674(00)81845-0

Villa-Rivera, M. G., Martínez, O., and Ochoa-Alejo, N. (2022). Putative transcription factor genes associated with regulation of carotenoid biosynthesis in chili pepper fruits revealed by RNA-Seq coexpression analysis. *Int. J. Mol. Sci.* 23, 11774. doi: 10.3390/ijms231911774

Waite, J. M., Kelly, E. A., Zhang, H., Hargarten, H. L., Waliullah, S., Altman, N. S., et al. (2023). Transcriptomic approach to uncover dynamic events in the development of mid-season sunburn in apple fruit. *G3* 13, jkad120. doi: 10.1093/g3journal/jkad120

Wang, L., Chen, Y., Wang, S., Xue, H., Su, Y., Yang, J., et al. (2018). Identification of candidate genes involved in the sugar metabolism and accumulation during pear fruit post-harvest ripening of 'Red Clapps's Favorite (*Pyrus communis* L.) by transcriptomic analysis. *Hereditas* 155, 11. doi: 10.1186/s41065-017-0046-0

Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* 10, 57–63. doi: 10.1038/nrg2484

Wu, Y., Xu, J., He, Y., Shi, M., Han, X., Li, W., et al. (2019). Metabolic profiling of pitaya (*Hylocereus polyrhizus*) during fruit development and maturation. *Molecules* 24, 1114. doi: 10.3390/molecules24061114

Wu, S., Zhang, B., Keyhaninejad, N., Rodríguez, G. R., Kim, H. J., Chakrabarti, M., et al. (2018). A common genetic mechanism underlies morphological diversity in fruits and other plant organs. *Nat. Commun.* 9, 4734. doi: 10.1038/s41467-018-07216-8

Xie, M., Huang, Y., Zhang, Y., Wang, X., Yang, H., Yu, O., et al. (2013). Transcriptome profiling of fruit development and maturation in Chinese white pear (*Pyrus bretschneideri* Rehd). *BMC Genomics* 14, 823. doi: 10.1186/1471-2164-14-823

Xu, J., Yan, J., Li, W., Wang, Q., Wang, C., Guo, J., et al. (2020). Integrative analyses of widely targeted metabolic profiling and transcriptome data reveals molecular insight into metabolomic variations during apple (*Malus domestica*) fruit development and ripening. *Int. J. Mol. Sci.* 21, 4797. doi: 10.3390/ijms21134797

Yang, L., Liu, L., Wang, Z., Zong, Y., Yu, L., Li, Y., et al. (2021). Comparative anatomical and transcriptomic insights into *Vaccinium corymbosum* flower bud and fruit throughout development. *BMC Plant Biol.* 21, 289. doi: 10.1186/s12870-021-03067-6

Yang, H., Tian, C., Ji, S., Ni, F., Fan, X., Yang, Y., et al. (2021). Integrative analyses of metabolome and transcriptome reveals metabolomic variations and candidate genes involved in sweet cherry (*Prunus avium* L.) fruit quality during development and ripening. *PloS One* 16, e0260004. doi: 10.1371/journal.pone.0260004

Ye, J., Hu, T., Yang, C., Li, H., Yang, M., Ijaz, R., et al. (2015). Transcriptome profiling of tomato fruit development reveals transcription factors associated with ascorbic acid, carotenoid and flavonoid biosynthesis. *PloS One* 10, e0130885. doi: 10.1371/journal.pone.0130885

Yu, X., Ali, M. M., Li, B., Fang, T., and Chen, F. (2021). Transcriptome data-based identification of candidate genes involved in metabolism and accumulation of soluble sugars during fruit development in 'Huangguan' plum. *J. Food Biochem.* 45, e13878. doi: 10.1111/jfbc.13878

Yun, D., Yan, Y., and Liu, J. (2022). Isolation, structure and biological activity of polysaccharides from fruits of *Lycium ruthenicum* Murr.: A review. *Carbohyd. Polym.* 291, 119618. doi: 10.1016/j.carbpol.2022.119618

Zamora, R., Alaiz, M., and Hidalgo, F. J. (2001). Influence of cultivar and fruit ripening on olive (*Olea europaea*) fruit protein content, composition, and antioxidative activity. *J. Agric. Food Chem.* 49, 4267–4270. doi: 10.1021/jf0104634

Zhang, M.-Y., Xue, C., Xu, L., Sun, H., Qin, M.-F., Zhang, S., et al. (2016). Distinct transcriptome profiles reveal gene expression patterns during fruit development and maturation in five main cultivated species of pear (*Pyrus* L.). *Sci. Rep.* 6, 28130. doi: 10.1038/srep28130

Zhang, A., Zheng, J., Chen, X., Shi, X., Wang, H., and Fu, Q. (2021). Comprehensive analysis of transcriptome and metabolome reveals the flavonoid metabolic pathway is associated with fruit peel coloration of melon. *Molecules* 26, 2830. doi: 10.3390/molecules26092830

Zhao, J., Li, H., Yin, Y., An, W., Qin, X., Wang, Y., et al. (2020). Transcriptomic and metabolomic analyses of *Lycium ruthenicum* and *Lycium barbarum* fruits during ripening. *Sci. Rep.* 10, 4354. doi: 10.1038/s41598-020-61064-5

Zhou, J., Li, M., Li, Y., Xiao, Y., Luo, X., Gao, S., et al. (2023). Comparison of red raspberry and wild strawberry fruits reveals mechanisms of fruit type specification. *Plant Physiol.* 193, 1016–1035. doi: 10.1093/plphys/kiad409

Zhou, X., Liu, S., Yang, Y., Liu, J., and Zhuang, Y. (2022). Integrated metabolome and transcriptome analysis reveals a regulatory network of fruit peel pigmentation in eggplant (*Solanum melongena* L.). *Int. J. Mol. Sci.* 23, 13475. doi: 10.3390/ijms232113475

Zhu, H., Dardick, C. D., Beers, E. P., Callanhan, A. M., Xia, R., and Yuan, R. (2011). Transcriptomics of shading-induced and NAA-induced abscission in apple (*Malus domestica*) reveals a shared pathway involving reduced photosynthesis, alterations in carbohydrate transport and signaling and hormone cross-talk. *BMC Plant Biol.* 11, 138. doi: 10.1186/1471-2229-11-138

Zhu, Q., Gao, P., Liu, S., Zhu, Z., Amanullah, S., Davis, A. R., et al. (2017). Comparative transcriptome analysis of two contrasting watermelon genotypes during fruit development and ripening. *BMC Genomics* 18, 3. doi: 10.1186/s12864-016-3442-3

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 15 August 2022
DOI 10.3389/fpls.2022.978013

OPEN ACCESS

EDITED BY
Neftali Ochoa-Alejo,
Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Mexico

REVIEWED BY
Xiaosan Huang,
Nanjing Agricultural University, China
Haiyan Shi,
Agricultural University of Hebei, China
Junfeng Guan,
Hebei Academy of Agriculture and Forestry Sciences (HAAFS), China

*CORRESPONDENCE
Hui Yuan
huiyuan@syau.edu.cn
Aide Wang
awang@syau.edu.cn

SPECIALTY SECTION
This article was submitted to Plant Development and EvoDevo, a section of the journal Frontiers in Plant Science

RECEIVED 25 June 2022
ACCEPTED 28 July 2022
PUBLISHED 15 August 2022

CITATION
Si Y, Lv T, Li H, Liu J, Sun J, Mu Z, Qiao J, Bu H, Yuan H and Wang A (2022) The molecular mechanism on suppression of climacteric fruit ripening with postharvest wax coating treatment *via* transcriptome. *Front. Plant Sci.* 13:978013. doi: 10.3389/fpls.2022.978013

The molecular mechanism on suppression of climacteric fruit ripening with postharvest wax coating treatment *via* transcriptome

Yajing Si[1], Tianxing Lv[2], Hongjian Li[2], Jiaojiao Liu[1], Jiamao Sun[1], Zhaohui Mu[1], Junling Qiao[1], Haidong Bu[3], Hui Yuan[1]* and Aide Wang[1]*

[1]Key Laboratory of Fruit Postharvest Biology (Liaoning Province), Key Laboratory of Protected Horticulture (Ministry of Education), National and Local Joint Engineering Research Center of Northern Horticultural Facilities Design and Application Technology (Liaoning), College of Horticulture, Shenyang Agricultural University, Shenyang, China, [2]Liaoning Institute of Pomology, Xiongyue, China, [3]Mudanjiang Branch, Heilongjiang Academy of Agricultural Sciences, Mudanjiang, China

Wax coating is an important means to maintain fruit quality and extend fruit shelf life, especially for climacteric fruits, such as apples (*Malus domestica*). Here, we found that wax coating could inhibit ethylene production, chlorophyll degradation, and carotenoid synthesis, but the molecular mechanism remains unclear. The regulatory mechanism of wax coating on apple fruit ripening was determined by subjecting wax-treated apple fruits to transcriptome analysis. RNA-seq revealed that 1,137 and 1,398 genes were upregulated and downregulated, respectively. These differentially expressed genes (DEGs) were shown to be related to plant hormones, such as ethylene, auxin, abscisic acid, and gibberellin, as well as genes involved in chlorophyll degradation and carotenoid biosynthesis. Moreover, we found that some genes related to the wax synthesis process also showed differential expression after the wax coating treatment. Among the DEGs obtained from RNA-seq analysis, 15 were validated by quantitative RT-PCR, confirming the results from RNA-seq analysis. RNA-seq and qRT-PCR of pear (*Pyrus ussuriensis*) showed similar changes after wax treatment. Our data suggest that wax coating treatment inhibits fruit ripening through ethylene synthesis and signal transduction, chlorophyll metabolism, and carotenoid synthesis pathways and that waxing inhibits endogenous wax production. These results provide new insights into the inhibition of fruit ripening by wax coating.

KEYWORDS

***Malus domestica*, *Pyrus ussuriensis*, fruit wax coating, fruit ripening, RNA-seq**

Introduction

Fruit ripening is a very important physiological process during fruit development and is often accompanied by physiological changes in fruit color, firmness, odor, and sugar-acid ratio (Lelièvre et al., 1997). These changes are affected by environmental factors, gene regulation, and biotic and abiotic stresses (Klee and Giovannoni, 2011; Larrainzar et al., 2014). Some fruits produce a large amount of ethylene during ripening and are called climacteric fruits, while others produce little ethylene and are called non-climacteric fruits (Adams-Phillips et al., 2004; Farcuh et al., 2018). The climacteric fruit undergoes a climacteric change and a peak of ethylene release during storage, leading to faster fruit decay and shortening the storage period (Kou and Wu, 2018).

The storage period of fruits is associated with the level of ethylene released during the ripening process (Osorio et al., 2013; Farneti et al., 2022). Reducing ethylene production delays fruit ripening and prolongs shelf life. 1-aminocyclopropane-1-carboxylic acid synthase (*ACS*) and 1-aminocyclopropane-1-carboxylic acid oxidase (*ACO*) genes are the key genes involved in ethylene synthesis. In the pear (*Pyrus ussuriensis*) fruit, ethylene treatment can induce four *ACS* genes and three *ACO* genes (Yuan et al., 2020). *MdACS1*-silenced transgenic apple (*Malus domestica*) fruits had little ethylene production during ripening and significantly prolonged storage (Dandekar et al., 2004). In addition to ethylene, many hormones are also involved in fruit ripening. The Gretchen GH3 (GH3) protein controls auxin homeostasis, and silencing of *AcGH3.1* delays fruit softening during postharvest kiwifruit (*Actinidia chinensis*) (Gan et al., 2019). Auxin application before the commercial harvest stage induces apple (*Malus domestica*) fruit ripening (Yue et al., 2020). Abscisic acid (ABA) is a ripening regulator in non-climacteric fruits and is associated with ripening in climacteric fruits. In the apricot (*Prunus armeniaca* L.) fruit, exogenous ABA treatment can significantly induce the expression of the ABA receptor gene *PaPYL9*, and the overexpression of *PaPYL9* in tomato fruit can promote early ripening (Jia et al., 2021). It has been reported that endogenous gibberellin content decreases during tomato (*Solanum lycopersicum*) fruit ripening, and the application of gibberellin (GA) biosynthesis inhibitors during the green ripening stage of fruit can accelerate fruit ripening (Li H. et al., 2019). Some researchers also applied GA combined with cytokinin to pineapple (*Ananas comosus*) fruits after flowering induction treatment and found that they could improve fruit quality and delay fruit ripening (Suwandi et al., 2016).

Chlorophyll I (ChI) degradation and carotenoid synthesis are common natural phenomena during fruit ripening. Ethylene can activate E3 ubiquitin ligase *MdPUB24* in the apple fruit to interact with and ubiquitinate *MdBEL7*, resulting in chlorophyll degradation during fruit storage (Wei et al., 2021). Exogenous ethylene treatment can strongly promote the expression of carotenoid biosynthesis genes in tomatoes (Cruz et al., 2018). Therefore, chlorophyll degradation and carotenoid synthesis are thought to be regulated by maturation.

Wax is a layer of wax film formed on the surface of the fruit, and is closely related to the quality and storage of the fruit as a natural protective layer. Removing the epidermal wax of postharvest blueberry fruit accelerates water loss and fruit rot and shortens the storage period (Chu et al., 2018). In tomato fruit, the key gene for wax synthesis, *CER1-1*, catalyzes the biosynthesis of waxy alkanes, which can improve the shelf life and drought tolerance of the fruit (Wu et al., 2022). In zucchini (*Cucurbitaceae Trichosanthes* L.) fruits, the alkane biosynthesis pathway and epidermal morphology play an important role in the postharvest fruit quality (Carvajal et al., 2021).

Artificial fruit wax is a preservative used to improve fruit quality. It is mainly made of animal and vegetable wax as film-forming agents or emulsified with animal and vegetable wax. Postharvest wax coating is an effective method to avoid the loss of fruit quality. Waxing is convenient to operate and easy to popularize, which cannot only improve fruit quality but also prolong fruit storage period and shelf life. Postharvest waxing of pineapple (*Ananas comosus*) can delay fruit color change, effectively improve fruit quality, and inhibit fruit ripening (Li X. et al., 2018). Wax coating treatment of jujube fruit (*Ziziphus jujuba* Mill) and passion fruit (*Passiflora edulis* Sims) also produced similar results (Rashwan et al., 2020; Sang, 2020). Researchers have found that wax coating together with low-temperature storage better prolongs the shelf life of cucumber (*Cucumis sativus*) fruit (Li J. et al., 2018). In mango (*Mangifera indica*) fruit, waxing combined with nitric oxide (NO) can maintain fruit firmness at 13°C for 18 days (Vázquez-Celestino et al., 2016).

These studies suggest that postharvest wax coating treatment can inhibit the ripening of climacteric fruits; however, corresponding changes in the gene expression and their molecular mechanisms are poorly understood. Hence, we aimed to undertake an in-depth analysis of waxed and unwaxed apple and pear fruits using transcriptome sequencing technology to gain new insights into the molecular basis of the inhibition of climacteric fruit ripening by wax coating.

Materials and methods

Plant materials and treatments

Apple (*Malus domestica* cv. Golden Delicious [GD]) fruits were harvested from the Experimental Orchard of Liaoning Pomology Institute (Xiongyue, China). Mature, pest-free, and uniform-sized GD fruits were collected on commercial harvest day (145 days after flowering, DAFB) and immediately transported to the laboratory. The apples were washed with

clean water and divided into two groups. The first group was treated with liquid wax that was applied evenly at a concentration of 80% to the peel and stem of the apple several times. The fruits were then dried naturally to form an even surface coating of the wax. The second group of apples remained untreated and was used as the control. The fruits in the control and treated groups were kept at 25°C for 25 days and sampled every 5 days. On each sampling day, five apples were chosen randomly from each group as biological replicates. The peel and pulp tissues of the fruit were separated, snap-frozen in liquid nitrogen, and stored at −80°C until further use.

Pear (*Pyrus ussuriensis* cv. Nanguo) fruits were harvested from an orchard farm in Anshan, Liaoning, China. After harvesting the fruits on commercial harvest day (135 DAFB), wax was applied as described above. Similar to the apple fruit, the pear fruits of the control and treatment groups were stored at 25°C for 25 days and sampled every 5 days. Since the fruit softened and could not be sampled at 25 days, the samples were taken to 20 days. On each sampling day, five pears were randomly selected from each group as biological replicates.

The liquid wax selected in this study was morpholine fatty acid salt fruit wax produced by the Beijing Institute of Chemical Industry (Beijing, China). Its main chemical components are natural palm wax (approximately 10–20%), morpholine fatty acid salt (2.5–3%), and water (85–87%). The preliminary studies found a concentration of 80% to be the optimal wax concentration for fruit treatment. Fruit wax (80%) was prepared by mixing 80 mL of morpholine fatty acid salt fruit wax with 20 mL of distilled water.

Measurement and analysis of physiological data

Ethylene production rate

Ethylene measurements were taken every 5 days. The ethylene production rate of the fruits was measured as previously described (Yuan et al., 2017). Five fruits per sample were used for measurement and analysis.

Flesh firmness

Fruit flesh firmness measurements were taken every 5 d. It was measured with a portable pressure tester (FT-327; Facchini, Italy), and the firmness of apple and pear fruits was measured with 11 and 8 mm probes, respectively. Four pericarps (approximately 2 cm in diameter) were removed from the opposite side of each fruit. The probe was then inserted into the tissue of the cut surface to the depth of the probe graduation mark vertically and quickly. At this point, the value on the dial was read and recorded. Five fruits per sample were used for measurement and analysis.

Chlorophyll and carotenoid contents

A total of 80% acetone and 95% ethanol were used in a 1:1 ratio to prepare an extraction solution. Peeled samples stored at −80°C were ground into a powder. The powdered sample (0.4 g) was weighed into a 10 mL-centrifuge tube, 8 mL of the extraction solution was added, and the mixture was vortexed. Experiments were performed in triplicate. After leaching at 25°C for 24 h in the dark, the mixture was centrifuged at 8,000 × g or 15 min at 25°C, and the supernatant was collected. A spectrophotometer was used to detect the absorbance of the supernatant at 470, 649, and 665 nm (zero-adjustment was done using the extraction solution), and the following formulae were used to calculate the total chlorophyll content and carotenoid content:

$$\text{Chlorophyll a (Ca)} = 13.95A_{665} - 6.8A_{649}, \quad (1)$$

$$\text{Chlorophyll b (Cb)} = 24.96A_{649} - 7.32A_{665}, \quad (2)$$

$$\text{Totalchlorophyll(Ct)} = \text{Ca} + \text{Cb}, \quad (3)$$

$$\text{Carotenoids(Cx)} = \frac{1000\ A_{470} - 2.05\ \text{Ca} - 114.8\ \text{Cb}}{248}, \quad (4)$$

$$\text{Pigmentcontent(mg/g)} = \frac{\text{Ct(Cx)} \times \text{V} \times \text{N}}{\text{W} \times 1000}, \quad (5)$$

where C is the pigment concentration (mg/L), A is the absorbance, V is the extract volume (mL), N is the dilution factor, and W is the sample weight (g).

RNA extraction, library preparation, and RNA sequencing

RNA-seq analysis was performed on the fruit mixture of peel and pulp stored for 10 d. The first analysis was performed in 2020 and the other two in 2021, forming three biological replicates of apple fruit. Pear fruit for RNA-seq analysis was performed in 2020.

Total RNA was isolated and purified using TRIzol reagent (Invitrogen, Carlsbad, CA, United States) following the manufacturer's procedure. The RNA amount and purity of each sample was quantified using NanoDrop ND-1000 (NanoDrop, Wilmington, DE, United States). The RNA integrity was assessed by Bioanalyzer 2100 (Agilent, CA, United States) with RIN number > 7.0, and confirmed by electrophoresis with denaturing agarose gel. Poly (A) mRNA purified from 2 μg total RNA using Dynabeads Oligo attached magnetic beads. The mRNA is fragmented into small pieces using divalent magnesium ions under elevated

temperatures. Then the cleaved RNA fragments were reverse-transcribed to create the final cDNA library according to the protocol (KC-Digital^{TM} Stranded mRNA Library Prep Kit for Illumina®). At last, we performed the 2 × 150 bp paired-end sequencing (PE150) on an Illumina NovaSeq^{TM} 6000 (Wuhan Seqhealth Co., Ltd., China) following the vendor's recommended protocol.

Bioinformatics analysis

Raw sequencing data were transformed into valid reads after data processing. The HISAT2 software was used to map the reads to the reference apple genome[1] and pear genome.[2] StringTie was used to perform expression level for mRNAs by calculating FPKM in 2020, and FeatureCounts was utilized to calculate RPKMs in 2021. Genes differentially expressed between control and treated were identified using the edgeR package. The differentially expressed transcripts were selected based on log2 (fold change) > 1 or log2 (fold change) < -1 and with statistical significance (p-value < 0.05). The pathway analyses of GO and KEGG enrichment of DEGs were based on the Gene Ontology Database[3] and KEGG pathway,[4] respectively.

qRT-PCR analysis

Total RNA was extracted according to the method of previously described (Li et al., 2014). Total RNA (800 ng) was used to synthesize first-strand cDNA using the PrimeScript First Strand cDNA Synthesis kit (TAKARA, Japan). qRT-PCR was performed as previously described (Tan et al., 2013). Specific primers for each gene were designed using Primer3.[5] The primers used in this study are listed in **Supplementary Table 1**. The apple actin gene (*MdActin*) was used as the internal control. Experiments were performed in triplicate.

Statistical analysis

The data of qRT-PCR and physiological assays were analyzed using Microsoft excel, and the results are shown as mean ± SE. Statistical significance among the means were determined with the Student's t-test, and $^{*}p < 0.05$, $^{**}p < 0.01$. All figures were prepared by Origin2018.

1 http://the-apple-genome.html
2 ftp://ftp.ncbi.nlm.nih.gov.html
3 http://geneontology.org.html
4 https://www.kegg.html
5 http://primer3.html

Results

Postharvest wax coating treatment inhibits the ripening of apple fruit

After harvested on commercial day, the apple fruits were treated with wax and stored at 25°C for 25 d (**Figure 1A**). The untreated apple fruits showed a climacteric peak when stored for 20 d, the ethylene release reached 0.31 $\mu L \cdot g^{-1} \cdot h^{-1}$, and the waxed fruit was about 11% of that of the control (**Figure 1B**). The hardness of the fruit tends to decrease with ripening. During the storage period, the wax coating treatment better maintained the hardness of the fruit, and its decline was reduced (**Figure 1C**). Wax coating delayed the yellowing of apple fruit skin (**Figure 1A**). To demonstrate this phenomenon more clearly, we measured the chlorophyll and carotenoid contents of the fruit peel and found that the decrease in chlorophyll content and the increase in carotenoid content were suppressed in the wax-treated fruits (**Figures 1D,E**). In 2021, fruits harvested at the commercial harvest stage were also treated with liquid wax, and we observed the same effect of ripening inhibition in 2021 as in 2020 (**Supplementary Figure 1**).

RNA-seq and analysis of differentially expressed genes in apple fruit

The waxed fruits showed differences in several indicators of ripeness compared to the untreated ones by day 10. To characterize DEGs associated with ripening under control and wax treatments, RNA-seq analysis was performed. A total of 36.58–50.24 million valid reads were obtained from the six samples (Control-1, Control-2, Control-3, Waxed-1, Waxed-2, and Waxed-3). A statistical summary of the three biological replicates is shown in **Table 1**. Six samples were prepared and sequenced as six independent RNA libraries, and 49,628,898, 57,559,864, 78,224,472, 48,782,016, 76,571,408, and 64,401,010 raw reads were obtained, respectively. The unreliable raw read data, including low-quality bases, sequencing linkers, and non-carrier errors, were excluded for obtaining high-quality, clear, and valid read data. The percentages of Q20, Q30, and GC content were more than 99.9, 98.2 and 47%, respectively. Furthermore, we used Hisat software to align the valid reads to the reference genome and found that the value of mapped reads was 93.15–97.04%, and the value of uniquely mapped reads was 81.31–97.51%. In all samples, according to the region information of the reference genome, the percentage of sequenced sequences located in exonic regions was the highest (**Supplementary Figure 2A**).

A total of 3,744 and 6,777 DEGs were identified between the control and wax-treated groups in the

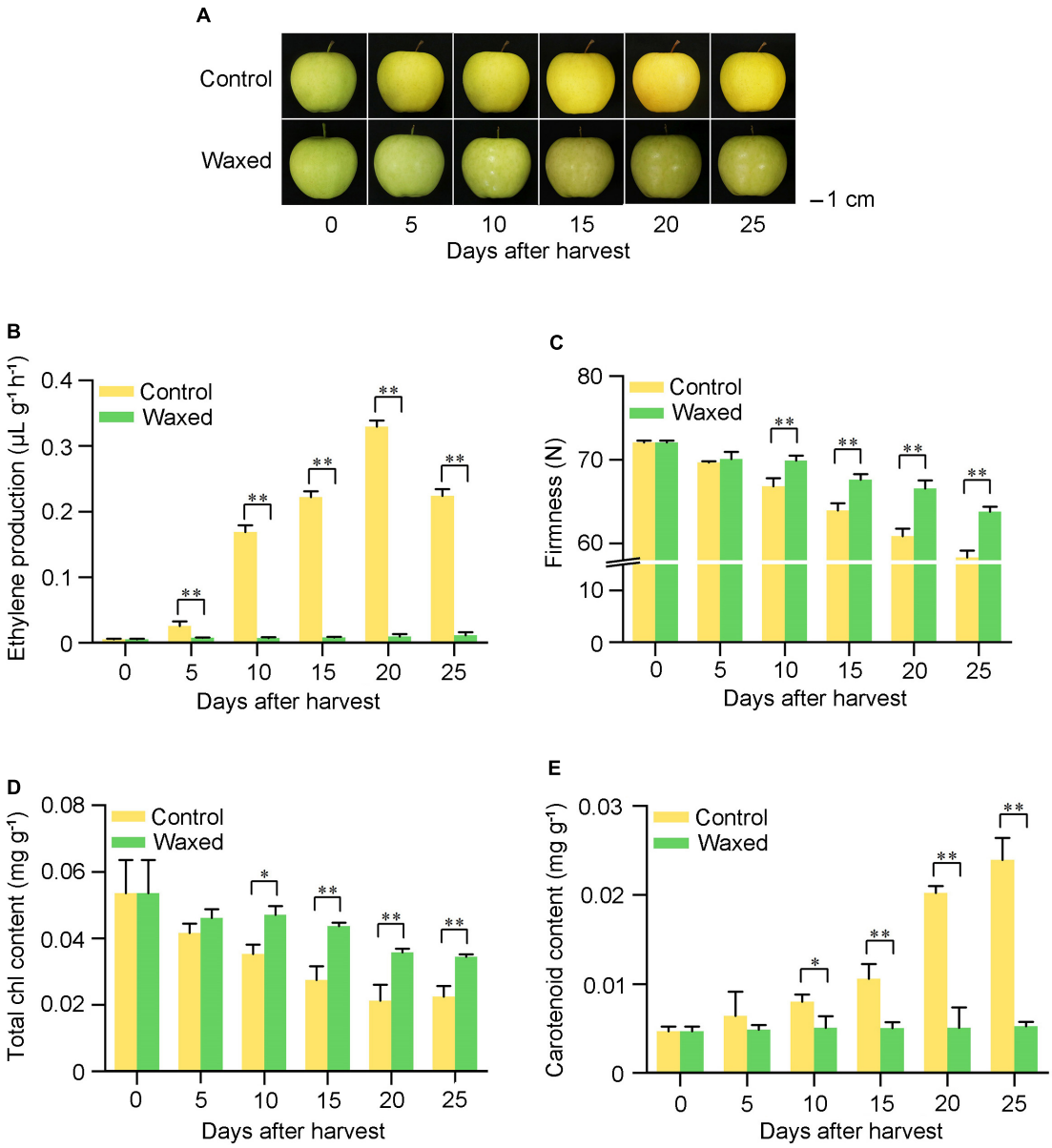


FIGURE 1
Postharvest wax coating treatment inhibits apple fruit ripening. Apple fruit were collected on the commercial harvest day (145 DAFB) in 2020 and storage at room temperature for 25 days **(A)**. After treatment, ethylene production **(B)**, fruit firmness **(C)**, total Chl content **(D)**, and carotenoid content **(E)** were measured. Control, apple fruit not receiving treatment; Waxed, fruit treated with morpholine fatty acid salt fruit wax. Scale bar = 1 cm. The x-axis represents days of storage at room temperature after harvest. Five biological replicates were analyzed for ethylene production, fruit firmness, and three biological replicates for total Chl content and carotenoid content. Values represent mean ± SE. Statistical significance was determined using a Student's *t*-test, and $*p < 0.05$, $**p < 0.01$.

RNA-seq analysis performed in 2020 and 2021, respectively (Supplementary Sheet 1). To identify specific genes that were associated with waxing, we identified the genes that were commonly upregulated and downregulated among the three biological replicates of the RNA-seq experiments performed in 2020–21 (Supplementary Sheet 2). The Venn diagrams showing these common gene numbers are presented in Supplementary Figures 2B,C.

Gene Ontology (Go) analysis was applied to analyze biological processes and functions enriched in common DEGs. The DEGs were classified into three GO categories: biological process, cellular component and molecular function. Among them, biological processes involve the most components, followed by molecular functions, and cellular components involve the least components. Among the up-regulated differential genes, nucleic acid binding, heterocyclic compound

TABLE 1 Summary of RNA-sequencing of apple fruit waxing treatment.

Sample	Control-1	Control-2	Control-3	Waxed-1	Waxed-2	Waxed-3
Raw reads	49,628,898	57,559,864	78,224,472	48,782,016	76,571,408	64,401,010
Valid reads	49,141,912	36,576,146	50,244,818	48,306,536	48,207,626	40,027,470
Q20 (%)	99.99	99.95	99.95	99.99	99.95	99.95
Q30 (%)	98.23	98.35	98.35	98.28	98.35	98.35
GC (%)	47	48	48	47	48	48
Total mapped reads (%)	46,139,503 (93.89%)	34,930,093 (95.50%)	47,999,689 (95.53%)	44,997,639 (93.15%)	46,644,607 (96.76%)	38,841,043 (97.04%)
Unique Mapped reads (%)	37,697,492 (81.71%)	34,014,756 (97.38%)	46,806,107 (97.51%)	36,586,784 (81.31%)	45,365,691 (97.26%)	37,787,652 (97.29%)
Multi Mapped reads (%)	8,442,011 (18.29%)	915,337 (2.62%)	1,193,582 (2.49%)	8,410,855 (18.69%)	1,278,916 (2.74%)	1,053,391 (2.71%)

Sample, sample name; Raw reads, the total number of reads of the offline data; Valid resds, the number of reads after removal of low-quality bases, sequencing adapters, and non-carrier errors; Q20, Percentage of bases with quality value ≥ 20; Q30, Percentage of bases with quality value ≥ 30; GC, The sum of base G and C accounts for the percentage of total base number; Total mapped read: the number of reads that can be aligned to the reference genome; Unique Mapped reads, the number of reads that can only be uniquely aligned to one position in the genome; Multi Mapped reads, the number of reads that can be aligned to multiple positions in the genome.

binding and organic cyclic compound binding were more abundant in molecular functions. However, among the down-regulated differential genes, lipid metabolic process, carboxylic acid metabolic process, oxoacid metabolic process and organic acid metabolic process were more abundant in biological processes (**Supplementary Figure 3**).

Analysis of plant hormone-related genes in apple fruit

Compared with the control, wax-coated apples had DEGs that showed the enrichment for hormone synthesis and signal transduction processes of auxin, abscisic acid, cytokinin, gibberellin, and ethylene (**Supplementary Sheet 3**).

The auxin influx carrier *AUX1* gene of the AUX1/LAX family was downregulated after fruit waxing. The auxin-responsive proteins IAA AUX/IAA showed an overall upregulated expression where four genes were found to be upregulated, and two were downregulated among auxin response factors *ARFs*. Moreover, the expression of *GH3.1* from the auxin-responsive *GH3* gene family was upregulated in wax-coated fruit, and some genes of the SAUR family proteins showed differential expression. This indicates that genes of the same family could have different functions in the context of fruit ripening (**Figure 2A**).

Transcriptome analysis showed that *NCED1 and NCED5* genes were downregulated after wax treatment. We found that some genes involved in ABA signaling were downregulated after treatment. These include the ABA receptor protein *PYL2*, two protein phosphatases *PP2C*, and an ABA response element-binding factor *ABF* (**Figure 2B**).

In the cytokinin signal transduction pathway, two CTK receptor *CRE* genes and three histidine-containing phosphotransfer protein *AHP* genes were differentially expressed following wax coating (**Figure 2C**).

The gibberellin receptor *GID1*, F-box protein *GID2*, and *DELLA* genes are components of the gibberellin signal transduction pathway. After waxing treatment, *GID2* and *DELLA* were upregulated; two of the three *GID1* genes were downregulated, while the other was upregulated (**Figure 2D**).

Ethylene is the most important hormone in the regulation of fruit ripening. In our study, 19 DEGs related to ethylene synthesis and signal transduction were identified, and a heatmap showing the differential expression of these genes is presented in **Figure 2E**. After wax coating, some rate-limiting enzyme genes associated with the ethylene biosynthesis pathway were identified, such as S-adenosylmethionine synthase (*metK*), 1-aminocyclopropane-1-carboxylate synthase (*ACS*), and 1-aminocyclopropane-1-carboxylate oxidase (*ACO*). This was consistent with the reduction in ethylene production after wax treatment. The ethylene receptor gene *ETR* was downregulated, and the expression of ethylene-responsive transcription factor *ERFs* showed different trends. This result indicates that wax treatment could affect the expression of ethylene-related genes and inhibit fruit ripening.

Analysis of chlorophyll and carotenoid-related genes in apple fruit

Chlorophyll and carotenoid-related differentially expressed genes are listed in **Supplementary Sheet 3**. After waxing, the chlorophyll synthesis-related genes magnesium-protoporphyrin IX monomethyl ester [oxidative] cyclase (*MPEC)*, protochlorophyllide reductase (*PORA)*, and divinyl chlorophyllide a 8-vinyl-reductase (*DVR)* were upregulated, and the decomposition genes probable chlorophyll(ide) b reductase (*NYC)* and chlorophyll(ide) b reductase (*NOL)* were

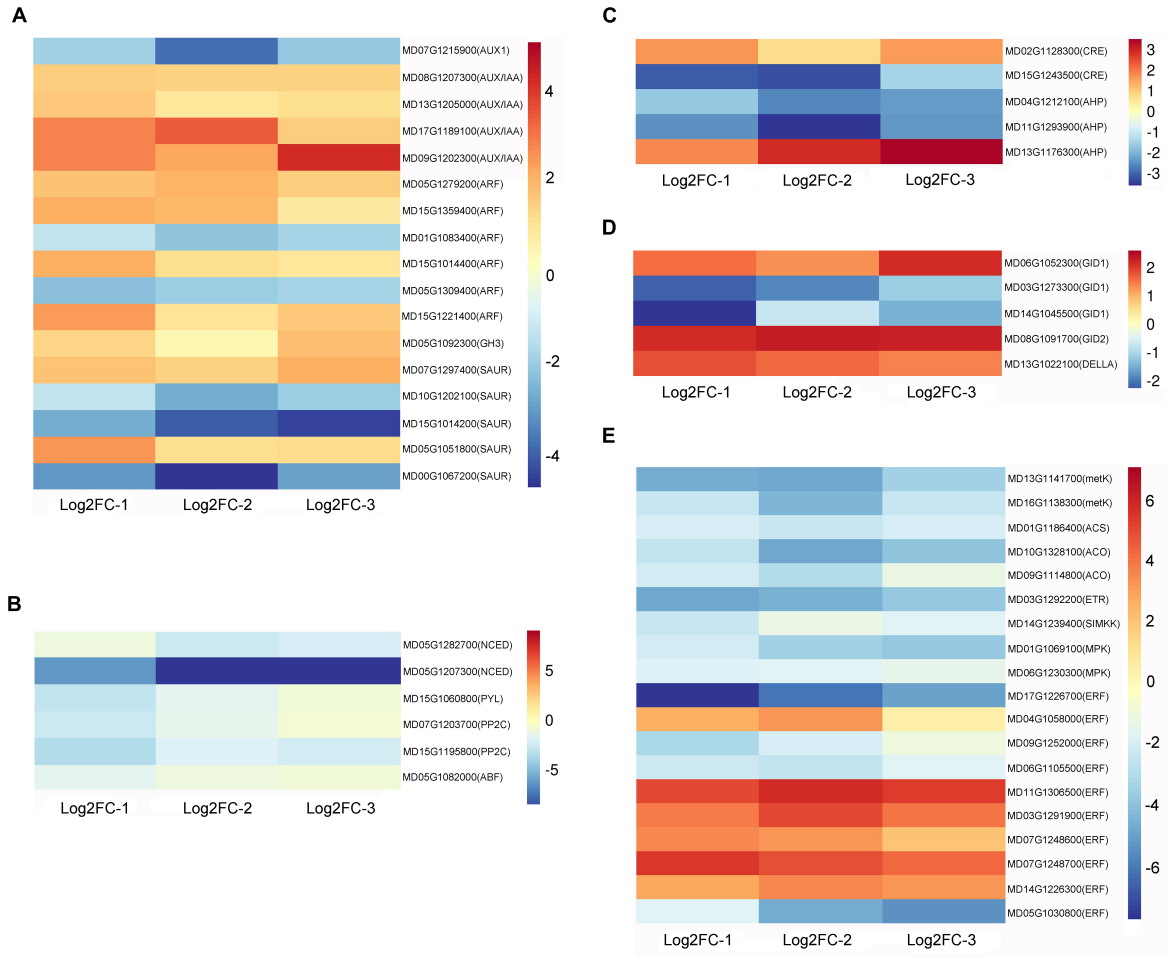


FIGURE 2
Heat maps of plant hormone-related DEGs after waxing treatment in apple fruit. Selected DEGs related to auxin **(A)**, abscisic acid **(B)**, cytokinin **(C)**, gibberellin **(D)**, and ethylene **(E)** from RNA-Seq date. Log2FC-1, Log2 (Fold Change) value of Waxed-1 and Control-1 in apple fruit RNA-Seq; Log2FC-2, Log2 (Fold Change) value of Waxed-2 and Control-2 in apple fruit RNA-Seq; Log2FC-3, Log2 (Fold Change) value of Waxed-3 and Control-3 in apple fruit RNA-Seq. The scale bar indicates upregulated (red) and downregulated (blue) DEGs.

downregulated (Figure 3A). This provides an explanation for the physiological phenomenon of maintenance of green color in wax-coated fruits. In the heat map, we can observe that, except for the violaxanthin de-epoxidase gene *VDE*, other differential genes were downregulated in wax-coated fruits (Figure 3B). These are structural genes encoding metabolic enzymes in the carotenoid pathway, including phytoene synthase (*PSY*), ζ-carotene isomerase (*Z-ISO*), zeta-carotene desaturase (*ZDS*), carotenoid isomerase (*Crt-ISO*), lycopene-β-cyclase (*LCY-β*), beta-carotene hydroxylase (*CHY-β*), carotenoid beta-ring hydroxylase (*LUT5*), zeaxanthin epoxidase (*ZEP*), capsanthin/capsorubin synthase (*CCS*), and 9-cis-epoxycarotenoid dioxygenase (*NCED*). They are responsible for the synthesis of many types of carotenoids, such as phytoene, zeta-carotene, lycopene, alpha-carotene, lutein, gamma-carotenoids, beta-carotene, zeaxanthin, and neoxanthin. This shows that the inhibition of fruit yellowing by wax coating is related to the reduction of carotenoid content and inhibition of related synthetic genes.

Analysis of fatty acid and wax-related genes in apple fruit

The expression trends of some genes in the fatty acid synthesis, elongation, degradation, and wax synthesis pathways were altered after wax coating (Supplementary Sheet 3). Based on the enrichment of genes in the KEGG database, genes involved in the fatty acid synthesis pathway, such as carboxylase *accA*, *accC*, transacylase *FabD*, synthase *FabF*, reductase *FabG*, *FabI*, dehydratase *FabZ*, thioesterase *FATA*, and long-chain acyl-CoA synthase *ACSL*, showed the highest enrichment among the total 22 downregulated genes. However, in the transcriptome data, we found that the

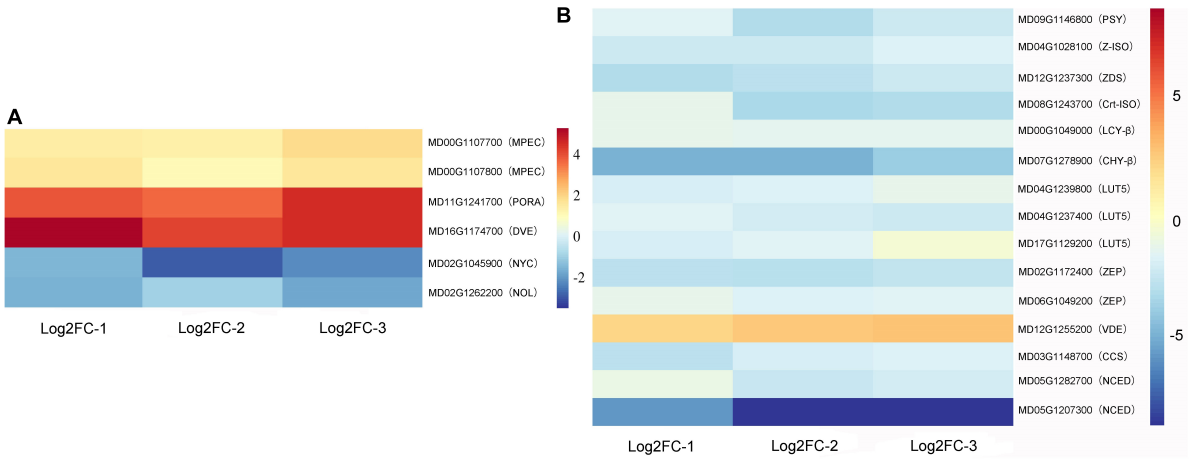


FIGURE 3
Heat maps of chlorophyll and carotenoid-related DEGs after waxing treatment in apple fruit. Selected DEGs related to chlorophyll **(A)** and carotenoid **(B)** from RNA-Seq date. Log2FC-1, Log2 (Fold Change) value of Waxed-1 and Control-1 in apple fruit RNA-Seq; Log2FC-2, Log2 (Fold Change) value of Waxed-2 and Control-2 in apple fruit RNA-Seq; Log2FC-3, Log2 (Fold Change) value of Waxed-3 and Control-3 in apple fruit RNA-Seq. The scale bar indicates upregulated (red) and downregulated (blue) DEGs.

stearoyl-[acyl-carrier-protein] 9-desaturase (*SAD*) gene was upregulated in wax-coated fruits. Interestingly, the octadecanoyl group during the synthesis of octadecanoic acid can be converted by the action of acyl-[acyl-carrier-protein] desaturase to octadecenoyl and then to octadecenoic acid. A heatmap of the related DEGs is shown in **Figure 4A**.

The extension of fatty acids is divided into two types: one occurs in the mitochondria, where two C atoms are added in each cycle of extension, the final number of C atoms is in the range of $4 < n < 16$, and fatty acid degradation will be carried out after the extension is completed. The other is in the endoplasmic reticulum, where the number of C atoms is 16 or more in the cycle extension, and the final product, i.e., long-chain acyl-CoA or long-chain fatty acid will enter the wax synthesis process. Moreover, 11 genes were downregulated during endoplasmic reticulum extension, including synthase *KCS*, reductase *KCR*, dehydratase *HACD*, and thioesterase *ACOT*. Nevertheless, *KCS6*, unlike other genes, is upregulated after waxing. Furthermore, six and four differentially expressed genes were involved in fatty acid degradation and wax biosynthesis, respectively. Among the former genes, *ADH*, *ALDH*, and *ACOX* were all upregulated, and among the latter, *FAR*, and *CER* were downregulated (**Figures 4B,C**).

Validation of apple fruit RNA-seq using qRT-PCR

To confirm the findings of RNA-seq, we randomly selected DEGs from the KEGG pathways as described above for qRT-PCR analysis. Fifteen DEGs were selected from different pathways such as ethylene, chlorophyll, carotenoid, and wax biosyntheses and fatty acid extension. The trend of up or downregulation of these selected genes in control and waxed apple fruits was consistent with the RNA-seq results, which confirmed the accuracy of RNA-seq (**Figure 5**).

Postharvest wax coating treatment inhibits pear fruit ripening and RNA-Seq analysis

To complement our research on apple fruit, we applied postharvest waxing to the pear fruit. Compared with the control fruit, the ethylene production of the wax-coated pear fruit was inhibited, and the firmness was better maintained (**Figures 6A–C**). Moreover, the yellowing of the epidermis of the treated fruit was inhibited, and the carotenoid and chlorophyll contents were consistent with phenotypic changes (**Figures 6D,E**).

Similarly, we performed RNA-seq analysis on wax-treated and -untreated pear fruit samples stored for 10 d. The two samples were named Pear-wax and Pear-con, respectively. The statistical summary are listed in **Supplementary Table 2**. In the transcriptome data, there were 3,247 DEGs, of which 1,349 genes were upregulated, and 1,898 genes were downregulated (**Supplementary Sheet 4**). We performed functional clustering and enrichment analysis for these DEGs using the KEGG database, and similar to our results obtained with the apple fruit, we identified enrichment for plant hormone synthesis and signal transduction, chlorophyll and carotenoid synthesis and degradation, and fatty acid with wax biosynthesis (**Supplementary Sheet 5**). Upon further analysis of the DEGs, we observed that wax coating resulted in the downregulation of genes involved in ethylene syntheses, such as *metK*, *ACS*,

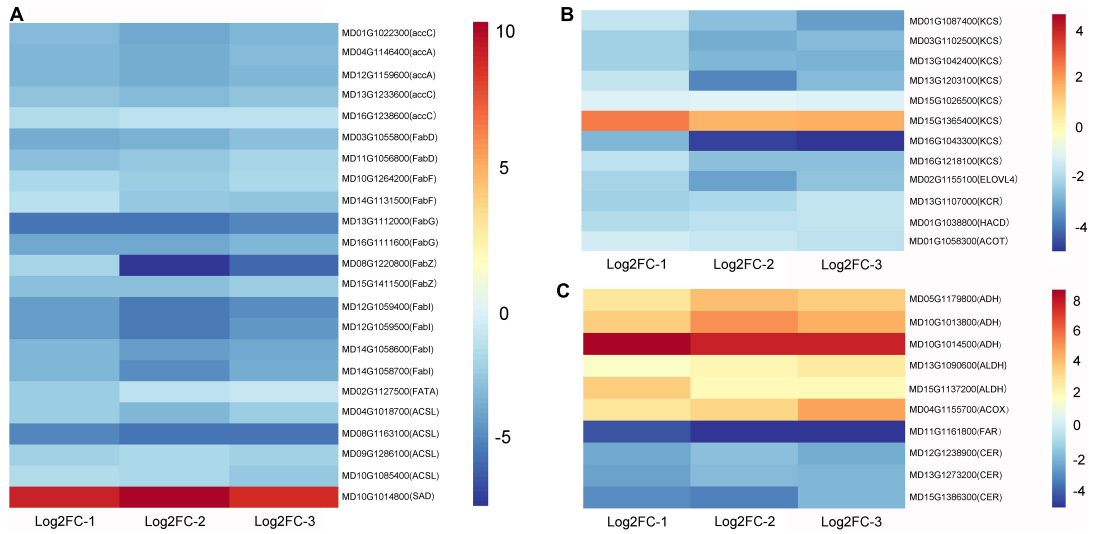


FIGURE 4
Heat maps of fatty acid and wax-related DEGs after waxing treatment in apple fruit. Selected DEGs related to fatty acid biosynthesis **(A)**, fatty acid elongation **(B)**, fatty acid degradation and wax biosynthesis **(C)** from RNA-Seq date. Log2FC-1, Log2 (Fold Change) value of Waxed-1 and Control-1 in apple fruit RNA-Seq; Log2FC-2, Log2 (Fold Change) value of Waxed-2 and Control-2 in apple fruit RNA-Seq; Log2FC-3, Log2 (Fold Change) value of Waxed-3 and Control-3 in apple fruit RNA-Seq. The scale bar indicates upregulated (red) and downregulated (blue) DEGs.

and *ACO*, whereas the *ERFs* did not show a clear trend. The genes associated with chlorophyll syntheses, such as *ChlH*, *MPEC*, *PORA*, and *CAO*, were upregulated. Except for *VDE*, the remaining DEGs in the carotenoid pathway were downregulated, which was consistent with our observations for the apple transcriptome. In the process of fatty acid synthesis, elongation, and degradation, related DEGs, such as *accC*, *FabD*, *FabF*, *KCS*, *ADH*, and *ALDH*, were downregulated after wax treatment. However, unlike the apple transcriptome, *FATA*, *ACSL*, and *KCR* genes were upregulated. Moreover, in the process of wax biosynthesis, the DEGs of pears were different from those of apples, which may be caused by the different effects of wax coating on wax conversion in the two fruits.

To confirm these observations from the RNA-seq analysis of pear fruit, we randomly selected six DEGs for qRT-PCR analysis (**Figure 7**). The test results show that the RNA-seq data is reliable.

In conclusion, postharvest wax treatment of pear fruit can delay fruit ripening, and the genes associated with ripening and fatty acid pathways showed a very similar expression profile to that observed for apple fruit.

Discussion

In this study, morpholine fatty acid salt fruit wax at a concentration of 80% was used for postharvest waxing of fruits, and RNA-seq technology was used to analyze DEGs in apple and pear fruits after waxing. The KEGG database was used for functional clustering and enrichment analysis of the DEGs. The highest enrichment for DEGs was observed for plant hormones, chlorophyll, carotenoids, and fatty acid pathways.

Effects of postharvest wax coating treatment on ripening of climacteric fruits

In previous studies, the use of edible beeswax coating on mango fruit was shown to control the ripeness of the fruit and maintain the color and hardness of the peel (Sousa et al., 2021). Moreover, fucoidan coating can also effectively prolong the shelf life of mango fruits (Xu and Wu, 2021). According to our study, the postharvest wax-coated apple and pear fruits showed a peak of ethylene release during storage, a slow rate of decrease in firmness, and the fruits did not turn yellow during the later storage period. Compared with the control group, the chlorophyll content of the treated peel was higher, and the carotenoid content was lower (**Figures 1**, **6**). This indicates that the fruit wax treatment used in this study could inhibit the postharvest ripening of climacteric fruits.

Postharvest wax treatment and plant hormones

There have been many studies on the relationship between plant hormones and fruit ripening, and ethylene can regulate the ripening of climacteric fruit (Pech et al., 2012). The application of 1-methylcyclopropene (1-MCP) to durian fruit

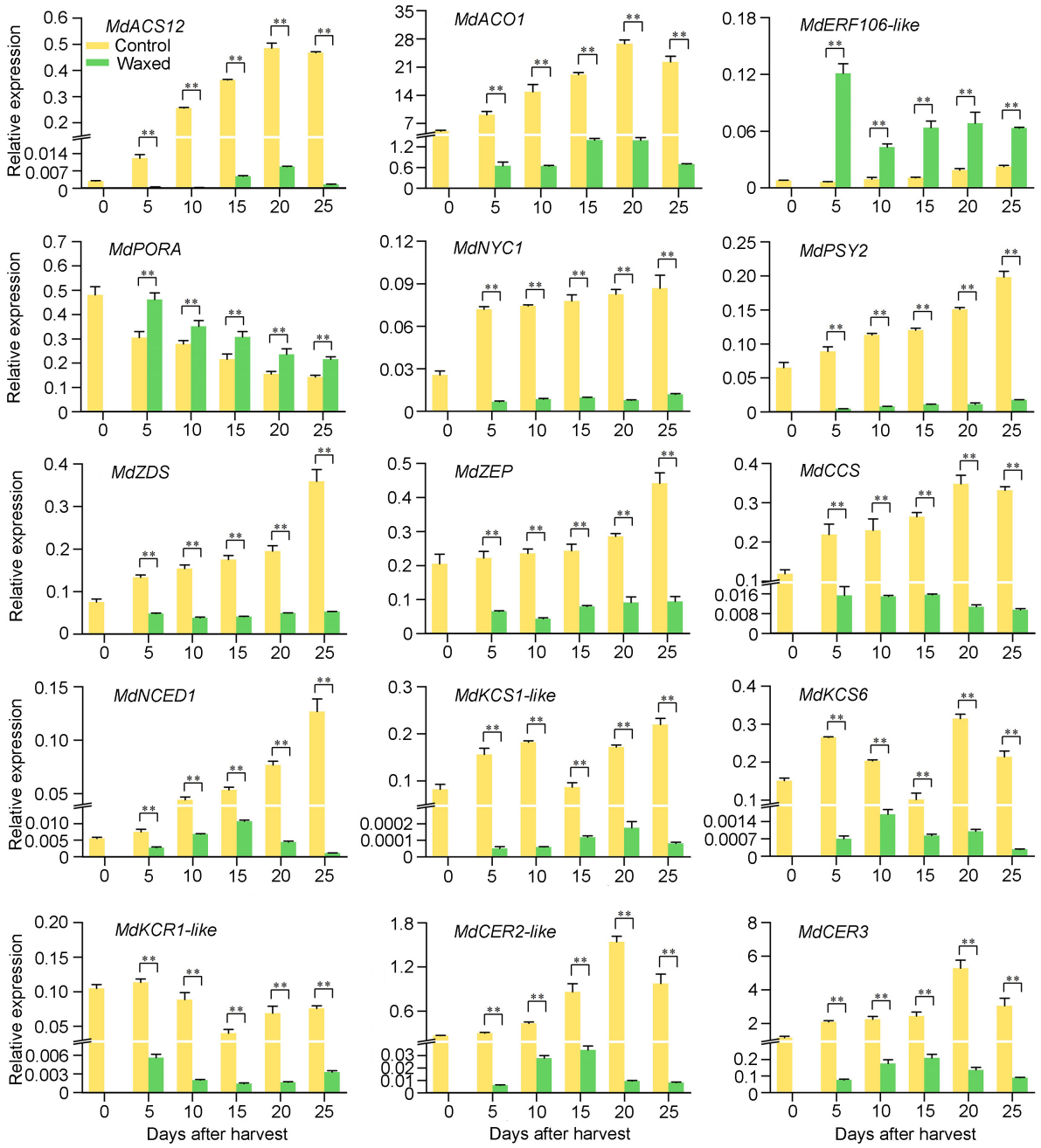


FIGURE 5
qRT-PCR validation of fifteen DEGs during apple fruit storage after wax coating treatment. Control, apple fruit not receiving treatment; Waxed, fruit treated with morpholine fatty acid salt fruit wax. The x-axis represents days of storage at room temperature after harvest. For qRT-PCR, three biological replicates were analyzed. The values represent means ± SE. Statistical significance was determined using a Student's *t*-test, and $^{**}p < 0.01$.

delayed the surge in ethylene production and inhibited the activity of *ACO* (Amornputti et al., 2016). Our transcriptome data showed that several genes involved in ethylene synthesis, such as *ACS* and *ACO*, were downregulated in wax-coated apple and pear fruits (**Figure 2E**). Wax treatment not only affected ethylene-related genes but also affected gene expression of several hormones such as auxin, ABA, cytokinin, and GA (**Supplementary Sheets 2, 5**).

In previous studies, plant hormones have been shown to be involved in regulating fruit development and ripening (Kou et al., 2021). Auxin, the first plant hormone to be discovered, controls plant developmental processes (Zhao, 2010;

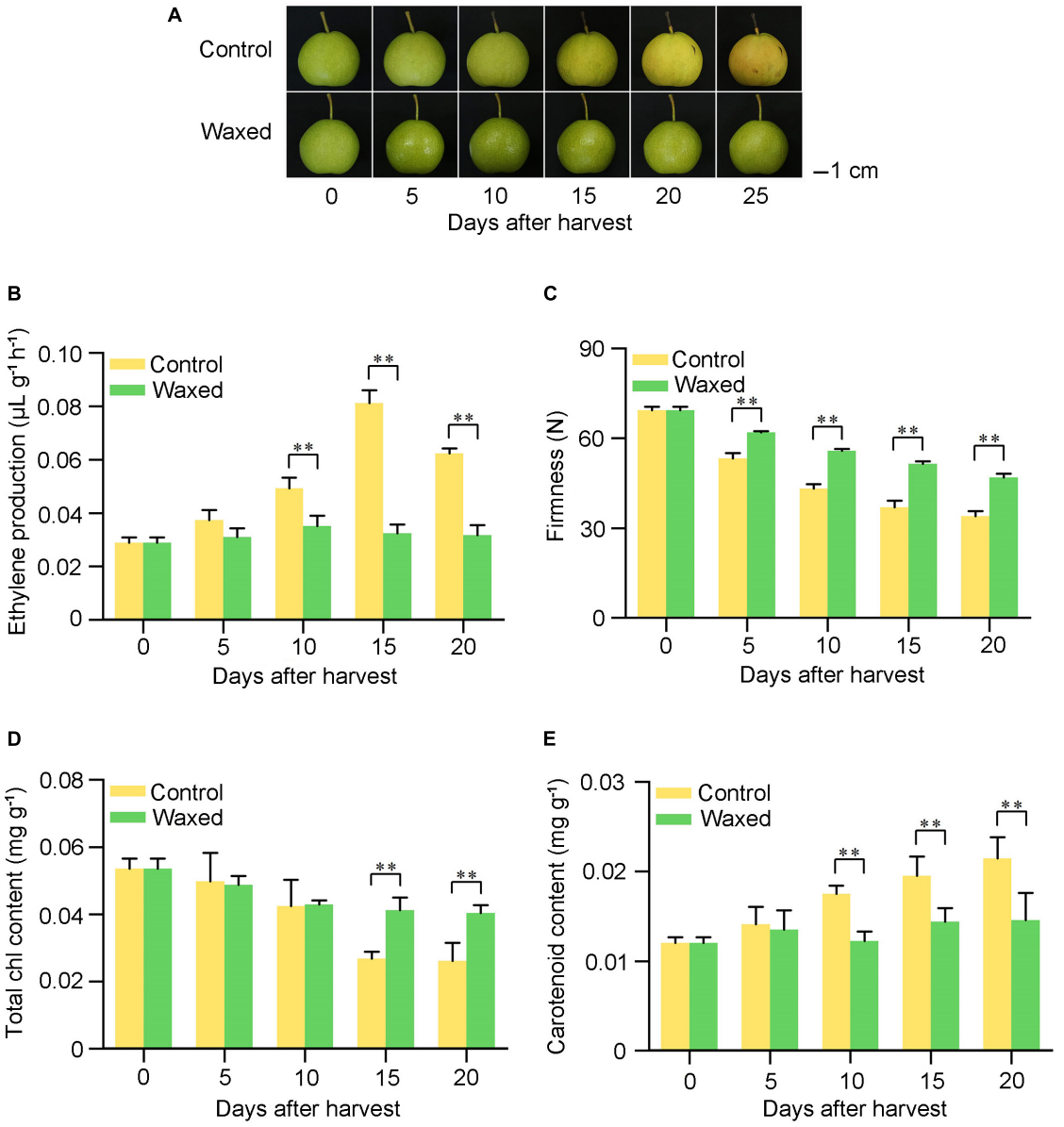


FIGURE 6
Postharvest wax coating treatment inhibits pear fruit ripening. Pear fruit were collected on the commercial harvest day (135 DAFB) in 2020 and storage at room temperature for 25 days **(A)**. After treatment, ethylene production **(B)**, fruit firmness **(C)**, total Chl content **(D)**, and carotenoid content **(E)** were measured. Control, pear fruit not receiving treatment; Waxed, fruit treated with morpholine fatty acid salt fruit wax. Scale bar = 1 cm. The x-axis represents days of storage at room temperature after harvest. Five bioligical replicates were analyzed for ethylene production, fruit firmness, and three bioligical replicates for total Chl content and carotenoid content. Values represent mean ± SE. Statistical significance was determined using a Student's *t*-test, and $^{**}p < 0.01$.

Krizek, 2011). In plants, natural auxins exist mainly in the form of IAA. Auxin is an essential plant hormone for fruit development, and endogenous auxin concentrations increase before fruit ripening (El-Sharkawy et al., 2014). *Aux/IAA* family members play an important role in suppressing ARF-activated gene expression levels (Dreher et al., 2006). After wax treatment, *Aux/IAA* genes were upregulated in the apple fruit; however, in the pear fruit, their differential expression trends were not consistent (**Figure 2A**). This indicated that auxin-related genes play different roles in the regulation of wax application to inhibit fruit ripening.

ABA has a regulatory effect on the ripening of climacteric and non-climacteric fruits (Chen et al., 2016). Moreover the exogenous application of ABA can accelerate the ripening of climacteric fruits (Mou et al., 2016). NCED is an essential enzyme in ABA biosynthesis, and mutations in the *SlNCED*

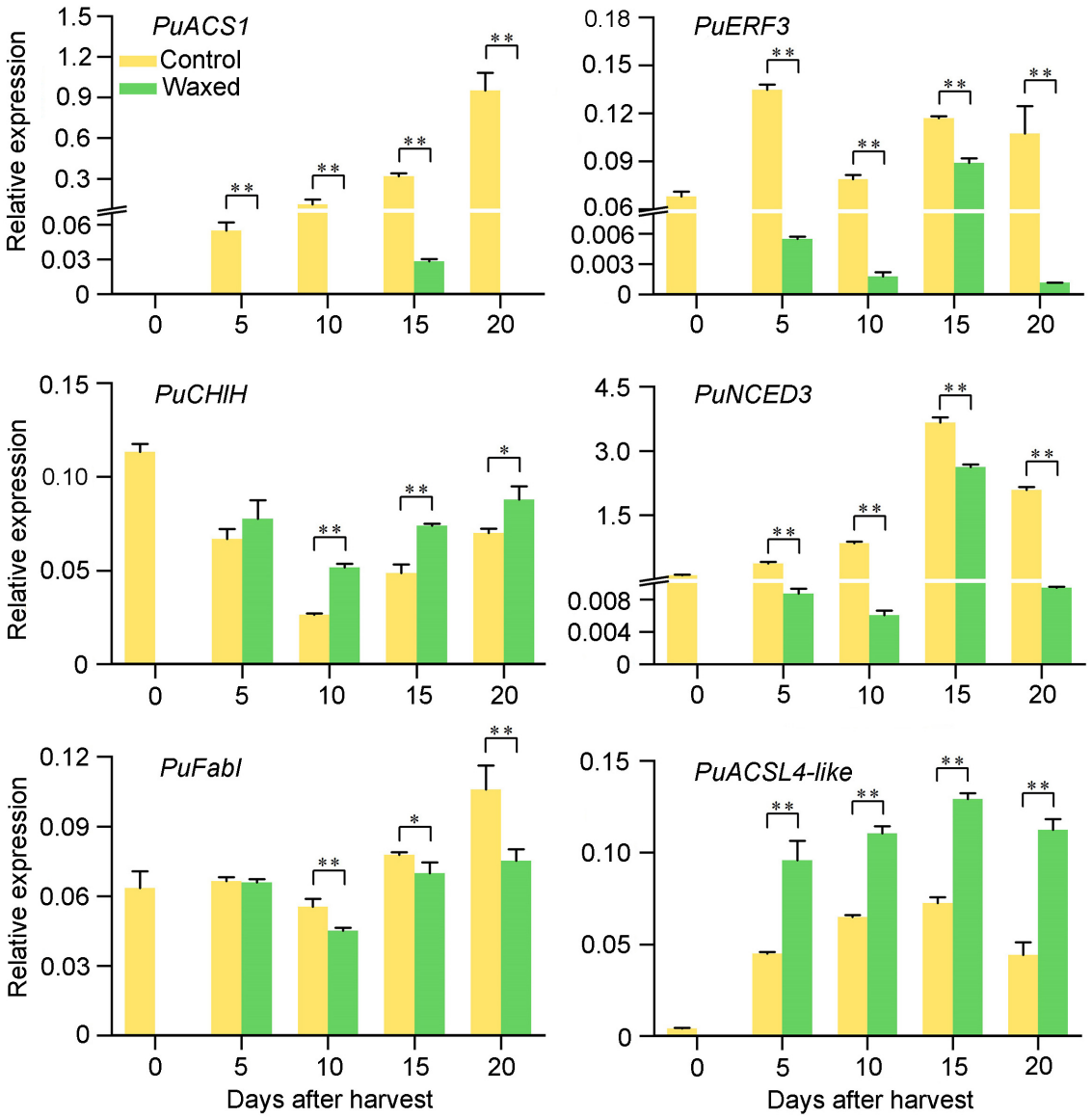


FIGURE 7
qRT-PCR validation of six DEGs during pear fruit storage after wax coating treatment. Control, apple fruit not receiving treatment; Waxed, fruit treated with morpholine fatty acid salt fruit wax. The x-axis represents days of storage at room temperature after harvest. For qRT-PCR, three biological replicates were analyzed. The values represent means ± SE. Statistical significance was determined using a Student's *t*-test, and $^{*}p < 0.05$, $^{**}p < 0.01$.

gene in tomatoes reduce ABA content and delay fruit ripening (Ji et al., 2014). In our study, the *NCED* gene was also found to be downregulated. Overexpression of *PYL9* in tomato fruits promotes early ripening. Moreover, one *PYL* gene was downregulated in our apple transcriptome data, which was consistent with the waxing inhibition of fruit ripening (**Figure 2B**). In summary, wax coating inhibits ABA biosynthesis, and ABA is involved in fruit ripening.

Cytokinins and GAs are important plant hormones. According to the KEGG pathway analysis, we found some DEGs in wax-coated fruit (**Figures 2C,D**). However, the DEGs in apple and pear fruits differed from each other; this may be due to the different effects of wax coating on hormonal signaling in the two species.

Chlorophyll and carotenoids in postharvest waxing treatment

During the natural ripening of the fruit, there is a change in the color of the epidermis, which is mainly caused by a decrease in the chlorophyll content and a concomitant increase

in the carotenoid content in the peel (**Figures 1**, **6**). After pre-treatment of pear fruit with 1-MCP, fruit ripening was inhibited, and the expression of the chlorophyll degradation gene *PcNOL* was downregulated in the treatment group (Zhao J. et al., 2020). Similarly, 1-MCP treatment delayed the breakdown of chlorophyll in cabbage (Hu et al., 2021). In our study, the chlorophyll degradation-related genes *NYC* and *NOL* were downregulated in the apple transcriptome (**Figure 3A** and **Supplementary Sheet 3**). Therefore, inhibiting fruit ripening delays the degradation of pericarp chlorophyll content, regardless of wax coating or 1-MCP treatment. Chlorophyll is synthesized from the abundant amino acid glutamate precursor in the chloroplast, which is a very complex enzymatic reaction process (Masuda, 2008). In Brassica napus, mutations in the coding sequence of BnaA03.ChlH lead to decreased gene activity, affecting chlorophyll biosynthesis (Zhao C. et al., 2020). Oleocellosis of citrus can lead to increased chlorophyll content and up-regulated expression of related genes, such as *MPEC, PORA,* and CAO (Xie et al., 2019). The blocked chlorophyll synthesis of snap bean is related to the *POR* gene (Liu et al., 2020). In rice yellow-green leaf mutants, the content of chlorophyll was reduced, which was found by DNA sequencing due to the substitution of the *DVR* gene in the mutants (Long et al., 2022). Above studies have investigated the expression of the genes associated with chlorophyll synthesis. Our transcriptome data showed that the chlorophyll synthesis-related genes *ChlH, MPEC, PORA, DVR,* and *CAO* were upregulated in apple and pear fruits after wax treatment, suggesting a novel use of wax to keep fruits green (**Figure 3A** and **Supplementary Sheets 3, 5**).

Carotenoids play important coloring roles in a variety of horticultural crops, such as apples, peaches, and mangoes. During the inhibition of fruit ripening by wax treatment, the carotenoid content in "Golden Delicious" apple and "Nanguo pear" fruits increased slowly, resulting in fruits staying greener for a longer duration (**Figures 1**, **6**). As per the transcriptome data, *PSY2*, *CHY*-β, *CCS*, and *NCED* genes were downregulated in apple and pear fruits. PSY2 catalyzes the first step in the carotenoid biosynthetic pathway and is a major rate-limiting enzyme (Zhou et al., 2022). Furthermore, MdPSY2 controls the carotenoid pathway in apples (Ampomah-Dwamena et al., 2015). In pepper fruits, PSY2 can compensate for the loss of PSY1, thereby biosynthesizing carotenoids (Jang et al., 2020). CHY-β converts β-carotene into zeaxanthin, and the expression of *CHY*-β is also increased in peach fruit with increased carotenoid content (Cao et al., 2017). In sweetpotato plants, *CHY*-β gene have the same effect (Park et al., 2015). CCS, one of the enzymes in carotenoid biosynthesis, is mainly expressed in pepper fruits (Arumingtyas et al., 2019). A study found that the expression of *CCS* genes was higher in peppers with high total carotenoid content, and two *CCS* genes were present in mature yellow peppers (Ha et al., 2007). Based on our enrichment results from the KEGG database, NCED can catalyze the conversion of violaxanthin into flavonoids and participate in ABA synthesis. The downregulation of the *NCED* gene after waxing may be due to the reduction of the catalyzed substrate because of the reduction in carotenoid content. Unlike the changing trends of other genes, *VDE* genes were upregulated in both apple and pear transcriptome data (**Figure 3B** and **Supplementary Sheets 3, 5**). VDE is a key enzyme in the lutein cycle and catalyzes the conversion of violaxanthin to zeaxanthin; therefore, we hypothesized that waxing treatment might promote the lutein cycle. Studies have shown that the expression of two EIN3 mature genes in papaya fruit promotes an increase in carotenoid content and the expression of *CpPDS2/4*, *CpZDS*, *CpLCY-e*, and *CpCHY-b* genes are also upregulated (Fu et al., 2017). An *AP2/ERF* transcription factor *MdAP2-34* can promote carotenoid accumulation within 150–170 DAFB (Dang et al., 2021). This indicates that ripening affects the accumulation of fruit carotenoids and the expression of associated genes. Consistent with our results, wax application inhibited fruit ripening and downregulated carotenoid-related gene expression.

Fatty acids and wax biosynthesis in postharvest waxing treatment

Fruit epidermis wax can prevent water evaporation and microbial tissue invasion and enhance the color appearance during storage (Yeats and Rose, 2013).

Typically, wax biosynthesis originates from saturated fatty acids containing 16 or 18 carbon atoms (C16 or C18), forming ultra-long-chain fatty acids in the endoplasmic reticulum. These are subsequently reduced to primary alcohols by acyl groups, followed by carbonyl formation of aldehydes, alkanes, ketones, and other components (Lee and Suh, 2015). The most important component of epidermal wax is a fatty acid. In the transcriptome data from apple and pear, we observed DEGs that are associated with the processes such as fatty acid elongation, wax biosynthesis, fatty acid biosynthesis, and fatty acid degradation (**Figure 4** and **Supplementary Sheets 3, 5**). During the storage of orange fruits, the wax content first increases and then decreases (Ding et al., 2018). Under cold storage conditions, ethephon treatment positively regulates epidermal wax accumulation in apples, whereas 1-MCP treatment inhibits this process (Li et al., 2017). Other researchers have reported that ethephon treatment increases the expression of *MdCER6*, *MdCER4*, and *MdWSD1*, the key genes for ultra-long-chain fatty acid synthesis in "Starkrimson apples, and 1-MCP decreases their expression (Li F. et al., 2019). Overexpression of *MdERF2*, a key factor in ethylene signaling, can upregulate the expression levels of Long-chain acyl-CoA synthetase 2 (*MdLACS2), Wax synthase1 (MdWSD1*), Eceriferum 4 (*MdCER4*), and Eceriferum 6 (*MdCER6*), which positively regulates the accumulation of total wax (Yingjie et al., 2022). Based on the previous studies,

fruit ripening affects the wax content and the expression of related genes, and ethylene increases the accumulation of wax and changes the structure and shape of wax. In our study, waxing treatment inhibited fruit ripening, so *LACS*, *KCS*, *KCR*, and *CER* were all downregulated as per the apple transcriptome data (**Figure 4**). However, in the pear fruit, *LACS*, *KCR*, and *WSD1* were upregulated after treatment (**Supplementary Sheet 5**). This may be related to the wax coating material we selected, or it may be that the content and composition of wax in apple and pear fruits are different, resulting in differential gene expression after wax coating. Interestingly, we found that the differential genes involved in fatty acid degradation were upregulated in both apple and pear fruits (**Supplementary Sheets 3, 5**). Using transcriptomic data, we found that wax coating inhibited wax synthesis. To deeply explore the effect of wax coating on endogenous fruit wax, measurement of endogenous wax seems to be a promising strategy for future studies to obtain more convincing physiological data.

Conclusion

In the present study, the wax coating treatment delayed the ripening of climacteric fruits, including a decrease in ethylene production, an increase in chlorophyll content, and an increase in the carotenoid content. Moreover, we analyzed the genes that were differentially expressed during this process by transcriptome sequencing of apples and pears. Based on the functional clustering of the DEGs, we identified that, in addition to ethylene, chlorophyll, and carotenoids, several other plant hormones play a role in this process and participate in the regulation of maturation. Further, we observed that fatty acid and wax synthesis-related genes were differentially affected by the wax application, inhibiting fruit ripening. In conclusion, our study provides a comprehensive overview of the gene networks that orchestrate the process of ripening of climacteric fruits.

Data availability statement

The original contributions presented in this study are publicly available. This data can be found here: NCBI, PRJNA851428.

Author contributions

HY, AW, and YS conceived this study. YS performed most of the experiments. TL and HL provided the plant materials. JL and JS treated the apple and pear fruit and took the sample. ZM and JQ participated in RNA extraction. HB analyzed the data. YS and HY wrote the manuscript. All authors read and approved the submitted manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (32125034).

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.978013/full#supplementary-material

References

Adams-Phillips, L., Barry, C., and Giovannoni, J. (2004). Signal transduction systems regulating fruit ripening. *Trends Plant Sci.* 9, 331–338. doi: 10.1016/j.tplants.2004.05.004

Amornputti, S., Ketsa, S., and van Doorn, W. G. (2016). 1-Methylcyclopropene (1-MCP) inhibits ethylene production of durian fruit which is correlated with a decrease in ACC oxidase activity in the peel. *Postharvest Biol. Technol.* 114, 69–75. doi: 10.1016/j.postharvbio.2015.11.020

Ampomah-Dwamena, C., Driedonks, N., Lewis, D., Shumskaya, M., Chen, X., Wurtzel, E. T., et al. (2015). The Phytoene synthase gene family of apple (*Malus x domestica*) and its role in controlling fruit carotenoid content. *BMC Plant Biol.* 15:185. doi: 10.1186/s12870-015-0573-7

Arumingtyas, E. L., Fuadati, A., and Dwinianti, E. (2019). Study on the profile of capsanthin-capsurobin synthase (Ccs) gene responsible for carotenoid synthesis

in chili pepper (*Capsicum frutescens* L.) mutants G1M6 M2 generation. *IOP Conf. Series* 391:012019.

Cao, S., Liang, M., Shi, L., Shao, J., Song, C., Bian, K., et al. (2017). Accumulation of carotenoids and expression of carotenogenic genes in peach fruit. *Food Chem.* 214, 137–146. doi: 10.1016/j.foodchem.2016.07.085

Carvajal, F., Castro-Cegrí, A., Jiménez-Muñoz, R., Jamilena, M., Garrido, D., and Palma, F. (2021). Changes in Morphology, Metabolism and Composition of Cuticular Wax in Zucchini Fruit During Postharvest Cold Storage. *Front. Plant Sci.* 12:778745. doi: 10.3389/fpls.2021.778745

Chen, P., Sun, Y. F., Kai, W. B., Liang, B., Zhang, Y. S., Zhai, X. W., et al. (2016). Interactions of ABA signaling core components (SlPYLs, SlPP2Cs, and SlSnRK2s) in tomato (*Solanum lycopersicon*). *Plant Physiol.* 205, 67–74. doi: 10.1016/j.jplph.2016.07.016

Chu, W., Gao, H., Chen, H., Fang, X., and Zheng, Y. (2018). Effects of cuticular wax on the postharvest quality of blueberry fruit. *Food Chem.* 239, 68–74. doi: 10.1016/j.foodchem.2017.06.024

Cruz, A. B., Bianchetti, R. E., Alves, F. R. R., Purgatto, E., Peres, L. E. P., Rossi, M., et al. (2018). Light, ethylene and auxin signaling interaction regulates carotenoid biosynthesis during tomato fruit ripening. *Front. Plant Sci.* 9:1370. doi: 10.3389/fpls.2018.01370

Dandekar, A. M., Teo, G., Defilippi, B. G., Uratsu, S. L., Passey, A. J., Kader, A. A., et al. (2004). Effect of Down-Regulation of Ethylene Biosynthesis on Fruit Flavor Complex in Apple Fruit. *Transgenic Res.* 13, 373–384. doi: 10.1023/B:TRAG.0000040037.90435.45

Dang, Q., Sha, H., Nie, J., Wang, Y., Yuan, Y., and Jia, D. (2021). An apple (*Malus domestica*) AP2/ERF transcription factor modulates carotenoid accumulation. *Hortic. Res.* 8:223. doi: 10.1038/s41438-021-00694-w

Ding, S., Zhang, J., Wang, R., Ou, S., and Shan, Y. (2018). Changes in cuticle compositions and crystal structure of 'Bingtang'sweet orange fruits (*Citrus sinensis*) during storage. *Int. J. Food Prop.* 21, 2411–2427. doi: 10.1080/10942912.2018.1528272

Dreher, K. A., Brown, J., Saw, R. E., and Callis, J. (2006). The Arabidopsis Aux/IAA protein family has diversified in degradation and auxin responsiveness. *Plant Cell* 18, 699–714. doi: 10.1105/tpc.105.039172

El-Sharkawy, I., Sherif, S. M., Jones, B., Mila, I., Kumar, P. P., Bouzayen, M., et al. (2014). TIR1-like auxin-receptors are involved in the regulation of plum fruit development. *J. Exp. Bot.* 65, 5205–5215. doi: 10.1093/jxb/eru279

Farcuh, M., Rivero, R. M., Sadka, A., and Blumwald, E. J. P. B. (2018). Ethylene regulation of sugar metabolism in climacteric and non-climacteric plums. *Postharvest Biol. Technol.* 139, 20–30. doi: 10.1016/j.postharvbio.2018.01.012

Farneti, B., Khomenko, I., Ajelli, M., Emanuelli, F., Biasioli, F., and Giongo, L. (2022). Ethylene Production Affects Blueberry Fruit Texture and Storability. *Front. Plant Sci.* 13:813863. doi: 10.3389/fpls.2022.813863

Fu, C. C., Han, Y. C., Kuang, J. F., Chen, J. Y., and Lu, W. J. (2017). Papaya CpEIN3a and CpNAC2 co-operatively regulate carotenoid biosynthesis-related genes CpPDS2/4, CpLCY-e and CpCHY-b during fruit ripening. *Plant Cell Physiol.* 58, 2155–2165. doi: 10.1093/pcp/pcx149

Gan, Z., Fei, L., Shan, N., Fu, Y., and Chen, J. (2019). Identification and Expression Analysis of Gretchen Hagen 3 (GH3) in Kiwifruit (*Actinidia chinensis*) During Postharvest Process. *Plants* 8:473. doi: 10.3390/plants8110473

Ha, S.-H., Kim, J.-B., Park, J.-S., Lee, S.-W., and Cho, K.-J. (2007). 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. doi: 10.1093/jxb/erm132

Hu, H., Zhao, H., Zhang, L., Zhou, H., and Li, P. (2021). The application of 1-methylcyclopropene preserves the postharvest quality of cabbage by inhibiting ethylene production, delaying chlorophyll breakdown and increasing antioxidant capacity. *Sci. Hortic.* 281:109986. doi: 10.1016/j.scienta.2021.109986

Jang, S.-J., Jeong, H.-B., Jung, A., Kang, M.-Y., Kim, S., Ha, S.-H., et al. (2020). Phytoene synthase 2 can compensate for the absence of PSY1 in the control of color in Capsicum fruit. *J. Exp. Bot.* 71, 3417–3427. doi: 10.1093/jxb/eraa155

Ji, K., Kai, W., Zhao, B., Sun, Y., Yuan, B., Dai, S., et al. (2014). SlNCED1 and SlCYP707A2: Key genes involved in ABA metabolism during tomato fruit ripening. *J. Exp. Bot.* 65, 5243–5255. doi: 10.1093/jxb/eru288

Jia, M., Feng, J., Zhang, L., Zhang, S., and Xi, W. (2021). PaPYL9 is Involved in the Regulation of Apricot Fruit Ripening through ABA Signaling Pathway. *Hortic. Plant J.* 8, 461–473. doi: 10.1016/j.hpj.2021.11.012

Klee, H. J., and Giovannoni, J. J. (2011). Genetics and control of tomato fruit ripening and quality attributes. *Annu. Rev. Genet.* 45, 41–59. doi: 10.1146/annurev-genet-110410-132507

Kou, X., Feng, Y., Yuan, S., Zhao, X., Wu, C., Wang, C., et al. (2021). Different regulatory mechanisms of plant hormones in the ripening of climacteric and non-climacteric fruits: A review. *Plant Mol. Biol.* 107, 477–497. doi: 10.1007/s11103-021-01199-9

Kou, X., and Wu, M. (2018). "Characterization of climacteric and non-climacteric fruit ripening," in *Plant Senescence. Methods in Molecular Biology*, ed. Y. Guo (New York, NY: Humana Press), 89–102.

Krizek, B. A. (2011). Auxin regulation of Arabidopsis flower development involves members of the AINTEGUMENTA-LIKE/PLETHORA (AIL/PLT) family. *J. Exp. Bot.* 62, 3311–3319. doi: 10.1093/jxb/err127

Larrainzar, E., Molenaar, J. A., Wienkoop, S., Gil-Quintana, E., Alibert, B., Limami, A. M., et al. (2014). Drought stress provokes the down-regulation of methionine and ethylene biosynthesis pathways in *Medicago truncatula* roots and nodules. *Plant Cell Environ.* 37, 2051–2063. doi: 10.1111/pce.12285

Lee, S. B., and Suh, M. C. (2015). Advances in the understanding of cuticular waxes in *Arabidopsis thaliana* and crop species. *Plant Cell Rep.* 34, 557–572. doi: 10.1007/s00299-015-1772-2

Lelièvre, J. M., Latchè, A., Jones, B., Bouzayen, M., and Pech, J. C. (1997). Ethylene and fruit ripening. *Physiol. Plant.* 101, 727–739. doi: 10.1111/j.1399-3054.1997.tb01057.x

Li, F., Min, D., Ren, C., Dong, L., Shu, P., Cui, X., et al. (2019). Ethylene altered fruit cuticular wax, the expression of cuticular wax synthesis-related genes and fruit quality during cold storage of apple *(Malus domestica* Borkh. cv Starkrimson) fruit. *Postharvest Biol. Technol.* 149, 58–65. doi: 10.1016/j.postharvbio.2018.11.016

Li, H., Wu, H., Qi, Q., Li, H., Li, Z., Chen, S., et al. (2019). Gibberellins play a role in regulating tomato fruit ripening. *Plant Cell Physiol.* 60, 1619–1629. doi: 10.1093/pcp/pcz069

Li, F., Min, D., Song, B., Shao, S., and Zhang, X. (2017). Ethylene effects on apple fruit cuticular wax composition and content during cold storage. *Postharvest Biol. Technol.* 134, 98–105. doi: 10.1016/j.postharvbio.2017.08.011

Li, J., Li, Q., Lei, X., Tian, W., Cao, J., Jiang, W., et al. (2018). Effects of wax coating on the moisture loss of cucumbers at different storage temperatures. *J. Food Qual.* 2018:9351821. doi: 10.1155/2018/9351821

Li, X., Zhu, X., Wang, H., Lin, X., Lin, H., and Chen, W. (2018). Postharvest application of wax controls pineapple fruit ripening and improves fruit quality. *Postharvest Biol. Technol.* 136, 99–110. doi: 10.1016/j.postharvbio.2017.10.012

Li, T., Li, X., Tan, D., Jiang, Z., Wei, Y., Li, J., et al. (2014). Distinct expression profiles of ripening related genes in the 'Nanguo' pear (*Pyrus ussuriensis*) fruits. *Sci. Hortic.* 171, 78–82.

Liu, C., Li, Y., Liu, D., Yan, Z., Feng, G., and Yang, X. (2020). Blocked chlorophyll synthesis leads to the production of golden snap bean pods. *Mol. Genet. Genom.* 295, 1325–1337. doi: 10.1007/s00438-020-01699-1

Long, W., Long, S., Jiang, X., Xu, H., Peng, Q., Li, J., et al. (2022). A Rice Yellow-Green-Leaf 219 Mutant Lacking the Divinyl Reductase Affects Chlorophyll Biosynthesis and Chloroplast Development. *J. Plant Growth Regul.* 1–10. doi: 10.1007/s00344-021-10508-x

Masuda, T. (2008). Recent overview of the Mg branch of the tetrapyrrole biosynthesis leading to chlorophylls. *Photosynth. Res.* 96, 121–143. doi: 10.1007/s11120-008-9291-4

Mou, W., Li, D., Bu, J., Jiang, Y., Khan, Z. U., Luo, Z., et al. (2016). Comprehensive analysis of ABA effects on ethylene biosynthesis and signaling during tomato fruit ripening. *PLoS One* 11:e0154072. doi: 10.1371/journal.pone.0154072

Osorio, S., Scossa, F., and Fernie, A. R. (2013). Molecular Regulation of Fruit Ripening. *Front. Plant Sci.* 4:198. doi: 10.3389/fpls.2013.00198

Park, S.-C., Kim, S. H., Park, S., Lee, H.-U., Lee, J. S., Park, W. S., et al. (2015). Enhanced accumulation of carotenoids in sweetpotato plants overexpressing IbOr-Ins gene in purple-fleshed sweetpotato cultivar. *Plant Physiol. Biochem.* 86, 82–90. doi: 10.1016/j.plaphy.2014.11.017

Pech, J. C., Purgatto, E., Bouzayen, M., and Latché, A. (2012). Ethylene and fruit ripening. *Annu. Plant Rev. Plant Horm. Ethyl.* 44, 275–304. doi: 10.1002/9781118223086.ch11

Rashwan, A. K., Karim, N., Shishir, M. R. I., Bao, T., Lu, Y., and Chen, W. (2020). *Jujube* fruit: A potential nutritious fruit for the development of functional food products. *J. Funct. Foods* 75:104205. doi: 10.1016/j.jff.2020.104205

Sang, N. (2020). Effect of ratio of bees wax and carnauba wax in mixed wax on respiration rate, weight loss, fruit decay and chemical quality of vietnamese passion fruits during low temperature storage. *Pak. J. Biotechnol.* 17, 63–70. doi: 10.34016/pjbt.2020.17.2.63

Sousa, F. F., Junior, J. S. P., Oliveira, K. T., Rodrigues, E. C., Andrade, J. P., and Mattiuz, B. H. (2021). Conservation of 'Palmer'mango with an edible coating of hydroxypropyl methylcellulose and beeswax. *Food Chem.* 346:128925.

Suwandi, T., Dewi, K., and Cahyono, P. (2016). Pineapple harvest index and fruit quality improvement by application of gibberellin and cytokinin. *Fruits* 71, 209–214. doi: 10.1051/fruits/2016010

Tan, D., Li, T., and Wang, A. (2013). Apple 1-Aminocyclopropane-1-Carboxylic Acid Synthase Genes,MdACS1andMdACS3a, are Expressed in Different Systems of Ethylene Biosynthesis. *Plant Mol. Biol. Rep.* 31, 204–209. doi: 10.1007/s11105-012-0490-y

Vázquez-Celestino, D., Ramos-Sotelo, H., Rivera-Pastrana, D. M., Vázquez-Barrios, M. E., and Mercado-Silva, E. M. (2016). Effects of waxing, microperforated polyethylene bag, 1-methylcyclopropene and nitric oxide on firmness and shrivel and weight loss of 'Manila'mango fruit during ripening. *Postharvest Biol. Technol.* 111, 398–405. doi: 10.1016/j.postharvbio.2015.09.030

Wei, Y., Jin, J., Xu, Y., Liu, W., Yang, G., Bu, H., et al. (2021). Ethylene-activated MdPUB24 mediates ubiquitination of MdBEL7 to promote chlorophyll degradation in apple fruit. *Plant J.* 108, 169–182. doi: 10.1111/tpj.15432

Wu, H., Liu, L., Chen, Y., Liu, T., Jiang, Q., Wei, Z., et al. (2022). Tomato SlCER1-1 catalyzes the synthesis of wax alkanes which increases the drought tolerance and fruit storability. *Hortic. Res.* 9:uhac004. doi: 10.1093/hr/uhac004

Xie, J., Yao, S., Ming, J., Deng, L., and Zeng, K. (2019). Variations in chlorophyll and carotenoid contents and expression of genes involved in pigment metabolism response to oleocellosis in citrus fruits. *Food Chem.* 272, 49–57. doi: 10.1016/j.foodchem.2018.08.020

Xu, B., and Wu, S. (2021). Preservation of mango fruit quality using fucoidan coatings. *LWT Food Sci. Technol.* 143:111150. doi: 10.1016/j.lwt.2021.111150

Yeats, T. H., and Rose, J. K. (2013). The formation and function of plant cuticles. *Plant Physiol.* 163, 5–20. doi: 10.1104/pp.113.222737

Yingjie, S., Xinyue, Z., Yaping, J., Jihan, W., Bingru, L., Xinhua, Z., et al. (2022). Roles of ERF2 in apple fruit cuticular wax synthesis. *Sci. Hortic.* 301:111144. doi: 10.1016/j.scienta.2022.111144

Yuan, H., Yue, P., Bu, H., Han, D., and Wang, A. (2020). Genome-wide analysis of ACO and ACS genes in pear (*Pyrus ussuriensis*). *In vitro Cell. Dev.Plant.* 56, 193–199. doi: 10.1007/s11627-019-10009-3

Yuan, H., Zhang, L., Jiang, Z., and Wang, A. (2017). Characterization of ripening-related puarp4 in pear (*pyrus ussuriensis*). *J. Plant Growth Regul.* 36, 766–772. doi: 10.1007/s00344-017-9680-z

Yue, P., Lu, Q., Liu, Z., Lv, T., Li, X., Bu, H., et al. (2020). Auxin-activated MdARF5 induces the expression of ethylene biosynthetic genes to initiate apple fruit ripening. *New Phytol.* 226, 1781–1795. doi: 10.1111/nph.16500

Zhao, C., Liu, L., Safdar, L. B., Xie, M., Cheng, X., Liu, Y., et al. (2020). Characterization and fine mapping of a yellow-virescent gene regulating chlorophyll biosynthesis and early stage chloroplast development in Brassica napus. *G3* 10, 3201–3211. doi: 10.1534/g3.120.401460

Zhao, J., Xie, X., Wang, S., Zhu, H., Dun, W., Zhang, L., et al. (2020). 1-Methylcyclopropene affects ethylene synthesis and chlorophyll degradation during cold storage of 'Comice'pears. *Sci. Hortic.* 260, 108865. doi: 10.1016/j.scienta.2019.108865

Zhao, Y. (2010). Auxin biosynthesis and its role in plant development. *Annu. Rev. Plant Biol.* 61, 49–64. doi: 10.1146/annurev-arplant-042809-112308

Zhou, X., Rao, S., Wrightstone, E., Sun, T., Lui, A., Welsch, R., et al. (2022). Phytoene Synthase: The Key Rate-Limiting Enzyme of Carotenoid Biosynthesis in Plants. *Front. Plant Sci.* 13:884720. doi: 10.3389/fpls.2022.884720

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 22 August 2022
DOI 10.3389/fpls.2022.950084

OPEN ACCESS

EDITED BY
Neftali Ochoa-Alejo,
Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Mexico

REVIEWED BY
Walid Badawy Abdelaal,
Agricultural Research Center, Egypt
Shaimaa Abdelmohsen,
Princess Nourah Bint Abdulrahman University, Saudi Arabia

*CORRESPONDENCE
Jun Han
HJhan2022@163.com
Shuming Liu
liusm@nwafu.edu.cn

SPECIALTY SECTION
This article was submitted to Plant Development and EvoDevo, a section of the journal Frontiers in Plant Science

RECEIVED 01 June 2022
ACCEPTED 27 July 2022
PUBLISHED 22 August 2022

CITATION
Su K, Sun J, Han J, Zheng T, Sun B and Liu S (2022) Combined morphological and multi-omics analyses to reveal the developmental mechanism of *Zanthoxylum bungeanum* prickles. *Front. Plant Sci.* 13:950084. doi: 10.3389/fpls.2022.950084

Combined morphological and multi-omics analyses to reveal the developmental mechanism of *Zanthoxylum bungeanum* prickles

Kexing Su[1], Jiaqian Sun[2,3], Jun Han[4]*, Tao Zheng[1], Bingyin Sun[5] and Shuming Liu[1]*

[1]College of Science, Northwest Agriculture and Forestry University, Xianyang, China, [2]Powerchina Northwest Engineering Corporation Limited, Xi'an, China, [3]Shaanxi Union Research Center of University and Enterprise for River and Lake Ecosystems Protection and Restoration, Xi'an, China, [4]Forestry and Grassland Bureau of Xunhua County, Qinghai, China, [5]Department of Ecological Engineering, Yangling Vocational and Technical College, Xianyang, China

Zanthoxylum bungeanum Maxim. as an important economic forest, its epidermis bears prickles which complicate the harvesting process and increase the labor costs. To explore the developmental mechanism of prickles, three varieties of *Zanthoxylum bungeanum* (PZB, SZB, GSZB) were selected for morphological and multi-omics analyses. The absorption spectra of prickles and stems were detected using Fourier-transform infrared spectroscopy (FTIR), and they were found different at 1617, 1110, 3319, and 1999 cm^{-1}. The morphology of prickles and stems were observed using light microscopy and transmission electron microscopy (TEM). The growth direction of cells on the prickle side and stem side were perpendicular to each other, and there was a resembling abscission zone (RAZ) between them. The vacuolar deposits of prickle cells were much more than stem cells, indicating that the lignification degree of prickles was higher than stems. In addition, 9 candidate genes (*ZbYABBY2*, *ZbYABBY1*, *ZbYABBY5*, *ZbWRKY*, *ZbLOG5*, *ZbAZG2*, *ZbGh16*, *ZbIAA33*, and *ZbGh16X1*) were screened out and validated base on transcriptome and qRT-PCA. As well as, 30 key metabolites were found related to prickle development base on metabolome analysis. Among them, 4-hydroxy-2-oxopentanoate, trans-2-hydroxy-cinnamate, trans-cinnamate, polyhydroxy-fatty acid, 10,16-dihydroxypalmitate, cinnamic acid were related to the biosynthesis of cutin, suberine and wax. Indole-3-acetate, tryptamine, anthranilate, fromylanthranilate, N6-(delta2-isopentenyl)-adenine were related to plant hormone signal transduction. Generally, this is the first study to reveal the developmental mechanism of prickles. The results of this study lay the foundation for the breeding of non-prickle *Zanthoxylum bungeanum*.

KEYWORDS

***Zanthoxylum bungeanum*, prickles, resembling abscission zone, vacuolar deposition, transcriptome, metabolome**

Introduction

Plants with spines, thorns or prickles are widely spread in nature (Simpson, 2019; Airoldi and Glover, 2020). These sharp pointed structures protect them from herbivores (Grubb, 1992; Simcha, 2016). In botanical terms, spines refer to modified leaves, thorns refer to modified stems, while prickles refer to modified epidermal cells (Steeves and Sussex, 1989; Huchelmann et al., 2017). In this study, our research object is prickles. The representatives of prickle-bearing plants including *Girardinia suborbiculata*, *Chorisia speciosa*, *Zanthoxylum bungeanum*, *Rubus corchorifolius*, and *Rosa chinensis* which suggests that prickles have evolved many times and represent an example of convergent evolution (Sussex and Kerk, 2001; Zhang et al., 2020).

Zanthoxylum bungeanum (Rutaceae), a drought-tolerant shrub, has been widely cultivated for its medicinal and edible value (Yang et al., 2013; Ma et al., 2019; Liu et al., 2020). Its pericarps, branches, roots and leaves are all medicine (Sun et al., 2019, 2020), which has been used to treat asthma (Tang et al., 2014), obesity (Wang et al., 2019, 2020), colitis (Zhang et al., 2017), diabetes (Zhang et al., 2021), and eczema (Su et al., 2020). As the native of *Zanthoxylum bungeanum*, Chinese cultivation area and output are both the largest in the world (Li et al., 2015; Chen et al., 2019; Sheng et al., 2020). However, evolution and natural selection have led the epidermis of *Zanthoxylum bungeanum* to bear prickles (Coverdale, 2020), just as **Figure 1** shows. These prickles make cultivation management difficult and at the same time result in high harvest costs (Appelhans et al., 2018; Fei et al., 2020). The same problem also plagues other prickle-bearing cash crops such as *Rosa chinensis*, *Vitis vinifera*, and *Rosa roxburghii* (Rajapakse and Arumuganathan, 2001; Debener and Linde, 2009; Wang et al., 2021).

Early research suggests that prickles are modified glandular trichomes that differentiate into prickle morphologies upon lignification (Kellogg et al., 2011). Until recently, only a few studies had been published about the mechanism of prickle development, but great progress had been made in trichome formation. In *Arabidopsis thaliana*, more than twenty genes are required to control the development of trichomes (Szymanski et al., 1998). The initial selection of trichome precursors requires the activity of both *GL1* and *TTG* genes (Herman and Marks, 1989; Payne et al., 1999). The *GL2* gene is required for subsequent phases of trichome morphogenesis (Chopra et al., 2019). And gibberellins could promote trichome formation by up-regulating *GL1* (Chien and Sussex, 1996; Perazza et al., 1998; Matias-Hernandez et al., 2016). In addition, *CPC*, *TRY*, *TCL1*, *ETC1*, *ETC2*, and *ETC3* are negative regulators of trichome initiation (Wada et al., 1997; Glover et al., 1998; Kirik et al., 2005; Wang et al., 2007). "Activator-inhibitor model" and "activator-substrate model" can be used to explain the mechanism of trichome development (Balkunde et al., 2010).

Except *Arabidopsis thaliana*, research on prickles of *Rosa roxburghii*, *Vitis vinifera*, and *Rosa chinensis* has also made some progress. Huang and his team proved that RrGL1 of *Rosa roxburghii* is an R2R3 MYB homolog which regulates trichome formation by interacting with GL3/EGL3 protein (Huang et al., 2019; Yan et al., 2021). In *Vitis vinifera*, Yin and Barba's research validated and mapped a major QTL for trichome density (Barba et al., 2019; Yin et al., 2021). In *Rosa chinensis*, *RcGL1*, *RcMYB82*, *RcMYB61*, *RcCPC*, *RcTRY*, *RcGL3*, *RcTT8*, *RcMYC1*, *RcTTG1*, *RcTTG2*, *RcZFP5*, *RcGIS3*, *RcGIS2*, and *RcZFP1* genes were considered candidate genes to control prickle development (Bourke et al., 2018; Zhou et al., 2020). Although some progress has been made in the study of prickles, there is no research on *Z. bungeanum* prickles yet.

To explore the developmental mechanism of *Z. bungeanum* prickles, this study selected three varieties of *Z. bungeanum* as experimental materials. Using Fourier-transform infrared spectroscopy, optical microscopy, transmission electron microscopy, transcriptomics and metabolomics methods to explore the developmental mechanism of prickles from the perspective of morphological and multi-omics analysis. Our research not only revealed the submicroscopic structure of prickles, but also screened out candidate genes and metabolites, which provided a theoretical basis for the breeding of non-prickle *Zanthoxylum bungeanum*.

Materials and methods

Plant materials

Three varieties of *Z. bungeanum* used for experiment were: Wild Prickly *Zanthoxylum bungeanum* (PZB), Wild Smooth *Zanthoxylum bungeanum* (SZB), and Grafted Smooth *Zanthoxylum bungeanum* (GSZB). Previous studies have shown that grafting can inhibit the expression of *Z. bungeanum* prickles, and the inhibition effect of grafting with young stems is more obvious than that with old stems. Therefore, we collected both young and old stems as experiment materials (three biological replicates per tissue), as shown in **Figure 2**.

The "Grafted Smooth *Zanthoxylum bungeanum*" was planted in Didian village (109.227158 E; 34.985359 N), The "Wild Prickly *Zanthoxylum bungeanum*" and "Wild smooth *Zanthoxylum bungeanum*" were picked in the Qinling Mountains (108.329167 E; 33.801945 N). These plant tissues were immediately frozen in liquid nitrogen and stored at −80°C for subsequent analysis.

Fourier transform infrared spectroscopy

The plant tissues were divided into 14 groups (**Supplementary Table 1**). They were placed in an oven to dry (60°C, 24 h) and then ground to powder. The powder was mixed with KBr at a ratio of 1:100 and pressed into flakes. The

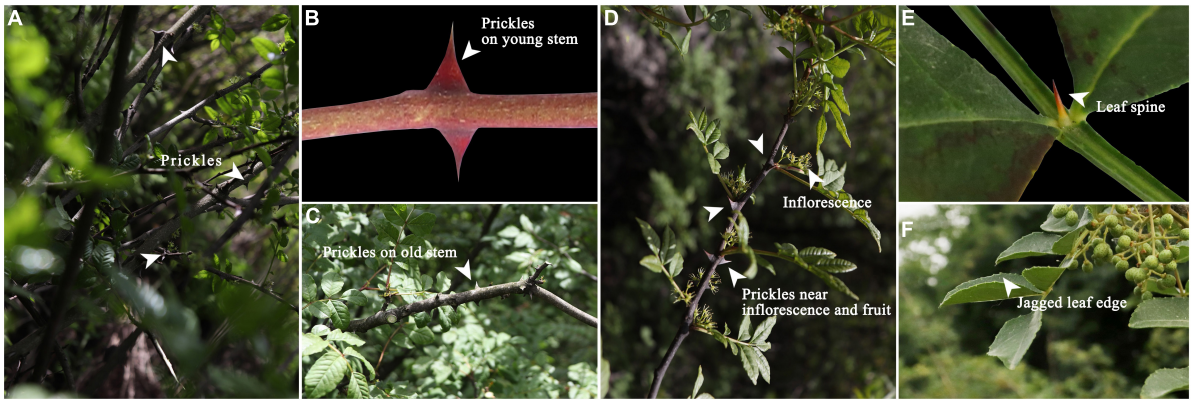


FIGURE 1
(A) Prickles; **(B)** Prickles on young stem; **(C)** Prickles on old stem; **(D)** Prickles near inflorescence; **(E)** Leaf spine; **(F)** Jagged leaf edge.

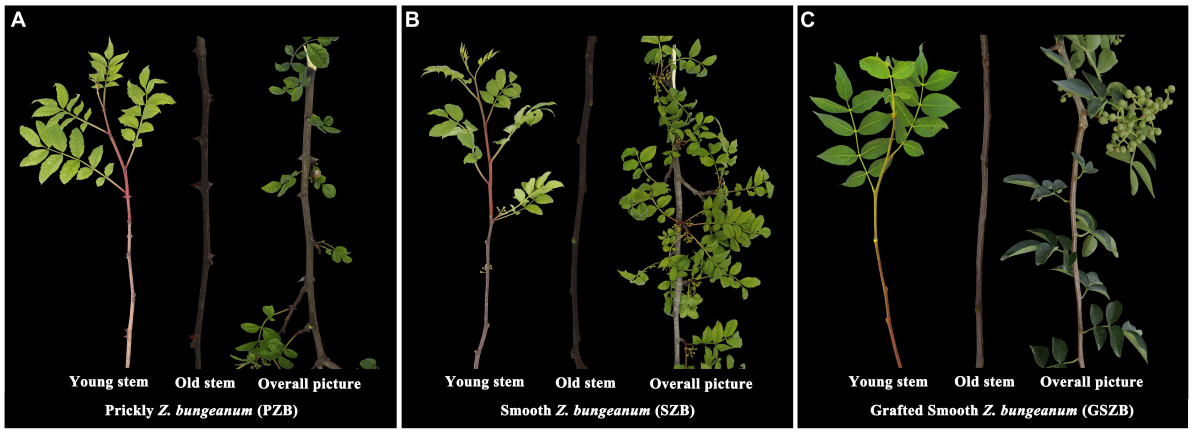


FIGURE 2
(A) Wild Prickly *Zanthoxylum bungeanum*; **(B)** Wild Smooth *Zanthoxylum bungeanum*; **(C)** Grafted Smooth *Zanthoxylum bungeanum*.

Fourier-transform infrared spectroscopy (vertex70, Germany) was used to determine the absorption spectra of plant tissues, each sample repeated three times. Finally, data analysis was performed using OPUS software (BRUKER, Munich, Germany) (Hu et al., 2021; Teimouri et al., 2021).

Preparation and observation of paraffin sections

The plant tissues were cut into small pieces (5mm × 5mm × 2mm) and fixed with FAA solution (38% Acetaldehyde 5 ml, 70% ethanol 90 ml, glacial acetic acid 5 ml, glycerol 5 ml) for 24 h. Then, they were subjected to the following dehydration procedure: in 70, 80, 90, 95, and 100% ethanol (each of 30 min), 1:1 mixture of xylene and absolute ethanol (30 min), xylene solution (twice, 20 min). Subsequently, they were embedded in paraffin wax and placed in a 38°C oven for 48 h. Then, they were cut into 8–10 μm sections using Leica microtome. The sections were stained with 0.05% toluidine blue (pH 4.3) and examined by Leica light microscope (Ma et al., 2021).

Observation of transmission electron microscope

The young stem segments were fixed with 4% glutaraldehyde at 4°C for 12 h and washed four times with 0.1 M phosphate buffer (pH 6.8). Afterward, they were post-fixed with 1% osmium tetroxide and subjected to the following dehydration procedure: in 30, 50, 70, 80, 90, and 100% alcohol (twice, each of 10 min), and then embedded in epoxy resin. Later, the material was cut into 80 nm ultrathin sections using ultrathin microtome (EMUC7). The ultrathin sections were transferred to copper grids and post-stained with

uranyl acetate and lead citrate. Observations were made with a TECNAI G2 SPIRIT BIO TEM at 120 kV (Naidoo et al., 2012; Munien et al., 2015).

cDNA library construction, sequencing, and data analysis

The young stem bark of PZB, SZB, and GSZB was sent to Sangon Biotech (Shanghai, China) for library construction, quality control, and paired-end sequencing with Illumina HiSeq^{TM} (Bolger et al., 2014). After sequencing, the clean reads were selected by removing low-quality sequences, and the transcripts were obtained by assembling clean reads using Trinity software (Wang et al., 2012; Okonechnikov et al., 2016). Then the transcripts were de-redundant to get unigenes. And unigenes were compared to public databases (NR, NT, KOG, Pfam, Swissprot, CDD, TrEMBL, GO, KEGG) using BLAST with an *e*-value threshold of 10^{-5}.

Identification and analysis of differentially expressed genes

To verify the transcription expression levels of all samples, transcripts per million (TPM) were used to quantify the expression level of genes. Differentially expressed genes (DEGs) between SZB, PZB and GSZB samples were identified using DESeq2, with $|\log_2$ fold change$| > 2$ and q-Value < 0.05. GO enrichment analyses were performed using clusterProfiler (Kanehisa and Goto, 2000; Tatusov et al., 2000; Marchler-Bauer et al., 2013; UniProt, 2015; Finn et al., 2016). Sequence alignment and phylogenetic analysis were performed on the homologous proteins sequences of candidate genes using MEGA software. And CD-Search was used to analyze conserved domains of homologous proteins.[1]

Quantitative real-time polymerase chain reaction verification

The specific quantitative primers (**Supplementary Table 2**) were designed using Primer 5.0 software, and the candidate gene sequences were shown in **Supplementary Table 3**. The total ribonucleic acid (RNA) was extracted from young stem bark of PZB, SZB, and GSZB using TRIzol reagent (Invitrogen, Carlsbad, CA, United States). First-strand cDNA was synthesized using the PrimeScript^{TM} RT Reagent Kit (TAKARA, Beijing, China) according to the manufacturer's instructions. TB Green® Premix Ex Taq^{TM} (TAKARA, China) was used to perform qRT-PCR on an ABI StepOne Plus (Applied Biosystems, Foster, CA, United States) (Qin et al., 2020). Relative quantification of specific mRNA levels was performed using the cycle threshold $2^{(-\Delta\Delta Ct)}$ method with the *Zanthoxylum bungeanum* 18S rRNA gene (HG002512.1) as an internal control.

Metabolomic profiling

The young stem bark of PZB, SZB, and GSZB were selected for metabolomic analysis. Extraction and analysis of differentially accumulated metabolites (DAMs) were performed by the Metware Biotechnology Co., Ltd. (Wuhan, China). Biological samples are freeze-dried by vacuum freeze-dryer (Scientz-100F) (Chen et al., 2013). Then the sample extracts were analyzed using an UPLC-ESI-MS/MS system (UPLC, SHIMADZU Nexera X2; MS, Applied Biosystems 4500 Q TRAP) (Chen et al., 2009; Fraga et al., 2010). DAMs were identified based on the thresholds $|\log_2$(fold change)$| \geq 1$ and VIP (variable importance in project) ≥ 1 (Fu et al., 2021).

Integrative analysis of transcriptome and metabolome

Genes and metabolites with a Pearson correlation coefficient (PCC) $> |0.8|$ and P-value < 0.05, were used to draw the nine-quadrant diagram and cluster heatmap (Jozefczuk et al., 2010).

Results

Differences in functional groups of prickle, bark, and stalk

The FTIR absorption spectrum of prickle was different from that of bark and stalk (**Figure 3**). The absorption spectrum of prickle had band at 1617 and 1110 cm^{-1}, while the bark and stalk had strong band at 3319 and 1999 cm^{-1}. According to the band correspondence, the associated -NH_2 or -NH bond could cause a band to appear at 3319 cm^{-1}, which means there were compounds contain such chemical bonds in bark and stalk. Besides the 1110 cm^{-1} band was caused by C-O-C stretching and symmetric vibration of the ester linkage, or -CH stretching in aromatic ring (Breub et al., 2010; Lupoi et al., 2015). In addition, the well-defined narrow band at 1617 cm^{-1} was attributed to C = C aromatic ring vibration, which indicated that the prickles contain a lot of aromatic compounds (Reyes-Rivera and Terrazas, 2017).

The FTIR absorption spectra of young and old stalk of PZB, SZB, and GSZB are shown in **Supplementary Figure 1A**, while the FTIR absorption spectra of young and old bark of

1 https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi

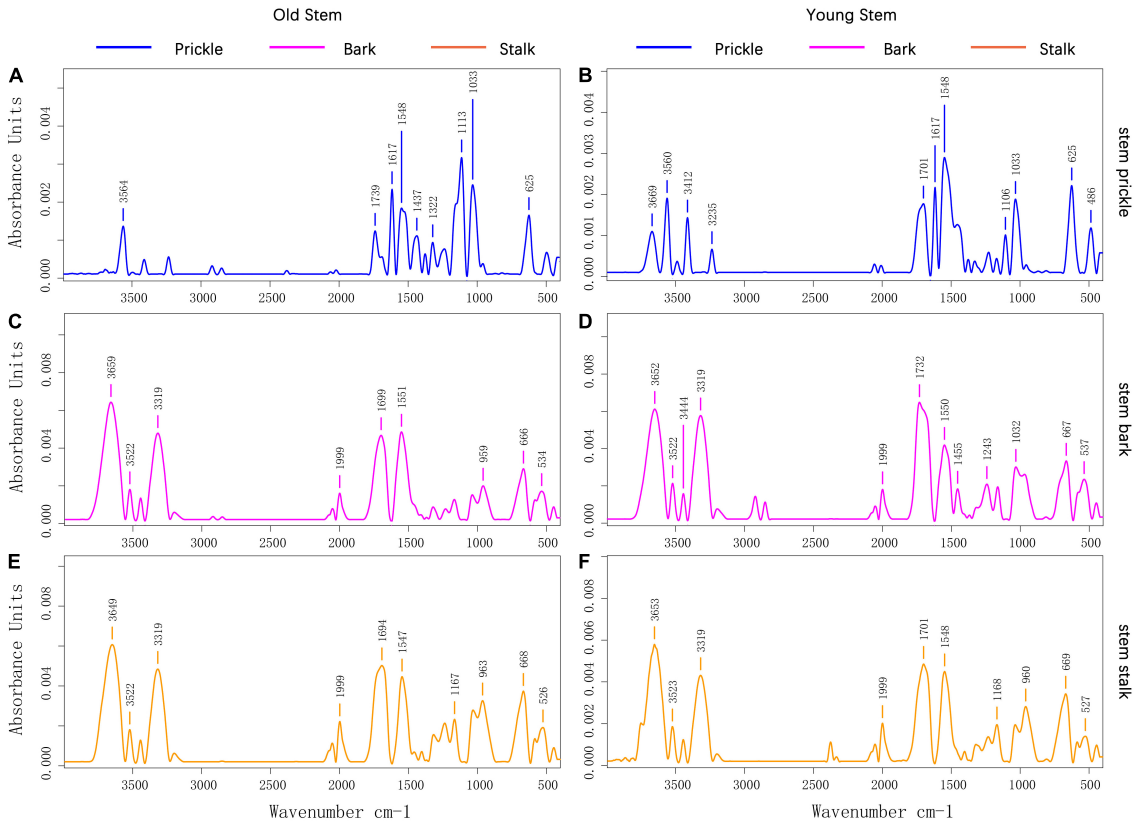


FIGURE 3
(A) Prickle's absorption spectrum on old stems of PZB; **(B)** Prickle's absorption spectrum on young stems of PZB; **(C)** Bark's absorption spectrum on old stems of PZB; **(D)** Bark's absorption spectrum on young stems of PZB; **(E)** Stalk's absorption spectrum on old stems of PZB; **(F)** Stalk's absorption spectrum on young stems of PZB.

PZB, SZB, and GSZB are shown in **Supplementary Figure 1B**. The results suggested that there was no significant difference in their functional groups between young and old stems, as well as between bark and stem. Therefore, we focused our research on prickle and young stems, which were significantly different, and explored them further later.

Microstructure of prickles and young stems

From **Figure 4A**, it can be seen that there was no vascular bundle in prickles, which means that the prickles of *Zanthoxylum bungeanum* were different from thorns, and belonging to the category of prickles (Airoldi and Glover, 2020). Besides, the growth direction of cells on prickle side and stem side were perpendicular to each other (**Figure 4B**), and there was an obvious dividing line between them, called resembling abscission zone (RAZ). The finding of RAZ provided a good explanation for why prickles were easily peeled off from stems.

The microstructure of prickles and young stems of PZB, SZB, and GSZB were shown in **Figure 4**. It can be found that the parts with high degree of lignification were deeply stained, while the parts with low degree of lignification were lightly stained. This difference was more obvious in pseudo-color images, that the spicule cells were similar in color to xylem cells and epidermal cells, while the stem cells were similar in color to parenchyma cells. This result indicated that prickle cells were lignified earlier than stem cells.

Ultra-microstructure of prickles and young stems

By comparing the cell morphology of PZB, SZB and GSZB (**Figures 5A–C**), it was found that there were far more crystalline substances (CS) in the vacuoles of prickle cells than stem cells. The detail structures of these crystalline substances (CS) were shown in **Figures 5E,F**, and the morphology of **Figure 5F** maybe evolve from **Figure 5E**. Besides, we combined this finding with that of microstructure, and speculated that this kind of crystalline substances (CS) might be a precursor for lignin synthesis.

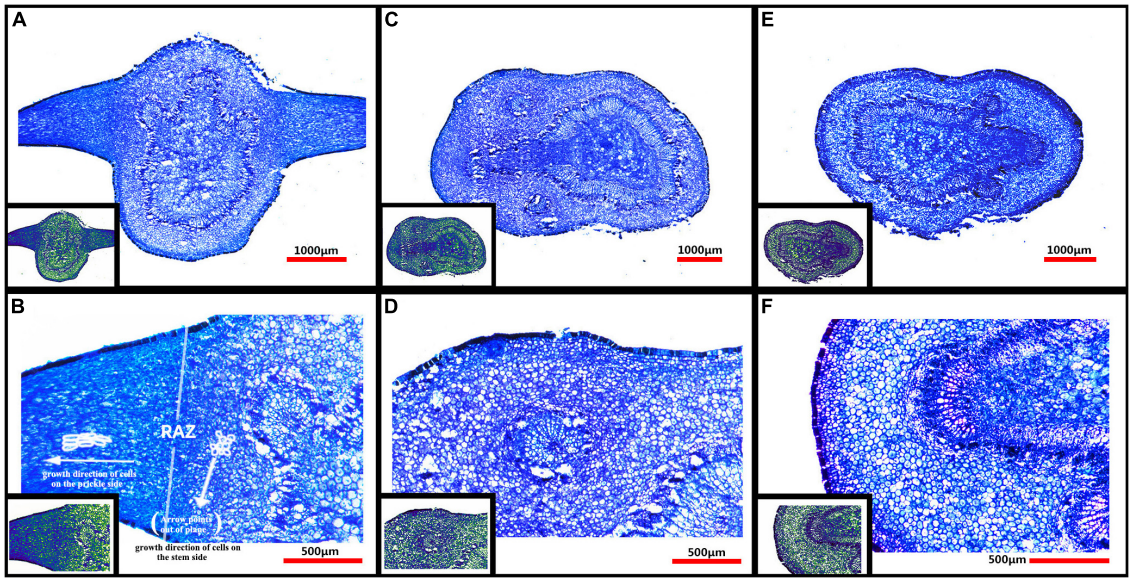


FIGURE 4
(A) Microstructure of PZB's young stem and prickles; **(B)** Partial enlarged view of panel **(A)**; **(C)** Microstructure of SZB's young stem; **(D)** Partial enlarged view of panel **(C)**; **(E)** Microstructure of GSZB's young stem; **(F)** Partial enlarged view of panel **(E)**. The lower left corner of each image is its corresponding pseudo-color image.

Figure 5G shows the epidermal cells of prickles. These epidermal cells were dead cells without any organelles, and their cell walls (CW) were thickened (Dominguez et al., 2011; Guzman et al., 2014; de la Fuente et al., 2017). Besides, there was a layer of epidermal wax (EW) on the outside of cell walls which means these cells were highly lignified (Jenks et al., 1994; Ma Z. Y. et al., 2016; Pereira, 2018).

Figures 5H,I show the chloroplasts (CP) structure of prickle cells. In addition to starch granules (SG) and thylakoids (Khan et al., 2020), a kind of dark crystalline substances (CS) were found in chloroplasts (CP), which means that during the process of cell lignification, certain changes have also occurred in chloroplasts (CP).

Figure 5J shows that there were plasmodesmata (PD) between cells in the resembling abscission zone (RAZ) (Li et al., 2012). In addition, it was found that the cells near the two sides of the resembling abscission zone (RAZ) had no obvious difference in ultrastructure.

Ribonucleic acid sequencing, assembly and functional annotation

In total, 9 cDNA libraries from the young stem bark of PZB, SZB, and GSZB were constructed and sequenced. The raw reads of the libraries were deposited in the NCBI Sequence Read Archive (SRA) with Bioproject accession number: PRJNA752915 and SRA accession number: SRR15440480-SRR15440497. After quality control and low-quality data screening, clean reads were obtained and assembled into 176,862 unigenes, with an average length of 596.7 bp. Among these unigenes, the maximum length was 16,917 bp, the minimum length was 201 bp (**Figure 6A**).

A total of 95,931 (54.24%) unigenes were annotated using nine public databases. In summary, 38,563 (21.80%), 27,755 (15.69%), 83,880 (47.43%), 73,846 (41.75%), 24,950 (14.11%), 56,697 (32.06%), 83,435 (47.18%), 70,937 (40.11%), and 6,421 (3.63%) unigenes were annotated in the CDD, KOG, NR, NT, PFAM, Swissprot, TrEMBL, GO and KEGG databases, respectively (**Figures 6B,C**). According to the Nr database, 37,367 (21.13%) unigenes exhibited significantly higher homology with sequences from *Citrus sinensis* than *Citrus clementina* and other species (**Supplementary Figure 2A**). Unigenes annotated in the KOG database were mainly distributed in "signal transduction mechanisms" and "general function prediction only" function classes (**Supplementary Figure 2B**). Besides, the annotation results of unigenes in GO gene function classification and KEGG pathway classification were shown in **Supplementary Figures 2C,D**.

Differentially expressed genes identification and enrichment analyses

The TPM values (Transcripts Per Million) were calculated for each unigene. Besides, $|\log_2(\text{fold change})| > 2$ and $q\text{-value} < 0.05$ were set as thresholds for significant DEGs

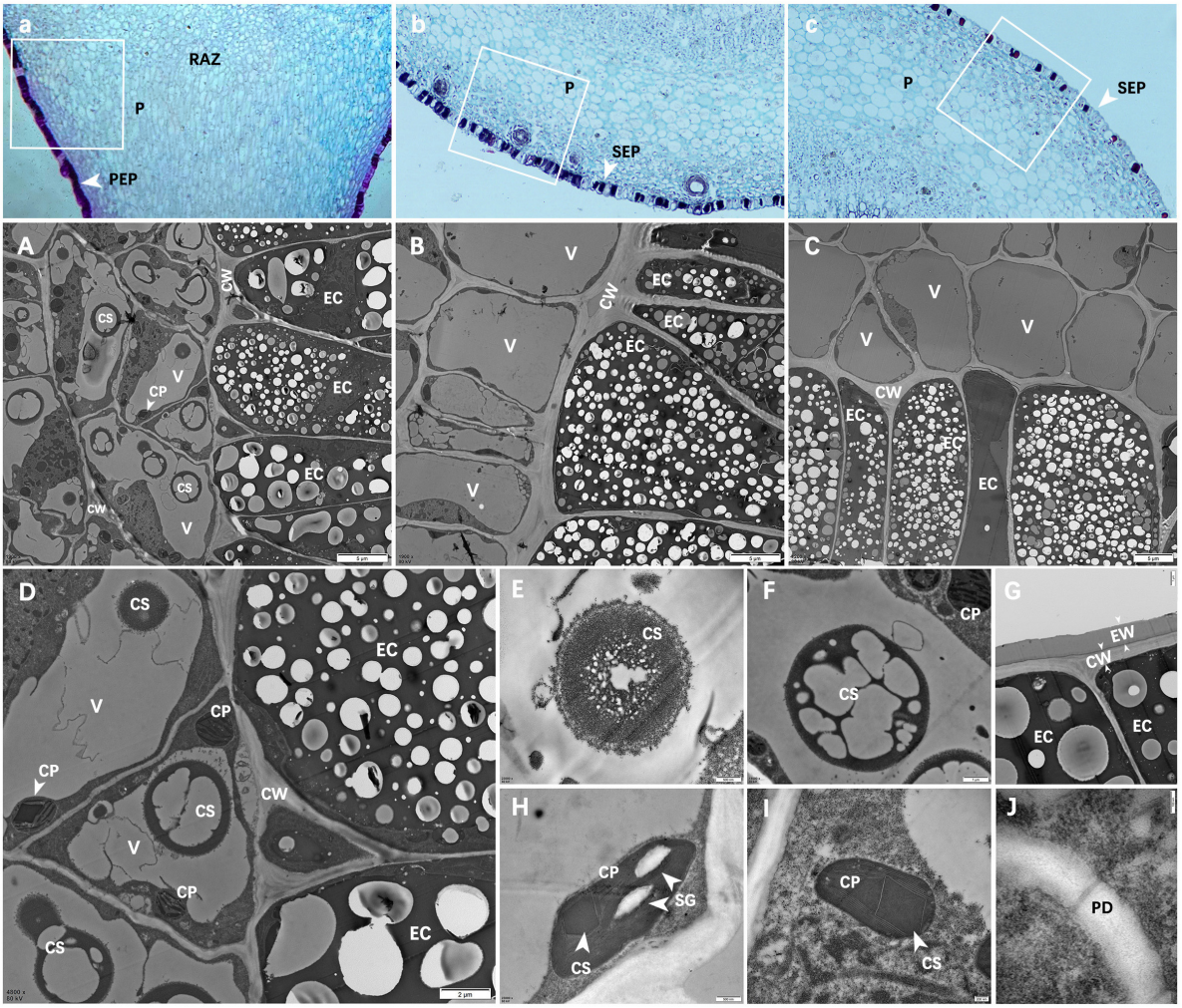


FIGURE 5
Ultra-microstructure of prickle and stem cells. **(A–C)** Are partial enlargements of the white boxes in panels **a–c**; **(A)** Epidermal cells of PZB prickle; **(B)** Epidermal cells of SZB stem; **(C)** Epidermal cell of GSZB stem; **(D)** Enlarged view of panel **(A)**; (**E–J)** Are enlargements of panel **(D)**; **(E,F)** Crystals in vacuoles; **(G)** Epidermal cells; **(H,I)** Chloroplasts; **(J)** Plasmodesmata. Scale bars: **(A–C)** = 5 μm; **(D)** = 2 μm; **(F,G)** = 1 μm; **(E,H)** = 500 nm; **(I)** = 200 nm; **(J)** = 100 nm. PEP, prickle epidermis; SEP, stem epidermis; P, parenchyma; RAZ, resembling abscission zone; V, vacuole; EC, epidermal cell; CP, chloroplast; CW, cell wall; CS, crystalline substance; EW: epicuticular waxes; SG, starch grains; PD, plasmodesmata.

selection. It can be seen from Figures 7A-C that a total of 27 DEGs were detected between PZB and SZB libraries (6 upregulated and 21 downregulated genes); 16,569 DEGs were detected between GSZB and PZB libraries (7,165 upregulated and 9,404 downregulated genes); 15,130 DEGs were detected between GSZB and SZB libraries (6,832 upregulated and 8,298 downregulated genes) (Supplementary Tables 4–6).

In order to screen for genes that affect prickle development, the DEGs shared by PZB_vs_SZB and PZB_vs_GSZB were selected, and the DEGs of SZB_vs_GSZB were excluded. As a result, 15 ungenes were selected as candidate genes (white slashed area in Figure 7A and Supplementary Table 3), and quantitative real-time PCR analysis was performed on them. The results of qRT-PCR and RNA-Seq were found to be in close agreement, validating the accuracy of sequencing (Supplementary Figure 3).

The hierarchical clustering results of DEGs in PZB, SZB, and GSZB groups are shown in Figure 7D. Green and red indicate high and low expression levels, respectively. From it can be seen that different groups of the same varieties can be clustered together, indicating that the data obtained were reliable and reproducible.

The significant DEGs from different groups were functionally categorized using GO (Gene Ontology) enrichment analyses (Supplementary Figures 4A–C and Supplementary Tables 7–9). The results found that the

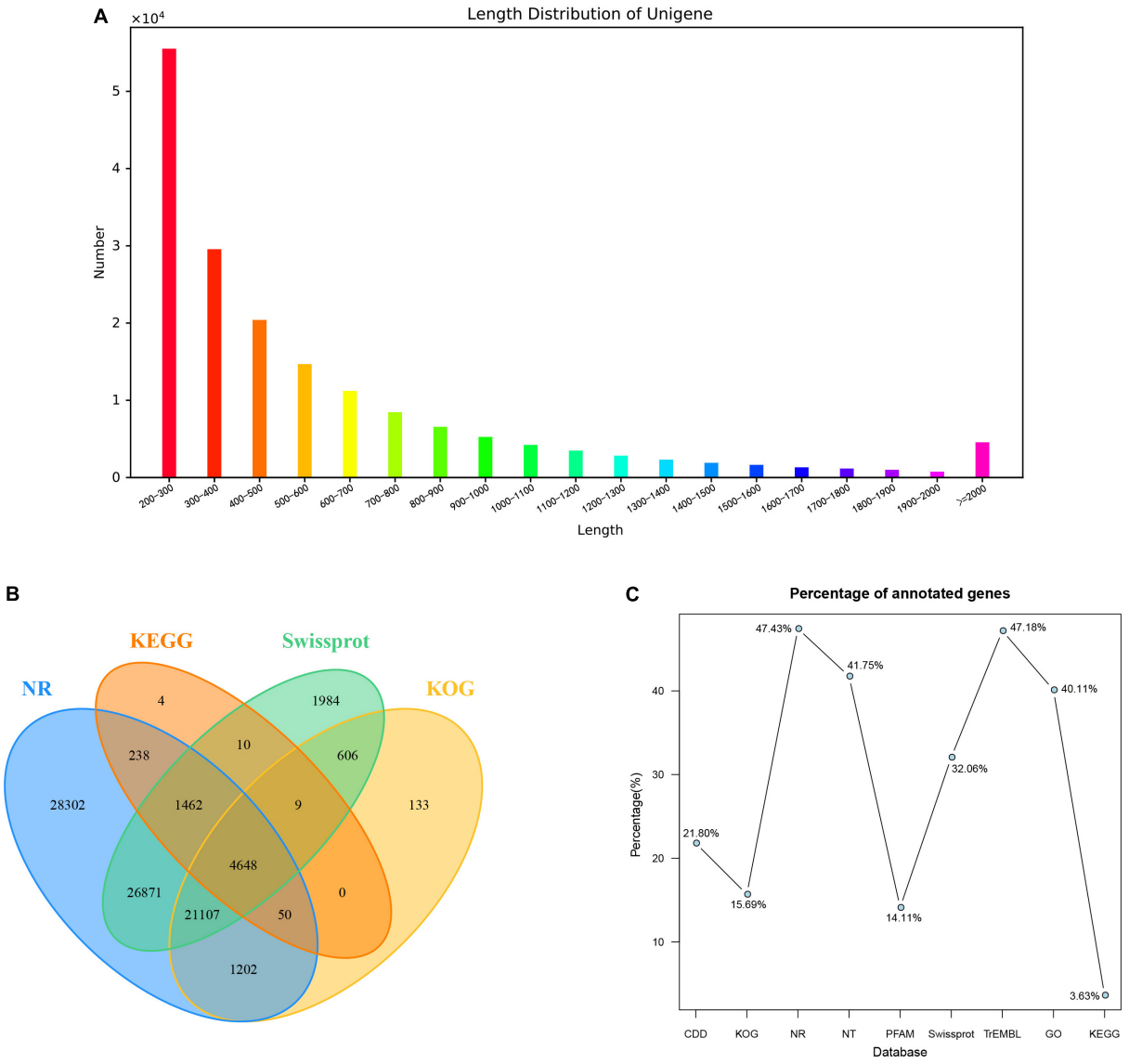


FIGURE 6
Summary of the assembly and annotations. **(A)** Length distribution of unigenes; **(B)** Venn diagram of annotation results; **(C)** Annotation ratio of different databases.

significant DEGs between the PZB and SZB libraries were mainly enriched in "regulation of shoot apical meristem development (GO:1902183)," "regulation of flower development (GO:0009909)," "specification of axis polarity (GO:0065001)," "polarity specification of adaxial/abaxial axis (GO:0009944)," "xyloglucosyl transferase activity (GO:0016762)," "xyloglucan metabolic process (GO:0010411)," and "plant hormone signal transduction (GO:0009739)" (**Supplementary Figure 4D** and **Supplementary Table 10**). In addition, the significant DEGs between SZB and GSZB libraries as well as between PZB and GSZB libraries are shown in **Supplementary Figures 4E,F** and **Supplementary Tables 11, 12**. To better display the results, the top 30 GO terms with the highest enrichment between PZB and SZB were selected to draw directed acyclic graphs (**Supplementary Figure 5**). The results of DAGs contain three parts: biological process, cellular component and molecular function, which are shown in **Supplementary Figures 5A–C**, respectively.

Identification of unigenes related to prickle development

In this section, detailed analyses of 15 candidate genes affecting the development of *Zanthoxylum bungeanum* prickles were carried out (*ZbYABBY2*, *ZbYABBY1*, *ZbYABBY5*,

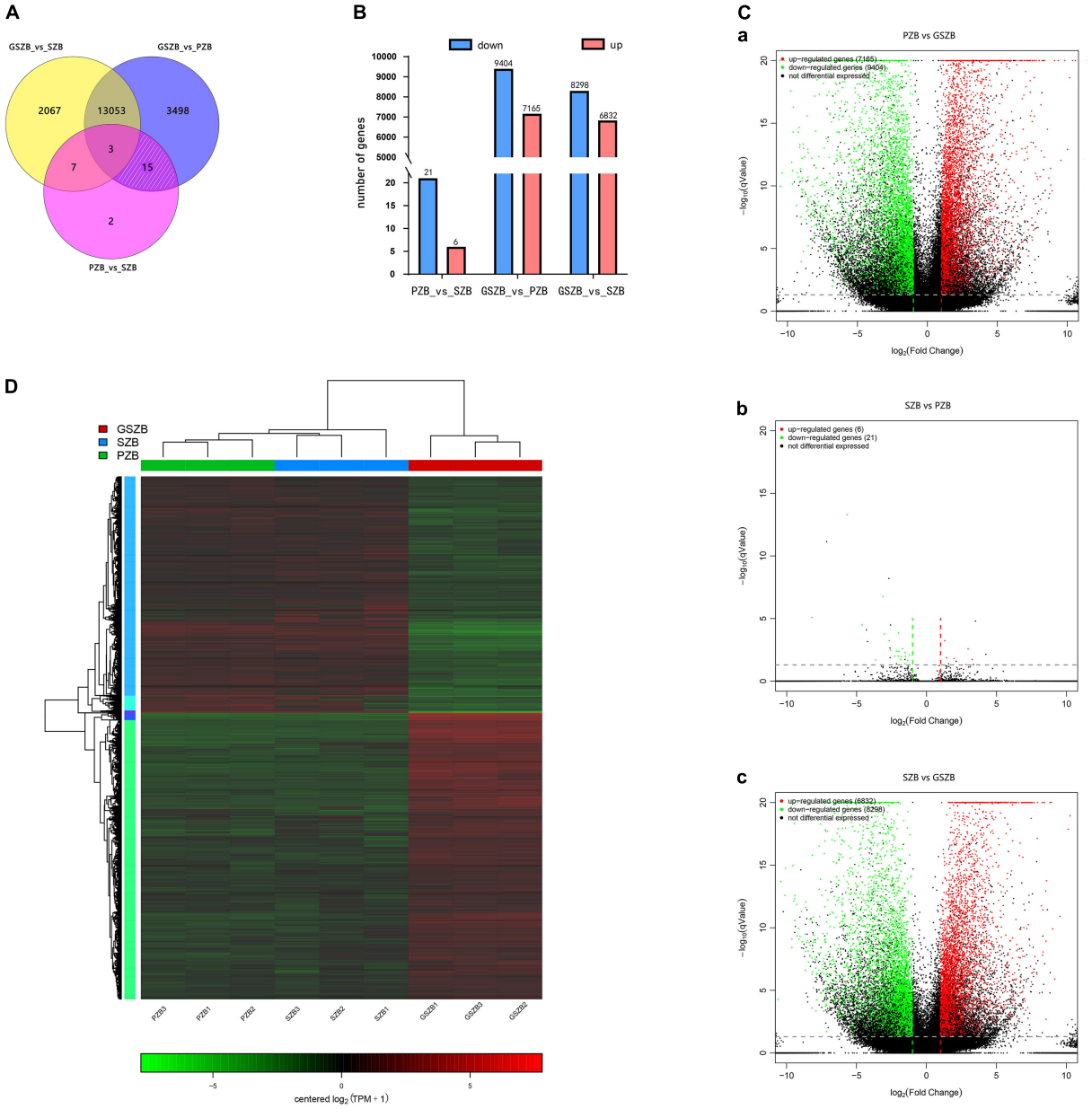


FIGURE 7
(A) Venn diagram of differentially expressed genes (DEGs) among two or three comparisons; **(B)** Total number of upregulated and downregulated DEGs between different groups; **(C)** Volcano plot of DEGs for PZB_vs_GSZB, SZB_vs_PZB and SZB_vs_GSZB; **(D)** Heatmap of DEGs based on hierarchical clustering analysis.

ZbWRKY28, *ZbAZG2*, *ZbLOG5*, *ZbIAA33*, *ZbGh16*, *ZbGh16X1*, *Zb19224c0g1*, *Zb33022c0g3*, *Zb36195c1g2*, *Zb51123c2g3*, *Zb40454c0g1*, and *Zb22746c0g1*) (**Supplementary Table 3**). Among them, the expression of *ZbYABBY2*, *ZbYABBY1*, *ZbYABBY5*, and *ZbWRKY28* were higher at PZB than at SZB and GSZB. The phylogenetic analysis of ZbYABBYs and its homologous proteins was shown in **Figure 8**. It can be found that ZbYABBYs was first clustered with CsYABBYs and CcYABBYs, and its conserved domain was HMG_box_2 (Siegfried et al., 1999; Bartholmes et al., 2012).

Bowman's research found that the YABBY gene family of *Arabidopsis thaliana* comprises six members that probably encode transcriptional regulators. Each member of the family is expressed in a polar manner in one or more above-ground lateral organs, and in every lateral organ at least one family member is expressed (Bowman, 2000). Besides, previous studies

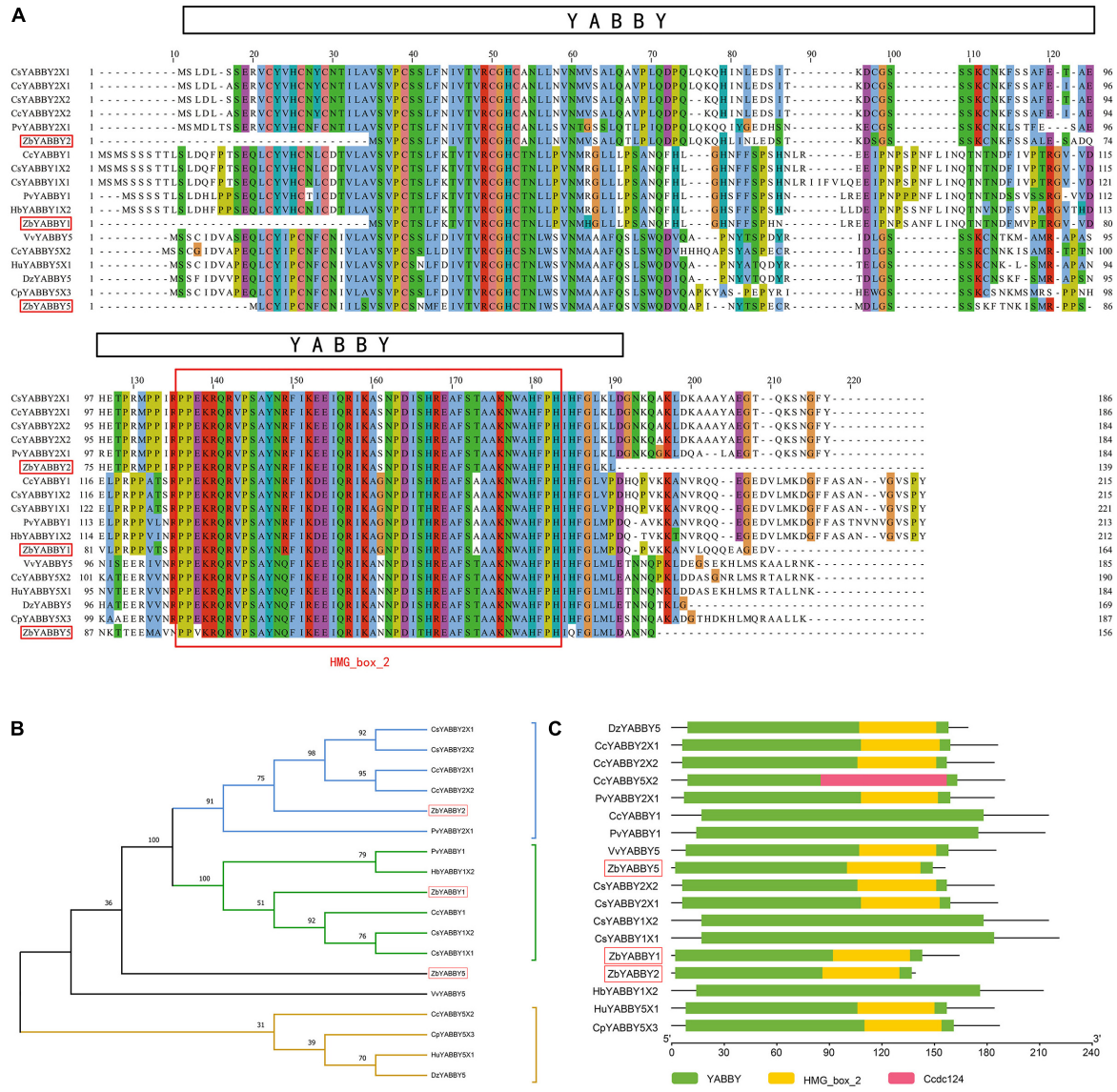


FIGURE 8
Phylogenetic analysis of ZbYABBYs and its homologous proteins. **(A)** Multiple alignments of deduced amino acid sequences of ZbYABBY2, ZbYABBY1, and ZbYABBY5 proteins with other functionally characterized YABBYs; **(B)** Phylogenetic analyses of ZbYABBY2, ZbYABBY1, and ZbYABBY5 in other plants; **(C)** Conserved domain analysis of ZbYABBY2, ZbYABBY1, and ZbYABBY5.

have shown that YABBYs are related to the establishment of polarity in angiosperm lateral organs (Bowman et al., 2002).

The phylogenetic analysis of ZbWRKY28 and its homologous proteins was shown in **Supplementary Figure 6**. The transcription factors of WRKY class have been proved to control trichome, seed coat, and root hair development in *Arabidopsis thaliana* (Eulgem et al., 2000; Johnson et al., 2002). Glandular trichome formation in *Gossypium* spp. and *Rosa chinensis* is also regulated by them (Ma D. et al., 2016; Hibrand Saint-Oyant et al., 2018). Therefore, *ZbYABBY2*, *ZbYABBY1*, *ZbYABBY5*, and *ZbWRKY28* may affect the development of *Zanthoxylum bungeanum* prickles.

The candidate genes of *ZbAZG2*, *ZbLOG5*, and *ZbIAA33* were related to plant hormone signal transduction. The results of multiple alignments, phylogenetic analyses and conserved domain analysis of their deduced amino acid sequences are respectively shown in **Supplementary Figures 7–9**. AZG1 and AZG2 proteins were shown to function as cytokinin transporters which can opening doors for cytokinin trafficking at the ER membrane (Romanov and Schmulling, 2021). Besides, AZGs were found can regulate lateral root emergence in *Arabidopsis thaliana* (Tessi et al., 2021). Kuroha's research suggested that LOGs (LOG1-LOG9) played a pivotal role in regulating cytokinin activity and worked in the direct activation

pathway in *Oryza sativa* shoot meristems (Kuroha et al., 2009; Tokunaga et al., 2012). Lv and his team found that IAA33 overexpression enhanced root distal stem cell differentiation (Audran-Delalande et al., 2012; Lv et al., 2020). In our research, the expression levels of *ZbAZG2*, *ZbIAA33*, and *ZbLOG5* in PZB were significantly higher than in SZB and GSZB. Combined with previous studies, it is speculated that the development of *Zanthoxylum bungeanum* prickles is regulated by plant hormones.

The phylogenetic analysis of ZbGh16Xs and its homologous proteins was shown in **Supplementary Figure 10**. Conserved domain analysis of ZbGh16X1 and ZbGh16 indicated that ZbGh16Xs possess a Glyco_hydro_16 central domain and an XET_C C-terminal domain. The activity of the XET_C domain is characterized by endolytic cleavage of the xyloglucan chain, which transfers the reducing end (donor) of the chain to the non-reducing end (acceptor) of a different xyloglucan chain (Atkinson et al., 2009).

Besides, the Gh16 (glycosyl hydrolase 16) family contains multiple glucanases and endo-acting galactanases (Strohmeier et al., 2004), which are involved in the Endotrans-glucosylation process. The process of endotrans-glucosylation allows cell expansion by temporarily loosening the cell wall in rapidly growing cells, and incorporates newly synthesized xyloglucan chains into the cell wall for reorganization in plants (Nishitani and Tominaga, 1992). Therefore, ZbGh16X1 and ZbGh16 may regulate the process of cell wall thickening and cell lignification. In our study, it was found that the expression levels of ZbGh16X1 and ZbGh16 in PZB were much higher than those in SZB and GSZB. This result explained exactly why the prickle cells had a higher degree of lignification than stem cells (**Figures 4**, **5**).

Metabolomic profiling

A widely targeted metabolomic analysis was performed to produce a metabolic profile. The total ion current (TIC) of all the samples showed high stability, large peak capacity, and good retention time, indicating the reliability and accuracy

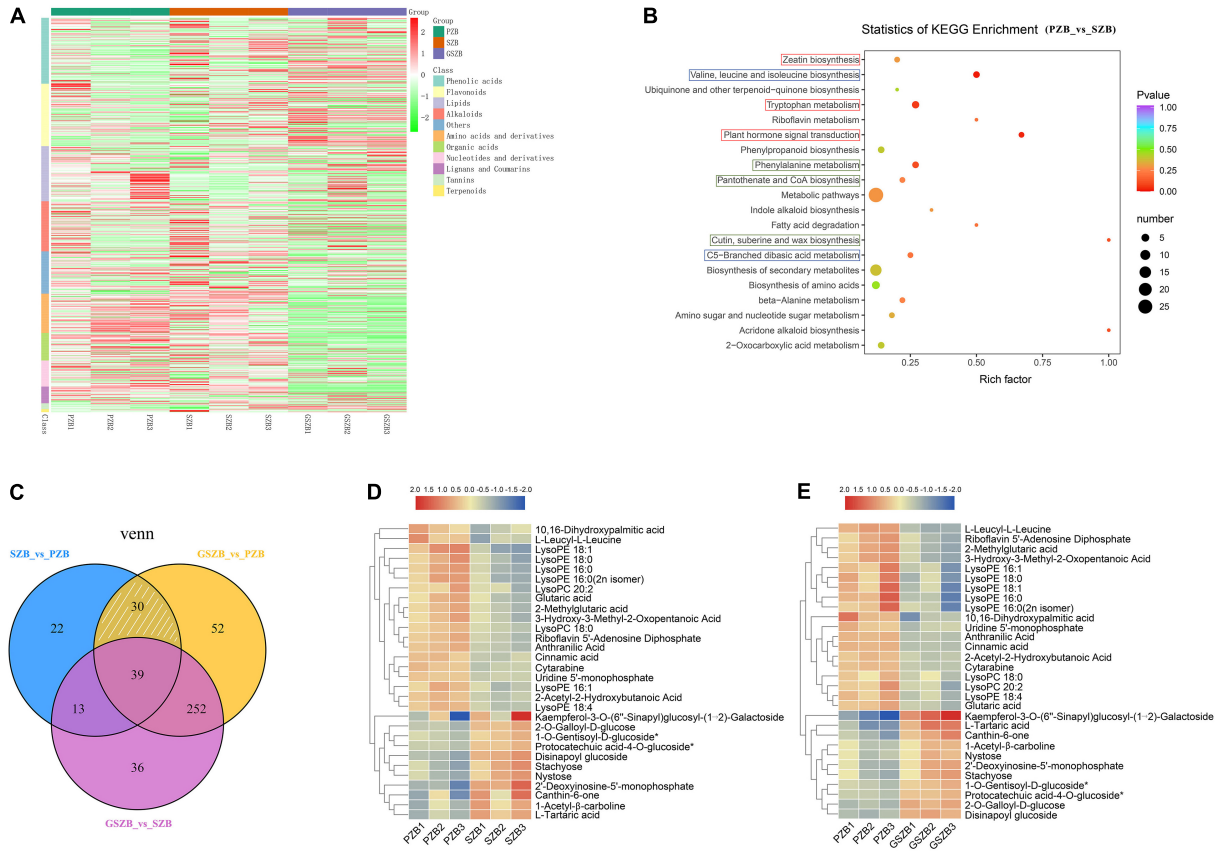


FIGURE 9
Metabolomic profiling results. **(A)** Heatmap of differentially accumulated metabolites (DAMs) based on hierarchical clustering analysis; **(B)** KEGG enrichment results between SZB and PZB. Abscissa represents Rich factor; Ordinate represents the name of pathways. The color of dots represents *P*-value; The size of dots represents the number of DAMs; **(C)** Venn diagram of DAMs among two or three comparisons; **(D)** Heatmap of 30 DAMs between SZB and PZB; **(E)** Heatmap of 30 DAMs between GSZB and PZB.

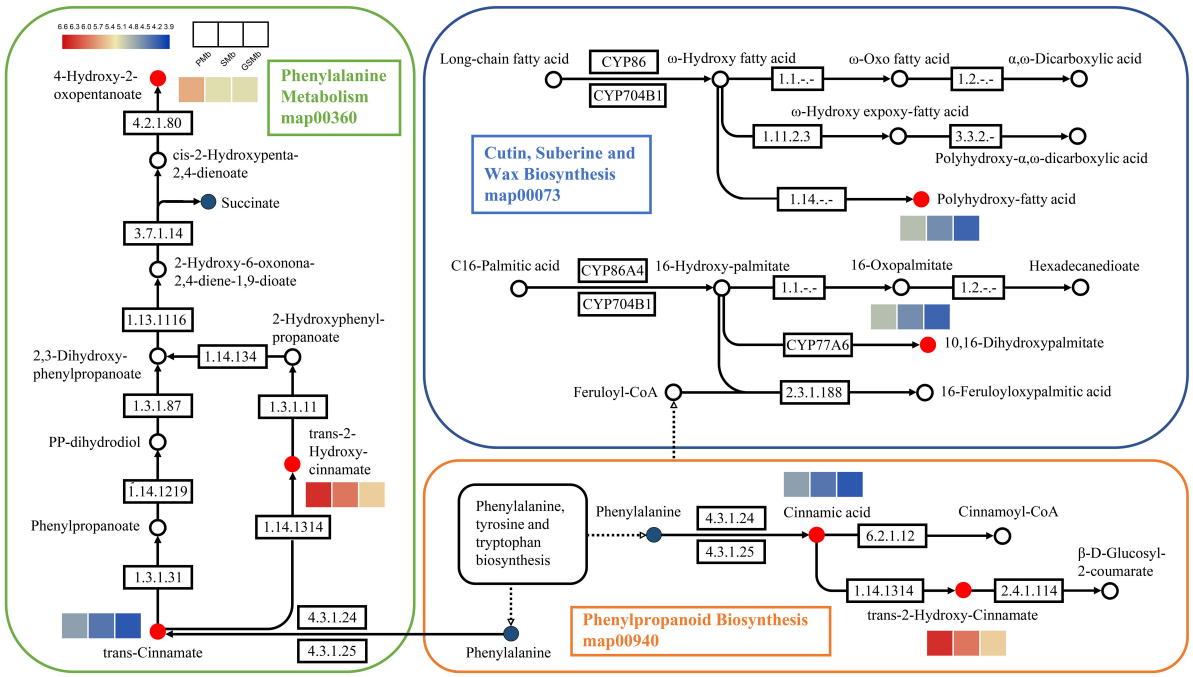


FIGURE 10
Differentially accumulated metabolites (DAMs) in the biosynthetic pathway of young stem bark of *Zanthoxylum bungeanum*. [KEGG map: The red dots represent up-regulation, the green dots represent down-regulation, and the blue dots represent no difference. Heatmap: The color scale from blue (low) to red (high) represents the fold change values].

of the data (**Supplementary Figure 11**). Besides, a total of 802 metabolites were identified in the young stem bark of *Zanthoxylum bungeanum* and were divided into 11 categories, primarily lipids, phenolic acids, lignans and coumarins, just as **Figure 9A** shows (**Supplementary Table 13**).

Identification of differentially accumulated metabolites

The DAMs of each comparison group were selected by setting VIP $\geq$ 1 and fold change $\geq$ 2 or $\leq$ 0.5 as thresholds. As a results, 373, 340, and 104 DAMs were detected between GSZB_vs_PZB, GSZB_vs_SZB, and SZB_vs_PZB, respectively (**Supplementary Tables 14–16** and **Supplementary Figures 12A–C**).

In order to screen for metabolites that affect prickle development, the DAMs shared by PZB_vs_SZB and PZB_vs_GSZB were selected, and the DAMs of SZB_vs_GSZB were excluded. Finally, 30 metabolites were selected as candidate metabolites (white slashed area in **Figure 9C**). Among them, the content of metabolites such as LysoPE, LysoPC, ainnamic acid, and 10,16-dihydroxypalmitic acid in PZB variety was higher than that in SZB and GSZB varieties. Besides, the content of metabolites such as stachyose, nystose, 2′-Deoxyinosine-5′-monophosphate, glutaric acid and Protocatechuic acid-4-O-glucoside in PZB variety was lower than that in SZB and GSZB varieties (**Figures 9D,E**).

The 10,16-dihydroxypalmitic acid is a precursor for the synthesis of cutin (Kolattukudy and Walton, 1972; Croteau and Kolattukudy, 1974), and its content in PZB was significantly higher than that in GSZB and SZB, indicating that prickle cells were lignified earlier than stem cells, which fits well with the results observed in transmission electron microscope (**Figures 5A–C**). In addition, Gao's research found that increased levels of the lipid metabolites such as lysoPC and lysoPE can help *Arabidopsis thaliana* resist extreme conditions (Wang et al., 2006; Gao et al., 2010). This phenomenon also found in *Zanthoxylum bungeanum*, which tends to grow more prickles under extreme drought conditions to protect themselves. And this may be the reason for why the content of lysoPC and lysoPE was higher in PZB than in SZB and GSZB varieties.

Enrichment analysis of the differentially accumulated metabolites

The KEGG pathway enrichment analysis of DAMs from each comparison group showed that they were enriched in multiple metabolic pathways, including cutin,

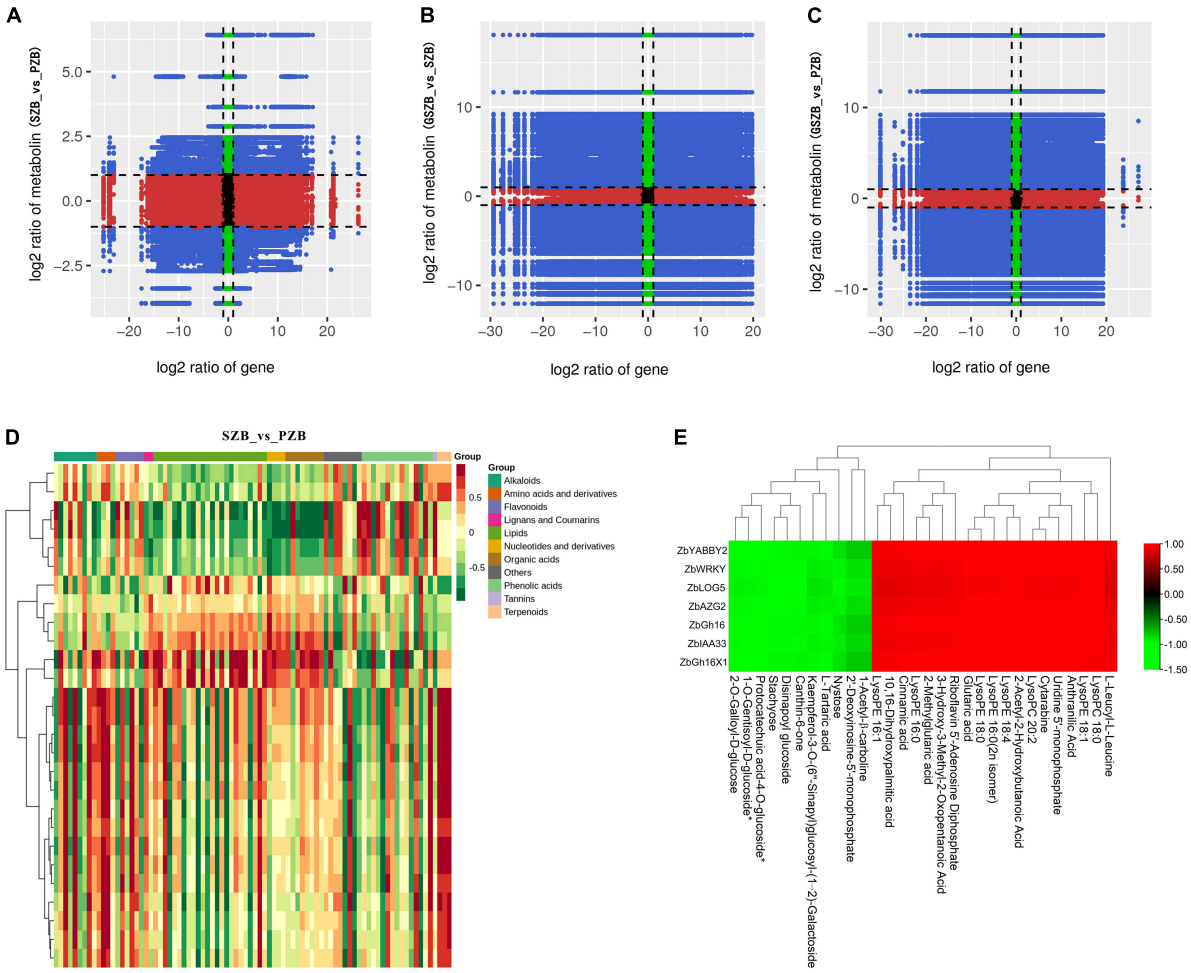


FIGURE 11
(A–C) Nine-quadrant diagram of correlations of differentially expressed genes (DEGs) and differentially accumulated metabolites (DAMs); (From left to right, top to bottom, are 1 to 9 quadrants respectively); **(D)** The correlation heatmap of DEGs and DAMs between SZB and PZB libraries; **(E)** The clustering heatmap of candidate DEGs and candidate DAMs.

suberine and wax biosynthesis; plant hormone signal transduction; zeatin biosynthesis; valine, leucine and isoleucine biosynthesis; etc. (Figure 9B and Supplementary Figures 12D–F).

Among them, the pathways of map00360, map00073, and map00940 were related to the biosynthesis of cutin, suberine and wax (Figure 10). And it can be found that the content of DAMs, such as 4-hydroxy-2-oxopentanoate, trans-2-hydroxy-cinnamate, trans-cinnamate, polyhydroxy-fatty acid, 10,16-dihydroxypalmitate, cinnamic acid, was significantly higher in PZB variety than in SZB and GSZB varieties. In addition, the pathways of map00380, map04075 and map00908 were related to plant hormone signal transduction, while the pathways of map0066 and map00290 were related to the biosynthesis of valine, leucine, and isoleucine. Their results were shown in Supplementary Figures 13, 14 respectively.

Correlation analysis between transcriptome and metabolome data

The accumulation of metabolites is controlled by many exogenous and endogenous factors. Therefore, metabolomics and transcriptomics need to be correlated to identify functional genes and metabolic pathways.

The relationship between DEGs and DAMs with a PCC (Pearson correlation coefficient) $\geq$ 0.8 was displayed in nine-quadrant graphs (Figures 11A–C). As well as, the correlation heatmap of DEGs and DAMs between SZB and PZB libraries was shown in Figure 11D (The similar results of GSZB_vs_SZB and GSZB_vs_PZB were shown in Supplementary Figures 15, 16).

For the 7 candidate DEGs and the 30 candidate DAMs, a separate clustering heatmap was drawn to intuitively reflect the relationship between them (Figure 11E). The genes

(*ZbYABBY2*, *ZbWRKY*, *ZbLOG5*, *ZbAZG2*, *ZbGh16*, *ZbIAA33*, and *ZbGh16X1*) were positively correlated with metabolites such as LysoPE, LysoPC, 10,16-dihydroxypalmitic acid, and negatively correlated with metabolites such as stachyose, nystose, canthin-6-one.

Discussion

This study was designed to uncover the morphological characteristics and molecular mechanism of prickle development of *Zanthoxylum bungeanum*. To this end, the tissues of "Wild Prickly *Zanthoxylum bungeanum* (PZB)," "Wild Smooth *Zanthoxylum bungeanum* (SZB)," and "Grafted Smooth *Zanthoxylum bungeanum* (GSZB)" were collected for experiments. And the techniques of Fourier-transform infrared spectroscopy (FTIR), optical microscopy, transmission electron microscopy (TEM), transcriptomic, quantitative real-time PCR (qRT-PCR), and metabolomic and multi-omics analysis were used for investigation.

Fourier-transform infrared spectroscopy (FTIR) results revealed that the absorption spectra of prickles and stems were different. Prickles had band at 1617 and 1110 cm^{-1}, while stems had strong band at 3319 and 1999 cm^{-1}. According to the band correspondence, the prickles of *Zanthoxylum bungeanum* contain unique aromatic compounds (Breub et al., 2010; Lupoi et al., 2015; Reyes-Rivera and Terrazas, 2017), while the stems contain compounds with -NH_2 or -NH bonds.

The results of morphological studies showed that the prickles of *Zanthoxylum bungeanum* had no vascular bundle. The growth directions of prickle and stem cells were at a vertical angle, and there was a resembling abscission zone (RAZ) between them. The existence of RAZ explains well why prickles tend to peel off the stems. Besides, prickles lignified earlier than stems, and the deposits in vacuoles of prickle cells were obviously more than stem cells. Combining with transcriptome, metabolome and FTIR results, these deposits may be compounds related to xyloglucan metabolism (Nishitani and Tominaga, 1992; Strohmeier et al., 2004; Atkinson et al., 2009) and cutin synthesis (Kolattukudy and Walton, 1972; Croteau and Kolattukudy, 1974; Wang et al., 2006; Gao et al., 2010).

Through transcriptomic analysis and qRT-PCR validation, 9 DEGs (*ZbYABBY2*, *ZbYABBY1*, *ZbYABBY5*, *ZbWRKY*, *ZbLOG5*, *ZbAZG2*, *ZbGh16*, *ZbIAA33*, and *ZbGh16X1*) related to prickle development were screened and validated. Their functional annotation, phylogenetic analyses, multiple sequence alignment, and conserved domain analyses revealed that *ZbYABBY1*, *ZbYABBY2*, *ZbYABBY5*, and *ZbWRKY* were involved in the shoot apical meristem development and the flower development (Siegfried et al., 1999; Bowman, 2000; Eulgem et al., 2000; Bowman et al., 2002; Johnson et al., 2002; Bartholmes et al., 2012; Ma D. et al., 2016; Hibrand Saint-Oyant et al., 2018); *ZbLOG5*, *ZbAZG2*, and *ZbIAA33* were involved in the plant hormone signal transduction and the response to gibberellin (Kuroha et al., 2009; Audran-Delalande et al., 2012; Tokunaga et al., 2012; Lv et al., 2020; Romanov and Schmulling, 2021; Tessi et al., 2021); *ZbGh16* and *ZbGh16X1* were involved in the cutin, suberin and wax biosynthesis (Nishitani and Tominaga, 1992; Strohmeier et al., 2004; Atkinson et al., 2009).

Metabolomic analysis identified 802 compounds, and screened out 30 candidate metabolites related to the development of prickles. The metabolites such as LysoPE, LysoPC and 10,16-dihydroxypalmitic were related to the biosynthesis of cutin, suberine and wax; The metabolites such as Indole-3-acetate, N6-(delta2-isopentenyl)-adenine and tryptamine were related to plant hormone signal transduction; The metabolites such as (S)-2-acetolactate, (R)-3-hydroxy-3-methyl-2-oxopentanoate and (S)-citramalate were related to valine, leucine and isoleucine biosynthesis (Kolattukudy and Walton, 1972; Croteau and Kolattukudy, 1974; Wang et al., 2006; Gao et al., 2010). In addition, through multi-omics analysis, DEGs and DAMs with PCC (Pearson correlation coefficient) ≥ 0.8 were screened out.

This study elucidated the developmental mechanism of prickles at morphological and molecular levels. The candidate genes and metabolites that affected prickle development were screened out, and the morphological characteristics of prickles were observed in detail. In conclusion, this study filled the gap in the research field of *Zanthoxylum bungeanum* prickles and provided a theoretical basis for the breeding of non-prickle *Zanthoxylum bungeanum*.

Data availability statement

The original contributions presented in this study are publicly available. This data can be found here: NCBI, PRJNA752915.

Author contributions

SL and KS conceived and designed the study. SL, KS, JS, BS, JH, and TZ participated in the coordination of the study. KS performed the experimental measurements, processed the experimental data, interpreted the data, and drafted and revised the manuscript. SL, KS, and JS reviewed and revised the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Central Finance Forestry and Grass Technology Promotion Demonstration Project (blue[2021]TG13) and the Xianyang Key Research and Development Project (2020k01-35).

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.950084/full#supplementary-material

References

Airoldi, C. A., and Glover, B. J. (2020). Evo-Devo: Tinkering with the Stem Cell Niche to Produce Thorns. *Curr. Biol.* 30:R873–R875. doi: 10.1016/j.cub.2020.06.019

Appelhans, M. S., Reichelt, N., Groppo, M., Paetzold, C., and Wen, J. (2018). Phylogeny and biogeography of the pantropical genus Zanthoxylum and its closest relatives in the proto-Rutaceae group (Rutaceae). *Mol. Phylogenet. Evol.* 126, 31–44. doi: 10.1016/j.ympev.2018.04.013

Atkinson, R. G., Johnston, S. L., Yauk, Y.-K., Sharma, N. N., and Schröder, R. (2009). Analysis of xyloglucan endotransglucosylase/hydrolase (XTH) gene families in kiwifruit and apple. *Postharvest Biol. Technol.* 51, 149–157. doi: 10.1016/j.postharvbio.2008.06.014

Audran-Delalande, C., Bassa, C., Mila, I., Regad, F., Zouine, M., and Bouzayen, M. (2012). Genome-wide identification, functional analysis and expression profiling of the Aux/IAA gene family in tomato. *Plant Cell Physiol* 53, 659–672. doi: 10.1093/pcp/pcs022

Balkunde, R., Pesch, M., and Hülskamp, M. (2010). Trichome patterning in *Arabidopsis thaliana*: From genetic to molecular models. *Plant Dev.* 91, 299–321. doi: 10.1016/S0070-2153(10)91010-7

Barba, P., Loughner, R., Wentworth, K., Nyrop, J. P., Loeb, G. M., and Reisch, B. I. (2019). A QTL associated with leaf trichome traits has a major influence on the abundance of the predatory mite Typhlodromus pyri in a hybrid grapevine population. *Hortic. Res.* 6:87. doi: 10.1038/s41438-019-0169-8

Bartholmes, C., Hidalgo, O., and Gleissberg, S. (2012). Evolution of the YABBY gene family with emphasis on the basal eudicot Eschscholzia californica (Papaveraceae). *Plant Biol.* 14, 11–23. doi: 10.1111/j.1438-8677.2011.00486.x

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Bourke, P. M., Gitonga, V. W., Voorrips, R. E., Visser, R. G. F., Krens, F. A., and Maliepaard, C. (2018). Multi-environment QTL analysis of plant and flower morphological traits in tetraploid rose. *Theor. Appl. Genet.* 131, 2055–2069. doi: 10.1007/s00122-018-3132-4

Bowman, J. L. (2000). The YABBY gene family and abaxial cell fate. *Curr. Opin. Plant Biol.* 3, 17–21. doi: 10.1016/S1369-5266(99)00035-7

Bowman, J. L., Eshed, Y., and Baum, S. F. (2002). Establishment of polarity in angiosperm lateral organs. *Trends Genet.* 18, 134–141. doi: 10.1016/S0168-9525(01)02601-4

Breub, M., Cazacu, G., and Chirila, O. (2010). Pyrolysis of lignin - A potential method for obtaining chemicals and/or fuels. *Cellul. Chem. Technol.* 45, 43–50.

Chen, W., Gong, L., Guo, Z., Wang, W., Zhang, H., Liu, X., et al. (2013). A novel integrated method for large-scale detection, identification, and quantification of widely targeted metabolites: application in the study of rice metabolomics. *Mol. Plant* 6, 1769–1780. doi: 10.1093/mp/sst080

Chen, X., Wang, W., Wang, C., Liu, Z., Sun, Q., and Wang, D. (2019). Quality evaluation and chemometric discrimination of Zanthoxylum bungeanum Maxim leaves based on flavonoids profiles, bioactivity and HPLC-fingerprint in a common garden experiment. *Ind. Crops Prod.* 134, 225–233. doi: 10.1016/j.indcrop.2019.04.017

Chen, Y., Zhang, R., Song, Y., He, J., Sun, J., Bai, J., et al. (2009). RRLC-MS/MS-based metabonomics combined with in-depth analysis of metabolic correlation network: finding potential biomarkers for breast cancer. *Analyst* 134, 2003–2011. doi: 10.1039/b907243h

Chien, J. C., and Sussex, L. M. (1996). Differential Regulation of Trichome Formation on the Adaxial and Abaxial Leaf Surfaces by Gibberellins and Photoperiod in Arabidopsis fhaliana (L.) Heynh. *Plant Physiol.* 111, 1321–1328. doi: 10.1104/pp.111.4.1321

Chopra, D., Mapar, M., Stephan, L., Albani, M. C., Deneer, A., Coupland, G., et al. (2019). Genetic and molecular analysis of trichome development in Arabis alpina. *Proc. Natl. Acad. Sci. U.S.A* 116, 12078–12083. doi: 10.1073/pnas.1819440116

Coverdale, T. C. (2020). Defence emergence during early ontogeny reveals important differences between spines, thorns and prickles. *Ann. Bot.* 124:iii–iv. doi: 10.1093/aob/mcz189

Croteau, R., and Kolattukudy, P. E. (1974). Biosynthesis of Hydroxyfatty Acid Polymers. Enzymatic Synthesis of Cutin from Monomer Acids by Cell-Free Preparations from the Epidermis of Vicia faba Leaves. *Biochemistry* 13, 3193–3202. doi: 10.1021/bi00712a030

de la Fuente, V., Rufo, L., Rodríguez, N., Franco, A., and Amils, R. (2017). Comparison of iron localization in wild plants and hydroponic cultures of Imperata cylindrica (L.) P. *Plant and Soil* 418, 25–35. doi: 10.1007/s11104-017-3251-8

Debener, T., and Linde, M. (2009). Exploring Complex Ornamental Genomes: The Rose as a Model Plant. *Crit. Rev. Plant Sci.* 28, 267–280. doi: 10.1080/07352680903035481

Dominguez, E., Heredia-Guerrero, J. A., and Heredia, A. (2011). The biophysical design of plant cuticles: an overview. *N. Phytol.* 189, 938–949. doi: 10.1111/j.1469-8137.2010.03553.x

Eulgem, T., Rushton, P. J., Robatzek, S., and Somssich, I. E. (2000). The WRKY superfamily of plant transcription factors. *Trends Plant Sci.* 5, 199–206. doi: 10.1016/S1360-1385(00)01600-9

Fei, X., Li, J., Kong, L., Hu, H., Tian, J., Liu, Y., et al. (2020). miRNAs and their target genes regulate the antioxidant system of Zanthoxylum bungeanum under drought stress. *Plant Physiol Biochem.* 150, 196–203. doi: 10.1016/j.plaphy.2020.01.040

Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., et al. (2016). The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* 44:D279–D285. doi: 10.1093/nar/gkv1344

Fraga, C. G., Clowers, B. H., Moore, R. J., and Zink, E. M. (2010). Signature-Discovery Approach for Sample Matching of a Nerve-Agent Precursor Using Liquid Chromatography-Mass Spectrometry. XCMS, and Chemometrics. *Anal. Chem.* 82, 4165–4173. doi: 10.1021/ac1003568

Fu, A., Wang, Q., Mu, J., Ma, L., Wen, C., Zhao, X., et al. (2021). Combined genomic, transcriptomic, and metabolomic analyses provide insights into chayote (Sechium edule) evolution and fruit development. *Hortic. Res.* 8:35. doi: 10.1038/s41438-021-00487-1

Gao, W., Li, H. Y., Xiao, S., and Chye, M. L. (2010). Acyl-CoA-binding protein 2 binds lysophospholipase 2 and lysoPC to promote tolerance to cadmium-induced oxidative stress in transgenic Arabidopsis. *Plant J.* 62, 989–1003. doi: 10.1111/j.1365-313X.2010.04209.x

Glover, B. J., Rodriguez, M. P., and Martin, C. (1998). Development of several epidermal cell types can be specified by the same MYB-related plant transcription factor. *Development* 125, 3497–3508. doi: 10.1242/dev.125.17.3497

Grubb, P. J. (1992). A positive distrust in simplicity - lessons from plant defences and from competition among plants and among animals. *J. Ecol.* 80, 585–610. doi: 10.2307/2260852

Guzman, P., Fernandez, V., Khayet, M., Garcia, M. L., Fernandez, A., and Gil, L. (2014). Ultrastructure of plant leaf cuticles in relation to sample preparation as observed by transmission electron microscopy. *Sci. World J.* 2014:963921. doi: 10.1155/2014/963921

Herman, P. L., and Marks, M. D. (1989). Trichome Development in Arabidopsis thaliana. II. lsolation and Complementation of the GLABROUSI Gene. *Plant Cell* 1, 1051–1055. doi: 10.2307/3869022

Hibrand Saint-Oyant, L., Ruttink, T., Hamama, L., Kirov, I., Lakhwani, D., Zhou, N. N., et al. (2018). A high-quality genome sequence of Rosa chinensis to elucidate ornamental traits. *Nat. Plants* 4, 473–484. doi: 10.1038/s41477-018-0166-1

Hu, Z. Z., Sha, X. M., Huang, T., Zhang, L., Wang, G. Y., and Tu, Z. C. (2021). Microbial transglutaminase (MTGase) modified fish gelatin-gamma-polyglutamic acid (gamma-PGA): Rheological behavior, gelling properties, and structure. *Food Chem.* 348:129093. doi: 10.1016/j.foodchem.2021.129093

Huang, X., Yan, H., Zhai, L., and Yi, Y. (2019). GLABROUS1 from Rosa roxburghii Tratt regulates trichome formation by interacting with the GL3/EGL3 protein. *Gene* 692, 60–67. doi: 10.1016/j.gene.2018.12.071

Huchelmann, A., Boutry, M., and Hachez, C. (2017). Plant Glandular Trichomes: Natural Cell Factories of High Biotechnological Interest. *Plant Physiol.* 175, 6–22. doi: 10.1104/pp.17.00727

Jenks, M. A., Joly, R. J., Peters, P. J., Rich, P. J., Axtell, J. D., and Ashworth, E. N. (1994). Chemically Induced Cuticle Mutation Affecting Epidermal Conductance to Water Vapor and Disease Susceptibility in Sorghum bicolor (L.) Moench. *Plant Physiol.* 105, 1239–1245. doi: 10.1104/pp.105.4.1239

Johnson, C. S., Kolevski, B., and Smyth, D. R. (2002). TRANSPARENT TESTA GLABRA2, a trichome and seed coat development gene of Arabidopsis, encodes a WRKY transcription factor. *Plant Cell* 14, 1359–1375. doi: 10.1105/tpc.001404

Jozefczuk, S., Klie, S., Catchpole, G., Szymanski, J., Cuadros-Inostroza, A., Steinhauser, D., et al. (2010). Metabolomic and transcriptomic stress response of *Escherichia coli*. *Mol. Syst. Biol.* 6:364. doi: 10.1038/msb.2010.18

Kanehisa, M., and Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 28, 27–30. doi: 10.1093/nar/28.1.27

Kellogg, A. A., Branaman, T. J., Jones, N. M., Little, C. Z., and Swanson, J.-D. (2011). Morphological studies of developing Rubus prickles suggest that they are modified glandular trichomes. *Botany* 89, 217–226. doi: 10.1139/b11-008

Khan, S., Zhou, J. L., Ren, L., and Mojiri, A. (2020). Effects of glyphosate on germination, photosynthesis and chloroplast morphology in tomato. *Chemosphere* 258:127350. doi: 10.1016/j.chemosphere.2020.127350

Kirik, V., Lee, M. M., Wester, K., Herrmann, U., Zheng, Z., Oppenheimer, D., et al. (2005). Functional diversification of MYB23 and GL1 genes in trichome morphogenesis and initiation. *Development* 132, 1477–1485. doi: 10.1242/dev.01708

Kolattukudy, P. E., and Walton, T. J. (1972). Structure and Biosynthesis of the Hydroxy Fatty Acids of Cutin in Vicia faba Leaves. *Biochemistry* 11, 1897–1907. doi: 10.1021/bi00760a026

Kuroha, T., Tokunaga, H., Kojima, M., Ueda, N., Ishida, T., Nagawa, S., et al. (2009). Functional analyses of LONELY GUY cytokinin-activating enzymes reveal the importance of the direct activation pathway in Arabidopsis. *Plant Cell* 21, 3152–3169. doi: 10.1105/tpc.109.068676

Li, J., Hui, T., Wang, F., Li, S., Cui, B., Cui, Y., et al. (2015). Chinese red pepper (Zanthoxylum bungeanum Maxim.) leaf extract as natural antioxidants in salted silver carp (Hypophthalmichthys molitrix) in dorsal and ventral muscles during processing. *Food Control* 56, 9–17. doi: 10.1016/j.foodcont.2015.03.001

Li, W., Zhao, Y., Liu, C., Yao, G., Wu, S., Hou, C., et al. (2012). Callose deposition at plasmodesmata is a critical factor in restricting the cell-to-cell movement of Soybean mosaic virus. *Plant Cell Rep.* 31, 905–916. doi: 10.1007/s00299-011-1211-y

Liu, X., Xu, L., Liu, X., Wang, Y., Zhao, Y., Kang, Q., et al. (2020).). Combination of essential oil from Zanthoxylum bungeanum Maxim. and a microemulsion system: Permeation enhancement effect on drugs with different lipophilicity and its mechanism. *J. Drug Deliv. Sci. Technol.* 55:101309. doi: 10.1016/j.jddst.2019.101309

Lupoi, J. S., Singh, S., Parthasarathi, R., Simmons, B. A., and Henry, R. J. (2015). Recent innovations in analytical methods for the qualitative and quantitative assessment of lignin. *Renew. Sustain. Energy Rev.* 49, 871–906. doi: 10.1016/j.rser.2015.04.091

Lv, B., Yu, Q., Liu, J., Wen, X., Yan, Z., Hu, K., et al. (2020). Non-canonical AUX/IAA protein IAA33 competes with canonical AUX/IAA repressor IAA5 to negatively regulate auxin signaling. *EMBO J.* 39:e101515. doi: 10.15252/embj.2019101515

Ma, D., Hu, Y., Yang, C., Liu, B., Fang, L., Wan, Q., et al. (2016). Genetic basis for glandular trichome formation in cotton. *Nat. Commun.* 7:10456. doi: 10.1038/ncomms10456

Ma, Z. Y., Wen, J., Ickert-Bond, S. M., Chen, L. Q., and Liu, X. Q. (2016). Morphology, Structure, and Ontogeny of Trichomes of the Grape Genus (Vitis, Vitaceae). *Front. Plant Sci.* 7:704. doi: 10.3389/fpls.2016.00704

Ma, W.-T., Lu, M., Ludlow, R. A., Wang, D.-J., Zeng, J.-W., and An, H.-M. (2021). Contrastive analysis of trichome distribution, morphology, structure, and associated gene expression reveals the formation factors of different trichome types in two commercial Rosa species. *Sci. Hortic.* 285:110131. doi: 10.1016/j.scienta.2021.110131

Ma, Y., Li, X., Hou, L.-X., and Wei, A.-Z. (2019). Extraction solvent affects the antioxidant, antimicrobial, cholinesterase and HepG2 human hepatocellular carcinoma cell inhibitory activities of Zanthoxylum bungeanum pericarps and the major chemical components. *Ind. Crops Prod.* 142:111872. doi: 10.1016/j.indcrop.2019.111872

Marchler-Bauer, A., Zheng, C., Chitsaz, F., Derbyshire, M. K., Geer, L. Y., Geer, R. C., et al. (2013). CDD: conserved domains and protein three-dimensional structure. *Nucleic Acids Res.* 41:D348–D352. doi: 10.1093/nar/gks1243

Matias-Hernandez, L., Aguilar-Jaramillo, A. E., Osnato, M., Weinstain, R., Shani, E., Suarez-Lopez, P., et al. (2016). TEMPRANILLO Reveals the Mesophyll as Crucial for Epidermal Trichome Formation. *Plant Physiol.* 170, 1624–1639. doi: 10.1104/pp.15.01309

Munien, P., Naidoo, Y., and Naidoo, G. (2015). Micromorphology, histochemistry and ultrastructure of the foliar trichomes of Withania somnifera (L.) Dunal (Solanaceae). *Planta* 242, 1107–1122. doi: 10.1007/s00425-015-2341-1

Naidoo, Y., Karim, T., Heneidak, S., Sadashiva, C. T., and Naidoo, G. (2012). Glandular trichomes of Ceratotheca triloba (Pedaliaceae): morphology, histochemistry and ultrastructure. *Planta* 236, 1215–1226. doi: 10.1007/s00425-012-1671-5

Nishitani, K., and Tominaga, R. (1992). Endo-xyloglucan transferase, a novel class of glycosyltransferase that catalyzes transfer of a segment of xyloglucan molecule to another xyloglucan molecule. *J. Biol. Chem.* 267, 21058–21064. doi: 10.1016/S0021-9258(19)36797-3

Okonechnikov, K., Conesa, A., and Garcia-Alcalde, F. (2016). Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics* 32, 292–294. doi: 10.1093/bioinformatics/btv566

Payne, T., Clement, J., Arnold, D., and Lloyd, A. (1999). Heterologous myb genes distinct from GL1 enhance trichome production when overexpressed in Nicotiana tabacum. *Development* 126, 671–682. doi: 10.1242/dev.126.4.671

Perazza, D., Vachon, G., and Herzog, M. (1998). Gibberellins Promote Trichome Formation by Up-Regulating GLABROUS1 in Arabidopsis. *Plant Physiol.* 117, 375–383. doi: 10.1104/pp.117.2.375

Pereira, C. (2018). *Plant Vacuolar Traffcking*. New York, NY: Humana. doi: 10.1007/978-1-4939-7856-4

Qin, G., Liu, C., Li, J., Qi, Y., Gao, Z., Zhang, X., et al. (2020). Diversity of metabolite accumulation patterns in inner and outer seed coats of pomegranate: exploring their relationship with genetic mechanisms of seed coat development. *Hortic. Res.* 7:10. doi: 10.1038/s41438-019-0233-4

Rajapakse, S., and Arumuganathan, K. (2001). Two genetic linkage maps of tetraploid roses. *Theor. Appl. Genet.* 103, 575–583. doi: 10.1007/PL00002912

Reyes-Rivera, J., and Terrazas, T. (2017). Lignin Analysis by HPLC and FTIR. *Methods Mol. Biol.* 1544, 193–211. doi: 10.1007/978-1-4939-6722-3_14

Romanov, G. A., and Schmulling, T. (2021). Opening Doors for Cytokinin Trafficking at the ER Membrane. *Trends Plant Sci.* 26, 305–308. doi: 10.1016/j.tplants.2021.02.006

Sheng, P., Zhou, H., Liu, J., and Jiang, H. (2020). Some like it hot: Sichuan pepper (Zanthoxylum bungeanum) and other spices from a late Bronze Age kingdom (Chu State) in Hubei, China. *Archaeol. Anthropol. Sci.* 12:249. doi: 10.1007/s12520-020-01201-3

Siegfried, K. R., Eshed, Y., Baum, S. F., Otsuga, D., Drews, G. N., et al. (1999). Members of the YABBY gene family specify abaxial cell fate in Arabidopsis. *Development* 126, 4117–4128. doi: 10.1242/dev.126.18.4117

Simcha, L. Y. (2016). *Defensive (anti-herbivory) Coloration in Land Plants*. Cham: Springer.

Simpson, M. G. (2019). *Plant Systematics*. Amsterdam: Elseiver.

Steeves, T. A., and Sussex, I. M. (1989). *Patterns in Plant Development*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511626227

Strohmeier, M., Hrmova, M., Fischer, M., Harvey, A. J., Fincher, G. B., and Pleiss, J. (2004). Molecular modeling of family GH16 glycoside hydrolases: potential roles for xyloglucan transglucosylases/hydrolases in cell wall modification in the poaceae. *Protein Sci.* 13, 3200–3213. doi: 10.1110/ps.04828404

Su, K., Zheng, T., Chen, H., Zhang, Q., and Liu, S. (2020). Climate Effects on Flavonoid Content of Zanthoxylum bungeanum Leaves in Different Development Stages. *Food Sci. Technol. Res.* 26, 805–812. doi: 10.3136/fstr.26.805

Sun, L., Yu, D., Wu, Z., Wang, C., Yu, L., Wei, A., et al. (2019). Comparative Transcriptome Analysis and Expression of Genes Reveal the Biosynthesis and Accumulation Patterns of Key Flavonoids in Different Varieties of Zanthoxylum bungeanum Leaves. *J. Agric. Food Chem.* 67, 13258–13268. doi: 10.1021/acs.jafc.9b05732

Sun, X., Zhang, D., Zhao, L., Shi, B., Xiao, J., Liu, X., et al. (2020). Antagonistic interaction of phenols and alkaloids in Sichuan pepper (Zanthoxylum bungeanum) pericarp. *Indus. Crops Prod.* 152:112551. doi: 10.1016/j.indcrop.2020.112551

Sussex, I. M., and Kerk, N. M. (2001). The evolution of plant architecture. *Curr. Opin. Plant Biol.* 4, 33–37. doi: 10.1016/S1369-5266(00)00132-1

Szymanski, D. B., Jilk, R. A., Pollock, S. M., and Marks, M. D. (1998). Control of GL2 expression in Arabidopsis leaves and trichomes. *Development* 125, 1161–1171. doi: 10.1242/dev.125.7.1161

Tang, W., Xie, Q., Guan, J., Jin, S., and Zhao, Y. (2014). Phytochemical profiles and biological activity evaluation of Zanthoxylum bungeanum Maxim seed against asthma in murine models. *J. Ethnopharmacol.* 152, 444–450. doi: 10.1016/j.jep.2014.01.013

Tatusov, R. L., Galperin, M. Y., Natale, D. A., and Koonin, E. V. (2000). The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* 28, 33–36. doi: 10.1093/nar/28.1.33

Teimouri, S., Dekiwadia, C., and Kasapis, S. (2021). Decoupling diffusion and macromolecular relaxation in the release of vitamin B6 from genipin-crosslinked whey protein networks. *Food Chem.* 346:128886. doi: 10.1016/j.foodchem.2020.128886

Tessi, T. M., Brumm, S., Winklbauer, E., Schumacher, B., Pettinari, G., Lescano, I., et al. (2021). Arabidopsis AZG2 transports cytokinins in vivo and regulates lateral root emergence. *New Phytol.* 229, 979–993. doi: 10.1111/nph.16943

Tokunaga, H., Kojima, M., Kuroha, T., Ishida, T., Sugimoto, K., Kiba, T., et al. (2012). Arabidopsis lonely guy (LOG) multiple mutants reveal a central role of the LOG-dependent pathway in cytokinin activation. *Plant J.* 69, 355–365. doi: 10.1111/j.1365-313X.2011.04795.x

UniProt, C. (2015). UniProt: a hub for protein information. *Nucleic Acids Res.* 43:D204–D212. doi: 10.1093/nar/gku989

Wada, T., Tachibana, T., Shimura, Y., and Okada, K. (1997). Epidermal Cell Differentiation in Arabidopsis Determined by a Myb Homolog. *CPC. Sci.* 277, 1113–1116. doi: 10.1126/science.277.5329.1113

Wang, D. J., Lu, M., Ludlow, R. A., Zeng, J. W., Ma, W. T., and An, H. M. (2021). Comparative ultrastructure of trichomes on various organs of Rosa roxburghii. *Microsc. Res. Tech.* 84, 2095–2103. doi: 10.1002/jemt.23765

Wang, L., Fan, W., Zhang, M., Zhang, Q., Li, L., Wang, J., et al. (2019). Antiobesity, Regulation of Lipid Metabolism, and Attenuation of Liver Oxidative Stress Effects of Hydroxy-alpha-sanshool Isolated from Zanthoxylum bungeanum on High-Fat Diet-Induced Hyperlipidemic Rats. *Oxid. Med. Cell Longeg.* 2019:5852494. doi: 10.1155/2019/5852494

Wang, L., Wang, S., and Li, W. (2012). RSeQC: quality control of RNA-seq experiments. *Bioinformatics* 28, 2184–2185. doi: 10.1093/bioinformatics/bts356

Wang, S., Kwak, S. H., Zeng, Q., Ellis, B. E., Chen, X. Y., Schiefelbein, J., et al. (2007). TRICHOMELESS1 regulates trichome patterning by suppressing GLABRA1 in Arabidopsis. *Development* 134, 3873–3882. doi: 10.1242/dev.009597

Wang, X., Li, W., Li, M., and Welti, R. (2006). Profiling lipid changes in plant response to low temperatures. *Physiol. Plant.* 126, 90–96. doi: 10.1111/j.1399-3054.2006.00622.x

Wang, Y., Yang, S. H., Zhong, K., Jiang, T., Zhang, M., Kwan, H. Y., et al. (2020). Network Pharmacology-Based Strategy for the Investigation of the Anti-Obesity Effects of an Ethanolic Extract of Zanthoxylum bungeanum Maxim. *Front. Pharmacol.* 11:572387. doi: 10.3389/fphar.2020.572387

Yan, H., Wu, Z., Liu, Y., Weng, Q., Yi, Y., and Huang, X. (2021). Functional divergence of RrGL3 and RrEGL3 from Rosa roxburghii in mediating trichome development. *Plant Cell Tissue Organ. Culture* 147, 313–324. doi: 10.1007/s11240-021-02125-z

Yang, L. C., Li, R., Tan, J., and Jiang, Z. T. (2013). Polyphenolics composition of the leaves of Zanthoxylum bungeanum Maxim. grown in Hebei, China, and their radical scavenging activities. *J. Agric. Food Chem.* 61, 1772–1778. doi: 10.1021/jf3042825

Yin, L., Karn, A., Cadle-Davidson, L., Zou, C., Underhill, A., Atkins, P., et al. (2021). Fine Mapping of Leaf Trichome Density Revealed a 747-kb Region on Chromosome 1 in Cold-Hardy Hybrid Wine Grape Populations. *Front. Plant Sci.* 12:587640. doi: 10.3389/fpls.2021.587640

Zhang, F., Rossignol, P., Huang, T., Wang, Y., May, A., Dupont, C., et al. (2020). Reprogramming of Stem Cell Activity to Convert Thorns into Branches. *Curr. Biol.* 30:2951-2961.e5. doi: 10.1016/j.cub.2020.05.068

Zhang, Y., Yu, X., Wang, M., Ding, Y., Guo, H., Liu, J., et al. (2021). Hyperoside from Z. bungeanum leaves restores insulin secretion and mitochondrial function by regulating pancreatic cellular redox status in diabetic mice. *Free Radic. Biol. Med.* 162, 412–422. doi: 10.1016/j.freeradbiomed.2020.10.320

Zhang, Z., Shen, P., Liu, J., Gu, C., Lu, X., Li, Y., et al. (2017). In Vivo Study of the Efficacy of the Essential Oil of Zanthoxylum bungeanum Pericarp in Dextran Sulfate Sodium-Induced Murine Experimental Colitis. *J. Agric. Food Chem.* 65, 3311–3319. doi: 10.1021/acs.jafc.7b01323

Zhou, N. N., Tang, K. X., Jeauffre, J., Thouroude, T., Arias, D. C. L., Foucher, F., et al. (2020). Genetic determinism of prickles in rose. *Theor. Appl. Genet.* 133, 3017–3035. doi: 10.1007/s00122-020-03652-7

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 07 September 2022
DOI 10.3389/fpls.2022.971230

OPEN ACCESS

EDITED BY
Neftali Ochoa-Alejo,
Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Mexico

REVIEWED BY
Tirthartha Chattopadhyay,
Bihar Agricultural University Sabour, India
Xuncheng Liu,
South China Botanical Garden (CAS), China

*CORRESPONDENCE
Guoju Chen
gjchen@ scau.edu.cn
Zhangsheng Zhu
zhuzs@scau.edu.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Plant Development and EvoDevo, a section of the journal Frontiers in Plant Science

RECEIVED 16 June 2022
ACCEPTED 09 August 2022
PUBLISHED 07 September 2022

CITATION
Cai Y, Xu M, Liu J, Zeng H, Song J, Sun B, Chen S, Deng Q, Lei J, Cao B, Chen C, Chen M, Chen K, Chen G and Zhu Z (2022) Genome-wide analysis of histone acetyltransferase and histone deacetylase families and their expression in fruit development and ripening stage of pepper (*Capsicum annuum*). *Front. Plant Sci.* 13:971230. doi: 10.3389/fpls.2022.971230

Genome-wide analysis of histone acetyltransferase and histone deacetylase families and their expression in fruit development and ripening stage of pepper (*Capsicum annuum*)

Yutong Cai[1†], Mengwei Xu[1†], Jiarong Liu[1†], Haiyue Zeng[2,3†], Jiali Song[1], Binmei Sun[1], Siqi Chen[1], Qihui Deng[1], Jianjun Lei[1], Bihao Cao[1], Changming Chen[1], Muxi Chen[4], Kunhao Chen[4], Guoju Chen[1*] and Zhangsheng Zhu[1*]

[1]Key Laboratory of Biology and Germplasm Enhancement of Horticultural Crops in South China, Ministry of Agriculture and Rural Areas, College of Horticulture, South China Agricultural University, Guangzhou, China, [2]Peking University Institute of Advanced Agricultural Sciences, Weifang, China, [3]School of Advanced Agricultural Sciences, Peking University, Beijing, China, [4]Guangdong Helinong Seeds Co., Ltd., Shantou, China

The fruit development and ripening process involve a series of changes regulated by fine-tune gene expression at the transcriptional level. Acetylation levels of histones on lysine residues are dynamically regulated by histone acetyltransferases (HATs) and histone deacetylases (HDACs), which play an essential role in the control of gene expression. However, their role in regulating fruit development and ripening process, especially in pepper (*Capsicum annuum*), a typical non-climacteric fruit, remains to understand. Herein, we performed genome-wide analyses of the HDAC and HAT family in the pepper, including phylogenetic analysis, gene structure, encoding protein conserved domain, and expression assays. A total of 30 HAT and 15 HDAC were identified from the pepper genome and the number of gene differentiation among species. The sequence and phylogenetic analysis of CaHDACs and CaHATs compared with other plant HDAC and HAT proteins revealed gene conserved and potential genus-specialized genes. Furthermore, fruit developmental trajectory expression profiles showed that *CaHDAC* and *CaHAT* genes were differentially expressed, suggesting that some are functionally divergent. The integrative analysis allowed us to propose CaHDAC and CaHAT candidates to be regulating fruit development

and ripening-related phytohormone metabolism and signaling, which also accompanied capsaicinoid and carotenoid biosynthesis. This study provides new insights into the role of histone modification mediate development and ripening in non-climacteric fruits.

KEYWORDS

pepper, fruit, development, ripening, HAT and HDAC, histone modification

Introduction

The dynamics of chromatin structure regulate DNA accessibility and DNA-templated processes and affect various biological processes of eukaryotes. The nucleosome is the basic unit of chromatin, and it compacts DNA nearly sevenfold, with about 146 bp of DNA wrapped around a histone octamer (Xu et al., 2015). The histone octamer is composed of two copies of H2A, H2B, H3, and H4 histone proteins (Arents et al., 1991). The unstructured amino-terminal tails and the globular domains determined that histones can be modified by dynamic post-translational modifications, including acetylation, methylation, ubiquitination, SUMOylation, phosphorylation, and less commonly by citrullination and ADP-ribosylation (Bannister and Kouzarides, 2011). Among these, histone acetylation is one of the well-characterized modifications; it has been shown that the acetylation of lysines is highly dynamic and regulated by the opposing action of two families of enzymes, histone acetyltransferases (HATs) and histone deacetylases (HDACs) (Stoll et al., 2018).

HATs and HDACs underlie a mechanism for reversibly modulating chromatin structure and transcriptional regulation (Bannister and Kouzarides, 2011). In plants, HATs have been distinctly divided into four groups, including (1) HAG for HATs of the GNAT (GCN5-related N-terminal acetyltransferases) superfamily; (2) HAM for HATs of the MYST superfamily; (3) HAC for HATs of the CREB-binding protein (CBP) family; (4) HAF for HATs of the TATA-binding protein-associated factor (TAFII250) family (Wang et al., 2014; Li et al., 2022). In parallel, plant HDACs have been classified into three families, including RPD3/HDA1 superfamily (HDA), Silent Information Regulator 2 (SRT), and HD2 (HDT) families (Chen X. et al., 2020). It was reported that changes in the acetylation level are associated with changes in gene expression by different endogenous and exogenous factors (Shen et al., 2015). HATs promote histone acetylation, leading to the generation of loose chromatin structures, facilitating the accessibility of promoters to components of the transcription machinery, and thus leading to transcriptional activation (Bannister and Kouzarides, 2011). Recently, the finding was that coactivator proteins are associated with HATs (Servet et al., 2010). HATs have been reported in the regulations of developmental transitions, integrations of hormone signals, and responses to environmental cues (Loidl, 2004). In contrast, the action of HDACs facilitating histone deacetylation is believed to be associated with a less accessible chromatin conformation (Yun et al., 2011), and corepressor proteins always form the complex with HDACs, resulting in gene silencing (Deng et al., 2022; Hu et al., 2022);. Studies have revealed they play key roles in regulating plant vegetative and reproductive development, stress responses, gene silencing, as well as cell death and cycle (Shen et al., 2015). These studies suggest that HATs and HDACs are essential for transcription regulation.

Fruit development comprises fruit set initiation, growth, maturation, and ripening; phytohormones auxins, gibberellins (GAs), cytokinins, abscisic acid (ABA), and ethylene have been implicated at various stages of the fruit growth cycle (White, 2002). Fruit development consists of cell division and expansion, the former shown to be influenced by auxin signaling, and cell expansion is believed to be synergistic regulation *via* both auxin and GAs (Fenn and Giovannoni, 2021). Fruit ripening requires a sequence of biochemical and physiological transformations. The highly dynamic development trajectories are tightly dependent on phytohormones (ABA and ethylene) crosstalk with epigenetic regulated gene expression (Giovannoni, 2004; Teyssier et al., 2015). Transcription factors started to emerge as recruitment of histone acetylation and deacetylation proteins by direct or indirect interaction to target gene promoters lead to changes in acetylation level, resulting alter gene expression (Servet et al., 2010).

During the fruit ripening process, based on the change of physiologies and in the action of ethylene, it is generally classified into climacteric fruit or non-climacteric fruit (Brumos, 2021). Ethylene plays a clear role in climacteric fruits during fruit ripening, whereas non-climacteric ripening is generally associated with ABA (Fenn and Giovannoni, 2021). Silencing of the HADC gene, *SlHDT3*, delays fruit ripening and suppresses carotenoid accumulation in tomato (Guo et al., 2017b). In tomato and apple, HDACs to deacetylation of ripening-related gene promoters, including ethylene biosynthetic genes to repress the ripening process (Deng et al., 2022; Hu et al., 2022), suggesting acetylation levels are associated with the onset of the climacteric fruit ripening process. However, whether HATs

and HDACs are involved in regulating the non-climacteric fruit ripening process is unclear.

Pepper (*Capsicum* sp.) is a typical non-climacteric fruit. Due to its desired pigments, flavor, and aroma, pepper is cultivated and distributed worldwide (Zhu et al., 2019; Villa-Rivera and Ochoa-Alejo, 2021; Sun et al., 2022), about 54.9 million tons of green pepper and 4.1 million tons of dry pepper (FAO). The development and ripening of pepper fruits mainly include three steps (i.e., fruit set, fruit development, and fruit ripening) which are believed to be tightly associated with auxins, GAs, and ABA (Villa-Rivera and Ochoa-Alejo, 2021). Besides alters in fruit shape, changes in metabolism that occur during the ripening process of peppers cause transformations in flavor, color, texture, and aroma (Jang et al., 2015; Liu et al., 2019). A characteristic quality of pepper fruits is pungency, and they produce and accumulate hot compounds called capsaicinoids, which are alkaloids found exclusively in *Capsicum* species (Sun et al., 2022). *Capsicum* fruit capsaicinoids are genetically determined and depend on the developmental stage, and it occurs in the fruits from the early stages to mature green (Zhu et al., 2019; Arce-Rodríguez et al., 2021; Sun et al., 2022). Pepper ripening is accompanied by an accumulation of carotenoids. The different colors of ripening pepper fruits are due to variations in carotenoid composition and content in the pericarp. The capsanthin is the major red pigment of the fruits and accounts for 80% of the carotenoids in the high red intensity pepper cultivars (Berry et al., 2019). The capsanthin biosynthesis process is strictly switched spatially and temporally, and the carotenoid biosynthetic genes mainly transcript in fruit from the break stage to the mature stage (Kim et al., 2014; Song et al., 2020). Through pepper fruit development and ripening process highly dynamic regulated at the transcription level, whether HATs and HDACs are associated with fruit development and ripening remains unknown.

Compared to Arabidopsis and climacteric fruit tomato, relatively few HATs and HDACs were characterized in other plant species. In contrast to the start to the emergence of HATs and HDACs in controlling climacteric fruit development and ripening (Guo et al., 2017b; Deng et al., 2022; Hu et al., 2022), rare studies reported on the process of these functions in non-climacteric fruits. To obtain more insight into HATs and HDACs associated with non-climacteric fruit development and ripening. In this study, the HAT and HDAC families were identified in the non-climacteric fruit pepper genomes. The phylogenetic relationships, gene structure organization, chromosome distribution, conserved domain analysis, and motif protein composition were studied. Furthermore, the expression analysis during pepper fruit development identified *HAT* and *HDAC* genes that may play important roles in fruit development and ripening. This work provided insights into the function of *HAT* and *HDAC* in non-climacteric fruit development and ripening processes, and some of them may hold promise for fruit trait improvement in peppers.

Materials and methods

CaHATs and CaHDACs sequence analyses and characteristics

CaHATs and *CaHDACs* were retrieved from the "Zunla-1" (*C. annuum*) genome, which is available in the Solanaceae Genomics Network,[1] by using the well-characterized HATs and HDACs from Arabidopsis and tomato as queries to search against the Solanaceae Genomics Network and NCBI database by the BLASTP program. The full-length amino acid sequences (length), theoretical isoelectric point (PI), and molecular weight (MW) instability index of ERF proteins were predicted by using the ExPASy server.[2]

Sequence alignment and phylogenetic analysis

For the sequence and phylogenetic analysis, the protein sequences of HATs and HDACs from four species (i.e., *Capsicum annuum*, *Solanum lycopersicum*, *Arabidopsis thaliana,* and *Oryza sativa*) were adopted for analysis. The CaHAT and CaHDAC amino acid sequences were aligned using MUSCLE with the default parameters, respectively. Maximum likelihood trees were constructed with 5,000 bootstrap replications using IQ-tree after trimming the results of multiple sequence comparisons using trimAL (V1.2).

Conserved motif and gene structure analysis

The conserved motif of CaHATs and CaHDACs was received by submitting protein sequences to the MEME Network[3] with the following parameters: The maximum number of motifs for CaHATs is 30, and CaHDACs is 20. The use of NCBI tools is to discover the determination of conserved protein domains. Then, the results are visualized through TBtools (Chen C. et al., 2020).

Genomic organization of CaHATs and CaHDACs

To determine the physical location of *CaHATs* and *CaHDACs*, the TBtools (Chen C. et al., 2020) was applied to locate the *CaHATs* and *CaHDACs* on "Zunla-1" pepper

1 https://solgenomics.net/
2 http://www.expasy.org/
3 https://meme-suite.org/meme/tools/meme

chromosomes according to their positions given in the genome database (Qin et al., 2014).

Expression analysis of CaHATs and CaHDACs

The pepper fruit nine development stages of "Zunla-1" (Qin et al., 2014) and 11 development stages of "6,421" (Liu et al., 2017) RNA-seq data were retrieved from the database. Heat maps showing the expression patterns of genes were constructed with the R Programming Language. The Pearson correlation coefficient was applied to determine the gene expression relationship.

Plant material

The "59" pepper inbred line with a high yield is resistant to a variety of pathogens and has strong resistance to abiotic stress. The seeds were sowed at the nursery site, and 35-day-old seedlings were transplanted into 35-cm non-woven pots. The plants were grown in a greenhouse with a daily temperature of 25–27°C, a nighttime temperature of 20–22°C, relative humidity of 60%, 16/8-h light/dark cycle, and light intensity of 6,500 Lux. During the pepper fruit was at the designed development stages, the fruits were sampled. The samples were frozen in liquid nitrogen and stored in a −80°C freezer.

RNA extraction and RT-qPCR analysis

To verify the transcription level of *CaHATs* and *CaHDACs,* the RT-qPCR analysis was performed. Total RNA was extracted from the 6, 14, 25, 33, 38, and 48 DPA pepper fruit samples by using the HiPure HP Plant RNA Mini Kit (Magen, China). A total of 500 ng RNA were reverse-transcribed with the HiScript III 1st Strand cDNA Synthesis Kit (+ gDNA wiper) (Vazyme Biotech, China).

Quantitative real-time RT-PCR analysis was performed on a CFX384 Touch detection system (Bio-Rad, United States). The reaction mix was 0.5 μL of cDNA template, 0.2 μL of each primer (10 μmol/μL), 5 μL of SYBR Green Master Mix (Vazyme Biotech, China), and added nuclease-free water to 10 μL. The PCR amplification conditions were as follows: 95°C for 5 min; then 40 cycles at 95°C for 5 s; and 60°C for 30 s. The *ubiquitin extension gene* (*CA12g20490*) served as the reference gene for expression analysis (Liu et al., 2021). The relative expression is expressed as the target gene to the *ubiquitin extension gene* (*CA12g20490*). Each value represents the mean of three biological replicates. The *CaHATs* and *CaHDACs* RT-qPCR primer, which enables amplified target genes PCR product length, ranged from 100 to 300 bp. The primers used in this study are listed in **Supplementary Table 2**.

Results

Identification of CaHATs and CaHDACs in pepper genome

We used the protein sequences of 12 AtHATs from Arabidopsis (*Arabidopsis thaliana*) and 32 SlHATs from tomato (*Solanum lycopersicum*); 18 AtHDACs from Arabidopsis, and 14 SlHDACs from tomato as queries to search against the Solanaceae Genomics Network and NCBI database by the BLASTP program. As a consequence, 30 CaHATs and 15 CaHDACs were identified from the "Zunla-1" pepper genome after searching for HAT and HDAC domain sequences, respectively. The HATs and HDACs were named according to their subfamily on the position of the chromosome (**Supplementary Table 1**). The HAT protein length ranged from 157 aa (CaHAG18) to 1,856 aa (CaHAF1), pI ranged from 4.19 (CaHAG23) to 9.56 (CaHAG11), and MW ranged from 18.15 kDa (CaHAG18) to 210.08 kDa (CaHDA3). Concerning HDAC proteins, the length ranged from 217 aa (CaHDA1) to 675 aa (CaHDA3), pI ranged from 4.57 (CaHD3) to 8.98 (CaSRT2), and MW ranged from 24.01 kDa (CaHDA1) to 75.98 kDa (CaHDA3). In contrast to the conserved number of *HDACs* in pepper (*Capsicum annuum*), Arabidopsis (*Arabidopsis thaliana*), tomato (*Solanum lycopersicum*), and rice (*Oryza sativa*), the *HATs* in these species seem subject to unequal duplication and evolution in which the number ranges from 8 (*Oryza sativa*) to 32 (*Solanum lycopersicum*).

Phylogenetic of CsHATs and CsHDACs in pepper

To further investigate the evolutionary relationships and classification of CaHAT and CaHDAC members, the protein sequences from the pepper (*Capsicum annuum*), Arabidopsis (*Arabidopsis thaliana*), tomato (*Solanum lycopersicum*), and rice (*Oryza sativa*) were used for phylogenetic analysis. The result indicates the HATs from the four species were divided into seven clades.

The phylogenetic tree showed that 15 HDACs protein members in pepper were clearly classified into three subfamilies, 10 members belong to HDA1/RPD3 (Reduced Potassium Dependence 3/Histone Deacetylase 1) subfamily, 3 proteins belong to the plant-specific HD2 (Histone Deacetylase 2) subfamily, and 2 members belong to SIR2 (Silent Information Regulator 2) subfamily. The phylogenetic results consistent with HDAC were divided into three subfamilies, RPD3/HDA1, SIR2, and plant-specific HD2 (Guo et al., 2017a; Yuan et al., 2020).

CsHATs and CsHDACs conserved domain

We identified the CaHAT and CaHDAC proteins through an online server (MEME). In total, 27 motifs were identified in 30 CaHATs and exhibited a high variation manner (Figure 1A). For the CaHATs, harboring motifs range from 1 to 18, whereas CaHAG14 did not present any motif. For these CaHATs, 19 conserved domains were identified, and CaHAT process conserved domains range from 1 to 6. Notably, most of the CaHATs presented the Acetyltransf_1 domain. The motifs of the 15 CaHDAC proteins displayed considerable variation, and 17 motifs were identified among these CaHDACs (Figure 1B). The motifs of the 15 CaHDAC proteins displayed considerable variation, in which motif number of CaHDAC proteins ranged from 1 to 10. Motif 4 was found in all CaHDA subgroup proteins. CaHDT subgroup was fond presented motifs 11, 12, and 19. CaSTR subgroup proteins only existed with one motif 17. Conserved domain analysis found HDAC domain exhibited in RPD3/HDA1 subgroup proteins, NPL domain presented in plant-specific CaHD2 subgroup proteins, and SIR2 and SIRT7 supper family domain only existed in SIR2 subgroup proteins. This result is highly consistent with the phylogenetic analysis.

Chromosomal location of CaHATs and CaHDACs

The identified pepper HATs and HDACs were mapped to chromosomes. However, four genes, including *CaHAG1*, *CaHAG2*, *CaHAG3,* and *CaHAC1,* were not anchored because the physical map of pepper was incomplete. Both *CaHATs* and *CaHDACs* were unevenly distributed on the chromosomes (Figure 2). There were no *CaHAT* genes on chromosomes 11, and chromosomes 1, 2, and 6 did not find *CaHDAC* genes. As to *CaHAG21*, *CaHAG22,* and *CaHAG23* were closely located on chromosome 12, while *CaHDA1*, *CaHDA2*, *CaHDA3,* and

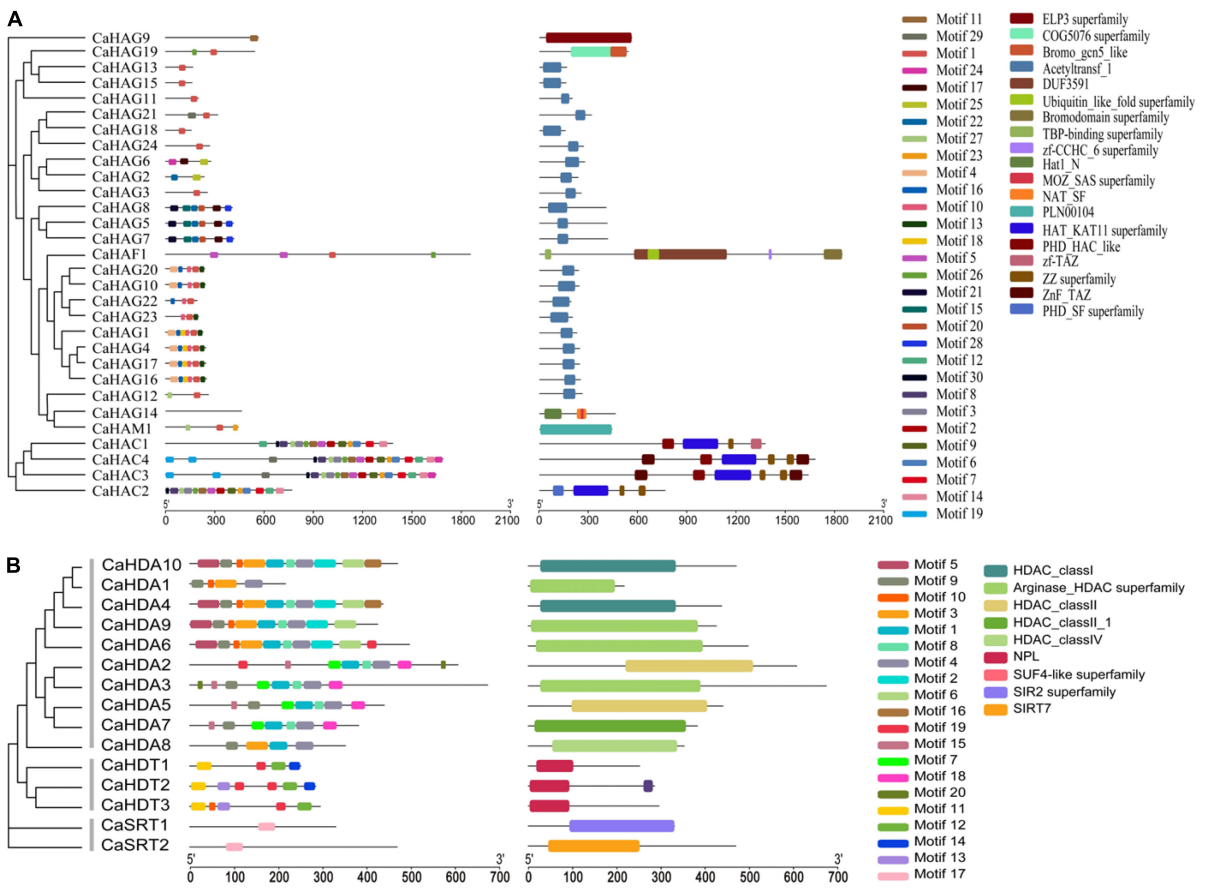


FIGURE 1
Phylogenetic tree of CaHAT **(A)** and CaHDAC **(B)** proteins associated with the motif compositions and conserved domain. The maximum likelihood method was applied for the construction phylogenetic tree. The number at each node represents the bootstrap value from 5,000 replicates. The motif composition and conserved domain of each CaHAT and CaHDAC protein are exhibited in the middle of the figure. Motifs and domains are displayed on the right of the figure. The length of the protein can be estimated from the scale provided at the bottom.

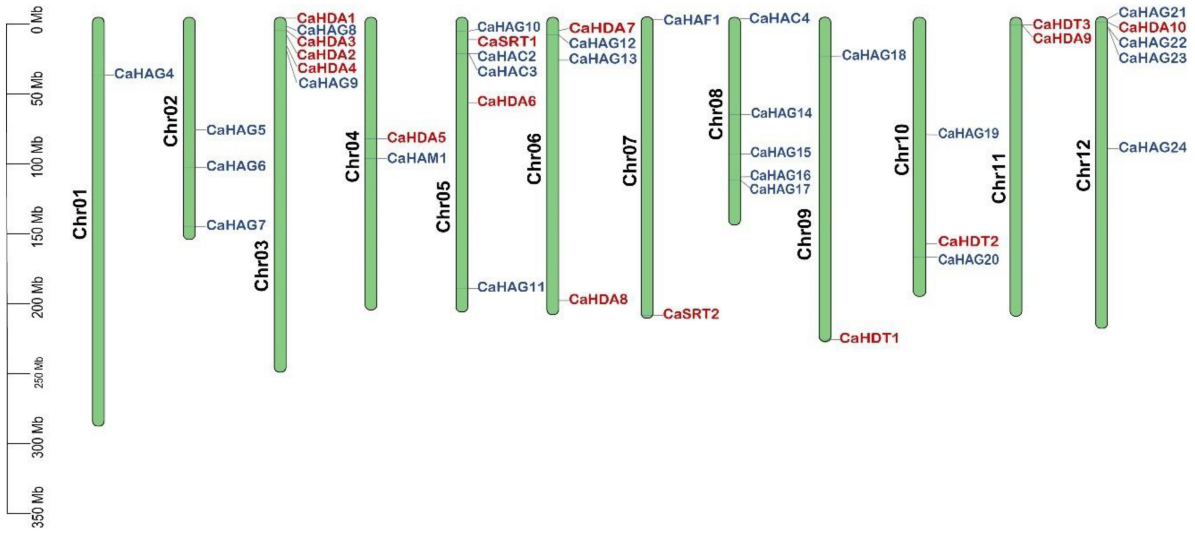


FIGURE 2
Chromosomal location of *CaHATs* and *CaHDACs* gene family. Four genes (*CaHAC1, CaHAG1, CaHAG2,* and *CaHAG3*) that did not assemble on the chromosome were omitted for display.

CaHDA4 were closely located on chromosome 3, suggesting the occurrence of tandem duplication.

Expression pattern of the CaHAT and CaHDAC gene family validates histone acetylation and deacetylation in fruit development

It has been reported that HAT and HDAC play key roles in plant growth and development (Pandey et al., 2002; Hu et al., 2009). To understand the expression profiles of *CaHAT* and *CaHDAC* genes in the pepper fruit developmental stage, the transcriptomic data of "Zunla-1" and "6,424" fruit's different developmental stages were retrieved from the public database (Qin et al., 2014; Liu et al., 2017). The results showed that *CaHAT* and *CaHDAC* gene family expression patterns exhibited considerable variations. According to the hierarchical clustering results, the expression pattern of two gene families was mainly classified into the fruit early development stage upregulation group and the fruit late development upregulation stage group. For the early stage upregulation group, the *CaHATs* (*CaHAM1*, *CaHAG7*, *CaHA14*, *CaHAG5,* and *CaHAC4*) and *CaHDACs* (*CaHAD7*, *CaHAD10,* and *CaSTR2*) might play a crucial role in fruit growth, especially by regulate and/or interaction with phytohormone biosynthetic, receptor, signal response components to govern fruit cell division and expansion. In addition, some gene expression profiles showed more complicated, and the expression levels were considerable in fruit growth and ripening stages (*CaHAG11*, *CaHAG24*, *CaHAG13*; CaHDA2, and CaHDA8), suggesting these proteins process multiple functions in fruit. Regarding the late-stage upregulation group, *CaHATs* (*CaHAG3*, *CaHAG6*, *CaHAG9*, *CaHAG12*, *CaHAG15*, *CaHAG21*, *CaHAF1*, and *CaHAC1*) and *CaHDACs* (*CaHDA7*, *CaHAD9,* and *CaSTR2*) expression tightly associate with the fruit ripening process. To further reveal whether *CaHAT* and *CaHDAC* expression patterns among different cultivars, transcriptome data of the inbred line "6,421" fruit 11 developmental stages, which include pericarp, placenta, and seeds, respectively, were retrieved from previous studies (Liu et al., 2017). The results indicate though including different tissues, these *CaHATs* and *CaHDACs* are also mainly grouped into early development stage upregulation group and the late development stage upregulation group (**Supplementary Figure 1**). However, some *CaHATs* and *CaHDACs* are preferentially expressed in certain tissues, such as *CaHAG4* and *CaHDA4,* mainly expressed in the placenta. The tissue preferential expressed *CaHATs* and *CaHDACs* are likely to govern the tissue-specific biological process.

The CaHAT and CaHDAC genes associated with fruit development

The auxin and gibberellic acid (GA) signaling play a crucial role in fruit development (Fenn and Giovannoni, 2021). Fruit development consists of cell division and expansion, cell division is shown to be influenced by auxin signaling, and cell expansion is believed to be synergistic regulation *via* both auxin and GAs and auxin-responsive Aux/IAA, and ARF proteins modulate crosstalk of auxin and GAs. To clarify the potential relationship of *CaHAT* and *CaHDAC* between auxin

and GAs biosynthetic, receptor, and signaling response genes, the expression and correlation of these genes were analyzed. The expression profile analysis indicates that most *CaAux/IAA* and *CaARF* are highly expressed in early development stage fruit (**Figure 3**), which exhibited similar profiles of early stage upregulation *CaHATs* and *CaHDACs* group (**Figures 3**, **4**). In addition, Pearson's correlation coefficient analysis indicates the early upregulation *CaHATs* (*CaHAG7* and *CaHAM1*) and *CaHDACs* (*CaHDA3*, *CaHDA6*, *CaHDA7,* and *CaHD2*) transcription levels were significantly correlated with the fruit development-related *Aux/IAA*, *ARF,* and GA signal gene (*CaGAI* and *CaGID1b.2*) (**Supplementary Figures 2–6**). These results suggest that early upregulation *CaHATs* (*CaHAG7* and *CaHAM1*) and *CaHDACs* might modulate these auxin-related and GAs-related genes coordinate to regulate fruit development.

CaHAT and CaHDAC are associated with pepper fruit capsaicinoid biosynthesis

The pepper characteristic metabolite capsaicinoid biosynthesis in the fruits from the early stage ZL.F.Dev3 reaches a maximum concentration at ZL.F.Dev6 (Qin et al., 2014). Consistent with capsaicinoid accumulation, the expression of the capsaicinoid biosynthetic genes is highly expressed from ZL.F.Dev3 to ZL.F.Dev4 (**Figure 5**). Moreover, the *CaHAG1* and *CaHAG19* of HATs and CaHDT1 and *CaHDT3* of HDACs also coexpressed with key capsaicinoid regulatory and biosynthetic genes (i.e., *CaMYB31*, *CaAT3*, *CaKasIa*, and *CaKasIIIb*) (**Figures 5A,B**). The results are also supported by Pearson's correlation coefficient analysis (**Supplementary Figures 7**, **8**). This result indicates these coexpressed *CaHATs* and CaHDACs are likely to regulate capsaicinoid biosynthetic gene histone acetylation levels to control gene expression.

CaHATs and CaHDACs associated with pepper fruit ripening

ABA is one of the main regulatory phytohormone in the ripening process of the non-climacteric fruits (Fenn and Giovannoni, 2021). Pepper is a typical non-climacteric fruit, in which fruit ripening is accompanied by the increased biosynthesis of the ABA (Xiao et al., 2020). To address the potential regulatory mechanism of *CaHATs* and *CaHDACs* in fruit ripening, the expression profiles of ABA biosynthesis, receptor, signal, and ripening-related genes, including carotenoid biosynthetic genes, were analyzed. The analysis of expression profiles uncovered that ABA-related (*CaNCED2/3, CaPLY4*, and *CaSnRK2.6*), ripening-related genes (*CaMADS-RIN*, *CaNR*, and *CaNAC-NOR*), and carotenoid biosynthetic genes (*CaPSY* and *CaCCS*) were mainly expressed in fruit's late

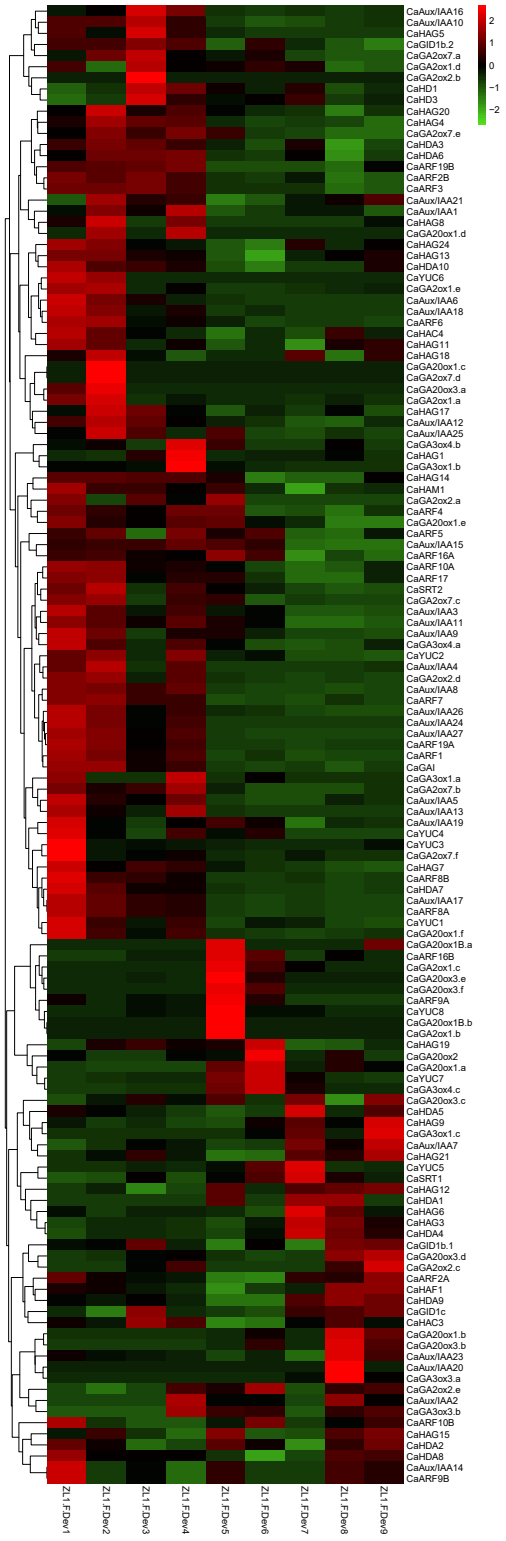


FIGURE 3
Hierarchical clustering analysis expression profiles of CaHATs, CaHDACs, auxin-, and GAs-related genes. The auxin-related genes include biosynthetic genes (*CaYUCs*) and signal transduction genes (*CaARFs* and *CaIAA/ARFs*).

(Continued)

FIGURE 3 (Continued)
The GA-related genes include biosynthetic genes (subfamily of *CaGA2ox*, *CaGA3ox,* and *CaGA20ox*) and signal transduction genes (*CaGAIs* and *CaGIDs*). The heat map was constructed by values of fragments per kilobase of exon per million fragments (FPKM). The name of each gene is shown at the right of the heat map. The expression data were retrieved from "Zunla-1" transcriptome (Qin et al., 2014).

development stages (**Figure 6**). Notably, the late upregulation group *CaHATs* and *CaHDACs* such as *CaHAG3*, *CaHAG9*, and *CaHAG12* of *CaHATs* and *CaHDA1*, *CaHDA3*, *CaHDA5*, and *CaHDA9* of *CaHDACs* are tightly coexpressed with these ABA-related, ripening-related genes and carotenoid biosynthetic genes (**Figure 6**), suggesting these histone acetylation modifiers mediate certain of these genes acetylation level to control fruit ripening.

RT-qPCR analysis validates CaHATs and CaHDACs expression in pepper fruit in different development stages

For validating the expression patterns of the histone acetylation modifiers consistent with pepper fruit development and ripening, in total, 13 of *CaHATs* (*CaHAG1*, *CaHAG5*, *CaHAG7*, *CaHAG9*, *CaHAG20*, and *CaHAG21*) and *CaHDACs* (*CaHDA4*, *CaHDA5*, *CaHDA7*, *CaHDA9*, *CaHDA10*, *CaHDT3,* and *CaSTR1*) were adopted for analysis, respectively. The total RNA was extracted from "59" inbred line fruit at the development stage of 6, 14, 25, 33, 38, and 48 DPA (corresponding to the "Zunla-1" development stage of ZL-1.F.Dev1, ZL-1.F.Dev3, ZL-1.F.Dev4, ZL-1.F.Dev6, ZL-1.F.Dev7, and ZL-1.F.Dev9). The RT-qPCR analysis indicates that these genes' relative expression levels were development-dependent (**Figure 7**). We found some *CaHATs* (*CaHAG1*, *CaHAG5*, *CaHAG7*, *CaHAG9*, and *CaHAG20*) and *CaHDACs* (*CaHDA7*, *CaHDT3,* and *CaSTR1*) mainly expressed in the fruit early stage, suggesting these are involved in fruit development. However, other *CaHATs* (*CaHAG19* and *CaHAG21*) and *CaHDACs* (*CaHDA4, CaHDA5, CaHDA9,* and *CaHDA10*) were mainly expressed in the fruit later stages, implying they are associated with the ripening process. The RT-qPCR analysis was consistent with the transcriptome results of pepper cultivar "Zunla-1" and "6,421" fruits, suggesting the gene expression was genotype-independent and strongly supported their association with fruit development and ripening.

Discussion

HAT and *HDAC* gene families have been identified in different species, including Arabidopsis (Pandey et al., 2002), rice (Hu et al., 2009), grapevine (Aquea et al., 2010), tomato (Aiese Cigliano et al., 2013), Citrus (Xu et al., 2015; Shu et al., 2021), litchi (Peng et al., 2017), tea plant (Yuan et al., 2020), and winter wheat (Li et al., 2021, 2022). In contrast relative conserved number of *HDAC* in different species, whereas *HAT* exhibited considerable variation. In this study, we identified 32 *CaHATs* and 15 *CaHDACs* in the pepper genome (**Table 1**). Notably, compared with other species, the *HATs* subfamily *HAGs* have remarkably expanded in pepper (24 *HAGs*), and tomato (26 *HAGs*) compared with dicotyledonous Arabidopsis (3 *HAGs*) and monocotyledonous rice (3 *HAGs*) (**Table 1**). In plants, the gene family expansion mainly arises from tandem duplications and segmental or whole-genome duplications (Freeling, 2009). Presumably, after being separated from their ancestors, *HATs* likely undergo unequal duplication events among species. Indeed, location analysis found some Ca*HATs* (such as *CaHAG21*, *CaHAG22,* and *CaHAG23*) were closely located on chromosomes, suggesting the expansion of these genes arises from tandem duplications. Apparently, before tomato and pepper speciation, tandem duplications and segmental occurred in common ancestors, which finally led to *HATs* significantly expanding in Solanaceae.

CaHATs and CaHDACs included subfamily-specific domains, and subfamily members also exhibited similar protein sequence length, motif composition, and gene structure, reflecting a close phylogenetic relationship. In this study, the HAT proteins could be divided into four families: HAC, HAF, HAM, and HAG (**Figures 1**, **8**), consistent with previous studies (Pandey et al., 2002; Aquea et al., 2010; Aiese Cigliano et al., 2013; Li et al., 2021; Shu et al., 2021). The largest subfamily of CaHATs contained 24 CaHAGs and was further divided into five groups. After analyzing their protein structures, we found that the structures of HAGs were quite different, indicating that they might execute different functions and support the proposed subgroups. The winter wheat genome identified nine HAGs, and their protein structures are highly variable, enabling these HAGs into three groups (Li et al., 2022). In this study, the CaHDACs were classified into three subfamilies: HDA/RPD3, HDT/HD2, and STR/SIR2 (**Figures 1**, **8** and **Table 1**). The results are consistent with the identified HDACs in the genome of Arabidopsis (Pandey et al., 2002), rice (Hu et al., 2009), grapevine (Aquea et al., 2010), tomato (Aiese Cigliano et al., 2013), Citrus (Xu et al., 2015), papaya (Fu et al., 2019), and tea plant (Yuan et al., 2020), but the HDT/HD2 subfamily orthologous seems to have been lost in the litchi genome during the duplication events (Peng et al., 2017). Whether these HDACs process similar functions among different species needs further study.

Histone modifications play a critical role in growth, development processes, flowering, and stress responses (Pandey et al., 2002; Hu et al., 2009). Recently, based on stress treatment-derived gene expression profiles, in tea plant (Yuan et al., 2020), *Citrus sinensis* (Shu et al., 2021), and winter wheat

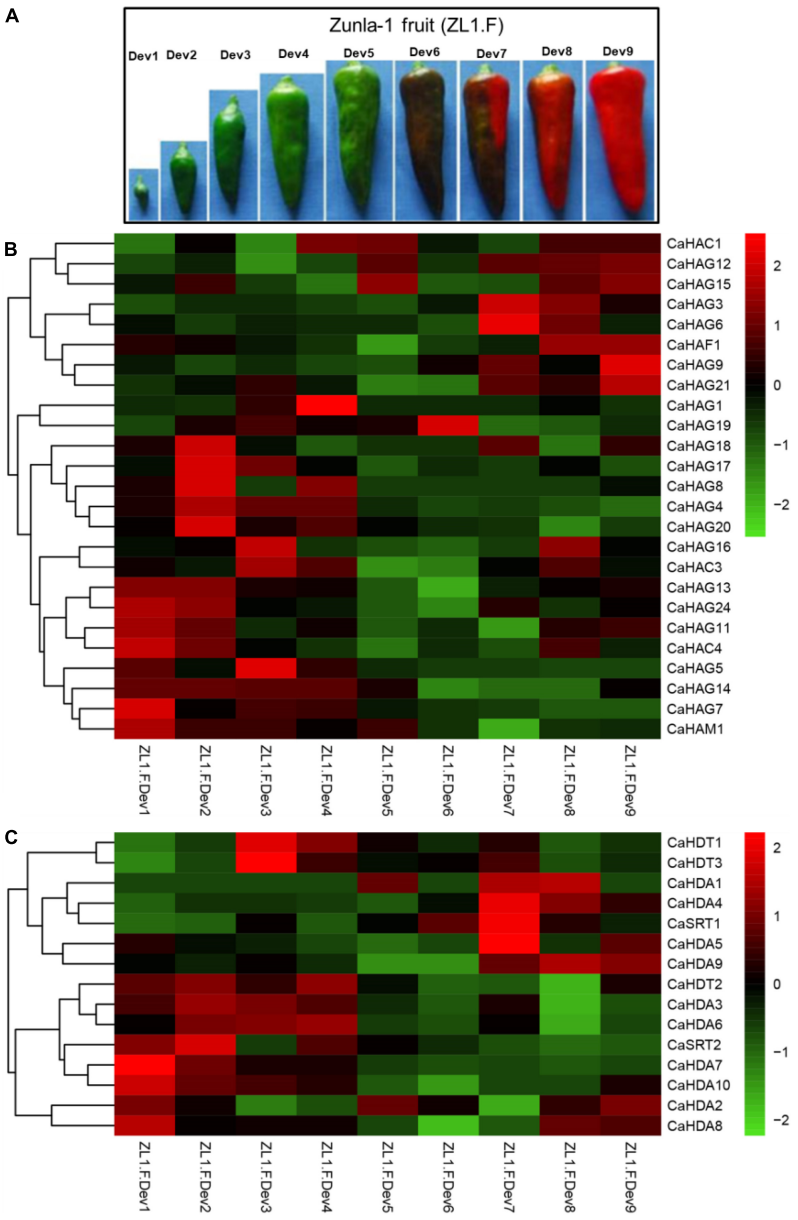


FIGURE 4
Expression pattern of CaHATs and CaHDACs in pepper nine fruit developmental stages. **(A)** "Zunla-1" (ZL-1) fruit at nine different stages, including the whole development stages of pepper fruit. The expression of *CaHATs* **(B)** and *CaHDACs* **(C)** at nine different development stages is shown in **(A)**. The expression data were retrieved from "Zunla-1" transcriptome (Qin et al., 2014).

(Li et al., 2022), we identified a set of stress-inducible *HATs* and *HDACs* and proposed these genes play an essential role in stress response. In this study, we identified 32 *CaHATs* and 15 *CaHDACs*, but the specific functions of HATs and HDACs in pepper remain unknown. Based on the expression profiles, the *CaHATs* and *CaHDACs* were mainly classified into the fruit early development stage upregulation group and the late development stage upregulation group (Figures 4, 9). The fruit's early un-regulation group histone modifiers were proposed to regulate fruit development (including cell division and expansion) for their highly coexpressed with auxin and GAs biosynthetic and signal response genes (Figure 3). In addition to changes in fruit size, pepper development also along with capsaicinoid biosynthesis. Notably, some *CaHATs* and *CaHDACs* display similar profiles with capsaicinoid biosynthetic genes (Figure 5). Perhaps, the mechanism of CaHATs and CaHDACs controls fruit development and capsaicinoid biosynthesis on the two sides. On the one hand, histone modifiers alter acetylation level to regulate development-related genes (including auxin and GAs pathway-related genes) and capsaicinoid biosynthetic

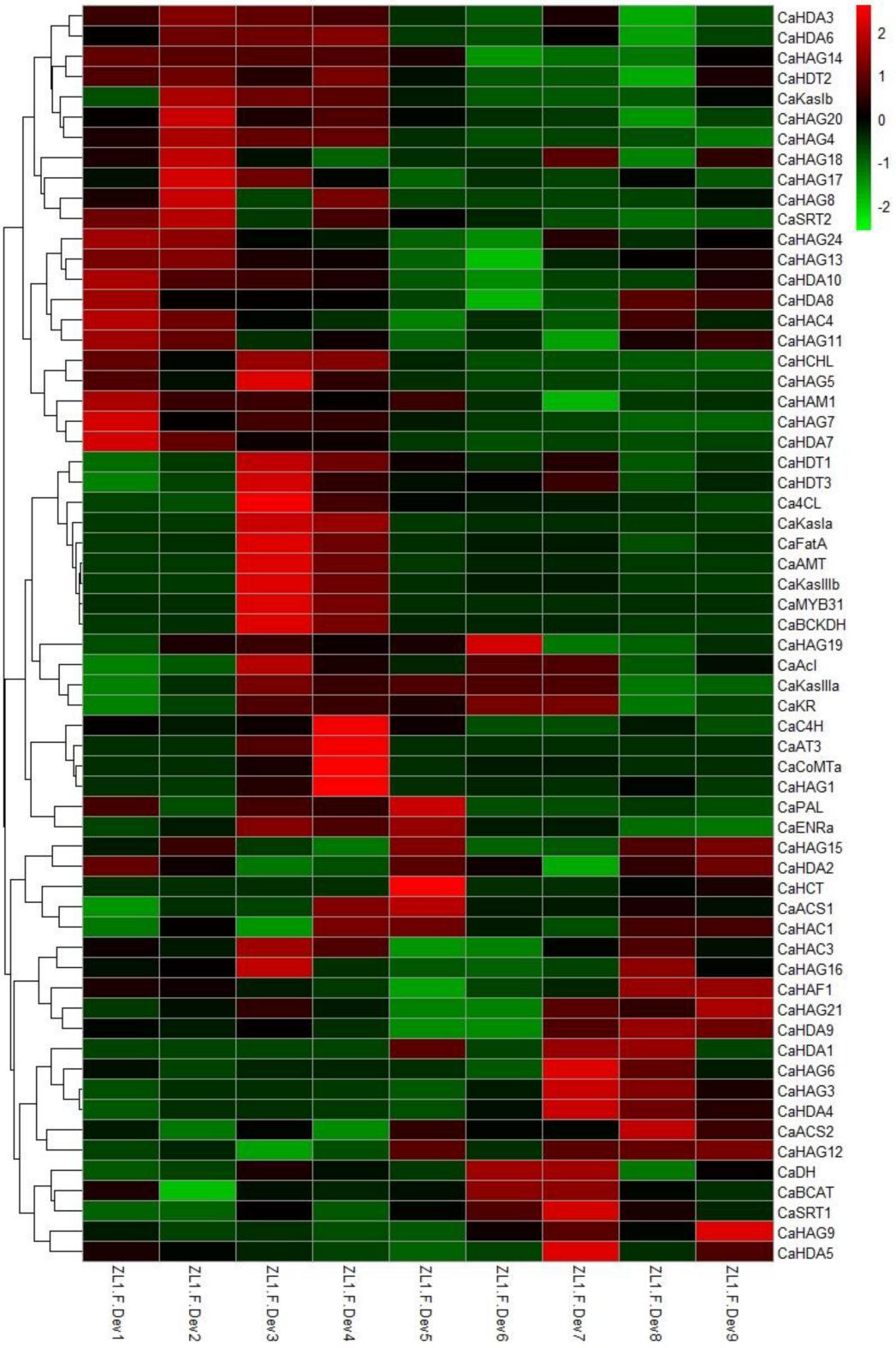


FIGURE 5
Expression profiles of *CaHATs*, *CaHDACs,* and capsaicinoid regulatory and biosynthetic genes. The heat map includes 19 capsaicinoid biosynthetic genes and one regulatory gene *CaMYB31*. The heat map was constructed by values of fragments per kilobase of exon per million fragments (FPKM). The expression data were retrieved from "Zunla-1" transcriptome (Qin et al., 2014).

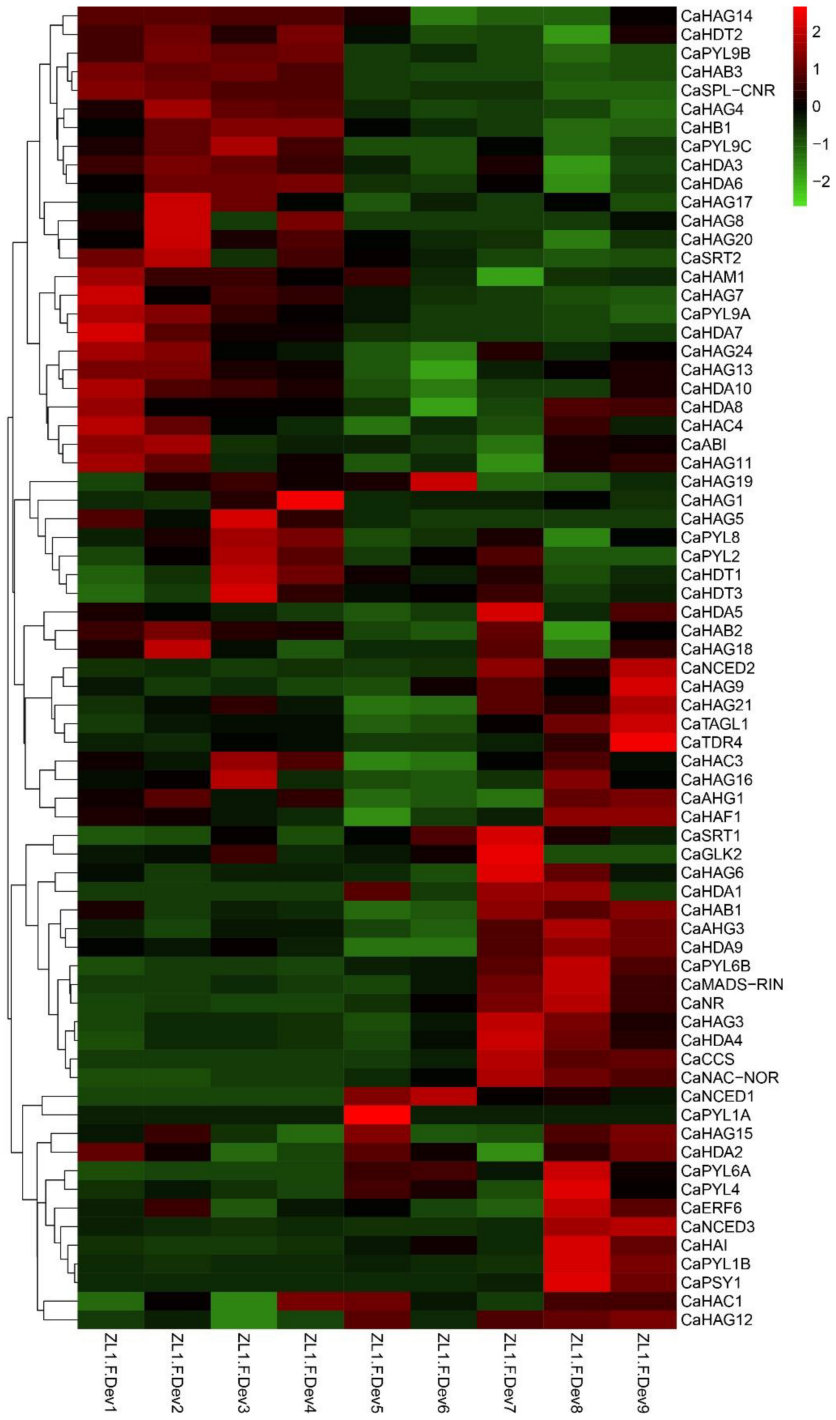


FIGURE 6
Expression profiles of *CaHATs*, *CaHDACs,* and ABA-related, ripening-related genes, and carotenoid biosynthetic genes. The ABA-related includes ABA biosynthetic (*CaNCEDs*) and signaling transduction genes (*CaPYLs*). Carotenoid biosynthetic genes include *CaPSY1* and *CaCCS*. The ripening-related genes were different transcription factors that govern the ripening process.

gene expression by changing the state of the chromatin; on the other hand, CaHATs and CaHDACs recruiting DNA-binding transcription factor to formate protein complexes affect target genes expression. It can be learned from Arabidopsis that a light-regulated histone deacetylase HDA15 interacts with PIF1to suppress the light-responsive genes during seed germination in the dark (Gu et al., 2017) and HDA15 also restricts chlorophyll biosynthesis- and photosynthesis-related

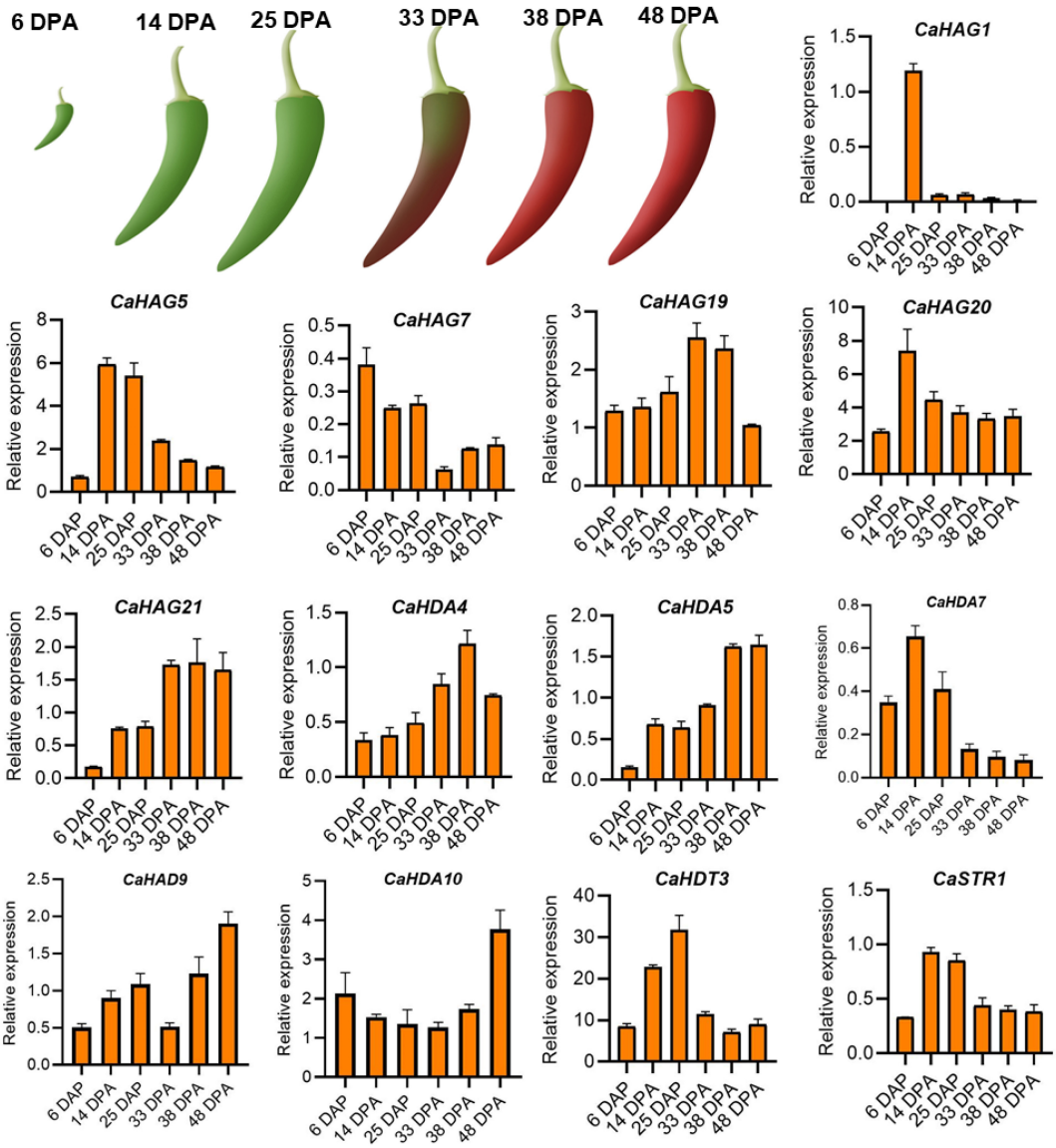


FIGURE 7
qRT-PCR analysis expression patterns of 13 *CaHAT* and *CaHDAC* genes in different developmental stages of the fruit. The upper figure indicates schematic diagram of "59" inbred line fruit phenotypes at six developmental stages: 6, 14, 25, 33, 38, and 48 DPA. Bottom of figure indicates the expression of 13 selected *CaHATs* and *CaHDACs* expression patterns in fruit six developmental stages. Data are mean ± SE (n = 3).

genes through its interaction with PIF3 (Liu et al., 2013). Likewise, HDA19 interacts with histone methyltransferase SUVH5 and regulates seed dormancy through ABA and GA signaling pathways by modulating overall histone H3 acetylation and H3K9me2 methylation on the promoter of the target genes *ABI3*, *RGA*, and *DOG1* (Zhou et al., 2020; Kumar et al., 2021). The Arabidopsis MYST histone acetyltransferase HAM1/2 enhances the expression of the negative regulator FLC gene through H4K5 acetylation and results in delayed flowering (Xiao et al., 2013). The cucumber (*Cucumis sativus*) *SF2* encodes a Histone Deacetylase Complex1 (HDC1) homolog and its role in regulating cell proliferation through the HDAC complex to promote histone deacetylation of key genes involved in multiple phytohormone pathways and cell cycle regulation (Zhang et al., 2020). A weak *sf2* allele impairs HDAC targeting to chromatin, resulting in elevated histone acetylation levels. Overall, these studies indicate histone acetylation plays a critical role in development, whereas the exact regulatory mechanism concerning CaHATs and CaHDACs control fruit development and capsaicinoid biosynthesis needs further investigation. In addition, CaHATs and CaHDACs cannot bind DNA directly, identifying the key transcription factors involved in plant development may interact, and recruiting these histone modifiers in the co-regulation of gene expression is needed.

TABLE 1 Number of HATs and HADCs in four plant species.

Types	Subfamily	Species			
		Capsicum annuum	*Solanum lycopersicum*	*Arabidopsis thaliana*	*Oryza sativa*
HATs	HAC	4	4	5	3
	HAG	24	26	3	3
	HAF	1	1	2	1
	HAM	1	1	2	1
Total		30	32	12	8
HDACs	HDA/RPD3	10	9	12	14
	HDT/HD2	3	3	4	2
	STR/SIR2	2	2	2	2
Total		15	15	18	18

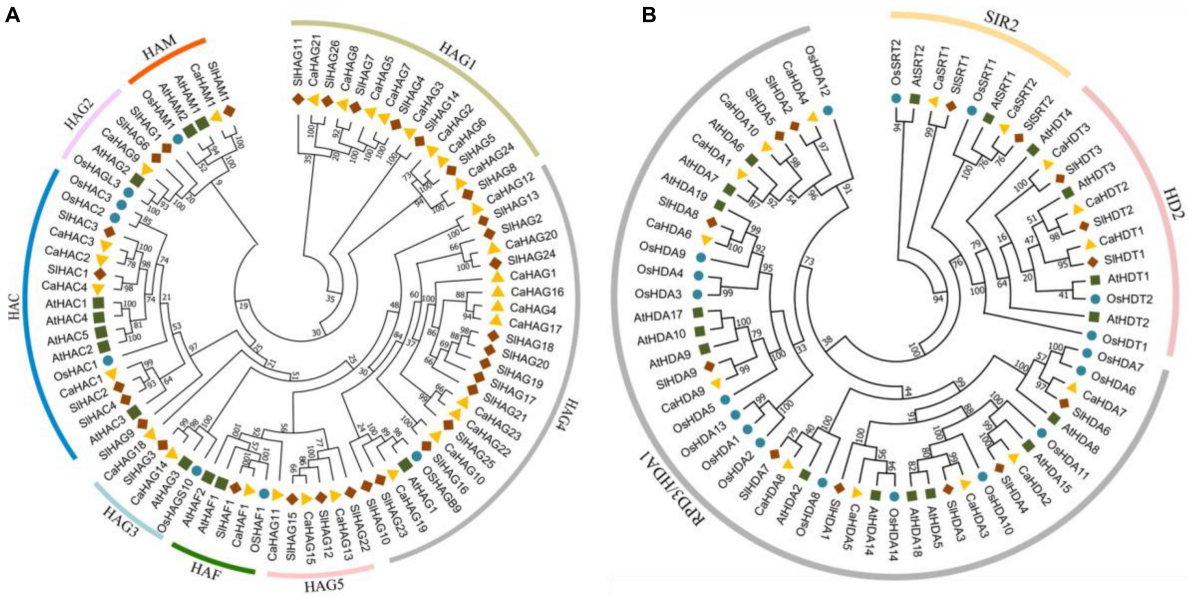


FIGURE 8
Phylogenetic tree of the HAT and HDAC proteins. The HAT **(A)** and HDAC **(B)** proteins of pepper (*Capsicum annuum*), Arabidopsis (*Arabidopsis thaliana*), tomato (*Solanum lycopersicum*), and rice (*Oryza sativa*) were adopted for analysis. The tree was constructed with the maximum likelihood method by adopting HAT and HDAC protein full-length protein sequences from pepper (*Capsicum annuum*), Arabidopsis (*Arabidopsis thaliana*), tomato (*Solanum lycopersicum*), and rice (*Oryza sativa*). The number at each node represents the bootstrap value from 1,000 replicates. The protein sequences used in this study were retrieved from Solanaceae Genomics Network and NCBI database.

Compared to the rare study focused on fleshy fruit development, numerous studies have reported histone acetylation involvement in the fruit ripening of tomato (Guo et al., 2017c, 2018; Deng et al., 2022), bananas (Han et al., 2016; Fu et al., 2018), papaya (Fu et al., 2019), and pear (Li X. et al., 2020) and apple (Hu et al., 2022). However, most of the reported studies were relevant to histone acetylation involvement in climacteric fruit ripening, while this process control in non-climacteric fruit ripening remains elusive. In this study, we identified fruit late upregulation *CaHATs* and *CaHDACs* and regarded them as ripening candidates (**Figures 6**, **7**, **9**). Expression analysis found that they are highly coexpressed with ABA, ripening, and carotenoid-related genes, which supported their association with maturation. The orthologous HAF in tomato (*SlHAF1*) (Xu et al., 2015) and Citrus (*CsHAF1/2*) has the strongest expression in mature fruit, suggesting an important role in ripening. Similarly, the identified *CaHAF1* also showed a high expression level in pepper breaking and ripening stage fruit (**Figure 4**), implying that they could have similar functions as tomato and Citrus. Tomato SlHDT1 and SlHDT3 positively regulate ripening by regulating carotenoid and ethylene content and expression of ethylene biosynthetic genes, ripening-associated genes, and cell wall metabolism genes (Guo et al., 2017b; Guo, 2022). The *CaHDT1*

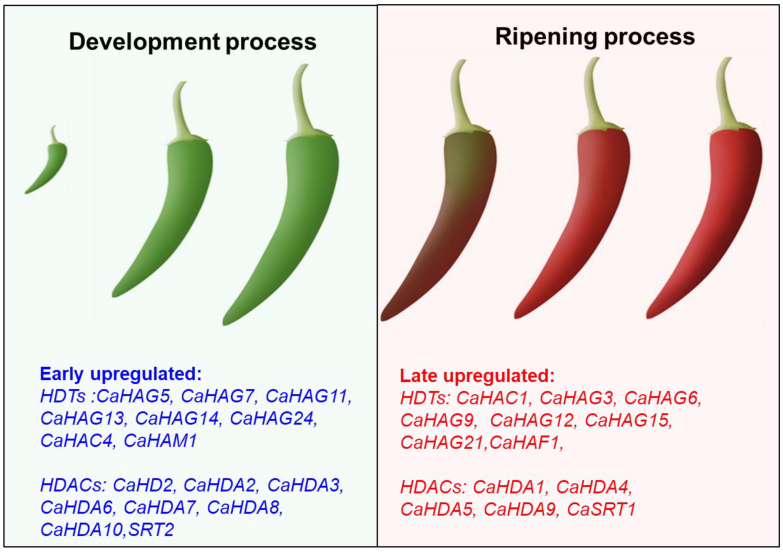


FIGURE 9
Proposal model of *CaHDTs* and *CaHDACs* candidates in pepper development and ripening. The leaf panel display early upregulated histone acetylation modifiers might be associated with the fruit development-related process. The right panel display late upregulated histone acetylation modifiers might be related to the fruit ripening-related process.

and *CaHDT3* were highly expressed in pepper immature and mature fruit (**Figure 4**), suggesting it might control fruit ripening but also process other functions. In contrast, SlHDA1 and SlHDA3 process redundancy function and play a role in the repression of fruit ripening and carotenoid accumulation (Guo et al., 2017c, 2018). While CaHDA3 is expressed in different development stages, HDA1 orthologous CaHDA1 is mainly expressed in pepper ripening fruit, implying the function suffers divergent evolution. Recently, studies found that HATs recruit coactivator proteins to activator gene transcription (Servet et al., 2010), whereas HDACs associated with corepressor proteins always form the complex to repress transcription (Kumar et al., 2021). Banana ripening negative regulator MaERF11 physically interacts with a histone deacetylase MaHDA1, and the interaction significantly strengthens the MaERF11-mediated transcriptional repression of MaACO1 and expansins (Han et al., 2016). Likewise, CpHDA3 is associated with CpERF9 in the suppression of CpPME1/2 and CpPG5 genes during papaya fruit ripening (Fu et al., 2019). Recently, the tripartite complex MdERF4–MdTPL4-MdHDA19 in apples (Hu et al., 2022) and SlERF.F12-SlTPL3-SlHDA1/3 in tomatoes (Deng et al., 2022) negatively regulate the onset of fruit ripening represses transcription of ripening genes by decreasing the level of the permissive histone acetylation marks H3K9Ac and H3K27Ac at their promoter regions. However, we should be kept in mind that all the mentioned above are climacteric fruit, and ethylene plays an essential role in the onset ripening process (Fenn and Giovannoni, 2021), whereas pepper is the non-climacteric fruit ripening, and ABA is responsible for ripening (Xiao et al., 2020). Whether orthologous climacteric fruit ripening-related *HDAC* genes process conserved functions in pepper needs further research. As another important modification, histone methylation, and demethylation have been evidenced to play an important role in fruit ripening by modulating H3K27 to regulate ripening-related gene expression (Li Z. et al., 2020; Ding et al., 2022); the exact function of histone methylation in pepper ripening also needs further study. In addition, fruit ripening along with genome-wide dynamic changes in DNA methylation has been identified in tomato (Zhong et al., 2013), strawberry (Cheng et al., 2018), and orange (Huang et al., 2019), and whether DNA methylation is associated with pepper fruit ripening should be dissected in future. Moreover, how histone modification and DNA methylation cooperate to regulate ripening-associated gene expression requires further investigation.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: NCBI, SRP018258. The names of additional repositories and accession number(s) can be found in the article/**Supplementary material**.

Author contributions

ZZ conceived and supervised the project, designed the study, and wrote the manuscript. GC provided some construction suggestions and revised the manuscript. YC, MX,

JRL, and HZ performed major experiments and analyzed the data. GC, CC, BC, BS, QD, JJL, MC, and KC revised the manuscript, provided some useful suggestions, and revised the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Guangdong Basic and Applied Basic Research Foundation (2022A1515012547), the Guangzhou Basic and Applied Basic Research Foundation (SL2023A04J00544), and the National Natural Science Foundation of China (32102380, U21A20230, 32070331, and 32072580).

Conflict of interest

MC and KC were employed by the Guangdong Helinong Seeds Co., Ltd.

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.971230/full#supplementary-material

References

Aiese Cigliano, R., Sanseverino, W., Cremona, G., Ercolano, M. R., Conicella, C., and Consiglio, F. M. (2013). Genome-wide analysis of histone modifiers in tomato: Gaining an insight into their developmental roles. *BMC Genom.* 14:57. doi: 10.1186/1471-2164-14-57

Aquea, F., Timmermann, T., and Arce-Johnson, P. (2010). Analysis of histone acetyltransferase and deacetylase families of *Vitis vinifera*. *Plant Physiol. Biochem.* 48, 194–199. doi: 10.1016/j.plaphy.2009.12.009

Arce-Rodríguez, M. L., Martínez, O., and Ochoa-Alejo, N. (2021). Genome-wide identification and analysis of the myb transcription factor gene family in chili pepper (*Capsicum* spp.). *Int. J. Mol. Sci.* 22:2229. doi: 10.3390/ijms22052229

Arents, G., Burlingame, R. W., Wang, B. I. C., Love, W. E., and Moudrianakis, E. N. (1991). The nucleosomal core histone octamer at 3.1 Å resolution: A tripartite protein assembly and a left-handed superhelix. *Proc. Natl. Acad. Sci. U. S. A.* 88, 10148–10152. doi: 10.1073/pnas.88.22.10148

Bannister, A. J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Res.* 21, 381–395. doi: 10.1038/cr.2011.22

Berry, H. M., Rickett, D. V., Baxter, C. J., Enfissi, E. M. A., and Fraser, P. D. (2019). Carotenoid biosynthesis and sequestration in red chilli pepper fruit and its impact on colour intensity traits. *J. Exp. Bot.* 70, 2637–2650. doi: 10.1093/jxb/erz086

Brumos, J. (2021). Gene regulation in climacteric fruit ripening. *Curr. Opin. Plant Biol.* 63:102042. doi: 10.1016/j.pbi.2021.102042

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., et al. (2020). TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol. Plant* 13, 1194–1202. doi: 10.1016/j.molp.2020.06.009

Chen, X., Ding, A. B., and Zhong, X. (2020). Functions and mechanisms of plant histone deacetylases. *Sci. China Life Sci.* 63, 206–216. doi: 10.1007/s11427-019-1587-x

Cheng, J., Niu, Q., Zhang, B., Chen, K., Yang, R., Zhu, J. K., et al. (2018). Downregulation of RdDM during strawberry fruit ripening. *Genome Biol.* 19:21. doi: 10.1186/s13059-018-1587-x

Deng, H., Chen, Y., Liu, Z., Liu, Z., Shu, P., Wang, R., et al. (2022). SlERF.F12 modulates the transition to ripening in tomato fruit by recruiting the co-repressor TOPLESS and histone deacetylases to repress key ripening genes. *Plant Cell* 34, 1250–1272. doi: 10.1093/plcell/koac025

Ding, X., Liu, X., Jiang, G., Li, Z., Song, Y., Zhang, D., et al. (2022). SlJMJ7 orchestrates tomato fruit ripening *via* crosstalk between H3K4me3 and DML2-mediated DNA demethylation. *New Phytol.* 233, 1202–1219. doi: 10.1111/nph.17838

Fenn, M. A., and Giovannoni, J. J. (2021). Phytohormones in fruit development and maturation. *Plant J.* 105, 446–458. doi: 10.1111/tpj.15112

Freeling, M. (2009). Bias in plant gene content following different sorts of duplication: Tandem, whole-genome, segmental, or by transposition. *Annu. Rev. Plant Biol.* 60, 433–453. doi: 10.1146/annurev.arplant.043008.092122

Fu, C., Chen, H., Gao, H., and Han, Y. (2019). Histone Deacetylase CpHDA3 Is Functionally Associated with CpERF9 in Suppression of CpPME1/2 and CpPG5 Genes during Papaya Fruit Ripening. *J. Agric. Food Chem.* 67, 8919–8925. doi: 10.1021/acs.jafc.9b03800

Fu, C. C., Han, Y. C., Guo, Y. F., Kuang, J. F., Chen, J. Y., Shan, W., et al. (2018). Differential expression of histone deacetylases during banana ripening and identification of MaHDA6 in regulating ripening-associated genes. *Postharvest Biol. Technol.* 141, 24–32. doi: 10.1016/j.postharvbio.2018.03.010

Giovannoni, J. J. (2004). Genetic regulation of fruit development and ripening. *Plant Cell* 16, S170–S180. doi: 10.1105/tpc.019158

Gu, D., Chen, C. Y., Zhao, M., Zhao, L., Duan, X., Duan, J., et al. (2017). Identification of HDA15-PIF1 as a key repression module directing the transcriptional network of seed germination in the dark. *Nucleic Acids Res.* 45, 7137–7150. doi: 10.1093/nar/gkx283

Guo, J. E. (2022). Histone deacetylase gene SlHDT1 regulates tomato fruit ripening by affecting carotenoid accumulation and ethylene biosynthesis. *Plant Sci.* 318:111235. doi: 10.1016/j.plantsci.2022.111235

Guo, J. E., Hu, Z., Li, F., Zhang, L., Yu, X., Tang, B., et al. (2017b). Silencing of histone deacetylase SlHDT3 delays fruit ripening and suppresses carotenoid accumulation in tomato. *Plant Sci.* 265, 29–38. doi: 10.1016/j.plantsci.2017.09.013

Guo, J. E., Hu, Z., Guo, X., Zhang, L., Yu, X., Zhou, S., et al. (2017a). Molecular Characterization of Nine Tissue-Specific or Stress-Responsive Genes of Histone Deacetylase in Tomato (*Solanum lycopersicum*). *J. Plant Growth Regul.* 36, 566–577. doi: 10.1007/s00344-016-9660-8

Guo, J. E., Hu, Z., Zhu, M., Li, F., Zhu, Z., Lu, Y., et al. (2017c). The tomato histone deacetylase SlHDA1 contributes to the repression of fruit ripening and carotenoid accumulation. *Sci. Rep.* 7:79307. doi: 10.1038/s41598-017-08512-x

Guo, J. E., Hu, Z., Yu, X., Li, A., Li, F., Wang, Y., et al. (2018). A histone deacetylase gene, SlHDA3, acts as a negative regulator of fruit ripening and carotenoid accumulation. *Plant Cell Rep.* 37, 125–135. doi: 10.1007/s00299-017-2211-3

Han, Y. C., Kuang, J. F., Chen, J. Y., Liu, X. C., Xiao, Y. Y., Fu, C. C., et al. (2016). Banana transcription factor MaERF11 recruits histone deacetylase MaHDA1 and represses the expression of MaACO1 and expansins during fruit ripening. *Plant Physiol.* 171, 1070–1084. doi: 10.1104/pp.16.00301

Hu, Y., Han, Z., Wang, T., Li, H., Li, Q., Wang, S., et al. (2022). Ethylene response factor MdERF4 and histone deacetylase MdHDA19 suppress apple fruit ripening through histone deacetylation of ripening-related genes. *Plant Physiol.* 188, 2166–2181. doi: 10.1093/plphys/kiac016

Hu, Y., Qin, F., Huang, L., Sun, Q., Li, C., Zhao, Y., et al. (2009). Rice histone deacetylase genes display specific expression patterns and developmental functions. *Biochem. Biophys. Res. Commun.* 388, 266–271. doi: 10.1016/j.bbrc.2009.07.162

Huang, H., Liu, R., Niu, Q., Tang, K., Zhang, B., Zhang, H., et al. (2019). Global increase in DNA methylation during orange fruit development and ripening. *Proc. Natl. Acad. Sci. U. S. A.* 116:201815441. doi: 10.1073/pnas.1815441116

Jang, Y. K., Jung, E. S., Lee, H. A., Choi, D., and Lee, C. H. (2015). Metabolomic Characterization of Hot Pepper (*Capsicum annuum* "cM334") during Fruit Development. *J. Agric. Food Chem.* 63, 9452–9460. doi: 10.1021/acs.jafc.5b03873

Kim, S., Park, M., Yeom, S. I., Kim, Y. M., Lee, J. M., Lee, H. A., et al. (2014). Genome sequence of the hot pepper provides insights into the evolution of pungency in *Capsicum* species. *Nat. Genet.* 46, 270–278. doi: 10.1038/ng.2877

Kumar, V., Thakur, J. K., and Prasad, M. (2021). Histone acetylation dynamics regulating plant development and stress responses. *Cell. Mol. Life Sci.* 78, 4467–4486. doi: 10.1007/s00018-021-03794-x

Li, H., Liu, H., Pei, X., Chen, H., Li, X., Wang, J., et al. (2021). Comparative Genome-Wide Analysis and Expression Profiling of Histone Acetyltransferases and Histone Deacetylases Involved in the Response to Drought in Wheat. *J. Plant Growth Regul.* 41, 1065–1078.

Li, H., Liu, H., Pei, X., Chen, H., Li, X., Wang, J., et al. (2022). Comparative Genome-Wide Analysis and Expression Profiling of Histone Acetyltransferases and Histone Deacetylases Involved in the Response to Drought in Wheat. *J. Plant Growth Regul.* 41, 1065–1078. doi: 10.1007/s00344-021-10364-9

Li, X., Guo, W., Li, J., Yue, P., Bu, H., Jiang, J., et al. (2020). Histone acetylation at the promoter for the transcription factor PUWRKY31 affects sucrose accumulation in pear fruit. *Plant Physiol.* 182, 2035–2046. doi: 10.1104/PP.20.00002

Li, Z., Jiang, G., Liu, X., Ding, X., Zhang, D., Wang, X., et al. (2020). Histone demethylase SlJMJ6 promotes fruit ripening by removing H3K27 methylation of ripening-related genes in tomato. *New Phytol.* 227, 1138–1156. doi: 10.1111/nph.16590

Liu, F., Yu, H., Deng, Y., Zheng, J., Liu, M., Ou, L., et al. (2017). PepperHub, an Informatics Hub for the Chili Pepper Research Community. *Mol. Plant* 10, 1129–1132. doi: 10.1016/j.molp.2017.03.005

Liu, R., Song, J., Liu, S., Chen, C., Zhang, S., Wang, J., et al. (2021). Genome - wide identification of the *Capsicum* bHLH transcription factor family?: Discovery of a candidate regulator involved in the regulation of species - specific bioactive metabolites. *BMC Plant Biol.* 21:262. doi: 10.1186/s12870-021-03004-7

Liu, X., Chen, C. Y., Wang, K. C., Luo, M., Tai, R., Yuan, L., et al. (2013). PHYTOCHROME INTERACTING FACTOR3 associates with the histone deacetylase HDA15 in repression of chlorophyll biosynthesis and photosynthesis in etiolated *Arabidopsis* seedlings. *Plant Cell* 25, 1258–1273. doi: 10.1105/tpc.113.109710

Liu, Z., Lv, J., Zhang, Z., Li, H., Yang, B., Chen, W., et al. (2019). Integrative Transcriptome and Proteome Analysis Identifies Major Metabolic Pathways Involved in Pepper Fruit Development. *J. Proteome Res.* 18, 982–994. doi: 10.1021/acs.jproteome.8b00673

Loidl, P. (2004). A plant dialect of the histone language. *Trends Plant Sci.* 9, 84–90. doi: 10.1016/j.tplants.2003.12.007

Pandey, R., Müller, A., Napoli, C. A., Selinger, D. A., Pikaard, C. S., Richards, E. J., et al. (2002). Analysis of histone acetyltransferase and histone deacetylase families of *Arabidopsis thaliana* suggests functional diversification of chromatin modification among multicellular eukaryotes. *Nucleic Acids Res.* 30, 5036–5055. doi: 10.1093/nar/gkf660

Peng, M., Ying, P., Liu, X., Li, C., Xia, R., Li, J., et al. (2017). Genome-wide identification of histone modifiers and their expression patterns during fruit abscission in litchi. *Front. Plant Sci.* 8:639. doi: 10.3389/fpls.2017.00639

Qin, C., Yu, C., Shen, Y., Fang, X., Chen, L., Min, J., et al. (2014). Whole-genome sequencing of cultivated and wild peppers provides insights into *Capsicum* domestication and specialization. *Proc. Natl. Acad. Sci. U. S. A.* 111, 5135–5140. doi: 10.1073/pnas.1400975111

Servet, C., Conde, E., Silva, N., and Zhou, D. X. (2010). Histone acetyltransferase AtGCN5/HAG1 is a versatile regulator of developmental and inducible gene expression in *Arabidopsis*. *Mol. Plant* 3, 670–677. doi: 10.1093/mp/ssq018

Shen, Y., Wei, W., and Zhou, D. X. (2015). Histone Acetylation Enzymes Coordinate Metabolism and Gene Expression. *Trends Plant Sci.* 20, 614–621. doi: 10.1016/j.tplants.2015.07.005

Shu, B., Xie, Y., Zhang, F., Zhang, D., Liu, C., Wu, Q., et al. (2021). Genome-wide identification of citrus histone acetyltransferase and deacetylase families and their expression in response to arbuscular mycorrhizal fungi and drought. *J. Plant Interact.* 16, 367–376. doi: 10.1080/17429145.2021.1934131

Song, J., Chen, C., Zhang, S., Wang, J., Huang, Z., Chen, M., et al. (2020). Systematic analysis of the *Capsicum* ERF transcription factor family: Identification of regulatory factors involved in the regulation of species-specific metabolites. *BMC Genom.* 21:573. doi: 10.1186/s12864-020-06983-3

Stoll, S., Wang, C., and Qiu, H. (2018). DNA methylation and histone modification in hypertension. *Int. J. Mol. Sci.* 19:1174. doi: 10.3390/ijms19041174

Sun, B., Chen, C., Song, J., Zheng, P., Wang, J., Wei, J., et al. (2022). The *Capsicum* MYB31 regulates capsaicinoid biosynthesis in the pepper pericarp. *Plant Physiol. Biochem.* 176, 21–30. doi: 10.1016/j.plaphy.2022.02.014

Teyssier, E., Boureauv, L., Chen, W., Lui, R., Degraeve-Guibault, C., Stammitti, L., et al. (2015). "Epigenetic regulation during fleshy fruit development and ripening," in *Applied Plant Genomics and Biotechnology*, eds P. Poltronieri and Y. Hong (Sawston: Woodhead Publishing), 132–151. doi: 10.1016/B978-0-08-100068-7.00008-2

Villa-Rivera, M. G., and Ochoa-Alejo, N. (2021). Transcriptional regulation of ripening in chili pepper fruits (*Capsicum* spp.). *Int. J. Mol. Sci.* 22:12151. doi: 10.3390/ijms222212151

Wang, Z., Cao, H., Chen, F., and Liu, Y. (2014). The roles of histone acetylation in seed performance and plant development. *Plant Physiol. Biochem.* 84, 125–133. doi: 10.1016/j.plaphy.2014.09.010

White, P. J. (2002). Recent advances in fruit development and ripening: An overview. *J. Exp. Bot.* 53, 1995–2000. doi: 10.1093/jxb/erf105

Xiao, J., Zhang, H., Xing, L., Xu, S., Liu, H., Chong, K., et al. (2013). Requirement of histone acetyltransferases HAM1 and HAM2 for epigenetic modification of FLC in regulating flowering in *Arabidopsis*. *J. Plant Physiol.* 170, 444–451. doi: 10.1016/j.jplph.2012.11.007

Xiao, K., Chen, J., He, Q., Wang, Y., Shen, H., and Sun, L. (2020). DNA methylation is involved in the regulation of pepper fruit ripening and interacts with phytohormones. *J. Exp. Bot.* 71, 1928–1942. doi: 10.1093/jxb/eraa003

Xu, J., Xu, H., Liu, Y., Wang, X., Xu, Q., and Deng, X. (2015). Genome-wide identification of sweet orange (*Citrus sinensis*) histone modification gene families and their expression analysis during the fruit development and fruit-blue mold infection process. *Front. Plant Sci.* 6:607. doi: 10.3389/fpls.2015.00607

Yuan, L., Dai, H., Zheng, S., Huang, R., and Tong, H. R. (2020). Genome-wide identification of the HDAC family proteins and functional characterization of CsHD2C, a HD2-type histone deacetylase gene in tea plant (*Camellia sinensis* L. O. Kuntze). *Plant Physiol. Biochem.* 155, 898–913. doi: 10.1016/j.plaphy.2020.07.047

Yun, M., Wu, J., Workman, J. L., and Li, B. (2011). Readers of histone modifications. *Cell Res.* 21, 564–578. doi: 10.1038/cr.2011.42

Zhang, Z., Wang, B., Wang, S., Lin, T., Yang, L., Zhao, Z., et al. (2020). Genome-wide target mapping shows histone deacetylase complex1 regulates cell proliferation in cucumber fruit. *Plant Physiol.* 182, 167–184. doi: 10.1104/pp.19.00532

Zhong, S., Fei, Z., Chen, Y. R., Zheng, Y., Huang, M., Vrebalov, J., et al. (2013). Single-base resolution methylomes of tomato fruit development reveal epigenome modifications associated with ripening. *Nat. Biotechnol.* 31, 154–159. doi: 10.1038/nbt.2462

Zhou, Y., Yang, P., Zhang, F., Luo, X., and Xie, J. (2020). Histone deacetylase HDA19 interacts with histone methyltransferase SUVH5 to regulate seed dormancy in *Arabidopsis*. *Plant Biol.* 22, 1062–1071. doi: 10.1111/plb.13158

Zhu, Z., Sun, B., Cai, W., Zhou, X., Mao, Y., Chen, C., et al. (2019). Natural variations in the MYB transcription factor MYB31 determine the evolution of extremely pungent peppers. *New Phytol.* 223, 922–938. doi: 10.1111/nph.15853

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 15 September 2022
DOI 10.3389/fpls.2022.967143

OPEN ACCESS

EDITED BY
Neftali Ochoa-Alejo,
Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Mexico

REVIEWED BY
Shi Liu,
Northeast Agricultural University, China
Lourdes Gómez-Gómez,
University of Castilla-La Mancha, Spain
Ralf Welsch,
University of Freiburg, Germany
Chengsheng Gong,
Jiangsu Academy of Agricultural Sciences (JAAS), China

*CORRESPONDENCE
Charles Ampomah-Dwamena
Charles.dwamena@plantandfood.co.nz

SPECIALTY SECTION
This article was submitted to Plant Development and EvoDevo, a section of the journal Frontiers in Plant Science

RECEIVED 12 June 2022
ACCEPTED 29 July 2022
PUBLISHED 15 September 2022

CITATION
Ampomah-Dwamena C, Tomes S, Thrimawithana AH, Elborough C, Bhargava N, Rebstock R, Sutherland P, Ireland H, Allan AC and Espley RV (2022) Overexpression of *PSY1* increases fruit skin and flesh carotenoid content and reveals associated transcription factors in apple (*Malus* × *domestica*). *Front. Plant Sci.* 13:967143. doi: 10.3389/fpls.2022.967143

Overexpression of *PSY1* increases fruit skin and flesh carotenoid content and reveals associated transcription factors in apple (*Malus* × *domestica*)

Charles Ampomah-Dwamena[1]*, Sumathi Tomes[1], Amali H. Thrimawithana[1], Caitlin Elborough[1,2], Nitisha Bhargava[1], Ria Rebstock[1], Paul Sutherland[1], Hilary Ireland[1], Andrew C. Allan[1] and Richard V. Espley[1]

[1]The New Zealand Institute for Plant and Food Research Ltd., Auckland, New Zealand, [2]BioLumic Limited, Palmerston North, New Zealand

Knowledge of the transcriptional regulation of the carotenoid metabolic pathway is still emerging and here, we have misexpressed a key biosynthetic gene in apple to highlight potential transcriptional regulators of this pathway. We overexpressed phytoene synthase (*PSY1*), which controls the key rate-limiting biosynthetic step, in apple and analyzed its effects in transgenic fruit skin and flesh using two approaches. Firstly, the effects of PSY overexpression on carotenoid accumulation and gene expression was assessed in fruit at different development stages. Secondly, the effect of light exclusion on PSY1-induced fruit carotenoid accumulation was examined. PSY1 overexpression increased carotenoid content in transgenic fruit skin and flesh, with beta-carotene being the most prevalent carotenoid compound. Light exclusion by fruit bagging reduced carotenoid content overall, but carotenoid content was still higher in bagged PSY fruit than in bagged controls. In tissues overexpressing *PSY1*, plastids showed accelerated chloroplast to chromoplast transition as well as high fluorescence intensity, consistent with increased number of chromoplasts and carotenoid accumulation. Surprisingly, the expression of other carotenoid pathway genes was elevated in PSY fruit, suggesting a feed-forward regulation of carotenogenesis when this enzyme step is mis-expressed. Transcriptome profiling of fruit flesh identified differentially expressed transcription factors (TFs) that also were co-expressed with carotenoid pathway genes. A comparison of differentially expressed genes from both the developmental series and light exclusion treatment revealed six candidate TFs

exhibiting strong correlation with carotenoid accumulation. This combination of physiological, transcriptomic and metabolite data sheds new light on plant carotenogenesis and TFs that may play a role in regulating apple carotenoid biosynthesis.

KEYWORDS

carotenoid biosynthesis, transgenic apple, transcriptome, plastids, gene expression, phytoene synthase, transcription factors

Introduction

Carotenoids play essential roles in photosynthesis as well as conferring coloration in flower and fruit tissues for pollination, seed dispersal and increased consumer appeal (Zhu et al., 2010). These pigments are also noted for their health associations, notably as precursors of vitamin A and as scavengers of free oxygen species (Weber and Grune, 2012; Eggersdorfer and Wyss, 2018; Rodríguez-Concepcíon et al., 2018). Despite the association of carotenoid pigments with health benefits, and their presence in apple skin, most commercial apple cultivars have little or no pigmentation in the fruit flesh at maturity, making the regulation of secondary metabolic pathways controlling these pigments in the fruit flesh an important area of research (Ampomah-Dwamena et al., 2012; Delgado-Pelayo et al., 2014). Apple fruit are widely consumed globally and increasing phytochemical content, such as carotenoids, could well improve their nutritional and commercial value. Although the long juvenile period in apple presents a significant challenge to breeding, knowledge of the carotenoid metabolic pathway and its regulation can contribute to accelerated breeding of a higher-carotenoid apple.

Carotenoid accumulation in tissues is a balance between biosynthesis and metabolite breakdown, controlled by the upstream pathway biosynthetic enzymes (regulating flux) and downstream breakdown enzymes (metabolite turnover), respectively (Brandi et al., 2011; Rodríguez-Concepcíon et al., 2018). The carotenoid enzymatic pathway is situated in the plastids and metabolically connected to the methylerythritol 4-phosphate (MEP) pathway, which generates the C_{20} compound geranylgeranyl pyrophosphate (GGPP), and which also is a common precursor for gibberellins, tocopherols, and chlorophylls (Okada et al., 2000). In the carotenoid pathway, GGPP is a substrate for the first committed enzymatic condensation of two GGPP molecules to give the colorless phytoene (Nisar et al., 2015; You et al., 2020). This step, catalyzed by phytoene synthase (PSY), is a well-regulated rate-limiting step in many plant species (Welsch et al., 2010) and controls the carotenoid pathway flux, as indicated by the amount of phytoene that accumulates in the presence of an attenuated phytoene desaturase (PDS) (Schaub et al., 2018). PDS catalyzes a two-step desaturation of phytoene into phytofluene and then (tri-*cis*) zeta-carotene, which becomes a substrate for zeta-carotene isomerase (Z-ISO) for conversion to di-*cis* zeta-carotene (Chen et al., 2010; Beltrán et al., 2015). Subsequent desaturation and isomerization by zeta-carotene desaturase (ZDS) and carotene isomerase (CRTISO), respectively, result in lycopene, providing the substrate for the lycopene cyclases to produce alpha-carotene and beta-carotene in a bifurcated pathway step (Isaacson et al., 2004). The conversion of alpha-carotene to lutein, and beta-carotene to zeaxanthin, is through the actions of carotene hydroxylases (Kim and DellaPenna, 2006; Quinlan et al., 2012; Niu et al., 2020).

Carotenoid content in tissues can thus be altered by manipulating the expression of rate-limiting steps such as PSY to intensify pathway flux ('Push' strategy) or by reducing downstream enzymatic turnover of accumulated compounds using a 'Block' approach (Diretto et al., 2010; Zeng et al., 2015). In the latter, carotenoid biosynthetic enzymes or degradation enzymes such as carotenoid cleavage dioxygenases (CCDs) or nine *cis* epoxy-carotenoid dioxygenases (NCEDs) can be downregulated to increase carotenoid content (Ko et al., 2018). Although the 'Push' strategy has been widely utilized to increase carotenoid content in various tissues (Ducreux et al., 2005; Zhou et al., 2022), its success depends on the efficiency of subsequent pathway steps and having the appropriate storage subcellular compartments in these tissues. For instance, the increased carotenoid content in tomato fruit as a result of *PSY1* overexpression was dependent on fruit stage, feed-forward effect on pathway and abundance of chromoplasts required for carotenoid storage (Fraser et al., 2007). Six apple PSY genes are present in the 'Golden Delicious' genome sequence and their homologs have been characterized in other apple cultivars (Velasco et al., 2010; Ampomah-Dwamena et al., 2012, 2015; Zhang et al., 2018; Cerda et al., 2020). These studies have mainly focused on gene expression analysis and functional characterization in heterologous systems. However, increased apple *PSY* gene expression has indirectly been associated with increased fruit carotenoid content. The overexpression in apple of the Arabidopsis gene AtDXR, which is an enzyme in the upstream MEP pathway, increased fruit carotenoid content with associated increased expression of apple *PSY*

genes (Arcos et al., 2020). More recently, the overexpression of *MdAP2-34* transcription factor in apple also increased fruit carotenoid content and upregulated expression of *MdPSY2* (Dang et al., 2021). This positive relationship between apple PSY gene expression and fruit carotenoid content indicates a 'push' strategy can be utilized to boost apple carotenoid concentration.

Plastid differentiation into chromoplasts offers another approach (expanding capacity) to increase carotenoid accumulation. The gain-of-function *Orange* (*OR*) gene mutation in cauliflower and subsequently the transgenic overexpression of *ORs* has increased carotenoid content in many plants (Li and Van Eck, 2007; Park et al., 2016). The OR protein, in addition to promoting plastid differentiation, post-transcriptionally regulates PSY protein content and activity, while the golden SNP OR^{HIS} mutation has been linked to downregulation of beta-carotene hydroxylase expression, resulting in increased beta-carotene content (Tzuri et al., 2015; Zhou et al., 2015; Park et al., 2016; Chayut et al., 2017). These observations of OR functions provide a snapshot of the linkages between the different mechanisms (both translational and post-translational) controlling carotenoid accumulation and further highlight the complex regulation of this metabolic pathway.

The carotenoid pathway is subject to various degrees of regulation, which affect carotenoid biosynthesis, and determine the accumulation in both photosynthetic and non-photosynthetic tissues (Sun and Li, 2020). Whereas transcriptional regulation of the pathway has received significant attention and led to the identification of many carotenoid transcription factors (TFs; Stanley and Yuan, 2019), post-transcriptional and post-translational mechanisms also exert some control on this process. Alternative splicing of carotenoid genes presents a way to exert control of the pathway, especially with pivotal genes such as *PSY*. Alternative splicing of Arabidopsis PSY has been linked with different translation efficiency, owing to the variation in their 5' untranslated regions, and different regulatory modules that respond to either carotenoid flux or abiotic stress (Álvarez et al., 2016). In crocus, spliced variants of *CsPSY1b* exhibited differential expression, with the presence of the intron-containing variant associated with tissues having reduced transcript levels (Ahrazem et al., 2019). Alternative splicing of wheat PSY-A1 resulted in four mRNA variants that affected the abundance of the wild-type transcript, which was the only one that produced an enzymatically active protein (Howitt et al., 2009). In tomato, alternative *trans*-splicing of *PSY1* resulted in a longer chimeric variant that was responsible for the yellow flesh *yft2* fruit phenotype (Chen et al., 2019). Additional post-translational control mechanisms play a role in altering the activity of carotenoid enzymes. For instance the abundance of PSY in the plastids is regulated by the OR protein with the amount of PSY protein significantly reduced in the *or* mutants of Arabidopsis and melon (Zhou et al., 2015; Chayut et al., 2017). While the mechanism for this OR-induced stability is not fully established, the holdase chaperone activity of OR plays a role by preventing PSY protein aggregation and subsequent degradation (Park et al., 2016). The activity of ATP-dependent serine type caseinolytic (Clp) protease, which targets protein for degradation, regulates the protein content and activity of PSY and other downstream pathway enzymes (Welsch et al., 2018). The silencing of the Clp protease activity in tomato resulted in increased expression of the *OR* gene and enhanced beta-carotene accumulation in fruit, highlighting the link between the holdase and protease activities of OR and Clp, respectively (D'Andrea et al., 2018; D'Andrea and Rodriguez-Concepcion, 2019).

Transcriptional regulation of the carotenoid pathway appears to vary among plant species, with a diversity of TFs reported to control the pathway. We previously identified a kiwifruit *MYB7* gene whose expression induced key pathway genes, implicating it has a role in carotenogenesis (Ampomah-Dwamena et al., 2019). MYB TFs such as CrMYB68 in citrus function as negative regulators of CrBCH2 and CrNCEDs (Zhu F. et al., 2017), whilst RCP1, another R2R3 MYB belonging to subgroup 21, is a positive regulator of carotenoid accumulation in *Erythranthe lewisii* flowers (Sagawa et al., 2016). Similarly, *Medicago truncatula* White Petal 1 protein (a homolog of AtMYB113 and in a different subgroup from RCP1) also transcriptionally activates carotenoid biosynthetic genes through its interaction with other proteins (Meng et al., 2019). Reports of potential transcriptional regulators of the carotenoid pathway belonging to TF classes, such as NACs, MADS, ERF, and bHLH, have been reviewed recently, highlighting that carotenoid pathway regulation may have evolved independently in different plant species (Stanley and Yuan, 2019). This apparent lack of conserved TF activity among plant species and the lack of pathway mutants makes it particularly important to adapt new strategies for identifying carotenoid TFs in discrete species. It is possible that the perturbations caused by transgene expression, for instance, could uncover gene network relationships controlling the observed phenotype. The overexpression of *PSY* in tomato, which increased carotenoid accumulation, also elicited various molecular and metabolic responses to uncover carotenoid pathway regulation in this species (Fraser et al., 2007). The molecular changes resulting from such induced expression could help outline the transcriptional regulatory networks involved in carotenogenesis and potentially highlight candidate TF genes involved in this metabolic process. In this study, we set out to understand the mechanisms controlling the apple carotenoid biosynthetic pathway by overexpressing the rate-limiting *phytoene synthase 1* (*PSY1*). Analyses of physiological, metabolic and transcriptional changes shed new light on the regulation of carotenoid metabolic pathway in this species.

Results

PSY1 expression increased fruit carotenoid content

We previously described six *PSY* genes present in the 'Golden Delicious' apple genome and characterized their homologs in 'Royal Gala' (Velasco et al., 2010; Ampomah-Dwamena et al., 2015). We selected *PSY1* (located on chromosome 17), whose expression strongly correlated with carotenoid accumulation in apple fruit skin and flesh, and expressed it constitutively using the CMV:35S promoter in 'Royal Gala' apple. To test if PSY1 was capable of creating enhanced pathway flux, we first generated transformed apple callus, with and without norflurazon (NFZ), an inhibitor of the next pathway step (PDS) and then measured phytoene content (Schaub et al., 2018; Koschmieder and Welsch, 2020). Without NFZ, the phytoene content in the PSY calli was up to 15-fold higher than in WT (**Supplementary Figure S1**). In the presence of NFZ, phytoene content increased a further two- to three-fold in PSY calli and was approximately six times higher than in WT callus on the same treatment. Other downstream compounds such as beta-carotene, alpha-carotene, lutein and zeaxanthin that accumulated in these calli, showed little to no change with NFZ treatment (**Supplementary Figure S1**). Next, we generated stably transformed *PSY1* overexpressed (OE) transgenic plants, with wild type (WT) as control, for analyses. During tissue culture regeneration and growth in soil, vegetative tissues of transgenic plants showed normal phenotype. However, during fruit development, the transgenic plants displayed a deep yellow fruit color phenotype (**Figure 1A**).

Carotenoid accumulation was analyzed in the fruit skin and flesh tissues of three transgenic PSY lines with sufficient fruit numbers on the tree (OE-1, OE-5, and OE-7), and WT as control at four developmental stages: 90, 120, 135, and 150 days after pollination (D).

The PSY fruit, at the early stages (90 and 120 D), had green skin and white flesh color and, in the later stages (135 and 150 D), both flesh and skin developed strong yellow pigmentation (**Figure 1A**). The control wild-type (WT) fruit, in contrast, developed red streaks of anthocyanin pigments on a green background skin at 120 D, as expected. After this stage, the fruit skin began to de-green to reveal a yellowish background skin color. The WT fruit flesh color was whitish during the early fruit stages and acquired a creamy color when fruit were at maturity (150 D).

High-performance liquid chromatography (HPLC) analysis of carotenoid content and composition showed that beta-carotene was the dominant compound accumulating in both skin and flesh tissues, with the content increasing with fruit development. We also observed increased amounts of esterified compounds in both PSY and WT fruit at 135 and 150 D. Carotenoid content in OE-1 fruit did not increase compared with WT at all the four fruit stages we examined while, a three-fold to eight-fold increase in total carotenoid content (TCC) was observed in OE-5 and OE-7 fruit skin. TCC in fruit flesh of WT, OE-5 and OE-7 were 21.2 ± 1.93, 185.8 ± 8.7, 50.9 ± 6.5 μg/g FW at 120 D; 17.9 ± 0.1, 141.3 ± 6.9, 56.6 ± 10.0 μg/g FW at 135 D and 95.3 ± 0.4, 155.9 ± 3.4, 160.8 ± 17.0 μg/g FW at 150 D, respectively (**Figure 1B**). TCC in OE-5 and OE-7 flesh thus increased by 8.7, 2.4-fold at 120 D; 7.9, 3.2-fold at 135 D; and by 1.6, 1.7-fold at 150 D, respectively when compared with the WT flesh. The concentration of beta-carotene, the predominant compound in these tissues, increased by 3.2-fold in OE-5 (42.9 ± 12.6 μg/g FW) and 5.9-fold in OE-7 (79.8 ± 4.5 μg/g FW) compared with WT (13.3 ± 3.6 μg/g FW) at 150 D. This represented 27% (OE-5) and 50% (OE-7) of total carotenoids, compared with 14% for WT at that stage. In contrast, chlorophyll in fruit skin or flesh decreased as the fruit developed, with no clear difference between the PSY lines and WT fruit. Total chlorophyll content in WT flesh was higher than in PSY fruit at most of the stages we monitored (**Figure 1B**).

Plastid changes in *PSY* transgenic fruit

Plastid transition in fruit was examined using confocal microscopy analysis of WT, OE-5, and OE-7 fruit at 120, 135, and 150 D. The chloroplasts, containing chlorophyll pigments, emitted red fluorescence under excitation, while the chromoplasts emitted green fluorescence. At 120 D, the WT fruit contained predominantly chloroplasts, which emitted red fluorescence, whilst in the two PSY lines at the same stage, there was a mixed population of chloroplasts and chromoplasts (green fluorescence), suggesting a hastened plastid transition. At 135 D, the WT had both plastid types in almost equal abundance, while mostly chromoplasts were present in the PSY fruit. At 150 D, which coincided with fruit ripening, mostly chromoplasts were observed in both PSY and WT tissues; however, in the PSY fruit these plastids were in abundance compared with their sparse distribution in WT fruit (**Figure 2**). The fluorescence emission spectra, a measure of carotenoid (550–650 nm) and chlorophyll (650–700 nm) content (Kilcrease et al., 2013; D'Andrea et al., 2014), further confirmed the increased plastid capacity in PSY fruit. Both the 120 D WT and PSY fruit showed a chlorophyll peak, but the PSY fruit showed additional carotenoid peaks, confirming early accumulation of carotenoids in the transgenic fruit. In both 135 D and 150 D stages, higher carotenoid fluorescent intensity was observed within PSY fruit than in the WT.

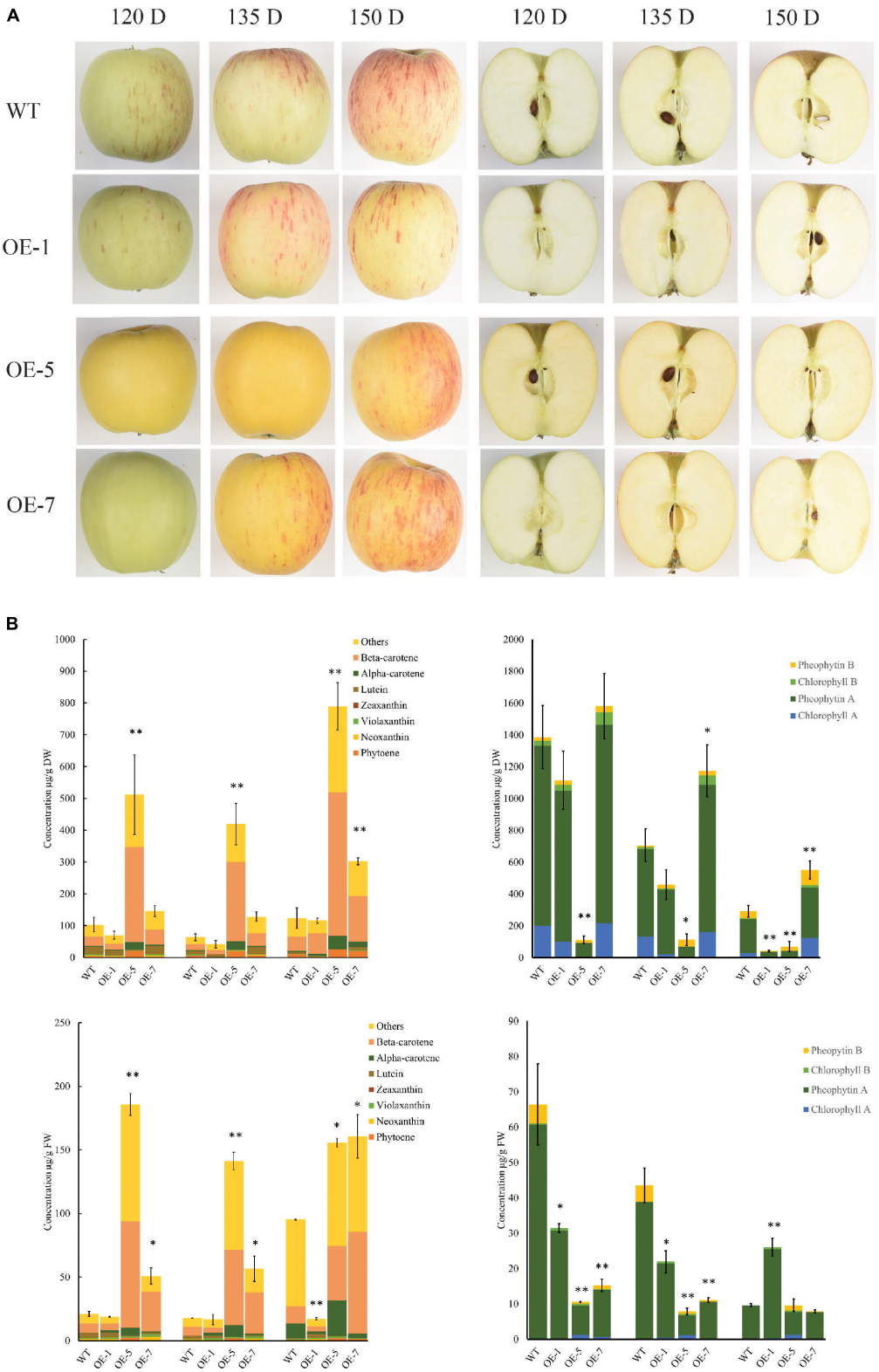


FIGURE 1
PSY1 expression increased carotenoid accumulation in 'Royal Gala' apple fruit. **(A)** Fruit of WT and PSY transgenic lines (OE-1, OE-5, and OE-7) at different developmental stages. **(B)** Bar graphs of carotenoid content measured by HPLC as beta-carotene equivalents (left column) and chlorophyll (right column) in fruit skin (top panel) and fruit flesh (bottom panel) in wild type and transgenic lines OE-1, OE-5, and OE-7. The error bars represent the standard errors of the mean of three biological replicates, with each replicate a pool of 5–7 fruit. Bar graph with asterisk show significant difference from WT at the same fruit stage using Dunnett's test (*$P < 0.05$; **$P < 0.01$).

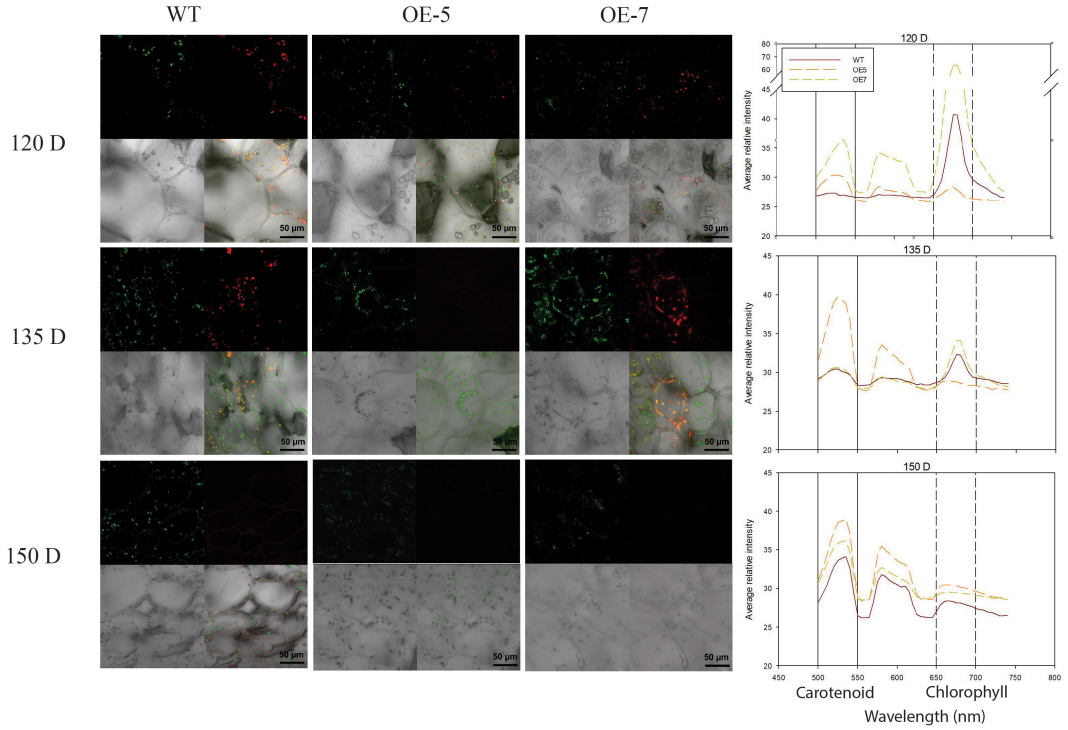


FIGURE 2
Visualization of chlorophyll and carotenoid autofluorescence in wild type (WT) and PSY 'Royal Gala' fruit. Fresh fruit tissues, at 120, 135, and 150 D, were analyzed by confocal microscopy to show plastids containing chlorophyll (red), carotenoid (green) and both pigments (yellow). Graph represents fluorescence emission spectra of WT and PSY transgenic tissue. Fluorescence emission between 500 and 600 nm represents carotenoids and 650–750, chlorophyll pigments.

Carotenoid gene expression in transgenic fruit

To understand the molecular mechanism underlying the increased carotenoid accumulation in PSY fruit, the relative transcript levels of carotenoid genes in fruit skin and flesh tissues from the OE-5, OE-7, and WT lines were assessed by real-time quantitative PCR (qPCR). *PSY1* expression in the transgenic lines was high (**Figure 3A**) and increased with development due to the cumulative contribution of both the endogenous gene and the 35S-driven transgene expression. In fruit skin, the expression of genes including *PSY2*, *ZDS1*, *ZDS2 CRTISO*, *LCB2* and *BCH1*, increased in the transgenic PSY lines compared with the WT and this was consistent across all developmental stages examined. Other genes however, including *PDS*, *ZISO*, *LCB1*, *BCH2* and *ZEP2*, showed higher expression in the WT fruit skin than in PSY lines at the 90 D and 120 D stages, but the trend was reversed at 135 and 150 D stages, with elevated expression in PSY lines. *LCYE* expression was higher in WT fruit than *PSY* at all four stages, while *ECH* showed higher expression in PSY fruit only at the 135 D stage (**Figure 3A** and **Supplementary Figure S2**). There was a similar trend for carotenoid gene expression in fruit flesh tissues, where transcript levels increased in the PSY fruit compared with the WT control, with the exception of the *LCB2* and *BCH2* genes, which showed higher transcript levels in the WT than in the PSY lines. The expression of *PDS*, *ZDS1*, *ZDS2*, *CRTISO* and *LCB1* was higher in the PSY lines than in the WT across all the fruit stages. In contrast, expression of *BCH1* and *BCH2* in PSY fruit reduced, compared with the WT, in the first three fruit stages but then increased at the 150 D stage (**Figure 3B**). Overall, PSY overexpression resulted in the upregulation of carotenoid pathway genes acting upstream of beta-carotene, and downregulation of genes downstream of beta-carotene. It appears that there is a coordinated regulatory response occurring in association with the mis-expression of PSY.

Transcriptomic analysis identifies differentially expressed genes associated with *PSY1* expression

Transcriptional changes associated with *PSY1*-induced carotenoid accumulation in apple fruit flesh were characterized by transcriptome sequencing of OE-7, which showed a more consistent fruit carotenoid accumulation pattern, and WT

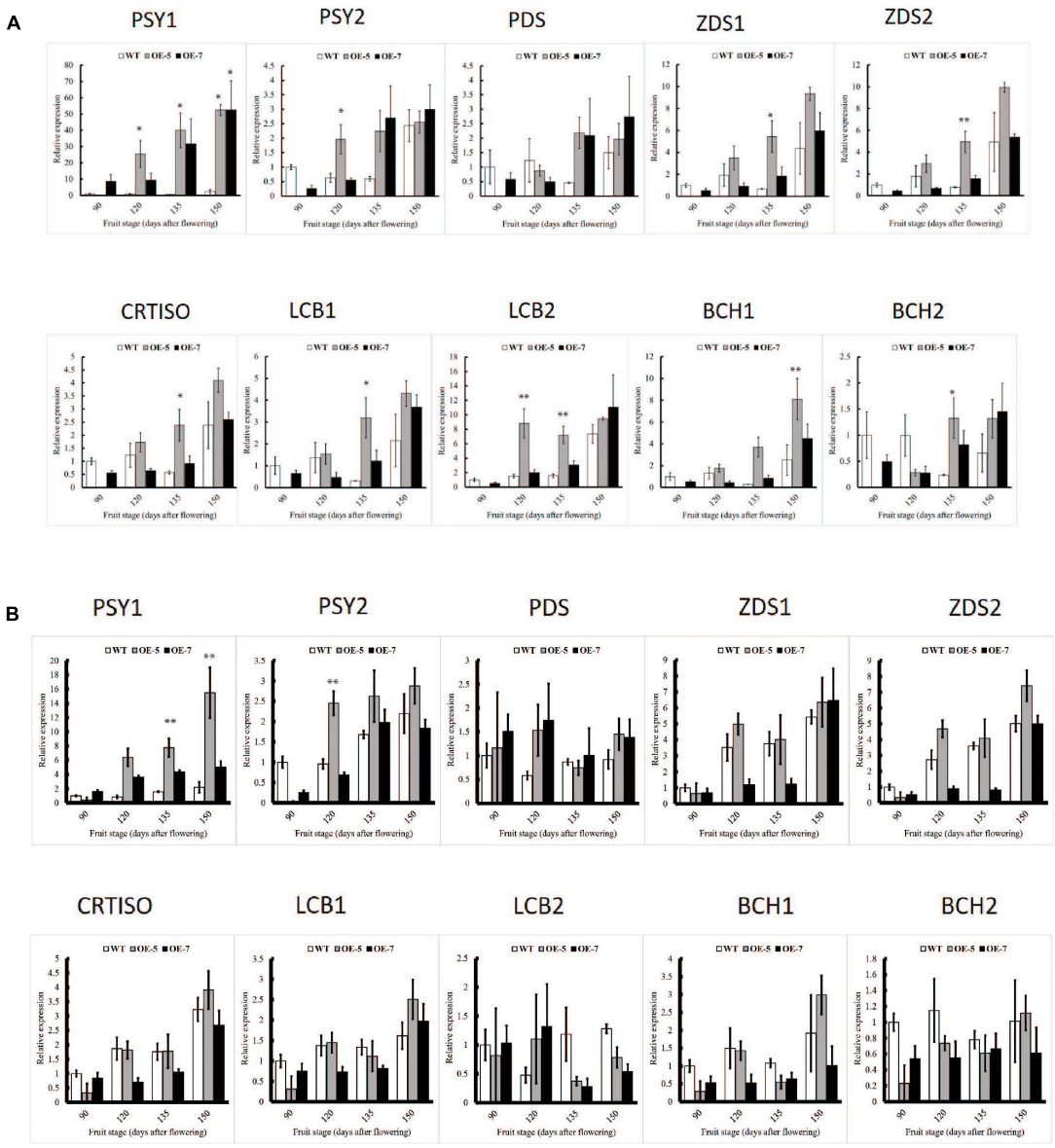


FIGURE 3
Comparison of carotenoid gene expression in wild-type (WT) and PSY 'Royal Gala' apple lines, OE-5 and OE-7, as determined by qPCR, relative to expression of housekeeping genes (*MdActin* and *MdEF1A*) in fruit skin **(A)** and fruit flesh **(B)**. Bars represent means and SE of four biological replicates. PSY, phytoene synthase; PDS, phytoene desaturase; ZDS, zeta carotene desaturase; CRTISO, carotene isomerase; LCB, lycopene beta-cyclase; BCH, beta-carotene hydroxylase. Bar graph with asterisk shows significant difference from WT at the same fruit stage using Dunnett's test (*$P < 0.05$; **$P < 0.01$).

at four stages (90, 120, 135, 150 D). A total sequence output of 180 GB from paired-end sequencing of 24 samples yielded an average of 25 million map-able reads per sample. Sequence reads were mapped to the 'Golden Delicious' double haploid (GDDH) genome (Daccord et al., 2017) to identify differentially expressed genes (DEGs), where the percentage of uniquely mapped reads was over 94%. Principal component analysis (PCA) of expression data showed that PC1 and PC2 explained 49% and 39% of variance, respectively, grouping samples according to treatment and fruit developmental stages (**Supplementary Figure S3**). For an overview of expressed carotenoid pathway genes in fruit, we analyzed reads from PSY fruit and WT that mapped to 54 gene models identified from the GDDH genome (**Supplementary Table S2**). Carotenoid genes that showed more than 1.5-fold increase ($p < 0.05$) in at least two PSY fruit stages included *GGPPS*, *PSY*, *Z-ISO*, *CRTISO* and *LCB*. Among the down-regulated pathway genes were *LCYE*, *BCH* and *ZEP* and showed

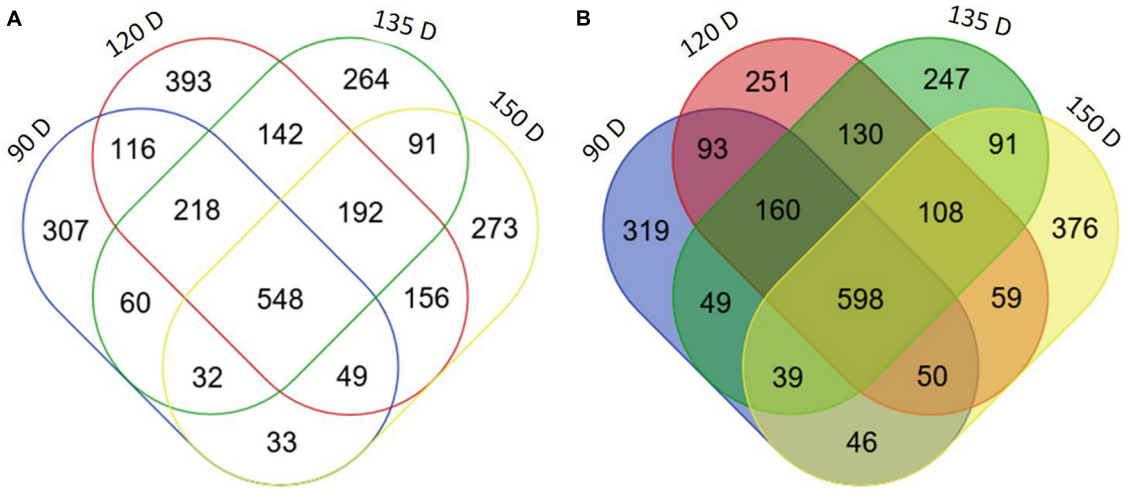


FIGURE 4
Differentially expressed genes in PSY transgenic apple fruit. Venn diagram of number of differentially expressed genes (log_2 fold change ≥ 2, $P < 0.05$) between WT and OE-7 fruit at four developmental stages (90, 120, 135, 150 D). Upregulated **(A)**, downregulated **(B)** in PSY fruit flesh.

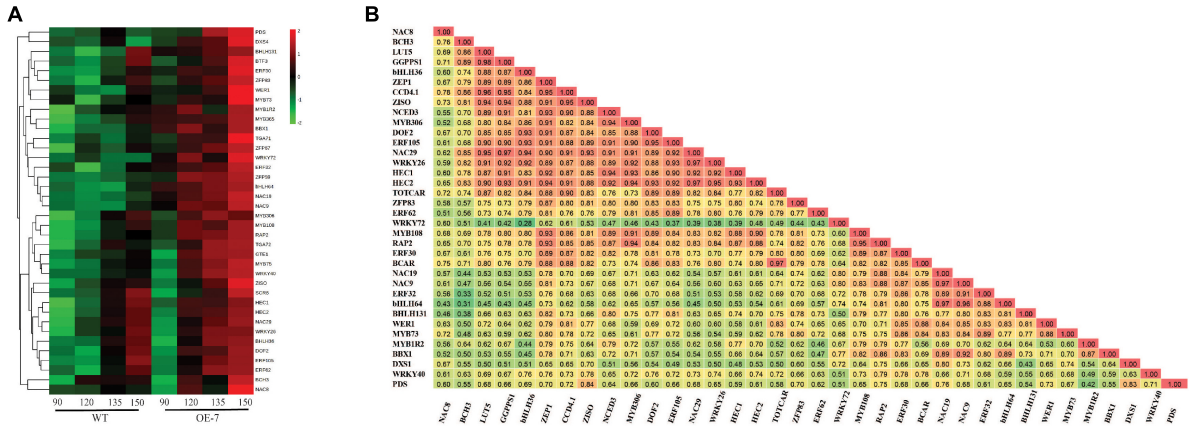


FIGURE 5
Analysis of carotenoid-associated transcription factor genes. **(A)** Heat map of differentially expressed transcription factor genes (**Supplementary Table S4**) in apple fruit flesh during development. The average gene expression of biological replicates at each of the four stages: 90, 120, 135, and 150 D, from WT and PSY transgenic line OE-7, were clustered using Euclidean distance relationships. Green to red color gradient indicates low to high relative gene expression. **(B)** Correlation matrix of gene transcripts and carotenoid metabolites (β-carotene, BCAR; total carotenoid content, TOTCAR) in transgenic PSY fruit. Values represent correlation coefficients with color gradient from green to red indicating weak to strong correlations at $P < 0.05$.

that *PSY* mis-expression affected the expression of the other carotenoid pathway genes.

With a log_2-fold cut-off ± 2 ($p < 0.05$) and a false discovery rate threshold of $p < 0.05$, the number of upregulated and downregulated DEGs in PSY fruit compared with WT control were: 90 D (1354 up, 1363 down), 120 D (1449 up, 1814 down), 135 D (1422 up, 1547 down) and 150 D (1367 up, 1374 down) (**Figures 4A,B**). In comparison, the number of DEGs common to two or more fruit stages was 1637 upregulated and 1423 downregulated. However, since DEGs common to all four fruit stages are likely to include genes specifically responding to the perturbation caused by *PSY1* mis-expression, as well as any subsequent metabolic changes, the 548 upregulated and 598 downregulated genes common to all four fruit developmental stages, referred to as *PSY1*-associated genes (PSYAGs; **Supplementary Table S3**), were further examined. The most upregulated PSYAG was a BAG family molecular chaperone regulator (MD01G1073600), a homolog of AtBAG7, implicated in unfolded protein responses. Upregulated genes included 1-aminocyclopropane-1-carboxylate oxidase 1 (MD15G1205100), squalene monooxygenase, and transcription factors such as

ethylene responsive factors (MdAP2D2, MD08G1060000), WRKY75 (MD09G1008800, similar to AtWRKY75, implicated in anthocyanin accumulation during stress), a homolog of AtMYB15, implicated in lignin biosynthesis, SEPALLATA1 (AGAMOUS-like 2) and a heat shock MYB close to BOS1. Among the downregulated PSYAGs were auxin-responsive proteins, beta-carotene isomerase D27, floral homeotic protein AGAMOUS and topless-related protein, transcriptional repressors, a senescence-related gene (SRG1) and MYB91, a homolog of Antirrhinum PHANTASTICA. To gain insight into the processes responding to the *PSY1* expression in apple fruit, we analyzed gene ontology (GO) term enrichment of the PSYAGs in the biological process (BP) and molecular function (MF) categories. The results showed enrichment for processes such as the carbohydrate metabolic process, lipid and fatty acid metabolic processes (BP category), and DNA binding transcription factor, catalytic and hydrolase activities in the MF category (**Supplementary Figure S4**).

Weighted gene co-expression network analysis identified clusters of genes associated with carotenoid accumulation in apple

Weighted gene co-expression network analysis (WGCNA) identified clusters of genes that were highly correlated with three key traits: TCC, beta-carotene content, and fruit development stage. Using a $p < 0.05$ cut-off, three clusters (Brown, Green, and Red modules) showed significant relationships across all the three traits. The Brown and Green-coded modules displayed positive gene significance, which is based on the correlation of the gene expression profile with sample trait (Langfelder and Horvath, 2008), and the Red-coded module negative gene significance, with respect to total carotenoid accumulation (**Supplementary Figure S5**). The Brown-coded module had 1504 genes showing gene significance (GS) $\geq$ 0.6, and contained the carotenoid genes *DXS*, *PDS* and *Z-ISO*, as well as *BCH*, *ZEP* and *NCED5*. The Green-coded module had 1770 members with a GS $\geq$ 0.6 and included carotenoid-related genes such as *ZDS*, *BCH*, *LUT5* and *OR*. The presence of known carotenoid genes in these modules suggests the other genes in these sets are likely to be associated with carotenoid metabolism and thus provide the opportunity to discern transcriptional regulation of this metabolic pathway.

To identify TFs that were associated with apple carotenoid biosynthesis, we performed a cluster analysis of the carotenoid and TF genes in the Brown-coded module, using the average expression values of biological replicates and applying a GS cut-off > 0.6, which identified 34 TFs (**Figure 5A** and **Supplementary Table S4**). A Pearson correlation matrix between these genes and carotenoid contents further highlighted the relationship between these candidate TFs and carotenoid metabolism (**Figure 5B**). The apple *PDS* is a single-copy gene in the duplicated *M. domestica* genome and expected to have an important limiting role in the apple carotenoid pathway (Velasco et al., 2010; Daccord et al., 2017). *PDS* expression in fruit flesh was strongly correlated with *ZISO* (r = 0.84) and *DXS* (r = 0.83), suggesting these genes may be co-regulated Among the TF genes, *PDS* showed highest correlation with *NAC19* (r = 0.76) and high correlations with *ZISO* (r = 0.69) and *DXS* (r = 0.80), respectively, suggesting that *DXS*, *PDS* and *ZISO* may be co-regulated in apple fruit.

PSY1 expression increased fruit carotenoid content in the absence of light

To gain insight into how *PSY1* overexpression interacts with the effects of light on carotenoid accumulation in apple, we selected another three PSY lines, OE-2, OE-3, and OE-4 (which had sufficient numbers of fruit), and WT control, and bagged their fruit at 30 D until 150 D. Both bagged and non-bagged PSY fruit showed yellow skin color, with little or no red anthocyanin pigmentation. The flesh of PSY fruit, irrespective of bagging treatment, also showed yellow pigmentation, suggesting increased carotenoid accumulation (**Figure 6A**). In contrast, the bagged WT fruit had reduced pigmentation in the skin and flesh compared with the non-bagged WT fruit, which had skin with red stripes on a yellow background and a creamy flesh (**Figure 6A**). HPLC data showed there were significant changes to fruit carotenoid content and profile because of the bagging treatment (**Figure 6B**). In fruit skin, bagging reduced TCC ~four-fold in the WT control (from 48 $\pm$ 2.3 to 12.1 $\pm$ 2.9 μg/g FW), and by ~two-fold in the PSY lines OE-2 (498.8 $\pm$ 88.6 to 264.9 $\pm$ 21.7 μg/g FW), OE-3 (577.3 $\pm$ 60.7 to 327.2 $\pm$ 6.5 μg/g FW), and OE-4 (153.3 $\pm$ 21.2 to 81.5 $\pm$ 1.2 μg/g FW). Similarly, bagging reduced TCC in fruit flesh by 2.5-fold in the WT control (from 35.1 $\pm$ 3.6 to 13.9 $\pm$ 2.3 μg/g FW), while in PSY lines, TCC reduced by 1.5- to 3-fold in OE-2 (131.8 $\pm$ 20.8 to 90.2 $\pm$ 4.4 μg/g FW), OE-3 (269.7 $\pm$ 34.7 to 120.5 $\pm$ 8.5 μg/g FW), and OE-4 (62.0 $\pm$ 4.4 to 20.6 $\pm$ 1.4 μg/g FW). Overall, bagging reduced TCC in fruit skin and flesh of both WT and PSY lines. However, *PSY1* overexpression still significantly induced carotenoid accumulation in these tissues even in the absence of light. The reduced carotenoid content in bagged fruit was consistent with the reduced numbers of plastids in these tissues. Bright-field microscopy analysis showed fewer plastids in bagged PSY and WT fruit compared with their respective non-bagged fruit, while in both bagged and non-bagged fruit, PSY tissues showed more plastids than the WT (**Figure 6C**).

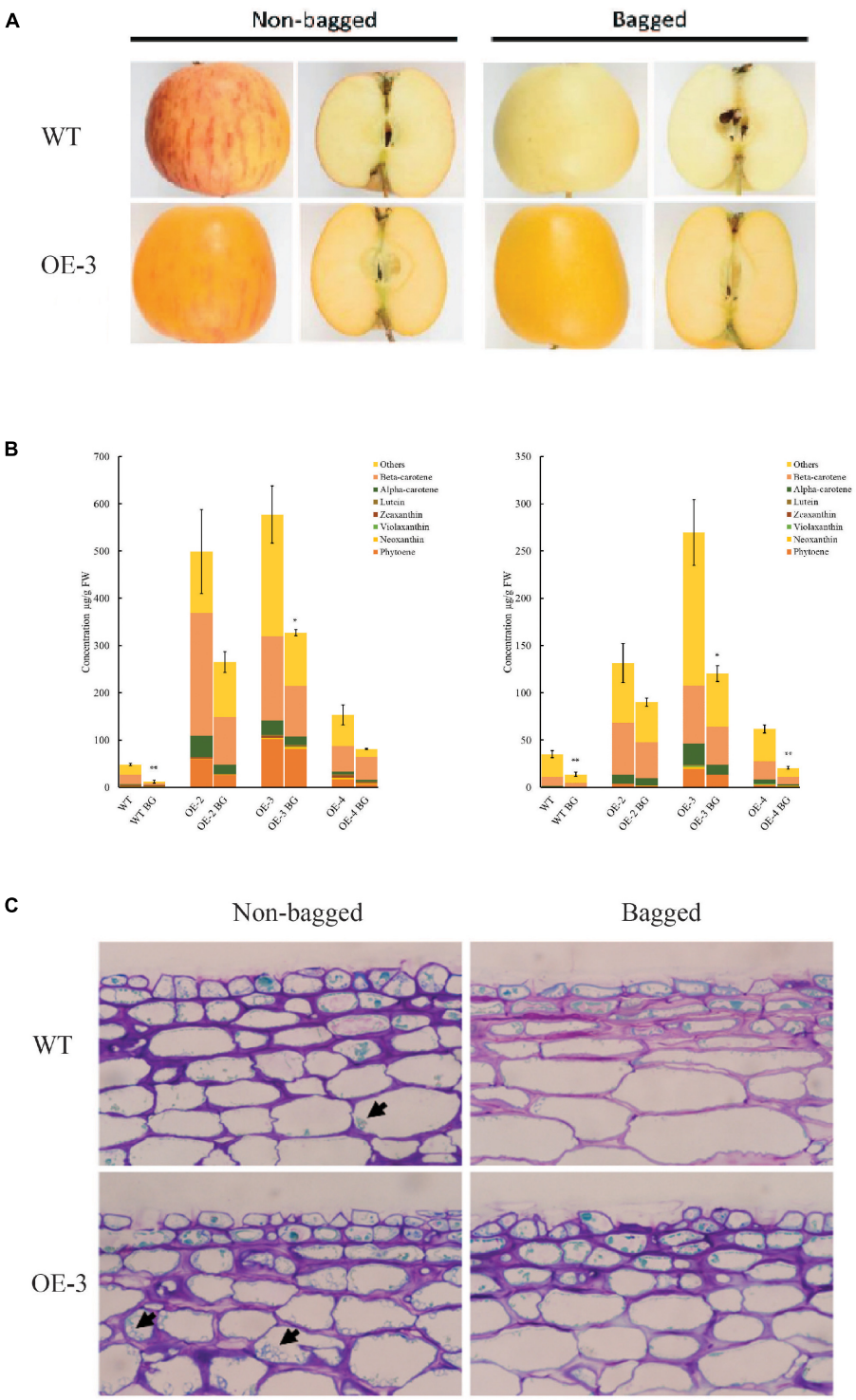


FIGURE 6
PSY1 increased fruit carotenoid content in apple in the absence of light. **(A)** Representative bagged and non-bagged fruit of WT and PSY line OE-3 at 150 D. **(B)** Graphs showing carotenoid content in fruit skin (left) and flesh (right) of different PSY lines (OE-2, OE-3, OE-4) and WT. Error bars represent standard error of total carotenoid contents of three biological replicates. Bar graph with asterisk show significant difference from the non-bagged fruit of the same line using Dunnett's test (*$P < 0.05$; **$P < 0.01$). **(C)** Stained fruit sections of WT and OE-3 displaying plastids (arrowed) in the cells.

Transcriptional changes in bagged fruit flesh

To ascertain transcriptional changes associated with the bagging treatment, the transcriptome of 150 D fruit flesh from three PSY lines OE-2, OE-3, OE-4, and the WT was analyzed. With a $\log_2$-fold cut-off $\pm$ 2 ($p < 0.05$) and a false discovery rate threshold of $p < 0.05$, bagged versus non-bagged WT fruit had 335 DEGs (152 up, 182 down), compared with 136 DEGs (34 up, 132 down) in bagged PSY fruit, present in at least two OE lines. Twenty-three of these DEGs (11 up, 12 down) were present in both bagged WT and PSY fruit. In contrast, when comparing WT and PSY fruit (bagged and non-bagged) we identified 530 upregulated and 1500 downregulated genes in PSY lines, a snapshot of a PSY-induced transcriptional response. Bagging also reduced expression of known light-regulated genes such as *ELONGATED HYPOCOTYL 5* (*HY5*), *chlorophyll a-b binding protein*, phytochrome a and *ribulose-l,5-bisphosphate carboxylase oxygenase*.

Assessment of the 54 selected carotenoid and related genes (**Supplementary Table S2**) in WT fruit revealed the expression of *DXS*, *GGPPS*, *ZDS*, *CRTISO*, *BCH*, *ZEP*, *VDE,* and *ECH* was downregulated in the bagged WT fruit. This downregulation of carotenoid genes was reversed in the PSY fruit. To ascertain whether *PSY1* modulated the expression of these genes in the absence of light, we compared bagged WT with bagged PSY fruit of the three independent OE lines. The expression of *DXS*, *GGPPS*, *PDS*, *ZDS*, *CRTISO*, *BCH,* and *ZEP* was upregulated in bagged PSY fruit of all or at least two transgenic lines, again revealing a feedforward effect on the pathway genes by *PSY1* mis-expression, independent of light. GO analysis of both WT and OE-3 DEGs showed enrichment for fatty acid biosynthesis (GO:0006633), photosynthesis (GO:0015979) and carotenoid biosynthesis (GO:0016117), all of these biological processes had negative z scores indicating their downregulation due to bagging treatment (**Figure 7A**).

To identify TFs associated with the *PSY1*-induced gene expression response in bagged fruit, differentially expressed TF genes (**Supplementary Table S5**) were selected for cluster analysis, using their average expression values. These divided into two clusters and clearly differentiated between WT and PSY fruit (**Figure 7B**). These TFs compared with those differentially expressed during fruit development of WT and PSY lines (**Figure 5A**) identified six TFs common to both data sets: *MdbHLH36* (MD09G1233000), *MdDOF2* (MD05G1018200), *MdRAP2* (MD17G1152400), *MdERF62* (MD04G1009000), and *MdMYB73* (MD15G1288600) and *MdNAC9* (MD01G1093500).

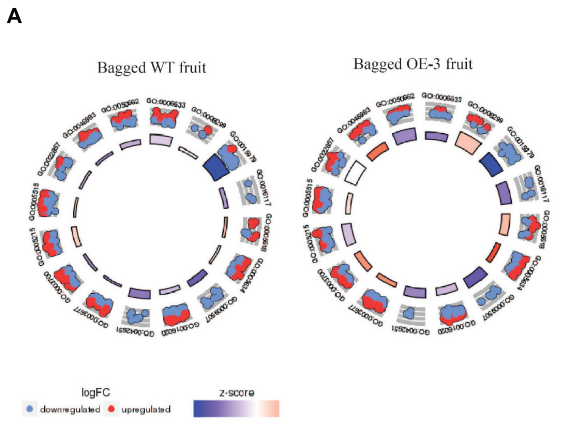
Expression of these TFs was validated using qPCR of fruit skin and flesh from different PSY transgenic lines. The expression of these genes, in both fruit skin and flesh, showed an increasing trend during fruit development that was consistent with the carotenoid accumulation in these tissues. Their expressions were also higher in the PSY lines compared with the WT at almost all the fruit stages examined (**Supplementary Figure S6**). Pearson correlation analysis showed strong correlation between the expression of these genes and carotenoid accumulation in both fruit skin and flesh tissues (**Figures 8A,B**). The gene expression of *MdbHLH36* was highly correlated with total carotenoid accumulation in the fruit flesh ($r = 0.93$), but had reduced correlation in the skin ($r = 0.52$). In contrast, *MdDOF2* and *MdRAP2* were highly correlated with carotenoid accumulation in both fruit skin ($r = 0.82$, 0.82) and flesh ($r = 0.93$, 0.90), respectively (**Supplementary Table S6**). Overall, the identification and expression of these TF genes suggest that they may have important roles during carotenogenesis in apple.

Discussion

Phytoene synthase controls carotenoid biosynthesis in apple

Phytoene synthase controls an important step, and acts as a limiting factor, in the carotenoid pathway (Rosas-Saavedra and Stange, 2016; Ahrazem et al., 2019). In this study, we analyzed *PSY1* overexpression in apple fruit to understand carotenoid regulation in this crop species. A deep yellow fruit pigmentation phenotype was attributed to increased carotenoid compounds in both skin and flesh tissues, with beta-carotene as the predominant compound. The accumulation of these compounds was probably due to an increased pathway flux, since the precursor GGPP is a common substrate for the chlorophyll, gibberellin and tocopherol pathways (Okada et al., 2000; Zhou et al., 2017). This altered pathway flux was illustrated with regenerated PSY leaf calli, which in the presence of the PDS inhibitor norflurazon, showed increased phytoene accumulation (Schaub et al., 2018). The increased phytoene content in PSY calli, with and without NFZ treatment, clearly indicated that *PSY1* overexpression in apple increased pathway flux in these tissues. Interestingly, apple has six *PSY* genes (compared with a single copy in Arabidopsis), with multiple *PSYs* expressed in fruit tissues. The increased carotenoid accumulation in response to *PSY1* overexpression confirms that the first committed enzyme step is a limiting factor in apple (Ampomah-Dwamena et al., 2015).

The enhanced carotenoid content in PSY fruit was associated with accelerated plastid transition and increased abundance of chromoplasts, as revealed by confocal microscopy. This metabolite-induced chromoplast abundance is probably mediated by OR, which was found in our weighted gene correlation network analysis (WGCNA), and has a role in plastid-induced carotenoid accumulation (Lu et al., 2006;



ID	Description
GO:0006633	fatty acid biosynthetic process
GO:0008299	isoprenoid biosynthetic process
GO:0015979	photosynthesis
GO:0016117	carotenoid biosynthetic process
GO:0005618	cell wall
GO:0005634	nucleus
GO:0009507	chloroplast
GO:0016020	membrane
GO:0042651	thylakoid membrane
GO:0003677	DNA binding
GO:0003700	DNA binding transcription factor activity
GO:0005215	transporter activity
GO:0005515	protein binding
GO:0022857	transmembrane transporter activity
GO:0046983	protein dimerization activity
GO:0050662	coenzyme binding

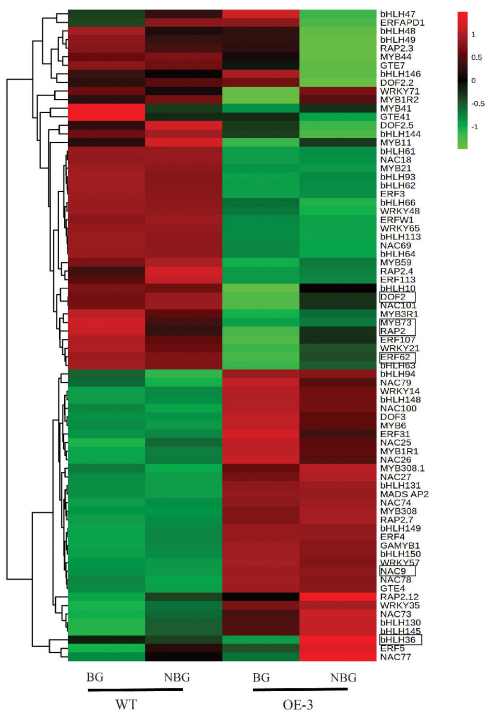


FIGURE 7
Differentially expressed genes (DEGs) in bagged apple fruit. **(A)** Go Plot R of DEGs in bagged fruit of WT and OE-3 fruit. The scatter plots show $\log_2$ fold change (logFC) for each gene under the gene ontology (GO) number. Red dots indicate upregulated genes and blue dots show downregulated genes. The *z*-score bars are a measure of the extent to which each identified process is upregulated or downregulated. **(B)** Heat map of differentially expressed transcription factor genes in bagged fruit. Clustering done using the average expression values of three biological replicates of bagged (BG) and non-bagged (NBG) fruit of WT and OE-3 transgenic line. Green to red color gradients indicate low to high relative gene expression. The highlighted TFs are common to the bagging and fruit development (**Figure 5A**) data sets.

Chayut et al., 2017; Feder et al., 2019). The effect of increased storage capacity on carotenoid accumulation is well established (Osorio, 2019). Plastid differentiation into chromoplasts stimulates carotenoid accumulation, as first revealed in the cauliflower orange mutant, where a gain-of-function mutation led to over-accumulation of beta-carotene (Lu et al., 2006). Overexpression of the *OR* gene has subsequently become an alternative approach to pathway engineering for increasing carotenoid content in crop plants like potato, rice, maize and tomato (Lopez et al., 2008; Bai et al., 2014; Park et al., 2016; Berman et al., 2017). The post-transcriptional regulation of PSY protein levels by OR and OR's effect on downregulation of beta-carotene hydroxylase further highlights the crosstalk between the various mechanisms controlling carotenogenesis. Our results show PSY-induced carotenoid biosynthesis was accompanied by accelerated chromoplast differentiation and increased expression of apple *OR* genes in the PSY fruit. This further reinforces the connection between plastid differentiation and carotenoid biosynthesis pathways in plants (Shumskaya et al., 2012; Park et al., 2016; Welsch et al., 2018).

Phytoene synthase overexpression increased carotenoid content in bagged fruit

Bagging of apple fruit reduced fruit carotenoid content and plastid concentration, highlighting the light regulating component of carotenoid accumulation and plastid development. Light plays a significant role in regulating fruit pigmentation, acting through signaling factors to regulate gene expression (Llorente et al., 2017). The light exclusion effect on the apple fruit tissues was associated with reduced gene expression of *HY5* and Phytochrome A (*PHYA*), which are two key light signaling regulators whose functions have been studied in tomato. The loss of function of SlHY5, either through genome editing or RNAi downregulation of HY5, impaired fruit ripening and resulted in reduced carotenoid content (Liu et al., 2004; Wang W. et al., 2021). Phytochromes, such as PHYA and PHYB, mediate light-induced ripening and carotenoid accumulation in tomato fruit with their expression level positively correlated with carotenoid biosynthesis, though such effect could be fruit stage specific or may involve complex

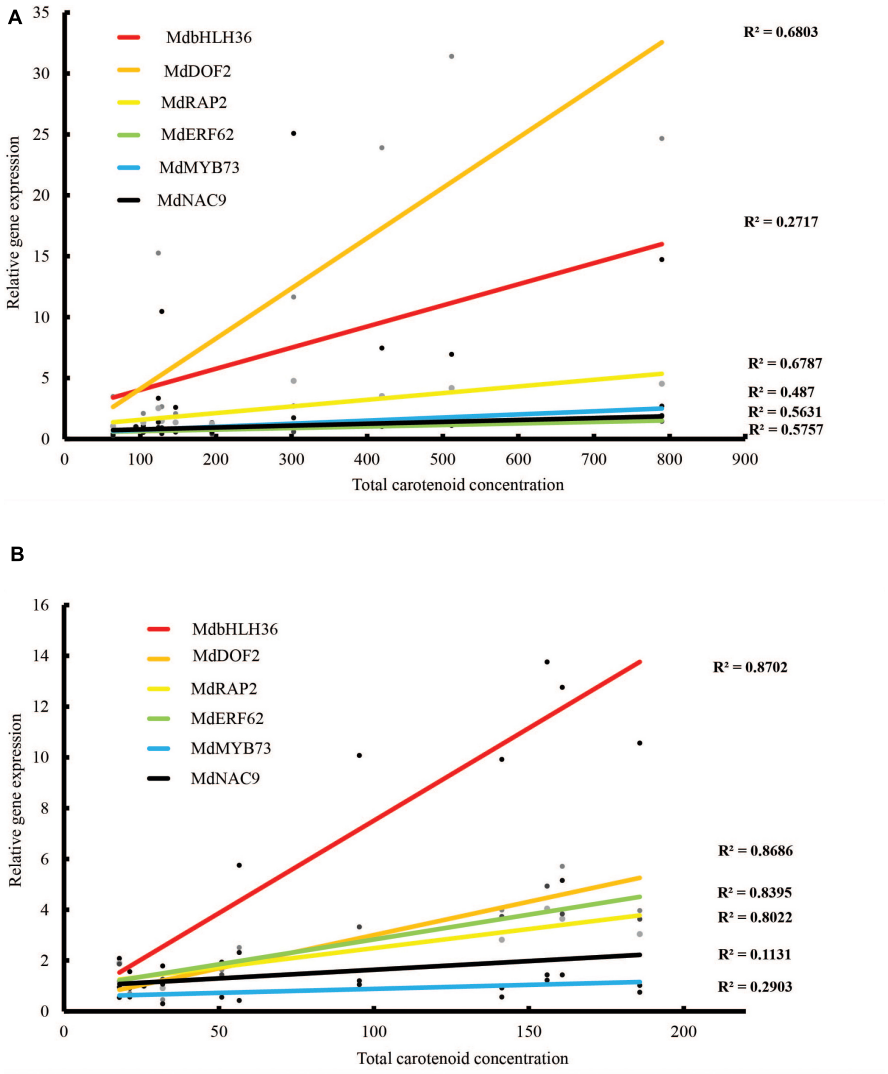


FIGURE 8
Correlations between carotenoid content and TF gene expression as determined by qRT-PCR in **(A)** fruit skin and **(B)** fruit flesh of WT and PSY fruit. R^2 values indicate coefficient of determination.

interactions with other phytochromes (Alba et al., 2000; Gupta et al., 2014; Bianchetti et al., 2018). While light exclusion negatively affected fruit carotenoid content, the bagged PSY apple fruit still accumulated higher carotenoid content than the bagged WT suggesting *PSY1* expression could compensate for the negative effect of light exclusion on the carotenoid pathway. This is consistent with PSY as a key target for light regulation of carotenoid biosynthesis, acting through transcription factors such as PIF1, which negatively regulates *PSY* (Toledo-Ortiz et al., 2010). Under low light conditions, PIF1 expression in tomato increases and binds the *PSY* promoter to repress gene expression (Llorente et al., 2016), implying increased *PSY* expression could override the inhibition of carotenogenesis by light exclusion.

Transcriptional response to *PSY1* expression in fruit

The transcriptional changes induced by *PSY1* overexpression provided an opportunity to identify related genes associated with the phenotypic changes. The observed up- and down-regulation of carotenoid pathway genes was consistent with the metabolite accumulation pattern, with beta-carotene as the dominant compound in PSY fruit. Genes encoding PDS, ZISO, ZDS, CRTISO and LCB showed increased gene expression in PSY fruit compared with the WT, and conversely the expression of *BCH*, *ECH*, and *ZEP* was reduced. The accumulation of beta-carotene, and not phytoene, in the transgenic fruit tissues clearly indicated an active apple

carotenoid pathway, with increased gene expression resulting in increased enzyme activity. *PDS* and *ZISO*, for example, are both single-copy genes in the apple genome and their increased activities would be required for the subsequent conversion of phytoene to zeta carotene (Chen et al., 2010; Beltrán et al., 2015; Rodrigo et al., 2019).

The changes in carotenoid gene expression suggested PSY1-induced feedforward regulation of the apple carotenoid pathway. The plant carotenoid pathway is characterized by metabolic feedforward and feedback regulation (McQuinn et al., 2015). *PSY* overexpression in Arabidopsis elevated *DXS* transcripts, which suggested a positive feedback regulation of the MEP pathway (Rodríguez-Villalón et al., 2009). In contrast, the over-accumulation of phytoene in the *pds3* Arabidopsis mutant caused negative feedback regulation of the carotenoid pathway, decreasing *GGPPS*, *PSY*, *ZDS,* and *LCYB* expression in the mutant (Qin et al., 2007). The expression of bacterial *CrtI* gene in tomato mutants elevated the pool of lycopene substrate in fruit and caused a feedforward transcriptional activation of the downstream cyclase genes (Enfissi et al., 2017). Similarly, overexpression of Arabidopsis *PDS* triggered a feedforward effect on carotenoid gene expression in tomato leaves and flowers (McQuinn et al., 2018). *LCY-B* overexpression increased *PSY* transcripts in transgenic carrot (Moreno et al., 2013); conversely, overexpression of carotene hydroxylase CYP97A3 reduced PSY protein content in carrot, suggesting a negative feedback regulation (Arango et al., 2014). Although mechanisms involved in these carotenoid feedback regulations are mainly uncharacterized, apocarotenoid-derived signaling molecules are involved in transcriptional regulation of carotenoid genes and inhibition of enzymatic activity (Avendaño-Vázquez et al., 2014; Enfissi et al., 2017). Whilst gene transcript levels are not always indicative of enzyme activity, it is possible *PSY1*-induced accumulation of phytoene exerted a metabolic feedforward regulation of carotenoid pathway genes that subsequently led to increased accumulation of downstream compounds, such as beta-carotene.

PSY1 overexpression uncovers carotenoid-associated regulatory genes

The transcriptional response to *PSY1* overexpression is likely to be mediated by transcription factors, whose gene expression would be expected to correlate with their target genes, as identified using gene co-expression analysis (Langfelder and Horvath, 2008; Zhang et al., 2020; Zogopoulos et al., 2021). Our analysis found six carotenoid-associated TF genes with their expression strongly correlated with fruit carotenoid accumulation during on-tree development as well as in the absence of light. These TF genes belong to the bHLH, DOF, ERF, MYB, and NAC families, some of which have members demonstrated to regulate carotenoid metabolism (Stanley and Yuan, 2019).

We showed previously that kiwifruit *AdMYB7* plays a significant role in controlling chlorophyll and carotenoid accumulation during fruit development through regulation of certain key genes (Ampomah-Dwamena et al., 2019). Other MYBs from *Erythranthe lewisii* (ElRCP1), citrus (CrMYB68), and Medicago (WP1) have been implicated in carotenoid regulation, although they are phylogenetically distant (Sagawa et al., 2016; Zhu F. et al., 2017; Meng et al., 2019). These examples suggest that MYBs from different clades have evolved to control carotenoid biosynthesis. *MdMYB73*, from this study, is an R2R3 MYB with similarity to Arabidopsis MYB70, MYB73, MYB77, and MYB44, which generally mediate responses to biotic and abiotic stresses (Stracke et al., 2001; Dubos et al., 2010). Apple *MdMYB73* controls malate metabolism through activation of pathway genes (Hu et al., 2017), while its Arabidopsis homolog, *AtMYB73*, regulates salt stress response (Kim et al., 2013; Wang L. et al., 2021). In addition to MYBs, bHLH TFs have a role in controlling carotenoid metabolism. BHLHs, such as the Phytochrome interacting factor 1 (PIF1), are key mediators of carotenogenesis in Arabidopsis, directly binding the promoter of *PSY* to repress gene expression under low light conditions, albeit this role was not required in roots (Toledo-Ortiz et al., 2010; Ruiz-Sola et al., 2014). In tomato, *SlPRE2*, a gibberellic acid-induced bHLH TF, negatively regulates carotenogenesis by downregulating *PSY* and *ZDS* gene expression (Zhu Z. et al., 2017; Zhu et al., 2019). Other bHLHs positively regulate carotenoid accumulation, such as papaya CpbHLH2, which activated expression of *CYCB* and *LCYB* in transient assays, and SlARANCIO, a tomato bHLH gene associated with a carotenoid QTL (D'Amelia et al., 2019; Zhou et al., 2019). Fine mapping of the Y2 locus controlling beta-carotene accumulation in carrot found a differentially expressed bHLH TF, suggesting a potential regulatory role in carotenoid metabolism (Ellison et al., 2017).

NAC domain (derived from the three type members, NO APICAL MERISTEM, ATAF, and CUP-SHAPED COTYLEDON; Aida et al., 1997) TFs regulate diverse traits such as stress responses, senescence, fruit ripening events and carotenoid biosynthesis in fruit. The tomato NAC ripening regulator NOR-like 1 was shown to directly regulate genes involved in ethylene biosynthesis as well as the carotenoid-related genes *GGPPS2* and *SGR1* (Gao et al., 2018). GGPPS is responsible for the synthesis of GGPP, which is the substrate for PSY, while SGR1 negatively regulates PSY to control carotenoid accumulation in tomato (Luo et al., 2013). The modulation of SlNAC1 transcript levels led to changes in carotenoid accumulation patterns, suggesting it was a key regulator. Overexpression of this gene resulted in delayed fruit ripening and reduced carotenoid content (Ma et al., 2014). Conversely, downregulation of SlNAC1 produced higher total carotenoid and lycopene in fruit, although fruit ripening was also delayed (Meng et al., 2016). This is in contrast to *SlNAC4*, which positively regulated carotenoid accumulation; RNAi repression

resulted in downregulation of *SlPSY1* transcripts and reduced carotenoid content (Zhu et al., 2014). In papaya, CpNAC1 activated the promoters of phytoene desaturase genes both *in vitro* and in transient assays (Fu et al., 2016). The AP2/ERFs also have roles in plant carotenoid metabolism. One of the two AP2/ERFs identified in this study, MdRAP2 (AP2D11), previously activated the *MdPSY2* promoter in a transient promoter assay screen suggesting that it is a positive regulator of the carotenoid pathway (Ampomah-Dwamena et al., 2015). Its Arabidopsis homolog, AtRAP2.2 interacted with *PSY* and *PDS* upstream regulatory sequences *in vitro*, indicating it may be part of a regulatory regime controlling carotenogenesis (Welsch et al., 2007). In a more recent study, overexpression of another AP2/ERF, *MdAP2-34*, in apple increased fruit carotenoid content, with the TF able to bind and activate the *MdPSY2* promoter (Dang et al., 2021). In tomato, SlAP2a and SlERF6 play important roles in carotenogenesis acting as negative regulators (Chung et al., 2010; Karlova et al., 2011; Lee et al., 2012). While the gene expression of the apple TFs in this study strongly correlated with fruit carotenoid accumulation, further experimentation will be required to characterize their functional roles in carotenogenesis.

In summary, overexpression of *PSY1* in apple produced fruit with increased carotenoid content, accelerated plastid differentiation, a feedforward effect in the pathway and a transcriptomic response that included altered expression of key TFs whose expression suggested roles in carotenogenesis and utility for developing biofortified apple cultivars.

Materials and methods

Cloning and plant materials

The apple *phytoene synthase 1* (MD17G1133400) was amplified from fruit complementary DNA (cDNA) using Hi-fidelity Taq polymerase (Life Technologies, Carlsbad, CA, United States) and primers PSY1-F and PSY1-R (**Supplementary Table S1**). The PCR amplicon was initially cloned into PCR8/GW TOPO vector (Life Technologies, Carlsbad, CA, United States) for sequence confirmation and then cloned into the binary vector pHEX2 using the Gateway cloning strategy as previously described (Ampomah-Dwamena et al., 2019). *Agrobacterium* strain LBA4404 carrying the resulting vector, pHEX2S-MdPSY1, was used to transform 'Royal Gala' apple as described previously (Yao et al., 1995). Transgenic 'Royal Gala' shoots from tissue culture were grafted onto 'Malling 9' apple rootstock for growth in soil in glasshouse. Flowers at anthesis were hand pollinated with 'Granny Smith' pollen. Fruit harvested at different development stages [90, 120, 135, and 150 days after pollination (D)] and separated into skin and flesh (cortex) tissues were frozen in liquid nitrogen for pigment analysis and gene expression. For the light exclusion experiment, fruit at 20 D were covered in paper bags, and left to grow on the tree until 150 D when they were harvested and analyzed. For callus culture, young leaves were collected from three PSY lines (OE-2, OE-3, OE-4), sterilized and cultured on callus induction medium (MS basal medium + 1 mg/L BA + 0.5 mg/L NAA, 7% agar), supplemented with 0.3 mg/L norflurazon, for 2 weeks.

High-performance liquid chromatography pigment analysis

Carotenoid and chlorophyll pigments were extracted from fruit tissues using acetone as described previously (Ampomah-Dwamena et al., 2015). Samples were weighed into tubes and solvent extracted overnight. HPLC analysis was performed using a Dionex Ultimate 3000 solvent delivery system (Thermo Scientific, Waltham, MA, United States) with a photodiode array detector was fitted with Acclaim C30 column (5 μm, 250 × 4.6 mm), coupled to a C30 guard column (Thermo Scientific, Waltham, MA, United States) as previously reported (Ampomah-Dwamena et al., 2019). Phytoene was monitored at 280 nm, colored carotenoids and Chlb were detected at 450 nm, while Chla and its derivatives were monitored at 400 nm. Carotenoid contents were expressed as b-carotene equivalents per gram dry weight (DW) of tissue. All *trans*-β-carotene, lutein and Chla standards were purchased from Sigma Chemicals (Sigma-Aldrich, St Louis, MO, United States). Other carotenoids were putatively identified by comparison with reported retention times and spectral data.

RNA isolation and complementary DNA synthesis

Total RNA was extracted from tissues using the Spectrum RNA isolation kit (Sigma-Aldrich, St Louis, MO, United States) using a modified manufacturer's protocol. One gram of homogenized tissue in liquid nitrogen was used as the starting material and the volumes of buffers were increased accordingly. cDNA was synthesized from total RNA using the Quantitec reverse transcription kit (Qiagen, Hilden, Germany) following the manufacturer's protocol. RNA samples were treated with the genomic DNA wipeout buffer followed by reverse transcription reaction as described earlier (Ampomah-Dwamena et al., 2019). The reactions were incubated at 42°C for 30 min followed by 95°C for 3 min for enzyme inactivation.

Quantitative real-time PCR analysis

First-strand cDNA samples from fruit skin and flesh tissues at 90, 120, 135, and 150 D were diluted 1:20 and

used as templates for quantitative real-time PCR according to methods described previously (Ampomah-Dwamena et al., 2015, Ampomah-Dwamena et al., 2019). PCR analysis was performed using the LightCycler 1.5 system and the SYBR Green master mix (Roche, Mannheim, Germany), following the manufacturer's protocol. Each reaction sample was analyzed from biological replicates, with a negative control using water as template. PCR conditions were as follows: pre-incubation at 95°C for 5 min followed by 40 cycles each consisting of 10 s at 95°C, 10 s at 60°C and 20 s at 72°C. Amplification was followed by a melting curve analysis with continuous fluorescence measurement during the 65–95°C melt (Ampomah-Dwamena et al., 2019). The relative expression was calculated using LIGHTCYCLER software v.4 and the expression of each gene was normalized to reference genes *MdACTIN* and *MdEF1A*, which have been shown to have stable expression in these tissues (Nieuwenhuizen et al., 2013; Ampomah-Dwamena et al., 2015, Ampomah-Dwamena et al., 2019). Primers were designed using PRIMER3 software (Rozen and Skaletsky, 2000) with respective EST sequences as templates. Primers were subjected to a stringent set of criteria, with a minimum melting temperature of 60°C and at least 22 nucleotide length.

Transcriptome analysis by RNA sequencing

Total RNA from fruit tissues with A260/280 and A260/230 absorbance ratios both greater than 1.8 were used in sequencing library preparation using the TruSeq mRNA library preparation kit (Illumina, NovogeneAIT Genomics, Singapore). Transcriptome was sequenced on a HiSeq2000 (Illumina) using 2 × 100 bp paired-end sequencing. Each treatment had three biological replicates, with each replicate being a pool of tissues from 5–7 fruit from individual plants. The multiplexed libraries were run on multiple lanes, generating between 40 and 56 million reads per sample. The resulting reads were then quality and adapter trimmed using bbduck from the BBMap suite (version 37.53), with a quality threshold cut-off of 20 and minimum nucleotide length of 35 bases, where k-mer size of 25 was used. Thereafter, reads were aligned to the 'Golden Delicious' double haploid annotated gene models (Daccord et al., 2017) using STAR (version 2.5.2b) and the '–quantMode GeneCounts' flag was used to extract read counts for the annotated genes. Differentially expressed genes identified using DESeq2 analysis, with a cut-off probability of $p < 0.05$, were subjected to Gene ontology (GO) enrichment analysis using goseq package (Young et al., 2012) in R and thereafter the Go Plot R package (Walter et al., 2015) for visualization. Weighted gene correlation network analysis was performed using R package (Langfelder and Horvath, 2008) to identify modules highly associated with selected traits, and the networks visualized using Cytoscape 3.8.2 (Saito et al., 2012).

Statistical analysis

Hierarchical clustering of gene expression and metabolite content data represented by the biological replicates was performed using MetaboAnalyst v 5.0[1] (Pang et al., 2021). Gene expression data were organized with gene IDs in the rows and biological replicates in the columns. Data were normalized by Log transformation where required. Heat maps were generated using Euclidean distance of the average values of biological replicates. Correlation tables with p-value cut-off < 0.05 were generated in MetaboAnalyst as above and color was formatted in Microsoft® Excel®.

Microscopy

Sections of 120, 135, and 150 D apple fruit tissue [150–250 nm cut using a VT1000S vibratome (Leica Microscopy Systems Ltd., Heerbrugg, Switzerland)] were taken from the equatorial region of each apple, mounted in 0.1 M phosphate buffer and imaged using an FV3000 laser scanning confocal (Olympus Optical Co. Ltd., Tokyo, Japan) on an IX83 inverted microscope platform (Olympus Optical Co. Ltd., Tokyo, Japan). To quantify differences in carotenoid and chlorophyll intensity between different fruit, lambda scans were performed using a 488 nm laser for excitation, collecting stepwise auto fluorescence emissions between 500 and 750 nm with a bandwidth of 10 nm and a step size 5 nm. To visualize carotenoid and chlorophyll accumulation in fruit sections, a 30-μm z-stack at 1-μm steps was collected for spectral bands between 500 and 550 nm, and 650–700 nm, respectively. Images were processed using Olympus Fluoview software to give a maximum intensity projection (Olympus Optical Co. Ltd., Tokyo, Japan).

Data availability statement

The datasets presented in this study and accession numbers are included in the article/**Supplementary material**. Further enquiries may be directed to the corresponding author.

Author contributions

CA-D designed the research and wrote the draft manuscript. ST transformed the apple plants. AT analyzed the transcriptome data and network analysis. CE and NB carried out total

1 https://www.metaboanalyst.ca/MetaboAnalyst/home.xhtml

RNA isolation, cDNA synthesis, and qPCR analysis. RR and PS carried out the light and confocal microscopy analysis. HI constructed mRNA libraries for transcriptome sequencing. AA and RE assisted with data analysis and provided helpful comments on manuscript. All authors contributed to the article and approved the submitted version.

Funding

The authors declare that this study received funding from The New Zealand Institute for Plant and Food Research's Technology Development Pipfruit program. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

Acknowledgments

We thank Dr. Robert Schaffer, Dr. Cyril Brendolise, Kui Lin-Wang for helpful discussions and advice, as well as the Plant and Food Research (PFR) Science Publication Office for critical comments and edits. We are also grateful to Monica Dragulescu for maintaining the plants in the glasshouse.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.967143/full#supplementary-material

SUPPLEMENTARY FIGURE S1
Graph of carotenoid pigments accumulating in regenerated calli from wild type (WT) and PSY transgenic (OE) leaf tissues. Bars represent the average of three biological replicates ± SE.

SUPPLEMENTARY FIGURE S2
Gene expression of carotenoid genes in fruit skin of WT, OE-5, and OE-7 lines as determined by real-time qPCR during fruit development. Bars represent the average of three biological replicates ± SE. LCE, lycopene epsilon-cyclase; ECH, epsilon carotene hydroxylase; ZEP, zeaxanthin epoxidase.

SUPPLEMENTARY FIGURE S3
PCA plot of RNA-sequencing data (RPKM) of WT and OE-7 apple fruit flesh sampled at 90 (S1), 120 (S2), 135 (S3), and 150 (S4) D.

SUPPLEMENTARY FIGURE S4
Gene ontology (GO) term enrichment of the PSYAGs in the biological process (BP) and molecular function (MF) categories. Enriched categories are highlighted in yellow.

SUPPLEMENTARY FIGURE S5
(A) Heatmap and eigengene expression of brown (left) and green (right) modules, across apple fruit development. **(B)** Module membership-gene significance correlation (left) and multi-dimensional plots (right) of brown (top panel) and green (bottom panel) modules.

SUPPLEMENTARY FIGURE S6
Gene expression analysis of the six carotenoid-associated TFs. Relative expression as determined by RT-PCR in fruit skin and flesh of WT, OE-5, and OE-7 lines during fruit development. Bars represent the average of three biological replicates ± SE.

SUPPLEMENTARY TABLE S1
List of PCR primers used.

SUPPLEMENTARY TABLE S2
List of 54 apple carotenoid genes identified from the "GDDH" apple sequence database (Daccord et al., 2017).

SUPPLEMENTARY TABLE S3
List of PSY-associated apple genes.

SUPPLEMENTARY TABLE S4
List of differentially expressed TFs (GS > 0.6) during apple fruit development.

SUPPLEMENTARY TABLE S5
List of differentially expressed TFs in bagged apple fruit. Highlighted in bold are the six TF genes common to both bagging and fruit development datasets.

SUPPLEMENTARY TABLE S6
Correlation coefficient (*r*) of the relationship between gene expression of carotenoid-associated TFs, apple carotenoid genes and carotenoid content in fruit skin and flesh. Color gradient from red to green, with arrows, indicate weak to strong correlations.

References

Ahrazem, O., Diretto, G., Argandoña Picazo, J., Fiore, A., Rubio-Moraga, Á, Rial, C., et al. (2019). The specialized roles in carotenogenesis and apocarotenogenesis of the phytoene synthase gene family in saffron. *Front Plant Sci.* 10:249. doi: 10.3389/fpls.2019.00249

Aida, M., Ishida, T., Fukaki, H., Fujisawa, H., and Tasaka, M. (1997). Genes involved in organ separation in Arabidopsis: An analysis of the cup-shaped cotyledon mutant. *Plant Cell* 9, 841–857. doi: 10.1105/tpc.9.6.841

Alba, R., Cordonnier-Pratt, M. M., and Pratt, L. H. (2000). Fruit-localized phytochromes regulate lycopene accumulation independently of ethylene production in tomato. *Plant Physiol.* 123, 363–370.

Álvarez, D., Voß, B., Maass, D., Wüst, F., Schaub, P., Beyer, P., et al. (2016). Carotenogenesis is regulated by 5' UTR-mediated translation of phytoene synthase splice variants. *Plant Physiol.* 172, 2314–2326.

Ampomah-Dwamena, C., Dejnoprat, S., Lewis, D. C., Sutherland, P., Volz, R., and Allan, A. (2012). Metabolic and gene expression analysis of apple (*Malus* × *domestica*) carotenogenesis. *J. Exp. Bot.* 63, 4497–4511.

Ampomah-Dwamena, C., Driedonks, N., Lewis, D., Shumskaya, M., Chen, X., Wurtzel, E., et al. (2015). The Phytoene synthase gene family of apple (*Malus* × *domestica*) and its role in controlling fruit carotenoid content. *BMC Plant Biol.* 15:185. doi: 10.1186/s12870-015-0573-7

Ampomah-Dwamena, C., Thrimawithana, A. H., Dejnoprat, S., Lewis, D., Espley, R. V., and Allan, A. C. (2019). A kiwifruit (*Actinidia deliciosa*) R2R3-MYB transcription factor modulates chlorophyll and carotenoid accumulation. *New Phytol.* 221, 309–325.

Arango, J., Jourdan, M., Geoffriau, E., Beyer, P., and Pand Welsch, R. (2014). Carotene hydroxylase activity determines the levels of both α-carotene and total carotenoids in orange carrots. *Plant Cell* 26, 2223–2233.

Arcos, Y., Godoy, F., Flores-Ortiz, C., Arenas, M. A., and Stange, C. (2020). Boosting carotenoid content in *Malus domestica* var. Fuji by expressing *AtDXR* through an *Agrobacterium*-mediated transformation method. *Biotechnol. Bioeng.* 117, 2209–2222.

Avendaño-Vázquez, A., Cordoba, E., Llamas, E., San Román, C., Nisar, N., Torre, D. L., et al. (2014). An uncharacterized apocarotenoid-derived signal generated in ζ-carotene desaturase mutants regulates leaf development and the expression of chloroplast and nuclear genes in Arabidopsis. *Plant Cell* 26, 2524–2537.

Bai, C., Rivera, S. M., Medina, V., Alves, R., Vilaprinyo, E., Sorribas, A., et al. (2014). An in vitro system for the rapid functional characterization of genes involved in carotenoid biosynthesis and accumulation. *Plant J.* 77, 464–475.

Beltrán, J., Kloss, B., Hosler, J. P., Geng, J., Liu, A., Modi, A., et al. (2015). Control of carotenoid biosynthesis through a heme-based cis-trans isomerase. *Nat. Chem. Biol.* 11, 598–605.

Berman, J., Zorrilla-López, U., Medina, V., Farré, G., Sandmann, G., Capell, T., et al. (2017). The Arabidopsis *ORANGE* (*AtOR*) gene promotes carotenoid accumulation in transgenic corn hybrids derived from parental lines with limited carotenoid pools. *Plant Cell Rep.* 36, 933–945.

Bianchetti, R. E., Lira, B. S., Monteiro, S. S., Demarco, D., Purgatto, E., Rothan, C., et al. (2018). Fruit-localized phytochromes regulate plastid biogenesis, starch synthesis, and carotenoid metabolism in tomato. *J. Exp. Bot.* 69, 3573–3586.

Brandi, F., Bar, E., Mourgues, F., Horváth, G., Turcsi, E., Giuliano, G., et al. (2011). Study of 'Redhaven' peach and its white-fleshed mutant suggests a key role of CCD4 carotenoid dioxygenase in carotenoid and norisoprenoid volatile metabolism. *BMC Plant Biol.* 11:24. doi: 10.1186/1471-2229-11-24

Cerda, A., Moreno, J. C., Acosta, D., Godoy, F., Cáceres, J. C., Cabrera, R., et al. (2020). Functional characterisation and in silico modelling of MdPSY2 variants and MdPSY5 phytoene synthases from *Malus domestica*. *J. Plant Physiol.* 249:153166.

Chayut, N., Yuan, H., Ohali, S., Meir, A., Sa'ar, U., Tzuri, G., et al. (2017). Distinct mechanisms of the ORANGE protein in controlling carotenoid flux. *Plant Physiol.* 173, 376–389.

Chen, L., Li, W., Li, Y., Feng, X., Du, K., Wang, G., et al. (2019). Identified trans-splicing of YELLOW-FRUITED TOMATO 2 encoding the PHYTOENE SYNTHASE 1 protein alters fruit color by map-based cloning, functional complementation and RACE. *Plant Mol. Biol.* 100, 647–658.

Chen, Y., Li, F., and Wurtzel, E. T. (2010). Isolation and characterization of the Z-ISO gene encoding a missing component of carotenoid biosynthesis in plants. *Plant Physiol.* 153, 66–79.

Chung, M. Y., Vrebalov, J., Alba, R., Lee, J., McQuinn, R., Chung, J. D., et al. (2010). A tomato (*Solanum lycopersicum*) APETALA2/ERF gene, *SlAP2a*, is a negative regulator of fruit ripening. *Plant J.* 64, 936–947.

Daccord, N., Celton, J., Linsmith, G., Becker, C., Choisne, N., Schijlen, E., et al. (2017). High-quality de novo assembly of the apple genome and methylome dynamics of early fruit development. *Nat. Genet.* 49, 1099–1106.

D'Amelia, V., Raiola, A., Carputo, D., Filippone, E., Barone, A., and Rigano, M. M. (2019). A basic Helix-Loop-Helix (SlARANCIO), identified from a *Solanum pennellii* introgression line, affects carotenoid accumulation in tomato fruits. *Sci. Rep.* 9, 1–10.

D'Andrea, L., Amenós, M., and Rodríguez-Concepcíon, M. (2014). Confocal laser scanning microscopy detection of chlorophylls and carotenoids in chloroplasts and chromoplasts of tomato fruit. *Methods Mol. Biol.* 1153, 227–232.

D'Andrea, L., and Rodriguez-Concepcion, M. (2019). Manipulation of plastidial protein quality control components as a new strategy to improve carotenoid contents in tomato fruit. *Front. Plant Sci.* 10:1071. doi: 10.3389/fpls.2019.01071

D'Andrea, L., Simon-Moya, M., Llorente, B., Llamas, E., Marro, M., Loza-Alvarez, P., et al. (2018). Interference with Clp protease impairs carotenoid accumulation during tomato fruit ripening. *J. Exp. Bot.* 69, 1557–1568.

Dang, Q., Sha, H., Nie, J., Wang, Y., Yuan, Y., and Jia, D. (2021). An apple (*Malus domestica*) AP2/ERF transcription factor modulates carotenoid accumulation. *Hortic. Res.* 8, 1–12.

Delgado-Pelayo, R., Gallardo-Guerrero, L., and Hornero-Méndez, D. (2014). Chlorophyll and carotenoid pigments in the peel and flesh of commercial apple fruit varieties. *Food Res. Int.* 65, 272–281.

Diretto, G., Al-Babili, S., Tavazza, R., Scossa, F., Papacchioli, V., Migliore, M., et al. (2010). Transcriptional-metabolic networks in β-carotene-enriched potato tubers: the long and winding road to the Golden phenotype. *Plant Physiol.* 154, 899–912.

Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C., and Lepiniec, L. (2010). MYB transcription factors in Arabidopsis. *Trends Plant Sci.* 15, 573–581.

Ducreux, L. J., Morris, W. L., Hedley, P. E., Shepherd, T., Davies, H. V., Millam, S., et al. (2005). Metabolic engineering of high carotenoid potato tubers containing enhanced levels of β-carotene and lutein. *J. Exp. Bot.* 56, 81–89.

Eggersdorfer, M., and Wyss, A. (2018). Carotenoids in human nutrition and health. *Arch. Biochem. Biophys.* 652, 18–26.

Ellison, S., Senalik, D., Bostan, H., Iorizzo, M., and Simon, P. (2017). Fine mapping, transcriptome analysis, and marker development for Y2, the gene that conditions β-carotene accumulation in carrot (*Daucus carota* L.). *G3* 7, 2665–2675.

Enfissi, E. M., Nogueira, M., Bramley, P. M., and Fraser, P. D. (2017). The regulation of carotenoid formation in tomato fruit. *Plant J.* 89, 774–788.

Feder, A., Chayut, N., Gur, A., Freiman, Z., Tzuri, G., Meir, A., et al. (2019). The role of carotenogenic metabolic flux in carotenoid accumulation and chromoplast differentiation: lessons from the melon fruit. *Front. Plant Sci.* 10:1250. doi: 10.3389/fpls.2019.01250

Fraser, P. D., Enfissi, E. M., Halket, J. M., Truesdale, M. R., Yu, D., Gerrish, C., et al. (2007). Manipulation of phytoene levels in tomato fruit: effects on isoprenoids, plastids, and intermediary metabolism. *Plant Cell* 19, 3194–3211.

Fu, C. C., Han, Y. C., Fan, Z. Q., Chen, J. Y., Chen, W. X., Lu, W. J., et al. (2016). The papaya transcription factor CpNAC1 modulates carotenoid biosynthesis through activating phytoene desaturase genes CpPDS2/4 during fruit ripening. *J. Agric. Food Chem.* 64, 5454–5463.

Gao, Y., Wei, W., Zhao, X., Tan, X., Fan, Z., Zhang, Y., et al. (2018). A NAC transcription factor, NOR-like1, is a new positive regulator of tomato fruit ripening. *Hortic. Res.* 5, 1–18.

Gupta, S. K., Sharma, S., Santisree, P., Kilambi, H. V., Appenroth, K., Sreelakshmi, Y., et al. (2014). Complex and shifting interactions of phytochromes regulate fruit development in tomato. *Plant Cell Environ.* 37, 1688–1702.

Howitt, C. A., Cavanagh, C. R., Bowerman, A. F., Cazzonelli, C., Rampling, L., Mimica, J. L., et al. (2009). Alternative splicing, activation of cryptic exons and amino acid substitutions in carotenoid biosynthetic genes are associated with lutein accumulation in wheat endosperm. *Funct. Integr. Genomics* 9, 363–376.

Hu, D. G., Li, Y. Y., Zhang, Q. Y., Li, M., Sun, C. H., Yu, J. Q., et al. (2017). The R2R3—MYB transcription factor *MdMYB73* is involved in malate accumulation and vacuolar acidification in apple. *Plant J.* 91, 443–454.

Isaacson, T., Ohad, I., Beyer, P., and Hirschberg, J. (2004). Analysis in vitro of the enzyme CRTISO establishes a poly-cis-carotenoid biosynthesis pathway in plants. *Plant Physiol.* 136, 4246–4255.

Karlova, R., Rosin, F. M., Busscher-Lange, J., Parapunova, V., Do, P. T., Fernie, A. R., et al. (2011). Transcriptome and metabolite profiling show that APETALA2a is a major regulator of tomato fruit ripening. *Plant Cell* 23, 923–941.

Kilcrease, J., Collins, A., Richins, R., Timlin, J., and O'Connell, M. (2013). Multiple microscopic approaches demonstrate linkage between chromoplast architecture and carotenoid composition in diverse *Capsicum annuum* fruit. *Plant J.* 76, 1074–1083.

Kim, J., and DellaPenna, D. (2006). Defining the primary route for lutein synthesis in plants: The role of Arabidopsis carotenoid beta-ring hydroxylase CYP97A3. *Proc. Natl. Acad. Sci. U.S.A.* 103, 3474–3479.

Kim, J. H., Nguyen, N. H., Jeong, C. Y., Nguyen, N. T., Hong, S. W., and Lee, H. (2013). Loss of the R2R3 MYB, *AtMyb73*, causes hyper-induction of the *SOS1* and *SOS3* genes in response to high salinity in Arabidopsis. *J. Plant Physiol.* 170, 1461–1465.

Ko, M. R., Song, M. H., Kim, J. K., Baek, S. A., You, M. K., Lim, S. H., et al. (2018). RNAi-mediated suppression of three carotenoid-cleavage dioxygenase

genes, OsCCD1, 4a, and 4b, increases carotenoid content in rice. *J. Exp. Bot.* 69, 5105–5116.

Koschmieder, J., and Welsch, R. (2020). "Quantification of carotenoid pathway flux in green and nongreen systems," in *Plant and Food Carotenoids. Methods in Molecular Biology 2083*, eds M. Rodríguez-Concepción and R. Welsch (New York, NY: Humana).

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9:559. doi: 10.1186/1471-2105-9-559

Lee, J. M., Joung, J. G., McQuinn, R., Chung, M. Y., Fei, Z., Tieman, D., et al. (2012). Combined transcriptome, genetic diversity and metabolite profiling in tomato fruit reveals that the ethylene response factor SlERF6 plays an important role in ripening and carotenoid accumulation. *Plant J.* 70, 191–204.

Li, L., and Van Eck, J. (2007). Metabolic engineering of carotenoid accumulation by creating a metabolic sink. *Transgenic Res.* 16, 581–585.

Liu, Y., Roof, S., Ye, Z., Barry, C., van Tuinen, A., Vrebalov, J., et al. (2004). Manipulation of light signal transduction as a means of modifying fruit nutritional quality in tomato. *Proc. Natl Acad. Sci. U. S. A.* 101, 9897–9902.

Llorente, B., D'Andrea, L., Ruiz−Sola, M. A., Botterweg, E., Pulido, P., Andilla, J., et al. (2016). Tomato fruit carotenoid biosynthesis is adjusted to actual ripening progression by a light−dependent mechanism. *Plant J.* 85, 107–119.

Llorente, B., Martinez-Garcia, J. F., Stange, C., and Rodriguez-Concepcion, M. (2017). Illuminating colors: regulation of carotenoid biosynthesis and accumulation by light. *Curr. Opin. Plant Biol.* 37, 49–55.

Lopez, A. B., Van Eck, J., Conlin, B. J., Paolillo, D. J., O'Neill, J., and Li, L. (2008). Effect of the cauliflower *Or* transgene on carotenoid accumulation and chromoplast formation in transgenic potato tubers. *J. Exp. Bot.* 59, 213–223.

Lu, S., Van Eck, J., Zhou, X., Lopez, A. B., O'Halloran, D. M., Cosman, K. M., et al. (2006). The cauliflower *Or* gene encodes a DNAJ cysteine-rich domain-containing protein that mediates high levels of β-carotene accumulation. *Plant Cell* 18, 3594–3605.

Luo, Z., Zhang, J., Li, J., Yang, C., Wang, T., Ouyang, B., et al. (2013). A STAY−GREEN protein SlSGR1 regulates lycopene and β−carotene accumulation by interacting directly with SlPSY1 during ripening processes in tomato. *New Phytol.* 198, 442–452.

Ma, N., Feng, H., Meng, X., Li, D., Yang, D., Wu, C., et al. (2014). Overexpression of tomato SlNAC1 transcription factor alters fruit pigmentation and softening. *BMC Plant Biol.* 14:351. doi: 10.1186/s12870-014-0351-y

McQuinn, R. P., Giovannoni, J. J., and Pogson, B. J. (2015). More than meets the eye: from carotenoid biosynthesis, to new insights into apocarotenoid signaling. *Curr. Opin. Plant Biol.* 27, 172–179.

McQuinn, R. P., Wong, B., and Giovannoni, J. J. (2018). *AtPDS* overexpression in tomato: exposing unique patterns of carotenoid self−regulation and an alternative strategy for the enhancement of fruit carotenoid content. *Plant Biotechnol. J.* 16, 482–494.

Meng, C., Yang, D., Ma, X., Zhao, W., Liang, X., Ma, N., et al. (2016). Suppression of tomato SlNAC1 transcription factor delays fruit ripening. *J. Plant Physiol.* 193, 88–96.

Meng, Y., Wang, Z., Wang, Y., Wang, C., Zhu, B., Liu, H., et al. (2019). The MYB activator WHITE PETAL1 associates with *MtTT8* and *MtWD40-1* to regulate carotenoid-derived flower pigmentation in *Medicago truncatula*. *Plant Cell* 31, 2751–2767.

Moreno, J. C., Pizarro, L., Fuentes, P., Handford, M., Cifuentes, V., and Stange, C. (2013). Levels of lycopene β-cyclase 1 modulate carotenoid gene expression and accumulation in *Daucus carota*. *PLoS One* 8:e58144. doi: 10.1371/journal.pone.0058144

Nieuwenhuizen, N. J., Green, S. A., Chen, X., Bailleul, E. J., Matich, A. J., Wang, M. Y., et al. (2013). Functional genomics reveals that a compact terpene synthase gene family can account for terpene volatile production in apple. *Plant Physiol.* 161, 787–804.

Nisar, N., Li, L., Lu, S., Khin, N., and Pogson, B. (2015). Carotenoid metabolism in plants. *Mol. Plant* 8, 68–82.

Niu, G., Guo, Q., Wang, J., Zhao, S., He, Y., and Liu, L. (2020). Structural basis for plant lutein biosynthesis from α-carotene. *Proc. Natl. Acad. Sci. U.S.A.* 117, 14150-14157.

Okada, K., Saito, T., Nakagawa, T., Kawamukai, M., and Kamiya, Y. (2000). Five geranylgeranyl diphosphate synthases expressed in different organs are localized into three subcellular compartments in Arabidopsis. *Plant Physiol.* 122, 1045–1056.

Osorio, C. E. (2019). The role of orange gene in carotenoid accumulation: manipulating chromoplasts toward a colored future. *Front. Plant Sci.* 10:1235. doi: 10.3389/fpls.2019.01235

Pang, Z., Chong, J., Zhou, G., de Lima Morais, D. A., Chang, L., Barrette, M., et al. (2021). MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic Acids Res.* 49, 388–396.

Park, S., Kim, H. S., Jung, Y. J., Kim, S. H., Ji, C. Y., Wang, Z., et al. (2016). Orange protein has a role in phytoene synthase stabilization in sweetpotato. *Sci. Rep.* 6:33563.

Qin, G., Gu, H., Ma, L., Peng, Y., Deng, X. W., Chen, Z., et al. (2007). Disruption of phytoene desaturase gene results in albino and dwarf phenotypes in Arabidopsis by impairing chlorophyll, carotenoid, and gibberellin biosynthesis. *Cell Res.* 17, 471–482.

Quinlan, R. F., Shumskaya, M., Bradbury, L. M., Beltrán, J., Ma, C., Kennelly, E. J., et al. (2012). Synergistic interactions between carotene ring hydroxylases drive lutein formation in plant carotenoid biosynthesis. *Plant Physiol.* 160, 204–214.

Rodrigo, M. J., Lado, J., Alós, E., Alquézar, B., Dery, O., Hirschberg, J., et al. (2019). A mutant allele of ζ-carotene isomerase (Z-ISO) is associated with the yellow pigmentation of the "Pinalate" sweet orange mutant and reveals new insights into its role in fruit carotenogenesis. *BMC Plant Biol.* 19:465. doi: 10.1186/s12870-019-2078-2

Rodríguez-Concepcíon, M., Avalos, J., Bonet, M., Boronat, A., Gómez-Gómez, L., Hornero-Méndez, D., et al. (2018). A global perspective on carotenoids: metabolism, biotechnology, and benefits for nutrition and health. *Progr. Lipid Res.* 70, 62–93.

Rodríguez-Villalón, A., Gas, E., and Rodríguez-Concepción, M. (2009). Phytoene synthase activity controls the biosynthesis of carotenoids and the supply of their metabolic precursors in dark-grown Arabidopsis seedlings. *Plant J.* 60, 424–435.

Rosas-Saavedra, C., and Stange, C. (2016). "Biosynthesis of carotenoids in plants: enzymes and color," in *Carotenoids in Nature. Subcellular Biochemistry*, Vol. 79, ed. C. Stange (Cham: Springer), 35–69.

Rozen, S., and Skaletsky, H. (2000). "Primer3 on the WWW for general users and for biologist programmers," in *Bioinformatics Methods and Protocols. Methods in Molecular Biology*, Vol. 132, eds S. Misener and S. A. Krawetz (Totowa, NJ: Humana Press), 365–386.

Ruiz-Sola, M. Á, Rodríguez-Villalón, A., and Rodríguez-Concepción, M. (2014). Light-sensitive Phytochrome-Interacting Factors (PIFs) are not required to regulate phytoene synthase gene expression in the root. *Plant Signal. Behav.* 9:e29248.

Sagawa, J. M., Stanley, L. E., LaFountain, A. M., Frank, H. A., Liu, C., and Yuan, Y. W. (2016). An R2R3−MYB transcription factor regulates carotenoid pigmentation in *Mimulus lewisii* flowers. *New Phytol.* 209, 1049–1057.

Saito, R., Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P. L., Lotia, S., et al. (2012). A travel guide to Cytoscape plugins. *Nat. Methods* 9, 1069–1076.

Schaub, P., Rodriguez-Franco, M., Cazzonelli, C. I., Álvarez, D., Wüst, F., and Welsch, R. (2018). Establishment of an Arabidopsis callus system to study the interrelations of biosynthesis, degradation and accumulation of carotenoids. *PLoS One* 13:e0192158. doi: 10.1371/journal.pone.0192158

Shumskaya, M., Bradbury, L. M., Monaco, R. R., and Wurtzel, E. T. (2012). Plastid localization of the key carotenoid enzyme phytoene synthase is altered by isozyme, allelic variation, and activity. *Plant Cell* 24, 3725–3741.

Stanley, L. E., and Yuan, Y. (2019). Transcriptional regulation of carotenoid biosynthesis in plants: So many regulators, so little consensus. *Front. Plant Sci.* 10:1017. doi: 10.3389/fpls.2019.01017

Stracke, R., Werber, M., and Weisshaar, B. (2001). The R2R3-MYB gene family in Arabidopsis thaliana. *Curr. Opin. Plant Biol.* 4, 447–456.

Sun, T., and Li, L. (2020). Toward the 'golden' era: the status in uncovering the regulatory control of carotenoid accumulation in plants. *Plant Sci.* 290:110331.

Toledo-Ortiz, G., Huq, E., and Rodríguez-Concepción, M. (2010). Direct regulation of phytoene synthase gene expression and carotenoid biosynthesis by phytochrome-interacting factors. *Proc. Natl. Acad. Sci. U. S. A.* 107, 11626–11631.

Tzuri, G., Zhou, X., Chayut, N., Yuan, H., Portnoy, V., Meir, A., et al. (2015). A 'golden'SNP in CmOr governs the fruit flesh color of melon (*Cucumis melo*). *Plant J.* 82, 267–279.

Velasco, R., Zharkikh, A., Affourtit, J., Dhingra, A., Cestaro, A., Kalyanaraman, A., et al. (2010). The genome of the domesticated apple (*Malus*× *domestica* Borkh.). *Nat. Genet.* 42, 833–839.

Walter, W., Sánchez-Cabo, F., and Ricote, M. (2015). GOplot: an R package for visually combining expression data with functional analysis. *Bioinformatics* 31, 2912–2914.

Wang, L., Qiu, T., Yue, J., Guo, N., He, Y., Han, X., et al. (2021). Arabidopsis ADF1 regulated by *MYB73* is involved in response to salt stress via affecting actin filaments organization. *Plant Cell Physiol.* 62, 1387–1395.

Wang, W., Wang, P., Li, X., Wang, Y., Tian, S., and Qin, G. (2021). The transcription factor SlHY5 regulates the ripening of tomato fruit at both the transcriptional and translational levels. *Hortic. Res.* 8:83.

Weber, D., and Grune, T. (2012). The contribution of β-carotene to vitamin A supply of humans. *Mol. Nutr. Food Res.* 56, 251–258.

Welsch, R., Arango, J., Bär, C., Salazar, B., Al-Babili, S., Beltrán, J., et al. (2010). Provitamin A accumulation in cassava (*Manihot esculenta*) roots driven by a single nucleotide polymorphism in a phytoene synthase gene. *Plant Cell* 22, 3348–3356.

Welsch, R., Maass, D., Voegel, T., DellaPenna, D., and Beyer, P. (2007). Transcription factor RAP2. 2 and its interacting partner SINAT2: stable elements in the carotenogenesis of Arabidopsis leaves. *Plant Physiol.* 145, 1073–1085.

Welsch, R., Zhou, X., Yuan, H., Álvarez, D., Sun, T., Schlossarek, D., et al. (2018). Clp protease and OR directly control the proteostasis of phytoene synthase, the crucial enzyme for carotenoid biosynthesis in Arabidopsis. *Mol. Plant* 11, 149–162.

Yao, J. L., Cohen, D., Atkinson, R., Richardson, K., and Morris, B. (1995). Regeneration of transgenic plants from the commercial apple cultivar Royal Gala. *Plant Cell Rep.* 14, 407–412.

You, M., Lee, Y., Kim, J., Baek, S., Jeon, Y., Lim, S., et al. (2020). The organ-specific differential roles of rice DXS and DXR, the first two enzymes of the MEP pathway, in carotenoid metabolism in *Oryza sativa* leaves and seeds. *BMC Plant Biol.* 20:167. doi: 10.1186/s12870-020-02357-9

Young, M. D., Wakefield, M. J., Smyth, G. K., and Oshlack, A. (2012). goseq: gene Ontology testing for RNA-seq datasets. *R Bioconductor* 8, 1–25.

Zeng, J., Wang, X., Miao, Y., Wang, C., Zang, M., Chen, X., et al. (2015). Metabolic engineering of wheat provitamin A by simultaneously overexpressing CrtB and silencing carotenoid hydroxylase (TaHYD). *J. Agric. Food Chem.* 63, 9083–9092.

Zhang, R., Fu, X., Zhao, C., Cheng, J., Liao, H., Wang, P., et al. (2020). Identification of the key regulatory genes involved in elaborate petal development and specialized character formation in *Nigella damascena* (Ranunculaceae). *Plant Cell* 32, 3095–3112.

Zhang, X. J., Gao, Y., Wen, P. F., Hao, Y. Y., and Chen, X. X. (2018). Cloning and expression analysis of the phytoene synthase gene in 'Granny Smith' apple (Malus × domestica Borkh.). *Biotechnol. Biotechnol. Equip.* 32, 1105–1112.

Zhou, D., Shen, Y., Zhou, P., Fatima, M., Lin, J., Yue, J., et al. (2019). Papaya CpbHLH1/2 regulate carotenoid biosynthesis-related genes during papaya fruit ripening. *Hortic. Res.* 6:80.

Zhou, F., Wang, C. Y., Gutensohn, M., Jiang, L., Zhang, P., Zhang, D., et al. (2017). A recruiting protein of geranylgeranyl diphosphate synthase controls metabolic flux toward chlorophyll biosynthesis in rice. *Proc. Natl. Acad. Sci. U. S. A.* 114, 6866–6871.

Zhou, X., Rao, S., Wrightstone, E., Sun, T., Lui, A., Welsch, R., et al. (2022). Phytoene synthase: the key rate-limiting enzyme of carotenoid biosynthesis in plants. *Front. Plant Sci.* 13:977. doi: 10.3389/fpls.2022.884720

Zhou, X., Welsch, R., Yang, Y., Álvarez, D., Riediger, M., Yuan, H., et al. (2015). Arabidopsis OR proteins are the major posttranscriptional regulators of phytoene synthase in controlling carotenoid biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* 112, 3558–3563.

Zhu, C., Bai, C., Sanahuja, G., Yuan, D., Farré, G., Naqvi, S., et al. (2010). The regulation of carotenoid pigmentation in flowers. *Archi. Biochem. Biophys.* 504, 132–141.

Zhu, F., Luo, T., Liu, C., Wang, Y., Yang, H., Yang, W., et al. (2017). An R2R3−MYB transcription factor represses the transformation of α−and β−branch carotenoids by negatively regulating expression of *CrBCH2* and *CrNCED5* in flavedo of Citrus reticulate. *New Phytol.* 216, 178–192.

Zhu, Z., Chen, G., Guo, X., Yin, W., Yu, X., Hu, J., et al. (2017). Overexpression of *SlPRE2*, an atypical bHLH transcription factor, affects plant morphology and fruit pigment accumulation in tomato. *Sci. Rep.* 7:5786.

Zhu, M., Chen, G., Zhou, S., Tu, Y., Wang, Y., Dong, T., et al. (2014). A new tomato NAC (N AM/A TAF1/2/C UC2) transcription factor, *SlNAC4*, functions as a positive regulator of fruit ripening and carotenoid accumulation. *Plant Cell Physiol.* 55, 119–135.

Zhu, Z., Liang, H., Chen, G., Li, F., Wang, Y., Liao, C., et al. (2019). The bHLH transcription factor *SlPRE2* regulates tomato fruit development and modulates plant response to gibberellin. *Plant Cell Rep.* 38, 1053–1064.

Zogopoulos, V. L., Saxami, G., Malatras, A., Angelopoulou, A., Jen, C. H., Duddy, W. J., et al. (2021). Arabidopsis Coexpression Tool: a tool for gene coexpression analysis in *Arabidopsis thaliana*. *Iscience* 24:102848.

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 20 September 2022
DOI 10.3389/fpls.2022.976901

OPEN ACCESS

EDITED BY
Octavio Martínez,
Centro de Investigación y Estudios Avanzados del IPN (CINVESTAV), Mexico

REVIEWED BY
Qing-Hua Gao,
Shanghai Academy of Agricultural Sciences, China
Gholamreza Khaksar,
Chulalongkorn University, Thailand

*CORRESPONDENCE
Raúl Herrera
raherre@utalca.cl
María Alejandra Moya-Leon
alemoya@utalca.cl

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Plant Development and EvoDevo, a section of the journal Frontiers in Plant Science

RECEIVED 23 June 2022
ACCEPTED 22 August 2022
PUBLISHED 20 September 2022

CITATION
Gaete-Eastman C, Stappung Y, Molinett S, Urbina D, Moya-Leon MA and Herrera R (2022) RNAseq, transcriptome analysis and identification of DEGs involved in development and ripening of *Fragaria chiloensis* fruit.
Front. Plant Sci. 13:976901.
doi: 10.3389/fpls.2022.976901

RNAseq, transcriptome analysis and identification of DEGs involved in development and ripening of *Fragaria chiloensis* fruit

Carlos Gaete-Eastman†, Yazmina Stappung†, Sebastián Molinett, Daniela Urbina, María Alejandra Moya-Leon* and Raúl Herrera*

Laboratorio de Fisiología Vegetal y Genética Molecular, Instituto de Ciencias Biológicas, Universidad de Talca, Talca, Chile

Fragaria chiloensis (Chilean strawberry) is a native species that produces fruit with an exotic pinkish color and a fruity aroma. It has a non-climacteric pattern of fruit ripening, and it is the mother of the commercial *Fragaria* x *ananassa*. The ripening of *F. chiloensis* fruit seems stimulated by ABA, and a complete set of genes participate in its softening, color, and aroma development. In addition, a set of transcription factors regulate the entire process, but few of them have been described. Over the last two decades, RNA-seq was used to identify genes at three fruit development/ripening stages, named C2 (unripe, large green) to C4 (full ripe), in whole fruit and fruit without achenes. A total of 204,754 contigs were assembled considering all samples, obtaining an N50 of 1.125 bp. Differentially expressed genes (DEGs) between two samples were identified, obtaining a total of 77,181 DEGs. Transcripts for genes involved in ABA biosynthesis present high and differential expression during the C2, C3, and C4 stages. Besides, contigs corresponding to ABA receptors, which interact with a regulatory network, are also differentially expressed. Genes associated with cell wall remodeling and those involved in flavonoid synthesis were also differentially expressed. An interaction network was built considering differentially expressed genes for the phenylpropanoid and flavonoid molecular pathways and having FcMYB1 as a transcription factor regulator. Identifying key genes could give an option to control the ripening of this non-climacteric fruit.

KEYWORDS

Illumina sequencing, Chilean strawberry fruit, fruit ripening, cell wall enzymes, ABA genes, flavonoids genes

Introduction

Fragaria chiloensis (L.), a mill, is a non-climacteric fruit with many valuable traits, such as a fruity aroma, a pinkish color, and high nutritional content (Letelier et al., 2020). The species has been established in different edaphoclimatic conditions in Chile. However, the major production is around the valleys close to Nahuelbuta mountain (cities of Purén and Contulmo) (Letelier et al., 2020). As a *Fragaria* species, the Chilean strawberry fruit is unique because the edible flesh is the enlarged receptacle tissue. The true fruit is the numerous dry achenes on the receptacle surface. The exchange of molecules between achenes and receptacles can coordinate the development and ripening of the fruit. In this sense, phytohormones produced by achenes are essential for the ripening of fruit receptacles, but little is known about the influence of achenes signals on fruit ripening in Chilean strawberries.

The suppression subtractive hybridization (SSH) strategy was used in our first attempt to unravel genes involved during the development and ripening of *Fragaria chiloensis* fruit (Pimentel et al., 2010). A set of 1,807 genes were isolated from six SSH libraries expressed during ripening, and 13 of them were analyzed by qPCR to validate their role in fruit development and ripening. Three genes, *FcPL*, *FcPG,* and *FcEG,* showed a high accumulation of transcripts at the late fruit ripening stages and are specifically expressed in fruits (Pimentel et al., 2010). Interestingly, auxin-related genes and the transcription factor MADS1 were also differentially expressed at the late fruit ripening stages but showed transcript accumulation in other tissues like flowers and runners.

Nowadays, transcriptome characterization through RNA-seq can achieve the expression profile of a larger number of genes, giving a more detailed insight into the molecular mechanism involved. Sequences can be efficient *de novo* assembled in organisms with reference genomes. A major interest has concentrated on studying the commercial *F. x ananassa* and the diploid *Fragaria vesca* (Hollender et al., 2014; Chen et al., 2016; Wang et al., 2017). These broad transcriptomic studies allowed the identification of differentially expressed genes during the development and ripening of fruit. The roles of plant hormones have been studied by applying the plant hormones abscisic acid (ABA) or auxin to the fruit (Li et al., 2015; Chen et al., 2016). Interestingly, different genes and metabolic pathways were identified when ABA and auxin were used during the fruit ripening of strawberries. The key role of ABA and auxins has been shown in the ripening of strawberry fruits, even though different experimental approaches have been used (Li et al., 2015; Chen et al., 2016).

Fragaria chiloensis (L.) Mill, or the Chilean strawberry, is a native fruit of Chile (Letelier et al., 2020). It is appreciated for its good organoleptic qualities, being its sweet and pleasant aroma, the main characteristic of this non-climacteric fruit, in addition to its big fruit size, resistance to pathogens, and better sustainability to soil salinity and low temperature (González et al., 2009, 2013). Furthermore, *F. chiloensis* is the maternal relative and, therefore, the gene source of the commercial strawberry (*Fragaria* × *ananassa*). *F. chiloensis* has the potential to be developed as a new exotic berry in the world market (Retamales et al., 2005). However, the fruit is highly perishable as its rapid softening alters the texture and negatively influences its postharvest life and quality (Perkins-Veazie, 1995; Figueroa et al., 2008).

In this study, RNAseq analysis was performed in *F. chiloensis* considering three different fruit stages and two different types of samples: whole fruit (receptacle with achenes) and only receptacles (without achenes). To understand the molecular mechanisms underlying the differential contribution of achenes and receptacles during the ripening of this strawberry fruit, we performed a comprehensive transcriptome analysis at three developing/ripening fruit stages: immature-large green (C2), turning (C3), and ripe (C4) stages. These results allow the understanding of the molecular regulation of fruit ripening in this non-climacteric species and provide pieces of evidence of the participation of achenes and receptacle tissues in ripening.

Materials and methods

Fruit samples

Fragaria chiloensis fruit samples were collected from commercial orchards in Purén (the Araucania region, Chile) and immediately transported to the laboratory. Fruit samples were segregated according to external phenotype into three developing and ripening fruit stages as described by Figueroa et al. (2008): C2, large green (large size fruit with red achenes); C3, turning (large size fruit with white receptacle and red achenes) and C4, ripe fruit (full-size fruit with pink receptacle and red/brown achenes). Receptacle fruit samples were obtained after the careful remotion of the achenes from the fruit surface with the help of tweezers. Complete fruit and deachened (receptacle) fruit samples were cut into pieces, frozen under liquid nitrogen, and stored at −80°C as bulks until their use.

RNA extraction, library construction, and Illumina sequencing

Total RNA was extracted from representative bulks of complete fruit stages (C2, C3, C4) and receptacle samples (RC2, RC3, RC4) using the methodology described by Pimentel et al. (2010). RNA samples were sent to the Next Generation Sequencing (NGS) division at Macrogen Inc. (Seoul, Republic of Korea). One RNA sample per condition was subjected to the

library construction protocol at Macrogen and then sequenced with the Illumina HiSeq 2000 platform; 100-bp paired-end reads were generated.

Sequence data analysis and assembly

Raw sequence data from high-throughput NGS sequencing services were subjected to quality control using the software FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The quality of each sample was checked based on the analysis of per base sequence quality (value >20), per sequence quality score (value of 38), per sequence G.C. content (normal distribution/46%), and percentage of over-represented sequences. Three sequences were identified over the threshold and identified as metallothionein-like proteins.

RNA-seq *de novo* assembly was performed at the NGS division of Macrogen Inc. by using Trinity (http://trinityrnaseq.sourceforge.net). The sequences for this project were loaded into BioProject (ID PRJNA849949).

Functional annotation

Several complementary approaches were employed to annotate the unigenes (Figure 1). The *de novo* obtained sequences were analyzed using the Basic Local Alignment Search Tool (BLAST) (Altschul et al., 1990), and the sequences were compared against reference genomes of four species: three from the Rosaceae family (apple, *Malus domestica 1.0*; peach, *Prunus persica 1.0*; woodland strawberry, *Fragaria vesca 1.1*) and poplar (*Populus trichocarpa 2.0*). Then the genes were functionally annotated under gene ontology terms (G.O., http://www.geneontology.org). Sequences with hits from the output analysis were used to obtain a tentative annotation, which was manually confirmed for a subset of sequences under study in the present work.

According to Trinity methodology, abundance estimation of transcripts was determined by the RSEM alignment-based quantification method, comparing two different sample conditions at a time and obtaining the relative abundance separately for each sample. The normalized expression metrics were reported as FPKM (fragments per Kilobase transcript length per million fragments mapped).

Two validation strategies were used with the assembled data. First, a pair-wise sequence alignment of 7 selected ripening-related genes from *F. chiloensis*, previously published and experimentally tested, was performed against the RNA seq data. Secondly, a sequence mapping of SSH libraries previously reported by Pimentel et al. (2010) against the RNA seq was performed.

Analysis of expression

The analysis of differentially expressed genes was performed by comparing the following pairs of sequenced libraries: C2 vs. C3; C3 vs. C4; C2 vs. C4; RC2 vs. RC3; RC3 vs. RC4; RC2 vs. RC4; C2 vs. RC2; C3 vs. RC3; C4 vs. RC4. In all these comparisons, the first library (underlined) was the control, and the second was the test library. Fold change and hierarchical clustering were the statistical methods used in the nine comparisons described, and up/down DEGs were counted.

RT-qPCR analysis

RNA extractions from intact (C samples) and de-achened fruit (R.C. samples) were followed by cDNA (complementary DNA) synthesis, as previously reported (Opazo et al., 2010). In short, 8 g of fruit samples detailed above were extracted using the CTAB method (Chang et al., 1993). Then, total RNA samples were treated with DNase I (Invitrogen) and cleaned up with the RNeasy Plant Mini Kit (Qiagen). Finally, cDNA synthesis was performed using the First Strand cDNA Synthesis kit (Fermentas). RT-qPCR analyses were performed as previously described (Opazo et al., 2013). Briefly, three biological replicates were used for each sample analyzed. All the primers were designed against (or at least one of them), 3'-UTR regions. qPCR reactions were performed using Maxima SYBR Green qRT-PCR Master Mix (Fermentas) following the manufacturer's recommendations in Stratagene Mx300P (Agilent Technologies). Relative expression levels of each gene under study, representing the mean between three biological replicates, were calculated using the method described by Pfaffl (2001) and employing *FcGAPDH1* as a normalizer.

Heatmap analysis

A color-coded two-dimensional mosaic describing the whole expression matrix (samples vs. gene targets) was built according to Sun and Li (2013). Each tile was colored with different intensity levels according to the preprocessed data expression. Gene expression values can be visualized with the color density ranging from the least expressed (blue) to the most expressed (red).

String interaction network

String is a database where known and predicted direct (physical) interactions, as well as indirect (functional) interactions, can be established based on co-expression, co-localization, text-mining, and others (Szklarczyk et al., 2017, 2019). The DEGs involved in synthesizing

10.3389/fpls.2022.976901

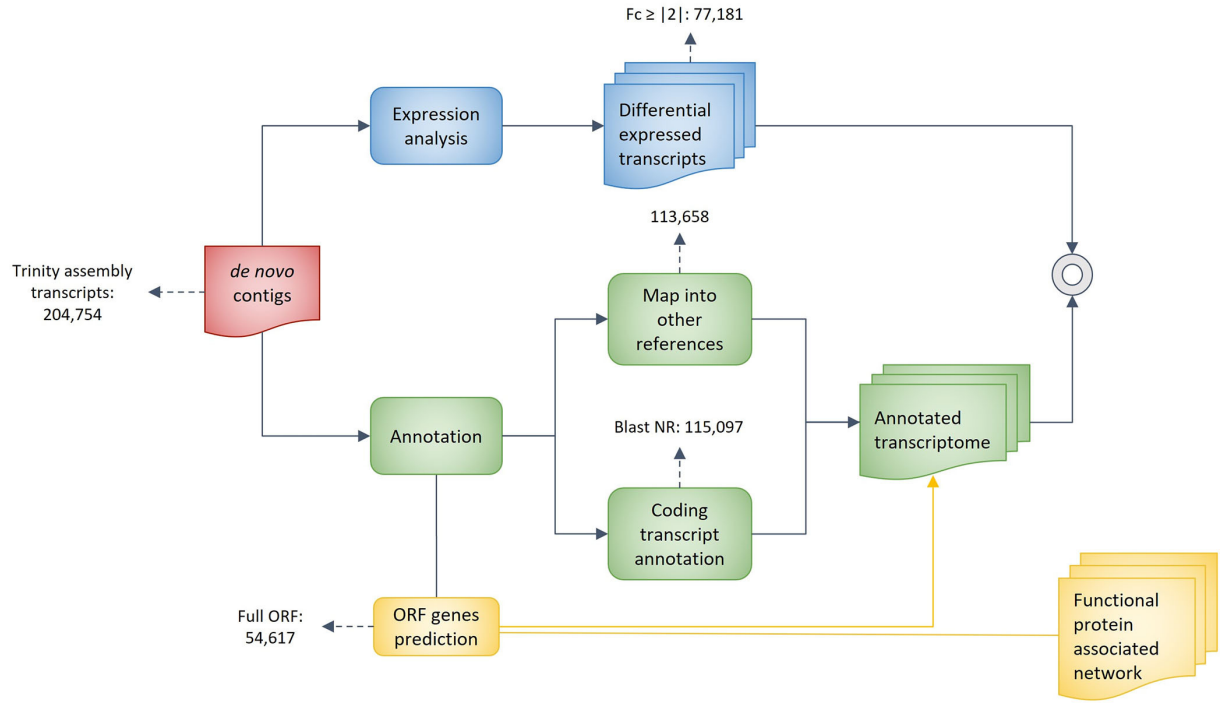


FIGURE 1
Flow diagram of gene annotation using transcriptomic reference genome. de novo assembled contigs were annotated using four reference genomes and NR database. Full ORF predicted genes from annotation stage were used for functional protein associated network.

flavonoids and FcMYB were picked, and the web server was interrogated to uncover potential protein-protein association networks. The database was interrogated for the last time in December 2021.

Results

Fragaria chiloensis fruit samples were harvested by local farmers in Purén, in the Araucania region of Chile. Fruits from different development and ripening stages were collected and separated into three different groups.

The NGS sequencing data performed transcriptomic analysis on the three fruit development stages. More than sixty million pair-end reads were generated for each sample (Supplementary Table 1). Because there is no reference genome for *Fragaria chiloensis*, a strict assembly pipeline for annotation and validation was applied (Figure 1). A total of 204,754 tentative contigs were assembled considering all libraries, with an N50 of 1,125 bp (Table 1). Bearing in mind that all libraries were sequenced on one line of Illumina HiSeq2000, the total throughput per library was homogeneous and near the maximum performance of 35 Gb per line. Paired-end reads of 101 bp were obtained for each sample constructed. Finally, read quality, trimming, and the check for no contaminant sequences were examined. The quality of each library was performed by: analysis of per base sequence quality (value >20), per sequence quality score (value of 38), per sequence G.C. content' (normal distribution fitting indicating no contamination with foreign DNA), and percentage of overrepresented sequences (3 sequences over the threshold: metallothionein-like protein). Two strategies were used to annotate all the sequences: first, four reference genomes were used, and second, all sequences were BLAST locally against the non-redundant (nr) database.

TABLE 1 Summary of *de novo* transcriptome assembly.

Assemble statistics	Merge
The total length of contigs	142,626,488
Total number of contigs	204,754
Max length bp	13,470
Min length bp	201
N90 bp	275
N80 bp	389
N70 bp	565
N60 bp	812
N50 bp	1,125

Information of total contigs and general features were obtained after de novo transcript assembly using Trinity software (2011-11-26 version).

The functional annotation stage was performed through the mapping comparison to three Rosaceae species (apple, peach, and woodland strawberry) and poplar, obtaining a reference of 113,658 contigs (Figure 2). Approximately 20 % of the total assembled contigs were assigned a corresponding cross-entry within the four reference genomes employed. The Fragaria genome contributes the largest of the annotations of the *F. chiloensis* transcriptome, contributing 55% of the total of sequenced contigs. On the other hand, 5,083 annotated contigs were unique between *F. vesca* and *F. chiloensis*.

The validation of assembled data was contrasted with seven selected ripening-related genes from *F. chiloensis*, indicating a complete coverage of the sequence and a very high sequence identity (Supplementary Table 2). Additionally, the alignment of the ESTs from the SSH library (Pimentel et al., 2010) with contigs from the RNA seq transcriptome allows the mapping of complete coverage sequences, in which 67% of ESTs share between 95 and 100% of sequence identity.

The annotation through the N.R. database provided 115,097 contigs, which corresponded to 56% of the total sequenced contigs (Figure 1). Also, an enrichment analysis by Gene Ontology (G.O.) assigned at least one G.O. term to 48.6% of the sequences (Figure 3). The G.O. categories were molecular function (44%), biological process (36%), and cellular component (20%). The "ion binding" subcategory is strongly represented in the molecular function category, followed by molecular function. In the cellular component category, some subcategories that were highly represented included cellular components, nuclei, and protein-containing complexes. In the case of the biological process category, the most represented subcategory is a biosynthetic process. Other important subcategories are the cellular protein modification process, cellular nitrogen compound metabolic process, and DNA metabolic process.

Analysis of differentially expressed genes

To perform an expression analysis, all those assembled contigs having at least one zero FPKM value were excluded from this study, leaving a final total of 77,181 contigs for the analysis (Supplementary Figure 1). The conditions were

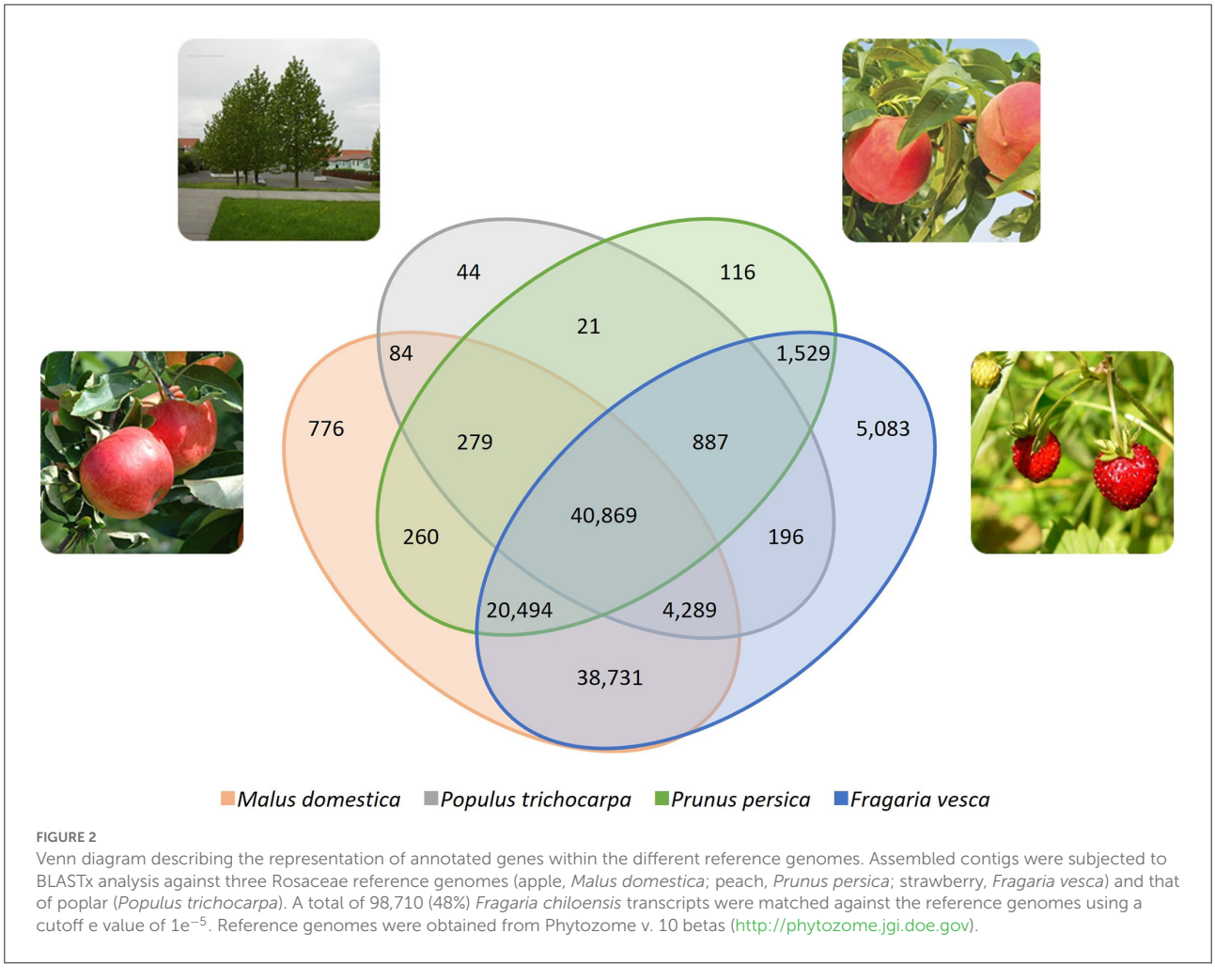


FIGURE 2
Venn diagram describing the representation of annotated genes within the different reference genomes. Assembled contigs were subjected to BLASTx analysis against three Rosaceae reference genomes (apple, *Malus domestica*; peach, *Prunus persica*; strawberry, *Fragaria vesca*) and that of poplar (*Populus trichocarpa*). A total of 98,710 (48%) *Fragaria chiloensis* transcripts were matched against the reference genomes using a cutoff e value of $1e^{-5}$. Reference genomes were obtained from Phytozome v. 10 betas (http://phytozome.jgi.doe.gov).

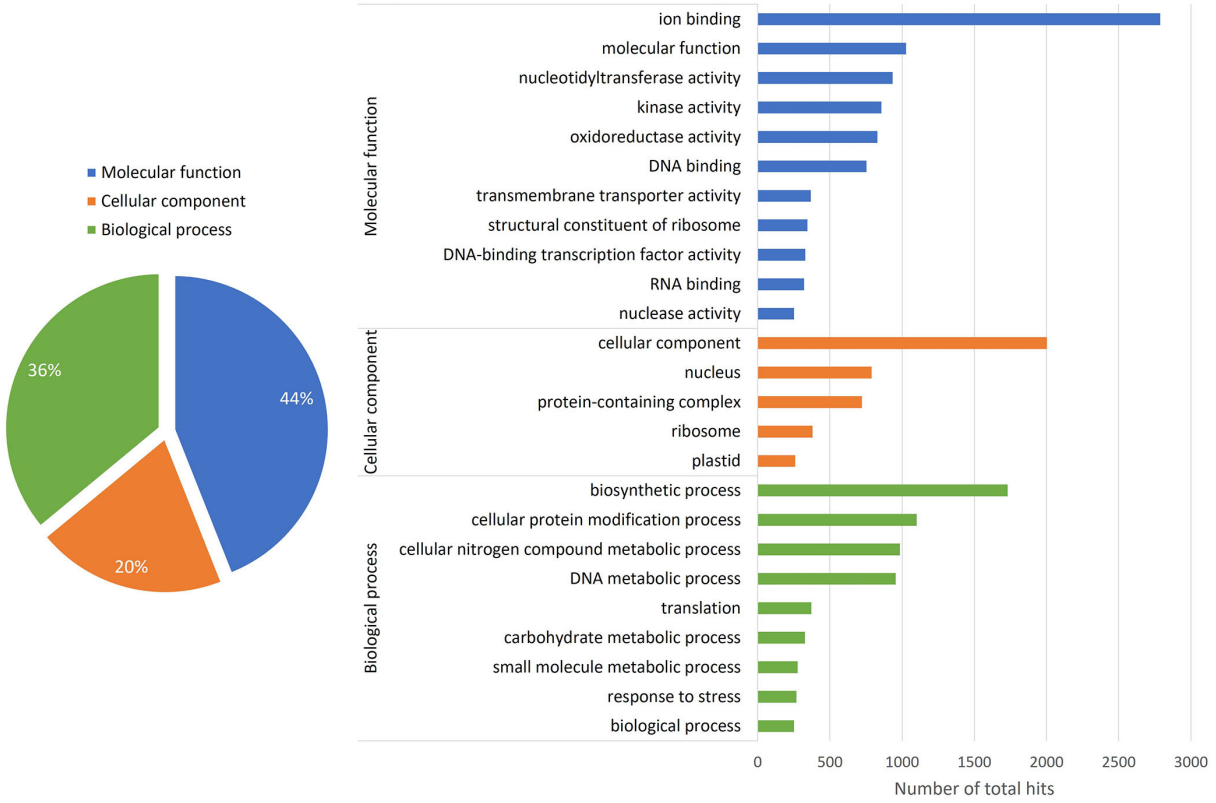


FIGURE 3
Classification of contigs under Gene Ontology terms. A total of 11,825 contigs were assigned with G.O. terms within the three main categories (Biological process, Cellular components, and Molecular function) and subdivided into their corresponding subcategories. Only relevant G.O. categories (≥ 100 hits per subcategory) are shown.

comparable after performing the distribution of normalized FPKM values before and after log2 transformation, and with or without quantile normalization. Then, hierarchical clustering was performed using the correlations between stages, which grouped more similar samples of this study (Supplementary Figure 2). Interestingly, the C3 stage (sample corresponding to whole fruit at turning stage) grouped with C2 and RC2 stage (samples corresponding to large green whole fruit and receptacle, respectively); while RC3 stage (sample corresponding to receptacle fruit at turning stage) grouped with C4 and RC4 stage (samples corresponding to a ripe whole and receptacle fruit, respectively).

The number of differentially expressed contigs between pairs of fruit development ripening stages was analyzed by a fold change ≥2 (Figure 4). In most of these comparisons, the number of significant up and downregulated contigs was similar. However, the highest number of differentially expressed contigs (DEG) was observed when the developmental stages of C4 vs. C2, RC3 vs. RC2, and RC4 vs. RC2 were compared. On the other hand, the lowest number of differentially expressed contigs was observed between the receptacle, and a whole fruit sample from stage 2 was compared.

Validation of RNAseq expression analysis

Two complementary strategies were used to validate the expression analysis data. First, the FPKM values of a random set of genes were compared with expression values determined by RT-qPCR in the same RNA samples employed for the RNA seq library ($r^2 = 0.90$) (Supplementary Table 3). Secondly, FPKM values of a selected group of ripening-related genes were compared to expression data previously reported for the specie (Salvatierra et al., 2010) ($r^2 = 0.78$) (Supplementary Figure 3). The correlation values between the changes in expression levels for different genes during the development and ripening stages of *F. chiloensis* using both strategies was $r^2 = 0.84$ (Figure 5), indicating a good correlation between FPKM values and RT-qPCR data.

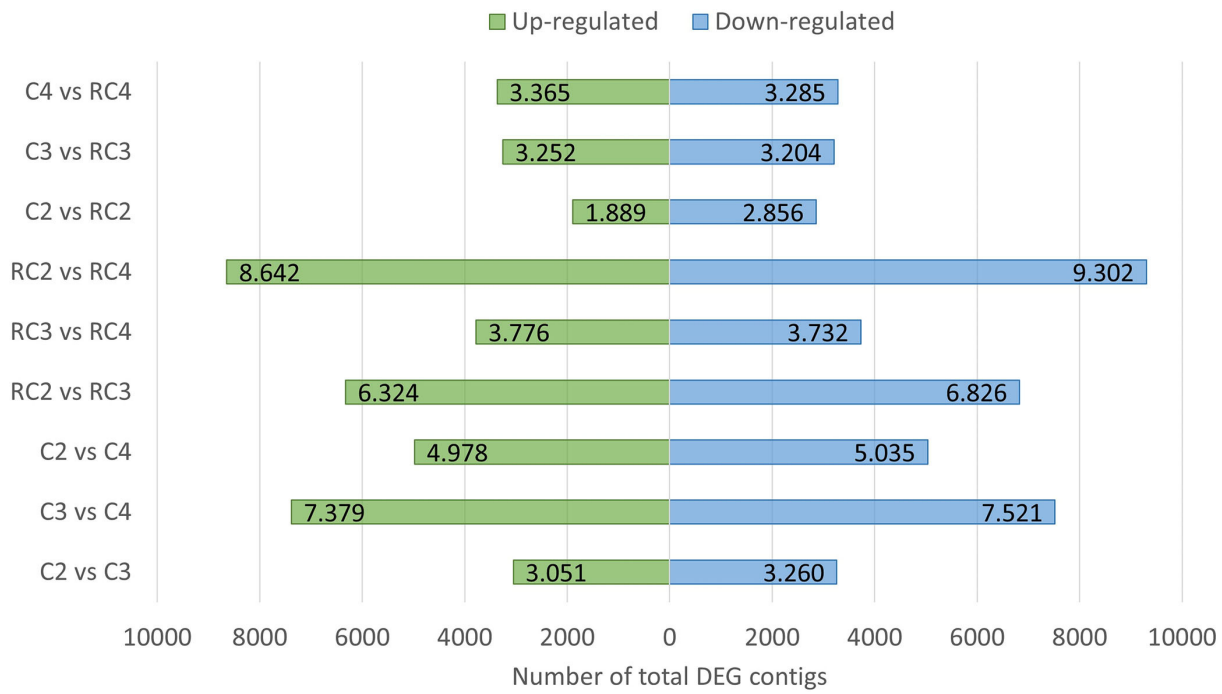


FIGURE 4
Up and down expressed genes in the different developmental and ripening stages. Contigs with a fold change ≥ [2] between the two samples are indicated. In each comparison, the first sample is the testing one, and the second is the control. The analysis was performed on 77,181 contigs, as 127,573 contigs were not considered for having at least one zero FPKM value in any of the libraries.

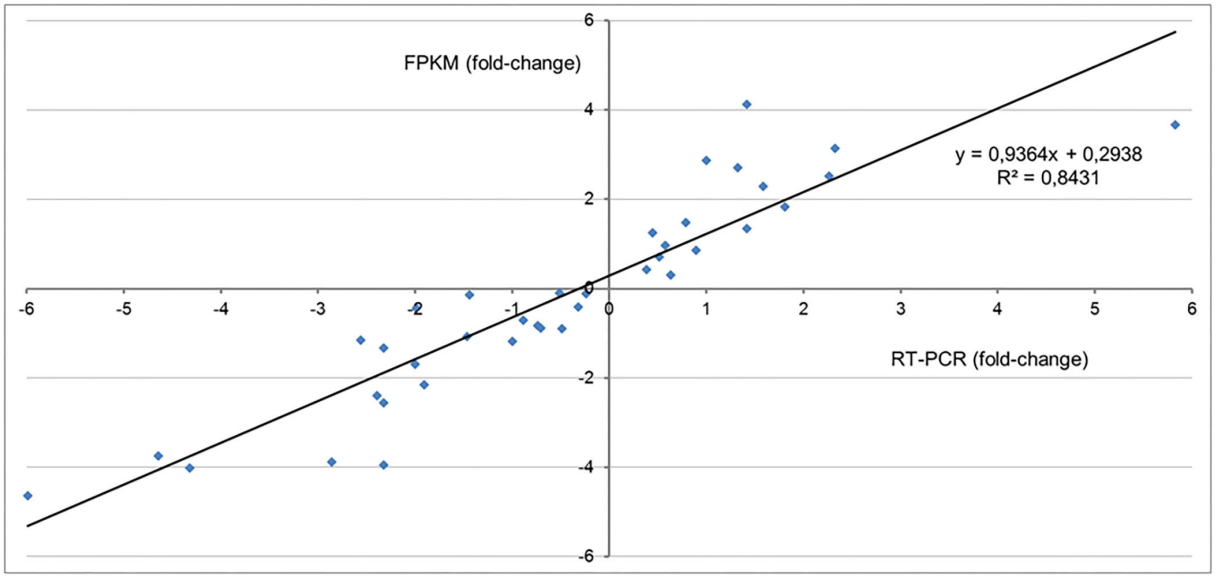


FIGURE 5
Validation of RNAseq analysis. The graph displays the relationship between the changes in expression determined by RT-qPCR analysis and FPKM data from the RNA seq analysis. A correlation value of 0.8 was determined. The ratio of changes in the transcriptional level of genes comparing two fruit ripening/developmental stages was obtained from the RNAseq data [log2(FPKM)] and compared to expression data quantified by RT-qPCR analysis or data previously reported for the specie (Salvatierra et al., 2010) (For more information see Supplementary Figures 3A,B, and Supplementary Table 3).

Data mining of differentially expressed genes

Data mining of differentially expressed genes was performed to identify relevant molecular pathways during the development and ripening of *F. chiloensis* fruit. In addition, receptacle and whole fruit samples were analyzed. Several molecular and biological processes involved in fruit ripening were determined as differentially expressed, which include the phenylpropanoid pathway, synthesis of flavonoids and anthocyanins, enzymes involved in cell wall metabolism, biosynthesis of aroma, and phytohormones. Transcription factors were also detected as differentially expressed in specific expression dynamics during fruit development and ripening. An apparent molecular fine-tuning is observed in terms of different isoforms of the same gene, as some are differentially expressed in different developmental stages and tissues, but in general, functional redundancy is observed.

Based on the FPKM value, 39 tentative contigs corresponding to genes related to ABA regulation were analyzed by comparing fruit development stages. RNA was extracted from the whole fruit or receptacle without achenes (Figure 6). A differential expression pattern between early (C2) and late developmental stages (C3 and C4) was observed. Two main cluster groups can be observed in the heatmap. Genes involved in the synthesis of ABA can be observed in one group, and those genes involved in ABA receptor genes are mostly grouped in the other. Interestingly, three genes, calmodulin, MAP kinase, and phosphatase 2CA, present the same level of transcripts accumulated within all samples. The other 36 genes showed differential expression within the different tissues analyzed. Contigs for genes involved in ABA biosynthesis present high and differential accumulation of transcripts during the C2, C3, and C4 stages. Besides, contigs corresponding to ABA-receptors (PYR family and GCR2 family), phosphatases (PP2C), and protein kinases (SnRK2, CDPKs, and MAPKs), which interact in a regulatory network, are also differentially expressed. Interestingly, some genes are expressed in one type of tissue sample but not the other. For example, four genes (SNF2-1; ABA-induced 1; ABA-induced 4; MCSU1) are expressed only in the whole fruit, and four genes (ARM8; Calmodulin 4; MAPkinase 1; AAO-1) are only expressed in a receptacle. Additionally, 16 genes (ARM_5; ABA induced_3; PYR_1; PYR_4; NCED_2; NCED_1; ARM_7; AAO_2; ARM_6; PYR_3; SDR_3; Calcium dependent_3; SNF1_1; SNF1_3; ZEP_1; Calmodulin_2) show similar expression patterns, but 17 different genes (ARM_10; bZIP_2; ARM_1; ARM_2; GTP-protein_3; Calcium dependent_2; AAO_1; ARM_8; ARM_3; SNF2_2; Phosphatase 2CA_3; Carotenoid_1; SDR_1; Carotenoid_2; MCSU_3; MCSU_2; Phosphatase 2CA_1) are quantitatively expressed in the different developmental and ripening stages.

When 36 DEG genes related to cell walls were analyzed (Figure 7), two cluster groups could be distinguished from the heatmap (Figure 7). Several members of the pectin lyase gene family are in one group, and members of the XTH gene family are in the other cluster group. Also, it can be observed that the two groups differed in the level of expression of the genes. One group is constituted by genes mostly over-expressed during fruit development and ripening, and the other group showed genes that are differentially expressed in one of the fruit stages. The expression pattern showed differences between early (C2) and late developmental stages (C3 and C4) but with divergence concerning fruit tissue. A higher expression level is observed for cell-wall-related genes in receptacles during the final stages of fruit development (RC3 and RC4).

Interestingly, 9 out of 36 genes showed no expression in all ripening stages. On the other hand, the EXP and P.L. grouped together showed similar expression patterns. However, they displayed a lower expression level than PME genes.

When genes involved in the phenylpropanoid and flavonoid pathways were analyzed, several genes showed a differential expression pattern (Figure 8). Key genes involved in the phenylpropanoid pathway are upregulated during the late stages of ripening (C3 and C4). The genes encoding for PAL isoforms, the first committed enzyme of this pathway, and COMT I, involved in lignin biosynthesis, demonstrated this expression profile. Besides, genes involved in the synthesis of flavonoids showed a remarkable accumulation of transcripts for genes associated with the biosynthesis of anthocyanins throughout ripening. Two main clusters can be observed in the heatmap, but one of them shows four sub-clusters. Interestingly, the genes involved in synthesizing flavonoids are on in one sub-cluster. When differentially expressed genes from receptacle were analyzed using STRING, there was an interaction between those genes (Figure 9). Interestingly, FcMYB was linked to FcANS and FcF3H.

Discussion

As the mother of the commercial F. x ananassa, the Chilean strawberry has become an interesting species for gene discovery. Fruit development and ripening in strawberries is an intricate process at the molecular level (Moya-León et al., 2019). Several genes have been identified and characterized for our group in this Chilean native species as involved in the softening process, which include pectin methylesterase (Figueroa et al., 2008), Expansins (Gaete-Eastman et al., 2015), Xyloglucan endotransglycosylase/hydrolase (XTH) (Opazo et al., 2010, 2013; Méndez-Yañez et al., 2017), and Rhamnogalacturonan I lyases (Méndez-Yañez et al., 2020). Other genes/enzymes have been described in the commercial *F.* x *ananassa*, which presented a high index of similarity with its maternal species (Civello et al., 1999; Harrison et al., 2001; Benítez-Burraco et al., 2003;

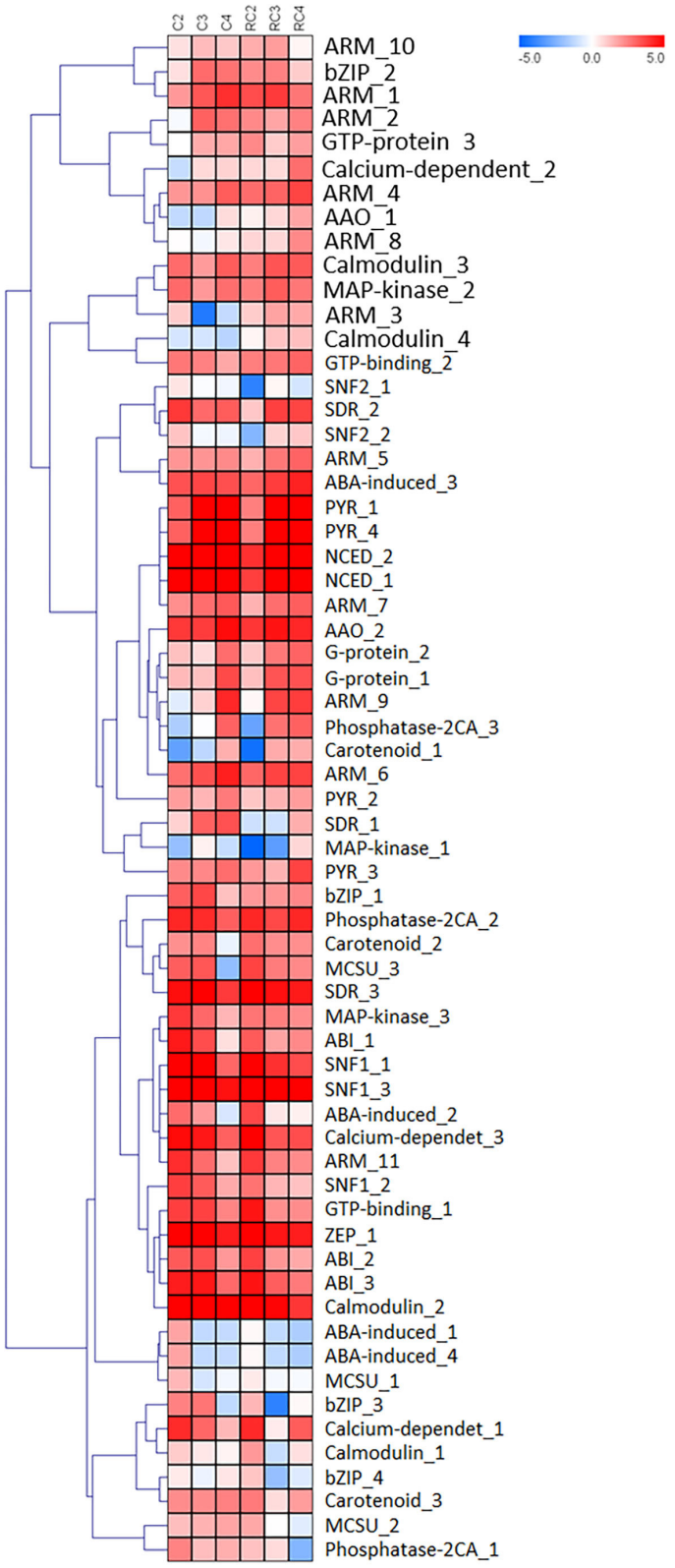


FIGURE 6
HeatMap for genes involved in the synthesis of Abscisic acid and receptors. Log2 is calculated to the FPKM value for each gene to delimit extreme values and used to scale expression. To perform hierarchical clustering for fruit developmental stage and tissue type (receptacle and receptacle without achene), we selected the Pearson Correlation distance metric and Average linkage clustering. Levels of decreased expression are shown in blue and increased expression is shown in red.

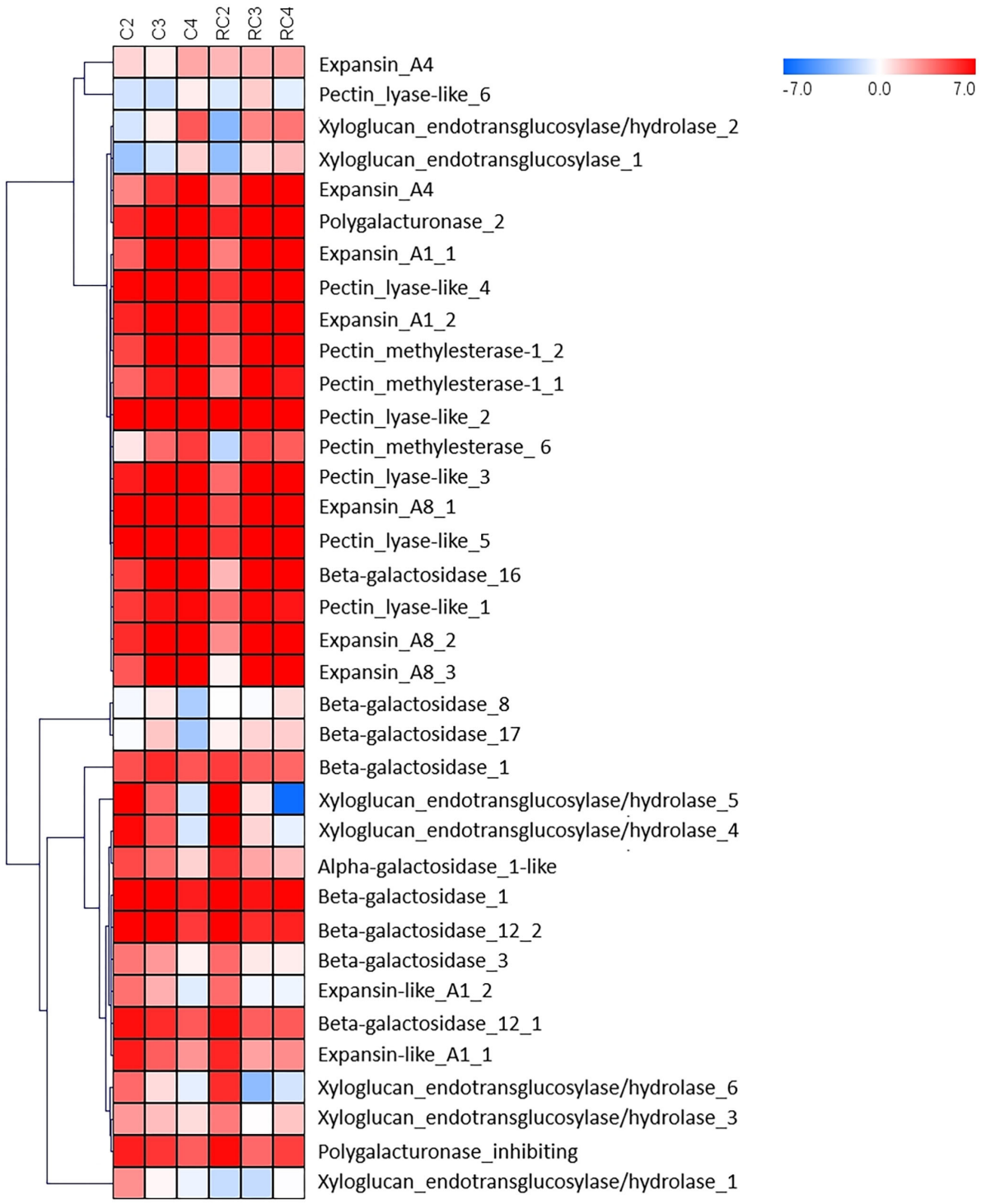


FIGURE 7
HeatMap for genes involved in the dynamics of the cell wall. Log2 is calculated to the FPKM value for each gene to delimit extreme values and used to scale expression. To perform hierarchical clustering for fruit developmental stage and tissue type (receptacle and receptacle without achene), we selected the Pearson Correlation distance metric and average linkage clustering. Levels of decreased expression are shown in blue and increased expression is shown in red.

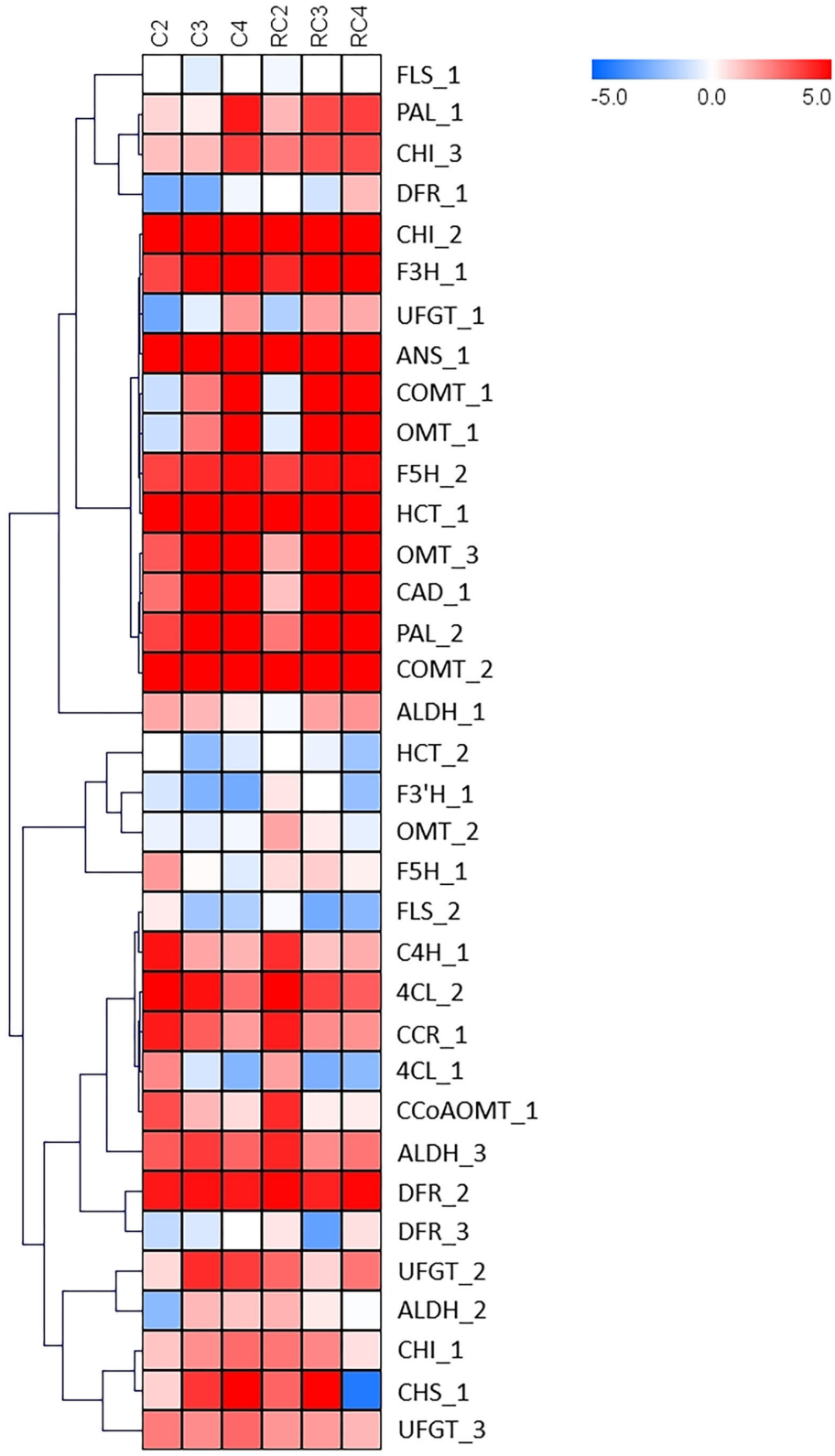


FIGURE 8
HeatMap for genes involved in the synthesis of flavonoids. Log2 is calculated to the FPKM value for each gene to delimit extreme values and used to scale expression. To perform hierarchical clustering for fruit developmental stage and tissue type (receptacle and receptacle without achene), we selected the Pearson Correlation distance metric and Average linkage clustering. Levels of decreased expression are shown in blue and increased expression is shown in red.

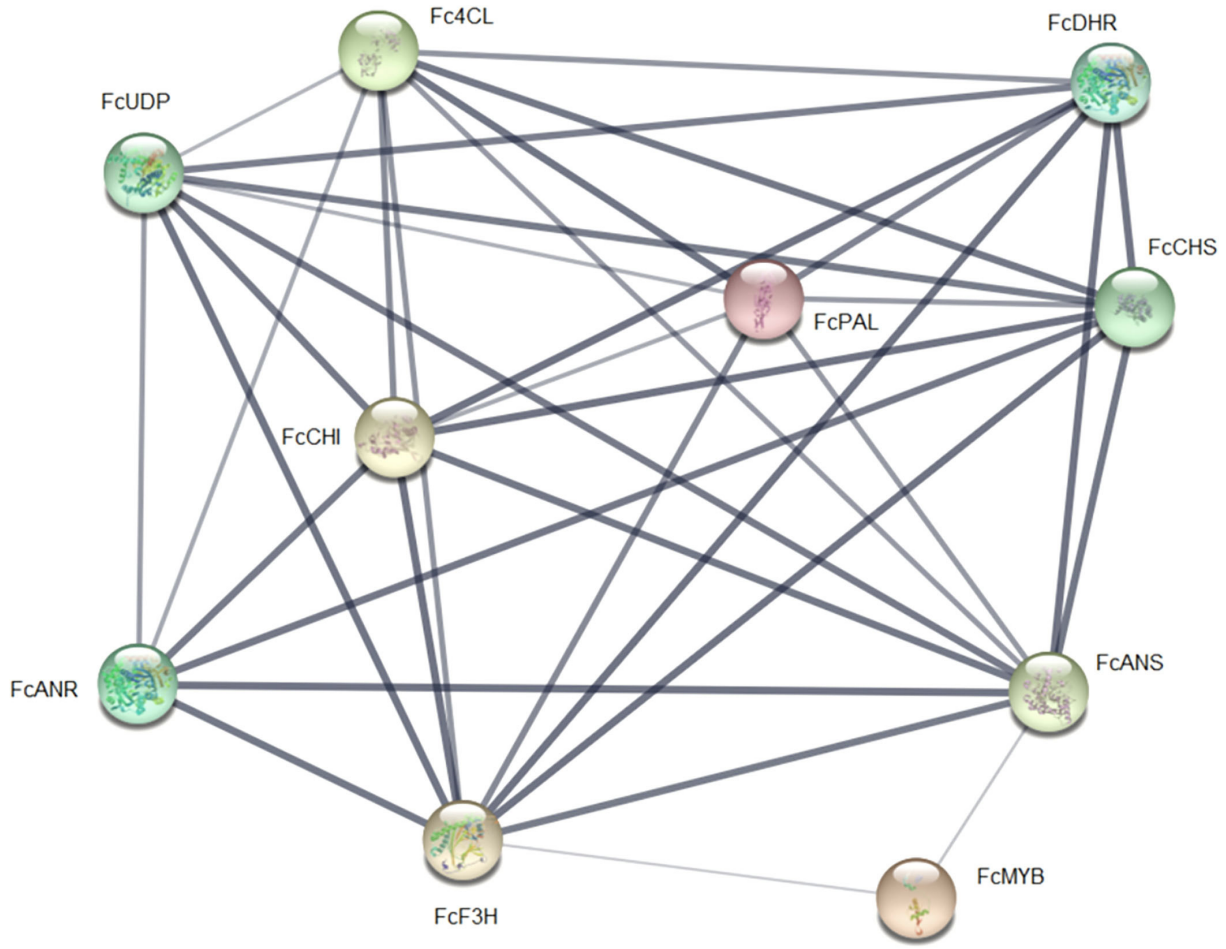


FIGURE 9
Network of gene interactions among genes differentially expressed from the phenylpropanoid and flavonoid pathways and FcMYB using STRING. Each node represents the interaction described in the literature between proteins from the phenylpropanoid and flavonoid pathways and FcMYB1, whose transcripts are differentially expressed in the RNAseq.

Castillejo et al., 2004; Rosli et al., 2004; Santiago-Doménech et al., 2008; Villarreal et al., 2008; García-Gago et al., 2009; Quesada et al., 2009; Molina-Hidalgo et al., 2013).

RNAseq has been used to unravel intricate processes at the molecular level, such as fruit development and ripening in commercial strawberries (Sanchez-Sevilla et al., 2017; Gu et al., 2019; Zhao et al., 2021). The description of a massive number of genes in the Chilean strawberry is described for the first time. The over-expression of metallothionein-like proteins was detected during the quality check of the NGS libraries. The same discovery has been reported in similar studies performed during the development and ripening of pineapple and other Fragaria fruit species (Nam et al., 1999; Moyle et al., 2005).

Gene ontology analyses provide a broad view of the transcriptional changes occurring during a particular biological event. In the case of *F. chiloensis*, most of the transcripts have been included within categories previously defined to be crucial for the development of fruit ripening. The "catalytic activity" and "binding" subcategories are strongly represented in the category of Molecular Function, which represents the active participation of multiple enzymes during fruit development and ripening, as well as the selective interaction between molecules and their specific ligands. In the same category, two other subcategories with high hits were "transporter" and nucleic acid binding transcription factor activities that predict an active transcriptional activity during fruit development and dynamic transport of molecules between compartments. In the category of cellular components, some subcategories were highly represented, including "organelle" and "membrane," which again represent the active participation of several organelles during the strong metabolic activity normally observed during fruit ripening. Similarly, the same categories were found when a suppression subtractive hybridization library was built (Pimentel et al., 2010).

In the case of the biological process category, the most represented subcategory is the metabolic process, and other important subcategories are the response to stimulus, biological regulation, cellular process, and developmental process. As a whole, this classification of the *F. chiloensis* transcripts perfectly represents the active metabolic activity developed during fruit ripening. This developmental process involves many cell compartments and the active movement of intermediates within organelles. The comparisons in the cluster analysis for the different sample stages showed the highest number of differentially expressed contigs (DEG) for the pair's conditions RC3 vs. RC2; RC4 vs. RC2; and C4 vs. C2, which could indicate that major changes in transcription are taking place between the C2 stage and the C3/C4 stages.

The massive number of transcripts obtained in RNAseq provides the opportunity to find other gene family members. In this study, several members of gene families related to cell walls were identified in the analysis of the RNAseq. Thus, members of the P.L., XTH, PME, β-Gal, and Exp family genes showed differences in the accumulation of transcripts during fruit ripening. A possible explanation of these results could be that many of these enzymes participate both in the biosynthesis and degradation of the cell wall, so some isoforms of genes associated with hemicellulose and pectin are early expressed. In contrast, other isoforms are lately expressed, modifying similar cell wall components but most probably breaking down glycosidic bonds. For instance, six new XTH isoforms were identified as differentially expressed during fruit development and ripening in *F. chiloensis*; one XTH showed high similarity to the gene previously reported by Opazo et al. (2010). Several members of gene families are differentially expressed simultaneously during fruit development and ripening. In this sense, a genome-wide analysis in the diploid strawberry *F. vesca* indicated 26 different XTHs, reported as differentially expressed in different tissues. However, only four displayed an accumulation of transcripts during fruit ripening (Opazo et al., 2017).

Interestingly, XTH presented two different types of activities. Different enzyme isoforms could reflect the need for hydrolysis or the modification in the extension of the xylan structure in the cell wall. Additionally, five new P.L.s, one P.G., and five new expansions were found as differentially expressed, but three new PME and eight β-Gal were also reported. There is no doubt that during fruit development and ripening, several events are influenced by the dynamics of the cell wall, which requires the participation of several different enzymes.

The plant hormones ABA and auxins play a key role during strawberry fruit development and ripening in *F. chiloensis* (Moya-León et al., 2019). Nicely, most of the genes involved in synthesizing both plant hormones are present in the RNAseq library and are differentially expressed. The results showed contigs corresponding to ABA-receptors (PYR family and GCR2 family), phosphatases (PP2C), and protein kinases (SnRK2, CDPKs, and MAPKs), which interact in a regulatory network where several components are differentially expressed. The genes SDR, MCSU, and AAO, which participate in the biosynthesis of ABA, were also found in the RNAseq library. The role of these plant hormones has been studied in commercial strawberries. A microarray chip was used, and several cell wall-related genes presented a high fold change after exogenous application of auxins and ABA (Medina-Puche et al., 2016). The study of the role of different plant hormones was analyzed in *F.* x *ananassa* cv. Seolhyang, and apart from the role of ABA in fruit ripening, it was proposed that *FaMYB10* is key in regulating strawberry ripening (Kim et al., 2019). In the late stage of fruit ripening, ABA plays a role, and in *F. chiloensis, it* also plays a role in softening (Mattus-Araya et al., 2022a) and color (Mattus-Araya et al., 2022b).

Another interesting molecular pathway for this species is the biosynthesis of flavonoids. Flavonoids are associated with the red fruit color, mainly pelargonidin-3-glucoside, but *F. chiloensis* is a white-pinkish fruit, even though it accumulates anthocyanins as cyanidin-3-glucoside (Simirgiotis et al., 2009). Interestingly, most of the genes presented similar expression patterns when the whole fruit was compared to the receptacle sample, but interestingly, a sub-cluster group containing all the genes for synthesizing flavonoids was all "on" at the late time of ripening (RC3 and RC4). Three different transcription factors have been reported to be involved in regulating fruit color (Shaart et al., 2012). *FcMYB* plays a role in the appearance of the red color in the Chilean strawberry during the orthologs in the commercial *F.* x *ananassa* cv. Camarosa (*FaMYB1*) is downregulated (Salvatierra et al., 2013). The transcription factors MYB5 and MYB11 were also identified in the RNAseq and MYB1. The interference in the expression of *FcMYB1* showed the downregulation of LAR and ANR and the upregulation of ANS and UFGT, which allowed the synthesis of proanthocyanidins (Salvatierra et al., 2013). String showed the close role of MYB with genes from the flavonoid molecular pathways during fruit ripening. Even though the two other members of this regulation cluster (bHLH and WD40) were also identified, in most of the samples, their FPKM values were below the cutoff.

Conclusion

The RNAseq library analyzed the massive transcriptional changes during the development and ripening of F. chiloensis fruit. The strategy generated a comprehensive database of expressed sequences during the process for the white Chilean strawberry fruit. Data mining of annotated sequences to G.O. terms classified into "cell wall," "phenylpropanoid-flavonoid," and "ABA-regulation" revealed a biologically meaningful reprogramming at different fruit tissues and ripening stages. Several members of different gene families were identified as differentially expressed, which may play a role in the cell wall's dynamics and fruit color development.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary material.

Author contributions

The present work was conceived and designed by RH, CG-E, and MM-L. SM prepared the RNA samples for sequencing. CG-E, YS, DU, SM, and MM-L contributed to collecting information. YS, CG-E, MM-L, and RH participated in the design of figures. YS prepared most of the tables and figures. MM-L and RH participated actively in the writing and discussion of the manuscript. All authors participated sufficiently in the work to take public responsibility for appropriate portions of the content, read, edited, and approved the final manuscript.

Funding

Thanks to ANID-FONDECYT Regular grants 1210948 (MM-L) and 1201011 (RH), and ANID-ACT210025 for financial support. The funders had no part in the design of the study or collection, analysis, interpretation of data, and in writing the manuscript.

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.976901/full#supplementary-material

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2

Benítez-Burraco, A., Blanco-Portales, R., Redondo-Nevado, J., Bellido, M. L., Moyano, E., Caballero, J. L., et al. (2003). Cloning and characterization of two ripening-related strawberry (*Fragaria x ananassa* cv. Chandler) pectate lyase genes. *J. Exp. Bot.* 54, 633–645. doi: 10.1093/jxb/erg065

Castillejo, C., de la Fuente, J. I., Iannetta, P., Botella, M. A., and Valpuesta, V. (2004). Pectin esterase gene family in strawberry fruit: study of FaPE1, a ripening-specific isoform. *J. Exp. Bot.* 55, 909–918. doi: 10.1093/jxb/erh102

Chang, S., Puryear, J., and Cairney, J. (1993). A simple and efficient method for isolating RNA from pine trees. *Plant Mol. Biol. Rep.* 11, 113–116. doi: 10.1007/BF02670468

Chen, J. X., Mao, L. C., Lu, W. J., Ying, T. J., and Luo, Z. S. (2016). Transcriptome profiling of postharvest strawberry fruit in response to exogenous auxin and abscisic acid. *Planta* 243, 183–197. doi: 10.1007/s00425-015-2402-5

Civello, P. M., Powell, A. L. T., Sabehat, A., and Bennett, A. B. (1999). An expansin gene expressed in ripening strawberry fruit. *Plant Physiol.* 121, 1273–1279. doi: 10.1104/pp.121.4.1273

Figueroa, C. R., Pimentel, P., Gaete-Eastman, C., Moya, M., Herrera, R., Caligari, P. D. S., et al. (2008). Softening rate of the Chilean strawberry (*Fragaria chiloensis*) fruit reflects the expression of polygalacturonase and pectate lyase genes. *Postharvest Biol. Technol.* 49, 210–220. doi: 10.1016/j.postharvbio.2008.01.018

Gaete-Eastman, C., Morales-Quintana, L., Herrera, R., and Moya-León, M. A. (2015). *In-silico* analysis of the structure and binding site features of an α-expansin protein from the mountain papaya fruit (VpEXPA2), through molecular modeling, docking and dynamics simulation studies. *J. Mol. Model.* 21, 1–12. doi: 10.1007/s00894-015-2656-7

García-Gago, J. A., Pose, S., Muñoz-Blanco, J., Quesada, M. A., and Mercado, J. A. (2009). The polygalacturonase FaPG1 gene plays a key role in strawberry fruit softening. *Plant Signal. Behav.* 4, 766–768. doi: 10.4161/psb.4.8.9167

González, G., Fuentes, L., Moya-León, M. A., Sandoval, C., and Herrera, R. (2013). Characterization of two PR genes from *Fragaria chiloensis* in response to *Botrytis cinerea* infection: a comparison with *Fragaria x ananassa*. *Physiol. Mol. Plant P.* 82, 73–80. doi: 10.1016/j.pmpp.2013.02.001

González, G., Moya, M., Sandoval, C., and Herrera, R. (2009). Genetic diversity in Chilean strawberry (*Fragaria chiloensis*): differential response to *Botrytis cinerea* infection. *Span. J. Agric. Res.* 7, 886–895. doi: 10.5424/sjar/2009074-1102

Gu, T., Jia, S., Huang, X., Wang, L., Fu, W., Huo, G., et al. (2019). Transcriptome and hormone analyses provide insights into hormonal regulation in strawberry ripening. *Planta* 250, 145–162. doi: 10.1007/s00425-019-03155-w

Harrison, E. P., Mason, S. J. M., and Manning, K. (2001). Expression of six expansin genes in relation to extension activity in developing strawberry fruit. *J. Exp. Bot.* 52, 1437–1446. doi: 10.1093/jexbot/52.360.1437

Hollender, C. A., Kang, C., Darwish, O., Geretz, A., Matthews, B. F., Slovin, J., et al. (2014). Floral transcriptomes in woodland strawberry uncover developing receptacle and anther gene networks. *Plant Physiol.* 165, 1062–1075. doi: 10.1104/pp.114.237529

Kim, J., Lee, J. G., Hong, Y., and Lee, E. J. (2019). Analysis of eight phytohormone concentrations, expression levels of ABA biosynthesis genes, and ripening related transcription factors during fruit development in strawberry. *J. Plant Physiol.* 239, 52–60. doi: 10.1016/j.jplph.2019.05.013

Letelier, L., Gaete-Eastman, C., Peñailillo, P., Moya-León, M. A., and Herrera, R. (2020). Southern species from the biodiversity hotspot of central Chile: A source of color, aroma, and metabolites for global agriculture and food industry in a scenario of climate change. *Front. Plant Sci.* 11, 1002. doi: 10.3389/fpls.2020.01002

Li, D., Li, L., Luo, Z., Mou, W., Mao, L., and Ying, T. (2015). Comparative transcriptome analysis reveals the influence of abscisic acid on the metabolism of

pigments, ascorbic acid and folic acid during strawberry fruit ripening. *Plos One* 10, e0130037. doi: 10.1371/journal.pone.0130037

Mattus-Araya, E., Guajardo, J., Herrera, R., and Moya-León, M. A. (2022b). ABA speeds up the progress of color in developing F. chiloensis fruit through the activation of PAL, CHS and ANS, key genes of the phenylpropanoid/flavonoid and anthocyanin pathways. *Int. J. Molec. Sci*. 23, 3854. doi: 10.3390/ijms23073854

Mattus-Araya, E., Stappung, Y., Herrera, R., and Moya-León, M. A. (2022a). Molecular actors involved in the softening of *Fragaria chiloensis* fruit accelerated by ABA treatment. *J. Plant Growth Regul*. 41, 1–16. doi: 10.1007/s00344-021-10564-3

Medina-Puche, L., Blanco-Portales, R., Molina-Hidalgo, F. J., Cumplido-Laso, G., Garcia-Caparros, N., Moyano-Cañete, E., et al. (2016). Extensive transcriptomic studies on the roles played by abscisic acid ans auxins in the development and ripening of strawberry. *Funct. Integr. Genomics*. 16, 671–692. doi: 10.1007/s10142-016-0510-3

Méndez-Yañez, A., Beltrán, D., Campano-Romero, C., Molinett, S., Herrera, R., Moya-León, M. A., et al. (2017). Glycosylation is important for FcXTH1 activity as judged by its structural and biochemical characterization. *Plant Physiol. Biochem*. 119, 200–210. doi: 10.1016/j.plaphy.2017.08.030

Méndez-Yañez, A., González, M., Carrasco-Orellana, C., Herrera, R., and Moya-León, M. A. (2020). Isolation of a rhamnogalacturonan lyase expressed during ripening of the Chilean strawberry fruit and its biochemical characterization. *Plant Physiol. Biochem*. 146, 411–419. doi: 10.1016/j.plaphy.2019.11.041

Molina-Hidalgo, F. J., Franco, A., Villatoro, C., Medina-Puche, L., Mercado, J. A., Hidalgo, M. A., et al. (2013). The strawberry (*Fragaria* × *ananassa*) fruit-specific rhamnogalacturonate lyase 1 (FaRGLyase1) gene encodes an enzyme involved in the degradation of cell-wall middle lamellae. *J. Exp. Bot*. 64, 1471–1483. doi: 10.1093/jxb/ers386

Moya-León, M. A., Mattus-Araya, E., and Herrera, R. (2019). Molecular events occurring during softening of strawberry fruit. *Front. Plant Sci*. 10:615. doi: 10.3389/fpls.2019.00615

Moyle, R., Fairbairn, D. J., Ripi, J., Crowe, M., and Botella, J. R. (2005). Developing pineapple fruit has a small transcriptome dominated by metallothionein. *J. Exp. Bot*. 56, 101–112. doi: 10.1093/jxb/eri015

Nam, Y. W., Tichit, L., Leperlier, M., Cuerq, B., Marty, I., and Lelièvre, J. M. (1999). Isolation and characterization of mRNAs differentially expressed during ripening of wild strawberry (*Fragaria vesca* L.) fruits. *Plant Mol. Biol*. 39, 629–636. doi: 10.1023/A:1006179928312

Opazo, M. C., Figueroa, C. R., Henríquez, J., Herrera, R., Bruno, C., Valenzuela, P. D. T., et al. (2010). Characterization of two divergent cDNAs encoding xyloglucan endotransglycosylase/hydrolase (XTH) expressed in *Fragaria chiloensis* fruit. *Plant Sci*. 179, 479–488. doi: 10.1016/j.plantsci.2010.07.018

Opazo, M. C., Lizana, R., Pimentel, P., Herrera, R., and Moya-León, M. A. (2013). Changes in the mRNA abundance of FcXTH1 and FcXTH2 promoted by hormonal treatments of *Fragaria chiloensis* fruit. *Postharvest Biol. Technol*. 77, 28–34. doi: 10.1016/j.postharvbio.2012.11.007

Opazo, M. C., Lizana, R., Stappung, Y., Davis, T. M., Herrera, R., and Moya-León, M. A. (2017). XTHs from *Fragaria vesca*: genomic structure and transcriptomic analysis in ripening fruit and other tissues. *BMC Gen*. 18, 852. doi: 10.1186/s12864-017-4255-8

Perkins-Veazie, P. (1995). Growth and ripening of strawberry fruit. *Hortic. Rev*. 17, 267–297. doi: 10.1002/9780470650585.ch8

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT PCR. *Nucl. Acids Res*. 29:e45. doi: 10.1093/nar/29.9.e45

Pimentel, P., Salvatierra, A., Moya-León, M. A., and Herrera, R. (2010). Isolation of genes differentially expressed during development and ripening of *Fragaria chiloensis* fruit by suppression subtractive hybridization. *J. Plant Physiol*. 167, 1179–1187. doi: 10.1016/j.jplph.2010.03.006

Quesada, M. A., Blanco-Portales, R., Pose, S., García-Cago, J. A., Jiménez-Bermúdez, S., Muñoz-Serrano, A., et al. (2009). Antisense down-regulation of the FaPG1 gene reveals an unexpected central role for polygalacturonase in strawberry fruit softening. *Plant Physiol*. 150, 1022–1032. doi: 10.1104/pp.109.138297

Retamales, J. B., Caligari, P., Carrasco, B., and Saud, G. (2005). The current Chilean strawberry status and the research need to convert the species into a commercial crop. *HortScience* 40, 1633–1634. doi: 10.21273/HORTSCI.40.6.1633

Rosli, H. G., Civello, P. M., and Martinez, G. A. (2004). Changes in cell wall composition of three *Fragaria x ananassa* cultivars with different softening rate during ripening. *Plant Physiol. Biochem*. 42, 823–831. doi: 10.1016/j.plaphy.2004.10.002

Salvatierra, A., Pimentel, P., Moya-León, M. A., Caligari, P. D. S., and Herrera, R. (2010). Comparison of transcriptional profiles of flavonoid genes and anthocyanin contents during fruit development of two botanical forms of *Fragaria chiloensis* ssp. chiloensis. *Phytochemistry* 71, 1839–1847. doi: 10.1016/j.phytochem.2010.08.005

Salvatierra, A., Pimentel, P., Moya-León, M. A., and Herrera, R. (2013). Increased accumulation of anthocyanins in *Fragaria chiloensis* fruits by transient suppression of FcMYB1 gene. *Phytochemistry* 90, 25–36. doi: 10.1016/j.phytochem.2013.02.016

Sanchez-Sevilla, J., Vallarino, J., Osorio, S., Bombarely, A., Pose, D., Merchant, C., et al. (2017). Gene expression atlas of fruit ripening and transcriptome assembly from RNAseq data in octoploid strawberry (*Fragaria x ananassa*). *Sci. Rep*. 7, 13737. doi: 10.1038/s41598-017-14239-6

Santiago-Doménech, N., Jimenez-Bermudez, S., Matas, A. J., Rose, J. K. C., Muñoz-Blanco, J., Mercado, J. A., et al. (2008). Antisense inhibition of a pectate lyase gene supports a role for pectin depolymerization in strawberry fruit softening. *J. Exp. Bot*. 59, 2769–2779. doi: 10.1093/jxb/ern142

Shaart, J., Dubos, C., Romero, I., van Houwelingen, A., de Vos, R., Jonker, H., et al. (2012). Identification and characterization of MYB-bHLH-WD40 regulatory complexes controlling proanthocyanidin biosynthesis in strawberry (*Fragaria* x *ananassa*) fruits. *New Phytol*. 197, 454–467. doi: 10.1111/nph.12017

Simirgiotis, M. J., Theoduloz, C., Caligari, P., and Schmeda-Hirschmann, G. (2009). Comparison of phenolic composition and antioxidant properties of two native Chilean and one domestic strawberry genotype. *Food Chem*. 113, 377–385. doi: 10.1016/j.foodchem.2008.07.043

Sun, X., and Li, J. (2013). Pairheatmap: comparing expression profiles of gene groups in heatmaps. *Comput. Methods Programs Biomed*. 112, 599–606. doi: 10.1016/j.cmpb.2013.07.010

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res*. 47, D607–D613. doi: 10.1093/nar/gky1131

Szklarczyk, D., Morris, J. H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., et al. (2017). The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res*. 45, D362–D368. doi: 10.1093/nar/gkw937

Villarreal, N. M., Rosli, H. G., Martínez, G. A., and Civello, P. M. (2008). Polygalacturonase activity and expression of related genes during ripening of strawberry cultivars with contrasting fruit firmness. *Postharvest Biol. Technol*. 47, 141–150. doi: 10.1016/j.postharvbio.2007.06.011

Wang, Q. H., Zhao, C., Zhang, M., Li, Y. Z., Shen, Y. Y., and Guo, J. X. (2017). Transcriptome analysis around the onset of strawberry fruit ripening uncovers an important role of oxidative phosphorylation in ripening. *Sci. Rep*. 7, 41477. doi: 10.1038/srep41477

Zhao, F., Song, P., Zhang, X., Li, G., Hu, P., Aslam, A., et al. (2021). Identification of candidate genes influencing anthocyanin biosynthesis during the development and ripening of red and white strawberry fruits via comparative transcriptome analysis. *Peer J*. 9, e10739. doi: 10.7717/peerj.10739

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 23 September 2022
DOI 10.3389/fpls.2022.952698

OPEN ACCESS

EDITED BY
Maria Carmen Gomez-Jimenez,
University of Extremadura,
Spain

REVIEWED BY
Iris F. Kappers,
Wageningen University and Research,
Netherlands
Aichun Zhao,
Southwest University,
China

*CORRESPONDENCE
Qinghua Pan
qinghua_pan@sina.com
Yuping Zhang
zhyptt@163.com

SPECIALTY SECTION
This article was submitted to Plant Development and EvoDevo, a section of the journal Frontiers in Plant Science

RECEIVED 25 May 2022
ACCEPTED 22 August 2022
PUBLISHED 23 September 2022

CITATION
Lu D, Zhang L, Wu Y, Pan Q, Zhang Y and Liu P (2022) An integrated metabolome and transcriptome approach reveals the fruit flavor and regulatory network during jujube fruit development.
Front. Plant Sci. 13:952698.
doi: 10.3389/fpls.2022.952698

An integrated metabolome and transcriptome approach reveals the fruit flavor and regulatory network during jujube fruit development

Dongye Lu[1], Lei Zhang[2], Yang Wu[1], Qinghua Pan[1]*, Yuping Zhang[1]* and Ping Liu[3]

[1]Beijing Academy of Agriculture and Forestry Sciences, Institute of Forestry and Pomology, Beijing, China, [2]State Key Laboratory of Tree Genetics and Breeding, Key Laboratory of Tree Breeding and Cultivation of National Forestry and Grassland Administration, Research Institute of Forestry, Chinese Academy of Forestry, Beijing, China, [3]Research Center of Chinese Jujube, Hebei Agricultural University, Baoding, China

The fruit flavor is a key economic value attribute of jujube. Here we compared metabolomes and transcriptomes of "Mazao" (ST) and "Ping'anhuluzao" (HK) with unique flavors during fruit development. We identified 437 differential metabolites, mainly sugars, acids, and lipids. Fructose, glucose, mannose and citric acid, and malic acid are the determinants of sugar and acid taste of jujube fruit. Based on the transcriptome, 16,245 differentially expressed genes (DEGs) were identified, which were involved in "glucosyltransferase activity," "lipid binding," and "anion transmembrane transporter activity" processes. Both transcriptome and metabolome showed that developmental stages 2 and 3 were important transition periods for jujube maturation. Based on WGCNA and gene-metabolite correlation analysis, modules, and transcription factors (*ZjHAP3*, *ZjTCP14*, and *ZjMYB78*) highly related to sugar and acid were identified. Our results provide new insights into the mechanism of sugar and acid accumulation in jujube fruit and provide clues for the development of jujube with a unique flavor.

KEYWORDS

***Ziziphus jujuba*, metabolome, transcriptome, fruit quality, regulatory network**

Introduction

Ziziphus jujuba belongs to *Ziziphus* of Rhamnaceae, which is native to China. It is rich in germplasm resources and has a long history of cultivation. There are many cultivars of jujube, which can be traditionally divided into three categories according to their uses: dried, fresh, dried, and fresh jujube. Among them, the production of dried fruits is the largest in China, while fresh jujube is abundant in nutrition and tastes crisply (Liu and Wang, 2009). Studies have shown that jujube is rich in carbohydrates, cyclic adenosine

monophosphate, triterpenoids, flavonoids, vitamin compounds, and inorganic salts such as phosphorus, calcium, and iron, which have high nutritional and medicinal value (Li et al., 2007; Gao et al., 2013). Understanding the differences and dynamic changes in nutritional components of jujube fruits during ripening will provide valuable information for the genetic improvement of jujube.

Flavor quality is an important economic attribute of fruits that affects people's choices (Barrett et al., 2010; Goldenberg et al., 2018). The acid, sugar composition and content of fruits determine important factors of fruit flavor (Zhu et al., 2018). In the process of jujube domestication, the sweetness/acidity of jujube fruit is based on the genetic selection that determines the content of acid and sugar (Huang et al., 2016). The dynamic analysis of sugar components in jujube fruits showed that fructose and glucose were the main accumulations in the early stages of fruit accumulation, while sucrose was dominant in the later stages (Zhang et al., 2021). Zhao et al. (2021) revealed the content characteristics of organic acid components in the fruits of 219 jujube germplasm and found that the contents of malic, quinic, and citric acids in jujube fruits were in the top three. Glucose metabolism produces pyruvate through glycolysis, which enters the tricarboxylic acid (TCA) cycle to form citric acid, malic acid, and others. The sugar content of jujube fruits is significantly higher than that of wild jujube and other fruit trees, such as apples, peaches, and grapes (Huang et al., 2021). Compared with Rosales fruit, the gene families involved in glucose metabolism in the jujube genome have a higher degree of expansion (Liu et al., 2014). The sugar content in fruits largely depends on the balance between the sugar source and the sink (Huang et al., 2021). Therefore, it is of great significance to reveal the metabolite contents of sugars and organic acids in jujube during different development processes as well as the biosynthetic pathways and regulatory mechanisms affecting their accumulation.

With the continuous development of omics technology, metabolome and transcriptome analysis have been successfully applied to study the regulatory mechanisms of leaf color, fruit anthocyanin, flavonoids, and other nutrients accumulation in jujube and apples (Shi et al., 2020; Xu et al., 2020; Li et al., 2021). In addition, the betaine biosynthetic pathway determines the pitaya fruit color formation including peel color (red and yellow) and the pulp color (Zhou et al., 2020). Gong et al. (2021b) revealed the differences in sugar accumulation between cultivated and wild watermelon through transcriptomics and metabolomics and found that UDP-glycosyltransferase was closely associated with glycosylation of cucurbitacin. By combining the results of WGCNA and metabolomics, Chu et al. (2022) identified genes and metabolites for flesh sweetness, bitterness, and color of watermelon. Xiong et al. (2020) analyzed the accumulation patterns of sugars, organic acids, ascorbic acid, and related genes throughout the development of yellow kiwifruit. Yang revealed the expression patterns of sugar, acid, flavonoid metabolites and genes during cherry ripening (Yang et al., 2021). Although we have studied fruit quality at the level of transcription and metabolism, jujube flavor, as a complex trait, still varies significantly among cultivars, so we need to explore its molecular mechanism.

In this study, we sequenced the metabolome and transcriptome datasets of "Mazao" (ST) and "Ping'anhuluzao" (HK) jujube cultivars at 30, 60, 80, 100, and 110 days after anthesis. These two cultivars have good flavor and rich nutrients. The dynamic accumulation patterns of sugars, organic acids, fatty acids and other nutrients at five developmental stages were analyzed for their primary metabolome, and gene expression patterns were analyzed by transcriptomics, to explore possible regulatory genes affecting jujube flavor by joint analysis. This study provided a rich genetic basis for further enriching the flavor of jujube fruits.

Materials and methods

Plant materials

"Mazao" (ST) and "Ping'anhuluzao" (HK) were excellent new cultivars selected in recent years for live breeding (Lu et al., 2022). Among them, ST jujube is flat and round while HK possesses constricted type (Figure 1), HK accumulates high total soluble sugar contents (22.68%) and low total organic acid contents (0.76 g/kg) at maturity, while ST was the opposite of its, with total soluble sugar contents of 14.71% and total organic acid contents of 1.04 g/kg. The trees were cultivated under normal field conditions, including irrigation, fertilization, and disease and pest control. The fruits of HK and ST were collected from the town of Qinglonghu (116°5′E, 39°47′N), Fangshan District, Beijing, China in 2021 at five different periods of 30 (young), 60 (enlarged), 80 (white-ripened), 100 (half-red), and 110 (full-red) days after anthesis. Fruits were pitted and chopped, then rapidly placed in liquid nitrogen and stored at −80°C until used for metabolomic analysis and transcriptomic sequencing. Three biological replicates were taken from each period of the two cultivars.

UPLC–MS/MS system-based widely targeted metabolomics analysis

The primary metabolites were extracted and identified by Metware Biotechnology Co., Ltd.[1] Biological samples were freeze-dried using a vacuum freeze-dryer (Scientz-100F), and 100 mg of the powder was dissolved in 1.2 ml of 70% methanol solution and kept at 4°C overnight. The filtered extracts were used for metabolite profiling by UPLC–MS/MS system (Applied Biosystems 4500 Q TRAP) analysis and quantification was performed by multiple reaction monitoring (MRM) in a triple quadrupole spectrometer (Chen et al., 2013). Metabolites were identified by comparing the exact mass, fragmentation patterns,

1 www.metware.cn

FIGURE 1
Phenotype of five developmental stages in "Mazao" **(A)** and "Ping'anhuluzao" **(B)**.

and retention times with the standards from a self-compiled database (MetWare, Wuhan, China) (Chen et al., 2013).

RNA extracted and RNA-sequencing

Total RNA was extracted from fruits (HK1, HK2, HK3, HK4, HK5, ST1, ST2, ST3, ST4, and ST5 with three biological replicates) using the RNAprep Pure Plant Plus Kit (TIANGEN, Beijing, China). A total amount of 1 μg RNA per sample was used for the sequenced library by NEBNext® UltraTM RNA Library Prep Kit for Illumina® (NEB, United States). The cDNA library products were sequenced by the Illumina Hiseq platform with 125 bp/150 bp paired-end reads. The raw data was filtered using fastp v 0.19.3 (Chen et al., 2018), mainly removing reads with adapters; when any sequencing read contained more than 10% of the bases of the read, the paired reads were removed; when any sequencing read contained more than 50% of the bases of the read with low quality ($Q \leq 20$), the paired reads were removed. Clean reads were compared to a reference genome (*Ziziphus jujuba* Mill. "Dongzao") using HISAT v2.1.0 (Liu et al., 2014). Novel gene prediction was performed using StringTie v1.3.4d (Pertea et al., 2015). Feature Counts v1.6.2 (Liao et al., 2014) was used to calculate the gene alignments and FPKM.

Differential metabolites and genes analysis

Unsupervised principal component analysis (PCA) was performed by statistics function "prcomp" within R v4.1.2. variable importance in projection (VIP) values were extracted from OPLS-DA result by using the R package "MetaboAnalystR" (Chong and Xia, 2018). Metabolites with VIP ≥ 1 and $\log_2$(fold change) ≥1 were considered significantly differential accumulation metabolites (DAMs) between groups (HK2 vs. HK1, HK3 vs. HK2, HK4 vs. HK3, HK5 vs. HK4, ST2 vs. ST1, ST3 vs. ST2, ST4 vs. ST3, and ST5 vs. ST4). To analyze the changing trend of metabolites, DAMs were standardized (*z*-score) and clustered by K-means.

DESeq2 v1.22.1 (Love et al., 2014) was used to analyze the differential expression genes (DEGs) with $|\log_2$FC(fold change)$| \geq 1$ and *p*-value <0.05 (Varet et al., 2016). The functions of the unigenes were annotated by the NR, KOG, SwissProt, GO, and KEGG databases (Ashburner et al., 2000; Bairoch and Apweiler, 2000; Kanehisa and Goto, 2000; Natale et al., 2000; Ogata et al., 2000; Wilke et al., 2012).

Combined metabolome and transcriptome analysis

The quantitative values of genes and metabolites in all samples were used for correlation analysis. The "cor" function in R was used to calculate the Pearson correlation coefficient of genes and metabolites with an absolute threshold larger than 0.85 and a *p*-value <0.05. The correlation analysis results of different genes and metabolites were selected. Differential genes and differential metabolites in each pathway were analyzed by CCA (canonical correlation analysis) (González et al., 2008). WGCNA v1.69 was used for weighted gene co-expression network analysis (WGCNA). Before WGCNA analysis, the genes with FPKM <0.1 were filtered out from all samples. Pearson's correlation, calculation of soft-power threshold (β), and the division of modules were performed according to previous studies (Chen et al., 2021; Lu et al., 2022), in this study, soft-power threshold (β)

was set to 7, the minimum number of genes contained in the modules was set to 50, while the threshold for merging similar modules was set to 0.25. Cytoscape 3.8 was used for visualization of the control network with default settings (Otasek et al., 2019).

qRT-PCR

Ten DEGs were selected for qRT-PCR analysis, and *ZjUBQ* was used as the internal reference gene. The primers were listed in Supplementary Table S2. The RNA was extracted from jujube fruit as described above. qRT-PCR was performed using TB Green® Premix Ex Taq™ II (Takara, Beijing, China). Three technical replicates and three biological replicates were performed. The relative expression levels were calculated using the $2^{-\Delta\Delta Ct}$method (Livak and Schmittgen, 2001).

Results

Overview of metabolite accumulation patterns during jujube fruit development

To define a comprehensive landscape of metabolite profile during fruit development of HK and ST, we performed metabolite profiling by using LC-MS. During fruit ripening, the pericarp changed from green to yellow and red, and flavonoids accumulated rapidly. A total of 508 metabolites of 10 categories were obtained at different development stages of jujube fruits, including organic acids, amino acids and derivatives, saccharides and alcohols, free fatty acids, nucleotides and derivatives, lysophosphatidyl cholines (LPCs), lysophosphatidyl ethanolamines (LPEs), vitamins, glycerol esters, and sphingolipids (Supplementary Table S1).

Principal component analysis was used to analyze the data for all compounds from five developmental stages for two cultivars with three biological replicates; the objective was to provide a preliminary understanding of the overall metabolic differences between groups of samples including different fruit development stages and cultivars and the magnitude of variability between samples within groups. PC1 and PC2 explained 33.68% and 25.94% of the variation, respectively (Figure 2A). The results showed that the variation between different fruit development stages was greater than the variation between the two cultivars. In addition, there was a large gap between the metabolomes of the third stage and the other four stages of development.

Analysis of metabolite differences and K-means analysis during fruit ripening in jujube

Heatmap and cluster analysis yielded an overview of dynamic metabolome changes during fruit development. To further explore the metabolic differences in the developmental stages and between cultivars, we conducted a different analysis. A total of 437 differential accumulation metabolites (DAMs) were identified (Figures 2B,C), with 308 DAMs in HK, 289 DAMs in ST, and 392 DAMs between HK and ST. There were more differential metabolites in HK3 vs. HK2 and ST3 vs. ST2, which was consistent with the results of PCA. In other words, the shift from stage 2 to 3 was an important transition in jujube fruit development. There were 33 common differential metabolites between the two cultivars at different periods, including 18 organic acids, eight amino acids and derivatives, two nucleotides and derivatives, two LPCs, one glycerol ester, one saccharide and alcohol, and one vitamin (Figure 2D).

To analyze the trends in metabolite content throughout fruit development, the relative contents of all the different metabolites identified in all group comparisons were standardized according to the screening criteria and then subjected to K-means cluster analysis (Figure 3A). Class1 contained 78 DAMs (organic acids, saccharides, alcohols, etc.) that accumulated mainly during the early stage (1, 2), and were reduced during the later stages (3, 4, 5) of fruit development (Figures 3B,C). The DAMs of Class 2 were mainly concentrated in stages 4 and 5 of HK. It contained a large number of organic acids and almost all lipids (LPC and LPE), which were specifically high in HK (Figures 3B,D) 0.77 metabolites, such as amino acids and derivatives and organic acids in Class 5, accumulated in large amounts in ST fruit at later stages (3, 4, 5), which is opposite to the metabolite accumulation mode in Class 2 (Figures 3B,E). This showed that the accumulation patterns of organic acids and saccharides were similar in both cultivars, but there were significant differences in lipids and amino acids, which may lead to the different nutritional value and taste of jujube.

Accumulation pattern and correlation analysis of sugars and organic acids in jujube fruits

Sugar and acids are important factors affecting fruit flavor. We found 50 saccharides and alcohol metabolites with differential accumulation in either developmental stage or cultivar, among which fructose, glucose, mannose, and galactose were the main soluble sugars of jujube. The contents of four sugars showed similar trends in ST and HK, with higher contents in stages 1 and 2 and a decreasing trend in the later stages (Figure 4A). The results indicated that sugar accumulation, which determines fruit sweetness, mainly occurred in the early stages of fruit development. Furthermore, 151 kinds of organic acids were found, including the common soluble acids citric acid, malic acid and quinic acid. Citric acid and malic acids were found to be the most abundant. Their accumulation pattern was opposite to the trend of sugar content, and the content increased with fruit ripening. In addition, the organic acid content of ST was higher than that of HK (Figure 4B). To further explain the relationship between sugar and organic acids, the correlation between organic

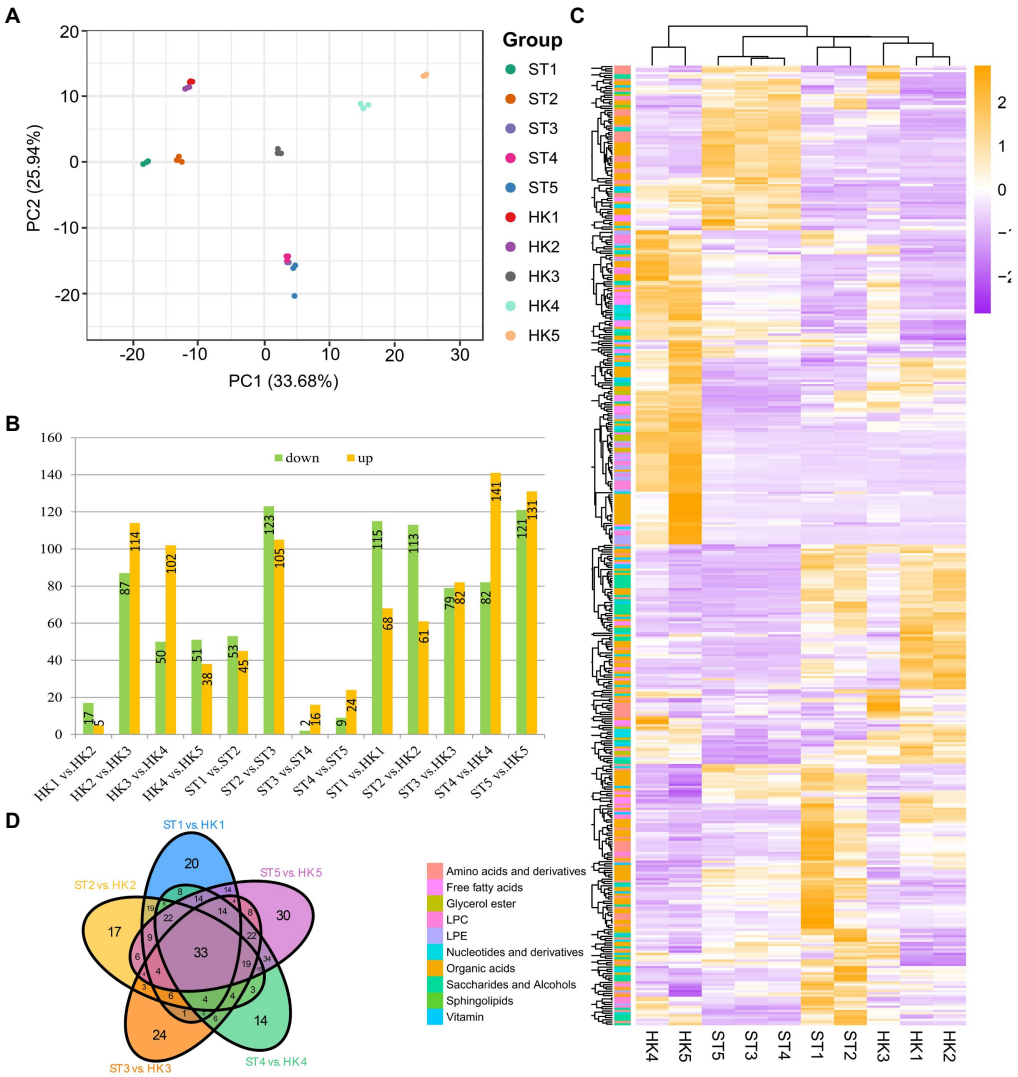


FIGURE 2
Comparison of metabolites in different developmental stages of jujube fruit. **(A)** Principal component analysis (PCA) score plot of all metabolites in 30 samples. HK1 (ST1), HK2 (ST2), HK3 (ST3), HK4 (ST4), and HK5 (ST5) represent the samples at 30, 60, 80, 100, and 110days after anthesis, respectively. **(B)** The number of differentially accumulation metabolites (DAMs) by comparing HK1 vs. HK2, HK2 vs. HK3, HK3 vs. HK4, HK4 vs. HK5, ST1 vs. ST2, ST2 vs. ST3, ST3 vs. ST4, ST4 vs. ST5, ST1 vs. HK1, ST2 vs. HK2, ST3 vs. HK3, ST4 vs. HK4, ST5 vs. HK5. **(C)** Overview of DAMs of two cultivars in five periods. **(D)** Venn diagram of the number of different developmental stages and cultivars.

acids and sugar content was analyzed. The results showed that there was a positive correlation among the four sugars except for maltose. Citric acid, malic acid, and quinic acid were negatively correlated with sugars, while succinic acid was positively correlated with sugars (Figure 4C).

Transcriptome analysis of the jujube fruit of two cultivars at different development stages

To further explore the possible regulatory genes affecting DAMs, we also sequenced the transcriptomes of jujube fruits at each stage. After removing the unknown reads, low-quality reads, and adaptor sequences, a total of 206.95 Gb clean data were obtained from 30 libraries with an average GC content of 43.88% (Supplementary Table S3). A total of 16,245 differentially expressed genes (DEGs) were identified from 13 differential comparisons (between different developmental stages and different cultivars). The number of DEGs between the comparison combinations ranged from 304 to 9,342, with ST3 vs. ST2 reaching a maximum of 9,342 DEGs (Supplementary Figure S1).

To reveal the molecular functions of DEGs, GO enrichment analysis indicated that they were more widely distributed in three categories of biological processes, molecular functions, and cell components. Multiple comparative combinations were enriched into the categories of "glucosyltransferase activity," "lipid binding," "anion transmembrane transporter activity," "anion transport,"

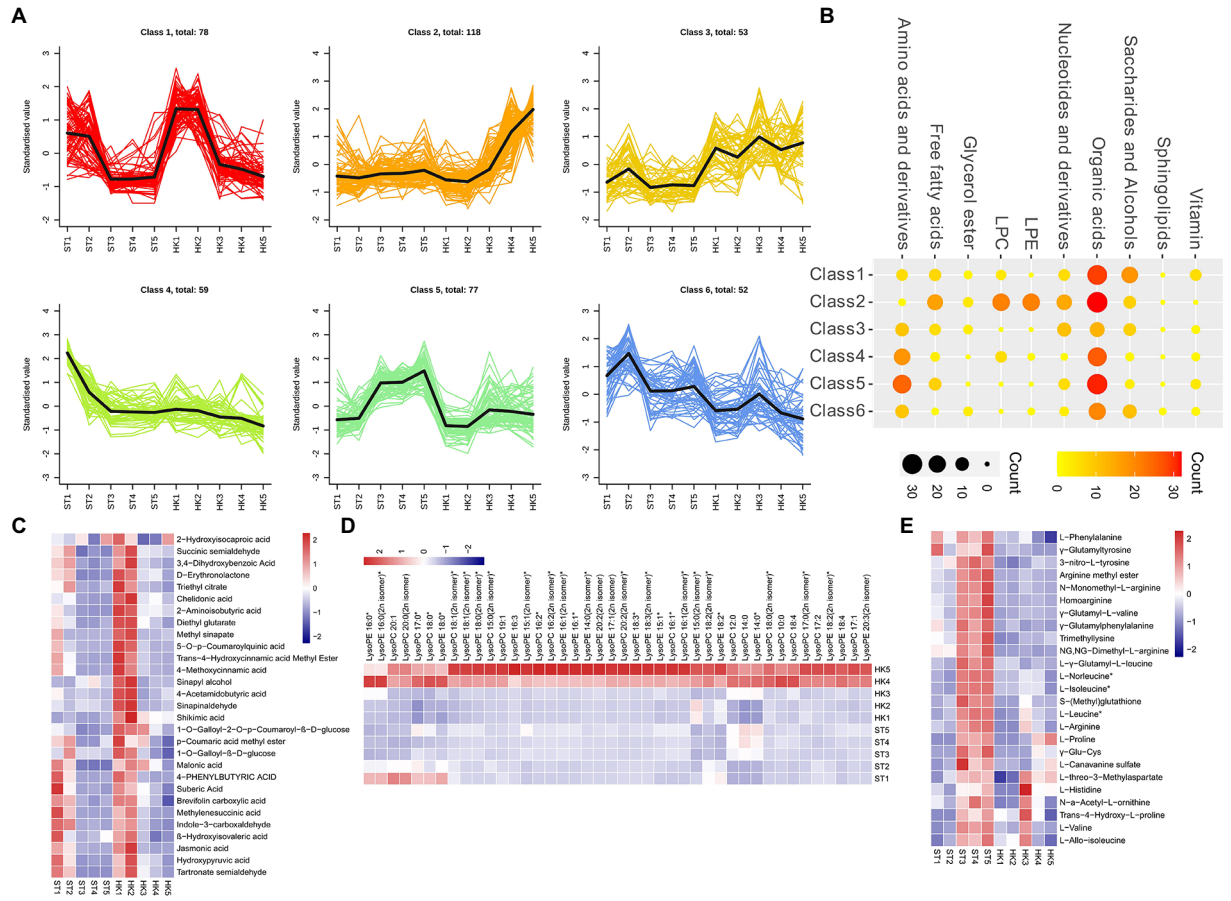


FIGURE 3
Cluster analysis of differential metabolites. **(A)** K-means cluster analysis of differential metabolites. **(B)** The number of metabolites in each category. Nodes from small to large and light to dark represent the number of metabolites. The accumulation pattern metabolites of class 1 **(C)**, 2 **(D)**, and 5 **(E)**.

"photosynthesis," "thylakoid" membrane" and "photosystem." ST3 vs. ST2 contained the most DEGs, which were significantly enriched in "phosphatase activity," "metal cluster binding," "ribonucleoside binding." DEGs of HK3 vs. HK2 were enriched in multiple cell structure related categories, including "supramolecular polymer," "polymeric cytoskeletal fiber," and "microtubule" (Supplementary Figure S2). KEGG enrichment indicated that DEGs were involved in starch and sucrose metabolism as well as secondary metabolites in periods 2 and 3. Among them, HK was uniquely enriched in fatty acid biosynthesis, metabolism, and degradation (Supplementary Figures S3, S4).

Identification of WGCNA modules associated with fruit quality

To reveal potential relationships between genes and fruit quality, we performed the WGCNA on DEGs. The differential genes were divided into 13 modules (Figure 5A). The modules were related to sugars and acids in fruit. The turquoise module was positively correlated with the four main sugars and succinic acids (0.41–0.93) but negatively correlated with the levels of citric, malic, and quinic acid (−0.75 to −0.69). In addition, the brown and red modules were negatively correlated with sugar content and positively correlated with organic acid content. This further suggests that there is a negative correlation between sugars and acids (Figure 5A). Further analysis showed that *ZjSKU5* (monocopper oxidase-like protein SKU5, LOC107403720), *ZjYABBY1* (C2C2-YABBY, LOC107403723), *ZjTOPP4* (serine/threonine-protein phosphatase PP1-like, LOC107412332), and *ZjMYB78* (LOC107426114) are the core genes of the module turquoise and red (Figures 5B,C).

Differentially expressed genes involved in sugar and organic acid metabolism

The analysis of genes related to sugar biosynthesis and transport and organic acid metabolism is of great significance to analyze and understand the accumulation of sugars and organic acids. During

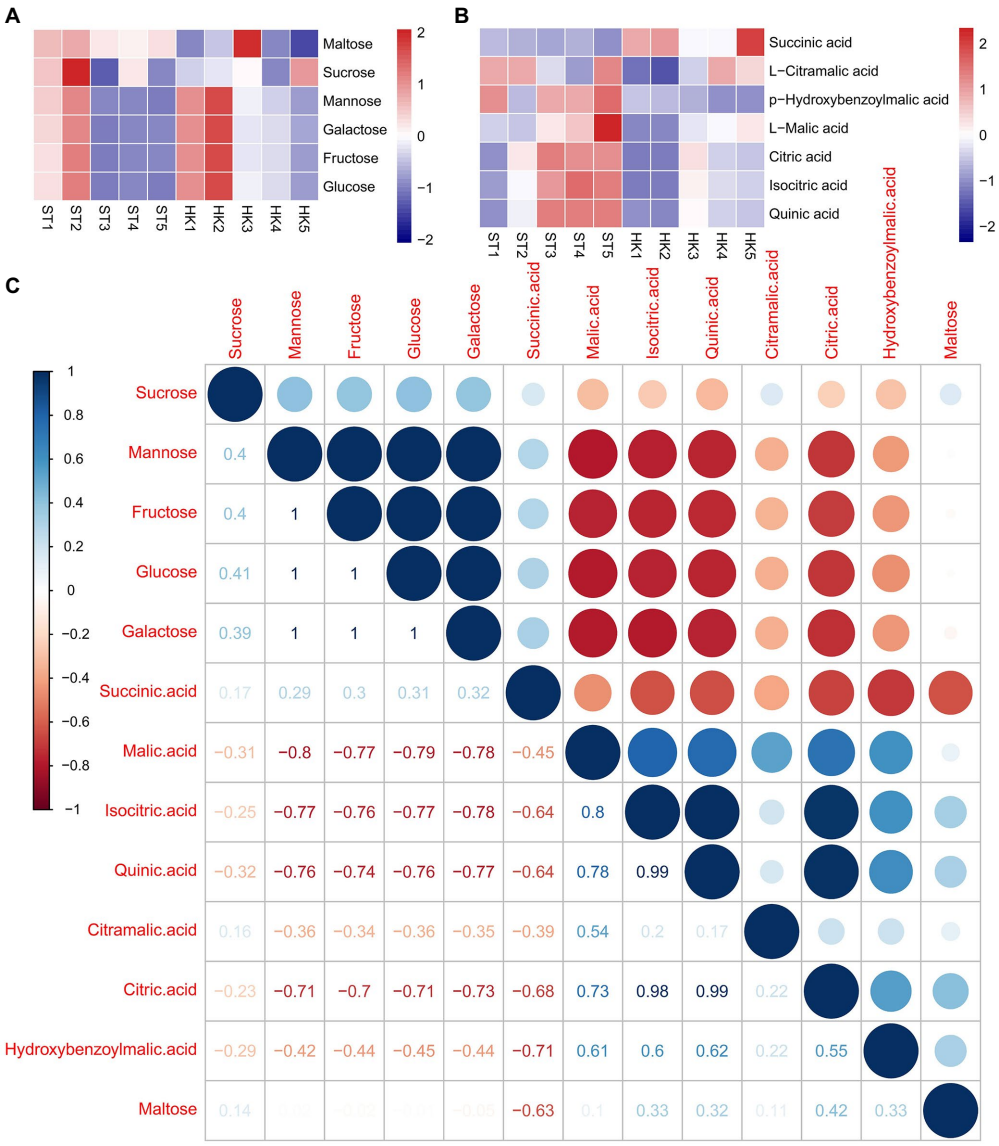


FIGURE 4
Patterns of sugar and acid accumulation and their correlation in jujube. Accumulation patterns of six major sugars **(A)** and seven major acids **(B)**. **(C)** Correlation between sugars and acids.

the 2/3 stage of jujube fruit development, more DEGs and DAMs were involved in the pathway of starch and sucrose metabolism and carbon metabolism (Supplementary Figure S4). In this study, eight fructokinases (*ZjFK*), nine sucrose synthases (*ZjSUSY*), 26 glucosidase-like (*ZjGLU*, four alpha-, 22 beta-), two hexokinases (*ZjHK*), six sucrose-phosphatases (*ZjSPS*), and six alactinol–sucrose galactosyltransferases (*ZjSIP*) were identified. Moreover, 45 sugar transporter genes were found, including three sucrose transport proteins (*ZjSUC*), three sugar transporters (*ZjSTP*), 10 ERD6-like sugar transporters (*ZjERD6*-like), eight SWEET sugar transporters (*ZjSWEET*), seven polyol transporters (*ZjPLT*), three inositol transporters (*ZjINT*), five phosphatidylinositol transfer proteins, four plastidic glucose transporters (*ZjpGlcT*), and two UDP-glucose transporters (Figures 6A,B).

Citric and malic acids are the main organic acids of jujube fruits, and they are also vital intermediates in the tricarboxylic acid cycle (TAC) downstream of glycolysis. In our analysis, 35 DEGs of the TAC pathway were found, including six aconitate hydratases (*ZjACO*), two ATP-citrate synthase alpha chain proteins (*ZjACLA*), one citrate synthase (*ZjCSY*), six dihydrolipoyl dehydrogenases (*ZjLPD*), two isocitrate dehydrogenase [NAD] catalytics (*ZjIDH*), three isocitrate dehydrogenase [NADP] (*ZjCICDH*), seven malate dehydrogenase (*ZjMDH*), three phosphoenolpyruvate carboxykinase (*ZjPCK*), and five pyruvate dehydrogenase E1 component subunit alpha (*ZjPDH*) (Figure 6C).

Transcriptional regulation is an important cause of gene expression and regulation of metabolite content. We jointly analyzed and screened transcription factors and structural genes

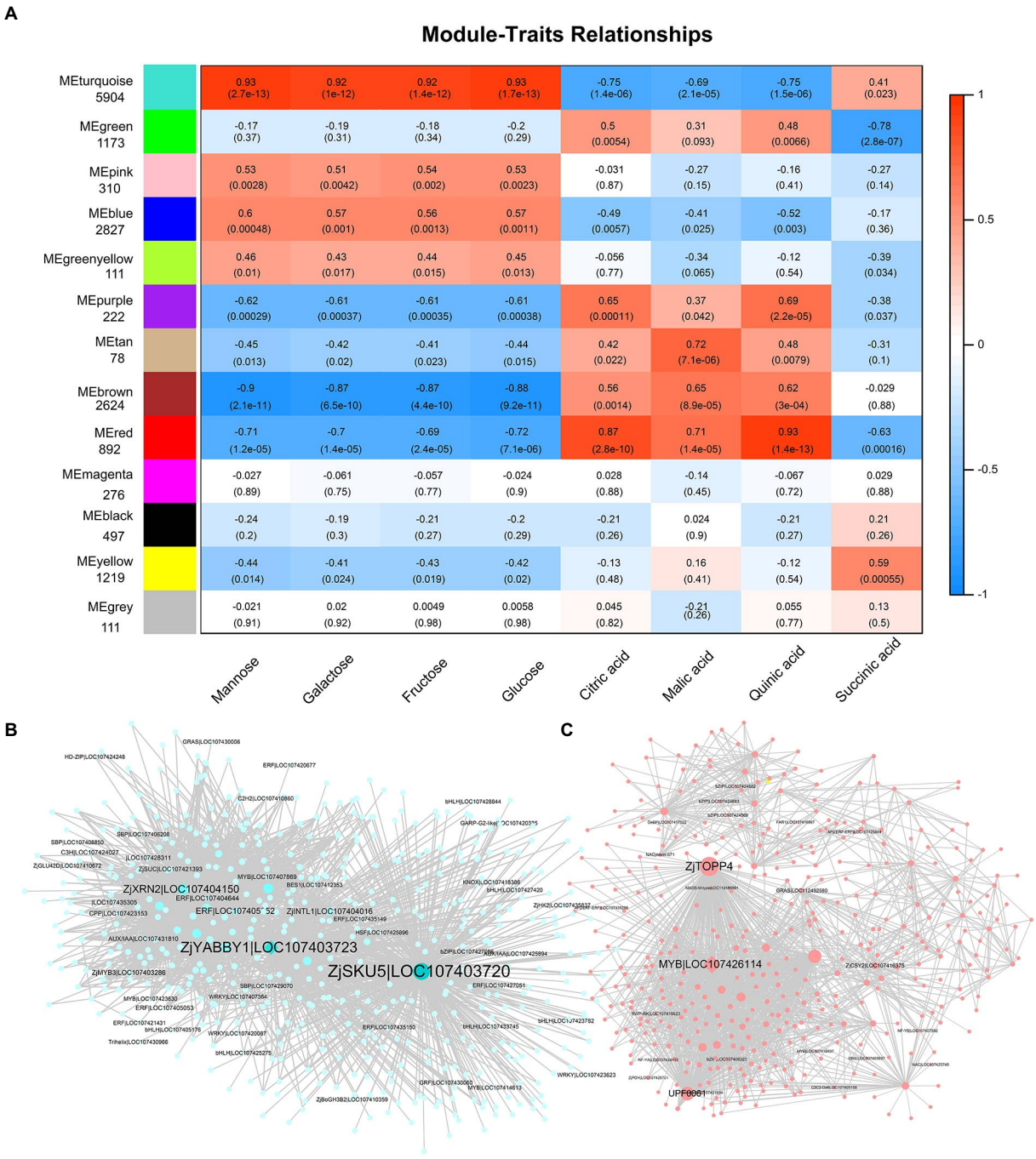


FIGURE 5
Weighted gene co-expression network analysis (WGCNA) of DEGs. **(A)** The relationship between modules and fruit quality. Weighted module-trait correlations and corresponding *p* values. The color scale on the right shows module-trait correlations from –1 (blue) to 1 (red). Cytoscape representation of co-expressed genes with edge weight≥0.50 in the "turquoise" module **(B)** and≥0.30 in the "red" module **(C)**.

related to sugars and organic acids. We found 3,118 genes associated with mannose, glucose, galactose, fructose, quinic acid, citric acid, and malic acid, including 202 transcription factors in 57 gene families distributed in modules turquoise, brown and red (Supplementary Table S4) ($|\text{coefficient}| \geq 0.85$, $p < 1.38 \times 10^{-12}$). Therefore, a co-expression network of sugar and acid metabolites with transcription factors and metabolic pathway structural genes was further constructed (Figure 6D). The results showed that candidate genes such as *ZjHAP3* (HEME ACTIVATOR PROTEIN (YEAST) HOMOLOG 3, LOC107409505), *ZjTCP14* (TEOSINTE BRANCHED, CYCLOIDEA AND PCF14, LOC107428978), and *ZjAGL61* (AGAMOUS-LIKE 61, LOC112488991) (At2g24840) may be involved in the accumulation of major sugars and acids in jujube fruits.

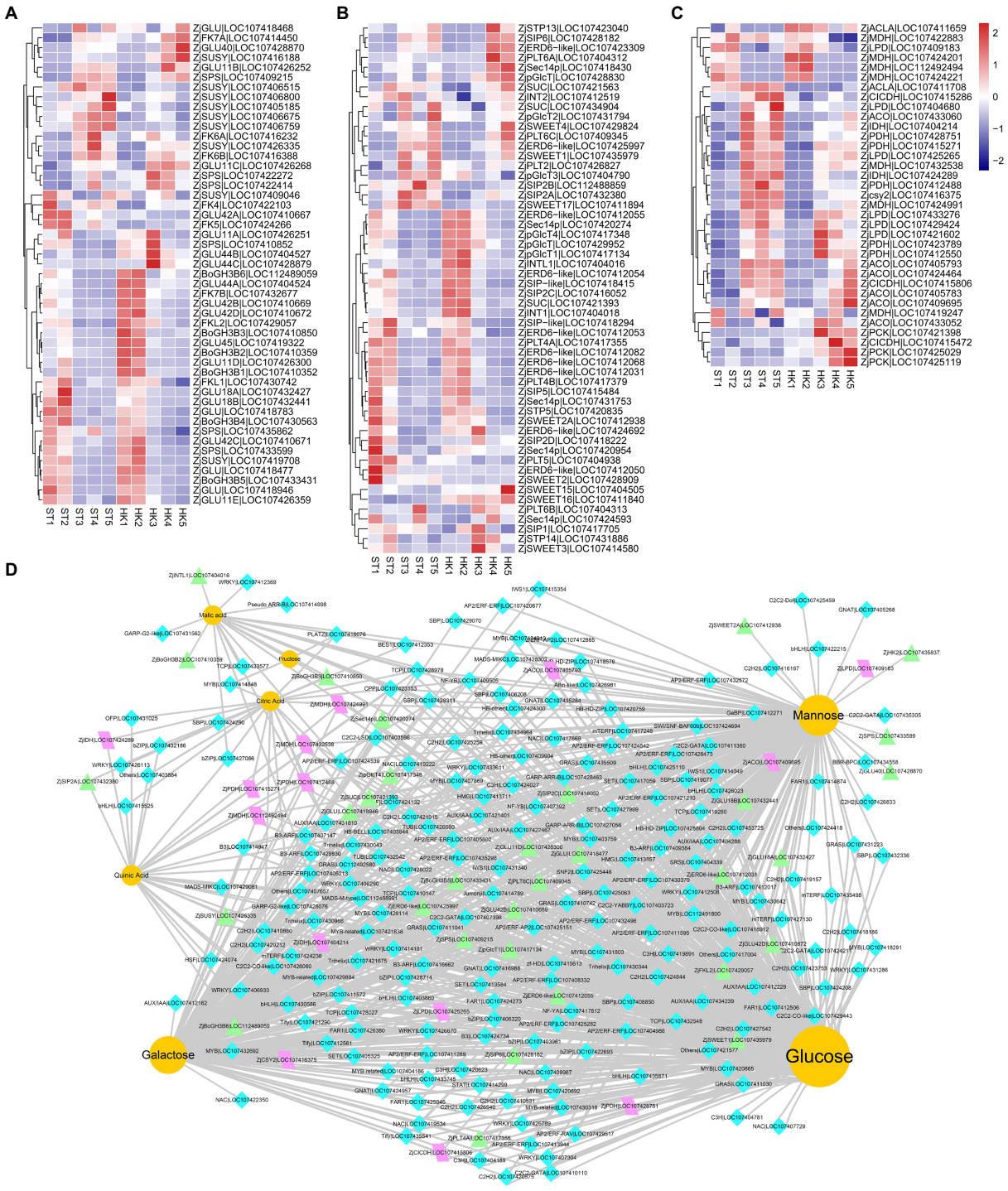


FIGURE 6
The expression profile of genes involved in organic acid and sugar biosynthetic pathways. Expression patterns of sugar-related kinases **(A)**, sugar transporters **(B)**, and tricarboxylic acid biosynthesis structural genes **(C)** during jujube ripening. The expression levels were standardized by *Z*-score. **(D)** The regulatory network of key flavor metabolites in jujube fruit. Yellow circles represent sugars and organic acids, turquoise diamonds represent transcription factors, purple parallelograms represent structural genes of the TCA cycle, and green triangles represent sugar biosynthesis-related genes.

To further verify the correctness of the transcriptome data, we selected 10 genes for qRT-PCR validation. The results showed that the transcriptome expression trends were consistent with the qRT-PCR results. *ZjTCP14*, *ZjYABBY1*, and *ZjSKU5* were highly expressed at the early stages of fruit development, while the other genes were highly expressed at late stages (Supplementary Figure S5).

Discussion

The combined analysis of metabolomics and transcriptomics are important technical tools for studying the flavor and nutrition of fruits such as watermelon, apple and jujube (Xu et al., 2020; Gong et al., 2021a), but there is still a lack of comprehensive understanding of the accumulation patterns of sugars, acids, and substances at different development stages of jujube. To reveal fruit flavor differences between the two cultivars that differed in flavor due to the different sugar and acid accumulation, this study constructed a global metabolome dataset of the two cultivars at five periods to provide a basis for studying the molecular accumulation of jujube metabolites. We identified 437 DAMs and 16,245 DEGs during fruit ripening. There were some special metabolites, such as LPC and LPE, that were highly accumulated between the two cultivars in the later stage of HK, with high levels of amino acids (leucine, arginine, and homoarginine) in ST (Figures 3D,E). This may have contributed to the different nutritional values of the two cultivars.

Soluble sugars, organic acids, and volatiles are important attributes that determine the color, flavor, and economic value of fruits (Gong et al., 2021a). Jujube is the largest economic tree species in China. Organic acids and soluble sugars change dramatically during the process of fleshy fruits from young to full maturity. Consistent with previous studies, the major sugar components in jujube fruits are fructose, glucose, and sucrose, and organic acids including citric acid and quinic acid (Zhang et al., 2021; Zhao et al., 2021). In this study, we found that both dominant sugars and organic acids were high in the early stages of development and decreased during later stages (Figure 4). Unlike other research that suggested that malic acid was dominant in the later stages of fruit development (Zhen et al., 2016), citric acid was the main content of both cultivars at all stages of fruit development, and the content of quinic acid was higher than that of malic acid. Citric acid and succinic acid are the main factors affecting acidity.

Elucidating the underlying molecular mechanisms of sugar and organic acid changes and their spatiotemporal interactions is a crucial step in understanding fruit development (Yu et al., 2021). The fruit flavor is controlled by the environmental signaling pathways, developmental signaling pathways, metabolic signaling pathways, and transcription factors play important roles in these processes (Hanson et al., 2008; Bastías et al., 2011). Overexpression of *SlAREB1* (ABA-response element binding factors) promoted levels of citric acid, malic acid, glutamic acid, glucose, and fructose in tomato (Bastías et al., 2011). While in apple (*Malus domestica* Borkh.), *MdAREB2* promoted sucrose and soluble sugar accumulation by activating *MdSUT2* (sugar transporter) (Ma et al., 2017). *AcERF182* regulated *AcBAM3.5*, a key structural gene involved in soluble sugar accumulation in kiwifruit (*Actinidia chinensis* Planch) (Wang et al., 2022). *MdbHLH3* directly activated *MdcyMDH* to promote malic acid accumulation in the apple. Additionally, overexpression of *MdbHLH3* increased photosynthetic capacity and carbohydrate content in apple leaves and also increased carbohydrate accumulation in fruits by regulating carbohydrate distribution from source to sink (Yu et al., 2021). Frank et al. (2018) reported that *BASIC LEUCINE ZIPPER63* (*bZIP63*) affects the circadian rhythm of *Arabidopsis* in response to sugar changes by regulating PSEUDO *RESPONSE REGULATOR7* (*PRR7*). In this study, we analyzed the genes that may be related to sugar and organic acid metabolites through WGCNA and Person's relation, and identified transcription factors such as *ZjYABBY1*, *ZjMYB78*, *ZjHAP3*, *ZjTCP14*, and *ZjAGL61*, which can be co-expressed with metabolites and related structural genes at the same time (Figures 5B,C, 6D).

In the present study, *ZjYABBY1*, *ZjHAP3* and *ZjAGL61*were identified as candidate genes regulating the accumulation and metabolism of sugars and organic acids, suggesting that they may participate in fruit development through the metabolic pathways of sugars and organic acids. It is known that fruit formation and ripening is a very complex process, many aspects of fruit size, shape, and further developmental changes depending on organ identities are determined at an early stage (Karlova et al., 2014), Therefore, genes that regulate the dynamic changes of sugar and acid contents during fruit ripening may also be related to fruit morphology. For example, the *ZjYABBY1* gene, which is related to sugar and acid metabolism in this study, has a homologue, *AtYABBY*, that functions in *Arabidopsis* flower as *CRABS CLAW* (*CRC*), which is involved in organ polarity in carpel and nectary development (Bowman and Smyth, 1999; Huang et al., 2013). Another *AtAGL61* regulates central cell development in *Arabidopsis*. MADS-domain proteins *TOMATO AGAMOUS-LIKE1* (*TAGL1*) and *MADS1* were found to be involved in fruit ripening in tomato (Itkin et al., 2009; Dong et al., 2013; Karlova et al., 2014). *ZjHAP3* is a homologous gene of *AtHAP3* (At2g38880), which controls the initiation and development of plant seed embryonic (Su et al., 2021). In contrast, previous studies have shown that *OsHAP3E* participated in the determination of meristem identity in both vegetative and reproductive developments of rice (Zhang and Xue, 2013). It was shown that *AtTCP14* (At3g47620) can break seed dormancy (Zhang et al., 2019; Ferrero et al., 2021). *ZjMYB78* functions in response to abscisic acid and plant drought stress (Dalal et al., 2018).

Conclusion

In general, this study identified the differences in gene expression and nutrient accumulation in different developmental stages of jujube through transcriptome and metabolome analysis. The accumulation of sugars and acids showed opposite trends. Several transcriptional regulators that may affect fruit flavor (sugar and acid) accumulation were identified by joint analysis. The mining of these candidate regulatory genes provides a basis for further improving the flavor and economic value of jujubes.

Data availability statement

The data presented in the study are deposited in the SRA repository, accession number PRJNA835207.

Author contributions

QP and YZ designed the research and revised the manuscript. DL and LZ conducted experiments and data analysis and wrote the manuscript. DL, LZ, YW, and PL performed data analysis. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the National Key R&D Program of China (2019YFD1001605), Beijing Postdoctoral Research Foundation (2022-ZZ-107), the Special Fund for the Construction of Scientific and Technological Innovation Capability (KJCX20200114), the Key R&D Program of Hebei Province (20326807D), and the Key Science and Technology Program of Inner Mongolia Autonomous Region (2021ZD0041-004).

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.952698/full#supplementary-material

References

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., and Cherry, J. M. (2000). Gene ontology: tool for the unification of biology: the gene ontology consortium. *Nat. Genet.* 25, 25–29. doi: 10.1038/75556

Bairoch, A., and Apweiler, R. (2000). The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Res.* 28, 45–48. doi: 10.1093/nar/28.1.45

Barrett, D. M., Beaulieu, J. C., and Shewfelt, R. (2010). Color, flavor, texture, and nutritional quality of fresh-cut fruits and vegetables: desirable levels, instrumental and sensory measurement, and the effects of processing. *Crit. Rev. Food Sci. Nutr.* 50, 369–389. doi: 10.1080/10408391003626322

Bastías, A., López-Climent, M., Valcárcel, M., Rosello, S., Gómez-Cadenas, A., and Casaretto, J. A. (2011). Modulation of organic acids and sugar content in tomato fruits by an abscisic acid-regulated transcription factor. *Physiol. Plant.* 141, 215–226. doi: 10.1111/j.1399-3054.2010.01435.x

Bowman, J. L., and Smyth, D. R. (1999). CRABS CLAW, a gene that regulates carpel and nectary development in *Arabidopsis*, encodes a novel protein with zinc finger and helix-loop-helix domains. *Development* 126, 2387–2396. doi: 10.1242/dev.126.11.2387

Chen, W., Gong, L., Guo, Z., Wang, W., Zhang, H., Liu, X., et al. (2013). A novel integrated method for large-scale detection, identification, and quantification of widely targeted metabolites: application in the study of rice metabolomics. *Mol. Plant* 6, 1769–1780. doi: 10.1093/mp/sst080

Chen, S., Yang, D., Liu, B., Chen, Y., Ye, W., Chen, M., et al. (2021). Identification of crucial genes mediating abdominal aortic aneurysm pathogenesis based on gene expression profiling of perivascular adipose tissue by WGCNA. *Ann. Transl. Med.* 9:52. doi: 10.21037/atm-20-3758

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. doi: 10.1093/bioinformatics/bty560

Chong, J., and Xia, J. (2018). MetaboAnalystR: an R package for flexible and reproducible analysis of metabolomics data. *Bioinformatics* 34, 4313–4314. doi: 10.1093/bioinformatics/bty528

Chu, S., Wang, S., Zhang, R., Yin, M., Yang, X., and Shi, Q. (2022). Integrative analysis of transcriptomic and metabolomic profiles reveals new insights into the molecular foundation of fruit quality formation in *Citrullus lanatus* (Thunb.) Matsum. & Nakai. *Food Qual. Saf.* 6:fyac015. doi: 10.1093/fqsafe/fyac015

Dalal, M., Sahu, S., Tiwari, S., Rao, A. R., and Gaikwad, K. (2018). Transcriptome analysis reveals interplay between hormones, ROS metabolism and cell wall biosynthesis for drought-induced root growth in wheat. *Plant Physiol. Biochem.* 130, 482–492. doi: 10.1016/j.plaphy.2018.07.035

Dong, T., Hu, Z., Deng, L., Wang, Y., Zhu, M., Zhang, J., et al. (2013). A tomato MADS-box transcription factor, SlMADS1, acts as a negative regulator of fruit ripening. *Plant Physiol.* 163, 1026–1036. doi: 10.1104/pp.113.224436

Ferrero, L. V., Gastaldi, V., Ariel, F. D., Viola, I. L., and Gonzalez, D. H. (2021). Class I TCP proteins TCP14 and TCP15 are required for elongation and gene expression responses to auxin. *Plant Mol. Biol.* 105, 147–159. doi: 10.1007/s11103-020-01075-y

Frank, A., Matiolli, C. C., Viana, A. J. C., Hearn, T. J., Kusakina, J., Belbin, F. E., et al. (2018). Circadian entrainment in *Arabidopsis* by the sugar-responsive transcription factor bZIP63. *Curr. Biol.* 28, 2597–2606.e6. doi: 10.1016/j.cub.2018.05.092

Gao, Q. H., Wu, C. S., and Wang, M. (2013). The jujube (*Ziziphus jujuba* Mill.) fruit: a review of current knowledge of fruit composition and health benefits. *J. Agric. Food Chem.* 61, 3351–3363. doi: 10.1021/jf4007032

Goldenberg, L., Yaniv, Y., Porat, R., and Carmi, N. (2018). Mandarin fruit quality: a review. *J. Sci. Food Agric.* 98, 18–26. doi: 10.1002/jsfa.8495

Gong, C., Diao, W., Zhu, H., Umer, M. J., Zhao, S., He, N., et al. (2021a). Metabolome and transcriptome integration reveals insights into flavor formation of 'Crimson' watermelon flesh during fruit development. *Front. Plant Sci.* 12:629361. doi: 10.3389/fpls.2021.629361

Gong, C., Zhu, H., Lu, X., Yang, D., Zhao, S., Umer, M. J., et al. (2021b). An integrated transcriptome and metabolome approach reveals the accumulation of taste-related metabolites and gene regulatory networks during watermelon fruit development. *Planta* 254:35. doi: 10.1007/s00425-021-03680-7

González, I., Déjean, S., Martin, P. G. P., and Baccini, A. (2008). CCA: an R package to extend canonical correlation analysis. *J. Stat. Softw.* 23, 1–14. doi: 10.18637/jss.v023.i12

Hanson, J., Hanssen, M., Wiese, A., Hendriks, M. M. W. B., and Smeekens, S. (2008). The sucrose regulated transcription factor bZIP11 affects amino acid metabolism by regulating the expression of *ASPARAGINE SYNTHETASE1* and *PROLINE DEHYDROGENASE2*. *Plant J.* 53, 935–949. doi: 10.1111/j.1365-313X.2007.03385.x

Huang, J., Chen, X., He, A., Ma, Z., Gong, T., Xu, K., et al. (2021). Integrative morphological, physiological, proteomics analyses of jujube fruit development provide insights into fruit quality domestication from wild jujube to cultivated jujube. *Front. Plant Sci.* 12:773825. doi: 10.3389/fpls.2021.773825

Huang, Z., van Houten, J., Gonzalez, G., Xiao, H., and van der Knaap, E. (2013). Genome-wide identification, phylogeny and expression analysis of *SUN, OFP* and *YABBY* gene family in tomato. *Mol. Gen. Genomics.* 288, 111–129. doi: 10.1007/s00438-013-0733-0

Huang, J., Zhang, C., Zhao, X., Fei, Z., Wan, K., Zhang, Z., et al. (2016). The jujube genome provides insights into genome evolution and the domestication of sweetness/acidity taste in fruit trees. *PLoS Genet.* 12:e1006433. doi: 10.1371/journal.pgen.1006433

Itkin, M., Seybold, H., Breitel, D., Rogachev, I., Meir, S., and Aharoni, A. (2009). TOMATO AGAMOUS-LIKE 1 is a component of the fruit ripening regulatory network. *Plant J.* 60, 1081–1095. doi: 10.1111/j.1365-313X.2009.04064.x

Karlova, R., Chapman, N., David, K., Angenent, G. C., Seymour, G. B., and De Maagd, R. A. (2014). Transcriptional control of fleshy fruit development and ripening. *J. Exp. Bot.* 65, 4527–4541. doi: 10.1093/jxb/eru316

Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27–30. doi: 10.1093/nar/28.1.27

Li, S., Deng, B., Tian, S., Guo, M., Liu, H., and Zhao, X. (2021). Metabolic and transcriptomic analyses reveal different metabolite biosynthesis profiles between leaf buds and mature leaves in *Ziziphus jujuba* mill. *Food Chem.* 347:129005. doi: 10.1016/j.foodchem.2021.129005

Li, J. W., Fan, L. P., Ding, S. D., and Ding, X. L. (2007). Nutritional composition of five cultivars of Chinese jujube. *Food Chem.* 103, 454–460. doi: 10.1016/j.foodchem.2006.08.016

Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930. doi: 10.1093/bioinformatics/btt656

Liu, M., and Wang, M. (2009). *Germplasm resource of Chinese jujube*. Beijing: China Forestry Publishing House.

Liu, M., Zhao, J., Cai, Q., Liu, G., Wang, J., Zhao, Z., et al. (2014). The complex jujube genome provides insights into fruit tree biology. *Nat. Commun.* 5:5315. doi: 10.1038/ncomms6315

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550–521. doi: 10.1186/s13059-014-0550-8

Lu, D., Wu, Y., Pan, Q., Zhang, Y., Qi, Y., and Bao, W. (2022). Identification of key genes controlling L-ascorbic acid during jujube (*Ziziphus jujuba* mill.) fruit development by integrating transcriptome and metabolome analysis. *Front. Plant Sci.* 13:950103. doi: 10.3389/fpls.2022.950103

Ma, Q. J., Sun, M. H., Lu, J., Liu, Y. J., Hu, D. G., and Hao, Y. J. (2017). Transcription factor AREB2 is involved in soluble sugar accumulation by activating sugar transporter and amylase genes. *Plant Physiol.* 174, 2348–2362. doi: 10.1104/pp.17.00502

Natale, D. A., Shankavaram, U. T., Galperin, M. Y., Wolf, Y. I., Aravind, L., and Koonin, E. V. (2000). Towards understanding the first genome sequence of a crenarchaeon by genome annotation using clusters of orthologous groups of proteins (COGs). *Genome Biol.* 1, RESEARCH0009–RESEARCH0019. doi: 10.1186/gb-2000-1-5-research0009

Ogata, H., Goto, S., Sato, K., Fujibuchi, W., Bono, H., and Kanehisa, M. (2000). KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, E104–E130. doi: 10.1093/nar/28.1.27

Otasek, D., Morris, J. H., Bouças, J., Pico, A. R., and Demchak, B. (2019). Cytoscape automation: empowering workflow-based network analysis. *Genome Biol.* 20:185. doi: 10.1186/s13059-019-1758-4

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290–295. doi: 10.1038/nbt.3122

Shi, Q., Du, J., Zhu, D., Li, X., and Li, X. (2020). Metabolomic andtranscriptomic analyses of anthocyanin biosynthesis mechanisms in the color mutant *Ziziphus jujuba* cv. Tailihong. *J. Agric. Food Chem.* 68, 15186–15198. doi: 10.1021/acs.jafc.0c05334

Su, L., Wan, S., Zhou, J., Shao, Q. S., and Xing, B. (2021). Transcriptional regulation of plant seed development. *Physiol. Plant.* 173, 2013–2025. doi: 10.1111/ppl.13548

Varet, H., Brillet-Guéguen, L., Coppée, J. Y., and Dillies, M. A. (2016). SARTools: a DESeq2- and EdgeR-based R pipeline for comprehensive differential analysis of RNA-Seq data. *PLoS One* 11:e0157022. doi: 10.1371/journal.pone.0157022

Wang, R., Shu, P., Zhang, C., Zhang, J., Chen, Y., Zhang, Y., et al. (2022). Integrative analyses of metabolome and genome-wide transcriptome reveal the regulatory network governing flavor formation in kiwifruit (*Actinidia chinensis*). *New Phytol.* 233, 373–389. doi: 10.1111/nph.17618

Wilke, A., Harrison, T., Wilkening, J., Field, D., Glass, E. M., Kyrpides, N., et al. (2012). The M5nr: a novel non-redundant database containing protein sequences and annotations from multiple sources and associated tools. *BMC Bioinformatics* 13, 1–5. doi: 10.1186/1471-2105-13-141

Xiong, Y., Yan, P., Du, K., Li, M., Xie, Y., and Gao, P. (2020). Nutritional component analyses of kiwifruit in different development stages by metabolomic and transcriptomic approaches. *J. Sci. Food Agric.* 100, 2399–2409. doi: 10.1002/jsfa.10251

Xu, J., Yan, J., Li, W., Wang, Q., Wang, C., Guo, J., et al. (2020). Integrative analyses of widely targeted metabolic profiling and transcriptome data reveals molecular insight into metabolomic variations during apple (*Malus domestica*) fruit development and ripening. *Int. J. Mol. Sci.* 21. doi: 10.3390/ijms21134797

Yang, H., Tian, C., Ji, S., Ni, F., Fan, X., Yang, Y., et al. (2021). Integrative analyses of metabolome and transcriptome reveals metabolomic variations and candidate genes involved in sweet cherry (*Prunus avium* L.) fruit quality during development and ripening. *PLoS One* 16:e0260004. doi: 10.1371/journal.pone.0260004

Yu, J. Q., Gu, K. D., Sun, C. H., Zhang, Q. Y., Wang, J. H., Ma, F. F., et al. (2021). The apple bHLH transcription factor MdbHLH3 functions in determining the fruit carbohydrates and malate. *Plant Biotechnol. J.* 19, 285–299. doi: 10.1111/pbi.13461

Zhang, W., Cochet, F., Ponnaiah, M., Lebreton, S., Matheron, L., Pionneau, C., et al. (2019). The MPK8-TCP14 pathway promotes seed germination in *Arabidopsis*. *Plant J.* 100, 677–692. doi: 10.1111/tpj.14461

Zhang, Y., Tong, P., Liang, F., Wu, C., and Wang, J. (2021). Dynamic changes of sugar content in jujube fruit and analysis of related gene expression. *Acta Agric. Jiangxi* 33, 25–31. doi: 10.19386/j.cnki.jxnyxb.2021.03.04

Zhang, J. J., and Xue, H. W. (2013). OsLEC1/OsHAP3E participates in the determination of meristem identity in both vegetative and reproductive developments of rice. *J. Integr. Plant Biol.* 55, 232–249. doi: 10.1111/jipb.12025

Zhao, A., Xue, X., Ren, H., Wang, Y., Li, D., and Li, Y. (2021). Analysis of composition and content characteristics of organic acids in jujube germplasm. *Acta Agric. Bor. Occidentalis Sin.* 30, 1185–1198. doi: 10.7606/j.issn.1004-1389.2021.08.009

Zhen, H., Zhang, Q., Li, W., Zhang, S., and Xi, W. (2016). Changes in soluble sugars and organic acids of Xinjiang apricot During fruit development and ripening. *Sci. Agric. Sin.* 49, 3981–3992. doi: 10.3864/j.issn.0578-1752.2016.20.012

Zhou, Z., Gao, H., Ming, J., Ding, Z., Lin, X. E., and Zhan, R. (2020). Combined transcriptome and metabolome analysis of pitaya fruit unveiled the mechanisms underlying peel and pulp color formation. *BMC Genomics* 21:734. doi: 10.1186/s12864-020-07133-5

Zhu, G., Wang, S., Huang, Z., Zhang, S., Liao, Q., Zhang, C., et al. (2018). Rewiring of the fruit metabolome in romato breeding. *Cell* 172, 249–261.e12. doi: 10.1016/j.cell.2017.12.019

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 28 September 2022
DOI 10.3389/fpls.2022.1021572

OPEN ACCESS

EDITED BY
Neftali Ochoa-Alejo,
Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Mexico

REVIEWED BY
David Horvath,
Agricultural Research Service (USDA), United States
Masamitsu Sato,
Waseda University, Japan

*CORRESPONDENCE
Sheng-Chao Yang
shengchaoyang@163.com
Yan-Li Liang
943029567@qq.com

†PRESENT ADDRESS
Qiu-Xiong Yang,
Kunning, China
Dan Chen,
Kunning, China

‡These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Plant Development and EvoDevo, a section of the journal Frontiers in Plant Science

RECEIVED 17 August 2022
ACCEPTED 08 September 2022
PUBLISHED 28 September 2022

CITATION
Yang Q-X, Chen D, Zhao Y, Zhang X-Y, Zhao M, Peng R, Sun N-X, Baldwin TC, Yang S-C and Liang Y-L (2022) RNA-seq analysis reveals key genes associated with seed germination of *Fritillaria taipaiensis* P.Y.Li by cold stratification. *Front. Plant Sci.* 13:1021572. doi: 10.3389/fpls.2022.1021572

RNA-seq analysis reveals key genes associated with seed germination of *Fritillaria taipaiensis* P.Y.Li by cold stratification

Qiu-Xiong Yang[1,2,3†‡], Dan Chen[1,2,3†‡], Yan Zhao[1,2,3], Xiao-Yu Zhang[1,2,3], Min Zhao[1,2,3], Rui Peng[4], Nian-Xi Sun[4], Timothy Charles Baldwin[5], Sheng-Chao Yang[1,2,3*] and Yan-Li Liang[1,2,3*]

[1]College of Agronomy & Biotechnology, Yunnan Agricultural University, Kunming, China, [2]Key Laboratory of Medicinal Plant Biology of Yunnan Province, Yunnan Agricultural Waseda University, Fengyuan , Kunming, China, [3]National & Local Joint Engineering Research Center on Germplasm Innovation & Utilization of Chinese Medicinal Materials in Southwestern China, Yunnan Agricultural University, Kunming, China, [4]Chongqing Academy of Chinese Materia Medica, Chongqing, China, [5]Faculty of Science and Engineering, University of Wolverhampton, Wolverhampton, United Kingdom

Seed dormancy is an adaptive strategy for environmental evolution. However, the molecular mechanism of the breaking of seed dormancy at cold temperatures is still unclear, and the genetic regulation of germination initiated by exposure to cold temperature requires further investigation. In the initial phase of the current study, the seed coat characteristics and embryo development of *Fritillaria taipaiensis* P.Y.Li at different temperatures (0°C, 4°C, 10°C & 25°C) was recorded. The results obtained demonstrated that embryo elongation and the dormancy-breaking was most significantly affected at 4°C. Subsequently, transcriptome analyses of seeds in different states of dormancy, at two stratification temperatures (4°C and 25°C) was performed, combined with weighted gene coexpression network analysis (WGCNA) and metabolomics, to explore the transcriptional regulation of seed germination in *F. taipaiensis* at the two selected stratification temperatures. The results showed that stratification at the colder temperature (4°C) induced an up-regulation of gene expression involved in gibberellic acid (GA) and auxin biosynthesis and the down-regulation of genes related to the abscisic acid (ABA) biosynthetic pathway. Thereby promoting embryo development and the stimulation of seed germination. Collectively, these data constitute a significant advance in our understanding of the role of cold temperatures in the regulation of seed germination in *F. taipaiensis* and also provide valuable transcriptomic data for seed dormancy for other non-model plant species.

KEYWORDS

seed dormancy, 4°C stratification, RNA-seq, phytohormone, WGCNA, *fritillaria taipaiensis* P.Y.Li

Introduction

Seed dormancy is an evolutionary adaption to habitat climate diversity, which results in a delay to seed germination (Willis et al., 2014; Finkelstein et al., 2008). It can also directly or indirectly impact the production of a wide variety of crops (Simsek et al., 2014). Therefore, seed dormancy plays a significant role in both plant ecology and agriculture. Seed dormancy is divided into four classes including: nondormancy (ND), physiological dormancy (PD), morphological dormancy (MD) and morphophysiological dormancy (MPD) (Baskin and Baskin, 2004). At present, research on seed dormancy has mainly focused on model plant species and seed physiological dormancy (PD) (Finkelstein and Reeves, 2018; Chahtane et al., 2017; Finch-Savage and Footitt, 2017) However, studies related to seed morphological dormancy (MD) are relatively uncommon (Walker et al., 2021). Our understanding of the molecular mechanism(s) of the breaking of seed dormancy at cold temperatures is even more limited.

Fritillaria taipaiensis P.Y.Li, a traditional Chinese medicinal plant, is important in Asia, both as a food crop and for its medicinal value. This species is a perennial alpine plant, indigenous to high-altitudes located in the mountainous regions of southwestern China. The process of seed dormancy in *F. taipaiensis* lasts for 200 days and is considered to be primarily due to morphophysiological dormancy (MPD). From our preliminary work, we found that cold temperatures may contribute to breaking the morphophysiological seed dormancy of this species. Therefore, we determined that the seed of *F. taipaiensis* provides good material for the investigation of seed germination initiated by exposure to cold temperatures.

It is widely known that both seed dormancy and germination are closely regulated by the crosstalk between the gene regulatory networks involved in the biosynthesis of the plant growth regulators ABA and GA (Gubler et al., 2005; Holdsworth et al., 2008; Graeber et al., 2012).

Expression of genes and transcription factors related to endogenous ABA biosynthesis, signal transduction and metabolism play a particularly important role in seed dormancy. The process whereby phytoxanthin is converted to xanthoxin, catalyzed by zeaxanthin cyclooxygenase (ZEP),9-cis-cyclooxygenase lyase (NCED), abscisic aldehyde oxidase (AAO) (Nambara and Marion-Poll, 2005; Arc et al., 2013) as a component of bioactive ABA synthesis *via* the carotenoid pathway is particularly important. Bioactive ABA can be glycosylated by UDP glucosyltransferase (UGT) to inactive ABA-GE (Liu et al., 2015). Conversely, β-glucosidases encoded by *AtBg1* and *AtBg2* can convert ABA-GE into bioactive ABA (Lee et al., 2006; Xu et al., 2012). The signal transduction of ABA is also important. PYR/PYL/RCAR receptors can bind PP2C and inhibit its activity,and cause the release of SnRKs. The phosphorylation of SnRKs can activate the downstream signal cascade, ABA-activated SnRK2s phosphorylate Raptor and inhibit TOR activity, TOR and ABA signaling balance plant growth (Cutler et al., 2010; Hubbard et al., 2010; Wang et al., 2018). Transcription factors *ABI3* and *ABI5* also play a key role in the downstream regulation of the ABA signalling network (Piskurewicz et al., 2008; Barrero et al., 2010; Kanai et al., 2010; Penfield, 2017).

GA promotes seed germination by antagonizing and inhibiting ABA (Graeber et al., 2012). Previous studies showed that GAs anabolic pathway is comprehensive, and 136 types of GAs have been identified, four of which (GA_1/GA_3/GA_4/GA_7) have biological activity (Yamaguchi, 2008). GA_{12} is an important precursor, which can be catalyzed to produce bioactive GA by GA 20-oxidase (GA20ox) and GA 3-oxidase (GA3ox) (Sasaki et al., 2002; Appleford et al., 2006). Bioactive GAs can be deactivated by GA 2-oxidases (GA2oxs) (Sakamoto et al., 2001; Schomburg et al., 2003). The DELLA protein is also an important negative regulator involved in the GA signal transduction pathway, responding to environmental changes together with the GA receptor GIBBERELLIN INSENSITIVE DWARF 1 (GID1) (Harberd et al., 2009; Resentini et al., 2015). *DELAY OF GERMINATION-1*(*DOG1*) and *EM6* are the key regulators in seed primary dormancy (Nishimura et al., 2018; Carrillo-Barral et al., 2020).

Many other transcription factors and signalling networks also impact upon seed germination. *MFT* encodes MOTHER of FT and TFL1 protein, being activated and regulated by ABI3 and ABI5 during seed embryo development (Xi et al., 2010). *ZOUPI (ZOU)* is specifically expressed in endosperm tissue with INDUCER OF CBP EXPRESSION 1 (ICE1), which mediates the degradation of the surrounding endosperm and embryonic epidermal development during embryo growth. The ICE-CBF-COR module regulates the cold adaptation process in the model species *Arabidopsis thaliana*, in which ICE1 binds to MYC recognition sites in the CBF3 promoter to positively regulate CBF expression. *ICE1* expression is induced by low temperature, ICE1 mutation impairs chilling and freezing tolerance (Denay et al., 2014; MacGregor et al., 2019). GASSHO1 and GASSHO2 (GSO1 and GSO2) receptors are located on the embryonic surface, being responsible for signal communication between embryo and endosperm (Doll and Ingram, 2022). The expansin family of cell wall-associated proteins also plays an important role in promoting endosperm rupture, seed embryo growth and germination (Chen and Bradford, 2000; Chen et al., 2001; Yan et al., 2014; Muthusamy et al., 2020). Late embryonic development protein (LEA) is considered to be a marker during seed development and maturity (Shih and Hsing, 2008), participating in the response to various abiotic stresses (Wang et al., 2018; Czernik et al., 2020).

Although studies on seed dormancy are abundant in the literature, the progress on seed morphophysiological dormancy and seed dormancy breaking induced under cold temperatures requires further investigation. In the current study, the seed coat characteristics and embryo growth and development were recorded at different temperatures. In addition we also

performed a transcriptome and endogenous hormone metabolome, with weighted gene coexpression network analysis (WGCNA), using the seeds of *F. taipaiensis* at nine different stages of dormancy. The main objectives of the study were (1) To identify differentially expressed genes and the response pathways of cold stratification. (2) To identify the key genes involved in the regulation of embryo elongation and seed germination under cold temperatures. (3) To explore the hub genes that link the cold temperature response to the regulation of embryo growth and development.

Materials and methods

Plant materials and growth conditions

Mature seeds of *F. taipaiensis* were surface sterilized with 20% (v/v) bleach for 5 min and then washed with sterile, distilled water. Turgid, shiny seeds were then selected for cultivation in petri dishes. Seeds were germinated on 2 layers of sterilized filter paper moistened under controlled conditions (0°C/4°C/10°C/25°C) for a period of two months, and each treatment was performed with three replicates (n=30, Supplementary Table 3). The rate of water absorption in the seed was assessed at various timepoints (0, 1, 3, 5, 10, 15, 20, 30,120, 240, 360, or 720min). The length of the embryo was counted each week, and the index of embryo elongation was calculated using the formula embryo rate (%) =seed embryo length/ endosperm length. The seed coat surface morphology was observed by scanning electron microscopy (Yu et al., 2014).

Sample collection and transcriptome sequencing

The mature seeds of *F. taipaiensis* used for transcriptome sequencing, were obtained from fresh seed capsules. The seeds were cultivated at 25°C in the controls, or at 4°C for the treatments.

The embryos treated with stratification at 4°C were observed to elongate rapidly. The seeds stratified at 4°C were then subdivided in to five A, B, C, D & E samples. Sample A consisted of fresh seeds whilst samples B, C, D and E comprised of seeds cultivated at 4°C in which the embryo rate reached 25%, 50%, more than 90% and the radicle had broken through the seed coat by 1-3mm (Figure 1).

The seeds cultivated at 25°C were used as the control and samples of which were labelled b, c, d and e. The time of control sampling was consistent with the treatments and the length of embryo were almost unchanged. Three replicates of each of the 9 samples, containing a total of 27 samples were analysed (Supplementary Table 1). Each sample was snap frozen in liquid nitrogen and stored at -80°C. After mRNA extraction, cDNA library construction and transcriptome sequencing were performed using an Illumina HiSeq platform, by Metware sequencing services (http://www.metware.cn/).

Transcriptome assembly and analysis

Trimmomatic v0.39 was used to filter the original sequence (Specific parameters: LEADING:20 TRAILING:20 SLIDINGWINDOW:4:20 MINLEN:50). RiboDetector v0.2.4 was utilized to remove rRNA sequences. *De novo* assembly was performed using Trinity v2.12.0 (k-mer:31, min_kmer_cov:2). The final unigenes were obtained from the cluster of assembled sequences using RapClust. The TransDecoder v5.5.0 was selected to predict the CDs region of unigenes, and BLAST v2.12.0 (The threshold of E-vaule = $1x10^{-5}$) was used to match the predicted CDs in the local KEGG, KOG, NR, NT, TrEMBL and Swissprot databases. The sequences were annotated to the Pfam database (version35.0) using the pfam-scan v1.6 program. GO and KO ortholog enrichment analyses were subsequently performed using interproscan v5.55-88.0 and kofam_scan v1.3.0. (The original data query accession numbers are: PRJNA874905, PRJNA874906, PRJNA874908, PRJNA874910, PRJNA874912, PRJNA874913, PRJNA874914, PRJNA874917, PRJNA874920. Fasta data in Supplementary Material 1, BUSCO 5.3.2,embryophyta: C: 92.0% [S:84.4%,D:7.6%], F:4.3%, M:3.7%, n: 1614).

RT-qPCR

Total RNA was extracted using a HiPure HP Plant RNA Mini Kit (Guangzhou Meiji Biotechnology Co., Ltd.). First-strand cDNA was synthesized from RevertAidTM First strand cDNA Synthesis (TransGen Biotech, Beijing, China). The PCR reactions were carried out on QuanstudioTM 5 Real-Time PCR Instrument (Thermo Fisher Scientific, Inc.), using ChamQ Universal SYBR qPCR Master Mix (Vazyme Biotech Co., Ltd.). A ubiquitin gene was selected as the internal reference, and primers were designed by Primer3 (https://primer3.ut.ee/). The amplification reaction mixture consisted of 10 μL of 2× ChamQ Universal SYBR qPCR Master Mix, 0.5 μL each of 10 mM forward and reverse primers, 1 μL of cDNA template, and 6 μL of ddH_2O in a final volume to 18 μL. The qPCR reaction conditions were: pre-denaturation: 95°C for 30 sec, cyclic reaction: 95°C for 10 sec, 60°C for 30 sec, 40 cycles, lysis curve reaction: 95°C 15 sec, 60°C 1 mine, 9°C 15 sec. Relative expression values were obtained using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). The primers used for fluorescence quantification are shown in the Supplementary Table 2.

Determination of endogenous hormones

The seeds of 5 developmental stages (A/B/C/D/E) were used to determine the endogenous hormone content. ABA, IAA, and

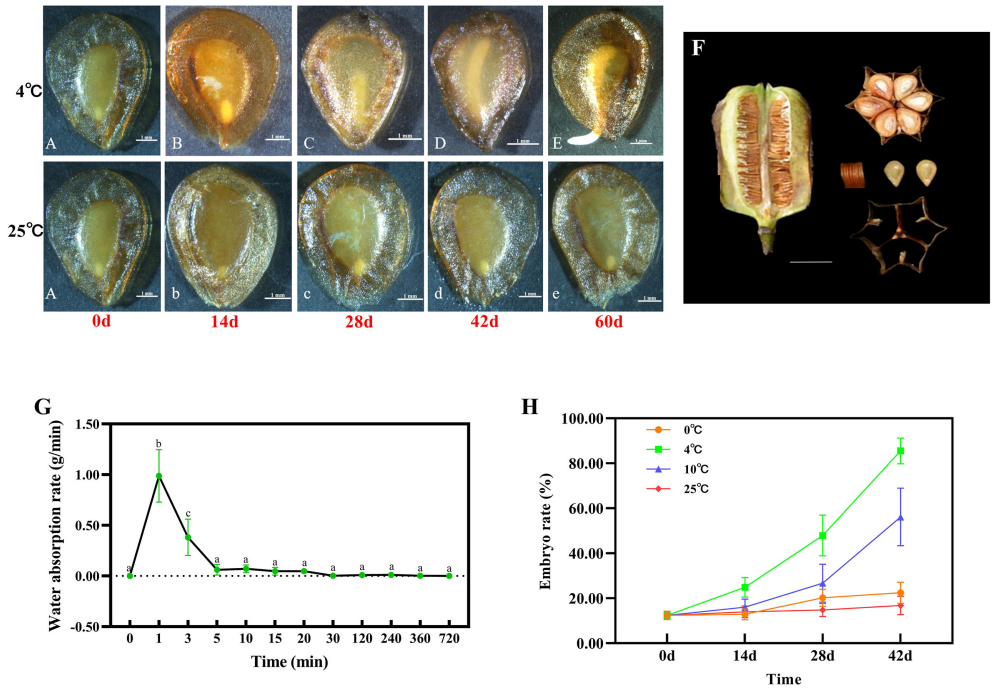


FIGURE 1
Phenotypic observation of the mature seed capsule and seed of *F. taipaiensis*. Seed embryo growth progression at 4°C **(A–E)** and 25°C **(b–e)**. The letters **(A–E)**, **(b–d)** and e represent 9 developmental stages of seed development, respectively (listed in Supplementary Table 1). Scale bar=1 mm. **(F)** Morphological and anatomical structure of mature seed capsule. Scale bar =1 cm. **(G)** The changes of seed water absorption rate. Values for each point were means ± SD (n = 5). Different letters indicate significant difference ($P < 0.05$). **(H)** The changes of embryo growth rate at different stratification temperatures of 0°C, 4°C, 10°C and 25°C. Values for each point were means ± SD (n = 30).

cytokinin were detected by MetWare (http://www.metware.cn/). Three replicates of each assay were performed.

WGCNA analysis

A phenotype weighted coexpression network was constructed in the R-package WGCNA with 27 samples and 9 phenotypes (Langfelder and Horvath, 2008), all differentially expressed genes were submitted to WGCNA, adjusting a batch using limma's removeBatchEffect (Ritchie et al., 2015), and the establishment of 9 module features (A/B/C/D/E/b/c/d/e). The soft threshold was calculated using the pickSoftThreshold function. The modules were obtained using the automatic network construction function blockwiseModules. The eigenvalues of each module were calculated and used to test the correlation of each trait. P values < 0.05 was considered significant. K_{ME} was used to express the connectivity of genes in specific modules. The network connectivity between nodes was analysed using 11 algorithms of the Cytoscape plugin cytoHubba (version 3.8.2) (Chin et al., 2014). The unigenes with high connectivity were selected as hub genes based on the results of Maximal Clique Centrality (MCC) analysis.

Results

Cold ambient air temperatures (4°C) may significantly promote embryo elongation in *F. taipaiensis*

Seeds of *F. taipaiensis* are flat, heart-shaped or obovate, and the observed 1000 grain weight in the current study was 4.45 ± 0.17g. The seed coat was shown to be highly permeable, and its epidermal cells were loosely arranged (Supplementary Figure 1). The water absorption rate of dry seeds reached 0.98 ± 0.25g/min, while the water was absorbed into seeds for 1-3min. Seeds were observed to become fully imbibed within 30min (Figure 1G).

The immature embryo was embedded in an abundant living endosperm, that occupied most of the seed's volume (Figure 1A). The length of the embryo accounted for 12.38% of the size of endosperm (Figure 1H, Supplementary Table 3). These seeds exhibit the characteristic features of morphological dormancy.

In order to investigate the effect of temperature upon seed germination, seeds were grown under the different temperature regimes (0, 4, 10 and 25°C), and the phenotype of germination was observed (Figure 1H). The results showed that the relatively cold temperature (4°C) had the most significant effect on embryo

elongation and dormancy-breaking. The embryo rate reached about 25% after cold stratification for 14 days at 4°C (Figure 1B). The embryo rate accounted for 50% and more than 85.5% of the endosperm (mature embryo) after 28 days and 42 days successively (Figures 1C, D). The seeds were then placed in warmer environmental conditions (15-20°C, 7-10 days) to complete their germination (Figure 1E, Supplementary Figure 2). The growth rate of embryos pretreated at 10°C was slower than those stratified at 4°C. The embryo rate accounted for 16%, 26% and 56% of the endosperm after 14 days, 28 days and 42 days (Figure 1H). The embryos treated at 0°C or 25°C showed no significant change at the same time. These data suggest that the best stratification temperature (of those examined) at which to break the seed dormancy of *F. taipaiensis* is 4°C.

RNA-sequencing analysis

To understand how cold temperature promotes seed germination, three biological replicates were selected and a total of 27 seed samples were used for transcriptomic analysis. The samples (named A, B, C, D, E, b, c, d and e) are shown in Figure 1. Approximately 177.09 Gb clean data were generated. An average of 93% raw reads had a quality score of Q30 (Supplementary Table 4, Sheet 1). A total of 67,784 unigenes were obtained and 61,574 unigenes could be matched to KEGG, NR, Swiss-Prot, KOG and TrEMBL databases (Supplementary Table 4, Sheet 2). The pearson correlation analysis of the RNA-seq data showed that the repeatability of the data was good (Supplementary Figure 3). PCA analysis indicated that high PC1 scores exhibited 17.94% of the trait variance for different time points, and PC2 explained 12.45% of the total variance for genotypes (Supplementary Figure 4).

To identify changes in our transcriptome data that occurred during the seed germination under cold treatment, the TPM (Transcripts Per Million) values were used to measure genes or transcripts expression levels. Differentially expressed genes (DEGs) (expression difference fold | log2FoldChange | > 1 and significance p-value < 0.05) were identified among samples A, B, C, D and E. Each was compared with samples b, c, d and e. The results showed that 374, 342, 968 and 3,220 DEGs were identified between A vs B, A vs C, A vs D, A vs E, respectively. There were 1,397, 2,211, 3,242 and 3,653 DEGs between B vs b, C vs c, D vs d, E vs e, respectively. It was found that the number of DEGs were enriched during the germination time points. 4,271 unigenes exhibited conserved expression at the 9 developmental stages (Figure 2A).

Functional enrichment analysis

The GO enrichment analysis was used to classify the function of the unigenes, which were differentially expressed during seed germination under cold stratification. A total of 37,327 unigenes were distributed into the different known GO terms. Of them, 32,246 unigenes were annotated in GO-MF terms, 9,565 unigenes in GO-CC terms, 29,095 unigenes in GO-BP terms (Figure 3A, B, Supplementary Figure 5). Among the "biological process" category, the most significantly enriched terms were "response to oxidative stress", "cell wall organization", "response to abscisic acid", "carbohydrate transport", "response to cold", "response to auxin". The top enriched GO terms in "molecular functions" were related to "hydrolase activity, hydrolyzing O-glycosyl compounds", "oxidoreductase activity", "transferase activity", "UDP-glycosyltransferase activity", "glutathione transferase activity", "signaling receptor activity", "abscisic acid binding", "pectate lyase activity". Moreover, GO analysis showed that more significant enrichment in developmental stage E than in stage A at different functional modules.

To identify the functions of DEGs related to seed germination at 4°C, KEGG pathway enrichment analysis was performed for 14 days, 28days and 48days germination at 4°C stratification, compared to fresh seeds. The top significant enriched pathways were "Phenylpropanoid biosynthesis", "Plant hormone signal transduction", "MAPK signaling pathway-plant", "Fructose and mannose metabolism", "Pentose and glucuronate interconversions", "Linoleic acid metabolism" The number of unigenes enriched in these pathways showed an upward trend in developmental stages A-E (Figure 3C). Therefore, based on the results of KEGG, the MAPK signaling, hormone signaling, and fructose and mannose metabolism related pathways were selected as being the most important pathways related to plants' perception and adaptation to cold temperature, and which are thereby implicated in the control of seed germination of *F. taipaiensis* seeds at 4°C.

Identification of DEGs involved in endogenous hormone regulation pathway

To elucidate the genetic regulation of the ABA pathway in the maintenance of seed dormancy and the breaking of seed dormancy, the DEGs were filtered and compared to those previously reported to be involved in ABA biosynthetic, transport, and signaling pathways. Heatmap analysis of the DEGs involved in ABA biosynthesis, revealed that their expression was inhibited at 4°C compared to 25°C, such as *ZEP (ZEP-FtLi.1, ZEP-FtLi.2, ZEP-FtLi.3)* and *NCED(NCED-FtLi.4)* (Figure 4A, Supplementary Table 5). Furthermore, the expression of the *PYR* receptors (*PYR-FtLi.1, PYR-FtLi.2*), kinases *SnRK2s* (SnRK2s-FtLi.1, SnRK2s-FtLi.2), transcription factors *ABI3(ABI3-FtLi.1, ABI3-FtLi.2, ABI3-FtLi.3, ABI3-FtLi.4)*, and transcription factors *bZIP67*(*bZIP67-FtLi.1, bZIP67-FtLi.2, bZIP67-FtLi.3, bZIP67-FtLi.4*) were also

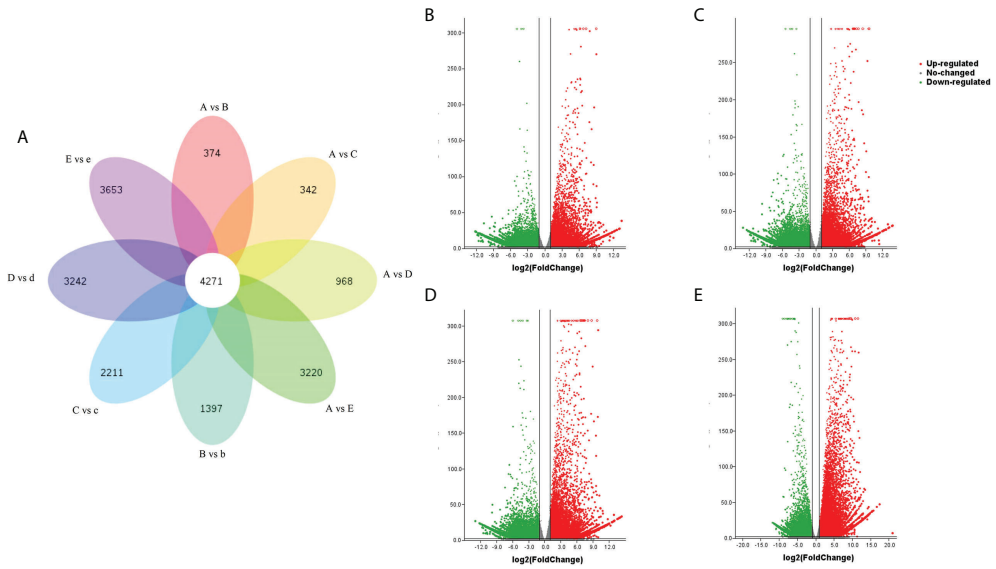


FIGURE 2
Transcriptomic analysis of the DEG distribution in the nine seed samples. **(A)** Venn diagram representation of the stage-specific DEGs distribution in the seed samples. A vs B, A vs C, A vs D, A vs E, B vs b, C vs c, D vs d and E vs e represent the DEGs in each pairwise comparison, respectively. **(B–E)** Differentially expressed gene volcano map of cold treatment for 14 d, 28 d, 42 d or germination. **(B–E)** were represent A vs B, A vs C, A vs D, and A vs E, respectively. Differentially expressed genes were: expression difference foldlog2FoldChange> 1 and significance p-value < 0.05.

significantly reduced. It was also found that the expression of *ABI5 (ABI5-FtLi.1, ABI5-FtLi.2, ABI5-FtLi.4)* was induced at 25°C. The *PP2C*(*PP2C-FtLi.4*) and *CYP707A* (*CYP707A-FtLi.1, CYP707A-FtLi.2, CYP707A-FtLi.3)* genes that are responsible for inactivating/inhibiting ABA biosynthesis were significantly upregulated in developmental stages A to E stages, but most significantly in sample E.

Many of the DEGs were genes involved in the *de novo* biosynthesis of active gibberellin (Figure 4B), such as *KAO* (*KAO-FtLi1, KAO-FtLi3, KAO-FtLi4, KAO-FtLi5*), *KO* (*KO-FtLi3, KO-FtLi4, KO-FtLi5, KO-FtLi6*), *GA3ox* (*GA3ox-FtLi1, GA3ox-FtLi2, GA3ox-FtLi3, GA3ox-FtLi4, GA3ox-FtLi8*) and *GA20ox* (*GA20ox-FtLi1, GA20ox-FtLi2)*. These genes were significantly expressed at 4°C compared to 25°C. In accordance with this biosynthetic process, the expression of the *GID1* (*GID1-FtLi6, GID1-FtLi7, GID1-FtLi8, GID1-FtLi9, GID1-FtLi16*) family of receptor genes was also upregulated. Stratification at 4°C was shown to inhibit the expression of the repressor gene *DELLA* (*DELLA-FtLi1, DELLA-FtLi2* and *DELLA-FtLi3*) and the inactivation gene *GAMT* (*GAMT-FtLi1, GAMT-FtLi2, GAMT-FtLi3, GAMT-FtLi4*).

We filtered DEGs for those previously reported to be involved in auxin biosynthesis, transport, and signaling pathways (Figure 4C). The auxin biosynthesis genes *YUCCA* (*YUCCA-FtLi.3*, *YUCCA-FtLi.4*), *TAA1*(*TAA1-FtLi.1*) and auxin efflux carrier genes *PIN* (*PIN-FtLi.2*), the influx transporter genes *AUX1*(*AUX1-FtLi.1, AUX1-FtLi.2, AUX1-FtLi.6, AUX1-FtLi.7*) showed the largest upregulation under cold stratification. *ARF* (*ARF-FtLi.1, ARF-FtLi.3, ARF-FtLi.5, ARF-FtLi.6*) and *SAUR* (*SAUR-FtLi.8, SAUR-FtLi.9, SAUR-FtLi.6*) were DEGs related to auxin signaling pathways, which were all shown to be highly expressed when the radicles ruptured the seed coat in the 4°C stratification treatments. The expression of the *GH3* (*GH3-FtLi.1, GH3-FtLi.2, GH3-FtLi.3, GH3-FtLi.4, GH3-FtLi.5*) gene family was shown to be significantly increased in stage B. The aminotransferase VAS1, negatively regulates IAA biosynthesis. Cold stratification significantly inhibited *VAS1-FtLi.6, VAS1-FtLi.7, VAS1-FtLi.8, VAS1-FtLi.9, VAS1-FtLi.10* expression.

ICE1 plays a positive regulatory role in the response to cold stress. Cold stratification upregulated the expression of *ICE1*, including *ICE1-FtLi.4, ICE1-FtLi.7, ICE1-FtLi.10, ICE1-FtLi.13* and *ICE1-FtLi.19*. By contrast, 4°C pretreatment inhibited the expression of several dormancy maintenance genes, including *DOG1* (*DOG1-FtLi3, DOG1-FtLi4, DOG1-FtLi5, DOG1-FtLi6, DOG1-FtLi7 and DOG1-FtLi8*), *EM6*(*EM6-FtLi1, EM6-FtLi2* and *EM6-FtLi3*) and *MFT* (*MFT-FtLi2*). In the current study, 24 genes were identified encoding for expansin proteins that play a pivotal role in primary cell wall loosening and cell expansion. Throughout seed embryo maturation, *EXPA-FtLi14* was highly expressed. The proteins (LEA-FtLi) play a critical role during seed maturation, and the genes that encode for it (*LEA-FtLi2, LEA-FtLi4, LEA-FtLi5, LEA-FtLi6, LEA-FtLi8*) showed significant upregulation in stage B. Furthermore, we identified several

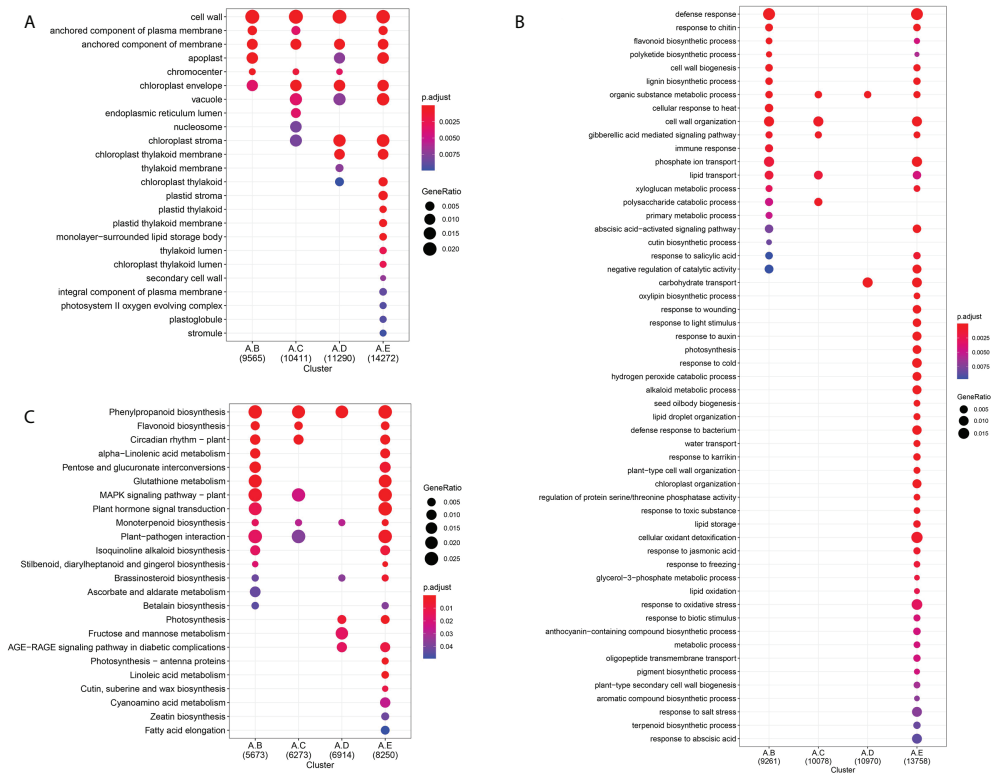


FIGURE 3
GO and KEGG enrichment analysis of DEGs related to seed germination by 4°C treatment for 0d, 14 d, 28 d, 42 d or germination. **(A)** GO analysis of the terms enriched in the DEGs among to the 4 comparisons (A vs B, A vs C, A vs D, A vs E). The top 24 terms of most enriched GO terms belong to cellular composition. padj < 0.01. **(B)** The top 55 terms of most enriched GO terms belong to biology process. padj < 0.01. **(C)** KEGG enrichment pathways. The left side of each figure represents the enriched function terms. The dot size indicates the number of GeneRatio. padj < 0.05.

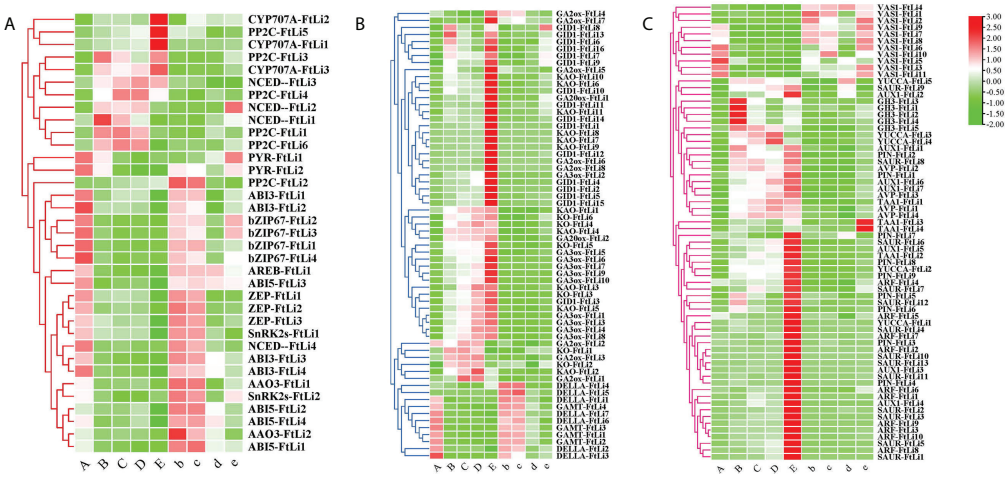


FIGURE 4
Heatmaps of the expression of ABA - **(A)**, GA- **(B)**, and auxin pathway **(C)** related genes between the treatments and controls. (A-E) represents the treatments and "b", "c", "d" "e" represents the controls, respectively. Red and green represent up- and downregulated transcripts, respectively. All genes are listed in detail in Supplementary Table 5.

genes involved in carbohydrate metabolism, including α-amylase (AMY), β-amylase (BAM) and mannan endo-1,4-beta-mannosidase (MAN). The enzyme genes *AMY-FtLi.3, AMY-FtLi.5, BAM-FtLi.7, BAM-FtLi.8, MAN-FtLi.4, MAN-FtLi.5* were highly expressed from stage B to stage E (Supplementary Figure 6).

Identification of key genes and modules during seed germination by WGCNA

A weighted gene co-expression network analysis (WGCNA) can construct gene interaction networks, as well as identify gene modules and hub genes within them. We performed a WGCNA using our RNA-seq data to identify genes that link the cold response to plant growth and development. For the purpose of reducing noise, we only included genes that were differentially expressed across at least one comparison (A vs. B, A vs. C, A vs. D, A vs. E, B vs. b, C vs. c, D vs. d or E vs. e). For the analysis of a network topology, 14 was set as the soft-thresholding power (Supplementary Figures 7A, B). Based upon pairwise correlation analysis of gene expression, two main branches could be identified among the 21 merged coexpression modules (Supplementary Figure 7C). According to the module-trait correlation analysis for the 27 samples, 4 modules ($R > 0.90$) were significant (purple, darkmagenta, mediumpurple4 and salmon), and which were significantly correlated with samples A, E, d, e. These modules had gene numbers that ranged from 1,426 (salmon) to 6,620 (darkmagenta) (Figure 5A, Supplementary Table 6), and the K_{ME} (eigengenes connectivity) value of each gene was calculated. We found that stage A was tightly associated with the MEpurple module ($R = 0.94$ and $p = 2\times10^{-13}$), and stage E with the MEdarkmagenta module ($R = 0.90$ and $p = 3\times10^{-10}$). Next, we annotated the unigenes in these two modules with GO and KEGG, and found that only unigenes in the medarkmagenta module were enriched by KEGG. The unigenes were shown to be enriched in the biosynthetic pathways such as starch and sucrose metabolism, riboflavin metabolism, isoflavonoid biosynthesis, fatty acid elongation, cutin, suberine and wax biosynthesis, flavonoid biosynthesis, carotenoid biosynthesis and photosynthesis (Supplementary Data Figure 10, *p.adjust*<0.04). Therefore, we suspected that the hub genes might be related to the carbohydrate metabolism and the synthesis and transport of endogenous hormones. Therefore, the hub genes could play a connective role to link seed germination and resultant seedling morphogenesis.

Seed dormancy and germination are closely linked to *ICE1, ARF, KAO, GA3ox, GID1, PP2C, YUC, SAUR, PIN, WOX, LEA, Expansin*, and *GOS1*. The MEdarkmagenta module was used to mine the genes connected with *ICE1, ARF, KAO, GA3ox, GID1, PP2C, YUC, SAUR, PIN, WOX, LEA, Expansin, GOS1* were mined (Supplementary Table 7). Twelve algorithms were used to evaluate the connectedness of these genes (Supplementary Table 8). Gene networks were constructed using MCC score rankings. Based on the node scores, the top 120 hub genes were identified in the MEdarkmagenta module (Figure 5B, Supplementary Table 9). In the top 120 hub genes, 107 unigenes were related to endogenous hormones and plant growth and development. The role of the remaining 13 unigenes is unclear. We clarified the functions of hub genes by homologous sequence alignment (TAIR, https://www.arabidopsis.org/Blast/index.jsp), these genes encoded some important proteins involved in endogenous hormone signalling (TTL3, ATGSTU17), transcription factors (bHLH and AHL24), carbohydrate and energy metabolism

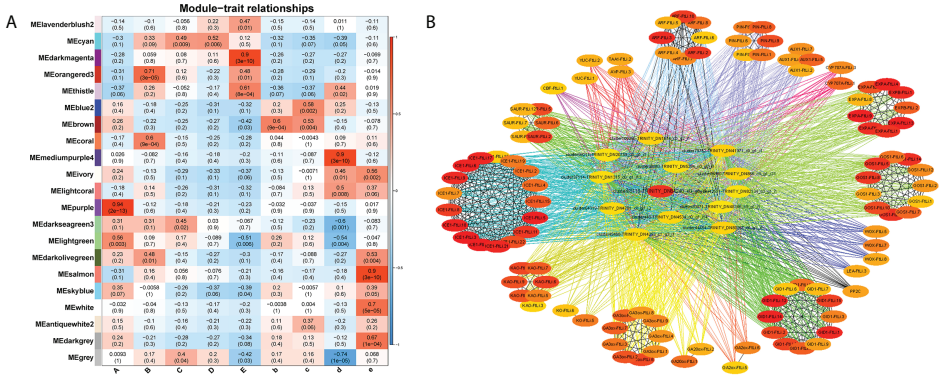


FIGURE 5
WGCNA of DEGs between the 9 development stages. **(A)** Module–trait association of the 9 sample groups (A, B, C, D, E, b, c, d and e). Each row corresponds to a module labeled with a colour. Each column corresponds to one sample group. **(B)** Interaction network analysis of the MEdarkmagenta module. The top 120 hub genes were used to build interaction network diagram in the MEdarkmagenta module. Dots represent hub genes. All details are listed in Supplementary Tables 9, 10.

(AT4G00905, NDPK3 and ARFA1B), the cell cycle, plant cell wall metabolism (PATATIN-LIKE PROTEIN 2 and AT5G02640), chloroplast biosynthesis (PHOTOSYSTEM II BY), and lignin biosynthesis (ATBCB) (Supplementary Table 10). Cold stratification progressively increased the expression levels of these 13 hub genes until the highest level was reached at germination stage E. Whereas at 25°C the expression of these genes was almost unchanged (Supplementary Figures 8).

RNA-seq validation using qRT-PCR

To assess the reliability of the differentially expressed transcripts, 10 candidate unigenes were selected and analyzed using qRT-PCR (primers listed in Supplementary Table 2). These genes are known to function in seed dormancy and germination processes, including GAs pathways (*GA20ox*, *GA3ox*, *GID1*), and ABA pathways (*ZEP*, *NCED*, *SnRK2*$_S$, *PP2C*), as well as auxin biosynthesis (*TAA1*) and response (*ARF*) (Figure 6). These changes in gene expression were consistent with the RNA-seq data. As a result of the cold treatment, *ZEP-FtLi.3*, *NCED-FtLi.1* and *SnRK2s-FtLi.1* gene expression was suppressed, but *PP2C-FtLi.1* expression was promoted. *GA20ox-FtLi.1*, *GA3ox-FtLi.*1, *GID1-FtLi.1*, *TAA1-FtLi.1* and *ARF-FtLi.1* were also positively promoted by 4°C stratification (Figures 6E–G). Furthermore, our transcriptome sequencing results were shown to be reliable.

The verification of transcriptome analysis data with regards to the endogenous hormone regulation pathway

According to the transcriptome data, we found that endogenous hormones regulate seed dormancy and germination. To verify the role of the genes of endogenous hormone regulation pathway, we measured the hormone concentrations in developmental stages A, B, C, D and E (Figure 7, Supplementary Table 11). The bioactive ABA concentrations were shown to sharply decline as a consequence of the stratification at 4°C, with the highest concentration recorded at developmental stage A (3.38 ± 0.77ng/g) and the lowest concentration at stage E (0.27 ± 0.02ng/g). In stages B and C, the concentrations of bioactive ABA were as low as 1.42 ± 0.18ng/g and 1.92 ± 0.09ng/g, respectively, suggesting that cold treatment indeed can promote ABA biodegradation or biological inactivation. It is of interest to note that the content of bioactive

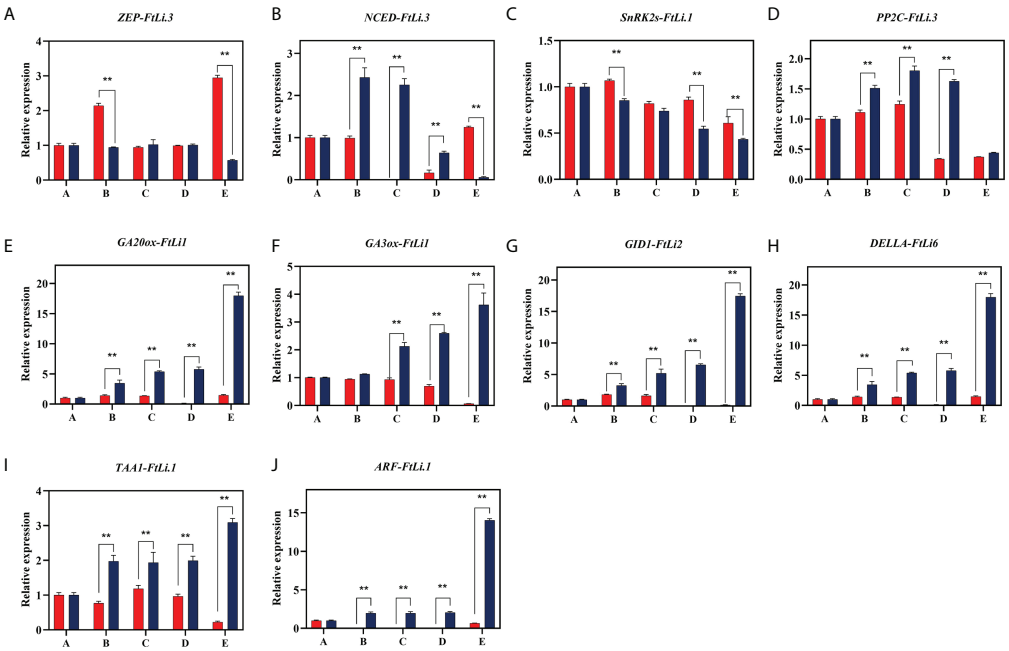


FIGURE 6
The RT-qPCR validation of the expression levels of genes related to endogenous hormones present in *F. taipaiensis*. The values presented are the means + SD of three biological replicates. The transcript IDs and the primers of each gene are listed in Supplementary Tables 2, 12. A(ZEP, zeaxanthin cyclooxygenase), B(NCED,9-cis-cyclooxygenase lyase), C(SnRK2s,Snf1-related protein kinases 2), D(PP2C,PP2C phosphatases), E (GA20ox,GA 20-oxidase), F(GA3ox,GA 3-oxidase), G(GID1,GIBBERELLIN INSENSITIVE DWARF 1), H(DELLA, DELLA protein), I(TAA1,tryptophan aminotransferase 1), J(ARF, AUXIN RESPONSE FACTOR). ** indicate significant difference ($P < 0.01$).

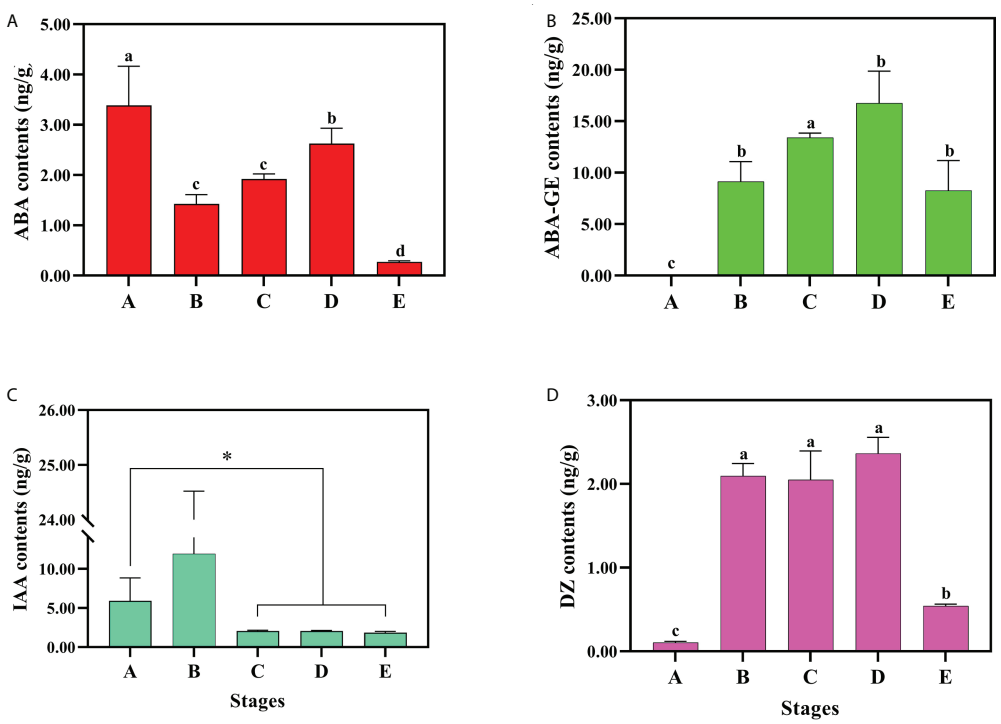


FIGURE 7
The endogenous hormone content There were contents changes of bioactive or inactive **(A)** ABA, **(B)** ABA-GE, **(C)** IAA and **(D)** DZ. Three biological replicates for each sample and analyzed by ANOVA. Values are reported as means ± SE (n = 3). * indicate significant difference ($P < 0.05$).

ABA increased to 2.62 ± 0.30ng/g at developmental stage D. In contrast to ABA, the ABA-GE did not show any significant change in content (Figures 7A, B). Meanwhile, the concentrations of bioactive auxin and cytokinin, that were observed to significantly increase in stage B showed the opposite trend with ABA. As the embryo began to elongate (developmental stage B), the content of IAA markedly increased, but decreased in the following stages (Figure 7C). The level of IAA-amino acid conjugates were significantly changed, the content of IAA-Trp, IAA-Asp and IAA-Glu decreased from stage A to E, but IAA-Glc was shown to increase (Supplementary Figure 9). Among the four bioactive forms of cytokinin, the content of DZ reached 2.0ng/g, and maintained a steady high level throughout developmental stages B, C and D, then fell back down to 0.54 ± 0.02ng/g in stage E (Figure 7C). These results suggest that DZ plays an important role in the process of embryo development.

Discussion

Globally speaking seed dormancy and germination are regulated by a diverse range of external environmental factors (Baskin and Baskin, 2004; Finkelstein et al., 2008; Rajjou et al., 2012). Temperature is a key environmental factor that affects the breaking of seed dormancy and subsequent seed germination (Simon et al., 1976; Yang et al., 2019). The majority of recent research on seed dormancy has mainly focused on the model species Arabidopsis, rice, *Brachypodium* (Hao et al., 2021), *Apium graveolens* (Walker et al., 2021), wheat (Gao et al., 2012), *Euphorbia esula* L. (Foley et al., 2010).

It is known that wheat seed can be released from a physiologically dormant state by after-ripening. Comparative transcriptomic analysis dormant (D) and post mature (AR) seeds in a dry state showed that the response elements of ABA increased during wheat dormancy, and that there was a linkage between wheat dormancy and the sensitivity of seeds to ABA. Studies on the regulation of *Euphorbia esula* seed germination under the alternate temperature of day and night also showed that genes related to abscisic acid signalling were the key regulators of this physiological process.

However, in recent years few reports on the seed dormancy of other plant species have been published. In the current study, we examined the role of cold temperature stratification (at 4°C) in breaking the dormancy of *F. taipaiensis* seed. Transcriptomic analysis was carried out to determine differentially expressed

pathways and genes, qRT-PCR was performed to confirm their reliability, endogenous hormones were quantified, and WGCNA was used to investigate hub genes. Our results provide a comprehensive analysis of the regulatory mechanisms that promote seed germination of *F. taipaiensis* under cold treatment.

Cold temperature is required to break dormancy of *F. taipaiensis* seed

Seed germination is a key and complex process in the plant life cycle, the optimum temperature for which is species specific. Cold stratification is widely used to break dormancy and trigger germination in many species (Kim et al., 2019). Generally, cold stratification to break seed dormancy is most effective between 0 and 10°C (Chen et al., 2020). However, it is still largely unknown as to how cold stratification breaks the dormancy in these species.

F. taipaiensis is a typical alpine plant, which requires precise seed dormancy adjustment to ensure seedling safety. According to Baskin, the type of seed dormancy exhibited by *F. taipaiensis* belongs to the morphological and physiological comprehensive dormancy grouping. Phenotypic analysis has shown that embryos from plants of this species treated at 10 °C develop quite slowly, and those treated at 0°C or 25°C did not develop hardly at all, but seeds placed at 4°C quickly broke their morphological dormancy (Figure 1). In this study we have demonstrated that cold temperature of 4°C plays an important role in the seed germination of *F. taipaiensis*.

To further investigate the molecular mechanism of cold regulation the seed germination of *F. taipaiensis*, RNA-Seq was performed. Altogether 177.09 Gb clean data was obtained, the clean reads were between 6.16-7.56G, base of Q30 percentage was 93% (Supplementary Table 4), and 61,574 unigenes were annotated. The enrichment analysis of GO and KEGG pathways showed that the most significant enriched pathways were "Phenylpropanoid biosynthesis", "Plant hormone signal transduction", "MAPK signaling pathway-plant", "Fructose and mannose metabolism", "Pentose and glucuronate interconversions", "Linoleic acid metabolism" (Figure 3C). Therefore, it was proposed that these pathways were significantly related to seed dormancy breaking in this species.

Hormonal regulation contributes to cold temperature germination of *F. taipaiensis* seed

A great deal of research has been performed related to how various genes function in the biosynthesis and signal transduction of hormones during seed dormancy and germination in both Arabidopsis and rice (Nonogaki, 2014).

From these studies it has been demonstrated that seed dormancy is determined not only by endogenous hormones ABA and GA, but that it is also indirectly influenced by cold temperatures (Finkelstein et al., 2008; Yamaguchi, 2008; Rajjou et al., 2012; Arc et al., 2013; Lv et al., 2021). In the study presented, several genes related to ABA, GA and auxin were shown to be differentially expressed in the process of seed dormancy breaking under 4°C stratification.

The endogenous hormone ABA is an important chemical compound in the induction and maintenance of seed dormancy (Kucera et al., 2005; Nonogaki, 2017). The expression of key enzymes for ABA synthesis *ZEP-FtLi.1* and *NCED-FtLi.4* were shown to be inhibited at 4°C compared to 25°C (Figure 4A, Figures 6, 7), as well as the changing trend of ABA content during seed germination. These results revealed that dynamic changes of ABA levels played an important role in the process of seed germination at the colder temperature (4°C).

ABA signaling networks are initiated and operate through the expression of receptor PYR/PYL genes (Di et al., 2018). According to SK Yadav, PYL genes (*AtPYL/2/4/5/8/9*) and *AtPYL6/13* have different functions, the former promotes seed germination, whereas the latter has the opposite effect (Yadav et al., 2020). In our study, the expression of *PYR-FtLi.1/2* was inhibited by cold temperature stratification, indicating its positive role in the seed dormancy of *F. taipaiensis*. The expression pattern of protein phosphatases (PPS) can also respond to stress (He et al., 2019), which is important for cell homeostasis (Singh et al., 2016). The SnRK2s and PP2C kinases play opposite roles in the ABA signaling pathway. The decreased expression of *SnRK2s-FtLi.1* and the increased expression of *PP2C-FtLi.4* showed that cold temperature inhibited positive ABA signal transduction. ABI5 is a key transcription factor that regulates ABA signaling and inhibits seed germination (Piskurewicz et al., 2008). The expression of *ABI5 (ABI5-FtLi.1, ABI5-FtLi.2, ABI5-FtLi.4)* was induced by higher temperatures in our experiment. These results strongly suggest that cold temperature played a role in the germination of *F. taipaiensis* by inhibiting ABA signaling pathway.

Gibberellins play a role in the promotion of germination (Finkelstein et al., 2008). GA bound to GID1, inhibites DELLA activity and thus promotes seed germination (Voegele et al., 2011; Hauvermale et al., 2015). In the current study, cold stratification of the seed was shown to positively regulate the GA signaling pathway by causing expressions of *GID1-FtLi6/7/8/9* and inhibiting *DELLA-FtLi1/2/3* under 4°C stratification (Figure 4B, Figure 6H). The GA 20-oxidase (GA20ox), GA 3-oxidase (GA3ox) proteins are the key rate-limiting enzymes for bioactive gibberellin synthesis (Sasaki et al., 2002; Appleford et al., 2006), and their expression was shown to be induced by cold, and remained low or even absent at 25°C stratification (Figure 4B). These results indicate a major role of GAs in the cold germination of *F. taipaiensis* seed.

Previous studies have shown that auxin is involved in seed germination (Mohamed et al., 2022). IAA is the main bioactive form of auxin in plants, while *TAA* and *YUCCA* encode the key enzymes involved in the auxin biosynthetic pathway (Stepanova et al., 2008; Won et al., 2011). Both *TAA1-FtLi.1* and *YUCCA-*

FtLi.3/4 were shown to be upregulated during the breaking of seed dormancy in *F. taipaiensis*. Moreover, as the embryo elongated at stage B (Figure 7C), the content of IAA was also observed to increase significantly, thus confirming the transcriptome results. Both *PIN* and *AUX1* were upregulated, encoding auxin efflux carriers and influx transporters under cold stratification. ARF and SAUR are known to play important roles in IAA signal transduction (Park et al., 2007; Pierdonati et al., 2019). *ARF* and *SAUR* expression levels were significantly improved during seed germination. GH3 can convert bioactive IAA to IAA-amino acid (Aoi et al., 2020), and the *GH3-FtLi.1/2/3* gene family were shown to be highly expressed at developmental stage B. The expression model of *GH3* was consistent with the variation of IAA-Glc (Figure 4C, Supplementary Figure 9E). In our data, cold temperature stratification at 4oC was demonstrated to significantly inhibit the expression of *VAS1*, and thus promote auxin catabolism. These results strongly indicate that auxin is an essential regulator of cold germination in seeds of this species.

Hub gene mining of integrated signals

The gene networks MEpurple and MEdarkmagent were identified as being related to seed germination and dormancy, respectively. *HEC3* (AT5G09750) was one of the hub genes in the MEdarkmagen module which can regulate tissue development (Riechmann et al., 2000; Gremski et al., 2007). It can be observed from our data that *HEC3* plays a central role in the regulation of seed germination (Figure 5B). The hub genes *ARFA1B* and *TTL3* are involved in auxin and brassinosteroid mediated signaling pathways (Ceserani et al., 2009). The hub gene *PSBY* was also identified, which is involved in the formation of the photosynthetic system (Sydow et al., 2016). The expression pattern of all the hub genes was similar, in that the levels of expression increased gradually with cold perception, reaching the highest level at the seed germination stage (Supplementary Figure 8). Therefore, it is clear that these hub genes play a key role in the breaking of seed dormancy in *F. taipaiensis* seed at cold temperatures, and that further studies on their molecular function(s) in this process are required.

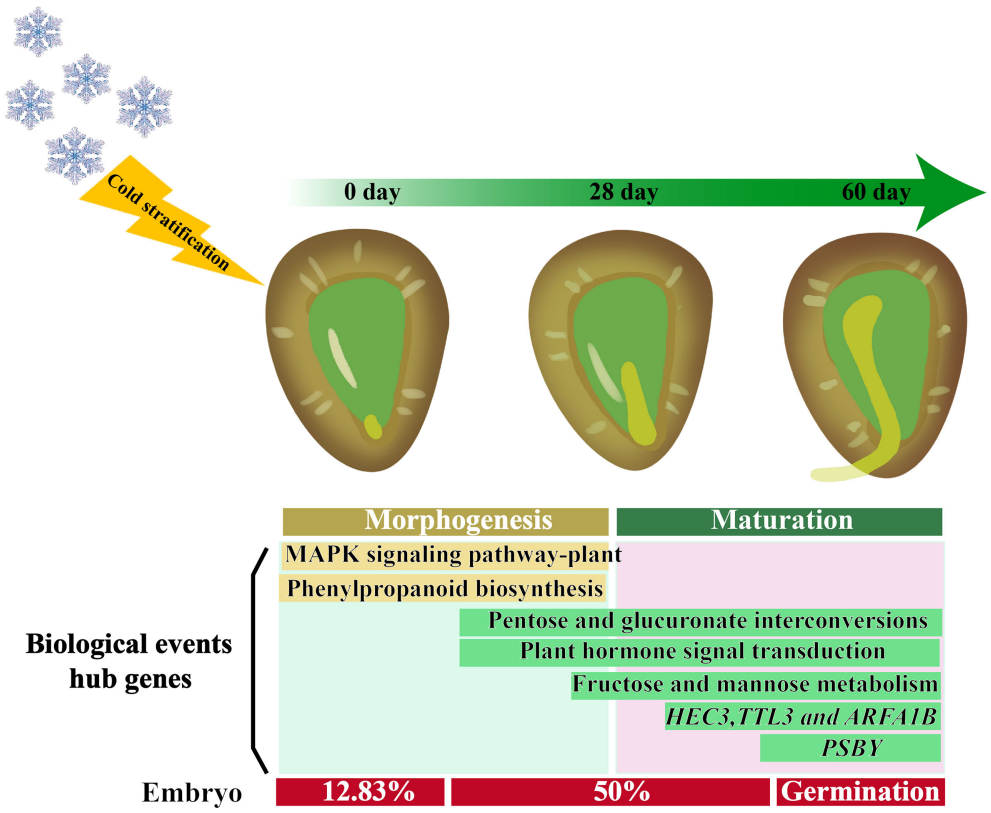


FIGURE 8
Proposed model of dormancy-breaking of *F. taipaiensis* seeds by cold stratification. The dormancy breaking process of *F. taipaiensis* seeds is defined as: When seeds 'sensed' cold stimulus, the pathways "MAPK signaling pathway of plant" and "Phenylpropanoid biosynthesis" in the seeds are quickly responded, and the pathways such as "Pentose and glucuronate interconversions", "Plant hormone signal transduction" and "Fructose and mannose metabolism" are activated successively and throughout the process of dormancy breaking. Besides, the hub genes *HEC3*, *TTL3*, *ARFA1B* and *PSBY* seems to play an important role in the seed germination progress.

Conclusions

The study presented provides an in-depth analysis of the transcriptomes of seeds of *F. taipaiensis* with different dormancy levels. We propose a model of the dormancy-breaking mechanism in *F. taipaiensis* seeds treated with cold temperature stratification which explains some key aspects the molecular mechanism of cold temperature germination in this species (Figure 8). The analyses of key metabolic pathways presented and of the differentially expressed genes, revealed that cold temperature stratification may up regulate GA and auxin pathways, inhibit the ABA pathway, and promote the embryo development and subsequently stimulate seed germination. Based on WGCNA, the hub genes in the core regulatory position were screened. In future experiments, we will investigate the function of these critical genes and identify their roles in regulating embryo expansion. Our research also provides valuable transcriptomic data for seed dormancy of non-model plants, which may enrich our understanding the mechanism of plant embryo expansion and seed germination.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Author contributions

Y-LL and S-CY directed the whole process of this project. They also assisted in the writing of the manuscript. Q-XY participated in the whole project, analyzed experimental data, and helped write the paper. DC conducted the most of experiments. YZ helped with the conception and writing of the manuscript. X-YZ, MZ, RP, N-XS and TB provided help and advice on the experimental design and data analysis and TB proofread the final draft of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by the National Natural Science Foundation of China (31971543) and by the Major Special Science and Technology Project of Yunnan Province (202102AA310031,202101BD070001-008, 202102AA310045, 202005AC160040).

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.1021572/full#supplementary-material

References

Aoi, Y., Tanaka, K., Cook, S. D., Hayashi, K. I., and Kasahara, H. (2020). GH3 auxin-amido synthetases alter the ratio of indole-3-Acetic acid and phenylacetic acid in arabidopsis. *Plant Cell Physiol.* 61 (3), 596–605. doi: 10.1093/pcp/pcz223

Appleford, N. E., Evans, D. J., Lenton, J. R., Gaskin, P., Croker, S. J., Devos, K. M., et al. (2006). Function and transcript analysis of gibberellin-biosynthetic enzymes in wheat. *Planta* 223 (3), 568–582. doi: 10.1007/s00425-005-0104-0

Arc, E., Sechet, J., Corbineau, F., Rajjou, L., and Marion-Poll, A. (2013). ABA crosstalk with ethylene and nitric oxide in seed dormancy and germination. *Front. Plant Sci.* 4. doi: 10.3389/fpls.2013.00063

Barrero, J. M., Millar, A. A., Griffiths, J., Czechowski, T., Scheible, W. R., Udvardi, M., et al. (2010). Gene expression profiling identifies two regulatory genes controlling dormancy and ABA sensitivity in arabidopsis seeds. *Plant J.* 61, 611–622. doi: 10.1111/j.1365-313X.2009.04088.x

Baskin, J. M., and Baskin, C. C. (2004). A classification system for seed dormancy. *Seed Sci. Res.* 14 (1), 1–16. doi: 10.1079/SSR2003150

Carrillo-Barral, N., Rodríguez-Gacio, M. D. C., and Matilla, A. J. (2020). Delay of germination-1 (DOG1): A key to understanding seed dormancy. *Plants* 9 (4), 480. doi: 10.3390/plants9040480

Ceserani, T., Trofka, A., Gandotra, N., and Nelson, T. (2009). VH1/BRL2 receptor-like kinase interacts with vascular-specific adaptor proteins VIT and VIK to influence leaf venation. *Plant J.* 57, 1000–1014. doi: 10.1111/j.1365-313X.2008.03742.x

Chahtane, H., Kim, W., and Lopez-Molina, L. (2017). Primary seed dormancy: a temporally multilayered riddle waiting to be unlocked. *J. Exp. Bot.* 68 (4), 857–869. doi: 10.1093/jxb/erw377

Chen, F., and Bradford, K. J. (2000). Expression of an expansin is associated with endosperm weakening during tomato seed germination. *Plant Physiol.* 124 (3), 1265–1274. doi: 10.1104/pp.124.3.1265

Chen, F., Dahal, P., and Bradford, K. J. (2001). Two tomato expansin genes show divergent expression and localization in embryos during seed development and germination. *Plant Physiol.* 127 (3), 928–936. doi: 10.1104/pp.127.3.928

Chen, D. L., Luo, X. P., Yuan, Z., Bai, M. J., and Hu, X. W. (2020). Seed dormancy release of halenia elliptica in response to stratification temperature, duration and soil moisture content. *BMC Plant Biol.* 20 (1), 1–8. doi: 10.1186/s12870-020-02560-8

Chin, C. H., Chen, S. H., Wu, H. H., Ho, C. W., Ko, M. T., and Lin, C. Y. (2014). CytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst. Biol.* 8 (4), 1–7. doi: 10.1186/1752-0509-8-S4-S11

Cutler, S. R., Rodriguez, P. L., Finkelstein, R. R., and Abrams, S. R. (2010). Abscisic acid: emergence of a core signaling network. *Annu. Rev. Plant Biol.* 61, 651–679. doi: 10.1146/annurev-arplant-042809-112122

Czernik, M., Fidanza, A., Luongo, F. P., Valbonetti, L., Scapolo, P. A., Patrizio, P., et al. (2020). Late embryogenesis abundant (LEA) proteins confer water stress tolerance to mammalian somatic cells. *Cryobiology* 92, 189–196. doi: 10.1016/j.cryobiol.2020.01.009

Denay, G., Creff, A., Moussu, S., Wagnon, P., Thévenin, J., Gérentes, M. F., et al. (2014). Endosperm breakdown in arabidopsis requires heterodimers of the basic helix-loop-helix proteins ZHOUPI and INDUCER OF CBP EXPRESSION 1. *Development* 141 (6), 1222–1227. doi: 10.1242/dev.103531

Di, F., Jian, H., Wang, T., Chen, X., Ding, Y., Du, H., et al. (2018). Genome-wide analysis of the PYL gene family and identification of PYL genes that respond to abiotic stress in brassica napus. *Genes* 9 (3), 156. doi: 10.3390/genes9030156

Doll, N. M., and Ingram, G. C. (2022). Embryo-endosperm interactions. *Annu. Rev. Plant Biol.* 73, 293–321. doi: 10.1146/annurev-arplant-102820-091838

Finch-Savage, W. E., and Footitt, S. (2017). Seed dormancy cycling and the regulation of dormancy mechanisms to time germination in variable field environments. *J. Exp. Bot.* 68 (4), 843–856. doi: 10.1093/jxb/erw477

Finkelstein, R., Reeves, W., Ariizumi, T., and Steber, C. (2008). Molecular aspects of seed dormancy. *Annu. Rev. Plant Biol.* 59 (1), 387–415. doi: 10.1146/annurev.arplant.59.032607.092740

Foley, M. E., Anderson, J. V., Chao, W. S., Doğramacı, M., and Horvath, D. P. (2010). Initial changes in the transcriptome of euphorbia esula seeds induced to germinate with a combination of constant and diurnal alternating temperatures. *Plant Mol. Biol.* 73 (1), 131–142. doi: 10.1007/s11103-009-9569-8

Gao, F., Jordan, M. C., and Ayele, B. T. (2012). Transcriptional programs regulating seed dormancy and its release by after-ripening in common wheat (Triticum aestivum l.). *Plant Biotechnol. J.* 10 (4), 465–476. doi: 10.1111/j.1467-7652.2012.00682.xCitatio

Graeber, K., Nakabayashi, K., Miatton, E., Leubner-Metzger, G., and Soppe, W. J. (2012). Molecular mechanisms of seed dormancy. *Plant Cell Environ.* 35 (10), 1769–1786. doi: 10.1111/j.1365-3040.2012.02542.x

Gremski, K., Ditta, G., and Yanofsky, M. F. (2007). The HECATE genes regulate female reproductive tract development in arabidopsis thaliana. *Development* 134 (20), 3593–3601. doi: 10.1242/dev.011510

Gubler, F., Millar, A. A., and Jacobsen, J. V. (2005). Dormancy release, ABA and pre-harvest sprouting. *Curr. Opin. Plant Biol.* 8 (2), 183–187. doi: 10.1016/j.pbi.2005.01.011

Hao, Z., Zhang, Z., Xiang, D., Venglat, P., Chen, J., Gao, P., et al. (2021). Conserved, divergent and heterochronic gene expression during brachypodium and arabidopsis embryo development. *Plant Reprod.* 34 (3), 207–224. doi: 10.1007/s00497-021-00413-4

Harberd, N. P., Belfield, E., and Yasumura, Y. (2009). The angiosperm gibberellin-GID1-DELLA growth regulatory mechanism: How an "inhibitor of an inhibitor" enables flexible response to fluctuating environments. *Plant Cell.* 21 (5), 1328–1339. doi: 10.1105/tpc.109.066969

Hauvermale, A. L., Tuttle, K. M., Takebayashi, Y., Seo, M., and Steber, C. M. (2015). Loss of arabidopsis thaliana seed dormancy is associated with increased accumulation of the GID1 GA hormone receptors. *Plant Cell Physiol.* 56 (9), 1773–1785. doi: 10.1093/pcp/pcv084

He, Z., Wu, J., Sun, X., and Dai, M. (2019). The maize clade a PP2C phosphatases play critical roles in multiple abiotic stress responses. *Int. J. Mol. Sci.* 20 (14), 3573. doi: 10.3390/ijms20143573

Holdsworth, M. J., Bentsink, L., and Soppe, W. J. J. (2008). Molecular networks regulating arabidopsis seed maturation, after-ripening, dormancy and germination. *New Phytol.* 179 (1), 33–54. doi: 10.1111/j.1469-8137.2008.02437

Hubbard, K. E., Nishimura, N., Hitomi, K., Getzoff, E. D., and Schroeder, J. I. (2010). Early abscisic acid signal transduction mechanisms: newly discovered components and newly emerging questions. *Genes Dev.* 24 (16), 1695–1708. doi: 10.1101/gad

Kanai, M., Nishimura, M., and Hayashi, M. (2010). A peroxisomal ABC transporter promotes seed germination by inducing pectin degradation under the control of ABI5. *Plant J.* 62 (2), 936–947. doi: 10.1111/j.1365-313X.2010.04205.x

Kim, S. Y., Warpeha, K. M., and Huber, S. C. (2019). The brassinosteroid receptor kinase, BRI1, plays a role in seed germination and the release of dormancy by cold stratification. *J. Plant Physiol.* 241, 153031. doi: 10.1016/j.jplph.2019.153031

Kucera, B., Cohn, M. A., and Leubner-Metzger, G. (2005). Plant hormone interactions during seed dormancy release and germination. *Seed Sci. Res.* 15 (4), 281–307. doi: 10.1079/SSR2005218

Langfelder, P., and Horvath, S. (2008). WGCNA: an r package for weighted correlation network analysis. *BMC Bioinf.* 9 (1), 1–13. doi: 10.1186/1471-2105-9-559

Lee, K. H., Piao, H. L., Kim, H., Y., Choi, S. M., Jiang, F., Hartung, W., et al. (2006). Activation of glucosidase *via* stress-induced polymerization rapidly increases active pools of abscisic acid. *Cell* 126 (6), 1109–1120. doi: 10.1016/j.cell.2006.07.034

Liu, S., Kracher, B., Ziegler, J., Birkenbihl, R. P., and Somssich, I. E. (2015). Negative regulation of ABA signaling by WRKY33 is critical for arabidopsis immunity towards botrytis cinerea 2100. *Elife* 4, e07295. doi: 10.7554/eLife.07295

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2– ΔΔCT method. *methods* 25 (4), 402–408. doi: 10.1006/meth.2001.1262

Lv, Y., Pan, J., Wang, H., Reiter, R. J., Li, X., Mou, Z., et al. (2021). Melatonin inhibits seed germination by crosstalk with abscisic acid, gibberellin, and auxin in arabidopsis. *J. Pineal Res.* 70 (4), e12736. doi: 10.1111/jpi.12736

MacGregor, D. R., Zhang, N., Iwasaki, M., Chen, M., Dave, A., Lopez-Molina, L., et al. (2019). ICE1 and ZOU determine the depth of primary seed dormancy in arabidopsis independently of their role in endosperm development. *Plant J.* 98 (2), 277–290. doi: 10.1111/tpj.14211

Mohamed, I. A., Shalby, N., El-Badri, A. M., Batool, M., Wang, C., Wang, Z., et al. (2022). RNA-Seq analysis revealed key genes associated with salt tolerance in rapeseed germination through carbohydrate metabolism, hormone, and MAPK signaling pathways. *Ind. Crops Prod.* 176, 114262. doi: 10.1016/j.indcrop.2021.114262

Muthusamy, M., Kim, J. Y., Yoon, E. K., Kim, J. A., and Lee, S. I. (2020). BrEXLB1, a brassica rapa expansin-like B1 gene is associated with root development, drought stress response, and seed germination. *Genes* 11 (4), 404. doi: 10.3390/genes11040404

Nambara, E., and Marion-Poll, A. (2005). Abscisic acid biosynthesis and catabolism. *Annu. Rev. Plant Biol.* 56, 165–185. doi: 10.1146/annurev.arplant.56.032604.144046

Nishimura, N., Tsuchiya, W., Moresco, J. J., Hayashi, Y., Satoh, K., Kaiwa, N., et al. (2018). Control of seed dormancy and germination by DOG1-AHG1 PP2C phosphatase complex *via* binding to hem. *Nat. Commun.* 9, 2132. doi: 10.1038/s41467-018-04437-9

Nonogaki, H. (2014). Seed dormancy and germination-emerging mechanisms and new hypotheses. *Front. Plant Sci.* 5. doi: 10.3389/fpls.2014.00233

Nonogaki, H. (2017). Seed biology updates - highlights and new discoveries in seed dormancy and germination research. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.00524

Park, J. E., Park, J. Y., Kim, Y. S., Staswick, P. E., Jeon, J., Yun, J., et al. (2007). GH3-mediated auxin homeostasis links growth regulation with stress adaptation response in arabidopsis. *J. Biol. Chem.* 282 (13), 10036–10046. doi: 10.1074/jbc.M610524200

Penfield, S. (2017). Seed dormancy and germination. *Curr. Biol.* 27 (17), R874–R878. doi: 10.1016/j.cub.2017.05.050

Pierdonati, E., Unterholzner, S. J., Salvi, E., Svolacchia, N., Bertolotti, G., Dello Ioio, R., et al. (2019). Cytokinin-dependent control of GH3 group II family genes in the arabidopsis root. *Plants* 8 (4), 94. doi: 10.3390/plants8040094

Piskurewicz, U., Jikumaru, Y., Kinoshita, N., Nambara, E., Kamiya, Y., and Lopez-Molina, L. (2008). The gibberellic acid signaling repressor RGL2 inhibits arabidopsis seed germination by stimulating abscisic acid synthesis and ABI5 activity. *Plant Cell.* 20 (10), 2729–2745. doi: 10.1105/tpc.108.061515

Rajjou, L., Duval, M., Gallardo, K., Catusse, J., Bally, J., Job, C., et al. (2012). Seed germination and vigor. *Annu. Rev. Plant Biol.* 63, 507–533. doi: 10.1146/annurev-arplant-042811-105550

Resentini, F., Felipo-Benavent, A., Colombo, L., Blázquez, M. A., Alabadí, D., and Masiero, S. (2015). TCP14 and TCP15 mediate the promotion of seed germination by gibberellins in arabidopsis thaliana. *Mol. Plant* 8 (3), 482–485. doi: 10.1016/j.molp.2014.11.018

Riechmann, J. L., Heard, J., Martin, G., Reuber, L., Jiang, C., Keddie, J., et al. (2000). Arabidopsis transcription factors: genome-wide comparative analysis

among eukaryotes. *Science* 290 (5499), 2105–2110. doi: 10.1126/science.290.5499.2105

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43 (7), e47. doi: 10.1093/nar/gkv007

Sakamoto, T., Kobayashi, M., Itoh, H., Tagiri, A., Kayano, T., Tanaka, H., et al. (2001). Expression of a gibberellin 2-oxidase gene around the shoot apex is related to phase transition in rice. *Plant Physiol.* 125 (3), 1508–1516. doi: 10.1104/pp.125.3.1508

Sasaki, A., Ashikari, M., Ueguchi-Tanaka, M., Itoh, H., Nishimura, A., Swapan, D., et al. (2002). A mutant gibberellin-synthesis gene in rice. *Nature* 416 (6882), 701–702. doi: 10.1038/416701a

Schomburg, F. M., Bizzell, C. M., Lee, D. J., Zeevaart, J. A., and Amasino, R. M. (2003). Overexpression of a novel class of gibberellin 2-oxidases decreases gibberellin levels and creates dwarf plants. *Plant Cell.* 15 (1), 151–163. doi: 10.1105/tpc

Shih, M. D., and Hsing, Y. I. C. (2008). Late embryogenesis abundant proteins. *Adv. Bot. Res.* 48, 211–255. doi: 10.1016/S0065-2296(08)00404-7

Simon, E. W., Minchin, A., McMenamin, M. M., and Smith, J. M. (1976). The low temperature limit for seed germination. *New Phytol.* 77 (2), 301–311. doi: 10.1111/j.1469-8137.1976.tb01519.x

Simsek, S., Ohm, J. B., Lu, H., Rugg, M., Berzonsky, W., Alamri, M. S., et al. (2014). Effect of pre-harvest sprouting on physicochemical changes of proteins in wheat. *J. Sci. Food Agric.* 94 (2), 205–212. doi: 10.1002/jsfa.6229

Singh, A., Pandey, A., Srivastava, A. K., Tran, L. S., and Pandey, G. K. (2016). Plant protein phosphatases 2C: From genomic diversity to functional multiplicity and importance in stress management. *Crit. Rev. Biotechnol.* 36 (6), 1023–1035. doi: 10.3109/07388551.2015.1083941

Stepanova, A. N., Robertson-Hoyt, J., Yun, J., Benavente, L. M., Xie, D. Y., Dolezal, K., et al. (2008). TAA1-mediated auxin biosynthesis is essential for hormone crosstalk and plant development. *Cell* 133 (1), 177–191. doi: 10.1016/j.cell.2008.01.047

Sydow, L. V., Schwenkert, S., Meurer, J., Funk, C., Mamedov, F., and Schröder, W. P. (2016). The PsbY protein of arabidopsis photosystem II is important for the redox control of cytochrome b559. *Biochim. Biophys. Acta* 1857 (9), 1524–1533. doi: 10.1016/j

Voegele, A., Linkies, A., Müller, K., and Leubner-Metzger, G. (2011). Members of the gibberellin receptor gene family GID1 (GIBBERELLIN INSENSITIVE DWARF1) play distinct roles during lepidium sativum and arabidopsis thaliana seed germination. *J. Exp. Bot.* 62 (14), 5131–5147. doi: 10.1093/jxb/err214

Walker, M., Pérez, M., Steinbrecher, T., Gawthrop, F., Pavlović, I., Novák, O., et al. (2021). Molecular mechanisms and hormonal regulation underpinning morphological dormancy: a case study using apium graveolens (Apiaceae). *Plant J.* 108 (4), 1020–1036. doi: 10.1111/tpj.15489

Wang, W., Gao, T., Chen, J., Yang, J., Huang, H., and Yu, Y. (2018). The late embryogenesis abundant gene family in tea plant (Camellia sinensis): Genome-wide characterization and expression analysis in response to cold and dehydration stress. *Plant Physiol. Biochem.* 135, 277–286. doi: 10.1016/j.plaphy.2018.12.009

Willis, C. G., Baskin, C. C., Baskin, J. M., Auld, J. R., Venable, D. L., Cavender-Bares, J., et al. (2014). The evolution of seed dormancy: environmental cues, evolutionary hubs, and diversification of the seed plants. *New Phytol.* 203 (1), 300–309. doi: 10.1111/nph.12782

Won, C., Shen, X., Mashiguchi, K., Zheng, Z., Dai, X., Cheng, Y., et al. (2011). Conversion of tryptophan to indole-3-acetic acid by TRYPTOPHAN AMINOTRANSFERASES OF ARABIDOPSIS and YUCCAs in arabidopsis. *Proc. Natl. Acad. Sci. U. S. A.* 108 (45), 18518–18523. doi: 10.1073/pnas

Xi, W., Liu, C., Hou, X., and Yu, H. (2010). MOTHER OF FT AND TFL1 regulates seed germination through a negative feedback loop modulating ABA signaling in arabidopsis. *Plant Cell.* 22 (6), 1733–1748. doi: 10.1105/tpc.109.073072

Xu, Z. Y., Lee, K. H., Dong, T., Jeong, J. C., Jin, J. B., Kanno, Y., et al. (2012). A vacuolar β-glucosidase homolog that possesses glucose-conjugated abscisic acid hydrolyzing activity plays an important role in osmotic stress responses in arabidopsis. *Plant Cell.* 24 (5), 2184–2199. doi: 10.1105/tpc.112.095935

Yadav, S. K., Santosh Kumar, V. V., Verma, R. K., Yadav, P., Saroha, A., Wankhede, D. P., et al. (2020). Genome-wide identification and characterization of ABA receptor PYL gene family in rice. *BMC Genomics* 21 (1), 1–27. doi: 10.1186/s12864-020-07083-y

Yamaguchi, S. (2008). Gibberellin metabolism and its regulation. *Annu. Rev. Plant Biol.* 59, 225–251. doi: 10.1146/annurev.arplant.59.032607.092804

Yang, M., Yang, J., Su, L., Sun, K., Li, D., and Liu, Y. (2019). Metabolic profile analysis and identification of key metabolites during rice seed germination under low-temperature stress. *Plant Sci.* 289, 110282. doi: 10.1016/j.plantsci.2019.110282

Yan, A., Wu, M., Yan, L., Hu, R., Ali, I., and Gan, Y. (2014). AtEXP2 is involved in seed germination and abiotic stress response in arabidopsis. *PloS One* 9 (1), e85208. doi: 10.1371/journal.pone.0085208

Yu, L., Shi, D., Li, J., Kong, Y., Yu, Y., and Chai, G. (2014). CELLULOSE SYNTHASE-LIKE A2, a glucomannan synthase, is involved in maintaining adherent mucilage structure in arabidopsis seed. *Plant Physiol.* 164 (4), 1842–1856. doi: 10.1104/pp.114.236596

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 03 November 2022
DOI 10.3389/fpls.2022.994159

OPEN ACCESS

EDITED BY
Neftali Ochoa-Alejo,
Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Mexico

REVIEWED BY
Junfeng Guan,
Hebei Academy of Agriculture and Forestry Sciences (HAAFS), China
Shaimaa Abdelmohsen,
Princess Nourah bint Abdulrahman University, Saudi Arabia

*CORRESPONDENCE
Mehar H. Asif
mh.asif@nbri.res.in

SPECIALTY SECTION
This article was submitted to Plant Development and EvoDevo, a section of the journal Frontiers in Plant Science

RECEIVED 14 July 2022
ACCEPTED 30 September 2022
PUBLISHED 03 November 2022

CITATION
Dhar YV and Asif MH (2022) Genome and transcriptome-wide study of carbamoyltransferase genes in major fleshy fruits: A multi-omics study of evolution and functional significance. *Front. Plant Sci.* 13:994159. doi: 10.3389/fpls.2022.994159

Genome and transcriptome-wide study of carbamoyltransferase genes in major fleshy fruits: A multi-omics study of evolution and functional significance

Yogeshwar V. Dhar[1,2] and Mehar H. Asif[1,2]*

[1]CSIR-National Botanical Research Institute (CSIR-NBRI), Lucknow, India, [2]Academy of Scientific and Innovative Research to Academy of Scientific and Innovative Research (AcSIR), Ghaziabad, India

The carbamoyltransferase or aspartate carbamoyltransferase (ATCase)/ornithine carbamoyltransferase (OTCase) is an evolutionary conserved protein family, which contains two genes, ATCase and OTCase. The ATCase catalyzes the committed step in the synthesis of UMP from which all pyrimidine molecules are synthesized. The second member, OTCase, catalytically regulates the conversion of ornithine to citrulline. This study traces the evolution of the carbomoyltransferase genes in the plant kingdom and their role during fruit ripening in fleshy fruits. These genes are highly conserved throughout the plant kingdom and, except for melon and watermelon, do not show gene expansion in major fleshy fruits. In this study, 393 carbamoyltransferase genes were identified in the plant kingdom, including 30 fleshy fruit representatives. Their detailed phylogeny, evolutionary patterns with their expression during the process of fruit ripening, was analyzed. The ATcase and OTcase genes were conserved throughout the plant kingdom and exhibited lineage-specific signatures. The expression analysis of the ATcase and OTcase genes during fruit development and ripening in climacteric and non-climacteric fruits showed their involvement in fruit ripening irrespective of the type of fruits. No direct role in relation to ethylene-dependent or -independent ripening was identified; however, the co-expression network suggests their involvement in the various ripening processes.

KEYWORDS

carbamoyltransferase, fleshy fruit, fruit ripening, methylation, evolution

Introduction

The carbamoyltransferase or ATCase/OTCase enzymes, member of transcarbamylase protein family (Couchet et al., 2021), are universal in their occurrence, from prokaryotes to eukaryotes (Couchet et al., 2021). They play an important role in an organism's biology by regulating the urea cycle (Smith and Garg, 2017), *de novo* pyrimidine biosynthesis (Schröder et al., 2005), and arginine biosynthesis (Winter et al., 2015; Urbano-Gamez et al., 2020), thus influencing the growth- and development-related processes. Their functional involvement in several important biological processes makes them crucial molecules for agronomic improvement of crop plants. The carbamoyltransferase contains two genes: aspartate carbamoyltransferase (ATCase) and ornithine carbamoyltransferase (OTCase). These two genes share different nucleotide and protein sequences but substantially conserved protein fold/structure. The ATCase genes participate in the enzymatic conversion of aspartate and carbamoyl phosphate into the N-carbamoyl-L-aspartate (CAA). This conversion is the initiation step of the pyrimidine biosynthetic pathway. The second member, OTCase, plays a crucial role in urea cycle by catalyzing the ornithine and carbamoyl phosphate to citrulline. OTCase is also known for its role in arginine biosynthesis and other nitrogenous compound in cytosol (Urbano-Gamez et al., 2020).

The ATCase consists of a catalytic homotrimer with three active sites in between the subunits, which can be allosterically regulated by association with other proteins. Many studies have been done on the biochemical and structural properties of prokaryotic, fungal, and animal ATCases (Wild et al., 1988; Kollman and Doolittle, 2000; Hong et al., 2004; Serre et al., 2004; Rabinowitz et al., 2008). There are limited studies in plants, and recently, the structural and functional analysis of *Arabidopsis* ATCase has been done (Bellin et al., 2021). They have shown a UMP-based regulation of the ATCases exclusively found in plants.

Apart from the functional role, the evolutionary aspect for the functional divergence and specificity in plant carbamoyltransferase genes is not well documented. Here, we identified members of carbamoyltransferase genes from various plant genomes with evolutionary divergence in their signature and focused on the representatives of fleshy fruits, with their transcriptomic and methylomic expression, to understand their behavior and influence during the process of fruit ripening. The stage-specific regulation of ethylene mediated fruit ripening, and its linkage with polyamine pathway is an important aspect of the fruit development process, where the carbamoyltransferase plays a crucial role with their ability to regulate polyamines. The involvement and essential role of polyamines in fruit ripening are well evidenced at physiological and molecular levels, and the regulation of polyamines shows the system's gene regulation for essential steps of fruit ripening in climacteric and non-climacteric fruits through the carbamoyltransferase machinery (Fortes and Agudelo-Romero, 2018; Gao et al., 2021).

Material and methods

Identification of carbamoyltransferase members in plant genomes

The carbamoyltransferase genes were identified by using the *Arabidopsis* and rice sequences as input seeds for major plant genomes used in study of Zhao et al. (Zhao and Schranz, 2019), applying a two-step identification process. In step 1, the BLAST program (Altschul et al., 1990) was used to find the carbamoyltransferase genes in plant genome sequence databases, including the genome-dedicated databases, along with phytozome (https://phytozome-next.jgi.doe.gov/) and plant ensemble (https://plants.ensembl.org/). In step 2, the HMMER profile was constructed, and sequences were searched on the basis of profile HMM (Eddy, 2011). All the retrieved sequences were checked by the SMART database (http://smart.embl-heidelberg.de/) (Letunic et al., 2015) and CDD (Marchler-Bauer et al., 2007) for the presence of conserved ATcase/OTcase domains. Sequences lacking the conserved domain were excluded from further analysis.

Phylogenetic analysis, microsynteny, and conserved motif search

Identified sequences of the ATcase/OTcase proteins from different plant groups were aligned using Clustal X (Thompson et al., 2002), with a gap opening penalty of 10 and a gap extension penalty of 0.1. Phylogenetic relations of ATcase/OTcase genes were analyzed using the IQ-TREE tool (Minh et al., 2020) maximum likelihood method with the JTT model for amino acids with ultrafast bootstrapping value of 1,000 and were considered as the parameters. To identify the signature splitting between two members of ATcase/OTcase, the split network analysis has been performed and visualized using the SplitsTree (Huson and Bryant, 2006) program. This complete process was used for kingdom-wide identified sequences and sequences identified in fruit representatives. The microsynteny analysis has been performed to identify the syntenic clusters between the different groups using the SynNet pipeline (Zhao et al., 2017). To understand the substitution, pressure with their time of divergence was estimated by calculating the synonymous substitution and non-synonymous substitution rate using the kS/kA calculator utility of tbtools (Chen et al., 2020). The speciation time in kS/kA analysis was obtained from timetree.org. To find out the conserved *de novo* sequential motif in ATcase/Otcase sequences, an offline MEME program (Bailey et al., 2009) was employed.

Transcriptome and methylome expression analysis

The transcriptomic meta-analysis was performed for nine different fruits. The expression of ATcase/Otcase genes from six fruits (apple, banana, grape, papaya, melon, and strawberry) were extracted from the GEO file of fruit encode project GSE116581 (Lu et al., 2018). For watermelon and pineapple, the expression values were extracted from their expression databases (http://cucurbitgenomics.org/rnaseq/ home for watermelon and http://pineapple.angiosperms.org/pineapple/html/index.html for pineapple). The *Citrus sinensis* fruit expression was extracted from GSE125726 (Feng et al., 2019). The differential expression (DE) of fpkm values were calculated using the DEseq (Anders and Huber, 2010) package and plotted using the MeV tool (Howe et al., 2011). The details of transcriptomic data used for analysis are mentioned in tabular form below (Table 1).

To understand the epigenetic influence of methylation and related regulation in ATcase genes, the bisulfite sequencing data, available in the fruit encode project (GSE116581), were considered for four representative fruits, namely, apple, banana, grape, and strawberry. For methylation analysis, the epigenetic analysis module of CLC genomics workbench (https://digitalinsights.qiagen.com/) was used with default values. The genomic coordinates were mapped on respective reference genome. The total methylation level with CG, CHG, and CHH contexts were calculated with a window length of 100 and *p*-value of 0.05, along with differently methylated regions using the epigenetic analysis module in CLC genomics workbench.

Gene co-expression network analysis

The relative co-expression profiles of ATcase/OTcase genes, for four fruits, were calculated using Pearson's correlation method implemented in the expression correlation module of the Cytoscape tool for their correlation value (R-value) using the fpkm values. The identified ATcase genes were used as bait genes for identification of co-expressing genes. The co-expression network was visualized and analyzed in the Cytoscape tool (Shannon et al., 2003).

Protein structural analysis

To understand and compare the structural features of ATcase protein, the protein structural modeling was performed using the Phyre2 tool (Kelley et al., 2015). The five plant representative sequences from algae (*Gonium pectorale*), bryophyte (*Marchantia polymorpha*), pteridophyte (*Selaginella moellendorffii*), gymnosperm (*Thuja plicata*), and angiosperm (*Arabidopsis thaliana*) were considered for structural modeling. The bacterial protein model was downloaded from the RCSB-PDB database (www.rcsb.org/) with ID number 5vmq.

Results and discussion

Identification of aspartate carbamoyltransferase/ornithine carbamoyltransferase proteins in plant groups

A total of 520 ATcase/OTcase genes (plant kingdom) were retrieved from the different plant groups, mentioned in Zhao et al. (Zhao and Schranz, 2019). The sequences were identified using blast and profile HMM method, using *Arabidopsis* and rice ATcase/OTcase sequences as seed. After removal of partial and redundant sequences, 393 sequences in total were considered as final sequences (Supplementary Table S1). From these sequences, 98 full-length sequences (Supplementary Table S2) with complete domain coverage in major plant species were considered as representative sequences and were used for further analysis. The occurrence of these genes in plant groups clearly indicated toward their conserved nature, as we observe two members of ATcase/OTcase in maximum plant groups and very few plant representatives showed presence of more carbamoyltransferase genes. In major fruit representatives (Table 2), the highest number was observed in melon (11 members) and watermelon (4 members), whereas apple, banana, grape, etc., showed two to three members only. The increased members could be due to either the whole genome duplication events or the expansion because of duplication. In a study in prokaryotic carbamoyltransferase, there was evidence of duplication of the genes and loss after duplication, thus maintaining the gene number. This could have also happened in plants where multiple rounds of whole genome duplications have taken place but the number of the genes is limited to two or three only.

Phylogenetic evolutionary analysis and structural comparison of aspartate carbamoyltransferase/ornithine carbamoyltransferase members

To study the phylogenetic relationship and evolutionary gains, the total identified 393 sequences and 98 representative sequences were aligned and subjected for maximum likelihood-based phylogeny using the IQ-TREE tool with best-fit model JTT + R6, chosen according to the Bayesian information criterion. The

TABLE 1 The fruit stage and sample wise detail of transcriptomic data, used for expression analysis.

Plant	GEO/data accession	Sample	Stage
Apple	GSE116581 (Fruit Encode Project)	Pulp	Unripe and ripe
Banana	GSE116581 (Fruit Encode Project)	Peel and pulp	Unripe green and ripe yellow
Grape	GSE116581 (Fruit Encode Project)	Fruit flesh	Immature green and ripe; Green and breaker
Papaya	GSE116581 (Fruit Encode Project)	Pulp	30 DPA and 150 DPA
Melon	GSE116581 (Fruit Encode Project)	Fruit flesh	10 DPA and 30 DPA
Strawberry	GSE116581 (Fruit Encode Project)	Receptacles	Unripe and ripe
Orange	GSE125726	Juice sac	Stage 1 green, stage 3 break, stage 6 ripe
Watermelon	SRP012849	Fruit flesh	10 DPA and 34 DPA
Pineapple	Pineapple genome database	Fruit	Stage 1 green, stage 3 golden breaker, stage 5 ripe golden

kingdom-wide phylogenetic tree (Supplementary Figure S1) showed a clear division between the ATCase and OTCase sequence signatures. This division defines the evolutionary conserved pattern of sequence signatures, where they retain the specific signature probably because of their functional diversity. To understand the division in signatures, the split phylogeny network analysis was performed using the SplitsTree4 tool. The obtained kingdom-wide phylogenetic network (Supplementary Figure S2) revealed the clear split between two different signatures and also showed the subsplit in both signatures rising from the basal signatures for both ATCase and OTCase. In both methods of phylogeny, tree construction and phylo-network, it was observed that there were different subbranches, following a mixed dicot–monocot pattern and a no climacteric versus non-climacteric pattern. To identify the pattern of similar signatures, the microsynteny analysis was performed using the SynNet pipeline. Interestingly, the microsynteny analysis (Figure 1) revealed the lineage-specific gene cluster formation in selected genes. This pattern of cluster formation indicates toward the shared lineage-specific signature. The lineage-specific pattern of shared sequence signatures reflects the plant clade–dependent evolutionary gain with maintaining the functional domain signature. A similar phylogenetic tree and split network (Figure 2) method was applied on 99 representative sequences, and similar representation was observed, where a distinct division was visible in ATCase and OTCase clades, supported and stabilized by mixed group subclades. This entire phylogeny and network analysis showed that both types of carbamoyltransferase are evolutionary-conserved and have lineage-specific sequence signatures, which have conserved amino acids to define sequence identity and variable amino acids for possible evolutionary and functional adaptation. The phylogeny of representative members reconfirmed that the carbamoyltransferase group maintained the functional identity in terms of domain composition and organization, with conserved amino acid sequence and gaining other features with variable amino acids. In microsynteny analysis, it was expected to get dicot- and monocot-specific clusters, but the clusters obtained were of mixed group and divided in shared sequence blocks according to aspartate and ornithine class. This shows that functional conservation of carbamoyltransferases are not affected by the dicot–monocot divergence. The fruit group members showed a similar observation, where no specific cluster was observed with higher number of fruit members or climacteric and non-climacteric gene clusters. To understand the evolutionary emergence of carbamoyltransferases in fruit groups, the sequence substitution analysis for synonymous (kS) and non-synonymous (kA) sites was performed (Figure 3). Initially, the substitution analysis was designed with *M. polymorpha* and *S. moellendorffii* as the evolutionary standard, but because of the highly fluctuating substitution divergence score, they were removed, and *Arabidopsis* was included as the evolutionary standard. In the fruit group, two climacteric fruits, apple and banana, and two non-climacteric fruits, grape and strawberry, were considered as fruit representative. The kS plot (Figure 3) was plotted with kS score against the gene pairs with speciation time of each fruit; the speciation time was obtained from timetree.org. The kS plot revealed that a significant amount of gene pairs of fruit groups emerges after the speciation time of selected fruits. The kS analysis also showed that the kS value is converged, which reflects the conserved pattern of substitution site selection. As all the analytical methods were indicating toward the conserved sequence signature, the conserved motif search was performed using MEME tool. The motif search result showed a 29-aa long motif (Figure 4), with conserve (M, F, P, S, R, T, R, S, F, E) and variable aa. The

TABLE 2 The fruits selected for study and number of carbamoyltransferase genes identified in them.

Fruit	Gene
Apple	3
Banana	2
Melon	11
Strawberry	2
Papaye	2
Citrus	2
Pineapple	2
Watermelon	4
Grape	2

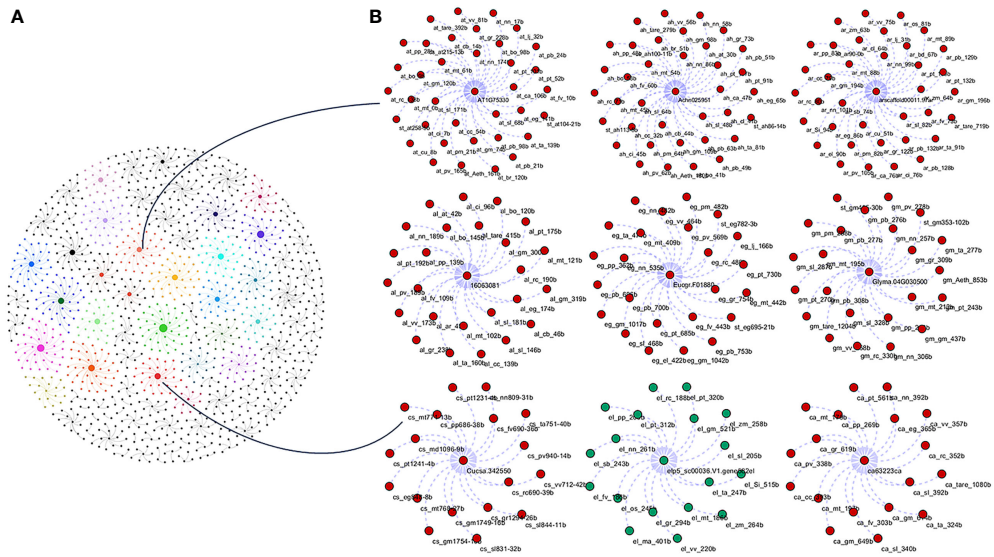


FIGURE 1
Kingdom-wide microsyntenic network of carbamoyltransferase genes, showing the major syntenic hubs in plant groups for aspartic and ornithine carbamoyltransferase genes. **(A)** The kingdom-wide microsyntenic network in globe net layout, where different colors represent different lineage-specific groups and node size represents connected syntenic genes. **(B)** The zoomed representation of lineage-specific groups showing major hubs, where the red color is for dicots and the green color is for monocot nodes.

reconfirmation phylogeny was constructed using the motif sequence, and again, the similar result was obtained in both (phylogeny and split network). This clearly shows the highly conserve nature of carbamoyltransferase genes where specific aa are involved in maintaining the functional identity, whereas other aa are replaced in accordance with evolutionary adaptation. The motif phylogeny also showed the clear plant group–based subclade formation. Thus, the ATCase and OTCase genes have evolved throughout the plant kingdom with minimal sequence divergence.

To understand the functional impact of conserved sequential evolution on the structures of ATcase/OTcase proteins, the bacterial protein model of ATcase (*Escherichia coli*, PDB ID:

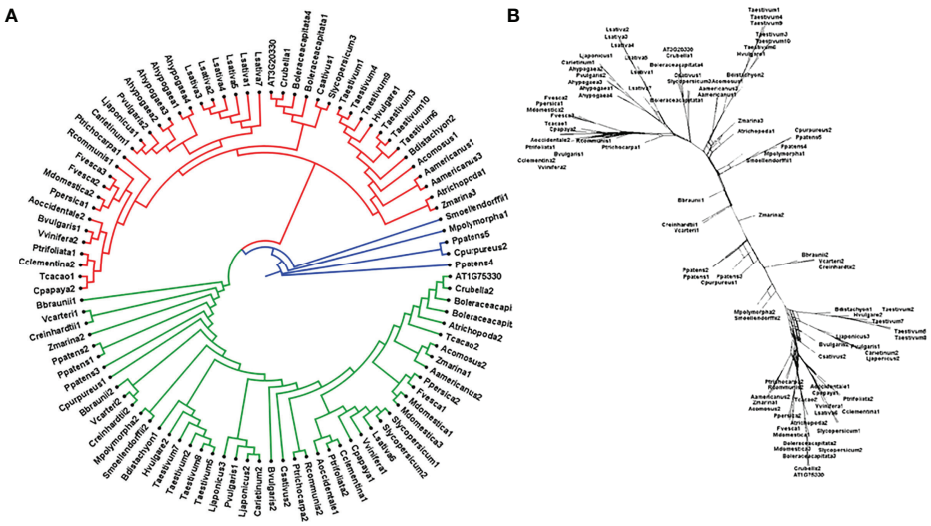


FIGURE 2
Phylogenetic **(A)** tree of carbamoyltransferase genes in plant groups including fruit representative members, where the red clade is aspartate carbamoyltransferase and the green clade is ornithine carbamoyltransferase. Basal groups are represented by the blue color. **(B)** The sequence signature splitting is represented by the phylogenetic split network.

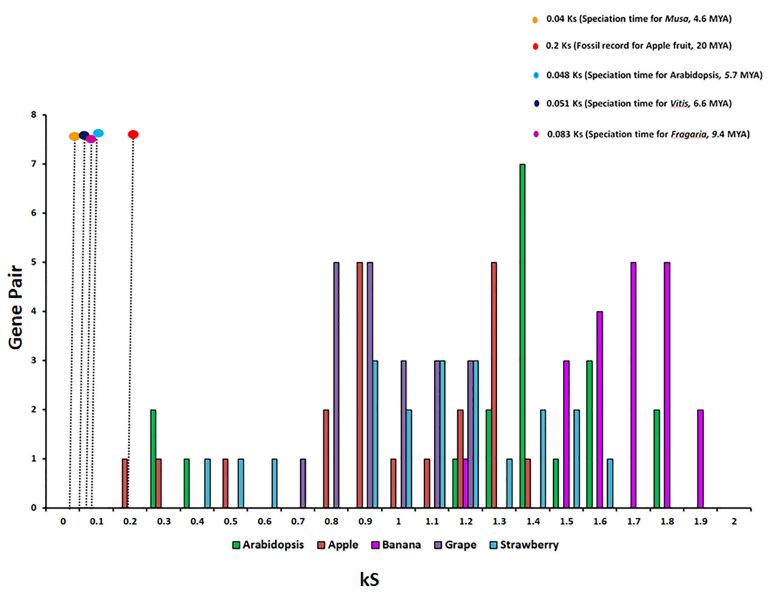


FIGURE 3
The relative substitution plot of kS score of gene pairs of fruit carbamoyltransferase genes, representing the synonymous substitution sites with speciation time of different fruits and *Arabidopsis*. The speciation time has been taken from timetree.org.

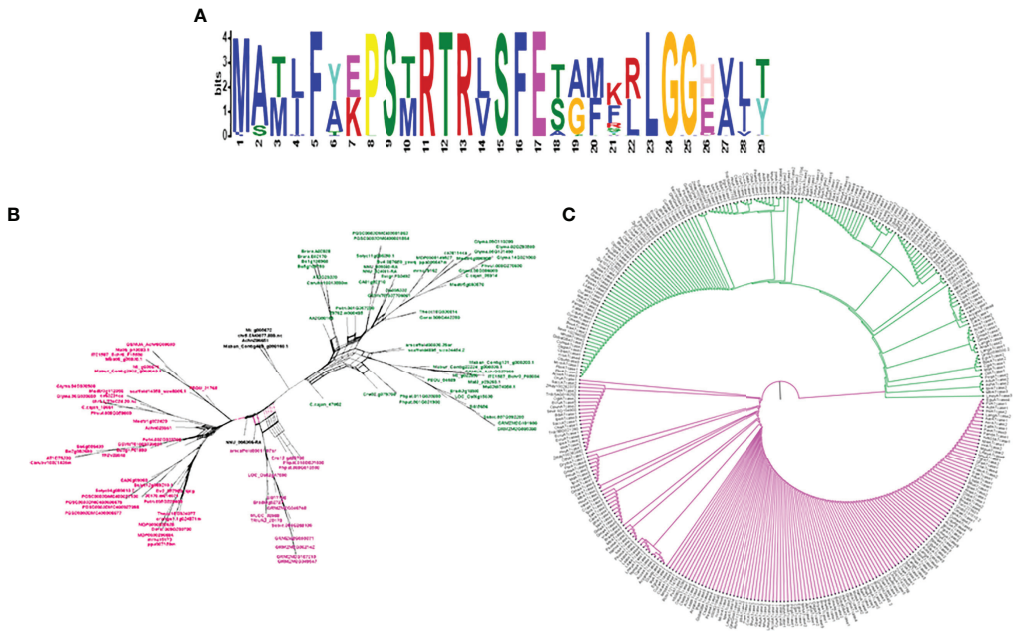


FIGURE 4
The logo representation of conserved motif **(A)** in carbamoyltransferase domain of plant groups and phylogeny **(C)** and network **(B)** of the same, representing the signature distribution of 2 different classes of carbamoyltransferase: aspartate carbamoyltransferase (Pink) and ornithine carbamoyltransferase (Green).

5vmq) was compared with five plant representative protein models of algae (*G. pectorale*), bryophyte (*M. polymorpha*), pteridophyte (*S. moellendorffii*), gymnosperm (*T. plicata*), and angiosperm (*A. thaliana*). The sequential comparison (Supplementary Figure S3) between the above-mentioned models showed significant variation in amino acid sequences with clear differences in bacterial conserved amino acids and plant conserved amino acids. The structural comparison (Supplementary Figure S4) shows highly conserved protein structure throughout the plant representatives. The comparison between bacterial and plant protein showed a significant structural enhancement, where the small helical structure of bacterial protein was replaced by the loop structure in plant protein models for functional adaptation. All the plant ATcases studied were highly conserved in sequence and structure.

Transcriptomic expression and methylation status of aspartate carbamoyltransferase/ornithine carbamoyltransferase genes during fruit development and ripening

To study the expression of ATcase/OTcase genes, the geo profiles of transcriptomic data from fruit encode initiative were downloaded for different fruit developmental stages. The DE was considered from the log2 of fpkm values using the DEseq package and plotted as heat map. The heat map (Figure 5) showed a contrasting pattern of expression where one half of genes were showing a upregulation while the other half were showing downregulation. In most of the cases, the ATCase genes

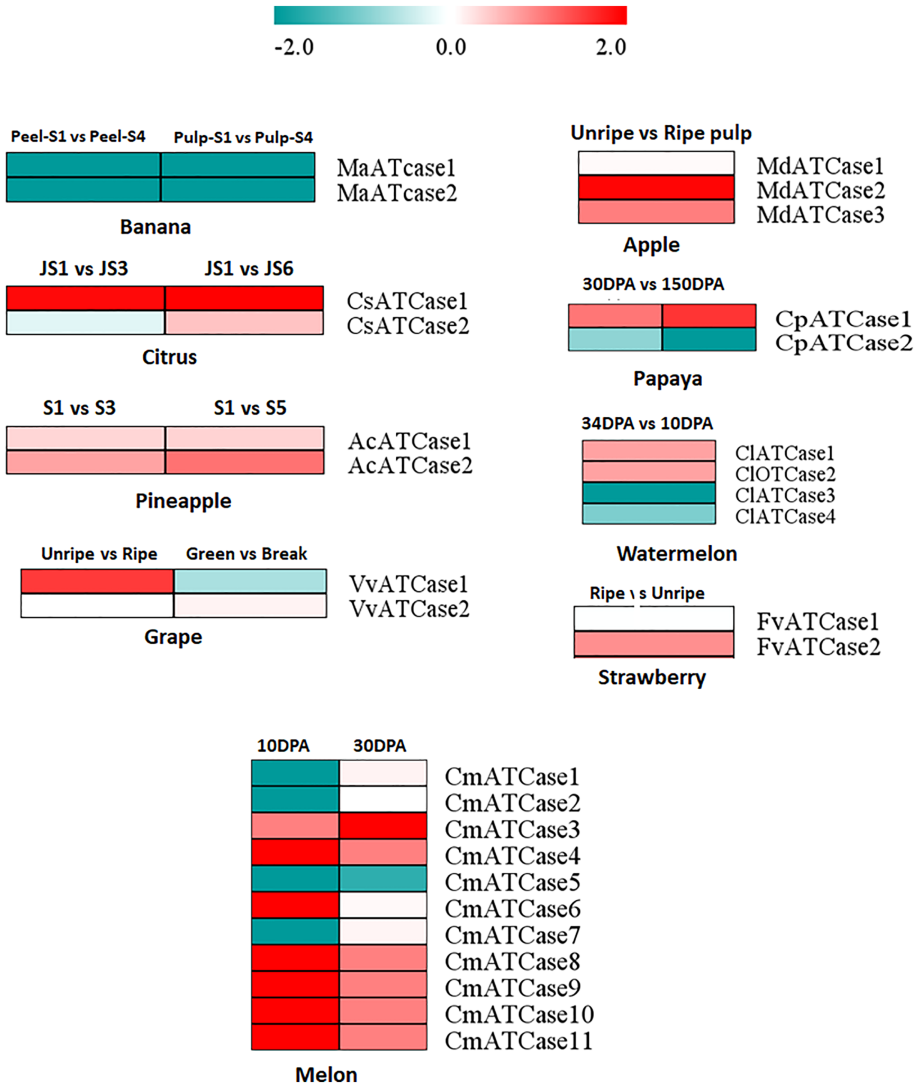


FIGURE 5
The heat map representation of the differential expression profile at the log2 fold change scale of carbamoyltransferase genes in different climacteric and non-climacteric fruits in different developmental and ripening conditions.

were upregulated, whereas the OTCase genes were downregulated. In case of banana, both genes showed downregulation. The expression analysis indicated the positive involvement of ATCase genes in fruit development and ripening without climacteric and non-climacteric specific expression grouping (Joshi et al., 2019).

To explore the role of cytosine methylation and related changes, the bs-seq data, provided by a fruitENCODE initiative, was used and mapped against the genomic coordinates. The obtained mapping result was further processed for the methylation status and methylation context for the ATcase/OTcase genes. The identified methylation context (CpG, CHG, and CHH) with significant total methylation levels was calculated for the unripe and ripe conditions using a CLC genomics workbench with default parameters. The methylation contexts and values were represented by a bar plot, along with the transcriptomic expression, to understand the correlation between them, if any (Figure 6). In methylation analysis, it was observed that, in case of climacteric representatives, banana (Figure 6) and apple (Figure 6), the CG methylation level was lower in the ripe condition, whereas in case of non-climacteric representatives, grape (Figure 6) and strawberry (Figure 6), the CG methylation level was higher during the ripe condition. Another interesting observation revealed that, in case of apple and grape, the CHH methylation level is much higher in the ripe condition in comparison with the unripe condition, with banana and strawberry. Though there are previous reports in tomato and citrus (Huang et al., 2019; Li et al., 2022; Liu et al., 2022), which showed the active involvement and global effect of methylation during fruit developmental stages and ripening, our analysis partially matched with them in case of banana apple and grape. In case of banana (Figure 6), the CG methylation level was equally higher for both genes in the unripe and ripe conditions in pulp and their transcriptomic expression was down, whereas in case of apple (Figure 6), again the overall CG methylation level was high and with a slight difference between the unripe and ripe conditions; the transcriptomic expression was upregulated. As mentioned earlier, that in case of apple and grape the CHH methylation level was higher, it is possible that

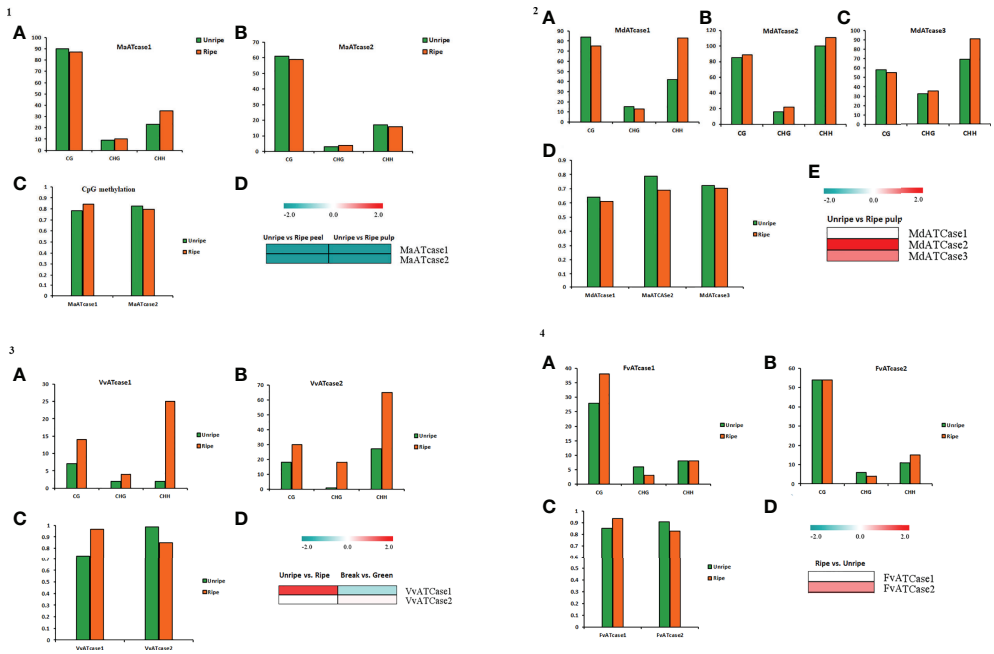


FIGURE 6
The panel representation for the status of CG, CHG, and CHH methylation sites in ATcase genes of banana (panel 1), apple (panel 2), grape (panel 3), and strawberry (panel 4). Status of CG, CHG, and CHH methylation sites in ATcase genes of banana, represented by bar plots **(A, B)** in unripe and ripe pulp conditions. Level of CpG methylation in banana Atcase genes **(C)**, represented by a bar plot, and compared with log2 fold change expression, represented by a heat map **(D)**. Status of CG, CHG, and CHH methylation sites in ATcase genes of apple, represented by bar plots **(A–C)** in unripe and ripe pulp conditions. Level of CpG methylation in banana Atcase genes **(D)**, represented by a bar plot, and compared with log2 fold change expression, represented by a heat map **(E)**. Status of CG, CHG, and CHH methylation sites in ATcase genes of grape, represented by bar plots **(A, B)** in unripe and ripe pulp conditions. Level of CpG methylation in grape Atcase genes **(C)**, represented by a bar plot, and compared with log2 fold change expression, represented by a heat map **(D)**. Status of CG, CHG, and CHH methylation sites in ATcase genes of strawberry, represented by bar plots **(A, B)** in unripe and ripe pulp conditions. Level of CpG methylation in strawberry Atcase genes **(C)**, represented by a bar plot, and compared with log2-fold change expression, represented by a heat map **(D)**.

the CHH methylation level is balancing or accompanying the CH levels and actively regulating the fruit ripening, which was earlier observed in case of pepper fruit (Xiao et al., 2020), citrus (Huang et al., 2019), and tomato (Liu et al., 2022). The methylation status of grape (Figure 6) showed a higher CG and CHH context, similar to apple, while the CG methylation level was higher in the VvATcase1 gene during the ripe condition, and the same gene was upregulated during the ripe condition; the other gene, VvATcase2, showed a lower CG level in the ripe condition and a very small increase in expression during the ripe condition. In case of strawberry, the methylation status of the CG context was much higher in comparison with the rest of the two contexts. The CG methylation level of the FvATcase1 gene (Figure 6) was higher in the ripe condition and transcriptomic expression was very low, whereas in case of FvATcase2, the negatively correlated pattern was observed where CG level was lower in the ripe condition and trancriptomic expression was positively higher. Overall, the methylation status of ATcase/OTcase genes during fruit development and ripening was partially following the patterns, in accordance with previous reports (Gallusci et al., 2016; Tang et al., 2020; Jia et al., 2022; Li et al., 2022).

Co-expression network during ripening

The co-expressing genes during ripening were identified by calculating the expression correlation matrix. The result was plotted in a force-directed layout using the correlation value as fusion value for the significant nodes, and ATcase genes were represented by the contrasting color balls. The co-expression network was calculated for the climacteric (banana and apple) and non-climacteric (grape and strawberry) representatives (Figure 7). The co-expression network were assumed to reflect the difference of climacteric and non-climacteric machinery in fruit ripening, but the retrieved co-expression network showed a conserved gene network pattern.

The co-expression network reflected the conserved functional nature of the carbamoyltransferase group in fruit ripening. The genes that participate in different stages and steps of fruit ripening were observed in co-expression networks of banana, apple, grape, and strawberry. Genes like expansin (Jiang et al., 2019), pectate lyase (Uluisik and Seymour, 2020), metallothionine (Liu et al., 2002), methyltransferase, beta-glucosidase, hydrolase (Yao et al., 2014), and alcohol dehydrogenase (Speirs et al., 1998) were observed in

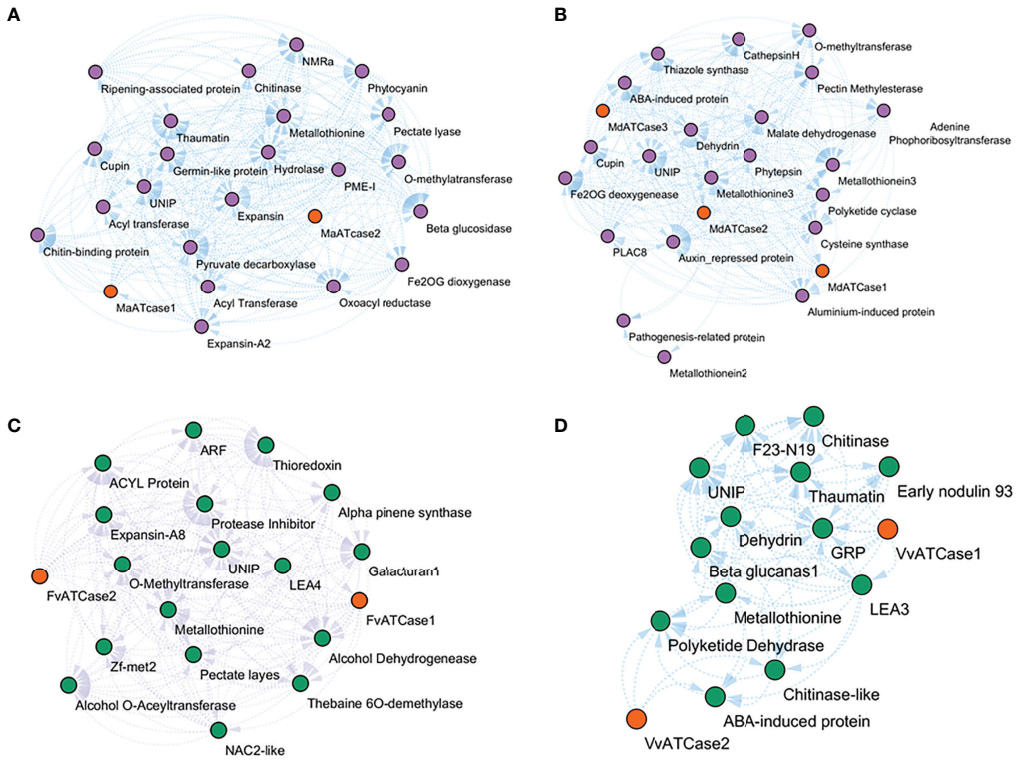


FIGURE 7
Relative co-expression network of aspartate and ornithine carbamoyltransferase genes during fruit ripening condition in banana **(A)**, apple **(B)**, grape **(C)**, and strawberry **(D)**.

correlation with ATcase/OTcase genes. This shows that the basic regulatory involvement of carbamoyltransferases has a potential role in fruit ripening. Also the ABA-induced protein LEA (Jeon et al., 2006) and dehydrin-like genes were observed, which might be an indication of the ABA-mediated regulation of fruit ripening and thus are co-expressed. No gene of polyamine pathway during ripening was observed in the co-expression network. Overall, the co-expression network revealed the conserved functional network of ATcase/OTcase genes during fruit development and ripening.

Conclusion

The carbamoyltransferase genes were identified from the large number of plant genomes focusing on fruit ripening. Though the members showed variation in amino acid sites, the main functional sequence signature was highly conserved for both types of representatives, along with their phylogenetic signature splitting and conserved motif. The microsynteny analysis showed the lineage-specific conservation of these genes. The expression of these genes showed, especially the aspartate carbamoyltransferase, positive regulation during fruit ripening, along with the influence of methylation and conserved co-expressing gene network. The carbamoyltransferase genes play a crucial role in various biological processes; thus, this study is important in understanding their evolutionary conserved and related functional features.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Author contributions

YD and MA conceptualized the study. YD collected the data and carried out the analysis. MA supervised the study. YD and MA wrote the manuscript. All authors contributed to the article and approved the submitted version.

Funding

The Council of Scientific and Industrial Research, New Delhi, is acknowledged for their financial support (GAP 3441 and MLP006).

Acknowledgments

The authors also acknowledge the CSIR-4PI Institute for providing the supercomputing facility. Manuscript number: CSIR-NBRI_MS/2022/09/01.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.994159/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Kingdom-wide phylogenetic tree of carbamoyltransferase genes, showing the major clades in plants groups for aspartic carbamoyltransferase (Pink) and ornithine carbamoyltransferase genes (Green). The basal groups are in represented by blue colour.

SUPPLEMENTARY FIGURE 2
Kingdom-wide phylogenetic network splitting of carbamoyltransferase genes, showing the major sequential splits in plants groups for aspartate carbamoyltransferase (Red and pink) and ornithine carbamoyltransferase genes (Green and blue). The basal plant sequence signatures are located in center.

SUPPLEMENTARY FIGURE 3
Multiple sequence alignment between the bacterial, algal, bryophyte, pteridophyte, gymnosperm and angiospermic protein model representatives, for conserve amino acid pattern.

SUPPLEMENTARY FIGURE 4
The tile representation to compare the protein structures of the bacterial, algal, bryophyte, pteridophyte, gymnosperm and angiospermic protein model representatives.

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biol.* 11, R106. doi: 10.1186/gb-2010-11-10-r106

Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., et al. (2009). MEME SUITE: Tools for motif discovery and searching. *Nucleic Acids Res.* 37, W202–W208. doi: 10.1093/nar/gkp335

Bellin, L., Del Cano-Ochoa, F., Velazquez-Campoy, A., Mohlmann, T., and Ramon-Maiques, S. (2021). Mechanisms of feedback inhibition and sequential firing of active sites in plant aspartate transcarbamoylase. *Nat. Commun.* 12, 947. doi: 10.1038/s41467-021-21165-9

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., et al. (2020). TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* 13, 1194–1202. doi: 10.1016/j.molp.2020.06.009

Couchet, M., Breuillard, C., Corne, C., Rendu, J., Morio, B., Schlattner, U., et al. (2021). Ornithine transcarbamylase - from structure to metabolism: An update. *Front. Physiol.* 12, 748249. doi: 10.3389/fphys.2021.748249

Eddy, S. R. (2011). Accelerated profile HMM searches. *PloS Comput. Biol.* 7, e1002195. doi: 10.1371/journal.pcbi.1002195

Feng, G., Wu, J., and Yi, H. (2019). Global tissue-specific transcriptome analysis of citrus sinensis fruit across six developmental stages. *Sci. Data* 6, 153. doi: 10.1038/s41597-019-0162-y

Fortes, A. M., and Agudelo-Romero, P. (2018). Polyamine metabolism in climacteric and non-climacteric fruit ripening. *Methods Mol. Biol.* 1694, 433–447. doi: 10.1007/978-1-4939-7398-9_36

Gallusci, P., Hodgman, C., Teyssier, E., and Seymour, G. B. (2016). DNA Methylation and chromatin regulation during fleshy fruit development and ripening. *Front. Plant Sci.* 7, 807. doi: 10.3389/fpls.2016.00807

Gao, F., Mei, X., Li, Y., Guo, J., and Shen, Y. (2021). Update on the roles of polyamines in fleshy fruit ripening, senescence, and quality. *Front. Plant Sci.* 12, 610313. doi: 10.3389/fpls.2021.610313

Hong, Y. S., Lee, D., Kim, W., Jeong, J. K., Kim, C. G., Sohng, J. K., et al. (2004). Inactivation of the carbamoyltransferase gene refines post-polyketide synthase modification steps in the biosynthesis of the antitumor agent geldanamycin. *J. Am. Chem. Soc.* 126, 11142–11143. doi: 10.1021/ja047769m

Howe, E. A., Sinha, R., Schlauch, D., and Quackenbush, J. (2011). RNA-Seq analysis in MeV. *Bioinformatics* 27, 3209–3210. doi: 10.1093/bioinformatics/btr490

Huang, H., Liu, R., Niu, Q., Tang, K., Zhang, B., Zhang, H., et al. (2019). Global increase in DNA methylation during orange fruit development and ripening. *Proc. Natl. Acad. Sci. U.S.A.* 116, 1430–1436. doi: 10.1073/pnas.1815441116

Huson, D. H., and Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* 23, 254–267. doi: 10.1093/molbev/msj030

Jeon, O.-S., Kim, C.-S., Lee, S.-P., Kang, S. K., Kim, C.-M., Kang, B.-G., et al. (2006). Fruit ripening-related expression of a gene encoding group 5 late embryogenesis abundant protein inCitrus. *J. Plant Biol.* 49, 403. doi: 10.1007/BF03178819

Jia, H., Jia, H., Lu, S., Zhang, Z., Su, Z., Sadeghnezhad, E., et al. (2022). DNA And histone methylation regulates different types of fruit ripening by transcriptome and proteome analyses. *J. Agric. Food Chem.* 70, 3541–3556. doi: 10.1021/acs.jafc.1c06391

Jiang, F., Lopez, A., Jeon, S., De Freitas, S. T., Yu, Q., Wu, Z., et al. (2019). Disassembly of the fruit cell wall by the ripening-associated polygalacturonase and expansin influences tomato cracking. *Hortic. Res.* 6, 17. doi: 10.1038/s41438-018-0105-3

Joshi, V., Joshi, M., Silwal, D., Noonan, K., Rodriguez, S., and Penalosa, A. (2019). Systematized biosynthesis and catabolism regulate citrulline accumulation in watermelon. *Phytochemistry* 162, 129–140. doi: 10.1016/j.phytochem.2019.03.003

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* 10, 845–858. doi: 10.1038/nprot.2015.053

Kollman, J. M., and Doolittle, R. F. (2000). Determining the relative rates of change for prokaryotic and eukaryotic proteins with anciently duplicated paralogs. *J. Mol. Evol.* 51, 173–181. doi: 10.1007/s002390010078

Letunic, I., Doerks, T., and Bork, P. (2015). SMART: recent updates, new developments and status in 2015. *Nucleic Acids Res.* 43, D257–D260. doi: 10.1093/nar/gku949

Liu, P., Goh, C.-J., Loh, C.-S., and Pua, E. C. (2002). Differential expression and characterization of three metallothionein-like genes in Cavendish banana (Musa acuminata). *Physiol. Plant.* 114 (2), 241–250. doi: 10.1034/j.1399-3054.2002.1140210.x

Liu, Z., Wu, X., Liu, H., Zhang, M., and Liao, W. (2022). DNA Methylation in tomato fruit ripening. *Physiol. Plant* 174, e13627. doi: 10.1111/ppl.13627

Li, X., Wang, X., Zhang, Y., Zhang, A., and You, C. X. (2022). Regulation of fleshy fruit ripening: From transcription factors to epigenetic modifications. *Hortic. Res.* 9. doi: 10.1093/hr/uhac013

Lu, P., Yu, S., Zhu, N., Chen, Y. R., Zhou, B., Pan, Y., et al. (2018). Genome encode analyses reveal the basis of convergent evolution of fleshy fruit ripening. *Nat. Plants* 4, 784–791. doi: 10.1038/s41477-018-0249-z

Marchler-Bauer, A., Anderson, J. B., Derbyshire, M. K., Deweese-Scott, C., Gonzales, N. R., Gwadz, M., et al. (2007). CDD: a conserved domain database for interactive domain family analysis. *Nucleic Acids Res.* 35, D237–D240. doi: 10.1093/nar/gkl951

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Von Haeseler, A., et al. (2020). IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37, 1530–1534. doi: 10.1093/molbev/msaa015

Rabinowitz, J. D., Hsiao, J. J., Gryncel, K. R., Kantrowitz, E. R., Feng, X. J., Li, G., et al. (2008). Dissecting enzyme regulation by multiple allosteric effectors: nucleotide regulation of aspartate transcarbamoylase. *Biochemistry* 47, 5881–5888. doi: 10.1021/bi8000566

Schröder, M., Giermann, N., and Zrenner, R. (2005). Functional analysis of the pyrimidine *de novo* synthesis pathway in solanaceous species. *Plant Physiol.* 138, 1926–1938. doi: 10.1104/pp.105.063693

Serre, V., Penverne, B., Souciet, J. L., Potier, S., Guy, H., Evans, D., et al. (2004). Integrated allosteric regulation in the s. cerevisiae carbamylphosphate synthetase - aspartate transcarbamylase multifunctional protein. *BMC Biochem.* 5, 6. doi: 10.1186/1471-2091-5-6

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303

Smith, L. D., and Garg, U. (2017). "Chapter 5 - urea cycle and other disorders of hyperammonemia," in *Biomarkers in inborn errors of metabolism*. Eds. U. Garg and L. D. Smith (San Diego: Elsevier), 103–123.

Speirs, J., Lee, E., Holt, K., Yong-Duk, K., Steele Scott, N., Loveys, B., et al. (1998). Genetic manipulation of alcohol dehydrogenase levels in ripening tomato fruit affects the balance of some flavor aldehydes and alcohols. *Plant Physiol.* 117, 1047–1058. doi: 10.1104/pp.117.3.1047

Tang, D., Gallusci, P., and Lang, Z. (2020). Fruit development and epigenetic modifications. *New Phytol.* 228, 839–844. doi: 10.1111/nph.16724

Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2002). Multiple sequence alignment using ClustalW and ClustalX. *Curr. Protoc. Bioinf.* 2, 3. doi: 10.1002/0471250953.bi0203s00

Uluisik, S., and Seymour, G. B. (2020). Pectate lyases: Their role in plants and importance in fruit ripening. *Food Chem.* 309, 125559. doi: 10.1016/j.foodchem.2019.125559

Urbano-Gamez, J. A., El-Azaz, J., Avila, C., de la Torre, F. N., and Canovas, F. M. (2020). Enzymes involved in the biosynthesis of arginine from ornithine in maritime pine (Pinus pinaster ait.). *Plants (Basel)* 9 (10), 1271. doi: 10.3390/plants9101271

Wild, J. R., Johnson, J. L., and Loughrey, S. J. (1988). ATP-liganded form of aspartate transcarbamoylase, the logical regulatory target for allosteric control in divergent bacterial systems. *J. Bacteriol.* 170, 446–448. doi: 10.1128/jb.170.1.446-448.1988

Winter, G., Todd, C. D., Trovato, M., Forlani, G., and Funck, D. (2015). Physiological implications of arginine metabolism in plants. *Front. Plant Sci.* 6, 534. doi: 10.3389/fpls.2015.00534

Xiao, K., Chen, J., He, Q., Wang, Y., Shen, H., and Sun, L. (2020). DNA Methylation is involved in the regulation of pepper fruit ripening and interacts with phytohormones. *J. Exp. Bot.* 71, 1928–1942. doi: 10.1093/jxb/eraa003

Yao, B. N., Tano, K., Konan, H. K., Bedie, G. K., Oule, M. K., Koffi-Nevry, R., et al. (2014). The role of hydrolases in the loss of firmness and of the changes in sugar content during the post-harvest maturation of carica papaya l. var solo 8. *J. Food Sci. Technol.* 51, 3309–3316. doi: 10.1007/s13197-012-0858-x

Zhao, T., Holmer, R., De Bruijn, S., Angenent, G. C., Van Den Burg, H. A., and Schranz, M. E. (2017). Phylogenomic synteny network analysis of MADS-box transcription factor genes reveals lineage-specific transpositions, ancient tandem duplications, and deep positional conservation. *Plant Cell* 29, 1278–1292. doi: 10.1105/tpc.17.00312

Zhao, T., and Schranz, M. E. (2019). Network-based microsynteny analysis identifies major differences and genomic outliers in mammalian and angiosperm genomes. *Proc. Natl. Acad. Sci. U.S.A.* 116, 2165–2174. doi: 10.1073/pnas.1801757116

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 04 November 2022
DOI 10.3389/fpls.2022.954929

OPEN ACCESS

EDITED BY
Octavio Martínez,
Centro de Investigación y Estudios Avanzados del IPN (CINVESTAV), Mexico

REVIEWED BY
Anja Christina Hoerger,
University of Salzburg, Austria
Rahul Kumar,
University of Hyderabad, India

*CORRESPONDENCE
Amy Litt
amy.litt@ucr.edu

†PRESENT ADDRESS
Alex Rajewski,
SEngine Precision Medicine, Seattle, WA, United States
Jessica Le,
Thomas Jefferson University, Center for Forensic Science Research & Education, Philadelphia, PA, United States

SPECIALTY SECTION
This article was submitted to Plant Development and EvoDevo, a section of the journal Frontiers in Plant Science

RECEIVED 27 May 2022
ACCEPTED 14 October 2022
PUBLISHED 04 November 2022

CITATION
Rajewski A, Maheepala DC, Le J and Litt A (2022) Multispecies transcriptomes reveal core fruit development genes. *Front. Plant Sci.* 13:954929. doi: 10.3389/fpls.2022.954929

Multispecies transcriptomes reveal core fruit development genes

Alex Rajewski†, Dinusha C. Maheepala, Jessica Le† and Amy Litt*

Department of Botany and Plant Sciences, University of California Riverside, Riverside, CA, United States

During angiosperm evolution there have been repeated transitions from an ancestral dry fruit to a derived fleshy fruit, often with dramatic ecological and economic consequences. Following the transition to fleshy fruits, domestication may also dramatically alter the fruit phenotype *via* artificial selection. Although the morphologies of these fruits are well documented, relatively less is known about the molecular basis of these developmental and evolutionary shifts. We generated RNA-seq libraries from pericarp tissue of desert tobacco and both cultivated and wild tomato species at common developmental time points and combined this with corresponding, publicly available data from Arabidopsis and melon. With this broadly sampled dataset consisting of dry/fleshy fruits and wild/domesticated species, we applied novel bioinformatic methods to investigate conserved and divergent patterns of gene expression during fruit development and evolution. A small set of 121 orthologous "core" fruit development genes show a common pattern of expression across all five species. These include key players in developmental patterning such as orthologs of *KNOLLE*, *PERIANTHIA*, and *ARGONAUTE7*. GO term enrichment suggests that these genes function in basic cell division processes, cell wall biosynthesis, and developmental patterning. We furthermore uncovered a number of "accessory" genes with conserved expression patterns within but not among fruit types, and whose functional enrichment highlights the conspicuous differences between these phenotypic classes. We observe striking conservation of gene expression patterns despite large evolutionary distances, and dramatic phenotypic shifts, suggesting a conserved function for a small subset of core fruit development genes.

KEYWORDS

dry fruit, fleshy fruit, transcriptome, Solanaceae, Arabidopsis, melon, tobacco, tomato

Introduction

Seed-bearing fruits are the hallmark feature uniting the angiosperms, and this innovation has contributed to the enormous success of the group in terms of both species richness and economic importance for humans. Indeed, 82% of daily calories eaten by humans are derived directly from angiosperm plants (FAO, 2017) and 80% of those calories are from the fruits themselves. When indirect sources are taken into account, nearly all calories eaten by humans derive from angiosperms.

From a diversity standpoint, angiosperms also represent an unparalleled evolutionary success story. Since their initial split with gymnosperms, angiosperms diversified prolifically to comprise approximately 90% of all extant land plant species and now occupy key positions in nearly every biome on the planet (Crepet and Niklas, 2009). Although the precise reasons for the evolutionary diversification and success of angiosperms are still debated (Armbruster, 2014), certainly the complex interplay between flowers and their pollinators and the ability to further use animals as a seed dispersal vectors has contributed significantly to this (Regal, 1977).

Although the molecular mechanisms underlying fruit development and evolution are not thoroughly understood, morphological changes are well documented and provide a conceptual framework to examine molecular mechanisms. Fruits can broadly be classified as either dry or fleshy. The true berry of cultivated tomato (*Solanum lycopersicum*) and the pepo of melon (*Cucumis melo*) are examples of fleshy fruits, while the capsules of desert tobacco (*Nicotiana obtusifolia*) and the silique of the model plant *Arabidopsis thaliana* (hereafter, Arabidopsis) are both dry fruits. Despite the very different appearances of these fruits, the developmental progression of each can be divided into common stages with similar processes occurring at each stage across all four species (Table 1) (Gillaspy et al., 1993).

All fruits are derived from one or multiple ovaries. The earliest stage of fruit development (Stage 1) occurs before the ovules have been fertilised and comprises a stage of ovary patterning that is common to all species. Although specific terminology differs, the ovaries of all four species previously mentioned are divided into multiple chambers. In the cases of desert tobacco, Arabidopsis, and the wild relative of tomato (*S. pimpinellifolium*), the ovary is divided into two chambers. The fruits of wild melon species have 2-5 chambers, while both cultivated melon and cultivated tomato have a variable number of chambers (Monforte et al., 2014). Following fertilisation of the ovules, the ovary transitions to a fruit and enters into a stage of rapid cell division (Stage 2). The length of this phase differs, with both Arabidopsis and desert tobacco undergoing cell division phases of 1-3 days, while tomato and melon cell division phases can occur over 1-2 weeks (Pabón-Mora and Litt, 2011; Chayut et al., 2015; Ripoll et al., 2019). Additionally, the orientation of these cell divisions in the pericarp (outer fruit wall) varies. Pericarp cell divisions in desert tobacco are primarily anticlinal and maintain 7-8 pericarp cell layers, but pericarp divisions in tomato, and likely melon, are both anticlinal and periclinal and increase the number of cell layers dramatically (Pabón-Mora and Litt, 2011; Shin et al., 2017).

Following this burst of cell division, the fruit enters a phase of cell differentiation (Stage 3). In this stage, the fruits of each species begin to morphologically diverge from one another more drastically. Among the dry-fruited species Arabidopsis and desert tobacco, Stage 3 is characterised primarily by the deposition of lignin in the secondary cell walls of the pericarp. Because both of these fruits are dehiscent, pericarp lignification is tightly spatially controlled to allow for the formation of dehiscence zones where the mature pericarp will split open to allow seed dispersal (Ferrándiz et al., 2000; Smykal et al., 2007). In tomato and melon, Stage 3 of fruit development is characterized by pronounced pericarp cell expansion and

TABLE 1 Developmental stages of fruit development in the study species and data sources.

Description of Developmental Stages				
	Desert Tobacco	***A. thaliana***	**Tomato**	**Melon**
Stage 1	Ovary patterning (0 DAP)	Ovary patterning (3 DAP)	Ovary patterning (1 DAP)	Ovary patterning (—)[1]
Stage 2	Transverse anticlinal cell division (3 DAP)	Cell division and expansion (6 DAP)	Anticlinal and periclinal cell division (3 DAP)	Cell division (10 DAP)
Stage 3	Beginning basipetal lignification (6 DAP)	Beginning basipetal lignification (9 DAP)	Cell expansion and endoreduplication (15 DAP)	Cell expansion (20 DAP)
Transition	Color change from green to brown (11 DAP)	Color change to from green to yellow (12 DAP)	Initial color change from green to red, often termed 'breaker' stage (35 DAP)	Increase in sugar content, maximum firmness (30 DAP)
Stage 4	Senescence and dehiscence (—)[1]	Senescence and dehiscence (—)[1]	Cell wall softening, increase in sugar content, and full accumulation of pigments (45 DAP)	Fruit softening, maximum sugar content (40 DAP)
Bioproject Accession	PRJNA646747 (This Study)	PRJEB25745 (Mizzotti et al, 2018)	PRJNA646747 (This Study)	PRJNA314069 (Chayut et al, 2017)

[1]Not sampled.
Study species are shown in bold in the first row with stage names in the first column. Intersections of species and stage provide a brief description of developmental hallmarks and the number of days after pollination (DAP) when this stage occurs and when sampling occurred. NCBI bioproject accession numbers for each data source are provided in the final row. Stage descriptions adapted from: Pabón-Mora and Litt, 2011; Mizzotti et al, 2018; and Zhang et al, 2016.

contributes strongly to the mature fruit size. Concomitant with the increase in cell volume is also an increase in cell ploidy, with endoreduplication up to 256x (Bourdon et al., 2010). Endoreduplication has also been reported in Arabidopsis pericarp cells undergoing cell expansion and may be a more general feature of Stage 3 across fruit types (Ripoll et al., 2019).

Having reached their final size, these fruits transition to physiological maturity (Stage 4). In the case of the dry fruits presented here, Stage 4 involves senescence, drying down, and dehiscence of the pericarp along the previously patterned dehiscence zones. During dehiscence, tension created by drying of the lignified pericarp and autolysis of certain cells in the dehiscence zone allow the pericarp to split open and seeds to be dispersed. In contrast, Stage 4 in fleshy fruits generally involves accumulation of sugars, volatile and flavour compounds, pigments, and nutrients in the pericarp, along with softening of pericarp cell walls. In the climacteric fruits tomato and melon, this process coincides with a burst in production of the gaseous hormone ethylene, but non-climacteric fruits undergo similar processes in an ethylene-independent manner. Especially in tomatoes, an initial transition or "breaker" stage is also recognized between Stages 3 and 4. Breaker stage is characterised by the initial colour change in the pericarp from green to pink or red.

The early morphological similarities and the similar early developmental processes occurring across these diverse fruit types are likely related to their shared evolutionary origin. In fact across angiosperm evolution, there have been repeated shifts from an ancestral dry fruit to a derived fleshy fruit (Cox, 1948; Bremer and Eriksson, 1992; Plunkett et al., 1997; Clausing et al., 2000; Spalik et al., 2001; Knapp, 2002; Weber, 2004; Givnish et al., 2005) The conservation of morphological, developmental, and evolutionary patterns led us to hypothesise that there might also be conservation of gene function and/or gene expression patterns in fruit development across species. Although many studies characterising gene expression during fruit development have dramatically advanced our understanding within single species or between closely related species, a comparison at higher taxonomic levels could provide evidence for a set of "core" fruit development genes and shed light on the conserved pathways necessary to build a fruit.

We examined pericarp transcriptomes of two dry- and three fleshy-fruited species across developmental time. Our results draw upon 42 pericarp RNAseq libraries of three members of the nightshade family (Solanaceae) generated for this study as well as data from 30 additional publically available pericarp libraries of more distantly related dry- and fleshy-fruited species (Table 1). Integrating information about orthologous genes and using nested models to call differential gene expression across developmental stages, we uncovered a set of 121 genes with conserved patterns of expression among these species. These genes participate in many biological processes and may constitute a core set of genes whose expression patterns are necessary (but not sufficient) for fruit development. In addition, we found a much larger set of 1,795 genes with patterns of expression conserved within, but divergent between, dry and fleshy fruits. These genes with divergent patterns between fruit types may represent accessory genes that act to specify the developmental patterns separating these fruit types.

Methods

Plant materials

Seeds for *Solanum lycopersicum* 'Ailsa Craig' and *Solanum pimpinellifolium* (LA 2547) were provided by the UC Davis Tomato Genetics Resource Center, and those for *Nicotiana obtusifolia* (TW143) were obtained from the New York Botanical Garden. We grew all plants in a temperature-controlled greenhouse at 26°C on the campus of the University of California, Riverside.

Developmental staging

For *Solanum* spp., we chose five developmental time points for sampling, corresponding to widely accepted stages in fruit development (Gillaspy et al., 1993): early ovary development until fruit set, initiation of cell division, initiation of cell differentiation, and ripening or maturity. For *Solanum* spp., we divided the ripening stage into a transition or "breaker" stage and true physiological maturity. The same schema was applied in the dry-fruited *N. obtusifolia*, except for physiological maturity, which is highly lignified and fully senesced. Because of the difficulty obtaining usable RNA from this stage, we did not include it for *N. obtusifolia* (Table 1).

To determine the timing of the early stages, we conducted serial sectioning and staining on a series of greenhouse-grown pericarps from each species. We collected fruit and ovary tissue from 0-15 days post anthesis (DPA) and trimmed them to roughly 1cm cubes as needed. We vacuum infiltrated (-0.08Mpa) these in FAA consisting of 10% formaldehyde, 50% ethanol, and 5% acetic acid in distilled water overnight and then stored them in 50% ethanol for later use. Before embedding the fixed tissue for sectioning, we first dehydrated it through an ethanol series ending with a final absolute ethanol dehydration overnight. Across two two-hour incubations at room temperature, we replaced the ethanol with 50% ethanol/50% Citrisolv (Decon Labs, King of Prussia, PA) followed by 100% Citrisolv. We then added paraffin chips, placed the samples in a 60°C oven, and replaced the solution with liquid paraffin approximately seven times over the next two days. After we could no longer smell the Citrisolv, we placed the tissue in aluminium crinkle dishes (VWR, Radnor, PA) to solidify before shaping and mounting them for sectioning the next day. We

sectioned the blocks into 8-10μM thick ribbons and affixed them to microscope slides.

We stained high-quality, representative sections with Safranin O and Astra Blue. To deparaffinize the tissue slides we washed them twice for five minutes each in xylene, and followed this by rehydration through an ethanol series. We first stained in Safranin O (1% w/v in water) for 60 minutes, rinsed them twice with deionized water and then counterstained with Astra Blue (1% w/v in a 2% tartaric acid solution) for 10 minutes. We then rinsed the slides twice in water, and dehydrated them through the same ethanol series before rinsing twice with xylene. We then affixed a coverslip with permount and dried the slides at 40° overnight. We imaged the slides to count cell layers and observe cell size increases in the case of *Solanum* spp. and to observe lignification in the case of *N. obtusifolia*.

To determine the timing of stage 2 (cell division) in *N. obtusifolia* we observed fruits for a conspicuous jump in size and a shift in fruit apical shape from conical to blunted.

RNA extraction and sequencing

For all three species, we hand-dissected pericarps on ice from developing fruits and, in the case of earlier developmental stages, pooled multiple pericarps from a single individual to obtain enough tissue for RNA isolation. Each biological replicate represents pericarps from a single plant. We snap froze dissected tissue in liquid nitrogen, ground each sample with a micropestle attached to a cordless drill, and isolated RNA with the RNeasy Plant Mini Kit (QIAGEN, Hilden, Germany) according to the manufacturer's protocol. For *N. obtusifolia* the lysis step of this protocol was modified to use buffer RLC instead of RLT and supplemented with 2.5% (w/v) polyvinylpyrrolidone (PVP). DNA contamination was removed with an on-column RNAse-Free DNAse kit (QIAGEN, Hilden, Germany) according to the manufacturer's protocol.

The UCR Institute for Integrative Genome Biology (IIGB) Genomics Core assessed the integrity of the isolated RNA using an Agilent 2100 Bioanalyzer. We prepared high-quality samples into Illumina RNA-sequencing libraries using the NEBNext Ultra II Directional RNA library Prep Kit for Illumina (New England BioLabs, Ipswich, MA, United States) and barcoded each library for multiplexing with the NEBNext Multiplex Oligos for Illumina (Index Primers Set 1) kit. Both protocols were undertaken according to the manufacturer's instructions.

Libraries for *S. lycopersicum*, *S. pimpinellifolium*, and *N. obtusifolia* were sequenced at the UCR IIGB Genomics Core. All *Solanum* libraries and the stage 1-3 libraries of *N. obtusifolia* were sequenced on the Illumina NextSeq V2 with a high-output 2x75bp run. The Stage 4 libraries were sequenced as part of an Illumina NextSeq 1x75bp run. Raw sequence reads for all 42 pericarp libraries are available under NCBI BioProject PRJNA646747.

Bioinformatic analyses

All scripts used to analyse RNA-seq data for this study are publically accessible in a GitHub repository (github.com/rajewski/SolTranscriptomes).

We downloaded the raw RNA-seq reads for the *Arabidopsis thaliana* and Cucumis melon experiments (PRJEB25745 and PRJNA314069, respectively, Table 1) from the Sequence Read Archive (Chayut et al., 2017; Mizzotti et al., 2018). We trimmed the demultiplexed RNA-seq data with TrimGalore (Krueger, 2012) and mapped reads using STAR v2.5.3a (Dobin et al., 2013). Because of the low continuity of the *S, pimpinellifolium* reference genome, we mapped RNA-seq reads for both *Solanum* species to the *S. lycopersicum* (SL4.0) genome assembly (Hosmani et al., 2019). For *N. obtusifolia*, we mapped the reads to version 1 of the *Nicotiana obtusifolia* reference genome assembly (Xu et al., 2017), for *Arabidopsis thaliana* data, we mapped reads to the TAIR10 assembly (Berardini et al., 2015), and for melon, we mapped read to the *Cucumis melo* cv. DHL92 genome (Garcia-Mas et al., 2012).

We used the program OrthoFinder2 (Emms and Kelly, 2019) to cluster the genes from the five species into orthologous groups based on protein sequence similarity. Within the framework of the OrthoFinder2 pipeline, we opted for gene tree estimation using multiple sequence alignments with MAFFT (Katoh and Standley, 2013) followed by IQ-Tree (Nguyen et al., 2015) instead of the default DendroBLAST algorithm (Kelly and Maini, 2013). To obtain a more tractable dataset for differential expression analyses, we eliminated orthologous groups with paralogs and filtered the results for single copy genes common to all species.

Because our experimental design contained several sequential timepoints and multiple species, pairwise comparisons with time points coded as unrelated categorical variables would fail to intuitively capture the dynamic nature of gene expression and would suffer from a severe multiple testing problem. Similarly, treating time as a linear predictor of gene expression would fail to identify transiently up-regulated genes. To avoid this problem, we opted instead to implement a natural cubic spline basis transform of the time coordinates, as outlined in the supplemental material of (Fischer et al., 2018). For differential expression testing, a gene (or orthogene) is determined to be differentially expressed if its expression profile is better fit by this spline model than by a model incorporating only noise, as determined by a likelihood ratio test. Additionally, for orthogene comparisons between fruit types, an orthogene may be differentially expressed if its expression profile is statistically significantly better fit by a model incorporating interaction between the fruit type (categorical) variable and the spline basis function coefficients than by a model with only the spline coefficients. We conducted these analyses in R using the DESeq2 and splines packages (Love

et al., 2014; R Core Team, 2019). We then clustered genes determined to be differentially expressed using the DIANA algorithm of divisive clustering (Kaufman and Rousseeuw, 2005) as implemented by the R package DEGreport (Pantano, 2019). We interrogated groups of similarly expressed genes using several methods. To test for enrichment of Gene Ontology (GO) terms, we queried all protein sequences extracted from the reference genomes against the PFAM, ProSiteProfiles, TIGRFAM, and PRINTS databases (Haft et al., 2001; Attwood et al., 2012; Sigrist et al., 2013; El-Gebali et al., 2019) and aggregated all associated GO terms for each protein using a custom bash script. We then used the R package topGO (Alexa and Rahnenfuhrer, 2016) to test for enrichment of GO terms using Fisher's Exact Test and the "weight01" algorithm against a background set of all GO terms in the genome (or in the set of orthologous genes) using a custom R script.

Results

Expression patterns for polyamine and isoprenoid biosynthesis are conserved between wild and cultivated tomato species

In our investigation, we began with the commonly studied cultivated tomato (*Solanum lycopersicum*) but also included its closest wild relative (*S. pimpinellifolium*). We reasoned that the intentional and unintentional changes during the domestication of cultivated tomato could have an impact on gene expression patterns in the fruit, whose ripening, flavour, and structure have been targets of artificial selection.

Using RNAseq data from five developmental stages from fruit of both tomato species (Table 1), we first asked which differentially expressed genes across fruit development showed a conserved pattern of expression between the two species. We aligned reads from both tomato species to the most recent annotation of the cultivated tomato genome. We chose to use the cultivated tomato reference genome for *S. pimpinellifolium* mapping because existing *S. pimpinellifolium* genomic resources lack the contiguity, thorough annotation, and/or data availability provided by the cultivated tomato reference genome. This also had the added benefit of simplifying ortholog determination between the two tomato species. Supplemental File 3 shows the mapping statistics for all sequencing libraries used in this study. Libraries from both tomato species showed nearly identical percentages of mapped reads indicating a negligible bias due to reference genome choice. We then called differential expression among developmental stages with a model that was blind to species (Sander et al., 2017; Fischer et al., 2018; Hosmani et al., 2019). This model required that the expression of a gene be statistically significantly different between at least two stages. We discovered 6,165 genes (of 34,075 total) with changes in pericarp expression level with the same pattern in cultivated and wild tomato. A GO term enrichment analysis of this cohort of genes revealed that they function in diverse general biological processes including glucose metabolism, transport, and responses to damage and stress (Figure 1A). In addition, several lower-level GO terms were also enriched among this set of genes including spermidine biosynthetic processes, which play a role in the synthesis of polyamine compounds related to flavour and timing of fruit senescence (Nambeesan et al., 2010).

To uncover more fine-scale patterns among these differentially expressed genes, we clustered them by their expression profiles during fruit development and performed GO analyses on each of the 20 resulting clusters (Supplementary Data Figure S1). Several of these clusters showed informative enrichments. Cluster 4 contained genes with low and steady expression in early fruit development, peaking at the transition to Stage 3 and remaining high through the red ripe stage (Figure 1B). This cluster showed enrichment for isoprenoid biosynthesis (GO:0008299), fatty acid biosynthesis, and potassium ion transport (Figure 1C). Given the peak expression of this cluster prior to the breaker stage, it is likely that these terms relate to the accumulation of pigment and flavour compounds before and during ripening (Adams et al., 1978; Tieman et al., 2012; Li et al., 2020) (Tieman et al., 2012; Li et al., 2020). This cluster also showed enrichment for genes related to cell wall modification, consistent with the prominent changes in cell wall composition as the fruit ripens and softens. Cluster 10 showed a nearly opposite pattern to cluster 4, with low expression in later fruit development, and high to moderate expression at Stages 1-3 (Figure 1D). These earlier stages of fruit development include bursts of cell division and DNA replication and this cluster contained significant hits for DNA replication, nucleotide biosynthesis and several cell wall biosynthetic terms (Figure 1E).

Wild and cultivated tomato show subtle differences in expression patterns

One of the most notable effects of artificial selection between cultivated and wild tomato is fruit size. As the pericarp makes up a substantial portion of the fruit, we wanted to know the extent to which pericarp gene expression patterns differ between the two species. We therefore called differentially expressed genes with a model that included the species as a covariate and used a likelihood ratio test to determine which genes showed a statistically significant difference in gene expression pattern between the two species. The resulting 1,472 genes that exhibited divergent expression patterns between cultivated and wild tomato showed GO term enrichment for plant-type cell wall organisation and lipid biosynthetic processes, with 11 genes assigned to each term, the maximum number of genes for any GO term in this analysis (Figure 1F). This enrichment likely reflects both the different flavour profiles of the two fruits as well

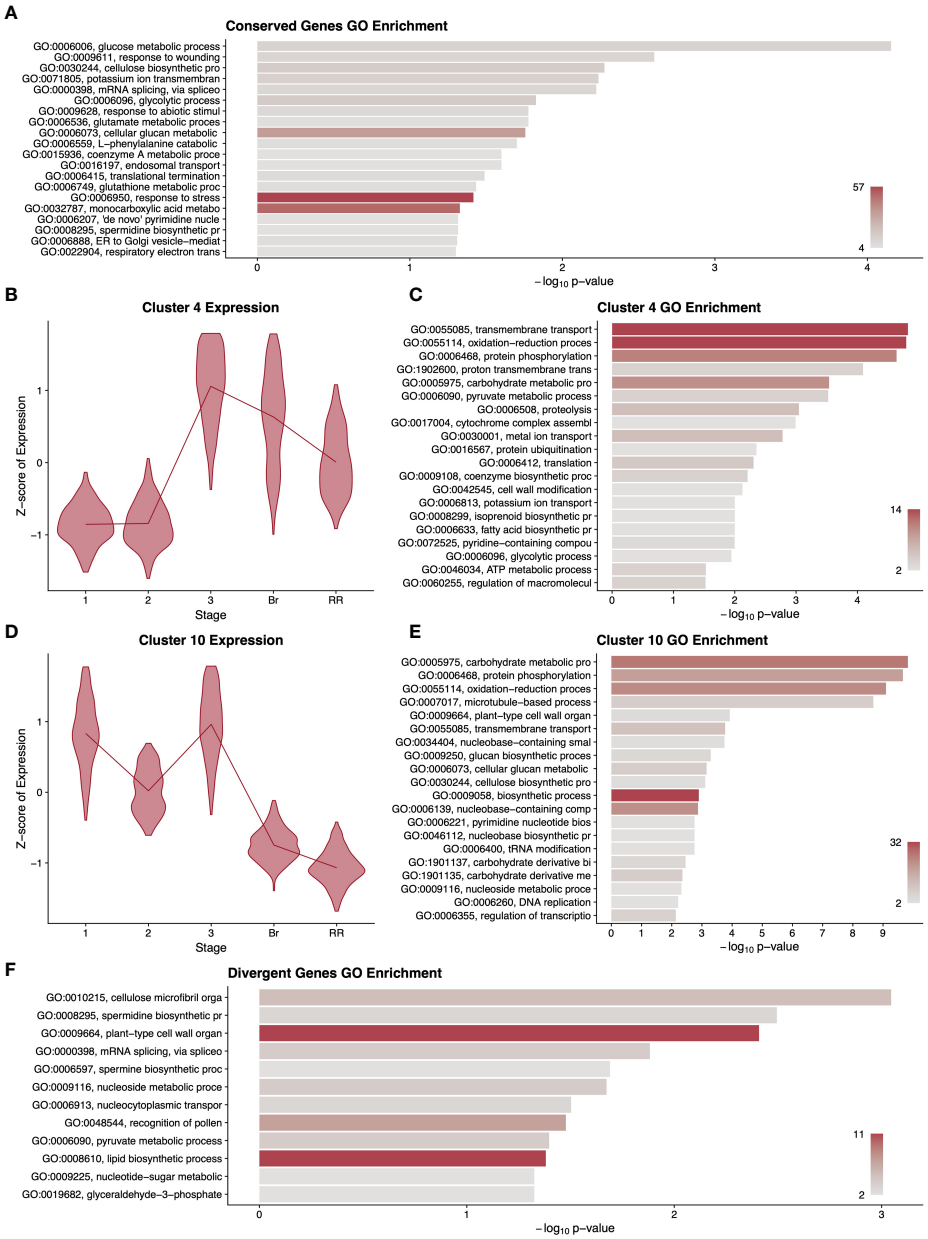


FIGURE 1
Summary of gene expression patterns conserved **(A–E)** or divergent **(F)** among cultivated and wild tomato. A gene ontology (GO) term enrichment analysis **(A)** performed on all differentially expressed genes without regard to species. Selected clusters of differentially expressed genes conserved among species are described with violin plots of normalised expression at each stage of development **(B, D)** and with GO enrichment analyses **(C, E)**, corresponding to 1415 and 1825 genes respectively. For differentially expressed genes with divergent expression between the species, we performed a GO enrichment analysis **(F)**. GO term descriptions to the left of the enrichment graphs are truncated for space and sorted by p-value. The bars are colored by the number of genes assigned to each GO term with legends in the lower right of each graph. Stages of fruit development in the axis of B and D are numbered sequentially followed by "Br" for breaker stage and "RR" for red ripe stage.

as their conspicuous differences in pericarp size. A clustering and GO analysis of these 1,472 genes produced clusters with only very subtle differences in gene expression profiles between species and no apparently informative GO terms (Supplementary Data Figure S2). Potentially the differences in fruit phenotype between wild and cultivated tomato involve a small number of genes with slight changes in expression pattern, but we cannot rule out that these differences involve changes in timing or expression domains that were not included in our sampling regime.

Divergence in expression of ethylene and secondary metabolite synthesis genes following domestication

Because cultivated tomato is routinely used as a model to study climacteric fruit ripening, many genes have been identified as playing a role in this process. We asked to what extent the expression patterns of these well-studied ripening genes have changed following domestication. We used our combined wild and cultivated tomato dataset to examine the expression of 21 structural genes involved in ethylene biosynthesis, pigment production, and flavour compound biosynthesis (Supplementary Data Figure S3). Among these structural genes, one ethylene-related gene and two flavour compound-related genes have a pattern of expression with statistically significant differences between cultivated and wild tomato (Figures 2A-C).

The gene *ACO6* encodes an ethylene biosynthesis enzyme whose role has not been well characterised during fruit development (Houben and Van de Poel, 2019). In our analysis, *ACO6* was the only structural gene related to ethylene synthesis or perception with a statistically significant difference in expression pattern between the two tomato species (Figure 2A). The other genes showed either no statistically significant change in expression across pericarp development or no statistically significant difference in pattern between the two species. In contrast, *SpACO6* has higher expression at every stage we sampled in wild tomato compared to *SlACO6* in

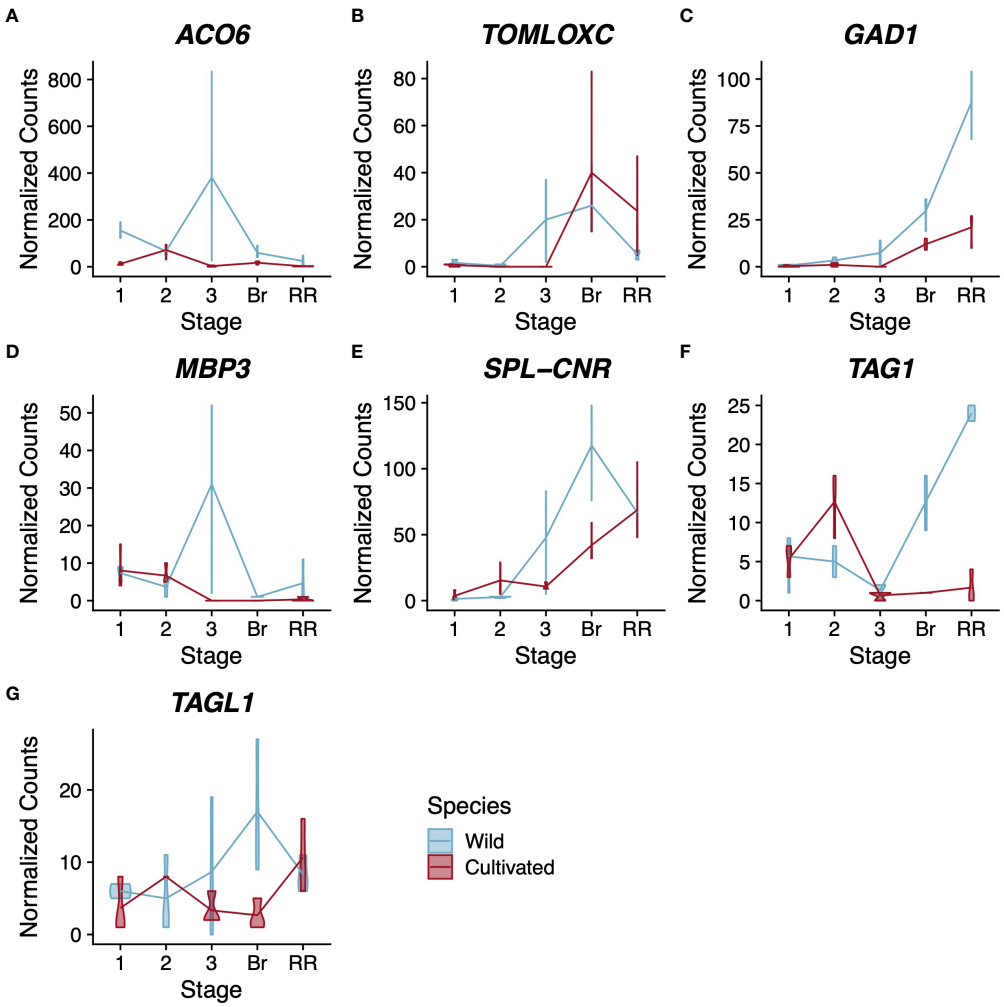


FIGURE 2
Expression profiles for ethylene-related **(A)**, flavour compound-related **(B–D)**, and regulatory **(D–G)** genes. Normalised counts of gene expression are represented by violin plots. Genes with statistically significant (FDR<0.01) differential expression across stages are shown in bold. Wild tomato is shown in blue and cultivated in red. Stages of fruit development on the X-axis are numbered sequentially followed by "Br" for breaker stage and "RR" for red ripe stage Note that panels have independent Y-axis to maximise readability.

cultivated tomato. Additionally, *SlACO6* reaches its maximum expression in cultivated tomato at stage 2, which is characterised primarily by cell division, whereas in wild tomato, peak expression of *SpACO6* is reached at stage 3, which is characterised largely by cell expansion (Table 1). The peak at stage 3 was not seen for any other *ACO* homologues, suggesting a divergent role for this enzyme during pericarp development (Supplementary Data Figures S3A-G).

TomLoxC encodes a lipoxygenase and contributes to desirable flavour in tomato fruit (Chen et al., 2004; Shen et al., 2014). In both species, expression was not detected in stages 1-2 of pericarp development (Figure 2B). In wild tomato, *SpTomLoxC* transcripts accumulated to moderate levels at stage 3 and breaker stage pericarps, but dropped to much lower levels in red ripe fruits. In cultivated tomato, however, we did not detect any *SlTomLoxC* transcripts until the breaker stage, where we observed maximum expression. The level dropped slightly at the red ripe stage, but still remained higher than the peak expression seen in wild tomato. Polymorphism in *TomLoxC* expression was recently observed in a large study of wild and cultivated tomato accessions, and found to correlate with a large deletion in the promoter of *TomLoxC* that was selected against during domestication (Gao et al., 2019).

Finally, *GAD1* encodes one of three known tomato glutamate decarboxylases, which are responsible for the production of γ-aminobutyric acid (GABA) (Akihiro et al., 2008). In our analysis both tomato species displayed a similar trend for *GAD1* expression during pericarp development, which was consistent with previous studies (Akihiro et al., 2008) (Figure 2C). However the two species showed a statistically significant difference in the magnitude of expression, with wild tomato showing approximately 3x higher peak expression of *SpGAD1* at the red ripe stage. GABA can accumulate to very high levels in tomato fruit and is thought to be involved with stress responses and defence (Bouché et al., 2003; MacGregor et al., 2003). Given that wild tomato is a widely recognized resource for introgression of stress tolerance, this difference in a key GABA biosynthesis enzyme represents a potential future avenue for plant breeders (Razali et al., 2018).

Fruit size-, firmness-, and lignification-related transcription factors differ in expression between wild and cultivated tomato

Because changes in the expression of transcription factors can influence the expression of many target genes simultaneously, we wanted to know the extent to which such regulatory genes differed in expression pattern between these two species. We selected 18 transcription factors with prominent roles in fruit and flower development and used our combined wild and cultivated tomato data set to ask if any of these genes showed statistically significant differences in expression between the two species (Supplementary Data Figure S4).

Although many of the selected genes showed statistically significant differential expression across pericarp development with a pattern common to both species, only four had statistically significant support for a difference in expression between the two species. This included three type-II MADS-box genes *MBP3*, *TAG1*, and *TAGL1*, along with the *SQUAMOSA* promoter-binding protein-like transcription factor, *SPL-CNR* (Figures 2D-G).

MBP3 and *AGL11* are orthologous to the Arabidopsis gene *SEEDSTICK*, which helps specify ovule identity (Pinyopich et al., 2003; Ocarez and Mejía, 2016). *AGL11* does not show statistically significant differential expression between tomato species; however its paralog, *MBP3*, does (Supplementary Data Figures S4A, S2D). Our dataset shows that in cultivated tomato, *SlMBP3* expression is low in stages 1 and 2 before becoming nearly undetectable for the rest of fruit development. In contrast, wild tomato *SpMBP3* is similar to cultivated tomato in expression at stages 1 and 2 but peaks at stage 3 with a roughly 3-fold increase compared to stage 1. Several functional characterizations suggest that *AGL11* helps specify ovule identity in tomato, but we could find no functional characterizations of *MBP3* (Ocarez and Mejía, 2016; Huang et al., 2017).

The tomato genes *TAG1* and *TAGL1* are orthologs of the Arabidopsis genes *AGAMOUS* and *SHATTERPROOF1/2*, respectively (Pnueli et al., 1994; Pan et al., 2010). Both tomato genes have been shown to control several aspects of fruit development and to help specify the identity of stamens and carpels (Pan et al., 2010; Gimenez et al., 2016). Comparing wild and cultivated tomato, *TAG1* shows a more extreme difference in expression than *TAGL1*, though both are statistically significant ($p<<0.01$, Figures 2F, G). In wild tomato, *SpTAG1* expression increases linearly nearly 25-fold between stage 3 and the red ripe stage; however, in cultivated tomato the increase in *SlTAG1* transcripts is barely detectable. For *TAGL1* the departure in expression is more subtle but most obvious at the breaker stage where wild tomato *SpTAGL1* expression peaks and cultivated tomato *SlTAGL1* expression is at its lowest levels. Previous silencing experiments in cultivated tomato suggest that both genes contribute positively to pericarp thickness (Gimenez et al., 2016). Our result is therefore counterintuitive as cultivated tomato generally has a thicker pericarp than wild tomato, but wild tomato showed consistently higher expression of both genes in the pericarp.

The *SQUAMOSA* promoter-binding protein-like transcription factor *SPL-CNR* is thought to be the causative gene for the *Cnr* mutation that affects ripe tomato fruit colour and firmness (Thompson et al., 1999; Eriksson et al., 2004; Manning et al., 2006; Lai et al., 2020). In our analysis, *SPL-CNR* showed a statistically significant difference in expression between the two tomato species ($p=3.2\text{x}10^{-4}$) with wild tomato showing higher expression in both stage 3 and breaker stage

pericarps (Figure 2E). Recently *SPL-CNR* expression has been shown to negatively affect cell-to-cell adhesion and to promote cell death (Lai et al., 2020), consistent with a model whereby low expression of *SPL-CNR* in the *Cnr* mutant could lead to a non-softening fruit due to increased cell adhesion or lower levels of cell death. The decreased firmness in mature wild tomato fruits coupled with their higher expression of *SlSPL-CNR* and the increased desirability of firmer cultivated tomato fruits suggests that the expression changes at the *SlSPL-CNR* locus could have been the result of domestication (Tanksley et al., 1996; Doganlar et al., 2002).

Desert tobacco pericarp transcriptome is enriched for secondary metabolite synthesis and shows fewer differentially expressed genes than tomato

In contrast to tomato, desert tobacco (*Nicotiana obtusifolia*) produces a dry capsular fruit. We extracted RNA from pericarps at stages 1-3 as well as a "transition" stage as the fruit is maturing, analogous to breaker stage in tomato (Table 1). Physiologically mature desert tobacco fruits are dry and highly lignified, and we were unable to extract RNA from this final stage.

Because fruit development in desert tobacco has not been molecularly characterised, we examined gene expression dynamics in desert tobacco pericarp development. We applied a similar model that required the expression of a gene be statistically significantly different between at least two stages in order to be considered differentially expressed. We uncovered 1,392 desert tobacco genes with differential expression across the four stages, much fewer than the 6,165 differentially expressed genes among the tomato stages. We performed a GO analysis on this cohort of genes and found that they largely relate to either DNA replication and synthesis or to the synthesis of secondary metabolites such as spermidine or terpenoids (Figure 3A). Interestingly, the set of genes with conserved expression among the two tomato species also showed an enrichment for secondary metabolites including the polyamine spermidine (Figure 1).

We performed an analysis to sort the differentially expressed genes into clusters with similar expression profiles over time. This unsupervised method produced six profiles, and for each profile we performed a GO analysis (Figures 3B-G and Supplementary Data Figure S5). Interestingly, clusters 1, 3, and 5 have roughly complementary patterns to clusters 2, 6, and 4, respectively. Clusters 1 and 3 both contain several terms related to protein modification or degradation, while cluster 5 is primarily enriched for lipid and fatty acid biosynthesis. Clusters 2, 4, and 6 generally have a pattern of decreasing expression over time, and these clusters are all enriched for very basic metabolic functions such as DNA replication, translation, and biological processes. This decrease in expression could reflect the beginning of senescence and a general cessation of active metabolic processes.

Solanaceae expression patterns align with prominent developmental processes

The tomato species differ in fruit type from desert tobacco, and we wanted to know the extent to which expression patterns are conserved (or not) among the fruit of these phenotypically diverse, but relatively closely related taxa. To answer this we used OrthoFinder2 to find single-copy orthologous genes from dry-fruited desert tobacco and both fleshy-fruited tomato species together (Emms and Kelly, 2019). Because we were unable to extract RNA from mature desert tobacco capsules, these datasets are sampled at four comparable developmental stages (Table 1). We then applied two nested statistical models to test for differential expression over time that was conserved among all species or divergent between fruit types.

Only 1,235 single-copy orthologs showed a statistically significant conservation of expression pattern across all three species. As a cohort, this comparatively small number of genes was enriched for five GO terms, including DNA replication and protein phosphorylation (Figure 4A). To examine finer scale patterns among these genes, we performed unsupervised clustering followed by a GO analysis of the genes in each cluster. This revealed seven profiles of gene expression patterns over time (Supplementary Data Figure S6). The expression patterns and GO term enrichments for the clusters largely agree with prominent developmental processes at various stages. For instance, cluster 3 has highest expression at stages 1 and 2 and is enriched for several terms related to DNA replication, which is known to occur early in fruit development (Figures 4B, C) (Gillaspy et al., 1993; Tanksley, 2004; Pabón-Mora and Litt, 2011).

Our search for single-copy orthologs that have statistically significant differences in expression pattern between fruit types yielded 4,647 genes. A GO term analysis of this set of genes revealed terms underlying known phenotypic differences between these two fruit types including terpenoid biosynthetic processes, which are likely related to flavour compound production, as well as polysaccharide catabolism, cellulose biosynthesis, glycolytic processes, and carbohydrate derivative metabolism, which could relate to the differential accumulation of sugars and/or cell wall composition between these fruit types (Figure 4D). Unsupervised clustering and GO analyses were also carried out on this dataset; however, this did not yield readily informative patterns or terms (Supplementary Data Figure S7).

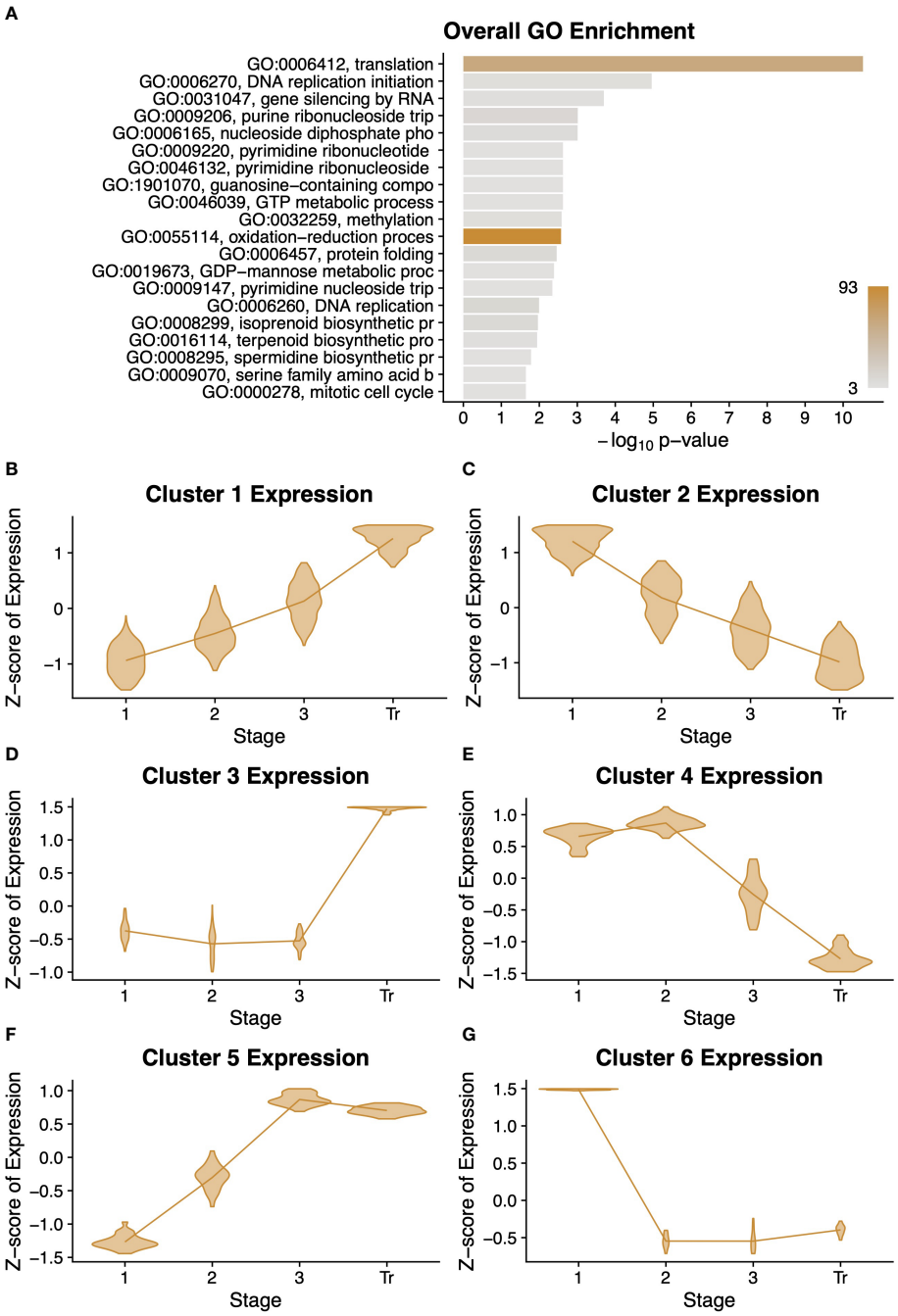


FIGURE 3
Summary of desert tobacco differentially expressed genes. A gene ontology (GO) term enrichment analysis **(A)** performed on all 1351 differentially expressed genes. All clusters of differentially expressed genes are described with violin plots of normalised expression at each stage of development **(B–G)** comprising. Stages of fruit development in the axis of **(B–G)** are numbered sequentially followed by "Tr" for transition to mature stage.

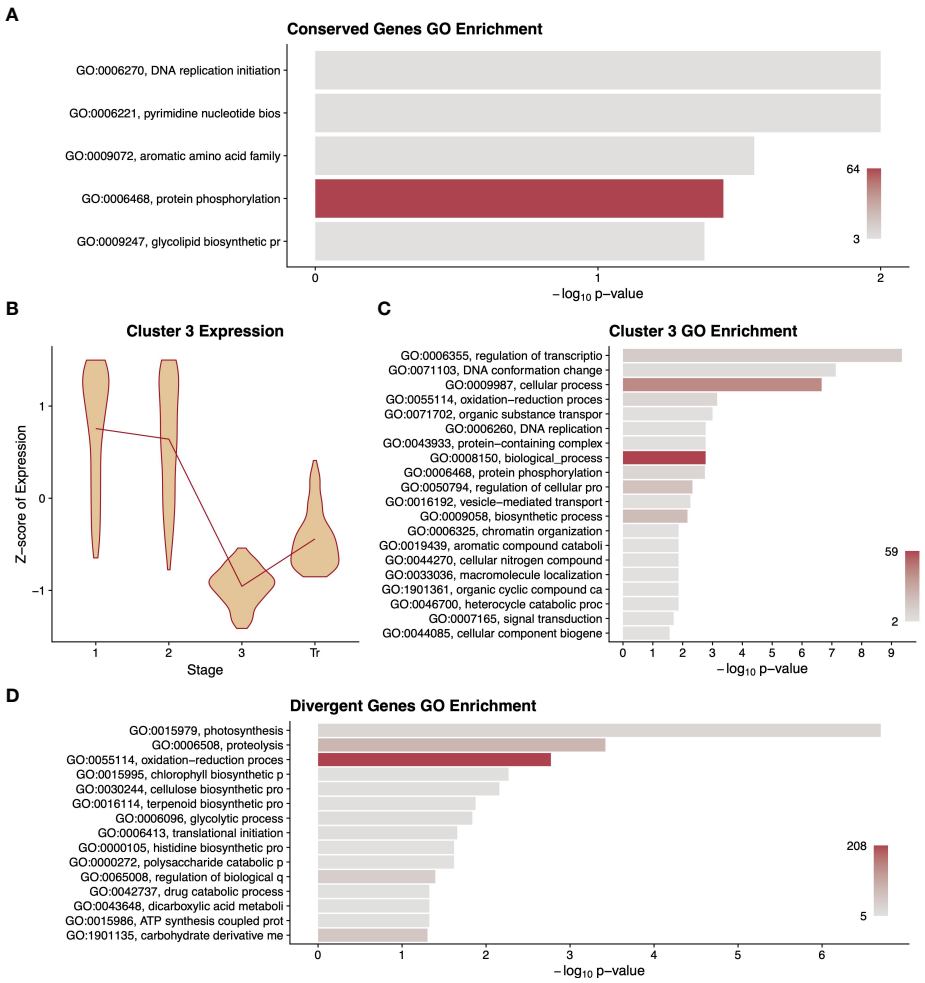


FIGURE 4
Summary of differentially expressed orthologous genes. A gene ontology (GO) term enrichment analysis **(A)** performed on differentially expressed genes that had conserved patterns among the three species. A representative cluster of 796 differentially expressed genes conserved among species is described with violin plots of normalised expression at each stage of development **(B)** along with a GO enrichment analysis **(C)** of the genes in that cluster. A gene ontology (GO) term enrichment analysis **(D)** performed on differentially expressed genes that had different patterns between fruit types. GO term descriptions to the left of the enrichment graphs are truncated for space and sorted by p-value. The bars are colored by the number of genes assigned to each GO term with legends in the lower right of the graph. Stages of fruit development in the axis of B-GD are numbered sequentially followed by "Tr" for transition to mature stage.

Solanaceae orthologs of ripening-related genes show fruit type-specific expression patterns

Given the interesting differences between wild and cultivated tomato in expression of the ripening related structural and regulatory genes, we asked to what extent the expression pattern of these genes has diverged between the fleshy-fruited tomato species and the dry-fruited desert tobacco. We restricted our analysis to genes that had a single unambiguous ortholog in all three species and found orthologs for four of 12 ethylene-related structural genes and five of 18 transcription factors (Table 2). We then pooled replicates from both tomato species as a single representative fleshy-fruited taxon and contrasted their expression values with those from desert tobacco. This effectively averages differences in expression that may have been apparent between wild and cultivated tomato but allows us to search for genes with strong signal of fruit-type specific expression over time. Using a likelihood ratio test, we were able to discern if the expression patterns show conservation between fruit types, within fruit types, or are divergent between fruit types.

Interestingly, all nine of the genes for which we determined orthology show a decrease in expression between stage 3 and the transition stage of the desert tobacco capsule (Figure 5). This result echoes that seen in desert tobacco clusters 2, 4, and 6 from

TABLE 2 Table showing the relationships between orthologous genes identified in this study.

Orthology and Abbreviations			
Gene Name	***Solanum* Gene ID**	***Nicotiana* Ortholog ID[1]**	***Nicotiana* Abbreviation**
ACO1	*Solyc07g049530.3.1*	—	—
ACO2	*Solyc12g005940.2.1*	—	—
ACO3	*Solyc07g049550.3.1*	—	—
ACO4	*Solyc02g081190.4.1*	*NIOBTv3_g13660.t1*	*NoACO4*
ACO5	*Solyc07g026650.3.1*	*NIOBTv3_g38689.t1*	*NoACO5*
ACO6	*Solyc02g036350.3.1*	*NIOBTv3_g02352.t1*	*NoACO6*
ACO7	*Solyc06g060070.3.1*	—	—
ACS2	*Solyc01g095080.3.1*	—	—
ACS4	*Solyc05g050010.3.1*	—	—
AGL11	*Solyc11g028020.3.1*	*NIOBTv3_g14436.t1*	*NoAGL11*
Cel2	*Solyc09g010210.3.1*	*NIOBTv3_g19880.t1*	*NoCel2*
Cel3	*Solyc07g005840.2.1*	*NIOBTv3_g12440.t1*	*NoCel3*
CHS-1	*Solyc09g091510.3.1*	—	—
CHS-2	*Solyc05g053550.3.1*	—	—
CTOMT1	*Solyc10g005060.4.1*	—	—
EIL1	*Solyc06g073720.3.1*	—	—
EIL2	*Solyc01g009170.4.1*	—	—
EIL4	*Solyc06g073730.2.1*	—	—
EJ2/MADS1	*Solyc03g114840.3.1*	—	—
EXP1	*Solyc06g051800.3.1*	*NIOBTv3_g17210.t1*	*NoEXP*
FUL1	*Solyc06g069430.3.1*	*NIOBTv3_g28929-D2.t1*	*NoFUL1*
FUL2	*Solyc03g114830.3.1*	*NIOBTv3_g39464.t1*	*NoFUL2*
FYFL	*Solyc03g006830.3.1*	*NIOBTv3_g10096.t1*	*NoFYFL*
GAD1	*Solyc03g098240.3.1*	*NIOBTv3_g11084.t1*	*NoGAD1*
GAD2	*Solyc11g011920.2.1*	—	—
GAD3	*Solyc01g005000.3.1*	—	—
J	*Solyc11g010570.2.1*	—	—
J2	*Solyc12g038510.2.1*	*NIOBTv3_g15806.t1*	*NoJ2*
MADS-RIN	*Solyc05g012020.4.1*	—	—
MBP10	*Solyc02g065730.2.1*	*NIOBTv3_g07845.t1*	*NoMBP10*
MBP20	*Solyc02g089210.4.1*	*NIOBT_gMBP20.t1*	*NoMBP20*
MBP3	*Solyc06g064840.4.1*	—	—
MC	*Solyc05g056620.2.1*	*NIOBTv3_g18077.t1*	*NoMC*
NAC-NOR	*Solyc10g006880.3.1*	*NIOBTv3_g08302.t1*	*NoNOR*
NR/ETR3	*Solyc09g075440.4.1*	*NIOBTv3_g10291.t1*	*NoETR3*
PGA2A	*Solyc10g080210.2.1*	—	—
PL1	*Solyc03g111690.4.1*	—	—
PSY1	*Solyc03g031860.3.1*	*NIOBTv3_g17569.t1*	*NoPSY1*
Solyc03g117740.3.1	*Solyc03g117740.3.1*	*NIOBTv3_g22270.t1*	*NIOBTv3_g22270.t1*
Solyc04g072038.1.1	*Solyc04g072038.1.1*	*NIOBTv3_g10008.t1*	*NIOBTv3_g10008.t1*
Solyc06g065310.3.1	*Solyc06g065310.3.1*	*NIOBTv3_g12238.t1*	*NIOBTv3_g12238.t1*
Solyc07g064300.3.1	*Solyc07g064300.3.1*	*NIOBTv3_g11662.t1*	*NIOBTv3_g11662.t1*
SPL-CNR	*Solyc02g077920.4.1*	*NIOBTv3_g27953.t1*	*NoSPL-CNR*
STM3	*Solyc01g092950.3.1*	—	—
TAG1	*Solyc02g071730.4.1*	*NIOBTv3_g22632-D2.t1*	*NoAG*
TAGL1	*Solyc07g055920.4.1*	*NIOBTv3_g13969.t1*	*NoSHP*
TM29	*Solyc02g089200.4.1*	*NIOBTv3_g14235.t1*	*NoSEP1*

(Continued)

TABLE 2 Continued

Orthology and Abbreviations			
Gene Name	***Solanum* Gene ID**	***Nicotiana* Ortholog ID[1]**	***Nicotiana* Abbreviation**
TM3	*Solyc01g093965.2.1*	—	—
TM5	*Solyc05g015750.3.1*	—	—
TOMLOXC	*Solyc01g006540.4.1*	—	—

[1]Only one-to-one and many-to-one orthologs.
Each row represents a single gene of interest with its abbreviated gene name in the first column, tomato gene identifier in the second column, desert tobacco gene identifier (if known) in the third column, and the desert tobacco abbreviated gene name (if known) in the fourth column.

the entire cohort of 1,392 differentially expressed genes, suggesting again that there may be a trend toward gradual ramping down of metabolic processes as the fruit begins to senesce.

Among the ethylene-related structural genes, we found orthologs for *ACO4, ACO5*, *ACO6*, and *NR/ETR3* (Figures 5A-D). *ACO4*, *ACO5*, and *NR/ETR3* each have statistically significant differences in their expression patterns between the fruit types (p=1.01x10^{-9}, 8,9x10^{-20}, and 1.8x10^{-5}, respectively). *ACO6* is differentially expressed over developmental time but this pattern is different in each of the three species. The lack of conservation for the *ACO6* expression pattern is likely due to the differences in expression among the two tomato species, which have nearly opposite patterns of expression over time. Interestingly, for *ACO5*, all desert tobacco timepoints show higher expression magnitudes than in either tomato species, and for *ACO6* desert tobacco shows higher expression than cultivated tomato. However desert tobacco capsules are non-climacteric fruits, and the high expression of these ethylene biosynthetic genes suggests that the involvement of ethylene in maturity of desert tobacco and other dry fruits deserves further study.

Among the transcription factors, we resolved unambiguous, single-copy orthologs across the three species for *AGL11, FYFL*, *SPL-CNR*, *TAG1*, and *TAGL1* (Figures 5E-I). Only *FYFL* and *TAG1* lacked statistically significant conservation of expression pattern among the three species (Figures 5F, H). In contrast to our tomato comparisons, *AGL11*, which did not show statistically significant differences between tomato species, does show statistically significant differences between fruit types (Figure 5E, p=6.5x10^{-3}, 5x10^{-5}, and 1.7x10^{-5}). As mentioned previously, the role of *AGL11* and its paralog *MBP3* in the pericarp is unclear at present, but the statistically significant divergence in expression pattern of *AGL11* between fruit types and of *MBP3* among tomato species highlights the need for further study of these gene functions following their duplication.

Orthologs of *SPL-CNR* and *TAGL1* both showed statistically significant conservation in their expression patterns by fruit types (Figures 5G, H, p=5.4x10^{-17} and 5.6x10^{-3}). The Arabidopsis ortholog of *TAGL1* promotes the formation of the dehiscence zone in the pericarp of that dry fruit (Ferrándiz et al., 2000). In our analysis, the pattern of expression for *TAGL1* is higher overall in dry fruited species and peaks at stage 3 as the dehiscence zone is forming. This provides some evidence for the functional conservation of this gene's role in dry fruit dehiscence. For *SPL-CNR*, we observe roughly opposing patterns of expression between dry and fleshy fruits. *SPL-CNR* increases in expression as fleshy fruits enter the breaker stage, before they have begun to soften. In contrast, we see a decrease in *SPL-CNR* expression as dry fruits approach dehiscence. Additional functional studies of this gene's role across dry-fruited species could help extend its established role in cell-cell adhesion and clarify its potential role in dry fruit maturity.

A small set of genes show conservation of expression pattern between dry and fleshy fruit

Our analysis of the tomato species and desert tobacco revealed a number of informative patterns, but all three species belong to the same family. As a result, we cannot tell if common patterns of gene expression are due to shared phylogenetic history or represent trends across angiosperm fruit development. We wanted to find generalizable trends in gene expression that might underlie the divergence between dry and fleshy fruit development or support conservation of certain gene expression patterns between these two phenotypically diverse fruits. We therefore chose to add *Arabidopsis thaliana*, which produces a dry silique and melon (*Cucumis melo*), which produces a type of berry with a leathery rind known as a pepo.

In order to enable expression comparisons between and among species, we used Orthofinder2 to group genes from these species into orthologous groups based on protein sequence similarity and phylogenetic relationships (Emms and Kelly, 2019). Due to their high degree of similarity, and because we had mapped wild tomato RNAseq using the cultivated tomato genome, we used cultivated tomato protein sequences in the orthology search to represent both cultivated and wild tomato. For subsequent gene expression analyses, however, the two

tomato species were not combined. We were able to group the genes from these species into 19,249 orthogroups (Figure 6A); however, many orthogroups were not shared among all species, and even among universally shared orthogroups, there were many cases of gene family expansion or loss within a single species. Because comparing transcript levels among unequal numbers of genes across species is not meaningful, we limited our interspecific expression analysis to only single-copy genes

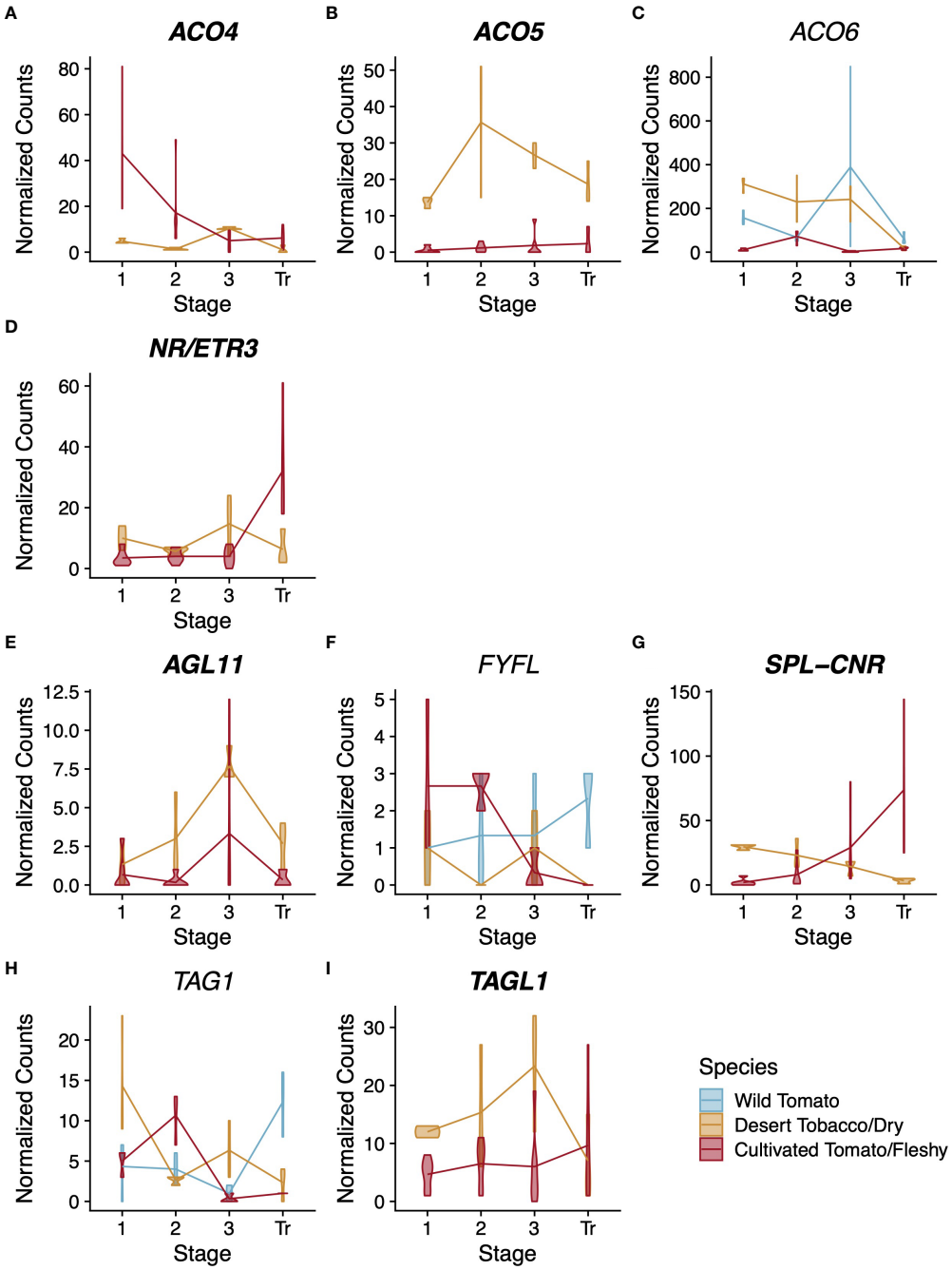


FIGURE 5
Expression profiles for ethylene-related **(A–D)** and regulatory **(E–I)** genes across the three solanaceous species. Normalised counts of gene expression are represented by violin plots. Genes with statistically significant (FDR<0.01) differential expression across stages are shown in bold. Dry-fruited desert tobacco values are shown in yellow. When the expression pattern is better described by individual species trends (based on a likelihood ratio test), wild tomato violin plots are shown in blue and cultivated tomato plots in red, otherwise both tomato species are shown together in red. Stages of fruit development on the X-axis are numbered sequentially followed by "Tr" for transition to maturity stage. Note that panels have independent Y-axis to maximise readability.

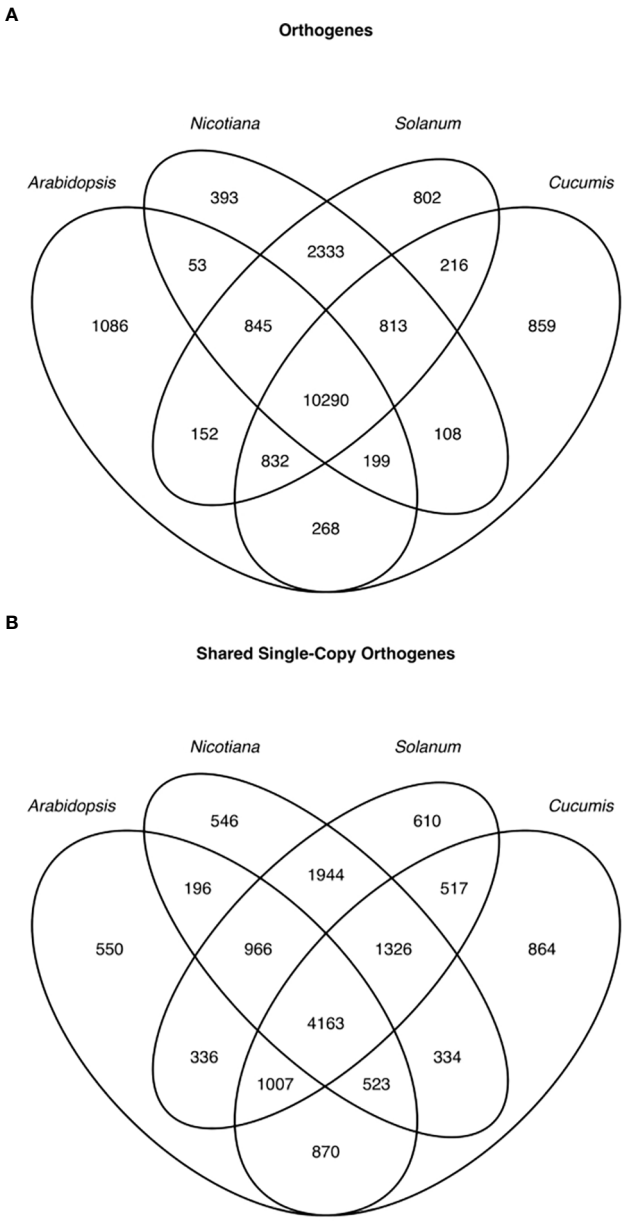


FIGURE 6
Venn diagram of orthologous genes (orthogenes) among the 4 genera used in this study. All genes across the 4 genera **(A)** and only single-copy genes **(B)**.

falling into universally present orthogroups. This filtering left 4,163 orthogenes for comparisons among both tomato species, desert tobacco, Arabidopsis, and melon (Figure 6B).

For these five species, we wanted to use comparable developmental stages to see if any orthologous genes shared similar expression dynamics over time among all species or among species with similar fruit types. After integrating the publically available Arabidopsis and melon pericarp RNAseq data with our own tomato and desert tobacco datasets, we had comparable data for stage 2, stage 3, and transition stage in all species (Table 1).

We first assessed the extent to which any of the 4,163 orthologous genes were differentially expressed over time and shared a conserved pattern across all five species. To call differential expression across the three stages, we used a model (Model 1) that is blind to species but requires a gene to have a statistically significant change in expression between at least two stages in order to be differentially expressed. Surprisingly, this

resulted in only 121 orthologous genes with a pattern of differential expression over time that is the same in all 5 species (Supplementary Data File S1). To determine if the expression data from these genes showed a detectable signal based on developmental stage, species, or fruit type, we conducted a principal component analysis using their expression values (Figures 7A-C). Model 1 did not consider species in calling differentially expressed genes, and in fact the variance explained by the first five principal components (PC) appears not to have strong signal for interspecific differences. The notable exception to this is PC2, which explains 15% of the variance and seems mostly to separate melon from the other four species; however, PC2 also separates stage 2 from later stages in both tomato species as well as stages 2 and 3 from the transition stage in Arabidopsis. PC1 explains 35% of the variance and largely distinguishes the breaker stage tomato samples from all other samples. PC3 serves to differentiate the three developmental stages of tomato from one another and also separates stage 2 samples from later stages in Arabidopsis. The developmental stages of melon are weakly distinguished by PC4 and more prominently by PC5, each of which explain 7% of the variance. PC5 also weakly separates the developmental stages of desert tobacco.

To categorise these 121 genes, we performed a GO term enrichment analysis and found a number of terms relating to prominent processes common across fruit development including cell proliferation, anatomical structure formation, cytokinesis, and cell wall modification (Figure 7D and Supplementary Data Figure S8). The shared expression patterns among these 121 genes showed only two clusters of expression profiles. Cluster 1 contains genes whose expression increases between stage 3 and the transition stage, while cluster 2 contains genes whose expression is generally decreasing during fruit development. The genes in cluster 1 are predicted to function in nucleotide metabolism, membrane and organelle structure and processes, lipid and carbohydrate metabolism, ion transport, and similar cellular processes that cannot be easily tied to any specific developmental outcomes. Given that these were pericarp transcriptomes, this is to be expected, because the pericarps of fleshy and dry fruit have little in common other than basic cellular processes like cell division and cellular metabolic processes. The features that distinguish dry and fleshy fruit, such as lignification of the former or cell softening of the latter, are likely to involve very different pathways. Cluster 2 genes are enriched for GO terms relating to DNA replication, cell division, and the phragmoplast, which forms late during cytokinesis and patterns the nascent cell wall. Enrichment for these terms is consistent with the observed cell divisions during stages 1 and 2 and the decline of cell division as fruit development proceeds into stages 3 and 4. This cluster also shows enrichment for several terms related to the thylakoid membrane of the chloroplast, which could be related to the developmental transition of fruits from photosynthetic sources to sinks.

The very small number of genes with conserved patterns across all five angiosperm species further suggests that it may be possible to define a core set of pericarp development-related genes that have a conserved function despite large divergences in both evolutionary time and in phenotype.

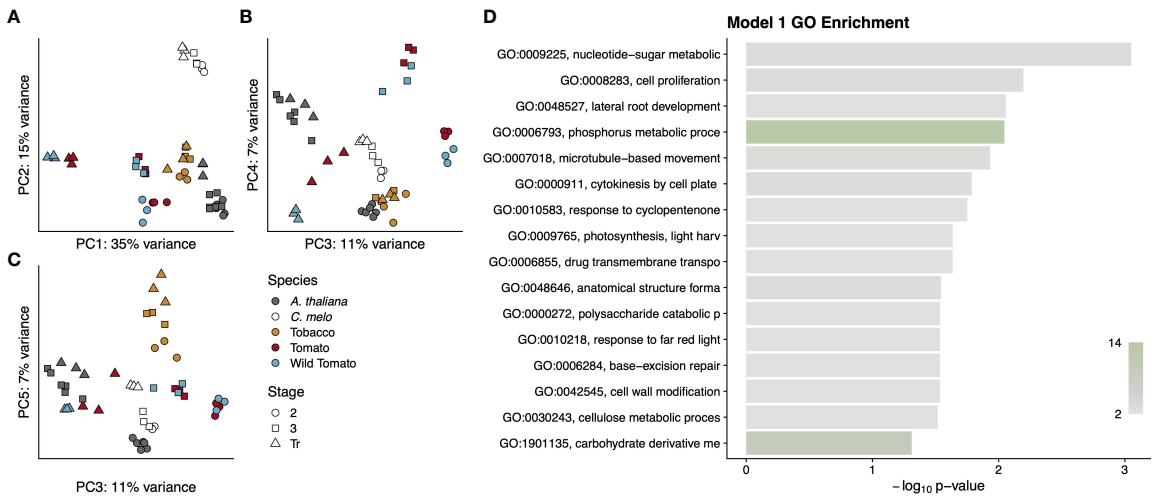


FIGURE 7
Summary of genes from Model 1. Principal components analysis **(A–C)** of gene expression values for each RNA-seq library. Points are colored by species and shaped by developmental stage as indicated in the legend. Principal components used for each graph are indicated on the axis along with the proportion of variance explained. A GO analysis **(D)** for the entire cohort of genes. GO term names to the left of the graph are truncated to available space. Terms are sorted by p-value, which is indicated by the bar height. Bars are colored by the number of genes annotated to that term, as indicated by the colour scale in the lower right.

Divergence in expression of genes related to cell division, plastid localization, and secondary cell wall composition between dry and fleshy fruits

Having established that few orthologous genes have conserved expression patterns across all five species, we next asked if and to what extent genes might show conservation of expression patterns within, but not between, fruit types. We reasoned that these fruit-type specific patterns could shed light on developmental processes shared by evolutionarily distant species with a common phenotype, dry or fleshy fruits. To answer this question, we created a model to call differentially expressed orthologous genes (Model 2) that is aware of fruit type for each of the five species but is blind to the species themselves. Like Model 1, which we used to find conserved patterns across all species, Model 2 also requires that a gene have a statistically significant change in expression between at least two of the three developmental stages. Because Models 1 and 2 are nested, genes are only differentially expressed by Model 2 if their expression pattern is better explained by Model 2 than by Model 1, as determined by a likelihood ratio test. This ensures that the difference in fruit type is driving the determination of differential expression.

Interestingly, Model 2 determined that nearly half of the 4,163 single-copy orthologous genes had divergent patterns of expression between dry and fleshy fruited species (Supplementary Data File S2). In contrast, only 202 (<5%) of these single-copy orthologous genes were differentially expressed when comparing between the wild and cultivated tomato species. We performed a principal component analysis to see if any grouping by species, developmental stage, fruit type, or evolutionary distance might be driving this large number of differentially expressed genes (Figures 8A-C). In this analysis, the first three principal components, which collectively explained 81% of the variance, served primarily to distinguish among the species. PC1 accounted for the majority of the variance (54%) and separated the dry and fleshy fruited species. On PC1, desert tobacco was separated from the two tomato species, but not as dramatically as Arabidopsis from melon, suggesting that PC1 might also incorporate some amount of variance due to phylogenetic distance in addition to fruit type. Similarly, PC2, which explained 19% of the variance, did not separate the two dry-fruited species but placed tomato and melon at two extremes. PC2 therefore combined both dry fruits but distinguished between two categories of fleshy fruits. PC3, which accounted for 8% of the variation, only seemed to separate desert tobacco from the other four species. PC4 and PC5 captured 3% and 2% of the variance, respectively, and showed a striking perpendicular separation of developmental stages in tomato and Arabidopsis but placed both melon and desert tobacco at their intersection, roughly overlapping with stage 3 of tomato (Figure 8C). Interestingly, in contrast to PC1-3, which primarily separated species, PC4 was the only principal component we examined that was able to separate the two tomato species, and even here the separation was only evident for the breaker stages samples.

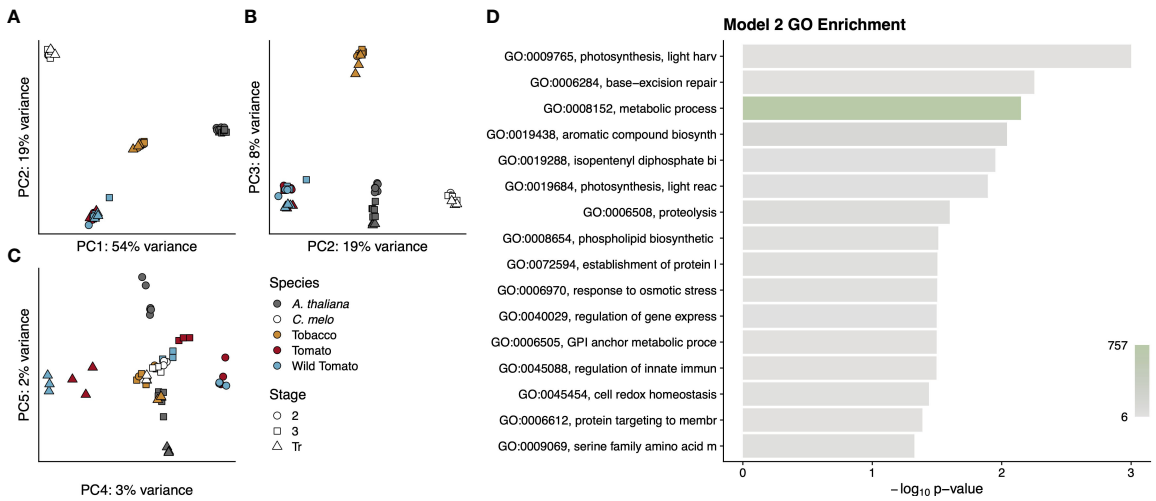


FIGURE 8
Summary of genes from Model 2. Principal components analysis **(A-C)** of gene expression values for each RNA-seq library. Points are colored by species and shaped by developmental stage as indicated in the legend. Principal components used for each graph are indicated on the axis along with the proportion of variance explained. A GO analysis **(D)** for the entire cohort of genes. GO term names to the left of the graph are truncated to available space. Terms are sorted by p-value, which is indicated by the bar height. Bars are colored by the number of genes annotated to that term, as indicated by the colour scale in the lower right.

To determine what sorts of genes were captured by this model, we performed a GO enrichment on all 1,795 genes (Figure 8D). In contrast to the very focused enrichment seen in Model 1, the genes from Model 2 were enriched for more diverse terms. In fact, the enrichment of the very high-level metabolic processes term with 757 associated genes highlights the diversity of functions that separate pericarp development in dry- and fleshy-fruited species. Even lower-level enriched terms fall into very disparate categories such as protein trafficking, secondary metabolite synthesis and regulation of gene expression.

Because of the diversity of functional terms in the GO analysis of the entire cohort of genes, we next asked in what ways the patterns of expression diverged between fruit types and what sorts of genes displayed these patterns. Our clustering analysis resulted in eight expression profiles, and we performed a GO analysis on each cluster (Supplementary Data Figure S9). Interestingly many, but not all, of these clusters showed distinctive expression profiles with more focused enrichments. In cluster 4 the relative expression diverges over time between dry and fleshy fruits, with fleshy fruits showing higher expression (Figure 9A). This cluster was enriched for several terms relating to glucose and polysaccharide synthesis, which could correspond to the accumulation of sugars in fleshy fruits as they begin to ripen (Figure 9B). Similarly, in cluster 6, dry fruits show the same pattern as cluster 4, but fleshy fruits show a slight drop in gene expression at stage 3 followed by a larger drop at the transition or breaker stage (Figure 9C). This cluster is enriched for terms relating to DNA replication and cytokinesis, likely related to the burst of cell division in stage 2 of fruit development followed by the endoreduplication that occurs in stage 3 of tomato pericarps (Figure 9D). At the transition or breaker stage of tomato fruit development, chloroplasts are known to reorganise and convert to chromoplasts, which store the conspicuous red pigments. This process is reflected in cluster 7 where dry fruits slowly drop in expression over time, but fleshy fruits show a jump in expression at the transition stage (Figure 9E). This cluster is enriched for a number of terms relating to plastid remodelling and trafficking (Figure 9F). Finally, cluster 8 highlights the key feature of dry fruit pericarps, which deposit lignin polymers in their secondary cell walls as they develop. In cluster 8, dry fruit expression remains moderate, while fleshy fruit expression values drop and remain low following stage 2 (Figure 9G). GO terms enriched in this cluster include a number of cell wall biogenesis terms (Figure 9H). Overall the profiles and enrichments seen in these clusters support a number of hypotheses regarding differential expression developmental processes separating dry and fleshy fruits and provide a basis for more direct studies of function divergence (or conservation) between these diverse fruit types.

Discussion

Across angiosperm evolution there have been repeated transitions from ancestral dry fruits to derived fleshy fruits, often with dramatic consequences. Although the morphological and developmental basis of these transitions have been well-documented, the underlying molecular and genetic mechanisms that enable or hinder these transitions have received less attention. Here we present evidence for a small set of "core" genes whose patterns of differential expression during pericarp development are conserved across several angiosperm taxa. We also show that a much larger set of "accessory" genes exists with patterns of differential expression during pericarp development that are similar within but different between dry- and fleshy-fruited species. The expression patterns of these core and accessory genes echo a number of phenotypic observations regarding differences in dry and fleshy fruit cell wall composition, cell division, and secondary metabolite production. Interestingly, these expression patterns also raise new questions about the role of ethylene in dry fruit maturity as well as the role of additional transcription factors in dry fruit dehiscence.

At lower taxonomic levels, our data also highlight a number of gene expression differences correlated with the domestication of tomato (*S. lycopersicum*) from its wild ancestor (*S. pimpinellifolium*) and provide further genetic support for previously noted phenotypic differences in fruit size, firmness, and lignification.

We also note that because our conclusions make use of externally generated datasets, and because of the variability in RNA-seq generally, a more thorough expression characterization of the genes highlighted here could be useful to control for variability in growth conditions, sampling times, and other potentially confounding variables.

Wild and domesticated tomato show differences in the expression of genes regulating domestication-related functions

Although wild and cultivated tomato species share a number of genetic and morphological similarities, cultivated tomato has undergone strong artificial selection (Blanca et al., 2015). The effects of this artificial selection are quite pronounced on the fruits, which are larger, sweeter, and firmer in cultivated than in wild tomato. We detected some potential consequences of this domestication in our pericarp gene expression dataset.

Profiling the expression of 21 ethylene- and flavour compound-related structural genes as well as 18 regulatory genes implicated in fruit ripening, we found a few key differences in expression pattern between wild and cultivated

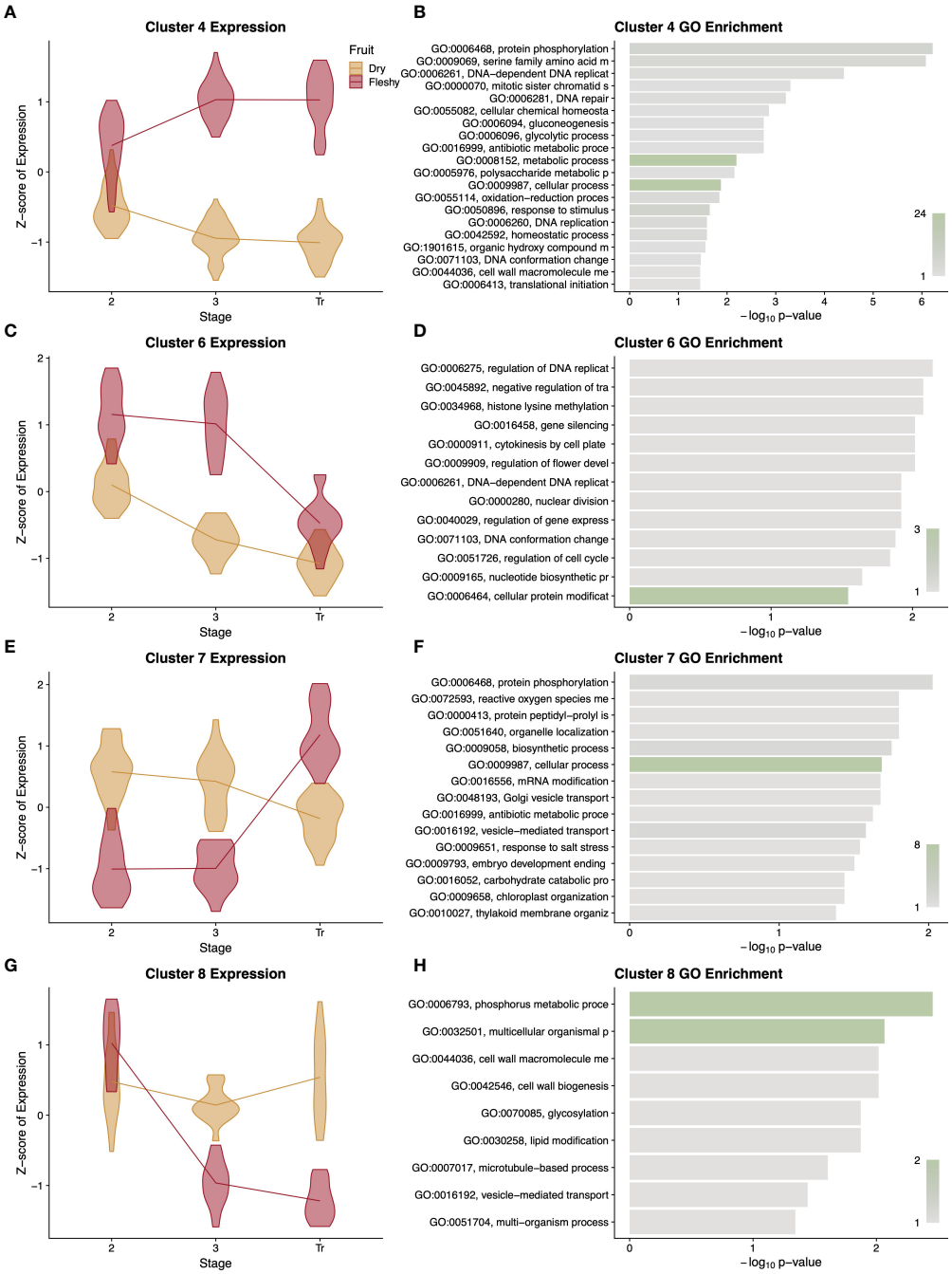


FIGURE 9
Summary of differentially expressed orthologous genes. Representative clusters of differentially expressed genes with patterns that differ between dry and fleshy fruited taxa are presented with violin plots of normalised expression at each stage of development **(A, C, E, G)** along with a GO enrichment analysis **(B, D, F, H)** of the genes in that cluster. Clusters 4, 6, 7, and 8 comprise 366, 102, 108, and 96 orthologous genes, respectively. GO term descriptions to the left of the enrichment graphs are truncated for space and sorted by p-value. The bars are colored by the number of genes assigned to each GO term with legends in the lower right of the graph. Stages of fruit development in the axis of **(A, C, E, G)** are numbered sequentially followed by "Tr" for transition to mature stage.

tomato (Figure 2 and Supplementary Data Figure S4.3, and Supplementary Data Figure S4.4). The gene *TomLoxC*, which encodes a lipoxygenase, contributes to desirable flavour in tomato fruit and showed different expression patterns between wild and cultivated tomato (Chen et al., 2004; Shen et al., 2014). This locus was previously identified as a target of selection during the domestication of tomato (Gao et al., 2019). The ethylene biosynthesis gene *ACO6* was the only ethylene-related

gene in our dataset that showed different patterns of expression between wild and cultivated tomato, with expression of this gene higher at all stages of pericarp development in wild tomato (Figure 2A). As we extended our analysis to include the dry-fruited desert tobacco pericarp transcriptome, we also saw comparatively high levels of *NoACO6* expression (Figure 5C). In fact, the levels of *NoACO6* expression were higher than in cultivated tomato throughout pericarp development and also higher than wild tomato at Stages 1 and 2, which are characterised by ovary patterning and cell division. We also saw higher expression across pericarp development for another ethylene biosynthetic enzyme *ACO5* in desert tobacco as compared to the two tomato species (Figure 5B). Higher expression of ethylene biosynthetic enzymes in this dry fruit is counterintuitive and highlights the need for further study of the roles these specific enzymes, and ethylene more generally, play in the ripening and maturity of dry fruits.

Among the regulatory genes, *MBP3* was expressed at higher levels in wild than cultivated tomato, following the stage of pericarp cell division (Figure 2D). The precise role of *MBP3* in tomato is unknown, but its paralog *AGL11* and their mutual ortholog in Arabidopsis both act to specify ovule identity (Pinyopich et al., 2003; Ocarez and Mejía, 2016; Huang et al., 2017). The role of these ovule identity genes in the pericarp is unclear at present, however the grape ortholog of these genes, *VvAGL11*, is adjacent to a QTL that controls both seedlessness and fruit size (Mejía et al., 2011). It could follow then that the differences in *MBP3* expression and in fruit size between wild and cultivated tomato represent possible subfunctionalization following the duplication that produced *AGL11* and *MBP3*.

We also detected species-specific patterns of expression for the transcription factors *TAG1* and *TAGL1* between wild and cultivated tomato (Figures 2F, G). Beyond their roles in organ identity, both *TAG1* and *TAGL1* have been shown to contribute positively to pericarp thickness; however, our results show higher expression for these genes in wild tomato, which has a thinner pericarp (Gimenez et al., 2016). Apart from this role in pericarp thickness, numerous orthologs of *TAGL1* are well documented to promote lignification of the pericarp (Ferrándiz et al., 2000; Giménez et al., 2010; Gimenez et al., 2016). We were curious if this difference in *TAGL1* expression between our two tomato species also correlated with changes in expression of structural genes involved in lignin biosynthesis. We queried our results for interspecific expression differences in the first three enzymatic steps of lignin biosynthetic (*SlPAL*: *Solyc09g007920*, *SlC4H*: *Solyc06g150137*, *Sl4CL.1*: *Solyc03g117870*, *Sl4CL.2*: *Solyc06g068650*, and *Sl4CL.3*: *Solyc12g042460*) as well as two enzymes at branch points of the pathway (*SlHCT*: *Solyc03g117600* and *SlF5H*: *Solyc02g084570*). We found that *SlHCT*, the first committed step in the formation of G- and S-type lignin, shows a statistically significant difference in expression pattern between wild and cultivated tomato (p=0.022, likelihood ratio test). This result suggests that, although neither fruit accumulates lignin to substantial levels, there may have been selection against pericarp lignification during tomato domestication. Extending the characterization of *TAGL1* to include desert tobacco, we also saw differences in expression for this gene between fruit types, with higher expression of the desert tobacco *TAGL1* ortholog, *NoSHP* from Stages 1 through 3 of fruit development (Figure 5I). This result supports potential conservation of the role *NoSHP* is expected to play in lignin patterning of the dehiscence zones across evolutionarily divergent dry fruits (Ferrándiz et al., 2000).

Finally, we found support for expression differences in *SPL-CNR* between wild and cultivated tomato (Figure 2E). Although the pattern of expression for both species shows an upward trend between Stage 2 and Breaker stage, the increase is more dramatic for wild tomato. *SPL-CNR* is believed to be the causative locus underlying the *Colorless non-ripening* (*Cnr*) mutant in tomato (Manning et al., 2006). Disruption of *SPL-CNR* in the *Cnr* mutant results in fruits that fail to soften or undergo colour change at the ripening stage, and this has been related to changes in cell wall composition and cell-cell adhesion (Eriksson et al., 2004; Lai et al., 2020). Although both species of tomato turn red and soften at maturity, that is, neither species displays the extreme *Cnr* phenotype normally, there are quantitative differences in fruit firmness between them. Two large-scale QTL mapping studies of wild and cultivated tomato advanced backcrosses discovered six QTL for fruit firmness, and wild tomato alleles at four of those QTL are shown to decrease fruit firmness (Tanksley et al., 1996; Doganlar et al., 2002). Because soft fruits are more easily damaged during harvest and less desirable to consumers, increasing fruit firmness for cultivated tomato is one target of breeding programs (Barrett et al., 2010). *SPL-CNR* might help increase fruit firmness through its role in cell-cell adhesion, and thus differences in *SPL-CNR* expression between these tomato species could be related to differences in fruit firmness, although many other loci are likely at play. Additionally, the established role of *SPL-CNR* in promoting cell-cell adhesion in tomato has led other authors to speculate that this gene might also play a role in dry fruit dehiscence (Eriksson et al., 2004). If this gene's function in cell-cell adhesion is conserved among diverse fruit types, then the difference in expression patterns for *SPL-CNR* between fruit types in our analysis is also suggestive of a potential role in dry fruit dehiscence. Including desert tobacco expression data, we observe roughly opposing patterns in *SPL-CNR* expression between dry and fleshy fruits (Figure 5G). *SPL-CNR* increases in expression as fleshy fruits enter the breaker stage, before they begin to soften. In contrast, we see a decrease in *SPL-CNR* expression as dry fruits approach dehiscence, where loss of cell adhesion allows the fruit to split open. Additional functional studies of this gene's role across dry-fruited species could help extend its established role in cell-cell adhesion and clarify confirm its potential role in dry fruit maturity dehiscence and the potential conservation of function across fruits.

In this study, we mapped RNAseq reads from both wild and cultivated tomato to the cultivated tomato reference genome (Hosmani et al., 2019), and in our searches for orthologous genes, we used cultivated tomato sequences as a proxy for both wild and cultivated tomato. This simplified our interspecific comparisons, and mitigated the fact that the genome assemblies of wild tomato are not thoroughly annotated (Razali et al., 2018). Although these decisions enabled better interspecific comparisons, it means we are unable to examine the role of gene duplications and mutations that may have arisen since wild and cultivated tomato split. However, this is unlikely to drastically affect our results since any gene duplications specific to a single species are filtered out of our interspecific comparisons.

Comparative transcriptome analysis reveals both core conserved fruit development genes, and dry- and fleshy-fruit-specific genes

By examining the expression patterns of single-copy, orthologous genes across wild and cultivated tomato, desert tobacco, Arabidopsis, and melon, we were able to find evidence for two groups of genes in dry and fleshy fruit development, which we have termed the core and accessory genes. The core genes comprise a set of 121 orthologs whose expression patterns in the pericarp are conserved among all five species, while the accessory genome includes 1,795 orthologs whose expression patterns are each similar within fruit types but which show difference between fruit types.

Not all of the 121 core genes have been thoroughly characterised, so at present it is not possible to give a full inventory of functions, but the list suggests common developmental mechanisms that may be necessary for pericarp development. Orthologs for many of these core genes have annotated functions in processes of cell division and cell wall synthesis including the gene *KNOLLE* (*AT1G08560*), which helps pattern the rate and plane of cell divisions (Lukowitz et al., 1996). However other prominent structural genes for cellulose synthase, pectin methylesterase, and pectin lyase, and microtubule organising proteins are also present (*CESA4*, *AT5G44030; PME5*, *AT5G47500; AT5G19730; CORD3*, *AT4G13370*; *CORD7*, *AT2G31920*; *FUSED*, *AT1G50240*). Other genes in this set have orthologs with annotated function in developmental patterning. For example, the Arabidopsis gene *ARABIDOPSIS CRINKLY4* (*ACR4*, *AT3G59420*) functions in pattern epidermal cells, root asymmetric cell divisions, and cuticle deposition, while *PERIANTHIA* (*AT1G68640*) helps determine floral organ number (Running and Meyerowitz, 1996; Watanabe et al., 2004; De Smet et al., 2008). Beyond the expected cell division and pattern genes we also found several brassinosteroid-related genes as well as *ARGONAUTE7* (AGO7, *AT1G69440*) in this set of core genes. *AGO7* is involved in tasiRNA formation and ultimately helps to regulate development progression from vegetative to reproductive stages as well as leaf morphology in an auxin dependent manner (Adenot et al., 2006; Montgomery et al., 2008). The genes *DWARF4* (*DWF4*, *AT3G50660*) and *TITAN-LIKE* (*TTL, AT4G24900*) are involved in brassinosteroid biosynthesis and growth-responses, respectively (Azpiroz et al., 1998; Lu et al., 2012). The dwarfed phenotype of *dwf4* mutants is related to reduced cell elongation but not cell division, whereas the *ttl* mutant was first characterised based on an endosperm nuclear division defect. The dry and fleshy fruits studied here differ in a number of ways from one another, but overall size, especially in the pericarp tissues we sampled is one very conspicuous difference (Gillaspy et al., 1993). The overall size of a plant organ can be decomposed into the number of cells present and their sizes, so it is interesting that the brassinosteroid related genes in the core set of genes have complementary effects, modulating cell size and nuclear divisions, respectively.

Although our dataset includes eudicot plants from phylogenetically distant families, we believe that the addition of more taxa could help refine this set of core and accessory genes. Because our method is based on patterns among shared, single-copy orthologs however, including additional very distantly related plants or plants with extremely reduced genomes would not be beneficial. We examined patterns of expression for approximately 5,000 orthologs in our five-species comparisons, and this number of orthologs is based not only on the presence of orthologs among all species, but also our ability to confidently identify orthologs. Including more taxa would likely reduce the number of true single-copy orthologs, but because the determination of orthology is based upon finding clusters of proteins with similar sequence and resolving a phylogenetic relationship among them, additional genes could produce more informative gene trees and help increase ortholog numbers.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/, PRJNA646747.

Author contributions

AR: Conceptualization, Methodology, Software, Formal Analysis, Investigation, Data Curation, Writing, Visualisation. DM: Conceptualization, Methodology. JL: Methodology. AL: Conceptualization, Writing - Review and Editing, Supervision,

Project Administration, Funding Acquisition. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the National Science Foundation (grant number IOS1456109) and the University of California, Riverside.

Acknowledgments

We thank the UCR Institute for Integrative Genome Biology (IIGB) Genomics Core for their assistance in various technical aspects of RNAseq library prep and the staff at the UCR High Performance Computing Cluster for their logistical and infrastructural support.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.954929/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Summary of clustered gene expression profiles for genes with conserved patterns between wild and cultivated tomato. Violin plots of normalised expression by developmental stage for each cluster are shown on the left and gene ontology (GO) enrichment plots for the genes in the corresponding cluster are shown on the right. GO term descriptions to the left of the enrichment graphs are truncated for space and sorted by p-value. The bars are colored by the number of genes assigned to each GO term with legends in the lower right of each graph. Stages of fruit development in the axis of **(B, D)** are numbered sequentially followed by "Br" for breaker stage and "RR" for red ripe stage.

SUPPLEMENTARY FIGURE 2
Summary of clustered gene expression profiles for genes with divergent patterns between wild and cultivated tomato. Violin plots of normalised expression by developmental stage for each cluster are shown on the left and gene ontology (GO) enrichment plots for the genes in the corresponding cluster are shown on the right. Profiles for wild tomato are shown in blue, while profiles for cultivated tomato are shown in red. GO term descriptions to the left of the enrichment graphs are truncated for space and sorted by p-value. The bars are colored by the number of genes assigned to each GO term with legends in the lower right of each graph. Stages of fruit development in the axis of **(B, D)** are numbered sequentially followed by "Br" for breaker stage and "RR" for red ripe stage.

SUPPLEMENTARY FIGURE 3
Summary of clustered gene expression profiles for genes with divergent patterns between wild and cultivated tomato. Violin plots of normalised expression by developmental stage for each cluster are shown on the left and gene ontology (GO) enrichment plots for the genes in the corresponding cluster are shown on the right. Profiles for wild tomato are shown in blue, while profiles for cultivated tomato are shown in red. GO term descriptions to the left of the enrichment graphs are truncated for space and sorted by p-value. The bars are colored by the number of genes assigned to each GO term with legends in the lower right of each graph. Stages of fruit development in the axis of **(B, D)** are numbered sequentially followed by "Br" for breaker stage and "RR" for red ripe stage.

SUPPLEMENTARY FIGURE 4
Expression profiles for selected regulatory genes. Normalised counts of gene expression are represented by violin plots. Genes with statistically significant (FDR<0.01) differential expression across stages are shown in bold. Where expression pattern is better described by individual species trends (based on a likelihood ratio test), wild tomato violin plots are shown in blue and cultivated tomato plots are shown in red, otherwise the common pattern is shown in red. Stages of fruit development on the X-axis are numbered sequentially followed by "Br" for breaker stage and "RR" for red ripe stage. Note that panels have independent Y-axis to maximise readability.

SUPPLEMENTARY FIGURE 5
GO Enrichment analysis for desert tobacco gene expression clusters in . GO term descriptions to the left of the enrichment graphs are truncated for space and sorted by p-value. The bars are colored by the number of genes assigned to each GO term with legends in the lower right of each graph. Stages of fruit development in the axis of B-GD are numbered sequentially followed by "Tr" for transition to mature stage.

SUPPLEMENTARY FIGURE 6
Summary of clustered gene expression profiles for genes with conserved patterns among the three solanaceous species. Violin plots of normalised expression by developmental stage for each cluster are shown on the left and gene ontology (GO) enrichment plots for the genes in the corresponding cluster are shown on the right. GO term descriptions to the left of the enrichment graphs are truncated for space and sorted by p-value. The bars are colored by the number of genes assigned to each GO term with legends in the lower right of each graph. Stages of fruit development in the axis are numbered sequentially followed by "Tr" transition to mature stage.

SUPPLEMENTARY FIGURE 7
Summary of clustered gene expression profiles for genes with divergent patterns by fruit tpe among the three solanaceous species. Violin plots of normalised expression by developmental stage for each cluster are shown on the left and gene ontology (GO) enrichment plots for the genes in the corresponding cluster are shown on the right. Profiles for dry fruits are shown in yellow, while profiles for both tomato species are shown in red. GO term descriptions to the left of the enrichment graphs are truncated for space and sorted by p-value. The bars are colored by the number of genes assigned to each GO term with legends in the lower right of each

graph. Stages of fruit development in the axis are numbered sequentially followed by "Tr" for transition to mature stage.

SUPPLEMENTARY FIGURE 8
Summary of clustered gene expression profiles for genes with conserved patterns among the five species. Violin plots of normalised expression by developmental stage for each cluster are shown on the left and gene ontology (GO) enrichment plots for the genes in the corresponding cluster are shown on the right. GO term descriptions to the left of the enrichment graphs are truncated for space and sorted by p-value. The bars are colored by the number of genes assigned to each GO term with legends in the lower right of each graph. Stages of fruit development in the axis are numbered sequentially followed by "Tr" transition to mature stage.

SUPPLEMENTARY FIGURE 9
Summary of differentially expressed orthologous genes. Representative clusters of differentially expressed genes with patterns that differ between dry and fleshy fruited taxa are presented with violin plots of normalised expression at each stage of development along with a GO enrichment analysis of the genes in that cluster. GO term descriptions to the left of the enrichment graphs are truncated for space and sorted by p-value. The bars are colored by the number of genes assigned to each GO term with legends in the lower right of the graph. Stages of fruit development in the axis of **(A, C, E, G)** are numbered sequentially followed by "Tr" for transition to mature stage.

SUPPLEMENTARY FILE 1
List of gene names for conserved orthologous genes.

SUPPLEMENTARY FILE 2
List of gene names for divergent orthologous genes.

SUPPLEMENTARY FILE 3
Mapping statistics for transcriptome data. Columns are the NCBI SRA Accession, species of origin, developmental stage, replicate number, number of mapped reads, percent of reads mapped, mean mapping coverage, percent of mapped reads hitting exons, percent of mapped reads hitting introns, and percent of mapped reads hitting other features.

SUPPLEMENTARY FILE 4
An Excel file with expression information for the expression profiles plotted in Figures 1, 3, 4, and 9. Data for Figures 1 and 3 show the single gene ID based on the reference annotation, whereas the data for Figure 4 shows both gene IDs in the orthologous pair. In the data for Figure 9 each orthogroup is listed twice, one line for fleshy fruits and another for dry fruits. Only the gene IDs for corresponding fleshy or dry fruited species are shown on a given line, but all genes in the orthogroup were considered orthologous. Expression is shown as a normalized Z-score as determined by DESeq2 and the given model of differential expression testing. Stages listed correspond to the figure labels and cluster identities correspond.

References

Adams, P., Davies, J. N., and Winsor, G. W. (1978). Effects of nitrogen, potassium and magnesium on the quality and chemical composition of tomatoes grown in peat. *J. Hortic. Sci.* 53, 115–122. doi: 10.1080/00221589.1978.11514805

Adenot, X., Elmayan, T., Lauressergues, D., Boutet, S., Bouché, N., Gasciolli, V., et al. (2006). DRB4-dependent TAS3 trans-acting siRNAs control leaf morphology through AGO7. *Curr. Biol.* 16, 927–932. doi: 10.1016/j.cub.2006.03.035

Akihiro, T., Koike, S., Tani, R., Tominaga, T., Watanabe, S., Iijima, Y., et al. (2008). Biochemical mechanism on GABA accumulation during fruit development in tomato. *Plant Cell Physiol.* 49, 1378–1389. doi: 10.1093/pcp/pcn113

Alexa, A., and Rahnenfuhrer, J. (2016). "topGO: Enrichment analysis for gene ontology," in *R package version 2.28. 0* (BioConductor). doi: 10.18129/B9.bioc.topGO

Armbruster, W. S. (2014). Floral specialization and angiosperm diversity: phenotypic divergence, fitness trade-offs and realized pollination accuracy. *AoB. Plants* 6, 1–24. doi: 10.1093/aobpla/plu003

Attwood, T. K., Coletta, A., Muirhead, G., Pavlopoulou, A., Philippou, P. B., Popov, I., et al. (2012). The PRINTS database: a fine-grained protein sequence annotation and analysis resource–its status in 2012. *Database* 2012, bas019. doi: 10.1093/database/bas019

Azpiroz, R., Wu, Y., LoCascio, J. C., and Feldmann, K. A. (1998). An arabidopsis brassinosteroid-dependent mutant is blocked in cell elongation. *Plant Cell* 10, 219–230. doi: 10.1105/tpc.10.2.219

Barrett, D. M., Beaulieu, J. C., and Shewfelt, R. (2010). Color, flavor, texture, and nutritional quality of fresh-cut fruits and vegetables: desirable levels, instrumental and sensory measurement, and the effects of processing. *Crit. Rev. Food Sci. Nutr.* 50, 369–389. doi: 10.1080/10408391003626322

Berardini, T. Z., Reiser, L., Li, D., Mezheritsky, Y., Muller, R., Strait, E., et al. (2015). The arabidopsis information resource: Making and mining the "gold standard" annotated reference plant genome. *Genesis* 53, 474–485. doi: 10.1002/dvg.22877

Blanca, J., Montero-Pau, J., Sauvage, C., Bauchet, G., Illa, E., Díez, M. J., et al. (2015). Genomic variation in tomato, from wild ancestors to contemporary breeding accessions. *BMC Genomics* 16, 257. doi: 10.1186/s12864-015-1444-1

Bouché, N., Lacombe, B., and Fromm, H. (2003). GABA signaling: a conserved and ubiquitous mechanism. *Trends Cell Biol.* 13, 607–610. doi: 10.1016/j.tcb.2003.10.001

Bourdon, M., Frangne, N., Mathieu-Rivet, E., Nafati, M., Cheniclet, C., Renaudin, J.-P., et al. (2010). ""Endoreduplication and growth of fleshy fruits,"," in *Progress in botany 71*. Eds. U. Lüttge, W. Beyschlag, B. Büdel and D. Francis (Berlin, Heidelberg: Springer Berlin Heidelberg), 101–132.

Bremer, B., and Eriksson, O. (1992). Evolution of fruit characters and dispersal modes in the tropical family rubiaceae. *Biol. J. Linn. Soc Lond.* 47, 79–95. doi: 10.1111/j.1095-8312.1992.tb00657.x

Chayut, N., Yuan, H., Ohali, S., Meir, A., Sa'ar, U., Tzuri, G., et al. (2017). Distinct mechanisms of the ORANGE protein in controlling carotenoid flux. *Plant Physiol.* 173, 376–389. doi: 10.1104/pp.16.01256

Chayut, N., Yuan, H., Ohali, S., Meir, A., Yeselson, Y., Portnoy, V., et al. (2015). A bulk segregant transcriptome analysis reveals metabolic and cellular processes associated with orange allelic variation and fruit β-carotene accumulation in melon fruit. *BMC Plant Biol.* 15, 274. doi: 10.1186/s12870-015-0661-8

Chen, G., Hackett, R., Walker, D., Taylor, A., Lin, Z., and Grierson, D. (2004). Identification of a specific isoform of tomato lipoxygenase (TomloxC) involved in the generation of fatty acid-derived flavor compounds. *Plant Physiol.* 136, 2641–2651. doi: 10.1104/pp.104.041608

Clausing, G., Meyer, K., and Renner, S. S. (2000). Correlations among fruit traits and evolution of different fruits within melastomataceae. *Botanical. J. Linn. Soc.* 133, 303–326. doi: 10.1111/j.1095-8339.2000.tb01548.x

Cox, H. T. (1948). Studies in the comparative anatomy of the ericales i. ericaceae-subfamily rhododendroideae. *Am. Midland. Nat.* 39, 220. doi: 10.2307/2421443

Crepet, W. L., and Niklas, K. J. (2009). Darwin's second "abominable mystery": Why are there so many angiosperm species? *Am. J. Bot.* 96, 366–381. doi: 10.3732/ajb.0800126

De Smet, I., Vassileva, V., De Rybel, B., Levesque, M. P., Grunewald, W., Van Damme, D., et al. (2008). Receptor-like kinase ACR4 restricts formative cell divisions in the arabidopsis root. *Science* 322, 594–597. doi: 10.1126/science.1160158

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. doi: 10.1093/bioinformatics/bts635

Doganlar, S., Frary, A., Ku, H.-M., and Tanksley, S. D. (2002). Mapping quantitative trait loci in inbred backcross lines of *Lycopersicon pimpinellifolium* (LA1589). *Genome* 45, 1189–1202. doi: 10.1139/g02-091

El-Gebali, S., Mistry, J., Bateman, A., Eddy, S. R., Luciani, A., Potter, S. C., et al. (2019). The pfam protein families database in 2019. *Nucleic Acids Res.* 47, D427–D432. doi: 10.1093/nar/gky995

Emms, D. M., and Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* 20, 238. doi: 10.1186/s13059-019-1832-y

Eriksson, E. M., Bovy, A., Manning, K., Harrison, L., Andrews, J., De Silva, J., et al. (2004). Effect of the colorless non-ripening mutation on cell wall biochemistry and gene expression during tomato fruit development and ripening. *Plant Physiol.* 136, 4184–4197. doi: 10.1104/pp.104.045765

FAO (2017) *Food balance sheet. FAO global statistical yearbook*. Available at: http://www.fao.org/faostat/en/#data/FBS (Accessed May 4, 2020).

Ferrándiz, C., Liljegren, S. J., and Yanofsky, M. F. (2000). Negative regulation of the *SHATTERPROOF* genes by FRUITFULL during *Arabidopsis* fruit development. *Science* 289, 436–438. doi: 10.1126/science.289.5478.436

Fischer, D. S., Theis, F. J., and Yosef, N. (2018). Impulse model-based differential expression analysis of time course sequencing data. *Nucleic Acids Res.* 46, e119. doi: 10.1093/nar/gky675

Gao, L., Gonda, I., Sun, H., Ma, Q., Bao, K., Tieman, D. M., et al. (2019). The tomato pan-genome uncovers new genes and a rare allele regulating fruit flavor. *Nat. Genet.* 51, 1044–1051. doi: 10.1038/s41588-019-0410-2

Garcia-Mas, J., Benjak, A., Sanseverino, W., Bourgeois, M., Mir, G., González, V. M., et al. (2012). The genome of melon (*Cucumis melo* l.). *Proc. Natl. Acad. Sci. U. S. A.* 109, 11872–11877. doi: 10.1073/pnas.1205415109

Gillaspy, G., Ben-David, H., and Gruissem, W. (1993). Fruits: A developmental perspective. *Plant Cell* 5, 1439–1451. doi: 10.2307/3869794

Gimenez, E., Castañeda, L., Pineda, B., Pan, I. L., Moreno, V., Angosto, T., et al. (2016). *TOMATO AGAMOUS1* and *ARLEQUIN/TOMATO AGAMOUS-LIKE1* MADS-box genes have redundant and divergent functions required for tomato reproductive development. *Plant Mol. Biol.* 91, 513–531. doi: 10.1007/s11103-016-0485-4

Giménez, E., Pineda, B., Capel, J., Antón, M. T., Atarés, A., Pérez-Martín, F., et al. (2010). Functional analysis of the arlequin mutant corroborates the essential role of the Arlequin/TAGL1 gene during reproductive development of tomato. *PloS One* 5, e14427. doi: 10.1371/journal.pone.0014427

Givnish, T. J., Pires, J. C., Graham, S. W., McPherson, M. A., Prince, L. M., Patterson, T. B., et al. (2005). Repeated evolution of net venation and fleshy fruits among monocots in shaded habitats confirms *a priori* predictions: evidence from an *ndhF* phylogeny. *Proc. Biol. Sci.* 272, 1481–1490. doi: 10.1098/rspb.2005.3067

Haft, D. H., Loftus, B. J., Richardson, D. L., Yang, F., Eisen, J. A., Paulsen, I. T., et al. (2001). TIGRFAMs: a protein family resource for the functional identification of proteins. *Nucleic Acids Res.* 29, 41–43. doi: 10.1093/nar/29.1.41

Hosmani, P. S., Flores-Gonzalez, M., van de Geest, H., Maumus, F., Bakker, L. V., Schijlen, E., et al. (2019). An improved *de novo* assembly and annotation of the tomato reference genome using single-molecule sequencing, Hi-c proximity ligation and optical maps. *Cold Spring Harbor Lab.*, 767764. doi: 10.1101/767764

Houben, M., and Van de Poel, B. (2019). 1-Aminocyclopropane-1-Carboxylic acid oxidase (ACO): The enzyme that makes the plant hormone ethylene. *Front. Plant Sci.* 10, 695. doi: 10.3389/fpls.2019.00695

Huang, B., Routaboul, J.-M., Liu, M., Deng, W., Maza, E., Mila, I., et al. (2017). Overexpression of the class d MADS-box gene *Sl-AGL11* impacts fleshy tissue differentiation and structure in tomato fruits. *J. Exp. Bot.* 68, 4869–4884. doi: 10.1093/jxb/erx303

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010

Kaufman, L., and Rousseeuw, P. J. (2005). *Finding groups in data: An introduction to cluster analysis* (Hoboken, NJ, USA: Wiley).

Kelly, S., and Maini, P. K. (2013). DendroBLAST: approximate phylogenetic trees in the absence of multiple sequence alignments. *PloS One* 8, e58537. doi: 10.1371/journal.pone.0058537

Knapp, S. (2002). Tobacco to tomatoes: a phylogenetic perspective on fruit diversity in the solanaceae. *J. Exp. Bot.* 53, 2001–2022. doi: 10.1093/jxb/erf068

Krueger, F. (2012) *Trim galore: a wrapper tool around cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files, with some extra functionality for MspI-digested RRBS-type (Reduced representation bisufite-seq) libraries*. Available at: https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/.

Lai, T., Wang, X., Ye, B., Jin, M., Chen, W., Wang, Y., et al. (2020). Molecular and functional characterization of the SBP-box transcription factor SPL-CNR in tomato fruit ripening and cell death. *J. Exp. Bot.* 71, 2995–3011. doi: 10.1093/jxb/eraa067

Li, X., Tieman, D., Liu, Z., Chen, K., and Klee, H. J. (2020). Identification of a lipase gene with a role in tomato fruit short-chain fatty acid-derived flavor volatiles by genome-wide association. *Plant J.* 104, 631–644. doi: 10.1111/tpj.14951

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. doi: 10.1186/s13059-014-0550-8

Lukowitz, W., Mayer, U., and Jürgens, G. (1996). Cytokinesis in the arabidopsis embryo involves the syntaxin-related KNOLLE gene product. *Cell* 84, 61–71. doi: 10.1016/S0092-8674(00)80993-9

Lu, X., Li, Y., Su, Y., Liang, Q., Meng, H., Li, S., et al. (2012). An arabidopsis gene encoding a C2H2-domain protein with alternatively spliced transcripts is essential for endosperm development. *J. Exp. Bot.* 63, 5935–5944. doi: 10.1093/jxb/ers243

MacGregor, K. B., Shelp, B. J., Peiris, S., and Bown, A. W. (2003). Overexpression of glutamate decarboxylase in transgenic tobacco plants deters feeding by phytophagous insect larvae. *J. Chem. Ecol.* 29, 2177–2182. doi: 10.1023/A:1025650914947

Manning, K., Tör, M., Poole, M., Hong, Y., Thompson, A. J., King, G. J., et al. (2006). A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. *Nat. Genet.* 38, 948–952. doi: 10.1038/ng1841

Mejía, N., Soto, B., Guerrero, M., Casanueva, X., Houel, C., Miccono, M., et al. (2011). Molecular, genetic and transcriptional evidence for a role of VvAGL11 in stenospermocarpic seedlessness in grapevine. *BMC Plant Biol.* 11, 57. doi: 10.1186/1471-2229-11-57

Mizzotti, C., Rotasperti, L., Moretto, M., Tadini, L., Resentini, F., Galliani, B. M., et al. (2018). Time-course transcriptome analysis of arabidopsis siliques discloses genes essential for fruit development and maturation. *Plant Physiol.* 178, 1249–1268. doi: 10.1104/pp.18.00727

Monforte, A. J., Diaz, A., Caño-Delgado, A., and van der Knaap, E. (2014). The genetic basis of fruit morphology in horticultural crops: lessons from tomato and melon. *J. Exp. Bot.* 65, 4625–4637. doi: 10.1093/jxb/eru017

Montgomery, T. A., Howell, M. D., Cuperus, J. T., Li, D., Hansen, J. E., Alexander, A. L., et al. (2008). Specificity of ARGONAUTE7-miR390 interaction and dual functionality in TAS3 trans-acting siRNA formation. *Cell* 133, 128–141. doi: 10.1016/j.cell.2008.02.033

Nambeesan, S., Datsenka, T., Ferruzzi, M. G., Malladi, A., Mattoo, A. K., and Handa, A. K. (2010). Overexpression of yeast spermidine synthase impacts ripening, senescence and decay symptoms in tomato: Polyamines enhance shelf life in tomato. *Plant J.* 63, 836–847. doi: 10.1111/j.1365-313X.2010.04286.x

Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/molbev/msu300

Ocarez, N., and Mejía, N. (2016). Suppression of the d-class MADS-box *AGL11* gene triggers seedlessness in fleshy fruits. *Plant Cell Rep.* 35, 239–254. doi: 10.1007/s00299-015-1882-x

Pabón-Mora, N., and Litt, A. (2011). Comparative anatomical and developmental analysis of dry and fleshy fruits of solanaceae. *Am. J. Bot.* 98, 1415–1436. doi: 10.3732/ajb.1100097

Pan, I. L., McQuinn, R., Giovannoni, J. J., and Irish, V. F. (2010). Functional diversification of *AGAMOUS* lineage genes in regulating tomato flower and fruit development. *J. Exp. Bot.* 61, 1795–1806. doi: 10.1093/jxb/erq046

Pantano, L. (2019) *DEGreport: Report of DEG analysis*. Available at: http://lpantano.github.io/DEGreport/.

Pinyopich, A., Ditta, G. S., Savidge, B., Liljegren, S. J., Baumann, E., Wisman, E., et al. (2003). Assessing the redundancy of MADS-box genes during carpel and ovule development. *Nature* 424, 85–88. doi: 10.1038/nature01741

Plunkett, G., Soltis, D., and Soltis, P. (1997). Clarification of the relationship between apiaceae and araliaceae based on matK and rbcL sequence data. *Am. J. Bot.* 84, 565. doi: 10.2307/2446032

Pnueli, L., Hareven, D., Rounsley, S. D., Yanofsky, M. F., and Lifschitz, E. (1994). Isolation of the tomato *AGAMOUS* gene *TAG1* and analysis of its homeotic role in transgenic plants. *Plant Cell* 6, 163–173. doi: 10.1105/tpc.6.2.163

Razali, R., Bougouffa, S., Morton, M. J. L., Lightfoot, D. J., Alam, I., Essack, M., et al. (2018). The genome sequence of the wild tomato *Solanum pimpinellifolium* provides insights into salinity tolerance. *Front. Plant Sci.* 9, 1402. doi: 10.3389/fpls.2018.01402

R Core Team (2019). *R: A language and environment for statistical computing*. The R Project for Statistical Computing

Regal, P. J. (1977). Ecology and evolution of flowering plant dominance. *Science* 196, 622–629. doi: 10.1126/science.196.4290.622

Ripoll, J.-J., Zhu, M., Brocke, S., Hon, C. T., Yanofsky, M. F., Boudaoud, A., et al (2019) Growth dynamics of the arabidopsis fruit is mediated by cell expansion *Proc. Natl. Acad. Sci. U. S. A.* 116 (50), 25333–25342. doi: 10.1073/pnas.1914096116

Running, M. P., and Meyerowitz, E. M. (1996). Mutations in the *PERIANTHIA* gene of arabidopsis specifically alter floral organ number and initiation pattern. *Development* 122, 1261–1269. doi: 10.1242/dev.122.4.1261

Sander, J., Schultze, J. L., and Yosef, N. (2017). ImpulseDE: detection of differentially expressed genes in time series data using impulse models. *Bioinformatics* 33, 757–759. doi: 10.1093/bioinformatics/btw665

Shen, J., Tieman, D., Jones, J. B., Taylor, M. G., Schmelz, E., Huffaker, A., et al. (2014). A 13-lipoxygenase, TomloxC, is essential for synthesis of C5 flavour volatiles in tomato. *J. Exp. Bot.* 65, 419–428. doi: 10.1093/jxb/ert382

Shin, A.-Y., Kim, Y.-M., Koo, N., Lee, S. M., Nahm, S., and Kwon, S.-Y. (2017). Transcriptome analysis of the oriental melon (*Cucumis melo* l. var. *makuwa*) during fruit development. *PeerJ* 5, e2834. doi: 10.7717/peerj.2834

Sigrist, C. J. A., de Castro, E., Cerutti, L., Cuche, B. A., Hulo, N., Bridge, A., et al. (2013). New and continuing developments at PROSITE. *Nucleic Acids Res.* 41, D344–D347. doi: 10.1093/nar/gks1067

Smykal, P., Gennen, J., De Bodt, S., Ranganath, V., and Melzer, S. (2007). Flowering of strict photoperiodic *Nicotiana* varieties in non-inductive conditions by transgenic approaches. *Plant Mol. Biol.* 65, 233–242. doi: 10.1007/s11103-007-9211-6

Spalik, K., Wojewódzka, A., and Downie, S. R. (2001). The evolution of fruit in scandiceae subtribe scandicinae (Apiaceae). *Can. J. Bot.* 79, 1358–1374. doi: 10.1139/b01-116

Tanksley, S. D. (2004). The genetic, developmental, and molecular bases of fruit size and shape variation in tomato. *Plant Cell* 16 Suppl, S181–S189. doi: 10.1105/tpc.018119

Tanksley, S. D., Grandillo, S., Fulton, T. M., Zamir, D., Eshed, Y., Petiard, V., et al. (1996). Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative l. pimpinellifolium. *Theor. Appl. Genet.* 92, 213–224. doi: 10.1007/BF00223378

Thompson, A. J., Tor, M., Barry, C. S., Vrebalov, J., Orfila, C., Jarvis, M. C., et al. (1999). Molecular and genetic characterization of a novel pleiotropic tomato-ripening mutant. *Plant Physiol.* 120, 383–390. doi: 10.1104/pp.120.2.383

Tieman, D., Bliss, P., McIntyre, L. M., Blandon-Ubeda, A., Bies, D., Odabasi, A. Z., et al. (2012). The chemical interactions underlying tomato flavor preferences. *Curr. Biol.* 22, 1035–1039. doi: 10.1016/j.cub.2012.04.016

Watanabe, M., Tanaka, H., Watanabe, D., Machida, C., and Machida, Y. (2004). The ACR4 receptor-like kinase is required for surface formation of epidermis-related tissues in arabidopsis thaliana. *Plant J.* 39, 298–308. doi: 10.1111/j.1365-313X.2004.02132.x

Weber, A. (2004). "Gesneriaceae," in *Flowering plants · dicotyledons*, 63–158. doi: 10.1007/978-3-642-18617-2_8

Xu, S., Brockmöller, T., Navarro-Quezada, A., Kuhl, H., Gase, K., Ling, Z., et al. (2017). Wild tobacco genomes reveal the evolution of nicotine biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* 114, 6133–6138. doi: 10.1073/pnas.1700073114

Zhang, H., Wang, H., Yi, H., Zhai, W., Wang, G., and Fu, Q. (2016). Transcriptome profiling of Cucumis melo fruit development and ripening. *Hortic Res* 3, 16014.

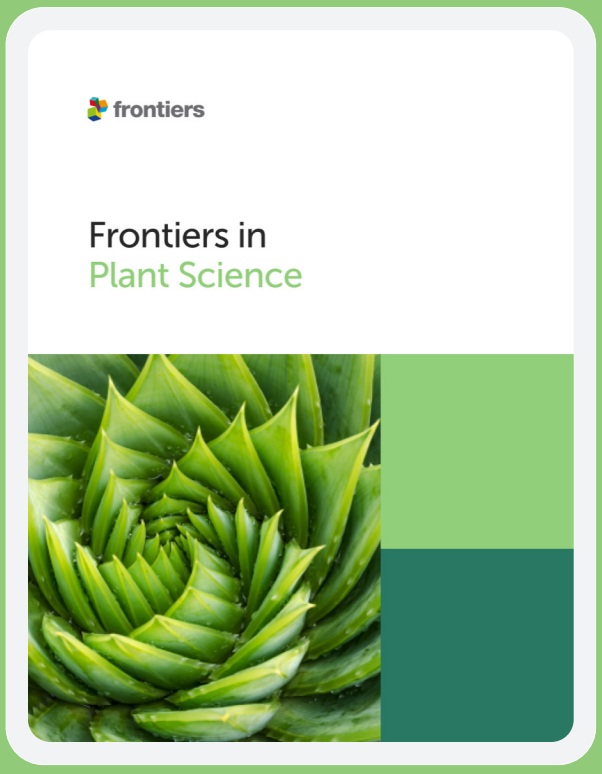
Frontiers in Plant Science

Cultivates the science of plant biology and its applications

The most cited plant science journal, which advances our understanding of plant biology for sustainable food security, functional ecosystems and human health.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact