the plant journal

SEB SOCIETY FOR EXPERIMENTAL BIOLOGY

The Plant Journal (2023) **116,** 1201–1217 doi: 10.1111/tpj.16404

RESOURCE

Diversity of the volatilome and the fruit size and shape in European woodland strawberry (*Fragaria vesca*)

María Urrutia[1], Victoriano Meco[1], José Luis Rambla[2,3], Carmen Martín-Pizarro[1], Jeremy Pillet[1], Javier Andrés[4], José F. Sánchez-Sevilla[5], Antonio Granell[2], Timo Hytönen[4] and David Posé[1,*]

[1]*Departamento de Mejora Genética y Biotecnología, Instituto de Hortofruticultura Subtropical y Mediterránea (IHSM), Universidad de Málaga - Consejo Superior de Investigaciones Científicas, Departamento de Biología Molecular y Bioquímica, Facultad de Ciencias, UMA, Málaga, Spain,*

[2]*IBMCP Institute for Plant Molecular and Cell Biology (CSIC-UPV), Valencia, Spain,*

[3]*Department of Biology, Biochemistry and Natural Sciences, Universitat Jaume I, Castellón de la Plana, Spain,*

[4]*Department of Agricultural Sciences, Viikki Plant Science Centre, University of Helsinki, Helsinki, Finland, and*

[5]*Junta de Andalucía, Unidad Asociada CSIC I+D+i Biotecnología & Mejora de Fresa, Instituto Andaluz de Investigación & Formación Agraria y Pesquera (IFAPA), Ctr. IFAPA Málaga, Málaga, Spain*

Received 21 March 2023; revised 30 June 2023; accepted 17 July 2023; published online 19 August 2023.
*For correspondence (e-mail dpose@uma.es).

SUMMARY

Woodland strawberry (*Fragaria vesca* subsp. *vesca*) is a wild relative of cultivated strawberry (*F.* × *ananassa*) producing small and typically conical fruits with an intense flavor and aroma. The wild strawberry species, *F. vesca*, is a rich resource of genetic and metabolic variability, but its diversity remains largely unexplored and unexploited. In this study, we aim for an in-depth characterization of the fruit complex volatilome by GC–MS as well as the fruit size and shape using a European germplasm collection that represents the continental diversity of the species. We report characteristic volatilome footprints and fruit phenotypes of specific geographical areas. Thus, this study uncovers phenotypic variation linked to geographical distribution that will be valuable for further genetic studies to identify candidate genes or develop markers linked to volatile compounds or fruit shape and size traits.

Keywords: ***Fragaria vesca*, strawberry, volatiles, VOC, aroma, fruit shape, natural population.**

Linked article: This paper is the subject of a Research Highlight article. To view this Research Highlight article visit https://doi.org/10.1111/tpj.16547.

INTRODUCTION

The woodland strawberry, *Fragaria vesca*, is a wild diploid species from the *Fragaria* genus (family *Rosaceae*), which includes 23 species with different levels of ploidy, including the commercial octoploid species *Fragaria* × *ananassa* (Liston et al., 2014), the most consumed berry crop worldwide (FAOSTAT http://faostat.fao.org/). *F. vesca* has a vast natural distribution, but with differences between its four subspecies. *F. vesca* subsp. *vesca* is widely distributed in the northern hemisphere, being naturally present across Eurasia (including Iceland) and eastern North America. However, the other three subspecies, that is, *F. vesca* subsp. *bracteata*, *americana*, and *californica*, are endemic to North America (Hilmarsson et al., 2017; Liston et al., 2014; Staudt, 2009).

Commonly, strawberry fruit refers to the fleshy part, which is derived from the enlargement of the floral receptacle, and the actual fruits, the achenes, which are embedded in this receptacle (Perkins-Veazie, 1995). The shape of the strawberry fruit is determined during the berry development by auxin and gibberellin (GA) phytohormones, which respectively promote horizontal and vertical growth (Liao et al., 2018; Wang et al., 2017). Furthermore, several QTLs for natural shape variation have been reported in both *F. vesca* and *F.* × *ananassa* (Cockerton et al., 2021; Lerceteau-Köhler et al., 2012; Rey-Serra et al., 2021; Urrutia et al., 2015).

Strawberry fruit ripening, as a non-climacteric fruit, is mainly promoted by the phytohormone abscisic acid (ABA) (Kim et al., 2019; Li et al., 2019; Liao et al., 2018). This process involves a series of complex and tightly

coordinated physiological and biochemical changes in the enlarged receptacle, including color, texture, flavor, and aroma, as a strategy for seed dispersal (Fait et al., 2008; Li et al., 2020; Wang et al., 2021). In particular, the changes in flavor and aroma are mainly due to the accumulation of volatile (VOCs) and non-volatile secondary metabolites. Plant volatile cues mediate plant communication with the environment carrying information about the identity of the emitter, being specific volatile ratio combinations heritable (Karban et al., 2014; Ninkovic et al., 2016).

Although strawberries are very valued for their fine and delicate flavor and aroma, as with many other fruit species such as tomato, these traits have not been historical priorities in *F.* × *ananassa* breeding programs (Klee & Tieman, 2018). However, this tendency is changing toward programs that include sensory quality in their selection process, being the VOCs essential to achieve a pleasant flavor and aroma (Fan et al., 2021; Folta & Klee, 2016). Thus, in order to counteract the reduction in genetic diversity resulting from traditional crop breeding, current programs include wild close relatives of crop species in the search for new sources of genetic variability (Warschefsky et al., 2014). This phenomenon has also occurred in the cultivated strawberry, where a small but significant genetic diversity reduction has been reported (Gil-Ariza et al., 2009). The study of the wild woodland strawberry, and particularly its Eurasian subpopulation, which is the most diverse among the species (Hilmarsson et al., 2017), is expected to reveal new genetic and metabolic diversity in the *Fragaria* genus.

Despite the complexity of the strawberry fruit volatilome, with more than 979 identified compounds (Ulrich et al., 2018), only around 20 of them have been found to date to contribute to fruit aroma and flavor perception from the human consumers' perspective. These contributing compounds include furanones that add sweet-caramel notes (mesifurane and furaneol), lactones contributing with peach-like notes (γ-decalacatone and γ-dodecalactone), the phenylpropene eugenol, with a spicy-nutmeg aroma, the acetic, butanoic, and hexanoic acid esters, which provide fruity and ester-like notes, and some terpenoids such as linalool, with a flowery-sweet touch (Du et al., 2011; Fan et al., 2021; Jetti et al., 2007; Klee & Tieman, 2018; Nuzzi et al., 2008; Schieberle & Hofmann, 1997; Schwieterman et al., 2014; Ulrich et al., 1997, 2018).

Interestingly, woodland strawberry fruit aroma is the result of a richer volatilome in terms of quantity and diversity compared with cultivated *F.* × *ananassa* fruits (Pyysalo et al., 1979; Ulrich et al., 2007; Ulrich & Olbricht, 2014). Thus, *F. vesca* develops fruits with a more intense, floral, and fruity aroma, mainly due to the presence of compounds such as methyl 2-aminobenzoate, methyl cinnamate, and by a greater variety and quantity of terpenoids (Negri et al., 2015; Ulrich et al., 1997, 2007; Ulrich & Olbricht, 2014).

Beyond the volatile-individual contribution to the aroma, some studies have focused on the contribution of these compounds to human perception in terms of acceptability and liking. These studies have pointed to individual volatiles that are positively correlated with positive perception, such as linalool, nerolidol, γ-decalactone, γ-dodecalactone, furaneol, 2-pentenal, methyl butanoate, 2-methyl butanoate, 1-methylbutyl butanoate, 3-methylbutyl butanoate, pentyl butanoate, hexyl butanoate, hexyl acetate (*Z*)-3-hexenyl acetate, and 1-penten-3-one. On the contrary, other VOCs, that is, mesifurane, butyl acetate, pentyl acetate, ethyl 2-methyl butanoate, ethyl 3-methyl butanoate, pentenal, (*E*)-2-hexenal, and 2-pentanone, are negatively correlated to consumer's acceptance (Schwieterman et al., 2014; Ulrich & Olbricht, 2016). Furthermore, it has been described that a simultaneous high content of terpenoids and lactones is highly correlated with a positive consumers' perception (Ulrich & Olbricht, 2016).

In recent years, a number of enzymes involved in the biosynthetic pathway for different key strawberry VOCs have been identified. However, we are still far from a complete description of those pathways. Among them, an anthranilic acid methyl transferase (*FanAAMT*) catalyzing the final steps in the synthesis of methyl 2-aminobenzoate (also named methyl anthranilate) and two anthranilic acid methyl transferase-like (*FanAAMT-like*) have been identified (Barbey et al., 2021; Pillet et al., 2017). This compound is rarely detected in cultivated strawberry, being responsible for the typical grape note of the wild strawberries' aroma. The glucosylation of methyl cinnamate, also normally absent in cultivated strawberry and a contributor to the pleasant spicy notes to the fruit scent (Ulrich et al., 2007), has been shown to be catalyzed by a UDP-Glc: cinnamate glucosyltransferase (*FaGT2*) (Lunkenbein et al., 2006). The diversity of terpenoids in woodland strawberry aroma is a signature of the species (Ulrich & Olbricht, 2014). Among the structural genes involved in the biosynthesis of these compounds, a nerolidol synthase (*FaNES1*) has been associated with the synthesis of the sesquiterpene nerolidol (Aharoni, 2004). Furthermore, several candidate genes for the biosynthesis of multiple terpenes have been identified recently in a QTL region, including *FaNES1* (Barbey et al., 2021). Besides these typical woodland strawberry characteristic compounds, some genes have also been identified for the biosynthesis of VOCs normally present in *F.* × *ananassa*, including the methyltransferase *FaOMT* for the production of mesifurane (Barbey et al., 2021; Zorrilla-Fontanesi et al., 2012), and the fatty acid desaturase *FaFAD1* for γ-decalactone (Chambers et al., 2014; Sánchez-Sevilla et al., 2014). Finally, three biosynthetic and regulatory genes (*EGS1*, *EGS2*, and *EOBII*) have been described in the eugenol biosynthetic pathway in the octoploid species (Araguez et al., 2013; Medina-Puche et al., 2015; Molina-Hidalgo et al., 2017).

The fruit volatilome is a complex mix of secondary metabolites that tends to be lineage-specific (Ninkovic et al., 2021) and that evolves due to the duplication and subsequent diversification of genes encoding for their biosynthetic enzymes. This commonly results in changes in the preferred substrate or resultant volatile product with minimal changes in the primary amino acid sequence (Pichersky et al., 2006). Metabolic markers have been shown to be useful in order to discriminate between landraces and genotypic groups in several crops such as maize, sunflower, cassava, rapeseed, pepper, and grape (Billet et al., 2018; Fernandez et al., 2019; Lamari et al., 2018; Perez-Fons et al., 2020; Venkatesh et al., 2016; Wagner et al., 2012; Wahyuni et al., 2013). In addition, studies comparing the prediction power of genetic and metabolic markers in maize concluded that metabolites are condensed information of the genotype, with a set of 130 metabolites being almost as good predictors as 38 000 SNPs (Riedelsheimer, 2012). As previously mentioned, wild and domesticated strawberry species can be differentiated based on secondary metabolism. In addition, the relative abundance of secondary metabolites among wild strawberry accessions is very variable (Vallarino et al., 2018). Thus, a more comprehensive analysis of the woodland strawberry volatilome would enhance the comprehension of strawberry metabolic diversity and might provide suitable biomarkers to trace plant divergence.

Volatilome as well as fruit shape and size are rich, diverse, and variable characters in woodland strawberry and depend on genotypes and environments. Previous studies characterizing woodland strawberry volatilome and fruit shape have focused on interspecific biparental populations (Urrutia et al., 2015, 2017) or a few accessions (Dong et al., 2013; Li et al., 2020; Negri et al., 2015; Ulrich et al., 2007; Ulrich & Olbricht, 2013, 2014), limiting the range of studied diversity of the species. Thus, this work aims at an in-depth characterization of the woodland strawberry species (*F. vesca* sub. *vesca*) volatilome and fruit shape across a wide geographical area, therefore reflecting secondary metabolites adaptation to environmental conditions across the European continent. In addition to revealing a richness of resources for cultivated strawberry breeding, this work will open opportunities for the identification of genetic markers and candidate genes for VOCs of interest.

RESULTS

Natural volatilome diversity of woodland strawberry fruit across Europe

Berries at their stage of full ripeness were harvested from the European *F. vesca* collection at the University of Helsinki (Helsinki, Finland). A total of three biological replicates were harvested from 125 and 170 accessions in 2016 and 2017, respectively, of which 113 were common between both years. The volatilome profiling of fruit samples was performed by GC-MS, resulting in the relative quantification of 99 unambiguously identified metabolites, which were classified depending on their biosynthetic pathway. Thus, the great majority of the identified compounds (45 in total) belonged to the esters or benzenoid esters biosynthetic pathways. Other abundant VOCs biosynthetic pathways included those for fatty acid degradation (19 volatiles) and cuticular-wax biosynthesis (10 volatiles). Other pathways for different chemical families were less represented, that is, terpenoid synthesis (6), carotenoid cleavage (4), volatile benzenoid biosynthesis (8), furans (4), lactones (4), and carbohydrate degradation (1). In addition, there were two ketonic compounds that were not assigned to any described pathway (Table 1). Among all these compounds, we detected and quantified most key VOCs for cultivated strawberry aroma including a variety of lactones (γ-decalacatone, γ-dodecalactone, γ-octalactone, and δ-decalactone) and terpenoids (α-farnesene, α-pinene, α-terpineol, linalool, nerol, and myrtenol), several key esters (methyl butanoate, 1-methylbutyl butanoate, hexyl butanoate, hexyl acetate, and (*Z*)-3-hexenyl acetate), furanes (mesifurane, furaneol), (*E*) 2-pentenal, and 1-penten-3-one. Furthermore, compounds rarely detected in commercial strawberry varieties, but expected in *F. vesca* fruits were also found, including VOCs that add the typical notes to their aroma, such as methyl 2-aminobenzoate and methyl cinnamate (Table 1, Table S1). All investigated compounds showed a quantitative distribution, with great differences detected between accessions, with ranges going from approximately eightfold difference, that is, for 3,4-dimethylbenzaldehyde or 6-methyl 5-hepten-2-one, to >1000 fold for mesifurane, methyl 2-aminobenzoate, γ-decalactone, ethyl octanoate, ethyl decanaoate, and ethyl dodecanoate (Table 1).

Genetic and environmental effects on the volatilome

VOCs differed in their stability between harvests as shown by Pearson's correlation analysis (Table 1). Interestingly, several terpenoids, lactones, and volatile esters, known to positively contribute to aroma perception and consumer's acceptance, showed a high correlation between the two harvest seasons (>0.5), revealing a steady distribution of these compounds in both seasons and supporting a high genetic effect on their accumulation (Table 1). Analysis of variance (ANOVA) between geographical origins revealed significant differences between subpopulations and between accessions for most of the quantified VOCs (96 and 92 out of the total 99 compounds, respectively) (Table S2), which indicates that there exist ample differences in the European woodland strawberry volatilome associated with their region of origin. The estimated effect size of the genotype (ω^2_G) was higher than the effect size of the harvest year (ω^2_E) for 79 of the VOCs identified. On the contrary, and consistently with their low correlation

Table 1 Diversity and variability of identified volatile compounds in fruits from a European woodland strawberry germplasm. Average (mean), standard deviation (SD), and range values in two independent harvest seasons, as well as Pearson correlation between harvests (corr), are provided. Relative quantification values are given in log2 scale. +, −, N/D: compounds with positive, negative or non-determined effect respectively over consumer's acceptance according to bibliography (Schwieterman et al., 2014; Ulrich & Olbricht, 2016)

		First season (H16)			Second season (H17)				
Compound	Pathway	Mean	SD	Range	Mean	SD	Range	Corr H16-H17	Effect on perception
γ-decalactone	Lactones	−2.07	2.44	10.52	−2.04	2.65	11.49	0.89	+
Methyl benzoate	Volatile benzenoid esters biosynthesis	−0.56	1.63	7.68	−1.37	1.29	7.29	0.81	N/D
Methyl cinnamate	Volatile benzenoid esters biosynthesis	−0.97	2.23	8.47	−1.74	2.33	8.5	0.8	N/D
Eugenol	Volatile benzenoid biosynthesis	−2.46	1.69	9.57	−1.84	2.17	10.72	0.77	N/D
2-undecanone	Cuticular-wax biosynthesis	−0.6	1.66	6.4	−1.43	1.73	7.48	0.76	N/D
α-farnesene	Terpenoid synthesis	−0.02	1.43	5.22	−1.81	1.19	4.9	0.73	N/D
Methyl dodecanoate	Volatile esters biosynthesis	−0.72	2.02	9.76	−2.47	2.05	9.52	0.7	N/D
2-nonanone	Cuticular-wax biosynthesis	−0.13	0.81	3.88	−1.25	1.4	6.53	0.7	N/D
Methyl 2-aminobenzoate	Volatile esters biosynthesis	−3.15	3.1	11.28	−3.49	3.31	11.15	0.69	N/D
Mesifurane	Furanes	−3.06	3.47	12.31	−2.44	3.21	12.31	0.69	−
Myrtenyl acetate	Volatile esters biosynthesis	−0.57	1.1	5.03	−0.3	0.86	5.35	0.68	N/D
γ-dodecalactone	Lactones	0	0.72	2.75	0.24	1.39	6.45	0.68	+
Methyl 3-hydroxyoctanoate	Volatile esters biosynthesis	−1.14	1.33	6.12	−1.74	2.07	8.27	0.66	N/D
Methyl decanoate	Volatile esters biosynthesis	−0.2	1.39	8.17	−1.79	1.92	10.97	0.64	N/D
2-heptanone	Cuticular-wax biosynthesis	−0.32	0.6	3.09	−0.77	1.17	5.07	0.63	N/D
3-methyl 2-butenyl acetate	Volatile esters biosynthesis	−0.22	1.01	7.46	−0.73	1.16	6.32	0.61	N/D
2-nonanol	Cuticular-wax biosynthesis	−1.03	2.14	11.85	−2.13	2.13	9.77	0.61	N/D
Linalool	Terpenoid synthesis	−0.75	0.8	3.83	0.97	1.04	5.69	0.6	+
2-undecanol	Cuticular-wax biosynthesis	−1.58	2.15	8.92	−2.16	2.34	9.74	0.59	N/D
1-methylethyl butanoate	Volatile esters biosynthesis	−0.02	1.15	4.66	−0.66	0.96	4.44	0.58	N/D
Octyl butanoate	Volatile esters biosynthesis	−0.15	1.53	7.93	−2.15	1.87	8.49	0.58	N/D
2-pentanone	Cuticular-wax biosynthesis	−0.17	1.01	4.84	−0.65	1.57	7.41	0.58	−
β-ionone	Carotenoid cleavage	−0.37	0.56	3.44	0.25	0.68	3.88	0.58	N/D
1-methylbutyl butanoate	Volatile esters biosynthesis	−0.37	1.38	6.26	−2.22	1.8	8.11	0.57	+
Cinnamyl acetate	Volatile benzenoid esters biosynthesis	−1.31	2.04	9.06	−1.32	2.14	9.37	0.57	N/D
Acetophenone	Not assigned	0.1	0.68	4.78	−0.23	0.54	3.55	0.57	N/D
Octyl acetate	Volatile esters biosynthesis	−0.11	0.97	6.22	−1.16	1.5	7.07	0.56	N/D
α-pinene	Terpenoid synthesis	−0.07	1.27	6.7	−1.12	1.31	7.09	0.56	N/D
Methyl hexanoate	Volatile esters biosynthesis	−0.05	1.04	4.97	−1.61	1.71	9.38	0.55	N/D
Octyl hexanoate	Volatile esters biosynthesis	−0.74	1.99	9.17	−2.84	2.37	10.38	0.55	N/D
Ethyl methylthioacetate	Volatile esters biosynthesis	−0.94	1.58	7.64	−1.75	2.03	8.15	0.55	N/D
δ-decalactone	Lactones	−0.4	1.13	4.93	−0.85	1.38	6.77	0.55	N/D
2-pentylfuran	Furanes	−0.11	0.98	4.57	0.03	0.54	3.91	0.54	N/D
1-methyloctyl butanoate	Volatile esters biosynthesis	−0.62	1.73	8.74	−2.41	1.83	10.06	0.53	N/D
Hexyl butanoate	Volatile esters biosynthesis	−0.24	1.33	7.42	−0.96	1.36	7.06	0.53	+
α-ionone	Carotenoid cleavage	−0.41	0.59	2.47	0.29	0.81	4.38	0.53	N/D
Methyl octanoate	Volatile esters biosynthesis	0.08	0.94	4.99	−1.25	1.67	9.14	0.52	N/D
Benzyl acetate	Volatile benzenoid esters biosynthesis	−0.18	0.84	4.1	−0.11	0.75	5.37	0.52	N/D
Butyl hexanoate	Volatile esters biosynthesis	−0.17	1.28	6.55	−1.47	1.73	9.3	0.51	N/D
Myrtenol	Terpenoid synthesis	−0.6	1.56	10.87	−1.19	1.57	7.99	0.51	N/D
α-terpineol	Terpenoid synthesis	−0.17	0.81	3.89	−0.03	0.79	5.26	0.5	N/D
2-heptanol	Cuticular-wax biosynthesis	−0.6	1.81	9.67	−1.75	1.67	6.99	0.5	N/D
Ethyl dodecanoate	Volatile esters biosynthesis	−2.8	3.27	11.64	−6.39	2.91	11.88	0.49	N/D
Acetone	Not assigned	−0.06	1.43	6.53	0.24	1.17	5.29	0.49	N/D
2-heptyl acetate	Volatile esters biosynthesis	−0.14	0.98	5.76	−2.23	1.68	8.22	0.48	N/D
Butyl butanoate	Volatile esters biosynthesis	−0.38	1.44	8.81	−1.35	1.76	9.01	0.48	N/D
2-methylbutyl acetate	Volatile esters biosynthesis	0.1	0.87	5.56	−0.83	1.08	5.91	0.47	N/D
Hexyl hexanoate	Volatile esters biosynthesis	−0.23	1.48	7.15	−2.23	1.75	7.68	0.47	N/D
2-tridecanone	Cuticular-wax biosynthesis	−0.29	1.26	6.98	−1.43	1.41	7.67	0.45	N/D
Benzaldehyde	Volatile benzenoid biosynthesis	−0.15	0.56	3.19	0.24	0.42	2.65	0.44	N/D
Decyl acetate	Volatile esters biosynthesis	−0.3	1.23	8.01	−1.91	1.86	10.31	0.43	N/D
2-pentadecanone	Cuticular-wax biosynthesis	−0.45	1.4	7.78	−1.5	1.65	7.9	0.43	N/D

(continued)

Table 1. (continued)

Compound	Pathway	First season (H16)			Second season (H17)			Corr H16-H17	Effect on perception
		Mean	SD	Range	Mean	SD	Range		
1-methylethyl acetate	Volatile esters biosynthesis	0.1	0.97	5.16	0.77	1.02	5.04	0.42	N/D
Ethyl hexanoate	Volatile esters biosynthesis	−0.8	1.53	8.66	−3.7	2.56	10.27	0.41	N/D
Butyl acetate	Volatile esters biosynthesis	−0.12	0.88	6.06	−0.87	1.21	6.04	0.39	−
Furaneol	Furanes	−0.32	1.17	5.57	−0.38	1.23	6.61	0.39	+
3-methylbutyl acetate	Volatile esters biosynthesis	0.23	1.02	3.9	−0.95	1	5.17	0.38	N/D
1-decanol	Fatty acid degradation	−0.8	0.92	5.1	−1.33	1.61	8.76	0.38	N/D
Ethyl butanoate	Volatile esters biosynthesis	−0.13	1.02	6.21	−2.15	1.74	8.28	0.36	N/D
Ethyl decanoate	Volatile esters biosynthesis	−1.43	2.5	11.48	−5.35	2.89	12.17	0.36	N/D
Ethyl octanoate	Volatile esters biosynthesis	−1.17	2.29	11.2	−4.63	2.98	12.12	0.36	N/D
Propyl butanoate	Volatile esters biosynthesis	0.04	1.07	5.56	−1.27	1.45	7.47	0.36	N/D
2-tridecanol	Cuticular-wax biosynthesis	−0.61	1.16	6.83	−1.34	1.63	8.14	0.36	N/D
Ethyl 2-hexenoate	Volatile esters biosynthesis	−1.06	2.09	8.55	−4.1	1.91	7.84	0.35	N/D
Nonyl acetate	Volatile esters biosynthesis	0.08	0.81	4.13	−0.87	0.99	4.84	0.35	N/D
1-octanol	Fatty acid degradation	−1.16	1.5	7.05	−1.33	1.27	7.54	0.35	N/D
Methyl butanoate	Volatile esters biosynthesis	0.14	1.11	6.11	−0.79	1.75	9.25	0.34	+
(E)-2-hexenyl acetate	Volatile esters biosynthesis	−0.81	1.05	5.68	0.96	0.9	4.94	0.32	N/D
2,3-butanediyl diacetate	Volatile esters biosynthesis	−0.35	1.26	6.04	−2.12	1.34	5.87	0.32	N/D
Hexanal	Fatty acid degradation	−0.36	0.5	2.9	0.37	0.44	3.59	0.32	N/D
Benzyl alcohol	Volatile benzenoid biosynthesis	−0.85	0.76	3.12	−0.31	0.95	5.43	0.31	N/D
Octanal	Fatty acid degradation	−1.6	1.88	8.36	0.07	0.9	6.27	0.3	N/D
(Z)-3-hexenyl acetate	Volatile esters biosynthesis	−1	1.06	4.99	0.85	0.93	4.55	0.29	+
Hexyl acetate	Volatile esters biosynthesis	0.01	0.47	2.38	−0.38	0.6	4.8	0.29	+
Ethyl acetate	Volatile esters biosynthesis	−1.18	2.12	9.73	−2.96	2.22	10.43	0.28	N/D
γ-octalactone	Lactones	0.09	0.63	3.83	−0.64	0.75	4.69	0.26	N/D
(Z)-3-hexenal	Fatty acid degradation	−0.19	0.52	2.56	1.1	0.72	3.64	0.26	N/D
Geranylacetone	Carotenoid cleavage	−0.61	0.99	5.1	−0.04	0.64	4.11	0.24	N/D
Nonanal	Fatty acid degradation	−0.92	1.34	10.15	−0.63	1.12	8.76	0.23	N/D
Decanal	Fatty acid degradation	−0.67	0.92	4.52	0.24	0.57	4.83	0.22	N/D
(E)-2-nonenal	Fatty acid degradation	−0.21	0.57	3.01	−0.42	0.49	4.63	0.21	N/D
(E)-2-octenal	Fatty acid degradation	−0.42	0.75	4.97	0	0.66	4.94	0.2	N/D
(E, E)-2–4 heptadienal	Fatty acid degradation	−0.45	0.86	4.88	−0.1	0.85	4.86	0.2	N/D
(E)- 2-hexen 1-ol	Fatty acid degradation	−0.58	1	5.6	1.88	0.94	6.04	0.18	N/D
Pentyl acetate	Volatile esters biosynthesis	−0.12	0.87	5.51	−0.26	0.68	4.3	0.16	−
Heptanal	Fatty acid degradation	−0.63	1	6.24	0.19	0.51	3.67	0.16	N/D
Methyl acetate	Volatile esters biosynthesis	−0.29	1.3	6.17	0.69	1.66	8.72	0.14	N/D
3, 4-dimethylbenzaldehyde	Volatile benzenoid biosynthesis	−0.14	0.34	2.7	−0.47	0.36	2.68	0.1	N/D
2-(1-pentenyl)furan	Furanes	−0.66	0.8	3.51	0.71	0.79	5.42	0.09	N/D
6-methyl 5-hepten-2-one	Carotenoid cleavage	−0.33	0.5	2.59	0.04	0.42	2.68	0.08	N/D
Ethanol	Carbohydrate degradation	−2.54	1.7	10.16	−2.53	2.1	10.69	0.07	N/D
(E)-2-pentenal	Fatty acid degradation	−0.91	0.86	4.36	1.26	0.76	4.16	0.06	+
Pentanal	Fatty acid degradation	−0.18	0.82	3.9	0.43	0.59	4.94	0.06	−
Nerol	Terpenoid synthesis	−0.96	0.8	3.48	−2	1.67	8.18	0.05	N/D
1-hexanol	Fatty acid degradation	−1.28	1.23	5.7	1.15	1.01	5.76	0.05	N/D
(E)-2-heptenal	Fatty acid degradation	−0.36	0.72	3.44	0.1	0.64	5.34	0	N/D
(E)-2-hexenal	Fatty acid degradation	−0.19	0.37	1.89	0.31	0.49	4.91	−0.03	−
1-penten-3-one	Fatty acid degradation	−1.35	1.09	4.93	1.32	0.8	8.32	−0.05	+
1-penten-3-ol	Fatty acid degradation	−1.48	1.16	5.38	−0.06	0.85	5.02	−0.12	N/D

values between the two harvest seasons, the estimated effect size of the harvest year was higher for eight VOCs, that is, (*E*)-2-pentenal, 1-penten-3-one, (*E*)-2- and (*Z*)-3-hexenyl acetate, (*Z*)-3-hexenal, (*E*)-2-hexen-1-ol, 2-heptyl acetate, and linalool, supporting a significant effect on their content by the environmental conditions (Table S2). However, for the vast majority of the volatiles both the genotype and the environmental effects were significant. Therefore, least square means per genotype were estimated and used in further statistical analyses (Table S3).

Woodland strawberry chemotypes across Europe

In order to investigate the distribution of wild strawberry aroma across the European continent, we classified the accessions into chemotypes based on high (above the mean) or low (below the mean) content in total terpenoids

and lactones, since high content of these compounds has been described as one of the most correlated traits to consumer's liking (Ulrich & Olbricht, 2016). Thus, we combined the content for the most genetically controlled (correlation between harvest >0.5) terpenoids (α-farnesene, α-pinene, α-terpineol, linalool, and myrtenol) and lactones (γ-decalactone, γ-dodecalactone, and δ-decalactone). In order to give the same weight to all added variables, values were rescaled before the addition (Table S3). This classification resulted in four chemotypes: high terpenoids and high lactones (High_TL), high terpenoids and low lactones (High_T), low terpenoids and high lactones (High_L), and low terpenoids and low lactones (Low_TL) (Table S3), which geographical distribution is depicted in Figure 1. In addition, the accessions were assigned to seven major geographical groups covering main western European landforms and islands, named south to north: Iberian Peninsula, The Alps, Middle Europe, Great Britain, Baltic Sea Region, Scandinavian Peninsula, and Iceland (Table S4). Although all chemotypes are present in all geographical areas, High_TL chemotype was more frequent among Icelandic and Central Europe accessions, High_L and High_T chemotypes were common in the Iberian Peninsula and the Baltic Sea Region, respectively, while the Low_TL accessions were commonly found in the Scandinavian region (Figure 1).

A principal components analysis (PCA) computed on LSmeans by genotype revealed that four of the geographical groups were clustered based on their volatilome profile (Figure 2). Samples from Icelandic and Baltic Sea regions were the most different based on PC1 and PC2, while PC3 and PC4 separated Iberian and Scandinavian peninsula samples revealing that subpopulations from extreme European latitudes were distinguishable based on their volatilome. However, accessions from intermediate latitude regions (Middle Europe, Alpine, and Great Britain) were not distinguishable within the first four principal components.

We next grouped the *F. vesca* collection based on their volatilome profiles. For that purpose, we performed a bootstrapped hierarchical clustering analysis (HCA) on LSmean data per genotype. Accessions clustered into five

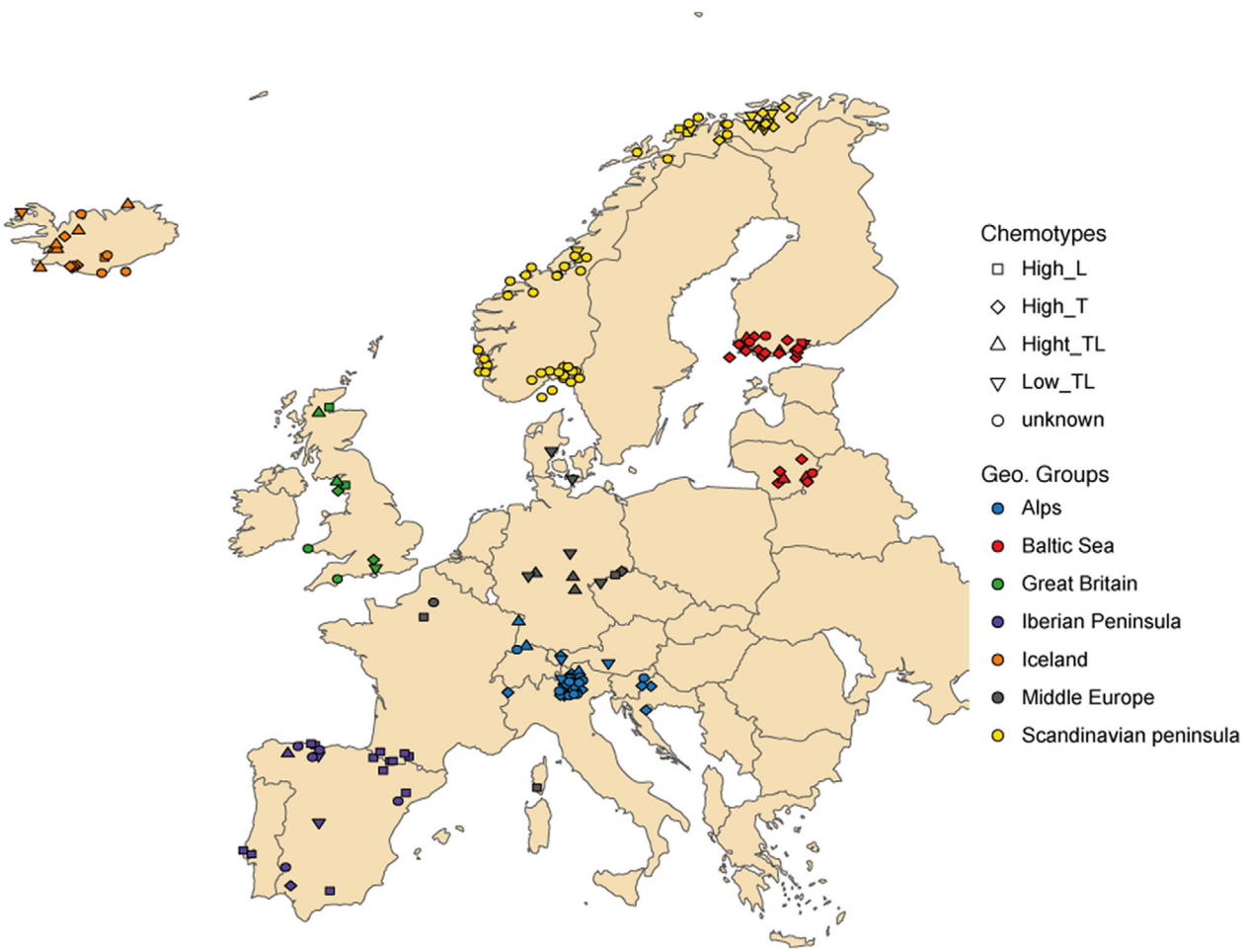

Figure 1. Chemotypes and geographical origins and grouping of the accessions from the European woodland strawberry collection.

main groups (A to E). For clarity, cluster E was divided into two subclusters (E1 and E2). This analysis showed that three of the clusters (C, E1, and E2) mostly corresponded to the four geographical groups that were distinguished by PCA. Thus, most Icelandic samples clustered in group C, most Iberian and some Alpine accessions grouped in cluster E1, while cluster E2 gathered the majority of the northern accessions, including those from the Baltic Sea Region and from the Scandinavian Peninsula. On the contrary, clusters B and D gathered a mix of accessions from different geographical origins, most of them being from the Alps, Middle Europe, and Great Britain. Finally, cluster A only comprised two accessions that came from distant populations, that is, UK6 from Scotland and NOR21 from northern Norway. The Great Britain geographical group was especially promiscuous considering their reduced sample size, with seven accessions that belonged to four different clusters.

Volatilome footprints in the European woodland strawberry collection

The bootstrapped HCA on the LSmeans volatiles values revealed volatilome footprints across the European continent. Volatiles were classified into two clear distinct clusters, named 1 and 2. Cluster 1 was also divided into three subclusters: 1.1, 1.2, and 1.3 (Figure 3). In many cases, compounds belonging to the same biosynthesis pathway clustered together. For instance, cluster 2 contains six ethyl esters and their precursor, ethanol. Interestingly, cluster 2 also groups mesifurane and methyl 2-aminobenzoate, both compounds associated with intense aroma and negatively correlated to consumer's acceptance. Lactones (except γ-decalactone) were tightly grouped within the cluster 1.2. Apocarotenoids (geranylacetone, α-ionone, β-ionone, and 6-methyl-5-hepten-2-one) were strongly correlated and grouped together

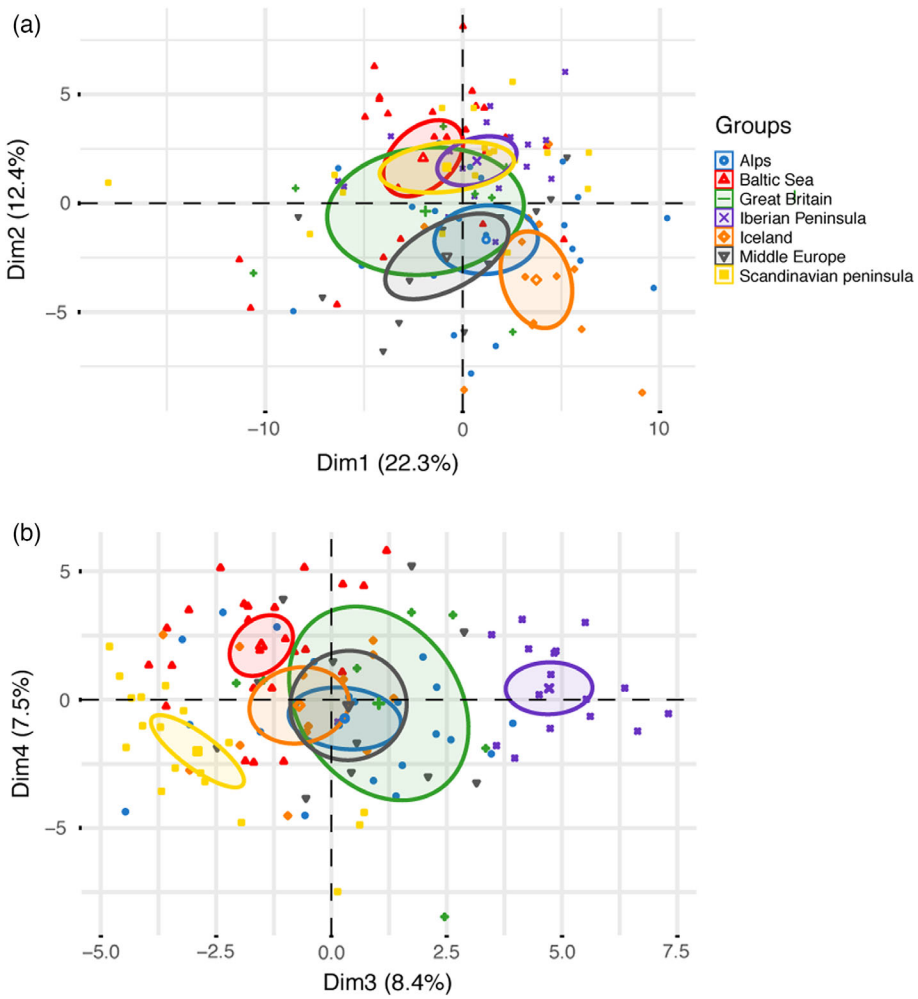

Figure 2. Principal component analysis (PCA) for the least square mean volatilome profile values of the *F. vesca* collection. Ellipses indicate 95% confidence intervals of geographical group means. (a) Principal components 1 and 2; (b) Principal components 3 and 4.

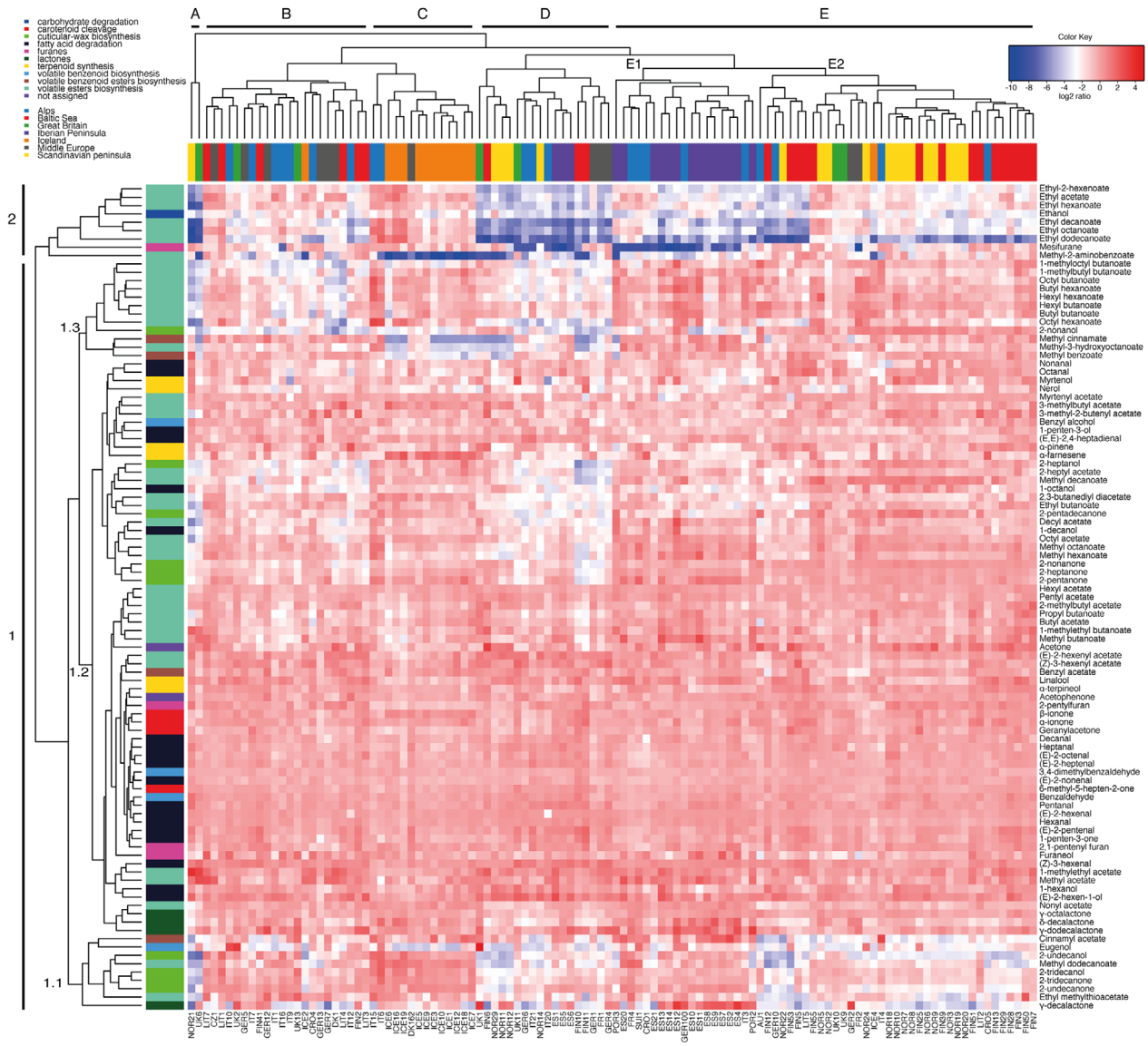

Figure 3. Hierarchical clustering analysis (HCA) and heatmap representation of the volatilome of 113 European *F. vesca* accessions. Data represent least square means of two harvests. Accessions are presented in rows and grouped according to their geographical origin (see legend). Volatile compounds are presented in columns and colored according to their biosynthesis pathway (see legend).

within the cluster 1.2, which is dominated by fatty acids derivates. Interestingly, methyl ketones (cuticular-wax biosynthesis pathway) were separated into two groups, one group gathering long chain (C11 and C13) methyl ketones and their secondary alcohols in cluster 1.1, and another group including short chain (C5, C7, and C9) and one long chain (C15) methyl ketones in cluster 1.2. Terpenoids clustered in three groups of two compounds within the cluster 1.2. Finally, the majority of the numerous volatile esters are grouped in clusters 1.2 and 1.3. It is noteworthy that cluster 1.3 gathers several long acyl chain esters (i.e., octyl hexanoate, octyl butanoate, and hexyl butanoate).

Volatilome signatures were found in those subpopulations separated by PCA and HCA. Thus, a pairwise comparison analysis revealed that the most southern accessions (Iberian Peninsula) were characterized by a higher content in γ-decalactone, cinnamyl acetate, and methyl hexanoate, while they showed lower levels of mesifurane, α-terpineol, α-pinene, and myrtenyl acetate, being this profile maintained in the 2016 and 2017 harvests (Figure 4a). In Icelandic samples, a higher accumulation of α-farnesene and several ethyl and methyl esters (e.g., ethyl 2-hexenoate, ethyl dodecanoate, and methyl dodecanoate) was found, while they contained lower levels of linalool, eugenol,

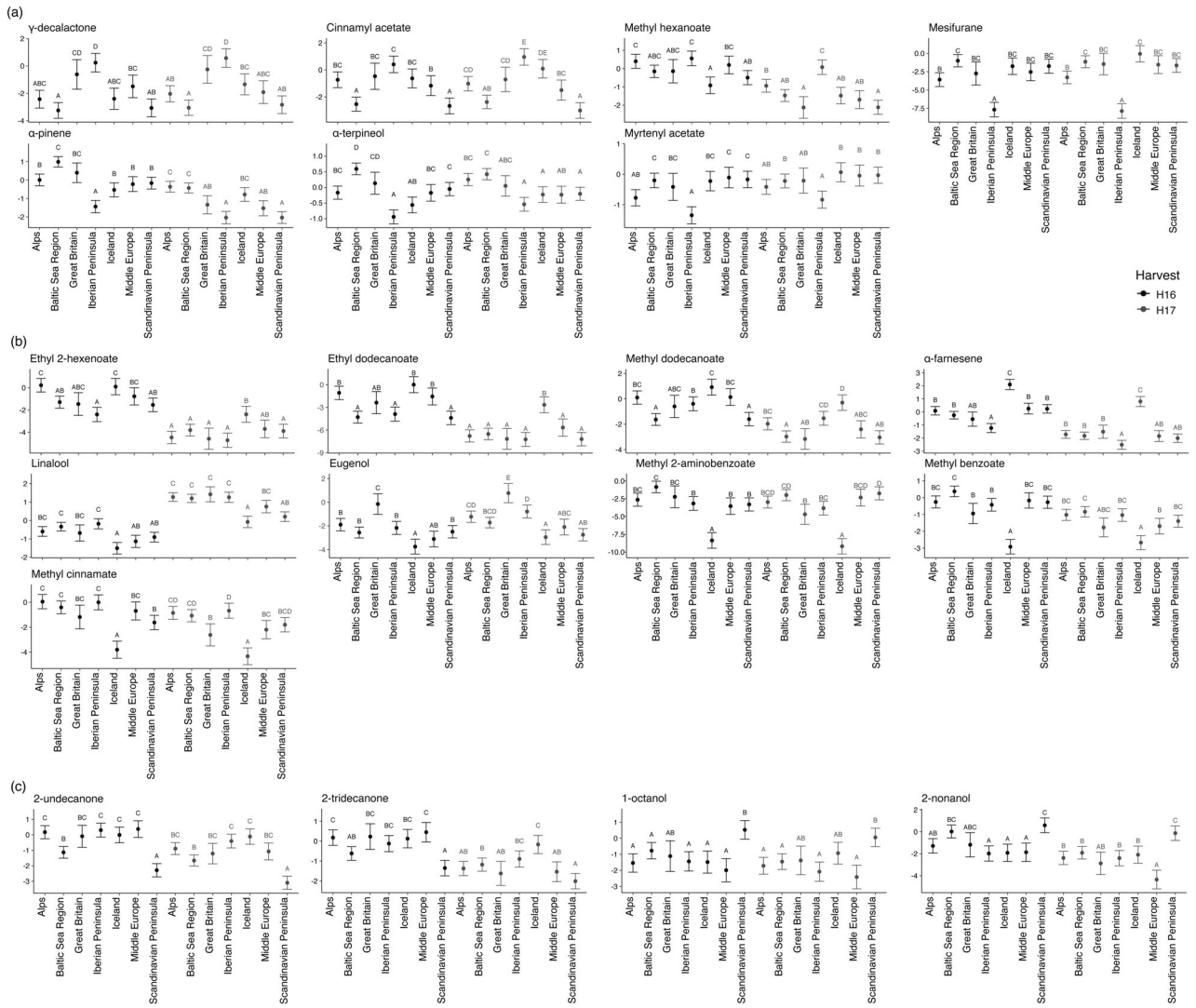

Figure 4. Volatile signatures of geographical areas. Tukey test graphical representation (*P*-value <0.05) of a selection of the most characteristic VOCs of specific geographical areas, such as Iberian Peninsula (a), Iceland (b), and Scandinavian Peninsula and Baltic Sea region (c) in two harvest seasons (H16 and H17).

methyl 2-aminobenzoate, methyl benzoate, and methyl cinnamate (Figure 4b). Scandinavian Peninsula accessions were characterized by a higher content in the alcohols 1-octanol and 2-nonanol together with a lower accumulation of the methyl ketones 2-undecanone and 2-tridecanone. In addition, both Scandinavian Peninsula and Baltic Sea region samples showed lower accumulation of cinnamyl acetate and γ-decalactone (Figure 4a,c). Finally, Great Britain accessions, especially UK1 and UK2, were characterized by a higher eugenol content (Figure 4b).

Besides individual VOCs that could be characteristic of a specific geographic area, we next evaluated, as a proof of concept, if a reduced set of these compounds could serve as a signature to classify an *a priori* unknown accession into a geographical origin category. For that purpose, we performed a sparse partial least square discriminant analysis (sPLS-DA) on groups separated by PCA, joining Middle Europe, Great Britain, and The Alps into a new group called 'Central Europe' (see experimental procedures). The model correctly predicted the origin of all Icelandic samples and committed just one miss assignment on Iberian Peninsula and Baltic Sea Region samples. However, the accuracy was below 50% for Scandinavian Peninsula samples and very poor for Central Europe samples, as expected by their less diverse volatilome profile (Table 2). Variables that were consistently selected for the prediction model are reported in Table S5, revealing that the most important compounds for geographical group separation are cinnamyl acetate, methyl-2-aminobenzoate, methyl benzoate, 2-pentylfuran, methyl cinnamate, α-farnesene, α-pinene, methyl-3-hydroxyoctanoate, α- and β-ionone, geranylacetone, and 2-nonanone among others.

Table 2 sPLS-DA prediction performance on testing set

	Test set origin				
Test set prediction	Baltic Sea R.	Central Europe	Iberian Peninsula	Iceland	Scandinavian Peninsula
Baltic Sea R.	7	8	0	0	5
Central Europe	0	4	1	0	0
Iberian Peninsula	0	3	10	0	0
Iceland	0	3	0	4	1
Scandinavian Peninsula	1	6	0	0	5

Diversity in fruit size and shape

Despite strawberries are typically considered as conical berries, they are very diverse in size and shape, varying from oblate to long conic berries. We phenotyped the European *F. vesca* collection for these traits by measuring fruit size (length, width, and volume) from all collected fruits in both harvests. Furthermore, fruit shape was inferred by the length/width ratio, being those fruits with this ratio >1 classified as elongated fruits, while those with a ratio ≤1 as rounded. A selection of fruits showing the diversity in fruit size and shape is shown in Figure 5a. Average fruits were 15.8–11.8 mm long and 11.9–12.3 mm wide in the first and second seasons, respectively, with variations of approximately 10 mm in fruit length between the shortest and largest accessions, and of approximately 7 mm in fruit width between the narrowest and widest accessions (Figure 5b). Clear differences were also detected in fruit volume in both harvest seasons, ranging from 0.31 to 1.40 cm^3 in the 2016 harvest and from 0.16 to 1.03 cm^3 in 2017. In addition, we also detected a range of fruit shapes in the collection, varying the length/width ratio from 0.90 to 1.77, and from 0.73 to 1.39 in the 2016 and 2017 harvests, respectively (Figure 5b). Average data per genotype can be found in Table S6.

Analysis of variance considering genotype (g) and harvest (e) as factors revealed that fruit length, width, and volume significantly varied among the accessions, being the effect size of the genotype (ω^2_G) bigger than that of the environment (ω^2_E) (Table S7). Interestingly, fruit length but not width or volume significantly varied in the two harvests as well, being fruits collected during the first season longer, but not wider, than during the second harvest (Figure 5b, Table S7). However, when the geographical origin (geo) of the accessions was considered as a factor, we observed that all fruit length, width, volume, and ratio were affected by both the geographical origin and the harvest year (Table S7). Pairwise comparisons between geographical origins revealed clear differences between fruit shape and size (Figure 5c). Fruit volume was strongly linked to geographical origin, being fruits from the Iberian Peninsula the biggest, those from northern locations (Iceland, Baltic Sea Region, and Scandinavian Peninsula) the smallest, and those from Middle Europe, the Alps, and Great Britain intermediate in size. Interestingly, these differences in volume were mainly driven by the differences in fruit width. Thus, both fruit width and volume were strongly negatively correlated with latitude, being those fruits from southern accessions wider, and consequently, bigger (Figure S1). Fruit length also tended to be shorter in northern accessions, especially in Scandinavian samples (Figure 5b), but the pattern was not so clear and its correlation with latitude was weaker (Figure S1). Regarding fruit shape, Iberian and Icelandic samples were the roundest and the most elongated, respectively, in both harvests, but other geographical origins were not so consistent with fruit shape. Thus, as found for the volatilome profile, fruit size seems to be associated with latitude, decreasing across the south–north axis.

DISCUSSION

The richness and complexity of woodland strawberry fruit volatilome have been revealed in previous studies (Negri et al., 2015), especially, in comparison with its cultivated relative, the commercial strawberry *F.* × *ananassa* (Dong et al., 2013; Pyysalo et al., 1979; Ulrich et al., 2007; Ulrich & Olbricht, 2013, 2014). However, those studies have been limited to a few accessions (Dong et al., 2013; Ulrich & Olbricht, 2013) or a segregating population (Urrutia et al., 2017). To our knowledge, this work constitutes the unprecedented volatilome analysis of a large-scale natural population covering one of the main areas of distribution of *F. vesca*, the European continent, and islands.

A review work by Ulrich *et al.* (Ulrich et al., 2018) highlighted that there is little consensus between the different published reports on strawberry (*F.* × *ananassa*) fruit volatile composition, since almost 70% of the total VOCs identified in this species were detected in just one of the 27 reviewed studies, whereas only around 30 compounds were commonly reported in the literature. This issue points to both biological (i.e., genotypes, growing conditions, and ripening stage) and technical causes (sampling preparation, extraction method, chromatography,

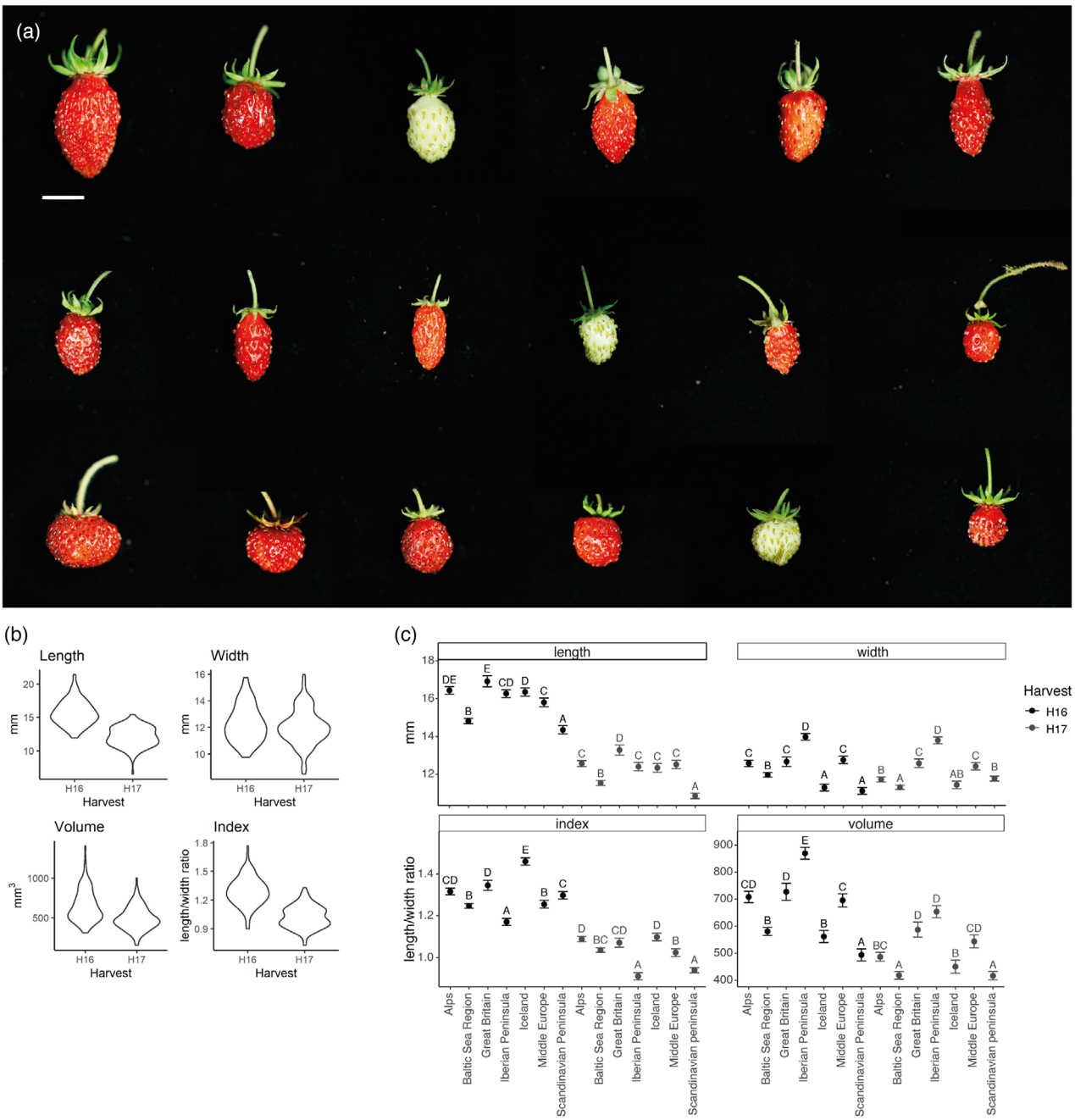

Figure 5. Diversity of fruit sizes and shapes in *F. vesca*. (a) Picture of a representative set of ripe berries from the *F. vesca* European collection. From left to right: Top row: ES2, IT15, FIN12, LIT2, IT7, and UK2. Middle row: ICE3, DK1, NOR22, GER2, ICE16, and GER10. Bottom row: POR3, ES12, UK1, FIN28, UK4, and NOR11. Scale bar, 1 cm. (b) Violin plots of average fruit size and shape. (c) Tukey test graphical representation of fruit size and shape (*P*-value <0.05).

and spectrometry techniques) behind these disagreements. In this study, we followed the experimental approaches described in Zorrilla-Fontanesi et al. (2012) and Urrutia *et al.* (Urrutia et al., 2017), identifying the same set of 99 compounds previously reported in *F. vesca* and 75% coincident with the compounds reported in *F.* × *ananassa* cv. Camarosa, which supports the reliability of the methodology used. Furthermore, the set of detected metabolites includes 22 of the 30 frequently reported VOCs described in Ulrich et al. (2018).

The expression of a plant phenotypic trait is the sum of its genetic potential, the particular environmental conditions where it is grown, and the interaction between them. Our results reveal a relevant genotypic effect on the studied traits, supported by a consistent and rather stable phenotype for most fruit VOCs, as well as fruit sizes and

shapes in two harvest years (Tables S4 and S7). In addition, genotypes from the same or closely located distribution areas showed remarkable phenotypic similarities, supporting a close genetic relationship between them and probable phenotypic adaptations to the environment (Karban et al., 2014; Ninkovic et al., 2016). Our study suggests that the phenotypes observed in plants grown in a common garden reflect the genetic potential of the plant. In fact, previous studies performed with different traits and species concluded that the geographical origin was determinant for environmental adaptations that remained expressed when plants were cultivated out of their geographical distribution of origin. For instance, a study carried out with European blueberries studying the association between their anthocyanin content and geographical origin concluded that differences between northern and southern varieties in anthocyanin content were maintained when they were cultivated together in the same environment (Åkerström et al., 2010). Similarly, studies with different varieties of *F.* × *ananassa* cultivated in several growing conditions detected genotype-specific secondary metabolites that allowed to differentiate between the cultivars, highlighting a prevalence of genetic effects over the environmental or year-to-year effects (Akhatou et al., 2016, 2017; Cocco et al., 2015; Josuttis et al., 2012). In addition, a study with five *F.* × *ananassa* varieties grown at different locations in Italy (with a 5° difference in latitude) reported strong dependence of fruit size on the genotype rather than the location (Cocco et al., 2015).

The stronger effect of the genotype over the environment for most VOCs is consistent with a previous volatilome study in ripe fruits from an interspecific near-isogenic line collection (NILs) with the genetic background of *F. vesca* and introgressions of *F. bucharica* (Urrutia et al., 2017). Similarly to our results, this study detected a high or moderate genetic effect on the accumulation of methyl 2-aminobenzoate, methyl cinnamate, myrtenyl acetate, eugenol, α-pinene, α-farnesene, methyl decanoate, methyl dodecanoate, benzyl acetate and methyl benzoate, γ-decalactone, 2-undecanone, cinnamyl acetate, and butyl butanoate. However, they also reported a high genetic effect in the accumulation of (*E*)-2-hexenal, (*Z*)-3-hexenal, and (*Z*)-3-hexenyl acetate in the NIL population (Urrutia et al., 2017), while these VOCs were among the least stable compounds in our population. Thus, we hypothesize that the effects of *F. bucharica* alleles of key genes on the biosynthetic pathways of these compounds resulted in drastic differences in their accumulation, enhancing the genetic effect. In the same way, previous volatilome studies with F1 progenies of *F.* × *ananassa* cultivars pointed to both genetic and environmental factors affecting volatile accumulation, detecting different degrees of stability between VOCs in samples harvested in different seasons (Olbricht et al., 2011; Zorrilla-Fontanesi et al., 2012). In agreement with our results, Olbricht and collaborators (2011) found 2-undecanone and methyl 2-aminobenzoate among the most stable compounds in their population. On the contrary, they reported γ-decalactone as the least stable compound in their population, while it was among the most stable in our study. Therefore, these differences highlight that choosing an appropriate population for a specific trait is key to the success of genetic mapping.

European woodland strawberry was divided into four chemotypes according to their total terpenoids and lactone accumulation. According to previous studies by Ulrich and Olbricht (2016), the group with high terpenoids and high lactones content (High_TL) is likely to present a more pleasant aroma profile from human's perspective. Thus, further organoleptic and molecular genetic studies of these chemotypes could uncover interesting genetic variants for aroma selection in the *Fragaria* genus.

In addition, European woodland strawberry subpopulations presented significant differences in fruit volatilome relative composition, being those differences increased with the geographical distances. Thus, subpopulations from extreme latitudes (Iberian Peninsula from the south, Iceland, Baltic Sea Region, and Scandinavian Peninsula from the North) showed distinctive volatilome footprints suggesting a low gene flow between those morphologically distant populations and an adaptation to their specific environmental conditions (Ninkovic et al., 2021). On the contrary, subpopulations from intermediate latitudes, such as The Alps, Middle Europe, and Great Britain, did not present clear distinctive volatilome footprints, likely due to a higher gene flow among them.

Icelandic accessions, the most isolated and northern group of samples, constituted the most distinctive cluster (cluster C in HCA), close to some Central European and Baltic accessions (cluster B) (Figure 3). However, the samples from the other island covered in this study, Great Britain, did not form a uniform cluster. These samples had diverse volatile profiles, which might be the result of a frequent gene flow enabled by the proximity of Great Britain to continental Europe. These results are consistent with a recent genetic analysis of 56 SSR markers in a global *F. vesca* collection, which divided European samples into two clusters, the Icelandic group and the rest of the Eurasian accessions (including Great Britain) and pointed to a closer genetic relation between Icelandic and Central European accessions than to those from Scandinavia (Hilmarsson et al., 2017). However, although no subpopulations were detected with molecular markers in the *F. vesca* from the Eurasian continent, PCA performed on volatilomes in this study identified three distinctive geographical groups based on volatile footprints on continental Europe, two from Northern areas (Baltic Sea Region and the Scandinavian Peninsula) and one from Southern area (Iberian Peninsula). Moreover, the volatile signature from the latter

was shared by some of the alpine accessions (CRO1, FR4, GER100, IT3, and SUI1) (Figure 2), suggesting a lineage continuity between those subpopulations. A sPLS-DA model confirmed that levels of few VOCs would be sufficient to classify Icelandic, Iberian Peninsula, and Baltic Sea Region samples with high accuracy. However, samples from Scandinavian Peninsula and Central Europe areas (The Alps, Great Britain, and Middle Europe) were more difficult to model, suggesting that further improvements to this model are needed.

The study of our natural collection in two independent harvests allowed the identification of compounds with a stable between-seasons pattern and a wide between-accessions range of accumulation. These compounds are likely more genetically regulated than others and are suitable for the search for candidate genes involved in their biosynthesis. Examples of these are γ-decalactone, methyl-2-aminobenzoate, methyl benzoate, methyl cinnamate, mesifurane or 2-undecanone among others. Interestingly, γ-decalactone showed a remarkable variable pattern among the geographical groups, being highly present in Iberian Peninsula accessions (Figure 4a) and, interestingly, in a few accessions of Middle Europe (i.e., GER2) and Great Britain (i.e., UK1, UK2, and UK6) as well (Figure 3). The higher γ-decalactone content profile in these geographically distinct accessions suggests either a genetic flow from Iberian populations or a convergent evolution in the biosynthesis of this compound. Further studies might facilitate the identification of new variants of the fatty acid desaturase *FaFAD1* previously identified in *F.* × *ananassa* (Chambers et al., 2014; Sánchez-Sevilla et al., 2014; Zorrilla-Fontanesi et al., 2012) and/or novel genes involved in the biosynthesis of this compound.

Fruit morphological differences were also driven by geographical distribution, being fruits from southern accessions significantly bigger and rounder, while berries from northern origin were smaller and more elongated. In addition, a general decline in fruit size was detected in the second harvest (2017) compared with the first (2016). Strawberry fruit growth and ripening are coordinated and regulated by the combined action of the phytohormones auxin, GA, and ABA. First, in early fruit developmental stages, auxin promotes fruit growth in both length and width in a GA-dependent and independent pathways, respectively, being GA responsible for promoting fruit elongation only (Liao et al., 2018). Then, when the fruit has reached its final size, ABA inhibits fruit growth and triggers ripening (Fenn & Giovannoni, 2021; Liao et al., 2018). Interestingly, average temperatures during the harvesting period in 2017 (June and July) were colder by 1.7 and 1.8°C than the same period in 2016 (Finnish Meteorological Institute, n.d.). Since it is known that lower temperatures repress GA biosynthesis (Zhou et al., 2017), the colder summer in 2017 might have contributed to the shorter fruits reported during the 2017 harvest (H17). Furthermore, it is suspected that both endogenous and exogenous environmental signals trigger the shift from growth to ripening stages (Durán-Soria et al., 2020; Kou et al., 2021). Recently, some QTLs have been found to explain variations in fruit size and shape (Cockerton et al., 2021; Rey-Serra et al., 2021; Urrutia et al., 2015). However, the genetic and environmental effects on these traits need further studies to understand underlying regulatory mechanisms, being this collection of European *F. vesca* accessions an excellent resource for that purpose.

In conclusion, this work summarizes the biochemical and physical diversity of woodland strawberry fruit across Europe. The genetic flux rate and exposure to different environmental threats and conditions may have shaped strawberry fruit volatilome across the continent, which has been shown to have a strong genetic basis based on our data. We demonstrate that there are phenotypic differences between fruits associated with their region of origin, allowing us to differentiate specific European subpopulations and pinpoint specific metabolic signatures that are characteristic of different areas. These results indicate that metabolites have the potential to be used, under certain circumstances, as an alternative to molecular markers as proposed in previous studies (Fernandez et al., 2016; Riedelsheimer et al., 2012). However, their potential needs to be further explored. Furthermore, this work highlights the European *F. vesca* collection as a valuable tool for future genetic studies to identify candidate genes controlling the contents of VOCs of interest as well as fruit size and shape and develop markers linked to traits of interest. In addition, future studies could address the interaction between the characteristic VOCs of each specific geographical origin and their environment, helping to elucidate adaptations to specific pests, diseases, seed dispersal fauna or the activation of specific biochemical pathways by climatic conditions such as temperature or light periods.

EXPERIMENTAL PROCEDURES

Plant material and harvest

The collection of *Fragaria vesca* used in this study comprises 182 accessions representing European genetic diversity. These accessions were collected from 18 countries covering all the main distribution areas of the species in Europe from the south of the Iberian Peninsula to Kiolen Mountains, Northern Plain, and Lapland in the Scandinavian Peninsula, from Iceland and Great Britain islands to the Alps and the central plateau in continental Europe. Geographical coordinates of sampling sites spanned the latitude and longitude ranges of 37.2614° N-70.1848° N and 23.1833° W-104.3050° E, respectively (Table S8). All *F. vesca* accessions were cultivated at the Department of Agricultural Sciences, University of Helsinki, in Helsinki, Finland. During the autumn of 2015 and 2016, three clones of each genotype were propagated from runner cuttings in 10 × 10 cm square pots. Two-month-old plants were subjected to 12 h short days at 11°C for 6 weeks in a greenhouse

to induce flowering. Following the flower induction, plants were acclimated at 6°C for 6 weeks prior to cold storage at −2°C during winter. In the spring, plants were transplanted into 3-liter pots. Plants were divided into three blocks with one clone per accession in each block and cultivated in a cage with polycarbonate roof for berry production during the summer of 2016 and 2017. Fertilized peat (Kekkilä, Finland) was used as a growing media and plants were regularly fertilized using liquid fertilizer (Kekkilä, N-P-K: 17–4-25, Finland). For the volatilome analysis, three biological replicates of pools with at least five fully ripe berries each were harvested from 125 and 170 *F. vesca* accessions in the first (June–July 2016) and second season (June–July 2017), respectively, being 113 accessions common between both years. In addition, samples from var. 'Hawaii 4' (H4), 'Reine des Vallées' (RV), and 'Yellow Wonder' (YW) were collected in one harvest and treated equally. Fruits were immediately frozen in liquid nitrogen, and the receptacles, after the removal of achenes, ground to a fine powder and stored at −80°C until the analysis was performed.

Seeds of all the accessions studied in this work are available upon request to Dr. Timo Hytönen.

Volatile quantification

VOCs were determined in a similar way as described in previous studies (Zorrilla-Fontanesi et al., 2012). Briefly, aliquots of 500 mg of three independent biological replicates per accession and year were analyzed as independent samples. VOCs were sampled by HS-SPME (headspace solid-phase microextraction) with a 65 μm PDMS/DVB (polydimethylsiloxane/divinylbenzene) fiber (Supelco, PA, USA). A Combi-PAL autosampler (CTC Analytics, Zwingen, Switzerland) was used for incubation, VOC extraction, and desorption. GC–MS was performed in a 6890 N gas chromatograph coupled to a 5975B mass spectrometer (Agilent Technologies, CA, USA). Compounds were monitored over the mass/charge ratio (mz^{-1}) range of 35–250. Chromatograms and mass spectra were analyzed using the Enhanced ChemStation software (Agilent Technologies, CA, USA). VOCs were unambiguously identified by comparison of both retention time and mass spectra to those of commercial standards (Sigma-Aldrich, MO, USA) run under the same conditions. For quantification, a specific ion was selected for integration of the area of each of the identified compounds. Areas were normalized by comparison with the peak area of the same compound in the flanking reference samples. Reference sample consisted of a homogeneous mix of all analyzed samples and was injected regularly each of five to six samples in order to correct for variations in sensitivity and fiber aging.

Fruit size, shape, and volume

Berry sizes were scored right before harvesting using a slide gauge. The length of the berries was measured using their largest axis from sepals to the tip, while the width was measured at their widest part. Fruit shape was inferred as a ratio between length and width, being indexes >1 an indication of elongated fruits, while indexes ≤1 indicate rounded fruits. Fruit volume was approximated by considering woodland strawberry berries as conical and estimating their volume by the formula vol = πrh/3, where r is half fruit width and h is fruit length.

Data and statistical analyses

Volatile quantification ratios were normalized by log2 transformation. Statistical analysis was carried out using R software basic functions unless otherwise is specified (R Core Team, 2022). The chemotypes groups were defined based on the sum of terpenoids (total terpenoids) and lactones (total lactones). Only terpenoids and lactones showing between harvest correlation values >0.5 were considered. Before summing up the relative quantification values, they were rescaled using the formula: function(x) = x-min(x))/(max(x)-min(x)). Chemotypes were defined as high (above the mean) or low (below the mean) for total terpenoids and total lactones respectively. Additive ANOVA models with error type II were estimated for volatile traits taking genotype and harvest (year) factors and their effect sizes were expressed as omega squared values (ω^2) calculated from ANOVA table as follows: $\omega^2_x = (SS_x\text{-}df_x * MS_{error})/(SS_t + MS_{error})$. For fruit size (length width and volume) and shape (index) traits, and for VOCs taking geographical origin as a factor, an interaction ANOVA model with error type III correction was fitted. In both cases, we used *Anova* function from the package car (Fox & Weisberg, 2019).

Volatile least square means (LSmean) per genotype were estimated using *emmeans* function emmeans R from package emmeans (Lenth, 2020). Principal component analysis (PCA) was performed on LSmeans data and plotted using function PCA from FactoMine package (Lê et al., 2008). The bootstrapped hierarchical clustering analysis (HCA) was performed on least square mean data computing Euclidean distances and complete clustering method with 1000 iterations using function *Bclust* from shipunov R package (Shipunov, 2020). Pairwise differences for VOCs, fruit size, and fruit shape traits were estimated using *emmeans* and *CLD* functions from emmeans package (Lenth, 2020; Piepho, 2004). Sparse partial least squares discriminant analysis (sPLS-DA) was performed with functions from mixOmics package (Rohart et al., 2017). VOCs dataset (average values per genotype in two harvests) was randomly subset for training (80% of samples) and testing (20% of samples) sets. sPLS-DA model was estimated with *splsda* function. Optimal number of components estimated with *tune.splsda* function (validation = 'Mfold', folds = 3, nrepeat = 100) was set to 5 and optimal number of variables per component was set to 10, 9, 9, 8, and 9, respectively. Model performance and stability of selected variables were estimated with function *perf* (validation = 'Mfold', folds = 3, nrepeat = 100). Fitted model was applied on testing set using *predict* function.

ACKNOWLEDGMENTS

We thank Francisco Durán for the maintenance of the European woodland strawberry collection at the Andalusian Institute for Research and Training in Agriculture, Fishery, Food and Ecological Production (IFAPA) facilities in Campanillas, Málaga, Spain, and José Duarte for sample grinding support. Volatile profiles were obtained at the Metabolomics Lab at the IBMCP, Spain. This work was supported by the European Research Council (grant number ERC Starting Grant ERC-2014-StG 638134), the Junta de Andalucía (grant numbers PAIDI DOC_00278 (2020), POST-DOC_21_00893, PAIDI P20_00385, and PR.CRF.CRF202200.002. FEADER-CAPADS), the Spanish Ministry of Science and Innovation (grant numbers RTI2018-097309-A-I00, PID2021-123677OB-I00, and IJC2020-045612-I) and The University of Helsinki doctorate program. Open access charge was funded by the Universidad de Málaga and CBUA. [Corrections added on 30 August 2023, after first online publication: Names of funding projects in the Acknowledgements section have been updated in this version.]

CONFLICT OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

AUTHOR CONTRIBUTIONS

TH and DP conceived and designed the project. VM, JP, and JA conducted the fieldwork in Finland. VM, JP, and CMP processed samples for analysis. JLR and TG performed metabolomic quantification. JFS provided greenhouse facilities in Málaga (IFAPA). MU conducted all biostatistical and data analyses. MU and DP wrote the manuscript. All authors are involved in final manuscript editing.

DATA AVAILABILITY STATEMENT

Raw metabolomic data and experimental details have been uploaded to MetaboLights database and are accessible through www.ebi.ac.uk/metabolights/MTBLS7433.

Transformed and averaged metabolomic data are provided in the supplementary information files.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. Correlation of accessions geographical coordinates versus fruit size traits. (a). Latitude versus fruit length, width, index, and volume. (b). Longitude versus fruit length, width, index, and volume.

Table S1. Volatile quantification data. Normalized (log2) ratio genotypic average of VOCs detected by GC–MS in harvest 2016 and 2017.

Table S2. ANOVA results of VOCs fitting the model genotype (g) + harvest year (e) or the model geographical origin (Geo) x harvest year (e).

Table S3. Least square means (LSmeans) per genotype and accession chemotypes classification.

Table S4. Geographical origin of the European *F. vesca* germplasm collection indicating latitude and longitude coordinates, country, and grouping area.

Table S5. Variables selected in each of the five components of sPLS-DA predictive model and their stability tested by M-fold permutation in 100 repeats.

Table S6. Mean and SD values of agronomic traits (fruit length, width, index, and volume) per genotype and harvest year.

Table S7. ANOVA results for fruit size and shape values.

REFERENCES

Aharoni, A. (2004) Gain and loss of fruit flavor compounds produced by wild and cultivated strawberry species. *The Plant Cell Online*, **16**, 3110–3131. Available from: https://doi.org/10.1105/tpc.104.023895

Åkerström, A., Jaakola, L., Bång, U. & Jäderlund, A. (2010) Effects of latitude-related factors and geographical origin on Anthocyanidin concentrations in fruits of *Vaccinium myrtillus* L. (bilberries). *Journal of Agricultural and Food Chemistry*, **58**, 11939–11945. Available from: https://doi.org/10.1021/jf102407n

Akhatou, I., González-Domínguez, R. & Fernández-Recamales, Á. (2016) Investigation of the effect of genotype and agronomic conditions on metabolomic profiles of selected strawberry cultivars with different sensitivity to environmental stress. *Plant Physiology and Biochemistry*, **101**, 14–22. Available from: https://doi.org/10.1016/j.plaphy.2016.01.016

Akhatou, I., Sayago, A., González-Domínguez, R. & Fernández-Recamales, Á. (2017) Application of targeted metabolomics to investigate optimum growing conditions to enhance bioactive content of strawberry. *Journal of Agricultural and Food Chemistry*, **65**, 9559–9567. Available from: https://doi.org/10.1021/acs.jafc.7b03701

Araguez, I., Osorio, S., Hoffmann, T., Rambla, J.L., Medina-Escobar, N., Granell, A. *et al.* (2013) Eugenol production in achenes and receptacles of strawberry fruits is catalyzed by synthases exhibiting distinct kinetics. *Plant Physiology*, **163**, 946–958. Available from: https://doi.org/10.1104/pp.113.224352

Barbey, C.R., Hogshead, M.H., Harrison, B., Schwartz, A.E., Verma, S., Youngjae, O. *et al.* (2021) Genetic analysis of methyl anthranilate, Mesifurane, linalool, and other flavor compounds in cultivated strawberry (*Fragaria* × *ananassa*). *Frontiers in Plant Science*, **12**, 17.

Billet, K., Houillé, B., Dugé de Bernonville, T., Besseau, S., Oudin, A., Courdavault, V. *et al.* (2018) Field-based metabolomics of *Vitis vinifera L.* stems provides new insights for genotype discrimination and polyphenol metabolism structuring. *Frontiers in Plant Science*, **9**, 798. Available from: https://doi.org/10.3389/fpls.2018.00798/full

Chambers, A.H., Pillet, J., Plotto, A., Bai, J., Whitaker, V.M. & Folta, K.M. (2014) Identification of a strawberry flavor gene candidate using an integrated genetic-genomic-analytical chemistry approach. *BMC Genomics*, **15**, 1–15.

Cocco, C., Magnani, S., Maltoni, M.L., Quacquarelli, I., Cacchi, M., Antunes, L.E.C. *et al.* (2015) Effects of site and genotype on strawberry fruits quality traits and bioactive compounds. *Journal of Berry Research*, **5**, 145–155. Available from: https://doi.org/10.3233/JBR-150098

Cockerton, H.M., Karlström, A., Johnson, A.W., Li, B., Stavridou, E., Hopson, K.J. *et al.* (2021) Genomic informed breeding strategies for strawberry yield and fruit quality traits. *Frontiers in Plant Science*, **12**, 724847. Available from: https://doi.org/10.3389/fpls.2021.724847/full

Dong, J., Zhang, Y., Tang, X., Jin, W. & Han, Z. (2013) Differences in volatile ester composition between *Fragaria* × *ananassa* and *F. vesca* and implications for strawberry aroma patterns. *Scientia Horticulturae*, **150**, 47–53. Available from: https://doi.org/10.1016/j.scienta.2012.11.001

Du, X., Plotto, A., Baldwin, E. & Rouseff, R. (2011) Evaluation of volatiles from two subtropical strawberry cultivars using GC–Olfactometry, GC-MS odor activity values, and sensory analysis. *Journal of Agricultural and Food Chemistry*, **59**, 12569–12577. Available from: https://doi.org/10.1021/jf2030924

Durán-Soria, S., Pott, D.M., Osorio, S. & Vallarino, J.G. (2020) Sugar Signaling During Fruit Ripening. *Frontiers in Plant Science*, **11**, 564917. Available from: https://doi.org/10.3389/fpls.2020.564917/full

Fait, A., Hanhineva, K., Beleggia, R., Dai, N., Rogachev, I., Nikiforova, V.J. *et al.* (2008) Reconfiguration of the achene and receptacle metabolic networks during strawberry fruit development. *Plant Physiology*, **148**, 730–750. Available from: https://doi.org/10.1104/pp.108.120691

Fan, Z., Hasing, T., Johnson, T.S., Garner, D.M., Colquhoun, T.A., Sims, C.A. *et al.* (2021) Strawberry sweetness and consumer preference are enhanced by specific volatile compounds. *Horticulture Research*, **8**, 66.

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

Fernandez, O., Urrutia, M., Bernillon, S., Giauffret, C., Tardieu, F., Le Louis, J. *et al.* (2016) Fortune telling: metabolic markers of plant performance. *Metabolomics*, **12**, 158. Available from: https://doi.org/10.1007/s11306-016-1099-1

Fernandez, O., Urrutia, M., Berton, T., Bernillon, S., Deborde, C., Jacob, D. *et al.* (2019) Metabolomic characterization of sunflower leaf allows discriminating genotype groups or stress levels with a minimal set of metabolic markers. *Metabolomics*, **15**, 56.

Finnish Meteorological Institute. Temperature and precipitation statistics from 1961 onwards. ilmatieteenlaitos. Available from: https://en.ilmatieteenlaitos.fi/statistics-from-1961-onwards [Accessed 27th July 2023]

Folta, K.M. & Klee, H.J. (2016) Sensory sacrifices when we mass-produce mass produce. *Horticulture Research*, **3**, 16032. Available from: https://doi.org/10.1038/hortres.2016.32

Fox, J. & Weisberg, S. (2019) *An R companion to applied regression*, Third edition. Los Angeles: SAGE.

Gil-Ariza, D.J., Amaya, I., López-Aranda, J.M., Sánchez-Sevilla, J.F., Ángel Botella, M. & Valpuesta, V. (2009) Impact of plant breeding on the genetic diversity of cultivated strawberry as revealed by expressed sequence tag-derived simple sequence repeat markers. *Journal of the American Society for Horticultural Science*, **134**, 337–347. Available from: https://doi.org/10.21273/JASHS.134.3.337

Hilmarsson, H.S., **Hytönen, T.**, **Isobe, S.**, **Göransson, M.**, **Toivainen, T.** & **Hallsson, J.H.** (2017) Population genetic analysis of a global collection of *Fragaria vesca* using microsatellite markers D. D. Fang, ed. *PLoS One*, **12**, e0183384. Available from: https://doi.org/10.1371/journal.pone.0183384

Jetti, R.R., **Yang, E.**, **Kurnianta, A.**, **Finn, C.** & **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. *Journal of Food Science*, **72**, S487–S496. Available from: https://doi.org/10.1111/j.1750-3841.2007.00445.x

Josuttis, M., **Carlen, C.**, **Crespo, P.**, **Nestby, R.**, **Toldam-Andersen, T.B.**, **Dietrich, H.** ***et al.*** (2012) A comparison of bioactive compounds of strawberry fruit from Europe affected by genotype and latitude. *Journal of Berry Research*, **2**, 73–95.

Karban, R., **Wetzel, W.C.**, **Shiojiri, K.**, **Ishizaki, S.**, **Ramirez, S.R.** & **Blande, J.D.** (2014) Deciphering the language of plant communication: volatile chemotypes of sagebrush. *The New Phytologist*, **204**, 380–385. Available from: https://doi.org/10.1111/nph.12887

Kim, J., **Lee, J.G.**, **Hong, Y.** & **Lee, E.J.** (2019) Analysis of eight phytohormone concentrations, expression levels of ABA biosynthesis genes, and ripening-related transcription factors during fruit development in strawberry. *Journal of Plant Physiology*, **239**, 52–60. Available from: https://doi.org/10.1016/j.jplph.2019.05.013

Klee, H.J. & **Tieman, D.M.** (2018) The genetics of fruit flavour preferences. *Nature Reviews. Genetics*, **19**, 347–356. Available from: https://doi.org/10.1038/s41576-018-0002-5

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

Lamari, N., **Zhendre, V.**, **Urrutia, M.**, **Bernillon, S.**, **Maucourt, M.**, **Deborde, C.** ***et al.*** (2018) Metabotyping of 30 maize hybrids under early-sowing conditions reveals potential marker-metabolites for breeding. *Metabolomics*, **14**, 132. Available from: https://doi.org/10.1007/s11306-018-1427-8

Lê, S., **Josse, J.** & **Husson, F.** (2008) FactoMineR: an *R* package for multivariate analysis. *Journal of Statistical Software*, **25**, 1–18. Available from: https://doi.org/10.18637/jss.v025.i01

Lenth, R. (2020) Emmeans: estimated marginal means, aka least-squares means. package version 1.5.0, https://cran.r-project.org/web/packages/emmeans/index.html [Accessed 27th July 2023].

Lerceteau-Köhler, E., **Moing, A.**, **Guérin, G.**, **Renaud, C.**, **Petit, A.**, **Rothan, C.** ***et al.*** (2012) Genetic dissection of fruit quality traits in the octoploid cultivated strawberry highlights the role of homoeo-QTL in their control. *Theoretical and Applied Genetics*, **124**, 1059–1077. Available from: https://doi.org/10.1007/s00122-011-1769-3

Li, D., **Mou, W.**, **Xia, R.**, **Li, L.**, **Zawora, C.**, **Ying, T.** ***et al.*** (2019) Integrated analysis of high-throughput sequencing data shows abscisic acid-responsive genes and miRNAs in strawberry receptacle fruit ripening. *Horticulture Research*, **6**, 26. Available from: https://doi.org/10.1038/s41438-018-0100-8

Li, Z., **Wang, Z.**, **Wang, K.**, **Liu, Y.**, **Hong, Y.**, **Chen, C.** ***et al.*** (2020) Co-expression network analysis uncovers key candidate genes related to the regulation of volatile esters accumulation in woodland strawberry. *Planta*, **252**, 55. Available from: https://doi.org/10.1007/s00425-020-03462-7

Liao, X., **Li, M.**, **Liu, B.**, **Yan, M.**, **Yu, X.**, **Zi, H.** ***et al.*** (2018) Interlinked regulatory loops of ABA catabolism and biosynthesis coordinate fruit growth and ripening in woodland strawberry. *Proceedings. National Academy of Sciences. United States of America*, **115**, E11542–E11550. Available from: https://doi.org/10.1073/pnas.1812575115

Liston, A., **Cronn, R.** & **Ashman, T.** (2014) *Fragaria* : a genus with deep historical roots and ripe for evolutionary and ecological insights. *American Journal of Botany*, **101**, 1686–1699. Available from: https://doi.org/10.3732/ajb.1400140

Lunkenbein, S., **Bellido, M.**, **Aharoni, A.**, **Salentijn, E.M.J.**, **Kaldenhoff, R.**, **Coiner, H.A.** ***et al.*** (2006) Cinnamate metabolism in ripening fruit. Characterization of a UDP-glucose:cinnamate glucosyltransferase from strawberry. *Plant Physiology*, **140**, 1047–1058. Available from: https://doi.org/10.1104/pp.105.074955

Medina-Puche, L., **Molina-Hidalgo, F.J.**, **Boersma, M.**, **Schuurink, R.C.**, **López-Vidriero, I.**, **Solano, R.** ***et al.*** (2015) An R2R3-MYB transcription factor regulates eugenol production in ripe strawberry. *Fruit Receptacles*, **168**, 17.

Molina-Hidalgo, F.J., **Medina-Puche, L.**, **Cañete-Gómez, C.** ***et al.*** (2017) The fruit-specific transcription factor FaDOF2 regulates the production of eugenol in ripe fruit receptacles A. Vicente, ed. *Journal of Experimental Botany*, **68**, 4529–4543. Available from: https://doi.org/10.1093/jxb/erx257

Negri, A.S., **Allegra, D.**, **Simoni, L.**, **Rusconi, F.**, **Tonelli, C.**, **Espen, L.** ***et al.*** (2015) Comparative analysis of fruit aroma patterns in the domesticated wild strawberries "Profumata di Tortona" (*F. moschata*) and "Regina delle Valli" (*F. vesca*). *Frontiers in Plant Science*, **6**, 56. Available from: https://doi.org/10.3389/fpls.2015.00056/full

Ninkovic, V., **Markovic, D.** & **Dahlin, I.** (2016) Decoding neighbour volatiles in preparation for future competition and implications for tritrophic interactions. *Perspectives in Plant Ecology, Evolution and Systematics*, **23**, 11–17 Available from: https://doi.org/10.1016/j.ppees.2016.09.005

Ninkovic, V., **Markovic, D.** & **Rensing, M.** (2021) Plant volatiles as cues and signals in plant communication. *Plant, Cell & Environment*, **44**, 1030–1043. Available from: https://doi.org/10.1111/pce.13910

Nuzzi, M., **Lo Scalzo, R.**, **Testoni, A.** & **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 Analytical Methods*, **1**, 270–282. Available from: https://doi.org/10.1007/s12161-008-9039-y

Olbricht, K., **Ulrich, D.**, **Weiss, K.** & **Grafe, C.** (2011) Variation in the amounts of selected volatiles in a model population of *Fragaria* × *ananassa* Duch. As influenced by harvest year. *Journal of Agricultural and Food Chemistry*, **59**, 944–952. Available from: https://doi.org/10.1021/jf1034948

Perez-Fons, L., **Ovalle, T.M.**, **Maruthi, M.N.**, **Colvin, J.**, **Lopez-Lavalle, L.A.B.** & **Fraser, P.D.** (2020) The metabotyping of an east African cassava diversity panel: a core collection for developing biotic stress tolerance in cassava M. Rahimi, ed. *PLoS One*, **15**, e0242245. Available from: https://doi.org/10.1371/journal.pone.0242245

Perkins-Veazie, P. (1995) Growth and ripening of strawberry fruit. In: Janick, J. (Ed.) *Horticultural Reviews*. Oxford, UK: John Wiley & Sons, Inc., pp. 267–297. Available from: https://doi.org/10.1002/9780470650585.ch8

Pichersky, E., **Noel, J.P.** & **Dudareva, N.** (2006) Biosynthesis of plant volatiles: nature's diversity and ingenuity. *Science*, **311**, 808–811. Available from: https://doi.org/10.1126/science.1118510

Piepho, H.-P. (2004) An algorithm for a letter-based representation of all-pairwise comparisons. *Journal of Computational and Graphical Statistics*, **13**, 456–466. Available from: https://doi.org/10.1198/1061860043515

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 Biology*, **17**, 147. Available from: https://doi.org/10.1186/s12870-017-1088-1

Pyysalo, T., **Honkanen, E.** & **Hirvi, T.** (1979) Volatiles of wild strawberries, *Fragaria vesca* L., compared to those of cultivated berries, *Fragaria* × *ananassa* cv Senga Sengana. *Journal of Agricultural and Food Chemistry*, **27**, 19–22.

R Core Team. (2022) R: A language and environment for statistical computing. https://www.R-project.org/ [Accessed 27th July 2023].

Rey-Serra, P., **Mnejja, M.** & **Monfort, A.** (2021) Shape, firmness and fruit quality QTLs shared in two non-related strawberry populations. *Plant Science*, **311**, 111010. Available from: https://doi.org/10.1016/j.plantsci.2021.111010

Riedelsheimer, C. (2012) Genomic and metabolic prediction of complex heterotic traits in hybrid maize. *Nature Genetics*, **44**, 6.

Riedelsheimer, C., **Lisec, J.**, **Czedik-Eysenberg, A.**, **Sulpice, R.**, **Flis, A.**, **Grieder, C.** ***et al.*** (2012) Genome-wide association mapping of leaf metabolic profiles for dissecting complex traits in maize. *Proceedings. National Academy of Sciences. United States of America*, **109**, 8872–8877. Available from: https://doi.org/10.1073/pnas.1120813109

Rohart, F., **Gautier, B.**, **Singh, A.** & **Lê Cao, K.-A.** (2017) mixOmics: an R package for 'omics feature selection and multiple data integration. *PLoS Computational Biology*, **13**, 1–19. Available from: https://doi.org/10.1371/journal.pcbi.1005752

Sánchez-Sevilla, J.F., **Cruz-Rus, E.**, **Valpuesta, V.**, **Botella, M.A.** & **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. Available from: https://doi.org/10.1186/1471-2164-15-218

Schieberle, P. & **Hofmann, T.** (1997) Evaluation of the character impact odorants in fresh strawberry juice by quantitative measurements and sensory studies on model mixtures. *Journal of Agricultural and Food Chemistry*, **45**, 227–232. Available from: https://doi.org/10.1021/jf960366o

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 M. Kusano, ed. *PLoS One*, **9**, e88446. Available from: https://doi.org/10.1371/journal.pone.0088446

Shipunov, A. (2020) Visual statistics use R! Available from: https://cran.r-project.org/doc/contrib/Shipunov-visual_statistics.pdf [Accessed 27th July 2023].

Staudt, G. (2009) Strawberry biogeography, genetics and systematics. *Acta Horticulturae*, **842**, 71–84. Available from: https://doi.org/10.17660/ActaHortic.2009.842.1

Ulrich, D., **Hoberg, E.**, **Rapp, A.** & **Kecke, S.** (1997) Analysis of strawberry flavour–discrimination of aroma types by quantification of volatile compounds. *Zeitschrift für Lebensmitteluntersuchung und-Forschung A*, **205**, 218–223. Available from: https://doi.org/10.1007/s002170050154

Ulrich, D., **Kecke, S.** & **Olbricht, K.** (2018) What do we know about the chemistry of strawberry aroma? *Journal of Agricultural and Food Chemistry*, **66**, 3291–3301. Available from: https://doi.org/10.1021/acs.jafc.8b01115

Ulrich, D., **Komes, D.**, **Olbricht, K.** & **Hoberg, E.** (2007) Diversity of aroma patterns in wild and cultivated *Fragaria* accessions. *Genetic Resources and Crop Evolution*, **54**, 1185–1196. Available from: https://doi.org/10.1007/s10722-006-9009-4

Ulrich, D. & **Olbricht, K.** (2013) Diversity of volatile patterns in sixteen *Fragaria vesca* L. accessions in comparison to cultivars of *Fragaria* × *ananassa*. *Journal of Applied Botany and Food Quality*, **86**, 37–46. Available from: https://ojs.openagrar.de/index.php/JABFQ/article/view/2392

Ulrich, D. & **Olbricht, K.** (2014) Diversity of metabolite patterns and sensory characters in wild and cultivated strawberries. *Journal of Berry Research*, **4**, 11–17.

Ulrich, D. & **Olbricht, K.** (2016) A search for the ideal flavor of strawberry – Comparison of consumer acceptance and metabolite patterns in *Fragaria* × *ananassa* Duch. *Journal of Applied Botany and Food Quality*, **89**, 223–234. Available from: https://doi.org/10.5073/JABFQ.2016.089.029

Urrutia, M., **Bonet, J.**, **Arús, P.** & **Monfort, A.** (2015) A near-isogenic line (NIL) collection in diploid strawberry and its use in the genetic analysis of morphologic, phenotypic and nutritional characters. *Theoretical and Applied Genetics*, **128**, 1261–1275. Available from: https://doi.org/10.1007/s00122-015-2503-3

Urrutia, M., **Rambla, J.L.**, **Alexiou, K.G.**, **Granell, A.** & **Monfort, A.** (2017) Genetic analysis of the wild strawberry (*Fragaria vesca*) volatile composition. *Plant Physiology and Biochemistry*, **121**, 99–117.

Vallarino, J.G., **de Abreu e Lima, F.**, **Soria, C.**, **Tong, H.**, **Pott, D.M.**, **Willmitzer, L.** ***et al.*** (2018) Genetic diversity of strawberry germplasm using metabolomic biomarkers. *Scientific Reports*, **8**, 14386. Available from: https://doi.org/10.1038/s41598-018-32212-9

Venkatesh, T.V., **Chassy, A.W.**, **Fiehn, O.**, **Flint-Garcia, S.**, **Zeng, Q.**, **Skogerson, K.** ***et al.*** (2016) Metabolomic assessment of key maize resources: GC-MS and NMR profiling of grain from B73 hybrids of the nested association mapping (NAM) founders and of geographically diverse landraces. *Journal of Agricultural and Food Chemistry*, **64**(10), 2162–2172.

Wagner, G., **Charton, S.**, **Lariagon, C.**, **Laperche, A.**, **Lugan, R.**, **Hopkins, J.** ***et al.*** (2012) Metabotyping: a new approach to investigate rapeseed (Brassica napus L.) genetic diversity in the metabolic response to Clubroot infection. *Molecular Plant-Microbe Interactions*, **25**, 1478–1491. Available from: https://doi.org/10.1094/MPMI-02-12-0032-R

Wahyuni, Y., **Ballester, A.-R.**, **Tikunov, Y.**, **de Vos, R.C.H.**, **Pelgrom, K.T.B.**, **Maharijaya, A.** ***et al.*** (2013) Metabolomics and molecular marker analysis to explore pepper (*Capsicum* sp.) biodiversity. *Metabolomics*, **9**, 130–144. Available from: https://doi.org/10.1007/s11306-012-0432-6

Wang, J., **Yang, E.**, **Chaurand, P.** & **Raghavan, V.** (2021) Visualizing the distribution of strawberry plant metabolites at different maturity stages by MALDI-TOF imaging mass spectrometry. *Food Chemistry*, **345**, 128838. Available from: https://doi.org/10.1016/j.foodchem.2020.128838

Wang, S., **Li, W.**, **Liu, Y.**, **Li, H.**, **Ma, Y.** & **Zhang, Z.** (2017) Comparative transcriptome analysis of shortened fruit mutant in woodland strawberry (*Fragaria vesca*) using RNA-Seq. *Journal of Integrative Agriculture*, **16**, 828–844. Available from: https://doi.org/10.1016/S2095-3119(16)61448-X

Warschefsky, E., **Penmetsa, R.V.**, **Cook, D.R.** & **von Wettberg, E.J.B.** (2014) Back to the wilds: tapping evolutionary adaptations for resilient crops through systematic hybridization with crop wild relatives. *American Journal of Botany*, **101**, 1791–1800. Available from: https://doi.org/10.3732/ajb.1400116

Zhou, M., **Chen, H.**, **Wei, D.**, **Ma, H.** & **Lin, J.** (2017) Arabidopsis CBF3 and DELLAs positively regulate each other in response to low temperature. *Scientific Reports*, **7**, 39819. Available from: https://doi.org/10.1038/srep39819

Zorrilla-Fontanesi, Y., **Rambla, J.-L.**, **Cabeza, A.**, **Medina, J.J.**, **Sanchez-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 Physiology*, **159**, 851–870. Available from: https://doi.org/10.1104/pp.111.188318