METABOLISM OF FRUIT VOLATILE ORGANIC COMPOUNDS

EDITED BY: Jinhe Bai, María José Jordán and Jian Li

PUBLISHED IN: Frontiers in Plant Science

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88976-444-0
DOI 10.3389/978-2-88976-444-0

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

METABOLISM OF FRUIT VOLATILE ORGANIC COMPOUNDS

Topic Editors:
Jinhe Bai, Horticultural Research Laboratory (USDA-ARS), United States
María José Jordán, Instituto Murciano de Investigación y Desarrollo Agrario y Alimentario (IMIDA), Spain
Jian Li, Beijing Technology and Business University, China

Citation: Bai, J., Jordán, M. J., Li, J., eds. (2022). Metabolism of Fruit Volatile Organic Compounds. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88976-444-0

Table of Contents

05 ***Editorial: Metabolism of Fruit Volatile Organic Compounds***
Jinhe Bai, María J. Jordán and Jian Li

08 ***Volatiles Influencing Sensory Attributes and Bayesian Modeling of the Soluble Solids–Sweetness Relationship in Strawberry***
Zhen Fan, Anne Plotto, Jinhe Bai and Vance M. Whitaker

21 ***Huanglongbing and Foliar Spray Programs Affect the Chemical Profile of "Valencia" Orange Peel Oil***
Xiuxiu Sun, Huqing Yang, Wei Zhao, Elise Bourcier, Elizabeth A. Baldwin, Anne Plotto, Mike Irey and Jinhe Bai

34 ***Dynamic Analyses of Transcriptome and Metabolic Profiling: Revealing Molecular Insight of Aroma Synthesis of Mango (*Mangifera indica *L. Var. Tainong)***
Ming Xin, Changbao Li, Hock Eng Khoo, Li Li, Xuemei He, Ping Yi, Yayuan Tang and Jian Sun

48 ***Metabolome and Transcriptome Integration Reveals Insights Into Flavor Formation of 'Crimson' Watermelon Flesh During Fruit Development***
Chengsheng Gong, Weinan Diao, Hongju Zhu, Muhammad Jawad Umer, Shengjie Zhao, Nan He, Xuqiang Lu, Pingli Yuan, Muhammad Anees, Dongdong Yang, M. O. Kaseb and Wenge Liu

61 ***Genetic Analysis of Methyl Anthranilate, Mesifurane, Linalool, and Other Flavor Compounds in Cultivated Strawberry (*Fragaria × ananassa*)***
Christopher R. Barbey, Maxwell H. Hogshead, Benjamin Harrison, Anne E. Schwartz, Sujeet Verma, Youngjae Oh, Seonghee Lee, Kevin M. Folta and Vance M. Whitaker

78 ***Comparative Study of Fig Volatile Compounds Using Headspace Solid-Phase Microextraction-Gas Chromatography/Mass Spectrometry: Effects of Cultivars and Ripening Stages***
Kahina Zidi, Djamel Edine Kati, Mostapha Bachir-bey, Manon Genva and Marie-Laure Fauconnier

93 ***Genome-Wide Identification and Functional Characterization of the Trans-Isopentenyl Diphosphate Synthases Gene Family in* Cinnamomum camphora**
Zerui Yang, Chunzhu Xie, Ting Zhan, Linhuan Li, Shanshan Liu, Yuying Huang, Wenli An, Xiasheng Zheng and Song Huang

109 ***Two Main Biosynthesis Pathways Involved in the Synthesis of the Floral Aroma of the Nacional Cocoa Variety***
Kelly Colonges, Juan-Carlos Jimenez, Alejandra Saltos, Edward Seguine, Rey Gastón Loor Solorzano, Olivier Fouet, Xavier Argout, Sophie Assemat, Fabrice Davrieux, Emile Cros, Renaud Boulanger and Claire Lanaud

133 ***Comparison of the Fruit Volatile Profiles of Five Muscadine Grape Cultivars (*Vitis rotundifolia *Michx.) Using HS-SPME-GC/MS Combined With Multivariate Statistical Analysis***
Honghong Deng, Runmei He, Meicun Long, Yanmei Li, Yuanyuan Zheng, Lijin Lin, Dong Liang, Xiaoai Zhang, Ming'an Liao, Xiulan Lv, Qunxian Deng and Hui Xia

147 ***Transcriptome Analysis Revealed the Mechanism by Which Exogenous ABA Increases Anthocyanins in Blueberry Fruit During Veraison***
Tianyu Han, Wenlong Wu and Weilin Li

160 ***Ethylene and Auxin: Hormonal Regulation of Volatile Compound Production During Tomato (*Solanum lycopersicum *L.) Fruit Ripening***
Eric de Castro Tobaruela, Bruna Lima Gomes, Vanessa Caroline de Barros Bonato, Elis Silva de Lima, Luciano Freschi and Eduardo Purgatto

frontiers | Frontiers in Plant Science

EDITORIAL
published: 02 June 2022
doi: 10.3389/fpls.2022.873515

Editorial: Metabolism of Fruit Volatile Organic Compounds

Jinhe Bai[1]*, *María J. Jordán*[2] and *Jian Li*[3]

[1] Horticultural Research Laboratory, USDA-ARS, Fort Pierce, FL, United States, [2] Research Group on Rainfed Agriculture for Rural Development, Department of Rural Development, Oenology and Sustainable Agriculture, Murcia Institute of Agri-Food Research and Development (IMIDA), Murcia, Spain, [3] Engineering and Technology Research Center, Beijing Technology and Business University, Beijing, China

Keywords: genotype, biotic, abiotic, development, ripening, plant hormone, flavor, aroma

Editorial on the Research Topic

Metabolism of Fruit Volatile Organic Compounds

Volatile organic compounds (VOCs) are essential for fruit flavor, and along with sugars and acids, play a key role in the perception and acceptability of fruits by consumers. Over 1,700 VOCs have been identified from 90 different plant families (Knudsen et al., 2006). Attending to the biosynthesis pathways, these components can be divided into several classes: terpenoids, phenylpropanoids/benzenoids, fatty acid derivatives, and amino acid derivatives (**Figure 1**). Volatile blends in plants are complex and differ among species and even cultivars (Barbey et al.; Colonges et al.; Deng et al.; Fan et al.; Zidi et al.). These volatile profiles can be affected by biotic and abiotic stresses (Sun et al.), plant hormone/regulator applications (Han et al.; Tobaruela et al.), and fruit mature/ripening stages (Gong et al.; Xin et al.). Some volatile compounds are not only important to aroma quality but also enhance or suppress sweetness perception in fresh fruit (Baldwin et al., 1998). Fan et al. identified 20 volatiles that enhance sweetness independently of sugars, such as benzaldehyde, (*E*)-2-pentenal, non-anal, mesifurane, ethyl butanoate, and ethyl hexanoate.

In the scope of applications of molecular markers in VOC research, two studies go in-depth to identify putative candidate genes involved in major volatile biosynthesis pathways using a genome-wide association study (GWAS). Barbey et al. integrated GWAS, transcriptomic, and metabolomic analyses to identify QTL and candidate genes associated with multiple aroma compounds in strawberry breeding populations. Novel fruit volatile QTLs were discovered for methyl anthranilate, methyl 2-hexenoate, methyl 2-methylbutyrate, mesifurane, and a shared QTL on Chr 3 was found for nine monoterpene and sesquiterpene compounds. These QTLs present new opportunities in breeding for improved flavor in commercial strawberries. Colonges et al. studied the genetic and biochemical determinism of the floral aroma of the modern *Nacional cocoa* variety from Ecuador. GWAS was conducted on a population of 152 genotypes of cocoa trees using simple sequence repeats (SSR), single-nucleotide polymorphisms (SNP) markers, biochemical compounds (in roasted and unroasted beans), and sensory evaluations. The results showed that the monoterpene biosynthesis pathway and the L-phenylalanine degradation pathway are mainly related to the floral note.

A VOC profile is the key quality attribute in fruit, and is frequently compared between species, cultivars, and hybrids. Deng et al. identified 44 VOCs from 5 muscadine grape cultivars. The significant VOC differences among cultivars, particularly the bronze color and the purple-black color, are due to the quantitative differences of 29 volatiles. The following VOCs were characterized for each cultivar: geraniol and cinnamyl alcohol for Alachua, ethyl (*E*)-2-butenoate and propyl acetate for Noble, green leaf volatiles for Carlos, Fry, and Granny Val. Zidi et al. identified 29

OPEN ACCESS

Edited and reviewed by:
Shan Lu,
Nanjing University, China

***Correspondence:**
Jinhe Bai
jinhe.bai@usda.gov

Specialty section:
This article was submitted to Plant Metabolism and Chemodiversity, a section of the journal Frontiers in Plant Science

Received: *10 February 2022*
Accepted: *27 April 2022*
Published: *02 June 2022*

Citation:
Bai J, Jordán MJ and Li J (2022) Editorial: Metabolism of Fruit Volatile Organic Compounds. Front. Plant Sci. 13:873515. doi: 10.3389/fpls.2022.873515

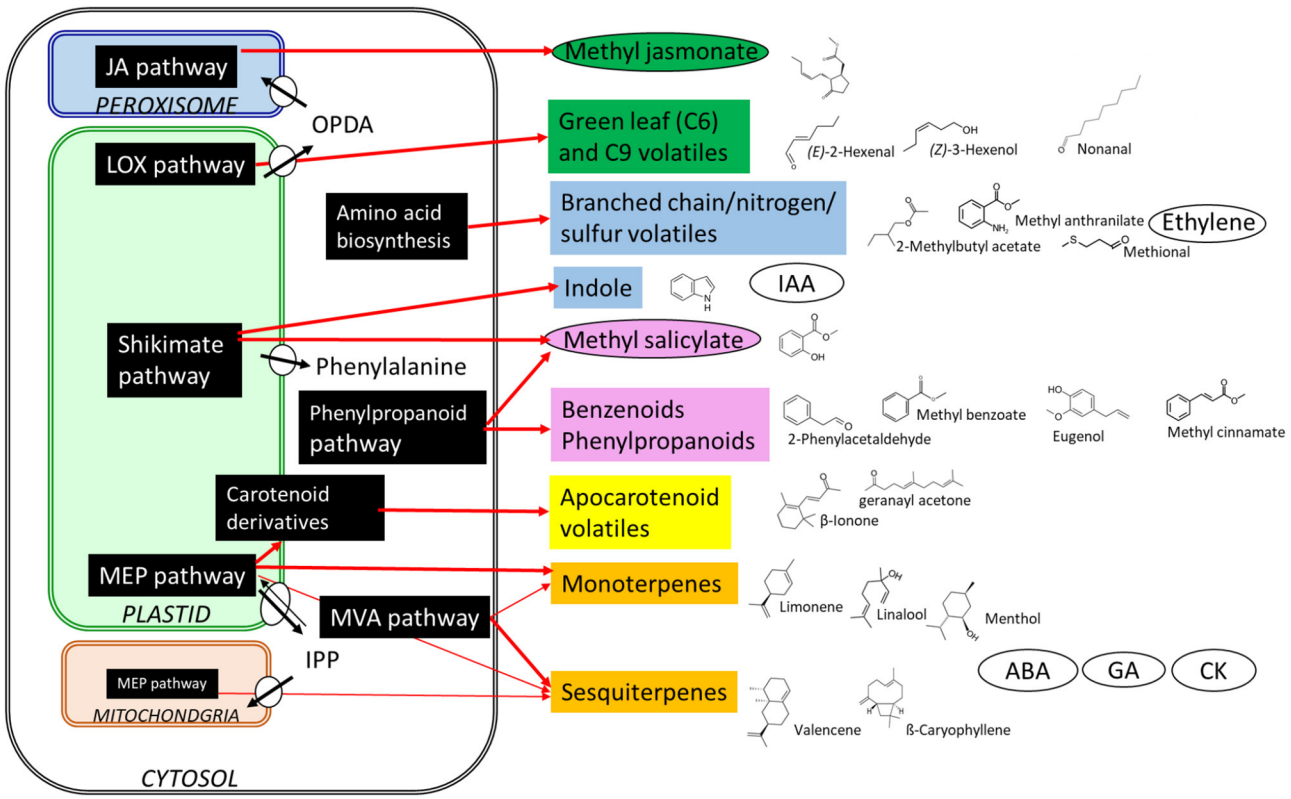

FIGURE 1 | Biosynthesis of different volatile organic compounds and hormones in plants. Shawn are structures and the biosynthetic origin. MVA, mevalonic acid; MEP, methylerythritol phosphate; JA, jasmonic acid; LOX, lipoxygenase; OPDA, cis-(+)-12-oxo-phytodienoic acid; IPP, isopentenyl diphosphate; IAA, indole-3-acetic acid; ABA, abscisic acid; GA, gibberellic acid; CK, cytokinin. Plant hormones are in the oval frames.

VOCs from 3 fig cultivars. Two out of the three had 20 VOCs, but only 14 compounds in the third cultivar. Aldehydes comprised the most abundant VOCs identified, while alcohols, ketones, and terpenes comprised the minor compounds found in all cultivars. Specific VOCs in each cultivar were identified and those can be used as fingerprints to distinguish cultivars. Yang et al. identified 10 trans-isopentenyl diphosphate synthase (TIDS) genes, which are involved in terpene diversity and the accumulation of terpenoids, in the genome of the *Cinnamomum camphora* borneol chemotype that was unevenly distributed on chromosomes. Synteny analysis revealed that the TIDS gene family in this species likely expanded through segmental duplication events.

Changes in VOCs during fruit development and ripening are another relevant issue for the integration of transcriptomic and metabolomic levels. Gong et al. generated the metabolomic and transcriptomic datasets during watermelon fruit growth and development with a total of 517 metabolites. Major VOCs, such as (*E,Z*)-2,6-nonadienal, (*E*)-2-nonenal, and (*Z*)-3-nonen-1-ol, were accumulated during fruit development with sugar increase and acid decline. They found that five alcohol dehydrogenase (ADH) genes were differentially expressed which were closely related to C9 alcohol/aldehyde metabolism. The integration of important genes in the glycolytic and LOX pathways provides molecular insights into flavor formation. Xin et al. identified 181 VOCs from mongo fruits throughout the development and ripening stages. These components, especially ethanol and (*E,Z*)-2,6-nonadienal, were the key aroma active compounds, and raised during ripening. RNA sequencing revealed 53,384 transcripts in mango, and catalytic activity, transferase activity, adenosine diphosphate binding, transcription factor activity, and oxidoreductase activity were differentially expressed. The content of α-pinene and the expression of genes and enzyme activities in terpenoid metabolic pathways gradually increased after harvest and generally plateaued when reaching the ripe stage. The integrative analyses also revealed potential molecular insights into fruit ripening and aroma biosynthesis. Zidi et al. analyzed the changes of VOC profiles in three fig cultivars. Different aroma descriptors were identified during the postharvest fruit development stages; fruity and green aromas were dominant in all cultivars, while a fatty aroma barely occurred. They demonstrated that certain VOCs differentiate the ripening stages in figs.

Fruit ripening is associated with hormone interaction (McAtee et al., 2013), and a complicated change of VOC profile occurs during ripening. However, Tobaruela et al. provided direct evidence that ethylene and auxin affect VOC

profiles independent of ripening. Two tomato cultivars, Micro-Tom and Sweet Grape, had different VOC profiles when reaching the ripening stage. However, both were similarly affected by exogenous ethylene and auxin treatments, with contrary responses to the hormones. Both hormones affect the aroma metabolism overlap, and the hormonal regulation is not cultivar-dependent. The hormones played essential roles in VOC metabolism, especially regarding carotenoid and fatty acid-derived compounds. Han et al. sprayed abscisic acid (ABA) at the late green stage of blueberries resulting in a substantial increase in anthocyanins. The direct reason for the result is due to the high expression of *CHS, CHI, DFR,* and *LDOX/ANS* during ripening. Among the transcription factors whose expression changed, *VcMYBA* had the greatest correlation with the changes of anthocyanin and corresponding synthetic structural genes.

Sun et al. demonstrated how biotic and abiotic factors impact VOC profiles in orange peel oil. Orange trees infected or not infected with huanglongbing disease were sprayed with fungicides/chemical agents. The infected fruit (HLB+) peel had lower concentrations of typical peel oil components, including valencene, octanal, and decanal, and was abundant in oxidative/dehydrogenated terpenes, such as carvone and limonene oxide. One of the enhanced foliar spray programs altered the VOC profile of HLB+ samples, resulting in peel oil similar to that of HLB- samples.

In summary, the work reported here gathered new information on the metabolism of VOCs at the genetic, transcriptional, hormonal, and metabolic levels and their impact on fruit flavor quality. We hope that this Research Topic will encourage scientists to make further efforts to reveal the mechanism of VOC changes and enhance fruit flavor quality.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

This work was partially supported by the U.S. National Institute of Food and Agriculture (NIFA; Grant Number 2018-70016-27453).

REFERENCES

Baldwin, E. A., Scott, J. W., Einstein, M. A., Malundo, T. M. M., Carr, B. T., Shewfelt, R. L., and Tandon, K. S. (1998). Relationship between sensory and instrumental analysis for tomato flavor. *J. Amer. Soc. Hort. Sci.* 123, 906–915. doi: 10.21273/JASHS.123.5.906

Knudsen, J. T., Eriksson, R., Gershenzon, J., and Ståhl, B. (2006). Diversity and distribution of floral scent. *Bot. Rev.* 72, 1–120. doi: 10.1663/0006-8101(2006)721:DADOFS2.0.CO;2

McAtee, P., Karim, S., Schaffer, R. J., and David, K. (2013). A dynamic interplay between phytohormones is required for fruit development, maturation, and ripening. *Front. Plant Sci.* 4:79. doi: 10.3389/fpls.2013.00079

Author Disclaimer: The mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture. USDA is an equal opportunity employer.

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 17 March 2021
doi: 10.3389/fpls.2021.640704

Volatiles Influencing Sensory Attributes and Bayesian Modeling of the Soluble Solids–Sweetness Relationship in Strawberry

*Zhen Fan[1], Anne Plotto[2], Jinhe Bai[2] and Vance M. Whitaker[1]**

[1] Horticultural Sciences Department, IFAS Gulf Coast Research and Education Center, University of Florida, Wimauma, FL, United States, [2] Horticultural Research Laboratory, USDA-ARS, Fort Pierce, FL, United States

Descriptive analysis *via* trained sensory panels has great power to facilitate flavor improvement in fresh fruits and vegetables. When paired with an understanding of fruit volatile organic compounds, descriptive analysis can help uncover the chemical drivers of sensory attributes. In the present study, 213 strawberry samples representing 56 cultivars and advanced selections were sampled over seven seasons and subjected to both sensory descriptive and chemical analyses. Principal component analysis and K-cluster analyses of sensory data highlighted three groups of strawberry samples, with one classified as superior with high sweetness and strawberry flavor and low sourness and green flavor. Partial least square models revealed 20 sweetness-enhancing volatile organic compounds and two sweetness-reducing volatiles, many of which overlap with previous consumer sensory studies. Volatiles modulating green, sour, astringent, overripe, woody, and strawberry flavors were also identified. The relationship between soluble solids content (SSC) and sweetness was modeled with Bayesian regression, generating probabilities for sweetness levels from varying levels of soluble solids. A hierarchical Bayesian model with month effects indicated that SSC is most correlated to sweetness toward the end of the fruiting season, making this the best period to make phenotypic selections for soluble solids. Comparing effects from genotypes, harvest months, and their interactions on sensory attributes revealed that sweetness, sourness, and firmness were largely controlled by genetics. These findings help formulate a paradigm for improvement of eating quality in which sensory analyses drive the targeting of chemicals important to consumer-desired attributes, which further drive the development of genetic tools for improvement of flavor.

Keywords: descriptive analysis, flavor, fruit chemical analysis, sensory, sugars, volatile organic compounds

OPEN ACCESS

Edited by:
Miyako Kusano, University of Tsukuba, Japan

Reviewed by:
Melissa Hamner Mageroy, Norwegian Institute of Bioeconomy Research (NIBIO), Norway
Yonathan Asikin, University of the Ryukyus, Japan

****Correspondence:***
Vance M. Whitaker
vwhitaker@ufl.edu

Specialty section:
This article was submitted to Plant Metabolism and Chemodiversity, a section of the journal Frontiers in Plant Science

Received: *11 December 2020*
Accepted: *01 February 2021*
Published: *17 March 2021*

Citation:
Fan Z, Plotto A, Bai J and Whitaker VM (2021) Volatiles Influencing Sensory Attributes and Bayesian Modeling of the Soluble Solids–Sweetness Relationship in Strawberry. Front. Plant Sci. 12:640704. doi: 10.3389/fpls.2021.640704

INTRODUCTION

The garden strawberry (*Fragaria* × *ananassa*) is popular for its pleasant aroma and sweet taste. High levels of sweetness and intense flavor are the leading factors driving frequent strawberry purchases (Colquhoun et al., 2012). This is consistent with consumer sensory studies that identify sweetness intensity and flavor intensity as the top sensory attributes associated with consumer liking

(Jouquand et al., 2008; Schwieterman et al., 2014). Therefore, sweetness and flavor must be essential criteria during all stages of strawberry breeding in order to develop cultivars that are successful in the marketplace.

While a consumer panel is useful for revealing relationships between some sensory attributes and hedonics (Oliver et al., 2018a), it requires a large number of panelists due to variation from diverse demographic backgrounds (Knee, 2002). In contrast, a descriptive analysis (DA) only requires eight to 12 trained panelists. A DA begins with creation of descriptors for the product and a process to calibrate the descriptors with reference standards until precise and specific descriptors are achieved (Lawless and Heymann, 2010). While a DA does not directly quantify hedonic responses, it can be used to interpret consumer liking when the same samples are tested by consumer panels (Lawless and Heymann, 2010). Trained DA panels have been widely implemented for sensory evaluations of fruits and vegetables under different storage conditions, maturity stages, spans of postharvest storage, or cultural practices (Cliff et al., 1998; Varela et al., 2005; Kårlund et al., 2015).

To identify the chemical drivers of sensory attributes, DA can be combined with chemical analysis. Combined analyses have led to discoveries of the relationships among sensory intensities and flavor-active compounds. In tomato, sweetness, sourness, bitterness, astringency, and saltiness intensities were found to be correlated to sugars, acids, and volatiles (Baldwin et al., 1998). In peach, sweetness and aroma are influenced by organic acids, sugars, and acids (Colaric et al., 2005). In jujube and many other fruits, multiple volatiles contribute to fruity flavor (Stavang et al., 2015; Plotto et al., 2017).

Flavor is a complex of inputs from multiple senses. The range of chemicals contributing to the flavor of fruits includes those interacting with taste buds like sugar and acids and those interacting with olfactory receptors like volatile organic compounds (Lim and Johnson, 2012; Klee and Tieman, 2018). Among all sensory attributes, sweetness is the predominant driver for consumer preference in strawberry (Schwieterman et al., 2014). However, the majority of strawberry consumer samples fail to meet consumers' expectations of sweetness according to our recent consumer sensory study (Fan et al., 2020). It is well established that congruent odors increase taste sensations (Frank and Byram, 1988; Lim and Johnson, 2012), and specific volatiles can enhance sweetness perception (Baldwin et al., 2004, 2008; Tieman et al., 2012). Therefore, increasing sweetness-enhancing volatile content in fruit through horticultural practices or genetic manipulations is an attractive alternative to increasing sugar, as it should not have a detrimental effect on agronomic traits like yield (Whitaker et al., 2012). In the same consumer study (Fan et al., 2020), we found 20 volatiles enhancing sweetness perception independently of sugars. Adding volatiles to the predictive model explained 28% more variability in sweetness than a model with sugars and acids alone (Fan et al., 2020). Additional sensory-chemical studies are needed to validate the effects of sweetness-modulating volatiles and evaluate volatiles' effects on additional sensory attributes (Whitaker et al., 2020).

Sugars, acids, and minerals are the major soluble components in strawberry fruit. Sucrose, glucose, and fructose together account for 99% of total sugars (Montero et al., 1996). Soluble solids content (SSC) has long been considered a good approximation of total sugars in strawberry fruit (Pelayo-Zaldívar et al., 2005; Whitaker et al., 2011). The high-throughput capability of refractometers allows SSC to be routinely used for quality assessment in breeding programs and other industry applications. However, the link between SSC and perceived sweetness may be complex. Discrepancies in the degree of sweetness perception as explained by SSC have been observed across studies (Jouquand et al., 2008; Whitaker et al., 2011; Perez et al., 2016). In addition to genetics, sugar concentration is influenced by harvest dates, locations, maturity, and even fruit-to-fruit variability (Shaw, 1990; Gunness et al., 2009; Hasing et al., 2013). Thus, using ambiguous or arbitrary SSC thresholds may lead to inaccurate conclusions about fruit quality. An ideal predictive framework would take into account uncertainty due to the dynamics of the biological system to construct confidence-based SSC scales. Probabilistic models should facilitate better understanding of biological pathways and mechanisms behind human perception and cognition (Heath et al., 2008; Ma, 2012). In particular, Bayesian models have the advantages of providing hierarchical structures of uncertainty and posterior probability distributions for sample prediction instead of point estimates of means. Advanced computational algorithms that make the implementation of Bayesian models feasible, like Markov chain Monte Carlo simulations, are now available (Van Boekel, 2004; Wilkinson, 2007).

The main objectives of the present study were to (1) combine descriptive sensory analysis and chemical analysis to explore the volatile drivers of sweetness and sourness as well as astringency, green, strawberry, overripe, and woody flavors in fresh strawberries, (2) construct Bayesian models to better define the relationship between sweetness and SSC, and (3) utilize a complex set of strawberry genotypes and environments to better understand the genetic and environmental effects underlying sensory attributes. Each of the three objectives aims to facilitate future flavor breeding by better quantifying factors impacting sensory qualities. Quantifying the effects of individual volatiles will allow us to narrow down the volatile candidates for genetic improvement. Probabilistic SSC evaluation criteria and a better understanding of genotype and environment effects on sensory attributes will inform breeding strategies for sweetness and other sensory qualities.

MATERIALS AND METHODS

Fruit Sampling

In 2009–2010 and 2015–2019, strawberry (*Fragaria* × *ananassa*) samples from 56 cultivars and advance selections (advance selections were elite breeding lines selected by the breeder, which have been evaluated for other agronomic and quality traits for more than 2 years) (**Supplementary Table 1**) were harvested two to four times a year, totaling 213 genotypes/harvest date combinations. All samples were harvested from strawberry

breeding research plots established at the University of Florida (UF) Gulf Coast Research and Education Center (Balm, FL) or the Florida Strawberry Growers Association headquarters in Dover, FL. All fruiting field trials were arranged in randomized complete block designs and were managed based on recommended commercial practices for Florida strawberry annual plasticulture (Whitaker et al., 2019). At each harvest date, one to five clamshells, depending on fruit availability, of fully ripe fruit from five replicate plots for each genotype were collected and transported to the US Department of Agriculture laboratory in Winter Haven (2009–2010) or Fort Pierce (2015–2019), FL. The fruits were stored at 5°C upon arrival and evaluated 1 day (2009–2010) or 3 days (2015–2019) after harvest. Fruits from each replicate plot were kept separate for chemical analyses but were combined for sensory evaluations.

Fruit Quality Analysis

Sensory descriptive analysis, SSC, and titratable acidity (TA) methods were previously described (Plotto et al., 2013). In brief, 10 to 12 panelists (three to four males and seven to eight females, age ranging from 25 to 65 years old, with mean range of 41–50 years old) trained to evaluate fresh fruits and fruit products reconvened each year to review descriptors and reference standards used for strawberry evaluation. The fruits were rated on a structured line scale with intervals from 0 to 10, with definitions as follows: 1 to 2 = low, 5 = medium, and 8 to 9 = high intensity of the rated attribute. Reference standards were served at each panel and were for sweet (sucrose 1–5% + citric acid 0.025–0.05%), sour (sucrose 1% + citric acid 0.05–0.15%), astringent (alum 0.125%), strawberry flavor (frozen strawberry puree), green flavor [(*Z*)-3-hexenal in water, 0.5–3.0 ppm], musty/woody (methyl isoborneol, 50 ppb, a drop on filter paper), and fermented/overripe (overripe strawberry left at 25°C overnight) (Plotto et al., 2013). Firmness was not evaluated in 2009 but was added in 2010 and thereafter; no reference was provided for firmness, but the scale was anchored with the words "soft" for ratings 1 and 2 and "firm" for ratings 8 and 9. The fruits were prepared by washing the individual strawberries under running water, drying, and serving as whole fruit (2009–2010) or cut into quarters (2015–2019). Two to three whole fruits or four to eight strawberry quarters were served in individual 4-oz cups with lids (Solo® cup Company, Urbana, IL, United States), making sure that each piece was from a different fruit. The fruits were served at room temperature, in isolated booths under red lighting. Four to six fruit samples (genotypes) were served in one session, at two sessions per day, with up to 12 genotypes randomly distributed across both sessions. Compusense® Five and Compusense Cloud (Compusense Inc., Guelph, ON, Canada) were used to assign a sample presentation following a William's design pattern and record panelist ratings.

The fruits from each replicate for chemical analyses were different from those used in the sensory panels. Up to 10 fruits (depending on availability) per field replication were cut, with tissue taken for volatile analysis (see below), and the remaining fruits were pureed and frozen at −20°C for later SSC and TA analysis, as described in Plotto et al. (2013).

Volatile Identification and Quantification

In 2009–2010, 30 g from about 10 fruits per genotype and replication was homogenized for 20 s. Saturated $CaCl_2$ was added (w/w) to reduce enzymatic activity (Buttery et al., 1987) right after or at the same time as when homogenizing. Internal standard 3-hexanone (Sigma-Aldrich) was added to a final concentration of 1 ppm. Finally, 5 ml of the mixture was transferred to 20-ml glass vials, crimped with magnetic caps, and stored at −20°C. From 2015 to 2019, 6-g wedges from multiple fruits per genotype were frozen in liquid nitrogen and immediately processed or stored at −80°C. Frozen tissue was grounded with pre-cooled mortar and pestle. Three grams of frozen fruit powder was transferred to 20-ml glass vials (Gerstel) with 3 ml saturated NaCl and 6 μl of internal standard, 3-hexanone at 1,000 ppm, to a final concentration of 1 ppm. The vials were crimped with magnetic caps and stored at −20°C. $CaCl_2$ was replaced by NaCl in 2015 after realizing that the calcium from $CaCl_2$ might interfere with pectin from the strawberry fruit.

Volatiles were sampled from headspace with a 2-cm tri-phase solid-phase micro-extraction (SPME) fiber (50/30 μm DVB/Carboxen/PDMS; Supelco, Bellefonte, PA, United States) and injected into a gas chromatography–mass spectrometry (GC/MS) system (a model 6890 GC coupled with a model 5973 N MS, Agilent Technologies, Palo Alto, CA, United States) as described by Bai et al. (2014). Briefly, a homogenized sample in the vial was incubated for 30 min at 40°C; the SPME fiber was then exposed to the headspace for 30 min at 40°C. After exposure, the SPME fiber was inserted into the injector of GC to desorb for 15 min at 250°C. A DB-5 (60-m length, 0.25-mm i.d., 1.00-μm film thickness; J&W Scientific, Folsom, CA, United States) column was used, with the oven programmed to increase at 4°C min^{-1}, from the initial 40°C to 230°C, and then ramped up at 100°C min^{-1} to 260°C and held for 11.70 min for a total run time of 60 min. Helium was used as the carrier gas at a flow rate of 1.5 ml min^{-1}. The settings for MS were inlet, ionizing source, and transfer line temperatures at 250, 230, and 280 °C, respectively. The mass units were monitored from 40 to 250 m/z and ionized at 70 eV.

Volatile identification and quantification of peak areas were conducted with MassHunter Workstation software (Version 10.0; Agilent Technologies). Initial identification was done by mass spectra searches with the NIST library (Version 14, match score > 0.9). The identification was then confirmed by comparing the retention indices generated by running standard C6-C17 alkane mixture under the same conditions as the samples with online resources (NIST Chemistry WebBook and Flavornet.org).

Bayesian Models of Sweetness Predicted by SSC

Here we described the structure of a robust linear mixed model (Rosa et al., 2003; Svensén and Bishop, 2005):

$$p\left(\beta_0, \beta_{ssc}, \{\tau_i\}, \nu, \sigma^2 \mid \{y_i\}, \{x_i\}\right) \propto \left[\prod_i N\left(y_i \,\middle|\, \beta_0 + \beta_{ssc} x_i, \frac{\sigma^2}{\tau_i}\right)\right.$$

$$\text{Gamma}(\tau_i|\frac{\nu}{2},\frac{\nu}{2})\Big]$$

$$\times N(\beta_0|0,10)\,N(\beta_{ssc}\mid 0,10)\,Unif\,(\nu\mid 0,100)\,\text{Gamma}(\frac{1}{\sigma^2}|1,1)$$

where β_0 is the interception, β_{ssc} is the slope of SSC, y_i is the sweetness rating for each sample, x_i is the sample SSC, and σ^2 is the variance of y_i. Since the linear model is vulnerable to extreme outliers, robust inference was modeled *via* hyper-parameter τ_i for changing residual variance in order to reduce the outliers' influence. All priors for parameters ($\beta_0, \beta_{ssc}, \{\tau_i\}, \nu, \sigma^2$) in the second line were set to random priors.

The above model was extended to a mixed hierarchical model incorporating varying slopes across harvest months.

$$p\left(\beta_0,\{\beta_j\},\{\tau_{ij}\},\nu,\sigma^2,\mu,\sigma_1^2\,\middle|\,\{y_{ij}\},\{x_{ij}\}\right)\propto\Big[\prod_j$$

$$\prod_i N\left(y_{ij}\,\middle|\,\beta_0\beta_j x_{ij},\frac{\sigma^2}{\tau_{ij}}\right)\text{Gamma}(\tau_{ij}|\frac{\nu}{2},\frac{\nu}{2})\Big]$$

$$\times[\prod_j N\left(\beta_j\mid\mu,\sigma_1^2\right)]N(\beta_0|0,10)\,N(\mu|0,10)\,Unif\,(\nu\mid 0,100)$$

$$\text{Gamma}(\frac{1}{\sigma^2}|1,1)\,\text{Gamma}(\frac{1}{\sigma_1^2}|1,1)$$

where β_j is the slope of SSC at harvest month j, and y_{ij} and x_{ij} are sweetness score and SSC level of the i_{th} sample at harvest month j, respectively. $\prod_j N\left(\beta_j\mid\mu,\sigma_1^2\right)$ allows different priors for β_j. The detailed model setup in Just Another Gibbs Sampler 4.3.0 and R software (R version 3.6.3) script to simulate data, run models, and visualize results can be found in the **Supplementary Presentation 1**.

Statistical Evaluation

Mean sample sensory attributes were averaged across 10 to 12 panelists (**Supplementary Table 3**). Chemical data were pooled across field replicates for each sample (**Supplementary Table 2**). Radar plots were drawn with the "fmsb" package in R software to visualize sensory changes among genotypes and harvest months in 2018 and 2019. Principal component analysis (PCA) and K-clustering ($k = 3$) were conducted on sensory attributes using the "prcomp" and the "kmeans" functions in R for the purpose of visualizing sample and sensory attribute relationships and PCA biplots constructed with the "factoextra" package. In order to find important chemicals that influence sensory attributes, partial least square (PLS) models were built for each year with all chemical data and sensory firmness from the DA to account for mouth feel since fruit firmness was not instrumentally measured, and pH was excluded due to its high correlation with TA. PLS (number of components = 5) was analyzed with the "plsr" function and "PLSVarSel" package in R. A chemical was deemed important if the variable importance for the projection (VIP) (Chong and Jun, 2005) was larger than 1. The importance index used to compare importance among chemicals was built such that importance index = (number of years with VIP > 1 and positive effect) – (number of years with VIP > 1 and negative effect). The range of index was anchored at ±7, allowing chemicals to have significant positive or negative effects for all 7 years of data. To investigate genetic and environmental effects on sensory and physicochemical attributes, individual multivariate models with fixed effects, $Y \quad GMG \quad Me$, were built for each year and attribute, where Y is a sensory attribute, G is the genotype effect, M is the harvest month effect, and $G \times M$ represents their interaction. The P-values for all effects were then extracted with the ANOVA function in R. Negative log transformed p-values, after Bonferroni correction, were plotted to show the significance level for each effect.

RESULTS AND DISCUSSION

Descriptive Analysis of Fresh Strawberry Fruits

In the 2009, 2010, 2015, 2016, 2017, 2018, and 2019 seasons, a total of 213 genotype/harvest date combinations from 56 cultivars and advanced selections were subjected to DA. The mean sweetness of all samples was 4.2, with a range from 2.5 to 5.9, on a 0-to-10 scale (**Supplementary Table 3**). Among all descriptors, the smallest range was observed for woody flavor, from 0.2 to 2.2 (**Supplementary Table 3**). Substantial eating quality differences were observed among genotypes. In 2018 and 2019, "Florida Beauty" had the highest sweetness (average 5.0), while 'Florida Brilliance' had the highest sensory firmness (average 5.8) (**Figure 1**). Decreasing average sweetness was observed from January to March in each year, in line with previous findings (Jouquand et al., 2008; Fan et al., 2020). This decline in sweetness was reflected by a similar decline in SSC due to rising temperatures during the Florida fruiting season (MacKenzie and Chandler, 2009; Hasing et al., 2013). Since changes in temperature alter enzymatic activity during fruit development, volatiles also exhibited great variability over the harvest months. In extreme cases, presence-or-absence changes were observed for some volatiles (**Supplementary Table 2**). Thus, a similar decline was also observed for strawberry flavor (**Figure 1**). The large variability of volatiles emphasizes the importance to evaluate the sensory qualities of strawberry over multiple harvests and seasons. PCA with two components using sensory descriptors, SSC, TA, and pH explained 50% of the total variation (**Figure 2**). Sweetness and strawberry flavor, which are desirable sensory attributes, as well as SSC, were negatively correlated with PC1, in contrast with sourness and green flavor. PC2 was positively correlated with TA, overripe, and woody flavors, which were undesirable traits and were negatively correlated with firmness. Similar PCA patterns have been observed for multiple DA studies using fresh strawberries (Plotto et al., 2013; Oliver et al., 2018a). The samples could be grouped into three clusters using K-means clustering (**Figure 2** and **Supplementary Table 1**). The first cluster was mostly confined to the 4th quadrant with high

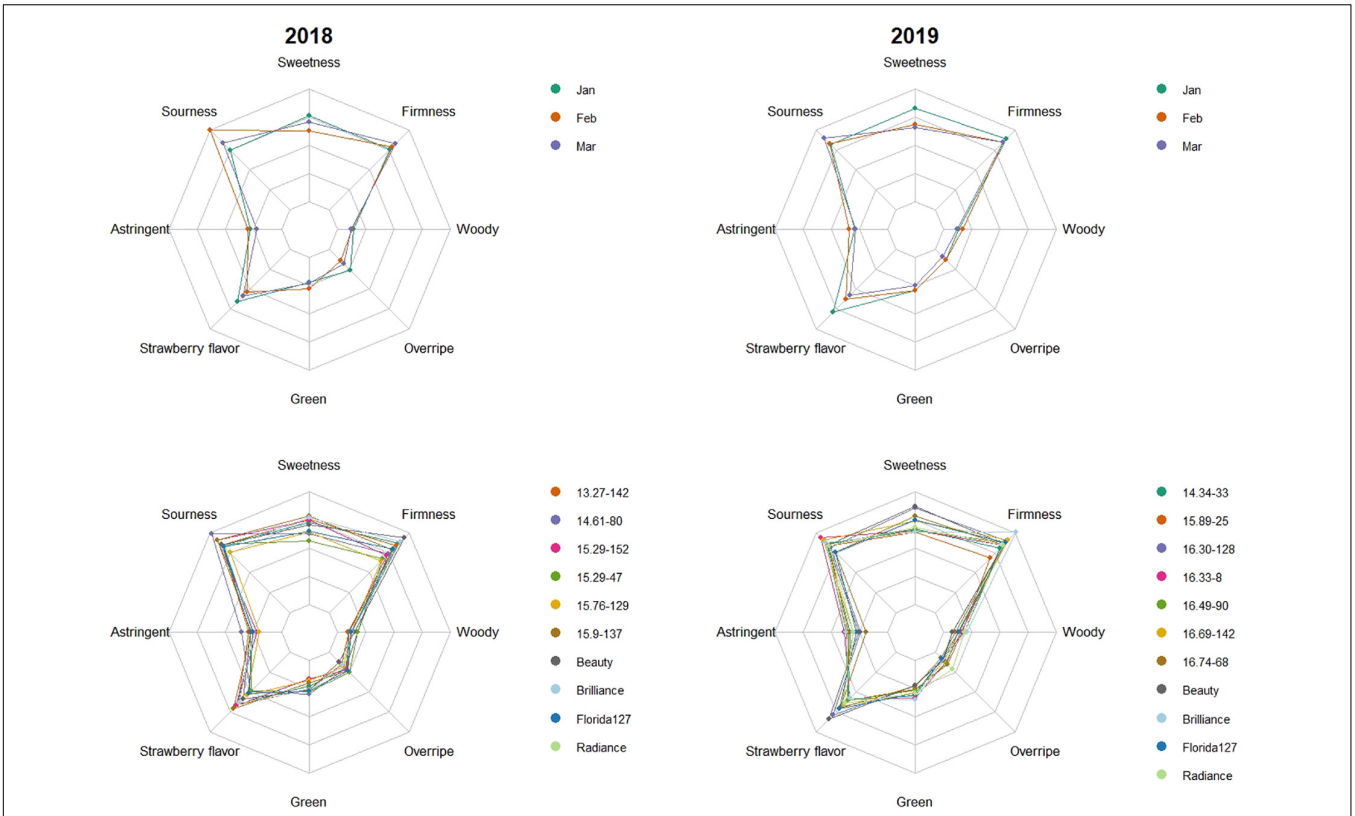

FIGURE 1 | Genotype means across harvest months and harvest month means across genotypes in 2018 and 2019 for eight descriptive sensory attributes. The top two plots show harvest month means across genotypes. The differently colored lines represent months. The bottom two plots show genotype means across three harvests. Colored lines represent genotypes. Scale from 0 to 10 was used for all attributes.

sourness and green flavor. The second cluster was along the negative side of PC1, classified with high sweetness, strawberry flavor, and firmness. This cluster appeared to comprise samples with the most desirable attributes. Our recent releases, "Florida Beauty" and Sweet Sensation® "Florida127," were consistently clustered in group 2. Specifically, "Florida Beauty" was rated group 2 in 10 out of 14 evaluations and Sweet Sensation® 'Florida127' was rated group 2 in 13 out of 16 evaluations (**Supplementary Table 1**). The last cluster was correlated with higher levels of woody flavor, overripe flavor, and astringent mouthfeel. While there is fluctuation over the years, sweetness and strawberry flavor have always been strongly correlated. In contrast, sourness has been correlated with astringent mouthfeel, green, woody, or overripe flavors, depending on the season (**Supplementary Presentation 2** and Plotto et al., 2013).

During the breeding process, many criteria are weighted to select the most well-rounded genotypes. Improvement of eating quality is generally subject to the taste preferences and limited sampling by the breeding team. DA provides a quantitative tool to finely discriminate the eating quality of different genotypes. Since consumer acceptance can be inferred from desirable attributes like sweetness, strawberry flavor, and firmness (Azodanlou et al., 2003), a DA-based quality index has been developed and implemented for strawberry growers and breeding programs (Ares et al., 2009). However, caution must be taken when extrapolating the index across programs because of differences in understanding of the scale and consumer expectations. To maximize the impact of DA, the process of identifying descriptors, developing quality indices, and establishing benchmarks must be carefully undertaken.

Volatiles Modulating Sensory Attributes

The chemical analysis of 213 samples yielded 71 volatiles, covering 17 of 22 aroma-active volatiles reported in Nuzzi et al. (2008) and 21 of 29 aroma-active volatiles evaluated with GC–olfactometry in 'Du et al. (2011). Volatiles with significant influence on sensory attributes were identified with PLS models. In order to minimize the effect of sample preparation and environments, the PLS model was analyzed independently for each year, and the summary statistic of importance index was used to gauge the overall effect on sensory attributes across 7 years. Models using only 5 years of data from 2015 to 2019 identified a consistent but smaller set of important chemicals. A significant effect of a volatile for a sensory quality had to be consistent over at least 2 years to be considered as influential.

Twenty volatiles were shown to enhance sweetness independent of SSC (**Figure 3** and **Supplementary Table 4**). While only 14 genotypes overlapped with our previous consumer study (**Supplementary Table 5** and Fan et al., 2020) and the volatile sampling methods were different, our top three

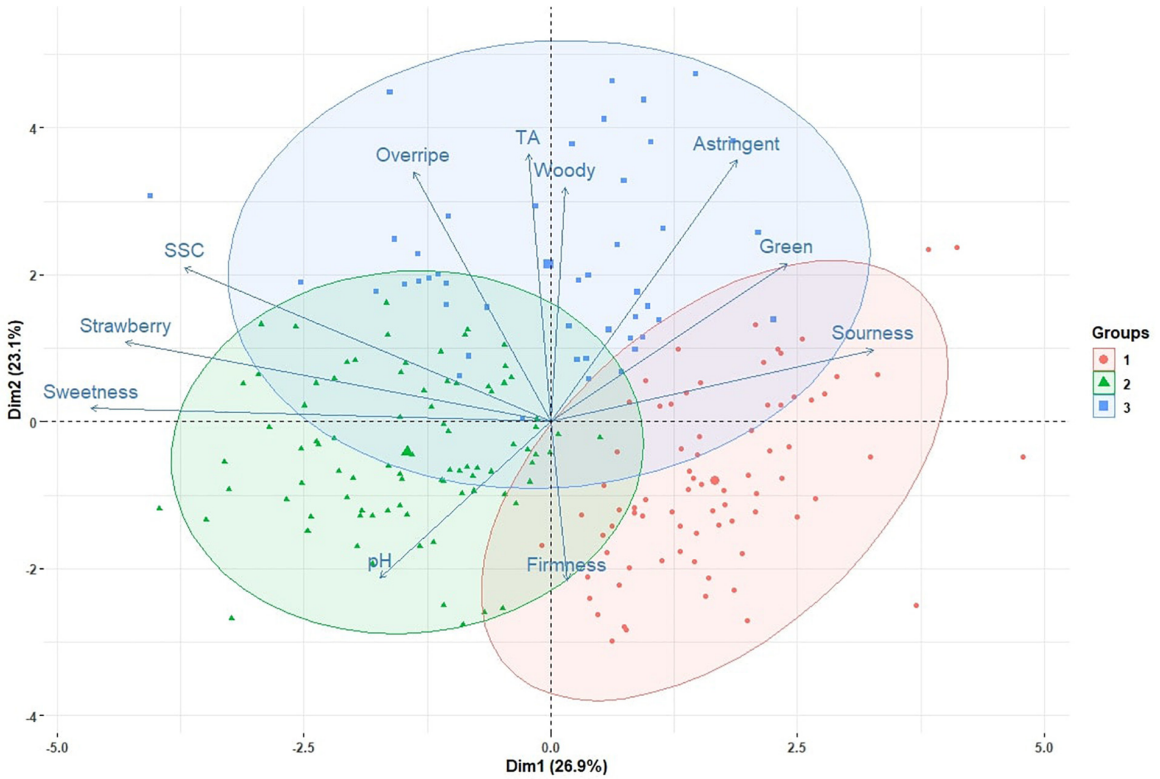

FIGURE 2 | Biplot with the first two components from a principal component analysis. Vectors represent sensory and physiochemical attributes. Points represent individual samples. Three colored eclipses indicate three clusters from a K-cluster analysis.

sweetness-enhancing volatiles [benzaldehyde, 2-pentanal, (E) and nonanal] were identified with both consumer and DA panels. 2,5-Dimethyl-4-methoxy-3(2H)-furanone (mesifurane), butanoic acid/ethyl ester, and hexanoic acid/ethyl ester exhibited a sweetness-enhancing ability in multiple years (**Figure 3**), in line with findings based on odor threshold and odor active values (Pyysalo et al., 1979; Larsen and Poll, 1990; Ulrich et al., 1997). In agreement with the previous consumer study, two medium-chain butanoic acid esters (butanoic acid, hexyl ester and butanoic acid, and octyl ester) had increased sweetness, as well as three medium-chain hexanoic acid esters (hexanoic acid, hexyl ester; hexanoic acid, octyl ester; and hexanoic acid, propyl ester). Butanoic acid, 3- methyl-, and two derived esters (butanoic acid, 3- methyl-, ethyl ester, and butanoic acid, 3-methylbutyl ester) had positive effects on sweetness and have been identified in multiple GC-O studies (Ulrich et al., 1997; 'Du et al., 2011; Cannon et al., 2015). Surprisingly, hexyl acetate, one of the most potent esters ('Du et al., 2011) in strawberry, had a negative effect on sweetness, possibly related to its green apple aroma at higher concentrations (Young et al., 1996). Volatiles influencing strawberry flavor mostly overlapped with sweetness influencers (**Figure 3**). Furanones, esters, and lactones generally not detected in white or half-red fruit undergo dramatic increases in the late stages of ripening (Ménager et al., 2004), shaping the strawberry flavor in ripe fruit. A major determinant in peach aroma, γ-decalactone, contributed to strawberry flavor, but not sweetness, in this study. The only volatile found to decrease both sweetness and strawberry flavor was acetic acid, 1-methylethyl ester.

Soluble solids content unsurprisingly had the highest correlations with both sweetness and strawberry flavor. Sugar content is also the main determinant for consumer liking (Schwieterman et al., 2014). Given the potential yield cost imposed by breeding varieties with higher sugar level (Whitaker et al., 2012), manipulating volatiles may provide a better alternative for the enhancement of eating quality. In our previous consumer sensory study, we found that most sweetness enhancers overlapped strongly with liking enhancers (Fan et al., 2020). In the present study, the prior list of sweetness-enhancing volatiles based on consumer studies is expanded with new compounds such as mesifurane; butanoic acid, ethyl ester and hexanoic acid, ethyl ester; and butanoic acid, 3- methyl-, which have been historically considered as important volatiles for strawberry flavor, highlighting the diverse flavor profiles existing in commercial germplasm.

In this study, alpha-terpineol; methanethiol; acetic acid, butyl ester; linalool; butanoic acid, 3- methyl-, methyl ester; pentanal; and propanoic acid, ethyl ester were identified as green flavor contributors (**Figure 4** and **Supplementary Table 4**). Aldehydes like pentanal, hexanal, and 2-hexenal, (*E*)- have typically been linked to green notes in immature fruit (Perez et al., 1992; Ménager et al., 2004) but are much reduced and suppressed

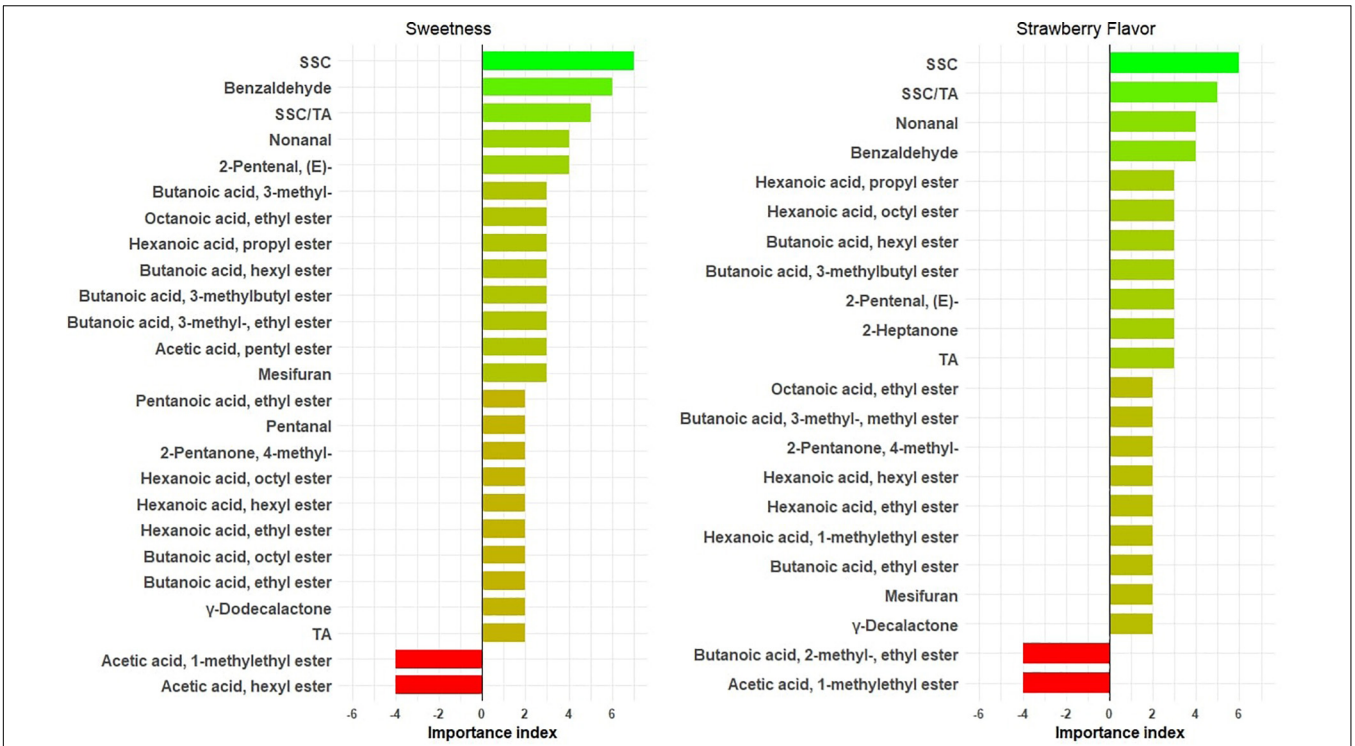

FIGURE 3 | Important physicochemical measures/volatiles for sweetness and strawberry flavor. The *X* axis shows importance index = [number of years with variable importance in projection (VIP) > 1 and positive effect] – (number of years with VIP > 1 and negative effect). The bars are annotated with chemical names.

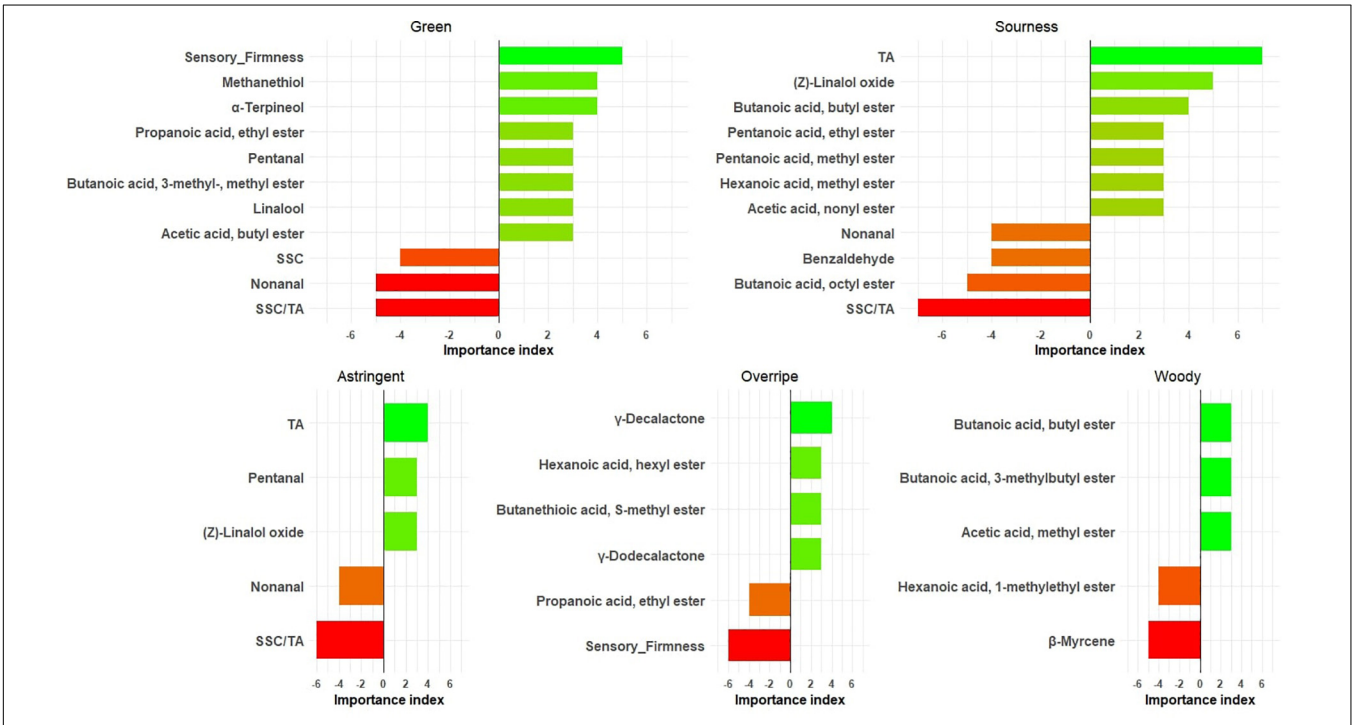

FIGURE 4 | Important physicochemical measures/volatiles for five sensory attributes. The *X* axis shows the importance index [number of years with variable importance in projection (VIP) > 1 and positive effect] – (number of years with VIP > 1 and negative effect). The bars are annotated with chemical names.

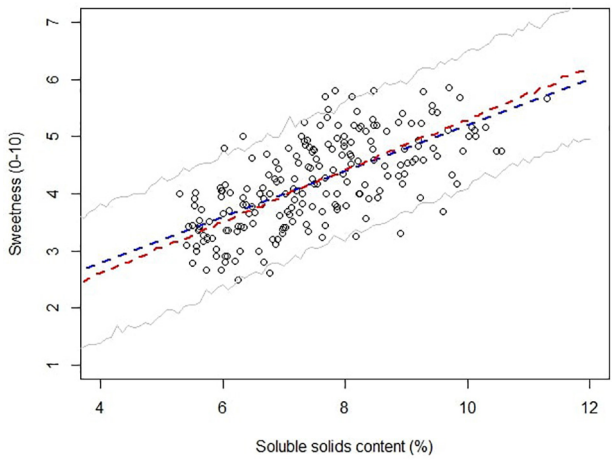

FIGURE 5 | Scatter plot of sweetness against soluble solids content. The blue line represents least square linear regression, and the red line represents the mean Bayesian robust regression. Gray lines represent 95% confidence intervals for posterior distributions.

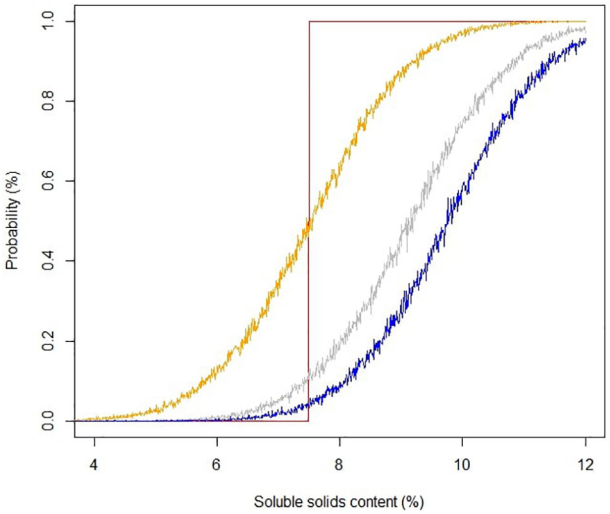

FIGURE 6 | Probability of sweetness perception based on soluble solids content. Yellow, gray, and blue lines represent accumulated probability distributions of above average, top 20%, and top 10% sweetness, respectively. The red line represents the inference from a least square linear model.

by other strawberry aromas upon ripening (Jetti et al., 2007). Unsurprisingly, firmness was found to be related to green notes (**Figure 4**). TA was the major contributor for sourness in all 7 years (**Figure 4**). Besides TA, (Z)-linalool oxide appeared to have positive effects on sourness, as opposed to benzaldehyde and nonanal. TA was also highly correlated with astringent mouthfeel, but not as strongly as for the sourness–TA relationship (**Figure 4**). Perceived astringency of fruit is mainly associated with phenolic compounds (Joslyn and Goldstein, 1964) and acids (Lawless et al., 1996). Two lactones (γ-decalactone and γ-dodecalactone) and butanethioic acid, S-methyl ester were associated with overripe flavor (**Figure 4**). During post-harvest storage, volatile compositions undergo large changes (Lu et al., 2018), such that the higher relative abundances of γ-decalactone may be observed in overripe strawberries. Since sulfur esters are usually described as giving undesirable aromas (Du et al., 2011), it is not surprising that butanethioic acid, S-methyl ester appears to be one of the compounds contributing to overripe flavor, which was also defined as "fermented." Because our study includes fruit from the narrow genetic pool of the UF strawberry breeding program, it does not present an exhaustive list of volatiles influencing sensory attributes from strawberry. Future identification of new sensory-modulating volatiles should embrace a wider germplasm and the wild relatives of cultivated strawberry.

Modeling Sweetness Based on Physiochemical Parameters

Soluble solids content is a robust analytical measurement that strongly correlated with total sugars in strawberry (Jouquand et al., 2008; Gunness et al., 2009). The initial implementation of SSC in strawberry quality control and breeding can be traced back to the late 1980s, and it is still routinely used in the strawberry industry (Alavoine and Crochon, 1988; Ares et al., 2009). In Alavoine's report, SSC-based thresholds were proposed to distinguish medium taste quality from high taste quality, with an arbitrary cutoff for the highest taste quality placed at SSC of 8%. Thirty years have passed, and commercial strawberry quality has seen an improvement. Some current cultivars like "Florida Beauty" had an average SSC of 8.2% across all harvests in this study. To guide future benchmarks, we utilized a Bayesian robust model to fit SSC against sweetness perception due to its ability to incorporate sample uncertainty and parameter uncertainty. Our model successfully converged with average $\hat{R}$ smaller than 1.1. The mean of the marginal posterior distribution slope β_{ssc} was 0.45, with SD of 0.03 and 95% credibility from 0.40 to 0.50, indicating a strong positive correlation between SSC and sweetness. The 95% credible marginal posterior distribution of y (**Figure 5**) captured 196 of 207 samples (94.7%). Eleven outliers were weighted less in the model with τ_i smaller than 1, ranging from 0.91 to 0.96. The main goal of adopting a Bayesian model was to give probabilistic-based predictions of sample sweetness from SSC data for breeding and quality control purposes. Striking differences for the probability of greater-than-average sweetness was obtained after incorporating uncertainty into the model. Predictions without uncertainty indicated that samples with SSC higher than 7.5% would have greater-than-average sweetness intensity (**Figure 6**). However, the Bayesian model indicated a 50% chance of greater-than-average sweetness at SSC of 7.5%. To achieve 80% confidence, SSC must be higher than 8.6%. A simplified probability chart with SSC ranging from 6 to 12 is shown in **Table 1**.

During winter and spring production in a subtropical climate, strawberry SSC exhibits large within-season variations due to changes in plant physiological and environmental conditions (MacKenzie and Chandler, 2009). Differences in SSC stability over the fruiting season among genotypes add

TABLE 1 | Sweetness probability chart for levels of soluble solids content (SSC).

SSC (%)	Sensory sweetness probability		
	Above average	Top 20%	Top 10%
6.00	0.11[a]	0.01	0.01
6.50	0.22	0.04	0.01
7.00	0.35	0.04	0.02
7.50	0.50	0.08	0.06
8.00	0.64	0.20	0.08
8.50	0.77	0.30	0.17
9.00	0.86	0.48	0.29
9.50	0.94	0.58	0.42
10.00	0.97	0.74	0.58
10.50	0.99	0.83	0.72
11.00	1.00	0.92	0.83
11.50	1.00	0.96	0.89
12.00	1.00	0.98	0.96

[a]*The probability of above-average sweetness perception at a soluble solids content of 6% is 11%.*

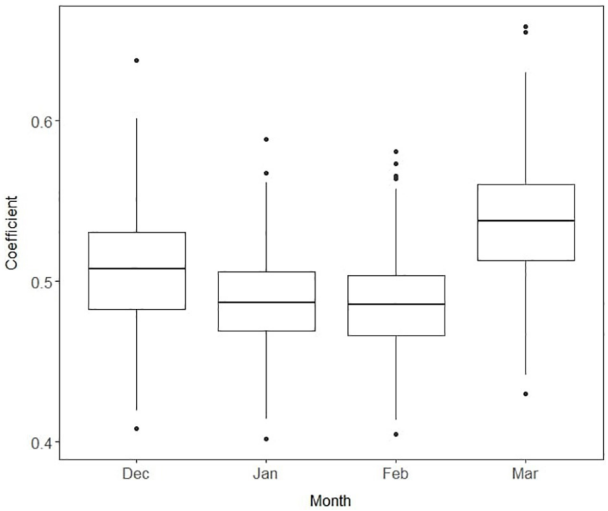

FIGURE 7 | Boxplots of the slope coefficients for soluble solids content by month. Each box represents median and interquartile range from 15,000 iterations. Higher values imply higher correlations of soluble solids content with sensory sweetness.

additional complications to the selection criteria for sweetness (Hasing et al., 2013). Furthermore, the relationship between SSC and sweetness can be strongly influenced by environmental conditions (Azodanlou et al., 2003). To examine the stability of relationship between SSC and sweetness during the season, we introduced varying slopes β_j to the Bayesian model, which allowed heterogeneity of the sweetness–SSC relationship. The highest coefficient of the slope was found in March, with a mean of 0.54 and 95% confidence interval of 0.47 to 0.61 (**Figure 7**). Lower coefficients were found in January and February, around 0.49. Our results suggest that March was the best month to make selections for SSC due to its higher correlation with perceived sweetness. This is also consistent with the objective of maintaining fruit quality at the conclusion of the season when conditions are more unfavorable for eating quality (Hasing et al., 2013). Breeders may therefore consider reducing the number of SSC phenotyping events early in the season to save labor and resources or more heavily weighting late-season observations.

Genetic and Environmental Factors Influencing Sensory and Physicochemical Attributes

The chemical compositions of strawberries are strongly affected by both genotypes and growing conditions (Forney et al., 2000), leading to eating quality differences (Jouquand et al., 2008). However, the existing literature exploring the fluctuation of sensory characteristics based on genotypes or environments only encompasses a few genotypes grown in one or two seasons (Jetti et al., 2007; Jouquand et al., 2008). In the present study, effects from genotype, harvest month, and their interactions were compared across 7 years (**Figure 8**). Sweetness was controlled by both harvest month and genotype, with no significant interaction between the two. Genotype was significant in 5 years, and harvest month was significant in 4 years. This result is corroborated by significant genotype and month effects for SSC in six of 7 years. Fewer genotype and month effects were observed for strawberry flavor. Sourness was mainly controlled by genotype, consistent with greater significance for genotype than for month for TA and pH (**Figure 8**). Sensory firmness also appears to be mainly controlled by genotype. Overripe flavor showed significance for genotype in 3 years and for harvest month in 2 years. Astringent, green, and woody flavors had little influence from genotype or month. There were no consistent G × E interactions for any sensory attributes. Although no previous sensory studies have included such large numbers of genotypes and environments (Ulrich et al., 2018), metabolite surveys have shown moderate to high heritability for SSC and volatile abundance (Whitaker et al., 2012; Gezan et al., 2017; Urrutia et al., 2017). A large portion of variability in most volatiles found in wild strawberry is explained by one or two major QTLs (Urrutia et al., 2017). Consistent with these studies, genetics was the primary force driving variation in volatile abundances among our strawberry samples (**Supplementary Presentation 3**). As expected, environmental effect is also pervasive for most volatiles over the harvest months (**Supplementary Presentation 3**). Our results indicate the strong genetic control of firmness, sourness, sweetness, and strawberry flavor, implying further potential for genetic improvement of these sensory attributes in the germplasm tested.

A Paradigm for Sensory Quality Improvement in Strawberry

Increasing consumer liking is one of the main goals of strawberry breeding. Descriptive analysis provides an objective evaluation of sensory characteristics that are strongly associated with consumer preference. Sweetness and strawberry flavor contribute to liking, while sour, astringent, overripe, and green flavors detract from liking (Oliver et al., 2018b). At the intermediate stage of cultivar

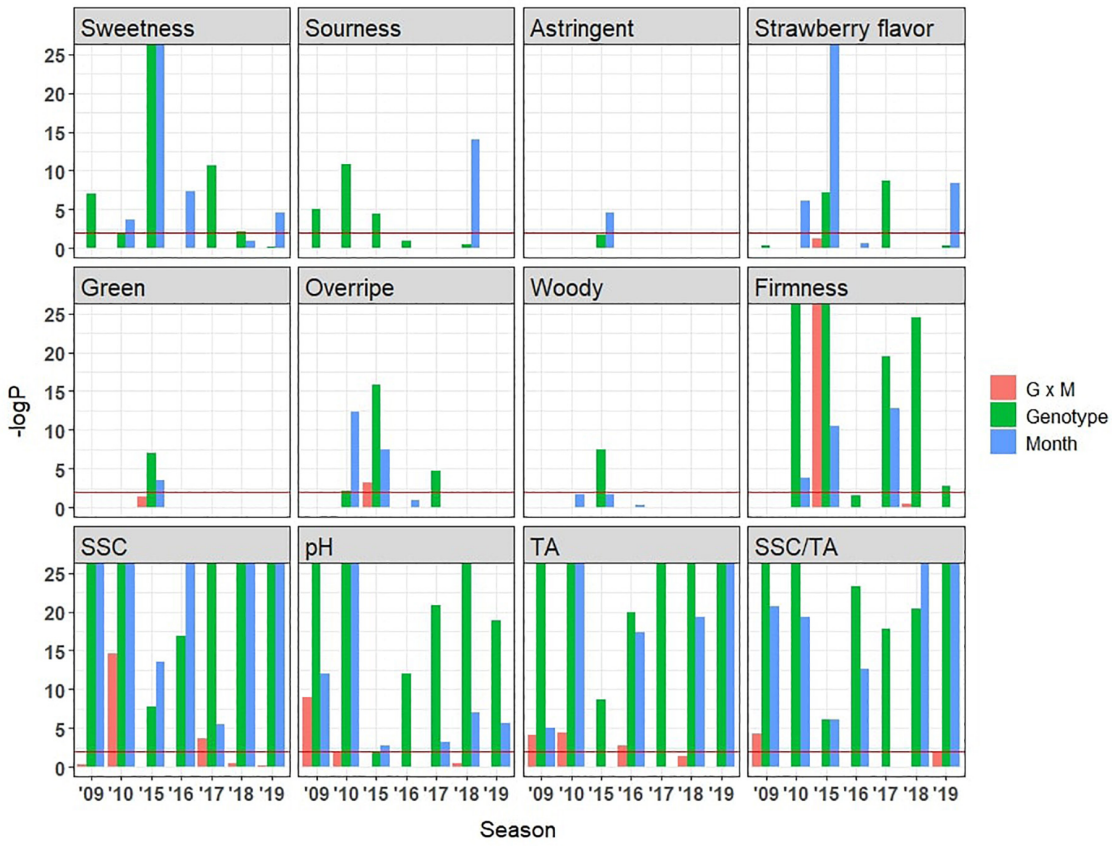

FIGURE 8 | Influence of genotype, harvest month, and month by genotype interaction effects on sensory and physiochemical attributes as indicated by –logP values. Horizonal lines indicate –logP values of 2 (Bonferroni-corrected *p*-value = 0.01). Taller bars were observed for physicochemical attributes compared to sensory attributes due to low variance among field replicates compared to panelist variance.

development, elite selections are grown in small, replicated trials for quality evaluation. Fruit yield from those trials is too small to supply large consumer panels (Fan et al., 2020), but small trained panels are well suited for this purpose. Importantly, these data also allow sensory evaluations to drive targeted genetic solutions for improving eating quality, for example, γ-decalactone was identified as a flavor target based on sensory evidence (Schwieterman et al., 2014) and the identification of a candidate gene (*FaFAD1*) responsible for the presence and abundance of this volatile (Chambers et al., 2014; Sánchez-Sevilla et al., 2014; Noh et al., 2017). An improved codominant DNA maker for *FaFAD1* was recently designed, validated, and implemented in a marker-assisted selection (Oh et al., 2019). Thus, a fruit quality improvement paradigm that begins with consumer-desired sensory attributes and progresses to chemical targets and finally to genetic tools is now feasible in strawberry. This paradigm is applicable not only in strawberry but also to a broad array of crops valued for their sensory qualities.

In conclusion, we identified additional volatiles that enhance sweetness independently of sugars, such as benzaldehyde; 2-pentenal, (E)-; nonanal; mesifurane; butanoic acid, 3- methyl-, ethyl ester; butanoic acid, 3-methyl-; γ-dodecalactone; butanoic acid, ethyl ester; and hexanoic acid, ethyl ester, many of which have been historically identified as important to strawberry flavor and in our consumer study (Fan et al., 2020). Thus, the identification of these volatiles will allow us to narrow down to a smaller number of flavor breeding targets. Sweetness prediction based on SSC has been updated using a probabilistic approach, better informing strategies for improving strawberry sweetness. Lastly, our quantification of genotype and environment effects and their interaction on various sensory attributes provides a statistical basis for the strategic improvement of strawberry sensory quality.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the Local Legislation

and Institutional Requirements. Written informed consent for participation was not required for this study in accordance with the National Legislation and the Institutional Requirements.

AUTHOR CONTRIBUTIONS

ZF performed data analyses and composed the manuscript. VW and AP designed the study and oversaw the whole project. JB organized and performed chemical analyses. AP organized and performed descriptive sensory panels. All the authors read and approved the final manuscript.

FUNDING

The authors are grateful for the support provided by the Florida Agricultural Experiment Station, the Florida Department of Agricultural and Consumer Services through the USDA Block Grant program, and the Florida Strawberry Research and Education Foundation.

ACKNOWLEDGMENTS

The authors acknowledge the significant efforts of the UF strawberry breeding program technical staff in growing and maintaining breeding trials, harvests, and fruit transportation.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.640704/full#supplementary-material

Supplementary Presentation 1 | R-script to simulate data, generate JAGS models and visualize results.

Supplementary Presentation 2 | PCA plots of sensory attributes for individual year.

Supplementary Presentation 3 | Influence of genotype, harvest month, and month by genotype interaction effects on volatile abundance as indicated by −logP values.

Supplementary Table 1 | Sample basic information and cluster based on sensory attributes. Four columns list sample genotype, harvesting year, month, and cluster number. NA means the sample was either not sent for sensory evaluation, or not quantified for TA and SSC.

Supplementary Table 2 | Chemical data for each sample. CAS numbers were provided.

Supplementary Table 3 | Average sensory attributes across panelists and average SSC, pH, and TA across sample replicates. NA means not measured.

Supplementary Table 4 | Importance index of all volatiles. Other tabs include PLS results for individual sensory attributes, including VIP scores and coefficients for each volatile.

Supplementary Table 5 | Summary of common genotypes used in this study and the previous consumer study (Fan et al., 2020).

REFERENCES

Alavoine, F., and Crochon, M. (1988). Taste quality of strawberry. *Int. Strawberry Symp*. 265.

Ares, G., Barrios, S., Lareo, C., and Lema, P. (2009). Development of a sensory quality index for strawberries based on correlation between sensory data and consumer perception. *Postharvest Biol. Technol.* 52, 97–102. doi: 10.1016/j.postharvbio.2008.11.001

Azodanlou, R., Darbellay, C., Luisier, J., Villettaz, J., and Amadò, R. (2003). Quality assessment of strawberries (*Fragaria* Species). *J. Agric. Food Chem.* 51, 715–721. doi: 10.1021/jf0200467

Bai, J., Baldwin, E., Hearn, J., Driggers, R., and Stover, E. (2014). Volatile profile comparison of USDA sweet orange-like hybrids versus 'Hamlin'and 'Ambersweet'. *HortScience* 49, 1262–1267.

Baldwin, E. A., Goodner, K., and Plotto, A. (2008). Interaction of volatiles, sugars, and acids on perception of tomato aroma and flavor descriptors. *J. Food Sci.* 73, S294–S307. doi: 10.1111/j.1750-3841.2008.00825.x

Baldwin, E. A., Goodner, K., Plotto, A., Pritchett, K., and Einstein, M. (2004). Effect of Volatiles and their Concentration on Perception of Tomato Descriptors. *J. Food Sci.* 69, S310–S318. doi: 10.1111/j.1750-3841.2004.tb18023.x

Baldwin, E. A., Scott, J. W., Einstein, M. A., Malundo, T., Carr, B. T., Shewfelt, R. L., et al. (1998). Relationship between sensory and instrumental analysis for tomato flavor. *J. Am. Soc. Hortic. Sci.* 123, 906–915.

Buttery, R. G., Teranishi, R., and Ling, L. C. (1987). Fresh tomato aroma volatiles: a quantitative study. *J. Agric. Food Chem.* 35, 540–544. doi: 10.1021/jf00076a025

Cannon, R. J., Agyemang, D., Curto, N. L., Yusuf, A., Chen, M. Z., and Janczuk, A. J. (2015). In depth analysis of Ciflorette strawberries (Fragaria× ananassa 'Ciflorette') by multidimensional gas chromatography and gas chromatography olfactometry. *Flavour Fragr. J.* 30, 302–319.

Chambers, A. H., Pillet, J., Plotto, A., Bai, J., Whitaker, V. M., and Folta, K. M. (2014). Identification of a strawberry flavor gene candidate using an integrated genetic-genomic-analytical chemistry approach. *BMC Genomics* 15:217. doi: 10.1186/1471-2164-15-217

Chong, I., and Jun, C. (2005). Performance of some variable selection methods when multicollinearity is present. *Chemometr. Intell. Lab. Syst.* 78, 103–112. doi: 10.1016/j.chemolab.2004.12.011

Cliff, M. A., Lau, O. L., and King, M. C. (1998). Sensory characteristics of controlled atmosphere- and air-stored 'Gala' Apples1. *J. Food Qual.* 21, 239–249. doi: 10.1111/j.1745-4557.1998.tb00519.x

Colaric, M., Veberic, R., Stampar, F., and Hudina, M. (2005). Evaluation of peach and nectarine fruit quality and correlations between sensory and chemical attributes. *J. Sci. Food Agric.* 85, 2611–2616. doi: 10.1002/jsfa.2316

Colquhoun, T. A., Levin, L. A., Moskowitz, H. R., Whitaker, V. M., Clark, D. G., and Folta, K. M. (2012). Framing the perfect strawberry: an exercise in consumer-assisted selection of fruit crops. *J. Berry Res.* 2, 45–61. doi: 10.3233/JBR-2011-027

'Du, X., Plotto, A., Baldwin, E., and Rouseff, R. (2011). Evaluation of volatiles from two subtropical strawberry cultivars using GC–olfactometry, GC-MS odor activity values, and sensory analysis. *J. Agric. Food Chem.* 59, 12569–12577. doi: 10.1021/jf2030924

Du, X., Song, M., and Rouseff, R. (2011). Identification of new strawberry sulfur volatiles and changes during maturation. *J. Agric. Food Chem.* 59, 1293–1300. doi: 10.1021/jf104287b

Fan, Z., Hasing, T., Johnson, T. S., Garner, D. M., Barbey, C. R., Colquhoun, T. A., et al. (2020). Strawberry sweetness and consumer preference are enhanced by specific volatile compounds. *bioRxiv* [preprint] doi: 10.1101/2020.12.04.410654 bioRxiv: 2020.12.04.410654,

Forney, C. F., Kalt, W., and Jordan, M. A. (2000). The composition of strawberry aroma is influenced by cultivar, maturity, and storage. *HortScience* 35, 1022–1026. doi: 10.21273/HORTSCI.35.6.1022

Frank, R. A., and Byram, J. (1988). Taste–smell interactions are tastant and odorant dependent. *Chem. Senses* 13, 445–455. doi: 10.1093/chemse/13.3.445

Gezan, S. A., Osorio, L. F., Verma, S., and Whitaker, V. M. (2017). An experimental validation of genomic selection in octoploid strawberry. *Hortic. Res.* 4:16070. doi: 10.1038/hortres.2016.70

Gunness, P., Kravchuk, O., Nottingham, S. M., D'Arcy, B. R., and Gidley, M. J. (2009). Sensory analysis of individual strawberry fruit and comparison with instrumental analysis. *Postharvest Biol. Technol.* 52, 164–172. doi: 10.1016/j.postharvbio.2008.11.006

Hasing, T. N., Osorio, L. F., and Whitaker, V. M. (2013). Within-season stability of strawberry soluble solids content. *J. Am. Soc. Hortic. Sci.* 138, 190–197.

Heath, J., Kwiatkowska, M., Norman, G., Parker, D., and Tymchyshyn, O. (2008). Probabilistic model checking of complex biological pathways. *Theor. Comput. Sci.* 391, 239–257. doi: 10.1016/j.tcs.2007.11.013

Jetti, R. R., Yang, E., Kurnianta, A., Finn, C., and Qian, M. C. (2007). Quantification of selected aroma-active compounds in strawberries by headspace solid-phase microextraction gas chromatography and correlation with sensory descriptive analysis. *J. Food Sci.* 72, S487–S496. doi: 10.1111/j.1750-3841.2007.00445.x

Joslyn, M. A., and Goldstein, J. L. (1964). Astringency of fruits and fruit products in relation to phenolic content. *Adv. Food Res.* 13, 179–217.

Jouquand, C., Chandler, C., Plotto, A., and Goodner, K. (2008). A sensory and chemical analysis of fresh strawberries over harvest dates and seasons reveals factors that affect eating quality. *J. Am. Soc. Hortic. Sci.* 133, 859–867.

Kårlund, A., Hanhineva, K., Lehtonen, M., Karjalainen, R. O., and Sandell, M. (2015). Nontargeted metabolite profiles and sensory properties of strawberry cultivars grown both organically and conventionally. *J. Agric. Food Chem.* 63, 1010–1019. doi: 10.1021/jf505183j

Klee, H. J., and Tieman, D. M. (2018). The genetics of fruit flavour preferences. *Nat. Rev. Genet.* 19, 347–356. doi: 10.1038/s41576-018-0002-5

Knee, M. (2002). *Fruit Quality and Its Biological Basis*. Sheffield, UK: CRC Press.

Larsen, M. D., and Poll, L. (1990). Odour thresholds of some important aroma compounds in strawberries. *Z. Lebensm. Unters. Forsch.* 195, 120–123. doi: 10.1007/BF01201770

Lawless, H. T., and Heymann, H. (2010). *Sensory Evaluation of Food: Principles and Practices*. New York, NY: Springer Science & Business Media.

Lawless, H. T., Horne, J., and Giasi, P. (1996). Astringency of organic acids is related to pH. *Chem. Senses* 21, 397–403. doi: 10.1093/chemse/21.4.397

Lim, J., and Johnson, M. B. (2012). The role of congruency in retronasal odor referral to the mouth. *Chem. Senses* 37, 515–522. doi: 10.1093/chemse/bjs003

Lu, H., Wang, K., Wang, L., Li, D., Yan, J., Ban, Z., et al. (2018). Effect of superatmospheric oxygen exposure on strawberry (Fragaria × ananassa Fuch.) volatiles, sensory and chemical attributes. *Postharvest Biol. Technol.* 142, 60–71. doi: 10.1016/j.postharvbio.2018.04.007

Ma, W. J. (2012). Organizing probabilistic models of perception. *Trends Cogn. Sci.* 16, 511–518. doi: 10.1016/j.tics.2012.08.010

MacKenzie, S. J., and Chandler, C. (2009). The late year decline in strawberry fruit soluble solid content observed in Florida is caused by rising temperatures. *Acta Hortic.* 842, 843–846. doi: 10.17660/ActaHortic.2009.842.186

Ménager, I., Jost, M., and Aubert, C. (2004). Changes in physicochemical characteristics and volatile constituents of strawberry (Cv. Cigaline) during maturation. *J. Agric. Food Chem.* 52, 1248–1254. doi: 10.1021/jf0350919

Montero, T. M., Mollá, E. M., Esteban, R. M., and López-Andréu, F. J. (1996). Quality attributes of strawberry during ripening. *Sci. Hortic.* 65, 239–250. doi: 10.1016/0304-4238(96)00892-8

Noh, Y., Lee, S., Whitaker, V. M., Cearley, K. R., and Cha, J. (2017). A high-throughput marker-assisted selection system combining rapid DNA extraction high-resolution melting and simple sequence repeat analysis: strawberry as a model for fruit crops. *J. Berry Res.* 7, 23–31. doi: 10.3233/JBR-160145

Nuzzi, M., Scalzo, R. L., Testoni, A., and Rizzolo, A. (2008). Evaluation of fruit aroma quality: comparison between Gas Chromatography–Olfactometry (GC–O) and Odour Activity Value (OAV) aroma patterns of strawberries. *Food Anal. Methods* 1, 270–282. doi: 10.1007/s12161-008-9039-y

Oh, Y., Zurn, J. D., Bassil, N., Edger, P. P., Knapp, S. J., Whitaker, V. M., et al. (2019). The strawberry DNA testing handbook. *HortScience horts.* 54, 2267–2270. doi: 10.21273/HORTSCI14387-19

Oliver, P., Cicerale, S., Pang, E., and Keast, R. (2018a). A comparison of Temporal Dominance of Sensation (TDS) and Quantitative Descriptive Analysis (QDA^{TM}) to identify flavors in strawberries. *J. Food Sci.* 83, 1094–1102. doi: 10.1111/1750-3841.14096

Oliver, P., Cicerale, S., Pang, E., and Keast, R. (2018b). Identifying key flavors in strawberries driving liking via internal and external preference mapping. *J. Food Sci.* 83, 1073–1083. doi: 10.1111/1750-3841.14109

Pelayo-Zaldívar, C., Ebeler, S. E., and Kader, A. A. (2005). Cultivar and harvest date effects on flavor and other quality attributes of California strawberries. *J. Food Qual.* 28, 78–97. doi: 10.1111/j.1745-4557.2005.00005.x

Perez, A. G., Rios, J. J., Sanz, C., and Olias, J. M. (1992). Aroma components and free amino acids in strawberry variety Chandler during ripening. *J. Agric. Food Chem.* 40, 2232–2235. doi: 10.1021/jf00023a036

Perez, Y., Sargent, S., Cecilia do, N. N., and Whitaker, V. (2016). Composition of commercial strawberry cultivars and advanced selections as affected by season, harvest, and postharvest storage. *HortScience horts.* 51, 1134–1143. doi: 10.21273/HORTSCI10996-16

Plotto, A., Bai, J., and Baldwin, E. (2017). "Fruits," in *Springer Handbook of Odor*, ed. A. Buettner (Cham: Springer International Publishing).

Plotto, A., Baldwin, E., Bai, J., Narciso, J., Whitaker, V., and Chandler, C. (2013). Update on sensory evaluation of University of Florida strawberry selections. *Proc. Fla. State Hortic. Soc.* 126, 247–250.

Pyysalo, T., Honkanen, E., and Hirvi, T. (1979). Volatiles of wild strawberries, *Fragaria vesca* L., compared to those of cultivated berries, *Fragaria*. times. ananassa cv Senga Sengana. *J. Agric. Food Chem.* 27, 19–22. doi: 10.1021/jf60221a042

Rosa, G., Padovani, C. R., and Gianola, D. (2003). Robust linear mixed models with normal/independent distributions and Bayesian MCMC implementation. *Biom. J.* 45, 573–590.

Sánchez-Sevilla, J. F., Cruz-Rus, E., Valpuesta, V., Botella, M. A., and Amaya, I. (2014). Deciphering gamma-decalactone biosynthesis in strawberry fruit using a combination of genetic mapping, RNA-Seq and eQTL analyses. *BMC Genomics* 15:218. doi: 10.1186/1471-2164-15-218

Schwieterman, M. L., Colquhoun, T. A., Jaworski, E. A., Bartoshuk, L. M., Gilbert, J. L., Tieman, D. M., et al. (2014). Strawberry flavor: diverse chemical compositions, a seasonal influence, and effects on sensory perception. *PLoS One* 9:e88446. doi: 10.1371/journal.pone.0088446

Shaw, D. V. (1990). Response to selection and associated changes in genetic variance for soluble solids and titratable acids contents in strawberries. *J. Am. Soc. Hortic. Sci.* 115, 839–843.

Stavang, J. A., Freitag, S., Foito, A., Verrall, S., Heide, O. M., Stewart, D., et al. (2015). Raspberry fruit quality changes during ripening and storage as assessed by colour, sensory evaluation and chemical analyses. *Sci. Hortic.* 195, 216–225. doi: 10.1016/j.scienta.2015.08.045

Svensén, M., and Bishop, C. M. (2005). Robust Bayesian mixture modelling. *Neurocomputing* 64, 235–252. doi: 10.1016/j.neucom.2004.11.018

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

Ulrich, D., Hoberg, E., Rapp, A., and Kecke, S. (1997). Analysis of strawberry flavour–discrimination of aroma types by quantification of volatile compounds. *Z. Lebensm. Unters. Forsch. A* 205, 218–223.

Ulrich, D., Kecke, S., and Olbricht, K. (2018). What do we know about the chemistry of strawberry aroma? *J. Agric. Food Chem.* 66, 3291–3301. doi: 10.1021/acs.jafc.8b01115

Urrutia, M., Rambla, J. L., Alexiou, K. G., Granell, A., and Monfort, A. (2017). Genetic analysis of the wild strawberry (*Fragaria vesca*) volatile composition. *Plant Physiol. Biochem.* 121, 99–117. doi: 10.1016/j.plaphy.2017.10.015

Van Boekel, M. (2004). *Bayesian Solutions for Food-Science Problems?*. Bangkok: Frontis, 17–27.

Varela, P., Salvador, A., and Fiszman, S. (2005). Shelf-life estimation of 'Fuji' apples: sensory characteristics and consumer acceptability. *Postharvest Biol. Technol.* 38, 18–24. doi: 10.1016/j.postharvbio.2005.05.009

Whitaker, V. M., Hasing, T., Chandler, C. K., Plotto, A., and Baldwin, E. (2011). Historical trends in strawberry fruit quality revealed by a trial of University of Florida cultivars and advanced selections. *HortScience* 46, 553–557. doi: 10.21273/HORTSCI.46.4.553

Whitaker, V. M., Knapp, S. J., Hardigan, M. A., Edger, P. P., Slovin, J. P., Bassil, N. V., et al. (2020). A roadmap for research in octoploid strawberry. *Hortic. Res.* 7, 1–17. doi: 10.1038/s41438-020-0252-1

Whitaker, V. M., Osorio, L. F., Hasing, T., and Gezan, S. (2012). Estimation of genetic parameters for 12 fruit and vegetative traits in the University of

Florida strawberry breeding population. *J. Am. Soc. Hortic. Sci.* 137, 316–324. doi: 10.21273/JASHS.137.5.316

Whitaker, V. M., Peres, N. A., Osorio, L. F., Fan, Z., do Nascimento Nunes, M. C., Plotto, A., et al. (2019). 'Florida Brilliance'Strawberry. *HortScience* 54, 2073–2077. doi: 10.21273/HORTSCI14327-19

Wilkinson, D. J. (2007). Bayesian methods in bioinformatics and computational systems biology. *Brief. Bioinform.* 8, 109–116.

Young, H., Gilbert, J. M., Murray, S. H., and Ball, R. D. (1996). Causal effects of aroma compounds on Royal Gala apple flavours. *J. Sci. Food Agric.* 71, 329–336. doi: 10.1002/(SICI)1097-0010(199607)71:33.0.CO;2-8

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 06 April 2021
doi: 10.3389/fpls.2021.611449

Huanglongbing and Foliar Spray Programs Affect the Chemical Profile of "Valencia" Orange Peel Oil

*Xiuxiu Sun[1], Huqing Yang[1,2], Wei Zhao[1], Elise Bourcier[1], Elizabeth A. Baldwin[1], Anne Plotto[1], Mike Irey[3] and Jinhe Bai[1]**

[1] USDA/ARS Horticultural Research Laboratory, Fort Pierce, FL, United States, [2] Zhejiang A & F University, Hangzhou, China, [3] Southern Gardens Citrus Nursery, Clewiston, FL, United States

OPEN ACCESS

Edited by:
Zhi-Yan (Rock) Du,
University of Hawaii at Manoa,
United States

Reviewed by:
Riccardo Lo Bianco,
University of Palermo, Italy
Georgios Vidalakis,
University of California, Riverside,
United States
Mo-Xian Chen,
Chinese Academy of Sciences (CAS),
China

***Correspondence:**
Jinhe Bai
jinhe.bai@usda.gov

Specialty section:
This article was submitted to Plant Metabolism and Chemodiversity, a section of the journal Frontiers in Plant Science

Received: 16 October 2020
Accepted: 16 February 2021
Published: 06 April 2021

Citation:
Sun X, Yang H, Zhao W, Bourcier E, Baldwin EA, Plotto A, Irey M and Bai J (2021) Huanglongbing and Foliar Spray Programs Affect the Chemical Profile of "Valencia" Orange Peel Oil. Front. Plant Sci. 12:611449. doi: 10.3389/fpls.2021.611449

Florida orange trees have been affected by huanglongbing (HLB) for more than a decade. To alleviate disease-caused tree decline, maintain fruit productivity, and reduce disease transmission, enhanced foliar spray programs combining vector control and nutritional supplementation have been applied to healthy and diseased trees. The aim of this research was to discover if the various foliar sprays affect fruit peel oil chemical components. In this study, "Valencia" orange trees, with or without HLB (HLB±), were treated with the grower standard program (control, C) or one of four proprietary enhanced foliar spray programs (N1, N2, N3, and N4) over 16 months. Compared with HLB−, HLB+ samples had lower concentrations of typical peel oil components, including valencene, octanal, and decanal, and were abundant in oxidative/dehydrogenated terpenes, such as carvone and limonene oxide. However, limonene, the dominant component, was not affected by any treatment. Control and three out of four enhanced foliar spray programs, N2, N3, and N4, had very little influence on the chemical profiles of both HLB− and HLB+ samples, while N1 treatment greatly altered the chemical profile of HLB+ samples, resulting in peel oil similar to that of HLB− samples.

Keywords: citrus greening disease, nutritional spray, insect vector control, cold pressed oil, volatile organic compounds

INTRODUCTION

Huanglongbing (HLB, or citrus greening), associated with the bacterium *Candidatus Liberibacter* asiaticus (CLas), is a very destructive citrus disease and is vector-transmitted by the Asian citrus psyllid (ACP; *Diaphorina citri* Kuwayama). Once infected, trees rapidly decline and eventually die (Bové, 2006; Gottwald et al., 2020). Before the severe decline, HLB-infected trees still produce fruits, however, with decreased yield and poor quality. Fruits are typically small, irregularly shaped, with a thick, pale peel that remains green at the stylar end (Baldwin et al., 2014; Chen et al., 2016; McCollum et al., 2016). Fruits and juice from HLB-diseased trees are also associated with low soluble sugar and ethyl butanoate, as well as high acid, limonoid, and flavonoid content, resulting in less sweet and fruity flavor with more sour and bitter taste (Dala-Paula et al., 2018).

Furthermore, HLB causes an increase in pre-harvest fruit drop, which contributes to the reduction in yield (Baldwin et al., 2014; McCollum et al., 2016). This premature fruit drop may be exacerbated by *Lasiodiplodia theobromae* (previously known as *Diplodia natalensis*), the causal organism of citrus stem end rot (SER), which infects citrus fruits under the calyx abscission zone (Zhao et al., 2015, 2016). HLB has spread over all major citrus growing regions of Florida and has been the primary cause of $7.8 billion in lost revenue and more than 7,500 jobs in the Florida citrus industry since 2006 (Hodges and Spreen, 2012). The orange production of Florida was estimated at 45 million boxes for the 2017–2018 season, which represents a decline of more than 80% from the state's historic peak citrus production.

Previous research showed remission of citrus decline is more likely if trees remain vigorous by reducing stress caused by nutrient deficiencies, with or without application of plant growth regulators, to help host defense against HLB (Spann and Schumann, 2009; Stansly et al., 2014; Li et al., 2016). Several reports indicated that the application of mineral fertilizers, such as Zn, Cu, and Fe, alleviated the symptoms of HLB symptomatic trees and restored their productivity (Nwugo et al., 2013; Zhang et al., 2016). These favorable results were attributed to plant health maintenance and the promotion of root growth. In certain nutrient/systemic acquired resistance (SAR) programs, salicylic acid (SA) and/or its analogs were applied as foliar amendments to help the host defend against HLB by activating the SAR pathway. A combination of psyllid management and foliar nutrition/SAR sprays maintained tree health and productivity despite the presence of HLB (Rouse et al., 2012). However, any positive effects these treatments have on disease expression or fruit yield remain to be demonstrated (Stansly et al., 2014). In fact, no conclusive study has been conducted demonstrating how to control HLB by inducing plant defense (Li et al., 2016).

Citrus peel oil is the first byproduct obtained during the processing of citrus fruits and is widely used in foods, perfumery, and cosmetics (Dharmawan et al., 2009; Gonzalez-Mas et al., 2019). Citrus oil is also used as a cleaner, solvent, fungicide, and even aromatherapy material for humans (Dugo and Mondello, 2011). Citrus oil possesses strong insecticidal and biochemical activities (Oyedeji et al., 2020), as well as antimicrobial and antioxidant activities (Fancello et al., 2020; Oikeh et al., 2020). The essential oil from Mandarin revealed the inhibitory effects against *Staphylococcus aureus* (Song et al., 2020). A commercial citrus essential oil showed antimicrobial activity against *Escherichia coli* (Ambrosio et al., 2020). In addition, the Blanco peel essential oil exhibited potential for the treatment of skin acne (Hou et al., 2019). The volatile composition of citrus peel oil has also been extensively studied (Minh Tu et al., 2002; Njoroge et al., 2003; Dugo and Mondello, 2011; Liu et al., 2012; Lin et al., 2019). Aldehyde composition is the most important factor to evaluate citrus peel oil quality, and a better-quality oil usually has a higher concentration of aldehydes (Xu et al., 2017b). Aldehydes are known for their distinctly potent fragrances and are often main contributors to the overall flavor of an essential oil, and limonene is the major volatile compound in citrus peel oils (Spadaro et al., 2012). The quality of citrus oil is frequently affected by the climatic condition, disease, and harvest maturity (Vekiari et al., 2002; Zouaghi et al., 2019). However, there has been little research concerning the impact of HLB and nutritional/insecticide sprays on citrus peel oil volatiles. The objective of this study was to investigate the effect of enhanced foliar spray programs on volatile components of cold pressed oil from HLB+ and HLB− Valencia orange peel.

MATERIALS AND METHODS

Chemicals

Standards were obtained from the following sources: octanal, nonanal, decanal, terpinyl aldehyde, (*E*,*E*)-2,4-decadienal, geranial, neral, octanol, linalool, citronellol, nerol, α-pinene, myrcene, and β-phellandrene were purchased from Aldrich (Milwaukee, WI, United States); valencene was bought from Bedoukian (Danbury, CT, United States); undecanal from Analabs (North Haven, CT, United States); sabinene was obtained from Treatt (Lakeland, FL, United States); δ-3-carene was from K&K (Royse City, TX, United States); terpinen-4-ol from Advanced Biotech (Totowa, NJ, United States); and α-humulene from Fluka (Buchs, Switzerland).

Field Management and Foliar Spray Programs

Experiments were carried out on a commercial block of *Citrus sinensis* (L.) Osbeck cv. Valencia orange on "Swingle" citrumelo (*Citrus paradisi* × *Poncirus trifoliata*) rootstock planted in 2,000 at 3.7 m between trees and 7.3 m between rows at the Southern Garden Citrus Nursery grove in Clewiston, FL (26°45′N; 80°56′W). The training system was round shaped, and average tree height was about 4.75 m. At least one border tree between two different foliar spay programs was provided. The block consisted of 99 trees, and about half of the trees were HLB affected based on a visual diagnose in 2010 by the leaf, fruits and canopy symptoms (Bové, 2006). The trees had received common cultural practices and the grower's standard pest and disease management before the initiation of the trial in January 2012.

HLB+ and HLB− trees, 30 each, uniform in canopy size, were selected for the study by using a split-plot design with three replicates. Each plot consisted of two subplots, HLB+ and HLB−, and each subplot consisted of two trees. Each plot was treated with one of the following five foliar spray programs for over 16 months from the beginning of 2012 to May 2013, each program included multiple applications, and each application contained macro- (N, P, and K) and micro- (Ca, Mg, Zn, B, Fe, and Mo) nutrients and vector control agents (chlorpyrifos, imidacloprid, abamectin, fenpropathrin, malathion, ζ-cypermethrin, and mineral oil), independently or in combination. The control (C) was the grower standard (**Table 1**) with eight sprays from April 4, 2012 to May 14, 2013. Each foliar spray program (N1, N2, N3, and N4) was featured by recommendation of one or more nutrient suppliers or grower groups. The exact composition of each foliar spray is proprietary. The common enhancement was an increase in spray times from 8× in the control to 10× in N1 and N4, 11× in N3, and 12× in

N2. All N1–N4 had a late dormant/spring flush in Feb 2012, in comparison with the earliest spray for control, which was April 4, 2012 (**Table 1**).

Fruit Harvest and Sample Processing

Fruits were harvested on April 28, 2014. For each replicate, approximately 38.5 kg (56 fruits) were harvested from the two trees. Fruits were washed and processed by using a standard processing method with a JBT extractor (JBT[§] FoodTech, Lakeland, FL, United States) at premium setting (Bai et al., 2013). Juice samples were collected for *C*Las titer testing; and "frit" peel tissues, which are located in around of the stem end and are rich in peel oil, were also collected for peel oil extraction. Briefly, the frit peel tissues were cold pressed with a manual oil extractor, the collected emulsion allowed to settle for 30 min, and then the top oil layer was collected and centrifuged at 3,500 × *g* at 25°C for 15 min. Finally, the peel oil supernatant was collected and stored at −20°C for analysis. To protect the oil samples from oxidation, the headspace of the sample vial was flushed with nitrogen gas before air-tight capping.

DNA Extraction and qPCR Detection of *Candidatus Liberibacter* asiaticus in Leaves and Fruit Juice

For leaf samples, 10 leaves from each tree were randomly taken on the same day as fruit harvest, and DNA was extracted from 100 mg of the midrib tissues by following Li et al.'s (2006) procedures. Primers targeting *C*Las 16S rDNA (Li primers) were used (Li et al., 2006; Baldwin et al., 2018); and TaqMan qPCR was performed in a 7500 Real-Time PCR system (Applied Biosystems, Inc., Carlsbad, CA, United States).

For fruit juice samples, DNA was extracted from 500 μl of orange juice sample using a modified cetyl trimethylammonium bromide (CTAB) method (Zhao et al., 2018), and *C*Las level was quantified by qPCR as previously described (Zhao et al., 2018) using primers targeting *C*Las hyv1 (LJ primers) (Morgan et al., 2012); SYBR Green qPCR was performed in a 7,500 Real-Time PCR system (Applied Biosystems). The default melt curve (disassociation) stage was continued after the 40 cycles of PCR. Quantification cycle (Cq) values were analyzed using ABI 7,500 Software version 2.0.6 (Applied Biosystems) with a threshold setting of 0.02 and automated baseline settings.

Peel Oil Volatile Composition Analysis

The volatile composition analysis of the oil was carried out using a gas chromatography–mass spectrometry (GC-MS, 6890N GC and 5975 MS, Agilent Technologies, Santa Clara, CA, United States) system equipped with a split/splitless injector and a DB-5 capillary column (60-m length, 0.25-mm diameter, and 1-μm film thickness; J&W Scientific, Folsom, CA, United States). The injector and detector temperature was 260°C. The injection volume was 1 μl with a split ratio of 40:1. The oven conditions were 40°C (0.5 min) and then 4°C·min^{-1} to 225°C (held for 13.25 min) for a total run time of 60 min. Helium was used as carrier gas at flow rate of 1.5 ml·min^{-1}. Inlet, ionizing source, and transfer line were kept at 250, 230, and 280°C, respectively. The mass spectrometry data were recorded in the scan mode from 40 to 400 m/z at 2 scans·s^{-1} with an ionization energy of 70 eV.

Identification of Volatile Compounds

Data were collected using the ChemStation G1701 AA data system (Hewlett-Packard, Palo Alto, CA, United States). A mixture of C-5 to C-18 *n*-alkanes was run at the beginning of each day to calculate retention indices (RIs) (Bai et al., 2014). The volatile components were identified by matching their spectra with those from NIST/EPA/NIH Mass Spectral Library (NIST 14) and authentic volatile compound standards, as well as by comparing their RIs with corresponding literature data (Lota et al., 2002; Deng et al., 2017). Quantification of major peel oil volatile components (limonene, hexanol, hexanal, linalool, etc.) was conducted by using a peak size vs. concentration curve built by a series of diluted standard solutions (Bai et al., 2002).

Statistical Analyses

Data presented were the mean values of three biological replicates. Statistical analysis was performed with JMP (version 11.2.2; SAS Institute, Cary, NC, United States). Differences were tested for significance by using a one-way analysis of variance (ANOVA), incorporating a split-plot design. Mean separations were examined by Tukey's HSD tests with the significance level at 0.05. For both principal component analysis (PCA) and hierarchical cluster analysis, the complete dataset including all replicates was performed. To visualize quantitative results, a heatmap was designed to present all volatile compounds detected in 10 combinations of foliar spray programs × tree types with the relative amounts in each chemical compound.

RESULTS

Candidatus Liberibacter asiaticus Infection Severity of Leaves and Juice

The qPCR analysis of leaves confirmed the *C*Las uninfected status of "healthy" trees (HLB−), with high Cq values (>38.29), while diseased trees (HLB+) had low Cq values (<21.76;**Table 2**). The cutoff Cq value between HLB+ and HLB− for Li primers has been determined to be 32 (Zhao et al., 2018).

Because the *C*Las titer in orange juice is much lower than in leaf tissue, we used another pair of primers targeting at *C*Las hyv1, which has more copies in the genome than that of rDNA (Morgan et al., 2012). With these primers, a cutoff Cq value of 29 was determined to differentiate between HLB+ and HLB− (Zhao et al., 2018). Similar to healthy leaf tissue, healthy juice had Cq values higher than the cutoff value, with N4 HLB− right at the cusp (Cq = 29.03) (**Table 1**). For HLB+ juice, N1 treatment resulted in a higher Cq value (29.63) than N2 (27.24) and N3 (27.90) treatments and slightly higher than cutoff value, indicating that N1 may have some effects on reducing *C*Las titer in juice (**Table 2**).

TABLE 1 | Chemicals/agents used for the grower standard/control (C) foliar spray program and the application regime.

Spray product	Active ingredient	Function	Company
A. Characteristics of chemicals/agents used for foliar application			
Nutri-Phite Magnum 2-40-16	N, P, K	Macronutrients	Verdesian Life Sciences, Cary, NC, United States
Dual Phos 14-7-7	N, P, K	Macronutrients	Griffin Fertilizer Company, Frostproof, FL, United States
Potassium nitrate 3-0-11	N, K	Macronutrients	Haifa North America, Altamonte Springs, FL, United States
Citrite 779	N, Mg, Zn, Fe	Macro/Micronutrients	Chemical Dynamics, Inc., Plant City, FL, United States
Soar Bloom	Ca, Mg, B, Mo	Micronutrients	Chemical Dynamics, Inc., Plant City, FL, United States
Solubor	B	Micronutrient	U.S. Borax, Boron, CA, United States
Nitro-Gold Nut # 1110243	Fe, Mn, Zn	Micronutrient	Southern Ag, Palmetto FL, United States
Whirlwind	Chlorpyrifos	Insecticide	Helena Chemical Co., Collierville, TN, United States
Drexel 80/20 surfactant	Alcohol ethoxylates/glycols	Adjuvant	Drexel Chemical, Memphis, TN, United States
Humac/Surfactant	Alcohol ethoxylates/glycols	Adjuvant	Drexel Chemical, Memphis, TN, United States
435 Spray Oil	Mineral oil	Insecticide	Drexel Chemical, Memphis, TN, United States
Imidacloprid 2F	Imidacloprid	Insecticide	Control Solutions, Inc., Pasadena, TX, United States
Agri-Mek SC	Abamectin	Insecticide	Syngenta, LLC, Vero Beach, FL, United States
Danitol	Fenpropathrin	Insecticide	Valent United States Corporation, Longwood, FL, United States
Malathion	Malathion	Insecticide	Griffin Fertilizer Company, Frostproof, FL, United States
Portal Miticide/Insecticide	Mitochondrial Electron Transport Inhibitor	Insecticide	Nichino America, Inc., Wilmington, DE, United States
Mustang	Zeta-cypermethrin	Insecticide	FMC Corporation, Philadelphia, PA, United States
Nu Cop	Cu	Fungicide	Albaugh Inc., Ankeny, IA, United States
Gem 500 SC	Trifloxystrobin	Fungicide	Bayer, Pittsburgh, PA, United States
Kocide 3000	Cu	Fungicide	Kocide LLC, Houston, TX, United States
B. Foliar application regime			
Application time (phenologic phase)	**Spray product**	**Dose (amount per ha)**	
4-Apr-2012	Nutri-Phite	2.34	L
(bloom spray)	Soar Bloom	4.68	L
	Kocide 3000	2.24	kg
	Humac/Surfactant	0.56	kg
19-Jun-2012	435 Spray oil	28.06	L
(1st summer oil spray)	Kocide 3000	2.24	kg
	Liquid Potassium Nitrate	42.09	L
	Nutri-Phite Magnum 2-40–16	2.34	L
	Nitro-Gold Nut #1110243	7.02	L
	Imidacloprid 2F	0.56	kg
	Agri-Mek SC	0.15	kg
2-Aug-2012	Gem 500 SC	0.18	kg
(2nd summer oil spray)	Kocide 3000	2.24	kg
	Liquid potassium nitrate	42.09	L
	Citrite 779	7.02	L
	Whirlwind	5.85	L
	Drexel 80/20 Surfactant	0.56	kg
20-Sep-2012	435 Spray oil	28.06	L
(3rd summer oil spray)	Kocide 3000	2.24	kg
	Liquid potassium nitrate	42.09	L
	Nutri-Phite Magnum 2-40-16	2.34	L
	Nitro-Gold Nut #1110243	7.02	L
	Imidacloprid 2F	1.12	kg
	Agri-Mek SC	0.30	kg
17-Dec-2012	Danitol	1.12	kg
(dormant spray)	Drexel 80/20 surfactant	0.21	kg
25-Jan-2013	Dual Phos 14-7-7	65.48	L

(Continued)

TABLE 1 | Continued

Application time (phenologic phase)	Spray product	Dose (amount per ha)	
(flash spray)	Malathion	1.75	L
	80/20	0.56	kg
5-Mar-2013	Soar Bloom	4.68	L
(prebloom spray)	Nutri-Phite	2.34	L
	Potassium Nitrate 3-0-11	42.09	L
	Nu Cop	2.24	kg
	Portal	4.68	L
	80/20	0.56	kg
14-May-2013	Kocide 3000	2.24	kg
(postbloom spray)	Potassium Nitrate 3-0-11	42.09	L
	Nutri-Phite	2.34	L
	Nitro-Gold Nut #1110243	7.02	L
	Mustang	0.30	kg
	80/20	0.56	kg
	Solubor	1.12	kg

TABLE 2 | Quantification cycle (Cq) values of *C*Las in "Valencia" orange leaf midrib and fruit tissue samples affected by Huanglongbing (HLB) and foliar spray programs.

Foliar spray program [z]	HLB−	HLB+
*C*Las 16S rDNA in leaf midrib		
C	$38.62 \pm 3.38^{a,y}$	20.75 ± 0.68^{b}
N1	40.00 ± 0.00^{a}	21.76 ± 0.66^{b}
N2	38.66 ± 1.99^{a}	21.24 ± 0.71^{b}
N3	39.36 ± 1.57^{a}	21.72 ± 0.73^{b}
N4 ANOVA significance	$38.29 \pm 2.00^{a***}$	21.33 ± 0.20^{b}
*C*Las hyv1 in orange juice		
C	$33.45 \pm 1.82^{b,c}$	$28.97 \pm 0.97^{c,d}$
N1	40.00 ± 0.00^{a}	29.63 ± 0.52^{c}
N2	32.07 ± 2.99^{c}	27.24 ± 1.37^{d}
N3	$38.41 \pm 2.75^{a,b}$	27.90 ± 0.95^{d}
N4 ANOVA significance	$29.03 \pm 0.97^{c***}$	27.57 ± 2.52^{d}

*[z]Spray treatments: C, control; N1–N4, four experimental foliar nutrient spray programs. [y]Means ± standard deviations within the same block with the same letter are not significantly different by Tukey's honestly significant difference (HSD) ($p < 0.05$). ***indicates significant at $p < 0.001$.*

The Effect of Huanglongbing on the Volatile Profile of "Valencia" Citrus Peel Oil

A total of 53 volatile compounds were identified in peel oil samples (**Table 3**), including seven monoterpenes, 14 sesquiterpenes, 13 terpene oxides, 10 aldehydes, 8 alcohols, and 1 ketone. Monoterpenes were predominant in all the "Valencia" citrus peel oil samples with limonene accounting for the major constituent (89.74–90.48%, **Table 3**). In general, there were more differences among all the treatments within HLB+ fruits (23 volatile compounds total) than among all the treatment within HLB− fruits (hexanal only).

Huanglongbing disease dramatically affected the volatile profile of citrus fruits (**Figures 1**, **2**). Cluster analysis (**Figure 1**) showed which compounds were more abundant in HLB− (in red) versus HLB+ (in green) and how the various treatments affected the profile. Hierarchical clustering shows that all the HLB+ N1 samples (dark pink/green) are clustered with the HLB− samples (red) (**Figure 2**) along with one replicate of HLB+ N4 sample (light purple/green). Analysis by PCA discriminated HLB+ from HLB− samples in both PC1 and PC2 (**Figure 3B**), explaining 52.8% of the variation in the first two components (**Figure 3A**). Most of peel oil volatiles from HLB+ samples were on the negative side of PC1 and positive side of PC2 (**Figures 3A,B**).

The Effect of Nutritional Spray Treatments on the Volatile Profile of "Valencia" Citrus Peel Oil

The treatment N1 showed the strongest impact on the volatile profile, especially for the HLB+ samples (**Table 3** and **Figures 1**, **2**). Treatment N1, for example, reduced (less green) compounds prevalent in HLB+ (e.g., sabinene) and enhanced (more red) some compounds prevalent in HLB− (e.g., α-terpinyl acetate), thus resulting in a profile more closely resembling peel oil from HLB− than oil from HLB+ trees subjected to other treatments. The N1 treatment, which had the highest fruit juice Cq values in both HLB+ and HLB− fruit tissue (**Table 2**), significantly reduced the concentration of hexanol, sabinene, δ-3-carene, (*E*)-limonene oxide, citronellal, *p*-menth-1-en-9-al, (*Z*)-carvone, and (*E*)-β-farnesene but significantly increased the concentration of (*Z*)-*p*-mentha-2,8-dien-1-ol, decanal, (*E*,*E*)-2,4-decadienal, α-terpinyl acetate, α-copaene, (*E*)-caryophyllene, valencene, α-selinene, δ-cadinene, and caryophyllene oxide (**Table 3** and **Figure 3B**) for HLB+ samples compared with the HLB+ control. Comparing HLB+ with HLB− within N1 treatment showed widespread similarity, with the noticeable exceptions of octadecanal, δ-3-carene, and β-phellandrene, which were more abundant in HLB− than in HLB+ peel oil. All four replicate samples of HLB+

TABLE 3 | Chemical compositions of peel oils of "Valencia" oranges affected by Huanglongbing (HLB) and foliar spray programs N1, N2, N3, and N4, with C as control fertilization program.

Peak	Compound	RI	HLB−					HLB+					ANOVA significance
			C[z]	N1	N2	N3	N4	C	N1	N2	N3	N4	
1	Hexanal	795	0.02^{a,by}	0.01^{b,c}	0.01^{a,b,c}	0.01^{b,c}	0.02^{a}	0.01^{c}	0.01^{c}	0.01^{c}	0.01^{c}	0.01^{c}	***
2	Hexanol	868	0.01^{b}	0.01^{a,b}	0.01^{a,b}	0.01^{a}	0.01^{a,b}	0.01^{a}	0.01^{a,b}	0.01^{a,b}	0.01^{a,b}	0.01^{a,b}	*
3	α-Pinene	954	1.11	1.14	1.10	1.11	1.16	1.15	1.11	1.08	1.07	1.11	NS
4	Sabinene	993	0.66^{a,b}	0.48^{b}	0.63^{a,b}	0.74^{a,b}	0.73^{a,b}	0.99^{a}	0.61^{a,b}	0.88^{a,b}	0.73^{a,b}	0.87^{a,b}	*
5	Myrcene	1,002	2.56	2.59	2.54	2.55	2.68	2.59	2.49	2.45	2.41	2.51	NS
6	Octanal	1,012	0.72	0.60	0.69	0.62	0.70	0.58	0.67	0.47	0.56	0.54	NS
7	α-Phellandrene	1,025	0.11	0.11	0.10	0.10	0.11	0.10	0.10	0.10	0.09	0.09	NS
8	δ-3-Carene	1,032	0.14	0.13	0.11	0.10	0.15	0.13	0.08	0.14	0.11	0.13	NS
9	Limonene	1,050	90.07	90.02	89.96	89.99	89.74	89.74	89.95	90.19	90.48	90.23	NS
10	β-Phellandrene	1,063	0.21	0.25	0.22	0.20	0.21	0.21	0.17	0.23	0.18	0.20	NS
11	Octanol	1,076	0.15	0.17	0.17	0.18	0.15	0.19	0.17	0.16	0.19	0.15	NS
12	Linalool	1,109	0.83	0.85	0.83	0.84	0.88	0.92	0.90	0.84	0.88	0.87	NS
13	Non-anal	1,112	0.15	0.14	0.15	0.14	0.15	0.14	0.16	0.13	0.13	0.13	NS
14	(*E*)-*p*-Mentha-2,8-dien-1-ol	1,139	0.01^{c}	0.01^{c}	0.01^{b,c}	0.01^{b,c}	0.01^{b,c}	0.01^{b,c}	0.01^{b,c}	0.02^{a}	0.01^{a,b}	0.01^{a,b}	***
15	(*Z*)-*p*-Mentha-2,8-dien-1-ol	1,153	0.01^{a,b}	0.01^{a,b}	0.01^{a,b}	0.01^{a,b}	0.01^{a,b}	0.00^{c}	0.01^{a,b}	0.01^{a}	0.01^{a,b}	0.00^{b,c}	***
16	(*E*)-Limonene oxide	1,157	0.02^{a}	0.02^{a}	0.02^{a}	0.02^{a}	0.02^{a}	0.02^{a}	0.02^{a}	0.02^{a}	0.02^{a}	0.02^{a}	*
17	Citronellal	1,161	0.13	0.12	0.12	0.12	0.13	0.14	0.11	0.12	0.13	0.14	NS
18	Non-anol	1,173	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	NS
19	Camphor	1,175	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	NS
20	Terpinen-4-ol	1,199	0.02^{c}	0.02^{c}	0.02^{b,c}	0.02^{b,c}	0.02^{b,c}	0.02^{a,b,c}	0.02^{a,b,c}	0.03^{a}	0.02^{a,b,c}	0.03^{a,b}	**
21	Decanal	1,210	0.95	0.93	0.97	0.92	0.94	0.85	1.00	0.82	0.82	0.80	NS
22	Nerol	1,230	0.04	0.04	0.04	0.04	0.04	0.05	0.04	0.05	0.05	0.05	NS
23	*p*-Menth-1-en-9-al	1,235	0.05^{a,b}	0.05^{a,b}	0.05^{a,b}	0.05^{a,b}	0.04^{b}	0.06^{a,b}	0.05^{a,b}	0.07^{a}	0.06^{a,b}	0.06^{a,b}	*
24	(*Z*)-Carvone	1,237	0.01^{a,b,c}	0.01^{a,b,c}	0.01^{b,c}	0.01^{b,c}	0.01^{b,c}	0.01^{a,b,c}	0.01^{c}	0.01^{a,b}	0.01^{a,b,c}	0.01^{a}	**
25	Neral	1,249	0.16	0.15	0.16	0.15	0.16	0.16	0.16	0.14	0.14	0.15	NS
26	Geraniol	1,256	0.03	0.04	0.04	0.04	0.03	0.04	0.04	0.04	0.04	0.03	NS
27	Carvone	1,261	0.01^{b}	0.01^{b}	0.01^{b}	0.01^{b}	0.01^{b}	0.02^{a,b}	0.01^{b}	0.03^{a}	0.02^{a,b}	0.02^{a,b}	**
28	(*E*,*E*)-Decenal	1,265	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	NS
29	Geranial	1,275	0.23	0.21	0.23	0.22	0.22	0.23	0.22	0.20	0.20	0.21	NS
30	Perilla aldehyde	1,295	0.05	0.04	0.05	0.05	0.06	0.05	0.04	0.05	0.05	0.05	NS
31	(*E*,*E*)-2,4-Decadienal	1,321	0.02^{a,b,c}	0.02^{a,b,c}	0.02^{a}	0.02^{a,b}	0.02^{a,b,c}	0.01^{c}	0.02^{a}	0.01^{b,c}	0.01^{a,b,c}	0.01^{a,b,c}	**
32	α-Terpinyl acetate	1,358	0.04	0.05	0.03	0.04	0.04	0.02	0.03	0.04	0.03	0.02	NS
33	α-Copaene	1,401	0.09^{a,b}	0.12^{a}	0.10^{a,b}	0.09^{a,b}	0.09^{a,b}	0.07^{b}	0.11^{a,b}	0.08^{b}	0.09^{a,b}	0.07^{b}	**
34	Dodecanal	1,405	0.23	0.21	0.23	0.21	0.21	0.21	0.24	0.19	0.19	0.19	NS
35	β-Cubebene	1,412	0.14B	0.13	0.15	0.17	0.15	0.19A	0.16	0.21	0.17	0.20	NS
36	(*Z*)-Carveol	1,416	0.04	0.05	0.04	0.05	0.05	0.03	0.04	0.04	0.03	0.03	NS
37	(*E*)-Caryophyllene	1,453	0.06^{b}	0.08^{a,b}	0.10^{a,b}	0.08^{a,b}	0.07^{b}	0.09^{a,b}	0.12^{a}	0.11^{a}	0.09^{a,b}	0.09^{a,b}	**
38	(*E*)-β-Farnesene	1,455	0.02^{a,b,c}	0.01^{c}	0.02^{a,b,c}	0.03^{a,b,c}	0.03^{a,b,c}	0.03^{a,b}	0.02^{b,c}	0.04^{a}	0.02^{a,b,c}	0.04^{a}	**

(Continued)

TABLE 3 | Continued

Peak	Compound	RI	HLB−					HLB+				ANOVA significance
			C[z]	N1	N2	N3	N4	C	N1	N2	N3	N4	
39	β-Copaene	1,459	0.16	0.16	0.16	0.17	0.16	0.15	0.17	0.16	0.16	0.15	NS
40	α-Humulene	1,488	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.02	0.02	NS
41	γ-Cadinene	1,501	0.02	0.02	0.02	0.02	0.02	0.01	0.02	0.01	0.02	0.01	NS
42	α-Farnesene	1,506	0.06B	0.06	0.06	0.07	0.07	0.08A	0.06	0.07	0.07	0.08	NS
43	Germacrene D	1,513	0.06	0.06	0.04	0.06	0.05	0.05	0.06	0.05	0.05	0.05	NS
44	γ-Selinene	1,515	0.00	0.01	0.03	0.01	0.01	0.01	0.02	0.01	0.01	0.01	NS
45	Valencene	1,524	$0.17^{b,c}$	0.30^{a}	$0.26^{a,b,c}$	$0.21^{a,b,c}$	$0.18^{a,b,c}$	0.14^{c}	$0.29^{a,b}$	$0.22^{a,b,c}$	$0.18^{a,b,c}$	0.15^{c}	**
46	α-Selinene	1,529	$0.02^{a,b,c}$	0.02^{a}	$0.02^{a,b}$	$0.02^{a,b,c}$	$0.02^{a,b,c}$	0.01^{c}	$0.02^{a,b}$	$0.02^{a,b,c}$	$0.02^{b,c}$	0.01^{c}	***
47	δ-Cadinene	1,545	$0.11^{a,b}$	0.13^{a}	$0.12^{a,b}$	$0.11^{a,b}$	$0.11^{a,b}$	0.09^{b}	0.13^{a}	$0.10^{a,b}$	$0.10^{a,b}$	0.09^{b}	**
48	Germacrene D-4-ol	1,612	0.02^{a}	$0.02^{a,b}$	$0.01^{a,b}$	$0.01^{a,b}$	$0.02^{a,b}$	$0.01^{a,b}$	$0.01^{a,b}$	$0.01^{a,b}$	$0.01^{a,b}$	0.01^{b}	*
49	Caryophyllene oxide	1,629	0.02^{a}	$0.03^{a,b}$	$0.02^{a,b}$	$0.02^{a,b}$	$0.02^{a,b}$	$0.02^{a,b}$	0.03^{a}	$0.02^{a,b}$	$0.02^{a,b}$	$0.02^{a,b}$	*
50	β-Gurjunene	1,663	$0.01^{a,b}$	$0.01^{a,b}$	0.01^{b}	0.02^{a}	$0.01^{a,b}$	$0.01^{a,b}$	$0.01^{a,b}$	$0.01^{a,b}$	0.01^{a}	$0.01^{a,b}$	*
51	α-Sinensal	1,773	0.10	0.10	0.10	0.11	0.11	0.11	0.09	0.10	0.10	0.10	NS
52	Octadecanal	1,828	0.03	0.10	0.04	0.05	0.03	0.05	0.04	0.04	0.04	0.04	NS
53	Nootkatone	1,873	0.04	0.05	0.04	0.04	0.04	0.03	0.04	0.04	0.04	0.03	NS

[z]*Spray treatments: C, control; N1–N4, four experimental foliar nutrient spray programs.* [y]*Means within the same row with the same letter are not significantly different by Tukey's honestly significant difference (HSD) ($p < 0.05$). *, **, and *** indicate significance at $p < 0.05$, 0.01, and 0.001, respectively; no significant effect is denoted by NS.*

N1 and one replicate sample of HLB+ N4 were positioned on the negative side of PC2, together with most of HLB− samples (**Figure 3A**).

DISCUSSION

Infection Severity of Fruit

A strict ACP control program had been thoroughly enforced since HLB was detected in the citrus orchard throughout the entire foliar spray research period. The qPCR results indicate all HLB− trees remained *C*Las-negative for the duration of the experiments (**Table 2**). All HLB+ trees remained *C*Las-positive (**Table 2**), indicating that the enhanced nutritional and ACP vector control applications (such as in **Table 1**) did not kill *C*Las. However, the qPCR tests using juice samples provided more complex results: in the HLB− trees with N4 foliar spray programs, the Cq value was 29.03 in the juice samples (close to the 29 cutoff value for juice samples determining *C*Las infection by LJ primers), while it was 38.29 in leaf midribs (well above the cutoff value of 32 for leaf *C*Las infection by Li primers) (**Table 2**). On the other hand, in the HLB+ N1 juice samples, a relatively high Ct value of 29.63 (just above the cutoff value for juice *C*Las infection by LJ primers) was determined, indicating the infection was less severe than in the other HLB+ samples (**Table 2**). The possible interpretations are as follows: although *C*Las is a phloem-restricted bacterium, and leaf midribs are rich in phloem vessels and thus high in *C*Las titers (Li et al., 2006; Bai et al., 2013), due to the uneven distribution of the pathogen organisms in the tree, leaf samples taken from a tree may test negative for *C*Las, even if some branches are already infected, especially at the early infection stages when only few branches are infected (Bai et al., 2013; McCollum et al., 2016). Although *C*Las titers in fruit juice are much lower than in leaves (Bai et al., 2013; Zhao et al., 2018), juice samples were from fruit harvested from the entire tree, including the infected branches, and thus represent the entire tree status better than a leaf sample. The difficulty of detecting *C*Las in juice samples was overcome by using a pair of primers with high copy numbers in the genome. It has been reported that qPCR using LJ primers targeting *C*Las hyv1 DNA improved the *C*Las detection accuracy in comparison with Li primers (Morgan et al., 2012), because there are more copies of hyv1 gene in the *C*Las genome than 16s rDNA, and therefore, the relative detectable threshold by LJ primers can be reduced by 7–11 cycles for leaf samples (Morgan et al., 2012) and 2–5 cycles for fruit juice samples (Zhao et al., 2018).

The Effect of Huanglongbing on the Volatile Profile of "Valencia" Citrus Peel Oil

The volatile profiles for both HLB− and HLB+ in this study are similar to what has been reported in previous research (Lota et al., 2002; Njoroge et al., 2003; Huang et al., 2017). Previous research showed that HLB resulted in a significant reduction in

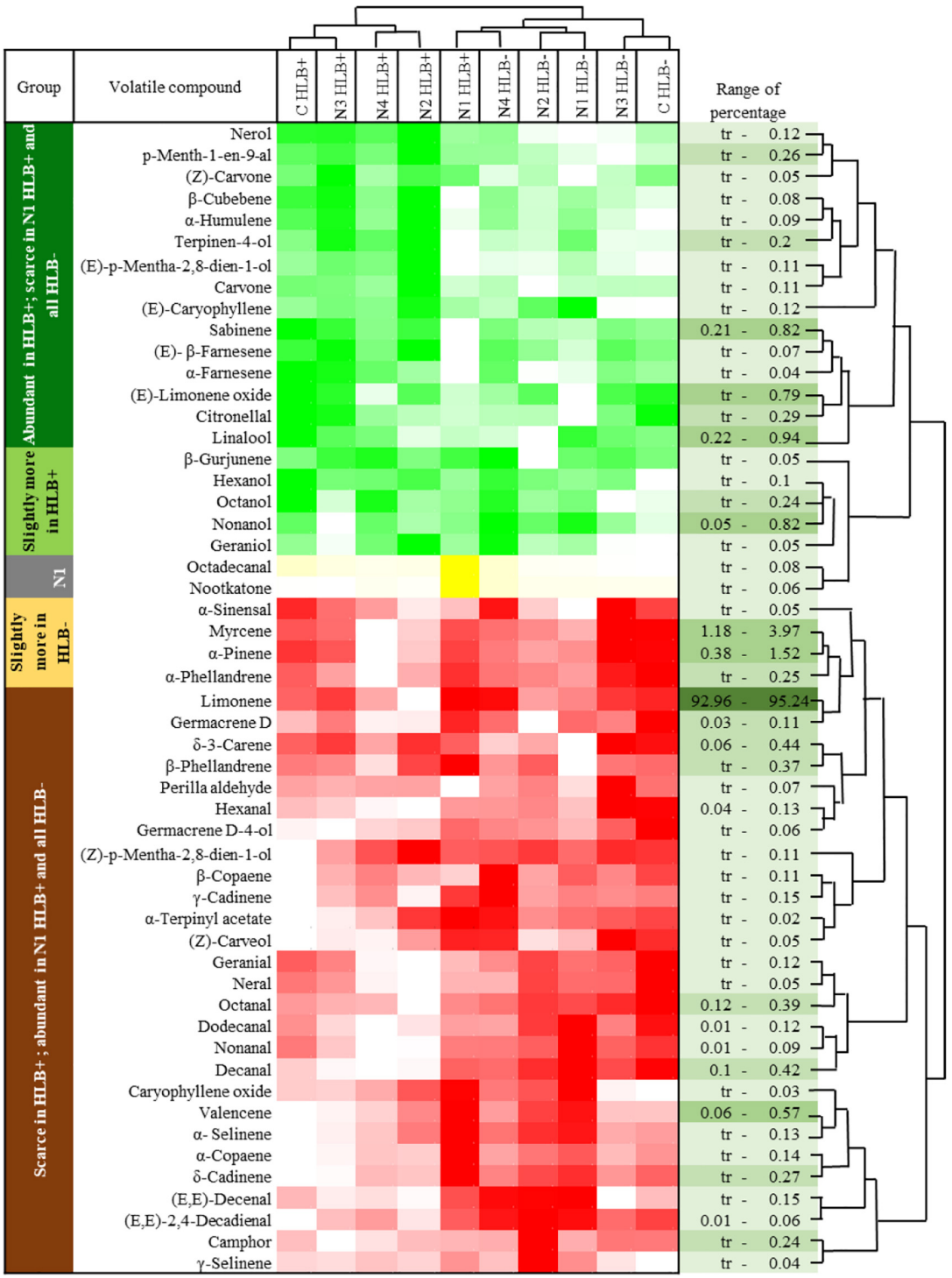

FIGURE 1 | Chemical groups enhanced and suppressed by huanglongbing (HLB) determined by average concentration of volatiles in peel oils extracted from "Valencia" oranges affected by HLB and foliar spray programs (C, control; N1–N4, four different foliar nutrient spray programs). Green and red represent enhanced and suppressed volatiles, respectively. Yellow represents volatiles only enhanced by N1 HLB+ combination. Color density indicates the relative content.

aldehydes, peel oil aroma volatiles formed during the normal ripening of HLB− fruits (Kiefl et al., 2018). In this research, the content of aldehydes hexanal, octanal, nonanal, decanal, and dodecanal was also significantly reduced in the HLB+ samples (**Table 3** and **Figure 1**), confirming that HLB negatively affects peel oil quality. Furthermore, similar to previous research, citrus

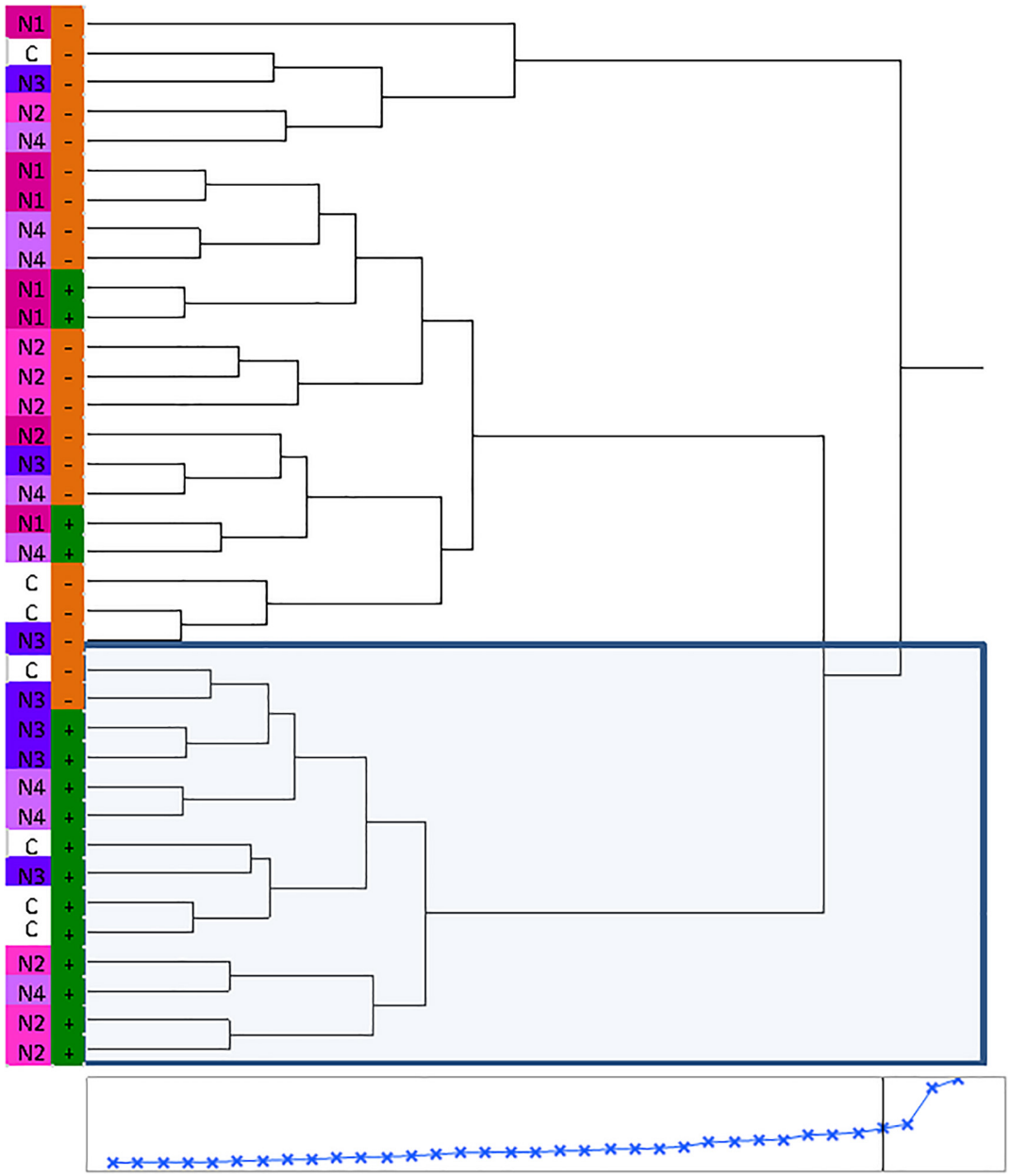

FIGURE 2 | Hierarchical clustering of volatile chemicals in peel oils extracted from "Valencia" oranges affected by Huanglongbing (HLB) and foliar spray programs (C, control; N1–N4, four different foliar nutrient applications; + in green, HLB+; and – in orange, HLB–).

peel oil derived from HLB+ fruits had lower concentrations of citronellal and geranial than oil derived from HLB− samples (**Table 3**; Xu et al., 2017a). Additionally, some esters, such as α-terpinyl acetate, decreased in HLB+ compared with HLB− samples (Plotto et al., 2008; Baldwin et al., 2010). Conceivably, severe HLB infection could substantially inhibit host secondary metabolism and volatile formation (Xu et al., 2017a).

The oxidative/dihydrogen compounds typical of terpenes, such as carvone and limonene oxide, were significantly higher in HLB+ samples. It is possible that these compounds increased because of the stress induced by HLB. These compounds may have all increased due to oxidative stress, which occurred in trees with severe HLB symptoms, creating greater concentrations of these volatiles (Xu et al., 2017a).

The Effect of Nutritional/Insecticidal Spray Treatments on the Volatile Profile of "Valencia" Citrus Peel Oil

Florida growers have been using foliar nutritional spray products that often contain micro- and macro-nutrients to compensate for lack of nutrient assimilation due to the HLB disease and

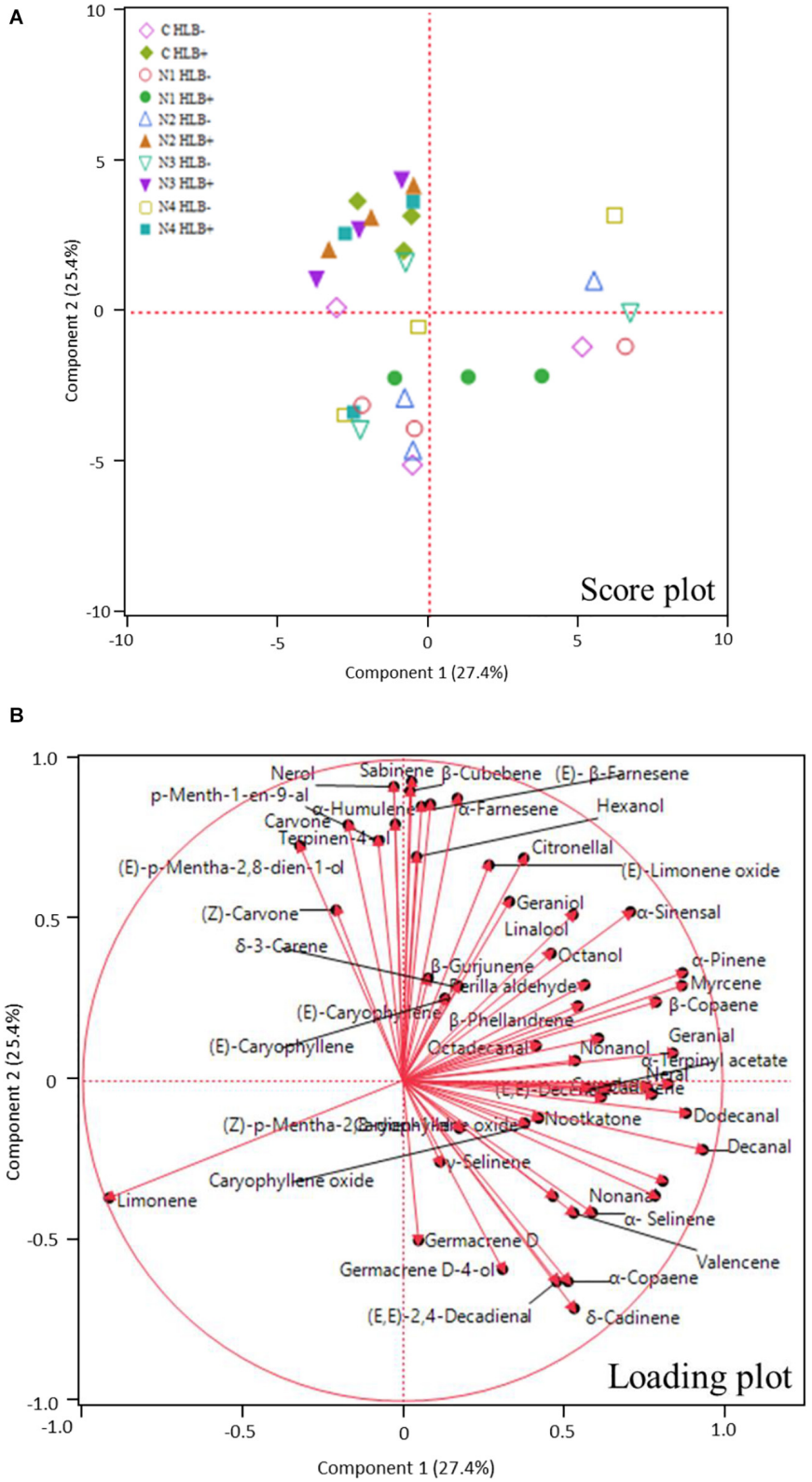

FIGURE 3 | Principle component analysis (PCA) score plot **(A)** and loading plot **(B)** of volatile compounds in peel oils of "Valencia" oranges affected by Huanglongbing (HLB) and foliar spray programs (C, control; N1–N4, four different foliar nutritional sprays).

compounds that are believed to activate SAR pathways in plants to increase tree defense response (Masaoka et al., 2011; Gottwald et al., 2012). The results in this experiment showed that one of the experimental foliar spray programs, N1, substantially altered the chemical profiles in HLB+ peel oil, and interestingly, the chemical profile turned out to be similar to that of HLB− peel oil. In N1 sprayed HLB+ peel oil, the most significant changes were recovery of valencene and other sesquiterpenes, which were suppressed by HLB (**Table 3** and **Figure 3**). Phylogenetic analysis of plant terpene synthase genes localized *Cstps1* to the group of angiosperm sesquiterpene synthases (Sharon-Asa et al., 2003). Within this group, *Cstps1* belongs to a subgroup of citrus sesquiterpene synthases (Sharon-Asa et al., 2003; Yu et al., 2019). *Cstps1* was found to be developmentally regulated: transcripts were found to accumulate only during fruit maturation, which corresponds to the timing of valencene accumulation in fruits (Sharon-Asa et al., 2003). HLB causes deficiency of valencene and other sesquiterpenes most likely due to the delay of fruit maturation (Dala-Paula et al., 2018), and the recovery of valencene and other sesquiterpenes indicates that the N1 spray program improved tree defense response, through either nutritional improvement, vector control preventing ACP mediated exacerbation of HLB, or a combination of both (Stansly et al., 2014; Li et al., 2016). **Table 2** clearly shows that the N1 spray program significantly reduced the *C*Las titer in orange juice sample. Further research is required to confirm if the presence of HLB results in a downregulation of *Cstps1*, and if a N1-like spray program can enhance the recovery.

As the experimental foliar sprays are of proprietary formulation, we do not know the exact difference between N2, N3, and N4. N1, which used two types of Zn and three types of Mg containing chemicals, was the only treatment that showed an improvement in the chemical profile of the peel oil (**Table 3** and **Figure 3**).

Peel oil quality is determined by the proper chemical combination of volatile compounds. Several low−abundant sesquiterpenes, such as valencene, α−sinensal, and β−sinensal, stand out in citrus as important flavor and aroma compounds (Dugo and Mondello, 2011). The profile of terpenoid volatiles in various citrus species and their importance as aroma compounds have been studied in detail (Njoroge et al., 2003; Xu et al., 2017a; Lin et al., 2019), but much is still lacking in our understanding of the physiological, biochemical, and genetic regulation of their production. Proteomic approaches, such as SWATH-MS, could be used to facilitate functional analysis in plant research (Zhu et al., 2020). Proteogenomics combines proteomics, genomics, and transcriptomics and has considerably improved genome annotation in poorly investigated phylogenetic groups for which homology information is lacking and may be a fruitful approach for elucidating the host–pathogen relationship between citrus and HLB with respect to the biosynthesis of volatile components of peel oil (Blank-Landeshammer et al., 2019; Chen et al., 2020).

CONCLUSION

In this research, the influence of HLB disease and foliar spray programs on the chemical composition of "Valencia" orange peel oil was investigated. HLB disease altered the volatile profile of "Valencia" orange peel oil in that many terpene compounds were accumulated at a higher level in the HLB+ peel oil, indicating that disease stress up-regulated the terpenoid pathways. In contrast, some key aldehydes in peel oil were suppressed in the HLB+ samples, which may negatively impact peel oil quality. Of the four proprietary foliar spray programs tested in this research, only N1 shifted the chemical profile of HLB+ peel oil to resemble that of HLB− samples.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

EAB, MI, AP, and JB: conceptualization. JB, MI, WZ, XS, EAB, and AP: methodology. HY, WZ, EB, and XS: conducting experiments. XS, WZ, and JB: statistical analysis of the results. XS, JB, EAB, MI, and HY: writing—original draft preparation. XS, JB, EAB, WZ, AP, MI, HY, and EB: writing—review and editing. All authors approved the submission.

REFERENCES

Ambrosio, C. M. S., Contreras-Castillo, C. J., and Da Gloria, E. M. (2020). In vitro mechanism of antibacterial action of a citrus essential oil on an enterotoxigenic *Escherichia coli* and *Lactobacillus rhamnosus*. *J. Appl. Microbiol.* 129, 541–553. doi: 10.1111/jam.14660

Bai, J., Baldwin, E., Hearn, J., Driggers, R., and Stover, E. (2014). Volatile profile comparison of USDA sweet orange-like hybrids versus 'Hamlin' and 'Ambersweet'. *HortScience* 49, 1262–1267. doi: 10.21273/HORTSCI.49.10.1262

Bai, J., Hagenmaier, R., and Baldwin, E. (2002). Volatile response of four apple varieties with different coatings during marketing at room temperature. *J. Agricult. Food Chem.* 50, 7660–7668. doi: 10.1021/jf020543n

Bai, J., Manthey, J. A., Ford, B. L., Luzio, G., Cameron, R. G., Narciso, J., et al. (2013). Effect of extraction, pasteurization and cold storage on flavonoids and other secondary metabolites in fresh orange juice. *J. Sci. Food Agric.* 93, 2771–2781. doi: 10.1002/jsfa.6097

Baldwin, E., Plotto, A., Bai, J., Manthey, J., Zhao, W., Raithore, S., et al. (2018). Effect of abscission zone formation on orange (*Citrus sinensis*) fruit/juice quality for trees affected by huanglongbing (HLB). *J. Agric. Food Chem.* 66, 2877–2890. doi: 10.1021/acs.jafc.7b05635

Baldwin, E., Plotto, A., Manthey, J., Mccollum, G., Bai, J. H., Irey, M., et al. (2010). Effect of liberibacter infection (Huanglongbing Disease) of citrus on orange fruit physiology and fruit/fruit juice quality: chemical and physical analyses. *J. Agric. Food Chem.* 58, 1247–1262. doi: 10.1021/jf9031958

Baldwin, E. A., Bai, J., Plotto, A., and Ritenour, M. A. (2014). Citrus fruit quality assessment; producer and consumer perspectives. *Stewart Postharvest Rev.* 2:7.

Blank-Landeshammer, B., Teichert, I., Marker, R., Nowrousian, M., Kuck, U., and Sickmann, A. (2019). Combination of proteogenomics with peptide de novo

sequencing identifies new genes and hidden posttranscriptional modifications. *mBio* 10:e02367-19. doi: 10.1128/mBio.02367-19

Bové, J. M. (2006). Huanglongbing: a destructive, newly-emerging, century-old disease of citrus. *J. Plant Pathol.* 88, 7–37.

Chen, H., Mccollum, G., Baldwin, E., and Bai, J. (2016). Impacts of huanglongbing symptom severity on fruit detachment force and mechanical properties of sweet oranges (*Citrus sinensis*). *HortScience* 51, 356–361. doi: 10.21273/HORTSCI.51.4.356

Chen, M. X., Zhu, F. Y., Gao, B., Ma, K. L., Zhang, Y., Fernie, A. R., et al. (2020). Full-length transcript-based proteogenomics of rice improves its genome and proteome annotation. *Plant Physiol.* 182, 1510–1526. doi: 10.1104/pp.19.00430

Dala-Paula, B. M., Plotto, A., Bai, J., Manthey, J. A., Baldwin, E. A., Ferrarezi, R. S., et al. (2018). Effect of Huanglongbing or greening disease on orange juice quality, a review. *Front. Plant Sci.* 9:1976. doi: 10.3389/fpls.2018.01976

Deng, G., Craft, J. D., Steinberg, K. M., Li, P. L., Pokharel, S. K., and Setzer, W. N. (2017). Influence of different isolation methods on chemical composition and bioactivities of the fruit peel oil of *Citrus medica* L. var. sarcodactylis (Noot.) Swingle. *Medicines (Basel)* 4:1. doi: 10.3390/medicines4010001

Dharmawan, J., Kasapis, S., Sriramula, P., Lear, M. J., and Curran, P. (2009). Evaluation of aroma-active compounds in Pontianak orange peel oil (*Citrus nobilis* Lour. Var. microcarpa Hassk.) by gas chromatography-olfactometry, aroma reconstitution, and omission test. *J. Agric. Food Chem.* 57, 239–244. doi: 10.1021/jf801070r

Dugo, G., and Mondello, L. (2011). *Citrus Oil: Composition, Advanced Analytical Techniques, Contaminants, and Biological Activity*. Boca Raton, FL: CRC Press. doi: 10.1201/b10314

Fancello, F., Petretto, G. L., Marceddu, S., Venditti, T., Pintore, G., Zara, G., et al. (2020). Antimicrobial activity of gaseous Citrus limon var pompia leaf essential oil against Listeria monocytogenes on ricotta salata cheese. *Food Microbiol.* 87:103386. doi: 10.1016/j.fm.2019.103386

Gonzalez-Mas, M. C., Rambla, J. L., Lopez-Gresa, M. P., Blazquez, M. A., and Granell, A. (2019). Volatile compounds in citrus essential oils: a comprehensive review. *Front. Plant Sci.* 10:12. doi: 10.3389/fpls.2019.00012

Gottwald, T., Poole, G., Mccollum, T., Hall, D., Hartung, J., Bai, J., et al. (2020). Canine olfactory detection of a vectored phytobacterial pathogen, *Liberibacter asiaticus*, and integration with disease control. *Proc. Natl. Acad. Sci. U.S.A.* 117, 3492–3501. doi: 10.1073/pnas.1914296117

Gottwald, T. R., Graham, J. H., Irey, M. S., Mccollum, T. G., and Wood, B. W. (2012). Inconsequential effect of nutritional treatments on huanglongbing control, fruit quality, bacterial titer and disease progress. *Crop Protection* 36, 73–82. doi: 10.1016/j.cropro.2012.01.004

Hodges, A. W., and Spreen, T. H. (2012). *Economic Impacts of Citrus Greening (HLB) in Florida, 2006/07Ű2010/11. EDIS 2012*. Available online at: https://journals.flvc.org/edis/article/view/119504

Hou, H. S., Bonku, E. M., Zhai, R., Zeng, R., Hou, Y. L., Yang, Z. H., et al. (2019). Extraction of essential oil from Citrus reticulate Blanco peel and its antibacterial activity against Cutibacterium acnes (formerly Propionibacterium acnes). *Heliyon* 5:e02947. doi: 10.1016/j.heliyon.2019.e02947

Huang, M., Valim, M. F., Shi, F., Reuss, L., Yao, L., Gmitter, F., et al. (2017). Characterization of the major aroma-active compounds in peel oil of an HLB-tolerant mandarin hybrid using aroma extraction dilution analysis and gas chromatography-mass spectrometry/olfactometry. *Chemosensory Perception* 10, 161–169. doi: 10.1007/s12078-017-9221-y

Kiefl, J., Kohlenberg, B., Hartmann, A., Obst, K., Paetz, S., Krammer, G., et al. (2018). Investigation on key molecules of huanglongbing (HLB)-induced orange juice off-flavor. *J. Agric. Food Chem.* 66, 2370–2377. doi: 10.1021/acs.jafc.7b00892

Li, J., Trivedi, P., and Wang, N. (2016). Field evaluation of plant defense inducers for the control of citrus huanglongbing. *Phytopathology§* 106, 37–46. doi: 10.1094/PHYTO-08-15-0196-R

Li, W., Hartung, J. S., and Levy, L. (2006). Quantitative real-time PCR for detection and identification of Candidatus Liberibacter species associated with citrus huanglongbing. *J. Microbiol. Methods* 66, 104–115. doi: 10.1016/j.mimet.2005.10.018

Lin, L. Y., Chuang, C. H., Chen, H. C., and Yang, K. M. (2019). Lime (*Citrus aurantifolia* (Christm.) Swingle) essential oils: volatile compounds, antioxidant capacity, and hypolipidemic effect. *Foods* 8:398. doi: 10.3390/foods8090398

Liu, C., Cheng, Y., Zhang, H., Deng, X., Chen, F., and Xu, J. (2012). Volatile constituents of wild citrus Mangshanyegan (*Citrus nobilis* Lauriro) peel oil. *J. Agric. Food Chem.* 60, 2617–2628. doi: 10.1021/jf2039197

Lota, M. L., De Rocca Serra, D., Tomi, F., Jacquemond, C., and Casanova, J. (2002). Volatile components of peel and leaf oils of lemon and lime species. *J. Agric. Food Chem.* 50, 796–805. doi: 10.1021/jf010924l

Masaoka, Y., Pustika, A., Subandiyah, S., Okada, A., Hanundin, E., Purwanto, B., et al. (2011). Lower concentrations of microelements in leaves of citrus infected with 'Candidatus Liberibacter asiaticus'. *Jarq-Japan Agric. Res. Q.* 45, 269–275. doi: 10.6090/jarq.45.269

McCollum, G., Baldwin, E., Gradziel, T. M., Mitchell, C. A., and Whipkey, A. L. (2016). Huanglongbing: devastating disease of citrus. *Horticult. Rev.* 44, 315–361. doi: 10.1002/9781119281269.ch7

Minh Tu, N. T., Onishi, Y., Choi, H. S., Kondo, Y., Bassore, S. M., Ukeda, H., et al. (2002). Characteristic odor components of *Citrus sphaerocarpa* Tanaka (Kabosu) cold-pressed peel oil. *J. Agric. Food Chem.* 50, 2908–2913. doi: 10.1021/jf011578a

Morgan, J. K., Zhou, L., Li, W., Shatters, R. G., Keremane, M., and Duan, Y. P. (2012). Improved real-time PCR detection of 'Candidatus Liberibacter asiaticus' from citrus and psyllid hosts by targeting the intragenic tandem-repeats of its prophage genes. *Mol. Cell. Probes* 26, 90–98. doi: 10.1016/j.mcp.2011.12.001

Njoroge, S. M., Ukeda, H., and Sawamura, M. (2003). Changes of the volatile profile and artifact formation in Daidai (*Citrus aurantium*) cold-pressed peel oil on storage. *J. Agric. Food Chem.* 51, 4029–4035. doi: 10.1021/jf021215q

Nwugo, C. C., Lin, H., Duan, Y., and Civerolo, E. L. (2013). The effect of 'Candidatus Liberibacter asiaticus' infection on the proteomic profiles and nutritional status of pre-symptomatic and symptomatic grapefruit (*Citrus paradisi*) plants. *BMC Plant Biol.* 13:59. doi: 10.1186/1471-2229-13-59

Oikeh, E. I., Ayevbuomwan, M., Irabor, F., Oikeh, A. O., Oviasogie, F. E., and Omoregie, E. S. (2020). Evaluation of the phenolic content, antioxidant and antimicrobial activities of oil and non-oil extracts of *Citrus sinensis* (L.) Osbeck Seeds. *Prev. Nutr. Food Sci.* 25, 280–285. doi: 10.3746/pnf.2020.25.3.280

Oyedeji, A. O., Okunowo, W. O., Osuntoki, A. A., Olabode, T. B., and Ayo-Folorunso, F. (2020). Insecticidal and biochemical activity of essential oil from *Citrus sinensis* peel and constituents on *Callosobruchus maculatus* and *Sitophilus zeamais*. *Pestic. Biochem. Physiol.* 168:104643. doi: 10.1016/j.pestbp.2020.104643

Plotto, A., Margaria, C. A., Goodner, K. L., and Baldwin, E. A. (2008). Odour and flavour thresholds for key aroma components in an orange juice matrix: esters and miscellaneous compounds. *Flavour Fragrance J.* 23, 398–406. doi: 10.1002/ffj.1888

Rouse, B., Irey, M., Gast, T., Boyd, M., and Willis, T. (2012). Fruit production in a southwest Florida citrus grove using the Boyd Nutrient/SAR foliar spray. *Proc. Florida State Hortic. Soc.* 125, 61–64.

Sharon-Asa, L., Shalit, M., Frydman, A., Bar, E., Holland, D., Or, E., et al. (2003). Citrus fruit flavor and aroma biosynthesis: isolation, functional characterization, and developmental regulation of Cstps1, a key gene in the production of the sesquiterpene aroma compound valencene. *Plant J.* 36, 664–674. doi: 10.1046/j.1365-313X.2003.01910.x

Song, X., Liu, T., Wang, L., Liu, L., Li, X., and Wu, X. (2020). Antibacterial effects and mechanism of mandarin (*Citrus reticulata* L.) essential oil against *Staphylococcus aureus*. *Molecules* 25:4956. doi: 10.3390/molecules25214956

Spadaro, F., Costa, R., Circosta, C., and Occhiuto, F. (2012). Volatile composition and biological activity of key lime *Citrus aurantifolia* essential oil. *Nat. Prod. Commun.* 7, 1523–1526. doi: 10.1177/1934578X1200701128

Spann, T. M., and Schumann, A. W. (2009). The role of plant nutrients in disease development with emphasis on citrus and huanglongbing. *Proc. Fla. State Hort. Soc.* 122, 169–171.

Stansly, P. A., Arevalo, H. A., Qureshi, J. A., Jones, M. M., Hendricks, K., Roberts, P. D., et al. (2014). Vector control and foliar nutrition to maintain economic sustainability of bearing citrus in Florida groves affected by huanglongbing. *Pest Manage. Sci.* 70, 415–426. doi: 10.1002/ps.3577

Vekiari, S. A., Protopapadakis, E. E., Papadopoulou, P., Papanicolaou, D., Panou, C., and Vamvakias, M. (2002). Composition and seasonal variation of the essential oil from leaves and peel of a cretan lemon variety. *J. Agric. Food Chem.* 50, 147–153. doi: 10.1021/jf001369a

Xu, B. M., Baker, G. L., Sarnoski, P. J., and Goodrich-Schneider, R. M. (2017a). A comparison of the volatile components of cold pressed Hamlin and Valencia (*Citrus sinensis* (L.) Osbeck) orange oils affected by huanglongbing. *J. Food Quality* 2017, 1–20. doi: 10.1155/2017/6793986

Xu, B. M., Sims, C. A., Etxeberria, E., and Schneider, R. M. G. (2017b). Physicochemical and sensory properties of cold pressed oils from florida hamlin and valencia oranges affected by huanglongbing. *J. Food Sci.* 82, 2158–2166. doi: 10.1111/1750-3841.13814

Yu, Q., Huang, M., Jia, H., Yu, Y., Plotto, A., Baldwin, E. A., et al. (2019). Deficiency of valencene in mandarin hybrids is associated with a deletion in the promoter region of the valencene synthase gene. *BMC Plant Biol.* 19:101. doi: 10.1186/s12870-019-1701-6

Zhang, M. Q., Guo, Y., Powell, C. A., Doud, M. S., Yang, C. Y., Zhou, H., et al. (2016). Zinc treatment increases the titre of 'Candidatus Liberibacter asiaticus' in huanglongbing-affected citrus plants while affecting the bacterial microbiomes. *J. Appl. Microbiol.* 120, 1616–1628. doi: 10.1111/jam.13102

Zhao, W., Bai, J., Mccollum, G., and Baldwin, E. (2015). High incidence of preharvest colonization of huanglongbing-symptomatic citrus sinensis fruit by *Lasiodiplodia theobromae* (*Diplodia natalensis*) and exacerbation of postharvest fruit decay by that fungus. *Appl. Environ. Microbiol.* 81, 364–372. doi: 10.1128/AEM.02972-14

Zhao, W., Baldwin, E. A., Bai, J., Plotto, A., and Irey, M. S. (2018). *Method for Assessing Juice/Cider Quality and/or Safety. Google Patents*.

Zhao, W., Gottwald, T., Bai, J., Mccollum, G., Irey, M., Plotto, A., et al. (2016). Correlation of Diplodia (*Lasiodiplodia theobromae*) infection, huanglongbing, ethylene production, fruit removal force and pre-harvest fruit drop. *Sci. Hortic.* 212, 162–170. doi: 10.1016/j.scienta.2016.09.032

Zhu, F. Y., Song, Y. C., Zhang, K. L., Chen, X., and Chen, M. X. (2020). Quantifying plant dynamic proteomes by SWATH-based mass spectrometry. *Trends Plant Sci.* 25, 1171–1172. doi: 10.1016/j.tplants.2020.07.014

Zouaghi, G., Najar, A., Aydi, A., Claumann, C. A., Zibetti, A. W., Ben Mahmoud, K., et al. (2019). Essential oil components of Citrus cultivar 'MALTAISE DEMI SANGUINE'(*Citrus sinensis*) as affected by the effects of rootstocks and viroid infection. *Int. J. Food Properties* 22, 438–448. doi: 10.1080/10942912.2019.1588296

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 07 May 2021
doi: 10.3389/fpls.2021.666805

Dynamic Analyses of Transcriptome and Metabolic Profiling: Revealing Molecular Insight of Aroma Synthesis of Mango (*Mangifera indica* L. Var. Tainong)

Ming Xin[1†], Changbao Li[1†], Hock Eng Khoo[2], Li Li[1], Xuemei He[1], Ping Yi[3], Yayuan Tang[3] and Jian Sun[1,3*]*

[1] Agro-food Science and Technology Research Institute, Guangxi Academy of Agricultural Sciences, Nanning, China, [2] College of Chemistry and Bioengineering, Guilin University of Technology, Guilin, China, [3] Guangxi Key Laboratory of Fruits and Vegetables Storage-processing Technology, Guangxi Academy of Agricultural Sciences, Nanning, China

OPEN ACCESS

Edited by:
Jian Li,
Beijing Technology and Business University, China

Reviewed by:
Jinhua Zuo,
Beijing Vegetable Research Center, China
Wei Shan,
South China Agricultural University, China
Luo Cong,
Guangxi University, China
Libin Wang,
Nanjing Agricultural University, China

***Correspondence:**
Changbao Li
changbaoli@gxaas.net
Jian Sun
jiansun@gxaas.net

[†]These authors share first authorship

Specialty section:
This article was submitted to Crop and Product Physiology, a section of the journal Frontiers in Plant Science

Received: 11 February 2021
Accepted: 09 April 2021
Published: 07 May 2021

Citation:
Xin M, Li C, Khoo HE, Li L, He X, Yi P, Tang Y and Sun J (2021) Dynamic Analyses of Transcriptome and Metabolic Profiling: Revealing Molecular Insight of Aroma Synthesis of Mango (Mangifera indica L. Var. Tainong). Front. Plant Sci. 12:666805. doi: 10.3389/fpls.2021.666805

This study aimed to evaluate the changes in aromatic components and other chemical properties of Tainong mango during fruit development, ripening, and storage. As the volatiles of Tainong mango and their related molecular mechanisms remain unclear, volatile profile, metabonomics, and transcriptome analyses were applied to investigate the molecular determinants of the synthesis of aroma components in mango during fruit development and storage. Total acids, total sugar, total carotenoids, enzyme activities of the mango pulp samples were also determined. Volatile components of the mango pulp samples were identified using a gas chromatography-mass spectrometric method. Ribonucleic acid (RNA) sequences of the samples were analyzed by real-time polymerase chain reaction. The results showed that 181 volatiles were isolated and identified in the fruit at seven stages. Compared to the other stages, mango collected on day 8 and day 12 had higher concentrations of 17 volatile components, especially (*E*,*Z*)-2,6-nonadienal, 53384 transcripts were also detected through RNA sequencing. The differentially expressed genes analyses included catalytic activity, transferase activity, adenosine diphosphate binding, transcription factor activity, and oxidoreductase activity. α-Pinene content and expression of the differentially expressed genes involved in terpenoid metabolism and enzyme activities in the terpenoid metabolic pathways gradually increased during the maturity of the fruit, and had maximum values at day 8 of storage. Moreover, the integrative analyses revealed potential molecular insights of mango development and aroma formation in the fruit.

Keywords: aromatic compound, de novo transcriptome assembly, gene ontology, RT-PCR, volatile profile

INTRODUCTION

Mango (*Mangifera indica* L.), the king of tropical fruit, is native to South Asia (Munafo et al., 2014). It is particularly rich in β-carotene. The compound is a precursor of vitamin A, which is rare in many fruits. Mango pulp is a typical source for the productions of fruit jam, canned food, pickled, sour, spicy pickles, and beverages, and more, besides being suitable for immediate consumption. It also exhibits tempting fragrance derived from the volatile components (San et al., 2017). The

quality and acceptability of mango pulp and juice are mainly assessed based on the flavor (Zhang et al., 2019a). The volatile profile varies considerably among different mango cultivars.

To date, a broad scope of investigations on mango volatiles has been implemented, resulting in the isolating and identification of more than 400 compounds in different varieties of mango, most of which are esters, ketones, aldehydes, alcohols, and terpenes (Andrade et al., 2000; Pino and Mesa, 2006; Shivashankara et al., 2006). Literature has shown that a relatively high amount of α-terpinolene has been determined in Kensington Pride, Chana, Bacuri, Coquinho, Gojoba, Cametá, Cheiro, Comum, and Carlota mangoes (Andrade et al., 2000; Lalel et al., 2003). It is one of the main volatiles in these fruits. A previous study also reported that 4-methoxy-2,5-dimethyl-3(2H)-furanone (MDMF), β-damascenone, ethyl butanoate, ethyl hexanoate, (*R*)-linalool, (*E*,*Z*)-2,6-nonadienal, and terpinolene are the essential aromatic compounds determined in yellow Thai Keaw mangoes (Boonbumrung et al., 2001). Hence, ethyl butanoate is one of the important aromatic compounds found in different cultivars of mango grown in Brazil (Tommy Atkins, Coracao de Boi, Carlota, Rubi, Haden, and Espada). Moreover, (*E*)-2-decenal, hexanal, (*E*)-β-ocimene, (*E*)-2-hexenal, γ-terpinene and (*Z*)-3-hexenal have been detected as the main aromatic components in Thai Khieo Sawoei mangoes (Tamura et al., 2001).

Based on the odor activity values, which represents the ratio of concentration to odor threshold, MDMF, ethyl butanoate, methyl benzoate, ethyl 2-methylpropanoate, decanal, (*E*,*Z*)-2,6-nonadienal, (*E*)-β-ionone, and (*E*)-2-nonenal have been potentially identified as the most important aroma compounds in the 20 cultivars of mango (Pino and Mesa, 2006). Literature also shows that terpinolene, (*E*,*Z*)-2,6-nonadienal, (*E*)-β-ionone, β-damascenone, (*E*)-2-nonenal, ethyl butanoate, and ethyl 2-methylpropanoate are the odor-active compounds of Corazon mangoes (Pino, 2012). In addition to the small number of these aromatic compounds identified in the mango pulp samples, many other volatiles could not be quantified due to their low detection limits (Zhang et al., 2019a).

The maturity of mango had greatly affected the aroma profile of the fruit besides the geographical origin, which was related to the growing environment and the variation in cultivar (Lebrun et al., 2008; Pandit et al., 2009b; Chauhan et al., 2010; Kulkarni et al., 2012). A previous study identified limonene and *p*-cymene, α-terpinene, and ethyl octanoate as the dominant volatile components in Kensington Pride mango samples collected during the pre-climacteric, climacteric stage, and fully ripe stages (Lalel et al., 2003). Although many volatile components have been isolated and identified in mangoes, studies on the molecular mechanism of aroma compound biosynthesis are limited. Acyl-CoA-oxidase (MiACO), 9-lipoxygenase (Mi9LOX), hydroperoxide lyase (MiHPL), peroxygenase (MiPGX1), and epoxide hydrolase 2 (MiEH2) genes involved in lactone biosynthesis have also been isolated from mango, and the transcript profiling of these genes was analyzed during various developmental stages in fruits of three mango cultivars (Kent, Pairi, and Alphonso) with different levels of lactones (Deshpande et al., 2017b). The transcriptome analysis has also been used to explore the distinct aroma characteristics in Alphonso mango, where transcripts for the biosynthesis of furanones, sesquiterpenes, lactones, monoterpenes, and diterpenes were identified.

Aroma is the most critical organoleptic quality of a mango. Volatile components give aroma to the fruit. These compounds are subjected to changes during fruit ripening and postharvesting. Although volatile components of fruits are widely studied, the associated molecular mechanisms remain unclear. It undergoes rapid and substantial changes during ripening and storage. Although some of the aromatic components have been isolated and identified in different cultivars of mango, little is known about the volatile profile in the "Tainong" variety of mango, which is widely grown in southern China. The ripe fruit of this variety has a special aroma because it contains certain aromatic compounds that give it a flavor. In this study, an integrative analysis of volatile profile and transcriptome was employed to identify molecular mechanism, as well as the changes in volatile components during the stages of fruit development and storage.

MATERIALS AND METHODS

Plant Materials and Growth Conditions

Tainong mango trees have been planted in the Tiandong National Mango Germplasm Resources Nursery in Guangxi (Tiandong County, Baise City, Guangxi Province, China, 23°16′N, 107°26′E) since 6 years ago. The ambient temperatures of the nursery ranged between 25 and 28°C, and the altitude is 110 m. Mango samples were collected from the nursery at different stages of fruit development (40, 60, 80, and 90 days after the flowering stage began). Fruit samples harvested at 90 days after the flowering were also kept for 4, 8, and 12 days in the laboratory at a controlled temperature of 25°C and a humidity level of 95%. Seven mangoes of an average weight of 100 ± 10 g were sampled. The fruits were picked from different trees at the same positions on the day of fruit collection. Mango peels were removed, and the pulps were directly used for volatile analysis. All pulp samples were stored at −80°C before extraction. Data of triplicate analyses were obtained for each analysis.

Analysis of Total Acid Content, Total Sugar Content, and Carotenoid Content

Total acid content, total sugar content, and carotenoid content of the mango pulp samples were measured according to the methods described by Shi et al. (2013); Liao et al. (2019), and Zhang et al. (2019a), respectively. The total acid content of mango pulp samples was determined by following the Official Methods of Analysis for Vinegar in China (GB/T 5009.41-2003), while total sugars and total carotenoids of the pulp were determined using HPLC.

Determination of Enzyme Activities

Enzyme activities of 1-deoxy-D-xylose-5-phosphate reductase (DXS), 1-deoxyxylose-5-phosphate synthase (DXR), geranyl pyrophosphate synthetase (GPPS), geranylgeranyl pyrophosphate synthetase (GGPPS), pyruvate carboxylase (PC),

diacylglyceryl transferase (DGAT), farnesyl diphosphate synthase (FPS), and hydroxymethyl glutarate monoacyl CoA reductase (HGMR) were measured using plant ELISA kits, where the all the ELISA kits were purchased from a local chemical supplier (Jianglai Biotechnology Co., Ltd, Shanghai, China). A microplate reader (BioTek Instruments, Inc., Winooski, VT, United States) was used to obtain the absorbance readings of each test; the analyses were performed according to the instructions provided by the manufacturers. The results were calculated based on the formula provided in the manufacturers' instructions.

Volatile Profile Analysis

Exactly 2.5 g of homogenized mango pulp was added with 2.5 mL of saturated sodium chloride and 100 μL of internal standard solution (32.88 μg/mL 2-octanol, Sigma-Aldrich) in a 20 mL headspace bottle (ANPEL Laboratory Technologies Inc., Shanghai, China). The bottle was sealed with a crew cap fitted with a PTFE/silicone septum. After a 15-min agitation at 50°C and 250 rpm, volatile compounds were extracted with 50/30 μm DVB/CAR/PDMS as the extraction fiber, which was later exposed to the headspace for 40 min.

The gas chromatography-mass spectrometry (GC-MS) analysis was performed using a GC-mass spectrometer (7890B-5977B, Agilent Technologies) to quantify the volatile components of mango pulp samples. Separation of the volatile compounds was performed using a DB-Wax column (30 m × 0.25 mm × 0.25 μm, Agilent Technologies, Shanghai, China). The extract was injected in splitless mode and desorbed at 260°C for 5 min. Helium was used as the carrier gas with a constant flow rate of 1 mL/min. The initial oven temperature was 40°C, and then programmed at 5°C/min to 220°C, followed by an increase of 20°C/min to 250°C, and finally held at 250°C for another 2.5 min. Electron ionization (EI+) was set at 70 eV, and the data were recorded in scan mode of m/z 20–400.

Based on the MS fragmentation patterns and linear retention indices, volatile compounds were identified and quantified through comparisons with the NIST14 library. The differential volatiles in each group were screened according to the following criteria: fold change $\geq$ 1.5 or fold change $\leq$ 0.67; variable importance in project (VIP) $\geq$ 1.

Real-Time PCR Analysis

The real-time quantitative PCR (qRT-PCR) was performed using the fluorescent intercalating dye SYBR Green in a detection system (MJ Research, Opticon 2), and MiACT was used as a standard control (Luo et al., 2013). The two-step RT-PCR procedure was performed according to the method described by Li et al. (2005).

Ribonucleic Acid Sequencing

RNAprep pure plant plus kit (TIANGEN Biotech Co., Ltd., Beijing China) was used to purify total ribonucleic acid (RNA) in mango pulp samples. The purification steps were done in accordance with the manufacturer's instructions. After a quality checked using NanoPhotometer® spectrophotometer and Agilent 2100 bioanalyzer, high-quality mRNA was enriched by poly-T oligo-attached magnetic beads. The library was constructed using the NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, United States). After the fragmentation of the purified mRNA by divalent cations at elevated temperature, the first-strand complementary DNA (cDNA) was generated using a random hexamer primer. The second strand of cDNA was generated using DNA polymerase I of M-MuLV reverse transcriptase (RNase H-) and RNase H (Sigma-Aldrich, Shanghai, China). After methylating the 3' ends of DNA fragments and ligating the adaptor for hybridization, the library fragments were purified using AMPure XP beads (Beckman Coulter, Beverly, United States). Under the action of high-fidelity DNA polymerase, Universal PCR primers, and Index (X) primer, PCR was performed, and the PCR products were purified with AMPure XP system. The list of primers used is showed in **Supplementary Table 1**. Then, match quality of the library was assessed using the Agilent 2100 bioanalyzer system (Waldbronn, Germany). The library sequencing was performed using the Illumina HiSeq 2500™ platform (PE125, paired-end).

De novo Transcriptome Assembly, Gene Expression, and Differential Expression Analysis

After the quality check and adaptor trimming, clean reads were assembled using the Trinity software, and the transcripts were generated. The reads contained unknown and over 50% low-quality nucleotides (Qphred $\leq$ 20) were removed. Quality of the transcripts was evaluated using the Benchmarking Universal Single-Copy Orthologs. The coding sequences (CDs) were predicted through comparisons with the NR and Swissprot protein libraries, and using ESTScan v3.0.3 software.

The clean reads were mapped to the transcripts using RSEM v1.1.17 software (Li and Dewey, 2011). The gene expression level was quantified using FPKM (fragments per kilobase of transcript per million fragments mapped) as an indicator. FPKM was calculated as follows:

$$\mathrm{FPKM} = \frac{\text{Mapped fragments of transcript}}{\text{Total count of mapped fragments (millions)} \times \text{Length of transcript (kb)}} \quad (1)$$

Differential expression analysis was done using the DEGSeq2, which was according to the criteria as follows: $\log_2$ (fold change) $\geq$ 1 and $p$adj $<$ 0.05.

Gene Ontology Enrichment Analysis

The gene ontology (GO) enrichment analysis of differentially expressed transcripts in mango pulp samples was performed using the KOBAS 2.0, GOseq, and GO database[1].

Correlation Analysis of Volatiles and Transcriptome Profile

The correlation analyses of differential volatiles and differential expression transcripts are achieved using the Pearson correlation analysis (SPSS version 15.0).

[1]http://www.geneontology.org/

RESULTS

Volatile Profiles in Mango at Different Development and Storage Stages

The mango morphology was observed. Total acid, total sugar, and carotenoid contents of the fruit pulp were determined at different stages of fruit development and storage. As shown in **Figure 1A**, the mangoes were immature and green in color at 40, 60, and 80 days after flowering (DAF). Days after postharvest storage, the peel turned yellow, especially after 4 days of storage. Moreover, the pulps of mango harvested after 40 and 60 days of flowering looked light-green in color. The inner mesocarp of mango started to turn yellow on day 80 (**Figure 1A**). During the postharvest storage period, especially at 8 and 12 days after picking (DAP), the mango pulp turned orange (**Figure 1A**). The color hues were determined using a colorimeter (data not shown).

As shown in **Figure 1B**, there is a gradual decline in total acid content during the fruit development. The total acid content started to reduce after 60 DAF. The total acid content of the pulp samples of 80 DAF dropped to almost half compared to the value determined for pulp samples of 60 DAF. A linear decrease in total acid content was found for the mango pulp samples during the 12 days of storage. Also, there was a gradual decline in total acid content in the developing fruits after 60 DAF (**Figure 1B**). The total acid content was higher than 5.6% at 40 and 60 DAF. However, the total acid content was 0.63% at 12 DAP (**Figure 1B**). Moreover, the content of total sugar and total carotenoid gradually increased with the fruits ripening. There was a higher increase in total acid content of the pulp samples of 8 DAP (**Figure 1B**).

Differential Analysis of Volatiles Among Different Development and Storage Stage

Volatile profiles of mango pulp samples were determined based on the headspace solid-phase microextraction method and GC-MS. The differential volatiles in mango pulp samples of different groups were screened according to the following criteria: fold change $\geq$ 1.5 or fold change $\leq$ 0.67; VIP$\geq$ 1. The numbers of significantly different volatiles in each group are shown in **Figure 2A**. The result showed that significant differences were found for the numbers of differential volatiles between the two groups of samples ($p < 0.05$). However, the highest significant difference in the numbers of differential volatiles was determined between 12 DAP and 0 DAP samples ($p < 0.01$).

The data determined for the 12 DAP and 0 DAP group included ten up-regulated volatiles and four down-regulated volatiles. As shown in **Figure 2B**, three up-regulated volatiles determined for the 60 DAF vs. 40 DAF group were *D*-limonene, (*E*)-2-hexenal, and hexanal; three up-regulated volatiles found in the 80 DAF vs. 40 DAF group were (*E*)-2-hexenal, hexanal, and 1,3-bis(1,1-dimethylethyl)-benzene; ten up-regulated volatiles determined in the 12 DAP and 0 DAP group were (2*E*)-nonenal, (2*E*)-(2-pentenyl)furan, geraniol, (*E*,*Z*)-2,6-nonadienal, ethanol, butanoic acid, *p*-mentha-1,5,8-triene, (*E*)-3,7-dimethyl-2,6-octadienal, (*Z*)-3,7-dimethyl-2,6-octadienal, and (6*Z*)-nonen-1-ol; seven up-regulated volatiles determined in the 8 DAP vs. 0 DAP group were (*E*)-butanoic acid 3,7-dimethyl-2,6-octadienyl ester, geraniol, linalool, ethanol, (*E*,*Z*)-2,6-nonadienal, α-pinene, and (*E*)-2-hexen-1-ol. These up-regulated volatiles might be the dominating aromatic components of Tainong mango.

As shown in Dataset 1, 181 volatiles were isolated and identified. The major volatiles detected in the ripen fruit samples are α-pinene, β-phellandrene, β-ocimene, *D*-limonene, β-myrcene, γ-terpinene, 2-carene, 3-carene, 4-carene, copaene hexanal, and (*E*)-2-hexenal (Dataset 1). In this study, the amounts of (*E*)-2-nonenal, (2*E*)-(2-pentenyl)furan, geraniol, (*E*,*Z*)-2,6-nonadienal, ethanol, butanoic acid, *p*-mentha-1,5,8-triene, (*E*)-3,7-dimethyl-2,6-octadienal, (*Z*)-3,7-dimethyl-2,6-octadienal, and (6*Z*)-nonen-1-ol in the mango pulp samples of 12 DAP were higher than that of the sample at baseline (0 DAP) (**Figure 2B**), whereas the amounts of (*E*)-3,7-dimethyl-2,6-octadienyl-butanoate, geraniol, linalool, ethanol, (*E*,*Z*)-2,6-nonadienal, α-pinene, and (*E*)-2-hexen-1-ol in the mango pulp samples of 8 DAP were higher than the sample at baseline.

RNA Sequencing Analysis of Mango During Different Development and Storage Stages

Molecular insight of volatile biosynthesis in mango pulp during the fruit development and storage can be obtained from RNA sequencing (RNA-Seq) analysis. Seven cDNA of the fruit samples, collected after 40, 60, 80, and 90 DAF, and at 4, 8, and 12 days of storage were constructed and large-scale sequenced. 47.29-79.61 million raw reads and 46.48-78.33 million cleaned reads were generated in each sample (**Table 1**). The Q20 (the percentage of bases with a Phred score greater than 20) and Q30 (the percentage of bases with a Phred score greater than 30) were higher than 95% (**Table 1**).

After assembled using the Trinity software, the transcripts of each sample were acquired. The frequency and numbers of transcripts and unigene in the corresponding length are shown in **Supplementary Figure 1A**. The CDs were predicted by comparing them with NR protein library, Swissprot protein library, and ESTScan v3.0.3 software. The counts of CDs with different lengths were also predicted based on the corresponding method (**Supplementary Figure 1B**).

Hierarchical clustering analysis of Pearson correlation according to the level of gene expression levels revealed that a high correlation was found for the gene expression in mango pulp samples between different development stages (**Supplementary Figure 1C**). Besides the three replicates of mango samples collected after 40 days of flowering, the Pearson correlation coefficients among the three repetitions of the other samples collected were higher than 0.95, which showed good reliability and repeatability of RNA-Seq data. Verification of

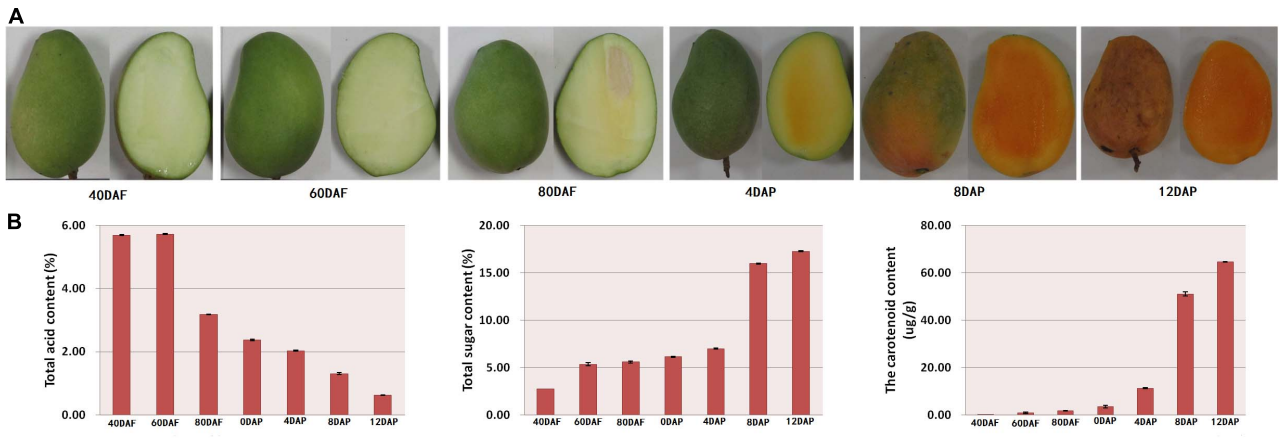

FIGURE 1 | (A) Morphology of mango at different developmental stages and **(B)** total acid content, total sugar content, and total carotenoid content of mango peel. Samples labeled 40 DAF, 60 DAF, and 80 DAF are the peel samples of mango harvested at 40, 60, 80, and 90 days after the flowering, respectively; samples labeled 0 DAP, 4 DAP, 8 DAP, and 12 DAP are the peel samples of mango harvested at 90 days after flowering which kept for 0, 4, 8, and 12 days, respectively. DAF, days after flowering; DAP, days after picking.

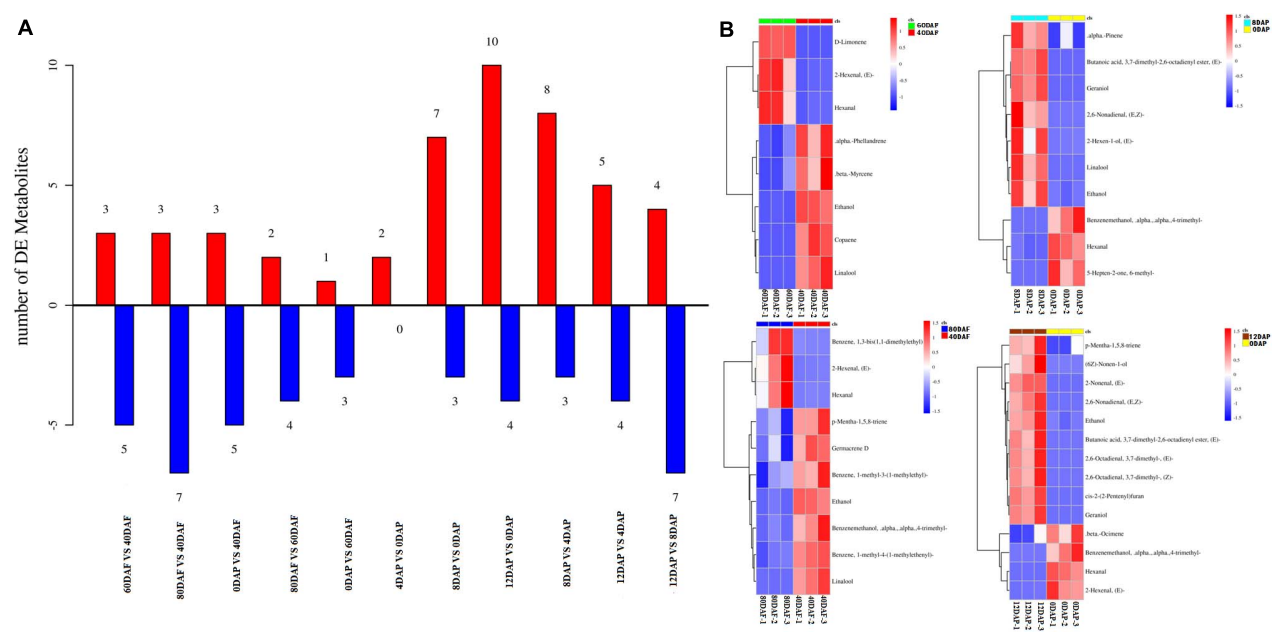

FIGURE 2 | (A) The numbers of volatiles (metabolites) in mango pulp samples at different developmental and ripening stages and **(B)** Hierarchical clustering analysis of the differential volatiles in the samples. Samples labeled 40 DAF, 60 DAF, and 80 DAF are the peel samples of mango harvested at 40, 60, and 80 days after the flowering, respectively; samples labeled 0 DAP, 4 DAP, 8 DAP, and 12 DAP are the peel samples of mango harvested at 90 days after flowering which kept for 0, 4, 8, and 12 days, respectively. DAF, days after flowering; DAP, days after picking.

the RNA-Seq data was done based on the nine selected genes (**Figure 3** and Dataset 2).

Differentially Expressed Transcripts and Gene Ontology Enrichment Analysis

The differentially expressed transcripts in the groups of 60 DAF vs. 40 DAF, 80 DAF vs. 40 DAF, 90 DAF vs. 40 DAF, 80 DAF vs. 60 DAF, 90 DAF vs. 60 DAF, 90 DAF vs. 80 DAF, 4 DAP vs. 0 DAP, 8 DAP vs. 0 DAP, 12 DAP vs. 0 DAP, 8 DAP vs. 4 DAP, 12 DAP vs. 4 DAP, and 12 DAP vs. 8 DAP were identified based on the DEGSeq analysis for obtaining an overview of the interesting genes. The numbers of differentially expressed transcripts in each group are shown in **Figure 4A** and Dataset 2. The results showed that the most differentially expressed transcripts were found in the group of 12 DAP vs. 0 DAP (including 6,520 up-regulated transcripts and 7,400 down-regulated transcripts), where the data are in line with the results obtained from the determination of

TABLE 1 | Reads numbers, Q20 values, Q30 values, and gas chromatography (GC) content of mango pulp based on the RNA sequencing (RNA-Seq) data in the libraries of A, B, C, D, E, F, and G.

Sample	Raw reads	Clean reads	Q20 (%)	Q30 (%)	GC content (%)	Total mapped reads
A_1	54542030	53752304	98.71	95.68	45.04	41320416 (76.87%)
A_2	79607262	78327428	98.68	95.59	44.67	60458266 (77.19%)
A_3	66307584	65005808	98.69	95.6	44.67	50076660 (77.03%)
B_1	61825070	60940514	98.55	95.17	44.91	47642222 (78.18%)
B_2	61776038	60874120	98.71	95.66	44.81	47623800 (78.23%)
B_3	62783660	61896096	98.72	95.65	44.75	48219096 (77.90%)
C_1	65584430	64443166	98.74	95.73	44.65	50675552 (78.64%)
C_2	66911796	65761656	98.67	95.52	44.68	51653690 (78.55%)
C_3	70930988	69608456	98.58	95.29	44.2	54910004 (78.88%)
D_1	64323328	63325728	98.66	95.46	44.22	48845080 (77.13%)
D_2	53106480	52366662	98.73	95.72	44.51	40206360 (76.78%)
D_3	61134368	60176382	98.66	95.47	44.35	46130222 (76.66%)
E_1	64246102	63464652	98.69	95.53	44.45	49456492 (77.93%)
294E_2	47298170	46481888	98.68	95.48	44.36	36320930 (78.14%)
E_3	58093350	57090472	98.75	95.69	44.49	44561666 (78.05%)
F_1	59865570	58791448	98.55	95.12	44.19	45797898 (77.90%)
F_2	67186956	66240678	98.75	95.71	44.1	51789204 (78.18%)
F_3	60845708	59572020	98.72	95.6	44.13	46554538 (78.15%)
G_1	73520470	72003528	98.78	95.73	43.98	56418766 (78.36%)
G_2	54733094	53300870	98.79	95.77	44.01	42180466 (79.14%)
G_3	68004602	67124008	98.65	95.35	43.96	-

A, 40 DAF; B, 60 DAF; C, 80 DAF; D, 0 DAP; E, 4 DAP; F, 8 DAP; G, 12 DAP. _1, _2, and _3 represent the triplicate analyses of each sample. Q20 is the percentage of bases with a Phred score greater than 20, whereas Q30 is the percentage of bases with a Phred score greater than 30.

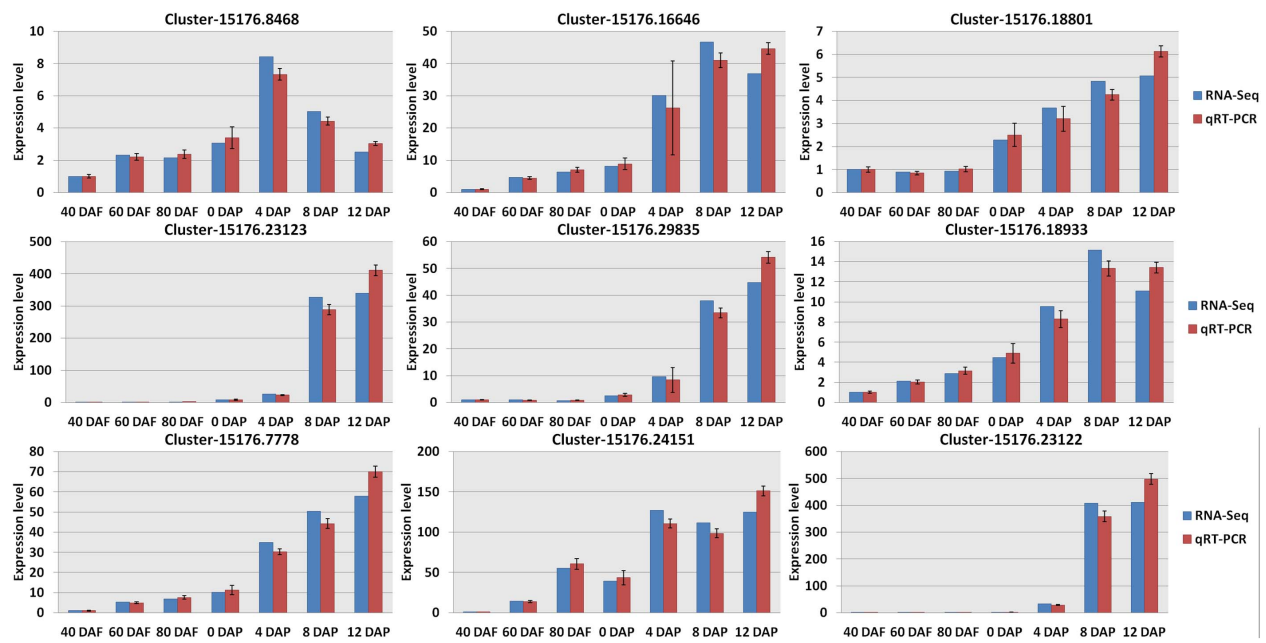

FIGURE 3 | Verification the RNA sequencing (RNA-Seq) results by real-time quantitative PCR (qRT-PCR). All data were presented as mean ± standard error of the mean of triplicate analyses. DAF, days after flowering; DAP, days after picking. Gene names (**Supplementary Table 2**): Cluster-15176.8468 (wrbA); Cluster-15176.23123 (LOX1_5); Cluster-15176.7778 (accA); Cluster-15176.16646 (4CL); Cluster-15176.29835 (SXD1); Cluster-15176.24151 (fabI); Cluster-15176.18801 (VTE1); Cluster-15176.18933 (TAT); and Cluster-15176.23122 (LOX1_5).

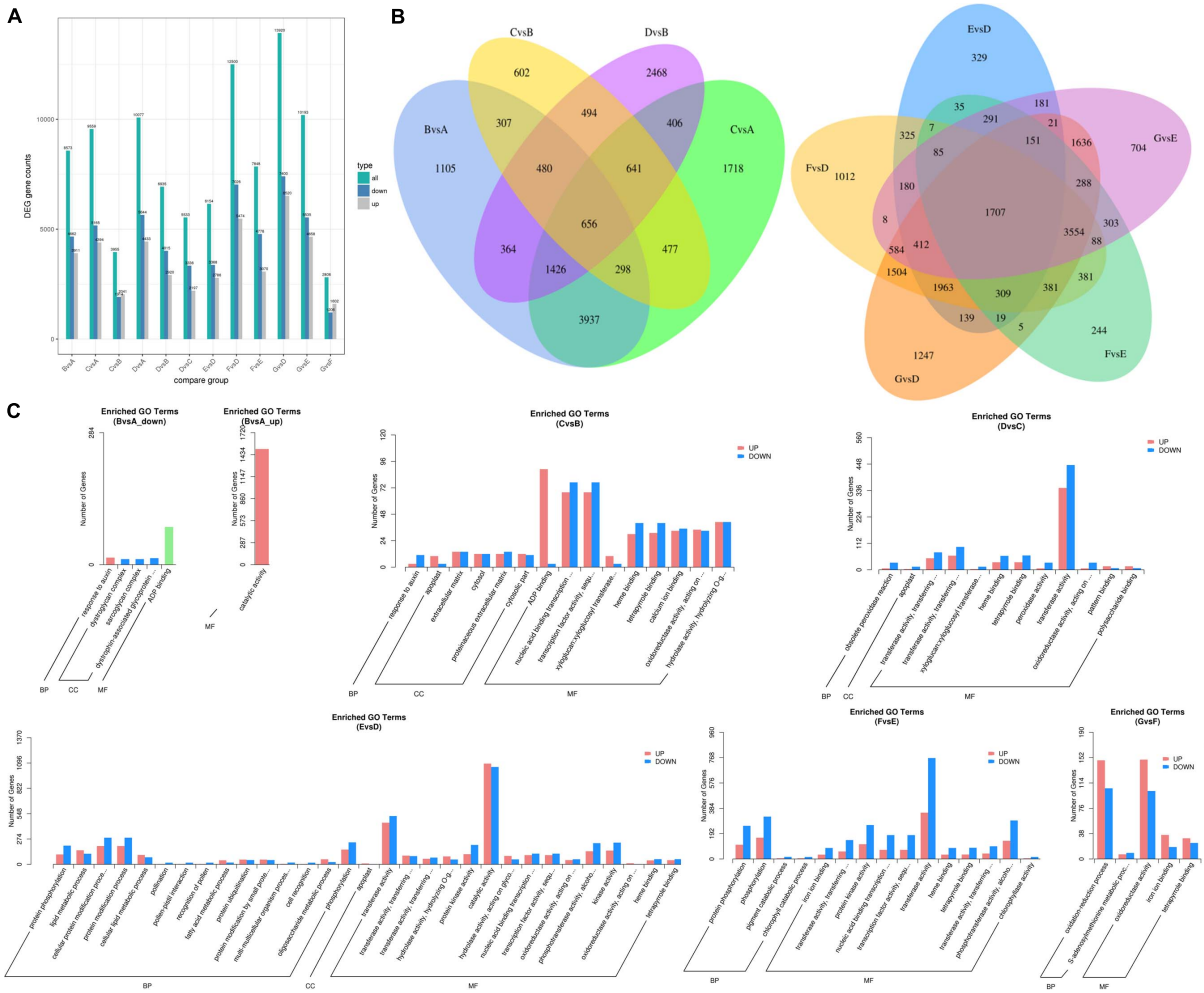

FIGURE 4 | Differentially expressed transcripts of mango pulp samples at different developmental and ripening stages. **(A)** Total numbers of the differentially expressed genes (DEGs), up-regulated genes, and down-regulated genes between different samples, B vs. A, C vs. A, D vs. A, C vs. B, D vs. B, D vs. C, E vs. D, F vs. D, G vs. D, G vs. E, F vs. E, and G vs. F; **(B)** Venn diagram of the DEGs between different samples, B vs. A, C vs. A, C vs. B, D vs. B (left) and E vs. D, F vs. D, G vs. D, G vs. E, F vs. E (right); **(C)** GO classification of the up-regulated genes and down-regulated genes between different samples, B vs. A, C vs. B, D vs. C, E vs. D, F vs. E, and G vs. F; A–C are the samples collected after 40, 60, and 80 days of flowering, respectively; D–G are the samples stored for 0, 4, 8, and 12 days which collected after 90 days of flowering.

differential volatiles (**Figure 2A**). The Venn diagram also shows that 24 mutually differentially expressed transcripts existed in 60 DAF vs. 40 DAF, 80 DAF vs. 40 DAF, 80 DAF vs. 60 DAF, and 90 DAF vs. 60 DAF, whereas the number of commonly differentially expressed transcripts in 4 DAP vs. 0 DAP, 8 DAP vs. 0 DAP, 12 DAP vs. 0 DAP, 8 DAP vs. 4 DAP, and 12 DAP vs. 4 DAP (**Figure 4B**). These differentially expressed transcripts might be relevant to the mango development and formation of aromatic compounds.

The GO analysis was performed to show the comprehensiveness of functions of differentially expressed transcripts. As shown in **Figure 4C**, the gene functions are described based on the cellular components, molecular functions, and biological processes. During the fruit development, the most enriched GO terms for the groups of 60 DAF vs. 40 DAF and 90 DAF vs. 80 DAF were catalytic activity and transferase activity, respectively. Adenosine diphosphate (ADP) binding and transcription factor activity were the two most enriched GO terms for 80 DAF vs. 60 DAF. During storage of mango, the most enriched GO terms found were catalytic activity (4 DAP vs. 0 DAP), transferase activity (8 DAP vs. 4 DAP), and oxidoreductase activity (12 DAP vs. 8 DAP).

Correlation Analysis of Differential Volatiles and Differentially Expressed Transcripts

Correlation analysis was performed to explore the transcripts related to aroma biosynthesis during the development and storage of Tainong mango (**Figure 5**, Dataset 3). The results showed that 304 differentially expressed transcripts determined in the group of 8 DAP vs. 0 DAP were either positively or

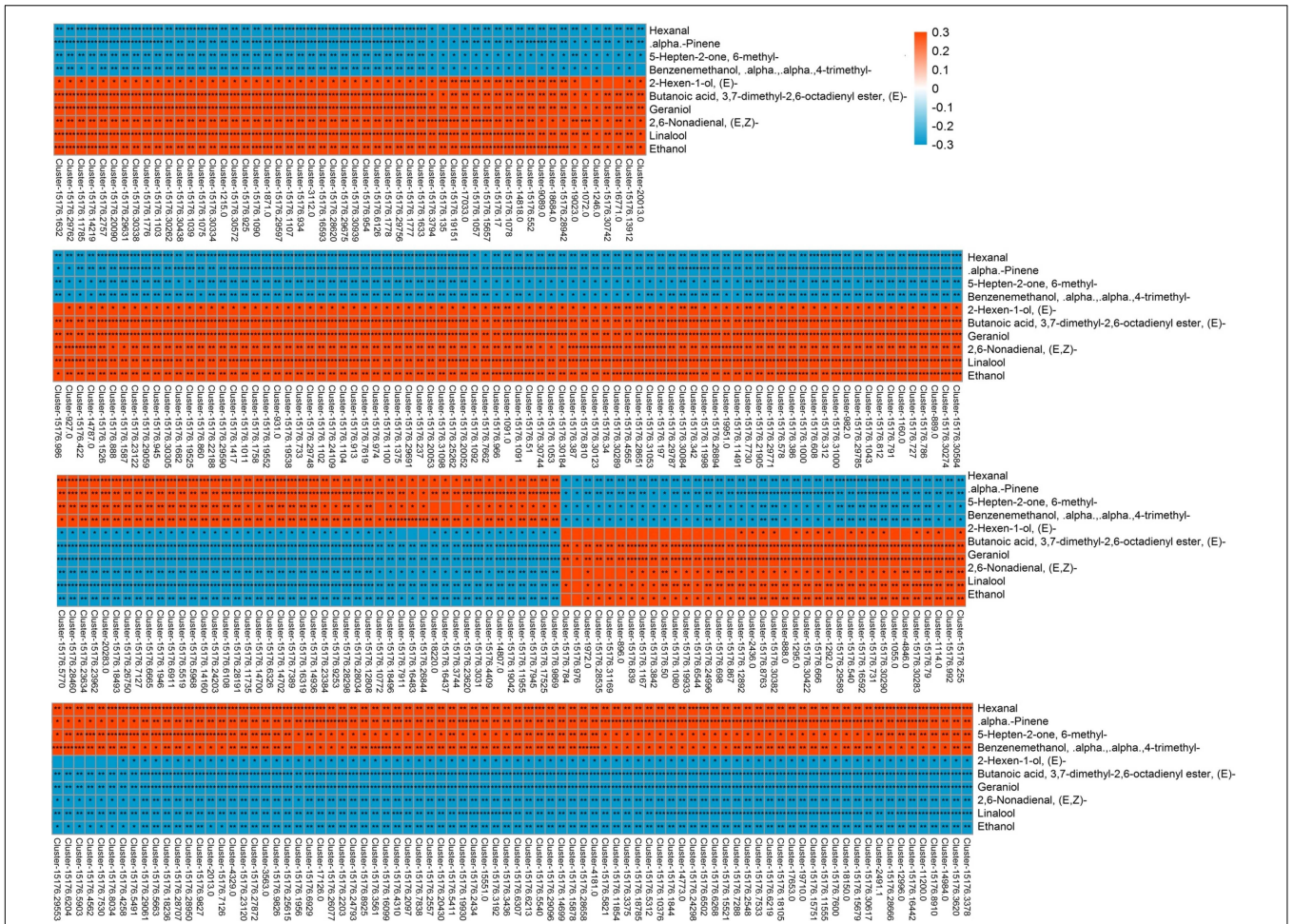

FIGURE 5 | Correlation analysis of transcriptome and volatile profiles of mango pulp samples from the 8 DAP vs. 0 DAP group. The X-coordinate is the transcript name, and the Y-coordinate is the metabolite name. Each square represents the correlation and significance of the transcript and the metabolite. The correlation coefficient is presented based on different colors and shades. Red represents a positive correlation and blue represents a negative correlation. The darker the color, the higher the correlation. The names of the identified compounds and their related unigene-id are tabulated in **Supplementary Table 1**.

negatively correlated with 3,7-dimethyl-2,6-octadienyl ester, (*E*)-butanoic acid, geraniol, hexanal, 6-methyl-5-hepten-2-one, linalool, ethanol, α,α,4-trimethyl-benzenemethanol, (*E,Z*)-2,6-nonadienal, α-pinene, and (2*E*)-hexen-1-ol; 420 differentially expressed transcripts determined in the group of 4 DAP vs. 0 DAP were either positively or negatively correlated with α,α,4-trimethyl-benzenemethanol and (*E,Z*)-2,6-nonadienal; 601 differentially expressed transcripts determined in the group of 12 DAP vs. 0 DAP were either positively or negatively correlated with (2*E*)-nonenal, hexanal, (2*E*)-(2-pentenyl)furan, geraniol, (2*E*)-hexenal, (*E,Z*)-2,6-nonadienal, ethanol, (*E*)-3,7-dimethyl-2,6-octadienal, (*Z*)-3,7-dimethyl-2,6-octadienal, butanoic acid, α,α,4-trimethyl-benzenemethanol, (6*Z*)-nonen-1-ol, p-mentha-1,5,8-triene, and β-ocimene; 277 differentially expressed transcripts determined in the group of 80 DAF vs. 60 DAF were either positively or negatively correlated with α,α,4-trimethyl-benzenemethanol, 1-(4-methylphenyl)- ethanone, 1-methyl-4-(1-methylethenyl)-benzene, p-mentha-1,5,8-triene, alloaromadendrene, and 3-hexenal; 381 differentially expressed transcripts determined in the group of 60 DAF vs. 40 DAF were either positively or negatively correlated with D-limonene, ethanol, copaene, linalool, α-phellandrene, (2*E*)-hexenal, hexanal, and β-myrcene (Dataset 3).

Genes and Enzymes Related to the Metabolism of Aromatic Compounds

As shown in **Figure 2B**, the amount of α-pinene in the fruit sample stored on day 8 was significantly higher than that on the other days ($p < 0.05$). It indicated that α-pinene is one of the key aroma components of mango. The results showed that a total of eight genes involved in the diterpenoid metabolism pathways (**Figure 6** and **Supplementary Table 3**). The transcriptome analysis manifested that the genes, such as Cluster-15176.332 (E5.5.1.13) and Cluster-15176.12278 (KAO), were expressed at the highest levels in the mango pulp samples of 8 DAP. The high expressions of Cluster-15176.3381 (E1.14.11.13), Cluster-15176.3380 (E1.14.11.13), Cluster-19253.0

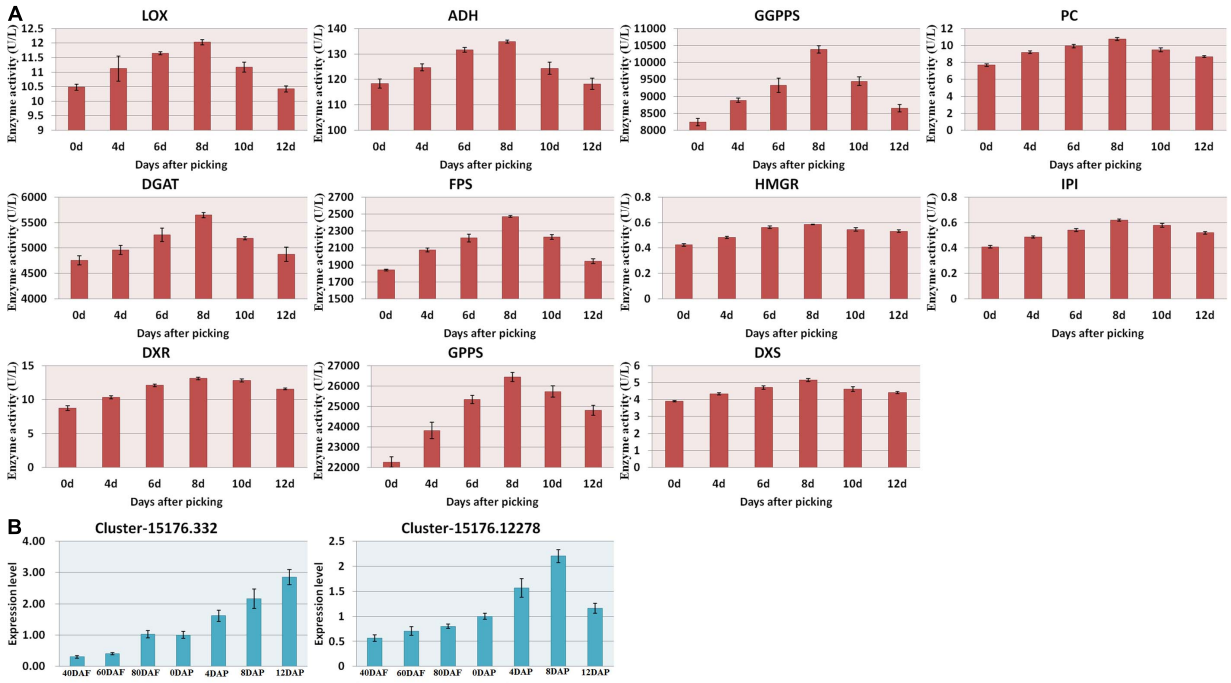

FIGURE 6 | (A) Enzyme activities of the mango pulp samples and **(B)** gene expressions of E5.5.1.13 (Cluster-15176.332) and KAO (Cluster-15176.12278) in the samples during the fruit development and ripening. The enzyme activities of 1-deoxyd-D-xylose-5-phosphate reductase (DXS), 1-deoxyxylose-5-phosphate synthase (DXR), geranyl pyrophosphate synthetase (GPPS), geranyl geranyl pyrophosphate synthetase (GGPPS), pyruvate carboxylase (PC), diacylglyceryltransferase (DGAT), farnesyl diphosphate synthase (FPS), and hydroxymethyl glutarate monoacyl CoA reductase (FPS) in the mango pulp samples stored at different days were determined using ELISA kits. All data are presented as means ± standard errors of the means. Triplicate analyses were performed for each sample. DAF, day after flowering; DAP, days after picking.

(GA3,CYP701), Cluster-15991.0 (E1.14.11.13), Cluster-3324.0 (E1.14.11.13), and Cluster-15176.1075 (CYP82G1) were also detected in mango pulp samples during the fruit development and ripening. These genes involved in diterpenoid biosynthesis. The qRT-PCR analysis also verified the highest expressions of E5.5.1.13 (ent-copalyl diphosphate synthase) and KAO (ent-kaurenoic acid hydroxylase) in the samples of 8 DAP compared with the other stages. These enzymes have been reported to regulate the biosynthesis of gibberellin (Su et al., 2016; Szymczyk et al., 2020). Only one gene was found to be involved in monoterpenoid biosynthesis. It was K15095 (Cluster-4351.0). The gene is related to the biosynthesis of nerolidol in the fruit. However, nerolidol was not detected in the fruit samples.

The gene names of the respective UniGene IDs are depicted in **Table 2**. Ent-kaurene oxidase, gibberellin 2-β-dioxygenase, and gibberellin 2-β-dioxygenase 8 isoform X3 were highly expressed during the fruit development and ripening. Only ent-kaurenoic acid oxidase 1-like was highly expressed in the matured fruit, and it was the main enzyme expressed in the terpenoid metabolism. Moreover, 18 others genes detected were known to be involved in ubiquinone and other terpenoid-quinone biosynthesis. Gibberellin is known to be synthesized via terpenoid biosynthesis pathway (Boba et al., 2020), and kaurene oxidase catalyzes the gibberellin biosynthesis. Besides, ent-kaurene is a tetracyclic hydrocarbon precursor for gibberellins. Moreover, the enzyme activities measured using ELISA kits, including DXS, DXR, GPPS, GGPPS, PC, DGAT, FPS, and HGMR were found to be increased gradually during the storage, and generally attained maximum levels at 8 DAP (**Figure 6A**).

DISCUSSION

Volatile Components of Mango at Different Stages

Volatile components in mango account for the tempting aroma of the fruit, which is also the most important quality of mango pulp and processed mango products. In common with nutritional quality, texture, and color, and the aroma is generally shaped by the coordination of biochemical and developmental pathways (Fujisawa et al., 2013). Many volatile components have been isolated and identified in mature mango and its juice of different cultivars, and some aroma-contributing compounds have also been confirmed in the previous studies (Andrade et al., 2000; Lalel et al., 2003; Pino and Mesa, 2006; Shivashankara et al., 2006; Munafo et al., 2014; Zhang et al., 2019).

In line with a previous study, volatile profiles of Tainong mango of different stages of harvesting and storage, including three stages during the fruit development and four stages during storage, were determined. Literature shows that the transcriptomes of Alphonso mango pulp and flower collected from the seven stages of fruit development and ripening were

TABLE 2 | UniGene ID of the transcripts and compounds identified in the mango pulp samples.

UniGene ID	Gene details
Cluster-15176.1075	GAY40054.1| hypothetical protein
Cluster-15176.3381	KDP37976.1| hypothetical protein
Cluster-15176.3380	XP_015574955.1| gibberellin 2-beta-dioxygenase
Cluster-15991.0	PSS05904.1| gibberellin 2-beta-dioxygenase
Cluster-19253.0	XP_006431588.1| ent-kaurene oxidase
Cluster-15176.332	ESR57297.1| ent-copalyl diphosphate synthase
Cluster-3324.0	XP_006467978.1| gibberellin 2-beta-dioxygenase 8 isoform X3
Cluster-15176.12278	XP_021808327.1| ent-kaurenoic acid oxidase 1-like,
Cluster-15176.13875	GAV73636.1| patatin domain-containing protein
Cluster-15176.7685	PIA28198.1| hypothetical protein
Cluster-15176.7686	EEF47638.1| diacylglycerol cholinephosphotransferase
Cluster-8978.0	XP_006433476.1| phospholipase D zeta 2 isoform X2
Cluster-15176.21081	XP_022878380.1| choline..ethanolaminephosphotransferase 1-like
Cluster-15176.17421	OAY57973.1| hypothetical protein
Cluster-15176.14147	XP_024036163.1| phospholipase D alpha 1
Cluster-15176.27952	KDO81980.1| hypothetical protein
Cluster-15176.22073	ESR39921.1| hypothetical protein
Cluster-15176.23120	AQZ55551.1| 9-lipoxygenase
Cluster-15176.23122	AQZ55556.1| linoleate 9S-lipoxygenase
Cluster-15176.23123	ANF89411.1| linoleate 9S-lipoxygenase
Cluster-15176.19505	ESR33138.1| hypothetical protein
Cluster-15176.23124	AQZ55551.1| 9-lipoxygenase
Cluster-15176.13300	ESR59166.1| hypothetical protein
Cluster-15176.23657	GAV90434.1| adh_short domain-containing protein
Cluster-15176.4077	ESR36401.1| hypothetical protein
Cluster-15176.7778	GAV68539.1| acetyl-CoA carboxylase carboxyl transferase subunit alpha
Cluster-15176.5176	XP_021897331.1| acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha
Cluster-15176.21236	AIU99499.1| hydroxyacyl-ACP dehydrase
Cluster-15176.15013	GAY44796.1| hypothetical protein
Cluster-15176.19615	KDO86915.1| hypothetical protein
Cluster-15176.19806	KRG93630.1| hypothetical protein
Cluster-15176.19807	OMP10121.1| biotin..lipoyl
Cluster-15176.13810	XP_006488591.1| malonyl CoA-acyl carrier protein transacylase
Cluster-15176.23708	GAY52806.1| hypothetical protein
Cluster-15176.8090	KVI00409.1| Fatty acid desaturase
Cluster-15176.15658	ESR56644.1| hypothetical protein
Cluster-15176.20399	OWM86900.1| hypothetical protein
Cluster-15176.6573	GAY39911.1| hypothetical protein
Cluster-15176.17613	XP_006492369.1| long chain acyl-CoA synthetase 1
Cluster-15176.1088	PON38544.1| fatty acid desaturase
Cluster-15176.8360	XP_006482734.1| biotin carboxyl carrier protein of acetyl-CoA carboxylase 2
Cluster-15176.27073	KJB64462.1| hypothetical protein
Cluster-15176.27077	XP_006434188.1| long chain acyl-CoA synthetase 9
Cluster-15176.27074	XP_008784201.1| long chain acyl-CoA synthetase 9
Cluster-15176.27075	GAY41784.1| hypothetical protein
Cluster-15176.11597	AOR17397.1| palmitoyl-acyl carrier protein thioesterase
Cluster-15176.16995	XP_024045530.1| 3-oxoacyl-[acyl-carrier-protein] synthase III
Cluster-15176.22688	ESR59166.1| hypothetical protein
Cluster-15176.13123	GAY44882.1| hypothetical protein
Cluster-15176.16081	XP_022757277.1| acetyl-coenzyme A carboxylase carboxyl transferase
Cluster-15176.21237	XP_006492623.1| 3-hydroxyacyl-[acyl-carrier-protein] dehydratase
Cluster-15176.15551	AKA09592.1| stearoyl-ACP desaturase

(Continued)

TABLE 2 | Continued

UniGene ID	Gene details
Cluster-15176.15554	GAY44796.1| hypothetical protein
Cluster-15176.17196	GAV92051.1| acyl-ACP_TE domain-containing protein
Cluster-15176.7526	XP_021284424.1| stearoyl-[acyl-carrier-protein] 9-desaturase 6
Cluster-15176.24109	ESR56805.1| hypothetical protein
Cluster-15176.16301	PON38544.1| fatty acid desaturase
Cluster-15176.24027	PPS18419.1| hypothetical protein
Cluster-15176.24028	XP_006480952.1| acetyl-coenzyme A carboxylase carboxyl transferase
Cluster-15176.18609	XP_009413949.1| long chain acyl-CoA synthetase 6
Cluster-15176.18608	PIN10227.1| long-chain acyl-CoA synthetases
Cluster-15176.15400	CBI26222.3| unnamed protein product
Cluster-15176.24150	XP_024443812.1| enoyl-[acyl-carrier-protein] reductase [NADH] 1
Cluster-15176.24151	XP_023914015.1| enoyl-[acyl-carrier-protein] reductase [NADH] 1
Cluster-15176.24152	AIS93131.1| enoyl-acyl carrier protein reductase [NAD+]
Cluster-15176.24153	POF25828.1| enoyl-[acyl-carrier-protein] reductase [NADH]
Cluster-4351.0	XP_006467945.1| (3S,6E)-nerolidol synthase 1-like
Cluster-15176.10743	PHU22294.1| acyl-[acyl-carrier-protein] desaturase
Cluster-15176.23658	GAY44036.1| hypothetical protein
Cluster-15176.14875	OMO53044.1| thiolase
Cluster-19268.0	XP_021679005.1| very-long-chain enoyl-CoA reductase-like
Cluster-15176.19912	ESR59974.1| hypothetical protein
Cluster-13603.0	XP_013702088.1| palmitoyl-monogalactosyldiacylglycerol delta-7 desaturase
Cluster-15176.9059	XP_021736703.1| palmitoyl-monogalactosyldiacylglycerol delta-7 desaturase
Cluster-15176.16917	AQZ55555.1| acyl-CoA-oxidase-1
Cluster-15176.6546	GAY56458.1| hypothetical protein
Cluster-15176.20306	AIE48274.1| omega-6 fatty acid desaturase
Cluster-15176.12892	AIC34706.1| omega-6 fatty acid desaturase
Cluster-15176.24984	KYP77364.1| chalcone synthase 1
Cluster-15176.2365	XP_006437739.2| leucoanthocyanidin reductase
Cluster-15176.2366	POE95046.1| leucoanthocyanidin reductase
Cluster-15176.17061	AIY25000.1| flavanone 3'-hydroxylase
Cluster-15176.4728	AIY24979.1| cinnamate 4-mono-oxygenase
Cluster-15176.15897	APR63682.1| p-coumarate 3-hydroxylase 3
Cluster-15176.9748	XP_021664632.1| 4-coumarate–CoA ligase
Cluster-15176.14112	OMO89673.1| oxoglutarate..iron-dependent dioxygenase
Cluster-15176.13889	AIY24995.1| flavanone 3-hydroxylase 2
Cluster-15176.6688	AIY25002.1| dihydroflavonol 4-reductase
Cluster-15176.2805	BAX25479.1| alkyldiketide-CoA synthase
Cluster-15176.19426	XP_022862158.1| 4-coumarate–CoA ligase-like 7
Cluster-15176.1022	KDO60238.1| hypothetical protein
Cluster-15176.4870	XP_021294689.1| 4-coumarate–CoA ligase 1-like
Cluster-19931.0	KDO46519.1| hypothetical protein
Cluster-15176.18933	TAT| tyrosine aminotransferase
Cluster-15176.29835	AAP97931.1| tocopherol cyclase
Cluster-15176.18801	XP_018842576.1| tocopherol cyclase
Cluster-15176.8468	PSR96325.1| NAD(P)H dehydrogenase FQR1-like
Cluster-15176.1733	PRQ41134.1| putative 1,4-dihydroxy-2-naphthoyl-CoA synthase
Cluster-15176.12108	XP_011082392.1| probable NAD(P)H dehydrogenase (quinone) FQR1-like 3 isoform X3
Cluster-15176.6725	ESR60477.1| hypothetical protein
Cluster-15176.6726	GAY65474.1| hypothetical protein
Cluster-15176.6727	XP_022732769.1| probable aminotransferase
Cluster-15176.26169	PSR95337.1| nicotianamine aminotransferase
Cluster-15176.16646	XP_006488032.1| 4-coumarate–CoA ligase-like 7
Cluster-15176.7245	XP_006469904.1| tyrosine aminotransferase
Cluster-18980.0	XP_006484288.1| protein PHYLLO

determined (Deshpande et al., 2017a). The volatile profiling and transcriptome were integrated for the identification of genes related to the metabolism of volatiles.

A previous study reported that 4-hydroxy-2,5-dimethyl-3(2H)-furanone was an important aromatic compound detected in mango cultivars Haden, White Alfonso, Praya Sowoy, Royal Special, and Malindi (Munafo et al., 2014). However, the GC-MS data showed that the mango pulp samples had a low amount of 4-hydroxy-2,5-dimethyl-3(2H)-furanone. Ethyl octanoate, 3-carene, limonene, α-terpinene, α-terpinolene, hexanal, and *p*-cymene were also the key volatiles detected in the different cultivars of Australian mango (San et al., 2017). Some of these compounds, such as ethyl octanoate, terpinolene, and cymene, were not detected in Tainong mango. A total of 12 components, including 2,4-dimethylstyrene, were identified as the major aroma active compounds in Keitt mango juice (Zhang et al., 2019).

The mango stored at these two stages had a more intense aroma than the other five stages. It could be due to a gradual increase in the total sugar content and a gradual decrease in total acid content during fruit maturity (**Figure 1B**). The results also showed that 17 volatile components were the key aroma active compounds in Tainong mango, especially ethanol and (*E*,*Z*)-2,6-nonadienal. Moreover, three up-regulated volatiles in both groups of 60 DAF vs. 40 DAF and 80 DAF vs. 40 DAF, and about 180 volatile components identified in the mango pulp samples might be the potential aromatic compounds.

Genes and Enzymes Involved in Metabolism of Aromatic Components

The transcriptome studies have put forth important information concerning the development of mango of different cultivars, such as Zill (Wu et al., 2014), Langra (Azim et al., 2014), Kent (Dautt-Castro et al., 2015), Dashehari (Srivastava et al., 2016), and Alphonso (Deshpande et al., 2017a). Most of these studies reported the general metabolic pathways involved in the biosynthesis of metabolites in the mango. The genes encoding multiple enzymes related to gluconeogenesis from carbohydrate metabolism, glycolysis, fatty acid biosynthesis and beta-oxidation, salicylic acid biosynthesis, citrate cycle, ethylene biosynthesis, amino acids biosynthesis and degradation, β-carotene biosynthesis, α-tocopherol biosynthesis, flavonoid biosynthesis, and terpenoid backbone synthesis have also been reported in the literature.

In this study, RNA-Seq was performed to explore the molecular mechanism of aroma compounds biosynthesis in mango during the fruit development, ripening, and storage. A large number of differentially expressed transcripts involved in multiple pathways have been identified among these samples. Quantitative RT-PCR analysis also showed that the relative expression patterns of eleven genes were consistent with the RNA-Seq data. The results indicate that the transcriptome data are reliable. The number of differentially expressed transcripts in the samples consisted of over 50,000 up-regulated and down-regulated transcripts. These genes are related to the development of mango. Only three genes were highly expressed and these genes are closely related to the biogenesis of aromatic components during the fruit development. These two genes, E5.5.1.13 and KAO, have never been reported in the transcriptome analysis and metabolic profiling of mango. Another gene that was also found to be highly expressed in the mango pulp samples of 8 DAP. K15095 was the only gene found to be highly expressed during the storage of mango.

A large number of transcription factors (TFs) has been identified in the fruit samples as these factors are known to regulate the fruit development and formation of the aroma of the fruit (Bastías et al., 2011; Hong et al., 2012; Shen et al., 2016; Liu et al., 2017; Lü et al., 2018; Zhang et al., 2018). These TFs might have been participated in controlling the mango development and the biosynthesis of aromatic components of the fruit. The contents of terpenes in mango have been quantified and reported in the literature (Pandit et al., 2009a,b). The expression of genes involved in GPP, FPP, and GGPP synthesis have also been studied (Azim et al., 2014; Wu et al., 2014; Dautt-Castro et al., 2015; Srivastava et al., 2016; Deshpande et al., 2017a). As reported by Ma et al. (2006), DXS, DXR, GPPS, GGPPS, PC, DGAT, FPS, and HGMRs are the key enzymes for terpenoid-isoprenoid biosynthesis. The results obtained from ELISA assays revealed that the activities of these enzymes increased gradually after harvested and reached a maximum level on day 8. On the contrary, the PCR analysis did not show expression of the genes related to these enzymes in the mango pulp samples. In this study, the two highly expressed genes are known to be involved in the pathways of diterpenoid biosynthesis.

CONCLUSION

Mango is favorable due to its pleasant sensory quality and high nutritional values. Although some of the aromatic components have been identified in different cultivars of mango, little is known about the volatile profile in the Tainong mango. 181 volatiles were isolated and identified in fruits collected at seven stages. These components, especially ethanol and (*E*,*Z*)-2,6-nonadienal, were the key aroma active compounds in Tainong mango. RNA-Seq and comparative analysis showed a large number of DEGs during development and after picking. These involved in catalytic activity, transferase activity, ADP binding, transcription factor activity, and oxidoreductase activity. The content of α-pinene, expression of genes involved in terpenoid metabolism, and enzyme activities in terpenoid metabolic pathways gradually increased after picking and generally attained their maximum levels on day-8. The integrative analyses also revealed potential molecular insights into fruit development and aroma formation. This study provides important cues for future work on mango quality improvement.

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/), and the data accession number is PRJNA697524.

AUTHOR CONTRIBUTIONS

CL, MX, PY, and YT performed the experiments. CL conceived and designed the research. LL and JS analyzed data. MX and XH prepared the manuscript. CL, JS, and HEK checked and revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was financially supported by the Guangxi Natural Science Foundation (2018GXNSFAA281149), the Major Program of Science and Technology in Guangxi (Guike AA17204038 and Guike AA17204042), the Key Research and Development Programs of Guangxi (Guike AB18221110 and Guike AB18294027), and Special Fund for "Bagui Scholars" of Guangxi ([2016]21).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.666805/full#supplementary-material

Supplementary Figure 1 | Overview the transcriptome analysis of mango pulp samples at different developmental and ripening stages. **(A)** The frequency and numbers of transcripts with different length; **(B)** The count of CDs with different length predicted by comparing with NR protein library and Swissprot protein library (left) or by estscan (3.0.3) software (right); **(C)** The hierarchical clustering analysis of Pearson correlation according to the level of gene expression levels in fruits of A, B, C, D, E, F, and G. A, 40 DAF; B, 60 DAF; C, 80 DAF; D, 0 DAP; E, 4 DAP; F, 8 DAP; G, 12 DAP. _1, _2 and _3 represent the Triplicate analyses of each sample.

REFERENCES

Andrade, E. H. A., Maia, J. G. S., and Maria das Graças, B. Z. (2000). Aroma volatile constituents of Brazilian varieties of mango fruit. *J. Food Compos. Anal.* 13, 27–33. doi: 10.1006/jfca.1999.0841

Azim, M. K., Khan, I. A., and Zhang, Y. (2014). Characterization of mango (*Mangifera indica* L.) transcriptome and chloroplast genome. *Plant Mol. Biol.* 85, 193–208. doi: 10.1007/s11103-014-0179-8

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

Boba, A., Kostyn, K., Kozak, B., Wojtasik, W., Preisner, M., Prescha, A., et al. (2020). Fusarium oxysporum infection activates the plastidial branch of the terpenoid biosynthesis pathway in flax, leading to increased ABA synthesis. *Planta* 251, 1–14. doi: 10.1007/s00425-020-03339-9

Boonbumrung, S., Tamura, H., Mookdasanit, J., Nakamoto, H., Ishihara, M., Yoshizawa, T., et al. (2001). Characteristic aroma components of the volatile oil of yellow keaw mango fruits determined by limited odor unit method. *Food Sci. Technol. Res.* 7, 200–206. doi: 10.3136/fstr.7.200

Chauhan, O. P., Raju, P. S., and Bawa, A. S. (2010). "Mango flavor," in *Handbook of Fruit and Vegetable Flavors*, ed. Y. H. Hui (Hoboken, NJ: Wiley), 319–343. doi: 10.1002/9780470622834.ch19

Dautt-Castro, M., Ochoa-Leyva, A., Contreras-Vergara, C. A., Pacheco-Sanchez, M. A., Casas-Flores, S., Sanchez-Flores, A., et al. (2015). Mango (*Mangifera indica* L.) cv. Kent fruit mesocarp de novo transcriptome assembly identifies gene families important for ripening. *Front. Plant Sci.* 6:62. doi: 10.3389/fpls.2015.00062

Deshpande, A. B., Anamika, K., Jha, V., Chidley, H. G., Oak, P. S., Kadoo, N. Y., et al. (2017a). Transcriptional transitions in Alphonso mango (*Mangifera indica* L.) during fruit development and ripening explain its distinct aroma and shelf life characteristics. *Sci. Rep.* 7:8711. doi: 10.1038/s41598-017-08499-5

Deshpande, A. B., Chidley, H. G., Oak, P. S., Pujari, K. H., Giri, A. P., and Gupta, V. S. (2017b). Isolation and characterization of 9-lipoxygenase and epoxide hydrolase 2 genes: Insight into lactone biosynthesis in mango fruit (*Mangifera indica* L.). *Phytochemistry* 138, 65–75. doi: 10.1016/j.phytochem.2017.03.002

Fujisawa, M., Nakano, T., Shima, Y., and Ito, Y. (2013). A large-scale identification of direct targets of the tomato MADS box transcription factor RIPENING INHIBITOR reveals the regulation of fruit ripening. *Plant Cell* 25, 371–386. doi: 10.1105/tpc.112.108118

Hong, G. J., Xue, X. Y., Mao, Y. B., Wang, L. J., and Chen, X. Y. (2012). Arabidopsis MYC2 interacts with DELLA proteins in regulating sesquiterpene synthase gene expression. *Plant Cell* 24, 2635–2648. doi: 10.1105/tpc.112.098749

Kulkarni, R. S., Chidley, H. G., Pujari, K. H., Giri, A. P., and Gupta, V. S. (2012). Geographic variation in the flavour volatiles of Alphonso mango. *Food Chem.* 130, 58–66. doi: 10.1016/j.foodchem.2011.06.053

Lalel, H. J. D., Singh, Z., and Tan, S. C. (2003). Aroma volatiles production during fruit ripening of 'Kensington Pride' mango. *Postharvest Biol. Technol.* 27, 323–336. doi: 10.1016/S0925-5214(02)00117-5

Lebrun, M., Plotto, A., Goodner, K., Ducamp, M. N., and Baldwin, E. (2008). Discrimination of mango fruit maturity by volatiles using the electronic nose and gas chromatography. *Postharvest Biol. Technol.* 48, 122–131. doi: 10.1016/j.postharvbio.2007.09.010

Li, B., and Dewey, C. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12:323. doi: 10.1186/1471-2105-12-323

Li, X. B., Fan, X. P., Wang, X. L., Cai, L., and Yang, W. C. (2005). The cotton ACTIN1 gene is functionally expressed in fibers and participates in fiber elongation. *Plant Cell* 17, 859–875. doi: 10.1105/tpc.104.029629

Liao, L., Dong, T., Qiu, X., Rong, Y., Wang, Z., and Zhu, J. (2019). Nitrogen nutrition is a key modulator of the sugar and organic acid content in citrus fruit. *PLoS One* 14:e0223356. doi: 10.1371/journal.pone.0223356

Liu, H. Y., Cao, X. M., Liu, X. H., Xin, R., Wang, J. J., Gao, J., et al. (2017). UV-B irradiation differentially regulates terpene synthases and terpene content of peach. *Plant Cell Environ.* 40, 2261–2275. doi: 10.1111/pce.13029

Lü, 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

Luo, C., He, X. H., Chen, H., Hu, Y., and Ou, S. J. (2013). Molecular cloning and expression analysis of four actin genes (MiACT) from mango. *Biol. Plant.* 57, 238–244. doi: 10.1007/s10535-012-0278-9

Ma, L., Ding, P., Yang, G., and He, G. (2006). Advances on the plant terpenoid isoprenoid biosynthetic pathway and its key enzymes. *Biotechnol. Bull. 2006, S* 1, 22–30.

Munafo, J. P. Jr., Didzbalis, J., Schnell, R. J., Schieberle, P., and Steinhaus, M. (2014). Characterization of the major aroma-active compounds in mango (*Mangifera indica* L.) cultivars Haden, White Alfonso, Praya Sowoy, Royal Special, and Malindi by application of a comparative aroma extract dilution analysis. *J. Agric. Food Chem.* 62, 4544–4551. doi: 10.1021/jf5008743

Pandit, S. S., Chidley, H. G., Kulkarni, R. S., Pujari, K. H., Giri, A. P., and Gupta, V. S. (2009a). Cultivar relationships in mango based on fruit volatile profiles. *Food Chem.* 114, 363–372. doi: 10.1016/j.foodchem.2008.09.107

Pandit, S. S., Kulkarni, R. S., Chidley, H. G., Giri, A. P., Pujari, K. H., Koöllner, T. G., et al. (2009b). Changes in volatile composition during fruit development and ripening of 'Alphonso' mango. *J. Sci. Food Agr.* 89, 2071–2081. doi: 10.1002/jsfa.3692

Pino, J. A. (2012). Odour-active compounds in mango (*Mangifera indica* L. cv. Corazoìn). *Int. J. Food Sci. Technol.* 47, 1944–1950. doi: 10.1111/j.1365-2621.2012.03054.x

Pino, J. A., and Mesa, J. (2006). Contribution of volatile compounds to mango (*Mangifera indica* L.) aroma. *Flavour Fragr. J.* 21, 207–213. doi: 10.1002/ffj.1703

San, A. T., Joyce, D. C., Hofman, P. J., Macnish, A. J., Webb, R. I., Matovic, N. J., et al. (2017). Stable isotope dilution assay (SIDA) and HS-SPME-GCMS quantification of key aroma volatiles for fruit and sap of Australian mango cultivars. *Food Chem.* 221, 613–619. doi: 10.1016/j.foodchem.2016.11.130

Shen, S. L., Yin, X. R., Zhang, B., Xie, X. L., Jiang, Q., Grierson, D., et al. (2016). CitAP2.10 activation of the terpene synthase CsTPS1 is associated with the synthesis of (+)-valencene in 'Newhall' orange. *J. Exp. Bot.* 67, 4105–4115. doi: 10.1093/jxb/erw189

Shi, J., Zou, X., Huang, X., Zhao, J., Li, Y., Hao, L., et al. (2013). Rapid detecting total acid content and classifying different types of vinegar based on near infrared spectroscopy and least-squares support vector machine. *Food Chem.* 138, 192–199. doi: 10.1016/j.foodchem.2012.10.060

Shivashankara, K. S., Isobe, S., Horita, H., Takenaka, M., and Shiina, T. (2006). Volatile aromatic constituents of tree ripened and mature green 'Irwin' mango fruits during low temperature storage. *J. Jpn. Soc. Hortic. Sci.* 75, 209–212. doi: 10.2503/jjshs.75.209

Srivastava, S., Singh, R. K., Pathak, G., Goel, R., Asif, M. H., Sane, A. P., et al. (2016). Comparative transcriptome analysis of unripe and mid-ripe fruit of *Mangifera indica* (var."Dashehari") unravels ripening associated genes. *Sci. Rep.* 6:32557. doi: 10.1038/srep32557

Su, P., Tong, Y., Cheng, Q., Hu, Y., Zhang, M., Yang, J., et al. (2016). Functional characterization of ent-copalyl diphosphate synthase, kaurene synthase and kaurene oxidase in the *Salvia miltiorrhiza* gibberellin biosynthetic pathway. *Sci. Rep.* 6:23057.

Szymczyk, P., Szymańska, G., Lipert, A., Weremczuk-Je/.zyna, I., and Kochan, E. (2020). Computer-aided saturation mutagenesis of *Arabidopsis thaliana* ent-copalyl diphosphate synthase. *Interdiscip. Sci.* 12, 32–43. doi: 10.1007/s12539-019-00342-x

Tamura, H., Boonbumrung, S., Yoshizawa, T., and Varanyanond, W. (2001). The volatile constituents in the peel and pulp of a green Thai mango, Khieo Sawoei cultivar (*Mangifera indica* L.). *Food Sci. Technol. Res.* 7, 72–77. doi: 10.3136/fstr.7.72

Wu, H. X., Jia, H. M., Ma, X. W., Wang, S. B., Yao, Q. S., Xu, W. T., et al. (2014). Transcriptome and proteomic analysis of mango (*Mangifera indica* L.) fruits. *J. Proteomics* 105, 19–30. doi: 10.1016/j.jprot.2014.03.030

Zhang, L., Zhang, Q., Li, W., Zhang, S., and Xi, W. (2019). Identification of key genes and regulators associated with carotenoid metabolism in apricot (*Prunus armeniaca*) fruit using weighted gene coexpression network analysis. *BMC Genom.* 20:876. doi: 10.1186/s12864-019-6261-5

Zhang, W., Dong, P., Lao, F., Liu, J., Liao, X., and Wu, J. (2019a). Characterization of the major aroma-active compounds in Keitt mango juice: comparison among fresh, pasteurization and high hydrostatic pressure processing juices. *Food Chem.* 289, 215–222. doi: 10.1016/j.foodchem.2019.03.064

Zhang, Y., Yin, X., Xiao, Y., Zhang, Z., Li, S., Liu, X., et al. (2018). An ETHYLENE RESPONSE FACTOR-MYB transcription complex regulates furaneol biosynthesis by activating QUINONE OXIDOREDUCTASE expression in strawberry. *Plant Physiol.* 178, 189–201. doi: 10.1104/pp.18.00598

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 12 May 2021
doi: 10.3389/fpls.2021.629361

Metabolome and Transcriptome Integration Reveals Insights Into Flavor Formation of 'Crimson' Watermelon Flesh During Fruit Development

*Chengsheng Gong, Weinan Diao, Hongju Zhu, Muhammad Jawad Umer, Shengjie Zhao, Nan He, Xuqiang Lu, Pingli Yuan, Muhammad Anees, Dongdong Yang, M. O. Kaseb and Wenge Liu**

Zhengzhou Fruit Research Institute, Chinese Academy of Agricultural Sciences, Zhengzhou, China

Metabolites have been reported as the main factor that influences the fruit flavor of watermelon. But the comprehensive study on the dynamics of metabolites during the development of watermelon fruit is not up-to-date. In this study, metabolome and transcriptome datasets of 'Crimson' watermelon fruit at four key developmental stages were generated. A total of 517 metabolites were detected by ultrahigh-performance liquid chromatography–electrospray ionization–tandem mass spectrometry and gas chromatography–solid-phase microextraction–mass spectrometry. Meanwhile, by K-means clustering analysis, the total differentially expressed genes were clustered in six classes. Integrating transcriptome and metabolome data revealed similar expression trends of sugars and genes involved in the glycolytic pathway, providing molecular insights into the formation of taste during fruit development. Furthermore, through coexpression analysis, we identified five differentially expressed ADH (alcohol dehydrogenase) genes (*Cla97C01G013600*, *Cla97C05G089700*, *Cla97C01G001290*, *Cla97C05G095170*, and *Cla97C06G118330*), which were found to be closely related to C9 alcohols/aldehydes, providing information for the formation of fruit aroma. Our findings establish a metabolic profile during watermelon fruit development and provide insights into flavor formation.

Keywords: watermelon, transcriptome, metabolome, sugars, volatile organic compounds, coexpression

OPEN ACCESS

Edited by:
Jinhe Bai,
Horticultural Research Laboratory (USDA-ARS), United States

Reviewed by:
Pengxiang Fan,
Zhejiang University, China
Shaogui Guo,
Beijing Academy of Agricultural and Forestry Sciences, China

****Correspondence:***
Wenge Liu
liuwenge@caas.cn

Specialty section:
This article was submitted to Plant Metabolism and Chemodiversity, a section of the journal Frontiers in Plant Science

Received: *14 November 2020*
Accepted: *29 March 2021*
Published: *12 May 2021*

Citation:
Gong C, Diao W, Zhu H, Umer MJ, Zhao S, He N, Lu X, Yuan P, Anees M, Yang D, Kaseb MO and Liu W (2021) 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

INTRODUCTION

Flavor quality is one of the most important criteria for consumers to select vegetables and fruits (Barrett et al., 2010) and promotes customers' desire for reconsumption. However, as flavor phenotyping is complex, expensive, and not suitable for high-throughput analysis, most breeders tend to focus more on yield, appearance quality, and plant resistance than fruit flavor quality (Tieman et al., 2017). The flavor is typically described as a sum of the interactions between taste and aroma. Generally speaking, the formation of flavor is the result of the interaction of multiple metabolites. The taste is activated mainly through taste receptors in the mouth. Sugars and acids are an essential basis for evaluating the taste of fruits such as tomatoes (Zhu et al., 2018). The aroma is the sense of smell through the nose to perceive volatile organic compounds (VOCs). Different fruits have different aromas that may result from a combination of multiple metabolites,

for instance, 2,6-nonadienal is the characteristic VOC in cucumber (Forss et al., 1962), but it is also detected in other fruits with different fruit aromas. As the market demand for high-quality fruit is a feature of modern consumer society, a more comprehensive understanding of fruit flavor-related metabolites and their improvement should be an important objective in breeding programs (Kyriacou and Rouphael, 2018).

Watermelon (*Citrullus lanatus*) is among the top five most consumed fresh fruits globally, with unique flavor, rich nutrition, and health benefits to humans (Guo et al., 2013; Zhang et al., 2017). Ramirez et al. (2020) conducted a sensory profile of seven watermelon varieties, and the result indicates that flavor had an important effect on the refreshing perception of watermelon. Undoubtedly, the formation of watermelon flavor is a combination of different metabolites, although the current understanding of flavor metabolites is still insufficient. Sweetness is a crucial factor for the commercial value of watermelon, but studies on sweetness mainly focus on sucrose, fructose, and glucose (Gao et al., 2018b), and research on watermelon acidity has focused on malic acid, citric acid, and oxalic acid (Umer et al., 2020). In parallel, although the acidity in watermelon is relatively low, it may regulate the feeling of sweetness (Soteriou et al., 2014). Progress has been made in watermelon breeding by improving metabolites. Gao et al. (2018a) developed a new watermelon variety 'SW' with a sweet and sour taste, which is an innovation in the breeding of watermelon taste. VOCs play a key role in the formation of fruit flavor, Bianchi et al. (2020) detected a total of 20 VOCs at the ripening stage of two mini watermelons 'Rugby' and 'Cuoredolce®', of which 13 were C9 aldehydes and alcohols [a metabolite with nine C atoms in the lipoxygenase (LOX) pathway, such as 2,6-nonadienal, (E,Z)-]. Existing studies have promoted the study of watermelon flavor; however, it is necessary to conduct a comprehensive and extensive metabolomic study on the development process of watermelon to establish watermelon flavor. Related research has been found in melon and other fruits; for example, approximately 2,000 metabolite signatures and 15 mineral elements were detected in melon, which provided valuable information for the study of melon flavor and nutrition (Moing et al., 2011). In recent years, with the continuous improvement of detection technology, the widely targeted metabonomics method developed by Chen et al. (2013) can realize the high-throughput determination of metabolites, and solid-phase microextraction (SPME)–gas chromatography (GC)–MS has been widely used in the detection of VOCs. The progress of detection methods provides a powerful tool for us to carry out watermelon flavor research.

Transcriptional regulation is the most important regulation mode of organisms. With the continuous development of sequencing technology, transcriptome has been widely used. For instance, Guo et al. (2015) systematically explained the molecular basis of differences in important traits, such as sugar content, flesh texture, and color through transcriptome profiles analysis of wild and cultivated watermelon; Gao et al. (2018b) identified the key genes of soluble sugars and organic acids through integrated analysis in a subsequent study. Notably, integrating of transcriptome data with metabolic profiles is an effective method to parse metabolic pathways and plant gene function (Luo, 2015); however, related research in watermelon is extremely lacking. In plants, metabolites associated with taste are synthesized and degraded through critical biological processes such as glycolysis and the tricarboxylic acid cycle (Nativ et al., 2017). VOCs are synthesized mainly through methylerythritol phosphate or mevalonate pathway to generate terpenoids; phenylpropane/benzene compounds through aromatic amino acids; and C6 and C9 alcohols/aldehydes formed from unsaturated fatty acids, such as linoleic acid and linolenic acid (Dudareva et al., 2013; Wei et al., 2016); the study of metabolic pathways has increased the understanding of the links between genes and metabolites. In conclusion, data integration platform is a helpful approach to exploring the hidden diversity in flavor formation by using transcriptome and metabolome profiles.

In this study, we generated the metabolome and transcriptome datasets from the 'Crimson' watermelon cultivar at 10, 18, 26, and 34 days after pollination (DAP). 'Crimson' watermelon cultivar is famous for its nutritionally rich flavor and is widely used in breeding programs. Ultrahigh-performance liquid chromatography–electrospray ionization–tandem mass spectrometry (UPLC–ESI–MS/MS) and GC–SPME–MS detected a total of 517 metabolites. The accumulation patterns of flavor-related metabolites, such as sugars, organic acids, and VOCs, were analyzed. The differentially expressed genes (DEGs) were clustered into six classes by K-mean clustering analysis. Furthermore, the metabolites and genes related to flavor enriched into related metabolic pathways were identified, which provides more comprehensive information for the formation of flavor during the development of watermelon fruit. This study provides molecular insights into the metabolites associated with formation of watermelon flesh flavor and lays a foundation for further studies on metabolomics.

MATERIALS AND METHODS

Experimental Material and Sampling

The experimental materials, watermelon cultivar 'Crimson' seeds, were collected from the National Mid-term Genebank for Watermelon and Melon, Zhengzhou Fruit Research Institute, Chinese Academy of Agricultural Sciences (Zhengzhou, Henan, China). 'Crimson' has a spherical fruit with red color flesh, resistance to disease, nutritionally rich taste, and aroma. Plants of watermelon were planted in the greenhouse (Xinxiang, 35°18′ N, 113°55′ E) in Henan Province, China. All the standard agronomic practices were followed during the developments, such as irrigation, weeding, and fertilization. The seedling was raised and transplanted to the field on April 13, 2019. The experimental design was a completely randomized block design and plant–plant spacing of 0.8 m and a row–row spacing of 1.5 m. And pollination work starts on May 11, 2019. After transplanting, the temperature and humidity of the greenhouse were artificially adjusted: in the daytime, the temperature was generally 25–35°C, and the humidity was 55–70%; in the evening, the temperature was generally more than 15°C, and the humidity was 75–80%. Pollination was done manually: every a second female flower

was artificially self-pollinated and tagged the pollination date, only one watermelon fruit allowed to grow on one plant. Flesh samples were preserved immediately in liquid nitrogen; from all four key developmental stages 10 DAP (immature white), 18 DAP (white-pink), 26 DAP (red), and 34 DAP (ripe) (Zhang et al., 2017), the samples were represented by KL1 (10 DAP), KL2 (18 DAP), KL3 (26 DAP), and KL4 (34 DAP), respectively. The well-grown watermelon fruit, uniform in appearance and without visual defects, was cut in the longitudinal direction, and the center flesh was taken, avoiding seeds. Three biological replicates were collected for each stage, and a total of 12 samples were collected. All the flesh samples were grounded in powder and preserved in liquid nitrogen for further transcriptome and metabolome analysis. A balanced mixture of 12 collected samples was used as the mixed sample for quality control.

Measurement of Fruit Weight, Flesh Firmness, Total Soluble Solids Content, and pH of Watermelon Flesh

At four key development stages, watermelon fruit weight was measured by an electronic scale. When the fruit was cut lengthwise, the center firmness of flesh was determined by FT 011 fruit firmness tester. And after the center flesh was homogenized, a laboratory refractometer (HC-112ATC, Shanghai LICHENKEYI, China) and a PHB-4 (Shanghai LICHENKEYI, China) instrument was used to determine total soluble solids content (TSS, Brix%) and pH, respectively (Gao et al., 2018b). Three separate well-grown fruits were measured and recorded at each developmental stage. The taste of watermelon was evaluated by human mouth and the nose to perceive the aroma of the fruit.

LC–ESI–MS/MS System-Based Widely Targeted Metabolomics Analysis

Freeze-dried watermelon flesh samples were grided evenly by using a mixer mill (MM 400, Retsch) for 1.5 min at 30 Hz. The preparation of watermelon fruit extract was performed as previously described (Qin et al., 2020). In brief, 100-mg sample powder was extracted overnight with 1.2 mL 70% methanol solution at 4°C. After centrifugation at 10,000 *g* for 10 min, the extracts were absorbed (CNWBond Carbon-GCB SPE Cartridge; Anpel, Shanghai, China) and filtered with an SCAA-104 membrane (Anpel, Shanghai, China) for further UPLC–MS/MS analysis.

UPLC–ESI–MS/MS system (UPLC, Shim-pack UFLC SHIMADZU CBM30A system; MS, Applied Biosystems 6500 Q TRAP) was used to analyze the sample extracts. The analytical conditions were as follows: chromatographic column, Waters ACQUITY UPLC HSS T3 C18 (1.8 μm, 2.1 × 100 mm). The column temperature was 40°C, the injection volume was 2 μL, and the procedure for sample measurements by UPLC was as follows: the initial condition was that 95% solvent A (pure water with 0.04% acetic acid), 5% solvent B (acetonitrile with 0.04% acetic acid); after 10 min, the linear gradient to 5% solvent A and 95% solvent B was programmed and kept 1 min, and then solvent B changed to 5% in 0.1 min and kept for 2.9 min, with a flow rate of 0.35 mL/min. Triple quadrupole-linear ion trap mass spectrometer (Q TRAP), equipped with an ESI turbo ion-spray interface, ran in positive and negative ion mode, and was controlled by analyst 1.6.3 software (AB SCIEX). The operating parameters of the ESI were as follows: source temperature was 550°C, ion spray voltage was 5,500 V in positive ion mode and −4,500 V in negative ion mode, and curtain gas (CUR) was set at 50, 60, and 30.0 psi, respectively, and the collision gas (CAD) was high. The instrument was tuned and mass calibrated with 10 and 100 μmol/L propanediol solutions in QQQ and LIT modes, respectively. QQQ scans were used as MRM experiments, and the collision gas (nitrogen) was set to as 5 psi. Further declustering potential and collision energy optimizations were performed for individual MRM transitions. Based on the metabolites eluted during this period, a specific set of MRM transitions was monitored in each period.

Qualitative analysis of MS data was based on the comparison of the retention time (RTS) and fragmentation pattern, accurate precursor ion (Q1) value, product ion (Q3) value, data of which were obtained the from standard injection under the same conditions (Sigma–Aldrich, United States); if no standard, the analysis was performed using the MWDB database (Metware Biotechnology Science and Technology Co., Ltd.), the Kyoto Encyclopedia of Genes and Genomes (KEGG) database[1], and the publicly available metabolite databases online at PLANTCYC[2], etc. MRM mode was used for quantitative analysis. The characteristic ions of each metabolite were screened by QQQ MS, and the signal intensity was obtained. The chromatographic peaks were integrated and corrected using Multiquant version 3.0.2 (AB SCIEX, Concord, Canada). The corresponding relative metabolite content was expressed by chromatographic peak area integral. In addition, metaX software[3] was used for screening and quantitative analysis of the differences of metabolites.

Extraction of Volatiles and SPME–GC–MS Analysis

Fresh samples were gridded into fine powder in liquid nitrogen, and 1 g of powder was immediately transferred to a 20-mL headspace vial (Agilent, Palo Alto, CA, United States) with NaCl saturated solution. The vials were sealed using crimp-top with TFE-silicone headspace septa (Agilent). During SPME analysis, each vial was placed at 60°C for 10 min. A 50/30 μm divinylbenzene/carboxen/polydimethylsilioxane fiber (Sigma) was exposed to the sample's headspace for 20 min at the temperature of 60°C. After sampling, desorption was performed at 250°C for 5 min at the injection port of the GC apparatus (Model 7890B; Agilent) in a shunt-free mode. The model of the mass spectrum was 7000D (Agilent); capillary column: 30 m × 0.25 mm × 0.25 μm DB-5 ms (5% phenyl polymethylsiloxane); carrier gas: helium, with a linear velocity of 1.0 mL/min; injector temperature: 250°C; and detector temperature: 280°C. The oven temperature was programmed from 40°C for 5 min and increased at 6°C/min to 280°C for

[1]http://www.kegg.jp/

[2]http://www.Plantcyc.org/

[3]http://metax.genomics.cn/

5 min; electron impact ionization mode: 70 eV; quadrupole mass detector temperature: 150°C; ion source temperature: 230°C; transfer line temperature: 280°C. The scanning range of MS was m/z 50–500 AMU, and the interval was 1 s. The volatile compounds were identified by comparing MS with the data system library (MWGC or NIST) and linear retention index.

Constriction of RNA-Sequencing Libraries

The total RNA was extracted by RNA extraction kit (TIANGEN, Beijing, Biotech, China) following the manufacturer's instructions. Furthermore, 1% agarose gels electrophoresis was used to analyze the integrity of RNA and any other contamination. The purity of RNA was detected by Nanodrop Nano Photometer (IMPLEN, GmbH, Munich, Germany); RNA concentration was measured by Qubit 2.0 Fluorometer, and the integrity of RNA was checked accurately by Agilent 2100 biological analyzer. Sequencing libraries were established using the NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, United States) and sequenced on an Illumina HiSeq platform, and 125-bp/150-bp paired-end reads were generated. After the libraries' detection, the sequencing was performed on the Illumina HiSeq platform. Raw reads were filtered to get good quality reads. Data were thoroughly checked to identify any false reads; moreover, GC content was also monitored, so that clean reads were obtained and can be used for the subsequent steps. Hisat2 was used to sequence clean reads with the reference genome (97103 V2)[4]. Transcripts or gene expression levels were measured by calculating FPKM (fragments per kilobase of transcript per million fragments mapped) (Mortazavi et al., 2008). Deseq2 (Love et al., 2014; Varet et al., 2016) was used to analyze differential gene expression among different samples. After that, multiple hypothesis tests were used to correct the hypothesis test probability (*P*-value) by Benjamin Hochberg to obtain the false discovery rate (FDR), and the Bioconductor package CusterProfller was used for KEGG enrichment analysis.

Differential Metabolites and Genes Analysis

For the analysis of differentially accumulated metabolites (DAMs), the two developmental stages' screening criteria were as follows: (1) The metabolite accumulation levels were that fold change ≥ 2 or fold change ≤ 0.5; (2) based on the OPLS-DA model analysis results, metabolites with VIP (variable importance in project) ≥ 1 (Yuan et al., 2018). The relative contents of differential metabolites were standardized and centralized before analysis. For DEGs between two developmental stages, the screening criteria were $|\log_2\text{fold change}| \geq 1$ and FDR < 0.05. Principal component analysis (PCA) diagram, hierarchical clustering analysis diagram, and Venn diagram were drawn by R software package, and the Pearson correlation coefficient (PCC) of genes and metabolites was calculated by the COR program in R. Coexpression analysis network diagram drawn by Cytoscape software (version 3.7.1) (Kohl et al., 2011).

[4]http://cucurbitgenomics.org/organism/21

RESULTS

The Ripeness of the Watermelon Fruit Is Adapted to the Formation of the Flavor

The most notable difference was that as watermelon fruits develop and ripen, the color of the flesh changes from light white to bright red (**Figure 1A**); at the same time, through taste evaluation, it was found that at 10 DAP, the taste of watermelon fruit was light, but at 34 DAP, the taste of watermelon became sweet and delicious; this corresponds to olfactory perception; at 10 DAP, the intensity of aroma was not much pleasant, but with a fresh and pleasant aroma at 34 DAP. Notably, significant phenotypic variation was obtained. Average fruit weight increased rapidly from 1.57 kg at 10 DAP to 4.04 kg at 18 DAP, and no significant difference was observed from 26 to 34 DAP (**Figure 1B**). An increasing trend was observed for TSS content at all development stages, i.e., 4.97, 7.87, 9.77, and 10.67 at 10, 18, 26, and 34 DAP (**Figure 1C**), respectively, whereas watermelon flesh's firmness was measured as 2.72, 2.05, 1.68, and 1.27 kg/cm^2 at 10, 18, 26, and 34 DAP (**Figure 1D**), which has an opposite trend compared with fruit weight and TSS. pH can reflect the change of watermelon acidity, high pH is consistent with low acidity. pH values were recorded as 4.95, 5.14, 5.33, and 5.66 at 10, 18, 26, and 34 DAP (**Figure 1E**), respectively, indicating a decrease in acidity during fruit development. As the fruit grows heavier and softer during watermelon development and ripening, the taste and aroma of the fruit become better. Changes of TSS and pH correspond to an increase in sweetness and a decrease in acidity, often caused by metabolite changes.

Overview of Metabolite Accumulation Patterns During Watermelon Fruit Development and Analysis of Key Metabolites Related to Taste

With the development of watermelon fruit, a variety of metabolites were synthesized and degraded. As taste results from the interaction of multiple metabolites, enhancing the information of metabolite changes is conducive to a more comprehensive understanding of watermelon taste formation. Therefore, we have carried out a widely targeted metabolomics method to detect metabolites to recognize metabolite accumulation patterns on a global level. A total of 443 metabolites were detected at different developmental stages, i.e., 10, 18, 26, and 34 DAP by UPLC–ESI–MS/MS. The PCA results showed that the first three principal components explain 35.81, 19.94, and 12.22% of the samples' variance, respectively (**Figure 2A**), and the heat map of metabolites can clearly distinguish the watermelon samples at different developmental stages (**Supplementary Figure 1**). These metabolites contain 135 flavonoids, 60 amino acids and derivatives, 56 lipids, 54 phenolic acids, 53 others, 35 organic acids, 33 nucleotides and derivatives, and 17 alkaloids (**Figure 2B** and **Supplementary Table 1**), which

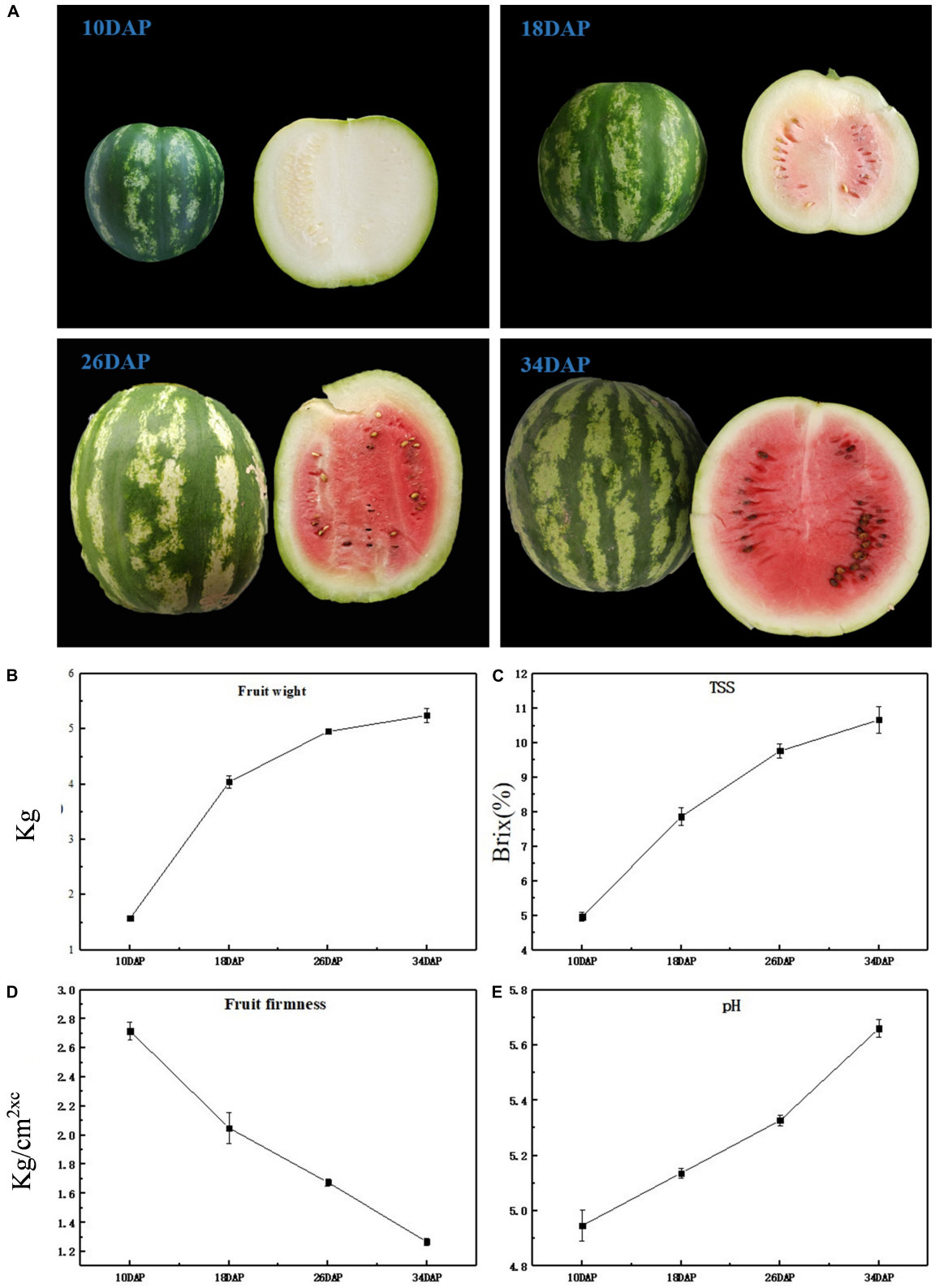

FIGURE 1 | Physiological characteristics of 'Crimson' watermelon fruit. **(A)** The appearance of 'Crimson' at four key development stages (10, 18, 26, and 34 DAP). **(B–E)** Fruit weight, TSS content (Brix%), flesh firmness, and pH during different development stages of 'Crimson' watermelon fruit, respectively. And each stage of these four agronomic traits took three biological replicates.

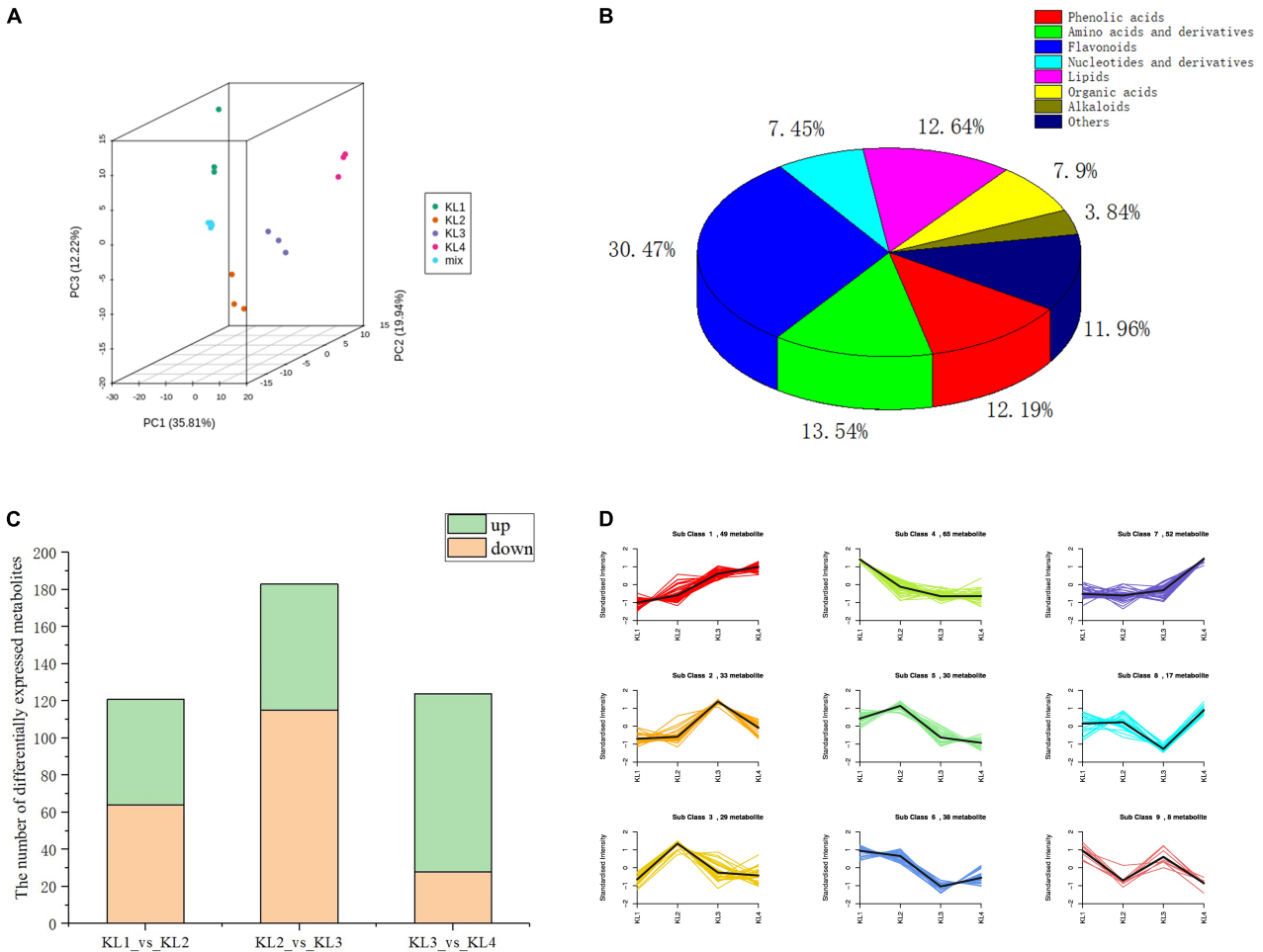

FIGURE 2 | The analysis results of the metabolic profiles obtained by LC–MS during the development of watermelon fruit. **(A)** Principal component analysis (PCA) score plot of all metabolites in 12 samples. KL1, KL2, KL3, and KL4 represent the samples at 10, 18, 26, and 34 days after pollination, respectively (DAP). The samples with the same color represent three biological replications of the same period. **(B)** Pie chart of the classification of 443 metabolites. **(C)** The number of differentially accumulation metabolites (DAMs) by comparing KL1 vs. KL2, KL2 vs. KL3, and KL3 vs. KL4. Green represents the number of up-regulated metabolites, On the contrary, orange stands for down-regulation. **(D)** Dynamics of Metabolite during four development stages. The bold black line represents the average pattern of all candidate metabolites in each class; the different colored lines represent different classes, and each line represents the average relative value of the metabolite over three biological replicates.

shows the diversity of metabolites during watermelon fruit flesh development.

To further clarify the accumulation pattern of different metabolites during watermelon fruit development, K-means clustering algorithm was performed. Three hundred twenty-one DAMs were obtained by comparing the metabolites at different developmental stages, and 121, 183, and 124 DAMs were screened by comparing KL1 (10 DAP) vs. KL2 (18 DAP), KL2 vs. KL3, and KL3 (26 DAP) vs. KL4 (34 DAP), respectively (**Supplementary Table 2**). Significantly, KL2 vs. KL3 had the most significant down-regulated metabolites (115), and KL3 vs. KL4 had the largest number of up-regulated metabolites (96) (**Figure 2C**), indicating that there was a more active metabolic activity from 18 to 34 DAP, which is the key period of watermelon flavor formation. Nine classes were obtained through the analysis of all the DAMs, among which the content of candidate metabolites in classes 1 and 7 presented an increasing trend, with 49 and 52 DAMs, respectively, such as sucrose citrulline and histidine; in contrast, 65 candidate DAMs in class 4 showed a decreasing trend, similarly, in classes 5 and 6, the metabolites also showed a downward trend during the watermelon flesh development, and there were 30 and 38 candidate DAMs, such as naringenin chalcone, vanillin, salicylic acid, and adenosine (**Figure 2D** and **Supplementary Table 1**). Watermelon flavor is derived from the interaction of multiple metabolites, and the watermelon fruit has better flavor during its development, so the metabolites in classes 1 and 7 can be used as potential flavor-related metabolites.

Sugars and organic acids play an important role in the taste formation of watermelon fruits. The K-means cluster analysis included 9 different sugars and 21 different organic

acids. Among them, all nine sugars were enriched in class 1 (**Supplementary Table 1**), such as glucose, trehaloseanhydrous, trehalose 6-phosphate, and sucrose, and showed the same trend with TSS, indicating that these sugars were accumulated during the development of fruit and were responsible for the sweetness of the mature watermelon. The 11 organic acids, including shikimic acid and citridic acid, have a negative accumulation trend and were enriched in classes 4, 5, and 6, which may be adapted to the increase of pH during watermelon fruit development. These key metabolites are thought to be related to the formation of watermelon taste.

As fatty acids are important prerequisites for the generation of alcohols/aldehydes and other VOCs, we analyze the accumulation pattern of fatty acids here to increase our understanding of the synthesis of aroma-related metabolites. It is worth noting that of 16 differential free fatty acids, all but pentadecanoic acids were enriched in class 4, suggesting that the content of free fatty acids decreases during the development of watermelon fruits. However, the aroma of watermelon increased during ripening, so the degradation of free fatty acids was considered related to the synthesis of VOCs. Subsequently, VOCs, especially the metabolites of the LOX pathway, were further determined to understand the formation of watermelon aroma during fruit development.

Eleven of the 74 VOCs Were Significantly Related to the Formation of Aroma During the Development of Watermelon Fruit

To provide insights into the formation of watermelon flesh aroma, VOCs were measured by GC–SPME–MS. Here, a total of 74 VOCs were detected, including 27 hydrocarbons, 15 ketones, 7 aldehydes, 5 benzene and its derivatives, 5 esters, 3 alcohols, and 12 others (**Figure 3A** and **Supplementary Table 3**). By comparing KL1 vs. KL2, KL2 vs. KL3, and KL3 vs. KL4, a total of 23 differentially accumulated VOCs were detected, and there were 12 (11 up-regulated and 1 down-regulated), 14 (11 up-regulated and 3 down-regulated), and 14 metabolites (all up-regulated), respectively (**Figure 3B** and **Supplementary Table 4**), among which, decane, 5-methyl-; 2-nonenal, (E)-; and 3-ethyl-3-methylheptane were the different common metabolites.

Furthermore, hierarchical clustering analysis showed that the accumulation of DAMs was developmental stage-specific (**Figure 3C**). It can be clearly observed that 2-hexenal, (E)- has a higher accumulation at 10 DAP, whereas 2-heptanone, 4,6-dimethyl-; undecane, 3,9-dimethyl-; 5,9-undecadien-2-one, 6,10- dimethyl-,(E)-; 2,6-nonadienal,(E,Z)-; 2-nonenal, (E)-; 3-nonen-1-ol, (Z)-; ethanone, 1-(4-methylphenyl)-; 3-buten-2-one,4-(2,6,6-trimethyl-1-cyclohexen-1-yl)-; 3,6-nonadien-1-ol, (E,Z)-; and 1-(Furan-2-yl)-2-methylpentan-1-one were significantly enriched at 34 DAP. These 11 VOCs were defined as the top abundant components of aroma. Interestingly, we found that the important VOCs such as 2-hexenal, (E)-; 3,6-nonadien-1-ol, (E,Z)-; 2,6-nonadienal, (E,Z)-; 3-nonen-1-ol, (Z)-; and 2-nonenal, (E)-, which were synthesized by the LOX pathway, were significantly enriched in different development stages.

K-Means Clustering Analysis Divides the Expression Pattern of DEGs Into Six Classes

To further investigate changes in gene expression during watermelon fruit at four developmental stages, transcriptome sequencing was performed on 12 samples, and a total of 98.22-Gb clean data were obtained. The clean data of each sample reached six Gb, and the Q30 base percentage was 90% or greater. FPKM was calculated as an index to measure the expression level of transcripts or genes (**Supplementary Table 5**). Correlation analysis was conducted by calculating PCC, and the results showed that for the three samples in the same period, R^2 was all greater than 0.9 (**Supplementary Figure 2A**), indicating good reproducibility between test samples. Overall, the obtained sequencing data were of high quality and could be performed for further analysis. The transcriptome data of watermelon flesh at different developmental stages were separated by PCA, and PC1 and PC2 were 45.63 and 16.6%, respectively (**Supplementary Figure 2B**). These results suggest that our samples have good reproducibility, and there are significant differences in gene expression levels of samples in different development stages.

Different genes in an organism perform different biological functions through interaction, and the pathway annotation analysis of DEGs helps to further interpret the functions of genes. Through KEGG annotation, it was found that the most abundant KEGG terms were metabolic pathways and biosynthesis of secondary metabolites (**Supplementary Figure 3**), implying that genes were involved in a variety of metabolic activities during watermelon fruit development.

Similar to metabolite analysis results, hierarchical clustering analysis showed that DEGs were characterized by specific expression at different developmental stages of watermelon fruit (**Supplementary Figure 4A**), suggesting a potential relationship between genes and metabolites. To understand gene expression patterns during watermelon fruit development, the K-means clustering algorithm was used. K-means cluster analysis results showed that DEGs were divided into six classes, from 10 DAP to 34 DAP (**Supplementary Figure 4B**). Genes of classes 1 and 2 have 2,376 and 3,199 genes, respectively, and showed a downward expression trend, and these genes may play a regulatory role in metabolite accumulation during fruit development; conversely, the DEGs of class 4 (1,549) and class 5 (1,531) was up-regulated during watermelon development, as metabolites associated with flavor formation are positively accumulated during the developmental process of watermelon; therefore, these genes may play an important role in regulating the metabolites associated with flavor-related metabolites. Transcription factors (TFs) may regulate the transcription of multiple genes during gene regulation; as shown here, it is worth noting that by comparing KL1 vs. KL2, KL2 vs. KL3, and KL3 vs. KL4, we found 300, 462, and 258 TFs were identified, respectively.

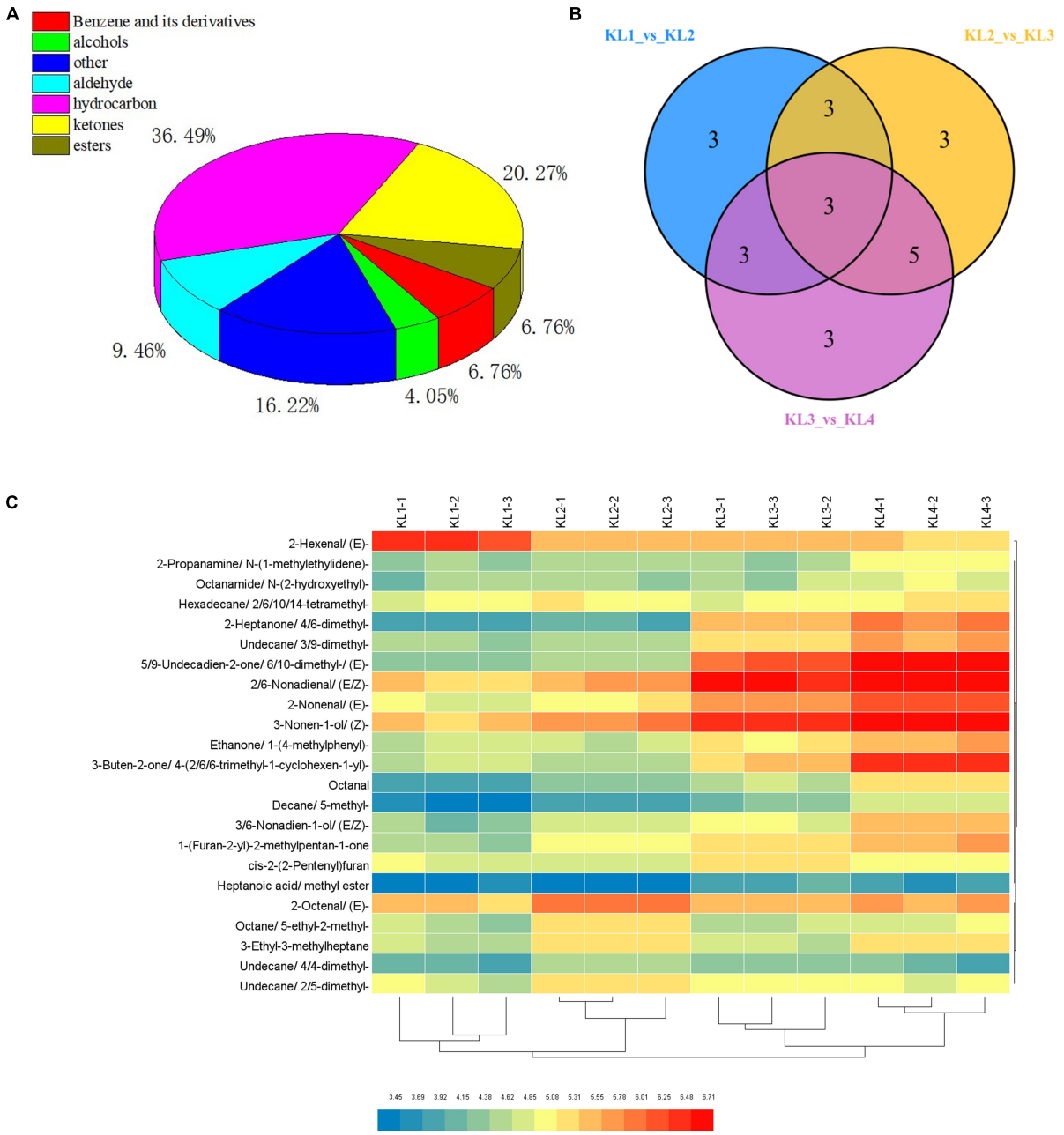

FIGURE 3 | Classification of volatile organic compounds (VOCs) detected by GC–MS and analysis of differential metabolites. **(A)** Pie chart of the classification of 74 VOCs. **(B)** Venn diagram of the number of differential accumulation VOCs by comparing KL1 vs. KL2, KL2 vs. KL3, and KL3 vs. KL4. **(C)** Hierarchical clustering heat map of 23 differential accumulation VOCs in 12 samples, KL1, KL2, KL3, and KL4 represent the samples at 10, 18, 26, and 34 DAP, and -1, -2, and -3 represent different samples from the same period.

Furthermore, the statistics show that there were 162, 271, 86, 48, 80, and 54 TFs with differential expression from class 1 to class 6 (**Supplementary Figure 4C**). The same expression trend was found in TFs and all genes, suggesting that TFs play an important role in regulating structural genes during watermelon fruit development. Similarly, genes with the same or opposite accumulation patterns are more likely to regulate the accumulation of metabolites.

Integrating Related Genes and Metabolites in the Glycolytic Pathway Provides Molecular Insights Into the Formation of Sweet Taste

Sweetness is the most important taste in fruit, and it is the most attractive characteristic of the cultivated watermelon fruit flesh (Guo et al., 2015). The synthesis and degradation of

sugars are usually dependent on glycolytic pathways, which provide energy for living organisms and provide substrates to synthesize many metabolites. To explore the potential links between the metabolome and the transcriptome data, sugars and DEGs were allocated to glycolysis pathways (**Figure 4**). Through data integration analysis, for genes, we observed that important genes in glycolytic pathway were highly expressed in 26 and 34 DAP. Most of these genes were enriched in classes 4 and 5 (with up-regulated expression trend, marked in **Figure 4**), suggesting that vigorous metabolic activities were carried out during fruit development; sugars, such as glucose and sucrose, positively accumulated with fruit development. As the expression of genes and the accumulation of sugars have the same trend, the accumulation of sugars may be mainly the regulation from the transcriptional level. In conclusion, sugars provide substrates for the active development of glycolytic pathway, and the positive expression of glycolytic pathway genes also promotes the accumulation of sugars; the formation of sweet taste in watermelon fruit was relatively consistent with the vigorous gene activity.

Coexpression Analysis of Lipid Metabolites Related Genes in LOX Pathway Provided Insights Into the Formation of Watermelon Aroma

C9 alcohols and aldehydes produced by the LOX pathway are important aroma metabolite components in watermelon development. LOX pathway uses linolenic acid as the precursor to synthesizing related VOCs under the regulation of a series of genes. LOX and ADH (alcohol dehydrogenase) have been shown to play a key role in the synthesis of alcohols and aldehydes (Dudareva et al., 2013).

For further clarification of this study, we screened DEGs by reference genome (see text footnote 4) search combined with FPKM having a period greater than at least 10. Thus, 3 *LOX* and 11 *ADH* were screened (**Figure 5**). We can observe that the contents of C6 VOCs (hexanal and 2-hexenal, (E)-) are higher when the watermelon was immature, whereas the contents of C9 VOCs (2,6-nonadienal, (E,Z); 2-nonenal, (E)-; 3,6-nonadien-1-ol, (E,Z)-; and 3-nonen-1-ol (Z)-) are higher when the watermelon was mature, indicating that the C9 alcohols and aldehydes were the main aroma metabolites of watermelon. LOX (Cla97C06G115570) and eight ADHs (Cla97C09G172570, Cla97C01G013600, Cla97C05G092340, Cla97C05G089700, Cla97C01G001290, Cla97C05G095170, Cla97C08G152400, Cla97C06G118330) in class 4 or 5 showed the same trend as the C9 VOCs. It is worth noting that the results of PCC analysis between genes and metabolites showed that five ADH genes, i.e., *Cla97C01G013600*, *Cla97C05G089700*, *Cla97C01G001290*, *Cla97C05G095170*, and *Cla97C06G118330* were associated with the metabolite 2,6-nonadienal, (E,Z); 2-nonenal, (E)-; 3,6-nonadien-1-ol, (E,Z)-; and 3-nonen-1-ol, (Z)- and had a highly positive correlation (**Supplementary Figure 5**), which suggests that these genes may play an important regulatory role in the synthesis of metabolites. In conclusion, four key C9 aldehyde/alcohol compounds to the formation of watermelon aroma are closely related to five ADHs with the same expression trend, providing new ideas and important data on the formation of watermelon aroma.

DISCUSSION

Metabolome is often thought of as a 'readout' of physiological states (Fang et al., 2019). Therefore, it is widely used in the study of plant growth and development, biological and abiotic stress, etc. The metabolome provides an essential means for the study of the nutrition and flavor of watermelon. In previous study, 18 kinds of lycopene, 3 kinds of sugars, 5 kinds of organic acids, and 16 kinds of VOCs were detected in 12 watermelon cultivars by HPLC, GC–flame-ionization detection, and GC–MS (Liu et al., 2012), which increased the understanding of the difference of metabolites in flesh of different flesh color. Although metabolome detection technology has made rapid progress, studies on metabolites in watermelon were usually focused only on a small number of metabolites, such as sucrose, glucose, and malic acid (Gao et al., 2018b; Umer et al., 2020). And the overall understanding of the metabolic profile during watermelon fruit development remains rare. Metabolic profiles of the four developmental stages showed that the levels of 321 of the 443 metabolites differed in at least one developmental stage by LC–MS analysis. Simultaneously, 23 of 74 VOCs had differences, which were analyzed by GC–MS. Such significant changes in multiple metabolites during development have also been found in fruits such as apples (Xu et al., 2020). Here, we have built a global map of the significant metabolic changes that occur during watermelon fruit development.

Flavor is the sum of taste and aroma; moreover, sweet and sour taste is an important index to measure the taste of fruit, and VOCs can cause the production of fruit aroma. The mature watermelon had better taste, so the metabolites in classes 1 and 7 were more likely to be taste related. Sugar accumulation is an important feature of watermelon development and maturation. All the nine different accumulated sugars were enriched in class 1, which was consistent with the accumulation trend of TSS by Guo et al. (2015). Furthermore, we found that most essential amino acids such as L-tryptophan, L-proline, L-histidine, and other metabolites were also enriched in these two classes, which combined with watermelon had better taste at maturity stage. The increase in pH indicates that the acidity of watermelon decreases during the development process, which may be related to the decrease of 11 organic acids obtained by our analysis. However, malic acid and citric acid were closely related to acidity formation (Gao et al., 2018b), with the former having no significant change and the latter increasing in content. Consider that the acidity may also be related to the concentration of hydrogen ions and the concentration of the buffer. Therefore, during the development of commercial watermelon fruit, the increase of acidity may result from the comprehensive action of many factors. VOCs are closely related to aroma, we screened 11 metabolites that may be more strongly associated with flavor from 23 different VOCs. Along with them, 2-hexenal, (E)- was thought to have the scent of leaves (Bianchi et al., 2020), and unsaturated C9 aliphatic aldehydes

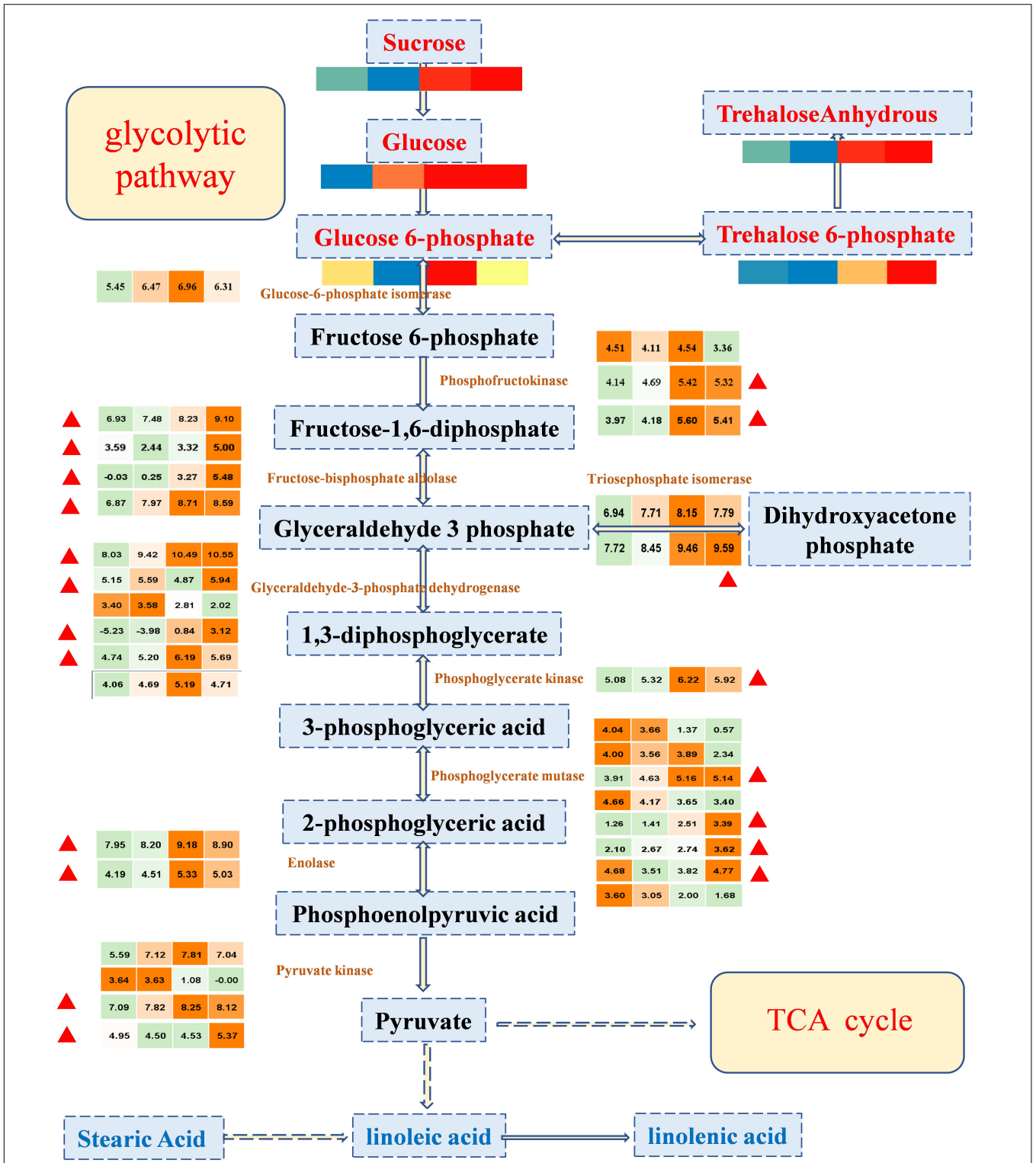

FIGURE 4 | The expression profile of important regulatory genes and the accumulation patterns of sugars in the glycolytic pathway during four important developmental periods. In the color scale, each cube metabolized the average of three replicates in the same period. The metabolite data were standardized by the logarithm of 10, and the gene was the logarithm of 2; red or orange represents a higher relative content, but blue or green represents a lower relative content, from left to right represents the average relative content of metabolites at 10, 18, 26, and 34 DAP, respectively. The red triangle represents that the enriched genes in class 4 and class 5 are in the K-means clustering analysis, showing a tendency of up-regulation.

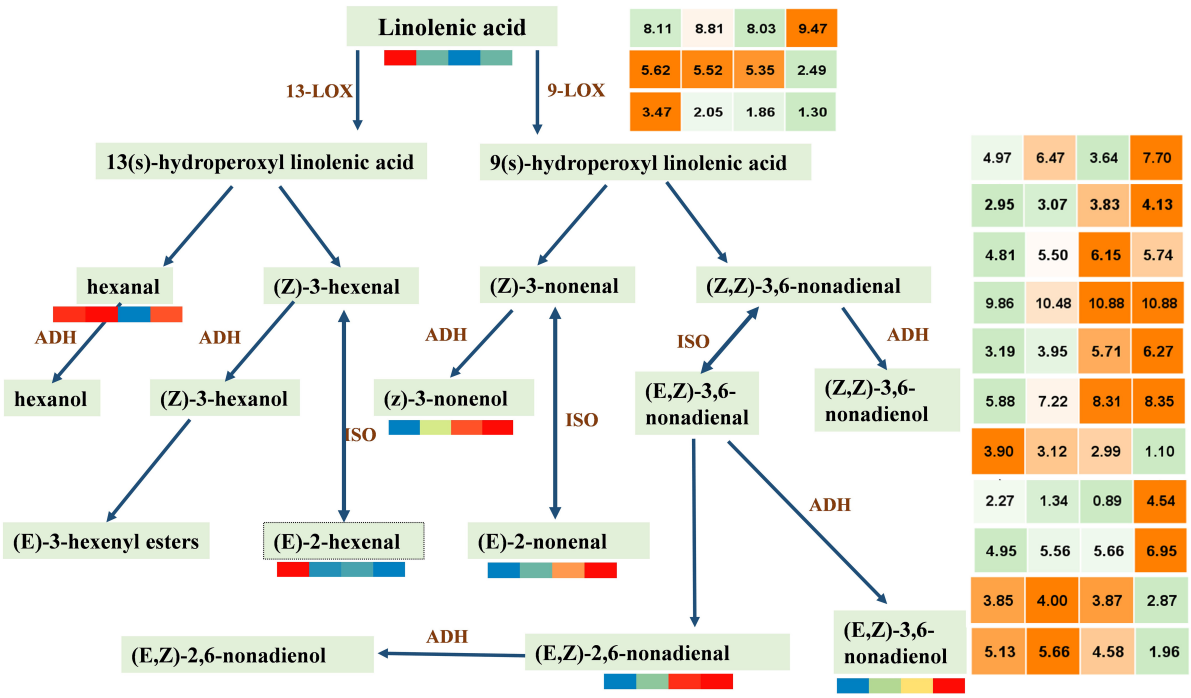

FIGURE 5 | Accumulation patterns of C6 and C9 alcohols/aldehydes synthesized by lipoxygenase pathway and expression patterns of key genes. LOX, lipoxygenase; ADH, alcohol dehydrogenase; ISO, isomerase. In the color scale, each cube metabolized the average of three replicates in the same period. The metabolite data were standardized by the logarithm of 10, and the gene was the logarithm of 2; red and orange represent a higher relative content, and blue represents a lower relative content. From left to right represents the average relative expression of metabolites at 10, 18, 26, 34 DAP, respectively. The gene expression profile information is shown in **Supplementary Table 6**.

and alcohol are considered to be the prevalent compounds in watermelon (Hatanaka et al., 1975). Besides, 5,9-undecadien-2-one,6,10- dimethyl-,(E)- and 3-buten-2-one,4-(2,6,6-trimethyl-1-cyclohexen-1-yl)- were thought to be related to the content of carotenoids (Liu et al., 2012). Our study provides a more comprehensive VOC profile of watermelon than before. Further, we could combine sensory evaluation, as in tomatoes (Tieman et al., 2017), to mine for metabolites associated with flavor, in line with consumer preferences.

Transcriptome has made progress in studies on watermelon flesh hardness, color, soluble solids, etc. (Guo et al., 2015; Zhu et al., 2017; Gao et al., 2018b; Sun et al., 2020). The transcriptional profiling analysis by K-means clustering has been applied to tomato and other crops earlier (Srivastava et al., 2010). Still, there is a lack of relevant research reports in watermelon. We obtained a total of six classes through cluster analysis, among which, the DEGs in classes 4 and 5 tended to be up-regulated. In order to confirm the reliability of the clustering results, we found that an α-galactosidase, *Cla97C04G070460*, can metabolize stachyose and raffinose into disaccharides and monosaccharides, located in class 4 (**Supplementary Figure 6A**; Guo et al., 2019). A *PSY* gene, *Cla97C01G008760*, belonging to class 5, was the first rate-limiting enzyme in the carotenoid pathway and showed a consistent expression pattern with the accumulation of carotenoids (**Supplementary Figure 6B**; Sun et al., 2018; Guo et al., 2019). Similarly, *Cla97C10G20570* is a phosphate transporter located in class 5, which is essential for carotenoid accumulation in fruits, showing a continuous accumulation pattern (**Supplementary Figure 6C**; Zhang et al., 2017). Besides this, we found the *LCYB* gene, *Cla97C04G070940*, a vital gene controlling carotenoid synthesis, had no significant difference, which was consistent with previous studies (**Supplementary Figure 6D**; Zhang et al., 2020). The results can reflect that our analysis method is effective; genes with the same trend have a regulatory effect on metabolites.

Combined transcriptome and metabolome analysis provides an important tool for the mining of metabolic networks and key genes. For example, in mulberry, the key genes regulating anthocyanins and proanthocyanidins were mined through integration analysis, and further verification showed that the abnormal expression of *bHLH3* could disrupt the homeostasis network of flavonoids and lead to differential accumulation of pigments in mulberry with different colors (Li et al., 2020). Wei et al. (2016) discovered the key genes that regulate VOCs in cucumber through integrated analysis and confirmed that functional heteromeric geranyl pyrophosphate synthase played a regulatory role in the accumulation of monoterpenes. In this study, we established the relationship between carbohydrate and glycolysis pathway genes through integration analysis, and the coexpression analysis found that five ADH genes regulate the synthesis of volatile C9 alcohols/aldehydes. ADH is an important gene involved in the conversion of alcohols and aldehydes;the expression of *CMADH3* and *CMADH12* was positively correlated with the key volatile compounds in melon

(Chen et al., 2016). Our research has provided important data support for the formation of watermelon fruit flavor. Next, we will further combine biochemical experiments to verify the function of candidate genes.

CONCLUSION

In summary, our results provide a global map of the transcriptome and metabolome during watermelon development. Five hundred seventeen metabolites were detected to be specifically accumulated at different developmental stages, and 344 metabolites were differentially accumulated, indicating that active metabolic activities occurred during development. Besides, genes with similar expression trends were identified by K-means clustering analysis. With the increase of sweetness and the decrease of acidity during the development process, watermelon tastes better, and 9 sugars with positive accumulation mode and 11 negative accumulation mode organic acids with negative accumulation mode are thought to play important roles. Eleven different VOCs, such as 5,9-undecadien-2-one,6,10-dimethyl- and 2,6-nonadienal, (E,Z)-, were positively accumulated during watermelon fruit development and were identified as key metabolites for watermelon aroma formation. Further, the integration of important genes in the glycolytic and LOX pathways provides molecular insights into the formation of flavor formation. In conclusion, we constructed a global map of transcriptome and metabolome changes during watermelon fruit flavor formation and provide valuable data resources for further studies on multi-omics linked with metabolomics.

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: NCBI BioProject, accession no: PRJNA703434.

AUTHOR CONTRIBUTIONS

WL and CG conceived and designed the experiments. WL, SZ, NH, and XL collected the watermelon accessions and participated in the material preparation. CG, WD, HZ, PY, and DY performed the experiments and analyzed the data. WL and CG wrote the manuscript. MU, MA, and MK proofread manuscripts. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the Agricultural Science and Technology Innovation Program (CAAS-ASTIP-2021-ZFRI), the China Agriculture Research System (CARS-25-03), the National Key R&D Program of China (2018YFD0100704), and the National Nature Science Foundation of China (31672178 and 31471893).

ACKNOWLEDGMENTS

We sincerely thank the editors and reviewers for their contributions. We thank the staff of Wuhan Metware Biotechnology Co., Ltd. (Wuhan, China), for their support of the determination and analysis of metabolites. We also thank Henan Key Laboratory of Fruit and Cucurbit Biology for providing transcriptome detection and analysis in our experiment.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.629361/full#supplementary-material

Supplementary Figure 1 | Hierarchical clustering analysis of all 443 metabolites.

Supplementary Figure 2 | Correlation analysis of transcriptome data in watermelonflesh at four key developmental stages.

Supplementary Figure 3 | KEGG pathway statistical analysis of differentially expressed genes (DEGs).

Supplementary Figure 4 | Gene expression profile during the development of 'Crimson' watermelon fruit.

Supplementary Figure 5 | Network representation of genes related to lipids.

Supplementary Figure 6 | The FPKM of the gene has been previously reported.

Supplementary Table 1 | The relative expression of 443 metabolites detected by UPLC-ESI-MS/MS of 12 samples.

Supplementary Table 2 | The information of differential accumulation of metabolites detected by UPLC-ESI-MS/MS.

Supplementary Table 3 | Relative expression levels of volatile organic compounds in 12 samples.

Supplementary Table 4 | Differential expression of volatile organic compounds at four development stages.

Supplementary Table 5 | The FPKM value of the gene was found in 12 samples.

Supplementary Table 6 | The Pearson correlation coefficient between lipidS and genes.

REFERENCES

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

Bianchi, G., Provenzi, L., and Rizzolo, A. (2020). Evolution of volatile compounds in 'Cuoredolce§' and 'Rugby' mini- watermelons (*Citrullus lanatus* (Thunb.) Matsumura and Nakai) in relation to ripening at harvest. *J. Sci. Food Agric.* 100, 945–952. doi: 10.1002/jsfa.10023

Chen, H., Cao, S., Jin, Y., Tang, Y., and Qi, H. (2016). The relationship between CmADHs and the diversity of volatile organic compounds of three aroma

types of melon (*Cucumis melo*). *Front. Physiol.* 7:254. doi: 10.3389/fphys.2016.00254

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

Dudareva, N., Klempien, A., Muhlemann, J. K., and Kaplan, I. (2013). Biosynthesis, function and metabolic engineering of plant volatile organic compounds. *New Phytol.* 198, 16–32. doi: 10.1111/nph.12145

Fang, C., Fernie, A. R., and Luo, J. (2019). Exploring the diversity of plant metabolism. *Trends Plant Sci.* 24, 83–98. doi: 10.1016/j.tplants.2018.09.006

Forss, D. A., Dunstone, E. A., Ramshaw, E. H., and Stark, W. (1962). The flavor of cucumbers. *J. Food Sci.* 27, 90–93. doi: 10.1111/j.1365-2621.1962.tb00064.x

Gao, L., Zhao, S., Lu, X., He, N., and Wenge, L. (2018a). 'SW', a new watermelon cultivar with a sweet and sour flavor. *HortScience* 53, 895–896. doi: 10.21273/HORTSCI12857-18

Gao, L., Zhao, S., Lu, X., He, N., Zhu, H., Dou, J., et al. (2018b). Comparative transcriptome analysis reveals key genes potentially related to soluble sugar and organic acid accumulation in watermelon. *PLoS One* 13:e0190096. doi: 10.1371/journal.pone.0190096

Guo, S., Sun, H., Zhang, H., Liu, J., Ren, Y., Gong, G., et al. (2015). Comparative transcriptome analysis of cultivated and wild watermelon during fruit development. *PLoS One* 10:e0130267. doi: 10.1371/journal.pone.0130267

Guo, S., Zhang, J., Sun, H., Salse, J., Lucas, W. J., Zhang, H., et al. (2013). The draft genome of watermelon (*Citrullus lanatus*) and resequencing of 20 diverse accessions. *Nat. Genet.* 45, 51–58. doi: 10.1038/ng.2470

Guo, S., Zhao, S., Sun, H., Wang, X., Wu, S., Lin, T., et al. (2019). Resequencing of 414 cultivated and wild watermelon accessions identifies selection for fruit quality traits. *Nat. Genet.* 51, 1616–1623. doi: 10.1038/s41588-019-0518-4

Hatanaka, A., Kajiwara, T., and Harada, T. (1975). Biosynthetic pathway of cucumber alcohol: trans-2,cis-6-nonadienol via cis-3,cis-6-nonadienal. *Phytochemistry* 14, 2589–2592. doi: 10.1016/0031-9422(75)85230-7

Kohl, M., Wiese, S., and Warscheid, B. (2011). Cytoscape: software for visualization and analysis of biological networks. *Methods Mol. Biol.* 696, 291–303. doi: 10.1007/978-1-60761-987-1_18

Kyriacou, M. C., and Rouphael, Y. (2018). Towards a new definition of quality for fresh fruits and vegetables. *Sci. Horticult.* 234, 463–469. doi: 10.1016/j.scienta.2017.09.046

Li, H., Yang, Z., Zeng, Q., Wang, S., Luo, Y., Huang, Y., et al. (2020). Abnormal expression of bHLH3 disrupts a flavonoid homeostasis network, causing differences in pigment composition among mulberry fruits. *Hortic. Res.* 7:83. doi: 10.1038/s41438-020-0302-8

Liu, C., Zhang, H., Dai, Z., Liu, X., Liu, Y., Deng, X., et al. (2012). Volatile chemical and carotenoid profiles in watermelons [*Citrullus vulgaris* (Thunb.) Schrad (Cucurbitaceae)] with different flesh colors. *Food Sci. Biotechnol.* 21, 531–541. doi: 10.1007/s10068-012-0068-3

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

Luo, J. (2015). Metabolite-based genome-wide association studies in plants. *Curr. Opin. Plant Biol.* 24, 31–38. doi: 10.1016/j.pbi.2015.01.006

Moing, A., Aharoni, A., Biais, B., Rogachev, I., Meir, S., Brodsky, L., et al. (2011). Extensive metabolic cross-talk in melon fruit revealed by spatial and developmental combinatorial metabolomics. *New Phytol.* 190, 683–696. doi: 10.1111/j.1469-8137.2010.03626.x

Mortazavi, A., Williams, B. A., Mccue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* 5:621. doi: 10.1038/nmeth.1226

Nativ, N., Hacham, Y., Hershenhorn, J., Dor, E., and Amir, R. (2017). Metabolic investigation of Phelipanche aegyptiaca reveals significant changes during developmental stages and in its different organs. *Front. Plant Sci.* 8:491. doi: 10.3389/fpls.2017.00491

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

Ramirez, J. L., Du, X., and Wallace, R. W. (2020). Investigating sensory properties of seven watermelon varieties and factors impacting refreshing perception using quantitative descriptive analysis. *Food Res. Int.* 138:109681. doi: 10.1016/j.foodres.2020.109681

Soteriou, G. A., Kyriacou, M. C., Siomos, A. S., and Gerasopoulos, D. (2014). Evolution of watermelon fruit physicochemical and phytochemical composition during ripening as affected by grafting. *Food Chem.* 165, 282–289. doi: 10.1016/j.foodchem.2014.04.120

Srivastava, A., Gupta, A., Datsenka, T., Mattoo, A., and Handa, A. (2010). Maturity and ripening-stage specific modulation of tomato (*Solanum lycopersicum*) fruit transcriptome. *GM Crops* 1, 237–249. doi: 10.4161/gmcr.1.4.13737

Sun, L., Zhang, Y., Cui, H., Zhang, L., Sha, T., Wang, C., et al. (2020). Linkage mapping and comparative transcriptome analysis of firmness in watermelon (*Citrullus lanatus*). *Front. Plant Sci.* 11:831. doi: 10.3389/fpls.2020.00831

Sun, T., Yuan, H., Cao, H., Yazdani, M., Tadmor, Y., and Li, L. (2018). Carotenoid metabolism in plants: the role of plastids. *Mol. Plant* 11, 58–74. doi: 10.1016/j.molp.2017.09.010

Tieman, D., Zhu, G., Resende, M. F. R., Lin, T., Nguyen, C., Bies, D., et al. (2017). A chemical genetic roadmap to improved tomato flavor. *Science* 355, 391–394. doi: 10.1126/science.aal1556

Umer, M. J., Bin Safdar, L., Gebremeskel, H., Zhao, S., Yuan, P., Zhu, H., et al. (2020). Identification of key gene networks controlling organic acid and sugar metabolism during watermelon fruit development by integrating metabolic phenotypes and gene expression profiles. *Hortic. Res.* 7:193. doi: 10.1038/s41438-020-00416-8

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

Wei, G., Tian, P., Zhang, F., Qin, H., Miao, H., Chen, Q., et al. (2016). Integrative analyses of nontargeted volatile profiling and transcriptome data provide molecular insight into VOC diversity in cucumber plants (*Cucumis sativus*). *Plant Physiol.* 172, 603–618. doi: 10.1104/pp.16.01051

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

Yuan, H., Zeng, X., Shi, J., Xu, Q., Wang, Y., Jabu, D., et al. (2018). Time-course comparative metabolite profiling under osmotic stress in tolerant and sensitive Tibetan Hulless Barley. *BioMed Res. Int.* 2018:9415409. doi: 10.1155/2018/9415409

Zhang, J., Guo, S., Ren, Y., Zhang, H., Gong, G., Zhou, M., et al. (2017). High−level expression of a novel chromoplast phosphate transporter ClPHT4;2 is required for flesh color development in watermelon. *New Phytol.* 213, 1208–1221. doi: 10.1111/nph.14257

Zhang, J., Sun, H., Guo, S., Ren, Y., and Xu, Y. (2020). Decreased protein abundance of lycopene β-cyclase contributes to red flesh in domesticated watermelon. *Plant Physiol.* 183, 1171–1183. doi: 10.1104/pp.19.01409

Zhu, G., Wang, S., Huang, Z., Zhang, S., Liao, Q., Zhang, C., et al. (2018). Rewiring of the fruit metabolome in tomato breeding. *Cell* 172, 249–261.e12. doi: 10.1016/j.cell.2017.12.019

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

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 19 May 2021
doi: 10.3389/fpls.2021.615749

Genetic Analysis of Methyl Anthranilate, Mesifurane, Linalool, and Other Flavor Compounds in Cultivated Strawberry (*Fragaria* × *ananassa*)

Christopher R. Barbey[1], Maxwell H. Hogshead[2], Benjamin Harrison[2], Anne E. Schwartz[2], Sujeet Verma[1], Youngjae Oh[1], Seonghee Lee[1], Kevin M. Folta[2] and Vance M. Whitaker[1]*

[1] Gulf Coast Research and Education Center, University of Florida, Wimauma, FL, United States, [2] Horticultural Sciences Department, University of Florida, Gainesville, FL, United States

OPEN ACCESS

Edited by:
María José Jordán, Murcian Institute for Agrarian and Food Research and Development (IMIDA), Spain

Reviewed by:
José L. Rambla, Polytechnic University of Valencia, Spain
Juan Muñoz-Blanco, University of Córdoba, Spain

***Correspondence:**
Christopher R. Barbey
cbarbey@ufl.edu

Specialty section:
This article was submitted to Plant Metabolism and Chemodiversity, a section of the journal Frontiers in Plant Science

Received: 20 October 2020
Accepted: 30 March 2021
Published: 19 May 2021

Citation:
Barbey CR, Hogshead MH, Harrison B, Schwartz AE, Verma S, Oh Y, Lee S, Folta KM and Whitaker VM (2021) Genetic Analysis of Methyl Anthranilate, Mesifurane, Linalool, and Other Flavor Compounds in Cultivated Strawberry (Fragaria × ananassa). Front. Plant Sci. 12:615749. doi: 10.3389/fpls.2021.615749

The cultivated strawberry (*Fragaria* × *ananassa*) is an economically important fruit crop that is intensively bred for improved sensory qualities. The diversity of fruit flavors and aromas in strawberry results mainly from the interactions of sugars, acids, and volatile organic compounds (VOCs) that are derived from diverse biochemical pathways influenced by the expression of many genes. This study integrates multiomic analyses to identify QTL and candidate genes for multiple aroma compounds in a complex strawberry breeding population. Novel fruit volatile QTL was discovered for methyl anthranilate, methyl 2-hexenoate, methyl 2-methylbutyrate, mesifurane, and a shared QTL on Chr 3 was found for nine monoterpene and sesquiterpene compounds, including linalool, 3-carene, β-phellandrene, α-limonene, linalool oxide, nerolidol, α-caryophellene, α-farnesene, and β-farnesene. Fruit transcriptomes from a subset of 64 individuals were used to support candidate gene identification. For methyl esters including the grape-like methyl anthranilate, a novel *ANTHANILIC ACID METHYL TRANSFERASE*–like gene was identified. Two mesifurane QTL correspond with the known biosynthesis gene *O-METHYL TRANSFERASE 1* and a novel *FURANEOL GLUCOSYLTRANSFERASE*. The shared terpene QTL contains multiple fruit-expressed terpenoid pathway-related genes including *NEROLIDOL SYNTHASE 1* (*FanNES1*). The abundance of linalool and other monoterpenes is partially governed by a co-segregating expression-QTL (eQTL) for *FanNES1* transcript variation, and there is additional evidence for quantitative effects from other terpenoid-pathway genes in this narrow genomic region. These QTLs present new opportunities in breeding for improved flavor in commercial strawberry.

Keywords: aromas, eQTL analysis, fruit volatiles, QTL analysis, terpenes, transcriptomics

INTRODUCTION

The dessert strawberry (*Fragaria* × *ananassa*) is a widely celebrated fruit with increasing consumption. For decades, consumers have reported the desire for improved flavor in commercial strawberry (Fletcher, 1917; Chambers, 2013). The aroma intensity of modern cultivars is lower than in wild strawberries (Ulrich and Olbricht, 2014), and breeding efforts seek to reclaim these qualities. Today, flavor and aroma are central priorities of strawberry breeding programs (Faedi et al., 2002; Whitaker et al., 2011; Vandendriessche et al., 2013). However, breeders face a significant challenge in the recapture and consolidation of genetics contributing to favorable flavors and aromas. Genetic and genomic analysis has been used to identify these elements and contribute to the breeding of new cultivars with improved sensory qualities.

Strawberry flavor and aroma are dictated by several factors, including sugars and acids, but it is the trace volatile organic compounds (VOCs) that shape the sensory experience (Bood and Zabetakis, 2002). VOCs, represented broadly as esters, alcohols, terpenoids, furans, and lactones, are a substantial portion of the fruit secondary metabolome and contribute to aroma, flavor, disease resistance, pest resistance, and overall fruit quality (Ulrich et al., 1997; Arroyo et al., 2007). Various studies have helped to identify human preferences for individual strawberry aroma and flavor compounds (Larsen and Poll, 1992; Schieberle and Hofmann, 1997; Ulrich et al., 1997; Schwieterman et al., 2014). Of hundreds of strawberry VOCs, these studies agree on fewer than 10 that clearly influence human preference (Schwieterman et al., 2014). Introgressing important compounds into commercially viable cultivars has been aided by efforts in volatilomic QTL detection (Urrutia et al., 2017), multiomic identification of VOC candidate genes (Chambers et al., 2014; Pillet et al., 2017), integration of sensory and consumer preference data (Sánchez-Sevilla et al., 2014; Schwieterman et al., 2014), and ultimately introgression of genes via marker-assisted selection (Eggink et al., 2014; Folta and Klee, 2016; Rambla et al., 2017).

Only a few strawberry genes controlling desirable aroma compounds have been identified with confidence. These include biosynthesis genes for linalool (Aharoni et al., 2004), mesifurane (Wein et al., 2002), γ-decalactone (Chambers et al., 2014; Sánchez-Sevilla et al., 2014), and methyl anthranilate (Pillet et al., 2017). The gene *ANTHRANILIC ACID METHYL TRANSFERASE* (*FanAAMT*), located on octoploid chromosome group 4, is a necessary-but-not-sufficient gene for catalyzing the methylation of anthranilate into the grape-like aroma compound methyl anthranilate (Pillet et al., 2017). Methyl anthranilate production has been long regarded as a complex trait, governed by multiple genes and strong environmental influences. Methyl anthranilate is produced abundantly in the fruit of the diploid strawberry sp. *Fragaria vesca*, but it is reported in only a few octoploid varieties including "Mara des Bois" and "Mieze Schindler" (Ulrich and Olbricht, 2016). For terpenoid biosynthesis, strawberry *NEROLIDOL SYNTHASE 1 (NES1)* was identified by comparing diploid and octoploid species, which are enriched respectively for nerolidol or linalool. A truncated plastid-targeted signal in the octoploid *FanNES1* gene retargets the enzyme to the cytosol, where there is abundant precursor for linalool biosynthesis (Aharoni et al., 2004). Recent reports have complicated this story somewhat, as three octoploid *F. virginiana* lines, hexaploid *F. moschata*, and some diploid strawberries produce linalool without this truncation (Ulrich and Olbricht, 2013). In mesifurane biosynthesis, the gene *O-METHYL TRANSFERASE 1* (*FanOMT1*) catalyzes the methylation of furaneol to create mesifurane (Wein et al., 2002; Zorrilla-Fontanesi et al., 2012). Mesifurane abundance is affected by a common *FanOMT1* promoter loss-of-function allele, which both eliminates gene expression and mesifurane production. Only one copy of the competent *FanOMT1* allele is reportedly sufficient for robust production; however, a lack of production is sometimes observed even in the homozygous positive state (Cruz-Rus et al., 2017). An octoploid gene encoding *QUINONE REDUCTASE* (*FanQR*) can produce furaneol *in vitro*; however, no natural variants of this gene have been established, which vary mesifurane levels *in vivo* (Raab et al., 2006). Similarly, glucosylation of both furaneol and mesifurane are known to occur in strawberry; however, genetic variation has not been established for this step. Several furaneol glucosyltransferases have been cloned and characterized *in vitro* from *F.* × *ananassa* (Song et al., 2016; Yamada et al., 2019).

This research integrates high-density genotyping and non-targeted fruit volatile metabolomics from eight pedigree-connected octoploid crosses ($n = 213$) (**Supplementary Figure 1**). Fruit transcriptomes from a subset of individuals ($n = 61$) were used to identify fruit-expressed candidate genes within QTL regions (**Supplementary Figure 1**). Three maps were utilized independently in this analysis, as less than one-third of octoploid subgenome-specific markers are incorporated in any single octoploid genetic map (van Dijk et al., 2014; Anciro et al., 2018). These are the "Holiday" × "Korona" (van Dijk et al., 2014) and FL_08-10 × 12.115-10 (Verma et al., 2017) genetic maps, and the *F. vesca* physical map. The correspondence of "Holiday" × "Korona" linkage groups to the recent octoploid "Camarosa" reference genome (Edger et al., 2019) helped specify the subgenomic identity of QTL (Hardigan et al., 2020). To correspond QTL markers to specific candidate gene regions, marker nucleotide sequences from the IStraw35 SNP genotyping platform were aligned by sequence to the octoploid genome. However, the very high sequence identity between homoeologous chromosomes limited the specificity of this approach. Evidence from all of these resources were integrated to specify candidate gene regions.

Fruit transcriptomes were used to identify expressed genes within QTL regions and to associate trait/transcript levels. Genotypic data were associated with fruit transcriptomics data via expression-QTL (eQTL) analysis. Transcript eQTL analysis identifies genetic variants associated with heritable transcript level variation. Transcript eQTL often correspond to the locus of the originating gene (*cis*-eQTL), and often signify gene promoter mutation or gene presence/absence variation. In cases where a trait is governed by simple genetic control of transcript levels of a causal gene, an eQTL should be detected, which co-segregates with trait QTL markers. This approach can help specify the casual mechanisms behind trait

QTL. Previous eQTL analyses in these same fruit RNA-seq populations identified hundreds of fruit eQTL, the vast majority of which were proximal to the originating gene locus (Barbey C. et al., 2020). Global transcriptomes from tissues throughout the octoploid cultivar "Camarosa" were used to correlate candidate genes and transcript abundance with ripening-associated volatile biosynthesis (Sánchez-Sevilla et al., 2017).

MATERIALS AND METHODS

Plant Materials

Eight controlled crosses were made among octoploid cultivars and also elite breeding lines from the University of Florida, totaling 213 individuals including nine unique parents. These were "Florida Elyana" × "Mara de Bois" (population 10.113; $n = 13$), "Mara des Bois" × "Florida Radiance" (population 13.75; $n = 21$), "Strawberry Festival" × "Winter Dawn" (population 13.76; $n = 20$), 12.115-10 × 12.121-5 (population 15.89; $n = 25$), 12.22-10 × 12.115-10 (population 15.91; $n = 19$), 12.115-10 × 12.74-39 (population 15.93; $n = 60$), "Florida Elyana" × 12.115-10 (population 16.11; $n = 26$), and "Mara des Bois" × "Mara des Bois" (population 16.85; $n = 20$). Seedlings were clonally propagated by runners in a summer nursery to generate multiple plants (clonal replicates), and two to four plants representing original seedlings were established in single plots in the fruiting field (Barbey C. et al., 2020).

Field Collection

All fruits were harvested from a field maintained under commercial growing practices during winter growing seasons at the Gulf Coast Research and Education Center (GCREC) in Wimauma, Florida. Fruit from populations 15.89, 15.91, 15.93 were harvested on February 19, March 15, and April 14, 2016. Populations 16.11 and 16.85 were harvested on February 16 and March 2, 2017. Populations 13.75 and 13.76 were harvested throughout the winter of 2014. Population 10.133 was sampled on January 20, February 11, February 25, and March 18, 2011 (Chambers, 2013). Harvest days were selected based on dry weather and moderate temperature, both on the day of harvest and for several days preceding harvest to maximize volatile production and capture. Three mature fruit per genotype at each harvest timepoint were cleaned, collected into a single sample bag, crushed, and immediately flash frozen in liquid nitrogen, in the field. Samples were transported to Gainesville, Florida, and maintained at −80°C.

Sample Processing and Preparation

Crushed frozen fruit samples were equilibrated from −80°C to liquid nitrogen temperature before being pureed in an electric blender. Fine frozen puree was collected into a 50-ml sterile collection tube and stored at −80°C. For volatile sample processing, 3 g of frozen puree from three combined fruits per harvest timepoint were aliquoted into two technical replicate 20-ml headspace vials and combined with 3 ml of 35% NaCl solution containing 1 ppm of 3-hexanone as an internal standard. Prepared vials were stored at −80°C, thawed at room temperature, and vortexed prior to GC-MS analysis.

Volatile Metabolomic Profiling and Analysis

Samples were equilibrated to 40°C for 30 min in a 40°C heated chamber. A 2-cm tri-phase SPME fiber (50/30 μm DVB/Carboxen/PDMS, Supelco, Bellefonte, PA, United States) was exposed to the headspace for 30 min at 40°C for volatile collection and concentration. The fiber was then injected into an Agilent 6890 GC (for 5 min at 250°C for desorption of volatiles. Inlet temperatures were maintained at 250°C, ionizing sources at 230°C, and transfer line temperatures at 280°C. The separation was performed via DB5ms capillary column (60 m × 250 μm × 1.00 μm) (J&W, distributed by Agilent Technologies) at a constant flow (He: 1.5 ml per min). The initial oven temperature was maintained at 40°C for 30 s, followed by a 4°C per min increase to a final temperature of 230°C, then to 260°C at 100°C per min, with a final hold time of 10 min. Data were collected using the Chemstation G1701 AA software (Hewlett-Packard, Palo Alto, CA, United States).

Chromatograms were processed using the Metalign metabolomics preprocessing software package (Lommen and Kools, 2012). Baseline and noise corrections were performed using a peak slope factor of 1 × noise, and a peak threshold factor of 2 × noise. Autoscaling and iterative pre-alignment options were not selected. A maximum shift of 100 scans before peak identification and 200 scans after peak identification was used. In later validation steps, these search tolerances were determined to be sufficiently inclusive while also limiting to false positives. The MSClust software package was then used for statistical clustering of ions based on retention time and co-variance across the population using default parameters (Tikunov et al., 2012). Clusters were batch queried against the NIST08 reference database using Chemstation G1701 AA software (Hewlett-Packard, Palo Alto, CA, United States). Library search outputs were parsed using a custom Perl script prior to multivariate analysis. Chromatograms were batched by GC-MS sampling year to mitigate ion misalignments caused by system-dependent retention time shifts. VOCs from different seasonal datasets were consolidated manually based on elution order, NIST identification, and rerunning of sample standards. VOC relative abundances between seasons were normalized based on the relative abundance of the 1-ppm 3-hexanone internal standard. Internal standard renormalization was not performed on within-season data as technical variation was low and the spike-in tended to introduce more variation than it resolved, which is a known issue in non-targeted analyses (Wehrens et al., 2016). All within-season technical and biological replicate VOC relative abundances were averaged.

Genotyping of Flavor and Aroma Populations

Individuals from all populations were genotyped using the IStraw90 (Bassil et al., 2015) platform, except populations 16.11 and 16.89, which were genotyped using the IStraw35 platform

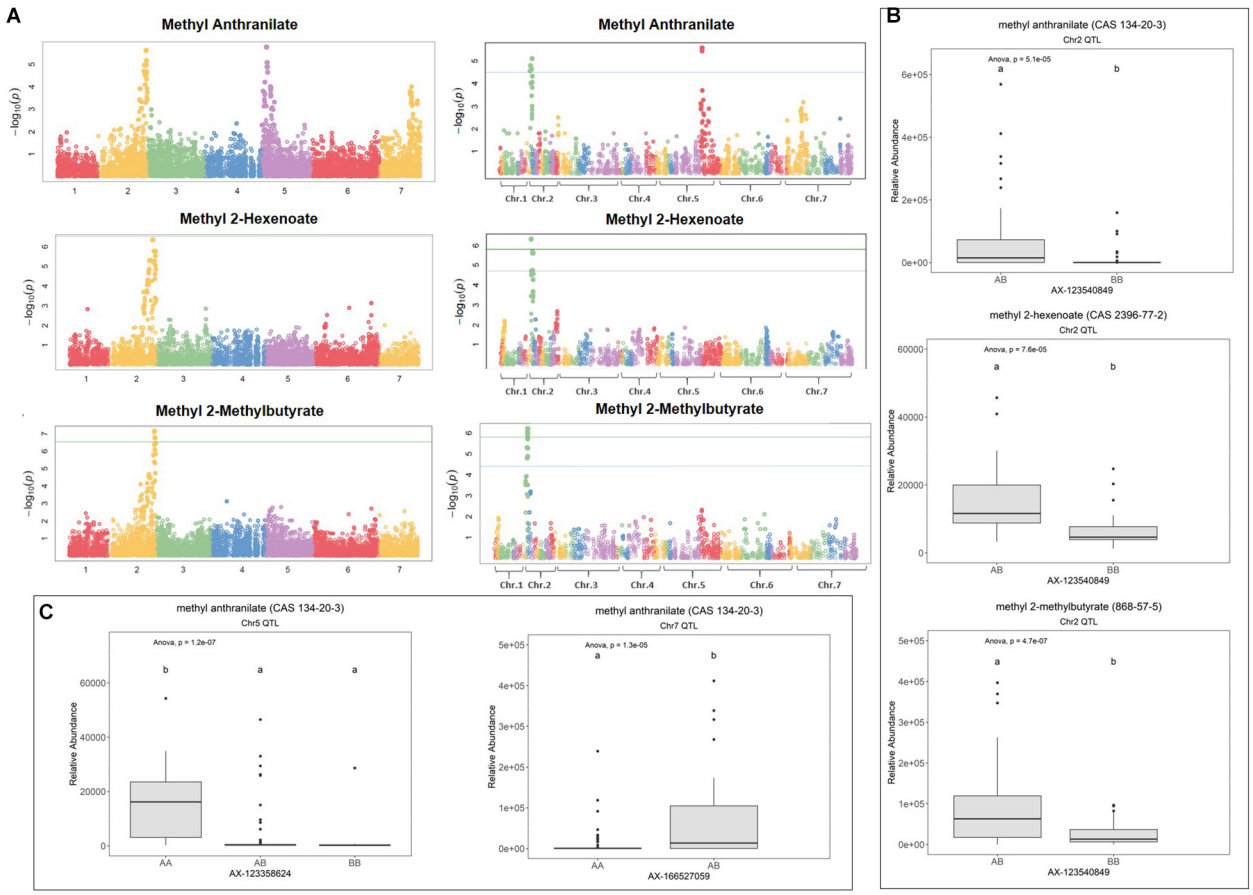

FIGURE 1 | Methyl ester QTL in strawberry fruit. **(A)** QTL Manhattan plots for methyl anthranilate, methyl 2-hexenoate, and methyl 2-methylbutyrate using the *F. vesca* physical map (left column) and the FL_08-10 × 12.115-10 genetic map (right column). **(B)** Effect sizes of a shared methyl ester QTL marker on Chr 2. Lowercase letters indicate statistically significance mean differences at $p < 0.05$ (ANOVA). **(C)** Effect sizes for the methyl anthranilate QTL on Chr 5 and the putative Chr 7 QTL.

(Verma et al., 2017). All parents and 204 progenies were selected for genotyping based on the segregation of desirable fruit volatiles. Sequence variants belonging to the poly high resolution (PHR) and no minor homozygote (NMH) marker classes were included for association mapping. Mono high resolution (MHR), off-target variant (OTV), call rate below threshold (CRBT), and other marker quality classes, were discarded and not used for mapping. Individual marker calls inconsistent with Mendelian inheritance from parental lines were removed.

Fruit Transcriptome Assembly and Analysis

Mature fruits from 61 parents and progeny from the biparental populations 10.113, 13.75, and 13.76 were sequenced via Illumina paired-end RNA-seq (average 65 million, 2 × 100-bp reads) and used for transcript eQTL analysis via the same samples and methodology reported for R-genes (Barbey et al., 2019) and other high-value fruit transcripts (Barbey C. et al., 2020). Briefly, RNA-seq reads were assembled based on the *Fragaria* × *ananassa* octoploid "Camarosa" annotated genome, with reads mapping equally well to multiple loci discarded from the analysis. Separately, raw RNA-seq reads from the "Camarosa" strawberry gene expression atlas study (Sánchez-Sevilla et al., 2017) were assembled via the same previously reported methodology and represent the average of three biological replicates. Transcript abundances were calculated in transcripts per million (TPM). Fruit eQTL analysis was performed using the mixed linear model method implemented in GAPIT v3 (Tang et al., 2016) as described in Barbey C. et al. (2020).

Genetic Association of Fruit Volatiles

Relative volatile abundance values were rescaled using the Box–Cox transformation algorithm (Box and Cox, 1964) performed in R (R. Development Core Team, 2014) using R-studio (Racine, 2011) prior to genetic analysis. GWAS on fruit volatiles was performed using the mixed linear model method implemented in GAPIT v3 (Tang et al., 2016) in R, using marker positions oriented to the *F. vesca* diploid physical map. Significantly associated volatiles were then reanalyzed in GAPIT using the "Holiday" × "Korona" and FL_08-10 × 12.115-10 genetic maps. Metabolomic associations were evaluated for significance based

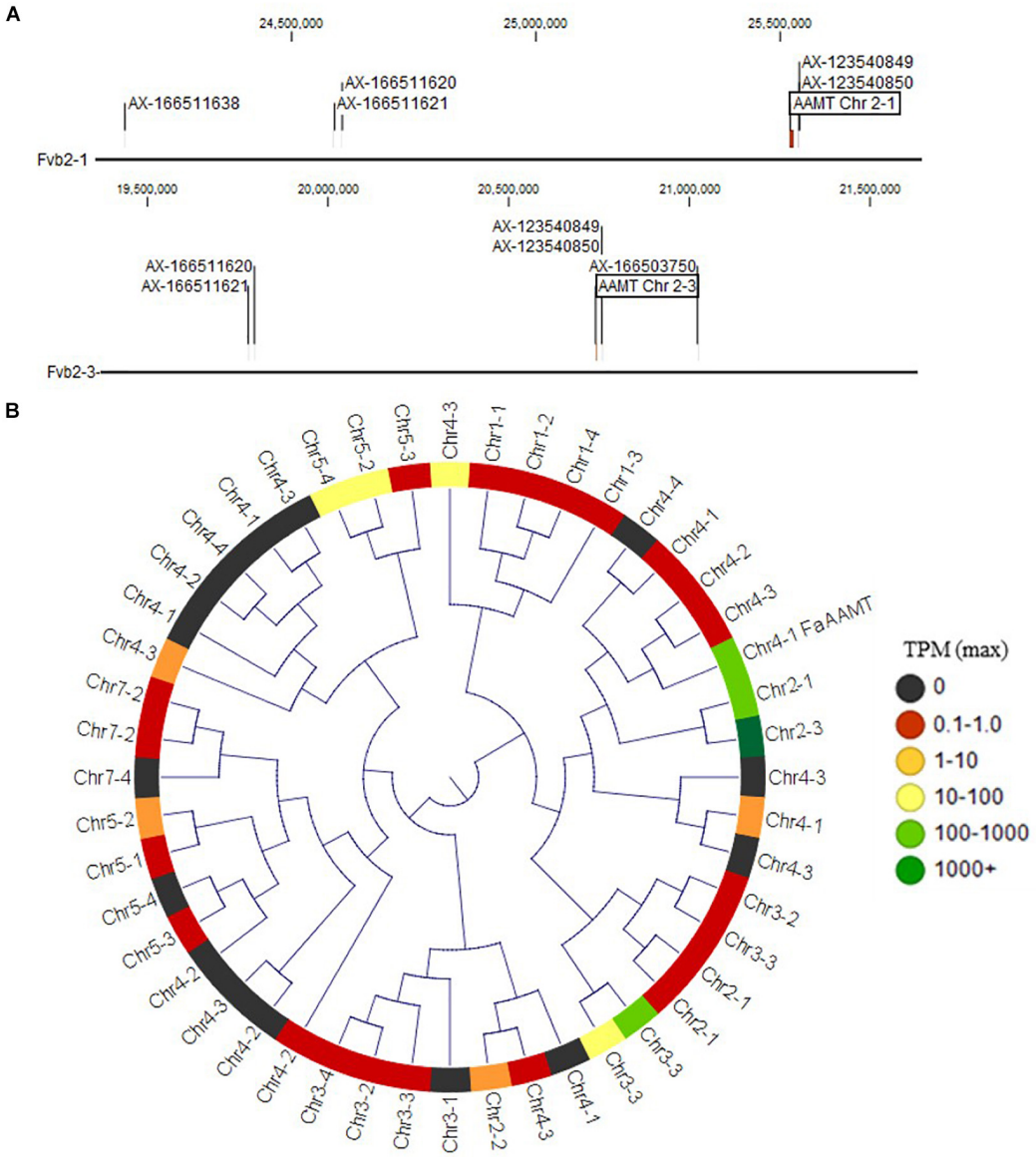

FIGURE 2 | Methyl ester and methyl anthranilate candidate genes. **(A)** The Chr 2 methyl ester QTL markers correspond to two homoeologous physical regions containing anthranilic acid methyl transferase-like (*AAMT-like*) genes on Chr 2-1 (top) and Chr 2-3 (bottom). **(B)** *AAMT-like* deduced proteins in the "Camarosa" genome are shown in a neighbor-joining cladogram, with transcript abundance heatmaps representing the highest TPM detected among the fruit transcriptomes. The Chr 2-1 and Chr 2-3 AAMT-like candidates are highly identical to the published AAMT-like "Camarosa" homolog (Chr 4-1 *FanAAMT*) and are highly abundant transcripts in the fruit.

on the presence of multiple co-locating markers of p-value < 0.05 after FDR multiple comparisons correction (Benjamini and Hochberg, 1995). Narrow-sense heritability (h^2) estimates were derived from GAPIT v3, while single-marker analysis was performed via ANOVA in R to investigate allelic effects.

Analysis of Candidate Genes

All gene models in the "Camarosa" genome were analyzed with the BLAST2GO pipeline and the Pfam protein domain database. Genes with significant homology to known volatile biosynthesis genes including *FanOMT* and *FanAAMT* were collected from the "Camarosa" genome using BLAST with inclusive criteria. This process was replicated for candidate genes including anthranilate synthase alpha subunit (*FanAS-α*) and others not presented in this analysis. Deduced protein sequences from transcripts were aligned using the slow progressive alignment algorithm in the CLC Genomics Workbench 11 (Gap Open cost = 10; Gap Extension = 1). Tree construction was performed using the neighbor joining method with Jukes–Cantor distance measuring with 1,000 bootstrapping replicates. Fruit transcript heatmaps were added to the cladogram to show the maximum transcript level detected among the 61 fruit transcriptomes.

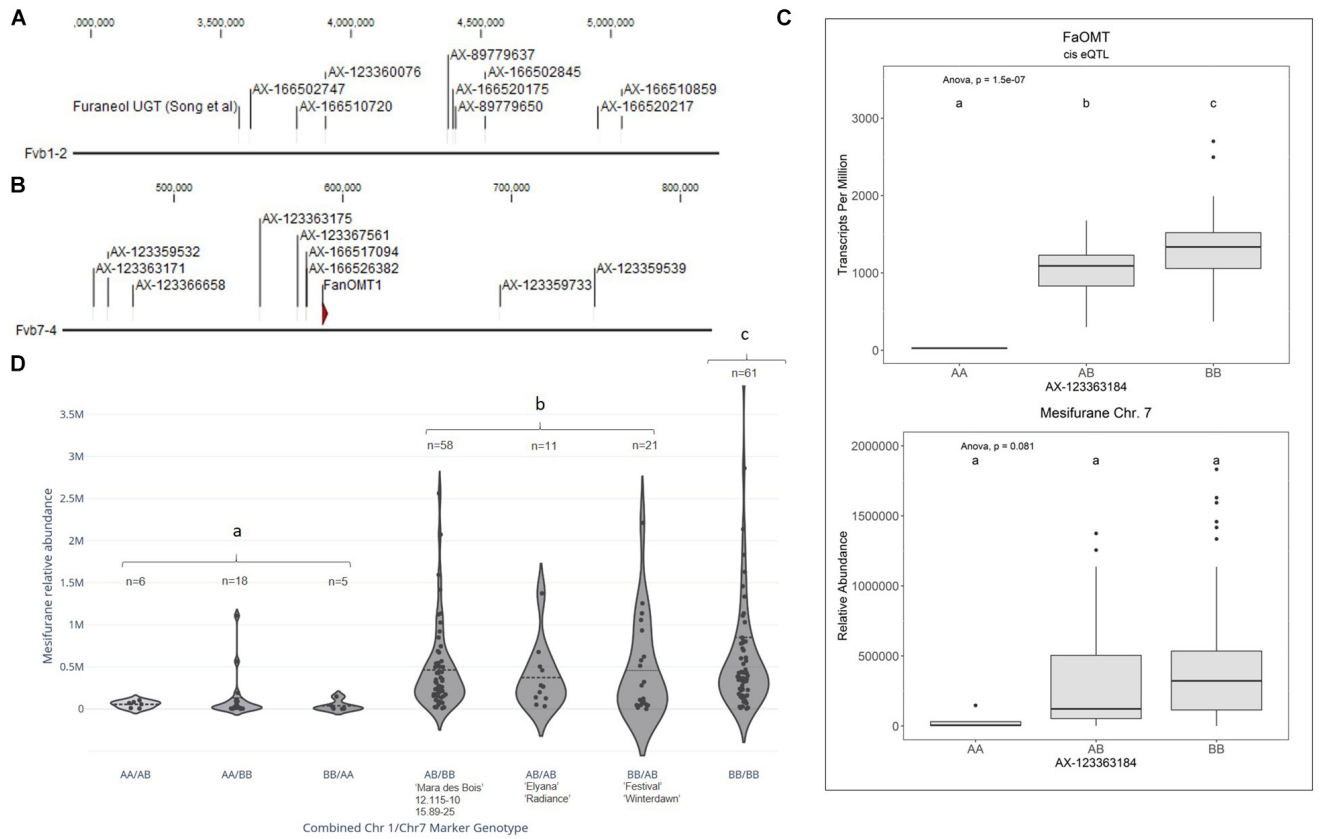

FIGURE 3 | Mesifurane QTL and candidate genes. **(A)** The Chr 1 mesifurane QTL markers correspond to a region containing a characterized furaneol UGT homolog. **(B)** The Chr 7 mesifurane QTL markers correspond to a region containing the published mesifurane *FanOMT1* biosynthesis gene. **(C)** A *cis*-eQTL was detected for *FanOMT1* (top), a gene known to be transcriptionally variable due to a common allele. The *FanOMT1* eQTL co-segregates with mesifurane (bottom) but is confounded by additional factors. Lowercase letters indicate statistically significance mean differences at $p < 0.05$ (ANOVA). **(D)** Allelic combinations at the Chr 1 and Chr 7 markers demonstrate an epistatic effect between the two loci. At least one competent allele at each loci is required for robust production of mesifurane (AA/BB vs. AB/BB Tukey HSD $p = 2.4e{-}7$), while the double-homozygous allelic state produces statistically elevated mesifurane levels (BB/BB vs. AB/AB, AB/BB, and BB/AB ANOVA) [$F(1,149) = 4.38$, $p = 0.038$]. The Chr 1/Chr 7 genotypes in parental lines is shown with the *n* of each allelic category. Dotted lines represent the population means. Lowercase letters indicate statistical significance mean differences at $p < 0.05$ (ANOVA).

Genes putatively belonging to published volatile biosynthesis gene families were selected for fruit transcript eQTL analysis, using methods described previously (Barbey et al., 2019; Barbey C. et al., 2020). The 200 genes surrounding the most-correlated volatile QTL markers were also analyzed for eQTL and compared for co-segregation with volatile QTL.

High-Resolution Melting Marker Test

For the marker test of mesifurane, two validation crosses were created consisting of "Florida Beauty" × 15.89-25 (population 18.50, $n = 27$) and 15.34-82 × 15.89-25 (population 18.51, $n = 44$). Total genomic DNA was extracted using the simplified cetyltrimethylammonium bromide (CTAB) method described by Noh et al. (2018) with minor modifications. To develop HRM markers for mesifurane, two probes were selected (AX-166520175 and AX-166502845) that were highly associated with mesifurane on Chr 1 QTL. The primers 5′-CCCTTGGCATCAATATTTGTGAAT-3′ and 5′-GAACTCC ATTAGAAATCAAGTTATCA GC-3′ were designed for AX-16 6520175, and the 5′-CTGATCCTGCTTCAAGTACAAG-3′ and 5′-TCAATGAAGACACTTGATCGAC-3′ were designed for AX-166502845 using IDT's PrimerQuest Software (San Jose, CA, United States). PCR amplifications were performed in a 5-µl reaction containing 2 × AccuStart™ II PCR ToughMix® (Quantabio, MA, United States), 1 × LC Green® Plus melting dye (BioFire, UT, United States), 0.5 µM of each HRM primer sets and 1 µl of DNA. The PCR and HRM analysis were performed in a LightCycler® 480 system II (Roche Life Science, Germany) using a program consisting of an initial denaturation at 95°C for 5 min, 45 cycles of denaturation at 95°C for 10 s, annealing at 62°C for 10 s, and extension at 72°C for 20 s. After PCR amplification, the samples were heated to 95°C for 1 min and cooled to 40°C for 1 min. Melting curves were obtained by melting over the desired range (60–95°C) at a rate of 50 acquisitions per 1°C. Melting data were analyzed using the Melt Curve Genotyping and Gene Scanning Software (Roche Life Science, Germany). Analysis of HRM variants was based on differences in the shape of the melting curves and in T_m values.

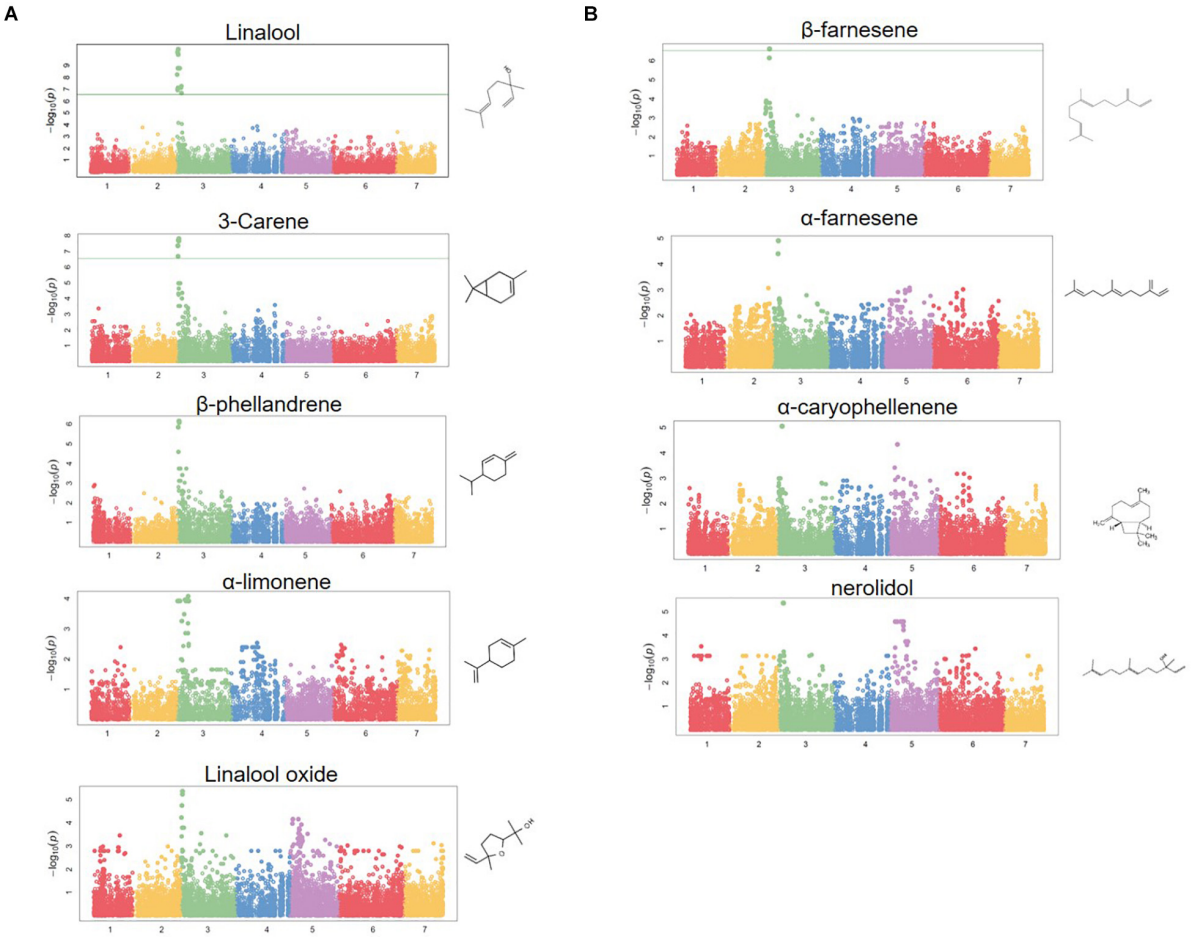

FIGURE 4 | Terpene volatile QTL in strawberry fruit. **(A)** QTL Manhattan plots for the monoterpenes linalool (r^2 = 0.15, p = 4.2e–11), 3-carene (r^2 = 0.09, p = 1.2e–8), β-phellandrene (r^2 = 0.21, p = 7.0e–7), α-limonene (r^2 = 0.06, p = 7.0e–7), linalool oxide (r^2 = 0.08, p = 4.4e–6). **(B)** The sesquiterpenes β-farnesene (r^2 = 0.12, p = 2.4e–7), α-farnesene (r^2 = 0.16, p = 1.6e–6), α-caryophellene (r^2 = 0.10, p = 8.9e–6), and nerolidol (r^2 = 0.09, p = 4.3e–6). Chemical structures are shown for each compound.

RESULTS

Strawberry Fruit Flavor and Aroma QTL

Volatile aroma QTL were discovered for methyl anthranilate (CAS 134-20-3), methyl 2-methylbutyrate (CAS 868-57-5), methyl 2-hexenoate (CAS 2396-77-2) (**Figures 1**, **2**), mesifurane (CAS 4077-47-8) (**Figure 3**), and nine mono- and sesquiterpene compounds (**Figures 4**, **5**). These terpenes include linalool (CAS 78-70-6), 3-carene (CAS 498-15-7), β-phellandrene (CAS 555-10-2), α-limonene (CAS 5989-27-5), linalool oxide (CAS 60047-17-8), nerolidol (CAS 7212-44-4), α-caryophellene (CAS 6753-98-6), α-farnesene (CAS 502-61-4), and β-farnesene (CAS 77129-48-7). Many of these compounds are consensus determinants of preferred strawberry flavor in human sensory trials, including grape-like methyl anthranilate (Ulrich et al., 1997), fruity-sweet methyl 2-methylbutanoate (Schieberle and Hofmann, 1997), fruity linalool (Larsen and Poll, 1992; Schieberle and Hofmann, 1997; Ulrich et al., 1997; Schwieterman, 2013), and sherry/caramel-like mesifurane (Larsen and Poll, 1992; Schieberle and Hofmann, 1997; Ulrich et al., 1997; Schwieterman, 2013). Population-wide fruit transcript-level data for all candidate genes in the following analysis are provided in **Supplementary Table 1**, and transcript levels throughout the "Camarosa" plant are provided in **Supplementary Table 2**.

Methyl Anthranilate and Methyl Ester QTL and Candidate Genes

Methyl anthranilate (h^2 = 0.59) QTL were identified on octoploid linkage groups (LGs) 2A and 5A of the FL_08-10 × 12.115-10 map (**Figure 1A** and **Table 1**). The LG 2A QTL is shared with methyl 2-hexenoate (h^2 = 0.43) and methyl 2-methylbutyrate (h^2 = 0.79) (**Figure 1A** and **Table 1**). This shared methyl ester QTL accounts for 11.7% of methyl anthranilate variance (p = 5.1e−5), 22.8% of methyl 2-hexenoate variance (p = 7.6e−5), and 18.1% of methyl 2-methylbutyrate variance (p = 4.7e−7) (AX-123540849, single-marker analysis) (**Figure 1B**). The QTL on LG 5D explains 19.7%

TABLE 1 | Methyl anthranilate QTL and candidate gene positions.

Chr 2 QTL	*Methyl anthranilate p*-value	Methyl 2-hexenoate p-value	Methyl 2-methylbutyrate *p*-value	Holiday × Korona map		14.95 map		"Camarosa" genome physical position			
				LG	Position	LG	Position	Chr 2-1	Chr 2-2	Chr 2-3	Chr 2-4
AAMT 2-1 candidate	–	–	–	–	–	–	–	25,519,186	–	–	–
AAMT 2-3 candidate	–	–	–	–	–	–	–	–	–	3,328,115	–
AX-123540849	2.46E−06	1.72E−06	7.08E−08	–	–	2A	9.799	25,537,654	–	3,313,464	25,836,027
AX-166503750	6.82E−06	1.72E−06	8.49E−07	–	–	2A	7.05	–	–	3,048,471	26,031,849
AX-123540850	1.07E−05	2.47E−05	2.99E−04	–	–	2A	10.409	25,537,654	–	3,313,464	25,836,027
AX-123360428	1.09E−05	6.36E−06	6.74E−05	–	–	–	–	–	–	–	25,369,790
AX-166511620	1.09E−05	6.36E−06	6.74E−05	2A	78.514	2A	1.851	24,602,373	–	4,275,047	25,558,703
AX-166511621	1.09E−05	6.36E−06	6.74E−05	–	–	–	–	24,586,556	–	4,292,083	25,573,913
AX-166503618	2.05E−05	4.50E−07	2.23E−04	–	–	2A	0.304	–	–	–	–
AX-166511638	2.05E−05	4.50E−07	2.23E−04	–	–			24,157,920	–	–	–
AX-166511640	2.05E−05	4.50E−07	2.23E−04	–	–	2A	0.304	–	–	–	–
AX-166511795	2.73E−05	2.95E−06	4.88E−05	–	–	2A	9.647	–	–	–	26,103,640
Chr 5 QTL	**Methyl anthranilate *p*-value**			**LG**	**Position**	**LG**	**Position**	**Chr 5-1**	**Chr 5-2**	**Chr 5-3**	**Chr 5-4**
GTP 5-4 candidate	–			–	–	–	–	–	–	–	1,182,596
AX-123358608	3.63E−06			5D	6.757	–	–	2,613,648	–	24,857,145	2,876,834
AX-166506768	3.63E−06			5D	6.757	–	–	2,524,864	3,118,449	–	–
AX-166524167	3.63E−06			5D	6.757	5D	12.114	2,568,242	3,075,907	–	–
AX-123361979	5.35E−06			5D	8.109	5D	12.114	2,212,147	3,426,121	24,503,708	
AX-166514937	9.60E−05			5D	9.484	–	–	3,950,769	–	24,414,731	3,354,090
AX-89892771	9.60E−05			5D	9.484	–	–	3,984,739		24,382,881	3,363,253
AX-166506799	1.51E−04			5D	10.802	–	–	–	–	22,754,788	4,769,438
AX-166518037	2.97E−04			5D	1.351	5D	8.12	–	–	–	1,615,784
AX-166505352	3.28E−04			5D	9.484	–	–	4,160,836	–	–	
AX-123358650	1.59E−03			5D	9.484	–	–	5,695,643	4,386,116	–	–

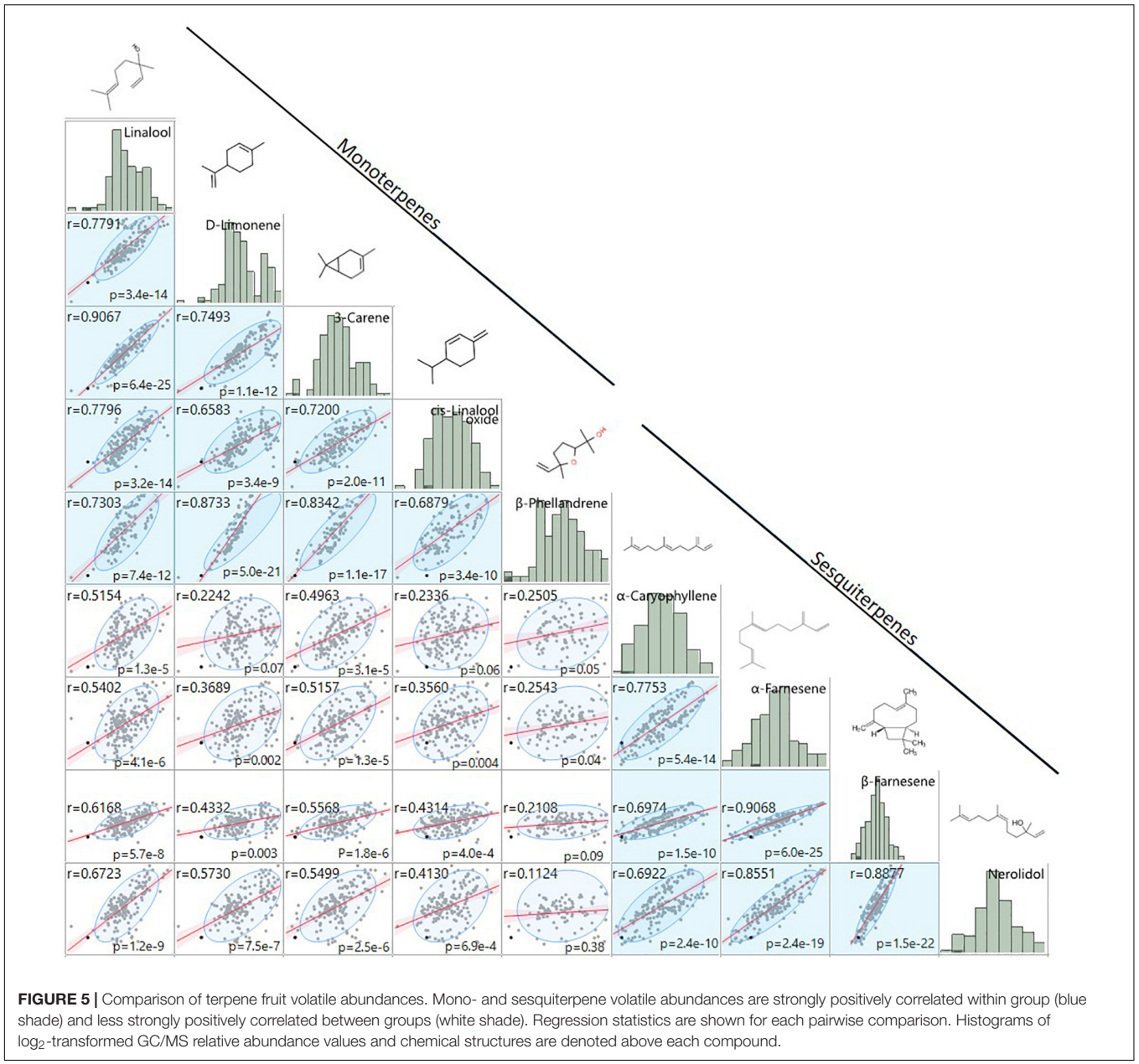

FIGURE 5 | Comparison of terpene fruit volatile abundances. Mono- and sesquiterpene volatile abundances are strongly positively correlated within group (blue shade) and less strongly positively correlated between groups (white shade). Regression statistics are shown for each pairwise comparison. Histograms of log_2-transformed GC/MS relative abundance values and chemical structures are denoted above each compound.

of methyl anthranilate variance (AX-123358624, p = 6.324e–07). A diagram of known and hypothesized methyl anthranilate pathway components is provided in **Supplementary Figure 2**.

The QTL on LG 2A broadly corresponds to the *F. vesca*-like Chr 2-2 of the "Camarosa" octoploid genome, based on chromosome-wide genetic–genomic connections (Hardigan et al., 2020). However, nucleotide BLAST of IStraw35 methyl anthranilate marker sequences align non-specifically to all "Camarosa" Chr 2 homoeologs except Chr 2-2 (**Table 1**). The corresponding syntenic regions of Chr 2-1 and Chr 2-3 each contain an *ANTHRANILIC ACID METHYL TRANSFERASE*-like homoeolog. The most-correlated QTL marker (AX-123540849) aligns 10 kb (two genes) from the *AAMT-like* gene on Chr 2-1 (maker-Fvb2-1-snap-gene-255.58) and 14.6 kb (four genes) from the *AAMT-like* gene on Chr 2-3 (maker-Fvb2-3-snap-gene-33.59) (**Figures 2A,B** and **Table 1**). These two *FanAAMT-like* genes have the highest sequence identity to the published Chr 4 *FanAAMT* gene except the highly expressed *FanAAMT* gene on Chr 4-1 and its non-expressed homoeologs (**Figure 2**). Reference-based RNA-seq shows that both candidate *FanAAMT*-like transcripts are highly abundant in some fruit transcriptomes, but relatively low or almost absent in others (Chr 2-1, 3-278 TPM; Chr 2-3, 15-1324 TPM) (**Supplementary Table 1**). The RNA-seq read alignments for both gene models show atypically high degrees of sequence disagreement with the "Camarosa" reference, which suggests these transcript reads could be derived from an alternative locus, such as a hypothesized deletion on Chr 2-2 (**Supplementary Figure 3**).

In the LG 5D/Chr 5-4 QTL region, no candidate genes were found, which correspond to known or hypothesized strawberry methyl anthranilate pathway components. However, a co-segregating transcript eQTL on Chr 5-4 was identified for a putative glutathione peroxidase gene (maker-Fvb5-4-augustus-gene-12.41) (transcript level h^2 = 1.0) (**Table 1**). Single-marker analysis explains 51.1% of the transcript variations observed (AX-123358624, p = 2.8-e9) (**Supplementary Figure 4A**). Accumulation of this transcript positively correlates with methyl anthranilate production (**Supplementary Figure 4B**).

A possible third methyl anthranilate signal on Chr 7 corresponds with the position of two *ANTHRANILATE SYNTHASE ALPHA* (*FanAS*-α) homoeologs (**Supplementary Figure 5A** and **Supplementary Table 3**). Both genes represent the only *FanAS*-α transcripts abundant in the fruit (**Supplementary Figure 5B**), and presence/absence variation of the Chr 7-4 *FanAS*-α transcript is governed by a *cis*-eQTL, which co-segregates with the methyl anthranilate signal (**Supplementary Figures 5C,D**).

Mesifurane QTL and Candidate Genes

Two mesifurane (h^2 = 0.72) QTLs were identified on LGs 1A and 7B of the FL_08-10 × 12.115-10 map (**Figures 3A,B** and **Table 2**). The mesifurane volatile QTL on LG 7B co-segregates with a transcript eQTL for the published mesifurane biosynthesis gene *O-METHYL TRANSFERASE 1* (*FanOMT1*; maker-Fvb7-4-augustus-gene-6.44*)* (h^2 of transcript accumulation = 0.80) (**Figure 3B**). Homozygosity of the mesifurane LG 7B minor allele (AX-123363184) eliminates both *FanOMT1* transcript and mesifurane production (**Figure 3C**). The *FanOMT1* allelic states correspond with stepwise increases in transcript abundance (**Figure 3C**). A novel mesifurane QTL was detected on LG 1A of the FL_08-10 × 12.115-10 genetic map (**Figure 3A** and **Table 2**). Alignments of probe nucleotide sequences to the "Camarosa" genome are not subgenome specific (**Table 2**).

An epistatic interaction was detected between the two QTLs [F(6,173) = 13.78, p = 9.0e−13] (**Figure 3D**). Homozygosity of the Chr 1-group QTL (AA genotype) strongly diminishes mesifurane abundance even when the competent Chr 7-4 *FanOMT1* allele is homozygous (BB genotype), and transcript abundance is highest (AA/BB vs. AB/BB Tukey HSD p = 2.4e−7). At least one competent allele at each locus is required for robust mesifurane production (**Figure 3D**). Mesifurane abundance is somewhat higher when both alleles are homozygous [BB/BB vs. AB/AB, AB/BB, and BB/AB ANOVA F(1,149) = 4.38, p = 0.038].

Because the mesifurane Chr 1 QTL probe sequences align equally well to multiple subgenomes, all Chr 1 homoeologous regions were considered for candidate gene identification. The published furaneol biosynthesis gene *FanQR* (maker-Fvb6-3-augustus-gene-21.50) is not located in Chr group 1 nor were genes of similar function found in the region. The published *F.* × *ananassa* furaneol glucosyltransferase gene from Yamada et al. (2019) (UGT85K16) has two putative homologs located in the "Camarosa" Chr 1 group; however, they are over 10 Mb from the mesifurane Chr 1 QTL (100% nucleotide identity, maker-Fvb1-4-augustus-gene-115.43; 97% nucleotide identity, maker-Fvb1-2-augustus-gene-137.19) (**Table 2**). The two most active *F.* × *ananassa* furaneol glucosyltransferases from Song et al. (2016) (UGT71K3a/b and UGT73B23/4) have putative orthologs outside of the Chr 1 group (98% nucleotide identity, augustus_masked-Fvb3-4-processed-gene-50.19; 100% nucleotide identity; augustus_masked-Fvb2-2-processed-gene-195.2). However, UGT73B23/4 has a highly identical second homolog in Chr 1-2 (augustus_masked-Fvb1-2-processed-gene-35.18; 96% nucleotide identity, 100% coverage) located 164 genes (0.9 Mb) from the most significant mesifurane marker (AX-166520175, p = 2.4e−6) and 49 genes (0.2 Mb) from the marker AX-166510720 (p = 5.6e−05) (**Figure 3A** and **Table 2**).

In diverse tissues of the "Camarosa" plant (AX-166520175 = BB), the candidate Chr 1-2 furaneol glucosyltransferase transcript levels are high in roots but low in the ripe fruit, with fruit expression somewhat increasing with ripening series (**Supplementary Table 3**). In the mature fruit RNA-seq populations, the Chr 1-2 candidate is modestly expressed in the fruit (5.3 ± 2.2 TPM).

Using a high-resolution melting (HRM) assay in two separate crosses (n = 72), two HRM markers targeting SNPs in the mesifurane Chr 1-2 QTL (AX-166520175 and AX-166502845) were tested for association with mesifurane abundance in marker-assisted seedling selection. Both Chr 1 markers were confirmed to predict the segregation of mesifurane abundance [F(1,71) = 8.25823, p = 0.0006] (**Supplementary Figure 6**).

Terpene QTL and Candidate Genes

A shared QTL for the production of nine terpene compounds was discovered corresponding to Chr 3 of the *F. vesca* physical map. Only two shared terpene markers are positioned in an octoploid genetic map (AX-166504318, AX-166521725), both of which correspond to LG 3B in "Holiday" × "Korona" (**Figure 4A** and **Table 3**), which represents Chr 3-3 in the "Camarosa" genome (Hardigan et al., 2020). For linalool (h^2 = 0.749), the most-correlated QTL marker (AX-166513106, p = 4.2e−11) explains 14.8% of the observed variance in linalool abundance. This represents a large absolute difference as linalool is among the most abundant volatiles in strawberry fruit, commonly exceeding 100 $ng^1\ gFW^{-1}\ h^{-1}$ in the cultivars used as parental lines for these populations (Schwieterman, 2013). The linalool QTL is shared with the monoterpenes 3-carene (R^2 = 0.09, p = 1.2e−8), β-phellandrene (R^2 = 0.21, p = 7.0e−7), α-limonene (R^2 = 0.06, p = 7.0e−7), and linalool oxide (R^2 = 0.08, p = 4.4e−6), and the sesquiterpenes nerolidol (R^2 = 0.09, p = 4.3e−6), α-caryophellene (R^2 = 0.10, p = 8.9e−6), α-farnesene (R^2 = 0.16, p = 1.6e−6), and β-farnesene (R^2 = 0.12, p = 2.4e−7) (**Figures 4A,B** and **Table 4**). The four sesquiterpene QTL are comprised of two shared significant markers, which are also common to the monoterpenes (**Table 4**). Mono- and sesquiterpene volatile abundances are very strongly correlated within group, and moderately correlated between groups (**Figure 5**).

The LG 3B terpene QTL IStraw35 probe sequences align non-specifically to all four Chr 3 homoeologs (**Table 3**). These corresponding genomic regions contain putative terpenoid biosynthesis gene clusters, which together contain three annotated copies of *(3S,6E)-NEROLIDOL SYNTHASE*, three copies of *(E,E)-ALPHA-FARNESENE SYNTHASE*, and

TABLE 2 | Mesifurane QTL and candidate gene positions.

Chr 1 candidate genes			Annotation			Fruit TPM	"Camarosa" genome physical position			
							Chr 1-1	Chr 1-2	Chr 1-3	Chr 1-4
augustus_masked-Fvb1-2-processed-gene-35.18			UGT73B23/4; 96% identical			5.3 ± 2.2	–	3,570,716	–	–
maker-Fvb1-2-augustus-gene-137.19			UGT85K16; 97% identical			6.3 ± 10.3	–	13,720,348	–	–
maker-Fvb1-4-augustus-gene-115.43			UGT85K16 100% identical			16.1 ± 1.3	–	–	–	11,555,530
Chr 1 QTL	***Mesifurane* QTL *p*-value**		**Holiday × Korona map**		**14.95 map**		**"Camarosa" genome physical position**			
			LG	**Position**	**LG**	**Position**	**Chr 1-1**	**Chr 1-2**	**Chr 1-3**	**Chr 1-4**
AX-166520175	1.3E−06		–	–	1A	14.522	26,511,819	4,390,575	1,142,702	1,929,385
AX-166502845	2.4E−05		–	–	–	–	26,626,204	4,514,320	1,030,480	2,093,705
AX-166520217	3.2E−05		–	–	1A	12.229	27,171,481	4,947,347	560,320	–
AX-166502747	3.2E−05		–	–	–	–	25,629,689	–	1,890,371	1,053,881
AX-123360076	3.8E−05		–	–	–	–	25,905,230	–	–	1,374,796
AX-166510551	5.2E−05		–	–	1A	16.117	–	–	–	809,104
AX-166510720	5.6E−05		–	–	1A	14.587	25,788,362	3,789,710	–	1,250,466
AX-89779650	6.1E−05		–	–	–	–	–	4,400,404	–	1,966,919
AX-166510859	7.2E−05		–	–	1A	12.381	27,316,667	–	–	–
AX-166502836	1.3E−04		–	–	–	–	–	–	–	1,961,919
Chr 7 Candidate Gene			**Annotation**			**Fruit TPM**	**"Camarosa" genome physical position**			
							Chr 7-1	**Chr 7-2**	**Chr 7-3**	**Chr 7-4**
maker-Fvb7-4-augustus-gene-6.44			FanOMT1; 100% identical			1,099 ± 579	–	–	–	587,687
Chr 7 QTL	***Mesifurane* QTL *p*-value**	**FanOMT1 eQTL p-value**	**Holiday × Korona map**		**14.95 map**		**Chr 7-1**	**Chr 7-2**	**Chr 7-3**	**Chr 7-4**
			LG	**Position**	**LG**	**Position**				
AX-166526335	9.4E−04	1.1E−05	–	–	–	–				42,976
AX-166526460	6.6E−04	2.6E−05	7B	60.9343	–	–	30,801,212	28,013,548		1,689,914
AX-123359530	9.4E−04	1.1E−05	7B	60.9343	–	–	–	–	–	447,154
AX-123359532	9.4E−04	1.1E−05	7B	60.9343	–	–	30,210,474		1,526,586	460,640
AX-123359539	9.4E−04	1.1E−05	–	–	–	–			1,824,585	748,784
AX-123359733	9.4E−04	1.1E−05	–	–	–	–	–	–	–	692,839
AX-123363167	9.4E−04	1.1E−05	7B	60.9343	–	–	32,004,806		1,170,841	230,562
AX-123363171	9.4E−04	1.1E−05	7B	60.9343	–	–	30,218,969	–	1,518,160	452,092
AX-123363175	9.4E−04	1.1E−05	–	–	–	–	30,120,861	–	1,675,218	550,652
AX-123366658	9.4E−04	1.1E−05	7B	60.9343	–	–	30,190,895	–	1,546,336	475,412

TABLE 3 | Terpene QTL and candidate gene positions.

Chr 3 candidate genes	Annotation	Fruit TPM	Fruit TPM with NES1 reference†	"Camarosa" genome physical position			
				Chr 3-1	Chr 3-2	Chr 3-3	Chr 3-4
maker-Fvb3-1-augustus-gene-304.55	(3S,6E)-nerolidol synthase	213.3 ± 91	2.1 ± 0.4	30,433,804		–	–
maker-Fvb3-1-augustus-gene-304.61	(E,E)-alpha-farnesene synthase	22.9 ± 41.9	23.7 ± 46.6	30,466,137	–	–	–
maker-Fvb3-1-augustus-gene-292.46	Solanesyl diphosphate synthase	5.3 ± 1.2	4.7 ± 1	29,233,943	–	–	–
maker-Fvb3-2-augustus-gene-19.46	(E,E)-alpha-farnesene synthase	13.2 ± 18.9	12.7 ± 18.7	–	1,973,778	–	–
maker-Fvb3-2-augustus-gene-19.40	(3S,6E)-nerolidol synthase	491.9 ± 215.1	0.4 ± 0.37	–	1,955,222	–	–
maker-Fvb3-2-augustus-gene-19.39	(3S,6E)-nerolidol synthase	270.9 ± 124.8	0.2 ± 0.15	–	1,931,177	–	–
maker-Fvb3-2-augustus-gene-20.45	(E,E)-alpha-farnesene synthase	0 ± 0	0 ± 0	–	2,025,126	–	–
maker-Fvb3-3-augustus-gene-10.48	Solanesyl diphosphate synthase	6.4 ± 1.3	6.9 ± 1.5	–	–	1,002,523	–
maker-Fvb3-4-augustus-gene-266.38	Solanesyl diphosphate synthase	7.8 ± 1.5	8.1 ± 1.6	–	–	–	26,679,507
FanNES1 (*ad hoc* contig)	Nerolidol Synthase 1 (KX450224)	–	2167.7 ± 955.9	–	–	–	–

Monoterpene markers	*Linalool* QTL *p*-value	Holiday × Korona map		14.95 map		"Camarosa" genome physical position			
		LG	Position	LG	Position	Chr 3-1	Chr 3-2	Chr 3-3	Chr 3-4
AX-166513106	4.2E−11	–	–	–	–	30,972,069	2,393,480	–	–
AX-89828071	7.3E−11	–	–	–	–	31,259,033	2,621,813		
AX-89884609	1.2E−10	–	–	–	–	30,989,345	2,413,322	–	–
AX-89863467	1.8E−09	–	–	–	–	31,178,162	2,542,736	–	–
AX-89786367	1.8E−09	–	–	–	–	30,801,101	–	–	–
AX-89880889	5.9E−09	–	–	–	–	–	–	–	–
AX-166504368	5.6E−08	–	–	–	–	29,214,600	900,770	1,263,036	26,634,939
AX-89785560	7.7E−08	–	–	–	–	–	1,072,295	866,163	26,820,703
AX-89787482	7.9E−08	–	–	–	–	–	2,510,805	–	–
AX-89885089	1.2E−07	–	–	–	–	–	2,602,215	–	–
AX-89788064	1.1E−07	–	–	–	–	31,473,593	2,747,017	–	–
AX-166504318	2.1E−07	3B*	6.95	–	–	–	–	–	26,516,095
AX-166521725	2.1E−07	3B*	6.95	–	–	–	–	–	26,516,247
AX-89827597	7.7E−08	–	–	–	–	31,088,976	2,467,624	–	–
Monoterpene and sesquiterpenes									
AX-166504368	5.6E−08	–	–	–	–	29,214,600	900,770	1,263,036	26,634,939
AX-89785560	7.7E−08	–	–	–	–	–	1,072,295	866,163	26,820,703

†*Represents TPM values after the addition of a FanNES1 reference sequence followed by RNA-seq reanalysis.*
**LG 3B in Holiday × Korona corresponds broadly to "Camarosa" genome Chr 3cccc-3.*

TABLE 4 | Shared terpene QTL and candidate transcript eQTL markers.

	Monoterpenes					Sesquiterpenes			Transcript eQTL	
Terpene markers[†]	***Linalool***	***3-Carene***	**β *-Phellandrene***	**α *-Limonene***	***Linalool oxide***	**β *-Farnesene***	**α *-Farnesene***	**α *-Caryophellene***	***Nerolidol***	***FanNES1***	***FanSPS1***
AX-166513106	4.18E−11	2.12E−08	7.00E−07	3.99E−07	4.39E−06	4.31E−04	1.18E−02	1.90E−03	7.10E−04	1.34E−03	2.53E−04
AX-89828071	7.34E−11	1.98E−08	7.00E−07	5.29E−07	1.82E−05	5.21E−04	1.23E−02	2.04E−03	1.11E−03	4.68E−03	3.33E−04
AX-89884609	1.17E−10	1.61E−08	8.63E−07	6.18E−07	5.96E−06	3.40E−04	5.46E−03	1.52E−03	7.94E−04	1.34E−03	2.53E−04
AX-89786014	1.78E−09	1.10E−05	7.78E−04	1.27E−04	1.62E−04	1.61E−04	1.40E−02	1.03E−03	6.50E−02	1.34E−03	2.53E−04
AX-89786367	1.78E−09	1.10E−05	7.78E−04	1.27E−04	1.62E−04	1.61E−04	1.40E−02	1.03E−03	1.19E−03	1.34E−03	2.53E−04
AX-89826999	1.78E−09	1.10E−05	7.78E−04	1.27E−04	1.62E−04	1.61E−04	1.40E−02	1.03E−03	6.50E−02	1.34E−03	2.53E−04
AX-89863467	1.78E−09	1.10E−05	7.78E−04	1.27E−04	1.62E−04	1.61E−04	1.40E−02	1.03E−03	1.19E−03	1.34E−03	2.53E−04
AX-89880889	5.90E−09	4.52E−08	1.46E−06	1.05E−05	5.85E−05	6.04E−04	5.29E−03	7.08E−03	1.12E−03	4.68E−03	3.33E−04
AX-166504368*	5.62E−08	1.02E−04	8.75E−03	1.49E−01	1.93E−02	2.40E−07	1.25E−05	8.86E−06	4.33E−06	2.69E−02	2.30E−03
AX-89827597	7.71E−08	2.32E−05	2.46E−03	5.37E−03	1.69E−03	1.27E−04	5.82E−03	3.93E−03	6.38E−02	1.34E−03	2.53E−04
AX-89785560*	7.74E−08	1.02E−04	1.27E−02	1.49E−01	2.59E−02	7.16E−07	4.01E−05	8.86E−06	4.33E−06	2.05E−02	8.19E−04
AX-89787482	7.85E−08	5.74E−05	2.46E−03	2.16E−02	1.44E−03	2.67E−04	7.31E−03	3.29E−03	1.84E−03	1.34E−03	2.53E−04
AX-89788064	1.12E−07	2.07E−07	2.58E−05	2.23E−05	3.79E−04	4.15E−03	2.51E−02	1.58E−02	1.83E−03	4.68E−03	4.86E−04

[†] *Terpene markers are ordered according to significance in linalool.*

**Markers AX-166504368 and AX-89785560 significances are elevated among sesquiterpenes.*

three copies of *SOLANESYL DIPHOSPHATE SYNTHASE* (**Table 3**). The characterized *FanNES1* deletion responsible for linalool biosynthesis in octoploids was not detected among the "Camarosa" *FanNES1*-like gene sequences; however, this gene appears on Chr 3-3 in the updated "Camarosa" v2 genome (Hardigan, personal communication).

To assess terpenoid-related transcript levels in fruit, the published *FanNES1* gene sequence was added *ad hoc* to the "Camarosa" v1 genome as an independent contig prior to reference-based RNA-seq re-assembly. The resulting *FanNES1* transcript was the only *NES1*-like transcript abundant in fruit (**Table 3**). Two eQTL signals co-segregated with terpene QTL markers (**Table 4**). These eQTL correspond to the ambiguously located *FanNES1* gene and to a *cis*-eQTL for a novel *SOLANESYL DIPHOSPHATE SYNTHASE* (*FanSPS*) gene on Chr 3-3 (maker-Fvb3-3-augustus-gene-10.48). Linalool abundance (n = 61) is statistically correlated with *FanNES1* transcript levels (R^2 = 0.31, p = 0.017), but a significant relationship was not detected with *FanSPS* (R^2 = 0.20, p = 0.14). An additional *cis*-eQTL was detected for an *(E,E)-ALPHA-FARNESENE SYNTHASE* gene on Chr 3-2 (maker-Fvb3-2-augustus-gene-19.46) (AX-166522353; R^2 = 0.50, p = 1.9e−7). While this terpene biosynthesis gene is positioned closely with potential IStraw35 physical positions (**Table 3**), these markers do not genetically co-segregate with the terpene QTL, and transcript levels are not correlated with linalool abundance (R^2 = 0.004, p = 0.98).

DISCUSSION

Many QTLs were discovered for strawberry flavor and aroma compounds known to influence the human sensory experience. These QTLs are derived from eight biparental crosses phenotyped across multiple seasons under a commercial cultural system in central Florida and are likely to be useful for making genetic gains in related germplasm. Markers correlated with these traits may be used to guide breeding decisions and identify and select for alleles mediating flavor and aroma. Potential causal genes were identified via a multiomic approach, and provide a foundation for possible gene-editing-based approaches to improved strawberry flavor. These genetic discoveries represent new opportunities for improving flavor in commercial strawberry, and advance the basic understanding of the molecular mechanisms driving fruit flavor and aroma.

Methyl Anthranilate

Consistent with the long-standing polygenic hypothesis for methyl anthranilate production in octoploid strawberry, multiple QTLs were identified for this trait. Multiomic analysis of QTL regions implicated several likely causal genes within distinct QTL. Because many discrete loci affect methyl anthranilate levels, and because environmental interactions transiently induce wide phenotypic swings including trait presence/absence, the interactions between loci could not be reliably measured in this sample size and diverse set of crosses. However, no loci could be identified as singularly required for production. The published *FanAAMT* gene on Chr 4, which did not emerge as

a QTL in this analysis, was identified solely in the context of the biparental population "Florida Elyana" × "Mara de Bois" (population "10.133"), which contained only 13 analyzed progeny from one cross (Pillet et al., 2017). It is possible that these differences are due to segregating genetic factors becoming fixed or lost in subsequent populations. This hypothesis is supported by the low positive rates resulting from F1 backcrosses to "Mara des Bois" (Chambers, 2013) and the fact that population "10.133" does not independently support the identified QTL regions. This QTL analysis is mostly comprised of populations using the parent "12.115-10," which is a descendant of "Mara des Bois" that produces more methyl anthranilate than its ancestor. It is likely that this breeding line has been enriched for favorable methyl anthranilate genetics. These findings might relate more to quantitative differences in methyl anthranilate abundance, rather than the genetic presence/absence, which historically defines this rare trait among strawberry cultivars.

The methyl anthranilate LG2A QTL is positively correlated with the production of two other methyl ester volatiles, namely, methyl 2-hexenoate and methyl 2-methylbutyrate. Consistent with historical segregation ratios, which implicate methyl anthranilate as a polygenic trait, less methyl anthranilate variance is explained by this QTL compared with the other two methyl ester volatiles. As their precursors are not closely related, a single promiscuous methyl transferase offers a parsimonious explanation. In Pillet et al. (2017), moderate methyl anthranilate levels were occasionally detected in the near absence of the published *FanAAMT* transcript, which is suggestive of the possibility of additional methyl transferases.

While hundreds of methyl transferase genes exist in the octoploid genome, only the published *FanAAMT* has experimentally demonstrated affinity for anthranilate. Four *FanAAMT*-like transcripts were abundantly detected in mature octoploid fruit transcriptomes. Two of these expressed *FanAAMT*-like genes correspond to the QTL on Chr 2, located within two genes (Chr 2-1) and four genes (Chr 2-3) from the most-correlated QTL markers. The expressed Chr 4-1 *AAMT*-like sequence in the "Camarosa" genome is the most similar to the published Chr 4 *AAMT* sequence, whose subgenomic identity was not established (Pillet et al., 2017). This gene on Chr 4-1 might be the *FanAAMT* gene in "Camarosa," particularly as RNA-seq reads from fruit transcriptomes have high sequence fidelity with this gene reference (**Supplementary Figure 3**).

Genetic mapping suggests only a single methyl anthranilate QTL for chromosome group 2, which should be located on "Camarosa" Chr 2-2. However, this QTL marker region in the Chr 2-2 physical sequence is completely absent. As "Camarosa" is not capable of producing methyl anthranilate, one or more required genetic elements are expected to be missing in this reference genome. Poor RNA-seq sequence agreement with the *FanAAMT*-like Chr2-1 and Chr 2-3 homoeologs suggests that the correct position of these transcript reads is not in the "Camarosa" genome. It is unlikely that RNA-seq reads corresponding to the published Chr 4 *FanAAMT* transcript would map falsely to the Chr 2 candidate loci, as the published sequence is the most identical to the Chr 4-1 *FanAAMT* gene, and the RNA-seq mapping criteria excludes all non-specific reads. A comparative pan-genome analysis using a methyl anthranilate-producing individual would be highly informative and will be undertaken in the future.

No candidate genes belonging to the hypothesized methyl anthranilate pathway are located in the Chr 5-4 region of the "Camarosa" reference. However, a co-segregating transcript *cis*-eQTL was detected for a putative glutathione peroxidase gene. Many, but not all, significant markers were shared between the trait QTL and transcript *cis*-eQTL, since methyl anthranilate levels are influenced at multiple loci, while the candidate transcript is under strong single locus control. In microbes, there is precedent for heme peroxidase activity catalyzing methyl anthranilate biosynthesis (Van Haandel et al., 2000); however, this reaction is unlikely to proceed via a glutathione peroxidase. It is possible that this *cis*-eQTL is simply in close linkage with the actual causal gene, which was either not correctly identified or is not present in the "Camarosa" reference genome.

A possible third methyl anthranilate QTL corresponds with two Chr 7 *ANTHRANILATE SYNTHASE ALPHA* (*FanAS*-α) homoeologs. Presence/absence variation of the Chr 7-4 *FanAS*-α transcript is governed by a *cis*-eQTL, which co-segregates with the putative methyl anthranilate markers at this locus. The Chr 7-2 *FanAS*-α transcript also demonstrates transcript presence/absence variation, but this is apparently due mostly to non-heritable factors that are uncorrelated with Chr 7-4 transcript level variation. Although there are few methyl anthranilate-positive individuals among 61 fruit transcriptomes, none of the 10 individuals with zero combined *FanAS*-α expression shows methyl anthranilate production. This pathway mechanism is consistent with previous findings implicating *FanAAMT* as necessary-but-not-sufficient for methyl anthranilate production (Pillet et al., 2017). The absence of anthranilate substrate in the mature fruit would help explain the observed absence of methyl anthranilate production even when *FanAAMT* transcript levels are high. Further efforts to validate this potential QTL signal are underway.

Mesifurane

Mesifurane (2,5-dimethyl-4-methoxy-3(2H)-furanone or DMMF) is derived from the methylation of furaneol (4-hydroxy-2,5-dimethyl-3(2H)-furanone, or HDF) by *FanOMT1* (Wein et al., 2002;Zorrilla-Fontanesi et al., 2012). Mesifurane variance is influenced by a loss-of-function mutation in the *FanOMT1* promoter, which eliminates transcription and mesifurane production. This model was validated by the detection of a *cis*-eQTL for the published *FanOMT1* gene, which co-segregates with the Chr 7B mesifurane trait QTL. A novel mesifurane QTL was detected on Chr 1A, which is in epistasis with the Chr 7B QTL. This QTL region contains a fruit-expressed furaneol glucosyl transferase, which is 95% identical to a characterized furaneol glucosyl transferase from *F.* × *ananassa*. A substrate-restricting glucosyltransferase candidate is consistent with the epistatic interaction detected with the *FanOMT1* locus. Depletion of substrate via glucosylation would limit mesifurane biosynthesis regardless of high *FanOMT1* transcript levels. Conversely, elimination of the *FanOMT1* transcript would eliminate mesifurane production regardless of substrate availability. The

LG 1A mesifurane QTL was subsequently confirmed using two validation populations, providing robust support for this QTL. This two-gene model for mesifurane biosynthesis in cultivated strawberry can be exploited for genetic gain via marker-assisted selection. Moderate mesifurane levels can be maintained via dual selection for heterozygous/heterozygous allelic states, and somewhat elevated mesifurane levels can be achieved via double-homozygote selection. These findings may resolve some of the outstanding questions in mesifurane genetics posed by Cruz-Rus et al. (2017).

Terpenes

Homoeologous terpene gene arrays were detected for nine strawberry mono- and sesquiterpene QTL, including the desirable compound linalool. In citrus, monoterpenes and sesquiterpenes co-locate to single genomic QTL containing paralogous terpene synthases (Yu et al., 2017). We identify a similar phenomenon in cultivated strawberry. This terpene hotspot contains clusters of multiple terpenoid synthase classes, in addition to homoeologous genes on three of four subgenomes. The known biosynthesis gene *FanNES1* was associated with terpene levels via trait/transcript level correlations and trait QTL/eQTL co-segregation. Solanesyl diphosphate synthase may contribute to terpene abundances as well. The *cis*-eQTL/QTL genetic association with solanesyl diphosphate synthase on Chr 3-3 helps support the subgenomic location of *FanNES1* and the shared terpenoid QTL in the "Camarosa" genome, despite only two markers being genetically mapped and probe nucleotide sequences aligning to multiple subgenomes. It is possible that the influence of other terpene-related genes in this array remains undetected due to limitations in genome completeness, marker subgenome ambiguity, and/or presence–absence variation among genomes. With additional octoploid strawberry genomes for comparison and improved subgenomic genotyping tools, complex associations in octoploid strawberry will become more robust.

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/, SRP039356; https://www.ebi.ac.uk/ena, PRJEB12420.

AUTHOR CONTRIBUTIONS

CB, SV, KF, and VM contributed to study conception and design. CB and SV performed QTL analysis. CB, MH, BH, AS prepared DNA, RNA, and GC/MS samples, performed GC/MS data analysis, eQTL analysis, and eQTL/QTL co-localization analysis. SL and YO designed and executed the mesifurane marker validation experiment via HRM. CB wrote the manuscript and all authors contributed substantially to the editing process.

ACKNOWLEDGMENTS

We gratefully acknowledge Aristotle Koukoulidis for the genomic DNA isolation, Natalia Salinas for the assistance with GC/MS sample collection and preparation, Denise Tieman for the assistance with GC/MS operation, Nadia Mourad for the assistance compiling eQTL results, Andrew Hanson and Harry Klee for the project discussion and guidance, Alan Chambers and Jeremy Pillet for the RNA isolation and RNA-seq line selection, Zhen Fan for the manuscript editing and discussion, and Angelita Arredondo and Kelsey Cearley for the field expertise and assistance with fruit collection. This manuscript has been released as a pre-print at bioRxiv (Barbey C.R. et al., 2020).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.615749/full#supplementary-material

Supplementary Figure 1 | Pedigree of eight interrelated strawberry families segregating for flavor and aroma. Families used in volatile Qtl analysis (light gray) are indicated with the number of analyzed progeny in parenthesis. Families used in both volatile Qtl analysis and fruit Rna-seq analysis are also shown (dark gray).

Supplementary Figure 2 | The known methyl anthranilate pathway in strawberry. Methyl anthranilate is conditionally derived from the methylation of anthranilate (bold arrow) in the mature fruit. Anthranilate (also referred to as anthranilic acid) is derived from chorismate via the anthranilate synthase enzyme complex, and is a substrate in tryptophan biosynthesis.

Supplementary Figure 3 | Rna-seq mapping of *Fan*Aamt-like genes in 'Mara des Bois' fruit. Rna-seq read-map assemblies are shown for Aamt-like coding sequences in the 'Camarosa' octoploid genome (yellow arrows) with predicted Snp variants (red marker). **(A)** The Chr 4-1 *Fan*Aamt genes shows no predicted coding sequence polymorphisms, while **(B)** the Chr 2-1 *Fan*Aamt candidate gene and the **(C)** Chr 2-3 *Fan*Aamt candidate gene references show poor agreement with transcript data.

Supplementary Figure 4 | Methyl anthranilate Ch5 Qtl and candidate genes. **(A)** The methyl anthranilate Qtl on chromosome 5-4 (Lg 5A) is shared with a *cis*-eQtl for a putative *glutathione peroxidase* transcript. **(B)** The range of both methyl anthranilate (r^2 = 0.181, p = 1.5e−5) abundance and *Glutathione Peroxidase* (r^2 = 0.511, p = 2.8e−9) transcript abundance is shown for the shared marker Ax-123358624.

Supplementary Figure 5 | Methyl anthranilate Chr7 candidate genes. **(A)** The putative Chr 7 methyl anthranilate signal corresponds to two homoeologous regions containing anthranilate synthase genes on Chr 7-2 (top) and Chr 7-4 (bottom). **(B)** Anthranilate synthase-like deduced proteins in the 'Camarosa' genome are shown in a neighbor-joining cladogram, with transcript abundance heatmaps representing the highest Tpm detected among the fruit transcriptomes. The anthranilate synthase alpha subunit candidate genes on Chr 7-2 and Chr 7-4 are highly abundant in the fruit, as are two corresponding beta subunits gene. **(C)** Variable transcript levels of one anthranilate synthase candidate (Chr 7-4) are governed by a transcript *cis*-eQtl. **(D)** The eQtl for the anthranilate synthase alpha candidate, which governs transcript presence/absence in the fruit (left), also co-segregates with the methyl anthranilate Chr 7 putative Qtl (righ).

Supplementary Figure 6 | High-resolution melting (Hrm) curves for two Chr 1 mesifurane Qtl markers. Ten individuals were initially confirmed to be either homozygous negative (red), heterozygous (green), or homozygous positive (blue) for the markers **(A)** Ax-166520175 and **(B)** Ax-166502845 based on melting curve properties. **(C)** Anova test statistics of fruit mesifurane abundance levels among 72 additional individuals tested by Hrm confirm the Chr 1 mesifurane Qtl markers.

REFERENCES

Aharoni, A., Giri, A. P., Verstappen, F. W. A., Bertea, C. M., Sevenier, R., Sun, Z., et al. (2004). Gain and loss of fruit flavor compounds produced by wild and cultivated strawberry species. *Plant Cell* 16, 3110–3131. doi: 10.1105/tpc.104.023895

Anciro, A., Mangandi, J., Verma, S., Peres, N., Whitaker, V. M., and Lee, S. (2018). FaRCg1: a quantitative trait locus conferring resistance to Colletotrichum crown rot caused by *Colletotrichum gloeosporioides* in octoploid strawberry. *Theor. Appl. Genet.* 131, 2167-2177. doi: 10.1007/s00122-018-3145-z

Arroyo, F. T., Moreno, J., Daza, P., Boianova, L., and Romero, F. (2007). Antifungal activity of strawberry fruit volatile compounds against colletotrichum acutatum. *J. Agric. Food Chem.* 55, 5701–5707. doi: 10.1021/jf0703957

Barbey, C. R., Hogshead, M. H., Harrison, B., Schwartz, A. E., Verma, S., Oh, Y., et al. (2020). Genetic analysis of Methyl anthranilate, mesifurane, linalool and other flavor compounds in cultivated strawberry *Fragaria ×ananassa*. *bioRxiv* [preprint] 2020.2010.2007.330001. doi: 10.1101/2020.10.07.330001

Barbey, C., Hogshead, M., Schwartz, A. E., Mourad, N., Verma, S., Lee, S., et al. (2020). The genetics of differential gene expression related to fruit traits in strawberry (*Fragaria ×ananassa*). *Front. Genet.* 10:1317. doi: 10.3389/fgene.2019.01317

Barbey, C., Lee, S., Verma, S., Bird, K. A., Yocca, A. E., Edger, P. P., et al. (2019). Disease resistance genetics and genomics in octoploid strawberry. *G3 Genes Genomes Genetics* g3, 400597.402019. doi: 10.1534/g3.119.400597

Bassil, N. V., Davis, T. M., Zhang, H., Ficklin, S., Mittmann, M., Webster, T., et al. (2015). Development and preliminary evaluation of a 90 K Axiom® SNP array for the allo-octoploid cultivated strawberry Fragaria× ananassa. *BMC Genomics* 16:155. doi: 10.1186/s12864-015-1310-1

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B (Methodol.)* 57, 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x

Bood, K. G., and Zabetakis, I. (2002). The biosynthesis of strawberry flavor (II): biosynthetic and molecular biology studies. *J. Food Sci.* 67, 2–8. doi: 10.1111/j.1365-2621.2002.tb11349.x

Box, G. E. P., and Cox, D. R. (1964). An analysis of transformations. *J. R. Stat. Soc. Ser. B (Methodol.)* 26, 211–243. doi: 10.1111/j.2517-6161.1964.tb00553.x

Chambers, A. H. (2013). *Strawberry Flavor: From Genomics to Practical Applications*. Gainesville, FL: University of Florida.

Chambers, A., Pillet, J., Plotto, A., Bae, J., Whitaker, V., and Folta, K. (2014). Identification of a strawberry flavor gene candidate using an integrated genetic-genomic-analytical chemistry approach. *BMC Genomics* 15:217. doi: 10.1186/1471-2164-15-217

Cruz-Rus, E., Sesmero, R., Ángel-Pérez, J. A., Sánchez-Sevilla, J. F., Ulrich, D., and Amaya, I. (2017). Validation of a PCR test to predict the presence of flavor volatiles mesifurane and γ-decalactone in fruits of cultivated strawberry (Fragaria × ananassa). *Mol. Breed.* 37:131. doi: 10.1007/s11032-017-0732-7

Edger, P. P., Poorten, T. J., Vanburen, R., Hardigan, M. A., Colle, M., Mckain, M. R., et al. (2019). Origin and evolution of the octoploid strawberry genome. *Nat. Genet.* 51, 541–547. doi: 10.1038/s41588-019-0356-4

Eggink, P. M., Tikunov, Y., Maliepaard, C., Haanstra, J. P. W., De Rooij, H., Vogelaar, A., et al. (2014). Capturing flavors from *Capsicum baccatum* by introgression in sweet pepper. *Theor. Appl. Genet.* 127, 373–390. doi: 10.1007/s00122-013-2225-3

Faedi, W., Mourgues, F., and Rosati, C. (2002). *Strawberry Breeding and Varieties: Situations and Perspectives*. Leuven: International Society for Horticultural Science (ISHS), 51–59. doi: 10.17660/ActaHortic.2002.567.1

Fletcher, S. W. (1917). *Strawberry in North America*. New York, NY: The Macmillan Company

Folta, K. M., and Klee, H. J. (2016). Sensory sacrifices when we mass-produce mass produce. *Hortic. Res.* 3:16032. doi: 10.1038/hortres.2016.32

Hardigan, M. A., Feldmann, M. J., Lorant, A., Bird, K. A., Famula, R., Acharya, C., et al. (2020). Genome synteny has been conserved among the octoploid progenitors of cultivated strawberry over millions of years of evolution. *Front. Plant Sci.* 10:1789. doi: 10.3389/fpls.2019.01789

Larsen, M., and Poll, L. (1992). Odour thresholds of some important aroma compounds in strawberries. *Z. Lebensm. Unters. Forsch.* 195, 120–123. doi: 10.1007/BF01201770

Lommen, A., and Kools, H. J. (2012). MetAlign 3.0: performance enhancement by efficient use of advances in computer hardware. *Metabolomics* 8, 719–726. doi: 10.1007/s11306-011-0369-1

Noh, Y.-H., Oh, Y., Mangandi, J., Verma, S., Zurn, J. D., Lu, Y.-T., et al. (2018). High-throughput marker assays for FaRPc2-mediated resistance to Phytophthora crown rot in octoploid strawberry. *Mol. Breed.* 38:104. doi: 10.1007/s11032-018-0861-7

Pillet, J., Chambers, A. H., Barbey, C., Bao, Z., Plotto, A., Bai, J., et al. (2017). Identification of a methyltransferase catalyzing the final step of methyl anthranilate synthesis in cultivated strawberry. *BMC Plant Biol.* 17:147. doi: 10.1186/s12870-017-1088-1

R. Development Core Team (2014). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Raab, T., López-Ráez, J. A., Klein, D., Caballero, J. L., Moyano, E., Schwab, W., et al. (2006). FaQR, required for the biosynthesis of the strawberry flavor compound 4-hydroxy-2,5-dimethyl-3(2H)-furanone, encodes an enone oxidoreductase. *Plant Cell* 18, 1023–1037. doi: 10.1105/tpc.105.039784

Racine, J. S. (2011). RStudio: a platform-independent IDE for R and Sweave. *J. Appl. Econom.* 27, 167–172. doi: 10.1002/jae.1278

Rambla, J. L., Medina, A., Fernández-Del-Carmen, A., Barrantes, W., Grandillo, S., Cammareri, M., et al. (2017). Identification, introgression, and validation of fruit volatile QTLs from a red-fruited wild tomato species. *J. Exp. Bot.* 68, 429–442. doi: 10.1093/jxb/erw455

Sánchez-Sevilla, J. F., Cruz-Rus, E., Valpuesta, V., Botella, M. A., and Amaya, I. (2014). Deciphering gamma-decalactone biosynthesis in strawberry fruit using a combination of genetic mapping. RNA-Seq and eQTL analyses. *BMC Genomics* 15:218. doi: 10.1186/1471-2164-15-218

Sánchez-Sevilla, J. F., Vallarino, J. G., Osorio, S., Bombarely, A., Posé, D., Merchante, C., et al. (2017). Gene expression atlas of fruit ripening and transcriptome assembly from RNA-seq data in octoploid strawberry (Fragaria × ananassa). *Sci. Rep.* 7:13737. doi: 10.1038/s41598-017-14239-6

Schieberle, P., and Hofmann, T. (1997). Evaluation of the character impact odorants in fresh strawberry juice by quantitative measurements and sensory studies on model mixtures. *J. Agric. Food Chem.* 45, 227–232. doi: 10.1021/jf960366o

Schwieterman, M. L. (2013). *Metabolite Analysis, Environmental Factors, and a Transgenic Approach to Understanding Strawberry (Fragaria× Ananassa) Flavor*. Gainesville, FL: University of Florida.

Schwieterman, M. L., Colquhoun, T. A., Jaworski, E. A., Bartoshuk, L. M., Gilbert, J. L., Tieman, D. M., et al. (2014). Strawberry flavor: diverse chemical compositions, a seasonal influence, and effects on sensory perception. *PLoS One* 9:e88446. doi: 10.1371/journal.pone.0088446

Song, C., Hong, X., Zhao, S., Liu, J., Schulenburg, K., Huang, F.-C., et al. (2016). Glucosylation of 4-Hydroxy-2,5-Dimethyl-3(2H)-Furanone, the key strawberry flavor compound in strawberry fruit. *Plant Physiol.* 171:139. doi: 10.1104/pp.16.00226

Tang, Y., Liu, X., Wang, J., Li, M., Wang, Q., Tian, F., et al. (2016). GAPIT Version 2: an enhanced integrated tool for genomic association and prediction. *Plant Genome* 9. doi: 10.3835/plantgenome2015.11.0120

Tikunov, Y. M., Laptenok, S., Hall, R. D., Bovy, A., and De Vos, R. C. H. (2012). MSClust: a tool for unsupervised mass spectra extraction of chromatography-mass spectrometry ion-wise aligned data. *Metabolomics* 8, 714–718. doi: 10.1007/s11306-011-0368-2

Ulrich, D., and Olbricht, K. (2013). Diversity of volatile patterns in sixteen Fragaria vesca L. accessions in comparison to cultivars of Fragaria× ananassa. *J. Appl. Bot. Food Qual.* 86, 37–46.

Ulrich, D., and Olbricht, K. (2014). Diversity of metabolite patterns and sensory characters in wild and cultivated strawberries. *J. Berry Res.* 4, 11–17. doi: 10.3233/JBR-140067

Ulrich, D., and Olbricht, K. (2016). A search for the ideal flavor of strawberry-Comparison of consumer acceptance and metabolite patterns in Fragaria× ananassa Duch. *J. Appl. Bot. Food Qual.* 89, 223–234.

Ulrich, D., Hoberg, E., Rapp, A., and Kecke, S. (1997). Analysis of strawberry flavour – discrimination of aroma types by quantification of volatile compounds. *Z. Lebensm. Unters. Forsch. A* 205, 218–223. doi: 10.1007/s002170050154

Urrutia, M., Rambla, J. L., Alexiou, K. G., Granell, A., and Monfort, A. (2017). Genetic analysis of the wild strawberry (*Fragaria vesca*) volatile composition. *Plant Physiol. Biochem.* 121, 99–117. doi: 10.1016/j.plaphy.2017.10.015

van Dijk, T., Pagliarani, G., Pikunova, A., Noordijk, Y., Yilmaz-Temel, H., Meulenbroek, B., et al. (2014). Genomic rearrangements and signatures of breeding in the allo-octoploid strawberry as revealed through an allele dose based SSR linkage map. *BMC Plant Biol.* 14:55. doi: 10.1186/1471-2229-14-55

Van Haandel, M. J. H., Sarabèr, F. C. E., Boersma, M. G., Laane, C., Fleming, Y., Weenen, H., et al. (2000). Characterization of different commercial soybean peroxidase preparations and use of the enzyme for N-Demethylation of Methyl N-Methylanthranilate to produce the food flavor methylanthranilate. *J. Agric. Food Chem.* 48, 1949–1954. doi: 10.1021/jf9909656

Vandendriessche, T., Geerts, P., Membrebe, B. N., Keulemans, J., Nicolaï, B. M., and Hertog, M. L. A. T. M. (2013). Journeys through aroma space: a novel approach towards the selection of aroma-enriched strawberry cultivars in breeding programmes. *Plant Breed.* 132, 217–223. doi: 10.1111/pbr.12028

Verma, S., Bassil, N. V., Van De Weg, E., Harrison, R. J., Monfort, A., Hidalgo, J. M., et al. (2017). *Development and Evaluation of the Axiom® IStraw35 384HT Array for the allo-octoploid Cultivated Strawberry Fragaria ×Ananassa*. Leuven: International Society for Horticultural Science (ISHS), 75–82. doi: 10.17660/ActaHortic.2017.1156.10

Wehrens, R., Hageman, J. A., Van Eeuwijk, F., Kooke, R., Flood, P. J., Wijnker, E., et al. (2016). Improved batch correction in untargeted MS-based metabolomics. *Metabolomics* 12:88. doi: 10.1007/s11306-016-1015-8

Wein, M., Lavid, N., Lunkenbein, S., Lewinsohn, E., Schwab, W., and Kaldenhoff, R. (2002). Isolation, cloning and expression of a multifunctional O-methyltransferase capable of forming 2,5-dimethyl-4-methoxy-3(2H)-furanone, one of the key aroma compounds in strawberry fruits. *Plant J.* 31, 755–765. doi: 10.1046/j.1365-313X.2002.01396.x

Whitaker, V. M., Hasing, T., Chandler, C. K., Plotto, A., and Baldwin, E. (2011). Historical trends in strawberry fruit quality revealed by a trial of University of Florida cultivars and advanced selections. *HortScience* 46, 553–557. doi: 10.21273/HORTSCI.46.4.553

Yamada, A., Ishiuchi, K. I., Makino, T., Mizukami, H., and Terasaka, K. (2019). A glucosyltransferase specific for 4-hydroxy-2,5-dimethyl-3(2H)-furanone in strawberry. *Biosci. Biotechnol. Biochem.* 83, 106–113. doi: 10.1080/09168451.2018.1524706

Yu, Y., Bai, J., Chen, C., Plotto, A., Yu, Q., Baldwin, E. A., et al. (2017). Identification of QTLs controlling aroma volatiles using a 'Fortune' x 'Murcott' (*Citrus reticulata*) population. *BMC Genomics* 18:646–646. doi: 10.1186/s12864-017-4043-5

Zorrilla-Fontanesi, Y., Rambla, J.-L., Cabeza, A., Medina, J. J., Sánchez-Sevilla, J. F., Valpuesta, V., et al. (2012). Genetic analysis of strawberry fruit aroma and identification of O-Methyltransferase FaOMT as the locus controlling natural variation in mesifurane content. *Plant Physiol.* 159:851. doi: 10.1104/pp.111.188318

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 02 July 2021
doi: 10.3389/fpls.2021.667809

Comparative Study of Fig Volatile Compounds Using Headspace Solid-Phase Microextraction-Gas Chromatography/Mass Spectrometry: Effects of Cultivars and Ripening Stages

Kahina Zidi[1]*, *Djamel Edine Kati*[1], *Mostapha Bachir-bey*[1], *Manon Genva*[2] and *Marie-Laure Fauconnier*[2]

[1] Laboratoire de Biochimie Appliquée, Faculté des Sciences de la Nature et de la Vie, Université de Bejaia, Bejaia, Algeria,
[2] Laboratory of Chemistry of Natural Molecules, Gembloux Agro-Bio Tech, University of Liège, Gembloux, Belgium

OPEN ACCESS

Edited by:
María José Jordán,
Murcian Institute for Agrarian and Food Research and Development (IMIDA), Spain

Reviewed by:
Joaquín J. Salas,
Instituto de la Grasa (IG), Spain
Barbara Sgorbini,
University of Turin, Italy

***Correspondence:**
Kahina Zidi
zidikahinaa@gmail.com

Specialty section:
This article was submitted to Plant Metabolism and Chemodiversity, a section of the journal Frontiers in Plant Science

Received: 14 February 2021
Accepted: 07 June 2021
Published: 02 July 2021

Citation:
Zidi K, Kati DE, Bachir-bey M, Genva M and Fauconnier M-L (2021) Comparative Study of Fig Volatile Compounds Using Headspace Solid-Phase Microextraction-Gas Chromatography/Mass Spectrometry: Effects of Cultivars and Ripening Stages. Front. Plant Sci. 12:667809. doi: 10.3389/fpls.2021.667809

Aroma is one of the essential parameters that determine fruit quality. It is also an important feature of varietal characterization and so valuable for agro-biodiversity identification and preservation. In order to characterize changes in the aroma fingerprint through fig development, the main objective of the present research was to study the volatile organic compound (VOC) profiles of figs (*Ficus carica* L.) from three cultivars, Taamriwthe (TH), Azegzaw (AZ), and Averkane (AV), at three ripening stages (unripe, ripe, and fully ripe). Analyses was performed using Headspace Solid-phase Microextraction and gas chromatography coupled with mass spectrometry. Results revealed the presence of 29 compounds that were grouped into different chemical classes. Aldehydes comprised the most abundant VOCs identified in all the studied figs, while alcohols, ketones, and terpenes comprised the minor compounds found in TH, AZ, and AV figs, respectively. Different aroma descriptors were identified throughout the ripening stages of figs; fruity and green aromas were dominant in all cultivars, while a fatty aroma scarcely occurred in figs. A gallery plot representation demonstrated that certain VOCs differentiate the studied cultivars and the different ripening stages of figs. Principal component analysis findings demonstrated characteristic VOCs of distinct ripening stages and cultivars, those VOCs can be used as fingerprints to distinguish different cultivars and/or ripening stages.

Keywords: ***Ficus carica* L., volatile organic compounds, ripening stages, head-space solid-phase microextraction, aroma, agrobiodiversity**

INTRODUCTION

Figs (*Ficus carica* L.), which have always been appreciated for their sweet taste and high nutritional value, are an essential constituent of the Mediterranean diet alongside olive oil, cereals, and vegetables. The global fig production is around 1 million tons per year, with Turkey being the largest producer, followed by Egypt, Morocco, and Algeria (FAO, 2018). Figs are characterized by an inflorescence called a syconium and are rich in minerals, vitamins, and fiber. They are

also cholesterol and fat-free (Solomon et al., 2006; Palmeira et al., 2019) and have antioxidant, antimicrobial, antispasmodic, and anticancer properties (Chawla et al., 2012; Barolo et al., 2014; Palmeira et al., 2019). Figs are also an important source of multiple bioactive compounds that have several uses in nutrition and therapy (Palmeira et al., 2019). They are consumed fresh, dried, or transformed into numerous other products such as jams, beverages, and confectionery (Zidi et al., 2020).

Several changes occur during fruit ripening, including physiological, biochemical, and organoleptic modifications. As examples, ethylene production increases, chlorophyll is degraded, anthocyanin content changes, and flavor and aroma modifications are observed throughout the ripening process. Also, a degradation of the cell wall structure during fruit ripening leads to the solubilization of polysaccharides such as pectin and cellulose and induces an increase in fruit sweetness (Prasanna et al., 2007).

Volatile organic compounds (VOCs) are involved in a wide range of functions. For example, VOCs contribute to fruit aroma and to protection against herbivores, microbial growth, and abiotic stress (Schwab et al., 2008). In addition, these compounds attract specific fig pollinators, such as *Blastophaga psenes*, to ensure flower pollination, enabling the development and growth of the fruit (Ware et al., 1993). Fruit VOCs also attract animals that contribute to seed dissemination. Volatile organic compounds that contribute to fruit aroma and flavor are generally present in both free and bound forms. The free VOCs directly contribute to fruit aroma while the bound form is considered to be a non-volatile precursor linked by a β-glycosidic linkage to mono or disaccharides (Hjelmeland and Ebeler, 2015). VOCs are influenced by fruit growth and ripening stages, leading to qualitative and quantitative changes (Iban et al., 2000), which are closely related to the release of VOCs that are often sequestered in glycosylated form (Chen et al., 2021). Thus, an aroma pattern might be used as a marker to discriminate between fruit maturity stages, but it can also be used to identify different fruit origins (genetic or geographic) (Khalil et al., 2017).

Even though figs are one of the oldest cultivated fruits, the changes that occur during the ripening process remain poorly studied. The most widespread methods of fig characterization are phenotypic and genotypic analyses. However, metabolomics, which considers secondary metabolites as the final stage of plant physiological response in a given environment, may be a more appropriate approach. In this context, this study aims to investigate the natural dynamics of the VOC profile throughout the figs' maturation and to demonstrate changes in the key aroma compounds between cultivars during the ripening stages. We thus developed a performant method to analyze fig VOCs, as this fruit emits low quantities of VOCs. Among the different methods tested, Headspace Solid-phase Microextraction (HS-SPME) was chosen for its high sensitivity and extraction efficiency (Elmore et al., 2000; Merkle et al., 2015). In the present study, Principal Component Analysis (PCA) was also conducted in order to determine the VOCs that characterize each cultivar, as well as those determining each ripening stage (unripe, ripe, and fully ripe).

MATERIALS AND METHODS

Plant Material

Fig samples were harvested in September 2018 from an orchard in Beni Maouche, Bejaia Province, in the north of Algeria. Three different fig cultivars were collected for this investigation: Taamriwthe (TH), Azegzaw (AZ), and Averkane (AV). The three ripening stages (unripe, ripe, and fully ripe) were distinguished by visual maturity characteristics based on color change and fruit firmness (**Figure 1**).

For TH figs, the unripe stage (TH1) was identified by its yellow color and the firmness of the fruit, whereas the ripened fruit (TH2) tended to be softer and yellow on the surface. Fully ripe (TH3) figs tended to lose their firmness and the surface cracked slightly. For AZ figs, unripe (AZ1) fruit were always firm with green-yellow skin, while ripe ones (AZ2) were soft with a dark-green color. Fully ripe figs (AZ3) became softer with cracking skin. For AV figs, unripe fruit (AV1) were identified as having purple-green skin and being hard; as they began to ripen (AV2) they slightly lost their firmness and became purple. When these figs reached the fully ripe stage (AV3), the fruit turned dark and lost its firmness.

Fig fruits were selected according to their color, size, and absence of physical damage. Then they were frozen (−20°C). Samples were homogenized with a blender (IKA A11 Basic) before analyses.

Headspace Solid-phase Microextraction (HS-SPME)

Several methods can be used for sample comparison by SPME, including in-fiber standardization and traditional calibration methods using external and internal standards (Wang et al., 2005; Bicchi et al., 2011). In the present research, the relative contents of the identified VOCs were presented as normalized peak areas.

Optimization of the parameters affecting SPME efficiency was realized according to Pawliszyn (2000). The sample weight, salt concentration, equilibrium time, extraction time, and extraction temperature were all optimized in order to reach the linear headspace conditions.

SPME Fiber

A Divinylbenzene/Carboxen/Polydimethylsiloxane (DVB/CAR/PDMS) stationary phase 50/30 μm, 1 cm length fiber was used for SPME (Supelco, Bellefonte, PA, USA). This fiber was chosen because of its efficiency in extracting a wide range of VOCs according to their molecular weight (40–275 g/mol) and polarity (Elmore et al., 2000).

The fiber was conditioned at 270°C for 30 min before first use. The equilibration step was performed in an HS-SPME vial of 20 mL containing 3 g of blended fruit samples for 25 min at 40°C. Then the fiber was exposed to the headspace for another 25 min at the same temperature for the adsorption of VOCs. In order to optimize the VOC volatility and the extraction yield, 10% of NaCl (w/w) was added with agitation (250 rpm) in the same conditions (Fiorini et al., 2015). 1,2,3-trichloropropane (3 μL) was used as an internal standard. The HS-SPME procedure was achieved using a GERSTEL MSP 2 Multipurpose Sampler (GERSTEL GmbH &

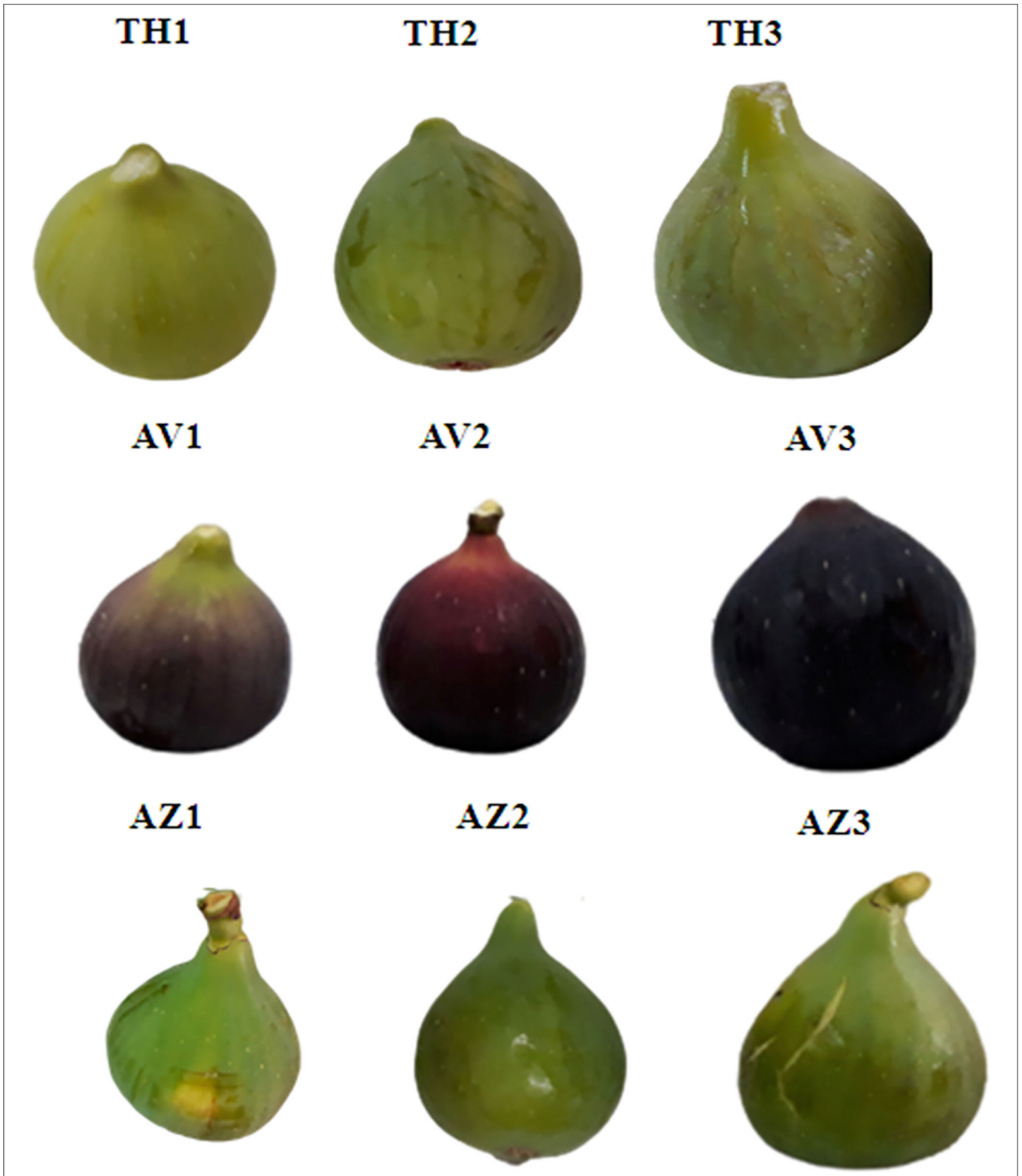

FIGURE 1 | Photographs of fig cultivars at the different ripening stages. Photograph credit: Zidi Kahina, Abderrahmane Mira University. (TH) Taamriwthe, (AV) Averkane, and (AZ) Azegzaw figs; 1, unripe figs; 2, ripe figs; 3, fully ripe figs.

Co. KG, Mülheim, Germany). The fiber was then desorbed for 1 min into the GC injector at 250°C.

Gas Chromatography Coupled With Mass Spectrometry

The volatile compound analyses was performed according to the method developed by Tanoh et al. (2020) with some modification, using gas chromatography coupled with mass spectrometry on an Agilent 7890A GC system (Agilent, Santa Clara, CA, USA) equipped with a 5975C Inert XL EI/CI/MSD detector at 70 eV. Volatile molecules were separated using a capillary column, VF-WAX ms 30 × 0.25 mm and 0.25 μm film thickness (Agilent Technologies, Santa Clara, CA, USA), with 1.3 mL/min at a constant flow rate of carrier gas (helium), in splitless injection mode.

The oven temperature was set at 40°C for 1 min then increased by 5°C/min up to 220°C, 10°C/min to 250°C, and held at 250°C for 5 min. The temperatures of the source and the quadrupole were, respectively, 230 and 150°C, with a scanned mass between 35 and 500 amu. Each component was identified by comparing the obtained mass spectra with the Wiley (Pal 600K®) data library. Further identification was realized by calculating the retention indices (RIs) using a standard mixture of *n*-alkanes, C_7-C_{30} 1,000 μg/mL in hexane (Supelco, Bornem, Belgium), in the same chromatographic conditions. Retention indices were determined according to the equation described by Babushok et al. (2011). The calculated RI was compared with those in the literature on the NIST, PubChem, or Pherobase databases. The data were established by Mass Hunter Workstation Software (Version B.08.00, Agilent Technologies, Inc. 2016, Santa Clara, CA, USA). Results were expressed as relative content of the identified compounds that were normalized by the internal standard. The analyses were performed on three independent biological replicates for each cultivar and each ripening stage. A representative chromatogram was presented in **Figure 2**.

Statistical Analyses

Differences in the evolution of fig volatile profiles during the three ripening stages were assessed by analyses of variance with one factor (ANOVA), performed using STATISTICA Software 7.1 (Stat Soft, Maisons-Alfort, France). Mean values were analyzed by LSD (least significant difference) at a significant level of $P < 0.05$. Principal components analysis (PCA) was analyzed using XLSTAT 5.01. A spider chart, drawn with Microsoft Office Excel 2013 (Microsoft Corporation, Redmond, WA, USA), was used to illustrate aroma profiles.

RESULTS

Evolution of VOC Profile During Ripening

As shown in **Table 1**, HS-SPME-GC/MS analyses detected 29 compounds in TH, AZ, and AV cultivars during the different ripening stages. Among these, seven alcohols, eight aldehydes, five esters, six terpenes, and three ketones were found to differ significantly, depending on cultivar and ripening stage. The evolution of VOCs during ripening for each fig cultivar is represented by chemical classes in **Figure 3**.

The results showed qualitative differences between the different chemical classes depending on the fig cultivars and ripening stages. Aldehydes were the most abundant compounds in TH figs while large quantities of both aldehydes and terpenes were present in AZ figs. Aldehydes were distinctive compounds

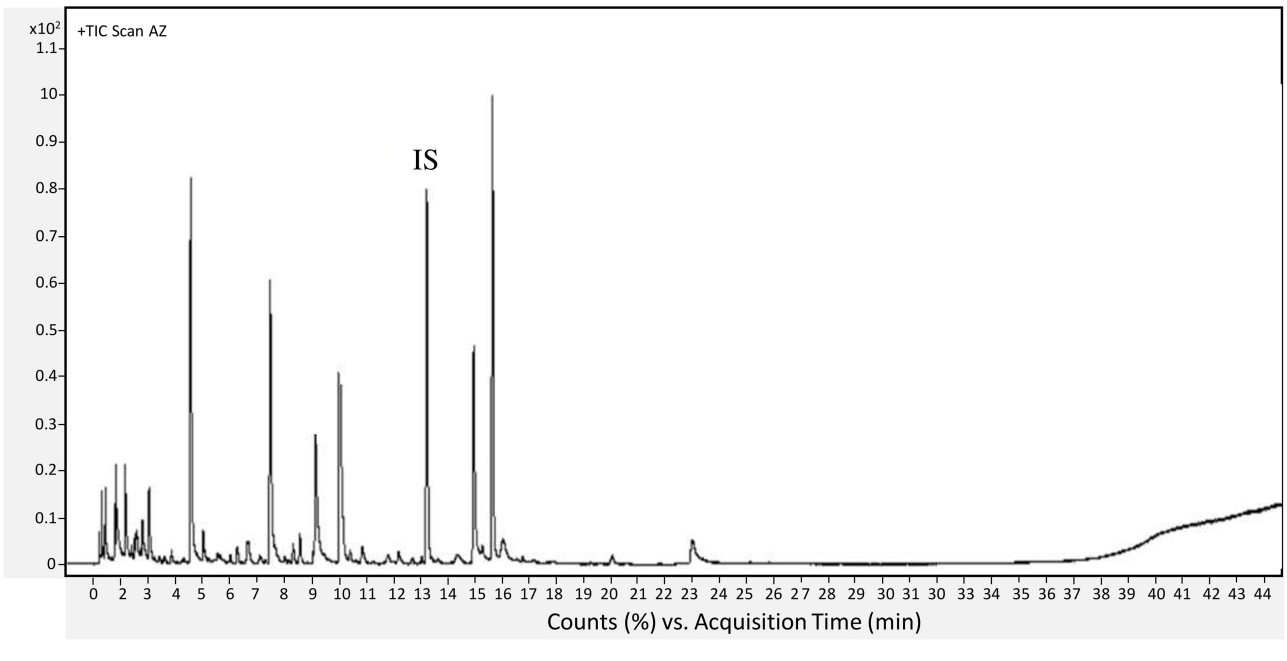

FIGURE 2 | Representative chromatogram of VOCs obtained form AZ figs at unripe stage. IS, internal standard.

TABLE 1 | Volatile organic compounds (VOCs) identified by HS/SPME coupled with GC/MS during three ripening stages of different fig cultivars.

	Volatile compounds	CAS number	Aroma description[A]	RI[B]	RI[C]	Odor threshold in water (μg/kg)[D]	Taamriwthe			Azegzaw			Averkane		
							Unripe	Ripe	Fully ripe	Unripe	Ripe	Fully ripe	Unripe	Ripe	Fully ripe
Alcohols															
1	Pent-1-en-3-ol	616-25-1	Pungent, like, green vegetable and tropical fruity nuances	1,155	1,155	400[1]	nd	0.034 ± 0.006	nd	0.035 ± 0.002[a]	0.026 ± 0.005[b]	0.020 ± 0.002[b]	nd	nd	0.035 ± 0.013
2	Hexan-1-ol	111-27-3	Pungent, ethereal, fruity and alcoholic, sweet with a green top note	1,348	1,347	500[2]	nd	nd	nd	0.026 ± 0.002[a]	0.024 ± 0.003[a]	0.024 ± 0.002[a]	nd	nd	nd
3	(*E*)-hex-2-en-1-ol	928-95-0	Fresh fatty green, fruity, vegetative, with leafy and herbal nuances	1,401	1,401	100[2]	nd	nd	nd	nd	0.020 ± 0.002	nd	nd	nd	nd
4	Benzyl alcohol	100-51-6	Sweet, floral, fruity with chemical nuances	1,868	1,868	10,000[3]	nd	nd	nd	0.166 ± 0.043	nd	nd	nd	nd	nd
5	Pentan-1-ol	71-41-0	Pungent, fermentedbready, yeasty, winey	1,249	1,249	nd	nd	nd	nd	nd	nd	nd	nd	0.020 ± 0.007	nd
6	3-methylbutan-1-ol	123-51-3	Fusel, alcoholic, pungent, ethereal, cognac, fruity	1,245	1,240	nd	nd	0.050 ± 0.009[b]	0.085 ± 0.017[a]	0.028 ± 0.002[a]	nd	0.037 ± 0.007[a]	nd	nd	0.032 ± 0.003
7	2-ethylhexan-1-ol	104-76-7	Citrus fresh, floral oily sweet	1,494	1,494	nd	0.020 ± 0.003	nd	nd	0.050 ± 0.011[a]	0.054 ± 0.007[a]	0.028 ± 0.005[b]	nd	nd	0.033 ± 0.011
Aldehydes															
8	Acetaldehyde	75-07-0	Pungent, ethereal, fresh, lifting, penetrating, fruity	nd	714	15[4]	0.124 ± 0.03[b]	0.148 ± 0.032[b]	0.273 ± 0.044[a]	0.075 ± 0.002[b]	0.064 ± 0.005[b]	0.126 ± 0.030[a]	nd	0.058 ± 0.011[b]	0.164 ± 0.043[a]
9	Heptanal	111-71-7	Fresh, green, citrus odor	nd	1,181	3[4]	nd	nd	0.032 ± 0.003	nd	nd	nd	nd	nd	nd
10	Nonanal	124-19-6	Citrus, with a fresh slightly green lemon	1,385	1,385	1–98[5]	nd	0.020 ± 0.018	nd	0.026 ± 0.007[a]	nd	0.026 ± 0.007[a]	nd	nd	nd
11	Benzaldehyde	100-52-7	Almond, fruity, powdery, nutty	1,512	1,513	350[4]	nd	0.093 ± 0.007[a]	0.075 ± 0.020[a]	0.937 ± 0.22[a]	0.424 ± 0.124[b]	0.350 ± 0.07[b]	nd	nd	0.090 ± 0.028
12	Hexanal	66-25-1	Green, fatty, leafy, vegetative, fruity and clean with a woody nuance	1,074	1,074	5[4]	0.311 ± 0.04[b]	0.553 ± 0.042[a]	0.359 ± 0.077[b]	0.788 ± 0.054[a]	0.330 ± 0.056[b]	0.258 ± 0.066[b]	0.226 ± 0.009[a]	0.152 ± 0.022[b]	0.141 ± 0.024[b]
13	(*E*)-2-methylbut-2-enal	497-03-0	Strong green fruit	1,084	1,090	nd	0.102 ± 0.031[ab]	0.157 ± 0.008[a]	0.078 ± 0.011[b]	nd	nd	nd	nd	nd	nd
14	(*E*)-hex-2-enal	6728-26-3	Green banana aldehydic fatty cheesy	1,217	1,217	17–316[5]	0.115 ± 0.001[b]	0.214 ± 0.039[ab]	0.311 ± 0.073[a]	nd	nd	nd	nd	0.050 ± 0.011[b]	0.175 ± 0.050[a]

(Continued)

TABLE 1 | Continued

	Volatile compounds	CAS number	Aroma description[A]	RI[B]	RI[C]	Odor threshold in water (μg/kg)[D]	Taamriwthe			Azegzaw			Averkane		
							Unripe	Ripe	Fully ripe	Unripe	Ripe	Fully ripe	Unripe	Ripe	Fully ripe
15	(*E*)-oct-2-enal	2548-87-0	Fresh cucumber fatty green herbal banana waxy green leaf	1,421	1,421	3[6]	nd	nd	nd	0.016 ± 0.002	nd	nd	nd	nd	nd
Esters															
16	Ethyl acetate	141-78-6	Ethereal, fruity, sweet, grape and rum-like	nd	899	5,000[6]	0.216 ± 0.038[a]	0.159 ± 0.039[a]	0.135 ± 0.098[a]	0.137 ± 0.015	nd	nd	nd	0.198 ± 0.041[b]	0.333 ± 0.090[a]
17	Methyl hexanoate	106-70-7	Fusel, fermented, fruity, banana, ethereal	1,176	1,177	84[7]	nd	0.093 ± 0.020	nd	nd	0.035 ± 0.003	nd	nd	nd	nd
18	Methyl octanoate	111-11-5	Green, sweet, orange with vegetative nuances	1,383	1,383	nd	nd	nd	nd	nd	0.026 ± 0.002	nd	nd	nd	0.067 ± 0.022
19	Ethyl 2-methylbutanoate	7452-79-1	Fruity, and berry with fresh tropical nuances	1,041	1,041	nd	nd	nd	0.037 ± 0.008	nd	nd	nd	nd	nd	nd
20	Ethyl hexanoate	123-66-0	Sweet fruity pineapple	1,246	1,246	14[2]	nd	nd	0.034 ± 0.007	nd	nd	nd	nd	nd	nd
Terpenes															
21	Limonene	5989-27-5	Citrus orange fresh sweet	1,180	1,180	10[2]	nd	nd	nd	nd	0.215 ± 0.032[a]	0.109 ± 0.024[b]	nd	nd	nd
22	Linalool	78-70-6	Citrus, orange, floral	1,541	1,541	6[7]	nd	0.021 ± 0.002	nd	1.049 ± 0.18[a]	0.498 ± 0.113[b]	0.892 ± 0.120[a]	nd	nd	0.043 ± 0.013
23	Epoxylinalol	14049-11-7	Floral honey	1,728	1,729	nd	nd	0.025 ± 0.004	nd	nd	nd	0.028 ± 0.009	nd	nd	nd
24	α-Santalene	512-61-8	Woody	1,615	1,608	nd	nd	0.033 ± 0.002	nd	nd	nd	nd	nd	nd	nd
25	1,8-cineol	470-82-6	Eucalyptus herbal camphor medicinal	1,189	1,190	1,3–12[8]	nd	nd	nd	nd	nd	nd	0.030 ± 0.011	nd	nd
26	ß-caryophyllene	87-44-5	Sweet woody spice clove dry	1,579	1,579	150[1]	0.292 ± 0.045[a]	0.254 ± 0.013[a]	nd	nd	nd	nd	nd	nd	nd
Ketones															
27	Acetone	67-64-1	Solvent ethereal apple pear	nd	813	nd	nd	0.059 ± 0.016[b]	0.100 ± 0.017[a]	0.034 ± 0.002[a]	0.050 ± 0.004[a]	0.045 ± 0.016[a]	0.116 ± 0.016[a]	0.079 ± 0.040[ab]	0.043 ± 0.018[b]
28	Heptan-2-one	110-43-0	Cheese, fruity, ketonic, green banana, with a creamy nuance	1,173	1,173	140[4]	nd	nd	0.073 ± 0.027	nd	nd	nd	0.107 ± 0.016[b]	nd	0.152 ± 0.018[a]
29	6-methylhept-5-en-2-one	110-93-0	Fruity, apple, musty, ketonic and creamy with slight cheesy and banana nuances	1,330	1,330	160[1]	nd	nd	nd	0.022 ± 0.003[a]	0.022 ± 0.003[a]	nd	nd	nd	nd

[A]Aroma description was obtained from literature data (http://www.thegoodscentscompany.com/); [B]RI, calculated retention indices; [C]RI, theoretical retention indices (Pubchem, NIST, and the Pherobase); nd, not detected; [D]All the odor thresholds were obtained from: [1]Tamura et al. (2001); [2]Yang et al. (2019); [3]Noguerol-Pato et al. (2012); [4]Buttery et al. (1988); [5]Buttery et al. (1969); [6]Wu et al. (2016); [7]Takeoka et al. (1990); [8]Czerny et al. (2008).
Results are expressed as normalized peak areas. Volatiles are listed in different chemical classes. Values are means ± SD of three samples analyzed in triplicate; results in the same line of each cultivar in the three ripening stages represented with different letters are significantly ($P < 0.05$) different ($a > b > c$).

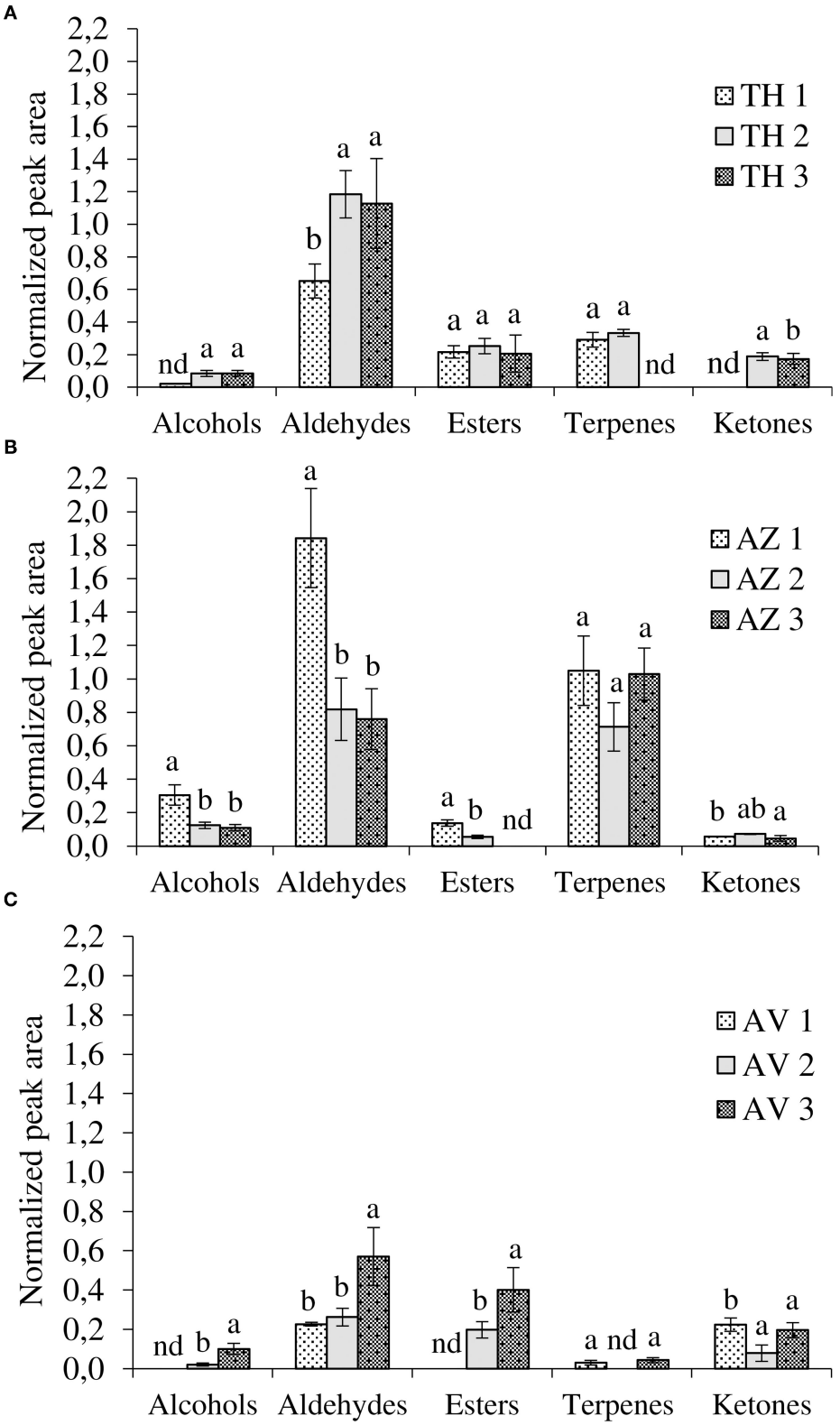

FIGURE 3 | Evolution of volatile compounds during ripening according to chemical classes. **(A)** Taamriwthe, **(B)** Azegzaw, and **(C)** Averkane; 1, unripe figs; 2, ripe figs; 3, fully ripe figs; nd, not detected. Results for each VOC group from the same cultivar with different letters are statistically different (ANOVA-LSD test; $P < 0.05$; a > b).

of AV figs. Moreover, significant differences in the proportion of chemical classes in the same cultivar, depending on fig ripening state, were also noted.

Aldehyde Profile

Aldehydes were the most abundant VOCs, demonstrated by their proportion of 50 ± 0.59%, 40.26 ± 0.27%, and 29 ± 1.20% of total VOCs for TH, AZ, and AV figs, respectively. Both TH and AV figs showed a significant increase in aldehyde content during the three ripening stages (**Figures 3A,C**), whereas AZ figs demonstrated a significant decline in aldehydes ($P < 0.05$). Six aldehydes were identified in cultivar TH, of which (*E*)-2-methylbut-2-enal contributes to the strong green fruit notes detected (**Table 1**). This aldehyde was only extracted from TH figs, and clearly decreased ($P < 0.05$) during the ripening process. During the development of TH figs the level of (*E*)-hex-2-enal increased 3-fold compared to the early ripening stage. Almost the same aldehydes were found in cultivar AZ with the exception of heptanal and (*E*)-oct-2-enal. In the case of the AV cultivar, only four aldehydes were extracted and a high content of (*E*)-hex-2-enal (31%) was detected.

Hexanal and benzaldehyde were the most abundant aldehydes in TH and AZ figs. All figs showed a significant decrease ($P < 0.05$) in hexanal during the studied stages. Acetaldehyde was identified in all fig cultivars and demonstrated a significant ($P < 0.05$) 2-fold increase from the beginning to the final stage of ripening.

Terpene Profile

The general trend for terpene was constant during fig ripening. The terpenes extracted from cultivar TH were present only at the ripe stage. However, ß-caryophyllene content increased significantly ($P < 0.05$) from the unripe to the ripe stage, when it was 76% higher. Limonene, epoxylinalol, and linalool were identified in cultivar AZ; it was observed that linalool was the most abundant terpene in AZ figs (87 ± 9.50%) and increased significantly ($P < 0.05$) throughout the ripening process. Only two terpenes were extracted from AV figs (linalool and 1,8-cineole). 1,8-cineole has an herbal eucalyptus odor with an odor threshold that varies from 1.3 to 12 μg/kg (**Table 1**).

Ester Profile

A significant increase ($P < 0.05$) in esters was recorded during AV fig ripening. The TH and AV cultivars had higher total esters content, with proportions of 19 ± 0.48% and 25 ± 2.34%, respectively, whereas AZ figs contained lower proportions of esters (around 3%). The esters identified differed between fig cultivars. Four esters were detected in cultivar TH; ethyl acetate was the most abundant with 76 ± 10.60% of total esters and was present at a constant proportion from the unripe to the fully ripe stage. Ethyl 2-methylbutanoate and ethyl hexanoate were only detected in TH figs and made an essential contribution to the fruity aroma (**Table 1**). AV figs contained a low number of esters as only two compounds were identified (methyl octanoate, and ethyl acetate). Ethyl acetate was present in large quantities and significantly increased ($P < 0.05$) during the ripening of AV figs, leading to the perception of a fruity aroma with high odor thresholds (5,000 μg/kg) (**Table 1**).

Alcohol Profile

There were significant differences ($P < 0.05$) in the alcohols produced by the three cultivars. Alcohol content significantly increased during the ripening of AV cultivars (**Figures 3A,C**). However, in the case of AZ figs, a significant decrease in alcohols was observed (**Figure 3B**).

Three different alcohols were identified in TH figs: pent-1-en-3-ol, 3-methylbutan-1-ol, and 2-ethylhexan-1-ol. The most abundant compound was 3-methylbutan-1-ol, which increased during fruit ripening. Four alcohols were extracted from the AV cultivar, among these pentan-1-ol was only identified at the ripe stage (**Table 1**). In the case of AZ figs, six alcohols were identified, among which were pent-1-en-3-ol and 2-ethylhexan-1-ol. Both of these compounds decreased significantly from the unripe to the fully ripe stage in AZ figs. Benzyl alcohol and (*E*)-hex-2-en-1-ol were only detected in cultivar AZ, and were present during the unripe and ripe stages, respectively.

Ketone Profile

Three ketones were identified in this study and two of them were found in all fig cultivars, with significant differences in occurrence. During the ripening period acetone significantly increased in TH figs and decreased in AV figs ($P < 0.05$). On the other hand, heptan-2-one significantly increased ($P < 0.05$) in both cultivars.

Changes in Aroma Descriptors During Ripening

A spider chart representation was created using odor thresholds obtained from the literature (Buttery et al., 1969, 1988; Takeoka et al., 1990; Tamura et al., 2001; Czerny et al., 2008; Noguerol-Pato et al., 2012; Wu et al., 2016; Yang et al., 2019) and aroma descriptors (fruity, green, sweet, floral, and fatty) [1], represented in **Figure 4**, in order to better visualize changes in aroma descriptors during ripening.

The gallery plot in **Figure 5** presents an overview of the intensity of VOC expression for the three analyzed figs that makes it possible to differentiate between the cultivars and to produce fingerprints of the ripening stages for each fig studied. It can be observed that hexanal was the only compound identified at high relative proportion in all cultivars at all the ripening stages.

Principal Component Analyses of Ripening Stages and Cultivar Characteristics

PCA was performed in order to better visualize the VOCs that are characteristic of each cultivar as well as those determining ripening stages (unripe, ripe, and fully ripe). The PCA results showed that the two first axes explained a total of 91.24 % (F1: 63.49% and F2: 27.75%) and 96.59% (F1: 89.66% and F2: 6.93%) of the total variance (**Figures 6A,B**, respectively).

[1]The Good Scents Company. Available online at: http://www.thegoodscentscompany.com (accessed Jan 30, 2020).

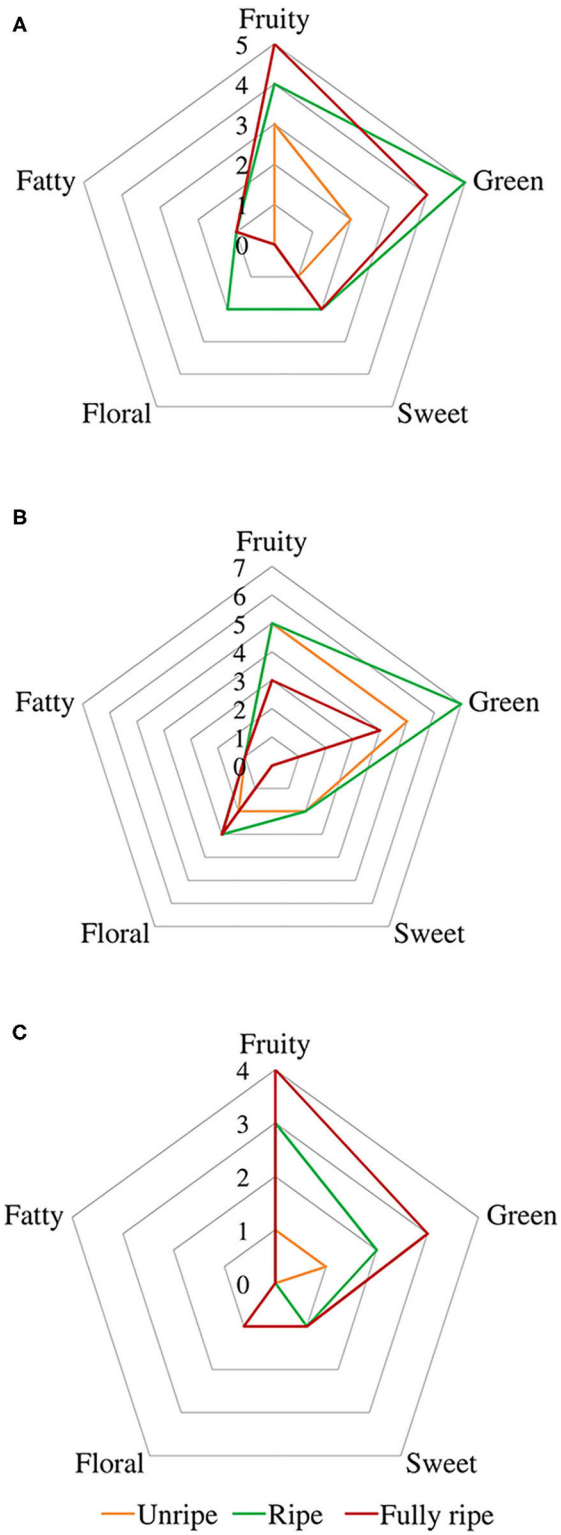

FIGURE 4 | Spider chart of aromatic descriptors from fig aroma based on their odor thresholds. Cultivars **(A)** Taamriwthe, **(B)** Azegzaw, and **(C)** Averkane. The presentation consists of equiangular spokes, where each spoke representing one descriptor, and the position of the spot on the spoke was proportional to the value of the corresponding descriptor.

DISCUSSION

While figs are highly appreciated worldwide for their sweet taste and their high nutritional value, their smell is very weak, which makes it difficult to analyze the volatile molecules emitted by these fruits. However, as VOC analyses is a parameter of interest that is commonly used to determine the quality of fruit, it is important to develop a sensitive method to characterize fig maturation based on the most relevant volatile molecules emitted. In the present study an efficient HS-SPME-GC-MS method has been developed, which identified 29 different VOCs, belonging to five different chemical classes, in the fig samples. This result is an improvement on the two previous methods (dynamic headspace and solvent extraction), which produced only a few results when compared to the SPME technique. Twenty VOCs were identified in both TH and AZ figs, it is notable that the TH fig is a dominant cultivar in Algerian orchards because it is highly valued in both fresh and dried form. Only 14 compounds were found in AV figs, highlighting the difference in the VOC profile of the different cultivars.

Several research already focused on the study of the volatile profile of fresh figs, dried figs, fig liquors, and fig spirits (Oliveira et al., 2010; Palassarou et al., 2017; Russo et al., 2017; Rodríguez-Solana et al., 2018). Nevertheless, to the best of our knowledge, no studies involving the VOCs changes observed during the ripening process of the studied cultivars have been previously performed.

In most fig samples, aldehydes represented the most abundant VOC detected, accounting for 22 ± 1.91–38.29 ± 2.17% of the total VOCs in TH figs, from 54.37 ± 1.03 to 22.10 ± 0.43% in AZ figs, and from 21.73 ± 4.02 to 53.52 ± 5.12% in AV figs. Acetaldehyde, benzaldehyde, hexanal, and (*E*)-hex-2-enal were the most important aldehyde molecules detected; these are important aroma compounds in figs, contributing to their fresh green odor and providing a fruity flavor due to their low odor thresholds (Schwab et al., 2008; Oliveira et al., 2010; Hou et al., 2020). The high aldehyde content in all the studied cultivars could be attributed to the abundance of unsaturated fatty acid precursors (Yang et al., 2011; Garcia et al., 2012b) in figs. Indeed, C_6 and C_9 aldehydes can be biosynthesized from unsaturated fatty acids via the lipoxygenase pathway (Dudareva et al., 2013). Unsaturated fatty acids are oxygenated and then cleaved into C_6 and C_9 aldehydes by hydroperoxide lyase (Fauconnier et al., 2008; Genva et al., 2019). As an example, it is known that (*E*)-hex-2-enal is biosynthesized from linolenic acid through the lipoxygenase pathway (Garcia et al., 2012a) and that this unsaturated fatty acid represents 53% of total lipids in some fig cultivars (Jeong and Lachance, 2001). (*E*)-hex-2-enal production is augmented due to the activity of lipoxygenase on membrane lipids during fruit ripening (Karabulut et al., 2018), which explains the increase of (*E*)-hex-2-enal observed in the present study in fully ripe TH and AV figs. Similar results have been observed by Song et al. (2018) during the maturation of jujube fruit. Benzaldehyde was also highly abundant in TH and AZ figs. This molecule is one of the major aromatic aldehydes that contribute greatly to fig aroma, and it is released from the oxidation of benzyl alcohol through the shikimic acid

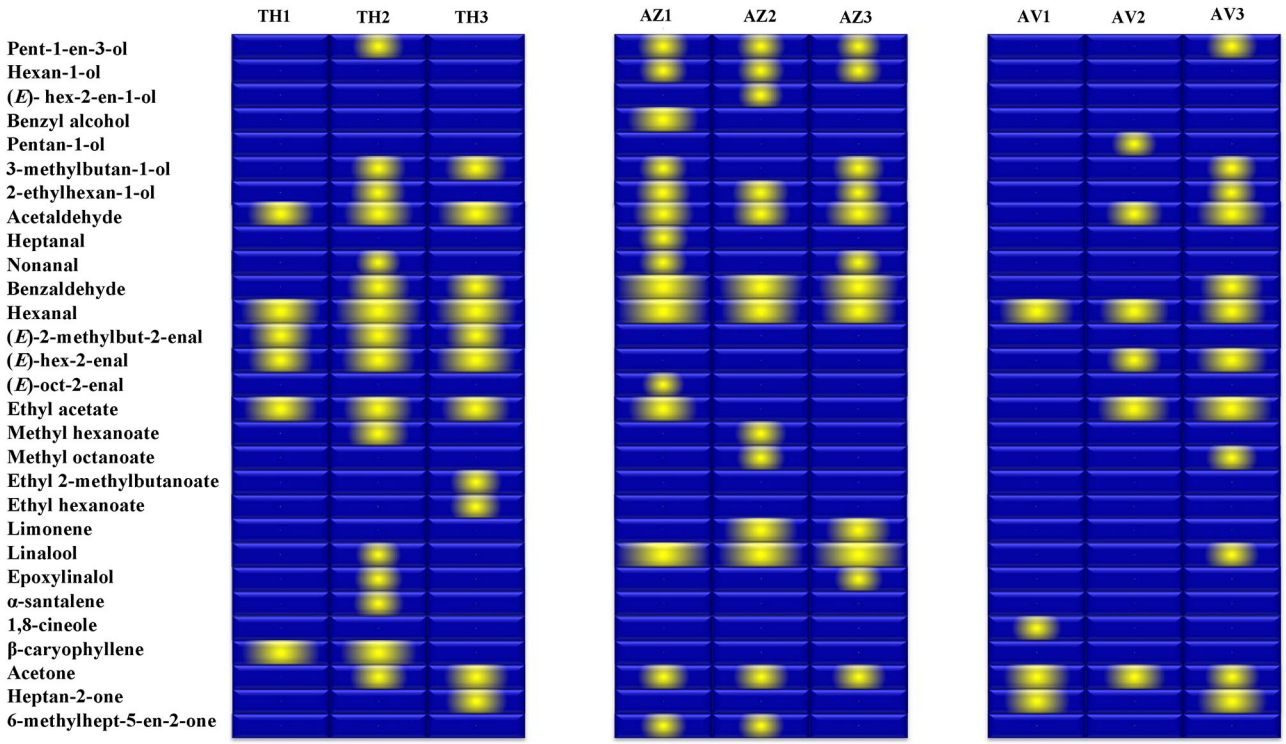

FIGURE 5 | Gallery plot of VOC expression intensity of TH (Taamriwthe), AZ (Azegzaw), and AV (Averkane) figs during ripening stages. 1, unripe figs; 2, ripe figs; 3, fully ripe figs.

pathway (Pereira et al., 2020). Similar abundance was reported by Palassarou et al. (2017) in the dried pulp of a Peloponnese fig cultivar. Also, a strong decrease in hexanal content was highlighted in all figs during ripening. This fall contributes to the diminution of the green aroma (Villalobos et al., 2018), and the same trend has been observed in avocado and kiwi fruit (Young and Paterson, 1995; Obenland et al., 2012). Moreover, for all studied cultivars, the acetaldehyde content increased from unripe to fully ripe figs. Pyruvate decarboxylase is an important enzyme responsible for the production of acetaldehyde, so the increase in this aldehyde can be explained by high levels of pyruvate decarboxylase enzyme activity (Echeverria et al., 2004). The increase in acetaldehyde during ripening in avocado fruits has also already been highlighted (Obenland et al., 2012) and, according to Dixon and Hewett (2000), acetaldehyde production generally occurs during the fruit ripening process. Gozlekci (2010) considers aldehydes to be among the most important chemical classes contributing to the aroma of *F. carica*. The observed decrease in total aldehydes during the later growth stage of AZ figs, compared to the other cultivars, could be due to genetic differences. Indeed, Oliveira et al. (2012) confirmed a correlation between genetic diversity and the fig metabolome. A similar decrease had already been observed by Young and Paterson (1985) and Wang et al. (2011), in kiwi fruit.

While terpenes were produced in low proportions by cultivars TH (46 ± 3.84%) and AZ (37.40 ± 1.37%), these molecules were produced in higher proportions in AV figs (58 ± 1.4%) which impacted AV figs aroma notes. Indeed, AV figs are characterized by aromatic notes provided by terpenes in comparison to the other cultivars. Terpenes are well-known for their contribution to pleasant sensory notes, as they are characterized by herbaceous, fruity, citrus, and floral scents (Arem et al., 2011; Russo et al., 2017). The molecules are synthesized by two independent pathways, the first being the mevalonate pathway, which produces volatile sesquiterpenes (C_{15}). The second is the methylerythritol phosphate pathway which contributes to the production of hemiterpenes (C_5), monoterpenes (C_{10}), and diterpenes (C_{20}) (Dudareva et al., 2013; El Hadi et al., 2013). During the ripening period, the terpene content of TH figs decreased markedly (46 ± 3.84–0%) while it remained constant ($P < 0.05$) in AZ (changing from 37.40 ± 1.37 to 37.06 ± 2.56%) and AV figs (changing from 41.11 ± 4.43 to 58.89 ± 10.40%) in AZ figs, which were characterized by high terpene contents. The major terpene molecule was linalool, which is recognized for its floral scents and has several biological properties (anti-inflammatory, antioxidant, and analgesic) (Aprotosoaie et al., 2014). This molecule is an important factor in the establishment of fig cultivar fingerprints, as all previous studies on fresh fig volatile molecules reported low terpene contents (Oliveira et al., 2010; Villalobos et al., 2018). The low terpene production seen in fully ripe TH figs may be due to genetic differences among the three different cultivars; for this reason, it would be interesting to perform a genomic comparison of the three cultivars studied in the present research.

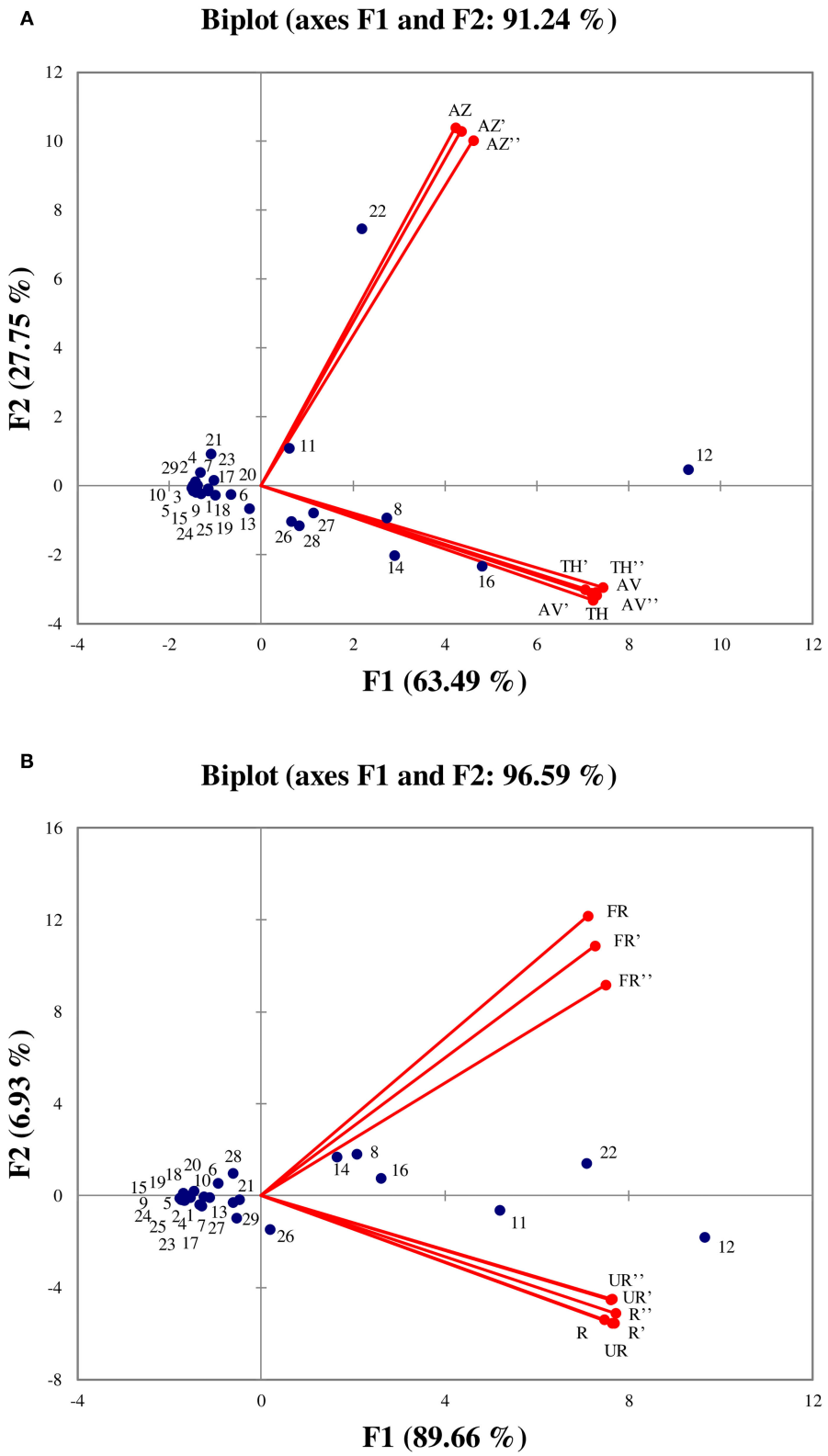

FIGURE 6 | Principal components analysis performed on all identified volatile organic compounds according to the three cultivars **(A)**, and between ripening stages **(B)**. (TH) Taamriwthe, (AV) Averkane, and (AZ) Azegzaw figs. The numbers correspond to the VOCs listed in **Table 1**. U, unripe figs; R, ripe figs; FR, fully ripe figs.

Ester molecules were produced in low quantities in all the studied fig cultivars. Volatile esters are derived from the esterification of alcohols and acyl-CoA resulting from both amino acids and fatty acids catalyzed by the enzyme alcohol-*O*-acyltransferase. The low ester content of the studied cultivars could also be due to the lack of available substrate for enzymes to act on (Echeverria et al., 2004). Esters were present in low proportions in TH (29.07 ± 2.78–37 ± 0.30%) figs. However, as the total VOC content was much lower in AV figs, ester molecules accounted for a high proportion of the total VOCs in ripe (34 ± 6.18%) and fully ripe (66.5 ± 5.20%). AV figs, with quantities increasing significantly during ripening. The most abundant ester molecule, which was detected in all fig cultivars, was ethyl acetate, which is probably due to the availability of the necessary alcohol precursors (Echeverria et al., 2004). Ethyl acetate is a product of fermentation reactions and is essentially known for its pleasant fruity note at a low concentration. However, it causes an undesirable odor at higher concentrations (Rodríguez-Solana et al., 2018). The high amount of ethyl acetate (88%) in the AV cultivar could be explained by the dark-colored skin of this cultivar, which is rich in this ester and contributes to the fig aroma (Villalobos et al., 2018). Alcohol molecules were detected in most fig samples in low proportions. Alcohols (C_6 and C_9) are biosynthesized from unsaturated fatty acids which are oxygenated by the lipoxygenase enzyme (Deboever et al., 2020) and then cleaved into C_6 and C_9 aldehydes by the hydroperoxide lyase enzyme. These are then reduced into their corresponding alcohols by an alcohol dehydrogenase (Dudareva et al., 2013; El Hadi et al., 2013; Genva et al., 2019). Alcohols primarily contribute to the flowery, green, and "herby" aromas in fruits and have previously been described as pertinent flavor volatiles in fig fruits (Janzantti et al., 2012; Villalobos et al., 2018). According to our results, the alcohol content trend differed among cultivars during ripening. In this context, Wang et al. (2011) reported an increase in alcohol content during the ripening of kiwi fruit while decreasing alcohol content was observed in peaches and nectarines during ripening (Visai and Vanoli, 1997). As alcohol molecules contribute positively to the sensory appeal of fig aroma, the decrease in pent-1-en-3-ol and 2-ethyl-hexan-1-ol emission in AZ figs as they develop from unripe to fully ripe figs probably leads to a fall in green and citrus flavors during fruit ripening (**Table 1**).

In the present study ketones were detected at low proportions (from 21.45 ± 3.35 to 41.86 ± 3.06%) in most samples, but reached 75.93 ± 4.40% in fully ripe TH figs. The finding of small quantities of ketones in fig aroma is similar to previous studies of fresh figs by Gozlekci et al. (2011) and Oliveira et al. (2010), and with studies of both dried and fresh figs by Russo et al. (2017) and Villalobos et al. (2018). These compounds are biosynthesized from unsaturated fatty acids through the ß-oxidation pathway (Schwab et al., 2008). An increasing trend in ketone content was observed during the ripening of TH figs.

Fruity and green odors were the two dominant aroma descriptors for all fig cultivars, which showed some changes from the early to the final stage of ripening. As shown in the spider chart, there was an equivalence between fruity and green aromas in unripe AZ figs and when they reached full ripeness an equal intensity of fruity and floral notes was observed. A fatty aroma was only produced by cultivars TH and AZ, which increased with the maturity of figs (**Figure 4**). Sweet and floral odor characteristics were mainly present in ripe TH and AZ figs, respectively. It should be noted that certain compounds characterize the studied cultivars, such as (*E*)-2-methybut-2-enal and β-caryophyllene, which are synthesized in considerable quantities by TH figs, as well as benzaldehyde and linalool, which are released at high levels by the AZ cultivar (**Figure 5**).

Some volatile compounds were particularly identified at different ripening stages of each cultivar; these VOCs could be considered to represent fingerprints of the ripening stages. Among these compounds, α-santalene was exclusively identified in fully ripe TH figs, and ethyl 2-methylbutanoate and ethyl hexanoate were the two esters that characterized the fully ripe state of TH figs. In addition, the unripe AZ cultivar was distinguished by the presence of benzyl alcohol, heptanal, and (*E*)-oct-2-enal while the ripe AZ fig was distinguished by only one alcohol [(*E*)-hex-2-en-1-ol]. Moreover, the two earlier ripening stages of AV figs were characterized by 1,8 cineol (AV1) and pentan-1-ol (AV2).

As shown in **Figure 6A**, the AZ cultivar is mainly characterized by linalool and benzaldehyde, whereas TH and AV figs have similar VOCs, as they are characterized by acetaldehyde, (*E*)-hex-2-enal, ethyl acetate, and acetone. From examination of **Figure 6A** it can be concluded that the TH and AV cultivars are very similar in terms of VOC composition and differ from AZ figs. **Figure 6B** shows the projection of VOCs grouped by different ripening stages. The axis characterizing unripe and ripe figs have been superimposed, showing that VOC content in these two ripening stages was not significantly different. The fully ripe figs were placed the other side of the axis from the other two ripening stages, and were characterized by ethyl acetate, acetaldehyde and (*E*)-hex-2-enal.

The increasing trend seen in VOC chemical classes during ripening may be due to the climacteric nature of figs. Indeed, Lalel et al. (2003) established a relationship between the biosynthesis of VOCs and ethylene production. The difference observed between VOC class evolution during the ripening process in the different cultivars could be explained by physiological pathways specific to ripening processes in the studied cultivars.

The VOC analyses performed on the three fig cultivars at three ripening stages revealed the presence of a total of 29 compounds belonging to five chemical classes. Overall, 20 VOCs were identified in both TH and AZ figs, showing that there is no difference between the total VOC content of a dominant cultivar (TH) and one which is endangered (AZ). However, only 14 compounds were found in AV figs. Fig aromas are represented by different aromatic descriptors including high fruity, green notes and a moderate sweet and floral aroma, as well as a slight note of fatty aroma. In this investigation, we demonstrated that cultivars and/or ripening stages can be distinguished by VOCs that can be considered as specific fingerprints and are also of interest

for identification in examining the agro-biodiversity of figs. Of course, further analyses is required to study the genetic and environmental factor that can affect the production of VOCs. The study of the bound aroma form of VOCs during the ripening process will be interesting of itself and help researchers to better understand the natural dynamic process of fig aroma biosynthesis.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

KZ, DEK, and M-LF conceptualized the research. KZ and MB-b contributed to the data creation and performance of the software. KZ wrote the original draft and conducted the formal analyses. DEK and M-LF contributed to the methodology, supervision, and validation of work. All authors reviewed and edited the manuscript.

FUNDING

This work was supported by the Algerian Ministry of Higher Education and Scientific Research, specifically the Directorate General for Scientific Research and Technological Development, Algeria. Support was also provided by the Laboratory of Chemistry of Natural Molecules, Gembloux Agro-Bio Tech, University of Liège, Belgium. This research was also funded by the Walinnov project OILPROTECT (1610128) granted by Wallonia via the SPF-Economie Emploi Recherche, Win2Wal program. This research was funded by the Education, Audiovisual and Culture Executive Agency (EACEA) through EOHUB project 600873-EPP-1-2018-1ES-EPPKA2-KA.

ACKNOWLEDGMENTS

The authors are grateful to the regional and local associations of fig producers (Bejaia, Algeria) for the supply of fig samples. A great deal of thanks is addressed to Tierry Kenne and Franck Michels for analytical support. We also wish to express our gratitude to Ramdane Haddouche for proofreading the English manuscript and Evelyne Amenan Tanoh for statistics help.

REFERENCES

Aprotosoaie, A. C., Hăncianu, M., Costache, I. I., and Miron, A. (2014). Linalool: a review on a key odorant molecule with valuable biological properties. *Flavour Fragr. J.* 29, 193–219. doi: 10.1002/ffj.3197

Arem, E., Guido, F., Emna, S., Manel, I., Nesrine, Z., Ali, F., et al. (2011). Chemical and aroma volatile compositions of date palm (*Phoenix dactylifera* L.) fruits at three maturation stages. *Food Chem.* 127, 1744–1754. doi: 10.1016/j.foodchem.2011.02.051

Babushok, V. I., Linstrom, P. J., and Zenkevich, I. G. (2011). Retention indices for frequently reported compounds of plant essential oils. *J. Phys. Chem. Ref. Data* 40, 1–47. doi: 10.1063/1.3653552

Barolo, M. I., Mostacero, N. R., and López, S. N. (2014). *Ficus carica* L. (Moraceae): an ancient source of food and health. *Food Chem.* 164, 119–127. doi: 10.1016/j.foodchem.2014.04.112

Bicchi, C., Ruosi, M. R., Cagliero, C., Cordero, C., Liberto, E., Rubiolo, P., et al. (2011). Quantitative analysis of volatiles from solid matrices of vegetable origin by high concentration capacity headspace techniques: determination of furan in roasted coffee. *J. Chromatogr. A* 1218, 753–762. doi: 10.1016/j.chroma.2010.12.002

Buttery, R. G., Seifert, R. M., Guadagni, D. G., and Ling, L. C. (1969). Characterization of some volatile constituents of bell peppers. *J. Agric. Food Chem.* 17, 1322–1327. doi: 10.1021/jf60166a061

Buttery, R. G., Turnbaugh, J. G., and Ling, L. C. (1988). Contribution of volatiles to rice aroma. *J. Agric. Food Chem.* 36, 1006–1009. doi: 10.1021/jf00083a025

Chawla, A., Kaur, R., and Sharma, A. K. (2012). *Ficus carica* Linn.: a review on its pharmacognostic, phytochemical and pharmacological aspects. *Int. J. Pharm. Phytopharm. Res.* 4, 215–232.

Chen, X., Fedrizzi, B., Kilmartin, P. A., and Quek, S. Y. (2021). Development of volatile organic compounds and their glycosylated precursors in tamarillo (*Solanum betaceum* Cav.) during fruit ripening: a prediction of biochemical pathway. *Food Chem.* 339:128046. doi: 10.1016/j.foodchem.2020.128046

Czerny, M., Christlbauer, M., Christlbauer, M., Fischer, A., Granvogl, M., Hammer, M., et al. (2008). Re-investigation on odour thresholds of key food aroma compounds and development of an aroma language based on odour qualities of defined aqueous odorant solutions. *Eur. Food Res. Technol.* 228, 265–273. doi: 10.1007/s00217-008-0931-x

Deboever, E., Deleu, M., Mongrand, S., Lins, L., and Fauconnier, M.-L. (2020). Plant-pathogen interactions: underestimated roles of phyto-oxylipins. *Trends Plant Sci.* 25, 22–34. doi: 10.1016/j.tplants.2019.09.009

Dixon, J., and Hewett, E. W. (2000). Factors affecting apple aroma / flavour volatile concentration : a review. *New Zeal. J. Crop Hortic. Sci.* 28, 155–173. doi: 10.1080/01140671.2000.9514136

Dudareva, N., Klempien, A., Muhlemann, K., and Kaplan, I. (2013). Tansley review Biosynthesis, function and metabolic engineering of plant volatile organic compounds. *New Phytol.* 198, 16–32. doi: 10.1111/nph.12145

Echeverria, G., Graell, J., López, M. L., and Lara, I. (2004). Volatile production, quality and aroma-related enzyme activities during maturation of 'Fuji' apples. *Postharvest Biol. Technol.* 31, 217–227. doi: 10.1016/j.postharvbio.2003.09.003

El Hadi, M. A. M., Zhang, F., Wu, F., Zhou, C., and Tao, J. (2013). Advances in fruit aroma volatile research. *Molecules* 18, 8200–8229. doi: 10.3390/molecules18078200

Elmore, J. S., Mottram, D. S., and Hierro, E. (2000). Two-fibre solid-phase microextraction combined with gas chromatography-mass spectrometry for the analysis of volatile aroma compounds in cooked pork. *J. Chromatogr. A* 905, 233–240. doi: 10.1016/S0021-9673(00)00990-0

FAO (2018). *Food and Agriculture Organisation of the United Nations*. Available online at: http://www.fao.org/faostat/en/#data/QC (accessed January 30, 2020).

Fauconnier, M.-L., Rojas-Beltran, J., Dupuis, B., Delaplace, P., Frettinger, P., and Gosset, V., et al. (2008). Changes in oxylipin synthesis after Phytophthora infestans infection of potato leaves do not correlate with resistance. *Plant Physiol. Biochem.* 46, 823–831. doi: 10.1016/j.plaphy.2008.04.010

Fiorini, D., Pacetti, D., Gabbianelli, R., Gabrielli, S., and Ballini, R. (2015). A salting out system for improving the efficiency of the headspace solid-phase microextraction of short and medium chain free fatty acids. *J. Chromatogr. A* 1409, 282–287. doi: 10.1016/j.chroma.2015.07.051

Garcia, C. V., Quek, S. Y., Stevenson, R. J., and Winz, R. A. (2012a). Characterisation of bound volatile compounds of a low flavour kiwifruit species: *Actinidia eriantha*. *Food Chem.* 134, 655–661. doi: 10.1016/j.foodchem.2012.02.148

Garcia, C. V., Quek, S. Y., Stevenson, R. J., and Winz, R. A. (2012b). Kiwifruit flavour: a review. *Trends Food Sci. Technol.* 24, 82–91. doi: 10.1016/j.tifs.2011.08.012

Genva, M., Obounou Akong, F., Andersson, M. X., Deleu, M., Lins, L., and Fauconnier, M.-L. (2019). New Insights into the biosynthesis of esterified oxylipins and their involvement in plant defense and developmental mechanisms. *Phytochemistry Rev*. 18, 343–358 doi: 10.1007/s11101-018-9595-8

Gozlekci, S. (2010). Selection studies on fig (*Ficus carica* L.) in Antalya Province of Turkey. *African J. Biotechnol.* 9, 7857–7861. doi: 10.5897/AJB10.1382

Gozlekci, S., Kafkas, E., and Ercisli, S. (2011). Volatile compounds determined by HS / GC-MS technique in peel and pulp of fig (*Ficus carica* L.) cultivars grown in Mediterranean Region of Turkey. *Not. Bot. Horti. Agrobo*. 39, 105–108. doi: 10.15835/nbha3926261

Hjelmeland, A. K., and Ebeler, S. E. (2015). Glycosidically bound volatile aroma compounds in grapes and wine: a review. *Am. J. Enol. Vitic.* 66, 1–11. doi: 10.5344/ajev.2014.14104

Hou, J., Liang, L., and Wang, Y. (2020). Volatile composition changes in navel orange at different growth stages by HS-SPME–GC–MS. *Food Res. Int.* 136:109333. doi: 10.1016/j.foodres.2020.109333

Iban, E., Reglero, G., Cano, M. P., and Bliss, A. (2000). Frozen storage effects on anthocyanins and volatile compounds of raspberry fruit. *J. Agric. Food Chem.* 48, 873–879. doi: 10.1021/jf990747c

Janzantti, N. S., Macoris, M. S., Garruti, D. S., and Monteiro, M. (2012). Influence of the cultivation system in the aroma of the volatile compounds and total antioxidant activity of passion fruit. *LWT - Food Sci. Technol.* 46, 511–518. doi: 10.1016/j.lwt.2011.11.016

Jeong, W. S., and Lachance, P. A. (2001). Phytosterols and fatty acids in fig (*Ficus carica*, var. Mission) fruit and tree components. *J. Food Sci.* 66, 278–281. doi: 10.1111/j.1365-2621.2001.tb11332.x

Karabulut, I., Gokbulut, I., Bilenler, T., Sislioglu, K., Ozdemir, I. S., Bahar, B., et al. (2018). Effect of fruit maturity level on quality, sensory properties and volatile composition of two common apricot (*Prunus armeniaca* L.) varieties. *J. Food Sci. Technol.* 55, 2671–2678. doi: 10.1007/s13197-018-3189-8

Khalil, M. N. A., Fekry, M. I., and Farag, M. A. (2017). Metabolome based volatiles profiling in 13 date palm fruit varieties from Egypt via SPME GC – MS and chemometrics. *Food Chem.* 217, 171–181. doi: 10.1016/j.foodchem.2016.08.089

Lalel, H. J. D., Singh, Z., and Tan, S. C. (2003). Aroma volatiles production during fruit ripening of "*Kensington Pride*" mango. *Postharvest Biol. Technol.* 27, 323–336. doi: 10.1016/S0925-5214(02)00117-5

Merkle, S., Kleeberg, K., and Fritsche, J. (2015). Recent developments and applications of solid phase microextraction (SPME) in food and environmental analysis—a review. *Chromatography* 2, 293–381. doi: 10.3390/chromatography2030293

Noguerol-Pato, R., González-Barreiro, C., Cancho-Grande, B., Martínez, M. C., Santiago, J. L., and Simal-Gándara, J. (2012). Floral, spicy and herbaceous active odorants in Gran Negro grapes from shoulders and tips into the cluster, and comparison with Brancellao and Mouratón varieties. *Food Chem.* 135, 2771–2782. doi: 10.1016/j.foodchem.2012.06.104

Obenland, D., Collin, S., Sievert, J., Negm, F., and Lu, M. (2012). Influence of maturity and ripening on aroma volatiles and flavor in 'Hass' avocado. *Postharvest Biol. Technol.* 71, 41–50. doi: 10.1016/j.postharvbio.2012.03.006

Oliveira, A. P., Baptista, P., Andrade, P. B., Martins, F., Pereira, J. A., Silva, B. M., et al. (2012). Characterization of *Ficus carica* L. cultivars by DNA and secondary metabolite analysis : is genetic diversity re fl ected in the chemical composition? *FRIN* 49, 710–719. doi: 10.1016/j.foodres.2012.09.019

Oliveira, A. P., Silva, L. R., Guedes, P., Pinho, D., Gil-izquierdo, A., Valentão, P., et al. (2010). Volatile profiling of Ficus carica varieties by HS-SPME and GC–IT-MS. *Food Chem.* 123, 548–557. doi: 10.1016/j.foodchem.2010.04.064

Palassarou, M., Melliou, E., Liouni, M., Michaelakis, A., Balayiannis, G., and Magiatis, P. (2017). Volatile profile of Greek dried white figs (*Ficus carica* L.) and investigation of the role of β -damascenone in aroma formation in fig liquors. *J. Sci. Food Agric.* 97, 5254–5270. doi: 10.1002/jsfa.8410

Palmeira, L., Pereira, C., Dias, M. I., Abreu, R. M. V., Corrêa, R. C. G., Pires, T. C. S. P., et al. (2019). Nutritional, chemical and bioactive profiles of different parts of a Portuguese common fig (*Ficus carica* L.) variety. *Food Res. Int.* 126, 1–10. doi: 10.1016/j.foodres.2019.108572

Pawliszyn (2000). Theory of solid-phase microextraction. *J. Chromatogr. Sci.* 38, 270–278. doi: 10.1093/chromsci/38.7.270

Pereira, C., Martín, A., López-Corrales, M., de Guía Córdoba, M., Galván, A. I., and Serradilla, M. J. (2020). Evaluation of the physicochemical and sensory characteristics of different fig cultivars for the fresh fruit market. *Foods* 9:50619. doi: 10.3390/foods9050619

Prasanna, V., Prabha, T. N., and Tharanathan, R. N. (2007). Fruit ripening phenomena-an overview. *Crit. Rev. Food Sci. Nutr.* 47, 1–19. doi: 10.1080/10408390600976841

Rodríguez-Solana, R., Galego, L. R., Pérez-Santín, E., and Romano, A. (2018). Production method and varietal source influence the volatile profiles of spirits prepared from fig fruits (*Ficus carica* L.). *Eur. Food Res. Technol.* 244, 2213–2229. doi: 10.1007/s00217-018-3131-3

Russo, F., Caporaso, N., Paduano, A., and Sacchi, R. (2017). Characterisation of volatile compounds in Cilento (Italy) figs (*Ficus carica* L.) cv. Dottato as affected by the drying process. *Int. J. Food Prop.* 20, 1366–1376. doi: 10.1080/10942912.2017.1344991

Schwab, W., Davidovich-Rikanati, R., and Lewinsohn, E. (2008). Biosynthesis of plant-derived flavor compounds. *Plant J.* 54, 712–732. doi: 10.1111/j.1365-313X.2008.03446.x

Solomon, A., Golubowicz, S., Yablowicz, Z., Grossman, S., Bergman, M., Gottlieb, H. E., et al. (2006). Antioxidant activities and anthocyanin content of fresh fruits of common fig (*Ficus carica* L.). *J. Agric. Food Chem.* 54, 7717–7723. doi: 10.1021/jf060497h

Song, J., Bi, J., Chen, Q., Wu, X., Lyu, Y., and Meng, X. (2018). Assessment of sugar content, fatty acids, free amino acids, and volatile profiles in jujube fruits at different ripening stages. *Food Chem.* 270, 344–352. doi: 10.1016/j.foodchem.2018.07.102

Takeoka, G. R., Flath, R. A., Mon, T. R., Teranishi, R., and Guentert, M. (1990). Volatile constituents of apricot (*Prunus Armeniaca*). *J. Agric. Food Chem.* 38, 471–477. doi: 10.1021/jf00092a031

Tamura, H., Boonbumrung, S., Yoshizawa, T., and Varanyanond, W. (2001). The volatile constituents in the peel and pulp of a Green Thai Mango, Khieo Sawoei Cultivar (*Mangifera indica* L.). *Food Sci. Technol. Res.* 7, 72–77. doi: 10.3136/fstr.7.72

Tanoh, E. A., Boué, G. B., Nea, F., Genva, M., Wognin, E. L., Ledoux, A., et al. (2020). Seasonal effect on thechemical composition, insecticidal properties and other biological activities of *Zanthoxylum leprieurii* Guill. and Perr. essential oils. *Foods* 9, 1–26. doi: 10.3390/foods9050550

Villalobos, M. C., Serradilla, M. J., Martín, A., Aranda, E., and López-corrales, M. (2018). Influence of modified atmosphere packaging (MAP) on aroma quality of fi gs (*Ficus carica* L.). *Postharvest Biol. Technol.* 136, 145–151. doi: 10.1016/j.postharvbio.2017.11.001

Visai, C., and Vanoli, M. (1997). Volatile compound production during growth and ripening of peaches and nectarines. *Sci. Hortic.* 70, 15–24. doi: 10.1016/S0304-4238(97)00032-0

Wang, M. Y., Macrae, E., Wohlers, M., and Marsh, K. (2011). Changes in volatile production and sensory quality of kiwifruit during fruit maturation in Actinidia deliciosa 'Hayward' and A. chinensis 'Hort16A.' *Postharvest Biol. Technol.* 59, 16–24. doi: 10.1016/j.postharvbio.2010.08.010

Wang, Y., O'Reilly, J., Chen, Y., and Pawliszyn, J. (2005). Equilibrium in-fibre standardisation technique for solid-phase microextraction. *J. Chromatogr. A* 1072, 13–17. doi: 10.1016/j.chroma.2004.12.084

Ware, A. B., Kaye, P. T., Compton, S. G., and Van Noort, S. (1993). Fig volatiles: their role in attracting pollinators and maintaining pollinator specificity. *Plant Syst. Evol.* 186, 147–156. doi: 10.1007/BF00940794

Wu, Y., Duan, S., Zhao, L., Gao, Z., Luo, M., Song, S., et al. (2016). Aroma characterization based on aromatic series analysis in table grapes. *Sci. Rep.* 6, 1–16. doi: 10.1038/srep31116

Yang, C., Wang, Y., Wu, B., Fang, J., and Li, S. (2011). Volatile compounds evolution of three table grapes with different flavour during and after maturation. *Food Chem.* 128, 823–830. doi: 10.1016/j.foodchem.2010.11.029

Yang, Y. N., Zheng, F. P., Yu, A. N., and Sun, B. G. (2019). Changes of the free and bound volatile compounds in *Rubus corchorifolius* L. f. fruit during ripening. *Food Chem.* 287, 232–240. doi: 10.1016/j.foodchem.2019.02.080

Young, H., and Paterson, V. J. (1985). The effects of harvest maturity, ripeness and storage on kiwifruit aroma. *J. Sci. Food Agric.* 36, 352–358. doi: 10.1002/jsfa.2740360507

Young, H., and Paterson, V. J. (1995). Characterisation of bound flavour components in kiwifruit. *J. Sci. Food Agric.* 68, 257–260. doi: 10.1002/jsfa.2740680218

Zidi, K., Kati, D. E., Benchikh, Y., Bey, M. B., Ouandjeli, D., and Yahiaoui, S. (2020). The use of modified atmosphere packaging as mean of bioactive compounds and antioxidant activities preservation of fresh figs (*Ficus carica* L.) from rare cultivars. *Ann. Univ. Dunarea Jos Galati, Fascicle VI Food Technol.* 44, 149–164. doi: 10.35219/foodtechnology.2020.1.09

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 13 September 2021
doi: 10.3389/fpls.2021.708697

Genome-Wide Identification and Functional Characterization of the Trans-Isopentenyl Diphosphate Synthases Gene Family in *Cinnamomum camphora*

Zerui Yang[1,2,3], Chunzhu Xie[1], Ting Zhan[1], Linhuan Li[1], Shanshan Liu[1], Yuying Huang[1], Wenli An[1], Xiasheng Zheng[1] and Song Huang[1,3]**

[1]School of Pharmaceutical Sciences, Guangzhou University of Chinese Medicine, Guangzhou, China, [2]National Engineering Research Center for Healthcare Devices, Institute of Medicine and Health, Guangdong Academy of Sciences, Guangzhou, China, [3]National Engineering Research Center for Modernization of Traditional Chinese Medicine, Guangzhou University of Chinese Medicine, Guangzhou, China

OPEN ACCESS

Edited by:
Jian Li,
Beijing Technology and Business University, China

Reviewed by:
Raimund Nagel,
University of Leipzig, Germany
Wajid Waheed Bhat,
Michigan State University, United States
Jinhua Zuo,
Beijing Vegetable Research Center, China

***Correspondence:**
Xiasheng Zheng
xszheng@gzucm.edu.cn
Song Huang
huangnn421@163.com

Specialty section:
This article was submitted to Plant Metabolism and Chemodiversity, a section of the journal Frontiers in Plant Science

Received: 12 May 2021
Accepted: 28 July 2021
Published: 13 September 2021

Citation:
Yang Z, Xie C, Zhan T, Li L, Liu S, Huang Y, An W, Zheng X and Huang S (2021) Genome-Wide Identification and Functional Characterization of the Trans-Isopentenyl Diphosphate Synthases Gene Family in Cinnamomum camphora. Front. Plant Sci. 12:708697. doi: 10.3389/fpls.2021.708697

Trans-isopentenyl diphosphate synthases (TIDSs) genes are known to be important determinants for terpene diversity and the accumulation of terpenoids. The essential oil of *Cinnamomum camphora*, which is rich in monoterpenes, sesquiterpenes, and other aromatic compounds, has a wide range of pharmacological activities and has therefore attracted considerable interest. However, the *TIDS* gene family, and its relationship to the camphor tree (*C. camphora* L. Presl.), has not yet been characterized. In this study, we identified 10 *TIDS* genes in the genome of the *C. camphora* borneol chemotype that were unevenly distributed on chromosomes. Synteny analysis revealed that the *TIDS* gene family in this species likely expanded through segmental duplication events. Furthermore, cis-element analyses demonstrated that *C. camphora* TIDS (CcTIDS) genes can respond to multiple abiotic stresses. Finally, functional characterization of eight putative short-chain TIDS proteins revealed that CcTIDS3 and CcTIDS9 exhibit farnesyl diphosphate synthase (FPPS) activity, while CcTIDS1 and CcTIDS2 encode geranylgeranyl diphosphate synthases (GGPPS). Although, CcTIDS8 and CcTIDS10 were found to be catalytically inactive alone, they were able to bind to each other to form a heterodimeric functional geranyl diphosphate synthase (GPPS) *in vitro*, and this interaction was confirmed using a yeast two-hybrid assay. Furthermore, transcriptome analysis revealed that the *CcTIDS3*, *CcTIDS8*, *CcTIDS9*, and *CcTIDS10* genes were found to be more active in *C. camphora* roots as compared to stems and leaves, which were verified by quantitative real-time PCR (qRT-PCR). These novel results provide a foundation for further exploration of the role of the *TIDS* gene family in camphor trees, and also provide a potential mechanism by which the production of camphor tree essential oil could be increased for pharmacological purposes through metabolic engineering.

Keywords: genome-wide identification, *Cinnamomum camphora*, trans-isopentenyl diphosphate synthases, functional characterization, gene family

INTRODUCTION

Cinnamomum camphora is a subtropical evergreen tree species that has been widely cultivated in southern China for over 1,500 years (Chen et al., 2020b). Studies have shown that essential oil extracted from the leaves of the camphor tree, which is rich in monoterpenes, sesquiterpenes, and other aromatic compounds, has a wide range of pharmacological activities, including antibacterial, antioxidant, and insecticidal properties (Yu et al., 2019).

Cinnamomum camphora subspecies can be grouped into five chemotypes according to the dominant component in their essential oils, which were extracted from their leaves: linalool, D-borneol, camphor, cineole, or nerolidol (Guo et al., 2017; Chen et al., 2018). These terpenoids have important industrial and pharmaceutical applications. For example, D-borneol is a well-established traditional Chinese medicine that is used to treat cardiovascular diseases, including stroke, coronary heart disease, and angina pectoris (Yang et al., 2020). D-Borneol has been documented in various versions of the Chinese Pharmacopoeia (Huang et al., 2016; Liang et al., 2018; Ren et al., 2018; Chai et al., 2019; Chen et al., 2019; Yang et al., 2020), and it is in high demand as a key ingredient in many traditional Chinese herbal formulas. This compound can also be used to relieve pain resulting from wounds, injuries, burns, and cuts (Yang et al., 2020). However, despite the high medicinal value of volatile terpenoids, limited plant resources, and low extraction efficiency limit the amounts of the essential oils from *C. camphora* that can be obtained for research or practical applications (Ma et al., 2021).

An increasing body of research has demonstrated that metabolic engineering consists of optimizing the genetic and regulatory mechanisms that govern cellular processes, and this is an effective method that can be used to increase the production of active natural products in microorganisms and plants (Choi et al., 2019; Li and Mutanda, 2019). Identifying the genes that govern production and accumulation of the essential oil in *C. camphora* could therefore enable the implementation of this approach to increase the essential oil yield, which would assist in meeting high demand.

Terpenoids are the largest category of plant specialized metabolites. More than 55,000 of these compounds have thus far been described, and are grouped into hemi- (C5), mono- (C10), sesqui- (C15), di- (C20), sester- (C25), tri- (C30), tetra- (C40), and poly- (C50) terpenoids according to the number of carbon atoms they contain (Dudareva et al., 2004; Pichersky et al., 2006; Ueoka et al., 2020). Despite diverse functions and structures, all terpenoids contain two universal C5 units: isopentenyl diphosphate (IPP) and its allylic isomer dimethylallyl diphosphate (DMAPP). The mevalonic acid (MVA) pathway gives rise to IPP, and *via* enzymatic isomerization, to DMAPP, whereas the methylerythritol phosphate (MEP) pathway directly produces both IPP and DMAPP. IPP and DMAPP can be condensed (head to tail) by prenyltransferase (PTS) or isoprenyl diphosphate synthase (IDS), resulting in a series of prenyl diphosphates with various chain lengths. These linear precursors are then catalyzed by terpene synthase (TPS) and other modifying enzymes to form a variety of different terpenoids (Tholl and Lee, 2011; Athanasakoglou and Grypioti, 2019; Johnson et al., 2019; Johnson and Bhat, 2019; Adal and Mahmoud, 2020; Hivert et al., 2020; Miller and Bhat, 2020).

Isoprenyl diphosphate synthase is located at the branch point of the terpenoid biosynthetic pathway, and plays a vital role in the formation of diverse terpenoid structures (Jia and Chen, 2016). Differences in terpenoid synthase gene expression and the supply of precursors determine the terpenoid composition produced by plants (Dudareva et al., 2000; Johnson et al., 2019; Johnson and Bhat, 2019; Adal and Mahmoud, 2020; Miller and Bhat, 2020). Both trans- and cis- isomers of the products of IDS exist (Liang et al., 2002). Trans-IDS (TIDS) enzymes synthesize isoprenyl diphosphates (Barja and Rodríguez-Concepción, 2020) and are classified as short-chain (SC-TIDS, C10–20), medium-chain (MC-TIDS, C25–35), or long-chain (LC-TIDS, C40–50), depending on the length of the isoprenyl diphosphates that they produce (Wang et al., 2019). Cis-IDSs (CIDSs) were initially predicted to synthesize long-chain isoprenyl diphosphates (>C50) for dolichol and polyprenol production (Jia and Chen, 2016). Although, TIDSs and CIDSs have similarities in substrate preference and reaction products, they utilize different catalytic mechanisms and may be readily distinguished from one another by their primary amino acid (AA) sequences (Akhtar et al., 2013).

Short-chain-trans-isopentenyl diphosphate synthase includes homodimeric or heterodimeric geranyl diphosphate synthase (GPPS), farnesyl diphosphate synthase (FPPS), and geranylgeranyl diphosphate synthase (GGPPS; Jia and Chen, 2016), and their products provide the precursors for monoterpenes, sesquiterpenes, and diterpenes, respectively. Geranylfarnesyl diphosphate synthase (GFPPS) and polyprenyl diphosphate synthase (PPPS) are MC-TIDSs, whereas solanesyl diphosphate synthase (SPPS) is classified as an LC-TIDS (Vandermoten et al., 2009; Nagel et al., 2015; Wang et al., 2016; Kopcsayová and Vranová, 2019).

Phylogenetic analysis has revealed that plant *TIDS* genes can be documented into five subfamilies according to sequence identity: *TIDS-a*, *-b*, *-c*, and *-e* including the genes encoding FPPS, SPPS, PPPS, and small subunits (SSUs) of GPPS, respectively (Jia and Chen, 2016). The *TIDS-d* subfamily is more complex, and it includes genes encoding GGPPS, GFPPS, and PPPS that share a high sequence identity of at least 40%, with some shared identities of 55% or greater (Jia and Chen, 2016; Cui et al., 2019). Identification of entire *TIDS* gene families in plant species is required to determine their functions and to understand their combined effect on the specific profile of terpenoids produced.

Thus far, this level of *TIDS* gene family characterization has only been achieved in the model plant *Arabidopsis thaliana* (Kopcsayová and Vranová, 2019), in which 16 *TIDS* genes were identified, and in the tomato plant *Solanum lycopersicum* (Zhou and Pichersky, 2020), in which 10 were characterized. Studies characterizing these *TIDS* genes have revealed that while most of the synthase types they encode are homodimeric, GPPS may exist as either a homodimer or a heterodimer containing a small subunit (SSU) and a large subunit (LSU),

or both (Rai et al., 2013; Chen et al., 2015; Adal and Mahmoud, 2020). The SSU can be further separated into two types (I and II), which are generally inactive alone. The LSU may be either inactive alone or possess GGPPS activity, while its heterodimer is an active GPPS (Burke et al., 2004; Rai et al., 2013).

Given that isoprenyl diphosphate biosynthesis is a crucial determinant for the formation of downstream terpenoid type and their yield, in-depth studies on TIDSs in the camphor tree would be instrumental in optimizing the production of these medicinally valuable products. Here, for the first time, we conducted a genome-wide analysis in the borneol chemotype of *C. camphora* to identify and characterize TIDS genes and the proteins they encode.

MATERIALS AND METHODS

Plant Materials

Borneol chemotype *C. camphora* plants were purchased from Ji'an Yu Feng Natural Species Co., Ltd. (Ji'an, China), and grown in an artificial climate box (Shanghai Yiheng Instrument Co., Ltd., Shanghai, China) at 25°C with a 12 h light and then 12 h dark photoperiod. We verified the origin of these plants through DNA barcodes.

Genome-Wide Identification of TIDS Genes in *C. camphora*

We downloaded TIDS protein sequences of *A. thaliana* and *S. lycopersicum* from The Arabidopsis Information Resource (TAIR)[1] and the Sol Genomics Network (Wijfjes and Smit, 2019).[2] We then used the Protein Basic Logical Alignment Search Tool (BLASTP, United States National Library of Medicine)[3] to compare these sequences with the *C. camphora* database (unpublished), using previously established thresholds for *e*-value ($\leq 1e^{-5}$) and identity (Cui et al., 2019). Any genes identified were then used as queries for a second round of BLASTP searches to ensure no putative *CcTIDS* genes were missed. We also conducted a Hidden Markov Model (HMM) search for sequence homologs using the HMMER 3.0 program and the polyprenyl synthase domain PF00348 as a bait, and previously established *e*-value and identity thresholds (Wong et al., 2015; Cui et al., 2019).

The BLASTP and HMM search results were then integrated to identify candidate TIDS genes. Their sequences were then submitted to the online Pfam database and NCBI conserved domains database (CDD) in order to verify the presence of the polyprenyl synthase domain (Zhu et al., 2020). The physicochemical parameters of each CcTIDS, including molecular weight and isoelectric point, were calculated using the ExPASy online tool (Zhu et al., 2020).[4] The MEME online tool[5] was used to discover conserved domains in the amino acid sequence of each CcTIDS (Zhu et al., 2020). Finally, we used the chloroP (Emanuelsson et al., 1999),[6] TargetP (Emanuelsson et al., 2000),[7] Wolfpsort (Horton et al., 2007),[8] and Plant-mPloc (Chou and Shen, 2010)[9] tools to predict the subcellular localization of each CcTIDS.

Phylogenetic Relationship, Exon-Intron Structure, Chromosomal Localization, and Cis-Acting Element Analysis

A maximum likelihood (ML) evolutionary tree was constructed using the identified *C. camphora* TIDS amino acid sequences and those of 10 other plant species (*A. thaliana*: Arabidopsis Genome Initiative, 2000; *Oryza sativa*: Yu et al., 2002; *Physcomitrella patens*: Rensing et al., 2008; *Selaginella moellendorffii*: Banks et al., 2011; *S. lycopersicum*: Tomato Genome Consortium, 2012; *Amborella trichopoda*: Amborella Genome Project, 2013; *Picea abies*: Nystedt et al., 2013; *Zea mays*: Jiao et al., 2017; and *Cinnamomum micranthum*: Chaw et al., 2019) and the Molecular Evolutionary Genetics Analysis (MEGA) tool (version 7.0; Kumar et al., 2016) with the best models of JTT + G. Bootstrap value, which indicates the reliability of each branch node, was set at 1,000 replicates. The resulting tree was visualized using EvolView v3 (Subramanian et al., 2019).

Exons and introns for each TIDS were identified using the gene transfer format (GTF) file from the *C. camphora* genome, which contains information regarding gene structure, and visualized using TBtools (Chen et al., 2020a). Finally, in order to identify cis-regulatory elements, TBtools was used to extract the 2,000 bp upstream sequence for each TIDS gene identified in the *C. camphora* genome (Chen et al., 2020a). We then compared these sequences with the PlantCARE database of plant cis-acting regulatory elements (Lescot et al., 2002). Circos graphs showing the chromosomal localization and results of the synteny analysis of the CcTIDS sequences were drawn using TBtools (Chen et al., 2020a).

RNA Extraction, cDNA Synthesis, and Gene Cloning

Total RNA was extracted from the root, stem, and leaf of *C. camphora* using the Plant Pure Plant RNA Kit (Aidlab, Beijing, China). cDNA was synthesized from 1 μg of high-quality total RNA (OD 260/280 = 1.8–2.2, OD 260/230 ≥ 2.0, RIN ≥ 6.5, and 28S:18S ≥ 1.0) using the TransScript One-Step gDNA Removal and cDNA Synthesis SuperMix kit (TransGen Biotech, Beijing, China) according to the manufacturer's instructions (Su et al., 2019).

The primers use to amplify eight putative SC-TIDSs were designed using the Primer Premier 5 program (**Supplementary Table S4**). Amplification by PCR was conducted using the 2 × TransStart FastPfu PCR SuperMix (TransGen Biotech, Beijing, China) and cDNA as the template. Purified PCR products were then cloned into the pEASY-blunt vector using previously described methods (Su et al., 2019).

[1] https://www.arabidopsis.org/index.jsp
[2] https://solgenomics.net/organism/Solanum_lycopersicum/genome/
[3] https://blast.ncbi.nlm.nih.gov/Blast.cgi
[4] https://web.expasy.org/protparam
[5] https://meme-suite.org/meme/tools/meme
[6] http://www.cbs.dtu.dk/services/ChloroP/
[7] http://www.cbs.dtu.dk/services/TargetP/
[8] https://wolfpsort.hgc.jp/
[9] http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/

Quantitative Real-Time PCR Analysis

Quantitative real-time PCR (qRT-PCR) was performed using the TransStar Tip Green qPCR SuperMix (TransGen Biotech, Beijing, China) and the CFX96 Touch Deep Well platform (Bio-Rad, United States; Yang et al., 2020). We used a previously described reaction system composition and qRT-PCR procedure (Yang et al., 2020). Primers are listed in **Supplementary Table S4**.

Subcellular Localization of CcTIDS Proteins

Full-length putative SC-TIDS gene sequences without stop codons were each fused with enhanced green fluorescent protein (EGFP) and ligated into the pAN580 vector (using the primers described in **Supplementary Table S4**). The recombinant vectors were then transformed into *Arabidopsis* protoplasts using polyethylene glycol (PEG; Zhou and Pichersky, 2020). Protoplast isolation and recombinant vector transformation were performed with protocols described in previous studies (Yoo et al., 2007; Wang et al., 2018). EGFP fluorescent signals were observed using a Zeiss laser scanning microscope (LSM) 800 (Zeiss, Germany) as previously described (Beck et al., 2013; Su et al., 2019).

Recombinant Expression and Enzymatic Assays

Truncated or full length versions of the eight putative SC-TIDSs were ligated into the pET32a [polyhistidine (6x His) tag, Wego, Guanghzou, China] or pMAL-C5X (MBP tag, Wego, Guanghzou, China) expression vectors using the pEASY-Basic Seamless Cloning and Assembly Kit (TransGen Biotech) following the manufacturer's instructions (Su et al., 2019), using the primers listed in **Supplementary Table S4**. Constructs were verified using Sanger sequencing. Upon sequence confirmation, recombinant plasmids were transferred into the expression strain *Escherichia coli* Rossetta (DE3; Huayueyang, Beijing, China). Heterologous protein expression in this strain and the purification of the fusion protein were conducted according to previously established methods (Su et al., 2019).

In order to establish the function of each SC-TIDS, a 200-μl *in vitro* enzymatic activity reaction containing 25 mmol/L MOPSO buffer with pH of 7.0, 10 mmol/L magnesium chloride, 10% glycerin, 20–50 μg protein, and substrates (150 μmol/L IPP + 40 μmol/L DMAPP or GPP or FPP, Sigma-Aldrich, United States), was conducted. Each mixture was first incubated at 30°C for 6 h, and then, 200 μl of 200 mmol/L Tris–HCl (pH 9.5), containing 2 units of bovine intestine alkaline phosphatase

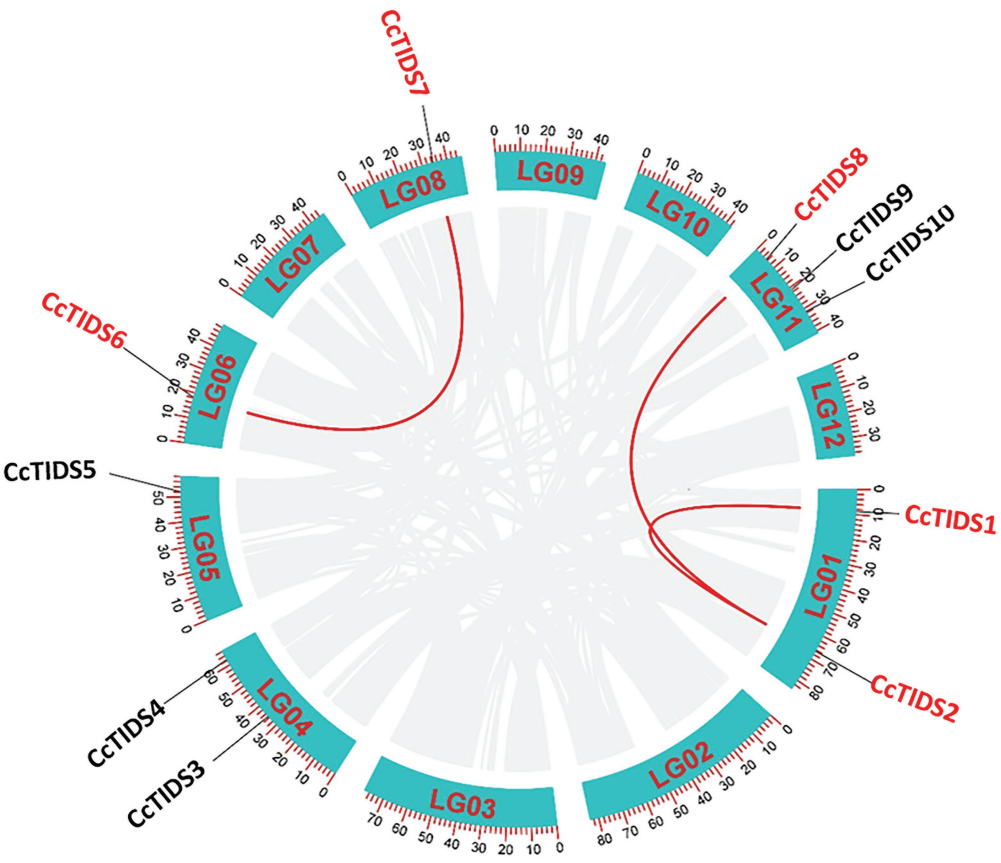

FIGURE 1 | Circos graph showing the chromosomal locations and duplicated gene pairs of trans-isopentenyl diphosphate synthases (*CcTIDS*) genes in the *C. camphora* genome. The blue segments represent the 12 chromosomes present in this species. Duplicated genes are indicated as red lines between each gene pair.

(18 units mg^{-1}; Sigma-Aldrich) and 2 units of potato apyrase (25.2 units mg^{-1}; Sigma-Aldrich) were added (Rai et al., 2013). An overnight hydrolysis reaction was carried out at 30°C. Ethyl acetate was then used to extract the enzymatic reaction buffer in 2 × 400 µl extractions (Su et al., 2019). The ethyl acetate extracts were concentrated to 100 µl under N_2, and these concentrated extracts were then used for gas chromatography–mass spectrometry (GC–MS) analysis using previously described methods (Athanasakoglou and Grypioti, 2019; Yang et al., 2020). The GPPS large and small subunits from *Catharanthus roseus* were also analyzed as a positive control (Rai et al., 2013).

Yeast 2-Hybrid Assay

The interactions between the CcGPPS small and large subunits were verified by the following experiments: the truncated versions of these subunits were amplified from the cDNA of *C. camphora*, followed by fusing to the activation domain of the pGADT7 vector or the binding domain of the pGBKT7 vector. Recombinant vectors were co-transformed into *Saccharomyces cerevisiae* AH109 yeast and sequentially cultivated on the synthetic dropout (SD) medium SD/-Trp/-Leu. The interaction between the two proteins was tested on SD/-Trp/-Leu/-His/-Ade medium supplemented with 150 mM 3-amino-1,24-triazole (3-AT).

RESULTS

Genome-Wide Identification of TIDS Genes in *C. camphora*

Based on the HMM scan and BLASTP search results, 10 full-length, protein-coding TIDS-like gene sequences were identified in *C. camphora*, which we numbered *CcPTPS1–10* according to their locations on the chromosomes. These genes included eight putative SC-TIDS (two FPPSs, *CcTIDS3* and *CcTIDS9*; two GPPSs, *CcTIDS4* and *CcTIDS5*; three GGPPSs, *CcTIDS1*, *CcTIDS2*, and *CcTIDS8*; and one GPPS small subunit, *CcTIDS10*) and two putative LC-TIDS (SPPS, and *CcTIDS6* and *CcTIDS7*), which were distributed along six of the 12 chromosomes (**Figure 1**). In addition, genome synteny analysis showed that three segmental duplications and no tandem duplication events likely occurred, suggesting that segmental duplication might be one of the reasons for TIDS gene family expansion (**Figure 1**).

The proteins encoded by the *CcTIDS* genes were found to have a minimum size of 298 AA and a maximum size of 428 AA. All of the CcTIDS proteins had an isoelectric point (pI) <7, indicating that these proteins are rich in acidic AAs. Specific information for all CcTIDS proteins is shown in **Table 1**.

Phylogenetic Relationships Within the *C. camphora* TIDS Gene Family

A total of 100 *TIDS*-like genes were identified in 10 other plant species (**Supplementary Tables S1** and **S2**). It can be concluded from the number of TIDS genes contained in each species that as the complexity of the species increases, the greater amount TIDS it contains (Coman et al., 2014).

TABLE 1 | Nomenclature and physicochemical characteristics of TIDS genes identified in *C. camphora*.

Gene ID	Rename	Chromosomal position	Putative function	Number of amino acids/AA	Theoretical pI	Molecular weight/KD	*In silico* subcellular localization prediction				Conserved motif
							chloroP	TargetP	Wolfpsort	Plant-mPloc	
evm.model.LG01.0506	CcTIDS1	LG01	GGPPS	383	5.98	41.34	Chloroplast	Chloroplast	Chloroplast	Chloroplast	FARM/SARM/CXXXC
evm.model.LG01.2465	CcTIDS2	LG01	GGPPS	387	6.33	41.7	Chloroplast	Chloroplast	Chloroplast	Chloroplast	FARM/SARM/CXXXC
evm.model.LG04.0985	CcTIDS3	LG04	FPPS	350	5.25	40.41	/	/	cytoplasm	Mitochondrion.	FARM/SARM
evm.model.LG04.2634	CcTIDS4	LG04	GPPS	421	5.67	46.45	/	Mitochondrion	Chloroplast	Chloroplast	FARM/SARM
evm.model.LG05.1976	CcTIDS5	LG05	GPPS	428	6.09	47.04	/	Mitochondrion	Mitochondrion	Chloroplast	FARM/SARM
evm.model.LG06.0936	CcTIDS6	LG06	SPPS	415	6.42	45.3	/	/	Chloroplast	Chloroplast	FARM/SARM
evm.model.LG08.1085	CcTIDS7	LG08	SPPS	401	5.50	43.68	/	Mitochondrion	Chloroplast	Chloroplast	FARM/SARM
evm.model.LG11.0380	CcTIDS8	LG11	GGPPS	383	5.78	41.39	Chloroplast	Chloroplast	Chloroplast	Chloroplast	FARM/SARM/CXXXC
evm.model.LG11.1115	CcTIDS9	LG11	FPPS	364	5.18	41.45	/	Chloroplast	Chloroplast	Mitochondrion	FARM/SARM
evm.model.LG11.1288	CcTIDS10	LG11	GPPS.SSU	298	5.24	32.13	Chloroplast	Chloroplast	Chloroplast	Chloroplast	CXXXC\CXXXC

In addition, fewer *CcTIDS* genes were found in *C. camphora* than those in *A. thaliana*, *Populus trichocarpa*, *O. sativa*, and *Z. mays*, but equal to those in *S. lycopersicum* and the stout camphor tree *C. micranthum*.

A ML evolutionary tree was built to determine the phylogenetic relationships between the *C. camphora TIDS* family and those found in the other plants (**Figure 2**). None of the genes apart from those in *S. lycopersicum* and *A. thaliana* has yet been functionally characterized. In order to better distinguish the functional attributes of the *TIDS* genes on each branch, we added several well-characterized TIDS genes when constructing the phylogenetic tree (**Supplementary Table S3**). The *TIDS* genes in the resulting tree clustered into recently described catalytically distinct subfamilies (Jia and Chen, 2016).

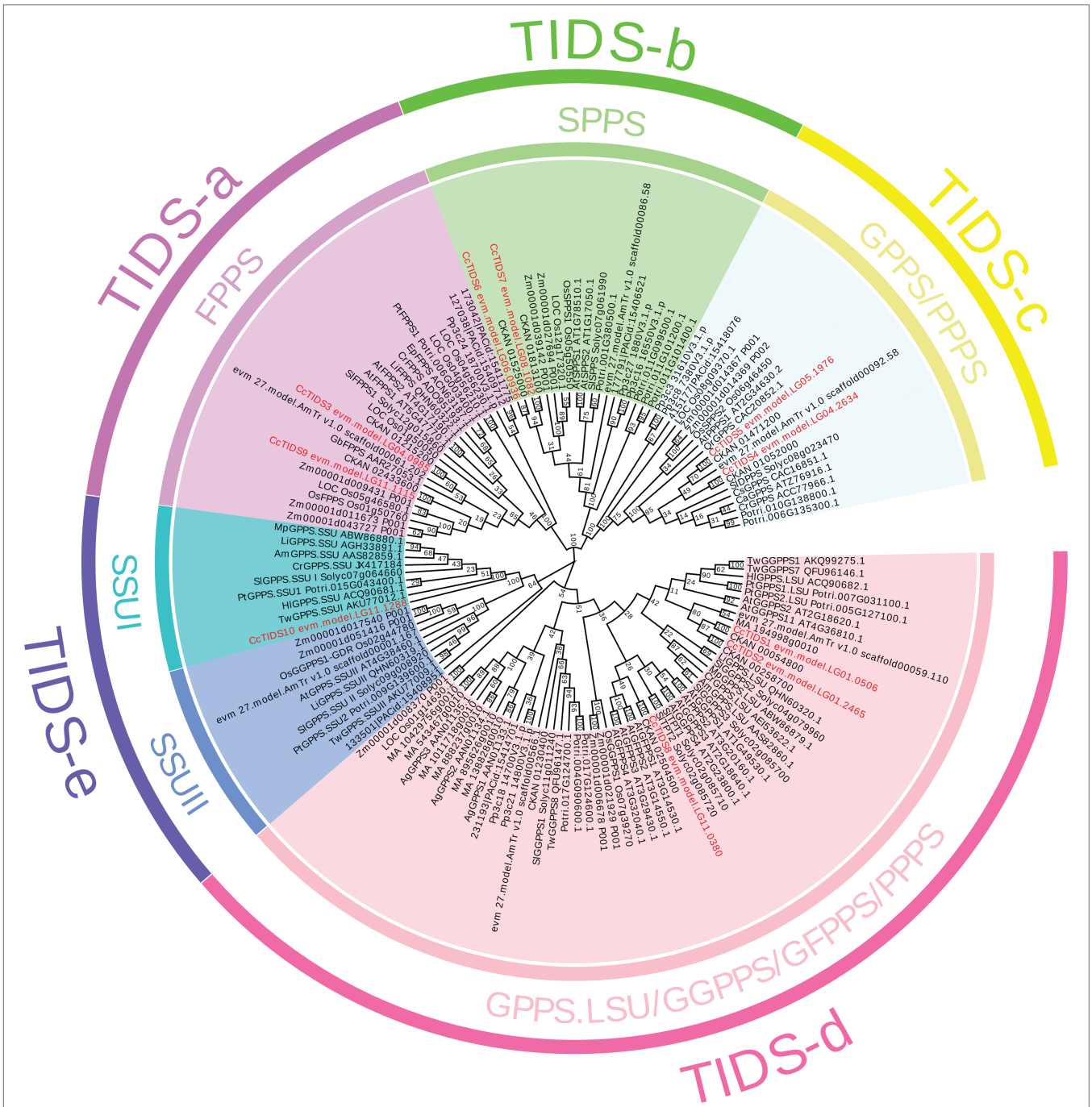

FIGURE 2 | A phylogenetic tree constructed using putative or characterized *TIDS* genes from 11 sequenced land plant genomes and a neighbor joining approach. GPPS, geranyl diphosphate synthase; FPPS, farnesyl diphosphate synthase; GGPPS, geranylgeranyl diphosphate synthase; LSU, large subunit; SSUI/II, small subunit I/II; GFPPS, geranylfarnesyl diphosphate synthase; PPPS, polyprenyl diphosphate synthase; and SPPS, solanesyl diphosphate synthase.

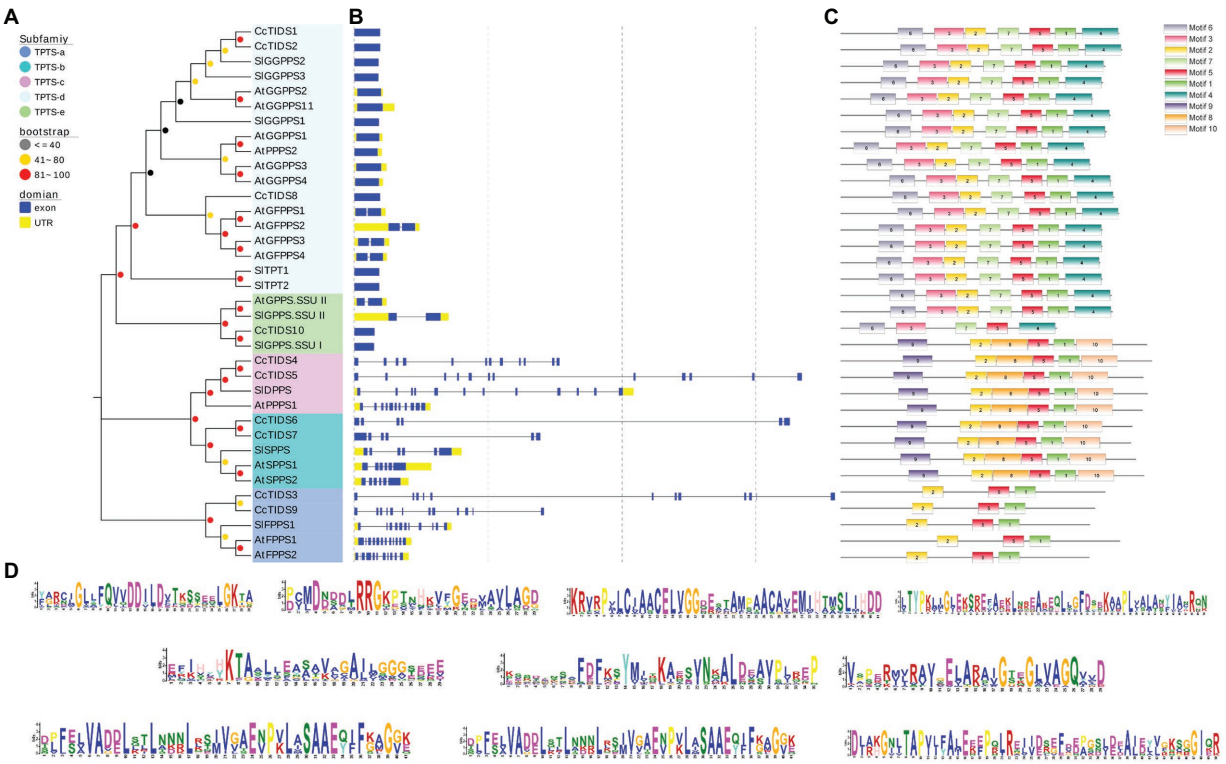

FIGURE 3 | Tran-prenyltransferase protein subfamily analysis based on phylogenetic relationships. **(A)** Intron-exon structures, **(B)** conserved motifs, **(C)** analysis of *C. camphora* (Cc)TIDS, *Arabidopsis thaliana* (At)TIDS, and *Solanum lycopersicum* (Sl)TIDS. **(D)** Legend depicting the protein sequence of the corresponding motif. GPPS, geranyl diphosphate synthase; FPPS, farnesyl diphosphate synthase; GGPPS, geranylgeranyl diphosphate synthase; LSU, large subunit; SSUI/II, small subunit I/II; GFPPS, geranylfarnesyl diphosphate synthase; PPPS, polyprenyl diphosphate synthase; and SPPS, solanesyl diphosphate synthase.

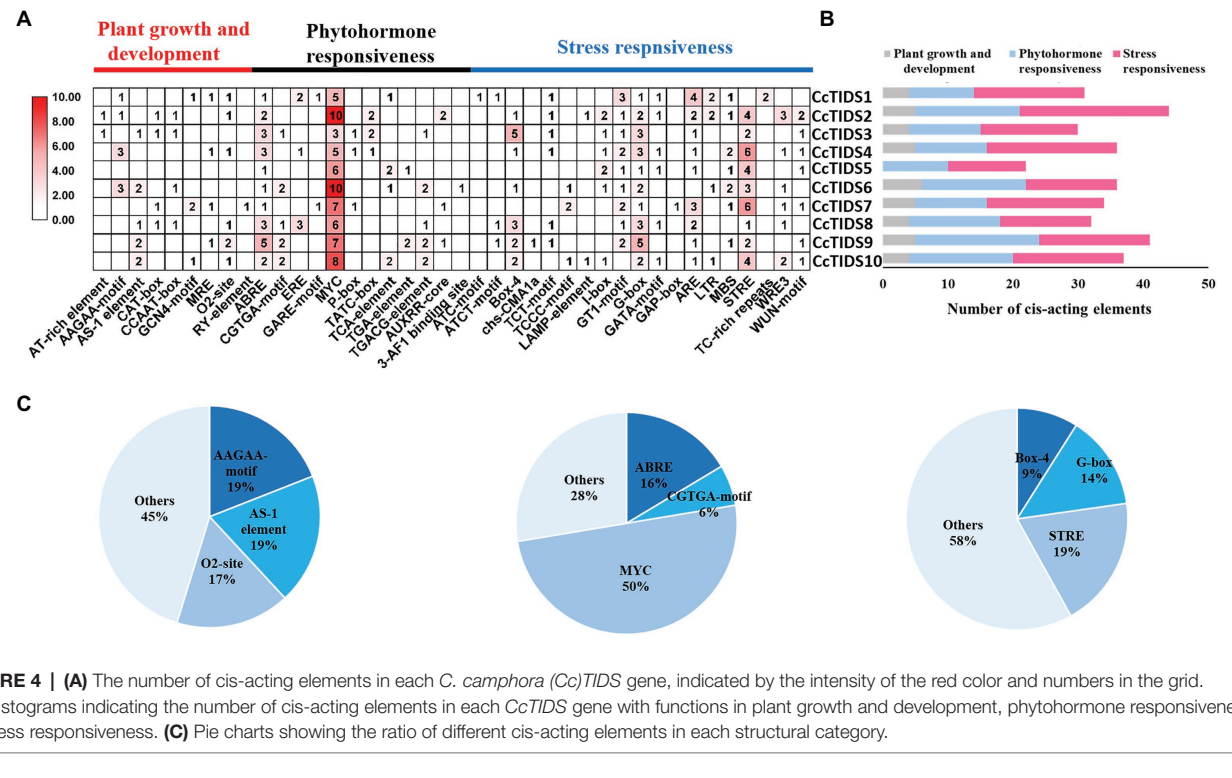

FIGURE 4 | (A) The number of cis-acting elements in each *C. camphora (Cc)TIDS* gene, indicated by the intensity of the red color and numbers in the grid. **(B)** Histograms indicating the number of cis-acting elements in each *CcTIDS* gene with functions in plant growth and development, phytohormone responsiveness, or stress responsiveness. **(C)** Pie charts showing the ratio of different cis-acting elements in each structural category.

According to enzymatic activity and the relevance of genetic evolution, the subfamily encoding TIDS-e proteins was subdivided into two clades, which contained genes encoding GPPS small subunits I and II. Among them, only one single GPPS small subunit-encoding *C. camphora* gene, *CcTIDS10*, was placed (within the small subunit I clade), indicating that the genome of *C. camphora* only contained GPPS small subunit genes from this class. *CcTIDS6* and *CcTIDS7* were clustered closely with the genes encoding the TIDS-b subfamily of proteins, indicating that these two genes may function as SPPS enzymes.

CcTIDS1, *CcTIDS2*, and *CcTIDS8* were found to belong to the TIDS-d subfamily, and therefore, may encode GGPPS, GFPPS, or PPPS enzymes. *CcTIDS3* and *CcTIDS9* were clustered with genes encoding TIDS-a proteins, indicating that they may encode FPPS enzymes. Although, the TIDS-c subfamily is known to contain mostly PPPS (Jia and Chen, 2016), genes encoding functionally characterized GPPS enzymes were included in this clade according to our analysis. Two *CcTIDSs*, *CcTIDS4*, and *CcTTPS5*, were clustered within this subfamily and were more closely related with the known GPPS-encoding genes, indicating that they may possess GPPS activity as well (**Figure 2**).

Conserved Motifs, Exon-Intron Structure, and Multiple Gene Alignment Analysis of *CcTIDS* in the Camphor Tree

The evolution of gene families may be accompanied by changes in gene structure, which may provide additional information from which the diversification of gene functions can be inferred (Ye et al., 2017; Cui et al., 2019; Li et al., 2020). For this reason, we further analyzed the organization of the exons and introns of *CcTIDS* genes and compared them with those from *A. thaliana* and *S. lycopersicum*. Our results showed that the *TIDS* genes within the same subgroup mostly exhibited similar arrangements of exons and introns. For instance, *CcTIDS6* and *CcTIDS7*, which were assigned to the TIDS-b-encoding subfamily, possessed six introns each. *CcTIDS1*, *CcTIDS2*, and *CcTIDS8* were all found to contain no introns, indicating that they do not encode GFPPS, which are known to contain two introns,

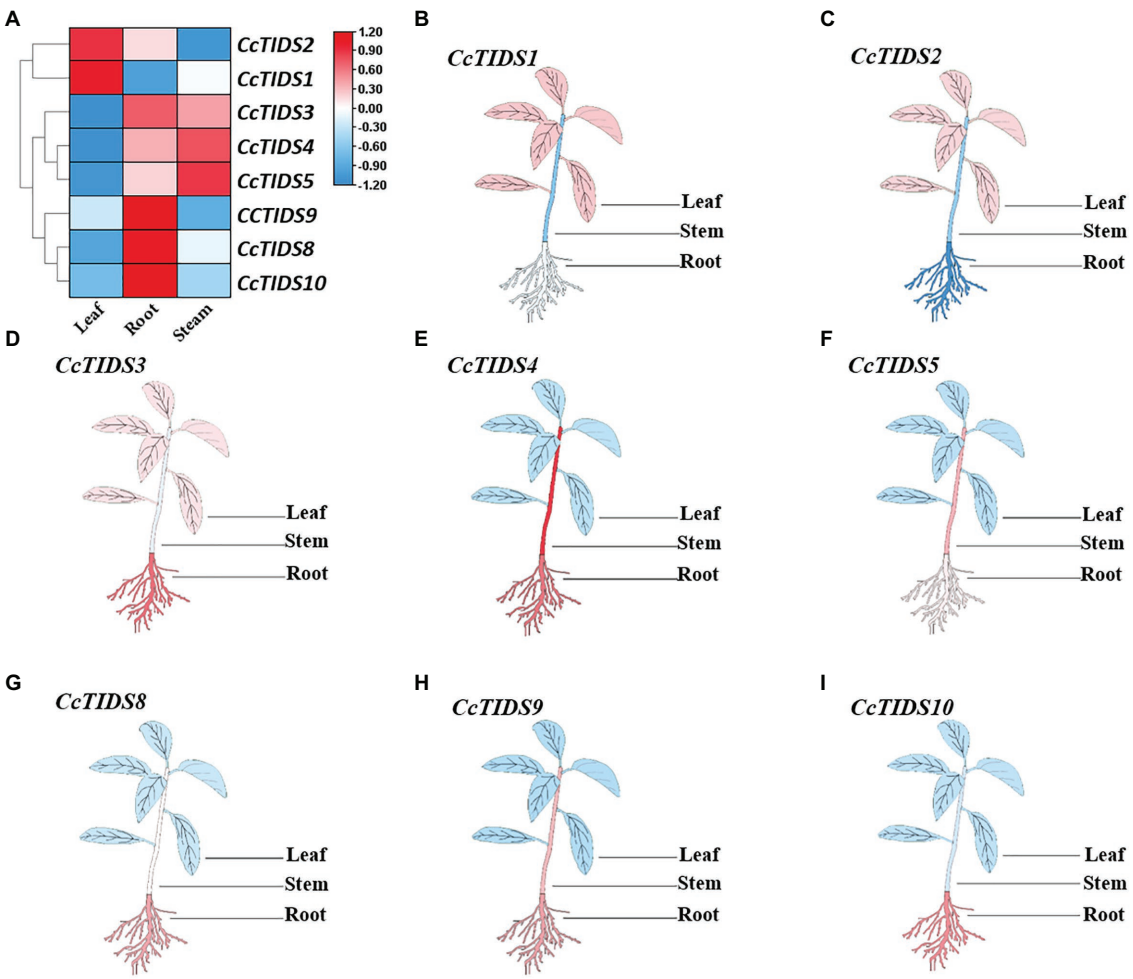

FIGURE 5 | Differential expression of *C. camphora (Cc)TIDS* genes in different tissues by transcriptome and quantitative real-time PCR (qRT-PCR). **(A)** FPKM value of the RNA sequencing data for each gene in leaf, root, and stem tissues. **(B–I)** Relative expression of each gene in the different plant tissues as calculated by qRT-PCR, with red and blue representing high and low expression levels, respectively.

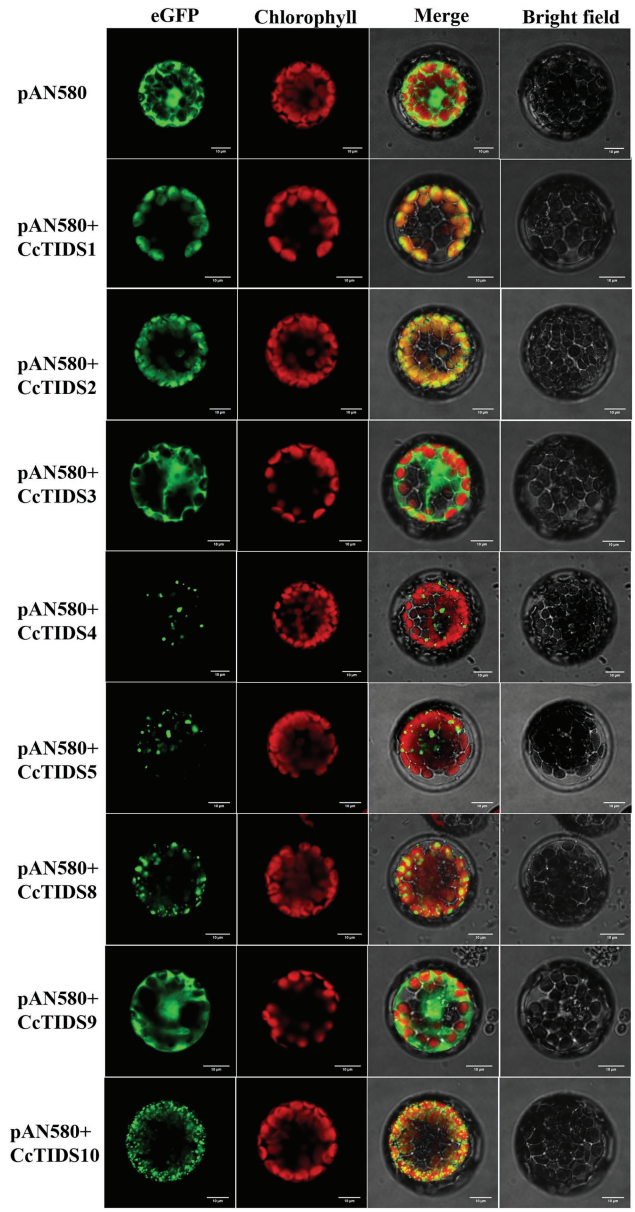

FIGURE 6 | Subcellular localization of eight putative *C. camphora* trans-prenyltransferase (CcTIDS) enzymes in *A. thaliana* protoplasts. eGFP, green fluorescent protein fluorescence image; chlorophyll, chlorophyll autofluorescence image; bright-field, transmission image; merged, all channels (eGFP, chlorophyll, and bright-field) combined.

although, they were closely clustered with known GFPPS genes, and belong to the same subfamily (**Figure 1**). However, it remains unclear whether these genes encode GPPS large subunits or GGPPS proteins.

Our analysis on conserved TIDS motifs showed that TIDS belonging to the same subfamily exhibited a similar motif composition and arrangement (**Figures 3A–D**). Catalytically important conserved motifs were recognized in every identified TIDS sequence after multiple sequence alignment of CcTIDS, *A. thaliana* TIDS (AtTIDS), and *S. lycopersicum* TIDS (SlTIDS). Two common aspartate-rich motifs, the first and second aspartate-rich motif (FARM and SARM, respectively), are known binding sites for DMAPP (Su et al., 2019), and these were present in sequences of all the TIDS sequences except for those of CcTIDS10, AtGPPS small subunit II, and SlGPPS small subunit I.

In addition, two conserved cysteine-rich (CxxxC) motifs were found in CcTIDS10, AtGPPS small subunit II, and SlGPPS small subunit I, and one each in CcTIDS1, CcTIDS2, and CcTIDS8. These CxxxC motifs play important roles in the interaction between the two subunits of the plant heterodimer GPPS. Our data suggest that CcTIDS10 is most likely a GPPS small subunit that forms a heterodimer with CcTIDS1, CcTIDS2, or CcTPS8 as a GPPS large subunit (**Supplementary Figure S1**).

Cis-Element Analysis of the TIDS Genes in *C. camphora*

We analyzed cis-elements in the promoter of the *CcTIDS* genes to better understand their potential regulation and function. In total, 343 cis-acting elements were identified, and they were grouped into three categories, which were responsible for plant growth and development, phytohormone responsiveness, and stress responsiveness, according to previous research (Abdullah et al., 2018). Of these, nine belong to the plant growth and development category, of which AAGAA motifs (involved in the endosperm) and AS-1 elements (involved in shoot expression) accounted for the highest proportion (19.05%).

The greatest proportion of cis-acting regulatory elements related to phytohormone response were myelocytomatosis (MYC) elements that are associated with methyl jasmonate (MeJA), accounting for 50%. All *CcTIDS* genes contained at least five MYC cis-acting elements, suggesting that *CcTIDS* gene expression might be moderated by MeJA. Nearly half of all the cis-acting elements were associated with stress responsiveness (167/343), of which the most common three cis-acting elements were the STRE motif related to stress (19%), G-box (14%), and box 4 (9%), which are associated with responsiveness to light (**Figure 4**). These results suggest that *CcTIDS* gene expression may be induced or suppressed by MeJA, and these genes may play roles in plant responses to a variety of abiotic stressors.

Expression Patterns of *CcTIDS* Genes in Different *C. camphora* Tissues

Our sequence alignment, evolution, and gene structure analyses all indicated that the *CcTIDS6* and *CcTIDS7* genes most likely encode SPPS enzymes. Hence, only the eight putative SC-TIDS-encoding genes were included in further analyses of expression patterns and subcellular localization, and an enzymatic assay.

To identify potential roles for *CcTIDS* genes in different tissues of *C. camphora*, we established a transcriptome database based on RNA sequencing of the leaves, roots, and stems (Accession number: PRJNA747104). qRT-PCR was used to verify these transcriptome data. We found that the expression profiles of *CcTIDS* genes were tissue-specific (**Figure 5**). For example, *CcTIDS1* and *CcTIDS2* were upregulated in the leaf tissue compared with the root and stem, whereas the expression

level of *CcTIDS3*, *9*, *8*, and *10* were upregulated in the root compared with the other tissues. This may indicate that there are functional differences in *CcTIDS* according to the different *C. camphora* organs from which the tissue was derived.

Subcellular Localization of Putative SC-TIDS Proteins

Subcellular localization predictions using four different online software programs revealed that CcTIDS1, 2, 8, and 10 are likely to be located in plastids. However, the results for CcTIDS3, 4, 5, and 9 were inconsistent (**Table 1**). In order to obtain exact subcellular location information for each candidate CcTIDS, we analyzed expression patterns using fluorescence microscopy (**Figure 6**). Consistent with all prediction programs, the GFP signals for CcTIDS1, 2, 8, and 10 were localized mainly within chloroplasts. Fluorescently tagged CcTIDS3 and CcTIDS9 proteins were both localized in the cytoplasm, while fluorescent signals for CcTIDS4 and CcTIDS5 recombinant proteins indicated that they may be located within mitochondria.

Heterologous Expression and *in vitro* Functional Characterization of the SC-TIDS Proteins

In vitro enzymatic assays were conducted using recombinant proteins extracted and purified from *E. coli* expression strains. The large subunit of CrGPPS was found to catalyze IPP and DMAPP to form geranyl diphosphate (GPP) and geranylgeranyl diphosphate (GGPP), whereas the small subunit of CrGPPS was found to be inactive alone. However, when both subunits were co-incubated, they catalyzed the formation of GPP (**Supplementary Figure S2**). This is accordance with the results reported in previous research, indicating that the method is valid for the verification of SC-TIDS *in vitro* enzymatic activity (Rai et al., 2013).

To characterize the two putative CcGPPS proteins (CcTIDS4 and 5), full-length and truncated versions of CcGPPSs were both analyzed. However, full-length recombinant CcGPPS protein was found to be completely insoluble. Therefore, soluble protein that was purified from the truncated version of each CcGPPS was used in the *in vitro* enzymatic activity assays. Unexpectedly, no catalytic products were detected for the truncated versions of both CcTIDS4 and CcTIDS5 under the conditions used here (**Table 2**).

Two full-length putative CcFPPS proteins (CcTIDS 3 and 9) lacking a transit peptide were generated. As expected, both recombinant proteins exhibited FPPS activity, producing FPP as the unique product, indicating that these two TIDS proteins are FPP synthases. Therefore, we reclassified these two TIDS proteins as CcFPPS1 and CcFPPS2 (**Figure 7**).

It was predicted that three putative CcGGPPS proteins (CcTIDS1, 2, and 8) and one putative CcGPPS small subunit (CcTIDS10) contained a transit peptide. Therefore, truncated versions of these four proteins were fused to His or MBP tags, resulting in the formation of soluble protein. The activity of these four soluble proteins was detected in all substrate combinations. As a result, CcTIDS1 and 2 emitted a prominent chromatographic signal corresponding to geranylgeraniol (C20), demonstrating that CcTIDS1 and 2 function individually as *bona fide* GGPPS enzymes. Hence, these two proteins were renamed CcGGPPS1 and CcGGPPS2 (**Figures 8A–F**). In contrast, CcTIDS8 and 10 did not exhibit any chromatographic peak signal for farnesol (C15), geraniol (C10), or geranylgeraniol (C20), suggesting that these proteins are inactive alone (**Figure 8G**).

CcTIDS8 Interacts With CcTIDS10 to Generate GPP

Because CcTIDS10 was identified as the only CcGPPS small subunit, we determined whether this protein was able to interact with CcGGPPS1/2 or CcTIDS8 to form active heterodimers. Intriguingly, when CcTIDS10 was co-incubated with CcTIDS8, but not CcGGPPS1 or 2, GPP was detected as a sole product (**Figures 8G-I**).

A yeast two-hybrid system (Y2H) was used to confirm the interaction of CcTIDS10 with CcTIDS8. As expected, CcTIDS10 was able to interact with CcTIDS8 and itself (**Figure 8J**). These results, along with the co-localization of CcTIDS10 and CcTIDS8 in the plastid, confirmed that the inactive CcTIDS10 and CcTIDS8 interact to form heteromeric functional GPPS, generating GPP as the sole product. Thus, CcTIDS10 was designated as CcGPPS.small subunit (SSU), whereas CcTIDS8 was renamed as CcGPPS.large subunit (LSU).

TABLE 2 | Functionally characterized enzymes in this study.

Genes	Full-length ORF/bp	Variants studied	Accepted substrates	Products	Rename
CcTIDS1	1,149	Ala62-end	DMAPP + IPP	GGPP	CcGGPPS1
CcTIDS2	1,161	Asp65-end	DMAPP + IPP	GGPP	CcGGPPS2
CcTIDS3	1,050	Full length	DMAPP + IPP GPP + IPP	FPP FPP	CcFPPS1
CcTIDS4	1,263	Full length	No activity	-	-
	1,263	M101-end	No activity	-	
CcTIDS5	1,284	Full length	No activity	-	-
	1,284	M108-end	No activity	-	
CcTIDS8	1,149	Lys56-end	No activity	-	CcGPPS.LSU
CcTIDS9	1,092	Full length	DMAPP + IPP GPP + IPP	FPP FPP	CcFPPS2
CcTIDS10	894	Asn65-end	No activity	-	CcGPPS.SSU
CcTIDS8&CcTIDS10	1,149&894	Lys56-end&Asn65-end	DMAPP + IPP	GPP	

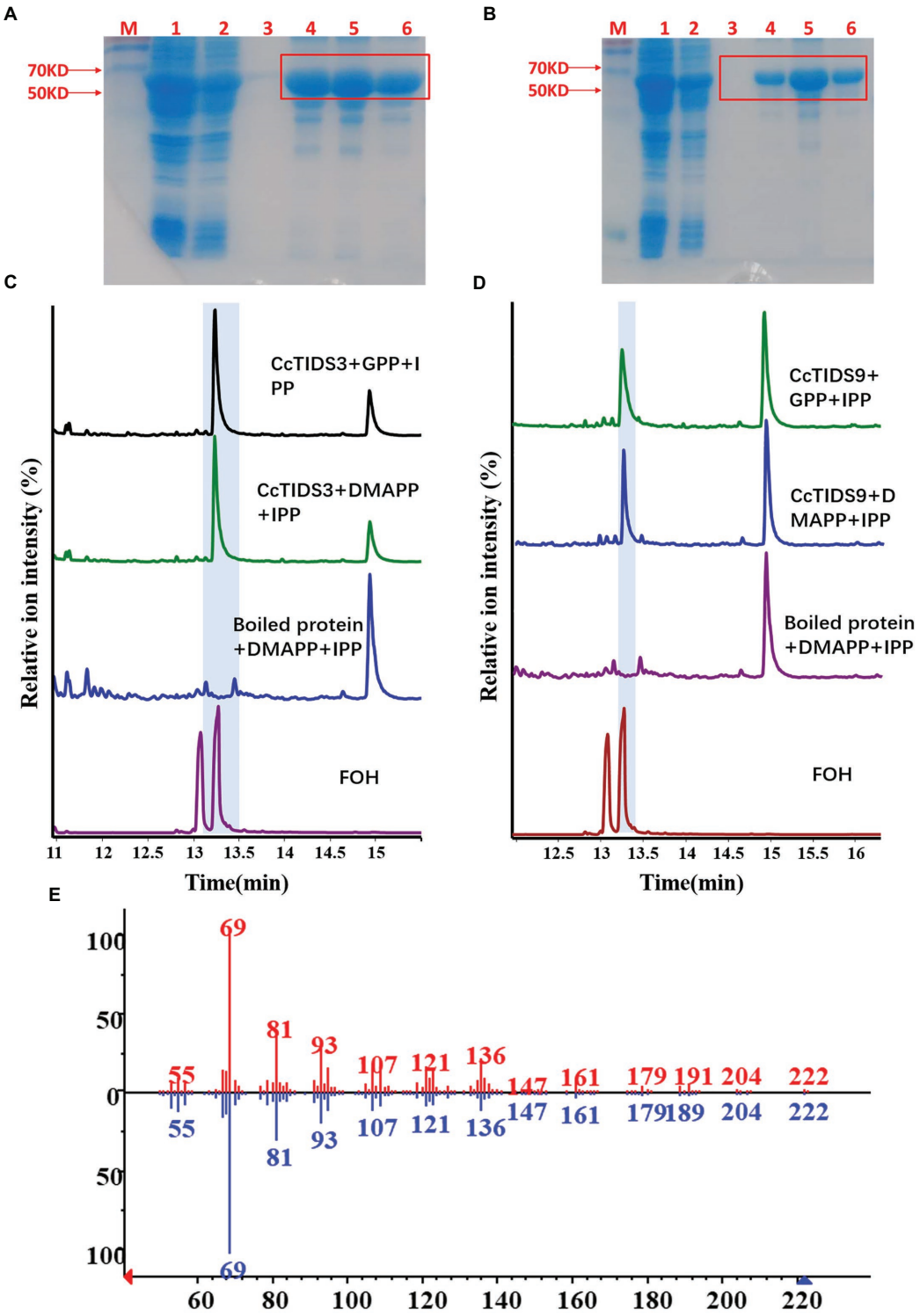

FIGURE 7 | Gas chromatography–mass spectrometry (GC–MS) profile showing *in vitro* reaction products of *C. camphora* trans-prenyltransferases CcTIDS3 and CcTIDS9, where different prenyldiphosphates were used as substrates. Expression and purification of **(A)** CcTIDS3 and **(B)** CcTIDS9 recombinant protein from *Escherichia coli* Rosetta (DE3) harboring pET32a(+)–CcTIDS3/CcTIDS9, where lane 1 shows the total protein after induction, 2 shows the soluble protein, and 3–6 shows the purified CcTIDS3 or CcTIDS9 recombinant protein from the first to the fourth collected tube. The GC–MS chromatogram of the reaction products generated by **(C)** CcTIDS3 and **(D)** CcTIDS9 and the acid hydrolysis products of farnesol (FOH) standards. **(E)** Mass spectra of farnesol standard (shown in red) and products in the NIST14/Wiley275 library (blue). GPP, geranyl diphosphate; DMAPP, dimethylallyl diphosphate; and IPP, isopentenyl diphosphate.

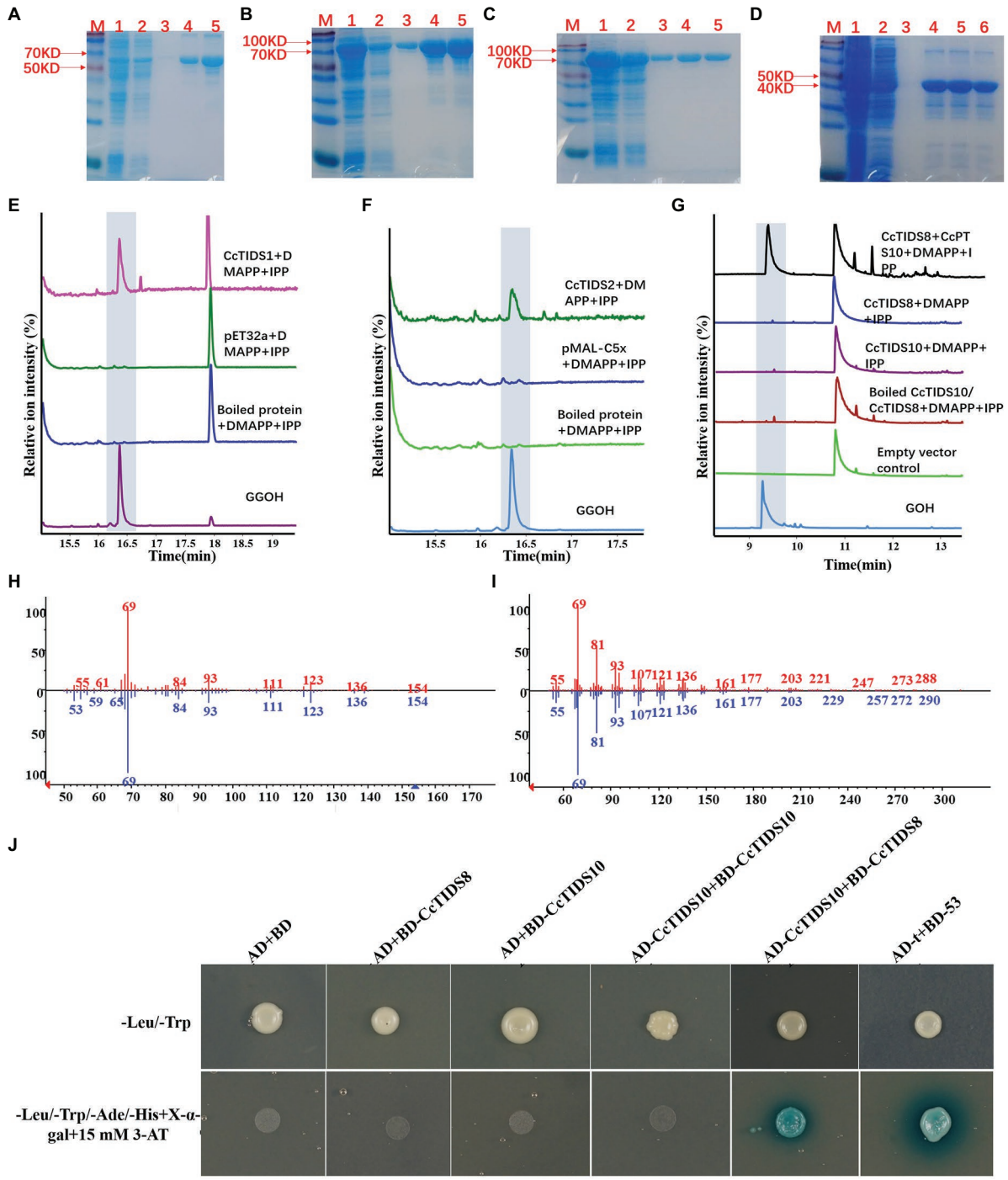

FIGURE 8 | Gas chromatography–mass spectrometry profile showing *in vitro* reaction products of *C. camphora* trans-prenyltransferases CcTIDS1, CcTIDS2, CcTIDS8, and CcTIDS9, where different prenyldiphosphates were used as substrates. Expression and purification of **(A)** CcTIDS1, **(B)** CcTIDS2, **(C)** CcTIDS8, and **(D)** CcTIDS10 recombinant protein from *E. coli* Rosetta (DE3) harboring pET32a(+)–CcTIDS1/CcTIDS10 or pMAL-C5x–CcTIDS2/CcTIDS8, where lane 1 shows the total protein after induction, 2 shows the soluble protein, 3–6 show the purified CcTIDS1 or CcTIDS2, CcTIDS8, or CcTIDS10 recombinant protein from the first to the fourth collected tube. The GC–MS chromatogram of the reaction products generated by **(E)** CcTIDS8 and CcTIDS10, **(F)** CcTIDS1, and **(G)** CcPTS2 and the acid hydrolysis products of geraniol (GOH) and gerylgeraniol (GGOH) standards. **(H)** Mass spectra of GOH standard (red) and products in the NIST14/Wiley275 library (blue). **(I)** Mass spectra of the GGOH standard (red) and products in the NIST14/Wiley275 library (blue). **(J)** Confirmation of the interaction between CcTIDS8 and CcTIDS10 using a yeast two-hybrid system, where blue color indicates an interaction. Yeast cells harboring both constructs were spotted on synthetic dropout (SD) medium lacking Ade, His, Leu, and Trp (SD/−4) to test for protein interactions. AD: pGADT7; BD: pGBDT7; AD-T: pGADT7::T; BD-53: pGBKT7-53. Cells co-transformed with pGADT7::T and pGBKT7-53 were included as a positive control, and pGADT7 and pGBDT7 as a vector control.

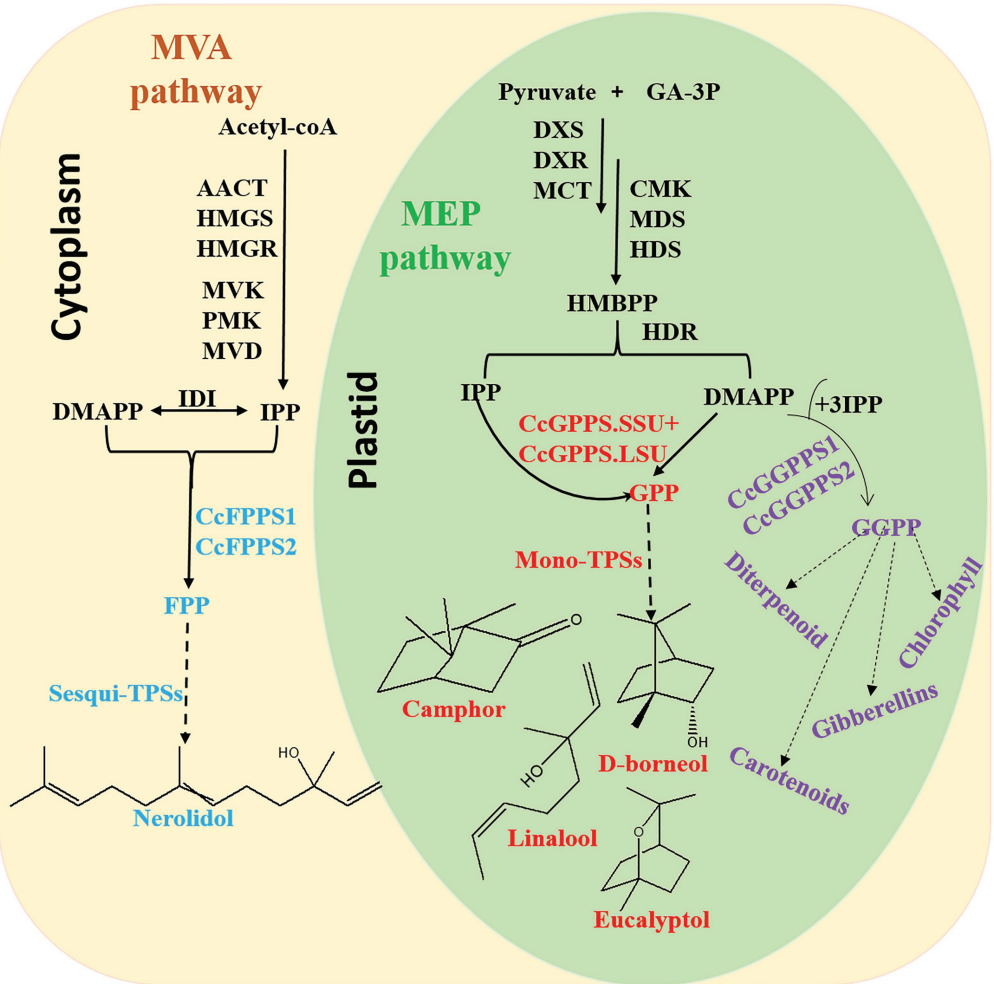

FIGURE 9 | Overview of the terpenoid biosynthesis pathway in *C. camphora*. AACT, acetoacetyl-coenzyme A thiolase; HMGS, 3-hydroxy-3-methylglutaryl coenzyme A synthase; HMGR, 3-hydroxy-3-methylglutaryl coenzyme A reductase; MVK, mevalonate kinase; PMK, 5-phospho mevalonate kinase; MVD, mevalonate diphosphate decarboxylase; IDI, isopentenyl diphosphate isomerase; FPPS, farnesyl diphosphate synthase; DXS,1-deoxy-D-xylulose-5-phosphate synthase; DXR,1-deoxy-D-xylulose-5-phosphate reductoisomerase; MCT, 2-C-methyl-D-erythritol-4-(cytidyl-5-diphosphate) transferase; CMK, 4-(cytidine50-diphospho)-2-C-methyl-D-erythritol kinase; MDS, 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase; HDS,1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate synthase; and HDR, 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate reductase.

DISCUSSION

In this study, we reported the characterization of the *C. camphora* TIDS gene family for the first time. Of the 10 identified *TIDS* genes, three (*CcGGPPS1*, *CcGGPPS2*, and *CcGPPS.LSU*) were clustered within the subfamily encoding TIDS-d enzymes, while one single gene, *CcTIDS10*, was closely related to the subfamily encoding TIDS-e enzymes. Apart from those in gymnosperms, TIDS proteins within the TIDS-d and TIDS-e subfamily have been classified as GGPPS paralogs in previous studies. As a result, many GGPPS homologs were predicted from the plant genomes (Coman et al., 2014).

The expansion of the GGPPS family in *A. thaliana* occurred at distinct evolutionary time points through different duplication mechanisms, including whole gene, tandem, and segmental genome duplications (Coman et al., 2014). The evolution of this *A. thaliana* GGPPS family likely involved neofunctionalization (with the duplicated gene developing a function that was not present in the ancestral gene), subfunctionalization (in which the duplicated and ancestral genes retain different parts of the original function of the ancestral gene), and pseudogenization (loss of function; Coman et al., 2014). Based on our fragment duplication and functional differentiation results for *CcGGPPS1*, *CcGGPPS2*, and *CcGPPS.LSU*, and the phylogenetic tree we constructed, these three genes may have been derived from the same ancestor, having undergone neo- or subfunctionalization.

Thus far, homodimeric GPPS has been functionally characterized in only a few plant species (Burke et al., 2004; Hsiao et al., 2008; Schmidt and Gershenzon, 2008; Gutensohn et al., 2013; Rai et al., 2013; Adal and Mahmoud, 2020). We found that the *C. camphora* genome contained two putative homodimeric *GPPS* genes (*CcTIDS4* and *CcTIDS5*).

Our evolutionary analysis suggested that these two genes were related to those encoding the functionally characterized homodimeric GPPS from *A. thaliana*, and subcellular localization experiments indicated that the two proteins encoded by these genes might localize to the mitochondria, consistent with earlier reports on *C. roseus* (Rai et al., 2013). However, there was no observable GPPS activity associated with the proteins *in vitro*.

Instead, CcGPPS.SSU (CcTIDS10) and CcGPPS.LSU (CcTIDS8), both identified in our study, were found to interact with one another to form a heterodimeric CcGPPS that produced GPP from IPP and DMAPP substrates *in vitro*. These proteins were found to be inactive alone. This is consistent with previous reports in both *Mentha*×*piperita* subunits (Burke et al., 2004), both *Lavendula*×*intermedia* subunits (Adal and Mahmoud, 2020), and *A. thaliana* GPPS.SSU2 (Wang and Dixon, 2009).

CcGPPS.SSU and CcGPPS.LSU were primarily expressed in roots compared with stems and leaves. Furthermore, both CcGPPS.SSU and CcGPPS.LSU were found to reside in plastids, which are the known site of monoterpene biosynthesis (Magnard et al., 2015; Yin et al., 2017). These results suggest that heteromeric CcGPPS likely participates in the biosynthesis of GPP in *C. camphora*.

Interestingly, all *CcTIDS* genes had orthologs of greater than 90% identity with the *C. micranthum* genome, with the sole exception of *CcGPPS.SSU* (**Supplementary Figure S3**). *Cinnamomum micranthum*, known as the stout camphor tree, is the sister species of *C. camphora*. The genome of *C. micranthum* with a contig N_{50} of 0.9 Mb has been published. However, contig N_{50} of the genome of *C. camphora* is 23.89 Mb (data not shown), indicating that the assembly quality of the *C. camphora* genome is significantly better than that of the *C. micranthum* genome. This lower quality may explain why the GPPS.SSU gene is missing from the *C. micranthum* genome.

In addition, although, CcTIDS4 and CcTIDS5 did not possess GPPS activity in our study, they may perform different functions. For example, At2g34630 from *A. thaliana* was initially classified as a homodimeric GPPS (Bouvier et al., 2000), but further study indicated that *At2g34630* silencing did not affect the production of monoterpenes, and At2g34630 was subsequently identified as a PPPS (Hsieh et al., 2011). Therefore, the functions of CcTIDS4 and 5 should be investigated in future studies.

In conclusion, this is the first comprehensive and systematic genome-wide analysis of TIDS gene families in *C. camphora*. In total, 10 TIDS genes in the borneol chemotype of *C. camphora* were identified. These genes likely expanded through segmental duplication events, and expression was found to respond to multiple abiotic stressors *via* cis-acting elements in the promoter region. Eight putative SC-TIDS were identified, six of which were catalytically active, including a heteromeric GPPS (composed of CcGPPS.SSU and CcGPPS.LSU), two CcFPPS (CcTIDS3 and 9), and two CcGGPPS (CcTIDS1 and 2), which catalyzed the biosynthesis of GPP, FPP, and GGPP, respectively (**Figure 9**).

Finally, *CcTIDS3*, *8*, *9*, and *10* were more active in roots compared with stems and leaves, which were confirmed by both transcriptome analysis and qRT-PCR experiments. These novel insights provide the basis for further investigation of the TIDS family in *C. camphora*, and a good foundation from which to develop a metabolic engineering approach to increase production of pharmacologically valuable essential oils from the camphor tree.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

ZY conceived, designed, and performed the experiments, analyzed the data, prepared the figures and tables, and authored and reviewed drafts of the paper. CX and LL conceived, designed, and performed the experiments, analyzed the data, and authored and reviewed drafts of the paper. SL, YH, and WA conceived, designed, and performed the experiments, and authored and reviewed drafts of the paper. SH and XZ conceived and designed the experiments, analyzed the data, prepared the figures and tables, and authored and reviewed drafts of the paper. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the National Natural Science Foundation of China (grant number: 81903741).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.708697/full#supplementary-material

REFERENCES

Abdullah, M., Cheng, X., Cao, Y., Su, X., Manzoor, M. A., Gao, J., et al. (2018). Zinc finger-homeodomain transcriptional factors (ZHDs) in upland cotton (*Gossypium hirsutum*): genome-wide identification and expression analysis in fiber development. *Front. Genet.* 9:357. doi: 10.3389/fgene.2018.00357

Adal, A. M., and Mahmoud, S. S. (2020). Short-chain isoprenyl diphosphate synthases of lavender (Lavandula). *Plant Mol. Biol.* 102, 517–535. doi: 10.1007/s11103-020-00962-8

Akhtar, T. A., Matsuba, Y., Schauvinhold, I., Yu, G., Lees, H. A., Klein, S. E., et al. (2013). The tomato cis-prenyltransferase gene family. *Plant J.* 73, 640–652. doi: 10.1111/tpj.12063

Amborella Genome Project (2013). The Amborella genome and the evolution of flowering plants. *Science* 342:1241089. doi: 10.1126/science.1241089

Arabidopsis Genome Initiative (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796–815. doi: 10.1038/35048692

Athanasakoglou, A., and Grypioti, E. (2019). Isoprenoid biosynthesis in the diatom Haslea ostrearia. *New Phytol.* 222, 230–243. doi: 10.1111/nph.15586

Banks, J. A., Nishiyama, T., Hasebe, M., Bowman, J. L., Gribskov, M., dePamphilis, C., et al. (2011). The Selaginella genome identifies genetic changes associated with the evolution of vascular plants. *Science* 332, 960–963. doi: 10.1126/science.1203810

Barja, M. V., and Rodríguez-Concepción, M. (2020). A simple in vitro assay to measure the activity of geranylgeranyl diphosphate synthase and other short-chain prenyltransferases. *Methods Mol. Biol.* 2083, 27–38. doi: 10.1007/978-1-4939-9952-1_2

Beck, G., Coman, D., Herren, E., Ruiz-Sola, M. A., Rodríguez-Concepción, M., Gruissem, W., et al. (2013). Characterization of the GGPP synthase gene family in *Arabidopsis thaliana*. *Plant Mol. Biol.* 82, 393–416. doi: 10.1007/s11103-013-0070-z

Bouvier, F., Suire, C., d'Harlingue, A., Backhaus, R. A., and Camara, B. (2000). Molecular cloning of geranyl diphosphate synthase and compartmentation of monoterpene synthesis in plant cells. *Plant J.* 24, 241–252. doi: 10.1046/j.1365-313x.2000.00875.x

Burke, C., Klettke, K., and Croteau, R. (2004). Heteromeric geranyl diphosphate synthase from mint: construction of a functional fusion protein and inhibition by bisphosphonate substrate analogs. *Arch. Biochem. Biophys.* 422, 52–60. doi: 10.1016/j.abb.2003.12.003

Chai, Y., Yin, Z., Fan, Q., Zhang, Z., Ye, K., Xu, Y., et al. (2019). Protective effects of angong niuhuang pill on early atherosclerosis in apoE(−/−) mice by reducing the inflammatory response. *Evid. Based Complement. Alternat. Med.* 2019:9747212. doi: 10.1155/2019/9747212

Chaw, S. M., Liu, Y. C., Wu, Y. W., Wang, H. Y., Lin, C. I., Wu, C. S., et al. (2019). Stout camphor tree genome fills gaps in understanding of flowering plant genome evolution. *Nat. Plants* 5, 63–73. doi: 10.1038/s41477-018-0337-0

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., et al. (2020a). 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, Q., Fan, D., and Wang, G. (2015). Heteromeric geranyl(geranyl) diphosphate synthase is involved in monoterpene biosynthesis in *Arabidopsis* flowers. *Mol. Plant* 8, 1434–1437. doi: 10.1016/j.molp.2015.05.001

Chen, J., Tang, C., Zhang, R., Ye, S., Zhao, Z., Huang, Y., et al. (2020b). Metabolomics analysis to evaluate the antibacterial activity of the essential oil from the leaves of *Cinnamomum camphora* (Linn.) Presl. *J. Ethnopharmacol.* 253:112652. doi: 10.1016/j.jep.2020.112652

Chen, Z. X., Xu, Q. Q., Shan, C. S., Shi, Y. H., Wang, Y., Chang, R. C., et al. (2019). Borneol for regulating the permeability of the blood-brain barrier in experimental ischemic stroke: preclinical evidence and possible mechanism. *Oxidative Med. Cell. Longev.* 2019:2936737. doi: 10.1155/2019/2936737

Chen, C., Zheng, Y., Zhong, Y., Wu, Y., Li, Z., Xu, L. A., et al. (2018). Transcriptome analysis and identification of genes related to terpenoid biosynthesis in *Cinnamomum camphora*. *BMC Genomics* 19:550. doi: 10.1186/s12864-018-4941-1

Choi, K. R., Jang, W. D., Yang, D., Cho, J. S., Park, D., and Lee, S. Y. (2019). Systems metabolic engineering strategies: integrating systems and synthetic biology with metabolic engineering. *Trends Biotechnol.* 37, 817–837. doi: 10.1016/j.tibtech.2019.01.003

Chou, K. C., and Shen, H. B. (2010). Plant-mPLoc: a top-down strategy to augment the power for predicting plant protein subcellular localization. *PLoS One* 5:e11335. doi: 10.1371/journal.pone.0011335

Coman, D., Altenhoff, A., Zoller, S., Gruissem, W., and Vranová, E. (2014). Distinct evolutionary strategies in the GGPPS family from plants. *Front. Plant Sci.* 5:230. doi: 10.3389/fpls.2014.00230

Cui, L., Yang, G., Yan, J., Pan, Y., and Nie, X. (2019). Genome-wide identification, expression profiles and regulatory network of MAPK cascade gene family in barley. *BMC Genomics* 20:750. doi: 10.1186/s12864-019-6144-9

Dudareva, N., Murfitt, L. M., Mann, C. J., Gorenstein, N., Kolosova, N., Kish, C. M., et al. (2000). Developmental regulation of methyl benzoate biosynthesis and emission in snapdragon flowers. *Plant Cell* 12, 949–961. doi: 10.1105/tpc.12.6.949

Dudareva, N., Pichersky, E., and Gershenzon, J. (2004). Biochemistry of plant volatiles. *Plant Physiol.* 135, 1893–1902. doi: 10.1104/pp.104.049981

Emanuelsson, O., Nielsen, H., Brunak, S., and von Heijne, G. (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol.* 300, 1005–1016. doi: 10.1006/jmbi.2000.3903

Emanuelsson, O., Nielsen, H., and von Heijne, G. (1999). ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci.* 8, 978–984. doi: 10.1110/ps.8.5.978

Guo, X., Cui, M., Deng, M., Liu, X., Huang, X., Zhang, X., et al. (2017). Molecular differentiation of five *Cinnamomum camphora* chemotypes using desorption atmospheric pressure chemical ionization mass spectrometry of raw leaves. *Sci. Rep.* 7:46579. doi: 10.1038/srep46579

Gutensohn, M., Orlova, I., Nguyen, T. T., Davidovich-Rikanati, R., Ferruzzi, M. G., Sitrit, Y., et al. (2013). Cytosolic monoterpene biosynthesis is supported by plastid-generated geranyl diphosphate substrate in transgenic tomato fruits. *Plant J.* 75, 351–363. doi: 10.1111/tpj.12212

Hivert, G., Davidovich-Rikanati, R., Bar, E., Sitrit, Y., Schaffer, A., Dudareva, N., et al. (2020). Prenyltransferases catalyzing geranyldiphosphate formation in tomato fruit. *Plant Sci.* 296:110504. doi: 10.1016/j.plantsci.2020.110504

Horton, P., Park, K. J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C. J., et al. (2007). WoLF PSORT: protein localization predictor. *Nucleic Acids Res.* 35, W585–W587. doi: 10.1093/nar/gkm259

Hsiao, Y. Y., Jeng, M. F., Tsai, W. C., Chuang, Y. C., Li, C. Y., Wu, T. S., et al. (2008). A novel homodimeric geranyl diphosphate synthase from the orchid *Phalaenopsis bellina* lacking a DD(X)2-4D motif. *Plant J.* 55, 719–733. doi: 10.1111/j.1365-313X.2008.03547.x

Hsieh, F. L., Chang, T. H., Ko, T. P., and Wang, A. H. (2011). Structure and mechanism of an *Arabidopsis* medium/long-chain-length prenyl pyrophosphate synthase. *Plant Physiol.* 155, 1079–1090. doi: 10.1104/pp.110.168799

Huang, J., Tang, X., Ye, F., He, J., and Kong, X. (2016). Clinical therapeutic effects of aspirin in combination with fufang danshen diwan, a traditional chinese medicine formula, on coronary heart disease: a systematic review and meta-analysis. *Cell. Physiol. Biochem.* 39, 1955–1963. doi: 10.1159/000447892

Jia, Q., and Chen, F. (2016). Catalytic functions of the isoprenyl diphosphate synthase superfamily in plants: a growing repertoire. *Mol. Plant* 9, 189–191. doi: 10.1016/j.molp.2015.12.020

Jiao, Y., Peluso, P., Shi, J., Liang, T., Stitzer, M. C., Wang, B., et al. (2017). Improved maize reference genome with single-molecule technologies. *Nature* 546, 524–527. doi: 10.1038/nature22971

Johnson, S. R., and Bhat, W. W. (2019). Promiscuous terpene synthases from *Prunella vulgaris* highlight the importance of substrate and compartment switching in terpene synthase evolution. *New Phytol.* 223, 323–335. doi: 10.1111/nph.15778

Johnson, S. R., Bhat, W. W., Bibik, J., Turmo, A., Hamberger, B., Evolutionary Mint Genomics, C., et al. (2019). A database-driven approach identifies additional diterpene synthase activities in the mint family (Lamiaceae). *J. Biol. Chem.* 294, 1349–1362. doi: 10.1074/jbc.RA118.006025

Kopcsayová, D., and Vranová, E. (2019). Functional gene network of prenyltransferases in *Arabidopsis thaliana*. *Molecules* 24:4556. doi: 10.3390/molecules24244556

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., et al. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* 30, 325–327. doi: 10.1093/nar/30.1.325

Li, J., Gao, K., Yang, X., Khan, W. U., Guo, B., Guo, T., et al. (2020). Identification and characterization of the CONSTANS-like gene family and its expression profiling under light treatment in *Populus*. *Int. J. Biol. Macromol.* 161, 999–1010. doi: 10.1016/j.ijbiomac.2020.06.056

Li, J., and Mutanda, I. (2019). Chloroplastic metabolic engineering coupled with isoprenoid pool enhancement for committed taxanes biosynthesis in *Nicotiana benthamiana*. *Nat. Commun.* 10:4850. doi: 10.1038/s41467-019-12879-y

Liang, P. H., Ko, T. P., and Wang, A. H. (2002). Structure, mechanism and function of prenyltransferases. *Eur. J. Biochem.* 269, 3339–3354. doi: 10.1046/j.1432-1033.2002.03014.x

Liang, C. J., Li, J. H., Zhang, Z., Zhang, J. Y., Liu, S. Q., and Yang, J. (2018). Suppression of MIF-induced neuronal apoptosis may underlie the therapeutic effects of effective components of Fufang Danshen in the treatment of Alzheimer's disease. *Acta Pharmacol. Sin.* 39, 1421–1438. doi: 10.1038/aps.2017.210

Ma, R., Su, P., Jin, B., Guo, J., Tian, M., Mao, L., et al. (2021). Molecular cloning and functional identification of a high-efficiency (+)-borneol

dehydrogenase from *Cinnamomum camphora* (L.) Presl. *Plant Physiol. Biochem.* 158, 363–371. doi: 10.1016/j.plaphy.2020.11.023

Magnard, J. L., Roccia, A., Caissard, J. C., Vergne, P., Sun, P., Hecquet, R., et al. (2015). Plant volatiles. Biosynthesis of monoterpene scent compounds in roses. *Science* 349, 81–83. doi: 10.1126/science.aab0696

Miller, G. P., and Bhat, W. W. (2020). The biosynthesis of the anti-microbial diterpenoid leubethanol in *Leucophyllum frutescens* proceeds via an all-cis prenyl intermediate. *Plant J.* 104, 693–705. doi: 10.1111/tpj.14957

Nagel, R., Bernholz, C., Vranová, E., Košuth, J., Bergau, N., Ludwig, S., et al. (2015). *Arabidopsis thaliana* isoprenyl diphosphate synthases produce the C25 intermediate geranylfarnesyl diphosphate. *Plant J.* 84, 847–859. doi: 10.1111/tpj.13064

Nystedt, B., Street, N. R., Wetterbom, A., Zuccolo, A., Lin, Y. C., Scofield, D. G., et al. (2013). The Norway spruce genome sequence and conifer genome evolution. *Nature* 497, 579–584. doi: 10.1038/nature12211

Pichersky, E., Noel, J. P., and Dudareva, N. (2006). Biosynthesis of plant volatiles: nature's diversity and ingenuity. *Science* 311, 808–811. doi: 10.1126/science.1118510

Rai, A., Smita, S. S., Singh, A. K., Shanker, K., and Nagegowda, D. A. (2013). Heteromeric and homomeric geranyl diphosphate synthases from *Catharanthus roseus* and their role in monoterpene indole alkaloid biosynthesis. *Mol. Plant* 6, 1531–1549. doi: 10.1093/mp/sst058

Ren, L., Wang, J., Feng, L., Wang, S., and Li, J. (2018). Efficacy of suxiao jiuxin pill on coronary heart disease: a meta-analysis of randomized controlled trials. *Evid. Based Complement. Alternat. Med.* 2018:9745804. doi: 10.1155/2018/9745804

Rensing, S. A., Lang, D., Zimmer, A. D., Terry, A., Salamov, A., Shapiro, H., et al. (2008). The Physcomitrella genome reveals evolutionary insights into the conquest of land by plants. *Science* 319, 64–69. doi: 10.1126/science.1150646

Schmidt, A., and Gershenzon, J. (2008). Cloning and characterization of two different types of geranyl diphosphate synthases from Norway spruce (Picea abies). *Phytochemistry* 69, 49–57. doi: 10.1016/j.phytochem.2007.06.022

Su, P., Gao, L., Tong, Y., Guan, H., Liu, S., Zhang, Y., et al. (2019). Analysis of the role of geranylgeranyl diphosphate synthase 8 from Tripterygium wilfordii in diterpenoids biosynthesis. *Plant Sci.* 285, 184–192. doi: 10.1016/j.plantsci.2019.05.013

Subramanian, B., Gao, S., Lercher, M. J., Hu, S., and Chen, W.-H. (2019). Evolview v3: a webserver for visualization, annotation, and management of phylogenetic trees. *Nucleic Acids Res.* 47, W270–W275. doi: 10.1093/nar/gkz357

Tholl, D., and Lee, S. (2011). Terpene specialized metabolism in *Arabidopsis thaliana*. *Arabidopsis Book* 9:e0143. doi: 10.1199/tab.0143

Tomato Genome Consortium (2012). The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* 485, 635–641. doi: 10.1038/nature11119

Ueoka, H., Sasaki, K., Miyawaki, T., Ichino, T., Tatsumi, K., Suzuki, S., et al. (2020). A cytosol-localized geranyl diphosphate synthase from *Lithospermum erythrorhizon* and its molecular evolution. *Plant Physiol.* 182, 1933–1945. doi: 10.1104/pp.19.00999

Vandermoten, S., Haubruge, E., and Cusson, M. (2009). New insights into short-chain prenyltransferases: structural features, evolutionary history and potential for selective inhibition. *Cell. Mol. Life Sci.* 66, 3685–3695. doi: 10.1007/s00018-009-0100-9

Wang, C., Chen, Q., Fan, D., Li, J., Wang, G., and Zhang, P. (2016). Structural analyses of short-chain prenyltransferases identify an evolutionarily conserved GFPPS clade in brassicaceae plants. *Mol. Plant* 9, 195–204. doi: 10.1016/j.molp.2015.10.010

Wang, G., and Dixon, R. A. (2009). Heterodimeric geranyl(geranyl)diphosphate synthase from hop (*Humulus lupulus*) and the evolution of monoterpene biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* 106, 9914–9919. doi: 10.1073/pnas.0904069106

Wang, J., Lin, H. X., Su, P., Chen, T., Guo, J., Gao, W., et al. (2019). Molecular cloning and functional characterization of multiple geranylgeranyl pyrophosphate synthases (ApGGPPS) from *Andrographis paniculata*. *Plant Cell Rep.* 38, 117–128. doi: 10.1007/s00299-018-2353-y

Wang, H., Ma, D., Yang, J., Deng, K., Li, M., Ji, X., et al. (2018). An integrative volatile terpenoid profiling and transcriptomics analysis for gene mining and functional characterization of AvBPPS and AvPS involved in the monoterpenoid biosynthesis in amomum villosum. *Front. Plant Sci.* 9:846. doi: 10.3389/fpls.2018.00846

Wijfjes, R. Y., and Smit, S. (2019). Hecaton: reliably detecting copy number variation in plant genomes using short read sequencing data. *BMC Genomics* 20:818. doi: 10.1186/s12864-019-6153-8

Wong, W. C., Yap, C. K., Eisenhaber, B., and Eisenhaber, F. (2015). dissectHMMER: a HMMER-based score dissection framework that statistically evaluates fold-critical sequence segments for domain fold similarity. *Biol. Direct* 10:39. doi: 10.1186/s13062-015-0068-3

Yang, Z., An, W., Liu, S., Huang, Y., Xie, C., Huang, S., et al. (2020). Mining of candidate genes involved in the biosynthesis of dextrorotatory borneol in *Cinnamomum burmannii* by transcriptomic analysis on three chemotypes. *PeerJ* 8:e9311. doi: 10.7717/peerj.9311

Ye, J., Yang, H., Shi, H., Wei, Y., Tie, W., Ding, Z., et al. (2017). The MAPKKK gene family in cassava: genome-wide identification and expression analysis against drought stress. *Sci. Rep.* 7:14939. doi: 10.1038/s41598-017-13988-8

Yin, J. L., Wong, W. S., Jang, I. C., and Chua, N. H. (2017). Co-expression of peppermint geranyl diphosphate synthase small subunit enhances monoterpene production in transgenic tobacco plants. *New Phytol.* 213, 1133–1144. doi: 10.1111/nph.14280

Yoo, S. D., Cho, Y. H., and Sheen, J. (2007). Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nat. Protoc.* 2, 1565–1572. doi: 10.1038/nprot.2007.199

Yu, J., Hu, S., Wang, J., Wong, G. K., Li, S., Liu, B., et al. (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. indica). *Science* 296, 79–92. doi: 10.1126/science.1068037

Yu, H., Ren, X., Liu, Y., Xie, Y., Guo, Y., Cheng, Y., et al. (2019). Extraction of *Cinnamomum camphora* chvar. Borneol essential oil using neutral cellulase assisted-steam distillation: optimization of extraction, and analysis of chemical constituents. *Ind. Crop. Prod.* 141:111794. doi: 10.1016/j.indcrop.2019.111794

Zhou, F., and Pichersky, E. (2020). The complete functional characterisation of the terpene synthase family in tomato. *New Phytol.* 226, 1341–1360. doi: 10.1111/nph.16431

Zhu, J.-H., Xia, D.-N., Xu, J., Guo, D., Li, H.-L., Wang, Y., et al. (2020). Identification of the bHLH gene family in *Dracaena cambodiana* reveals candidate genes involved in flavonoid biosynthesis. *Ind. Crop. Prod.* 150:112407. doi: 10.1016/j.indcrop.2020.112407

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 24 September 2021
doi: 10.3389/fpls.2021.681979

Two Main Biosynthesis Pathways Involved in the Synthesis of the Floral Aroma of the Nacional Cocoa Variety

Kelly Colonges[1,2,3,4], Juan-Carlos Jimenez[5], Alejandra Saltos[5], Edward Seguine[6], Rey Gastón Loor Solorzano[5], Olivier Fouet[1,2], Xavier Argout[1,2], Sophie Assemat[3,4], Fabrice Davrieux[3,4], Emile Cros[3,4], Renaud Boulanger[3,4]* and Claire Lanaud[1,2]**

[1] Cirad, UMR Amélioration Génétique et Adaptation des Plantes Méditerranéennes et Tropicales, Montpellier, France, [2] Amélioration Génétique et Adaptation des Plantes Méditerranéennes et Tropicales, Univ Montpellier, Cirad, INRAE, Institut Agro, Montpellier, France, [3] Cirad, UMR Qualisud, Montpellier, France, [4] Qualisud, Univ Montpellier, Avignon Université, Cirad, Institut Agro, IRD, Université de La Réunion, Montpellier, France, [5] Cocoa and Coffee Research Program, Instituto Nacional de Investigacion Agropecurias, Quito, Ecuador, [6] Guittard, Burlingame, CA, United States

OPEN ACCESS

Edited by:
María José Jordán, Instituto Murciano de Investigación y Desarrollo Agrario y Alimentario (IMIDA), Spain

Reviewed by:
Natasha Spadafora, University of Calabria, Italy
Jinhe Bai, Horticultural Research Laboratory (USDA-ARS), United States

***Correspondence:**
Kelly Colonges
kelly.colonges@cirad.fr
Renaud Boulanger
renaud.boulanger@cirad.fr
Claire Lanaud
claire.lanaud@cirad.fr

Specialty section:
This article was submitted to Crop and Product Physiology, a section of the journal Frontiers in Plant Science

Received: *17 March 2021*
Accepted: *30 August 2021*
Published: *24 September 2021*

Citation:
Colonges K, Jimenez J-C, Saltos A, Seguine E, Loor Solorzano RG, Fouet O, Argout X, Assemat S, Davrieux F, Cros E, Boulanger R and Lanaud C (2021) Two Main Biosynthesis Pathways Involved in the Synthesis of the Floral Aroma of the Nacional Cocoa Variety. Front. Plant Sci. 12:681979. doi: 10.3389/fpls.2021.681979

Theobroma cacao is the only source that allows the production of chocolate. It is of major economic importance for producing countries such as Ecuador, which is the third-largest cocoa producer in the world. Cocoa is classified into two groups: bulk cocoa and aromatic fine flavour cocoa. In contrast to bulk cocoa, fine flavour cocoa is characterised by fruity and floral notes. One of the characteristics of Nacional cocoa, the emblematic cocoa of Ecuador, is its aromatic ARRIBA flavour. This aroma is mainly composed of floral notes whose genetic and biochemical origin is not well-known. This research objective is to study the genetic and biochemical determinism of the floral aroma of modern Nacional cocoa variety from Ecuador. Genome-Wide Association Study (GWAS) was conducted on a population of 152 genotypes of cocoa trees belonging to the population variety of modern Nacional. Genome-Wide Association Study was conducted by combining SSR and SNP genotyping, assaying biochemical compounds (in roasted and unroasted beans), and sensory evaluations from various tastings. This analysis highlighted different areas of association for all types of traits. In a second step, a search for candidate genes in these association zones was undertaken, which made it possible to find genes potentially involved in the biosynthesis pathway of the biochemical compound identified in associations. Our results show that two biosynthesis pathways seem to be mainly related to the floral note of Nacional cocoa: the monoterpene biosynthesis pathway and the L-phenylalanine degradation pathway. As already suggested, the genetic background would therefore appear as largely explaining the floral note of cocoa.

Keywords: GWAS, cocoa aroma, floral, monoterpenes, phenolic compounds

INTRODUCTION

Theobroma cacao L. is native to the tropical rainforests of northern South America and is a member of the family *Malvaceae* (Bayer and Kubitzki, 2003). The cocoa tree is a diploid ($2n = 20$) with a small genome that is now sequenced and of which 96.7% of the assembly is anchored on all 10 chromosomes (Argout et al., 2011, 2017; Motamayor et al., 2013).

Cocoa farming represents an important economic issue for many tropical countries because it is the only source of chocolate supply. In 2018/2019, cocoa production represented more than 4,780 thousand tonnes worldwide. The three largest producers are Ivory Coast, Ghana, and Ecuador with, respectively, 1,964, 905, and 287 thousand tonnes produced (ICCO, 2020). Even if Africa remains the leading producer, America maintains its reputation thanks to the aromatic quality of its cocoa. Cocoa is classified into two types of products: bulk cocoa and fine flavour cocoa. Fine flavour cocoa is characterised by fruity and floral notes unlike bulk cocoa (Sukha et al., 2008). Bulk cocoa accounts for around 95% of world production compared to 5% for fine flavour cocoa. *Theobroma* cacao L. is highly diverse and has been classified into 10 genetic groups: Amelonado, Contamana, Criollo, Curaray, Guiana, Iquitos, Marañón, Nanay, Nacional, and Purùs (Motamayor et al., 2008).

Nowadays, three varieties are mainly capable to produce fine flavour cocoa: Criollo, Nacional, and Trinitario (hybrids between Criollo and Amelonado). Criollo is not widely cultivated because of its high susceptibility to diseases and low vigour (Cheesman, 1944). Nacional is native to Ecuador and is well-known for its Arriba floral flavour. It is for this reason that it is sought after by chocolate makers. It is characterised by floral and woody notes (Luna et al., 2002). Also, Nacional is known for its low astringency and bitterness (International Cocoa Organization, 2017). The first hypothesis explaining floral notes of Arriba flavour was suggested by Ziegleder (1990) who observed that linalool, a volatile compound (VOC) belonging to monoterpenes, was observed in higher concentration in Nacional cocoa.

Overall, fine flavours are often produced during the fermentation process (Rodriguez-Campos et al., 2011). The cocoa fermentation takes place in two stages: first, the alcoholic fermentation made by yeast thanks to the presence of sugar in the cocoa pulp, then, there is an acetic fermentation carried out by bacteria (Ho et al., 2014). Fermentations produce aroma precursors but also VOCs. An adaptation of fermentation conditions is required to improve cocoa beans fine flavour. Fermentation time has an important effect on the concentration of different VOCs, as for some alcohol concentrations, which decreases from 2 to 8 days of fermentation (Rodriguez-Campos et al., 2012; Hamdouche et al., 2019). The drying process occurs after fermentation, which allows stopping it. This step is very important for cocoa bean conservation. It allows moisture decrease from 80 to under 8% (Cros and Jeanjean, 1995; Afoakwa et al., 2008). The artificial drying temperature can also influence the aromatic fraction with a decrease in isobutyric acid and an increase in tri and tetramethylpyrazine at lower drying temperature (70 vs. 80°C) (Rodriguez-Campos et al., 2012).

Cocoa beans have been studied to understand how their specific flavour is synthesised. A study on unfermented dry cocoa beans showed that terpenes are already present and important for fruity and floral aromas, even without fermentation (Qin et al., 2017). Other scientists have also proven the importance of terpenes such as linalool or epoxylinalool in cocoa fine flavour after fermentation (Kadow et al., 2013; Cevallos-Cevallos et al., 2018). Kadow et al. (2013) demonstrated that the aroma specificity depends on the presence of VOCs and can be different depending on the genotype. The most important VOCs for the floral aroma of cocoa have been identified: they include terpenes mainly linalool, 2-phenylethanol (or phenylethyl alcohol), 2-phenylethyl acetate, and acetophenone (Ziegleder, 1990, 2009; Afoakwa et al., 2008; Kadow et al., 2013; Cevallos-Cevallos et al., 2018; Utrilla-Vázquez et al., 2020). Rottiers et al. (2019) also compared the compounds contained in cocoa beans from the modern Nacional (EET varieties) and a standard cocoa variety CCN51. They were able to identify 14 compounds known to have a floral taste by GC-MS. Only five of them were found during the analysis with an electronic nose: 2-phenylacetaldehyde, 2-phenylethyl acetate, 2-phenylethanol, acetophenone, and linalool. However, other VOCs could be responsible for floral aroma (Schwab et al., 2008).

The biosynthesis pathway of aromatic compounds has been studied. Linalool is a volatile floral compound present in various flowers as *Clarkia breweri*, rose, Chinese jasmine green tea, and wine (Dudareva et al., 1996; Ito et al., 2002; Genovese et al., 2007; Feng et al., 2014). Its biosynthesis pathway is very well-studied. Pichersky et al. (1994) highlighted the linalool biosynthesis pathway in *C. breweri* flowers. They observed the transformation of geranyl pyrophosphate (GPP) to linalool by linalool synthase (LIS). Subsequently, the linalool was transformed into 6,7-epoxylinalool. The 6,7-epoxylialool was then converted to pyranoid linalool oxide or furanoid linalool oxide. Other studies showed that cytochrome P450 is responsible for the transformation of linalool to 6,7-epoxylinalool and cyclases for the transformation of 6,7-epoxylinalool to pyranoid linalool oxide or furanoid linalool oxide (Kreck et al., 2003; Meesters et al., 2007; Chen et al., 2010).

2-phenylethanol (or phenylethyl alcohol) has been found in muscadine grape juice, wine, and roses (Baek et al., 1997; Helsper et al., 1998; Genovese et al., 2007). 2-phenetylethanol and 2-phenylethyl acetate were observed in the same biosynthesis pathway in roses (Roccia et al., 2019). L-phenylalanine is converted to 2-phenylacetaldehyde by phenyl acetaldehyde synthase (PAAS). Subsequently, 2-phenylacetaldehyde is reduced to 2-phenylethanol by phenyl acetaldehyde reductase (PAR). Next, 2-phenylethanol is acetylated to 2-phenylethyl acetate by acetyl-coenzyme a: geraniol/citronellol acetyl transferase (AAT) (Roccia et al., 2019).

Acetophenone has been found in muscadine grape juice and Camellia (Baek et al., 1997; Dong et al., 2012). It has the same precursor as 2-phenylethanol but has a parallel biosynthesis pathway identified in the fungus *Bjerkandera adusta*. The transformation of L-phenylalanine to 2-phenylethanol is due to the non-oxidative degradation pathway of L-phenylalanine, while L-phenylalanine transformation to acetophenone belongs to β-oxidation pathway (Lapadatescu et al., 2000). In Camellia, the acetophenone biosynthesis pathway has been characterised (Dong et al., 2012). First, L-phenylalanine (L-phe) is converted to cinnamic acid (CA). Next, CA is transformed into 3-hydroxy-3-phenylpropionic acid (HPPA). 3-Hydroxy-3-phenylpropionic acid is converted to 3-phenylpropionic acid (PPA) and PPA is transformed into acetophenone. The enzymes involved in these reactions have not yet been identified.

Few studies were carried out on the genetic determinants of cocoa qualities. The first were based on QTL analyses of some sensory traits and fat content (Lanaud et al., 2003) and also showed hotspots of VOCs co-located on the genome (Lanaud et al., 2012).

This study aims to contribute to the deciphering of the genetic and biochemical determinism of Nacional cocoa floral notes. To this end, we conducted a genome-wide association study (GWAS) on a modern cultivated Nacional population, composed of trees resulting from hybridizations between three contrasting main ancestors: Criollo, Amelonado, and the ancestral Nacional variety. This population was characterised by VOCs and sensory analyses and presented a high degree of variability. Thanks to the availability of the genome sequence and high-density SNP genotyping, candidate genes involved in key traits could be proposed.

MATERIALS AND METHODS

Vegetal Material

The plant material used for these experiments was composed of a collection of 152 cocoa trees from Ecuador conserved in the Pichilingue experimental station of the "Instituto Nacional de Investigaciones Agropecurias" (INIAP) and the "Colecion de Cacao de Aroma Tenguel" (CCAT) of Tenguel. This population represents the Nacional variety currently grown in Ecuador and has been described by Loor (2007).

Fermentation Processes

Micro-fermentations of cocoa beans were carried out in a wooden box in the most homogeneous way possible with a homogeneous cocoa mass. The process lasted 4 days with two turns at 24 and 72 h after the beginning of the fermentation. Each clone sample (152) was placed in a protective laundry bag and micro-fermented in a cocoa mass. After fermentation, the samples were put in a dry place. They were considered dried when their moisture content was below 8%.

Sensorial Analysis

One hundred and forty-six individuals were characterised by sensory analyses based on blind tastings carried out on three repetitions per sample. The tastings were carried out on cocoa liquor. The cocoa liquor corresponds to merchant cocoa (dried fermented beans) which have been roasted and crushed. Sixteen floral notes were judged with a score ranging from zero (no floral notes detected) to 10. We used the average of the three replicates for the phenotype of the GWAS analysis (ISCQF, 2020).

This study was managed by Mr. Edward Seguine, whose work consists of conducting sensory analyses of chocolate samples (see attached documents). This study does not require the approval of an ethics committee.

Volatile Compound Analysis by GC-MS

Preparation of Cocoa Samples

The analysis of VOCs was carried out on dried fermented beans and roasted beans. For each sample, 50 g of beans were taken. The beans were deshelled and crushed to obtain nibs. Then, nibs were put in liquid nitrogen and ground with a blender (SEB, France), to obtain cocoa powder, which was stored at −80°C until analysis. In a 10 ml vial, 2.85 g of powder, 1 ml of standard internal solution (butan-1-ol at a concentration of about 600 μg/ml), and 2 ml of distilled water were added.

Compounds Extraction

The VOCs of cocoa samples were extracted using the technique of solid-phase microextraction in the headspace (SPME-HS) using a 50/30-μm divinylbenzene/carboxene/polydimethylsiloxane (DVB/CAR/PDMS) fibre provided by Supelco to extract volatiles. The fibre was previously conditioned at 250°C for 3 min and then exposed to the sample headspace at 50°C for 45 min. Extracted aroma VOCs were analysed using an Agilent 6890 N gas chromatography–mass spectrometer (GC–MS) equipped with a Hewlett Packard capillary column DBWAX, 30 m length × 0.25 mm internal diameter × 0.25 μm film thickness (Palo Alto, CA, USA). The GC oven temperature was initially set at 40°C for 5 min, increased to 140°C at a rate of 2°C/min and then increased at a rate of 10 to 250°C for 66 min. The carrier gas was high-purity helium at 1 ml min^{-1}. Injection mode was split less at 250°C for 2 min. The selective mass detector was a quadrupole (Hewlett Packard, Model 5973), with an electronic impact ionisation system at 70 eV and at 230°C (Assi-Clair et al., 2019).

Compounds Identification

The identification was done by comparing the mass spectra with the commercial NIST Wiley 275L database. No deconvolution was applied. Co-eluted VOCs were excluded from this study, with the exception of cis-ocimene co-eluted with ethyl hexanoate (cis-ocimene + ethyl hexanoate) which showed interesting results.

DNA Extraction Protocol

DNA extraction was conducted according to Risterucci et al. (2000) protocol.

Genotyping by SSR

This population was genotyped using SSR markers by Loor (2007). SSR loci were scored individually and alleles were recorded by the presence of polymorphic DNA fragments (alleles) among the individuals of each population. Only those alleles that showed consistent amplification were used in the analysis of results and smeared or weak bands were ignored.

Genotyping by Sequencing

DNA samples were genotyped by sequencing (GBS) using DArTseq (Diversity Arrays Technology Sequencing) technology (Kilian et al., 2012). This method is based on enzymatic restriction of coding regions of the genome by the restriction enzymes: Pst1 and Mse1. The restriction generated many short fragments, with each locus represented more than 10 times. Then, illumina Hiseq2000 machine sequenced all the fragments and the result was analysed. Reads were aligned with the V2 sequence of the Criollo genome (Argout et al., 2017). Reads that have more than one location were discarded. Markers with unknown locations were discarded for analysis. All the markers used are available on http://tropgenedb.cirad.fr/tropgene/JSP/interface.

jsp?module=COCOAinthegenotypessectionandtheCocoa-Nacional-aromasub-section.

Population's Structure Analysis

The phylogenetic tree was generated using DARwin software (Perrier and Jacquemoud-Collet, 2006). The genetic distances were calculated using the Dice coefficient and the Neighbour-Joining method (Dice, 1945; Saitou and Nei, 1987).

Association Mapping

The graphic representation of the markers along the 10 chromosomes was made with the R package "CMplot" (Yin, 2020). Several analyses of associations with SNP or SSR markers have been performed:

SNP GWAS

First, we performed a GWAS analysis with SNP markers associated with biochemical (146 accessions × 5,195 markers) and sensory (144 accessions × 5,195 markers) traits using TASSEL v5.

For all the traits, we used a mixed model (MLM) on the one hand. The MLM was carried out with a structure matrix, determined by running a principal component analyses (PCA integrated with TASSEL v5 software), considered as a fixed effect, and also with a kinship matrix considered as a random effect as covariates to control the false-positive rate. The option of not compressing and re-evaluating the components of variance for each marker was chosen. The kinship matrix using the identity by state (IBS) pairwise method proposed by Tassel v5 was established.

On the other hand, we used a fixed-effect model (GLM) with a structure matrix, determined by running a PCA. The option of 500 permutations was chosen.

For both methods, quantile-quantile plots were used to graphically evaluate the false-positive numbers observed in the selected model, based on deviations from the uniform law. The threshold was determined using the Bonferonni correction formula as proposed by Gao et al. (2008) with the effective number of independent tests (Meff) used as the denominator and calculated by SimpleM R package (Gao et al., 2010). Meff was 2,796, which corresponds to a *P*-value of approximately $1.79e^{-05}$. The significance of all markers was plotted using Manhattan plots with the R QQman package.

SSR GWAS

We performed an analysis with SSR markers associated with biochemical (180 accessions × 180 markers) and sensory (197 accessions × 180 markers) traits using TASSEL v3. We used a fixed-effect model (GLM) with a structure matrix; the option of 500 permutations was chosen. The threshold was determined using the Bonferonni correction corresponding to a p-value about $2.78e^{-04}$.

The borders of the association zones were calculated using Haploview (Barrett et al., 2005). The haplotypic blocks were calculated with SNP data using Haploview with the association test, Family trio data, Standard TDT, and ignore pairwise comparison of markers above to 10,000 kb calculation parameters. The haplotypic block information was used to determine the confidence intervals of association areas.

The physical maps with the QTL representation were created using SpiderMap v1.7.1 software (Rami, 2007 unpublished). The size of the dots is correlated to the R2.

The identification of candidate genes was performed using the *Theobroma cacao* genome sequence (Argout et al., 2017).

Statistical Analysis

Principal component analyses analysis and visualisation were made with the "mixOmics" R package. Calculation of correlation was made with "agricolae" R package and visualisation of correlation matrix with "corrplot" R package.

RESULTS

Genetic Diversity and Population Structure

The population studied represents the modern population of the Nacional variety cultivated in Ecuador. It is the result of various crosses between three main ancestors: the Criollo, the Amelonado, and the ancient Nacional varieties (Loor, 2007). Using SNP markers, the structure of the genetic diversity of the population was studied. There was a continuous distribution of population trees between the three ancestors (Criollo, Nacional, and Amelonado varieties) as shown in **Figure 1**. Loor (2007) had also shown this distribution using microsatellite markers.

Characterisation of the Studied Traits

To identify the areas of *T. cacao* genome involved in the synthesis of typical Nacional floral aromas, a GWAS was conducted with two types of traits: the VOCs present in cocoa beans (before and after roasting) and sensory analysis data.

Sensorial Traits Analysis

Sixteen floral notes were determined by sensory analyses performed on cocoa liquor. A total of 16 sensorial traits were therefore used for this study (**Supplementary Table 1**).

Principal component analysis for sensory traits showed continuous variation in the population (**Supplementary Figure 1**). Axis 1 is mainly defined by the aromatic notes: browned flavour, floral bark woody and smoky. Axis 2 is mainly defined by the aromatic notes: floral tobacco, fruity acidity, and astringency. Correlation analyses between sensory traits showed strong positive and negative correlations (**Figure 2A**). These strong correlations suggest either that the correlated sensory notes are produced by the same compounds or that an interaction exists between the perceptions of the two sensory traits.

Analysis of Aroma Volatile Compounds

The biochemical characterisation was done on unroasted and roasted beans. Among 160 VOCs identified, 26 VOCs are known to have a floral taste or are involved in biosynthetic pathways of known floral compounds (**Table 1**). Eighteen of them were detected in unroasted beans and 17 in roasted beans such as

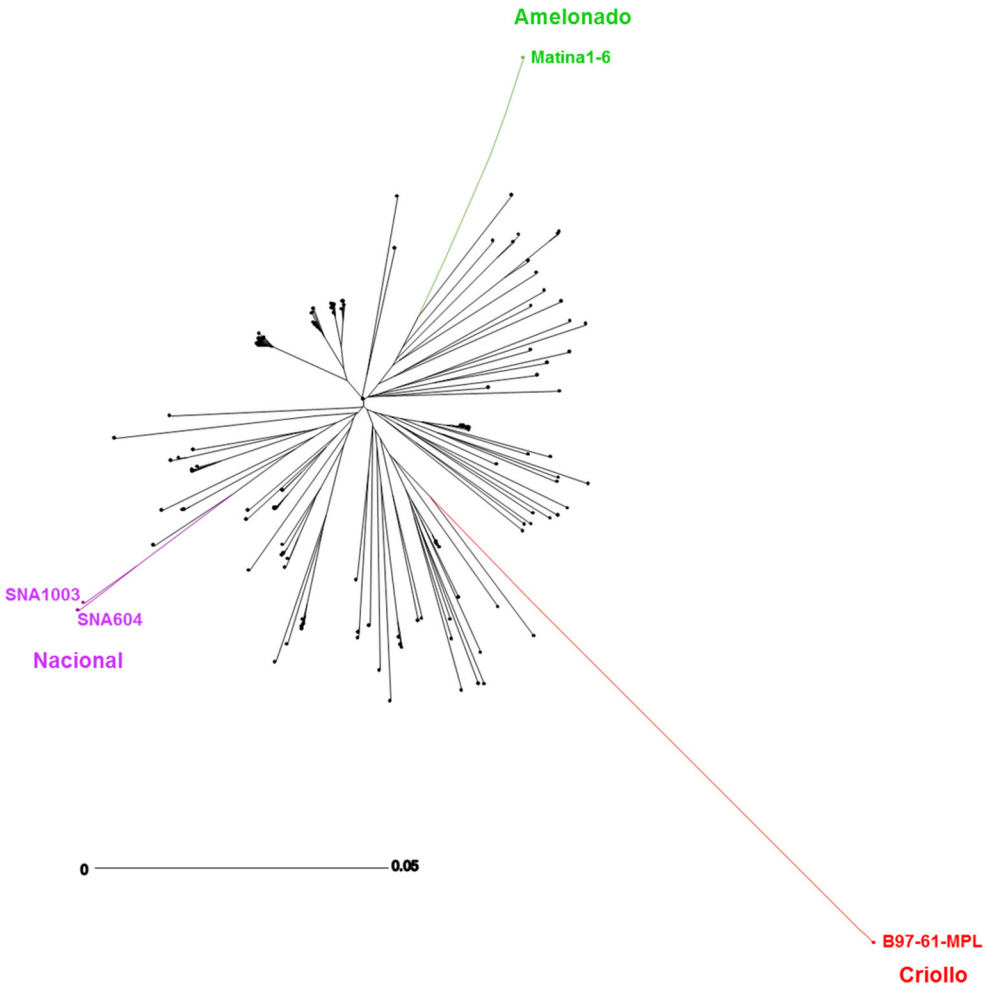

FIGURE 1 | Phylogenetic tree representing the modern Nacional population and its ancestors. Phylogenetic tree of the individuals of the studied population made with 4,130 SNPs and including the ancestor controls of the population: in red, the Criollo variety (B97-61-B2); in purple, the Nacional variety (SNA604, SNA1003); in green, the Amelonado variety (Matina 1–6); in black, the individuals of the studied population. The graph's scale represents the edge lengths which are proportional to the genetic distance.

linalool, acetophenone, or 2-phenylthanol. These VOC were used to conduct a GWAS analysis (**Table 1**).

Principal component analyses of aroma VOCs was made (**Supplementary Figures 2, 3**). Axis 1 of the PCA from analyses of biochemical compounds in unroasted beans is mainly defined by the linalool trans furanic oxide, meso-2,3-butan-di-yl diacetate, and linalool trans pyranic oxide. Axis 2 is mainly defined by ethyl acetate, ethyl-(2-methyl)-propionate, and benzaldehyde. Axis 1 of the PCA from analyses of biochemical compounds in roasted beans is mainly defined by epoxylinalool, 2-acetylpyrrole, and ethylphenyl acetate. Axis 2 is mainly defined by pentan-2-ol, pentan-2-one, and 1.2.5-trimethylbenzene. As with sensory traits, PCA of aroma VOCs showed that the distribution of traits showed a continuous variation within the population which can be explained by the great genetic diversity present in this group of individuals deriving from several generations of crosses.

Correlation analyses between the different traits showed positive correlations between several biochemical compounds in roasted and unroasted beans (**Figure 2B**). The highest correlations (>0.8) were observed in unroasted beans: between benzyl acetate and acetophenone; between 2-phenylethyl acetate and 2-pentylfuran co-eluted with ocimene; between guaiacol and trans furanic oxide linalool; between trans furanic oxide linalool and linalool. High correlations were also observed in roasted beans: between 1-phenylethyl acetate and epoxylinalool; between ethylphenyl acetate and guaiacol. A negative correlation between -0.4 and -0.6 was observed between linalool cis pyranic oxide and 2-pentylfuran co-eluted with ocimene in unroasted beans.

These various correlations between compounds can be partly explained by the fact that they belong to the same biosynthesis pathway. This is the case for the different terpenes which are strongly correlated or compounds resulting from the degradation of L-phenylalanine (acetophenone,

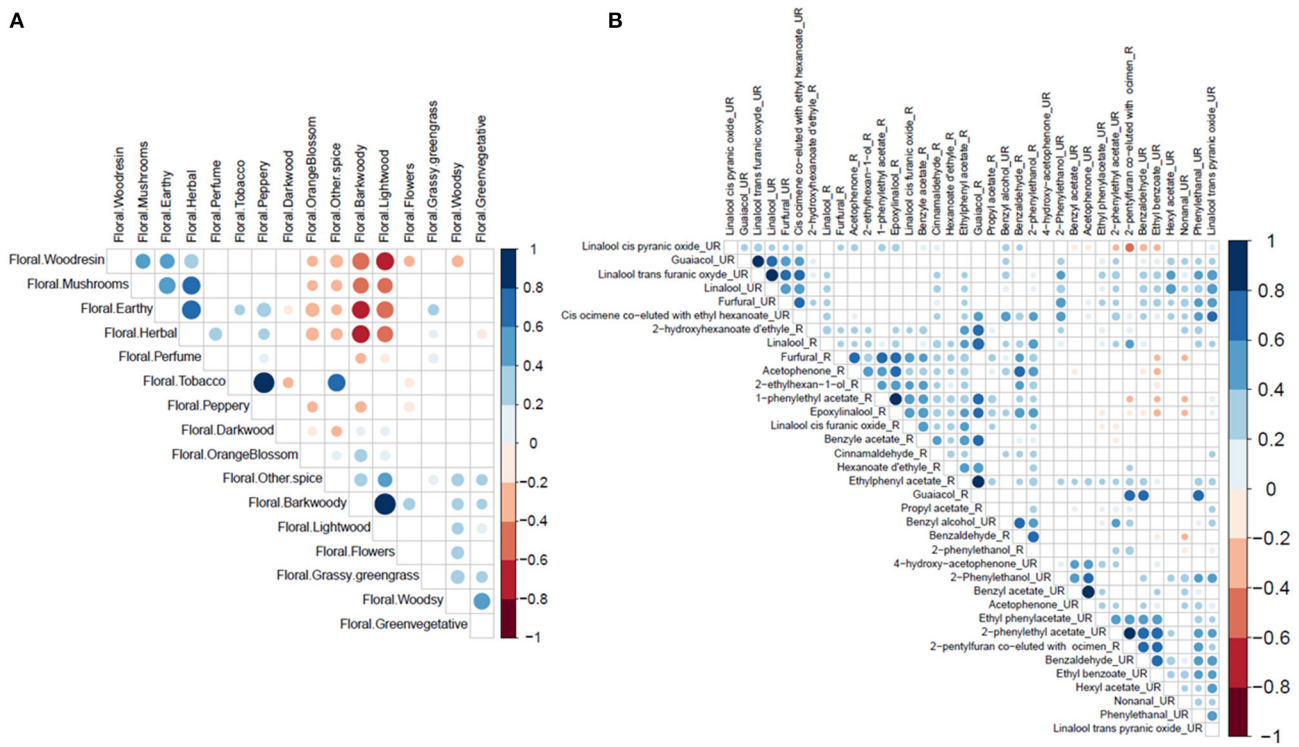

FIGURE 2 | Significant correlation matrix. **(A)** Correlation matrix between the sensorial profiles determined in cocoa liquor. **(B)** Correlation matrix between the biochemical compounds measured in unroasted (UR) and roasted beans (R). The correlations were calculated by the Pearson method. The white boxes represent no significant correlations. The colour of the circles corresponds to Pearson's correlation coefficient. The areas of circles correspond to a *p*-value of correlation coefficients. The *p*-value threshold for a significant correlation is 0.05. The different shades of blue represent a positive correlation coefficient while the different shades of red represent a negative correlation coefficient. The intensity of the colour depends on the strength of the R2 correlation coefficient. The scale on the right indicates the interpretations of different colours.

2-phenylethanol, and benzaldehyde). On the other hand, no strong correlation between biochemical and sensory traits was detected (**Supplementary Figure 4**).

Genome-Wide Association Study

The linkage disequilibrium observed in this population amounts to 15 cM (Loor, 2007). Genome-wide association study analyses were performed by different methods (GLM and MLM) and with different types of markers (SSR and SNP).

Marker Sorting

To limit the biases due to rare alleles, sorting by the frequencies of the minor alleles (MAF) was done at 5% (MAF5). The population being very heterozygous, the sorting by MAF allowed to eliminate the alleles with a total frequency lower than 5% but left homozygous genotypes very poorly represented (one individual per class). The hypothesis was that the low representation of genotypic classes could induce a bias in the analyses, in the same way as a minor allele. It was therefore undertaken to do a further sorting of markers by discarding markers for which genotype classes had <5% representation of the total population (Minor genotype frequencies, MGF). We conserved markers that had at least seven individuals per genotype class (G7). Several tests were performed such as the comparison of Q–Q plots or the comparison of *p*-values (Zhang et al., 2019) to determine which of the two sorting methods had the least bias (**Supplementary Figure 5**). None of the tests could determine which of the two was the most biassed. The results differed in some respects, so both marker sorting methods were retained for the GWAS studies.

SNP Marker Distribution

For the GWAS, SNPs were selected without missing data and with a genotype frequency above 5% or a MAF above 5%. The final data set consisted of 5,195 SNP markers for the G7 data set and 6,541 SNP markers for the MAF5 data set (Ruiz et al., 2017). The SNP markers are well spread over all 10 chromosomes of *T. cocoa*. However, a decrease in marker density is observed in the centromeric and peri-centromeric areas (**Figure 3**).

Determination of Confidence Intervals of Associations Based on Haplotypes

Haplotypes were calculated based on the known linkage disequilibrium of the population which is 15 cM, corresponding to 10,000 kb. A total of 681 haplotypic blocks were thus determined with a minimum of 42 haplotypic blocks present on chromosome 8 and a maximum of 96 haplotypic blocks present on chromosome 1. Confidence intervals were defined based on these haplotypic blocks. In this paper, each association zone,

TABLE 1 | List of biochemical compounds related to floral traits used for the GWAS analysis of unroasted (UR) and roasted (R) beans.

UR	R	Aroma volatile compounds	Compound family	Aroma
	X	1-Phenylethyl acetate	Ester	Fruity (Garg et al., 2018)
	X	2-Ethylhexan-1-ol	Alcohol	Leafy, rose (Garg et al., 2018)
X	X	2-Phenylethanol	Alcohol	Rose, honey (Jezussek et al., 2002; Genovese et al., 2007)
X		2-Phenylethyl acetate	Ester	Floral, underwood, rose (Guichard et al., 2003; Genovese et al., 2007; Wang et al., 2014)
X		4-Hydroxy-acetophenone	Ketone	
X	X	Acetophenone	Ketone	Acacia honey, floral, and fruity (Genovese et al., 2007; Wang et al., 2014)
X	X	Benzaldehyde	Aldehyde	Bitter, cherry, almond, fruity (Perestrelo et al., 2006; Pham et al., 2008; Wang et al., 2014)
X	X	Benzyl acetate	Ester	Floral, Jasmin (Ito et al., 2002)
X		Benzyl alcohol	Alcohol	Fruity (Ito et al., 2002)
	X	Cinnamaldehyde	Aldehyde	Spicy, cinnamon (Garg et al., 2018)
	X	Epoxylinalool	Terpene	Floral (Arn and Acree, 1998)
	X	Ethyl 2-hydroxyhexanoate	Ester	Floral (Wang et al., 2014)
X		Ethyl benzoate	Ester	Fruity, violet, candy (Ferreira et al., 1997)
	X	Ethyl dodecanoate	Ester	Floral, fruity, leafy (Garg et al., 2018)
	X	Ethyl hexanoate	Ester	Fruity (strawberry, green apple) and floral (Larsen and Poll, 1992; Ferreira et al., 1997; Genovese et al., 2007; Wang et al., 2014)
X	X	Ethylphenyl acetate	Ester	Rose, floral (Perestrelo et al., 2006)
X	X	Furfural	Furan	Incense, fruity, floral, toasted, sweet, and almond (Ferreira et al., 1997; Colahan-Sederstrom and Peterson, 2005; Wang et al., 2014)
X	X	Guaiacol	Aromatic hydrocarbon	Phenolic, floral, smoky, sweet, medicament (Ferreira et al., 1997; Arn and Acree, 1998; Genovese et al., 2007)
X		Hexyl acetate	Ester	Fruity (pear) and floral (Guichard et al., 2003; Wang et al., 2014)
X	X	Linalool	Terpene	Floral, citrus peel, orange flower (Ferreira et al., 1997; Genovese et al., 2007)
	X	Linalool cis furanic oxide	Terpene	Floral, woody (Arn and Acree, 1998)
X		Linalool cis pyranic oxide	Terpene	Fruity, citrus, green (Arn and Acree, 1998; Ito et al., 2002)
X		Linalool trans furanic oxide	Terpene	Citrus, leafy, floral (Arn and Acree, 1998; Ito et al., 2002)
X		Linalool trans pyranic oxide	Terpene	
X		Nonanal	Aldehyde	Orange-like, floral, soapy (Kumazawa and Masuda, 2002; Mahajan et al., 2004; Karagül-Yüceer et al., 2006)
X		Phenylethanal	Aldehyde	Floral, rose, honey (Perestrelo et al., 2006)
	X	Propyl acetate	Ester	Celery, floral, pear, red fruit (Garg et al., 2018)

UR, unroasted beans; R, roasted beans.
X, biochemical compounds find in unroasted or roasted beans.

thus corresponding to a haplotypic block, is represented by its association peak. The association peak corresponds to the marker for which the association is the most significant.

Comparison of the Four Different Methods Used for SNP Association Studies

The GLM method has made it possible to highlight more areas of association than the MLM method. In both cases, the use of the set of markers sorted according to a 5% MAF (MAF5) also made it possible to highlight more association zones: 333 against 295 for the GLM method and 152 against 94 for the MLM method. The MLM method, therefore, appears to be more stringent.

Some areas of the association are common for different methods. For example, in the case of terpene relatives' traits, 63 co-locations between positive associations for different methods for the same trait was found on all chromosomes except chromosome 4 and 8. A co-localisation between GLM_MAF5 and GLM_G7 methods for linalool cis pyranic oxide (UR) was observed on chromosome 2 as shown in **Figure 4A**. In the case of L-phenylalanine relatives' traits, co-locations of the association zones between the different methods for the same trait was observed on all chromosomes. This is the case for example on chromosome 5 where co-localisation of associations for GLM_MAF5, MLM_G7, and MLM_MAF5 linked to 4-hydroxyacetophenone (UR) was observed (**Supplementary Figure 6**).

Identification of Significant Associations for Sensorial Traits

Among all the associations, only 38 are related to the sensory data with floral notes. Out of a total of 16 floral perceptions, significant associations were detected for 11 of them, on all chromosomes

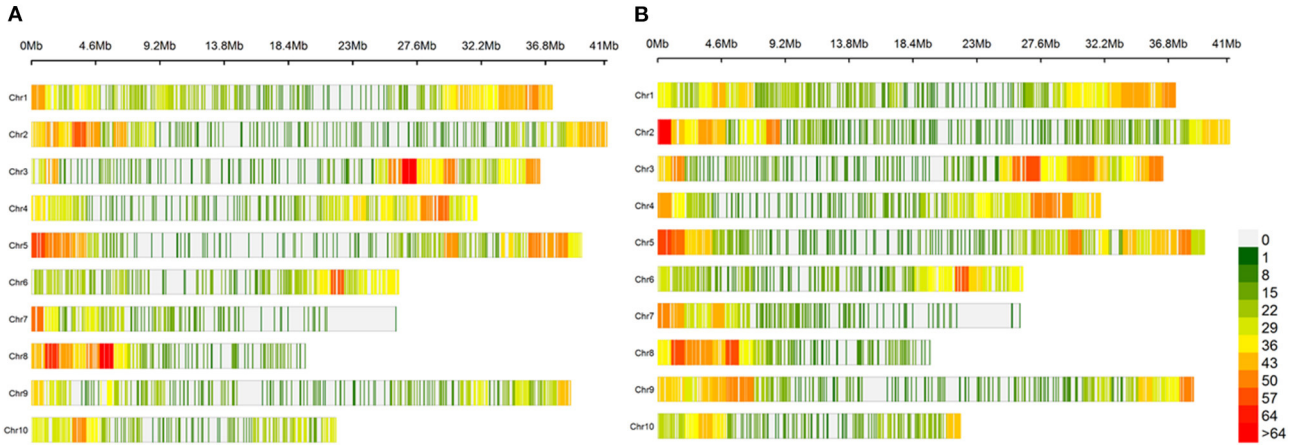

FIGURE 3 | Distribution of markers along the 10 chromosomes of *T. cacao*. **(A)** Distribution of markers from the G7 dataset along the 10 chromosomes of *T. cacao*. The graph shows the distribution of markers along the 10 chromosomes. The density is calculated on a 1 Mb window. The areas without markers are shown in grey. The weakly marked areas are in green and the strongly marked areas are in red. A colour gradient between green and red represents the marking gradient. **(B)** Distribution of markers from the MAF5 dataset along the 10 chromosomes of *T. cacao*.

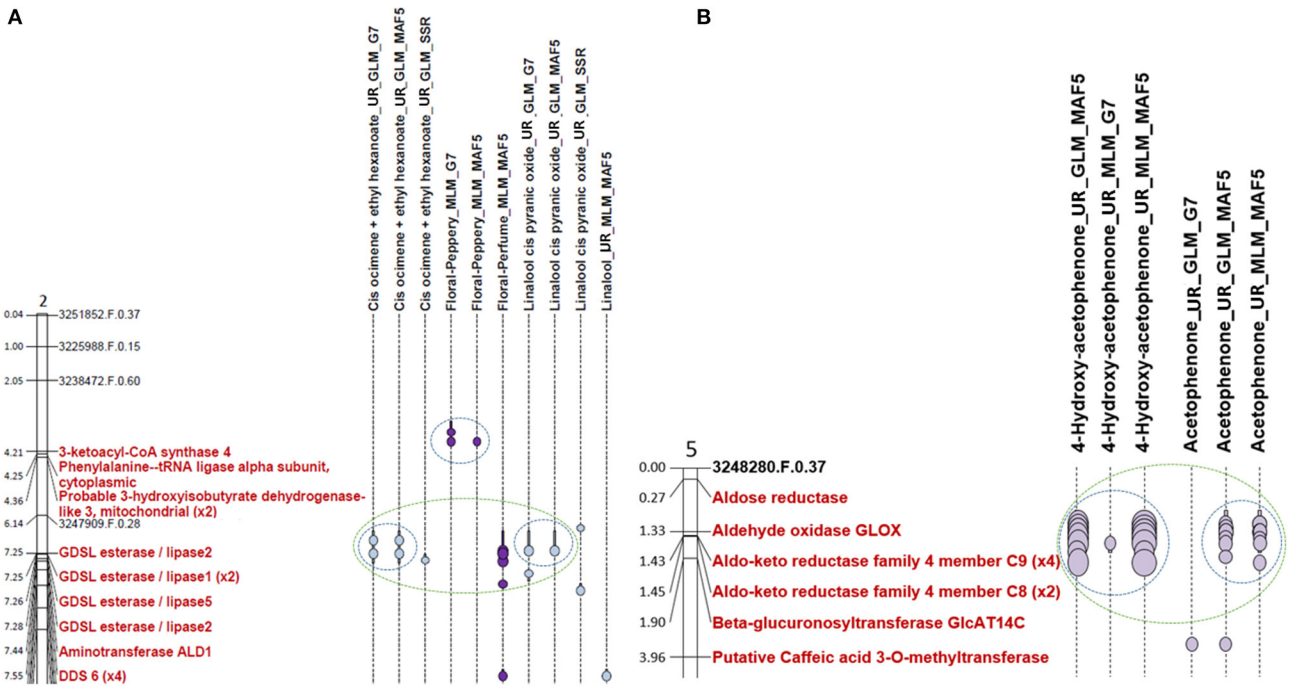

FIGURE 4 | Extract of chromosome 2 map and chromosome 5 map. **(A)** Extract from the chromosome 2 map representing the associations detected for compounds involved in the monoterpene biosynthetic pathway. **(B)** Extract from the chromosome 5 map representing the associations detected for compounds involved in the L-phenylalanine degradation pathway. The light blue dots represent the peaks of associations in relation to traits whose beans have not been roasted. The dark purple dots represent the peaks of association in relation to traits whose beans have been roasted. The bars around these points correspond to the confidence intervals of the association zone. Co-locations are represented by a blue circle for the associations co-localised for the same trait and identified by different methods. Co-locations are represented with a green circle for the co-locations between different biochemical compounds. Candidate genes are written in red. One scale unit on the chromosome corresponds to 1 Mb.

except chromosome 5 and chromosome 7. Only one area of association was revealed for each of the six floral notes: the floral notes bark woody, dark wood, mushrooms, orange blossom, other spice, and tobacco (**Table 2**). Four association zones were also detected for the floral note Lightwood on chromosome 1. The area of strongest association detected for the light wood

floral note and the tobacco floral note is in the same haplotypic block. The floral note that allowed detecting the most areas of association is the floral perfume where 13 areas were highlighted. The variation in the floral perfume note is the one that seems to be the most explained by the genetic variation observed, with an explanation rate for variation in the trait of 24%.

Identification of Significant Associations for Aroma Volatile Compounds

The GWAS analyses brought to light 393 association zones. Some of them were detected with several VOCs. All the associations found can be consulted in the **Supplementary Table 2**.

Significant associations for 18 VOCs in unroasted beans and 17 volatile compounds in roasted beans were identified (**Supplementary Table 2**). No association zones were detected for five VOCs, four of which were assayed in roasted beans: ethylphenyl acetate (UR), ethyl 2-hydroxyhexanoate (R), ethyl hexanoate (R), guaiacol (R), and cis linalool oxide (R).

Two major pathways for the biosynthesis of compounds known to have a floral taste, among those compounds for which a significant association was detected, seem to be particularly represented: the monoterpene biosynthesis pathway and, the L-phenylalanine degradation pathway that allows the synthesis of, among others, acetophenone and 2-phenylethanol.

The results obtained were mapped to visualise the areas of significant associations, their locations, as well as possible co-locations between them. Two maps were made. A map with the results of significant associations related to the compounds involved in the terpene biosynthesis pathway and the floral traits from the sensorial evaluation. A second map includes the results of the significant associations of floral tastes and of compounds involved in the degradation pathway of L-phenylalanine which allows, the synthesis of acetophenone and 2-phenylethanol known to have a floral taste. Some results differ between the different methods (GLM and MLM) or the sorting of SNP markers (MAF5 or G7) or between the type of SNP and SSR markers. All results are shown on the maps in **Supplementary Figures 6, 7**. Results that are repeatable between methods appear to be the most conclusive.

Significant Associations Identified for the Biochemical Compounds Involved in Terpene Biosynthetic Pathway

Among the 27 compounds related to the floral note, six VOCs derived from the terpene biosynthesis pathway: linalool (UR and R), trans furanic oxide linalool (UR), cis pyranic oxide linalool (UR), epoxylinalool (R), and cis ocimene co-eluted with ethyl hexanoate (UR) (**Figure 5**). Eighteen zones of association were revealed for the linalool in unroasted beans (UR) against two zones for linalool in roasted beans (R). The most significant association of linalool (UR) was found on chromosome 7 while that of linalool (R) was found on chromosome 6. Twenty-nine association zones were highlighted for the linalool trans furanic oxide (UR). The most significant association linked to linalool trans furanic oxide (UR) was detected on chromosome 7 which is in the same haplotypic bloc of the most significant association of cis ocimene co-eluted with ethyl hexanoate (UR). Twenty-seven associations were observed for linalool cis pyranic oxide (UR). Finally, thirty-eight areas of associations were revealed for the epoxylinalool (R) (**Table 3**; **Supplementary Table 2**).

The map with the results for terpenes (**Supplementary Figure 7**) shows several interesting results. Among a large number of associations, several co-locations can be observed between different biochemical compounds involved in the terpene pathway. For example, a co-localisation between the Linalool (UR), the Linalool cis-pyranic oxide (UR), and the Linalool trans-furanic oxide (UR) was observed in chromosome 6 (**Supplementary Figure 7**). This suggests the greater likelihood that most of these compounds already known for their floral notes are well-involved in floral notes of Nacional cocoa.

Co-locations Between Biochemical Compounds

Sixteen co-locations between different biochemical compounds were also observed on chromosomes 2, 4, 5, 7, 9, and 10, for

TABLE 2 | Most significant association detected for each of the sensory floral traits.

CH	Position of the association peak (bp)	N° hap. bloc	Floral note detected	*p-value* of the strongest association	Explanation rate of the trait of the strongest association (%)	Total number of associations for the character
1	4,079,457	NA	Floral-other spice	1,45E-05	14	1
1	4,129,759	10	Floral-lightwood	2,56E-06	16	4
1	4,131,970	10	Floral-tobacco	4,52E-06	16	1
2	3,606,270	NA	Floral-peppery	9,42E-07	15	5
2	7,476,546	36	Floral-perfume	1,87E-09	24	13
6	21,137,437	45	Floral-wood resin	8,74E-07	15	6
6	26,160,073	NA	Floral-dark wood	5,78E-06	13	1
8	15,196,137	37	Floral-orange blossom	0,000132	12	1
9	38,188,583	58	Floral-mushrooms	1,57E-08	22	1
9	4,248,470	17	Floral-bark woody	6,56E-06	15	1
9	6,245,108	21	Floral-green vegetative	7,45E-09	23	4

CH, chromosome; hap, haplotypic; bp, base pair.

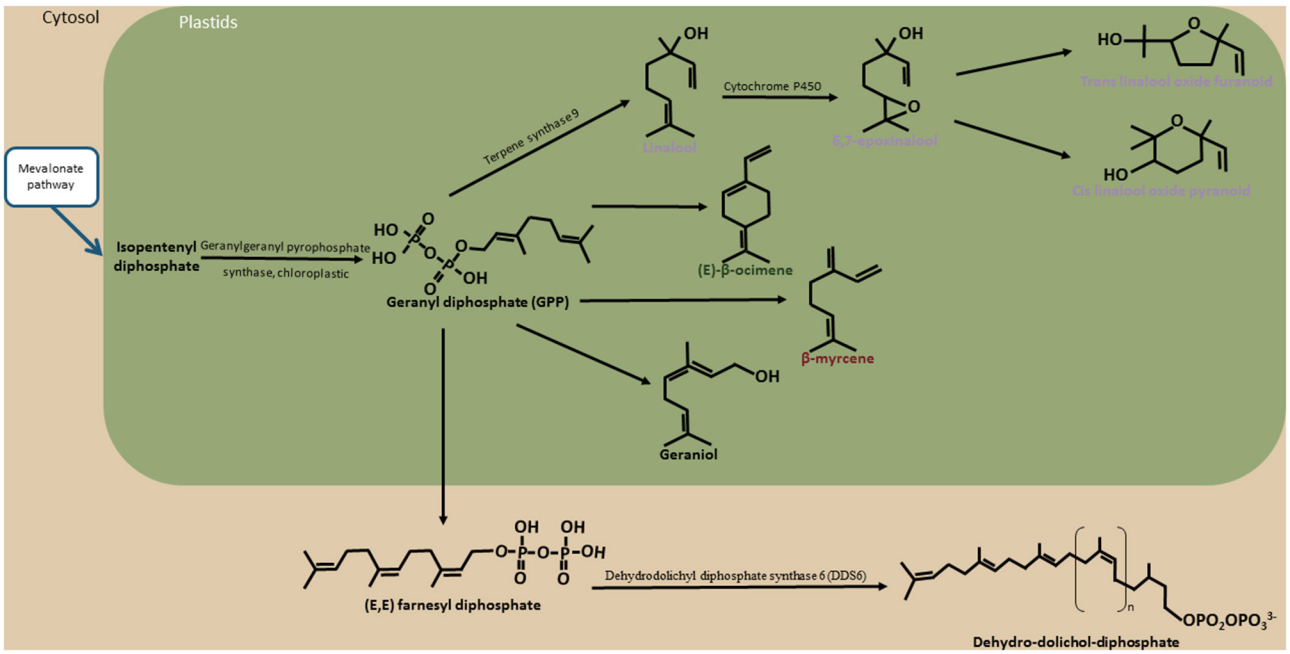

FIGURE 5 | Terpene biosynthesis pathway. The schema illustrates the different biosynthesis pathways of compounds belonging to the terpene biosynthesis pathway identified in cocoa. Compounds known to have a floral taste are noted in purple. The blue arrows represent the bridges between the terpene biosynthesis pathway and mevalonate pathways. The names of these other biosynthetic pathways are framed in blue. The black arrows represent the enzymatic actions. The names of the enzymes are indicated (when identified) around these arrows. In green represented the limit of the plastids. In light brown are represented the limits of cytosol.

TABLE 3 | Most significant associations for biochemical compounds related to terpene pathway.

CH	Position of the association peak (bp)	N° hap. bloc	Traits	*p-value* of the strongest association	Explanation rate of the trait of the strongest association (%)	Associations detected
1	6,448,063	23	Epoxylinalool_R *	3,23E-11	32	38
6	5,543,124	17	Linalool_ R *	5,57E-06	16	2
7	5,607,833	24	Linalool trans furanic oxyde_UR *	5,51E-10	29	29
7	5,607,833	24	Cis ocimene + ethyl hexanoate_UR	3,73E-12	35	42
7	10,459,413	31	Linalool_ UR *	3,53E-13	37	18
10	6,167,221	27	Linalool cis pyranic oxide_ UR *	3,27E-07	23	27

The most significant association detected for each compound is reported. CH, chromosome; hap., haplotypic; UR, unroasted beans; R, roasted beans.
**Biochemical compounds known for floral notes, bp, base pair.*

example on chromosome 2 between the linalool cis pyranic oxide (UR) and cis-ocimene co-eluted with ethyl hexanoate (**Figure 4A**). Various numbers of co-locations could be observed according to chromosomes. Only one co-location are observed on chromosome 9 and chromosome 10 and five co-locations were highlighted on chromosome 7 (**Supplementary Figure 7**). Co-localisations between association zones identified for different VOCs can be explained by their belonging to the same biosynthesis pathway such as for linalool trans furanic oxide (UR) and linalool (UR) on chromosome 3, or for cis pyranic oxide (UR) and epoxylinalool (R) on chromosome 4 (**Supplementary Figure 7**). It can then be thought that this zone of associations is due to the presence of a gene coding for an enzyme that is part of this biosynthetic pathway. To verify this hypothesis, we have begun to search for candidate genes at the level of the association zones.

Co-locations Between Biochemical Compounds and Sensorial Traits

Seven co-locations between at least one biochemical compound and a floral note were detected on chromosomes 1 and 2. On chromosome 1, two co-locations were observed between

epoxylinalool (R) and the floral note lightwood and one between epoxylinalool (R), floral notes lightwood and floral notes tobacco (**Supplementary Figure 7**). On chromosome 2, a co-localisation exists between cis ocimene co-eluted with ethyl hexanoate (UR), cis pyranic oxide linalool (UR), and floral scent (**Figure 4A**). A co-localisation is also observable between cis ocimene co-eluted with ethyl hexanoate (UR) and floral perfume. A co-localisation is also observable between linalool (UR) and the floral perfume note (**Supplementary Figure 7**).

Significant Associations Identified for the Biochemical Compounds Involved in the Degradation of L-Phenylalanine Pathway

Eighteen compounds for which significant associations have been identified appear to be involved in the degradation pathway of L-phenylalanine to either 2-phenylethanol or acetophenone (**Table 4**; **Figure 6**). Among these compounds for two of them, ethylphenyl acetate (R) and phenylethanal (UR), only one zone of the association was identified. The most significant association for phenylethanal (UR) co-localises with the strongest association detected for linalool (R) on chromosome 6. Thirty-six association zones were showed for acetophenone (UR) compared to 40 for acetophenone (R). The most significant association of acetophenone (UR) is on chromosome 2 while that of acetophenone (R) is on chromosome 6. Two hundred and six association zones were detected for cinnamaldehyde (R). Twelve zones of associations were revealed for 2-phenylethanol (UR) and three for 2-phenylethanol (R). The most significant association zones for 2-phenylethanol (UR) and (R) are located on chromosome 4 but at a different position. Two association zones were highlighted for ethyl benzoate (UR). Three areas of association were revealed for 2-phenylethyl acetate (UR). Two zones of associations were revealed for benzaldehyde (UR) against 72 with benzaldehyde (R). Benzaldehyde (UR) presents its most significant association on chromosome 7, while that of benzaldehyde (R) is located on chromosome 6. Thirty-eight association zones were revealed for benzyl acetate (UR) against two for benzyl acetate (R). Twenty-nine association zones were highlighted for 4-hydroxy acetophenone (UR). Seven regions of associations were revealed 2-ethylhexan-1-ol (R). Seventy-three association areas were highlighted for 1-phenylethyl acetate (R). The last two compounds involved in these biosynthetic pathways, benzyl acetate (R) and 1-phenylethyl acetate (R), have their most significant area of association co-locating and forming part of the same haplotypic block number 26 on chromosome 10. The variation of two biochemical compounds seems to be explained mainly by genetic variation. Indeed, the variation in the concentration of 4-hydroxy-acetophenone is explained at 79% by the strongest association zone as well as the variation in cinnamaldehyde which is explained at 65% by the association zone.

The map showing the results for compounds of the L-phenylalanine degradation pathway (**Supplementary Figure 6**) shows several interesting results.

TABLE 4 | Most significant associations for biochemical compounds related to L-phenylalanine degradation pathway.

CH	Position of the association peak (bp)	N° hap. bloc	Traits	*p-value of the strongest association*	Explanation rate of the trait of the strongest association (%)	Associations detected
10	5,308,832	26	1-Phenylethyl acetate_ R	1,17E-10	31	73
9	1,099,704	5	2-Ethylhexan-1-ol_ R *	2,61E-08	29	7
4	23,646,147	40	2-Phenylethanol_ R *	2,12E-06	16	3
4	17,349,904	NA	2-Phenylethanol_ UR *	9,89E-07	18	12
5	29,926,884	53	2-Phenylethyl acetate_ UR *	6,92E-06	18	3
7	2,815,797	11	4-Hydroxy-acetophenone_ UR	4,86E-43	79	29
6	23,097,785	58	Acetophenone_ R *	1,59E-08	25	40
2	8,389,914	39	Acetophenone_ UR *	2,53E-12	31	36
6	25,213,164	65	Benzaldehyde_ R	8,53E-09	26	72
7	10,459,413	31	Benzaldehyde_ UR	1,32E-05	17	2
7	2,092,063	9	Benzyl acetate_ UR *	2,36E-15	41	38
10	5,228,191	26	Benzyl acetate_ R *	8,15E-06	18	2
2	7,448,797	36	Benzylic alcohol_ UR	2,12E-08	25	42
3	31,503,427	58	Cinnamaldehyde_ R	1,39E-28	65	206
5	27,513,744	45	Ethyl benzoate_ UR *	2,33E-06	18	2
1	36,447,062	91	Ethylphenyl acetate_ R *	1,35E-05	17	1
6	5,543,124	17	Phenylethanal_ UR *	1,77E-05	14	1

The most significant association detected for each compound is reported. CH, chromosome; hap., haplotypic; UR, unroasted beans; R, roasted beans.
** Biochemical compounds known for floral notes, bp, base pair.*

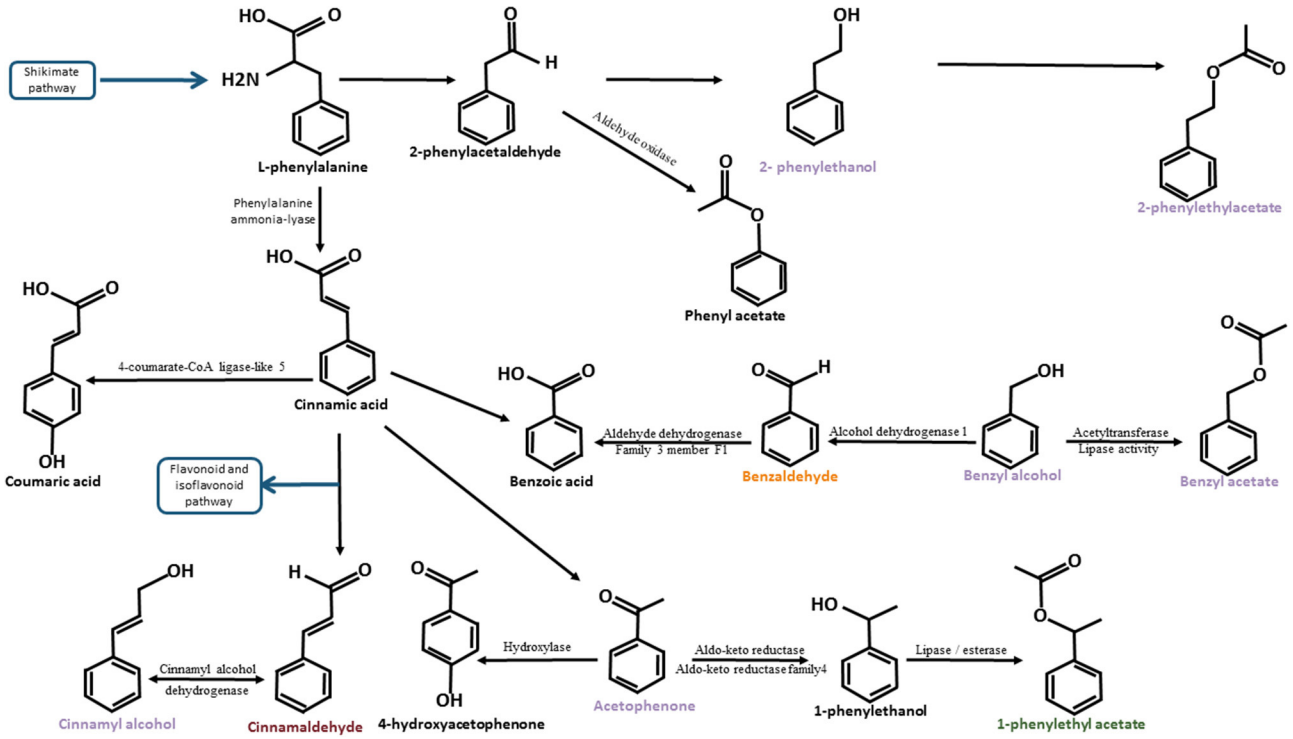

FIGURE 6 | Degradation pathway of L-phenylalanine adapted from Lapadatescu et al. (2000). The schema illustrates the different biosynthesis pathways of compounds belonging to the L-phenylalanine degradation pathway identified in cocoa. Compounds known to have a floral taste are noted in purple. Compounds known to have a fruity taste are noted in orange. Compounds known to have a spicy note are noted in dark red. The blue arrows represent the bridges between the L-phenylalanine degradation pathway and other biosynthetic pathways. The names of these other biosynthetic pathways are framed in blue. Black arrows represent the enzymatic actions. The names of the enzymes are indicated (when identified) around these arrows.

One hundred and eleven co-locations between different VOCs were also observed on all chromosomes. An example of co-localisation was observed between 4-hydroxyacetophenone (UR) and acetophenone (UR) on chromosome 5 (**Figure 4B**).

Thirteen co-locations between at least one aroma VOC and one sensory trait were observed on chromosomes 1, 2, 8, and 9 (**Supplementary Figure 6**).

Significant Associations Were Identified for the Biochemical Compounds Involved in Other Pathways

Several areas of association were highlighted for seven other compounds known also to have a floral taste: ethyl dodecanoate (R), guaiacol (UR and R), hexyl acetate (UR), furfural (UR and R), propyl acetate (R), and nonanal (UR). One hundred and seventeen association zones were detected for guaiacol (UR) against zero for guaiacol (R). Twelve association zones were observed for furfural (UR) compared to 30 for furfural (R) (**Table 5**). The variation in hexyl acetate concentration is very high compared to other compounds. On the other hand, the genetic explanation for the variation in the concentration of propyl acetate is very weak compared to the other characteristics of this study (4%).

Candidate Genes Potentially Involved in the Formation of the Floral Aroma

Of the 393 association zones exposed, 27 with candidate genes with predicted functions were identified.

Candidate Genes Linked to the Terpene Biosynthesis Pathway

Candidate genes related to the terpene biosynthetic pathway were found on chromosomes 1, 2, 5, 7, 9, and 10. The association zone number and candidate genes are reported in **Supplementary Figure 7**; **Supplementary Table 3**; and **Table 6**.

On chromosome 1, three association zones contain candidate genes. Association zone 1 (805,132–2,445,782 bp) linked to epoxylinalool (R) contains a gene coding for a "*Geranylgeranyl pyrophosphate synthase, chloroplastic*." This enzyme allows the synthesis of geranylgeranyl pyrophosphate in chloroplasts. This compound is a precursor of terpenes. As the monoterpene biosynthesis pathway is located in the plastids, the indication of chloroplastic synthesis seems to confirm the correspondence to another compound derived from linalool also synthesised in Chloroplast (Ying and Qingping, 2006; Feng et al., 2014). Association zone 2

TABLE 5 | Most significant associations for biochemical compounds related to other pathways.

CH	Position of the association peak (bp)	N° hap bloc	Traits	*p*-Value of the strongest association	Explanation rate of the trait of the strongest association (%)	Associations detected
3	31,273,182	57	Ethyl dodecanoate _ R *	6,60E-11	36	50
10	5,308,832	26	Furfural_ R *	3,15E-07	22	13
7	4,776,442	NA	Furfural_ UR *	3,70E-08	24	5
4	737,310	NA	Guaiacol_ UR *	1,60E-09	28	54
7	10,459,413	31	Hexyl acetate_ UR *	1,01E-138	67	94
9	37,557,289	54	Nonanal_ UR *	2,45E-07	22	9
9	3,653,985	14	Propyl acetate _ R *	2,42E-14	4	41

The most significant association detected for each compound is reported. CH, chromosome; hap, haplotypic; UR, unroasted beans; R, roasted beans.
**Biochemical compounds known for floral notes, bp, base pair.*

TABLE 6 | Candidate genes identified for terpene biosynthesis pathway.

N° asso	CH	Position of candidate gene (bp)	Position of the pic of association (bp)	Candidate Gene function	Trait in association
1	1	1,883,959	2,368,915	*Geranylgeranyl pyrophosphate synthase, chloroplastic*	Epoxylinalool-R
2	1	3,173,783	3,379,361	*Cytochrome P450 81E8*	Epoxylinalool (R), floral note lightwood
2	1	3,179,883	3,379,361	*Cytochrome P450 81E8*	Epoxylinalool (R), floral note lightwood
3	1	6,026,043	6,130,253	*Cytochrome P450 78A7*	Epoxilinalool (R)
4	2	7,549,475	7,324,500	*Dehydrodolichyl diphosphate synthase 6*	Cis ocimene co-eluted with ethyl hexanoate (UR), floral perfume
4	2	7,551,259	7,324,500	*Dehydrodolichyl diphosphate synthase 6*	Cis ocimene co-eluted with ethyl hexanoate (UR), floral perfume
4	2	7,553,795	7,324,500	*Dehydrodolichyl diphosphate synthase 6*	Cis ocimene co-eluted with ethyl hexanoate (UR), floral perfume
4	2	7,572,073	7,324,500	*Putative Dehydrodolichyl diphosphate synthase 6*	Cis ocimene co-eluted with ethyl hexanoate (UR), floral perfume
5	2	8,257,841	8,389,914	*Probable 3-hydroxyisobutyryl-CoA hydrolase 2*	Linalool cis pyranic oxide (UR)
6	5	32,749,861	33,303,465	*Cytochrome P450 89A2*	Linalool (UR) and linalool trans furanic oxide (UR)
6	5	33,057,632	33,303,465	*Cytochrome P450 89A9*	Linalool (UR) and linalool trans furanic oxide (UR)
6	5	33,064,477	33,303,465	*Cytochrome P450 89A2*	Linalool (UR) and linalool trans furanic oxide (UR)
6	5	33,073,996	33,303,465	*Cytochrome P450 89A2*	Linalool (UR) and linalool trans furanic oxide (UR)
6	5	33,094,200	33,303,465	*Cytochrome P450 89A2*	Linalool (UR) and linalool trans furanic oxide (UR)
6	5	33,099,009	33,303,465	*Cytochrome P450 89A2*	Linalool (UR) and linalool trans furanic oxide (UR)
7	7	6,346,577	6,181,185	*Putative Probable terpene synthase 9*	Linalool cis pyranic oxide (UR)
7	7	6,365,007	6,181,185	*Probable terpene synthase 9*	Linalool cis pyranic oxide (UR)
7	7	6,380,614	6,181,185	*Putative Probable terpene synthase 9*	Linalool cis pyranic oxide (UR)
8	9	791,968	749,365	*3-hydroxyisobutyryl-CoA hydrolase-like protein 2, mitochondrial*	Epoxylinalool (R)
9	10	6,317,543	6,167,221	*Probable terpene synthase 9*	Linalool cis pyranic oxide (UR)

asso., associations; CH, chromosome; R, roasted beans; UR, unroasted beans.

(3,083,032–3,398,183 bp) linked to epoxylinalool (R) and the floral note lightwood contains two candidate genes encoding a "*Cytochrome P450 81E8*." Cytochrome P450 has been identified to be responsible for the synthesis of epoxylinalool from linalool in kiwifruit (Chen et al., 2010). Association zone 3 (5,940,526–6,204,028 bp) linked to epoxylinalool (R) contains a candidate gene encoding a "*Cytochrome P450 78A7*."

On chromosome 2 (**Supplementary Figure 7**), two association zones contain candidate genes. Association zone 4 (7,324,500–7,617,242 bp) linked to cis ocimene co-eluted with ethyl hexanoate (UR) and floral perfume contains four genes encoding a "*Dehydrodolichyl diphosphate synthase 6*" (DDS 6) in **Figure 4A**. Dehydrodolichyl diphosphate synthase 6 allows the synthesis of dehydrodolichyl diphosphate, one of the precursors of which is geranyl diphosphate, the main precursor of the monoterpene biosynthesis pathway. The synthesis of dehydrodolichyl diphosphate could thus compete with the synthesis of cis-ocimene and explain the association with this compound as well as with the floral perfume, which is a taste attributed to several monoterpenes (linalool, epoxylinalool, ocimene). Association zone 5 (8,239,972–8,416,672 bp) linked to linalool cis pyranic oxide (UR) contains a gene encoding a "*Probable 3-hydroxyisobutyryl-CoA hydrolase 2*." The enzyme 3-hydroxyisobutyryl-CoA hydrolase 2 can enable the production of acetyl-CoA by releasing a CoA. Acetyl-CoA is a precursor

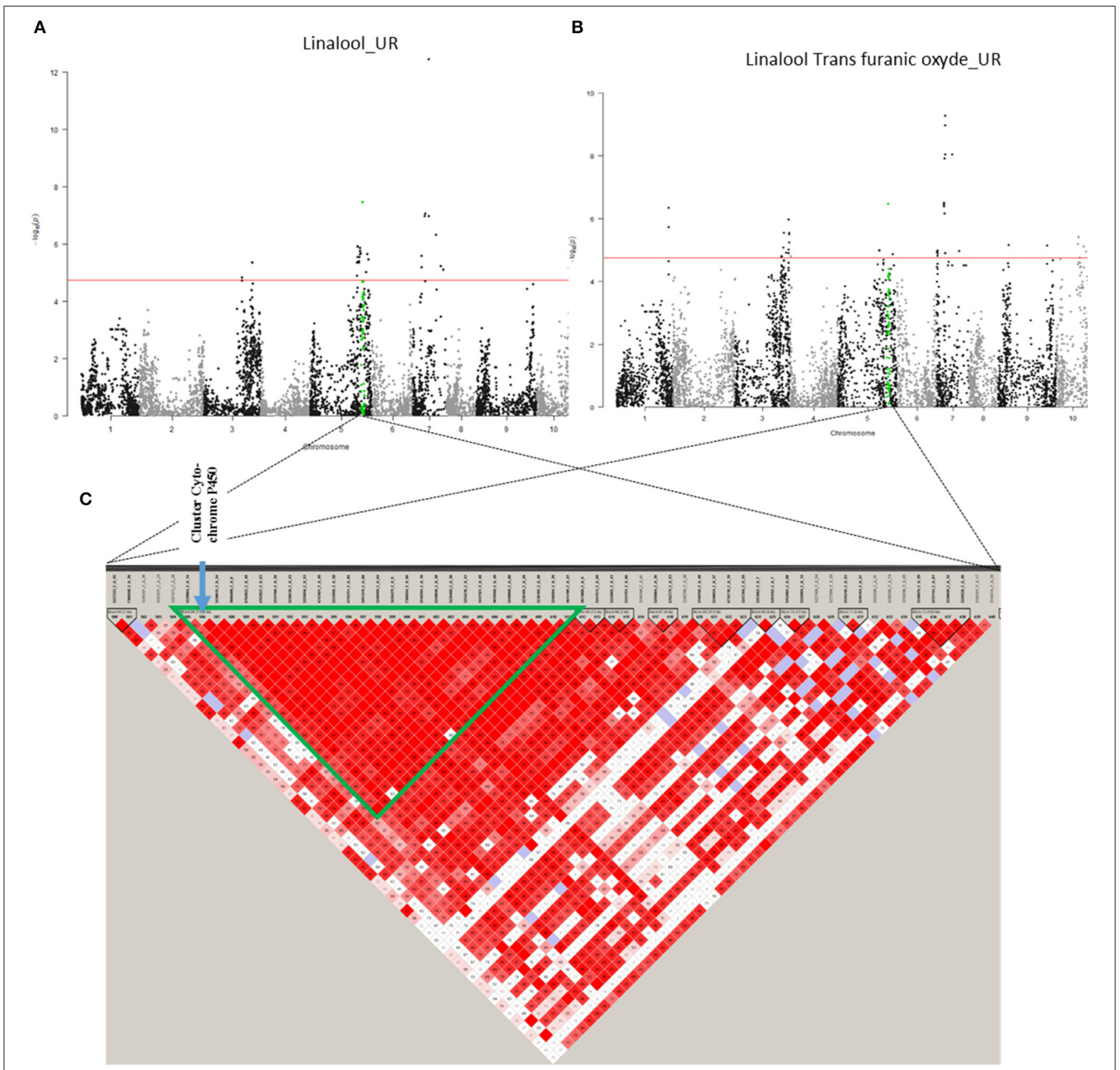

FIGURE 7 | Co-localisation between linalool (UR) and trans furanic oxide (UR) with candidate genes. **(A)** Manhattan plot representing association results for the Linalool (UR) trait, revealed by GLM-MAF5 method. **(B)** Manhattan plot representing association results for linalool trans furanic oxide (UR) trait, revealed by GLM-MAF5 method. **(C)** Heat map of a part of chromosome 5. The common region of association is represented by a green triangle.

of the mevalonate biosynthetic pathway that allows the production of geranyl diphosphate (Kreck et al., 2003; Miziorko, 2011).

On chromosome 5, only association region 6 (32,660,102–33,718,239 bp) contains candidate genes. It is linked to linalool (UR) and linalool trans furanic oxide (UR) and contains six candidate genes, five of which are known to code for "*Cytochrome P450 89A2*" and one for "*Cytochrome P450 89A9*" (**Figure 7**; **Supplementary Figure 7**). The presence of cytochrome P450 could explain the associations with linalool and trans furanic oxide linalool as they would allow the transformation of linalool into epoxylinalool (Chen et al., 2010).

On chromosome 7 (**Supplementary Figure 7**), only association zone 7 (6,128,106–6,410,151 bp) contains candidate genes. It is linked to linalool cis pyranic oxide (UR) and contains three genes encoding "*Probable terpene synthase 9*." Terpene synthases 9 are known to be involved in the synthesise of linalool, one of the precursors of linalool cis pyranic oxide (Cseke et al., 1998).

On chromosome 9 (**Supplementary Figure 7**), only association zone 8 (713,588–857,818 bp) contains a candidate gene. It is linked to epoxylinalool (R) and contains a gene encoding a "*3-hydroxyisobutyryl-CoA hydrolase-like protein 2, mitochondrial*." This enzyme is involved in the mevalonate biosynthetic pathway, one of the biosynthetic pathways leading to the formation of geranyl diphosphate, a key compound in the monoterpene biosynthetic pathway (Lamarti et al., 1994).

On chromosome 10 (**Supplementary Figure 7**), the association zone 9 (6,023,982–6,718,126 bp) linked to linalool cis pyranic oxide (UR) contains a gene coding for "*Probable terpene synthase 9*." This enzyme is known to synthesise linalool, which could enable the synthesis of linalool cis pyranic oxide.

Candidate Genes Linked to the L-Phenylalanine Degradation Pathway

In a second step, candidate genes linked to the L-phenylalanine degradation pathway were found on chromosomes 1, 2, 4, 5, 7, 8, 9, and 10. The association zone number and candidate genes are reported in **Supplementary Figure 6**; **Supplementary Table 3**; and **Table 7**.

On chromosome 1, four association zones contain candidate genes. Association zone 10 (805,132–2,445,782 bp) linked to 1-phenylethyl acetate (R), benzaldehyde (R), and cinnamaldehyde (R) contains a gene coding for an "*Aldehyde dehydrogenase family 3 member F1*." This enzyme could be responsible for the transformation of benzaldehyde into benzoic acid. The presence of this enzyme could compete with the production of cinnamaldehyde or 1-phenylethyl acetate (**Figure 6**; Lapadatescu et al., 2000). Association zone 11 (3,083,032–3,398,183 bp) linked to 1-phenylethyl acetate (R), phenylethyl acetate co-eluted with 2-ethylphenol (R), acetophenone (R), benzaldehyde (R), cinnamaldehyde (R), and the floral note lightwood, contains two candidate genes encoding a "*Probable cinnamyl alcohol dehydrogenase*." These enzymes are known to transform cinnamaldehyde into cinnamyl alcohol (Wyrambik and Grisebach, 1975). According to another study, "*Probable cinnamyl alcohol dehydrogenase*" has the ability to remove hydrogen from cinnamyl alcohol to convert it to cinnamaldehyde. Cinnamyl alcohol is known to have a floral, cinnamon, and balsamic taste (Steinhaus et al., 2009), which may be associated with the floral note lightwood. The association zone 12 (5,940,526–6,204,028 bp) linked to 1-phenylethyl acetate (R) and cinnamaldehyde (R) contains a gene encoding a "*Shikimate kinase 1, chloroplastic*." The shikimate biosynthesis pathway allows the synthesis of phenylalanine, a precursor of 1-phenylethyl acetate and cinnamaldehyde (Tohge et al., 2013). The association zone 13 (6,834,165–7,942,921), linked to 1-phenylethyl acetate (R), phenylethyl acetate co-eluted with 2-ethylphenol (R), acetophenone (R), benzaldehyde (R), and cinnamaldehyde (R), contains two genes coding for an "*Alcohol dehydrogenase 1*." Alcohol dehydrogenase is necessary for the degradation of benzaldehyde to benzyl alcohol or vice versa, which are both compounds with a fruity taste. The other compounds in association in this area are upstream of this degradation reaction, which could explain their associations (Lapadatescu et al., 2000).

On chromosome 2 (**Supplementary Figure 6**), only association region 14 (7,324,500–7,617,242 bp) contains candidate genes. It is linked to acetophenone (NT and R), benzaldehyde (R), benzyl alcohol (UR), cinnamaldehyde (R), and the floral perfume note and contains a candidate gene coding for an "*ALD1 Aminotransferase*." Several aminotransferases have been identified in the shikimate biosynthesis pathway that allows the synthesis of L-phenylalanine (Tohge et al., 2013).

On chromosome 4 (**Supplementary Figure 6**), four association zones contain candidate genes. Association region 15 (22,435,678–22,617,119 bp) linked to 1-phenylethyl acetate and cinnamaldehyde (R) contains a gene encoding an "*NSI acetyltransferase*." The acetyl transferase NSI has the function of acetylating histones. It is likely to play a role in regulating the expression of genes for the synthesis of 1-phenylethyl acetate or cinnamaldehyde. Association zone 16 (26,703,951–27,146,370 bp) linked to 1-phenylethyl acetate (R) contains two candidate genes coding for: a "*Chalcone synthase 2*" and a "*3-ketoacyl-CoA thiolase 2, peroxisomal*." Chalcone synthases participate in the flavonoid and isoflavonoid biosynthesis pathway that follows the degradation of phenylalanine to CA (Pyrzynska and Biesaga, 2009). A ketoacyl-Coa thiolase is required for the synthesis of benzoyl-CoA (Amano et al., 2018), which can be the basis for phenylbenzoate synthesis. The association zone 17 (27,507,597–27,608,727 bp) linked to the floral perfume contains two genes encoding a "*2-hydroxyisoflavanone dehydratase*." 2-hydroxyisoflavanone is part of the isoflavonoid biosynthesis pathway. Its transformation could compete with the synthesis of compounds known to have a floral taste such as acetophenone or 2-phenylethanol (Pyrzynska and Biesaga, 2009). The association zone 18 (28,257,730–28,352,788 bp) linked to 1-phenylethyl acetate (R) contains a gene coding for a "*Probable aldo-keto reductase 1*." An acetaldehyde reductase may be required for the synthesis of 1-phenylethanol from acetophenone, the probable precursor of 1-phenylethyl acetate (Dong et al., 2012).

On chromosome 5, five association zones contain candidate genes. Association region 19 (1,326,444–1,374,494 bp) linked to 4-hydroxy acetophenone (UR), acetophenone (UR), and

TABLE 7 | Candidate genes identified for L-phenylalanine degradation pathway.

N° asso	CH	Position of candidate gene (bp)	Position of the pic of association (bp)	Gene function	Trait in association
10	1	1,102,658	2,430,002	*Aldehyde dehydrogenase family 3 member F1*	1-Phenylethyl acetate (R), benzaldehyde (R), and cinnamaldehyde (R)
11	1	3,107,999	3,379,361	*Probable cinnamyl alcohol dehydrogenase 7/8*	1-Phenylethyl acetate (R), phenylethyl acetate co-eluted with 2-ethylphenol (R), acetophenone (R), benzaldehyde (R), cinnamaldehyde (R), and the floral note lightwood
11	1	3,112,047	3,379,361	*Probable cinnamyl alcohol dehydrogenase*	1-Phenylethyl acetate (R), phenylethyl acetate co-eluted with 2-ethylphenol (R), acetophenone (R), benzaldehyde (R), cinnamaldehyde (R), and the floral note lightwood
12	1	6,039,217	5,940,526	*Shikimate kinase 1, chloroplastic*	1-Phenylethyl acetate (R) and cinnamaldehyde (R)
13	1	7,124,110	6,855,567	*Alcohol dehydrogenase 1*	1-Phenylethyl acetate (R), phenylethyl acetate co-eluted with 2-ethylphenol (R), acetophenone (R), benzaldehyde (R), and cinnamaldehyde (R),
13	1	7,131,266	6,855,567	*Alcohol dehydrogenase 1*	1-Phenylethyl acetate (R), phenylethyl acetate co-eluted with 2-ethylphenol (R), acetophenone (R), benzaldehyde (R), and cinnamaldehyde (R),
14	2	7,436,835	7,453,377	*Aminotransferase ALD1*	Acetophenone (NT and R), benzaldehyde (R), benzyl alcohol (UR), cinnamaldehyde (R), and the floral perfume note
15	4	22,566,348	22,503,297	*Acetyltransferase NSI*	1-Phenylethyl acetate and cinnamaldehyde (R)
16	4	26,741,963	26,876,494	*3-ketoacyl-CoA thiolase 2, peroxisomal*	1-Phenylethyl acetate (R)
16	4	26,715,852	26,876,494	*Chalcone synthase 2*	1-Phenylethyl acetate (R)
17	4	27,604,955	27,507,597	*2-hydroxyisoflavanone dehydratase*	Floral perfume
17	4	27,608,704	27,507,597	*2-hydroxyisoflavanone dehydratase*	Floral perfume
18	4	28,270,492	28,285,175	*Probable aldo-keto reductase 1*	1-Phenylethyl acetate (R)
19	5	1,328,453	1,353,636	*Aldehyde oxidase GLOX*	4-Hydroxy acetophenone (UR), acetophenone (UR), and benzyl acetate (UR)
20	5	1,431,497	1,380,802	*Aldo-keto reductase family 4 member C9*	4-Hydroxy acetophenone (UR), acetophenone (UR), benzyl acetate (UR), and cinnamaldehyde (R)
20	5	1,435,043	1,380,802	*Aldo-keto reductase family 4 member C9*	4-Hydroxy acetophenone (UR), acetophenone (UR), benzyl acetate (UR), and cinnamaldehyde (R)
20	5	1,438,324	1,380,802	*Aldo-keto reductase family 4 member C9*	4-Hydroxy acetophenone (UR), acetophenone (UR), benzyl acetate (UR), and cinnamaldehyde (R)
20	5	1,441,067	1,380,802	*Aldo-keto reductase family 4 member C9*	4-Hydroxy acetophenone (UR), acetophenone (UR), benzyl acetate (UR), and cinnamaldehyde (R)
20	5	1,444,916	1,380,802	*Aldo-keto reductase family 4 member C8*	4-Hydroxy acetophenone (UR), acetophenone (UR), benzyl acetate (UR), and cinnamaldehyde (R)
20	5	1,450,213	1,380,802	*Aldo-keto reductase family 4 member C8*	4-Hydroxy acetophenone (UR), acetophenone (UR), benzyl acetate (UR), and cinnamaldehyde (R)
21	5	2,978,188	2,732,709	*Phenylalanine ammonia-lyase*	Cinnamaldehyde (R)
22	5	30,446,215	30,471,918	*Alcohol dehydrogenase-like 6*	Benzaldehyde (R)

(Continued)

TABLE 7 | Continued

N° asso	CH	Position of candidate gene (bp)	Position of the pic of association (bp)	Gene function	Trait in association
23	7	2,055,963	2,092,063	*GDSL esterase/lipase At1g28570*	4-Hydroxy acetophenone (UR), acetophenone (UR), and benzyl acetate (UR)
24	8	1,199,801	1,148,435	*3-ketoacyl-CoA synthase 4*	Floral note wood resin
25	8	2,170,173	2,251,806	*GDSL esterase/lipase EXL3*	Cinnamaldehyde (R)
26	8	6,559,785	6,751,843	*Acetyltransferase At1g77540*	Acetophenone (UR) and benzyl acetate (UR)
26	8	6,570,129	6,751,843	*Caffeic acid 3-O-methyltransferase*	Acetophenone (UR) and benzyl acetate (UR)
26	8	6,581,002	6,751,843	*Caffeic acid 3-O-methyltransferase*	Acetophenone (UR) and benzyl acetate (UR)
27	8	15,370,133	14,498,544	*Putative O-acyltransferase WSD1*	Benzaldehyde (R), benzyl acetate (UR), cinnamaldehyde (R), and orange blossom note
27	8	15,402,378	14,498,544	*Putative O-acyltransferase WSD1*	Benzaldehyde (R), benzyl acetate (UR), cinnamaldehyde (R), and orange blossom note
28	8	18,679,641	19,249,315	*Putative GDSL esterase/lipase At1g29670*	Benzyl acetate (UR)
29	9	5,605,457	6,010,658	*GDSL esterase/lipase EXL3, putative*	Benzyl alcohol (UR) and the floral note green vegetative
29	9	6,069,175	6,010,658	*3-hydroxyisobutyryl-CoA hydrolase-like protein 3, mitochondrial*	Benzyl alcohol (UR) and the floral note green vegetative
30	9	23,325,443	23,302,911	*Feruloyl CoA ortho-hydroxylase 2*	Acetophenone (R), benzaldehyde (R), and benzyl acetate (UR)
31	10	5,303,984	5,308,832	*Putative 4-coumarate–CoA ligase-like 5*	1-Phenylethyl acetate (R), benzyl acetate (R), phenylethyl acetate co-eluted with 2-ethylphenol (R), to acetophenone (R), benzaldehyde (R,) and cinnamaldehyde (R)
6	5	32,749,861	33,303,465	*Cytochrome P450 89A2*	2-Phenylethanol (UR)
6	5	33,057,632	33,303,465	*Cytochrome P450 89A9*	2-Phenylethanol (UR)
6	5	33,064,477	33,303,465	*Cytochrome P450 89A2*	2-Phenylethanol (UR)
6	5	33,073,996	33,303,465	*Cytochrome P450 89A2*	2-Phenylethanol (UR)
6	5	33,094,200	33,303,465	*Cytochrome P450 89A2*	2-Phenylethanol (UR)
6	5	33,099,009	33,303,465	*Cytochrome P450 89A2*	2-Phenylethanol (UR)

asso., Associations; CH, chromosome; R, Roasted beans; UR, unroasted beans.

benzyl acetate (UR) contains a candidate gene encoding a "*GLOX Aldehyde oxidase*." An aldehyde oxidase is in some cases responsible for the oxidation of phenylacetaldehyde to phenylacetate, both of which are part of the L-phenylalanine degradation pathway (Kücükgöze and Leimkühler, 2018). Association zone 20 (1,380,802–1,510,054 bp) linked to 4-hydroxy acetophenone (UR), acetophenone (UR), benzyl acetate (UR), and cinnamaldehyde (R) contains six candidate genes, four of which code for an *Aldo-keto reductase family 4 member C9* and two for an *Aldo-keto reductase family 4 member C8* (**Figure 8**; **Supplementary Figure 6**). An acetaldehyde reductase may be required for the synthesis of 1-phenylethanol from acetophenone, a probable precursor of 1-phenylethyl acetate (Dong et al., 2012). The association zone 21 (2,674,400–3,039,540 bp) linked to cinnamaldehyde (R) contains a gene coding for a *Phenylalanine ammonia-lyase*. This enzyme is known to transform L-phenylalanine into CA, which is the precursor of cinnamaldehyde (Lapadatescu et al., 2000). The association zone 22 (30,407,214–30,473,075 bp) linked to benzaldehyde (R) contains a gene coding for an *Alcohol dehydrogenase-like 6*. This enzyme could degrade benzaldehyde to benzyl alcohol. Association zone 6 (32,660,102–33,718,239 bp) is linked to 2-phenylethanol (UR) (the same to terpene association zone 6). It contains six genes, five of which code for *Cytochrome P450 89A2* and one for *Cytochrome P450 89A9*. Cytochrome P450 has redox activities. Several of these reactions are involved in the synthesis of 2-phenylethanol (Lapadatescu et al., 2000).

On chromosome 7 (**Supplementary Figure 6**), only association zone 23 (1,894,664–2,092,063 bp) contains a candidate gene. It is linked to 4-hydroxy acetophenone (UR), acetophenone (UR), and benzyl acetate (UR) and contains a gene encoding a *GDSL esterase/lipase At1g28570*. A lipase/esterase may be required for the formation of benzyl acetate from benzyl alcohol or the synthesis of 1-phenyl acetate from 1-phenyl ethanol (Mäki-Arvela et al., 2008; Melo et al., 2017).

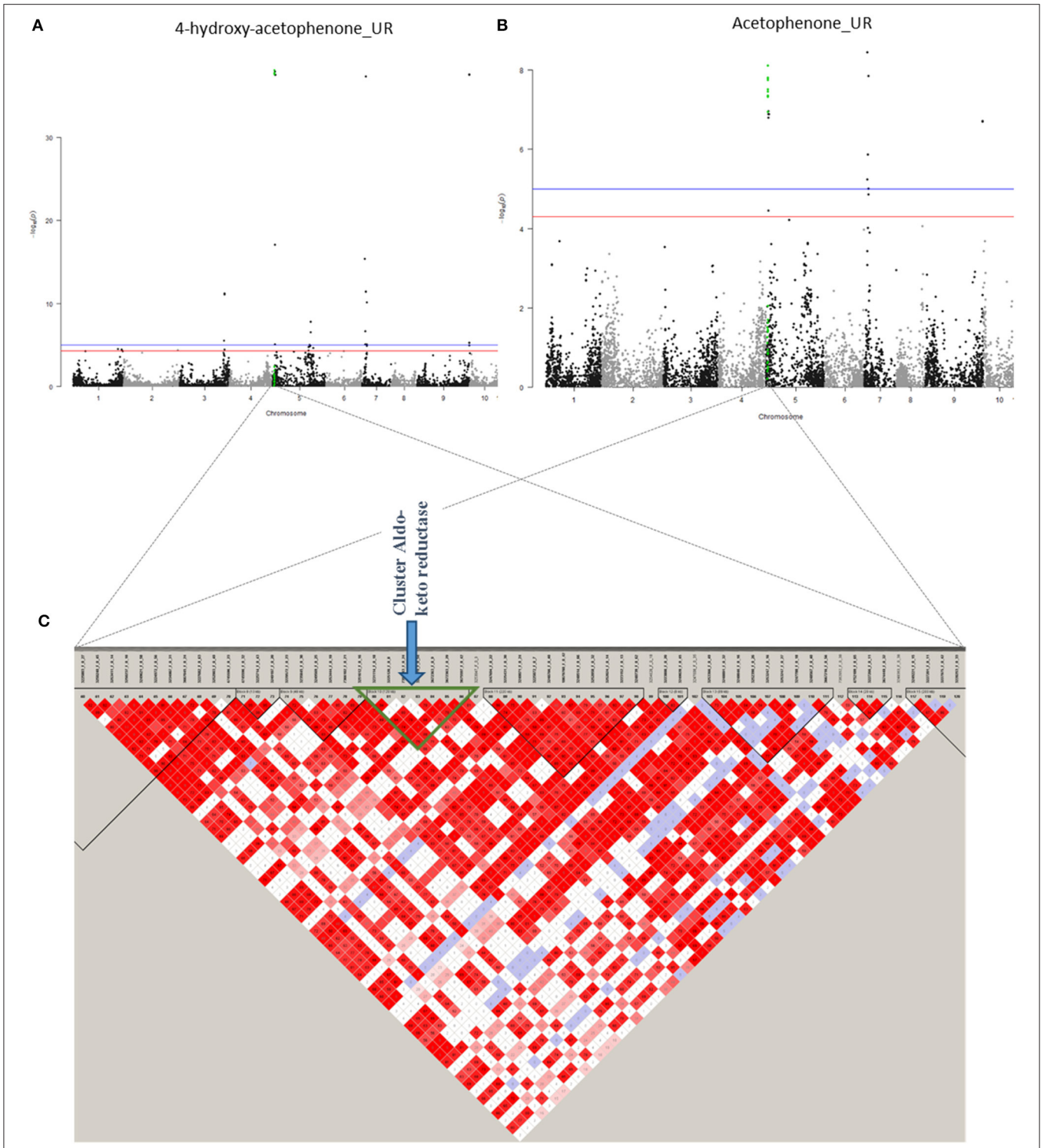

FIGURE 8 | Co-localisation between 4-hydroxy-acetophenone (UR) and acetophenone (UR) with candidate genes. **(A)** Manhattan plot representing association results for the trait 4-hydroxy-acetophenone (UR). **(B)** Manhattan plot representing association results for acetophenone (UR). **(C)** Heat map of a part of chromosome 5. The common region of association is represented by a green triangle.

On chromosome 8 (**Supplementary Figure 6**), five association zones contain candidate genes. Association zone 24 (1,121,979–1,520,555 bp) linked to the floral note wood resin contains a candidate gene encoding a *3-ketoacyl-CoA synthase 4*. This enzyme is involved in the transformation of a very long chain of acyl-CoA into acetyl-CoA which can itself

be transformed into ketones (Tong et al., 2006). Since this zone of associations is linked to the floral note wood resin, this gene can perhaps lead to the synthesis of ketones known to have a floral taste like acetophenone. Association zone 25 (2,021,946–2,268,116 bp) linked to cinnamaldehyde (R) contains a candidate gene encoding a *GDSL esterase/lipase EXL3*. An esterase/lipase may be required as previously discussed for the formation of benzyl acetate from benzyl alcohol or the synthesis of 1-phenylehtyl acetate (Mäki-Arvela et al., 2008; Melo et al., 2017). The synthesis of these compounds could compete with the synthesis of cinnamaldehyde. The association zone 26 (6,533,242–6,978,549 bp) linked to acetophenone (UR) and benzyl acetate (UR) is linked to three genes, two of which code for *Caffeic acid 3-O-methyltransferase* and one for *Acetyltransferase At1g77540*. Caffeic acid 3-O-methyltransferase has the role of transforming caffeic acid into ferulic acid and can thus compete with the synthesis of acetophenone or benzyl acetate (Tu et al., 2010). An acetyltransferase is required to convert benzyl alcohol to benzyl acetate (Hao et al., 2014). This function may explain the associations with acetophenone, which requires a common benzyl alcohol precursor for synthesis. Association zone 27 (14,444,953–15,439,624 bp) linked to benzaldehyde (R), benzyl acetate (UR), cinnamaldehyde (R), and orange blossom note contains two genes encoding a *Putative O-acyltransferase WSD1*. This enzyme allows the synthesis of a "wax ester" from long-chain fatty alcohol. It could allow the synthesis of a "wax ester" with a floral taste of orange blossom type or contribute to this aromatic note. The association zone 28 (17,816,898–19,249,315 bp) linked to benzyl acetate (UR) contains a candidate gene coding for a *Putative GDSL esterase/lipase At1g29670* that may play a role in the degradation of benzyl acetate (Mäki-Arvela et al., 2008; Melo et al., 2017).

On chromosome 9 (**Supplementary Figure 6**), two association zones contain candidate genes. Association zone 29 (5,327,028–6,165,415 bp) linked to benzyl alcohol (UR) and the floral note green vegetative contains two genes: one coding for *3-hydroxyisobutyryl-CoA hydrolase-like protein 3, mitochondrial* and one for *GDSL esterase/lipase EXL3, putative*. The 3-hydroxyisobutyryl-CoA hydrolase-like enzyme could lead to the synthesis of terpenes with floral tastes as described above. It could thus explain the association with the floral green vegetative taste. Lipase may be required for the formation of benzyl acetate from benzyl alcohol (Melo et al., 2017). The enzyme encoded by the GDSL esterase/lipase gene EXL3, putative could compete with the synthesis of benzyl alcohol. Association zone 30 (23,101,222–23,892,356 bp) linked to acetophenone (R), benzaldehyde (R), and benzyl acetate (UR) contains a gene encoding a *Feruloyl CoA ortho-hydroxylase 2*. Ferulic acid has CA as a precursor, as do acetophenone, benzaldehyde, and benzyl acetate. The activity of this enzyme could therefore compete with the synthesis of these compounds.

On chromosome 10 (**Supplementary Figure 6**), one association zone contains candidate genes. Association zone 31 (5,153,882–5,419,006 bp) linked to 1-phenylethyl acetate (R), benzyl acetate (R), phenylethyl acetate co-eluted with 2-ethylphenol (R), to acetophenone (R), benzaldehyde (R), and cinnamaldehyde (R) contains a candidate gene encoding a *Putative 4-coumarate–CoA ligase-like 5*. The activity of this enzyme could compete with the synthesis of compounds associated with this region as it could induce a transformation of CA to coumaric acid.

DISCUSSION

This study contributes to highlighting the importance of cocoa genetic background in the aroma composition of cacao products. The GWAS analyses revealed a large number of associations. Several are related to VOCs known for their floral aromas, others are related to compounds, without floral aroma, but involved in the biosynthesis of these aromatic compounds, and others are related to the perception of sensory notes.

Determination of Associations Area

The confidence interval of the association zones was determined using haplotypic blocks. This method gives an idea of the size of the association zone as a function of the linkage disequilibrium of the population, which seems biologically logical. However, in some cases, this limit may underestimate the true size of the association, as it is certainly the case on chromosome 1 for the epoxylinalool (R) trait (**Supplementary Figure 7**) where we see hot spots of associations extending over the first seven megabases. In cases where there is a cluster of very close association zones, it is legitimate to ask whether the method of determining the association zones is not too stringent.

Insights into the Genetic Architecture of Floral Aromas in Cocoa

Genome-Wide Association Study analysis, two main biosynthesis pathways of compounds known for their floral notes seem to be involved in cocoa floral aromas: the monoterpene synthesis pathway and the L-phenylalanine degradation pathway. These biosynthesis pathways have already been identified in other such as grapes or its derivative wine as important contributors to their floral aromas (Ferreira et al., 1997; Mateo and Jiménez, 2000). Some of the association zones contain candidate genes directly involved in the synthesis of the associated compound, or candidate genes involved upstream in the biosynthetic pathway. The presence of these genes increases the probability that the detected association is not a false positive. The GWAS analyses revealed several genes that appear to be involved in the synthesis of compounds known to have a floral taste and could thus be involved in the variation of floral tastes. Candidate genes coding for enzymes are the most obvious, but other types of genes may be involved in cocoa floral taste such as certain transcriptional factors that could activate or repress several biosynthetic pathways at the same time.

Some associations linked to compounds from the same biosynthesis pathway have been co-localised. Roasting has been suggested to play a role in the transformation of these compounds (Jinap et al., 1998). This could explain some of the co-localisation observed in this study, for example, in the terpene biosynthesis pathway the degradation of linalool to epoxylinalool or vice versa (co-localisation on chromosome 5), the transformation of

cis pyranic oxide linalool to epoxylinalool or the opposite (co-localisation on chromosomes 4 and 10). Roasting may also play a role in the transformation of compounds in the L-phenylalanine degradation pathway as, for example: 4-hydroxy acetophenone to acetophenone or vice versa (co-localisation on chromosomes 7 and 10), the transformation of benzyl acetate into benzaldehyde or the opposite (co-locations on chromosomes 2, 5, 7, 8, 9, and 10), and the transformation of benzyl alcohol into benzaldehyde or vice versa (co-locations on chromosomes 2, 3, 4, 5, 6, 8, and 10).

Other associations give information on a balance between the presence of aromatic and non-aromatic compounds of the same biosynthetic pathway: suggesting that an enzyme could be responsible for the transformation of one of these compounds into another and thus influence the flavour as observed in roses by Farhi et al. (2010). The presence of certain odours would thus depend on the activation or repression of the enzyme responsible for the synthesis of the compound with the floral aroma. This is the case, for example, for an area on chromosome 1 associated with cinnamaldehyde and the floral note lightwood containing a gene coding for a "Probable cinnamyl alcohol dehydrogenase." When this enzyme is active, it would allow the transformation of cinnamaldehyde into cinnamyl alcohol. There would then be a possible accumulation of cinnamyl alcohol known to have a floral note. When this enzyme is not active, cinnamaldehyde, which has a spicy (cinnamon) taste, would accumulate. Other areas of association suggest that a similar system has been put in place: this is the case for the co-locations between 1-phenylethyl acetate and acetophenone on chromosomes 1, 6, 9, and 10 where a gene coding for an esterase/lipase has been detected in nearby location for association zones in chromosome 1, 6, and 9 (**Supplementary Table 3**). If that gene would be active, an accumulation of 1-phenylethyl acetate known to have a fruity odour would be possible. Otherwise, a possible accumulation of acetophenone, also known to have a floral note would be obtained. This is also the case for the co-localisation between benzyl acetate and benzyl alcohol on chromosome 2. A cluster of genes coding for an esterase/lipase and a gene with an acetyltransferase function was detected close to co-location (**Supplementary Table 3**). In this case, if the enzyme is active, an accumulation of benzyl alcohol known to have a sweet taste could be observed. If the enzyme is inactive, a possible accumulation of benzyl acetate known to have a jasmine note could be observed. In the case of co-locations between 4-hydroxy acetophenone and acetophenone on chromosomes 5, 7, and 9 the enzyme transforming 4-hydroxy acetophenone into acetophenone has not been characterised. The candidate gene must have a hydroxylase function that allows the addition of the hydroxyl function on carbon number 4. Two genes (*2-nonaprenyl-3-methyl-6-methoxy-1, 4-benzoquinol hydroxylase,* and *Abscisic acid 8'-hydroxylase 2*) with this function been identified close to the association zones on chromosomes 7 and 9 (**Supplementary Table 3**).

The position of the most significant association zones for the same compound may be different if this compound has been detected in roasted or unroasted beans. This is the case for benzyl acetate, acetophenone, benzaldehyde, furfural, and linalool (**Tables 3–5**). This difference can be explained by the response to two different phenomena: during fermentation, the enzymes responsible for the synthesis of compounds would be activated. A "classical" synthesis would then be carried out in the bean. Whereas, during roasting, the thickness of the shell or the size of the bean could play a role in the chemical conditions of the bean such as temperature or pH and thus influence the degradation of certain aromatic compounds. In that case, the detection of association would depend also on the location of genes involved in the bean structure and size. It is also possible that the difference is due to the presence of precursors that allow the genesis of aromatic compounds during roasting.

This is not the case for all compounds. On the contrary, 2-phenylethanol dosed in roasted and unroasted beans has peaks of very close associations and there are also co-locations between acetophenone related associations dosed in roasted and unroasted beans on chromosomes 2, 6, and 9 confirming the importance of these areas in the genesis of these compounds.

The formation of an aroma as well as its perception depends on a large number of conditions. An aromatic note is generally composed of a combination of several VOCs at different concentrations (Pérez-Silva et al., 2006). Aromatic traits, therefore, have a high probability of being polygenic, which is consistent with the large number of associations that have been found in this study. The expression of an aromatic note also depends on the matrix in which VOCs are contained (Afoakwa et al., 2008). The production of these compounds by plants also depends on their environment (Baldwin, 2010). These factors therefore partly explain why large number of associations was found.

The synthesis of a flavour is therefore due to many external parameters but also the genetic background of the *T. cacao* trees (Luna et al., 2002; Afoakwa et al., 2008). Due to its multigenic determinism, the total variance of a compound is the result of many small associations, each of which would explain, a small part of the genetic variance. Once these small associations are combined, they could explain a large part of the genetic variance. In this case, some associations may contain only one associated marker, as is the case for linalool on chromosome 2. It is also possible that some associations do not cross the significance threshold and are therefore not identified. This hypothesis suggests that some associations with certain VOCs have not been revealed, explaining why the analysis of some compounds known to have a floral taste does not reveal an association zone as for guaiacol (R).

Role of Fermentative Micro-Organisms in Cocoa Flavour Synthesis

The analysis of three other compounds known to have a floral taste belonging to the family of esters did not detect zones of associations: ethyl 2-hydroxyhexanoate (R), ethylphenyl acetate (UR), and ethyl hexanoate (UR). These compounds present after fermentation and before roasting could also be synthesised by yeasts during fermentation (Soles et al., 1982). In this case, no area of association can be found as this would depend on the micro-organisms population and not on the cocoa seeds. The

non-detection of association zones can also be due partially to the pollination of the mother tree made by a mix of progenitors. While genotyping is done on the mother tree, phenotyping (VOC assay and sensory analysis) is done on the beans, hybrids between the mother tree and male pollinators, which could lead to a partial discrepancy between genetic and phenotypic data. Currently, it is not possible to genotype and phenotype individually each bean.

Volatile organic compounds (VOCs) produced by plants are involved in various processes and often released for defence, signalling, or pollinator attraction purposes (Baldwin, 2010). Volatile organic compounds belong to different biochemical families such as terpenes. They are notably involved in direct and indirect defence against insects (Martin et al., 2002) and micro-organisms (Pichersky et al., 1995). Compounds of the terpene family are recognised as a molecular signal in many interactions between plants and various other species, particularly in competition reactions, in the presence of herbivores or pathogenic microorganisms, but also the presence of beneficial insects (Langenheim, 1994; Bohlmann et al., 1998). The same is true for certain phenolic compounds such as acetophenone or 4-hydroxyacetophenone that could be involved in defence mechanisms (Parent et al., 2018), which has also been observed for furfural (Palmqvist et al., 1999; Miller et al., 2009).

During fermentation, the change in environment and chemical composition of the medium induced by yeasts and bacteria can be taken as a threat and cause the seed to react. Then, they could release VOCs to defend themselves and would be responsible for the synthesis of VOCs involved in fine flavour, as suggested by Sabau et al. (2006) who observed an increase in the expression of the gene coding for linalool synthase during fermentation. Also, a strong increase in the concentration of linalool, epoxylinalool, and 2-phenylethanol has also been observed during fermentation in aromatic fine cocoa beans by other authors (Cevallos-Cevallos et al., 2018).

If cocoa beans use VOCs as a defence mechanism against external microorganisms such as fermentative yeasts, lactic bacteria, or acetic bacteria, some questions remain unanswered: by which mechanisms do they detect such microorganisms? Knowing that different types of yeast have been identified according to the place of fermentation (Schwan and Wheals, 2004), we can also ask ourselves whether certain types of yeast or microorganisms are more favourable to this activation. Another hypothesis is that the presence of microorganisms and the transformations they induce (change in pH, synthesis of unknown compounds in the seed, etc.) induce the synthesis of VOCs. In this case, VOCs could be triggered in the absence of microorganisms.

CONCLUSIONS AND PERSPECTIVES

The perception of an aroma and the sensorial analyses is a difficult task. They, therefore, depends on a large number of conditions, including the perception threshold of aromatic molecules. The presence of a molecule is therefore not synonymous with the perception of its taste. Similarly, regions of the genome identified as being associated with the content of biochemical compounds do not mean that these compounds are involved in the flavour of cocoa. Additional analyses are necessary to validate the involvement of these molecules in the formation of taste such as gas chromatography coupled to olfactometry (GCO) analyses for example. Knowing the main molecules responsible for the floral taste as well as the mechanisms of synthesis and degradation of the compounds during fermentation and roasting could also, in the long term, allow the adaptation of the roasting process (temperatures and roasting time) to preserve the most fragile aromatic compounds. Knowledge of the biosynthesis pathway of cocoa aromatic compounds could provide a better mastering of the parameters of fermentations allowing the synthesis of these molecules.

The identification of these molecules and their biosynthetic pathway within the cocoa tree is complex. A genomic selection approach could allow early prediction of aroma traits for the search of cocoa trees having good aroma potential, especially as certain genetic variation could explain a large extend of biochemical compounds in the beans. In this case, a marker-assisted selection could be envisaged in the selected programmes to make it easier for the selection of the cocoa trees aromatic quality.

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**.

ETHICS STATEMENT

The studies involving human participants did not require approval in line with regional/national guidelines. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

EC, CL, and RL conceived the experiment. J-CJ and AS conducted biochemical analyses. ES carried out sensorial analyses. OF carried out DNA experiments. KC, J-CJ, AS, RB, CL, FD, SA, and XA analysed data. KC, RB, and CL wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was funded by the United States Department of State (U.S. Foreign Ministry), the U.S. Embassy, Quito, and the U.S. Department of Agriculture (USDA-ARS) with the agreement n° 58-4001-2-F128 and the MUSE Amazcacao project with the reference ANR-16-IDEX-0006.

ACKNOWLEDGMENTS

We thank the USDA and the I-Site MUSE for their financial support to this project. This work, part of the MUSE Amazcacao

project, was publicly funded through ANR (the French National Research Agency) under the Investissement d'avenir programme with the reference ANR-16-IDEX-0006. We are grateful to Eric Rosenquist for his support in the coordination of our project.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.681979/full#supplementary-material

REFERENCES

Afoakwa, E. O., Paterson, A., Fowler, M., and Ryan, A. (2008). Flavor formation and character in cocoa and chocolate: a critical review. *Crit. Rev. Food Sci. Nutr.* 48, 840–857. doi: 10.1080/10408390701719272

Amano, I., Kitajima, S., Suzuki, H., Koeduka, T., and Shitan, N. (2018). Transcriptome analysis of *Petunia axillaris* flowers reveals genes involved in morphological differentiation and metabolite transport. *PLoS ONE* 13, e0198936. doi: 10.1371/journal.pone.0198936

Argout, X., Martin, G., Droc, G., Fouet, O., Labadie, K., Rivals, E., et al. (2017). The cacao Criollo genome v2.0: an improved version of the genome for genetic and functional genomic studies. *BMC Genomics* 18, 730. doi: 10.1186/s12864-017-4120-9

Argout, X., Salse, J., Aury, J.-M., Guiltinan, M. J., Droc, G., Gouzy, J., et al. (2011). The genome of *Theobroma cacao*. *Nat. Genet.* 43, 101. doi: 10.1038/ng.736

Arn, H., and Acree, T. E. (1998). Flavornet: a database of aroma compounds based on odor potency in natural products. *Dev Food Sci* 40, 27. doi: 10.1016/S0167-4501(98)80029-0

Assi-Clair, B. J., Koné, M. K., Kouamé, K., Lahon, M. C., Berthiot, L., Durand, N., et al. (2019). Effect of aroma potential of *Saccharomyces cerevisiae* fermentation on the volatile profile of raw cocoa and sensory attributes of chocolate produced thereof. *Eur. Food Res. Technol.* 245, 1459–1471. doi: 10.1007/s00217-018-3181-6

Baek, H. H., Cadwallader, K. R., Marroquin, E., and Silva, J. L. (1997). Identification of predominant aroma compounds in muscadine grape juice. *J. Food Sci.* 62, 249–252. doi: 10.1111/j.1365-2621.1997.tb03978.x

Baldwin, I. T. (2010). Plant volatiles. *Curr. Biol.* 20, R392–R397. doi: 10.1016/j.cub.2010.02.052

Barrett, J. C., Fry, B., Maller, J., and Daly, M. J. (2005). Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21, 263–265. doi: 10.1093/bioinformatics/bth457

Bayer, C., and Kubitzki, K. (2003). "Malvaceae. Fam. Genera Vasc.," in *Plants Dicotyledons Malvales Capparales Non-Betalain Caryophyllales*, eds K. Kubitzki and C. Bayer (Berlin: Springer), 225–311. doi: 10.1007/978-3-662-07255-4_28

Bohlmann, J., Meyer-Gauen, G., and Croteau, R. (1998). Plant terpenoid synthases: molecular biology and phylogenetic analysis. *Proc. Natl. Acad. Sci. U.S.A.* 95, 4126–4133. doi: 10.1073/pnas.95.8.4126

Cevallos-Cevallos, J. M., Gysel, L., Maridueña-Zavala, M. G., and Molina-Miranda, M. J. (2018). Time-related changes in volatile compounds during fermentation of bulk and fine-flavor cocoa (*Theobroma cacao*) beans. *J. Food Qual.* 2018, 1758381. doi: 10.1155/2018/1758381

Cheesman, E. (1944). Notes on the nomenclature, classification and possible relationships of cacao populations. *Trop. Agric.* 21, 144–159.

Chen, X., Yauk, Y.-K., Nieuwenhuizen, N. J., Matich, A. J., Wang, M. Y., Perez, R. L., et al. (2010). Characterisation of an (S)-linalool synthase from kiwifruit (*Actinidia arguta*) that catalyses the first committed step in the production of floral lilac compounds. *Funct. Plant Biol.* 37, 232–243. doi: 10.1071/FP09179

Colahan-Sederstrom, P. M., and Peterson, D. G. (2005). Inhibition of key aroma compound generated during ultrahigh-temperature processing of bovine milk via epicatechin addition. *J. Agric. Food Chem.* 53, 398–402. doi: 10.1021/jf0487248

Cros, E., and Jeanjean, N. (1995). Qualité du cacao : influence de la fermentation et du séchage. *Plant Rech Dev* 2, 21–27.

Cseke, L., Dudareva, N., and Pichersky, E. (1998). Structure and evolution of linalool synthase. *Mol. Biol. Evol.* 15, 1491–1498. doi: 10.1093/oxfordjournals.molbev.a025876

Dice, L. R. (1945). Measures of the amount of ecologic association between species. *Ecology* 26, 297–302. doi: 10.2307/1932409

Dong, F., Yang, Z., Baldermann, S., Kajitani, Y., Ota, S., Kasuga, H., et al. (2012). Characterization of l-phenylalanine metabolism to acetophenone and 1-phenylethanol in the flowers of *Camellia sinensis*using stable isotope labeling. *J. Plant Physiol.* 169, 217–225. doi: 10.1016/j.jplph.2011.12.003

Dudareva, N., Cseke, L., Blanc, V. M., and Pichersky, E. (1996). Evolution of floral scent in Clarkia: novel patterns of S-linalool synthase gene expression in the *C. breweri* flower. *Plant Cell* 8, 1137–1148. doi: 10.1105/tpc.8.7.1137

Farhi, M., Lavie, O., Masci, T., Hendel-Rahmanim, K., Weiss, D., Abeliovich, H., et al. (2010). Identification of rose phenylacetaldehyde synthase by functional complementation in yeast. *Plant Mol. Biol.* 72, 235–245. doi: 10.1007/s11103-009-9564-0

Feng, L., Chen, C., Li, T., Wang, M., Tao, J., Zhao, D., et al. (2014). Flowery odor formation revealed by differential expression of monoterpene biosynthetic genes and monoterpene accumulation in rose (*Rosa rugosa Thunb.*). *Plant Physiol. Biochem.* 75, 80–88. doi: 10.1016/j.plaphy.2013.12.006

Ferreira, V., López, R., Escudero, A., and Cacho, J. F. (1997). The aroma of Grenache red wine: hierarchy and nature of its main odorants. *J. Sci. Food Agric.* 77, 259–267. doi: 10.1002/(SICI)1097-0010(199806)77:2<259::AIDJSFA36>3.0.CO;2-Q

Gao, X., Becker, L. C., Becker, D. M., Starmer, J. D., and Province, M. A. (2010). Avoiding the high Bonferroni penalty in genome-wide association studies. *Genet. Epidemiol.* 34, 100–105. doi: 10.1002/gepi.20430

Gao, X., Starmer, J., and Martin, E. R. (2008). A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. *Genet. Epidemiol.* 32, 361–369. doi: 10.1002/gepi.20310

Garg, N., Sethupathy, A., Tuwani, R., Nk, R., Dokania, S., Iyer, A., et al. (2018). FlavorDB: a database of flavor molecules. *Nucleic Acids Res.* 46, D1210–D1216. doi: 10.1093/nar/gkx957

Genovese, A., Gambuti, A., Piombino, P., and Moio, L. (2007). Sensory properties and aroma compounds of sweet Fiano wine. *Food Chem.* 103, 1228–1236. doi: 10.1016/j.foodchem.2006.10.027

Guichard, H., Lemesle, S., Ledauphin, J., Barillier, D., and Picoche, B. (2003). Chemical and sensorial aroma characterization of freshly distilled calvados. 1. Evaluation of quality and defects on the basis of key odorants by olfactometry and sensory analysis. *J. Agric. Food Chem.* 51, 424–432. doi: 10.1021/jf020372m

Hamdouche, Y., Meile, J. C., Lebrun, M., Guehi, T., Boulanger, R., Teyssier, C., et al. (2019). Impact of turning, pod storage and fermentation time on microbial ecology and volatile composition of cocoa beans. *Food Res. Int. Ott. Ont.* 119, 477–491. doi: 10.1016/j.foodres.2019.01.001

Hao, R., du, D., Wang, T., Yang, W., and Wang, J., Zhang, Q. (2014). A comparative analysis of characteristic floral scent compounds in *Prunus mume* and related species. *Biosci. Biotechnol. Biochem.* 78, 1640–1647. doi: 10.1080/09168451.2014.936346

Helsper, J. P. F. G., Davies, J. A., Bouwmeester, H. J., Krol, A. F., and van Kampen, M. H. (1998). Circadian rhythmicity in emission of volatile compounds by flowers of *Rosa hybrida* L. cv. Honesty. *Planta* 207, 88–95. doi: 10.1007/s004250050459

Ho, V. T. T., Zhao, J., and Fleet, G. (2014). Yeasts are essential for cocoa bean fermentation. *Int. J. Food Microbiol.* 174, 72–87. doi: 10.1016/j.ijfoodmicro.2013.12.014

ICCO. (2020). *Production of cocoa beans (thousand tonnes) year 2019/2020. Q. Bull. Cocoa Stat. XLVI.*

International Cocoa Organization. (2017). *Fine or Flavour Cocoa.*

ISCQF. (2020). *First Draft of the Protocol for Cocoa Liquor Sensory Evaluation: part of the International Standards for the Assessment of Cocoa Quality and Flavour (ISCQF).* Compiled by Bioversity International, in collaboration with the members of the ISCQF Working Group.

Ito, Y., Sugimoto, A., Kakuda, T., and Kubota, K. (2002). Identification of potent odorants in Chinese jasmine green tea scented with flowers of *Jasminum sambac*. *J. Agric. Food Chem.* 50, 4878–4884. doi: 10.1021/jf020282h

Jezussek, M., Juliano, B. O., and Schieberle, P. (2002). Comparison of key aroma compounds in cooked brown rice varieties based on aroma extract dilution analyses. *J. Agric. Food Chem.* 50, 1101–1105. doi: 10.1021/jf0108720

Jinap, S., Rosli, W. I. W., Russly, A. R., and Nordin, L. M. (1998). Effect of roasting time and temperature on volatile component profiles during nib roasting of cocoa beans (*Theobroma cacao*). *J. Sci. Food Agric.* 77, 441–448. doi: 10.1002/(SICI)1097-0010(199808)77:4<441::AID-JSFA46>3.0.CO;2-%23

Kadow, D., Bohlmann, J., Phillips, W., and Lieberei, R. (2013). Identification of main fine flavour components in two genotypes of the chocolate tree (*Theobroma cacao* L). *J. Appl. Bot. Food Qual. Bot.* 86, 90–98. doi: 10.5073/JABFQ.2013.086.013

Karagül-Yüceer, Y., Drake, M. A., and Cadwallader, K. R. (2006). Aroma-active components of liquid cheddar whey. *J. Food Sci.* 68, 1215–1219. doi: 10.1111/j.1365-2621.2003.tb09627.x

Kilian, A., Wenzl, P., Huttner, E., Carling, J., Xia, L., Blois, H., et al. (2012). "Diversity arrays technology: a generic genome profiling technology on open platforms," in *Data Production and Analysis in Population Genomics*, eds F. Pompanon and A. Bonin (Totowa, NJ: Humana Press), 67–89. doi: 10.1007/978-1-61779-870-2_5

Kreck, M., Püschel, S., Wüst, M., and Mosandl, A. (2003). Biogenetic studies in *Syringa vulgaris* L.: Synthesis and bioconversion of deuterium-labeled precursors into lilac aldehydes and lilac alcohols. *J. Agric. Food Chem.* 51, 463–469. doi: 10.1021/jf020845p

Kücükgöze, G., and Leimkühler, S. (2018). Direct comparison of the four aldehyde oxidase enzymes present in mouse gives insight into their substrate specificities. *PLoS ONE* 25, e0191819. doi: 10.1371/journal.pone.0191819

Kumazawa, K., and Masuda, H. (2002). Identification of potent odorants in different green tea varieties using flavor dilution technique. *J. Agric. Food Chem.* 50, 5660–5663. doi: 10.1021/jf020498j

Lamarti, A., Badoc, A., Deffieux, G., and Carde, J.-P. (1994). *Biogénèse des monoterpènes. ii - la chaîne isoprénique. Bull. Soc. Pharm. Bordeaux* (1994), 133, 79–99.

Lanaud, C., Boult, E., Clapperton, J., N'Goran, J., Cros, E., Chapelin, M., et al. (2003). "Identification of QTLs related to fat content, seed size and sensorial traits in *Theobroma cacao* L.," in *14th International Cocoa Research Conference* (Accra), 1119–1126.

Lanaud, C., Saltos, A., Jimenez, J. C., Lemainque, A., Pavek, S., Argout, X., et al. (2012). "Adding value to *T. cacao* germplasm collections combining GWAS and genome sequence analysis," *Plant and Animal Genome XX Conference W118* (San Diego).

Langenheim, J. H. (1994). Higher plant terpenoids: a phytocentric overview of their ecological roles. *J. Chem. Ecol.* 20, 1223–1280. doi: 10.1007/BF02059809

Lapadatescu, C., Giniès, C., Le Quéré, J. L., and Bonnarme, P. (2000). Novel scheme for biosynthesis of aryl metabolites from L-phenylalanine in the fungus *Bjerkandera adusta*. *Appl. Environ. Microbiol.* 66, 1517–1522. doi: 10.1128/AEM.66.4.1517-1522.2000

Larsen, M., and Poll, L. (1992). Odour thresholds of some important aroma compounds in strawberries. *Z. Lebensm. Unters. Forsch.* 195, 120–123. doi: 10.1007/BF01201770

Loor, S. R. G. (2007). *Contribution à L'étude de la Domestication de la Variété de Cacaoyer Nacional d'Equateur : Recherche de la Variété Native et de ses Ancêtres Sauvages*. INIAP Archivo Historico.

Luna, F., Crouzillat, D., Cirou, L., and Bucheli, P. (2002). Chemical composition and flavor of *Ecuadorian cocoa* liquor. *J. Agric. Food Chem.* 50, 3527–3532. doi: 10.1021/jf0116597

Mahajan, S. S., Goddik, L., and Qian, M. C. (2004). Aroma compounds in sweet whey powder. *J. Dairy Sci.* 87, 4057–4063. doi: 10.3168/jds.S0022-0302(04)73547-X

Mäki-Arvela, P., Sahin, S., Kumar, N., Heikkilä, T., Lehto, V.-P., Salmi, T., et al. (2008). Cascade approach for synthesis of R-1-phenyl ethyl acetate from acetophenone: effect of support. *J. Mol. Catal. Chem.* 285, 132–141. doi: 10.1016/j.molcata.2008.01.032

Martin, D., Tholl, D., Gershenzon, J., and Bohlmann, J. (2002). Methyl jasmonate induces traumatic resin ducts, terpenoid resin biosynthesis, and terpenoid accumulation in developing xylem of Norway spruce stems. *Plant Physiol.* 129, 1003–1018. doi: 10.1104/pp.011001

Mateo, J. J., and Jiménez, M. (2000). Monoterpenes in grape juice and wines. *J. Chromatogr. A* 881, 557–567. doi: 10.1016/S0021-9673(99)01342-4

Meesters, R. J. W., Duisken, M., and Hollender, J. (2007). Study on the cytochrome P450-mediated oxidative metabolism of the terpene alcohol linalool: indication of biological epoxidation. *Xenobiot. Fate Foreign Compd. Biol. Syst.* 37, 604–617. doi: 10.3109/00498250701393191

Melo, A. D. Q., Silva, F. F. M., Dos Santos, J. C. S., Fernández-Lafuente, R., Lemos, T. L. G., and Dias Filho, F. A. (2017). Synthesis of benzyl acetate catalyzed by lipase immobilized in nontoxic chitosan-polyphosphate beads. *Molecules* 22, 2165. doi: 10.3390/molecules22122165

Miller, E. N., Jarboe, L. R., Turner, P. C., Pharkya, P., Yomano, L. P., York, S. W., et al. (2009). Furfural inhibits growth by limiting sulfur assimilation in ethanologenic *Escherichia coli* strain LY180. *Appl. Environ. Microbiol.* 75, 6132–6141. doi: 10.1128/AEM.01187-09

Miziorko, H. M. (2011). Enzymes of the mevalonate pathway of isoprenoid biosynthesis. *Arch. Biochem. Biophys.* 505, 131–143. doi: 10.1016/j.abb.2010.09.028

Motamayor, J. C., Lachenaud, P., da Se Mota, J. W., Loor, R., Kuhn, D. N., Brown, J. S., et al. (2008). Geographic and genetic population differentiation of the amazonian chocolate tree (*Theobroma cacao* L). *PLoS ONE* 3, e3311. doi: 10.1371/journal.pone.0003311

Motamayor, J. C., Mockaitis, K., Schmutz, J., Haiminen, N., I. I. I. D. L., Cornejo, O., et al. (2013). The genome sequence of the most widely cultivated cacao type and its use to identify candidate genes regulating pod color. *Genome Biol.* 14, r53. doi: 10.1186/gb-2013-14-6-r53

Palmqvist, E., Almeida, J. S., and Hahn-Hägerdal, B. (1999). Influence of furfural on anaerobic glycolytic kinetics of *Saccharomyces cerevisiae* in batch culture. *Biotechnol. Bioeng.* 62, 447–454. doi: 10.1002/(SICI)1097-0290(19990220)62:4<447::AID-BIT7>3.0.CO;2-0

Parent, G. J., Giguère, I., Mageroy, M., Bohlmann, J., and MacKay, J. J. (2018). Evolution of the biosynthesis of two hydroxyacetophenones in plants. *Plant Cell Environ.* 41, 620–629. doi: 10.1111/pce.13134

Perestrelo, R., Fernandes, A., Albuquerque, F. F., Marques, J. C., and Câmara, J. S. (2006). Analytical characterization of the aroma of Tinta Negra Mole red wine: identification of the main odorants compounds. *Anal. Chim. Acta* 563, 154–164. doi: 10.1016/j.aca.2005.10.023

Pérez-Silva, A., Odoux, E., Brat, P., Ribeyre, F., Rodriguez-Jimenes, G., Robles-Olvera, V., et al. (2006). GC–MS and GC–olfactometry analysis of aroma compounds in a representative organic aroma extract from cured vanilla (*Vanilla planifolia* G. Jackson) beans. *Food Chem.* 99, 728–735. doi: 10.1016/j.foodchem.2005.08.050

Perrier, X., and Jacquemoud-Collet, J. P. (2006). *DARwin Software*. Available online at: http://darwin.cirad.fr/

Pham, A. J., Schilling, M. W., Yoon, Y., Kamadia, V. V., and Marshall, D. L. (2008). Characterization of fish sauce aroma-impact compounds using GC-MS, SPME-Osme-GCO, and Stevens' power law exponents. *J. Food Sci.* 73, C268–C274. doi: 10.1111/j.1750-3841.2008.00709.x

Pichersky, E., Lewinsohn, E., and Croteau, R. (1995). Purification and characterization of S-linalool synthase, an enzyme involved in the production of floral scent in *Clarkia breweri*. *Arch. Biochem. Biophys.* 316, 803–807. doi: 10.1006/abbi.1995.1107

Pichersky, E., Raguso, R. A., Lewinsohn, E., and Croteau, R. (1994). Floral scent production in Clarkia (Onagraceae) (I. Localization and developmental modulation of Monoterpene emission and Linalool synthase activity). *Plant Physiol.* 106, 1533–1540. doi: 10.1104/pp.106.4.1533

Pyrzynska, K., and Biesaga, M. (2009). Analysis of phenolic acids and flavonoids in honey. *TrAC Trends Anal. Chem.* 28, 893–902. doi: 10.1016/j.trac.2009.03.015

Qin, X.-W., Lai, J.-X., Tan, L.-H., Hao, C.-Y., Li, F.-P., He, S.-Z., et al. (2017). Characterization of volatile compounds in Criollo, Forastero, and Trinitario cocoa seeds (*Theobroma cacao* L.) in China. *Int. J. Food Prop.* 20, 2261–2275. doi: 10.1080/10942912.2016.1236270

Risterucci, A. M., Grivet, L., N'Goran, J. A. K., Pieretti, I., Flament, M. H., and Lanaud, C. (2000). A high-density linkage map of *Theobroma cacao* L. *Theor. Appl. Genet.* 101, 948–955. doi: 10.1007/s001220051566

Roccia, A., Oyant, L. H.-S., Cavel, E., Caissard, J.-C., Machenaud, J., Thouroude, T., et al. (2019). Biosynthesis of 2-phenylethanol in rose petals is linked

to the expression of one allele of RhPAAS. *Plant Physiol.* 179, 1064–1079. doi: 10.1104/pp.18.01468

Rodriguez-Campos, J., Escalona-Buendia, H. B., Contreras-Ramos, S. M., Orozco-Avila, I., Jaramillo-Flores, E., and Lugo-Cervantes, E. (2012). Effect of fermentation time and drying temperature on volatile compounds in cocoa. *Food Chem.* 132, 277–288. doi: 10.1016/j.foodchem.2011.10.078

Rodriguez-Campos, J., Escalona-Buendia, H. B., Orozco-Avila, I., Lugo-Cervantes, E., and Jaramillo-Flores, M. E. (2011). Dynamics of volatile and non-volatile compounds in cocoa (*Theobroma cacao* L.) during fermentation and drying processes using principal components analysis. *Food Res. Int.* 44, 250–258. doi: 10.1016/j.foodres.2010.10.028

Rottiers, H., Tzompa Sosa, D. A., Lemarcq, V., De Winne, A., De Wever, J., Everaert, H., et al. (2019). A multipronged flavor comparison of Ecuadorian CCN51 and Nacional cocoa cultivars. *Eur. Food Res. Technol.* 245, 2459–2478. doi: 10.1007/s00217-019-03364-3

Ruiz, M., Sempéré, G., and Hamelin, C. (2017). "Using TropGeneDB: a database containing data on molecular markers, QTLs, maps, genotypes, and phenotypes for tropical crops," in *Plant Genomics Databases Methods and Protocols*, ed A. D. J. van Dijk (New York, NY: Springer), 161–172. doi: 10.1007/978-1-4939-6658-5_8

Sabau, X., Loor, R. G., Boccara, M., Fouet, O., Jeanneau, M., Argout, X., et al. (2006). "Preliminary results on linalool synthase expression during seed development and fermentation of Nacional and Trinitario clones," in *15th International Cocoa Research Conference: Cocoa Productivity, Quality, Profitability, Human Health and the Environment* (San José).

Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425,

Schwab, W., Davidovich-Rikanati, R., and Lewinsohn, E. (2008). Biosynthesis of plant-derived flavor compounds. *Plant J.* 54, 712–732. doi: 10.1111/j.1365-313X.2008.03446.x

Schwan, R. F., and Wheals, A. E. (2004). The microbiology of cocoa fermentation and its role in chocolate quality. *Crit. Rev. Food Sci. Nutr.* 44, 205–221. doi: 10.1080/10408690490464104

Soles, R. M., Ough, C. S., and Kunkee, R. E. (1982). Ester concentration differences in wine fermented by various species and strains of yeasts. *Am. J. Enol. Vitic.* 33, 94–98.

Steinhaus, M., Sinuco, D., Polster, J., Osorio, C., and Schieberle, P. (2009). Characterization of the key aroma compounds in pink guava (*Psidium guajava* L.) by means of aroma re-engineering experiments and omission tests. *J. Agric. Food Chem.* 57, 2882–2888. doi: 10.1021/jf803728n

Sukha, D. A., Butler, D. R., Umaharan, P., and Boult, E. (2008). The use of an optimised organoleptic assessment protocol to describe and quantify different flavour attributes of cocoa liquors made from Ghana and Trinitario beans. *Eur. Food Res. Technol.* 226, 405–413. doi: 10.1007/s00217-006-0551-2

Tohge, T., Watanabe, M., Hoefgen, R., and Fernie, A. R. (2013). Shikimate and phenylalanine biosynthesis in the green lineage. *Front. Plant Sci.* 4:62. doi: 10.3389/fpls.2013.00062

Tong, M. K. H., Lam, C.-S., Mak, T. W. L., Fu, M. Y. P., Ng, S.-H., Wanders, R. J. A., et al. (2006). Very long-chain acyl-CoA dehydrogenase deficiency presenting as acute hypercapnic respiratory failure. *Eur. Respir. J.* 28, 447–450. doi: 10.1183/09031936.06.00139205

Tu, Y., Rochfort, S., Liu, Z., Ran, Y., Griffith, M., Badenhorst, P., et al. (2010). Functional analyses of caffeic acid O-methyltransferase and cinnamoyl-CoA-reductase genes from perennial ryegrass (*Lolium perenne*). *Plant Cell* 22, 3357–3373. doi: 10.1105/tpc.109.072827

Utrilla-Vázquez, M., Rodríguez-Campos, J., Avendaño-Arazate, C. H., Gschaedler, A., and Lugo-Cervantes, E. (2020). Analysis of volatile compounds of five varieties of Maya cocoa during fermentation and drying processes by Venn diagram and PCA. *Food Res. Int. Ott. Ont.* 129, 108834. doi: 10.1016/j.foodres.2019.108834

Wang, X., Fan, W., and Xu, Y. (2014). Comparison on aroma compounds in Chinese soy sauce and strong aroma type liquors by gas chromatography–olfactometry, chemical quantitative and odor activity values analysis. *Eur. Food Res. Technol.* 239, 813–825. doi: 10.1007/s00217-014-2275-z

Wyrambik, D., and Grisebach, H. (1975). Purification and properties of isoenzymes of cinnamyl-alcohol dehydrogenase from soybean-cell-suspension cultures. *Eur. J. Biochem.* 59, 9–15. doi: 10.1111/j.1432-1033.1975.tb02418.x

Yin, L. (2020). *CMplot: Circle Manhattan Plot*. Available online at: https://github.com/YinLiLin/CMplot

Ying, H., and Qingping, Z. (2006). Genetic manipulation on biosynthesis of terpenoids. *J Chin. Biotechnol.* 26, 60–64.

Zhang, Y.-M., Jia, Z., and Dunwell, J. M. (2019). Editorial: The applications of new multi-locus gwas methodologies in the genetic dissection of complex traits. *Front. Plant Sci.* 10:100. doi: 10.3389/fpls.2019.00100

Ziegleder, G. (1990). Linalool contents as characteristic of some flavor grade cocoas. *Z. Lebensm. Unters. Forsch.* 191, 306–309. doi: 10.1007/BF01202432

Ziegleder, G. (2009). "Flavour development in cocoa and chocolate," *Industrial Chocolate Manufacture and Use, 4th Edn.,* ed S. T. Beckett (Oxford: Wiley-Blackwell), 169–191. doi: 10.1002/9781444301588.ch8

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 25 October 2021
doi: 10.3389/fpls.2021.728891

Comparison of the Fruit Volatile Profiles of Five Muscadine Grape Cultivars (*Vitis rotundifolia* Michx.) Using HS-SPME-GC/MS Combined With Multivariate Statistical Analysis

Honghong Deng, Runmei He, Meicun Long, Yanmei Li, Yuanyuan Zheng, Lijin Lin, Dong Liang, Xiaoai Zhang, Ming'an Liao, Xiulan Lv, Qunxian Deng and Hui Xia**

Institute of Pomology and Olericulture, College of Horticulture, Sichuan Agricultural University, Chengdu, China

OPEN ACCESS

Edited by:
Jinhe Bai,
Horticultural Research Laboratory (USDA-ARS), United States

Reviewed by:
Liping Du,
Tianjin University of Science and Technology, China
Bo Zhang,
Zhejiang University, China

***Correspondence:**
Qunxian Deng
1324856299@qq.com
Hui Xia
susanxia_2001@163.com

Specialty section:
This article was submitted to Plant Metabolism and Chemodiversity, a section of the journal Frontiers in Plant Science

Received: *22 June 2021*
Accepted: *22 September 2021*
Published: *25 October 2021*

Citation:
Deng H, He R, Long M, Li Y, Zheng Y, Lin L, Liang D, Zhang X, Liao M, Lv X, Deng Q and Xia H (2021) Comparison of the Fruit Volatile Profiles of Five Muscadine Grape Cultivars (Vitis rotundifolia Michx.) Using HS-SPME-GC/MS Combined With Multivariate Statistical Analysis. Front. Plant Sci. 12:728891. doi: 10.3389/fpls.2021.728891

Fruit aromas are composed of a complex mixture of volatile organic compounds, which are essential attributes associated with the overall flavor and consumer preference. Muscadine grape (MG; *Vitis rotundifolia* Michx.) is an aroma-dense fruit crop. However, there is very scarce information on its volatile profiles. In this study, the volatile constituents of five newly introduced MG cultivars, including Alachua, Carlos, Fry, Granny Val, and Noble, were profiled using headspace solid-phase microextraction-gas chromatography-mass spectrometry (HS-SPME-GC/MS) combined with multivariate statistical analysis. A total of 44 compounds, including esters, aldehydes, alcohols, fatty acids, terpenes, ketones, and furan, were identified and relatively quantified. Principal component analysis (PCA) and partial least-squares discriminant analysis (PLS-DA) evidently discriminated against the five MG cultivars based on their volatile profiles. The specific volatiles that contributed the most to this discrimination were highlighted. Geraniol and cinnamyl alcohol were demonstrated to be essential for characterizing the Alachua MG cultivar, whereas ethyl trans-2-butenoate and propyl acetate were shown to be important compounds to characterize the Noble MG cultivar. The results further showed that 2-Ethyl-1-hexanol, (*Z*)-3-hexenal, and (*E*)-2-hexenol were closely related to Carlos, Fry, and Granny Val cultivars, respectively. This investigation is the first in-depth exploration of the volatile profiles of the aroma-dense muscadine grape, which is essential for future genetic or biotechnological improvements to attain a cultivar with the desired flavor.

Keywords: volatile organic compounds, headspace solid-phase microextraction, gas chromatography-mass spectrometry, principal component analysis, partial least-squares discriminate analysis

INTRODUCTION

Grapes (*Vitis vinifera* L.) are among the most economically important and the earliest domesticated fruit crops throughout the world (Reisch et al., 2012; Ramos-Madrigal et al., 2019). Similar to fresh fruit, grapes are an essential nutrient-dense food source, being part of the human diet, while as processed products, i.e., dried raisins, fruit preserves, primarily wine, and spirits, they have added

economic values and represent a globally traded commodity with significant economic value (Reisch et al., 2012; Ramos-Madrigal et al., 2019). Most of the grapes cultivated for fruit production are either of the species *V. vinifera*, hybrids of *V. vinifera*, belonging to the genus *Vitis* L., or, to a lower extent, to the closely related subgenus *Muscadinia* (Reisch et al., 2012; Lin et al., 2019).

Taxonomically, the muscadine grape (MG) (*Muscadinia rotundifolia* Michx., syn. *V. rotundifolia* Michx.) is closely related to *Vitis* species (Olien, 1990; Liu et al., 2016; Wen et al., 2018). Genetically, the basic chromosome of MG (x = 20) is interestingly one chromosome greater than the 19 basic chromosomes of the *V. vinifera* L. (Olien, 1990). Indigenous to the warm and humid climate of the southeastern United States, MG was the first American grape species to be cultivated (Olien, 1990) and has been cultivated for more than 400 years (Stanley, 1997; Conner, 2009). It can be found growing naturally from Delaware to Central Florida state, i.e., south-north distribution, and from Texas and Oklahoma to the eastern coastal plain, i.e., east-west span, in the United States, along the Mississippi River to Missouri and near the Appalachian Mountains from the east and west (Olien, 1990).

Muscadine grape has relatively high intrinsic resistance to disease, including downy and powdery mildews, gray mold, and anthracnose, especially Pierce's disease (*Xylella fasti diosa*) (Olien, 1990). Over the past few decades, MG has drawn increasing attention from consumers, growers, and breeders because it constitutes an important source of dietary fibers, amino acids, minerals, vitamins, antioxidant phytochemicals, such as polyphenolic compounds, and other health-promoting compounds (Sandhu and Gu, 2010; Wang et al., 2010). Since the resurgence of the interest in MG in the United States to date, more than 100 MG cultivars have been released (Hoffmann et al., 2020), and MG has also been extended to California, Chile, and China (Hoffmann et al., 2020). For example, five MG cultivars, including Alachua, Carlos, Fry, Granny Val, and Noble, have been successfully introduced and cultivated in southern China (Wei et al., 2017; Ren et al., 2020).

In addition to having significantly more phytonutrients than the average table grape cultivars (Marshall et al., 2012), MG has a distinct musky aroma and unique flavor (Kambiranda et al., 2016), which is essential for its current wine, juice, and table grape markets. Volatile organic compounds (VOCs), such as esters, ketones, terpenes, aldehydes, alcohols, C_{13} norisoprenoids, and benzenoids, are significant determinants of the grape aroma and perceived flavor (Gürbüz et al., 2014; Lin et al., 2019). The VOCs of fruits are complex and vary significantly depending on the analytical methods used and a plethora of factors, such as the cultivar, sample type, ripening stage, environmental conditions, abiotic and biotic stress (Sánchez-Palomo et al., 2005; Kalua and Boss, 2009; Wu et al., 2020). VOCs are the main contributors to the fresh and fruity note of wines (Perez-Coello and Diaz-Maroto, 2009) and are important nutritional constituents, shaping the sensory properties of foods and influencing the perception of consumers (Berger, 2009). Each fruit species has a specific combination of various VOCs with different concentration and perception thresholds (El Hadi et al., 2013). The differences of VOCs between the table and wine grapes have been well investigated, whereas only a few studies have reported the VOC compositions of MG (Lee et al., 2016).

Headspace solid-phase microextraction (HS-SPME) is a passive sampling approach that collects compounds present in the vapor phase above samples, which does not interfere with the samples and omits the compounds in the vapor phase above the matrix during heating (Lubes and Goodarzi, 2017). Compared with solvent extraction, HS-SPME is rapid, simple, and reproducible, and yields more realistic VOC discrimination results (Lubes and Goodarzi, 2017). HS-SPME coupled with gas chromatography-mass spectrometry (GC-MS) has been extensively employed to study the volatile profiles of many fruit varieties, vegetables, and beverages.

Muscadine grape is an aroma-dense fruit. However, relatively few publications have documented its aromatic VOC compositions. Therefore, a systematic, qualitative, and quantitative investigation of the MG VOCs is highly required and of great importance. In this study, HS-SPME-GC-MS was employed to compare, both qualitatively and quantitatively, the volatile constituents of five MG cultivars that were newly introduced in China. Multivariate statistical analysis methods, including principal component analysis (PCA), and partial least-squares discriminate analysis (PLS-DA), were utilized to highlight the differences among cultivars and identify the chemical biomarkers discriminating the five MG cultivars. Determining the VOCs of MG and exploring the differences in VOCs among cultivars is essential for future genetic or biotechnological improvements targeting aroma-dense grape cultivars.

MATERIALS AND METHODS

Fruit Materials

This study was conducted using 5-year-old (in 2020) newly introduced MG cultivars (Alachua, Carlos, Fry, Granny Val, and Noble). Trees were grown in the fields of the experimental station of the College of Horticulture, Sichuan Agricultural University (30°42′N, 103°51′E), China. In this area, all vines were subjected to identical and standard viticultural practices for table grape cultivars throughout the experiment, including winter pruning, pest and pathogen control, basal fertilizer, and irrigation.

For each cultivar, fruit samples (**Figure 1**) were randomly collected from three clusters of four different plants in September (the maturation stage), which is considered ideal for commercialization in Sichuan based on the observations during the previous years. Specifically, the solid soluble content of Alachua, Carlos, Fry, Granny Val, and Noble samples ranged from 16.28 to 17.7°Brix, from 14.2 to 16.18°Brix, from 14.62 to 16.3°Brix, from 14.9 to 18.03°Brix, and from 14.7 to 15.38°Brix, respectively. Each cluster was clipped off at the end of the peduncle. Samples were stored in a cold chamber and transported to the laboratory within 2 h. Three berries from the top, central, and bottom parts of each cluster were pooled together. A total of 36 berries were picked and randomly divided into three biological replicates. A 0.5 cm thickness slice was obtained

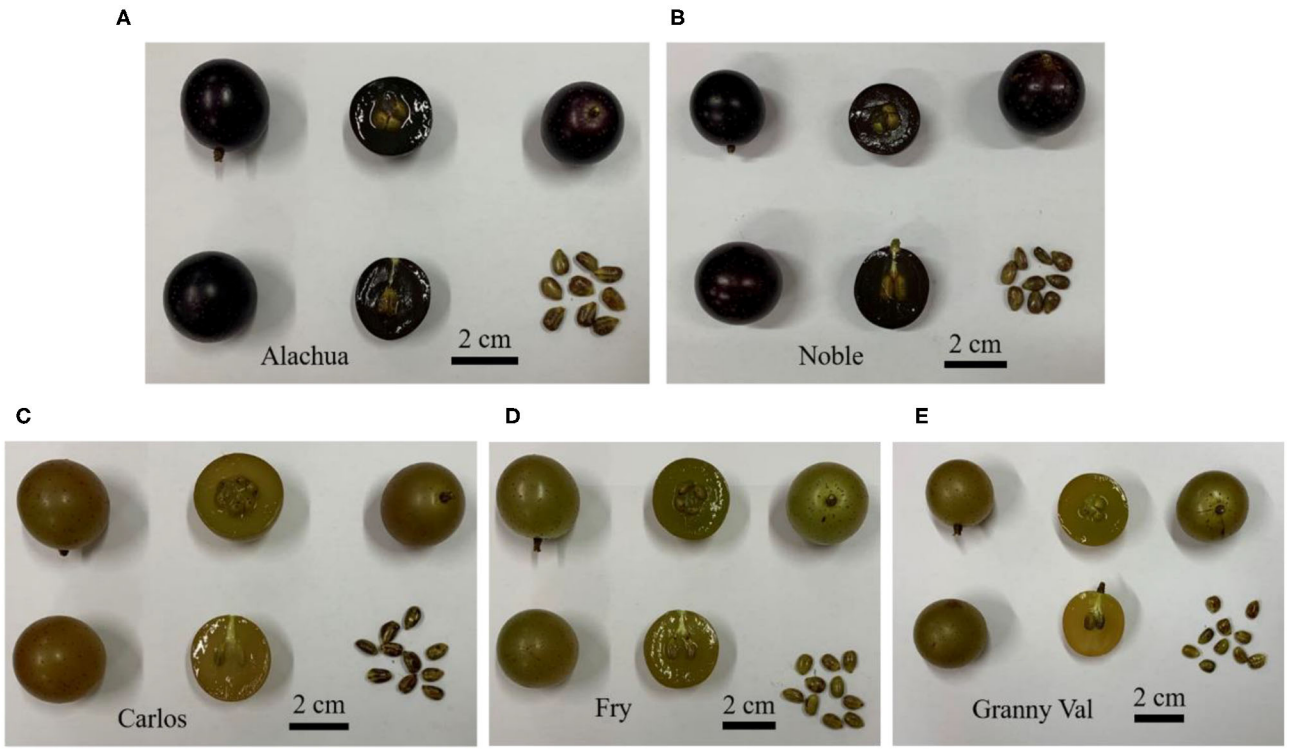

FIGURE 1 | Fruit phenotypes of the Alachua **(A)**, Noble **(B)**, Carlos **(C)**, Fry **(D)**, and Granny Val **(E)** muscadine grape cultivars. Scale bars denote 2 cm.

from the equatorial region of each fruit, immediately frozen in liquid nitrogen, and stored at −80°C until the further determination of VOCs.

Chemicals and Solvents

The study used *n*-Alkane (C_7-C_{30}) standards and the available authentic standards, including ethyl acetate (≥99.5%), propyl acetate (≥99.5%), butyl acetate (99.7%), ethyl trans-2-butenoate, ethyl hexanoate (≥99%), ethyl heptanoate (≥99%), ethyl octanoate (≥99%), ethyl 3-hydroxybutyrate (≥ 98%), hexyl hexanoate (≥ 98%), hexanal (≥95%), (*Z*)-3-hexenal (50% in triacetin), (*E*)-2-hexenal (≥97%), nonanal (≥99.5%), benzaldehyde (≥99.5%), (*E*)-2-decenal (≥95%), citral (≥95%), 1-butanol (≥99.4%), 1-hexanol (≥99.9%), (*E*)-2-hexenol (96%), 1-octen-3-ol (≥98%), 1-heptanol (≥99.5%), 2-ethylhexanol (≥99%), 1-octanol (≥99%), 2-octen-1-ol (97%), (*Z*)-5-octen-1-ol (≥97%), (*E*)-5-decen-1-ol (≥97%), phenylethyl alcohol (≥99%), cinnamyl alcohol (≥96%), acetic acid (≥99.7%), hexanoic acid (≥98%), heptanoic acid (≥99%), octanoic acid (99%), nonanoic acid (≥99.5%), limonene (mixture of D- and L-form at ratio of 1:1, ≥95%), linalool (≥99%), citronellol (≥95%), nerol (≥97%), geraniol (≥98.5%), 2-octanone (≥99.5%), acetophenone (≥99.5%), 2-pentylfuran (≥97%), 2-octanol (≥97%), and sodium chloride (NaCl, ≥99%) which were all purchased from Sigma-Aldrich (St. Louis, MO, USA). Geranic acid (sum of isomers, 98%) was purchased from Alfa Aesar Corporation (Tianjin, China). Ultrapure water was prepared using a Milli-Q water purification system (Millipore Corporation, Bedford, MA, USA) with a 0.22 μm filter.

Fruit Sample Preparation for HS-SPME

The VOCs from the whole fruit were extracted using HS-SPME. Fruit samples, including the seeds, flesh, and skins, were pooled together and fully ground into a fine powder in liquid nitrogen. For each extraction sample, 100 mg of powder, spiked with 10 μl of 2-octanol, was accurately weighed and transferred to a 20-ml glass sample container (Thermo Scientific, Bellefonte, PA, USA). Samples were overlaid with a 5 ml saturated sodium chloride (NaCl) solution to inhibit enzyme degradation. A tiny stirring bar was added to facilitate VOC release before the glass vial was capped. The homogenized samples were incubated for 15 min in a 60°C water bath with continuous agitation (125 rpm). Thereafter, the VOCs were collected using a 2 cm DVB/CAR/PMDS SPME fiber (50/30 μm, Supelco Inc., Bellefonte, PA, USA) by exposing the fiber to the headspace for another 30 min under the same conditions. The fibers were activated before sampling according to the instructions of the manufacturer. After this incubation step, the SPME fiber was inserted directly into the injection port of the GC system for thermal desorption (4 min at 250°C) in a splitless mode.

GC-MS Operating Conditions

Volatile organic compounds were analyzed using an Agilent 7890 gas chromatography system equipped with a 5977B mass spectrometer (Agilent Technologies Inc., Santa Clara, CA, USA).

TABLE 1 | The chemical compositions and relative concentrations of fruit volatiles of five muscadine grape cultivars.

Compound names	RI	Alachua	Carlos	Fry	Granny Val	Noble
Esters						
Ethyl Acetate	982	1.95 ± 0.38a	3.92 ± 0.67a	3.15 ± 0.95a	2.61 ± 0.51a	3.91 ± 1.56a
Propyl acetate	952	0.03 ± 0a	0.03 ± 0.01a	0.03 ± 0.01a	0.03 ± 0a	0.09 ± 0.04a
Butyl acetate	1,060	0.54 ± 0.12a	0.93 ± 0.35a	0.49 ± 0.18a	1.05 ± 0.27a	1.78 ± 0.7a
Ethyl trans-2-butenoate	1,158	2.55 ± 0.73ab	0.67 ± 0.46b	1.36 ± 0.19b	2.24 ± 0.8ab	4.3 ± 0.33a
Ethyl hexanoate	1,241	1.19 ± 0.29a	0.64 ± 0.06a	0.77 ± 0.14a	0.5 ± 0.02a	0.53 ± 0.11a
Ethyl heptanoate	1,331	0.05 ± 0.02a	0.04 ± 0a	0.04 ± 0.01a	0.03 ± 0.01a	0.02 ± 0a
Ethyl octanoate	1,412	0.24 ± 0.09ab	0.37 ± 0.09ab	0.55 ± 0.08a	0.15 ± 0.05b	0.05 ± 0.01b
Ethyl 3-hydroxybutyrate	1,482	0.16 ± 0.03c	0.74 ± 0.09b	0.79 ± 0.07b	1.31 ± 0.18a	0.21 ± 0.06c
Hexyl hexanoate	1,599	0.08 ± 0.05a	0.02 ± 0a	0.02 ± 0a	0.01 ± 0a	0.01 ± 0a
Subtotal		6.78 ± 0.24a	7.37 ± 2.65a	7.21 ± 2.45a	7.94 ± 1.19a	10.90 ± 4.57a
Aldehydes						
Hexanal	1,072	2.67 ± 0.3a	2.82 ± 0.67a	2.04 ± 0.62a	2.17 ± 0.24a	1.44 ± 0.09a
(Z)-3-Hexenal	1,124	0.1 ± 0.01a	0.14 ± 0.02a	0.12 ± 0.04a	0.07 ± 0.01a	0.04 ± 0a
(E)-2-Hexenal	1,215	6.4 ± 0.31abc	8.09 ± 1.21ab	9.19 ± 1.26a	4.87 ± 0.72bc	3.04 ± 0.33c
(Z)-2-Heptenal	1,319	0.15 ± 0.01a	0.1 ± 0.03ab	0.09 ± 0ab	0.09 ± 0.01ab	0.07 ± 0.02b
Nonanal	1,392	0.07 ± 0.01b	0.17 ± 0.02a	0.15 ± 0.01a	0.14 ± 0.01a	0.05 ± 0b
Benzaldehyde	1,520	0.17 ± 0.01bc	0.64 ± 0.15a	0.5 ± 0.08abc	0.52 ± 0.07ab	0.13 ± 0.03c
Benzeneacetaldehyde	1,630	1.07 ± 0.33a	2.14 ± 0.31a	1.8 ± 0.05a	2.39 ± 0.37a	0.95 ± 0.39a
(E)-2-Decenal	1,634	0.01 ± 0bc	0.02 ± 0ab	0.03 ± 0a	0.01 ± 0bc	0 ± 0c
Citral	1,717	0.76 ± 0.02a	0.12 ± 0.03a	0.64 ± 0.16a	0.27 ± 0.13a	1.01 ± 0.38a
Subtotal		14.41 ± 1.32ab	14.24 ± 3.66a	14.57 ± 3.28a	10.53 ± 2.31ab	6.73 ± 1.82b
Alcohols						
1-Butanol	1,138	0.48 ± 0.04a	0.3 ± 0.05a	0.52 ± 0.03a	0.3 ± 0.01a	0.65 ± 0.16a
1-Hexanol	1,034	1.45 ± 0.37b	3.95 ± 0.22ab	5.08 ± 0.15ab	6.53 ± 1.88a	2.58 ± 0.28ab
(E)-2-Hexenol	1,394	0.4 ± 0.09a	0.8 ± 0.15a	1.28 ± 0.24a	3.66 ± 1.98a	0.79 ± 0.16a
1-Octen-3-ol	1,420	0.32 ± 0.03a	0.51 ± 0.14a	0.33 ± 0.02a	0.49 ± 0.03a	0.26 ± 0.05a
1-Heptanol	1,454	0.06 ± 0.01b	0.23 ± 0.05a	0.15 ± 0ab	0.13 ± 0.03ab	0.05 ± 0.01b
2-Ethylhexanol	1,499	0.16 ± 0.02a	0.18 ± 0.01a	0.18 ± 0.03a	0.19 ± 0.01a	0.15 ± 0.01a
1-Octanol	1,548	0.92 ± 0.09c	3.45 ± 0.83ab	4.01 ± 0.37a	1.28 ± 0.35bc	1.27 ± 0.42bc
(E)-2-Octen-1-ol	1,620	0.14 ± 0.01ab	0.16 ± 0.03ab	0.21 ± 0.01a	0.14 ± 0ab	0.1 ± 0.02b
(Z)-5-Octen-1-ol	1,626	0.1 ± 0b	0.18 ± 0.02b	0.48 ± 0.01a	0.11 ± 0.04b	0.11 ± 0.03b
(E)-5-Decen-1-ol	1,767	1.48 ± 0.16b	5.22 ± 0.95ab	7.62 ± 0.88a	4.35 ± 1.41ab	2.03 ± 0.53b
Phenylethyl alcohol	1,902	0.22 ± 0.04c	4.89 ± 0.81abc	9.25 ± 1.02a	5.51 ± 2.12ab	0.34 ± 0.09bc
Cinnamyl alcohol	2,252	0.06 ± 0.02a	0.01 ± 0b	0.01 ± 0b	0 ± 0b	0 ± 0b
Subtotal		5.79 ± 0.74b	19.88 ± 3.63ab	29.11 ± 4.15a	22.69 ± 13.19ab	8.34 ± 1.86b
Acids						
Acetic acid	1,445	0.17 ± 0.02a	0.11 ± 0.01a	0.09 ± 0.03a	0.12 ± 0.04a	0.13 ± 0.08a
Hexanoic acid	1,840	2.6 ± 0.15a	2.31 ± 0.18a	2.92 ± 0.74a	4.15 ± 2.03a	2.3 ± 0.17a
Heptanoic acid	1,962	0.16 ± 0.01a	0.1 ± 0.03a	0.12 ± 0.03a	0.18 ± 0.11a	0.11 ± 0.01a
Octanoic acid	2,034	0.23 ± 0.04ab	0.18 ± 0.02ab	0.26 ± 0.02a	0.18 ± 0.02ab	0.15 ± 0.01b
Nonanoic acid	2,174	1.1 ± 0.12ab	0.93 ± 0.05ab	1.52 ± 0.14a	0.68 ± 0.34b	0.98 ± 0.06ab
Geranic acid	2,287	17.89 ± 1.32a	0.89 ± 0.2b	2.05 ± 0.09b	0.63 ± 0.16b	13.36 ± 3.14a
Subtotal		4.25 ± 0.51a	3.63 ± 0.39a	4.90 ± 1.27a	5.33 ± 4.20a	3.67 ± 0.10a
Terpenes						
D-Limonene	1,201	0.09 ± 0.01a	0.01 ± 0b	0.02 ± 0b	0.01 ± 0b	0.09 ± 0.02a
Linalool	1,553	1.04 ± 0.11a	0.03 ± 0.01b	0.14 ± 0.02b	0.03 ± 0b	1.08 ± 0.37a
Citronellol	1,755	1.17 ± 0.16a	0.03 ± 0.01c	0.14 ± 0.01bc	0.03 ± 0.01c	0.86 ± 0.33ab
Nerol	1,767	0.09 ± 0a	0.01 ± 0b	0.03 ± 0b	0 ± 0b	0.1 ± 0.03a
Geraniol	1,830	11.07 ± 3.16a	2.44 ± 0.41b	9.29 ± 0.49ab	1.96 ± 0.33b	9.55 ± 2.06ab
Subtotal		13.46 ± 5.80a	2.52 ± 0.73b	9.61 ± 0.88ab	2.03 ± 0.58b	11.68 ± 2.92a

(Continued)

TABLE 1 | Continued

Compound names	RI	Alachua	Carlos	Fry	Granny Val	Noble
Ketones						
2-Octanone	1,275	4.82 ± 0.02a	5.27 ± 0.13a	4.73 ± 0.75a	5.18 ± 0.14a	4.8 ± 0.1a
Acetophenone	1,645	0.1 ± 0.03a	0.06 ± 0ab	0.08 ± 0ab	0.04 ± 0ab	0.02 ± 0b
Subtotal		4.92 ± 0.02a	5.33 ± 0.24a	4.80 ± 1.30a	5.22 ± 0.25a	4.82 ± 0.18a
Furan						
2-Pentylfuran	1,229	0.08 ± 0c	0.11 ± 0bc	0.33 ± 0.02a	0.18 ± 0.02b	0.11 ± 0.03bc
Subtotal		0.08 ± 0c	0.11 ± 0bc	0.33 ± 0.02a	0.18 ± 0.02b	0.11 ± 0.03bc
Total		64.60 ± 5.90a	53.98 ± 4.87a	72.58 ± 7.49a	54.50 ± 20.2a	59.61 ± 1.34a

Data are expressed as M ± SD from three technologic replicates. 1 abundance unit is equal to 100 μg/g fresh weight. Numbers in the same row followed by the same letter within each cultivar are not significantly different by Tukey's HSD test (95% confidence).

Helium (99.999% purity) was used as the carrier gas with a front inlet purge flow rate of 3 ml min^{-1} and a constant gas flow rate through the column of 1 ml min^{-1}. VOC separation was carried out using an Agilent DB-Wax (30 m × 250 μm ×0.25 μm, Agilent Technologies Inc., Santa Clara, CA, USA) column. The oven temperature program was initiated at 40°C for 4 min, then ramped up to 245°C at a rate of 5 °C min^{-1}, and held for 5 min. The transfer line, ion source, and quadrupole mass detector temperature values were set to 250, 230, and 150°C, respectively. Mass spectra in electron impact ionization (ME-EI) mode were recorded at ionization energy of 70 eV. Data acquisition was performed using the mass spectrometer scanning mode from m/z 20 to 500. The solvent delay time was 0 min.

VOC Identification and Quantification

The row peak obtained from GC-MS was first processed using Chroma TOF 4.3X software (LECO Corporation, St Joseph, MI, USA). The parameters used for raw peak extraction, data baseline filtering and calibration of the baseline, peak alignment, deconvolution analysis, peak identification, integration, and spectrum match of the peak area were the same for all samples. VOCs were identified by matching the retention time in conjunction with the MS fragments with the data of previous studies using a similar chromatographic column and the built-in commercial MS databases, such as the NIST reference library (National Institute of Standards and Technology, Gaithersburg, MA, USA). The VOCs were checked against authentic standards when they were accessible. The VOC concentrations were obtained *via* peak normalization and semi-quantified to the internal reference standard in the same GC-MS run, which was added at the beginning of the VOC extraction step, as mentioned above. Correspondingly, the peak area of each VOC was converted into a relative concentration value for the following statistical analyses.

Statistical Analyses

The normalized data were exported to R free software v. 3.2.3 (R Core Team, 2020) for statistical analysis. The means and SEs were determined for all the detected variables. Significant variances were validated using one-way ANOVA and Tukey's honestly significant difference (HSD) test ($P < 0.05$). The data were then imported into SIMCA software version 14.1 (Umetrics, Umea, Sweden) for multivariate statistical analyses. PCA was first conducted to visualize the main correlations in the whole data matrix, followed by PLS-DA to discriminate the MG varieties further. A permutation test (100 times) was applied to validate the PLS-DA results and to avoid overfitting (Saccenti, 2012). Afterward, the variable importance in projection (VIP) was used to define which VOCs significantly contribute to discriminate the five MG cultivars (Chong and Jun, 2005).

RESULTS AND DISCUSSION

Characterization of the Fruit VOCs of the Five MG Cultivars

Headspace-SPME combined with GC-MS analysis and the comparison of mass spectra with the NIST17 library and the available authentic standards resulted in the identification of 44 metabolites in the five MG cultivars, whose chromatographic profiles are shown in **Supplementary Figure 1**. These metabolites include eight esters, ten aldehydes, twelve alcohols, six fatty acids, five terpenes, two ketones, and one furan based on their chemical nature (**Table 1**). We found that this chemical classification agreed with the major constituents for grape volatiles that have been reported so far (Lee et al., 2016; Mencarelli and Bellincontro, 2018; Lin et al., 2019; Wu et al., 2020; Golombek et al., 2021; Ju et al., 2021).

The level of each VOC was evaluated using its relative area toward the internal standard. Geranic acid, geraniol, 2-octanone, and (*E*)-2-hexenal components were found to be with the greatest concentration values in the Alachua and Noble MG cultivars. The main components quantified in the Carlos, Fry, and Granny Val MG cultivars were (*E*)-5-decen-1-ol, phenylethyl alcohol, 1-hexanol, 2-octanone, and (*E*)-2-hexenal (**Table 1**). A log transformation was performed to allow the in-depth analysis of all the VOCs detected (**Figure 2**). One-way ANOVA combined with Tukey's HSD test ($P < 0.05$) was used to estimate the significant differences in VOC content among the five MG cultivars, resulting in 29 significantly different ($P < 0.05$) metabolites (**Figure 2**).

The cultivar (Ju et al., 2021), cultural practices (Golombek et al., 2021), and postharvest biological control (Mencarelli and Bellincontro, 2018) were the most influential factors affecting

FIGURE 2 | Continued

the volatile compositions and the production of grapes. In the present study, all vines were grown under the same conditions using the same horticultural practices. In addition, strictly identical extraction conditions and analytical parameters were used for all samples. The influence of the environmental factors and technical parameters on the composition and production

FIGURE 2 | Box and whisker plots showing the significantly different volatiles identified in the five muscadine grape cultivars. Vertical bars represent the normalized relative concentration of volatiles, and the *x*-axis depicts the different muscadine grape cultivars. A, Alachua; C, Carlos; F, Fry; G, Granny Val; N, Noble.

of volatiles was negligible. Therefore, by bringing together these extensive factors, we demonstrated in our study that the main source of variance is closely related to the cultivar. The predominant compounds contributing to the MG volatile profiles agree with those found previously, among which the hexenal, (*E*)-2-hexenal, and 1-hexanol were reported as the major constituents

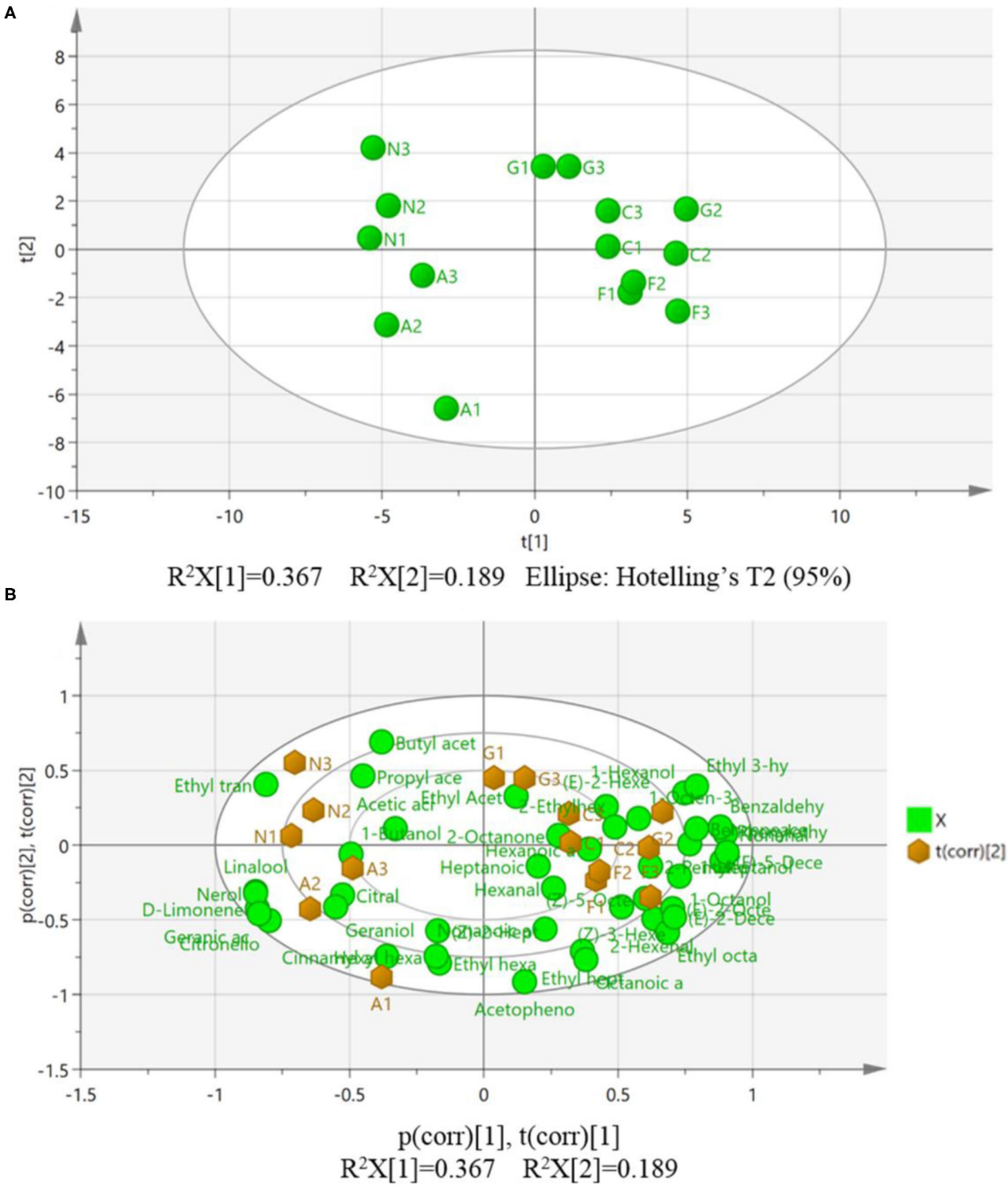

FIGURE 3 | Principal component analysis (PCA) showing the main differences in the fruit volatiles of the five muscadine grape cultivars. **(A)** PCA score plot of all the investigated samples; **(B)** PCA biplot of the variables associated with 15 investigated samples. A, Alachua; C, Carlos; F, Fry; G, Granny Val; N, Noble; t[1], the first principal component score; t[2], the second principal component score.

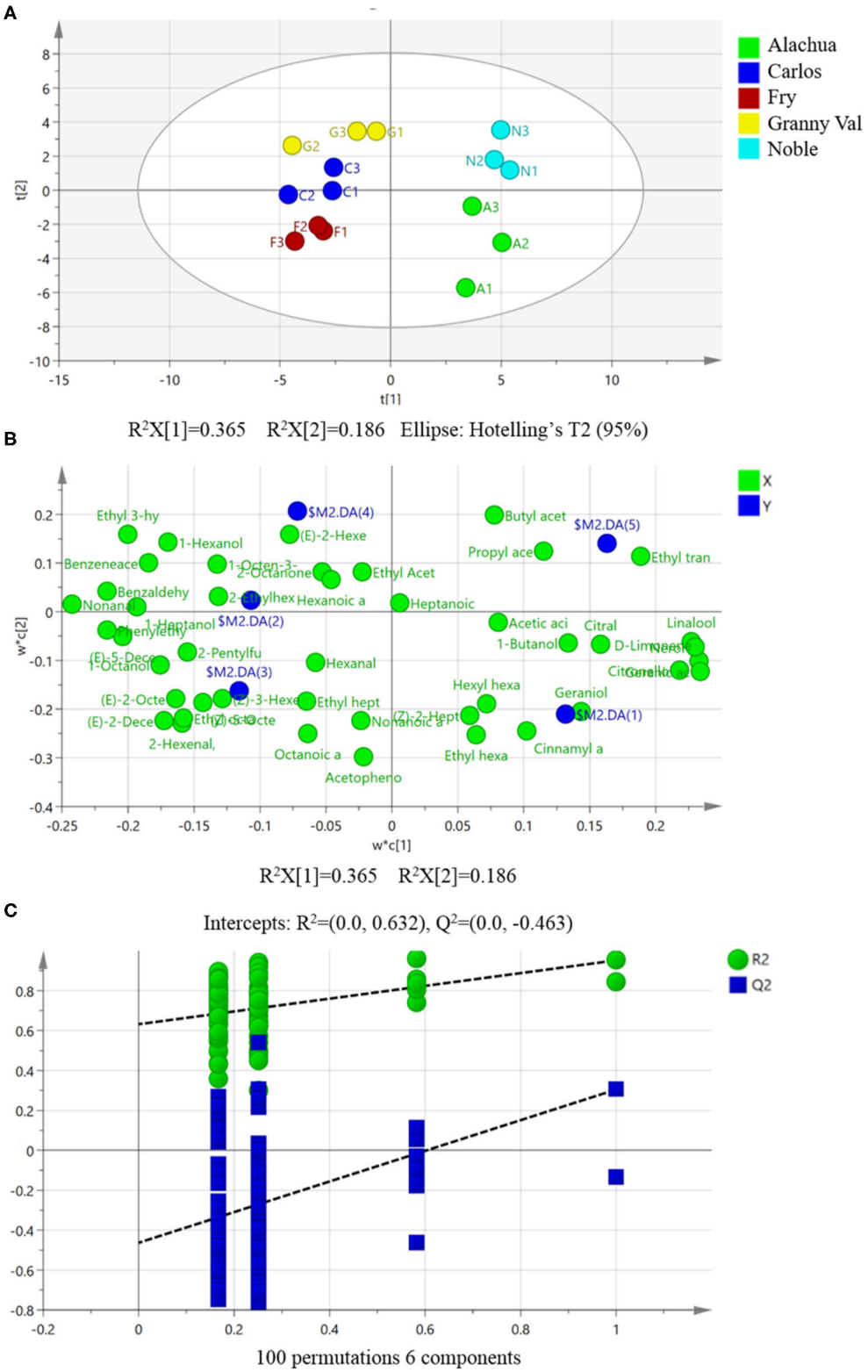

FIGURE 4 | Partial least-squares discriminate analysis (PLS-DA) showing a better separation among the five muscadine grape cultivars based on the fruit volatiles. **(A)** PLS-DA score plot of all the investigated samples; **(B)** loadings plot of variables associated with different muscadine grape cultivars; **(C)** permutation test for PLS-DA model validation. A, Alachua; C, Carlos; F, Fry; G, Granny Val; N, Noble; t[1], the first principal component score; t[2], the second principal component score.

of six clones of spine grape berries (Ju et al., 2021), and geraniol and ethyl acetate were reported as the main constituents of Cowart MG berries (Lee et al., 2016). The (*E*)-2-hexenal and 1-hexanol identified here have been reported as the dominant C_6 volatile compounds in *V. vinifera* cultivars. However, there was a wide variation in the concentration and percentage of the C_6 volatiles contributing to the total volatiles (Yang et al., 2009). This finding indicated that the composition and concentration of VOCs in MG cultivars could vary with the genetic background.

Multivariate PCA and PLS-DA Analyses of the HS-SPME-GC-MS Data

Multivariate analysis of the dataset using PCA was primarily performed to visualize the overall differentiation and intrinsic variation of VOCs among the five MG cultivars (**Figure 3A**). PCA is an unsupervised chemometric method that reduces dimensionality and visualizes the main correlations and variability of a complex dataset (Vidal et al., 2016). In the PCA score scatter plot, the cultivars Carlos, Fry, and Granny Val were located at the right side of the score plot towards the center, while the distribution of samples from the center to the left side of the score plot comprised the cultivars Alachua and Noble. A clear differentiation among the five MG cultivars can be observed, in particular, the bronze-colored cultivars are distributed far away from the purple-black colored cultivars, as depicted in **Figure 3A**. All the investigated samples were located within the 95% confidence interval, which indicated that no outliers existed in the whole dataset. The accumulated variance contribution rate R^2X reached 0.667, with the first and second principal components (PCs) carrying data variance of 36.7 and 18.9%, respectively. The Q^2 value of the PCA model was 0.254 (**Figure 3A**).

Since in a PCA model, the directions in the score plot correspond to the directions in the loading plot, a comparison of these two plots can be used to identify which variables (loadings) have the greatest influence on the class separation of the different samples (scores) (Lubes and Goodarzi, 2017). Based on these criteria, the individual volatiles attributed to the variation in the five MG cultivars were graphically illustrated in a biplot (**Figure 3B**).

Partial least-squares discriminant analysis is a supervised multivariate statistical analysis and a variant of the PLS regression approach, which is widely used to construct a multidimensional model to predict features, discriminating between different samples, and further potential biomarkers explorations (Kalivodová et al., 2015). Therefore, PLS-DA was subsequently applied to perform an even better separation of the MG cultivars. Similarly, the score scatters plot of PLS-DA showed that the clusters of bronze-colored cultivars were located far from the purple black-color cultivars. Compared with the PCA model (**Figure 3A**), a noticeable improvement in the distinction of the five MG cultivars was observed (**Figure 4A**). All data points were within the 95% confidence interval (**Figure 4A**) and consistent with the results of the PCA model.

The PLS-DA model revealed that the corresponding values of $R^2X(cum)$, $R^2Y(cum)$, and $Q^2(cum)$ were 0.823, 0.927, and 0.503, respectively (**Figure 4A**). In the PLS-DA model, the R^2X and R^2Y values were utilized to describe the total explained variation in X and Y, respectively, and were represented by the PCs. At the same time, the Q^2 parameter was used to assess

TABLE 2 | The variable importance in projection (VIP) scores within the partial least-squares discriminate analysis (PLS-DA) model.

Variable ID (Primary)	Variable importance in projection (VIP) scores
2-Pentylfuran	1.41
(Z)-5-Octen-1-ol	1.28
Cinnamyl alcohol	1.24
1-Heptanol	1.18
1-Octanol	1.17
(E)-2-Hexenol	1.16
Ethyl 3-hydroxybutyrate	1.14
Nonanoic acid	1.12
Citral	1.11
(Z)-2-Heptenal	1.09
Octanoic acid	1.08
Acetophenone	1.07
Hexanoic acid	1.03
Geraniol	1.03
Ethyl trans-2-butenoate	1.03
1-Butanol	1.02
(E)-2-Decenal	1.01
1-Hexanol	1.01
(E)-2-Hexenal	0.99
Ethyl hexanoate	0.99
(Z)-3-Hexenal	0.99
Nerol	0.97
Heptanoic acid	0.97
Phenylethyl alcohol	0.95
Ethyl octanoate	0.94
1-Octen-3-ol	0.94
Hexanal	0.94
D-Limonene	0.93
Ethyl Acetate	0.93
(E)-5-Decen-1-ol	0.93
Geranic acid	0.92
Citronellol	0.91
Linalool	0.90
Benzeneacetaldehyde	0.89
Ethyl heptanoate	0.89
Hexyl hexanoate	0.89
(E)-2-Octen-1-ol	0.88
Nonanal	0.88
Butyl acetate	0.88
Propyl acetate	0.85
Benzaldehyde	0.85
2-Octanone	0.82
2-Ethylhexanol	0.70
Acetic acid	0.66

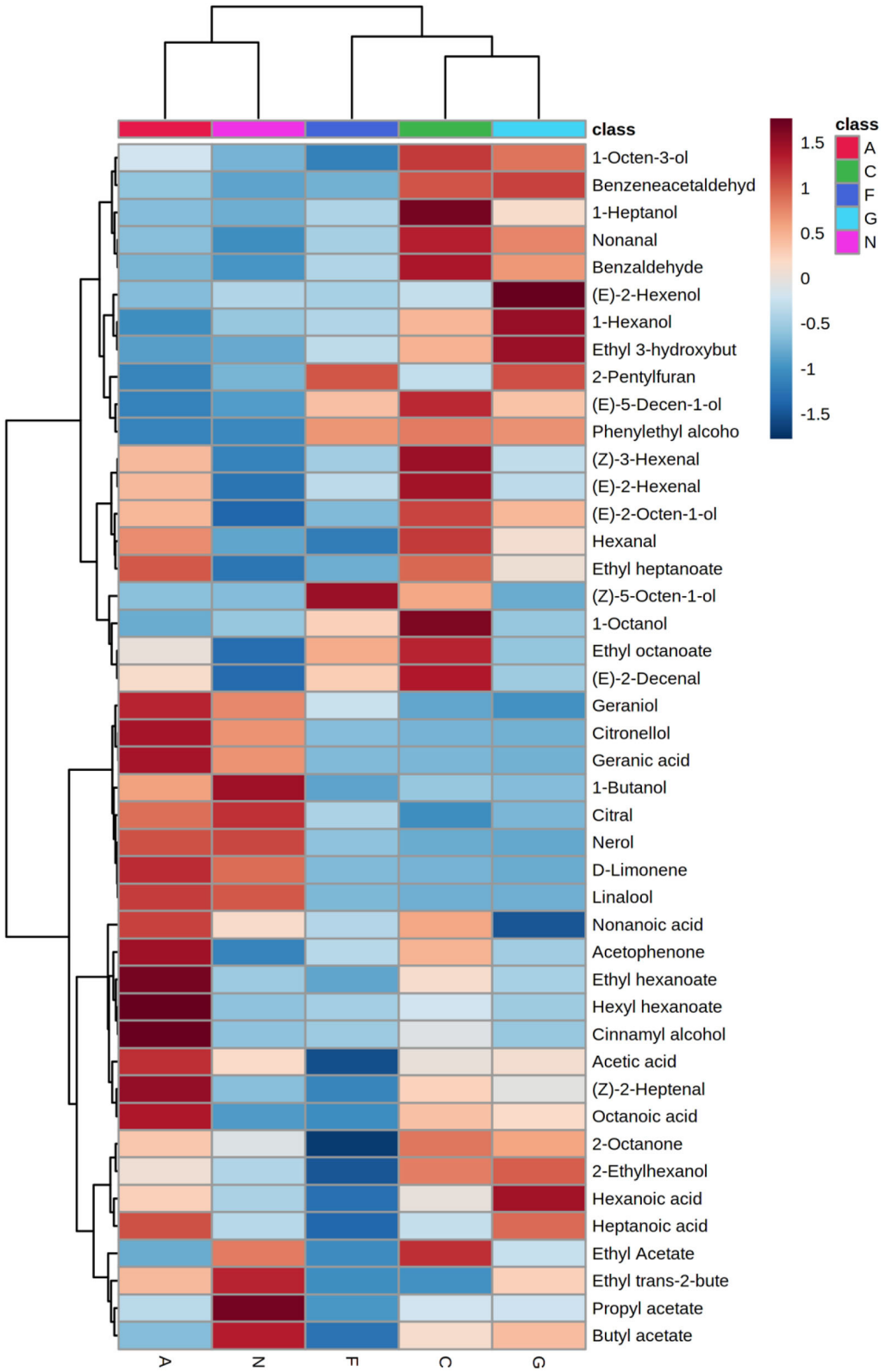

FIGURE 5 | Hierarchical clustering heatmap of the different volatiles in the five muscadine grape cultivars. Each colored cell on the map corresponds to a relative concentration of a volatile, with the samples in the rows and compounds in the columns. A, Alachua; C, Carlos; F, Fry; G, Granny Val; N, Noble.

the robustness of the model (Kalivodová et al., 2015). The R^2Y and Q^2 parameters of PLS-DA were significantly elevated (>0.5) (**Figure 4A**), indicating a valid and robust model (Kalivodová et al., 2015).

In this study, the PLS-DA model was shown to be a tool that could be used to investigate whether the five MG cultivars can be effectively discriminated (**Figure 4A**) and enable the visualization of the volatiles that contributed the most to the corresponding discrimination of the five MG cultivars (**Figure 4B**). The loadings of the variables on the two PLS-DA components are graphically illustrated in **Figure 4B**. More specifically, geraniol and cinnamyl alcohol were demonstrated to be essential for characterizing the Alachua MG cultivar in the loading plot of the PLS-DA. At the same time, ethyl trans-2-butenoate and propyl acetate were essential compounds to characterize the Noble MG cultivar. Furthermore, 2-Ethyl-1-hexanol, (*Z*)-3-hexenal, and (*E*)-2-hexenol were closely related to the Carlos, Fry, and Granny Val cultivars, respectively (**Figure 4B**).

The study further preformed 100 permutation tests to validate the PLS-DA model and are summarized in **Figure 4C**. The permutated models resulted in all R^2 (representing the explained variance) and Q^2 (representing the predictive capability) values (*Y*-axis data) on the left being lower than the original points on the right, indicating that the original model is statistically viable (**Figure 4C**).

Specific Volatile Markers of the Five MG Cultivars

The potential markers to distinguish different samples can be selected using the PLS-DA model (Lubes and Goodarzi, 2017). The VIP scores in the PLS-DA model estimate the importance of each *x* variable for each *x* variate in the prediction model and summarize the contribution that a variable makes to the model (Chong and Jun, 2005). Generally, a VIP score greater than 1 is considered a criterion for variable selection. A VIP score lower than 0.5 indicates that the variable is unimportant for the model classification and discrimination (Chong and Jun, 2005). In this study, by setting a threshold value of 1 for the VIP score in the PLS-DA model, 24 volatile metabolites were identified as crucial differential volatiles (**Table 2**). Combining the VIP scores in **Table 2** and the ANOVA results in **Figure 2**, five volatile markers that discriminate the five MG cultivars, namely 2-pentylfuran, 1-heptanol, ethyl hexanoate, (*Z*)-3-hexenal, and phenylethyl alcohol, were obtained.

Subsequently, a hierarchical cluster analysis (HCA) dendrogram was constructed to visualize the content differences of the potentially characterized VOCs (**Figure 5**). The HCA results revealed that the key volatiles were grouped into two classes. The first class comprised nine volatiles more abundant volatiles in the Carlos, Fry, and Granny Val MG cultivars. The second contained 15 volatiles correlated with the Alachua and Noble MG cultivars. This result was in accordance with the above multivariate analysis results.

Fruit aroma is highly influential on the overall flavor and consumer preference (Lin et al., 2019). As an aroma-dense fruit, the MG grape is a promising cultivar for future breeding efforts to attain the desired volatile aroma of table grapes and resulting wine. The characterization of the volatile profiles of the aroma-dense MG cultivars is the first step in elucidating the possible molecular mechanisms underlying volatile synthesis and the future genetic improvement of grape aroma. In this study, the volatile profiles of five newly introduced MG cultivars were comprehensively investigated and evaluated. The results opened the avenues to attain a cultivar with the desired flavor in the future.

CONCLUSIONS

Muscadine grape is an aroma-dense fruit and has been increasingly appreciated by growers, breeders, and consumers worldwide. Recently, five MG cultivars, namely Alachua, Carlos, Fry, Granny Val, and Noble, have been successfully introduced and cultivated in southern China. In the current study, the volatile constituents of the berries of the five MG cultivars were isolated using HS-SPME and analyzed using GC-MS. The results identified 44 different compounds in the studied cultivars by comparing the mass spectra and retention index with authentic standards, NIST libraries, and literature data. These volatiles were divided into esters, aldehydes, alcohols, fatty acids, terpenes, ketones, and furan based on their chemical nature. ANOVA, combined with Tukey's HSD test, revealed that the significant differences among the five cultivars are due to the quantitative differences of the 29 volatiles. Multivariate PCA and PLS-DA analyses showed a clear differentiation among the five cultivars, particularly the bronze-colored cultivars and the purple-black colored cultivars. The volatiles that contribute the most to the corresponding discrimination of five cultivars were highlighted. The implications of these findings were also discussed. This investigation is the first in-depth exploration of the volatile profiles of the aroma-dense MG.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

HD, QD, and HX conceived and designed the research and checked and revised the manuscript. HD, RH, ML, YL, and YZ performed the experiments. HD, LL, DL, XZ, ML, and XL analyzed the data. HD prepared and wrote the manuscript. All authors contributed to this article and approved the submitted version.

FUNDING

This study was financially supported by the Sichuan Science and Technology Programs (2020YFH0144 and 2020JDKP0039) and Sichuan Science Technology Resource Sharing Service Platform (Grape Resources).

ACKNOWLEDGMENTS

We thank all the members of HD's laboratory for so much help and so many great moments. Professor Jiang Lu is gratefully acknowledged for providing the five Muscadine grape cultivars. HD expresses her great and deep appreciation to her supervisor, Dr. Changyong Zhou, who gave her the greatest supporting in the work and life when she was a visiting scholar in Southwest University.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.728891/full#supplementary-material

REFERENCES

Berger, R. G. (2009). Biotechnology of flavours-the next generation. *Biotechnol. Lett.* 31, 1651–1659. doi: 10.1007/s10529-009-0083-5

Chong, I. G., and Jun, C. H. (2005). Performance of some variable selection methods when multicollinearity is present. *Chemom. Intell. Lab. Syst.* 78, 103–112. doi: 10.1016/j.chemolab.2004.12.011

Conner, P. J. (2009). Performance of Muscadine grape cultivars in Southern Georgia. *J. Am. Pomol. Soc.* 63, 101-107. Available online at: https://muscadines.caes.uga.edu/content/dam/caes-subsite/muscadine-grape-breeding/docs/Grape-Cultivar-Final.pdf (accessed June 21, 2021).

El Hadi, M. A. M., Zhang, F. J., Wu, F. F., Zhou, C. H., and Tao, J. (2013). Advances in fruit aroma volatile research. *Molecules* 18, 8200–8229. doi: 10.3390/molecules18078200

Golombek, P., Wacker, M., Buck, N., and Durner, D. (2021). Impact of UV-C treatment and thermal pasteurization of grape must on sensory characteristics and volatiles of must and resulting wines. *Food Chem.* 338, 128003. doi: 10.1016/j.foodchem.2020.128003

Gürbüz, O., Rouseff, J., and Rouseff, R. L. (2014). Comparison of Muscadine ester volatiles profiles of wines produced using pectinase pre-treatment and traditional methods. In: *Flavour Science Proceedings from XIII Weurman Favour Research Symposium*, eds. *V. Ferreira and R. Lopez (London: Elsevier Inc.)* pp. 429–433. doi: 10.1016/B978-0-12-398549-1.00080-5

Hoffmann, M., Conner Professor, P., Breeding Program, M., Brannen, P., Burrack, H., Mitchem, W., et al. (2020). Muscadine grape production guide for the Southeast, 2020. https://smallfruits.org/files/2020/07/muscadine-grape-production-guide-southeast.pdf. (accessed June 22, 2021)

Ju, Y., lun, Yue, X., feng, Cao, X., ying, Wei, X., feng, and Fang, Y., lin (2021). First study on the fatty acids and their derived volatile profiles from six Chinese wild spine grape clones (*Vitis davidii* Foex). *Sci. Hortic. (Amsterdam).* 275, 109709. doi: 10.1016/j.scienta.2020.109709

Kalivodová, A., Hron, K., Filzmoser, P., Najdekr, L., Janečková, H., and Adam, T. (2015). PLS-DA for compositional data with application to metabolomics. *J. Chemom.* 29, 21–28. doi: 10.1002/cem.2657

Kalua, C. M., and Boss, P. K. (2009). Evolution of volatile compounds during the development of cabernet sauvignon grapes (*Vitis vinifera* l.). *J. Agric. Food Chem.* 57, 3818–3830. doi: 10.1021/jf803471n

Kambiranda, D., Basha, S. M., Singh, R. K., He, H., Calvin, K., and Mercer, R. (2016). In Depth Proteome Analysis of Ripening Muscadine Grape Berry cv. Carlos reveals proteins associated with flavor and aroma compounds. *J. Proteome Res.* 15, 2910–2923. doi: 10.1021/acs.jproteome.5b01064

Lee, B., Lin, P., ching, Cha, H., soo, Luo, J., and Chen, F. (2016). Characterization of volatile compounds in Cowart muscadine grape (*Vitis rotundifolia*) during ripening stages using GC-MS combined with principal component analysis. *Food Sci. Biotechnol.* 25, 1319–1326. doi: 10.1007/s10068-016-0207-3

Lin, J., Massonnet, M., and Cantu, D. (2019). The genetic basis of grape and wine aroma. *Hortic. Res.* 6, 81. doi: 10.1038/s41438-019-0163-1

Liu, X. Q., Ickert-Bond, S. M., Nie, Z. L., Zhou, Z., Chen, L. Q., and Wen, J. (2016). Phylogeny of the Ampelocissus-Vitis clade in Vitaceae supports the New World origin of the grape genus. *Mol. Phylogenet. Evol.* 95, 217–228. doi: 10.1016/j.ympev.2015.10.013

Lubes, G., and Goodarzi, M. (2017). Analysis of volatile compounds by advanced analytical techniques and multivariate chemometrics. *Chem. Rev.*, 117, 6399–6422. doi: 10.1021/acs.chemrev.6b00698

Marshall, D. A., Stringer, S. J., and Spiers, J. D. (2012). Stilbene, ellagic acid, flavonol, and phenolic content of Muscadine grape (*Vitis rotundifolia* Michx.) cultivars. *Pharmac. Crops* 3, 69–77. doi: 10.2174/22102906012030010069

Mencarelli, F., and Bellincontro, A. (2018). Recent advances in postharvest technology of the wine grape to improve the wine aroma. *J. Sci. Food Agric.* 28, 303–325. doi: 10.1002/jsfa.8910

Olien, W. C. (1990). The Muscadine grape: botany, viticulture, history, and current industry. *HortScience* 25, 732–739. doi: 10.21273/HORTSCI.25.7.732

Perez-Coello, M. S., and Diaz-Maroto, M. S. (2009). Volatile and aroma compounds. In: Wine Chemistry and Biochemistry, eds. V. Moreno-Arribas and M. C. *Polo (New York: Springer)* pp. 249–436.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Available online at: http://www.r-project.org/index.html (accessed June 21, 2021).

Ramos-Madrigal, J., Runge, A. K. W., Bouby, L., Lacombe, T., Samaniego Castruita, J. A., Adam-Blondon, A. F., et al. (2019). Palaeogenomic insights into the origins of French grapevine diversity. *Nat. Plants* 5, 595–603. doi: 10.1038/s41477-019-0437-5

Reisch, B. I., Owens, C. L., and Cousins, P. S. (2012). Grape. In: Fruit Breeding, eds. *M. L. Badenes and D. H. Byrne (Boston, MA: Springer Science+Business Media, LLC)* pp. 225–262. doi: 10.1007/978-1-4419-0763-9_7

Ren, J., Huang, J., Shen, Y., Liang, D., and Xia, H. (2020). Comparative analysis of fruit quality and nutritional components of four muscadine grape cultivars. *Food Ferment. Ind.* 46, 158–163. doi: 10.13995/j.cnki.11-1802/ts.021453

Saccenti, E. (2012). Double-check: Validation of diagnostic statistics for PLS-DA models in metabolomics studies. *Metabolomics* 8, S3–S16. doi: 10.1007/s11306-011-0330-3

Sánchez-Palomo, E., Díaz-Maroto, M. C., and Pérez-Coello, M. S. (2005). Rapid determination of volatile compounds in grapes by HS-SPME coupled with GC-MS. *Talanta* 66, 1152–1157. doi: 10.1016/j.talanta.2005.01.015

Sandhu, A. K., and Gu, L. (2010). Antioxidant capacity, phenolic content, and profiling of phenolic compounds in the seeds, skin, and pulp of *Vitis rotundifolia* (Muscadine Grapes) as determined by HPLC-DAD-ESI-MSn. *J. Agric. Food Chem.* 58, 4681–4692. doi: 10.1021/jf904211q

Stanley, D. (1997). America's first grape-The muscadine. *Aricultural Res.* 45, 14–16.

Vidal, R., Ma, Y., and Sastry, S. S. (2016). Principal component analysis. In: Generalized Principal Component Analysis, Interdisciplinary Applied Mathematics 40, eds. *R. Vidal, Y. Ma, and S. Sastry (New York: Springer-Verlag)* 25–62. doi: 10.1007/978-0-387-87811-9

Wang, X., Tong, H., Chen, F., and Gangemi, J. D. (2010). Chemical characterization and antioxidant evaluation of muscadine grape pomace extract. *Food Chem.* 123, 1156–1162. doi: 10.1016/j.foodchem.2010.05.080

Wei, Z., Luo, J., Huang, Y., Guo, W., Zhang, Y., Guan, H., et al. (2017). Profile of polyphenol compounds of five muscadine grapes cultivated in the United States and in newly adapted locations in China. *Int. J. Mol. Sci.* 18, 1–18. doi: 10.3390/ijms18030631

Wen, J., Lu, L. M., Nie, Z. L., Liu, X. Q., Zhang, N., Ickert-Bond, S., et al. (2018). A new phylogenetic tribal classification of the grape family (Vitaceae). *J. Syst. Evol.* 56, 262–272. doi: 10.1111/jse.12427

Wu, Y., Zhang, W., Song, S., Xu, W., Zhang, C., Ma, C., et al. (2020). Evolution of volatile compounds during the development of Muscat grape 'Shine Muscat' (*Vitis labrusca* × *V. vinifera*). *Food Chem.* 309, 125778. doi: 10.1016/j.foodchem.2019.125778

Yang, C., Wang, Y., Liang, Z., Fan, P., Wu, B., Yang, L., et al. (2009). Volatiles of grape berries evaluated at the germplasm level by headspace-SPME with GC-MS. *Food Chem.* 114, 1106–1114. doi: 10.1016/j.foodchem.2008.10.061

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 11 November 2021
doi: 10.3389/fpls.2021.758215

Transcriptome Analysis Revealed the Mechanism by Which Exogenous ABA Increases Anthocyanins in Blueberry Fruit During Veraison

*Tianyu Han[1], Wenlong Wu[2] and Weilin Li[1]**

[1] Co-Innovation Center for Sustainable Forestry in Southern China, Forestry College, Nanjing Forestry University, Nanjing, China, [2] Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, Nanjing, China

Blueberry (*Vaccinium* spp.) is a popular healthy fruit worldwide. The health value of blueberry is mainly because the fruit is rich in anthocyanins, which have a strong antioxidant capacity. However, because blueberry is a non-model plant, little is known about the structural and regulatory genes involved in anthocyanin synthesis in blueberries. Previous studies have found that spraying 1,000 mg/L abscisic acid at the late green stage of "Jersey" highbush blueberry fruits can increase the content of anthocyanins. In this experiment, the previous results were verified in "Brightwell" rabbiteye blueberry fruits. Based on the previous results, the anthocyanin accumulation process in blueberry can be divided into six stages from the late green stage to the mature stage, and the transcriptome was used to systematically analyze the blueberry anthocyanin synthesis process. Combined with data from previous studies on important transcription factors regulating anthocyanin synthesis in plants, phylogenetic trees were constructed to explore the key transcription factors during blueberry fruit ripening. The results showed that ABA increased the anthocyanin content of blueberry fruits during veraison. All structural genes and transcription factors (MYB, bHLH, and WD40) involved in the anthocyanin pathway were identified, and their spatiotemporal expression patterns were analyzed. The expression of *CHS*, *CHI*, *DFR,* and *LDOX/ANS* in ABA-treated fruits was higher in the last two stages of maturity, which was consistent with the change in the anthocyanin contents in fruits. In general, six MYB transcription factors, one bHLH transcription factor and four WD40 transcription factors were found to change significantly under treatment during fruit ripening. Among them, *VcMYBA* plays a major role in the regulation of anthocyanin synthesis in ABA signaling. This result preliminarily explained the mechanism by which ABA increases the anthocyanin content and improves the efficiency of the industrial use of blueberry anthocyanins.

Keywords: blueberry, anthocyanins, transcriptome, ABA, *VcMYBA*

OPEN ACCESS

Edited by:
Jian Li,
Beijing Technology and Business University, China

Reviewed by:
Branka Salopek Sondi,
Rudjer Boskovic Institute, Croatia
Nicola Busatto,
Fondazione Edmund Mach, Italy
Cristian Balbontín,
Instituto de Investigaciones Agropecuarias, Chile

****Correspondence:***
Weilin Li
1923872828@qq.com

Specialty section:
This article was submitted to Plant Metabolism and Chemodiversity, a section of the journal Frontiers in Plant Science

Received: *13 August 2021*
Accepted: *13 October 2021*
Published: *11 November 2021*

Citation:
Han T, Wu W and Li W (2021) Transcriptome Analysis Revealed the Mechanism by Which Exogenous ABA Increases Anthocyanins in Blueberry Fruit During Veraison. Front. Plant Sci. 12:758215. doi: 10.3389/fpls.2021.758215

INTRODUCTION

Blueberry belongs to the genus *Vaccinium*, and the main cultivated species are northern highbush blueberry plants (*Vaccinium corymbosum* L.), southern highbush blueberry (primarily *V. corymbosum* L.), lowbush blueberry (*V. angustifolium* Aiton), and rabbiteye blueberry (*V. ashei* Reade) (Leisner et al., 2017). Blueberry fruit is one of the most popular healthy fruits worldwide.

This popularity is mainly because blueberries contain a variety of phytonutrients, the most representative of which are anthocyanins (Borges et al., 2010). With the in-depth study of blueberry anthocyanins, a growing number of clinical and animal experiments have proven that blueberry anthocyanins can effectively alleviate obesity (Prior et al., 2010) and cardiovascular disease (Zhu et al., 2013) and prevent type 2 diabetes (Burton-Freeman et al., 2019) and cancer (Faria et al., 2010). Therefore, the content and variety of anthocyanins are important characteristics of blueberry fruit. This study systematically explored the key transcription factors in anthocyanin synthesis and laid a foundation for future molecular breeding.

The structural genes of the anthocyanin pathway in plants are well understood, and the important functions of *CHS* (Schijlen et al., 2007), *CHI* (Guo et al., 2015), *DFR* (Katsu et al., 2017), *ANS* (Zhang et al., 2015), and *UFGT* (Boss et al., 1996) have also been verified. At present, three kinds of transcription factors, MYB, bHLH, and WD40, have been found to play a major regulatory role in the anthocyanin pathway (Dubos et al., 2010). They form the MBW complex and directly regulate the expression of structural genes, among which MYB transcription factors play a major role (Gonzalez et al., 2008). In *Arabidopsis thaliana*, the expression levels of *AtMYB75*, *AtMYB90* (Borevitz et al., 2000), *AtMYB113* and *AtMYB114* (Gonzalez et al., 2008) were positively correlated with changes in the anthocyanin content. Similarly, *GL3* and *EGL3* (Zhang et al., 2003) of the bHLH family identified in *Arabidopsis thaliana* also showed a positive correlation. The loss of *AtTTG1* in the WD40 family can affect the expression of DFR and other anthocyanin synthesis genes (Gonzalez et al., 2008). However, there are few studies on the structural genes and transcription factors regulating anthocyanin synthesis in blueberry. Only one MYB transcription factor, *VcMYBA,* has been found to increase anthocyanin synthesis by activating the promoter of *DFR* (Plunkett et al., 2018).

Exogenous ABA can increase the anthocyanin content of non-climacteric fruits such as strawberry (Jia et al., 2011; Li et al., 2011) and grape (Sandhu et al., 2011; Yamamoto et al., 2015). Exogenous ABA was also effective for non-climacteric fruit blueberry, and the northern highbush blueberries were found to have accelerated coloration and increased anthocyanin content 12 days after treatment with 1,000 mg/L exogenous ABA application (Oh et al., 2018). Therefore, the discovery that ABA treatment increased the total anthocyanin content of blueberry fruits can be used to study the mechanism of anthocyanin synthesis. With the development of high-throughput sequencing, the transcriptome has been widely used to study metabolic pathways and key genes. The genome of highbush blueberry (*V. corymbosum*) was assembled, which could provide a good reference for studies on its transcriptome (Colle et al., 2019). The mechanism of exogenous ABA application on grape berry ripening at 22 and 44 h was systematically illustrated by RNA-seq (Pilati et al., 2017). Similarly, the transcriptome can be used to systematically study the synthesis mechanism and explore the key regulatory genes in blueberry.

In this study, blueberry fruit was divided into six stages from late green to mature. We systematically analyzed the effect of ABA on whole transcripts during fruit ripening, especially those involved in the anthocyanin pathway. Combined with data from previous studies on anthocyanin biosynthesis in plants, key transcription factors involved in anthocyanin biosynthesis were identified, and their expression patterns were analyzed. Among the key transcription factors, the expression of *VcMYBA* was consistent with the increase in anthocyanins and ABA-responsive elements were found in the promoter of *VcMYBA*, suggesting that *VcMYBA* may play a role in ABA pathway. Transient silencing experiments showed that *VcMYBA* played an important role in anthocyanin synthesis.

Abbreviations: CHS, chalcone synthase; CHI, chalcone isomerase; F3H, flavonoid 3′- hydroxylase; F3′5′H, flavonoid 3′,5′-hydroxylase; DFR, dihydroflavonol 4-reductase; ANS, anthocyanidin synthase; UFGT, flavonoid-3-O-glucosyltransferase.

MATERIALS AND METHODS

Plant Materials and ABA Treatments

Four years old rabbiteye blueberry "Brightwell" were grown in experimental base of Baima district, Nanjing city, China. Blueberries of same size and growing condition were selected to do the treatment. Each treatment was arranged in a randomized complete block design with 6 replications and each block replication contains 6 shrubs.

The (+)-Abscisic Acid (Purity 95%, Coolaber company, China) was dissolved in double distilled H_2O containing 5% (v/v) ethanol and 0.1% Tween 80. When most of blueberries fruits were in a growing stage of "green mature" and the fruit at the top of the branch has just begun to turn red, 0, 500, and 1,000 mg/L ABA solutions were sprayed on fruit clusters. Tiny sprayer were used to spray the ABA solutions on the peels untill the peels are wetted. The fruit is guaranteed to be sprayed 2–3 times on all sides. Although the mock treatment was performed on blueberry shrubs, leaves, and branches were carefully avoided.

In different development stage of fruits, 30 fruits were immediately frozen in liquid nitrogen and stored at −80°C for later experiments.

Physiological Characterization

The Color of Fruit Peel

Fruit peel color was measured by colorimeter (Ci64, X-Rite, United States) and shown by the International Commission on Illumination a* and b* color space co-ordinates (Hunter and Harold, 1987). The a* value is negative for green and positive for red and the b* value is negative for blue and positive for yellow, both of the values range from −100 to 100. Due to the coloration of rabbiteye fruits starts from top to bottom, the top, side and bottom of fruit peel were separately measured in the same stage.

The Total Anthocyanins Content

The total anthocyanins content were determined by the double pH differential method (Pantelidis et al., 2007): absorbance of the extract was measured at 510 and 700 nm in buffers at pH 1.0 (hydrochloric acid–potassium chloride, 0.2 M) and 4.5 (acetate acid–sodium acetate, 0.2 M). Total anthocyanins content

was calculated using a molar extinction coefficient of 29,600 (cyanidin-3-glucoside) and absorbance of $A = [(A510 - A700)_{pH\ 1.0} - (A510 - A700)_{pH\ 4.5}]$.

Fruit Hardness

The fruit hardness were determined by Fruit hardness tester (Catno.9300, Takemura Electric Works Co., Japan). The cone type tip was used and the tip was perpendicularly applied on the side surface of blueberry fruits. The value was measured at the moment of tip intrusion to the surface.

Brix

The Brix were determined by saccharometer (PAL-1, Atago Co., Japan). The juice were left on the prim for 20 s to do the measurement.

Transcriptome Analysis

RNA Preparation and Sequencing

Total RNA was extracted from whole fruits using Trizol reagent (Invitrogen, Carlsbad, CA). The RNA integrity was analyzed on agarose gel. RNA concentration and integrity were measured by Qubit® RNA Assay Kit in Qubit® 2.0 Flurometer (Life Technologies, CA, United States) and RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system, respectively. Sequencing libraries were generated using NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, United States) following manufacturer's recommendations. Then, the libraries were sequenced on an Illumina Hiseq platform and 150 bp paired-end reads were generated.

Reads Mapping to the Reference Genome

Clean reads were obtained by in-house Perl scripts that remove reads containing adapter, reads containing ploy-N and low quality reads from raw reads. Reference genome and gene model annotation files of *V. corymbosum* (version 1.0) were downloaded from Giga science[1] (Colle et al., 2019). Paired-end clean reads were aligned to the reference genome using STAR (v2.5.1b) by the method of Maximal Mappable Prefix. The function of the transcript was annotated by UniProt database.[2]

Differential Expression Analysis

HTSeq v0.6.0 was used to count the reads numbers mapped to each gene and FPKM (fragments per kilobase per million reads) was calculated based on the length of the gene and reads count mapped to this gene. Then, differential expression analysis of two groups was performed using the DESeq2 R package (1.10.1). Genes with *P*-value using Benjamini and Hochberg's approach were assigned as differentially expressed. Enrichment analysis of differentially expressed genes of physiological processes was carried out by KEGG[3] and GO (clusterProfiler R package).

Construction of Phylogenetic Trees

The maximum likelihood phylogenetic trees was constructed by IQ-tree (Nguyen et al., 2015). The optimal alternative model is selected after calculation. The bootstrap value is 1,000. The following sequence was downloaded from GenBank accessions: **MYB:** *AtMYB75*(NC_003070.9), *AtMYB113*(NC_003070.9), *AtMYB90*(NC_003070.9), *AtMYB114*(NC_003070.9), *PhAN2*(AAF66727.1), *PhPHZ*(ADQ00388.1), *PhDPL*(ADQ00393.1), *VvMYBA1*(BAD18977), *VvMYBA2*(BAD18978), *MdMYB10*(ACQ45201), *MdMYBA*(NC_041797.1), *VcMYBA*(MH105054), *AcMYB10*(PSS35990), *AcMYB110*(AHY00342), *AtMYB4*(At4g38620), *PhMYB4*(ADX33331.1), *PhMYB27*(AHX24372), *VvMYBC2-L2*(ACX50288.2), *VuMYBC2*(AKR80571), *VvMYBC2-L1*(AFX64995.1), *AtMYB12*(ABB03913), *SlMYB12*(ACB46530.1), *VvMYBF1*(ACV81697), *AtMYB123*(CAC40021), *VvMYBPA2*(ACK56131.1), *VcMYB17*(ALP43798.1), *MdMYB11*(NC_041797.1), *MdMYB9*(NC_041796.1), *VuMYBPA1*(AKC94840.1), *VvMYBPA1*(CAJ90831.1), *AtMYB5*(NP_187963.1), *VvMYB5b*(AAX51291), *AtMYB56*(AT5G17800), *MdMYB1*(NC_041789.1). **bHLH:** *AtEGL3*(AT1G63650), *AtGL3*(AT1G11130), *AmDELILA*(Uniprot:Q38736), *PhJAF13*(Uniprot:O64908), *AtTT8*(AT4G09820), *MdbHLH3*(ADL36597.1), *PhAN1*(FJ227329.1), *VvbHLH1*(Uniprot:A0A438KI27), *MdBHLH33*(EI011581.1). **WD40**:*AtTTG1*(AT5G24520), *MdTTG1*(GU173814.1), *PhAN11*(Uniprot:O24514), *VvWDR1*(Uniprot:Q19N39), *VvWDR2*(Uniprot:Q19N38).

Quantitative Real-Time PCR Analysis

Eleven transcription factors with altered expression levels were selected for qRT-PCR verification. The specific primers for 11 genes are shown in **Supplementary Table 4**. Expression level is relative to the housekeeping gene *Actin* of highbush blueberry (Zifkin et al., 2012). qPCR was conducted using the TB Green® Fast qPCR Mix (Takara, Japan) according to manufacturer's requirements. The $2^{-\Delta\Delta Ct}$ method was used to calculate the relative gene expression (Livak and Schmittgen, 2001). The experiment carried out three biological repetitions.

Virus-Induced Gene Silencing in Blueberry Fruit

In order to identify the function of four copies of *VcMYBA*, transient silencing was performed by the tobacco rattle virus based VIGS technology. Because the other four sequences are similar and the three coding sequences are shifted, they cannot be translated into proteins. Therefore, interference sequences are designed in exon region and intron region, respectively. The length of the two sequences was about 330 bp, which was amplified by PCR with specific primers (**Supplementary Table 6**). The PCR products were cloned into the pTRV2 vector to produce TRV: VcMYBA-exon and TRV: VcMYBA-intron constructs. The *Agrobacterium tumefaciens*, strain GV3101, was used in transforming the recombinant plasmids pTRV1, pTRV2. Infection solution formula: 0.01 mM $MgCl_2$, 0.01 mM MES, 200 μM AS, 0.6OD Agrobacterium solution (pTRV1 and pTRV2 mixed in equal proportion). Slowly inject the infection solution to the side of the fruit. Observe the fruit after 10 days.

Analysis of Anthocyanin Changes in Fruits by UPLC/MS

UPLC-MS/MS analysis was performed to evaluate anthocyanin changes in transiently silenced fruits. The standards of petunidin,

[1] http://gigadb.org/dataset/100537

[2] https://www.uniprot.org/

[3] http://www.genome.jp/kegg/

cyanidin, delphinidin, pelargonidin, and peonidin (Coolaber company, China) were used to make the standard curve. 200 mg fruit powder was added 5 ml of extraction solution (Absolute ethanol: H_2O: HCL = 2:1:1) and ultrasonic extraction in water bath at 4°C for 0.5 h. Then, the samples were hydrolyzed in boiling water bath for 60 min. After that, the samples were taken out and cooled, and the volume of the samples was fixed to 5 ml. The samples were left standing and the supernatant was transferred to the fresh tubes. Finally, the supernatants were filtered through 0.22 μM filter membrane and loaded on instrument for analysis. LC-MS/MS analyses were performed using a LC-30AD system (Shimadzu) coupled with an Qtrap6500 mass spectrometer (AB SCIEX). Samples were injected onto an Agilent Zorbax sb-c18 column (150 × 4.6 mm, 1.9 μm) using a 16-min linear gradient at a flow rate of 1 m/min. The eluents were eluent A (0.1% formic acid) and eluent B (0.1% acetonitrile). The solvent gradient was set as follows: 18% B, 0 min; 20% B, 4 min; 35% B, 7 min; 95% B, 8 min; 95% B, 12 min; 18% B, 12.1 min; 18% B, 0 min, 16 min. Mass spectrometer was operated in ESI positive polarity mode with spray voltage of 5.5 kV, capillary temperature of 550°C, atomization gas pressure of 50 psi and auxiliary air pressure of 60 psi.

RESULTS

Fruit Coloration, Total Anthocyanin Content, and Fruit Quality

The developmental stage of cv "Brightwell" from ripening initiation to maturity was divided into six stages (**Figure 1A**) based on fruit color (Stage 1, the whole fruit was green; Stage 2, the top turned red; Stage 3, the side turned red; Stage 4, the whole fruit was red; Stage 5, the whole fruit turned purple; and Stage 6, the whole fruit was purple) and size (**Supplementary Table 1**). Application of 1,000 mg/L ABA resulted in a significant acceleration of fruit coloration (**Figure 1B**). From S3 to S4, the a* value suggests that the top, side and bottom of fruit from the 1,000 mg/L ABA group turned red earlier than those from the 0 to 500 mg/L ABA groups. At S5, the top and bottom of the fruits turned green earlier. From S5 to S6, the b* value suggests that the side and bottom of the fruit from the 1,000 mg/L ABA group turned blue earlier than those from the 0 to 500 mg/L ABA groups. From the a* and b* values, there were no notable differences among the three groups at S6. Anthocyanin accumulation was consistent with the change in fruit coloration (**Figure 1C**). At S1–S3, almost no anthocyanins were detected. At S4, although the average anthocyanin content of the 1,000 mg/L treatment was higher than that of the control group, there was no statistically significant difference. At S5 and S6, the total anthocyanin content of the 1,000 mg/L ABA treatment group was obviously higher than that of the 0 and 500 mg/L treatment groups. From S5 to S6, the anthocyanin content of the 0 mg/L treatment group nearly doubled but was still lower than that in the 1,000 mg/L group. Application of 1,000 mg/L ABA also changed the fruit quality (**Supplementary Figure 1**). Compared to 0 and 500 mg/L ABA treatment, fruits hardness of 1,000 mg/L ABA treatment declined more sharply from S4 and become softer in S6 (**Supplementary Figure 1A**). Consist with fruit hardness, fruits of 1,000 mg/L ABA treatment contains more soluble solid in S6 (**Supplementary Figure 1B**). Application of ABA did not significantly change the fruit size, which was within the range of **Supplementary Table 1**. Therefore, blueberry fruits treated with 0 and 1,000 mg/L ABA were selected for transcriptome analysis.

Assembly and Annotation of the Transcriptome

The fruits in Stages 1–6 with 0 and 1,000 mg/L ABA treatment were further subjected to transcriptome analysis. The raw bases of each sample varied from 7.04 G to 10.91 G with a 0.03% error ratio and 46% GC content. In total, 2218553864 clean reads and 2176923110 raw reads were obtained (**Supplementary Table 2**). The raw reads of 12 samples with three biological repeats were uploaded to the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) database (all samples and corresponding explanations can be found in PRJNA664011). For each sample, 87.43–89.19% of the reads were mapped to the highbush blueberry genome. Approximately 87% of reads were mapped to exonic regions, 7% of reads were mapped to intergenic regions, and 6% of reads were mapped to intronic regions. The clean reads were assembled into 128,559 unigenes. A total of 88,532 unigenes were annotated by the Swiss-Prot database, accounting for 68.86% of the total (**Supplementary Table 3**).

Profiling of the Transcriptome

Principal component analysis (PCA) was used to analyze the overall variation in transcripts among 32 samples. The 3D plot shows that the three repetitions are well clustered together, and the differences between the sample points are obvious (**Figure 2A**). The FPKM value of genes in the ABA-treated group was compared with that in the control group. A Log_2(fold change) value greater than 1.3 or less than −1.3 indicates upregulation and downregulation of gene expression, respectively. At S2, the number of genes with significant changes was approximately twice as high as that of the control. The number of differentially expressed genes decreased sharply at S3 and then gradually increased at S4–S6 (**Figure 2B**). Volcano maps were used to represent the overall changes in gene expression between two groups (**Figure 2C**). The results showed that the number of genes with drastic changes in expression decreased continuously from S1 to S6.

The total amount of anthocyanins in the ABA treatment group was higher than that in the control group from S4, so the transcripts of fruits at S4 were functionally categorized and analyzed using the GO and KEGG databases (**Figure 3**). The results showed that the anthocyanin synthesis pathway had changed significantly.

Structural Genes Involved in Anthocyanin Synthesis

The transcripts of all key enzymes involved in the anthocyanin pathway were analyzed (**Figure 4**). Green indicates that the expression of transcripts in the ABA-treated group was lower than that in the control group, while red indicates the

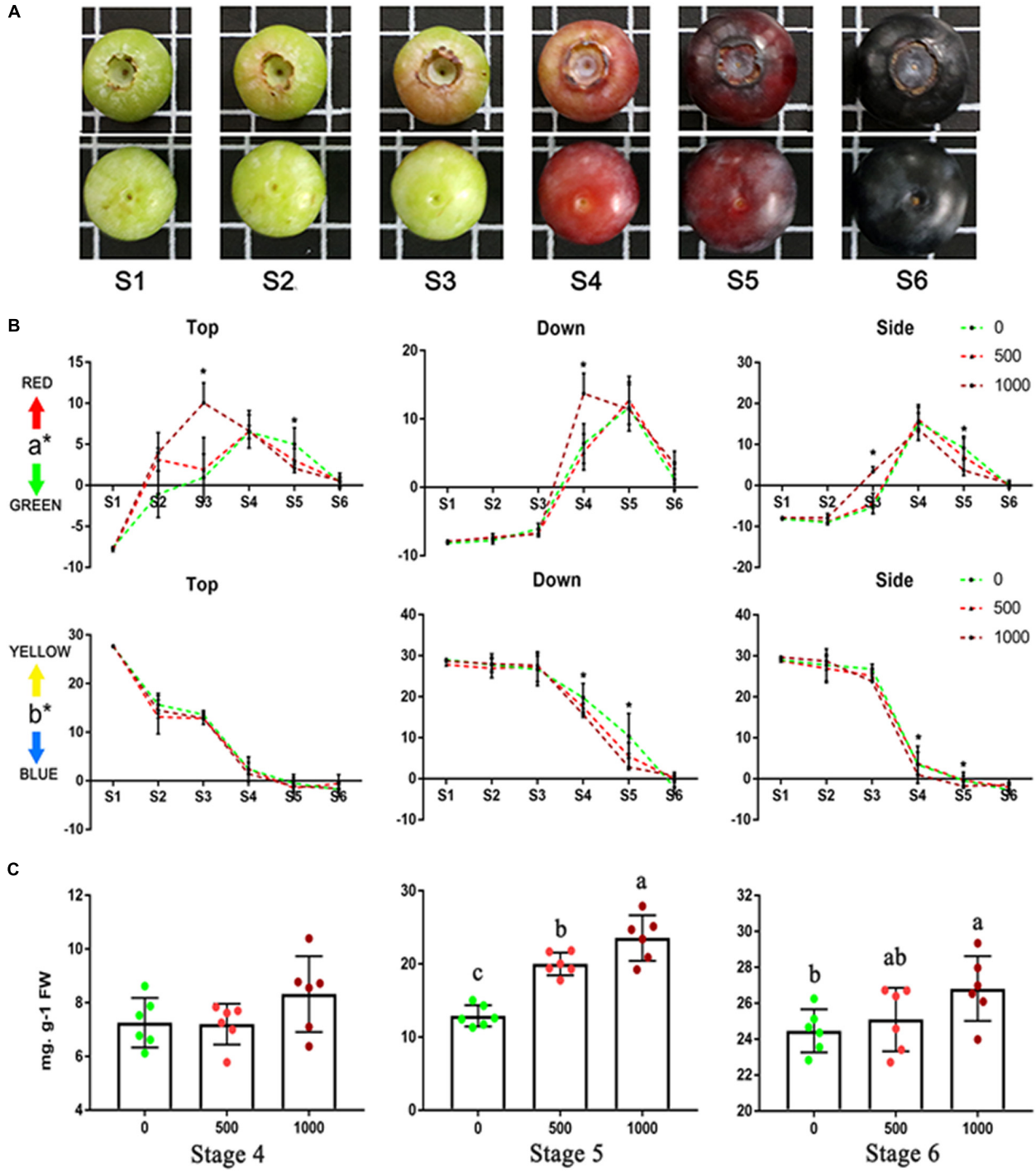

FIGURE 1 | Ripening progress of blueberry (cv "Brightwell") and coloration and anthocyanins content of fruits under 0, 500, and 1,000 mg/L ABA treatment. **(A)** The ripening progress is divided into six stages. **(B)** The color of top, side and bottom of fruits from S1 to S6. **(C)** Total anthocyanins content from S4 to S6. Different letters and asterisk indicate statistical significance ($P < 0.05$) as determined by a one-way ANOVA test (Duncan's multiple range).

opposite. In the whole process, the transcripts of CHS and CHI, two important enzymes upstream, all showed consistent downregulation at S1 and S2 and increased at S5 and S6. Another key enzyme, LDOX/ANS, showed a similar expression pattern and was highly expressed at S3 and S4. The transcripts of the three enzymes F3H, F3′H and F3′5′H, which are found on different branches, showed different expression patterns across the six stages. Similarly, the expression of UFGT transcripts was

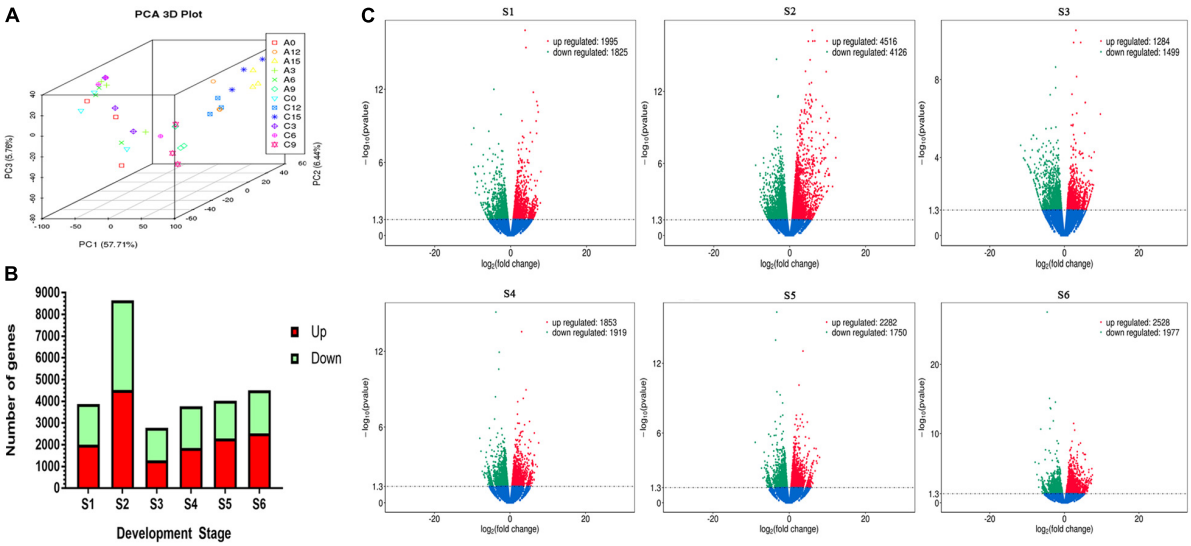

FIGURE 2 | Profiling of transcripts difference between the two groups in six stages. **(A)** Principal component analysis of each sample. A0-15 represents 1,000 mg/L ABA treatment S1–S6, C0-15 represents S1–S6 of control group **(B)** The total number of transcripts significantly up-regulated and down regulated in development stages. **(C)** Volcano maps show overall differences between treatments.

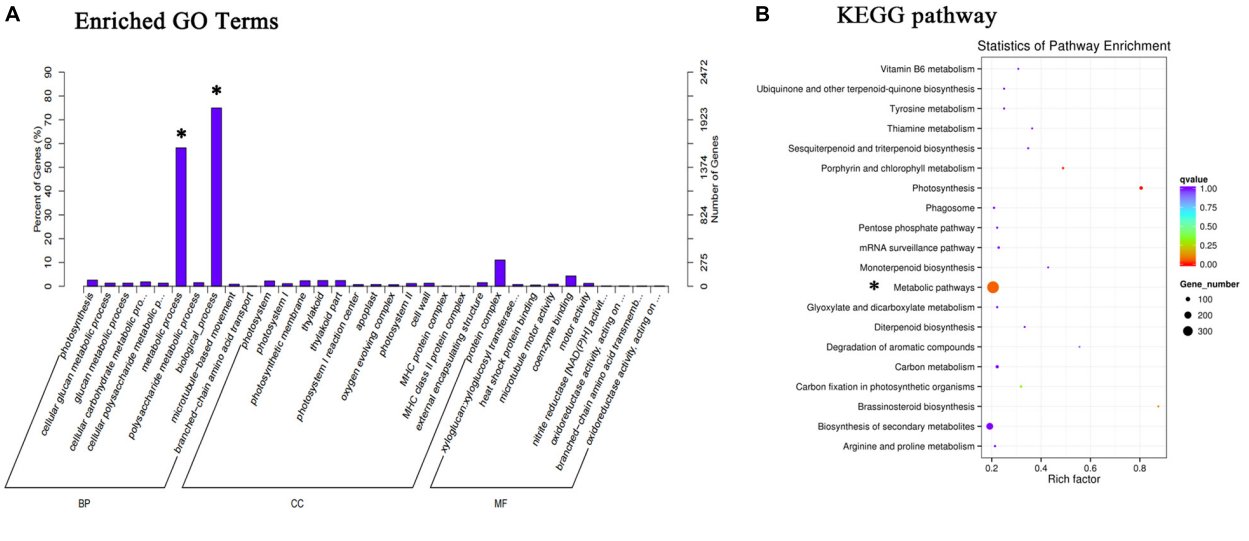

FIGURE 3 | GO and KEGG were used to analyze the pathway of significant changes in S4. **(A)** Different expression genes were annotated and enriched with the GO database. **(B)** Metabolic pathways with different changes annotated by KEGG database. Asterisk marks changes in metabolic pathways.

inconsistent. Although the transcripts of DFR differed to some extent across the six stages, they were generally highly expressed after S3. In general, the transcript expression patterns of CHS, CHI, DFR, and LDOX were consistent with the changes in the anthocyanin content in the fruits.

Identification of Key Transcription Factors Involved in Anthocyanin Synthesis

The expression pattern of several key structural genes was the same, which indicates that the expression of transcription factors may affect the expression of downstream genes. Published transcription factors from different species were used to construct a phylogenetic tree with homologous transcription factors in blueberry. According to the classification of MYB transcription factors (Jiang and Rao, 2020) and their functions, MYB transcription factors involved in anthocyanin synthesis in blueberry were identified (**Figure 5**). Five blueberry MYB transcripts belong to the branch SG VIII-e2κ, and the expression of MYB transcription factors on this branch has been proven to promote anthocyanin synthesis. A blueberry transcript belonging to the SG V δ branch that promotes anthocyanin synthesis was also identified. For the negatively regulated SG VIII-E e1β

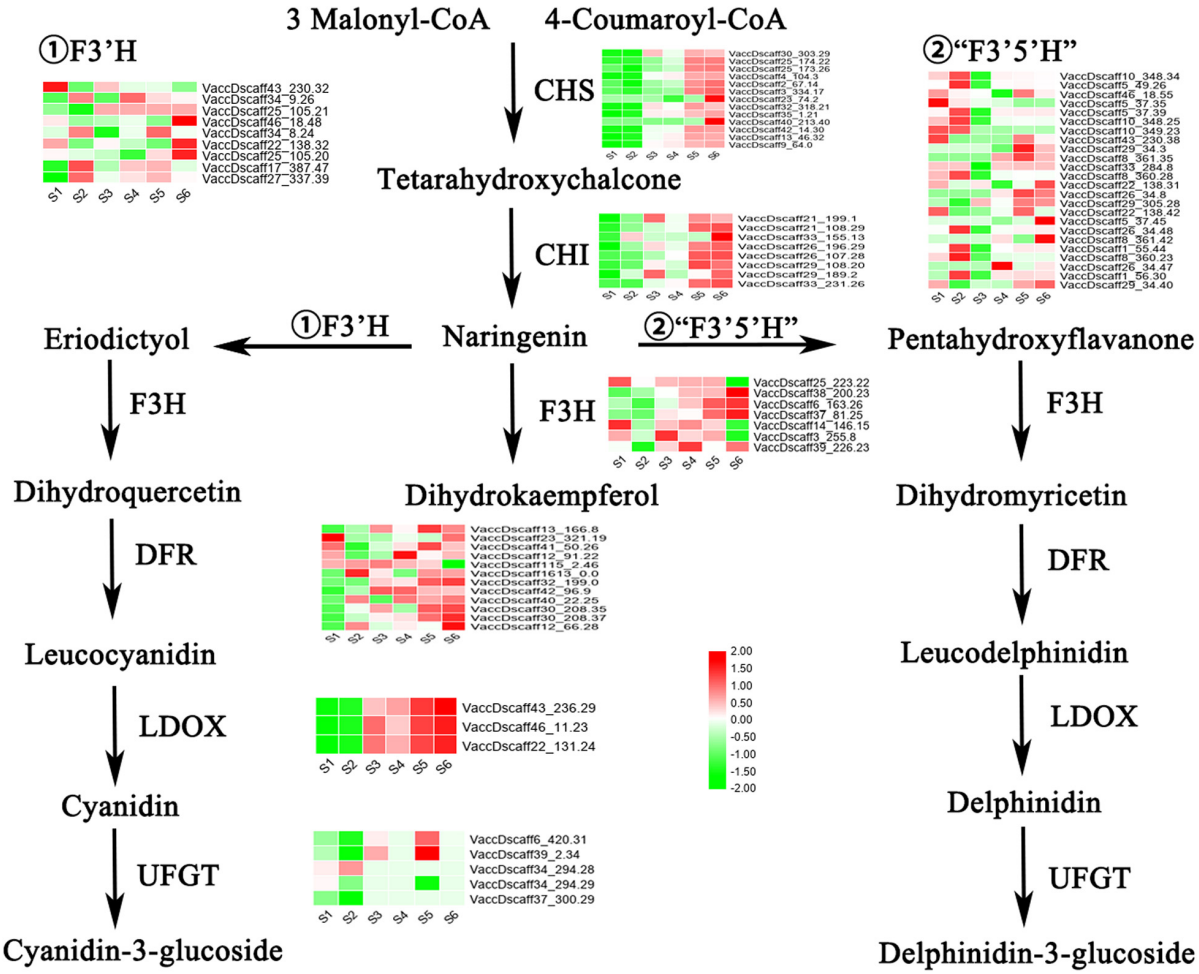

FIGURE 4 | Expression changes of key structural enzymes in anthocyanin pathway during S1–S6 after ABA treatment. Red represents up regulation of gene expression, while green represents down regulation of gene expression. The expression levels (FPKM) corresponding to the different colors ranges from –2 to 2.

branch, two transcripts were identified. For bHLH and WD40 transcription factors involved in anthocyanin regulation, four homologous genes were identified (**Supplementary Figures 2**, **3**).

Spatiotemporal Expression Analysis of Key Transcription Factors

Among the identified transcription factors, six MYB transcription factors, one bHLH transcription factor and four WD40 transcription factors showed significant differences in specific stages (**Figure 6**). The expression pattern of these transcription factors was verified by qRT-PCR (**Supplementary Figure 4**). The expression of two MYB transcription factors on the SG VIII-e2κ branch was positively correlated with anthocyanin content. At the same time, both transcripts are homologous genes of *VcMYBA* that have been proved to regulate anthocyanin synthesis. Two homologous negatively regulated MYB transcription factors showed upregulated and downregulated differential expression at S1. The transcript belonging to the SG V δ branch showed a high increase from S4 to S6, which was earlier than the change in the anthocyanin content. The bHLH transcript homologous to *AtGL3* was downregulated at S1–S2 but highly expressed at S4–S5. There was no correlation between the expression patterns of the four WD40 transcripts. The results showed that transcripts belonging to the SG VIII-e2κ branch and homologous to *VcMYBA* were most likely related to the increase in anthocyanins. Through promoter element analysis, 2-4 ABREs (ABA-responsive elements) was found to appear in the promoter region of four copies of *VcMYBA* (**Supplementary Table 5**). This shows that *VcMYBA* responds in the ABA pathway and may therefore be associated with the changes of *PYL* (Encodes regulatory components of ABA receptor), *SnRK2* (protein kinases involved in ABA signaling), and *ABF* (bZIP transcription factor with specificity for ABRE) gene families (**Supplementary Figure 6**). After ABA treatment, seven genes of *ABF* family were up-regulated in S1–S6, but the up-regulation degree decreased gradually. In S4, the expression of four ABF genes in the treatment group was 1.27–1.5 times higher than that in the control group. *PYL* and *SnRK2* gene families did not show an obvious overall change trend, but one gene in *PYL*

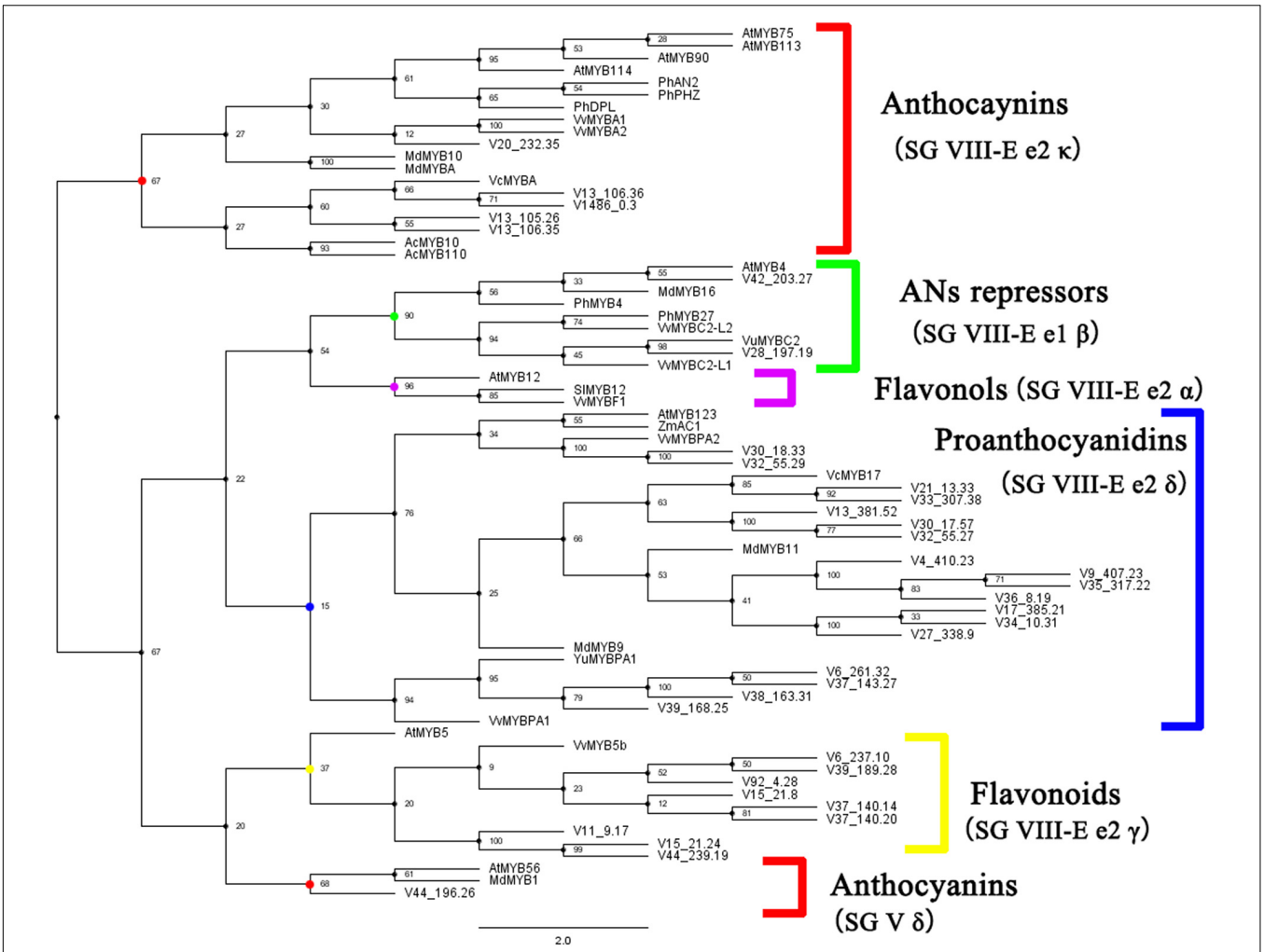

FIGURE 5 | Identification of known MYB transcription factors regulating anthocyanin synthesis in blueberry. Phylogenetic tree of MYB transcription factors was constructed in whole nucleic acid sequences using maximum likelihood tree with 1,000 bootstrap in IQ-tree. The optimal alternative model was selected after calculation by IQ-tree.

family changed 1.79 times in S3 and one gene in *SnRK2* family changed 1.59 times in S4. From the expression of these genes, *VcMYBA* has a positive correlation with their expression.

Transient Silencing of the Expression of Four Copies of *VcMYBA*

Through the genome, a total of five copies of *VcMYBA* were identified. One copy had transposon insertion, two copies had frameshift and could not be translated into protein, one copy of TCA base sequence was mutated into the termination codon TGA and only one copy could be fully translated into protein (**Supplementary Figure 5**). Because of their high sequence similarity, transient silencing was designed in exon and intron regions, respectively. The results showed that in the control group, the anthocyanin synthesis in the injection wound area increased, and the anthocyanin synthesis in the fruit treated with TRV: *VcMYBA*-exon decreased, while the anthocyanin accumulation in the wound treated with TRV: *VcMYBA*-intron was not obvious, and the anthocyanin synthesis in the fruit did not decrease significantly (**Figure 7A**). The expression of *VcMYBA* in TRV:*VcMYBA*-exon group was also significantly lower than that in control group and TRV:*VcMYBA*-intron group (**Figure 7B**). The contents of six main anthocyanins were detected by UPLC-MS/MS (**Figure 7C**). The contents of petunidin, cyanidin, delphinidin, and peonidin in TRV:*VcMYBA*-exon group were significantly lower than those in the other two groups (**Figure 7D**). There was no significant change in anthocyanin content in TRV:*VcMYBA*-intron compared with the control group.

DISCUSSION

Effect of ABA on Blueberry Fruit

Many studies have reported that exogenous ABA could increase the anthocyanin content in non-climacteric fruits

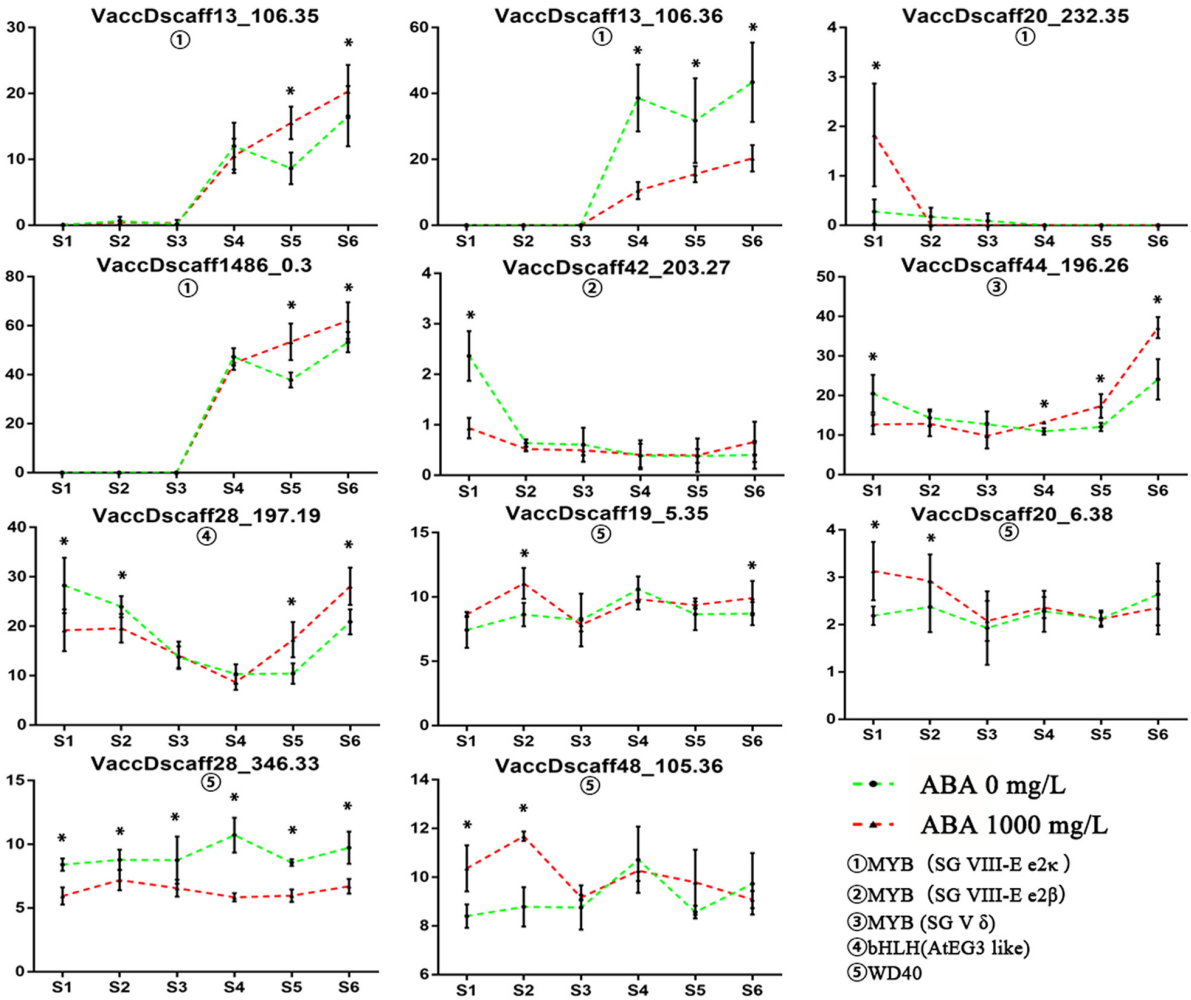

FIGURE 6 | MYB, bHLH, and WD40 transcription factors with different expression in S1–S6. FPKM is used to represent the level of gene expression. Asterisk indicates statistical significance ($P < 0.05$) as determined by a one-way ANOVA test (Duncan's multiple range). Values represent the means ± SD of three replicates.

(Jia et al., 2011; Sandhu et al., 2011). The anthocyanin content of plants will also increase under many abiotic stresses, such as drought (Nakabayashi et al., 2014), low temperature (Li et al., 2017), and nitrogen deficiency (Zhang et al., 2017). The response of plants to these stresses is also related to the ABA signaling pathway. Previous studies in blueberry only focused on the 12 days after ABA application at the late green fruit stage (Oh et al., 2018). The results showed that the content of anthocyanins increased and the color of blueberry fruits changed faster after 1,000 mg/L ABA treatment, which was consistent with the results of other non-climacteric fruits. There was no significant difference in anthocyanin content, fruit quality and size between 500 mg/L ABA treatment and the control group. The effective ABA concentration of increasing anthocyanin accumulation in blueberry is about three times higher than that of grape (*Vitis vinifera* L.) It may be that fruit pruina blocks part of ABA.

For the mechanism by which ABA increases the anthocyanin content, only transcriptome analysis at 24 and 48 h after ABA treatment was carried out (Pilati et al., 2017). It takes approximately 35 days for late green fruits to ripen, and the key genes involved in this process need to be studied in detail. Therefore, we divided blueberry fruits into six stages from late green to mature and analyzed the changes in gene expression throughout the whole process in detail. The transcriptome analysis results showed that the expression of structural genes involved in anthocyanin synthesis was inconsistent. In ABA-treated fruits, the transcripts of CHS, CHI, DFR, and LDOX/ANS, which were proven to be the key structural genes in anthocyanin biosynthesis, were generally expressed at low levels at S1–S2 and at high levels at S5–S6. Previous studies only involved the changes of several structural genes in anthocyanin pathway in mature fruits (Jia et al., 2011; Li et al., 2011). We analyzed the structural gene changes of the whole anthocyanin pathway through transcriptome system to clarify the direct reason for the increase of anthocyanin content during fruit color transformation.

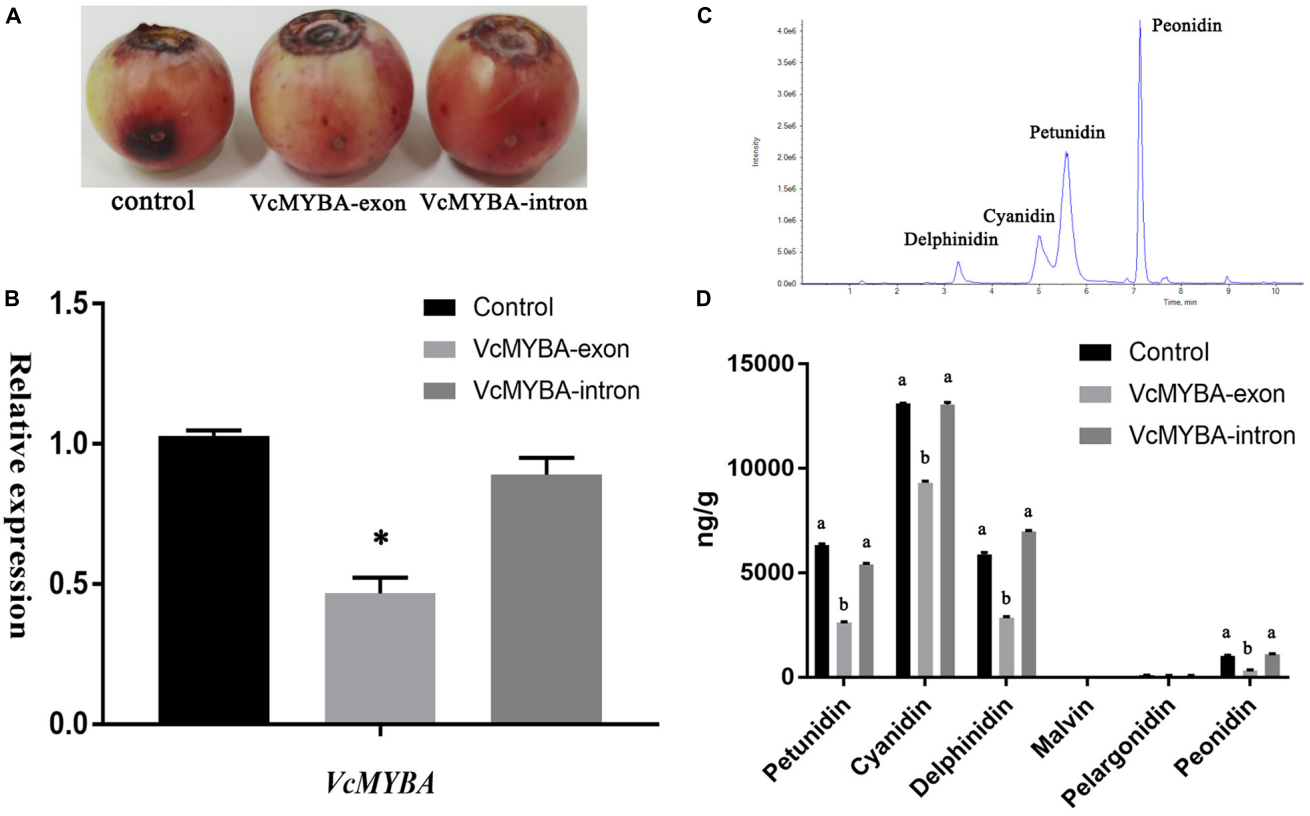

FIGURE 7 | Virus-induced gene silencing of *VcMYBA* in blueberry fruit. **(A)** Transgenic fruits 10 days after VIGS treatment. **(B)** qRT-PCR analysis of *VcMYBA* expression in treated fruits. **(C)** UPLC analysis of anthocyanins. **(D)** Contents of six anthocyanins in treated fruits. Values represent the means ± SD of three replicates. Different letters and asterisk indicates statistical significance ($P < 0.05$) as determined by a one-way ANOVA test (Duncan's multiple range).

Potential Key Transcription Factors in Anthocyanin Pathway

Combined with previous studies on transcription factors upstream of these structural genes (Gonzalez et al., 2008), we speculate that the MBW complex is involved in the anthocyanin regulation of the ABA pathway. Homologous genes of MBW transcription factors published in important plants were identified by database annotation and phylogenetic tree construction in blueberry. Two *VcMYBA* (*AtMYB75/90* homologous gene) copies, one *AtMYB56* homologous copy and one *AtGL3* homologous copy were highly correlated with changes in anthocyanin contents. Overexpression of *VcMYBA* (Plunkett et al., 2018) and *AtMYB75/90* (Gonzalez et al., 2008) can enhance anthocyanin synthesis. In addition, *DFR* and *CHS* were also proven to be downstream of *AtMYB75/90,* and *DFR* was downstream of *VcMYBA*. Our results are consistent with those of previous studies. The high expression of *VcMYBA* at S5–S6 was consistent with the change in anthocyanins, and the same structural genes, *DFR* and *CHS*, were highly expressed. *Atmyb56* needs sucrose induction, and its anthocyanin increasing effect is relatively low (Jeong et al., 2018). The homologous gene of *AtMYB56* in blueberry began to be highly expressed at S3, suggesting that this gene may not play a major role. *AtGL3* can enhance downstream anthocyanin expression by interacting with *AtMYB75/90* (Zhang et al., 2003). Therefore, *VcMYBA* may play a key role in ABA pathway.

Similarities and Differences Between *VcMYBA* and *VvMYBA*

Grape, which is also a non-climacteric fruit similar to blueberry, can increase anthocyanin accumulation under drought (Castellarin et al., 2007), low temperature (Cohen et al., 2012) and ABA treatment (Zhai et al., 2017). *DFR*, *LDOX, UFGT,* and *VvMYBA1* are upregulated under water deficit (Castellarin et al., 2007), and the promoters of both *VvMYBA1-2* and *VvMYBA2* can be activated by ABA, but only *VvMYBA2* can be activated by ethylene (Zhai et al., 2017). White grapes arose through the mutation of *VvMYBA1-2* and *VvMYBA2* (Walker et al., 2007). These findings indicate that *VvMYBA* plays a role in the ABA pathway of non-climacteric fruits and that *DFR*, *LDOX,* and *UFGT* are downstream of *VvMYBA*. For *VcMYBA*, heterologous expression of *VcMYBA* in *Nicotiana benthaminana* could induce anthocyanin accumulation (Plunkett et al., 2018). Through promoter element analysis, there are four promoter elements ABRE responding to ABA in the promoter region of *VcMYBA* (**Supplementary Table 5**). Transient silencing of the exon region of vcmyba in blueberry fruit significantly reduced the synthesis of anthocyanin in blueberry fruit, while silencing

the intron region did not significantly reduce the synthesis of anthocyanin in blueberry fruit. Therefore, it can be considered that the copy of *VcMYBA* that can be translated into protein can effectively regulate anthocyanin synthesis, and whether other copies play a role as lncRNA remains to be further studied.

This study is based on many previous studies, and the results are consistent with the previous conclusions. These results preliminarily explain the mechanism by which ABA increases the anthocyanin content in non-climacteric fruits and broadens the MYB pathway in response to ABA. As ABA is a stress response hormone, these results also provide a reference for explaining the stress-induced increase in the anthocyanin content in plants. This result can provide a basis and new target for anthocyanin breeding in blueberry.

CONCLUSION

The application of exogenous ABA in the late green period of blueberry fruit maturation resulted in an increase in anthocyanins. The most direct reason for this result is that *CHS*, *CHI*, *DFR,* and *LDOX/ANS* are highly expressed at S5–S6. Among the transcription factors whose expression changes, *VcMYBA* had the greatest correlation with the changes of anthocyanin and corresponding synthetic structural genes. The promoter of *VcMYBA* has ABRE element, and the up-regulated expression of *VcMYBA* is related to the changes of *ABF* gene family expression, indicating that *VcMYBA* may respond in the ABA pathway. Through VIGS mediated transient silencing, *VcMYBA* has been proved to play an important role in anthocyanin synthesis.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: National Center for Biotechnology Information (NCBI) BioProject database under accession number PRJNA664011.

AUTHOR CONTRIBUTIONS

TH and WL designed the experiments. TH performed the experiments and analyzed the data. WL and WW wrote the manuscript. All authors read and approved the final manuscript.

FUNDING

This research was supported by the Jiangsu Agriculture Science and Technology Innovation Fund [CX(17)2011 and CX(18)2018].

ACKNOWLEDGMENTS

We thanks for the experimental site and materials provided by institute of botany, Jiangsu province and Chinese academy of sciences.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.758215/full#supplementary-material

Supplementary Figure 1 | Hardness and Brix fruits under ABA treatment. **(A)** The hardness of fruits under 0, 500, and 1,000 mg/L ABA treatment from S4 to S6. **(B)** The Brix of fruits under 0, 500, and 1,000 mg/L ABA treatment from S4 to S6. Different letters indicates statistical significance ($P < 0.05$) as determined by a one-way ANOVA test (Duncan's multiple range).

Supplementary Figure 2 | Identification of known bHLH transcription factors regulating anthocyanin synthesis in blueberry. Phylogenetic tree of bHLH transcription factors was constructed in whole nucleic acid sequences using maximum likelihood tree with 1,000 bootstrap in IQ-tree. The optimal alternative model was selected after calculation.

Supplementary Figure 3 | Identification of known WD40 transcription factors regulating anthocyanin synthesis in blueberry. Phylogenetic tree of WD40 transcription factors was constructed in whole nucleic acid sequences using maximum likelihood tree with 1,000 bootstrap in IQ-tree. The optimal alternative model was selected after calculation.

Supplementary Figure 4 | Verification of transcriptional results by qRT-PCR. The $\triangle\triangle$Ct method was applied to each gene pair, and the sample with the highest Ct value smaller than 35 was chosen as the control. Different letters indicate statistical significance ($P < 0.05$) as determined by a one-way ANOVA test.

Supplementary Figure 5 | Gene location and sequence analysis of five copies of *VcMYBA*. **(A)** Gene location of five copies of *VcMYBA*. **(B)** Sequence alignment of four copies of *VcMYBA*. The box in the figure shows the code shift area caused by insertion mutation.

Supplementary Figure 6 | Heatmaps of the gene expression changes of *ABF*, *SnRK2,* and *PYL* gene families in 1,000 mg/L ABA treatment group compared with the control group during S1–S6. Gene changes were expressed by $\log_2$ fold change.

Supplementary Table 1 | Mean fresh fruit length, diameter and weight throughout ripening. Data are means ± SE ($n = 10$). Different letters indicates statistical significance ($P < 0.05$) as determined by a one-way ANOVA test (Duncan's multiple range).

Supplementary Table 2 | The overall statistical analysis of bases and reads of each samples.

Supplementary Table 3 | The overall statistical analysis of reads mapping of each samples.

Supplementary Table 4 | Primers used in qRT-PCR.

Supplementary Table 5 | Four copies of *VcMYBA* promoter region elements were predicted. The prediction is made by online database PLACE.

Supplementary Table 6 | Primers for VIGS, including enzyme digestion sites.

REFERENCES

Borevitz, J. O., Xia, Y., Blount, J., Dixon, R. A., and Lamb, C. (2000). Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. *Plant Cell* 12, 2383–23293. doi: 10.1105/tpc.12.12.2383

Borges, G., Degeneve, A., Mullen, W., and Crozier, A. (2010). Identification of flavonoid and phenolic antioxidants in black currants, blueberries, raspberries, red currants, and cranberries. *J. Agric. Food Chem.* 58, 3901–3909. doi: 10.1021/jf902263n

Boss, P. K., Davies, C., and Robinson, S. P. (1996). Analysis of the expression of anthocyanin pathway genes in developing vitis vinifera L. cv shiraz grape berries

and the implications for pathway regulation. *Plant Physiol.* 111, 1059–1066. doi: 10.1104/pp.111.4.1059

Burton-Freeman, B., Brzezinski, M., Park, E., Sandhu, A., Xiao, D., and Edirisinghe, I. (2019). A selective role of dietary anthocyanins and flavan-3-ols in reducing the risk of type 2 diabetes mellitus: a review of recent evidence. *Nutrients* 11:841. doi: 10.3390/nu11040841

Castellarin, S. D., Matthews, M. A., Di Gaspero, G., and Gambetta, G. A. (2007). Water deficits accelerate ripening and induce changes in gene expression regulating flavonoid biosynthesis in grape berries. *Planta* 227, 101–112. doi: 10.1007/s00425-007-0598-8

Cohen, S. D., Tarara, J. M., Gambetta, G. A., Matthews, M. A., and Kennedy, J. A. (2012). Impact of diurnal temperature variation on grape berry development, proanthocyanidin accumulation, and the expression of flavonoid pathway genes. *J. Exp. Bot.* 63, 2655–2665. doi: 10.1093/jxb/err449

Colle, M., Leisner, C. P., Wai, C. M., Ou, S., Bird, K. A., Wang, J., et al. (2019). Haplotype-phased genome and evolution of phytonutrient pathways of tetraploid blueberry. *Gigascience* 8:giz012. doi: 10.1093/gigascience/giz012

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. doi: 10.1016/j.tplants.2010.06.005

Faria, A., Pestana, D., Teixeira, D., De Freitas, V., Mateus, N., and Calhau, C. (2010). Blueberry anthocyanins and pyruvic acid adducts: anticancer properties in breast cancer cell lines. *Phytother. Res.* 24, 1862–1869. doi: 10.1002/ptr.3213

Gonzalez, A., Zhao, M., Leavitt, J. M., and Lloyd, A. M. (2008). Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in Arabidopsis seedlings. *Plant J.* 53, 814–827. doi: 10.1111/j.1365-313X.2007.03373.x

Guo, J., Zhou, W., Lu, Z., Li, H., Li, H., and Gao, F. (2015). Isolation and functional analysis of chalcone isomerase gene from purple-fleshed sweet potato. *Plant Mol. Biol. Rep.* 33, 1451–1463. doi: 10.1007/s11105-014-0842-x

Hunter, R. S., and Harold, R. W. (1987). *The Measurement of Appearance*. Hoboken, NJ: John Wiley & Sons.

Jeong, C. Y., Kim, J. H., Lee, W. J., Jin, J. Y., Kim, J., Hong, S. W., et al. (2018). AtMyb56 regulates anthocyanin levels via the modulation of AtGPT2 expression in response to sucrose in Arabidopsis. *Mol. Cells* 41, 351–361.

Jia, H. F., Chai, Y. M., Li, C. L., Lu, D., Luo, J. J., Qin, L., et al. (2011). Abscisic acid plays an important role in the regulation of strawberry fruit ripening. *Plant Physiol.* 157, 188–199. doi: 10.1104/pp.111.177311

Jiang, C. K., and Rao, G. Y. (2020). Insights into the diversification and evolution of R2R3-MYB transcription factors in plants. *Plant Physiol.* 183, 637–655. doi: 10.1104/pp.19.01082

Katsu, K., Suzuki, R., Tsuchiya, W., Inagaki, N., Yamazaki, T., Hisano, T., et al. (2017). A new buckwheat dihydroflavonol 4-reductase (DFR), with a unique substrate binding structure, has altered substrate specificity. *BMC Plant Biol.* 17:239. doi: 10.1186/s12870-017-1200-6

Leisner, C. P., Kamileen, M. O., Conway, M. E., O'connor, S. E., and Buell, C. R. (2017). Differential iridoid production as revealed by a diversity panel of 84 cultivated and wild blueberry species. *PLoS One* 12:e0179417. doi: 10.1371/journal.pone.0179417

Li, C., Jia, H., Chai, Y., and Shen, Y. (2011). Abscisic acid perception and signaling transduction in strawberry: a model for non-climacteric fruit ripening. *Plant Signal. Behav.* 6, 1950–1953. doi: 10.4161/psb.6.12.18024

Li, P., Li, Y. J., Zhang, F. J., Zhang, G. Z., Jiang, X. Y., Yu, H. M., et al. (2017). The Arabidopsis UDP-glycosyltransferases UGT79B2 and UGT79B3, contribute to cold, salt and drought stress tolerance via modulating anthocyanin accumulation. *Plant J.* 89, 85–103. doi: 10.1111/tpj.13324

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

Nakabayashi, R., Yonekura-Sakakibara, K., Urano, K., Suzuki, M., Yamada, Y., Nishizawa, T., et al. (2014). Enhancement of oxidative and drought tolerance in Arabidopsis by overaccumulation of antioxidant flavonoids. *Plant J.* 77, 367–379. doi: 10.1111/tpj.12388

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

Oh, H. D., Yu, D. J., Chung, S. W., Chea, S., and Lee, H. J. (2018). Abscisic acid stimulates anthocyanin accumulation in 'Jersey' highbush blueberry fruits during ripening. *Food Chem.* 244, 403–407. doi: 10.1016/j.foodchem.2017.10.051

Pantelidis, G. E., Vasilakakis, M., Manganaris, G. A., and Diamantidis, G. (2007). Antioxidant capacity, phenol, anthocyanin and ascorbic acid contents in raspberries, blackberries, red currants, gooseberries and cornelian cherries. *Food Chem.* 102, 777–783. doi: 10.1016/j.foodchem.2006.06.021

Pilati, S., Bagagli, G., Sonego, P., Moretto, M., Brazzale, D., Castorina, G., et al. (2017). Abscisic acid is a major regulator of grape berry ripening onset: new insights into ABA signaling network. *Front. Plant Sci.* 8:1093. doi: 10.3389/fpls.2017.01093

Plunkett, B. J., Espley, R. V., Dare, A. P., Warren, B. A. W., Grierson, E. R. P., Cordiner, S., et al. (2018). MYBA from blueberry (*Vaccinium* Section *Cyanococcus*) is a subgroup 6 type R2R3MYB transcription factor that activates anthocyanin production. *Front. Plant Sci.* 9:1300. doi: 10.3389/fpls.2018.01300

Prior, R. L., Wilkes, S. E., Rogers, T. R., Khanal, R. C., Wu, X., and Howard, L. R. (2010). Purified blueberry anthocyanins and blueberry juice alter development of obesity in mice fed an obesogenic high-fat diet. *J. Agric. Food Chem.* 58, 3970–3976. doi: 10.1021/jf902852d

Sandhu, A. K., Gray, D. J., Lu, J., and Gu, L. (2011). Effects of exogenous abscisic acid on antioxidant capacities, anthocyanins, and flavonol contents of muscadine grape (*Vitis rotundifolia*) skins. *Food Chemistry* 126, 982–988. doi: 10.1016/j.foodchem.2010.11.105

Schijlen, E. G., De Vos, C. H., Martens, S., Jonker, H. H., Rosin, F. M., Molthoff, J. W., et al. (2007). RNA interference silencing of chalcone synthase, the first step in the flavonoid biosynthesis pathway, leads to parthenocarpic tomato fruits. *Plant Physiol.* 144, 1520–1530. doi: 10.1104/pp.107.100305

Walker, A. R., Lee, E., Bogs, J., Mcdavid, D. A., Thomas, M. R., and Robinson, S. P. (2007). White grapes arose through the mutation of two similar and adjacent regulatory genes. *Plant J* 49, 772–785. doi: 10.1111/j.1365-313X.2006.02997.x

Yamamoto, L. Y., De Assis, A. M., Roberto, S. R., Bovolenta, Y. R., Nixdorf, S. L., García-Romero, E., et al. (2015). Application of abscisic acid (S-ABA) to cv. Isabel grapes (Vitis vinifera×Vitis labrusca) for color improvement: effects on color, phenolic composition and antioxidant capacity of their grape juice. *Food Res. Int.* 77, 572–583. doi: 10.1016/j.foodres.2015.10.019

Zhai, X., Zhang, Y., Kai, W., Liang, B., Jiang, L., Du, Y., et al. (2017). Variable responses of two VlMYBA gene promoters to ABA and ACC in Kyoho grape berries. *J. Plant Physiol.* 211, 81–89. doi: 10.1016/j.jplph.2016.12.013

Zhang, F., Gonzalez, A., Zhao, M., Payne, C. T., and Lloyd, A. (2003). A network of redundant bHLH proteins functions in all TTG1-dependent pathways of Arabidopsis. *Development* 130, 4859–4869. doi: 10.1242/dev.00681

Zhang, J., Han, Z.-Y., Tian, J., Zhang, X., Song, T.-T., and Yao, Y.-C. (2015). The expression level of anthocyanidin synthase determines the anthocyanin content of crabapple (Malus sp.) petals. *Acta Physiol. Plant.* 37:109. doi: 10.1007/s11738-015-1857-0

Zhang, Y., Liu, Z., Liu, J., Lin, S., Wang, J., Lin, W., et al. (2017). GA-DELLA pathway is involved in regulation of nitrogen deficiency-induced anthocyanin accumulation. *Plant Cell Rep.* 36, 557–569. doi: 10.1007/s00299-017-2102-7

Zhu, Y., Ling, W., Guo, H., Song, F., Ye, Q., Zou, T., et al. (2013). Anti-inflammatory effect of purified dietary anthocyanin in adults with hypercholesterolemia: a randomized controlled trial. *Nutr. Metab. Cardiovasc. Dis.* 23, 843–849. doi: 10.1016/j.numecd.2012.06.005

Zifkin, M., Jin, A., Ozga, J. A., Zaharia, L. I., Schernthaner, J. P., Gesell, A., et al. (2012). Gene expression and metabolite profiling of developing highbush blueberry fruit indicates transcriptional regulation of flavonoid metabolism

and activation of abscisic acid metabolism. *Plant Physiol.* 158, 200–224. doi: 10.1104/pp.111.180950

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 10 December 2021
doi: 10.3389/fpls.2021.765897

Ethylene and Auxin: Hormonal Regulation of Volatile Compound Production During Tomato (*Solanum lycopersicum* L.) Fruit Ripening

Eric de Castro Tobaruela[1,2†], Bruna Lima Gomes[1,2†], Vanessa Caroline de Barros Bonato[1,2†], Elis Silva de Lima[1,2], Luciano Freschi[3] and Eduardo Purgatto[1,2]*

[1] Department of Food and Experimental Nutrition, School of Pharmaceutical Sciences, University of São Paulo (USP), São Paulo, Brazil, [2] Food Research Center (FoRC), São Paulo, Brazil, [3] Department of Botany, Institute of Bioscience, University of São Paulo (USP), São Paulo, Brazil

As the auxin-ethylene interaction in climacteric fruit ripening has been highlighted, the hormonal regulation of aroma changes in climacteric fruits requires clarification. The influence of both phytohormones on the volatile organic compound (VOC) metabolism was evaluated during tomato (*Solanum lycopersicum* L.) fruit ripening. Tomato fruits cv. Micro-Tom and Sweet Grape at the mature green stage were randomly grouped according to treatment with ethylene (ETHY), auxin (IAA), or both (ETHY + IAA). At middle ripening, Micro-Tom ETHY + IAA fruits present VOC profiles similar to those of ETHY fruits, while Sweet Grape presents VOC profiles closer to those of IAA fruits. At full ripeness, Micro-Tom and Sweet Grape ETHY + IAA fruits show profiles closer to those of IAA fruits, suggesting that the auxin overlaps the ethylene effects. Aroma compounds positively correlated with consumer preferences (2-isobutylthiazole, 6-methyl-5-hepten-2-one, and others) are identified in both cultivars and have their contents affected by both hormone treatments. The transcription of genes related to the biosynthesis of important tomato VOCs that have fatty-acid and carotenoid precursors evidences their regulation by both plant hormones. Additionally, the results indicate that the observed effects on the VOC metabolism are not restricted to the Micro-Tom cultivar, as these are also observed in the Sweet Grape cultivar. In conclusion, ethylene and auxin directly regulate the metabolic pathways related to VOC formation, impacting tomato aroma formation during ripening since Micro-Tom fruits apparently at the same maturation stage have different aromas.

Keywords: climacteric fruit, crosstalk, carotenoids, aromatic volatiles, fruit quality

OPEN ACCESS

Edited by:
María José Jordán, Instituto Murciano de Investigación y Desarrollo Agrario y Alimentario (IMIDA), Spain

Reviewed by:
Denise Tieman, University of Florida, United States
Branka Salopek Sondi, Ruđer Bošković Institute, Croatia

***Correspondence:**
Eduardo Purgatto
epurgatt@usp.br

† These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Plant Metabolism and Chemodiversity, a section of the journal Frontiers in Plant Science

Received: *27 August 2021*
Accepted: *02 November 2021*
Published: *10 December 2021*

Citation:
Tobaruela EC, Gomes BL, Bonato VCB, Lima ES, Freschi L and Purgatto E (2021) Ethylene and Auxin: Hormonal Regulation of Volatile Compound Production During Tomato (Solanum lycopersicum L.) Fruit Ripening. Front. Plant Sci. 12:765897. doi: 10.3389/fpls.2021.765897

INTRODUCTION

Tomatoes (*Solanum lycopersicum* L.) are among the most cultivated and consumed horticultural crops worldwide[1]. They have been extensively studied as a reference for climacteric fruits and a model system for fleshy fruit development and ripening due to their advantages over other species of agronomic interest (Gapper et al., 2013; Zhu et al., 2018; Food and Agriculture Organization Corporate Statistical Database [FAOSTAT], 2021). In these fruits, ripening is associated with a

[1] http://www.fao.org/faostat

rapid spike in the biosynthesis of ethylene, a phytohormone that plays a critical role in controlling ripening (Tucker et al., 2017). During this process, climacteric fruits experience an initial auto-inhibitory phase (system 1) followed by a rapid increase in ethylene emission during the initiation of an autocatalytic phase (system 2; Fenn and Giovannoni, 2021). Without undermining the role of ethylene in fruit ripening, new perspectives have been introduced on ripening regulation that indicate coordinated action between ethylene and other phytohormones. Auxin was shown to significantly affect the interplay between other plant hormones, and it is now commonly accepted as a main regulator of fruit ripening (Muday et al., 2012; Kumar et al., 2014). Indole-3-acetic acid is the main auxin in plants (Korasick et al., 2013). Although the intricate mechanisms underlying auxin regulation remain largely unclear, fruit ripening studies have shown the roles of a few auxin-related genes in controlling fruit firmness (Guillon et al., 2008) and regulating pigment accumulation (Breitel et al., 2016; Hao et al., 2016). Furthermore, evidence suggests crosstalk between ethylene and auxin during ripening, with auxin having the opposite effect to that of ethylene (Su et al., 2015; Li et al., 2016). Elevated auxin levels delay the fruit ripening process, and their biosynthesis decreases with the transition to fruit ripening (Kumar et al., 2014). Auxin is crucial for triggering ripening and impacts the transition between the two ethylene production systems. Therefore, it does not necessarily inhibit fruit ripening (Ross et al., 2011; McAtee et al., 2013; Seymour et al., 2013).

The interplay between multiple phytohormones regulating the metabolic pathways involved in fruit ripening can occur at different levels, including through interaction between components of phytohormone signal transduction (Muday et al., 2012; Li et al., 2021). Regarding ethylene-auxin crosstalk, recent reports have elucidated antagonistic effects of the two plant hormones, including auxin acting as a ripening repressor and thereby opposing the known role of ethylene in inducing ripening in tomato fruit (Su et al., 2015; Li et al., 2016). However, since the auxin-ethylene relationship is extremely intricate, some of the phenotypic effects have yet to be assigned to either phytohormone (Wang et al., 2021).

Tomato aroma quality is strictly ripening-dependent once most of the volatile organic compounds (VOCs) are produced from fatty acids, carotenoids, and amino acids in metabolic processes that occur during ripening process. These compounds accumulate at the onset of tomato ripening and peak at either full ripening or shortly before, producing the tomato fruit aroma (Klee and Giovannoni, 2011; Karlova et al., 2014). In conjunction with hormonal factors, the VOC profile is also affected by cultivation conditions, ripening stage, and postharvest treatment (Baldwin et al., 2015; Lee et al., 2018; Zhao et al., 2019). Another increasingly studied aspect is the wide range of phenotypic variation and metabolic diversity of tomato cultivars. Several hundred VOCs were identified in the fruits of nearly 400 tomato cultivars. However, only 23 of these were identified as being responsible for the characteristic aroma of tomatoes, while 13 were found to be important in terms of consumer acceptance: methional; 3-methylbutanenitrile; 2-methyl-1-butanol; 3-methyl-1-butanol; 2-isobutylthiazole; 1-octen-3-one; β-ionone; 3-methylbutanoic acid; (E,E)-2,4-decadienal; 3-methylbutanal; (E)-2-heptenal; and 6-methyl-5-hepten-2-one (Tieman et al., 2017). Comparing the flavors of 71 tomato accessions in China, another study highlighted 15 VOCs as compounds that could be used as parameters for evaluating tomato flavor (Cheng et al., 2020). Evidencing this diversity in tomato aroma, Tieman et al. (2017) also identified some flavorful components that have been lost with breeding. Other recent studies have revealed substantial gene loss and intense negative selection of genes and promoters during tomato domestication and improvement. Gao et al. (2019) identified a rare allele of *TomLoxC* that may have undergone negative selection during domestication. This tomato lipoxygenase gene is involved in apocarotenoid production, which contributes to the desirable tomato aroma.

Few studies have evaluated the effects of phytohormone crosstalk on ripening parameters. Knowledge of aroma regulation is limited, and the role of ethylene and auxin in aroma formation during climacteric fruit ripening remains almost unknown. According to Fenn and Giovannoni (2021), VOC accumulation and the role of plant hormones in VOC production are intriguing and currently expanding areas of inquiry that are open to further exploration. In this context, this study explored the interactions between ethylene and auxin, as well as the role of each phytohormone in the VOC metabolism in two tomato cultivars. We explored the effects of exogenous ethylene and auxin treatments on the VOC profiles and the expression of key genes related to regulating the metabolism of tomato aroma compounds. Finally, our results provide insights into the contributions of ethylene, auxin, and their interactions to regulating aroma metabolism during tomato fruit ripening.

MATERIALS AND METHODS

Plant Materials and Phytohormone Treatments

Tomato plants (*S. lycopersicum* cv. Micro-Tom and Sweet Grape) were grown under standard greenhouse conditions. Fruit samples were harvested at the mature green stage and randomly separated into four groups (n = 100 fruits by group) according to phytohormone treatments: CTRL (without treatment), ETHY (ethylene treatment), IAA (indole-3-acetic acid treatment), and ETHY + IAA (both treatments). After harvested, tomato fruits were sterilization with 0.1% aqueous sodium hypochlorite solution for 15 min and hormonal treatments were immediately applied. During the experiments, fruits were left to ripen spontaneously in a 323 L chamber at 22°C, for a 16-h-day/8-h-night cycle at 80% relative humidity. Ethylene treatments were performed using a gaseous hormone at 10 $\mu L.L^{-1}$ for 12 h. The indole-3-acetic acid solutions were prepared at 100 μM in 10 mM MES buffer at pH 5.6 using 3% sorbitol and injected through the calyx end as described by Su et al. (2015). Fruits from the IAA group received indole-3-acetic acid solution, while fruits from the ETHY + IAA group were exposed to gaseous ethylene before being infiltrated with auxin solution. Fruits from the ETHY and CTRL groups were injected with a buffer solution only to maintain a consistent injection method for all samples.

Both treatments were optimized under our laboratory conditions and the phytohormone concentrations were selected as they were sufficient to change tomato ripening. Two experimental blocks were performed with Micro-Tom fruits, while one was performed with Sweet Grape fruits. In the first Micro-Tom and Sweet Grape experiments, fruits in all groups were frozen during the ripening stage of the CTRL group. The second Micro-Tom experiment had a different experimental design, with samples being frozen when their group reached the defined ripening stages. During all the experiments, the ethylene emission and peel color shift were evaluated daily as ripening parameters. For further analyses, at least five fruits from each group were collected at different points, covering three ripening stages: mature green, breaker, and red. Samples were then frozen in liquid nitrogen and stored at −80°C.

Ethylene Emission

Five fruits from each group were individually placed in airtight glass containers and left at 25°C. After 1 h, 1 mL samples were collected from the headspace with a gas-tight syringe through a rubber septum for ethylene analysis. This analysis was performed by gas chromatography (Agilent Technologies Inc., Santa Clara, United States, model HP-6890). A flame ionization detector was employed, and an HP-Plot Q column (30 m × 0.53 mm × 40 μm, Agilent Technologies) was used. The injector and detector temperatures were set at 250°C, and the oven temperature was set at 30°C. The injections were performed in pulsed splitless mode. Helium was used as the carrier gas (1 mL.min^{-1}).

Peel Color Characterization

The peel color was measured using a HunterLab ColorQuest XE instrument (Hunter Associates Laboratories) in terms of L (lightness), A (redness and greenness), and B (yellowness and blueness). The hue value was used to represent color variations and was defined as described by Fabi et al. (2007). At least five representative measurements were taken from each experimental group.

Extraction of Volatile Organic Compounds and Solid-Phase Microextraction–GC–MS Analysis

The headspace volatile production of tomato fruit was determined by solid-phase microextraction (SPME). A pool of five fresh fruits was homogenized with 30% sodium chloride solution (Merck). Aliquots of 10 g were placed in vials and frozen at −20°C. Samples were thawed under agitation at 40°C immediately before analysis. The headspace equilibrium time was 10 min, and the adsorption time was 50 min. The SPME fiber (PDMS/DVB/CAR 50 μm, Supelco Co.) was injected directly into an HP-6890 gas chromatograph (Agilent Technologies) coupled to a mass selective detector (HP-5973, Agilent Technologies) and held for 10 min for desorption of VOCs. The injector temperature was 200°C. Components were then separated using a Supelcowax 10 capillary column (30 m × 0.25 mm × 0.25 μm), and the oven temperature was programmed to increase from 40 to 150°C at 2°C.min^{-1}. These conditions were previously optimized and selected according to the higher number of peaks and total area of the chromatogram. VOCs were identified by comparison using the NIST (NIST98, version 2.0, Gaithersburg, United States) and confirmed with the spectral data available from the MassBank of North America (MoNA[2]). Identification was also based on comparing the retention indices (RIs) of a series of *n*-alkanes (C7–C30, Supelco) with those reported in the literature and the retention times and mass spectra with those of authentic external standards. A pool of the following volatile compound external standards (Sigma–Aldrich) was prepared for use in identifying the compounds by mass spectral comparison: (E)-2-hexenal, hexanal, 1-hexanol, (Z)-3-hexen-1-ol, (Z)-3-hexenal, 1-penten-3-one, 6-methyl-5-hepten-2-one, guaiacol, 1-pentanol, pentanal, geranyl acetone, β-ionone, citral, methyl salicylate, 2-isobutylthiazole, 2-phenylethanol, and (E)-2-octenal. All analyses were performed in triplicate.

Carotenoid Analysis

Representative samples for each experimental group were used for carotenoid extractions. Five frozen fruits were powdered, and analyses were performed according to the method developed by Sérino et al. (2009).

RNA Isolation and Quantitative RT-PCR Analysis

Gene expression analyses were performed according to Fabi et al. (2012), following the "Minimum Information for Publication of Quantitative Real-Time PCR Experiments – MIQE" (Bustin et al., 2009). All the primers used for amplification are listed in **Supplementary Table 1**.

The total RNAs from five individual fruits at each experimental stage were extracted using Concert™ Plant RNA Reagent (Invitrogen™) and treated with the "Ambion® DNA-free™ DNase Treatment and Removal Reagents" kit (Invitrogen™) for genomic DNA removal. RNAs were verified by agarose gel electrophoresis. For the cDNA synthesis, 1 μg of RNAs, measured spectrophotometrically, was used in reverse transcription reactions according to the instructions for the "ImProm-II™ Reverse Transcription System" kit (Promega). In total, 10 μL of real-time PCR reactions were set using the "Power SYBR Green Master Mix" kit (Applied Biosystems). Reactions were run in a QuantStudio 7 Flex system (Applied Biosystems) programmed to remain at 95°C for 5 min, 60 cycles of 95°C for 15 s, 60°C for 30 s, and 72°C for 30 s. The TIP41-like protein (TIP) and expressed unknown protein (EXP) genes were used as the internal controls, and the relative expressions were calculated according to the method developed by Pfaffl (2001).

Statistical Analysis

SPSS Version 19.0 (SPSS Inc., Chicago, IL, United States) was used to perform statistical analysis. Data were analyzed using Student's *T*-tests or a one-way ANOVA with a subsequent Tukey's test to evaluate the effects of phytohormone treatments on VOCs and other evaluated parameters at different ripening stages. Differences were considered significant at $p < 0.05$. VOC data

[2]http://mona.fiehnlab.ucdavis.edu/

were uploaded to MetaboAnalyst 5.0 (Pang et al., 2021) for a heatmap, hierarchical clustering analysis (HCA), partial least squares discriminant analysis (PLS-DA), and variable importance in projection (VIP) score after normalization by median, log transformation, and Pareto scaling. Linear regression analysis was applied to gene expression during fruit ripening, with the phytohormone treatment data being used as predictor variables (x) and transcript levels as response variables (y) and assuming a significance level of $p < 0.05$.

RESULTS

Overview of Ethylene and Auxin Effects on Primary Ripening Parameters

The ripening of tomato (*S. lycopersicum* L.) fruits treated with one phytohormone (ETHY or IAA) or a combination thereof (ETHY + IAA) was monitored daily using peel color characterization and ethylene emission as primary ripening parameters (**Figure 1**). For the first Micro-Tom and Sweet Grape experiments, CTRL fruits (without treatment) were used as reference of ripening stage, what means that fruits of all groups were frozen according to the ripening stage of the CTRL. Photographs of Micro-Tom and Sweet Grape fruits during ripening are presented in **Figures 1A,D**, respectively. As expected, Micro-Tom fruits showed a color shift (**Figure 1B**), and their ethylene emission (**Figure 1C**) was accelerated by ethylene treatment. ETHY fruits showed ethylene emission approximately four times higher than that of CTRL fruits on the 1st day after harvest (DAH) and showed higher concentrations between 2 and 4 DAH. The IAA and ETHY + IAA groups showed delays at the beginning of the color shift; however, the climacteric peak was advanced compared with that of the CTRL group. Fruits treated with both plant hormones showed less variation in ethylene levels over the evaluated ripening days. On the 14th DAH, CTRL and ETHY fruits had reached full ripeness, while auxin-treated fruits (IAA and ETHY + IAA) did not show the observed red color. Sweet Grape fruits presented the same response profiles when subjected to the phytohormone treatments. As observed in the Micro-Tom experiment, the color shift (**Figure 1E**) and ethylene emission (**Figure 1F**) were accelerated by ethylene treatment and delayed by auxin. However, Sweet Grape tomatoes treated with both hormones showed results closer to those of ETHY than of IAA fruits. On the 6th DAH, ETHY fruits had reached full ripeness, while the same stage was reached on the 8th DAH for the CTRL and ETHY + IAA fruits. Those treated only with auxin did not show the observed red color.

Volatile Organic Compound Profiles of Micro-Tom and Sweet Grape Fruits Differ During Ripening

In total, 55 VOCs were identified as contributing to the tomato aroma, of which 24 were common to both cultivars (**Figure 2**). These compounds were mainly aldehydes (13), alcohols (10), ketones (9), and terpenoids (8); however, esters (6), furans (5), sulfur compounds (2), a benzene compound (1), and a carboxylic acid (1) were also found. The VOC profile of Sweet Grape fruits was more similar to the profile common to both cultivars than that of Micro-Tom, based on the compound classes. Both cultivars had many aldehydes (22% of the VOC profile in Micro-Tom and 26% in Sweet Grape); however, Sweet Grape fruits presented a profile rich in alcohols (23%) while Micro-Tom fruits had a profile with higher ketone (20%) and terpenoid contents (18%). Regarding the metabolic precursors, 26 compounds had fatty-acid precursors, 14 were obtained through amino-acid degradation, 10 were directly derived from isoprenoids, and five were derived from carbohydrates. Approximately 56% of the compounds comprising the VOC profile of Sweet Grape fruits were derived from fatty acids, making this profile more similar to that common to both cultivars (58%). Micro-Tom fruits had almost three times more isoprenoid-derived compounds than Sweet Grape. Moreover, only 44% of their VOC profile was derived from fatty acids.

Multivariate analyses were performed to differentiate the two cultivars' fruits at the mature green, breaker, and red stages based on their VOC profiles to understand the changes in the VOC profile during ripening and identify the similarities and differences between untreated Micro-Tom and Sweet Grape fruits (**Figure 3**). The PLS-DA showed that the two cultivars were undiscriminated at the mature green stage, differed at the breaker stage, and were clearly discriminated at the red stage (**Figure 3A**), with compounds of different classes contributing significantly to explaining the PLS-DA model (**Figure 3B**). Furans (2-ethylfuran and 2-propylfuran) and ketones (6-methyl-5-hepten-2-one and 3-pentanone), in addition to (Z)-3-hexen-1-ol and 2-isobutylthiazole, were the most important compounds in distinguishing the ripening stages in the two cultivars. The heatmap in **Figure 3C** indicates that the VOC profiles of both cultivars are more similar at the mature green and breaker stages than at the red stage, besides the existence of four VOC clusters. This indicated that, although specific compounds presented significant differences at the mature green stage, the cultivars presented more similar VOC profiles at the beginning than at the end, the latter being responsible for the final specific aroma of each cultivar.

Hormonal Regulation of Volatile Organic Compound Profiles During Tomato Ripening

Volatile organic compound profiles were analyzed by SPME–GC–MS at the mature green, breaker, and red stages to investigate how each hormonal treatment, and their combination, affected the tomato fruit aroma. Additionally, these were analyzed the 1st day after the treatments (01 DAH). In total, 62 VOCs were identified and confirmed in Micro-Tom fruits, of which 36 contributed to the aroma through odor notes and varying odor thresholds (**Supplementary Table 2**). The number and relative abundance (%) of compounds changed during ripening, in addition to being affected by the hormonal treatments (**Supplementary Figure 1**). These compounds were mainly aldehydes (8), ketones (7), and terpenoids (7); however, esters (5), furans (4), alcohols (3), a carboxylic acid (1),

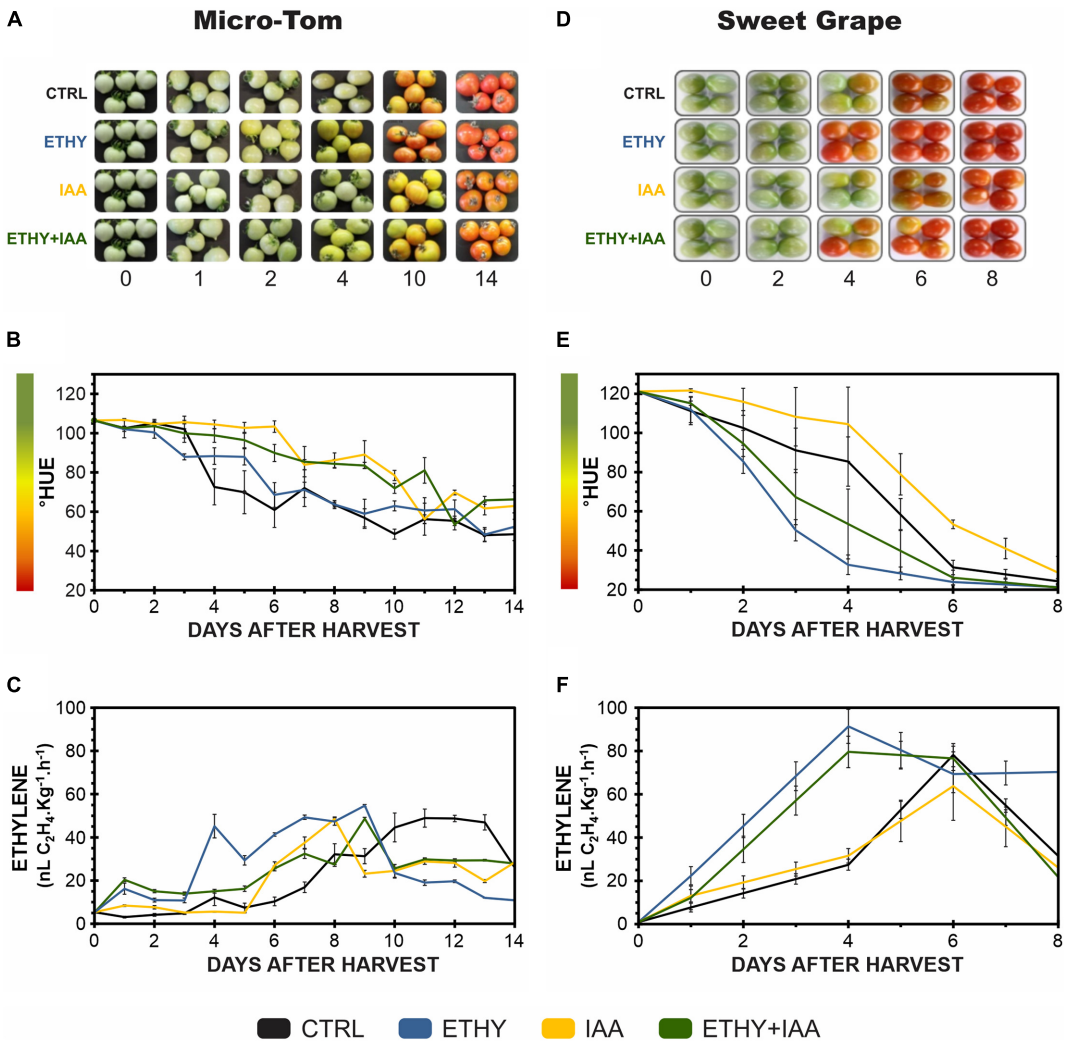

FIGURE 1 | Effects of ethylene, auxin, and both treatments on fruit color and ethylene emission during tomato (*Solanum lycopersicum* L.) cv. Micro-Tom **(A–C)** and Sweet Grape **(D–F)** fruit ripening. **(A,D)** Illustrative photographs of tomato fruits. **(B,E)** Peel color expressed by hue angle. **(C,F)** Ethylene emission. CTRL, control group. ETHY, ethylene-treated group. IAA, indole-3-acetic acid-treated group. ETHY + IAA, group treated with both hormones. Vertical bars represent the standard deviation of three replicates (n = 3).

and a sulfur compound (1) were also found (**Supplementary Figure 1A**). Regarding the metabolic precursors, sixteen compounds had fatty-acid precursors, nine were directly derived from isoprenoids, seven were obtained through amino-acid degradation, and four were derived from carbohydrates (**Supplementary Figure 1B**).

Multivariate analyses (PLS-DA, HCA, and VIP score) were performed to obtain an overall description of the effects of the plant hormones on VOC profiles at 01, 04, and 14 DAH. The PLS-DA results are shown in **Figure 4A** and **Supplementary Figure 2A**. At the beginning of ripening, Micro-Tom fruits presented VOC profiles comprising 28 of the 36 previously selected compounds, and 1 of these presented abundances with significant differences ($p < 0.05$) between at least two groups. PLS-DA explained 80.6% (component 1, 44.8%; component 2, 35.8%) of the total variance between the CTRL and treated samples. Volatile profile differences between ETHY and the other groups were evidenced primarily by component 1, while component 2 separated ETHY + IAA from the other groups. HCA reinforced differences visualized by PLS-DA, showing that the IAA and CTRL groups presented similar VOC profiles, with both being more similar to the ETHY + IAA than to the ETHY group. Based on PLS-DA, aldehydes, alcohols, and ketones were the most important in distinguishing the CTRL and treated samples. Four of the eight aldehydes found at the beginning of ripening [hexanal, 3-hexenal, (E)-2-hexenal, and (E)-2-octenal] presented abundances with significant differences ($p < 0.05$) between at least two groups. These compounds had fatty-acid precursors and were found with higher contents in the CTRL and ETHY + IAA fruits, derived from fatty acids, found with higher contents in the CTRL and ETHY fruits. Concerning the ketones, 2-heptanone and 1-hepten-3-one contents were higher

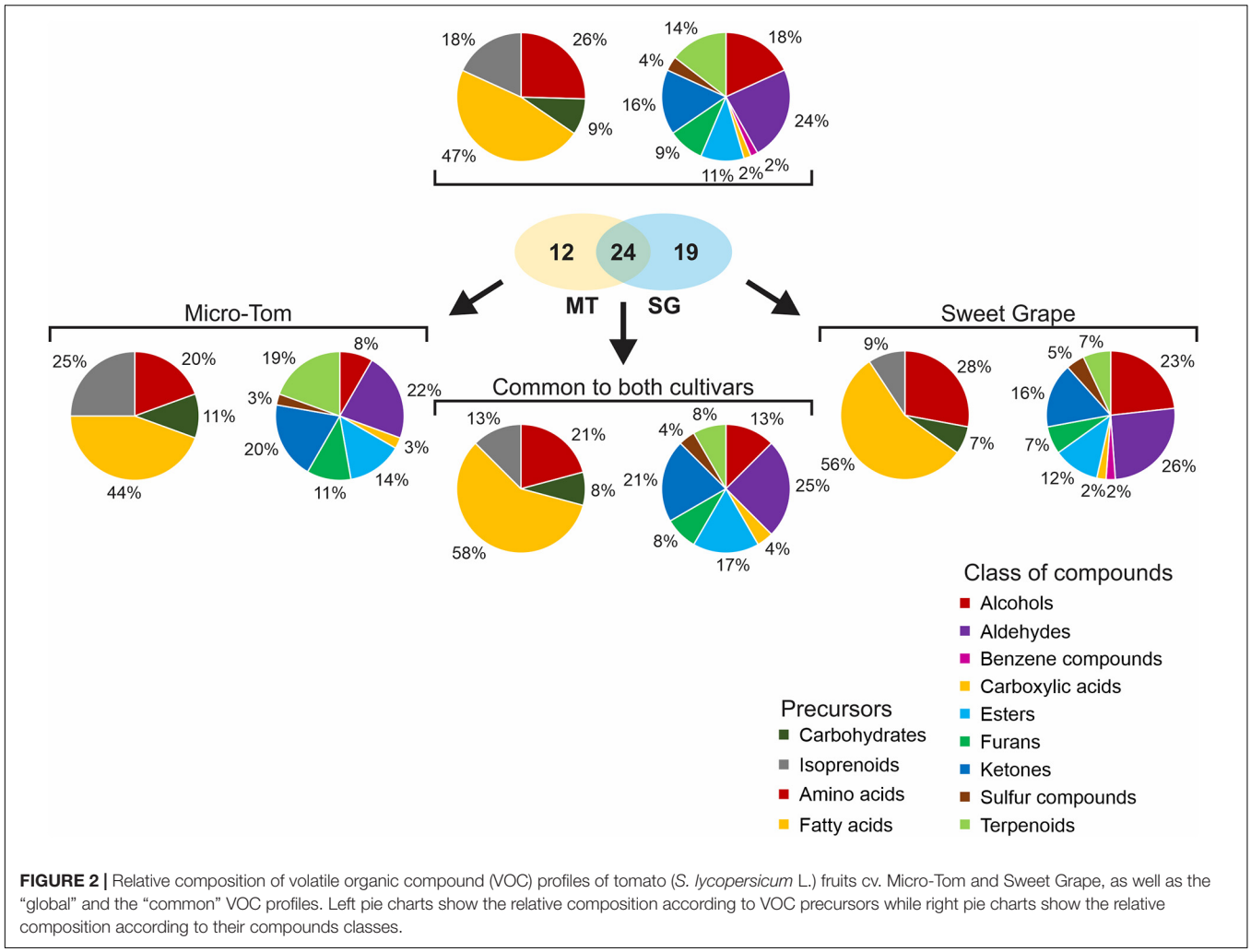

FIGURE 2 | Relative composition of volatile organic compound (VOC) profiles of tomato (*S. lycopersicum* L.) fruits cv. Micro-Tom and Sweet Grape, as well as the "global" and the "common" VOC profiles. Left pie charts show the relative composition according to VOC precursors while right pie charts show the relative composition according to their compounds classes.

in treated fruits and with significant difference between ETHY and CTRL groups.

After 3 DAH, when the CTRL fruits reached the breaker stage, their VOC profiles comprised 31 compounds, 26 of which presented abundances with significant differences ($p < 0.05$) between at least two groups. PLS-DA explained 85.2% of the total variation between samples. Component 1 (56.4%) separated the CTRL and IAA from ETHY and ETHY + IAA groups, and component 2 (28.8%) separated the ETHY and IAA groups from the CTRL and ETHY + IAA groups. In addition to PLS-DA, HCA showed that the ETHY and ETHY + IAA groups had more similar VOC profiles, as did the IAA and CTRL groups. Hexanal, (E)-2-hexenal, (E)-2-heptenal, 6-methyl-5-hepten-2-one, and 2-isobutylthiazole, aroma compounds important to consumer acceptance, were found with higher contents in ETHY and ETHY + IAA fruits. Other important VOCs, including 2-ethylhexan-1-ol and (E,E)-2,4-hexadienal, were found only in the CTRL fruits, while decanal and β-ionone were exclusive to ETHY fruits.

When the CTRL group was at the red stage, the VOC profiles comprised 27 compounds, only 11 of which presented abundances with significant differences ($p < 0.05$) between at least two groups. PLS-DA explained 62.1% of the total variation between samples. Component 1 (50.9%) separated the IAA from the other groups, while component 2 (11.2%) separated the IAA and ETHY from the CTRL and ETHY + IAA groups. Additionally, HCA showed that the IAA and ETHY + IAA fruits presented more similar VOC profiles, while ETHY presented profile more similar to the CTRL group. Regarding fatty-acid-derived volatiles, the accumulation of (Z)-3-hexen-1-ol, 2-propenal, hexanal, (E)-2-hexenal, (E)-2-octenal, 1-penten-3-one, and 1-hepten-3-one was affected by exogenous ethylene and auxin treatment during ripening. Hexanal and 2-propenal were found in higher contents in IAA and ETHY + IAA than in ETHY fruits. However, 1-penten-3-one, 1-hepten-3-one, (Z)-3-hexen-1-ol, and (E)-2-octenal were fatty-acid-derived volatiles with higher contents in ETHY than in IAA and ETHY + IAA fruits. Hexanal, (E)-2-hexenal, and (Z)-3-hexen-1-ol were fatty-acid-related volatiles found with higher contents in the tomato aroma profiles, at 10, three, and two times higher than the other VOCs, respectively. Other compounds, including citral, 6-methyl-5-hepten-2-one, and methyl salicylate, were found with

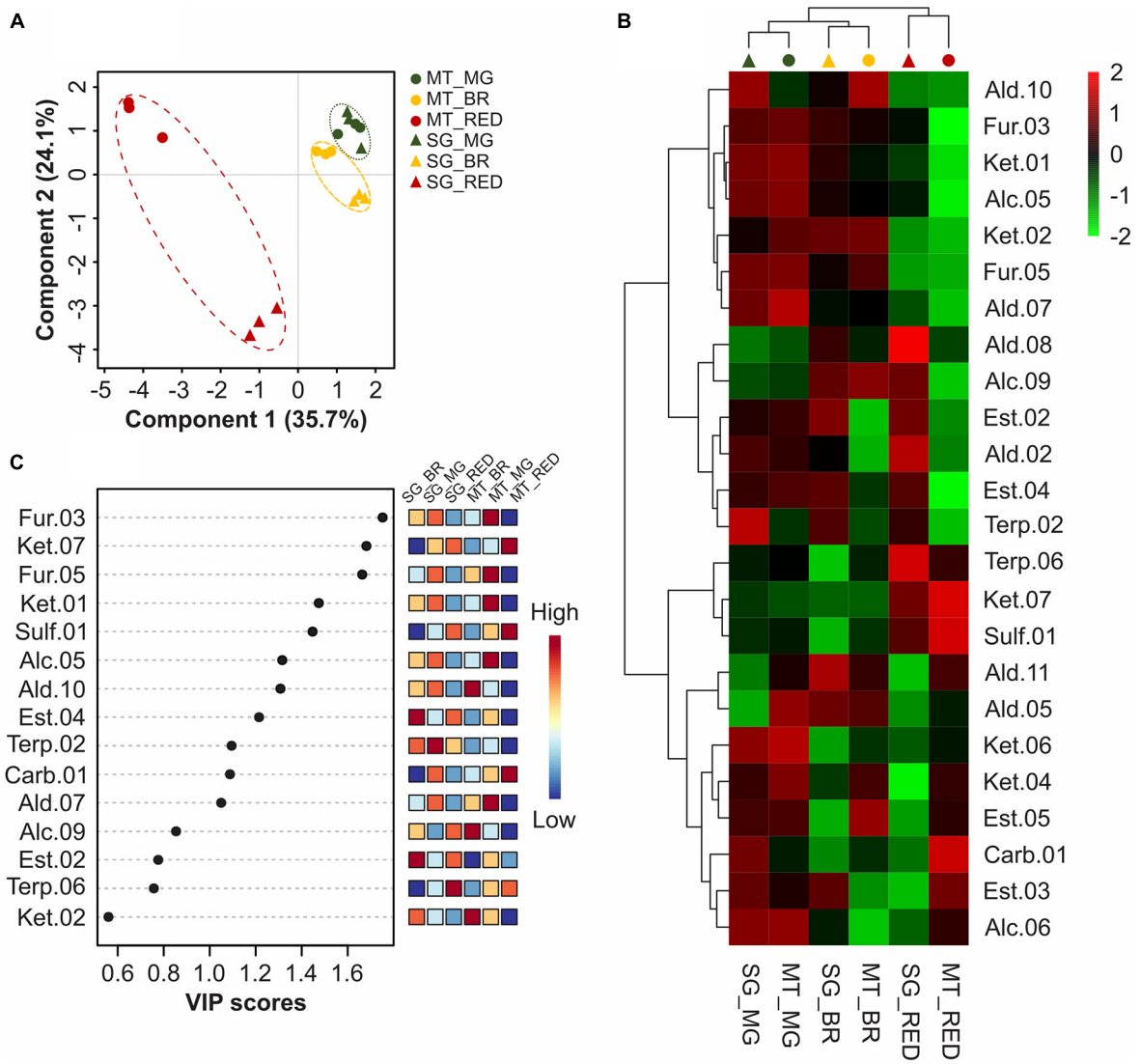

FIGURE 3 | Comparison of VOC profiles of tomato (*S. lycopersicum* L.) fruits cv. Micro-Tom and Sweet Grape at mature green, breaker, and red stages. **(A)** Partial least squares discriminant analysis (PLS-DA). **(B)** Heatmap with hierarchical clustering. **(C)** Variable importance in projection (VIP) score plot for the TOP15 most important VOCs identified by the PLS-DA. MT, Micro-Tom. SG, sweet grape. MG, mature green stage. BR, breaker stage. RED, red stage. Each value is presented as the mean (n = 3).

higher contents in ETHY and ETHY + IAA than in IAA fruits (**Supplementary Figure 3**).

Knowing that Micro-Tom and Sweet Grape fruits are highly similar at the onset of ripening but different when fully ripe, it was possible to identify the effects of the same hormonal treatments on Sweet Grape fruits and compare their results with the effects on Micro-Tom fruits. In total, 51 compounds were identified and confirmed in Sweet Grape fruits, of which 43 had the odoriferous note described in the literature (**Supplementary Table 3**). As observed in the Micro-Tom experiment, the number and the relative abundance (%) of compounds were affected by the ripening stage and hormonal treatment (**Supplementary Figure 4**). These compounds were mainly aldehydes (11), alcohols (10), and ketones (7); however, esters (5), terpenoids (3), furans (3), sulfur compounds (2), a benzene compound (1), and a carboxylic acid (1) were also found (**Supplementary Figure 4A**). Regarding the metabolic precursors, 24 compounds had fatty-acid precursors, 12 were obtained through amino-acid degradation, four were directly derived from isoprenoids, and three were derived from carbohydrates (**Supplementary Figure 4B**).

On the 1st DAH (**Figure 4B** and **Supplementary Figure 2B**), Sweet Grape fruits presented VOC profiles comprising 28 of the 43 previously selected compounds, only 10 of which presented abundances with significant differences ($p < 0.05$) between at least two groups. PLS-DA explained 78.2% (component 1, 45.3%; component 2, 32.9%) of the total variance between the CTRL and treated samples. Component 1 separated the

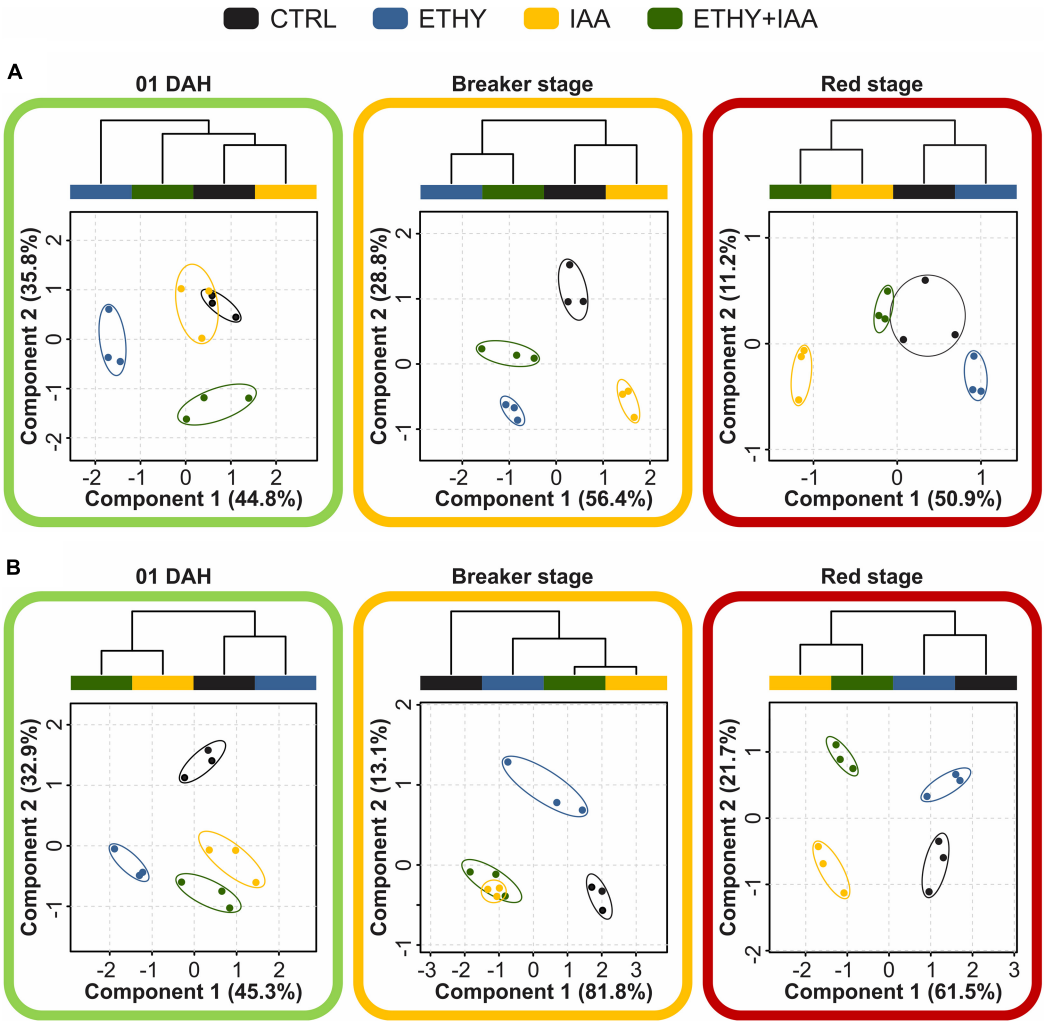

FIGURE 4 | Effects of ethylene, auxin, and both treatments on VOC profiles during tomato (*S. lycopersicum* L.) fruits cv. Micro-Tom and Sweet Grape ripening. **(A)** Partial least squares discriminant analysis (PLS-DA) and hierarchical clustering analysis (HCA) of VOCs identified in Micro-Tom fruits at 01, 04 (breaker), and 14 (red) days after harvest (DAH). **(B)** PLS-DA and HCA of VOCs identified in Sweet Grape fruits at 01, 04 (breaker), and 08 DAH (red).

ETHY from the other groups, while the VOC profile of the CTRL group was separated by component 2. HCA reinforced the differences visualized by PLS-DA, where the CTRL and ETHY groups presented more similar VOC profiles, as did the IAA and ETHY + IAA groups. At this stage, some important fatty-acid-derived volatiles, including 3-pentanone, hexanal, and (E)-2-hexenal, were affected by the hormone treatments and showed higher contents in the CTRL than in treated fruits. Other fatty-acid derivatives were found in higher levels only in the CTRL and ETHY (3-hexenal and 4-hexen-3-one) or in IAA and ETHY + IAA fruits (ethyl acetate). Treatment with auxin (IAA group) increased the contents of some compounds in the tomato fruits, including (Z)-2-penten-1-ol, 6-methyl-5-hepten-2-one, (E)-3-hexen-1-ol, and linalool. However, a different result was found for ETHY + IAA fruits, again suggesting that the ethylene and auxin responses overlapped within hours of treatment. Additionally, other compounds were important in distinguishing CTRL and treated fruits due to their presence in specific groups. Benzaldehyde was found only in CTRL fruits, while 2-methylbutanal was exclusive to ETHY and 1-hexanol to ETHY + IAA fruits; 1-penten-3-one and (E,E)-2,4-hexadienal were found only in CTRL and ETHY fruits.

When the CTRL fruits reached the breaker stage, the VOC profiles comprised 36 compounds; 28 of these presented abundances with significant differences ($p < 0.05$) between at least two groups. PLS-DA explained 94.9% of the total variation between samples. Component 1 (81.8%) separated the CTRL and ETHY from the IAA and ETHY + IAA groups, while component 2 (13.1%) separated the ETHY from the other groups. HCA showed that the IAA and ETHY + IAA fruits presented more similar VOC profiles, and both were more similar to the ETHY than to the CTRL fruits. Apart from 3-pentanone, which was inhibited by both exogenous treatments, fatty-acid-derived compounds were found in higher contents

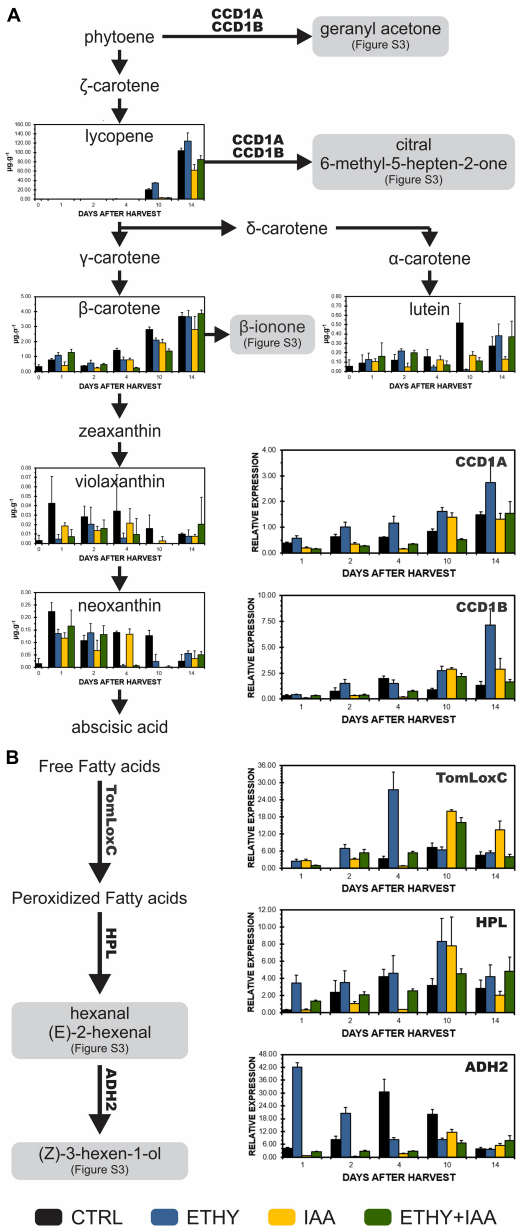

FIGURE 5 | Effects of ethylene, auxin, and both treatments on carotenoid contents and genes involved in the biosynthesis of VOCs during tomato (*S. lycopersicum* L. Micro-Tom) fruit ripening. **(A)** Carotenoid content and *CCD1A* and *CCD1B* gene expression. **(B)** *TomLoxC*, *HPL*, and *ADH2* gene expression. CTRL, control group. ETHY, ethylene-treated group. IAA, indole-3-acetic acid-treated group. ETHY + IAA, group treated with both hormones. CCD, carotenoid cleavage dioxygenase. LoxC, Lipoxygenase C. HPL, hydroperoxide lyase. ADH2, alcohol dehydrogenase-2. The contents of carotenoids are given in micrograms per gram of fruit on fresh weight basis. Vertical bars represent the standard deviation of three replicates (n = 3).

in treated fruits. The accumulation of hexanal, 3-hexenal, 4-hexen-3-one, (E)-2-hexenal, (E,E)-2,4-hexadienal, and (E)-2-octenal was induced by ethylene treatment, while (Z)-2-penten-1-ol, (E)-3-hexen-1-ol, and (Z)-3-hexen-1-ol were affected by auxin treatment and showed higher contents in the IAA and ETHY + IAA groups. Concerning amino-acid-derived volatiles, the accumulation of 2-methylbutanal, 2-ethylhexan-1-ol, benzaldehyde, and phenylethyl alcohol was more affected by exogenous ethylene treatment, and the (Z)-3-hexenyl and (E)-2-butenoate contents were higher in IAA and ETHY + IAA fruits. The accumulation of VOCs derived from carotenoids was separately affected by ethylene (6-methyl-5-hepten-2-one) and auxin (linalool), suggesting that each hormone regulated different sites in the geranyl diphosphate pathway. Regarding the other volatiles, heptanal, 2-propylfuran, and 2-phenylacetaldehyde were important in distinguishing ETHY fruits from the others due to their presence in this specific group. Ethyl acetate, (E)-2-pentenal, and 2-propylfuran were also important because they were exclusive to the CTRL and ETHY fruits, while 1-hexanol was exclusive to the CTRL and IAA fruits, and 1-pentanol was only found in the CTRL and ETHY + IAA fruits.

After the CTRL fruits fully ripened, their VOC profiles comprised 34 compounds, 18 of which presented abundances with significant differences ($p < 0.05$) between at least two groups. PLS-DA explained 83.2% (component 1, 61.5%; component 2, 21.7%) of the total variance between samples. The first component separated the CTRL and ETHY from the other groups, while the second separated the CTRL and IAA from ETHY and ETHY + IAA groups. As observed in the analysis with Micro-Tom fruits, HCA reinforced differences visualized by PLS-DA showing that the IAA and ETHY + IAA groups presented similar VOC profiles, as did the CTRL and ETHY groups. Fatty-acid-derived compounds had important roles in the VOC profiles of Sweet Grape tomato fruits; 3-hexenal, 2-propenal, and (E)-3-hexen-1-ol were affected by ethylene treatment, while the accumulation of propanoyl propanoate and (Z)-3-hexen-1-ol was induced by exogenous auxin. Hexanal was found in higher contents in CTRL and ETHY + IAA fruits, and 3-pentanone and 4-hexen-3-one accumulation was affected by both treatments, although these did not show higher contents in ETHY + IAA fruits. Regarding amino-acid-derived volatiles, 6-methyl-5-hepten-2-ol and 2-phenylacetaldehyde accumulation was more affected by exogenous ethylene treatment, and the 2-isobutylthiazole and 2-ethylhexan-1-ol contents were higher in IAA fruits. Additionally, other compounds were important in distinguishing the CTRL and treated fruits due to their presence in specific groups. Linalool was found only in CTRL fruits, while 2-methylbut-2-enal, (E)-2-pentenal, and benzaldehyde were exclusive to ETHY and (E,E)-2,4-hexadienal to ETHY + IAA fruits. Citral was found only in CTRL and ETHY fruits, while 1-pentanol and pentanoic acid were found only in ETHY and ETHY + IAA fruits.

Carotenoids and Aroma-Related Gene Changes Induced by Ethylene and Auxin

The expression profiles of key genes involved in the biosynthesis of VOCs during Micro-Tom fruit ripening, as well as the contents of tomato fruit carotenoids, were analyzed to investigate the origin of the observed changes in tomato fruit aromas and identify the effects of exogenous auxin and ethylene treatments

on the metabolic pathways of VOC precursors (**Figure 5**). *TomLoxC*, *HPL*, and *ADH2*, genes that code enzymes involved in the synthesis of C6 volatiles during tomato fruit ripening, were found to be upregulated by ethylene, especially at the breaker stage (04 DAH). Conversely, auxin delayed the increase in the transcriptional levels until 10 DAH. Carotenoid cleavage dioxygenase (CCD), specifically *CCD1A* and *CCD1B*, transcript levels were also downregulated by auxin and upregulated by ethylene. However, treatment with both hormones led to a similar expression pattern for IAA fruits. Lycopene contents on 10 and 14 DAH were lower in IAA and ETHY + IAA but higher in ETHY fruits, in accordance with the peel color measurements and photographs shown in **Figures 1A,D**. Additionally, no significant differences in the β-carotene levels were found between the groups.

Confirmation of Direct Hormonal Effects on Micro-Tom Volatile Organic Compound Profile

A third experiment was performed with Micro-Tom fruits to determine whether the changes observed in the VOC profiles after treatment occurred due to the direct effect of hormonal regulation or to a change in the "natural" ripening process. This new sampling directly compared the same maturation stages. The peel color shift and ethylene emission were analyzed daily (**Figures 6A,B**, respectively). Both parameters showed that ethylene treatment accelerated the ripening of the tomatoes, while the IAA and ETHY + IAA groups showed the opposite effect. Compared with the first Micro-Tom experiment, the CTRL fruits took on average 5 more days to reach the red stage (19 DAH), while ETHY fruits took 18 DAH and IAA and ETHY + IAA fruits took 21 DAH to fully ripen. The breaker stage was identified on the 8th DAH in CTRL fruits, the 6th day in ETHY fruits, and the 9th and 10th days in IAA and ETHY + IAA fruits.

In total, 92 VOCs were identified and confirmed at the mature green, breaker, and red stages; of these, 48 had the odoriferous note described in the literature (**Supplementary Table 4**). At the breaker stage (**Figure 6C**), VOC profiles comprised 41 compounds, 19 of which presented abundances with significant differences ($p < 0.05$) between at least two groups. PLS-DA explained 66.6% of the total variation between samples. Component 1 (44.1%) separated the CTRL and ETHY + IAA groups from the ETHY and IAA groups, while component 2 (22.5%) separated the IAA from the other groups. HCA made the differences visualized by PLS-DA clear, showing that the CTRL and ETHY + IAA groups presented similar VOC profiles, with both being more similar to the IAA group than to the ETHY group. Apart from methanethiol and 3-methylbutanal, which were inhibited by both exogenous treatments, the VOC profiles of ETHY fruits were characterized by high levels of fatty-acid-derived compounds, including hexanal, (Z)-2-hexenal, (E)-2-hexenal, and 1-octen-3-ol, besides the amino-acid derivatives 3-methyl-1-butanol and benzaldehyde. The accumulation of ethyl acetate, limonene, 3-methyl-1-pentanol, and nonanal was induced in the IAA fruits, while ETHY + IAA fruits only showed higher levels of o-guaiacol when compared with the other tomato fruits. Other fatty-acid and amino-acid derivatives had important roles in differentiating the VOC profiles. Pentanal and 1-pentanol were found in their lowest contents in IAA and ETHY + IAA fruits, in addition to 6-methyl-5-hepten-2-one, an important carotenoid-derived volatile. Normal accumulation of (E)-3-hexenal and 2-isobutylthiazole was observed in IAA fruits, with lower contents in ETHY and ETHY + IAA fruits. Ethylene and auxin treatment affected the accumulation of 2-methylfuran, 1-penten-3-one, and 2,3-pentanedione. However, these showed normal levels in ETHY + IAA fruits.

Upon reaching full ripening (**Figure 6D**), the VOC profiles comprised 44 compounds, 35 of which presented abundances with significant differences ($p < 0.05$) between at least two groups. PLS-DA explained 78.7% of the total variation between samples. Component 1 (52.2%) separated the ETHY from the other groups, while component 2 (26.5%) separated CTRL and ETHY + IAA fruits from those treated only with ethylene or auxin. HCA reinforced differences visualized by PLS-DA, where the CTRL and ETHY + IAA groups presented more similar VOC profiles, with both being more similar to the IAA group than to the ETHY group. Compared with treated tomatoes, CTRL fruits were better characterized by higher contents of ethyl acetate, 3-methylbutanal, 1-hexanol, and 2-isobutylthiazole. However, citral, methyl salicylate, geranyl acetone, and pentanoic acid levels were higher in ETHY and IAA fruits. Treated fruits showed altered contents of several compounds, mainly those derived from fatty acids and amino acids. The accumulation of 2,3-pentanedione, (E)-4-non-enal, and 1-octen-3-one was more affected by exogenous ethylene treatment, and the content of 3-methyl-1-pentanol was higher in IAA fruits. ETHY + IAA fruits showed higher levels of 1-penten-3-one and hexanal, two important fatty-acid-derived compounds. Benzene compounds derived from amino acids (benzaldehyde, and 1-phenylethanone) showed higher accumulation in fruits treated with ethylene and auxin, while aldehyde compounds derived from fatty acids were found in higher contents in ETHY and ETHY + IAA aromatic profiles, including (Z)-3-hexenal, heptanal, octanal, and nonanal. Additionally, methyl salicylate, 6-methyl-5-hepten-2-one, and 2-ethylfuran were also found in higher contents in ETHY and ETHY + IAA fruits.

DISCUSSION

Fruit ripening is a genetically coordinated process marked by significant biochemical changes in color, texture, flavor, aroma, and nutritional content that coincide with seed maturation. It is precisely regulated by a complex hormonal network. Ethylene plays critical roles in regulating fruit ripening, and the mechanisms involved in this regulatory process have been extensively studied. In contrast, however, even the exact role of auxin in ripening control remains unclear (Zhu et al., 2019). Regarding ethylene-auxin crosstalk in fruit ripening, recent studies have elucidated antagonistic effects of the two hormones. Auxin triggers ripening, acting on the transition between systems I and II of ethylene production, while ethylene acts as a positive

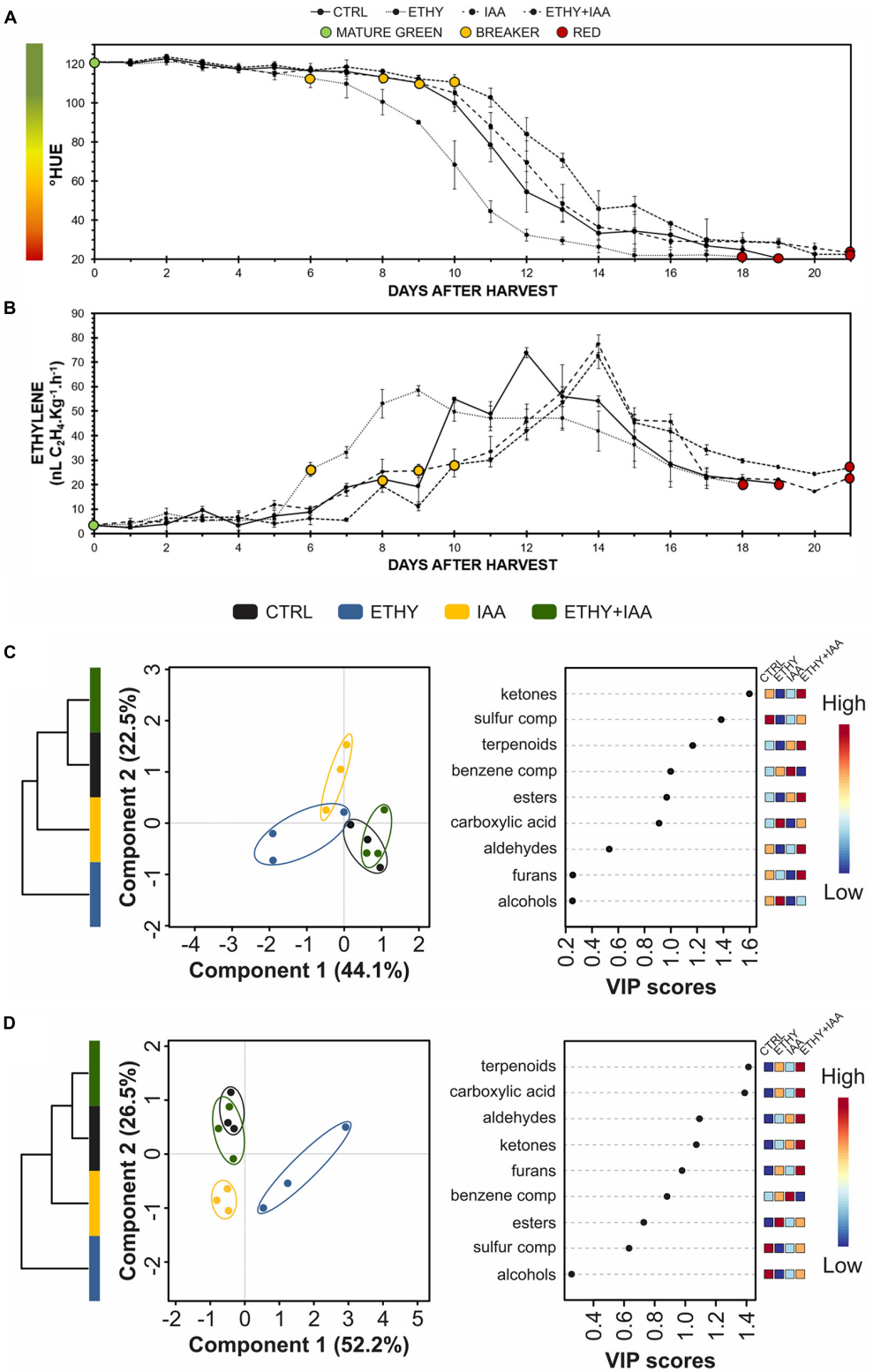

FIGURE 6 | Effects of ethylene, auxin, and both treatments on fruit color, ethylene emission, and VOC contents during Micro-Tom (*S. lycopersicum* L.) fruit ripening. **(A)** Peel color expressed by hue angle. **(B)** Ethylene emission. **(C,D)** Hierarchical clustering, partial least squares discriminant analysis (PLS-DA), and variable importance in projection (VIP) of VOCs produced at breaker **(C)** and red **(D)** stages. CTRL, control group. ETHY, ethylene-treated group. IAA, indole-3-acetic acid-treated group. ETHY + IAA, group treated with both hormones. Vertical bars represent the standard deviation of three replicates (n = 3).

regulator of auxin conjugation, affecting the transcript of SlGH3 genes besides the ethylene and auxin response factors (ERFs and ARFs, respectively) that regulate the biosynthesis of both hormones positively or negatively (Muday et al., 2012; McAtee et al., 2013; Kumar et al., 2014; Liu et al., 2018; Sravankumar et al., 2018; Yuan et al., 2019). In the present study, we evaluated the effects of exogenous ethylene and auxin treatments on primary ripening parameters, VOC profiles, carotenoid contents, and the expression of key genes related to regulating the metabolism of tomato aroma compounds. The color shift and ethylene emission, as primary ripening parameters, were evaluated in tomato fruits from two cultivars, Micro-Tom and Sweet Grape. Both parameters were accelerated by ethylene and delayed by auxin, while tomato fruits treated with both hormones presented intermediate responses. Previous studies reported that the tomato ripening process can be accelerated by exogenous ethylene treatment and delayed by exogenous auxin treatment (Li et al., 2016).

Tomato flavor is among the most important criteria for consumer preference (Xiao et al., 2017). The flavor of a fresh tomato is generally influenced by complex interactions of sugars, organic acids, and VOCs. During tomato ripening, several primary and secondary metabolisms are known to be controlled by regulatory mechanisms involving ethylene and auxin, including sugar and organic acid-related pathways. However, these mechanisms are not fully understood (Batista-Silva et al., 2018). VOC biosynthesis also appears to be regulated by both hormones, with most VOCs accumulating at the onset of ripening and peaking at full ripening (Klee and Giovannoni, 2011). Based on these aroma changes throughout ripening and the metabolic diversity of tomato cultivars, the VOC profiles of non-treated Micro-Tom and Sweet Grape fruits were compared at the mature green, breaker, and red stages. Multivariate analysis of the VOCs showed that the two cultivars were undiscriminated at the mature green stage, differed at the breaker stage, and clearly discriminated at the red stage. These changes observed during fruit ripening respond to hormonal regulation and several other intrinsic factors. This process promotes changes in the compound ratios and their interactions, producing changes in the tomato aroma (Klee and Giovannoni, 2011; Wang et al., 2016). Interestingly, based on their compound classes and precursors, Sweet Grape fruits presented VOC profiles more similar to the fully ripe tomato profile common to both cultivars than that of Micro-Tom.

Previous studies have shown the roles of ethylene, auxin, and other hormones (Wu et al., 2018) in regulating the VOC metabolism. However, in considering the crosstalk between the two hormones, our study unprecedentedly shows the effect of simultaneous treatment with ethylene and auxin on the tomato fruit VOC profile. Approximately 400 compounds were previously identified in tomato fruits. However, fewer than 10% are produced in sufficient quantities to be perceived by humans, being crucial in conferring the characteristic tomato aroma (Baldwin et al., 1998; Mathieu et al., 2009; Du et al., 2015). The other VOCs may provide background odors that impact the overall aroma quality (Klee and Giovannoni, 2011). Tieman et al. (2017) and Klee and Tieman (2018) described the principal contributors to the ripe tomato flavor that positively correlates with consumer preferences: 1-nitro-3-methylbutane; 1-penten-3-one; 1-octen-3-one; 2-isobutylthiazole; 2-methylbutanol; 3-methylbutanol; 2-phenylethanol; 3-pentanone; 6-methyl-5-hepten-2-ol; 6-methyl-5-hepten-2-one; benzaldehyde; heptaldehyde; nonyl aldehyde; phenylacetaldehyde; β-damascenone; citral; furaneol; geranyl acetone; guaiacol; β-ionone; isobutyl acetate; isovaleraldehyde; isovaleronitrile; isovaleric acid; methional; methyl salicylate; 1-pentanal; 1-pentanol; (E)-2-pentenal; 1-hexanal; 1-hexanol; (E)-3-hexen-1-ol; (Z)-3-hexen-1-ol; (E)-2-hexenal; (Z)-3-hexenal; (E)-2-heptenal; (E,E)-2,4-decadienal; and (Z)-4-decenal. These aroma compounds are derived from diverse precursors, including amino acids, fatty acids, and carotenoids (Klee and Tieman, 2013; Bauchet et al., 2017). In our study, Micro-Tom and Sweet Grape tomato fruits presented different VOC profiles but with similar numbers and relative areas when grouped by precursors and classes. Both cultivars had their volatile profiles changed by the hormonal treatments (ETHY, IAA, and ETHY + IAA) from the 1st day after treatment (01 DAH). VOCs derived from fatty acids [3-hexenal, (E)-2-hexenal, (Z)-3-hexen-1-ol, and others] and amino acids (2-isobutylthiazole, methyl salicylate, and others) were dominant (in number and relative area) in all the profiles of the two cultivars. Despite not representing a large chromatographic area in Sweet Grape fruits, compounds derived from carotenoids described by Tieman et al. (2017) and Klee and Tieman (2018) as positively correlated with consumer preferences (6-methyl-5-hepten-2-one and citral) were also identified in both cultivars. The multivariate analysis (PLS-DA and HCA) showed that hormonal treatments affected Micro-Tom and Sweet Grape fruits differently during the ripening process. At the breaker stage, ETHY + IAA fruits presented a VOC profile more similar to that of ETHY in Micro-Tom and that of IAA in Sweet Grape. However, the relationships between the experimental groups were the same in both cultivars when the CTRL reached the red stage. Fruits treated with both hormones presented VOC profiles more similar to that of fruits treated only with auxin, suggesting that the auxin and ethylene effects overlap.

MicroTom tomato is a mutant that expresses a truncated version of the DWARF gene, which encodes the SlCYP85A1 protein that catalyzes the oxidation of 6-deoxocastasterone to castasterone in brassinosteroid (BR) biosynthesis (Bishop et al., 1999). The mutation causes synthesis of BRs at reduced levels and results in the dwarf phenotype. The exogenous application of BRs in tomato fruits of non-defective varieties in the synthesis or response to BRs, leads to accelerated ripening, as well as the application of brassinazole, an inhibitor of BRs synthesis, delayed it (Vardhini and Rao, 2002; Zhu et al., 2015). Reports on other fruits, such as apples and pears, however, show an opposite effect of BRs on ripening, indicating different actions of this hormonal class in fruits (Ji et al., 2021). Overexpression of SlCYP90B3 in tomato fruit from Ailsa Craig cultivar resulted in higher levels of brassinolide in RED stage fruits compared to wild type, resulting in increased accumulation of carotenoids and volatile compounds characteristic of ripe fruit (Hu et al., 2020). The effect on such metabolisms is attributed to the influence of

brassinolide on the expression of ACS1 and ACO1 impacting ethylene synthesis, and it has been suggested that, in tomatoes and other fruits with similar responses to higher levels, BRs would be involved in the early stages of climacteric ethylene synthesis.

Despite the low levels of BRs synthesis, the MicroTom fruit can ripen, and responds to exogenous ethylene with increased expression of several genes related to ripening, including those associated to climacteric ethylene synthesis. This seems to indicate that BRs should act in more specific aspects of ripening and that are not yet completely clear. In our results, analyzing the ethylene synthesis profiles (**Figure 1**) in auxin-treated groups, it is possible to see different patterns between MicroTom and Sweet Grape fruits. To confirm whether such differences are due to the lower level of BRs in MicroTom fruits, further studies with designs directed toward this end would be needed.

In our results it is observed that the global volatile profile in fruits treated with auxin and auxin plus ethylene seems to be affected differently between MicroTom and Sweet Grape fruits, mainly in the RED stage (**Figure 4**). While in Sweet Grape fruits the Principal Component 1 (PC1) indicates a pronounced effect of auxin in groups IAA and IAA + ETHY, in MicroTom tomatoes this second group presented an overall profile that is closer to the control group. Other variables may have contributed to this different behavior; however, it is possible to suggest that the lower levels of BRs may somehow contribute to this effect. Studies point to crosstalk between BRs and auxins in root development and hypocotyl elongation (Nemhauser et al., 2004; Kim et al., 2006), suggesting that BRs act synergistically by increasing the expression of several auxin early responsive genes. Regarding fruit ripening, there are a lack of studies that can suggest what would be the interaction between BRs and auxins. However, if BRs and auxins act synergistically in fruits as well, the results observed for MicroTom tomatoes from the IAA + ETHY group suggest that normal levels of BRs would be necessary for the effects of auxin on the synthesis of volatile compounds.

Finally, several VOCs identified in Micro-Tom and Sweet Grape and affected by the ethylene and auxin treatments are derived from fatty acids or carotenoids and contribute significantly to the characteristic tomato aroma. The carotenoid-derived volatiles are produced by the cleavage of the fruit carotenoids, mainly lycopene and β-carotene (Klee and Tieman, 2018). In our study, we analyzed these main carotenoids in addition to violaxanthin, neoxanthin, and lutein. The expression profiles of *CCD1A* and *CCD1B* genes were also analyzed during Micro-Tom fruit ripening. Hexanal, (E)-2-hexenal, (Z)-3-hexen-1-ol, and other C6 volatiles are derived from fatty acids. They are formed during the cleavage of linoleic and linolenic acids by the action of lipoxygenase C (TomLoxC) followed by oxidative cleavage catalyzed by hydroperoxide lyase (HPL) to generate hexanal and (Z)-3-hexenal, which can be isomerized to (E)-2-hexenal. The alcohol dehydrogenase 2 (ADH2) catalyzes the reduction of hexanal, (Z)-3-hexenal, and (E)-2-hexenal to 1-hexanol, (Z)-3-hexen-1-ol, and (E)-2-hexen-1-ol, respectively (Speirs et al., 1998; Chen et al., 2004; Rambla et al., 2014). *TomLoxC*, *HPL*, and *ADH2* gene expression profiles were also analyzed during Micro-Tom fruit ripening. All these genes were found to be upregulated by ethylene and downregulated by auxin, which corroborates with their VOC levels produced in these pathways, lycopene contents, and appearances (see photographs).

The results described were obtained in experiments that applied the same experimental design, as most studies used exogenous hormonal treatment to investigate fruit ripening. Since the CTRL and treated fruits were sampled and analyzed on the same DAH, we compared fruit of the same age but at a different ripening stage, potentially leading to different interpretations. Therefore, a third experiment was performed with Micro-Tom fruits to determine whether the changes observed in the VOC profiles after treatment occurred due to the direct effect of hormonal regulation or as a secondary result of the changing ripening process. This new approach allowed us to observe that the fruits treated with both hormones presented a delay in the ripening process but maintained a VOC profile similar to that the CTRL fruits. Additionally, ETHY and ETHY + IAA fruits had higher contents of the compounds described by Klee and Tieman (2018) as positively correlated with consumer preferences (methyl salicylate, 6-methyl-5-hepten-2-one, and 2-ethylfuran). This result suggests that both hormones directly regulated pathways involved in VOC biosynthesis and highlights the antagonistic effects of ethylene and auxin on aroma metabolism.

CONCLUSION

Fruits of tomato cv. Micro-Tom and Sweet Grape had similar VOC profiles at the early ripening stage, with differences subsequently being observed during the ripening process. Although the two tomato fruit cultivars had distinct volatile profiles when fully ripe, Micro-Tom and Sweet Grape were similarly affected by exogenous ethylene and auxin treatments, with contrary responses to the hormones. Our results indicate that the auxin and ethylene effects on the aroma metabolism overlap and that this hormonal regulation is not cultivar-dependent. Both hormones played important roles in tomato VOC formation and regulate this metabolism differently, especially regarding carotenoid and fatty-acid-derived compounds. Moreover, the content of carotenoids and the expression of *CCD1A*, *CCD1B*, and other genes involved in forming VOCs important for the aroma of tomato fruits were also antagonistically affected by ethylene and auxin treatments. Furthermore, our results suggest that ethylene and auxin directly regulated the aroma-related pathways and VOC formation since fruits apparently at the same ripening stage had different VOC profiles.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

EP, ET, and BG: conceptualization. EP: funding acquisition, project administration, supervision, and writing—review and editing. EP and LF: resources. ET, BG, VB, and EL: investigation and formal analysis. ET: visualization and writing—original draft preparation. All authors contributed to the article and approved the submitted version.

FUNDING

This research was funded by São Paulo Research Foundation (FAPESP), grant number 2013/07914-8. The authors acknowledge the scholarships provided by FAPESP, the National Council for Scientific and Technological Development (CNPq), and the Coordination for the Improvement of Higher Education Personnel (CAPES).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.765897/full#supplementary-material

Supplementary Figure 1 | Number and relative area of volatile organic compounds (VOCs) in tomato (*Solanum lycopersicum* L. cv. Micro-Tom) fruits after ethylene, auxin, and both treatments at 0 (mature green), 04 (breaker), and 14 (red) days after harvest (DAH). **(A)** Number and relative area of VOCs according to the compound classes. **(B)** Number and relative area of VOCs obtained from the same precursors. CTRL, control group. ETHY, ethylene-treated group. IAA, indole-3-acetic acid-treated group. ETHY + IAA, group treated with both hormones. (n = 3).

Supplementary Figure 2 | Variable importance in projection (VIP) score plot for the most important volatile organic compound (VOC) classes identified by the PLS-DA shown in **Figure 4**. **(A)** VOCs identified in tomato (*S. lycopersicum* L. cv. Micro-Tom) fruits after ethylene, auxin, and both treatments at 01, 04 (breaker), and 14 (red) days after harvest (DAH). **(B)** VOCs identified in tomato (*S. lycopersicum* L. cv. Sweet Grape) fruits after ethylene, auxin, and both treatments at 01, 04 (breaker), and 08 DAH (red).

Supplementary Figure 3 | Normalized area of volatile organic compounds important to tomato (*S. lycopersicum* L. cv. Micro-Tom) fruits aroma after ethylene, auxin, and both treatments throughout ripening. CTRL, control group. ETHY, ethylene-treated group. IAA, indole-3-acetic acid-treated group. ETHY + IAA, group treated with both hormones. Each value is presented as the mean, and vertical bars represent the standard deviation of three replicates (n = 3).

Supplementary Figure 4 | Number and relative area of volatile organic compounds (VOCs) in tomato (*S. lycopersicum* L. cv. Sweet Grape) fruits after ethylene, auxin, and both treatments at 0 (mature green), 04 (breaker), and 08 (red) days after harvest (DAH). **(A)** Number and relative area of VOCs according to the compound classes. **(B)** Number and relative area of VOCs obtained from the same precursors. CTRL, control group. ETHY, ethylene-treated group. IAA, indole-3-acetic acid-treated group. ETHY + IAA, group treated with both hormones. (n = 3).

Supplementary Table 1 | Primers used in the RT-PCR.

Supplementary Table 2 | Identified volatile organic compounds in tomato (*Solanum lycopersicum* L. cv. Micro-Tom) fruit after ethylene, auxin, and both treatments throughout ripening. Data presented are p-values for mean comparisons between treated and control fruits (n = 3).

Supplementary Table 3 | Identified volatile organic compounds in tomato (*S. lycopersicum* L. cv. Sweet Grape) fruit after ethylene, auxin, and both treatments throughout ripening. Data presented are p-values for mean comparisons between treated and control fruits (n = 3).

Supplementary Table 4 | Identified volatile organic compounds in tomato (*S. lycopersicum* L. cv. Micro-Tom) fruits after ethylene, auxin, and both treatments at breaker and red stages. Data presented are p values for mean comparisons between treated and control fruits (n = 3).

REFERENCES

Bai, J., Baldwin, E. A., Imahori, Y., Kostenyuk, I., Burns, J., and Brecht, J. K. (2011). Chilling and heating may regulate C6 volatile aroma production by different mechanisms in tomato (*Solanum lycopersicum*) fruit. *Postharvest Biol. Technol.* 60, 111–120. doi: 10.1016/j.postharvbio.2010.12.002

Baldwin, E. A., Scott, J. W., and Bai, J. H. (2015). Sensory and chemical flavor analyses of tomato genotypes grown in florida during three different growing seasons in multiple years. *J. Am. Soc. Hortic. Sci.* 140, 490–503. doi: 10.21273/JASHS.140.5.490

Baldwin, E. A., Scott, J. W., Einstein, M. A., Malundo, T. M. M., Carr, B. T., Shewfelt, R. L., et al. (1998). Relationship between Sensory and Instrumental Analysis for Tomato Flavor. *J. Am. Soc. Hortic. Sci.* 123, 906–915. doi: 10.21273/JASHS.123.5.906

Batista-Silva, W., Nascimento, V. L., Medeiros, D. B., Nunes-Nesi, A., Ribeiro, D. M., Zsogon, A., et al. (2018). Modifications in Organic Acid Profiles During Fruit Development and Ripening: correlation or Causation?. *Front. Plant Sci.* 9:1689. doi: 10.3389/fpls.2018.01689

Bauchet, G., Grenier, S., Samson, N., Segura, V., Kende, A., Beekwilder, J., et al. (2017). Identification of major loci and genomic regions controlling acid and volatile content in tomato fruit: implications for flavor improvement. *New Phytol.* 215, 624–641. doi: 10.1111/nph.14615

Bishop, G. J., Nomura, T., Yokota, T., Harrison, K., Noguchi, T., Fujioka, S., et al. (1999). The tomato DWARF enzyme catalyses C-6 oxidation in brassinosteroid biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* 96, 1761–1766. doi: 10.1073/pnas.96.4.1761

Breitel, D. A., Chappell-Maor, L., Meir, S., Panizel, I., Puig, C. P., Hao, Y., et al. (2016). AUXIN RESPONSE FACTOR 2 intersects hormonal signals in the regulation of tomato fruit ripening. *PLoS Genet.* 12:e1005903. doi: 10.1371/journal.pgen.1005903

Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., et al. (2009). The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.* 55, 611–622. doi: 10.1373/clinchem.2008.112797

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

Cheng, G., Chang, P., Shen, Y., Wu, L., El-Sappah, A. H., Zhang, F., et al. (2020). Comparing the Flavor Characteristics of 71 Tomato (*Solanum lycopersicum*) Accessions in Central Shaanxi. *Front. Plant Sci.* 11:586834. doi: 10.3389/fpls.2020.586834

Du, X., Song, M., Baldwin, E., and Rouseff, R. (2015). Identification of ur volatiles and GC-olfactometry aroma profiling in two fresh tomato cultivars. *Food Chem.* 171, 306–314. doi: 10.1016/j.foodchem.2014.09.013

Exposito-Rodriguez, M., Borges, A. A., Borges-Perez, A., and Perez, J. A. (2008). Selection of internal control genes for quantitative real-time RT-PCR studies during tomato development process. *BMC Plant Biol.* 8:131. doi: 10.1186/1471-2229-8-131

Fabi, J. P., Cordenunsi, B. R., Barreto, G. P. M., Mercadante, A. Z., Lajolo, F. M., and Nascimento, J. R. O. (2007). Papaya fruit ripening: response to ethylene and 1-methylcyclopropene (1-MCP). *J. Agric. Food Chem.* 55, 6118–6123. doi: 10.1021/jf070903c.

Fabi, J. P., Seymour, G. B., Graham, N. S., Broadley, M. R., May, S. T., Lajolo, F. M., et al. (2012). Analysis of ripening-related gene expression in papaya using an Arabidopsis-based microarray. *BMC Plant Biol.* 12:242. doi: 10.1186/1471-2229-12-242

Food and Agriculture Organization Corporate Statistical Database [FAOSTAT] (2021). *FAOSTAT [WWW Document]*. Available online at: http://www.fao.org/faostat/en/#data (Accessed August 21, 2021)

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

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

Gapper, N. E., McQuinn, R. P., and Giovannoni, J. J. (2013). Molecular and genetic regulation of fruit ripening. *Plant Mol. Biol.* 82, 575–591. doi: 10.1007/s11103-013-0050-3

Guillon, F., Philippe, S., Bouchet, B., Devaux, M. F., Frasse, P., Jones, B., et al. (2008). Down-regulation of an auxin response factor in the tomato induces modification of fine pectin structure and tissue architecture. *J. Exp. Bot.* 59, 273–288. doi: 10.1093/jxb/erm323

Hao, Y., Hu, G., Breitel, D., Liu, M., Mila, I., Frasse, P., et al. (2016). Auxin response factor SlARF2 is an essential component of the regulatory mechanism controlling fruit ripening in tomato. *PLoS Genet.* 11:e1005649. doi: 10.1371/journal.pgen.1005649

Hu, S., Liu, L., Li, S., Zhiyong, S., Fanliang, M., Haoran, L., et al. (2020). Regulation of fruit ripening by the brassinosteroid biosynthetic gene SlCYP90B3 *via* an ethylene-dependent pathway in tomato. *Hortic. Res.* 7:163. doi: 10.1038/s41438-020-00383-0

Ilg, A., Bruno, M., Beyer, P., and Al-Babili, S. (2014). Tomato carotenoid cleavage dioxygenases 1A and 1B: relaxed double bond specificity leads to a plenitude of dialdehydes, mono-apocarotenoids and isoprenoid volatiles. *FEBS Open Bio* 4, 584–593. doi: 10.1016/j.fob.2014.06.005

Ji, Y. L., Qu, Y., Jiang, Z. Y., Yan, J. J., Chu, J. F., Xu, M. Y., et al. (2021). The mechanism for brassinosteroids suppressing climacteric fruit ripening. *Plant Physiol.* 185, 1875–1893. doi: 10.1093/plphys/kiab013

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

Kim, H., Park, P. J., Hwang, H. J., Lee, S. Y., Oh, M. H., and Kim, S. G. (2006). Brassinosteroid signals control expression of the AXR3/IAA17 gene in the crosstalk point with auxin in root development. *Biosci. Biotechnol. Biochem.* 70, 768–773. doi: 10.1271/bbb.70.768

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

Klee, H. J., and Tieman, D. M. (2013). Genetic challenges of flavor improvement in tomato. *Trends Genet.* 29, 257–262. doi: 10.1016/j.tig.2012.12.003

Klee, H. J., and Tieman, D. M. (2018). The genetics of fruit flavor preferences. *Nat. Rev.* 19, 347–356. doi: 10.1038/s41576-018-0002-5

Korasick, D. A., Enders, T. A., and Strader, L. C. (2013). Auxin biosynthesis and storage forms. *J. Exp. Bot.* 64, 2541–2555. doi: 10.1093/jxb/ert080

Kumar, R., Khurana, A., and Sharma, A. K. (2014). Role of plant hormones and their interplay in development and ripening of fleshy fruits. *J. Exp. Bot.* 65, 4561–4575. doi: 10.1093/jxb/eru277

Lee, J. H. J., Jayaprakasha, G. K., Rush, C. M., Crosby, K. M., and Patil, B. S. (2018). Production system influences volatile biomarkers in tomato. *Metabolomics* 14:99. doi: 10.1007/s11306-018-1385-1

Li, C., Hou, X., Qi, N., Liu, H., Li, H., Huang, D., et al. (2021). Insight into ripening-associated transcription factors in tomato: a review. *Sci. Hortic.* 288:110363. doi: 10.1016/j.scienta.2021.110363

Li, J., Tao, X., Li, L., Mao, L., Luo, Z., Khan, Z. U., et al. (2016). Comprehensive RNA-Seq Analysis on the Regulation of Tomato Ripening by Exogenous Auxin. *PLoS One* 11:e0156453. doi: 10.1371/journal.pone.0156453

Liu, M., Chen, Y., Chen, Y., Shin, J. H., Mila, I., and Audran, C. (2018). The tomato Ethylene Response Factor Sl-ERF.B3 integrates ethylene and auxin signaling *via* direct regulation of Sl-Aux/IAA27. *New Phytol.* 219, 631–640. doi: 10.1111/nph.15165

Mathieu, S., Cin, V. D., Fei, Z., Li, H., Bliss, P., Taylor, M. G., et al. (2009). Flavour compounds in tomato fruits: identification of loci and potential pathways affecting volatile composition. *J. Exp. Bot.* 60, 325–337. doi: 10.1093/jxb/ern294

McAtee, P., Karim, S., Schaffer, R., and David, K. (2013). A dynamic interplay between phytohormones is required for fruit development, maturation, and ripening. *Front. Plant Sci.* 4:79. doi: 10.3389/fpls.2013.00079

Muday, G. K., Rahman, A., and Binder, B. M. (2012). Auxin and ethylene: collaborators or competitors?. *Trends Plant Sci.* 17, 181–195. doi: 10.1016/j.tplants.2012.02.001

Nemhauser, J. L., Mockler, T. C., and Chory, J. (2004). Interdependency of brassinosteroid and auxin signaling in Arabidopsis. *PLoS Biol.* 2:e258. doi: 10.1371/journal.pbio.0020258

Pang, Z., Chong, J., Zhou, G., Morais, D., Chang, L., Barrette, M., et al. (2021). MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic Acids Res.* 49, W388–W396. doi: 10.1093/nar/gkab382

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 29, 2002–2007. doi: 10.1093/nar/29.9.e45

Qin, G., Wang, Y., Cao, B., Wang, W., and Tian, S. (2012). Unraveling the regulatory network of the MADS box transcription factor RIN in fruit ripening. *Plant J.* 70, 243–255. doi: 10.1111/j.1365-313X.2011.04861.x

Rambla, J. L., Tikunov, Y. M., Monforte, A. J., Bovy, A. G., and Granell, A. (2014). The expanded tomato fruit volatile landscape. *J Exp. Bot.* 65, 4613–4623. doi: 10.1093/jxb/eru128

Ross, J. J., Weston, D. E., Davidson, S. E., and Reid, J. B. (2011). Plant hormone interactions: how complex are they?. *Physiol. Plant.* 141, 299–309. doi: 10.1111/j.1399-3054.2011.01444.x

Sérino, S., Gomez, L., Costagliola, G., and Gautier, H. (2009). HPLC Assay of Tomato Carotenoids: validation of a Rapid Microextraction Technique. *J. Agric. Food Chem.* 57, 8753–8760. doi: 10.1021/jf902113n

Seymour, G. B., Chapman, N. H., Chew, B. L., and Rose, J. K. C. (2013). Regulation of ripening and opportunities for control in tomato and other fruits. *Plant Biotechnol. J.* 11, 269–278. doi: 10.1111/j.1467-7652.2012.00738.x

Speirs, J., Lee, E., Holt, K., Yong-Duk, K., Scott, N. S., and Loveys, B. (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

Sravankumar, T., Akash, Naik, N., and Kumar, R. (2018). A ripening-induced SlGH3-2 gene regulates fruit ripening *via* adjusting auxin-ethylene levels in tomato (*Solanum lycopersicum* L.). *Plant Mol. Biol.* 98, 455–469. doi: 10.1007/s11103-018-0790-1

Su, L., Diretto, G., Purgatto, E., Danoun, S., Zouine, M., Li, Z., et al. (2015). Carotenoid accumulation during tomato fruit ripening is modulated by the auxin-ethylene balance. *BMC Plant Biol.* 15:114. doi: 10.1186/s12870-015-0495-4

Tieman, D., Zhu, G., Resende, M. F. R., Lin, T., Nguyen, C., Bies, D., et al. (2017). A chemical genetic roadmap to improved tomato flavor. *Science* 355, 391–394. doi: 10.1126/science.aal1556

Tucker, G., Yin, X., Zhang, A., Wang, M. M., Zhu, Q., Liu, X., et al. (2017). Ethylene and fruit softening. *Food Qual. Saf.* 1, 253–267. doi: 10.1093/fqsafe/fyx024

Vardhini, B. V., and Rao, S. S. R. (2002). Acceleration of ripening of tomato pericarp discs by brassinosteroids. *Phytochemistry* 61, 843–847. doi: 10.1016/s0031-9422(02)00223-6

Wang, L. B., Baldwin, E. A., and Bai, J. H. (2016). Recent advance in aromatic volatile research in tomato fruit: the metabolisms and regulations. *Food Bioproc. Technol.* 9, 203–216. doi: 10.1007/s11947-015-1638-1

Wang, X., Meng, J., Deng, L., Wang, Y., Liu, H., Yao, J. L., et al. (2021). Diverse Functions of IAA-Leucine Resistant PpILR1 Provide a Genic Basis for Auxin-Ethylene Crosstalk During Peach Fruit Ripening. *Front. Plant Sci.* 12:655758. doi: 10.3389/fpls.2021.655758

Wu, Q., Tao, X. Y., Ai, X. Z., Luo, Z. S., Mao, L. C., Ying, T. J., et al. (2018). Contribution of abscisic acid to aromatic volatiles in cherry tomato (*Solanum lycopersicum* L.) fruit during postharvest ripening. *Plant Physiol. Biochem*. 130, 205–214. doi: 10.1016/j.plaphy.2018.06.039

Xiao, Z., Wu, Q., Niu, Y., Liu, Q., Chen, F., Ma, N., et al. (2017). Optimization of headspace solid-phase micro-extraction and its application in analysis of volatile compounds in cherry tomato by gas chromatography. *Food Anal. Methods* 10, 596–609. doi: 10.1007/s12161-016-0622-3

Yuan, Y., Xu, X., Gong, Z., Tang, Y., Wu, M., Yan, F., et al. (2019). Auxin response factor 6A regulates photosynthesis, sugar accumulation, and fruit development in tomato. *Hortic. Res.* 6:85. doi: 10.1038/s41438-019-0167-x

Zhao, J., Sauvage, C., Zhao, J., Bitton, F., Bauchet, G., Liu, D., et al. (2019). Meta-analysis of genome-wide association studies provides insights into genetic control of tomato flavor. *Nat. Commun*. 10:1534. doi: 10.1038/s41467-019-09462-w

Zhu, G., Wang, S., Huang, Z., Zhang, S., Liao, Q., Zhang, C., et al. (2018). Rewiring of the fruit metabolome in tomato breeding. *Cell* 172, 249–261. doi: 10.1016/j.cell.2017.12.019

Zhu, T., Tan, W. R., Deng, X. G., Zheng, T., Zhang, D. W., and Lin, H. H. (2015). Effects of brassinosteroids on quality attributes and ethylene synthesis in postharvest tomato fruit. *Postharvest Biol. Technol*. 100, 196–204. doi: 10.1016/j.postharvbio.2014.09.016

Zhu, X. Y., Ye, L. L., Ding, X. C., Gao, Q. Y., Xiao, S. L., Tan, Q. Q., et al. (2019). Transcriptomic analysis reveals key factors in fruit ripening and rubbery texture caused by 1-MCP in papaya. *BMC Plant Biol*. 19:309. doi: 10.1186/s12870-019-1904-x

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.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership