Fan et al. *Horticulture Research* (2021)8:66
https://doi.org/10.1038/s41438-021-00502-5

Horticulture Research
www.nature.com/hortres

ARTICLE Open Access

Strawberry sweetness and consumer preference are enhanced by specific volatile compounds

Zhen Fan[1], Tomas Hasing[2], Timothy S. Johnson[3], Drake M. Garner[3], Michael L. Schwieterman[4], Christopher R. Barbey[1], Thomas A. Colquhoun[3], Charles A. Sims[5], Marcio F. R. Resende[6] and Vance M. Whitaker[1]

Abstract

Breeding crops for improved flavor is challenging due to the high cost of sensory evaluation and the difficulty of connecting sensory experience to chemical composition. The main goal of this study was to identify the chemical drivers of sweetness and consumer liking for fresh strawberries (*Fragaria* × *ananassa*). Fruit of 148 strawberry samples from cultivars and breeding selections were grown and harvested over seven years and were subjected to both sensory and chemical analyses. Each panel consisted of at least 100 consumers, resulting in more than 15,000 sensory data points per descriptor. Three sugars, two acids and 113 volatile compounds were quantified. Consumer liking was highly associated with sweetness intensity, texture liking, and flavor intensity, but not sourness intensity. Partial least square analyses revealed 20 volatile compounds that increased sweetness perception independently of sugars; 18 volatiles that increased liking independently of sugars; and 15 volatile compounds that had positive effects on both. Machine learning-based predictive models including sugars, acids, and volatiles explained at least 25% more variation in sweetness and liking than models accounting for sugars and acids only. Volatile compounds such as γ-dodecalactone; 5-hepten-2-one, 6-methyl; and multiple medium-chain fatty acid esters may serve as targets for breeding or quality control attributes for strawberry products. A genetic association study identified two loci controlling ester production, both on linkage group 6 A. Co-segregating makers in these regions can be used for increasing multiple esters simultaneously. This study demonstrates a paradigm for improvement of fruit sweetness and flavor in which consumers drive the identification of the most important chemical targets, which in turn drives the discovery of genetic targets for marker-assisted breeding.

Introduction

The cultivated strawberry (*Fragaria* ×*ananassa*) is one of the most widely grown fruit crops in the world due to its sweet and aromatic flavor and health-associated compounds including anthocyanins, antioxidants, fiber and ellagic acid[1]. Breeding for flavor in strawberry has been challenging due to the chemical complexity of the trait and the variability of the production environment. Targeted selection of flavor-associated chemicals in strawberry has thus far been restricted to increasing sugar content. In spite of improvements in eating quality over time, most strawberries on the market still do not meet consumer expectations of sweetness and flavor[2]. Efforts at flavor improvement must be reinforced by investigating the relationships between fruit chemical diversity and consumer preference.

During strawberry fruit ripening, changing auxin levels drive the accumulation of sugars and their derivatives, as well as secondary metabolites[3,4]. Sucrose, glucose, and fructose comprise the major soluble sugars in strawberry and are rapidly translocated from photosynthesizing organs to fruit, in synchrony with reductions in other sugars including xylose, galactose, sugar phosphates and sugar alcohols[3]. Besides sugars, amino acids, phenolic compounds and volatiles are the main indicators of fruit

Correspondence: Vance M. Whitaker (vwhitaker@ufl.edu)
[1]Horticultural Sciences Department, University of Florida, IFAS Gulf Coast Research and Education Center, Wimauma, FL, USA
[2]Elo Life Systems, Durham, NC, USA
Full list of author information is available at the end of the article

© The Author(s) 2021, corrected publication 2021
Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

Downloaded from https://academic.oup.com/hr/article/doi/10.1038/s41438-021-00502-5/6446614 by guest on 05 August 2026

ripening. Strawberry aroma is determined by low-molecular-weight volatile compounds generated during ripening. Over 360 volatiles have been identified in the aroma of cultivated and wild strawberries, with about 280 of them reported in cultivated strawberry[5]. Esters, primarily methyl and ethyl esters, constitute 25–90% of the total abundance of volatiles in strawberry and provide fruity notes[6]. Lactones, cyclic esters providing aromas similar to those in peach, are prominent volatiles in some varieties[7]. Aldehydes such as 2-hexenal, (E)- and 3-hexenal, (Z)- contribute to green or fresh aromas[8]. Furanones such as furaneol and mesifurane are often associated with sweet aromas[9].

Sensory analyses are necessary to comprehensively characterize fruit flavors, which are highly influenced by retronasal olfaction[10]. Odor Activity Values (OAVs), which are ratios of concentration to odor thresholds, were adopted in early studies to identify potent volatile compounds in strawberry[11]. These studies identified several compounds as important to strawberry flavor including (Z)-3-hexenal (green), 4-hydroxy-2,5-dimethyl-3(2H)-furanone (sweet), methyl butanoate (fruity), and ethyl butanoate (fruity). However, OAVs do not account for interactions between the matrix and the volatile, since the denominator only reflects odor intensity in a water solution[12]. Furthermore, OAVs are only based on orthonasal olfaction which is a different sensory experience than retronasal olfaction. Gas chromatography-olfactometry (GC-O) is another powerful tool to identify potent volatiles in food. Comparable results were obtained by GC-O and OAV for strawberry, agreeing on the most intense odorants[8,13]. However, while these studies identified potent volatiles, synergy among volatile compounds to produce human sensory responses and interactions between taste and retronasal olfaction were unexplored.

Studies on tomato flavor have highlighted the flaws of using OAVs exclusively for determining volatiles important to sensory perceptions[14]. Large-scale sensory and chemical studies have facilitated identification of flavor- or taste- associated volatiles[15], allowing breeding efforts to focus on genetic improvement of a smaller set of important volatiles. Once natural genetic variations responsible for volatile biosynthesis within heirloom tomato populations were exploited, DNA marker-assisted breeding techniques were utilized to introgress desirable alleles for volatile biosynthesis genes into modern genetic backgrounds[16]. A recent study in blueberry identified genomic regions associated with those compounds and also explored correlations between sensory attributes and volatile compounds[17]. A previous study in strawberry by the present authors, utilizing approaches and techniques similar to Tieman et al.[15], linked sugars, acids, and volatile data to consumer panel data for 35 strawberry genotypes[18]. Significant correlations between volatiles and sweetness and flavor intensities were found. Due to the two-year time-frame of the study, total sample size was limited to 54, and chemicals influencing consumer liking were not explored[18]. Marker-assisted selection in strawberry breeding for some flavor-associated volatiles has recently become a reality[19]. There is now a need to identify more sensory-associated volatile compounds in order to breed more comprehensively for strawberry flavor.

Genetic background is a major factor influencing primary and secondary metabolites in fresh strawberry fruit. Multiple studies have found that volatile compounds vary both qualitatively and quantitatively among strawberry varieties, leading to flavor differences[5–7,20]. Notable differences in volatile profiles have been observed in popular cultivars grown in the US, with more than ten-fold difference observed in all classes of volatiles[7,21,22]. Distinguishable separations between summer-grown European strawberry cultivars and two winter-grown Florida cultivars were found for 12 key strawberry volatiles[8]. During domestication, some desirable aroma-active volatiles have been lost over time[23]. Methyl cinnamate and methyl anthranilate, which are abundant in wild ancestor *F. vesca* and elicit pleasant aromas, are not detectable in modern cultivars[24]. Temperature is another critical factor that influences strawberry quality. In a single winter production season, temperature fluctuations drove substantial changes in fruit metabolic composition[25]. Cooler nights enhanced both sucrose accumulation and production of aroma compounds in a greenhouse setting[26]. Other environmental factors influencing strawberry fruit chemical abundance include light[27], maturity, postharvest storage[6], cultivation system[28], and location[29].

Recent studies have found that specific volatiles are capable of enhancing, not just flavor perception, but also sweetness perception in fresh fruit[15,18]. Hence, quantification of volatiles, sugars and acids should allow some degree of predictive ability for sensory characteristics, including sweetness. The recent development of machine learning models has facilitated prediction of sensory attributes and hedonic experiences based on chemistry. Partial least square (PLS) and its derivatives have been tested and validated for prediction of sensory attributes in different fruits, vegetables, and drinks[30–34]. In one example, artificial neural networks were used to simulate non-linear relationships for improvement of prediction accuracy for flavor intensity in blackcurrant[35]. To date, no predictive models for sensory characteristics or liking have been reported for fresh strawberries.

The main objectives of present study were to (1) explore the chemical drivers of consumer preference in fresh strawberries; (2) identify volatile compounds enhancing sweetness perception independently of sugars; and (3) predict sweetness perception and liking from chemical

Downloaded from https://academic.oup.com/hr/article/doi/10.1038/s41438-021-00502-5/6446614 by guest on 05 August 2026

data. This work builds on previous research[18], increasing the sample size approximately three-fold and applying updated mathematical and predictive models.

Results

Relationships among sensory and hedonic consumer responses

Over the course of seven years, 384 consumer panelists with an average age of 31.7 participated in evaluating strawberry samples. Panelists were 35% male and 65% female. Significant correlations were found among all pair-wise comparisons of sensory and hedonic ratings (Supplementary Table S1A). Sweetness ($r = 0.714$*), texture liking ($r = 0.783$*) and flavor intensity ($r = 0.688$*) were highly correlated with overall liking. Sourness ($r = 0.298$*) had a significant but lower positive correlation with overall liking. Notably, sweetness had significant correlations with all other sensory ratings ($r = 0.411$*, 0.592* and 0.838* for sourness, texture and flavor intensity, respectively). To remove the confounding effect of sweetness, partial correlations were calculated between liking and other attributes while controlling for sweetness (Supplementary Table S1B). After controlling for sweetness, the correlation between sourness and liking reduced to an insignificant level ($r = 0.008$), which was consistent with previous findings[18]. Conversely, this result was at odds with negative effects of sourness on liking found in other fruits[34,36]. In hierarchical clustering, liking, strawberry flavor, and sweetness were grouped in one cluster with 100% probability, separate from the sourness cluster with 0% probability (Fig. 1).

Due to a diverse collection of strawberry samples from different genetic backgrounds and harvest dates, patterns in perceived sweetness and consumer liking were distinguishable among genotypes and harvest dates. The average liking score was 25.8 across all samples. The top five samples for consumer liking were 'Festival' (twice), Sensation® 'Florida127', 'Albion' and Proprietary 2 (Supplementary Table S2). However, a large amount of the variability in sensory attributes was driven by temperature. Declines in sweetness and consumer liking over increasing temperature (at different harvest dates) was illustrated with Sensation® 'Florida127', 'Florida Radiance' and 'Florida Beauty' in 2016 and 2017 seasons (Fig. 2). The same pattern was also observed for total sugars but not total volatiles (Fig. 2). Intersections among lines suggest some level of genotype by environment (G×E) interactions underlying fruit sensory attributes.

Key chemical compounds contributing to sweetness and liking

Chemical analysis of 148 samples yielded 118 compounds including 3 sugars, 2 acids, and 113 volatiles (Supplementary Table S3). Besides sugars and acids, 59 volatiles were common among all three datasets corresponding to three periods: 2011–2012, 2013–2015 and 2016–2017. Most of the 59 common volatiles were esters (35), followed by aldehydes (7) and ketones (7).

As a first step in identifying the chemical drivers of hedonic experience, correlations among all sensory and chemical attributes were examined. A total of 646 significant correlations ($p < 0.01$) were detected among all variables (Fig. 3, Supplementary Table S4). Glucose ($r = 0.63$ and $r = 0.55$) and sucrose ($r = 0.62$ and $r = 0.60$) had the highest correlations with both sweetness and liking. In addition to these two sugars, 2-pentenal, (E)- (V12); 1-penten-3-one (V3); heptanal (V37); fructose; butanoic acid, butyl ester (V43); butanoic acid, 3-methyl-, butyl ester (V51); 5-hepten-2-one, 6-methyl- (V42); and γ-dodecalactone (V81) were the compounds with the highest correlations ($r > 0.4$) to liking. The list of highly correlated volatiles ($r > 0.4$) with sweetness included those correlated with liking, plus acetic acid, hexyl ester (V46); butanoic acid, octyl ester (V69); acetic acid, butyl ester (V21); and butanoic acid, hexyl ester (V62). The number of volatiles negatively correlated with liking or sweetness was smaller. 2-Hexen-1-ol, acetate, (E)- (v47) exhibited the strongest negative correlations with both liking ($r = -0.27$) and sweetness ($r = -0.21$). In a network analysis, significant correlations with liking and sweetness were visualized with edges connecting to sugars; multiple esters; γ-dodecalactone (V81); 5-hepten-2-one, 6-methyl- (V42); 1-penten-3-one (V3); heptanal (V37); 2-pentenal, (E)- (V12) (Fig. 4).

Most volatiles significantly correlated with sweetness and liking were corroborated in the PLS models. PLS models were able to explain more than 80 percent of variation for sweetness in each of the data subsets ($R^2 = 82.93\%$, 80.96% and 93.37%, respectively). PLS is advantageous when analyzing small samples sizes, and the highest R^2 was obtained for the 3rd dataset which has the smallest sample size. VIP tests revealed 20 important volatiles (VIP > 1) common to the three subsets in addition to glucose, fructose and sucrose (Table 1 and Supplementary Table S5). These consist of 10 esters, 5 aldehydes, 2 lactones, 2 ketones, and 1 alcohol. Sixteen of these volatiles passed *t*-tests derived from linear models at Bonferroni corrected *p*-values < 0.05, indicating that they increased sensory sweetness independent of total sugars. Important esters included seven butanoic acid esters. Among them, butanoic acid, butyl ester (V43) has the highest *r* of 0.51. Two abundant lactones, γ-dodecalactone (V81) and γ-decalactone (V74) were both selected in the PLS, also showing high correlations with sweetness ($r = 0.45$ and 0.33). 5-Hepten-2-one, 6-methyl- (V42) with floral aroma was one of the two ketones showing enhancing effect on sweetness with high correlation ($r = 0.45$).

A slightly lower portion of variation in response was explained in the liking models (77.56%, 70.49% and 92.41%, respectively). Sucrose was the most important

Downloaded from https://academic.oup.com/hr/article/doi/10.1038/s41438-021-00502-5/6446614 by guest on 05 August 2026

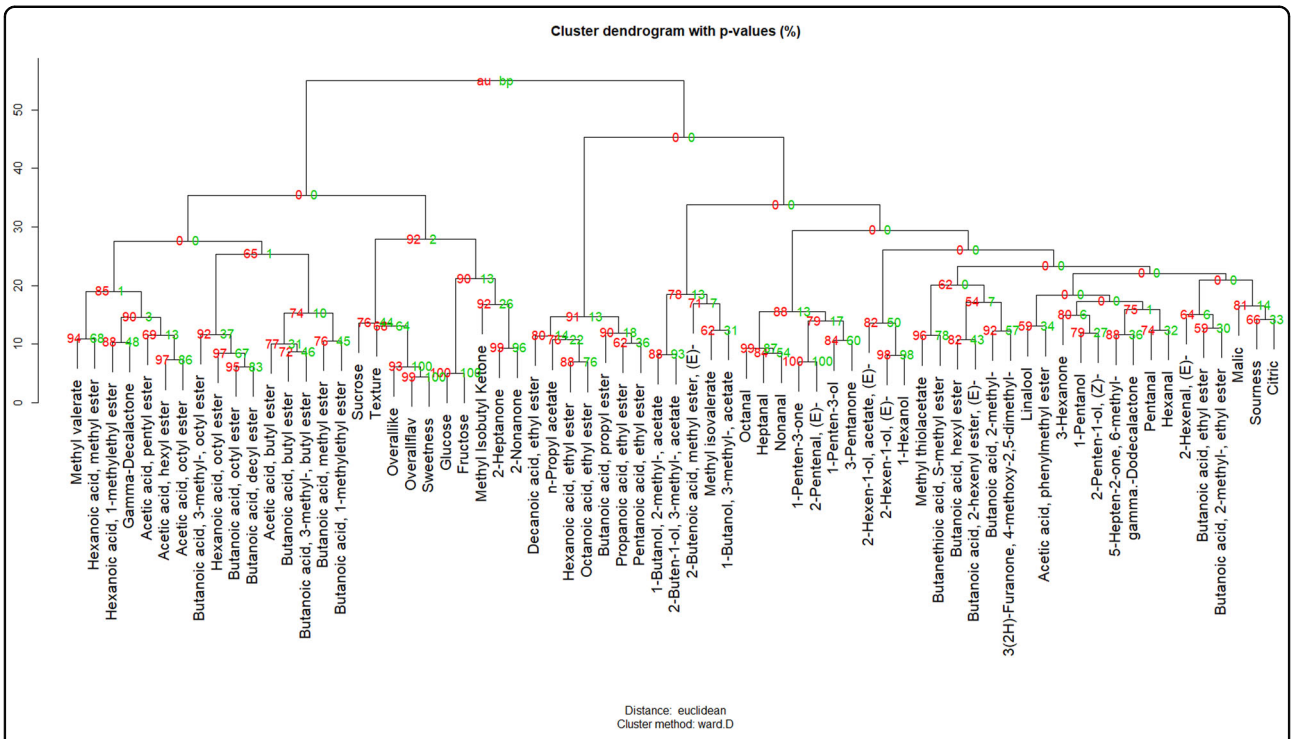

Fig. 1 Cluster dendrogram of sensory attributes, volatiles, sugars and acids. AU (approximately unbiased) *p*-values are in red and BP (bootstrap probability) *p*-values are in green

compound to consumer liking in the first two datasets. Fructose and glucose had lesser but important influence on liking in all three datasets. Important volatiles obtained for consumer liking included 10 esters, 3 aldehydes, 1 lactone, 3 ketones, and 1 alcohol. Eleven of these compounds were significant for the *t*-test (Table 1 and Supplementary Table S5), providing evidence that they had positive impact on consumer liking that was independent of total sugars. Butanoic acid, ethyl ester (V19), known to be one of the most abundant and important contributors to strawberry aroma, showed significant positive correlations ($r = 0.31$) with liking but did not pass the PLS threshold for two of the three datasets. Other abundant esters like butanoic acid, methyl ester (V9); hexanoic acid, methyl ester (V40) and hexanoic acid, ethyl ester (V44) all showed positive but low to moderate correlations ($r =$ 0.29, 0.17, and 0.13, respectively) with liking and did not reach PLS thresholds. Fifteen compounds were identified via PLS as important drivers of both sweetness and liking models, reflecting the high correlation between sweetness and liking (Table 1).

Predicting sweetness and liking with sugars, acids, and volatiles

Since sugars and acids are known to be important factors influencing sweetness perception and consumer liking, we first built a model with sugars and acids. This model produced an R^2 of 0.41 in testing CV (independent test dataset) for sweetness, indicating that 41% of variation in sweetness could be accounted for by sugars and acids. To further investigate the influences of volatiles on sweetness and liking, we included them as variables and tested prediction accuracies with a broad spectrum of algorithms (Fig. 5a), including decision-tree based Random Forest (RF), bayes GLM, GLM, GLMBOOST, and LASSO. With the inclusion of volatiles, LASSO and GLMBOOST models explained the most variation for sweetness in testing CV ($R^2 = 0.65$ and 0.69), outperforming the model with sugars and acids by 28% for the test dataset. In addition, LASSO and GLMBOOST generated the lowest standard deviation of R^2 among all models (Fig. 5a). Similar results were produced in prediction of liking but with a slightly lower total variance explained. LASSO model accounted for 55% in testing CV, compared to only 30% in models with sugars and acids alone (Fig. 5b). These results provide further strong evidence of volatile compound enhancement of strawberry sweetness perception and consumer liking independently of sugars.

Chemical networks in strawberry fruit

Broad variations in quantity of chemical compounds were measured for the 148 strawberry samples. For example, the most abundant ester (butanoic acid, methyl

Downloaded from https://academic.oup.com/hr/article/doi/10.1038/s41438-021-00502-5/6446614 by guest on 05 August 2026

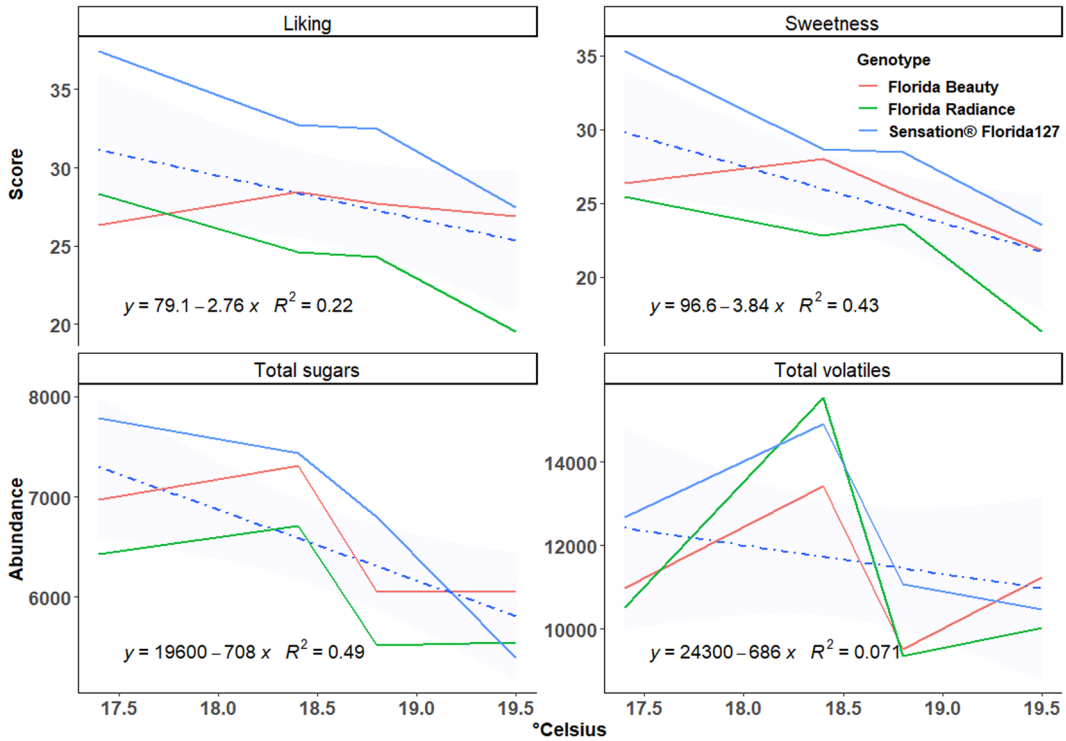

Fig. 2 Liking, sweetness, total sugars, and total volatiles of three cultivars as influenced by five-day average soil temperature prior to harvest. The trend line was modeled with linear regression. Fruits were harvested on 2/22/2016 (five-day average 17.4 °C), 3/7/2016 (18.8 °C), 2/7/2017 (18.4 °C), and 2/14/2017 (19.5 °C). Sizable differences among genotypes and among temperatures indicate strong genotype and temperature effects on desired sensory attributes and total sugars

ester; V9) ranged from 92.0 ng[1] $100gFW^{-1}h^{-1}$ in 'Radiance' in 2017 to 7359.6 ng[1] $100gFW^{-1}h^{-1}$ in a proprietary sample from 2011, an approximately 70-fold difference. Qualitative differences were also observed. For example, methyl anthranilate (V68) was only detectable in 3 samples and 2-nonanone (V56) was detected in 99 of 148 samples (Supplementary Table S3).

Collinearity among chemicals may reflect metabolic pathways, helping to narrow the number of genetic targets for flavor improvement. Fructose and glucose exhibited a very high correlation ($r = 0.91$). The highest positive correlations among volatiles were found between butanoic acid, octyl ester (V69) and butanoic acid, decyl ester (V80) ($r = 0.87$), followed by hexanoic acid, ethyl ester (V44) and octanoic acid, ethyl ester (V64) ($r = 0.83$); 1-penten-3-one (V3) and 2-pentenal, (E)- (V12) ($r = 0.83$); and acetic acid, hexyl ester (V46) and acetic acid, octyl ester (V65) ($r = 0.82$). Other correlations are detailed in Fig. 3 and Supplementary Table S4.

Group correlations and clusters among volatiles were better illustrated with hierarchical clustering and network analyses. Fatty acid esters derived from the same alcohol were grouped together in several clusters (Fig. 1). For example, acetic acid, butyl ester (V21); butanoic acid, butyl ester (V43); and butanoic acid, 3-methyl-, butyl ester (V51) were grouped with 81% confidence. Butanoic acid, 3-methyl-, octyl ester (V72); hexanoic acid, octyl ester (V79); butanoic acid, octyl ester (V69); and butanoic acid, decyl ester (V80) were clustered with 93% confidence. Other fatty acid ester clusters contained esters derived from the same fatty acid. For instance, acetic acid, pentyl ester (V38); acetic acid, hexyl ester (V46); and acetic acid, octyl ester (V65) were grouped with 62% confidence. A sulfur ester group was remote from the main ester group with a strong bond (96% confidence) between methyl thioacetate (V6) and butanethioic acid, S-methyl ester (V34). Besides esters, we observed an aldehyde group including octanal (V45), heptanal (V37) and nonanal (V59) with 99% confidence.

In the chemical networks, sixty-five nodes were connected by 299 significant edges (Fig. 4). A distinct butanoic acid ester cluster consisting of 8 esters including butanoic acid, hexyl ester (V62) and butanoic acid, octyl ester (V69). A second ester cluster centered on hexanoic acid, ethyl ester (V44) included multiple ethyl esters, such as octanoic acid, ethyl ester (V64), pentanoic acid, ethyl ester (V36), and propanoic acid, ethyl ester (V7). Aldehydes had their own cluster around heptanol (V37),

Downloaded from https://academic.oup.com/hr/article/doi/10.1038/s41438-021-00502-5/6446614 by guest on 05 August 2026

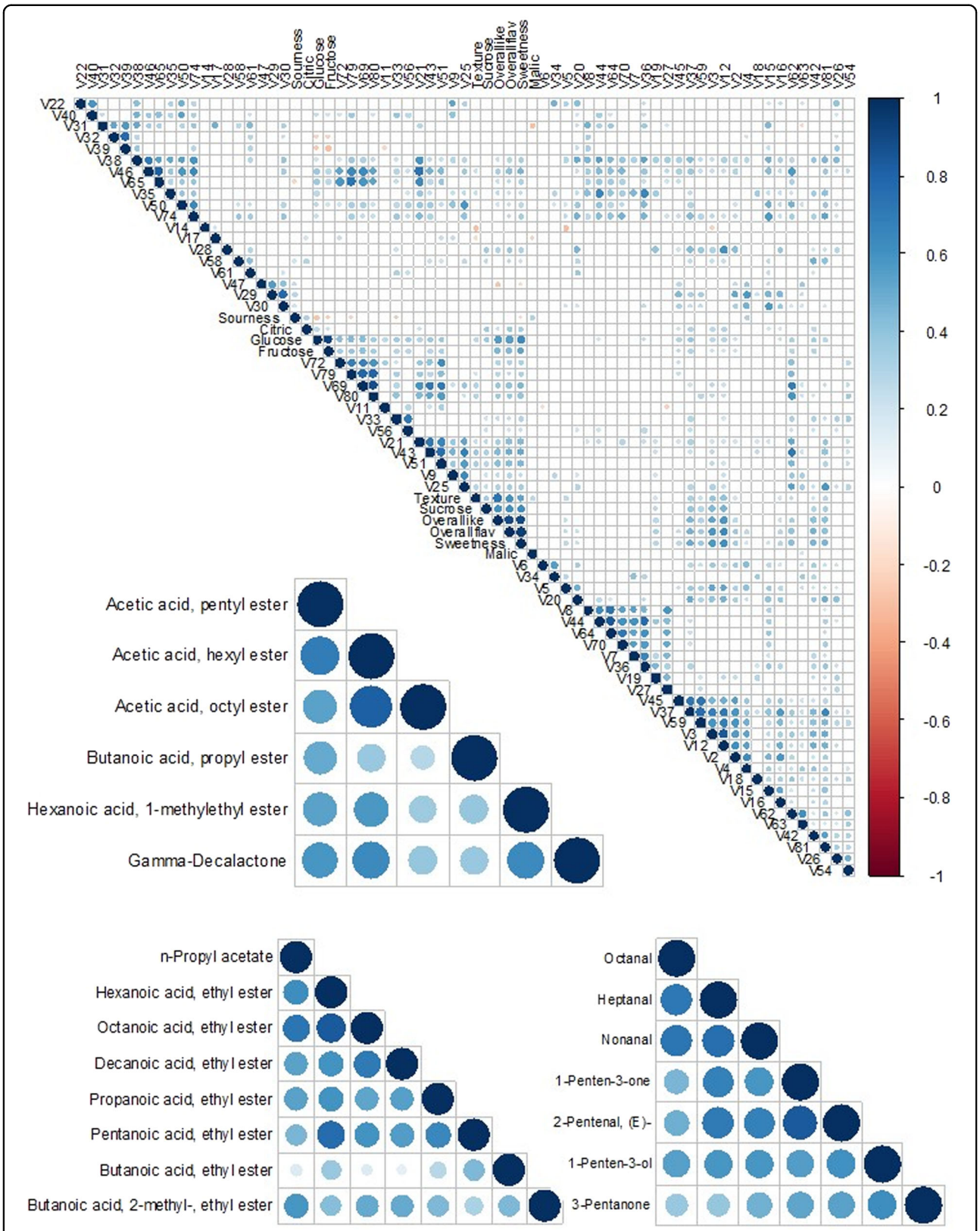

Fig. 3 Pairwise correlations of consumer attributes and chemical compounds. Dots represent significant correlations after false discovery rate (FDR) correction. The intensity of color and size of each dot is propotional to the absolute value of the correlation coeffcient (*r*). Three highly correlated regions of the plot are enlarged

Downloaded from https://academic.oup.com/hr/article/doi/10.1038/s41438-021-00502-5/6446614 by guest on 05 August 2026

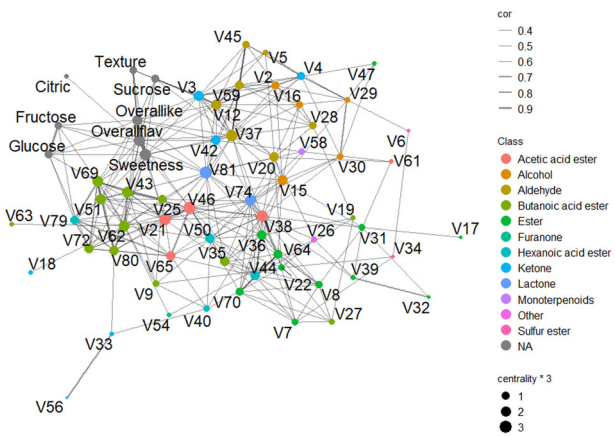

Fig. 4 A sensory-chemical network for strawberry. Significant correlations are connected by edges. Sizes of nodes are proportional to centrality scores. Centrality scores were calculated for each node to reveal importance of the volatile in the cluster. Volatiles are colored by chemical class

surrounded by 2-pentenal, (E)- (V12); nonanal (V59); hexanal (V20); pentanal (V5); octanal (V45) and 2-hexenal, (E)- (V28).

Genetic and chemical diversity among tested varieties

In order to explore the genetic diversity of the genotypes tested, we obtained SNP data via the Axiom® IStraw35 array[37] for 26 samples. The genotyped panel comprised 20 UF genotypes, 4 UC-Davis genotypes, 1 proprietary genotype and 1 genotype of European origin. The PCA individual plot showed a clear separation between UC and UF genotypes on PC2 (Fig. 6a). On the other hand, PC1 explained the most variation within UF genotypes. Combining PC1 and PC2 explained 22.1 percent of variation. In a complementary analysis we conducted PCA with sugars, acids and 59 common volatiles (Fig. 6b). No visible separation for genotypes from different origins was observed in the individual plot. Combining PC1 and PC2 explained 29 percent of variation.

Two major QTLs for production of esters

A chemical network analysis suggested the possibility of common genetic control for multiple esters (Fig. 4). To investigate this possibility, QTLs for ester production were mapped across eight experimental crosses. Medium-chain esters including acetic acid, decyl ester; acetic acid, hexyl ester; acetic acid, octyl ester; acetic acid, nonyl ester; butanoic acid, butyl ester; and butanoic acid, octyl ester shared a single QTL on linkage group 6 A (Supplementary Fig. S1). Single-marker analyses revealed that a shared maker (AX-123358920) explained 7.7% to 30.3% of the variation in ester abundances. Another QTL on the same linkage group but at a distinct location was discovered for both butanoic acid, ethyl ester and hexanoic acid, ethyl ester (Fig. 7). The single shared marker (AX-166508286) explained 15.0% of the observed variance in hexanoic acid, ethyl ester abundance ($p = 1.5e\text{-}5$) and 15.6% of butanoic acid, ethyl ester abundance ($p = 1.6e\text{-}5$).

Discussion

A large-scale sensory and chemical analysis of 148 fresh strawberry samples was conducted over seven years for the purposes of examining the sensory and chemical drivers of consumer preference. Our study largely expands the previous work[18] in terms of sample size, and provides new insights towards volatiles contributing to liking using updated statistical methods, chemical networks, and volatile genome-wide associations. In agreement with studies in tomato, blueberry and the previous strawberry work, perceived sweetness ($r = 0.714$*) was the major determinant for liking in strawberry[15,18,34]. A previous study using a glucose solution found that sweetness and hedonics followed a biphasic relationship in which the apex of liking was reached at glucose concentration of 500 mM[38]. In the present study the highest recorded total sugars concentration was 10056.2 $mg^{1}100gFW^{-1}$, roughly estimated to approximately 10 mM glucose solution, much lower than the equivalent optimal concentration. This was consistent with a low degree of satisfaction in sweetness in the sensory panel, where only 8.2% of strawberry samples reached the panelist's reported ideal sweetness (data not shown). This means that 91.8% of samples the sensory panels failed to achieve the panelist's ideal sweetness for a strawberry. This discrepancy between ideal sweetness and sample sweetness dramatically illustrates the need for breeding to increase sweetness in strawberries. One strategy to enhance perceived sweetness is to increase sugar concentration. The major sugars in strawberry fruit are glucose, fructose and sucrose[39]. All three sugars exhibited high correlations with sweetness (Table 1 and Fig. 3) and were the top drivers of liking. Nevertheless, breeding strawberries for higher sugar content may result in negative impacts on other critical traits. A trade-off between yield and soluble solids content (SSC) was found in UF strawberry breeding germplasm at the additive genetic level, demonstrating that genetic gains for SSC would likely result in yield losses[40]. In peach, higher SSC is associated with a longer fruit development period and reduced fruit weight[41]. In contrast, volatiles typically occur at 10^3 to 10^6 times lower concentration (uM to nM range) than sugars. Increasing sweetness-enhancing volatile concentrations even a small amount has the potential to increase the perception of sweetness and overall liking at a lower carbon cost to industry-demanded agronomic traits. Thus, non-sugar sources of sweetness would be highly advantageous for improving both flavor and agronomic traits simultaneously.

Downloaded from https://academic.oup.com/hr/article/doi/10.1038/s41438-021-00502-5/6446614 by guest on 05 August 2026

Table 1 Variable importance for predicting sweetness and liking

Chemical compound	VIPs in sweetness model				VIPs in liking model			
	2011–2012[a]	2013–2015	2016–2017	*p*-value[b]	2011–2012	2013–2015	2016–2017	*p*-value
Glucose	1.68	2.14	1.52	[c]	1.19	2.18	1.65	
Fructose	1.60	1.62	1.79		1.00	1.65	1.80	
Sucrose	2.17	2.10	0.27		2.19	2.32	0.20	
1-Penten-3-one	1.76	1.26	1.34	0.0000	1.87	1.31	1.23	0.0000
2-Pentenal, (E)-	1.52	1.53	1.76	0.0000	1.83	1.48	1.71	0.0000
2-Penten-1-ol, (Z)-	0.19	1.12	2.00	1.0000	0.33	1.02	1.94	1.0000
Heptanal	1.36	1.26	1.60	0.0000	1.43	1.13	1.38	0.0000
5-Hepten-2-one, 6-methyl-	1.16	1.42	0.49	0.0000	1.31	1.20	0.41	0.0000
Butanoic acid, butyl ester	1.38	1.48	0.88	0.0013	1.17	1.48	1.03	0.3781
Acetic acid, hexyl ester	1.39	1.24	0.39	0.0013	1.17	1.12	0.52	1.0000
Butanoic acid, 3-methyl-, butyl ester	1.23	1.33	0.77	0.0000	1.05	1.42	1.02	0.0033
Butanoic acid, pentyl ester	1.28	1.18	0.00	0.0000	1.28	1.12		0.0000
Nonanal	1.16	0.81	1.36	0.0000	1.33	0.77	1.30	0.0013
Butanoic acid, hexyl ester	1.50	0.90	1.71	0.0467	1.27	0.87	1.70	1.0000
Butanoic acid, octyl ester	1.01	1.30	1.30	0.0007	0.79	1.32	1.40	0.2430
Butanoic acid, decyl ester	1.06	1.03	1.23	0.0000	1.01	1.07	1.27	0.0005
γ-Dodecalactone	1.51	1.25	0.74	0.0000	1.54	1.21	0.84	0.0000
Hexanoic acid, butyl ester	[d]	1.34	1.44	0.0002		1.26	1.48	1.0000
γ-Decalactone	1.08	1.12	0.69	0.2637	—[e]	—	—	—
Acetic acid, butyl ester	1.16	1.34	0.44	0.0371	—	—	—	—
2-Hexenal, (E)-	1.17	0.52	1.61	0.0000	—	—	—	—
Benzaldehyde		1.01	1.99	1.0000	—	—	—	—
Butanoic acid, phenylmethyl ester		1.03	1.04	1.0000	—	—	—	—
Butanoic acid, propyl ester	—	—	—	—	1.18	1.09	0.86	1.0000
Methyl Isobutyl Ketone	—	—	—	—	0.79	1.07	1.47	1.0000
Butanoic acid, 3-methyl-, octyl ester	—	—	—	—	0.91	1.21	1.05	1.0000

Independent partial least square (PLS) models were fitted for 2011–2012, 2013–2015 and 2016–2017 datasets. Variable importance in projection (VIP) scores measure the importance of each variable in predicting the response variable. A *t*-test was conducted for each of the chemical compound to determine its significance to model sweetness or liking while controlling for total sugars. [a]Seasons. [b]*p*-Values were ajusted with bonferroni corrections. [c]*p*-Value was not obtained for sugars. [d]Compound not detected. [e]Compounds with no VIPs did not pass the threshold of 1 which is the typical cutoff for relevant variable selection

When fruits are chewed or swallowed, volatiles flow into the retronasal tract. Signals from retronasal olfaction interact with signals from the tongue in the brain, imparting the sensation of taste[15]. Evidence of integration of flavor and taste was adduced by a faster response time when drinking a flavored sugar solution[42]. While the effects of volatiles on flavor is widely known, the enhancement of sweetness by specific volatiles in sugar solutions or fruit puree is also a known phenomenon[43]. When accompanied by strawberry odor, perceived sweetness increases at a range of sugar concentrations in water solution[44]. In the present study, we demonstrate this phenomenon via consumer testing of actual strawberry fruit.

Strawberry volatile compounds are highly diverse and variable. A recent study consolidating information from 25 strawberry volatile studies from 1991 to 2016 found no consensus on the detected volatiles[5]. The most frequently identified volatile compounds have been distributed among the esters, acids, lactones, aldehydes, furans, alcohols, ketones, and terpenoids. In our chemical analyses, we captured 113 volatiles with 59 common to all years (Supplementary Table S3). In our previous study comprising 54 of these samples, six

Downloaded from https://academic.oup.com/hr/article/doi/10.1038/s41438-021-00502-5/6446614 by guest on 05 August 2026

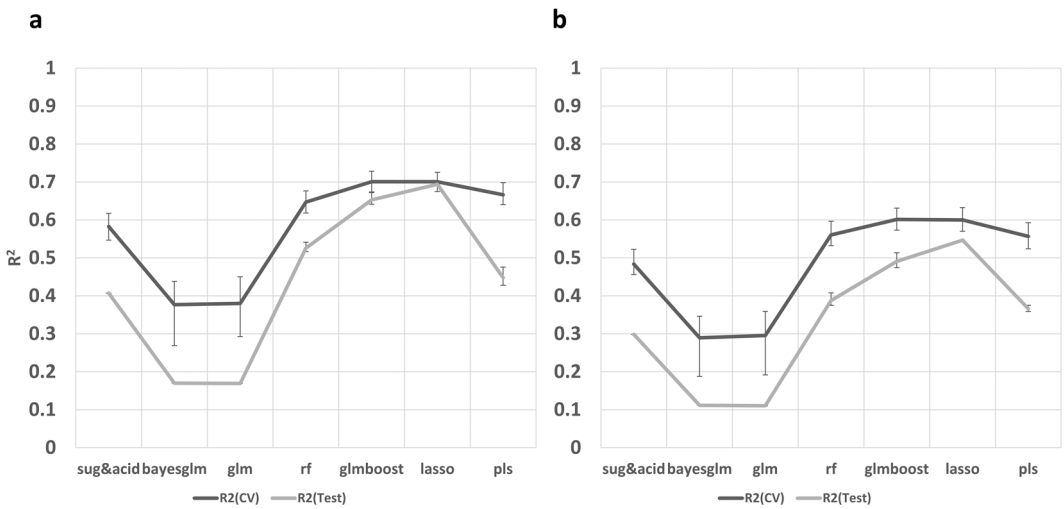

Fig. 5 Evaluation of models incorporating volatiles. Histograms of performance of six machine learning algorithms using sugars, acids, and volatiles to predict sweetness intensity (**a**) and consumer liking (**b**) compared to a model with sugars and acids only (leftmost in both figures). Black lines represent the average R^2 of nested cross validation while gray lines show the average R^2 when using a test dataset. Error bars represent standard deviations based on 100 iterations

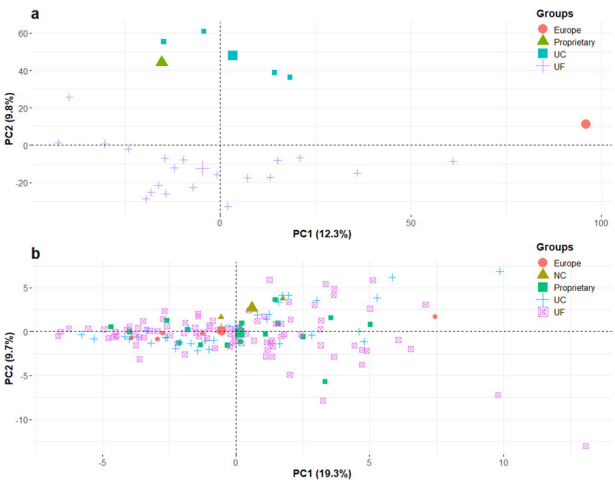

Fig. 6 Principal component analyses of genotypes. a PCA) plot of genotypes using Axiom® IStraw35 38000 single nucleotide polymorphisms (SNPs). Genotypes from different origins are grouped by color. Percentage of total variation explained by each component is shown in parentheses. **b** PCA plot of genotypes using volatiles, sugars and acids

volatiles (1-penten-3-one (V3); γ-dodecalactone (V81); butanoic acid, pentyl ester (V53); butanoic acid, hexyl ester (V62); acetic acid, hexyl ester (V46); and butanoic acid, 1-methylbutyl ester (V48)) were found significantly correlated with sweetness independent of sugars[18]. In the present study, five of these six volatiles except for butanoic acid, 1-methylbutyl ester were validated in additional datasets and fourteen additional sweetness-enhancing volatiles and eighteen liking-enhancing volatiles were reported. Among the nineteen sweetness enhancers, ten are esters with nine of the ten having a carbon number more than eight. To our surprise, the most abundant esters butanoic acid, methyl ester (V9) and butanoic acid, ethyl ester (V19) did not reach the VIP threshold for some datasets, even though significant correlations were found between butanoic acid, ethyl ester and liking and sweetness.

Strawberry varieties with high concentrations of lactones, which impart peach-like odor, have been implicated as preferable in previous consumer tests[45]. We show evidence that γ-decalactone (V74) and γ-dodecalactone (V81) are sweetness-enhancing. Nevertheless, only γ-dodecalactone was important for liking. Distributions of the two lactones were also disparate, with γ-decalactone following a bimodal distribution in our study, consistent with its presence being controlled by a single gene[19]. On the other hand, γ-dodecalactone followed a normal distribution suggestive of polygenic control. Three correlated volatile compounds, 1-penten-3-one (V3); 2-pentenal, (E)- (V12); and 2-penten-1-ol, (Z)- (V16) were highly associated with sweetness and liking. Although none of them elicit sweet aroma according to the literature[46], their interactions with sugars or other volatiles at the neurological level may lead to increased sweetness perception and hedonic experience. Interestingly, 1-penten-3-one and 2-pentenal, (E)- positively impacts liking in tomato[16], indicating that common contributors to liking may exist across a wide diversity of fruits. Two volatiles impacting liking in this study, nonanal (V59) and 5-hepten-2-one, 6-methyl- (V42) impart flowery notes to the strawberry, which are important building blocks in the diversity of

Downloaded from https://academic.oup.com/hr/article/doi/10.1038/s41438-021-00502-5/6446614 by guest on 05 August 2026

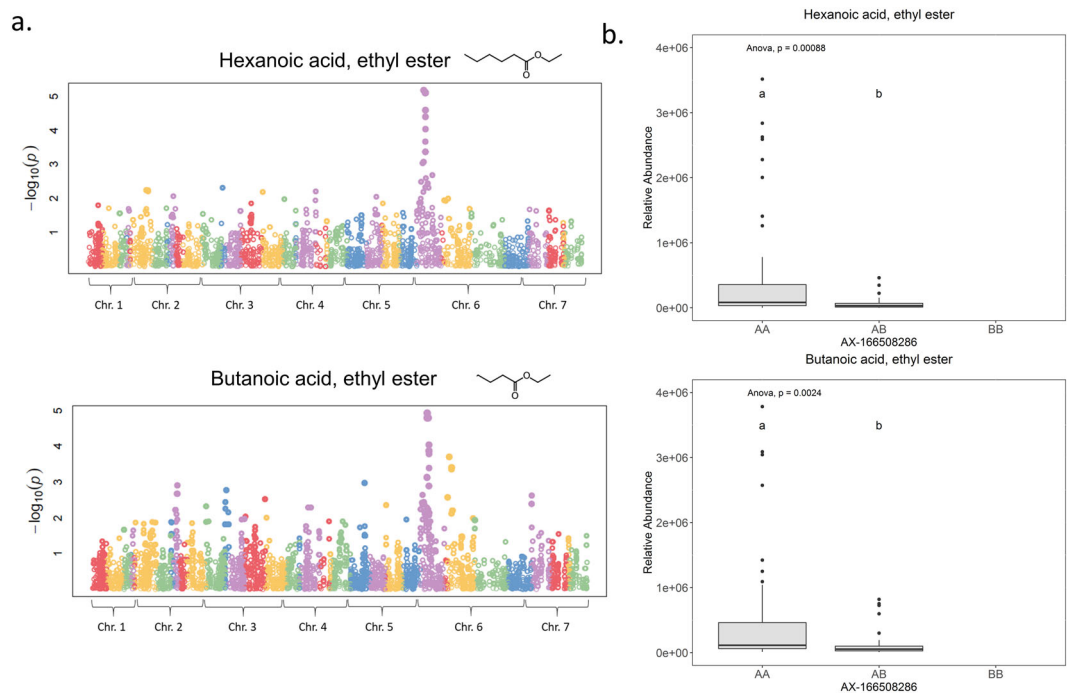

Fig. 7 Genome-wide association of esters in strawberry fruit. a Manhattan plots for hexanoic acid, ethyl ester and butanoic acid ethyl ester. A shared QTL was observed on linkage group 6 A. **b** A single shared marker explained a large portion of variation in the corresponding ester abundance. Different letters indicate significant differences at $p = 0.05$

strawberry aroma. Carotenoids including 5-hepten-2-one, 6-methyl- were also positively correlated with consumer acceptability in tomato[47]. Due to the isolation and the detection techniques employed in this study, furaneol was not quantified. Future experiments investigating the influence of furaneol on strawberry will need to harness other isolation approaches. In tomato, Buttery et al.[48] demonstrated the isolation of and detection furaneol with anhydrous sodium sulfate coupled with dynamic headspace sampling.

Fruit chemical data have great potential for predicting sensory responses and carry the advantage of accuracy and objectivity as targets for genetic improvement. Univariate models with soluble solid content or titratable acidity have often provided poor prediction of liking or sensory attributes[49]. Including additional sensory contributors like volatiles as variables can increase model prediction accuracy. However, a mathematical challenge of adding volatile data to prediction models is the depletion of degrees of freedom with high dimensional data. A few studies have addressed this problem by adopting different statistical models like multivariate models, PLS models, principle component regression models, and artificial neural networks[35,49,50]. In these studies, machine learning models always resulted in higher explained variance of sensory attributes than simple models with sugars only. However, these studies have limited applicability, suffering from small sample size or lack of cross-validation[50,51]. In this study, 90% samples were used to train the model and 10% was exclusively used for testing. Comparisons of multiple prediction methods revealed a 28% increase in variation explained for sweetness using a LASSO model compared to linear model with sugars and acids, as well as 25% in liking. GLMBOOST and LASSO models exhibited better performance than the traditional PLS model in all quality measurements. On the other hand, the large increasement in the R-square value after incorporating volatile data corroborates the importance of volatiles to enhance sweetness and consumer liking.

Strawberry quality is strongly influenced by both genetic and environmental factors[5,25,52]. In our analyses, genetic variation affected presence and absence of volatiles as well as their abundance (Supplementary Table S3). In addition, temperature fluctuated dramatically during the evaluated seasons, from periods of nearly freezing weather in December and January to an average over 26 °C in March and April, far above the optimal temperature for fruit quality[52]. Thirteen compounds including citric acid and total sugars were significantly affected by harvest date within season (data not shown). Similarly, multiple other studies found declines of total sugars or SSC during the course of strawberry growing seasons[25,53]. Although a seasonal decline was observed for total volatile abundance

Downloaded from https://academic.oup.com/hr/article/doi/10.1038/s41438-021-00502-5/6446614 by guest on 05 August 2026

(Fig. 2), it was not statistically significant, and not all volatiles were negatively affected by temperature. In blueberry, fructose and glucose abundance exhibited significant G×E or environmental effects, in contrast to many of the volatiles[34]. Perhaps volatiles may hold part of the key for stabilizing sensory qualities in the face of a changing environment and may attenuate the impact of declines in total sugars with increasing temperature. Network and clustering analyses of metabolites may provide evidence for common biosynthetic pathways or regulation mechanism to help narrow targets for genetic improvement. Here we used pair-wise correlations and hierarchical clustering approach to explore connections among volatiles. Fatty acid esters were clustered based on either alcohol substrate or acyl-CoA substrate. Butanoic acid esters were clustered by both approaches with high confidence (Figs. 1 and 4). Ethyl esters, such as octanoic acid, ethyl ester (V64); pentanoic acid, ethyl ester (V36); propanoic acid, ethyl ester(V7); and hexanoic acid, ethyl ester (V44) were highly correlated and clustered in the volatile network. Common genetic regulation for esters was further corroborated with the discovery of major QTLs shared among multiple medium-chain esters and the two most abundant ethyl esters. In a European strawberry mapping population, possibly the same QTL on linkage group 6-1 influenced the productions of multiple medium-chain esters[54], which supports broad applications for maker assisted selection at this locus to improve sensory sweetness and flavor. Based on evidence of significant makers and ripening-induced genes[55], we identified two candidate genes for the medium-chain esters QTL and the ethyl esters QTL. The candidate for the medium-chain esters QTL is a putative cinnamoyl-CoA reductase1 (maker-Fvb6-1-augustus-gene-164.26), upregulated more than five-fold in red fruit stage campared to the green fruit stage (Supplementary Fig. S2). This enzyme catalyzes the reduction of a substituted cinnamoyl-Coa, which may provide aldyhyde substrates for downstream esterification. A putative pyruvate decarboxylase-2 (maker-Fvb6-1-augustus-gene-297.79) was found within the ethyl esters QTL region. The expression of this gene increased 2.3 times from the green fruit stage to the red fruit stage (Supplementary Fig. S2). Pyruvate decarboxylase-2 is responsible for conversion of pyruvate to acetyl cozenzyme A (CoA), a essential substrate for ester biosynthesis[56]. Another important volatile class, aldehydes, synthesized via decarboxylases or hydroperoxide lyase[57,58], has not been well-characterized in strawberry. Our network analysis also predicts a common enzyme catalyzing a range of aldehydes. In wild relative *Fragaria vesca*, a volatile metabolomics study using near-isogenic lines also showed ester and aldehyde clusters[59]. Similar coordinated metabolic networks were demonstrated in peach, tomato, apple, and melon[15,60–62]. Together these analyses reveal concerted regulation within and between volatile classes in both climacteric and non-climacteric fruits.

While consumers desire many traits in fresh strawberries, including appearance and health attributes, flavor ranks at the top[2]. Flavor is a difficult trait to improve due its chemical complexity, strong environmental effects and negative relationships with grower-demanded agronomic traits. It is also not trivial to measure. Quantifying sensory attributes for every genotype in the breeding pool is not practical, nor is comprehensive volatile quantification. In this study we have identified a reduced set of chemical targets for consumer liking and sweetness improvement in strawberry. High-throughput and cost-effective DNA markers can be easily developed based on the significant makers identified from our association analysis and implemented in diverse strawberry breeding programs. We must also consider that this study intentionally samples the chemical diversity of commercial germplasm that is readily deployed in agriculture. Yet distinct flavors in the wild relatives of cultivated strawberry could be introgressed into the cultivated germplasm pool, expanding the flavor toolbox of modern strawberry in the long-term. Narrowing the non-sugar targets for sweetness-enhancement in strawberry is an important step toward a worthy goal: increased consumer satisfaction for fresh strawberries.

Materials and methods

Ethics statement

Consumer panels were conducted at the Food Science and Human Nutrition Department at the University of Florida (UF) in Gainesville, FL. The UF Institutional Review Board 2 (IRB2) chaired by Ira S. Fischler approved the protocol and written consent form (case #2003-U-0491) that sensory panel participants were required to complete.

Plant material and fruit production

From 2010-2017, 148 strawberry (*Fragaria ×ananassa*) samples from 48 cultivars and breeding lines (Supplementary Table S6) were harvested from strawberry research plots of the University of Florida (UF) Gulf Coast Research and Education Center (GCREC) (Balm, FL) or the Florida Strawberry Growers Association (FSGA) headquarters in Dover, FL. All fruiting field trials were managed according to commercial standards for Florida annual plasticulture[63]. Data from 54 samples harvested during 2010-2012 were analyzed and published in a previous paper[18]. These samples represented major commercial varieties grown in California, Florida and Europe. Ninety-four additional samples were collected during 2013-2017, representing a wider range of UF germplasm including varieties and advanced breeding lines. Thirty-three genotypes were harvested more than one time

Downloaded from https://academic.oup.com/hr/article/doi/10.1038/s41438-021-00502-5/6446614 by guest on 05 August 2026

across the seven years. The combined 7-year dataset captured variation produced by genetics, harvest dates and seasons (Supplementary Table S6).

At each harvest date, fully ripe fruit with at least 90% red surface coloration from three to five genotypes were harvested from the field in the early morning, transported to the Food Science sensory lab in Gainesville, FL, and stored at 4 °C in the dark overnight. Each sample consisting of at least 200 fruits was collected from a set of 500 plants. Seven or more fruits, each collected from a separate plant were randomly chosen for chemical quantification. The rest were used for sensory evaluation during the next morning. Temperature data were obtained from the Balm, FL station of the Florida Automated Weather Network located directly adjacent to the plots from which strawberries were sampled. Soil temperature was recorded at a depth of -10 cm. Temperature data were used to regress liking, sweetness, total sugar and total volatiles from three cultivars during the 2016 and 2017 seasons.

Sensory evaluation

In each session, an average of 100 consumers were divided into individual booths and utilized specialized sensory software to rate the samples (CompuSense® 5 Sensory Analyses Software, CompuSense, Guelph, Canada). One or two whole strawberries from each of three to five randomly coded and ordered genotypes were served per individual/session. All strawberries were served at room temperature.

Consumers were asked to rate hedonic attributes (consumer liking and texture liking) on the Global Hedonic Intensity Scale (GHIS), which was anchored with the most intense liking ever experienced at the top (+100) and the most intense disliking ever experienced at the bottom (-100). Intensity of individual sensory attributes (sweetness, sourness and strawberry flavor) were scored on the Global Sensory Intensity Scale (GSIS) which is anchored with the most intense sensation at the top (+100) and no sensation at the bottom (0)[64].

Volatile detection and quantification

At least seven fruits were removed from 4 °C storage and pooled in a blender prior to splitting into three 15 g replicates for immediate capture of volatile emissions. The remainder was frozen in liquid nitrogen and stored at −80 °C for subsequent quantification of sugars and acids. A two-hour collection in a dynamic headspace volatile collection system[65] allowed for concentration of emitted volatiles on HaySep 80–100 porous polymer adsorbent (Hayes Separations Inc., Bandera, TX, USA). Elution from the polymer was described by Schmelz[66]. Quantification of volatiles in an elution was performed on an Agilent 7890 A Series gas chromatograph (GC) (carrier gas; He at 3.99 ml min−1; splitless injector, temperature 220 °C, injection volume 2 µl) equipped with a DB-5 column ((5%-phenyl)-methylpolysiloxane, 30 m length ×250 µm i.d. × 1 µm film thickness; Agilent Technologies, Santa Clara, CA, USA). Oven temperature was programmed from 40 °C (0.5 min hold) at 5 °C min^{-1} to 250 °C (4 min hold). Signals were captured with a flame ionization detector (FID) at 280 °C. Peaks from FID signal were integrated manually with Chemstation B.04.01 software (Agilent Technologies, Santa Clara, CA). Volatile emissions (ng^1 $100gFW^{-1}$ h^{-1}) were calculated based on individual peak area relative to sample elution standard peak area. GC-mass spectrometry (MS) analyses of elutions were performed on an Agilent 6890 N GC in tandem with an Agilent 5975 MS (Agilent Technologies, Santa Clara, CA, USA), and retention times were compared with authentic standards (Sigma Aldrich, St Louis, MO, USA) for volatile identification[67].

Sugars and acids quantification

Titratable acidity, pH, and soluble solids content were averaged from four replicates of the supernatant of centrifuged and thawed homogenates[68]. An appropriate dilution of the supernatant from a separate homogenate (centrifugation of 1.5 ml at 16,000 *g* for 20 min) was analyzed using biochemical kits (per manufacturer's instructions) for quantification of citric acid, L-malic acid, D-glucose, D-fructose, and sucrose (CAT# 10-139-076-035, CAT# 10-139-068-035, and CAT# 10-716-260-035; R-Biopharm, Darmstadt, Germany) with absorbance measured at 365 nm on an Epoch Microplate Spectrophotometer (BioTek, Winooksi, VT, USA). Metabolite average concentration (mg^1 $100gFW^{-1}$) was determined from two to six technical replicates per pooled sample. Derived sucrose concentrations via D-glucose and D-fructose were mathematically pooled.

Statistical analyses and model parameters

For the purpose of comparing averaged sensory features between samples assessed by different panelists, a mixed linear model was constructed with panelist as a random effect. Least square means (Supplementary Table S2) of the sensory features for each sample were obtained using SAS software (Version 9.2.; SAS Institute, Cary, NC). Spearman's rank-order correlation was used for correlations among consumer attributes. Chemical data were averaged across replicates.

The whole dataset was divided into three subsets corresponding to three periods: 2011-2012, 2013-2015 and 2016-2017 since volatile quantification was performed by a different technician in each period and the sugars and acids quantifications were performed three times from frozen samples according to these periods. We observed significant differences in chemical abundance among periods. In order to compensate for technical differences,

Downloaded from https://academic.oup.com/hr/article/doi/10.1038/s41438-021-00502-5/6446614 by guest on 05 August 2026

the normalization method of autoscaling[69] was conducted for each chemical compound within each period such that periods centered to the same mean and standard deviation. Correlations, clustering and predictive models were built on merged data combining three auto-scaled datasets. Pearson correlations were conducted for all pairs of consumer attributes and chemical compounds.

After false discovery rate (FDR) correction, the correlation matrix was plotted with the "corrplot" package in R for visualization of the relationships. Correlations were filtered with $p < 0.01$ and order of variables was ranked via hierarchical clustering. Non-supervised hierarchical clustering analysis was conducted with the "pvclust" package in R using method.hclust = "ward". AU (Approximately Unbiased) *p*-values and BP (Bootstrap Probability) p-values were calculated via multiscale bootstrap resampling to measure the confidence of the clusters[70]. Sensory-chemical network was constructed with the "ggraph" package in R. Only significant correlations were retained after Bonferroni correction to reduce the number of edges for better visualization. Significant relationships between nodes were connected by edges. Kleinberg's authority centrality scores were calculated for each node to reveal importance of the volatiles in the network.

Partial least square models were built with the "PLS" package in R. Sweetness and consumer liking were regressed on sugars, acids, volatiles and texture liking. Three PLS models corresponded to the three datasets: 2011-2012, 2013-2015 and 2016-2017. The variable importance for the projection (VIP)[71] was calculated with the R package "PLSVarSel" to reveal important compounds for predictions. We selected 3 principal components for all models due to the smallest square root of mean square error (RMSE) for this number of components (data not shown). A chemical compound was deemed to be important to predict a sensory response if it had a minimum VIP score of 1.0[72] in two or more of the three datasets. For each of the compounds, a *t*-test on the slope of the volatile β was applied in a multiple linear regression containing the volatile and total sugars.

Machine learning algorithms were compared for prediction accuracy of liking and sweetness with the "caret" package in R. The whole dataset was split randomly into training (90%) and test (10%) sets for cross-validation (CV). Training dataset were then trained with Random Forest (RF), bayesGLM, GLM, GLMBOOST, LASSO, and PLS models. For comparison, a simple GLM model of individual sugars and acids was constructed using only sucrose, glucose, fructose, malic acid and citric acid. SVMlinear and xgbDART models were also tested but are not reported due to low performance and long computational time, respectively. Nested cross validation of five-fold was performed within 90% training data to tune the hyperparameters. A hundred iterations of nested cross validation were performed to evaluate the stability of model performance.

Genetic diversity analysis

Twenty-six cultivars and selections included in the sensory study were also genotyped (Supplementary Table S7). These included UF germplasm, UC-Davis cultivars and two cultivars from other origins. Genotyping was performed using the 38 K Axiom® IStraw35 384HT array[37]. Principal component analysis (PCA) was conducted to examine genetic relationships among samples from different origins. SNPs with missing data were omitted in the PCA analyses. A similar PCA was also conducted using chemical data. Individual plots were plotted with the "factoextra" package in R.

Genetic association analysis

Eight experimental populations, genotyping methods and volatile collection and quantification methods for this analysis are described previously[73]. In brief, eight outcrosses were made among octoploid cultivars and elite materials between year 2012 and 2016. Two to four clonal replicates of each seedling were established in single plots in the research field. Fruits were collected three times for each seedling during its growing season. Fruit samples were instantly frozen in liquid nitrogen upon harvest and stored at −80 °C prior to quantification with GC-MS. Relative abundances of individual esters were normalized via the Box-Cox transformation algorithm performed in R-studio prior to genetic correlation. Association analysis was performed using the mixed linear model method implemented in GAPIT v3[74] in R, using marker positions oriented to the linkage map derived from the 'Holiday' x 'Korona' mapping population[75]. Genetic associations were evaluated for significance based on the presence of multiple co-locating markers of *p*-value < 0.05 after FDR multiple comparisons correction.

Acknowledgements

The authors acknowledge the significant efforts of the UF strawberry breeding program technical staff in growing and maintaining breeding trials, harvests and fruit transportation, as well as the staff members of the Sims and Colquhoun labs in processing and analyzing such large amounts of samples.

Author details

[1]Horticultural Sciences Department, University of Florida, IFAS Gulf Coast Research and Education Center, Wimauma, FL, USA. [2]Elo Life Systems, Durham, NC, USA. [3]Department of Environmental Horticulture and Plant Innovation Center, Institute of Food and Agricultural Sciences, University of Florida, Gainesville, FL, USA. [4]Plenty Unlimited, Inc., South San Francisco, CA 94080, USA. [5]Food Science and Human Nutrition Department, University of Florida, Gainesville, FL, USA. [6]Horticultural Sciences Department, University of Florida, Gainesville, FL, USA

Author contributions

Z.F. and V.W. designed the study. Z.F. and T.H. performed data analyses. T.J., D.G., M.S. and T.C. performed chemical analyses. C.B. performed genetic association analysis. C.S. organized and performed consumer sensory panels. M.R. provided direction on data analyses. T.C., C.S. and V.W. oversaw the whole

Downloaded from https://academic.oup.com/hr/article/doi/10.1038/s41438-021-00502-5/6446614 by guest on 05 August 2026

project. Z.F. and V.W. composed the manuscript. All authors read and approved the final manuscript.

Conflict of interest

T.H. performed his part of this study while employed at UF/IFAS GCREC and was later employed by Elo Life Systems, Durham, NC, United States by the time of submission. T.J. performed his part of this study while employed at UF/IFAS Environmental Horticulture and was later employed by Driscoll's Strawberry Associates by the time of submission. The remaining authors declare that they have no conflict of interest.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41438-021-00502-5.

Received: 24 August 2020 Revised: 16 December 2020 Accepted: 6 February 2021
Published online: 01 April 2021

References

1. Hannum, S. M. Potential impact of strawberries on human health: a review of the science. *Crit. Rev. Food Sci. Nutr.* **44**, 1–17 (2010).
2. Colquhoun, T. A. et al. Framing the perfect strawberry: an exercise in consumer-assisted selection of fruit crops. *J. Berry Res.* **2**, 45–61 (2012).
3. Fait, A. et al. Reconfiguration of the achene and receptacle metabolic networks during strawberry fruit development. *Plant Physiol.* **148**, 730–750 (2008).
4. Vallarino, J. G. et al. Genetic diversity of strawberry germplasm using metabolomic biomarkers. *Sci. Rep.* **8**, 14386–13 (2018).
5. Ulrich, D., Kecke, S. & Olbricht, K. What do we know about the chemistry of strawberry aroma? *J. Agric. Food Chem.* **66**, 3291–3301 (2018).
6. Forney, C. F., Kalt, W. & Jordan, M. A. The composition of strawberry aroma is influenced by cultivar, maturity, and storage. *Hort. Sci.* **35**, 1022–1026 (2000).
7. Jetti, R. R., Yang, E., Kurnianta, A., Finn, C. & Qian, M. C. 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**, 487 (2007).
8. Du, X., Plotto, A., Baldwin, E. & Rouseff, R. 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 (2011).
9. Pyysalo, T., Honkanen, E. & Hirvi, T. 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 (1979).
10. Noble, A. C. Taste-aroma interactions. *Trends Food Sci. Technol.* **7**, 439–444 (1996).
11. Schieberle, P. & Hofmann, T. 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 (1997).
12. Zabetakis, I. & Holden, M. A. Strawberry flavour: analysis and biosynthesis. *J. Sci. Food Agric.* **74**, 421–434 (1997).
13. Nuzzi, M., Scalzo, R. L., Testoni, A. & Rizzolo, A. 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 (2008).
14. Baldwin, E. A. et al. Relationship between sensory and instrumental analysis for tomato flavor. *J. Am. Soc. Hort. Sci.* **123**, 906–915 (1998).
15. Tieman, D. et al. The chemical interactions underlying tomato flavor preferences. *Curr. Biol.* **22**, 1035–1039 (2012).
16. Klee, H. J. & Tieman, D. M. The genetics of fruit flavour preferences. *Nat. Rev. Genet.* **19**, 347–356 (2018).
17. Ferrão, L. F. V. et al. Genome-wide association of volatiles reveals candidate loci for blueberry flavor. *N. Phytologist* **226**, 1725–1737 (2020).
18. Schwieterman, M. L. et al. Strawberry flavor: diverse chemical compositions, a seasonal influence, and effects on sensory perception. *PloS ONE* **9**, e88446 (2014).
19. Chambers, A. H. et al. Identification of a strawberry flavor gene candidate using an integrated genetic-genomic-analytical chemistry approach. *BMC Genomics* **15**, 217 (2014).
20. Du, X., Song, M. & Rouseff, R. Identification of new strawberry sulfur volatiles and changes during maturation. *J. Agric. Food Chem.* **59**, 1293–1300 (2011).
21. Pelayo-Zaldívar, C., Ebeler, S. E. & Kader, A. A. Cultivar and harvest date effects on flavor and other quality attributes of California strawberries. *J. Food Qual.* **28**, 78–97 (2005).
22. Jouquand, C., Chandler, C., Plotto, A. & Goodner, K. A sensory and chemical analysis of fresh strawberries over harvest dates and seasons reveals factors that affect eating quality. *J. Am. Soc. Hort. Sci.* **133**, 859–867 (2008).
23. Olbricht, K. & Ulrich, D. *Domestication effects in European breeding history of strawberry demonstrated by aroma compound pattern* (VIII International Strawberry Symposium 1156, 2016).
24. Ulrich, D., Komes, D., Olbricht, K. & Hoberg, E. Diversity of aroma patterns in wild and cultivated Fragaria accessions. *Genet. Resour. Crop Evol.* **54**, 1185 (2007).
25. Perez, Y. & Sargent, S. Cecilia do, N. N. & Whitaker, V. Composition of commercial strawberry cultivars and advanced selections as affected by season, harvest, and postharvest storage. *Hort. Sci.* **51**, 1134–1143 (2016).
26. Sanz, C., Pérez, A. G. & Olías, R. *Effects of temperature on flavor components in strawberries* (IV International Strawberry Symposium 567, 2000).
27. Kasperbauer, M. J., Loughrin, J. H. & Wang, S. Y. Light reflected from red mulch to ripening strawberries affects aroma, sugar and organic acid concentrations. *Photochem. Photobiol.* **74**, 103–107 (2001).
28. Kårlund, A., Hanhineva, K., Lehtonen, M., Karjalainen, R. O. & Sandell, M. Non-targeted metabolite profiles and sensory properties of strawberry cultivars grown both organically and conventionally. *J. Agric. Food Chem.* **63**, 1010–1019 (2015).
29. Josuttis, M., Verrall, S., Stewart, D., Krüger, E. & McDougall, G. J. Genetic and environmental effects on tannin composition in strawberry (Fragaria × ananassa) cultivars grown in different European locations. *J. Agric. Food Chem.* **61**, 790–800 (2013).
30. Piggott, J. R., Paterson, A. & Clyne, J. Prediction of flavour intensity of blackcurrant (Ribes nigrum L.) drinks from compositional data on fruit concentrates by partial least squares regression. *Int. J. Food Sci. Tech.* **28**, 629–637 (1993).
31. Aznar, M., López, R., Cacho, J. & Ferreira, V. Prediction of aged red wine aroma properties from aroma chemical composition. Partial least squares regression models. *J. Agric. Food Chem.* **51**, 2700–2707 (2003).
32. François, I. M. et al. Predicting sensory attributes of different chicory hybrids using physico-chemical measurements and visible/near infrared spectroscopy. *Postharvest Biol. Technol.* **49**, 366–373 (2008).
33. Kreutzmann, S., Svensson, V. T., Thybo, A. K., Bro, R. & Petersen, M. A. Prediction of sensory quality in raw carrots (Daucus carota L.) using multi-block LS-ParPLS. *Food Qual. Preference* **19**, 609–617 (2008).
34. Gilbert, J. L. et al. Identifying breeding priorities for blueberry flavor using biochemical, sensory, and genotype by environment analyses. *PLoS ONE* **10**, e0138494 (2015).
35. Boccorh, R. K. & Paterson, A. An artificial neural network model for predicting flavour intensity in blackcurrant concentrates. *Food Qual. Preference* **13**, 117–128 (2002).
36. Predieri, S., Ragazzini, P. & Rondelli, R. *Sensory evaluation and peach fruit quality* (VI International Peach Symposium 713, 2005).
37. Verma, S. et al. *Development and evaluation of the Axiom® IStraw35 384HT array for the allo-octoploid cultivated strawberry Fragaria× ananassa* (VIII International Strawberry Symposium 1156, 2016).
38. Jayasinghe, S. N. et al. Is sweet taste perception associated with sweet food liking and intake?. *Nutrients* **9**, 750 (2017).
39. Kallio, H., Hakala, M., Pelkkikangas, A. & Lapveteläinen, A. Sugars and acids of strawberry varieties. *Eur. Food Res. Technol.* **212**, 81–85 (2000).
40. Whitaker, V. M., Osorio, L. F., Hasing, T. & Gezan, S. Estimation of genetic parameters for 12 fruit and vegetative traits in the University of Florida strawberry breeding population. *J. Am. Soc. Hort. Sci.* **137**, 316–324 (2012).
41. Cirilli, M., Bassi, D. & Ciacciulli, A. Sugars in peach fruit: a breeding perspective. *Hort. Res.* **3**, 1–12 (2016).
42. Veldhuizen, M. G., Shepard, T. G., Wang, M. & Marks, L. E. Coactivation of gustatory and olfactory signals in flavor perception. *Chem. Senses* **35**, 121–133 (2010).
43. Baldwin, E. A., Goodner, K., Plotto, A., Pritchett, K. & Einstein, M. Effect of volatiles and their concentration on perception of tomato descriptors. *J. Food Sci.* **69**, S310–S318 (2004).
44. Frank, R. A. & Byram, J. Taste–smell interactions are tastant and odorant dependent. *Chem. Senses* **13**, 445–455 (1988).
45. Ulrich, D. & Olbricht, K. A search for the ideal flavor of strawberry-comparison of consumer acceptance and metabolite patterns in Fragariax ananassa Duch. *J. Appl. Botany Food Qual.* **89**, 223–234 (2016).

Downloaded from https://academic.oup.com/hr/article/doi/10.1038/s41438-021-00502-5/6446614 by guest on 05 August 2026

46. Krumbein, A. & Auerswald, H. Characterization of aroma volatiles in tomatoes by sensory analyses. *Food* **42**, 395–399 (1998).
47. Vogel, J. T. et al. Carotenoid content impacts flavor acceptability in tomato (Solanum lycopersicum). *J. Sci. Food Agric.* **90**, 2233–2240 (2010).
48. Buttery, R. G., Takeoka, G. R., Naim, M., Rabinowitch, H. & Nam, Y. Analysis of furaneol in tomato using dynamic headspace sampling with sodium sulfate. *J. Agric. Food Chem.* **49**, 4349–4351 (2001).
49. Corrêa, S. C. et al. Prediction of the sensory acceptance of fruits by physical and physical–chemical parameters using multivariate models. *LWT - Food Sci. Technol.* **59**, 666–672 (2014).
50. López, M. L. et al. Volatile compounds, quality parameters and consumer acceptance of 'Pink Lady®' apples stored in different conditions. *Postharvest Biol. Technol.* **43**, 55–66 (2007).
51. Yu, P., Low, M. Y. & Zhou, W. Development of a partial least squares-artificial neural network (PLS-ANN) hybrid model for the prediction of consumer liking scores of ready-to-drink green tea beverages. *Food Res. Int.* **103**, 68–75 (2018).
52. Wang, S. Y. & Camp, M. J. Temperatures after bloom affect plant growth and fruit quality of strawberry. *Sci. Hort.* **85**, 183–199 (2000).
53. MacKenzie, S. J. & Chandler, C. The late year decline in strawberry fruit soluble solid content observed in Florida is caused by rising temperatures. *Acta Hort.* **842**, 843–846 (2009).
54. Zorrilla-Fontanesi, Y. et al. Genetic analysis of strawberry fruit aroma and identification of O-methyltransferase FaOMT as the locus controlling natural variation in mesifurane content[C][W][OA]. *Plant Physiol.* **159**, 851–870 (2012).
55. Sánchez-Sevilla, J. F. et al. Gene expression atlas of fruit ripening and transcriptome assembly from RNA-seq data in octoploid strawberry (Fragaria × ananassa). *Sci. Rep.* **7**, 1–13 (2017).
56. Aharoni, A. et al. Identification of the SAAT gene involved in strawberry flavor biogenesis by use of DNA microarrays. *Plant Cell* **12**, 647–662 (2000).
57. Pérez, A. G., Olías, R., Luaces, P. & Sanz, C. Biosynthesis of strawberry aroma compounds through amino acid metabolism. *J. Agric. Food Chem.* **50**, 4037–4042 (2002).
58. Leone, A. et al. Lipoxygenase involvement in ripening strawberry. *J. Agric. Food Chem.* **54**, 6835–6844 (2006).
59. Urrutia, M., Rambla, J. L., Alexiou, K. G., Granell, A. & Monfort, A. Genetic analysis of the wild strawberry (Fragaria vesca) volatile composition. *Plant Physiol. Biochem.* **121**, 99–117 (2017).
60. Moing, A. et al. Extensive metabolic cross-talk in melon fruit revealed by spatial and developmental combinatorial metabolomics. *N. Phytol.* **190**, 683–696 (2011).
61. Sánchez, G., Besada, C., Badenes, M. L., Monforte, A. J. & Granell, A. A non-targeted spproach unravels the volatile network in peach fruit. *PloS ONE* **7**, e38992 (2012).
62. Kumar, S. et al. Genome-wide scans reveal genetic architecture of apple flavour volatiles. *Mol. Breed.* **35**, 118 (2015).
63. Whitaker, V. M. et al. 'Florida brilliance strawberry. *Hort. Sci.* **54**, 2073–2077 (2019).
64. Bartoshuk, L. M. et al. Valid across-group comparisons with labeled scales: the gLMS versus magnitude matching. *Physiol. Behav.* **82**, 109–114 (2004).
65. Underwood, B. A. et al. Ethylene-regulated floral volatile synthesis in petunia corollas. *Plant Physiol.* **138**, 255–266 (2005).
66. Schmelz, E. A., Alborn, H. T., Banchio, E. & Tumlinson, J. H. Quantitative relationships between induced jasmonic acid levels and volatile emission in Zea mays during Spodoptera exigua herbivory. *Planta* **216**, 665–673 (2003).
67. Schmelz, E. A., Alborn, H. T. & Tumlinson, J. H. The influence of intact-plant and excised-leaf bioassay designs on volicitin-and jasmonic acid-induced sesquiterpene volatile release in Zea mays. *Planta* **214**, 171–179 (2001).
68. Whitaker, V. M., Hasing, T., Chandler, C. K., Plotto, A. & Baldwin, E. Historical trends in strawberry fruit quality revealed by a trial of University of Florida cultivars and advanced selections. *Hort. Sci.* **46**, 553–557 (2011).
69. van den Berg, RobertA., Hoefsloot, H. C., Westerhuis, J. A. & Smilde, A. K. & van der Werf, Mariët J. Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics* **7**, 142 (2006).
70. Suzuki, R. & Shimodaira, H. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* **22**, 1540–1542 (2006).
71. Chong, I. & Jun, C. Performance of some variable selection methods when multicollinearity is present. *Chemometrics Intell. Lab. Syst.* **78**, 103–112 (2005).
72. Wold, S., Sjöström, M. & Eriksson, L. Partial least squares projections to latent structures (PLS) in chemistry. *Encycl. Comput. Chem.* **3**, 523–550 (2002).
73. Barbey, C. R. et al. Genetic analysis of methyl anthranilate, mesifurane, linalool and other flavor compounds in cultivated strawberry (Fragaria ×ananassa). Preprint at https://www.biorxiv.org/content/10.1101/2020.10.07.330001v1.full (2020).
74. Tang, Y. et al. GAPIT version 2: an enhanced integrated tool for genomic association and prediction. *Plant Genome* **9** (2016).
75. van Dijk, T. et al. 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 (2014).

Downloaded from https://academic.oup.com/hr/article/doi/10.1038/s41438-021-00502-5/6446614 by guest on 05 August 2026