JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

pubs.acs.org/JAFC Article

Chemical Differentiation of Sugarcane Cultivars Based on Volatile Profile and Chemometric Analysis

Pedro Silva, Jorge Freitas, Fernando M. Nunes, and José S. Câmara*

Cite This: J. Agric. Food Chem. 2021, 69, 3548−3558 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Sugarcane (SC) is a perennial grass widely cultivated in tropical and subtropical regions. However, its cultivation in Europe is residual, where Madeira Island, Portugal, is the only region where SC continues to be extensively cultivated. For the first time, the volatile profiles of regional cultivars were established by solid-phase microextraction combined with gas chromatography−mass spectrometry. Different volatile profiles for each cultivar were recognized, identifying 260 volatile organic compounds belonging to 15 chemical classes, such as aldehydes, alcohols, ketones, hydrocarbons, esters, and terpenes. Chemometric analysis procedure, namely, one-way ANOVA with Tukey's test, principal component analysis, partial least-square analysis, linear discriminant analysis, and hierarchical clustering analysis, allowed the differentiation between all regional cultivars. This study represents an important contribution for the maintenance of biodiversity and subsistence of the SC industry in Europe. Furthermore, it is also a valuable contribution to establish the typicality of traditional SC-based products, such as SC honey.

KEYWORDS: *sugarcane, cultivar, volatile organic compound, differentiation, chemometric analysis*

INTRODUCTION

Sugarcane (SC) is a large perennial grass widely cultivated in tropical and subtropical regions. It is a high value crop due to brief yield production cycle (10−14 months) and high content of sucrose, which is the primary raw material for sugar production and biofuel production. Furthermore, SC is a multipurpose crop commonly used as the main ingredient in several food products, such as syrup, molasses, and spirits, and also for animal feed and biomaterial manufacturing.[1−4] Taxonomically, SC is positioned into the family Poaceae (or Gramineae), under the genus *Saccharum*, being constituted by the inter-breeding of six species, namely, *S. officinarum*, *S. barberi*, *S. sinense*, *S. robustum*, *S. spontaneum*, and *S. edule*.[2] Although the modern cultivars contain some genetic information from these six species, the main part of its genetic pool arises from hybridization between *S. officinarum* (high sucrose content) and *S. spontaneum* (high abiotic stress resistance). Nevertheless, *S. officinarum*, known as "noble cane", is responsible for 80% of the genome due to the continuous intra-breeding to increase the sucrose content, leading to a breeding process described as "nobilization". The comprehension of the SC pathway since ancient times to modern cultivars is complex and challenging due to its polyploid nature, huge genome size, high hybrid variability, and multi-specific origin.[2−5]

Although SC was a historically important crop in Southern Europe, mainly in the Mediterranean area, and later, in the Atlantic islands, such as the Canaries and Madeira archipelagos, since its introduction by navigators from Portugal and Spain at the early 16th century in the Caribbean, Central, and South America, a sharp decline has occurred on its cultivation until today, culminating almost in its disappearance in Europe.[6,7] The current cultivation of SC in Europe is limited to small areas in southern Spain and Portugal, where the Madeira Island (Portugal) represents the last cluster of the once thriving European SC industry. There, SC continues to be cultivated extensively since its introduction in the early 14th century, where the main purpose was sugar production until the 17th century. From there until now, SC was used in the manufacturing of two main traditional food products: mel-de-cana, a black syrup, also called SC honey (SCH), and aguardente, a SC rum. Subsequently, both products are used as the main ingredient in traditional pastry/confectionery and regional brandies/spirits, respectively.[8,9]

The centenary cultivation of SC in volcanic soils under the Atlantic climatic conditions of Madeira Island can lead to specific regional cultivars. However, these cultivars can be closely related to the primary cultivars brought to the American continent due to relative agronomic isolation common in small and distant islands. In fact, the first SC plants cultivated in Brazil were brought from Madeira Island, later being dispersed throughout all South America.[4,5] So, the analysis of SC cultivars grown in Madeira Island can provide valuable information of the ancient European cultivars, as well about the early cultivars planted in the America, being an important biodiversity pool in order to understand the genetic and taxonomic connections of SC between the two continents. Besides the contribution to biodiversity maintenance, the analysis of regional cultivars also has economic importance.

Received: November 30, 2020
Revised: February 22, 2021
Accepted: February 24, 2021
Published: March 15, 2021

The exclusive use of regional SC cultivars can give rise to unique organoleptic and nutritional proprieties in the traditional products previously described, contributing to the appreciation and authentication not only of the products themselves but principally for the maintenance of one of the last SC industries in Europe.

One of the most applied approaches for differentiating cultivars in agro-food science is the molecular characterization based on the volatile profile. There, a wide range of volatile organic compounds (VOCs) are identified based on their unique mass fragment spectrum, which is expected to form a specific volatile profile for each cultivar highly influenced by the predecessor cultivars, soil proprieties, agronomic practices, and climatic conditions. Moreover, the volatile profile gives useful information about the sensory characteristics, such as aroma and flavor, allowing the evaluation of the quality and postharvest life.[8] Recently, the VOC profiling was widely recognized as a valuable tool for cultivar differentiation of grass crops, such as barley,[10] wheat,[11] and rice.[12]

For the first time, we aim to differentiate chemically six SC cultivars from Madeira Island, Portugal, empirically named by regional farmers as *amarela* (AMA), *radiada* (RAD), *roxa* (ROX), *verde* (VER), *violeta* (VIO), and *canica* (CAN), throughout the establishment of its volatile profile by solid-phase microextraction in headspace mode (HS-SPME) combined with gas chromatography−mass spectrometry (GC−MS) method and further chemometric analysis. The HS-SPME/GC−MS method was previously developed, optimized, and fully validated in our previous study.[8] The chemometric analysis was based on previously developed procedures in our previous studies[8,9] with some modifications. The establishment of the volatile profile of regional cultivars represents an important contribution for the maintenance of the genetic pool and biodiversity heritage. Furthermore, this study can be a valuable input to establish the typicality of traditional SC-based products, such as SCH, guaranteeing its quality, authenticity, safety control, and, consequently, a potential application to the European Union (EU) certification, namely, the Protected Designation of Origin (PDO) and Protected Geographical Indication (PGI), promoting the subsistence of sugar industry not only in Madeira Island, Portugal, but also in Europe.

MATERIALS AND METHODS

SC Samples. SC samples (N = 12) were harvested at Madeira Island, Portugal, where AMA, RAD, ROX, VER, and VIO cultivars were picked in Ribeira Brava, the southern part of Madeira Island, and CAN cultivar were picked in Porto da Cruz, Machico, the northern part of Madeira Island. The samples of all cultivars were harvested in triplicate (three replicates, N = 13) for 2 years, 2015 and 2017, and stored under stable conditions at 4 °C until aliquoting process. After the end of harvest session, the SC samples were mechanically squeezed by cold pressing to obtain SC juice. Then, the SC juice was transferred to 50 mL amber glass bottles in 25 mL aliquots and stored at −80 °C until analysis. The identification (ID) replicate number, ID replicate code, ID sample code, ID variety code, variety empirical name, harvest year, and geographical are presented in Table S1 (Supporting Information).

Chemicals. Sodium chloride was acquired from Panreac (Barcelona, Spain). The internal standard (IS), 4-heptanone, and the reference standards (RS) described in Table S2 were purchased from Sigma-Aldrich (St. Louis, Missouri, USA).

Materials and Software. SPME holder for the manual sampling of SPME and the fiber divinylbenzene/carboxen/polydimethylsiloxane (DVB/CAR/PDMS) with 50/30 μm film thickness were purchased from Supelco (Bellefonte, PA, USA). Ultrapure deionized water (H_2O) purified with a Milli-Q ultra-pure water system from Millipore (Massachusetts, USA). The MAXI MIX Vortex Mixer was acquired from Thermo Scientific (Massachusetts, USA). All data analysis and statistical processing were performed using the STATSOFT STATISTICA 12.0 (2013) software (Tulsa, USA).

Solid-Phase Microextraction Procedure. The SPME procedure used for VOC extraction from SC samples was based on our analytical method previously developed, optimized, and validated.[8] Briefly, 20 mL of sample was daily placed into a 50 mL polytetrafluoroethylene centrifuge tube and mechanical homogenized for 1 min and stored at 4 °C in aliquots of 5 mL. Then, these aliquots were transferred to an 8 mL glass vial with 60 mg of NaCl and 5 μL of IS (1 μL.L^{-1}) previously added, which was placed in a thermostatic bath at 30 °C for 5 min for sample temperature equilibrium. The SPME was performed in HS, where the fiber DVB/CAR/PDMS was attached to a manual SPME holder and exposed in the sample headspace for 60 min at 30 °C under magnetic agitation (600 rpm). The VOCs extracted by SPME were thermally desorbed by the fiber direct insertion into the GC injector at 250 °C in splitless mode for 10 min. The fiber was daily conditioned for 15 min at 250 °C into a GC injector to avoid contamination by unwanted interferents. Triplicate experiments were performed for all samples under analysis. Blank experiments were performed before the analysis of each sample, where the fiber was directly placed into GC injector without being subjected to any SPME extraction procedure.

GC−MS Analysis. The analysis of extracted VOCs was carried on an Agilent Technologies 6890N Network gas chromatography system (Santa Clara, California, USA) equipped with a BP-20 fused silica capillary column with 60 m × 0.25 mm I.D. × 0.25 μm film thickness (SGE, Dortmund, Germany) and interfaced with an Agilent 5975 quadrupole inert mass selective detector (Santa Clara, California, USA). The protocol employed for column oven temperatures was 40 °C for 2 min and then was increased at 0.25 °C min^{-1} until 45 °C with a 2 min hold, then was increased at 4 °C min^{-1} to 70 °C with a 2 min hold, was increased again at 3 °C min^{-1} to 130 °C with a 2 min hold, and finally was increased at 3 °C min^{-1} to 220 °C. This final temperature was maintained for 7 min for a total GC run time of 91.25 min. Column flow was constant at 1 mL min^{-1} using He carrier gas at a purity of 99.999% (helium N60, Air Liquide, Portugal). The injection port was operated in the splitless mode and held at 250 °C. For the 5975 MS system, the operating temperatures of the transfer line, quadrupole, and ionization source were 270, 150, and 230 °C, respectively. Electron impact mass spectra were recorded at 70 eV ionization voltages and the ionization current was 10 μA. The electron multiplier was set to the autotune procedure. The acquisitions were performed in the scan mode (30−300 m/z). VOC identification was based on visual interpretation of spectra through the Agilent MS ChemStation Software and confirmed comparing each VOC mass spectra with the NIST14 Mass Spectral Library (2014), which is successfully identified when the similarity threshold was higher than 75%. In addition, some VOCs were also identified using a RS, where each VOC was individually analyzed by GC−MS. The total peak area values were obtained by target ion quantitation protocol. The VOCs were semiquantitate by dividing the total peak area value of each VOC by the total peak area value of the IS and expressed as relative peak areas (RPA).

Chemometric Analysis. The chemometric analysis was based on previously developed procedures in our previous studies,[8,9] with some modifications according to recommendations for analytical applications.[13,14] One-way ANOVA was performed to determine the VOCs with statistically significant differences (p < 0.05) based on its variance level between RPA values of all SC samples. Additionally, the Tukey post-hoc test was also performed to confirm the VOCs with statistically significant differences (p < 0.05) between RPA values of SC samples from two pairs of cultivars. Principal component analysis (PCA) and partial least squares (PLS) analysis were applied on RPA values of VOCs dataset in order to obtain the preview of differentiation/correlation structure based on the variance of samples from all cultivars during 2015 and 2017, without classification and

https://dx.doi.org/10.1021/acs.jafc.0c07554
J. Agric. Food Chem. 2021, 69, 3548−3558

with classification according to the cultivar, respectively. A V-fold with a V-value fixed in 7 was used for cross-validation. Linear discriminant analysis (LDA) was applied as a supervised pattern recognition method for variable selection in order to achieve a matrix composed by the lower number of VOCs that allows the correct classification of all samples under analysis for the group assignment (AMA, RAD, ROX, VER, VIO, and CAN). A backward selection method (p value of 0.05 to enter and 0.05 to remove) was used to determine the most predictive VOCs and remove the least predictive from analysis, where a classification structure is obtained based on canonical discriminant functions (CDFs). A V-fold with a leave-one-out strategy was applied for cross-validation. LDA is a highly recommended method to reduce the dimensionality of higher dimension matrix into a lower dimension matrix by removal of less predictive variables preserving the interclass separation (classification structure).[13,14] A matrix reduction procedure was applied prior to LDA, where the matrix was reduced to 20% of initial dimension based on F-value from one-way ANOVA between all samples. Alternatively, a matrix reduction procedure based on the variable importance in projection (VIP) scores from the PLS analysis was also applied. Thus, the 52 VOCs with higher F-value and VIP score were selected, respectively. Finally, a second PLS analysis and hierarchical clustering analysis (HCA) were applied on RPA values of the most predictive VOCs. PLS based on dataset from the reduced matrix was performed to verify the differentiation/correlation structure between all samples when classified according to SC cultivar and compare with structure obtained from PLS performed with the complete matrix. HCA was performed in order to determine the Euclidean linkage distances between all samples and complete an appropriate and visual measure of distance and linkage criterion between SC cultivars based only in the most predictive VOCs. All data analysis and statistical processing were performed using the STATSOFT STATISTICA 12.0 (2013) software (Tulsa, USA).

RESULTS AND DISCUSSION

Establishment of Volatile Profiles from SC Cultivars. Through the application of HS-SPME/GC−MS method to all SC samples under analysis in a 2-year study (2015 and 2017), it was possible to establish, for the first time, the volatile profile of six different SC cultivars from Madeira Island, Portugal. The ID number, retention time, target ion, match percent, IUPAC name, NIST database name, abbreviation, CAS number, molecular formula, and main chemical class of VOCs identified through analysis of SC samples during 2015 and 2017 harvest years are described in Table S2. Also, the mean RPA values and the respective relative standard deviation (RSD) of each VOC identified in AMA, RAD, and ROX and in VER, VIO and CAN cultivars are summarized in Table S3A,B, respectively. The typical GC−MS chromatograms for each SC cultivar are presented in Figure S1.

Number of VOCs. The analysis of all samples from six regional SC cultivars allowed the identification of 260 different VOCs. Although there may be a wide disparity in RPA values, all VOCs were recognized in all samples under analysis. This may be due to the procedure used for VOC identification, where all 260 VOCs were searched in all samples based on a stage-by-stage follow-up strategy. The mean, minimum, and maximum values of RPA for all VOCs identified in each SC cultivar are summarized in Table S4.

As expected, 88 (33.8%) of the 260 VOCs were also previously identified in other SC-based products, where 56 (21.5%) in juice, 34 (13.1%) in sugar, 16 (6.2%) in treacle, 20 (7.7%) in molasses, 45 (17.3%) in syrup, 42 (16.2%) in rum or spirits, 3 (1.2%) in infusion, and 4 (1.4%) in alcoholic fermented beverages. Also, 36 (14.2%) VOCs were previously recognized in SCH samples.[8] Nevertheless, 173 (66.5%) VOCs were identified for the first time in food products from SC. This information and respective references are summarized in Table S5.

Interestingly, some of the 260 VOCs were previously identified in raw juice and thermal-processed products (*i.e.*, sugar, treacle, molasses, and syrup), such as dimethyl sulfide (DMSULFI), 3-hydroxy-2-butanone (HXY3BT2ONE), 1-hydroxy-2-propanone (HXY1PP2ONE), dimethyl sulfoxide (DMSULFO), and furfuryl alcohol (FURFOL), while other VOCs were recognized in raw juice and fermented products (*i.e.*, rum, spirit, and wine-like beverage), namely, 2-heptanol (HPT2OL), 1-hexanol (HX1OL), 2-octanol (OCT2OL), 1-octen-3-ol (OCT3E1OL), and β-ionone (BIONNE). Moreover, some VOCs were recognized in most of the SC-based products, being identified in raw juice, thermal processed and fermented products, such as 3-methyl-butanal (M3BTAL), ethyl alcohol (ETOL), 2,3-butanedione (BT23DONE), hexanal (HXAL), 2-methyl-1-propanol (M2PP1OL), 3-methyl-1-butanol (M3BT1OL), nonanal (NONAL), furfural (FURAL), ethanoic acid (ETNOIC), mequinol (MEQNOL), and benzeneethanol (BENZETOL) and 2-methoxy-4-vinylphenol (MXY2VYL4PHEOL). These findings are indicative that some VOCs from raw SC are still present in thermal processed and/or fermented products, demonstrating that the volatile analysis of specific cultivars can be a useful tool for the determination of the typicality and authenticity of SC products based on these same cultivars. On the other hand, some VOCs are apparently sensitive to thermal or fermentation processing, being only identified in raw juice, namely, 2-propanone (PP2ONE), 1-penten-3-one (PT1E3ONE), 2-pentanol (PT2OL), 2-heptanone (HPT2ONE), *trans*-2-pentenol (TPT1E2OL), *cis*-2-pentenol (CPT1E2OL), 2-pentene (PT2ENE), *cis*-3-hexen-1-ol (CHX1E3OL), *trans*-2-octenal (TOCT2EAL), *trans*,*trans*-2,4-heptadienal (TTHPT22DEAL), 2-nonanol (NON2OL), β-myrcene (BMYRCNE), *trans*-2-octen-1-ol (TOCT1E2OL), α-terpineol (ATERPINOL), and nonanoic acid (NONOIC).

Main VOCs. Although each one of the 260 VOCs contributed for the establishment of the volatile profile of all SC cultivars, its RPA values had a wide range, ranging between 0.03 and 3000. Thus, the 20 VOCs with higher RPA values for each SC cultivar were selected in order to determine the main contributors to volatile profile. The RPA values of the 20 main VOCs for AMA, RAD, ROX, VER, VIO, and CAN cultivars during 2015 and 2017 are described in Figure 1A−F, respectively.

Based on the analysis of the 20 main contributors to the volatile profile for each SC cultivar, 43 different VOCs were selected. From these, only seven were common to all cultivars, namely, DMSULFI, HX1OL, 1,2-cyclopentanedione (CPT12DONE), HPT2OL, ETOL, PT2ENE, and CPT1E2OL. On the contrary, 18 VOCs were selected as main contributors only in a specific SC cultivar, such as cyclodecane (CDECANE) for AMA; ethyl acetate (EESTAA), 2-ethyl methyl ester pentanoic acid (E2MESTPTA), benzenemethanol (BENZMTOL), styrene (STYNE), and TOCT1E2OL for ROX; 1,3-dihydroxy-2-propanone (DHYPAONE), 2,2-dimethyl-3-heptanone (DM22HPT3ONE), MEQNOL, 1-hydroxy-2-propanone (HXY1PP2ONE), and 4-methyl-2-heptanone (M4HPT2ONE) for VER; and 2,4-dimethyl-1-heptene (DM24HPT1ENE), β-phellandrene (BPHELDNE), *tert*-butyl ethyl ether (BEETHR), 4-methylheptane (M4HPTANE), α-phellandrene (APHELDNE), heptanoic acid (HEPTOIC), and *trans*-*p*-2-menthen-1-ol

https://dx.doi.org/10.1021/acs.jafc.0c07554
J. Agric. Food Chem. 2021, 69, 3548−3558

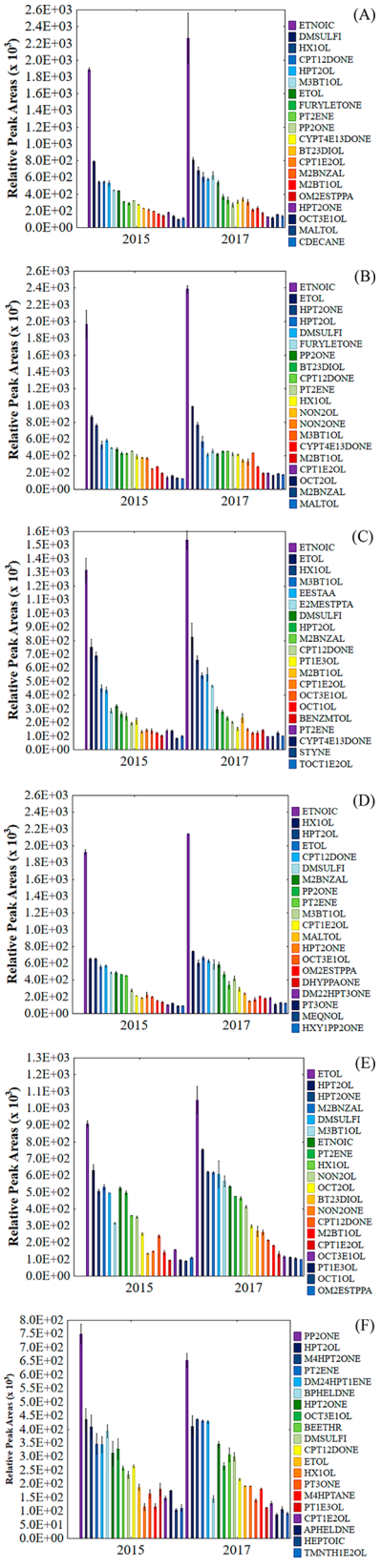

Figure 1. The RPA values of the 10 main VOCs for *amarela* (A), *radiada* (B), *roxa* (C), *verde* (D), *violeta* (E), and *canica* (F) SC cultivars during 2015 and 2017 harvest years.

(TMNTH1E2OL) for CAN. None of the main VOCs were exclusively found for RAD and VIO cultivars. Remarkably, three VOCs, ETNOIC, 3-methyl-1-butanol (M3BT1OL), and 2-methyl-benzaldehyde (M2BNZAL), were classified as the main contributor only in cultivars picked in Ribeira Brava, the southern part of Madeira Island, having potential as markers for this geographic area, which is characterized by higher temperatures and less precipitation than the northern part of the island. In fact, ETNOIC was the main contributor to the volatile profile of AMA, RAD, ROX, and VER cultivars, presenting RPA values considerably higher than the other main VOCs. Possibly, the southern climatic conditions are more favorable to the formation of ETNOIC through natural action of bacteria on the sugars and alcohols.[15]

Chemical Class Classification of VOCs. A wide diversity of chemical classes was recognized through the VOC profiling of regional SC cultivars, being fundamental its classification into different groups to characterize the volatile profile of each SC cultivar. The classification group based on the main chemical class of each VOC is described in Table S2. The number of VOCs, RPA, and TRPA values of main chemical classes identified in SC samples are summarized in Table S6. The contribution (RPA and TRPA values) of each classification group assigned according to the chemical class is presented in Figure 2A,B, respectively.

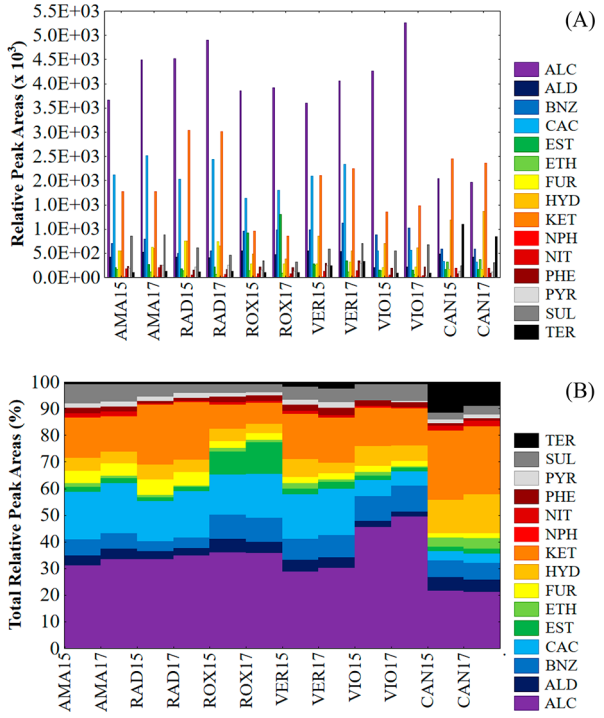

Figure 2. The contribution, RPA and total RPA (%) values, of each classification group assigned according to the chemical class for all SC cultivars during 2015 and 2017 harvest years.

Throughout the classification of all 260 VOCs, 15 different chemical classes were recognized, namely, alcohol (ALC), aldehyde (ALD), benzene (BNZ), carboxylic acid (CAC), ester (EST), ether (ETH), furan (FUR), hydrocarbon (HYD), ketone (KET), naphthalene (NPH), nitrogen (NIT), phenol (PHE), pyran (PYR), sulfur (SUL), and terpene/terpenoid (TER).

Surprisingly, VOCs from all these chemical classes were found in cultivars from other grass crops, such as barley,[10] wheat,[11] rice,[12] oat,[16] and rye.[16] Some of these chemical classes are commonly linked to plants, and therefore, their VOCs are

https://dx.doi.org/10.1021/acs.jafc.0c07554
J. Agric. Food Chem. 2021, 69, 3548−3558

expected to be found in grass cultivars. BNZ and TER derivatives play a key role in plant biochemistry, such as plant−environment interactions and abiotic stress response, being derived from well-known benzenoid and terpenoid pathways.[17] Although its synthesis is not well established, being probably derivatives from the oxidation of fatty acids, VOCs from ALC, ALD, CAC, EST, HYD, and KET were also commonly found in various plant tissues and contribute in diverse physiological processes.[18−20] On the other hand, for some classes, it is difficult to explain their origin and purpose in grass crops. FUR, PYR, NIT, NPH, and PHE were normally related to thermal degradation (*i.e.*, Maillard reactions) of plant components (*i.e.*, sugars, amino acids, and fatty acids)[21,22] and biomass burning.[23] Also, ALD, KET, and HYD can be related to these two processes. A potential explanation is the common use of fires to clean the land from weeds and waste in grass crops, which generated VOCs from pyrolysis of plant components (*i.e.*, sugars, amino acids, cellulose, lignin, starch, among others). Another explanation may be the occurrence of thermal degradation *in vivo* of SC components, such as sugars (high content) and amino acids, due to the high temperatures which the SC stalks are subjected to during the days with more sunlight incidence. For example, the occurrence of Maillard reactions *in vivo* during the ripening of fruits was previously detected.[24] In addition, VOCs from ALC, CAC, EST, ETH, and SUL classes could also be formed by the microbial activity of yeast and bacteria in plant tissues[20] or plant-based foods.[25]

Based on the results described in Figure 2, the contribution of all chemical classes to volatile profile differs widely from cultivar to cultivar, which obtained an exclusive volatile profile in terms of class contribution for each cultivar. Once again, the higher contribution differences were verified between the CAN cultivar and the southern cultivars. For example, the ALC class has the highest contribution for AMA, RAD, ROX, VER, and VIO cultivars while the KET class for CAN. Moreover, the contribution of TER, HYD, and ETH classes was considerably higher in CAN cultivar, while the CAC contribution was superior in the southern cultivars. However, some differences were also observed in the southern cultivars. For example, the ROX cultivar presented expressively higher TRPA value for EST class while showed the lower contribution of KET class. Likewise, the VIO cultivar presented TRPA values for PYR class almost 10 times lower than the other cultivars. There, results obtained suggest an apparent differentiation in the chemical class of volatile profiles among all regional SC cultivars.

Chemometric Analysis of Volatile Profiles from SC Cultivars. Chemometric analysis procedure, namely, one-way ANOVA with Tukey's test, PCA, PLS analysis, LDA, and HCA, was applied in order to achieve the differentiation between all regional SC cultivars. Furthermore, similar chemometric analysis procedures have been successfully applied for differentiation of cultivars based on volatile profile.[26−28]

One-Way ANOVA with Post-hoc Tukey's Test. One-way ANOVA was performed to assess the existence of statistically significant differences between all SC cultivars for each VOC based on its RPA value variance level, while the post-hoc Tukey's test was done to determine the statistically significant differences between RPA values of SC samples, comparing pairs of cultivars (*i.e.*, all combinations between one specific cultivar and each one of the other cultivars). The one-way ANOVA with post-hoc Tukey's test results, namely, probability (P) and Fischer (F) values, between all cultivars and for each pair of SC cultivars are summarized in Table S7.

Of the all 260 VOCs, only 3,5,5-trimethyl-2-cyclopenten-1-one showed no statistically significant difference ($P > 0.05$) in RPA values among all SC cultivars. Among the 259 VOCs with statistically significant differences ($P < 0.05$), 133 VOCs (51.2%) presented high differences (F values ≥ 100) between all cultivars, where eight VOCs (3.1%) demonstrated very high significant differences (F values ≥ 1000), namely, 4,4-dimethyl-2-pentanone (DM44PT2ONE), M4HPT2ONE, DM23HX2OL, 3-ethyl-2-methyl-1,3-hexadiene (E3M2HX13DENE), menthol (MNTHOL), benzeneacetaldehyde (BENZACETAL), 2-undecanol (UNDEC2OL), and 2-acetylpyrrole (ACTLPYROLE).

According to the results from post-hoc Tukey's test described in Table S7, a substantial dissimilarity was found in volatile profiles between all SC cultivars. Again, the CAN cultivar showed the higher dissimilarity level, the cultivar with the largest number of VOCs with significant differences in the RPA values for all pairs formed with each one of the other cultivars (CAN vs RAD, CAN vs ROX, CAN vs VER, and CAN vs VIO). That is, the CAN cultivar presented 115 (44.2%) VOCs with statistically significant differences for AMA, RAD, ROX, VER, and VIO cultivars simultaneously. Also, the ROX cultivar presented some dissimilarity from the remaining cultivars, in which 83 (31.9%) VOCs had significant differences. In minor extension, both AMA and VER cultivars presented 59 (22.7%) VOCs, followed by VIO cultivar with 51 (19.6%) VOCs and RAD cultivar with 49 (18.8%) VOCs. Unexpectedly, only TTHPT22DEAL demonstrated statistically significant differences for all combinations of pairs between the six cultivars. Contrariwise, some VOCs only presented significant differences for all combinations completed with one specific cultivar. For example, 15 VOCs only showed significant differences between the CAN cultivar and each one of the southern cultivars, namely, hexane (HXANE), BEETHR, M4HPTANE, 2-methyl-2-heptene (M2HPTENE), DM24HPT1ENE, 2,3,4-trimethyl-hexane (TM234HXANE), 2,6-dimethyl-nonane (DM26NNANE), 4-methyl-2-pentanone (M4PT2ONE), 2-methy-2-butanol (M2BT2OL), APHELDNE, BPHELDNE, 4,6-dimethyl-2-heptanone (DM46HPT2ONE), 1,3-bis(1,1-dimethylethyl)-benzene (BI13DME11BNZ), 5-methyl-2-heptanone (M5HPTAONE), and 1,5-cyclooctanedione (CYOCTD15ONE). For ROX cultivar, nine VOCs were specific for these cultivars, specifically ethyl propanoate (EESTPA), ethyl 2-methylpropanoate (M2EESTPA), propyl acetate (PESTAA), 2-methylpropyl acetate (M2PESTAA), 2-methyl-3-buten-2-ol (M3BT2OL), methyl 3,3-dimethylbutanoate (DM33MESTBA), 2-hexanone (HX2ONE), PT2OL, and E2MESTPTA. Interestingly, most of the compounds specific to the CAN cultivar belong to the chemical class groups of KET, HYD, and TER, while for the ROX cultivar, they belong mostly to the EST group. For remaining cultivars, 2-methyl-2-cyclopenten-1-one (M2CY2PT1EONE) and 2-propenoic acid (PPE2NOIC) were specific for AMA cultivar, while for VER cultivar, ethylene glycol ethyl ether (ENGOLEETHR) and *cis-p*-2-menthen-1-ol (CMNTH1E2OL) were specific, and only 3-octen-2-one (OCT2E3ONE) was specific for RAD cultivar. None of the VOCs have specificity for VIO cultivar combinations.

PCA and PLSs. PCA and PLS analyses were performed to preview the differentiation/correlation structure based on the variance of samples from all SC cultivars during 2015 and

https://dx.doi.org/10.1021/acs.jafc.0c07554
J. Agric. Food Chem. 2021, 69, 3548−3558

2017, without classification and with classification accordingly, respectively. The PCA and PLS information is summarized in Table S8. The loading results and VIP scores of PCA and PLS analysis for each variable (VOCs) are described in Table S9, while the loading results of all cultivar samples (PCA) and six cultivar centroids (PLS) are described in Table S10. The PCA and PLS score results of all samples under analysis are described in Table S11. The PCA loading line plots of all cultivar samples for PC1, PC2, and PC3 are shown in Figure S2A−C, respectively. The PLS loading line plots of six cultivar centroids for PLS1, PLS2, and PLS3 are shown in Figure S2DF, respectively. The PCA loading 3D plots of all cultivar samples and all variables for PC1, PC2, and PC3 are shown in Figure 3A,B, respectively. The PLS loading 3D plots of the six cultivar centroids and all variables for PLS1, PLS2, and PLS3 are shown in Figure 4A,B, respectively.

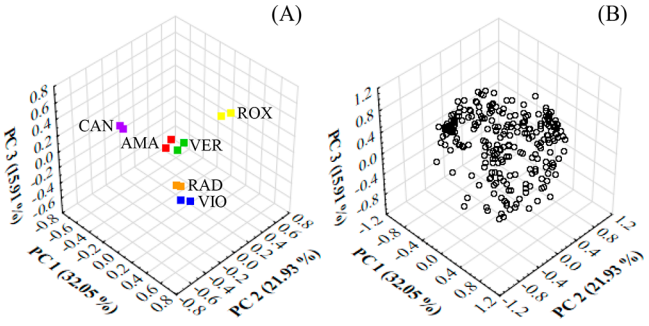

Figure 3. The PCA loading 3D plot of all cultivar samples (A) and the selected 259 VOCs (B) for the three main components.

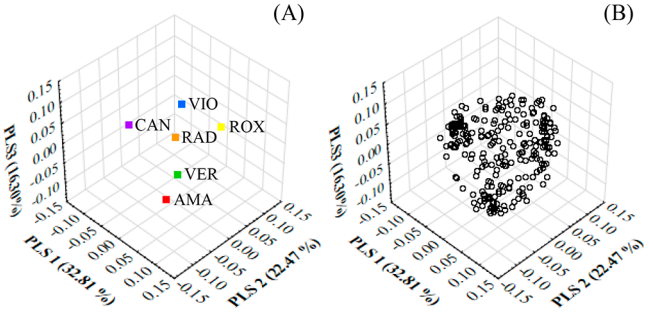

Figure 4. The PLS loading 3D plot of all cultivar centroids (A) and the selected 259 VOCs (B) for the three main components.

The PCA analysis based on PC1, PC2, and PC3 explained 69.9% of the total variance (TVA), where the sum of all the 14 components explained 99.5%. The projection of structure based on PC1, PC2, and PC3 loading results demonstrated the formation of precise sample groups (2015 and 2017) according to SC cultivar, namely, AMA, RAD, ROX, VER, VIO, and CAN groups. There, a visual differentiation between all cultivar groups was observed, where CAN and ROX cultivars are clearly separated from the other cultivars. In PC1 projection (32.0% TVA), CAN cultivar demonstrated a high variance from the southern cultivars, while ROX cultivar showed a slight variance for the other southern cultivars. For PC2 projection (21.9%), the ROX cultivar presented a high variance from all the other regional cultivars. Also, in PC2 projection, AMA cultivar exhibited a substantial variance from the RAD, VER, VIO, and CAN cultivars. Finally, in PC3 projection (15.9%), a considerable variance was observed between all cultivars, principally for RAD and VIO cultivars.

A PLS analysis was performed to evaluate the variance between samples when classified according to SC cultivar, being classified as centroid-AMA (C-AMA), centroid-RAD (C-RAD), centroid-ROX (C-ROX), centroid-VER (C-VER), centroid-VIO (C-VIO), and centroid-CAN (C-CAN). The PLS analysis based on PLS1, PLS2, and PLS3 explained 71.6% of TVA, where the sum of all the 18 components explained 99.7%. All cultivar centroids were clearly apart from each other, while the results of PLS analysis were very similar to those obtained in the PCA in terms of variance due to the reduced intra-variance of samples from the same cultivar. This low intra-variance for each cultivar can be verified through analysis of PCA and PLS scores of all samples described in Table S11. Thus, in PLS1 projection (32.8%), the higher variance was observed between CAN cultivar and the southern cultivars, while in PLS2 projection (22.5%), the higher variance was observed for the ROX cultivar. For PLS3 projection (16.3%), all cultivar centroids were well separated, where the higher variance was observed between VIO and RAD cultivars from the other cultivars. Both PCA and PLS projections indicated that the volatile profile can be an effective strategy to differentiate molecularly the six SC regional cultivars.

Although the PLS projection, as well the PCA projection, was based on all 259 VOCs, each one influenced differently the projection of centroids, being possible to recognize the contribution of each VOC for the projection of a specific cultivar centroid. The 3D plot constructed with three main component loading results for all VOCs under analysis showed that a high number of VOCs influenced the projections of CAN and ROX cultivars. Interestingly, the CAN projection was highly influenced by VOCs belonging to KET, HYD, and TER chemical classes, namely, by VOCs previously mentioned in the ANOVA test, such as M4PT2ONE, DM46HPT2ONE, M2HPTENE, DM24HPT1ENE, APHELDNE, and BPHELDNE. Likewise, ROX projection was very influenced by VOCs also referred in the ANOVA test, mainly belonging to EST chemical class, such as EESTPA, M2EESTPA, PESTAA, M2PESTAA, and E2MESTPTA. Also, VIO projection was influenced by some VOCs, namely, HPT2ONE, 5-methyl-4-hexen-3-one (M5HX3E4ONE), OCT2OL, NON2OL, and *cis*-5-decenol (CDEC5E1OL). Although the projections of the remaining regional cultivars were directly influenced by a smaller number of VOCs, some can be highlighted by their high influence on the projection of a specific cultivar, namely, 2,3-pentanedione (PTDONE) and MALTOL for AMA cultivar, 2-methyl-furan (M2FUR) for RAD cultivar, and 1-decanol (DEC1OL) for VER cultivar. Thus, the specificity of these VOCs for a single cultivar could represent a potential marker of this specific cultivar.

Linear Discriminant Analysis. LDA was performed as a supervised pattern recognition method to achieve classification rules for the cultivar assignment of all samples under analysis based only in the most predictive VOCs. The matrix reduction procedure based on the VIP scores presented poor results to PLS and HCA analysis (shown in Figure S5) and was therefore disregarded. The LDA information of the selected 52 VOCs based on the VIP scores are summarized in Table S14. The LDA information and CDF coefficients of the selected 52 VOCs based on *F*-value from one-way ANOVA test are summarized in Table 1. The CDF coefficients and highest probability classification results of all samples are summarized

https://dx.doi.org/10.1021/acs.jafc.0c07554
J. Agric. Food Chem. 2021, 69, 3548−3558

Table 1. Results of Variables from LDA after Matrix Reduction Method to 20% of Original Dimension According with Higher Fisher Values Obtained from One-Way ANOVA Test Based on the Relative Peak Areas of the Identified VOCs in SC Samples

			ANOVA			LDA		
						CDF[e]		
VOCs	abbreviations	ID type[a]	F[b]	W[c]	F[d]	1	2	3
menthol	MNTHOL	RS; MS	1781.56	1.13×10^{-4}	1.42×10^{4}	−0.0022	−0.0015	0.0015
4,4-dimethyl-2-pentanone	DM44PT2ONE	MS	1684.33	2.02×10^{-11}	1.12×10^{11}	−0.0025	−0.0017	0.0018
2-acetylpyrrole	ACTLPYROLE	MS	1508.90	removed from analysis				
2,3-dimethyl-2-hexanol	DM23HX2OL	MS	1243.14	2.24×10^{-11}	1.01×10^{11}	−0.0021	−0.0015	0.0017
2-undecanol	UNDEC2OL	MS	1153.59	removed from analysis				
3-ethyl-2-methyl-1,3-hexadiene	E3M2HX13DENE	MS	1111.23	4.51×10^{-12}	4.98×10^{11}	0.0009	−0.0032	−0.0021
4-methyl-2-heptanone	M4HPT2ONE	MS	1066.87	6.50×10^{-4}	2.46×10^{3}	−0.0019	−0.0014	0.0013
benzeneacetaldehyde	BENZACETAL	MS	1009.48	9.62×10^{-12}	2.34×10^{11}	−0.0002	0.0001	0.0052
4-amino-phenol	AMIPHEOL	MS	886.23	removed from analysis				
2-ethylhexyl octanoate	EHESTOA	MS	873.20	5.01×10^{-12}	4.49×10^{11}	0.0024	0.0011	−0.0059
cis-5-decenol	CDEC5E1OL	MS	706.67	removed from analysis				
2-nonanol	NON2OL	MS	695.37	2.43×10^{-12}	9.24×10^{11}	0.0012	−0.0014	−0.0021
2,4,5-trihydroxypyrimidine	THDXYPYMNE	MS	637.12	removed from analysis				
cis-piperitol	CPIPETOL	MS	596.97	removed from analysis				
2-phenoxy-ethanol	PHENXYOL	MS	592.27	removed from analysis				
propanoic acid	PPANOIC	MS	582.46	6.87×10^{-12}	3.28×10^{11}	−0.0009	0.0012	−0.0025
undecanoic acid	UNDECOIC	MS	561.56	removed from analysis				
1-(2-furanyl)-ethanone	FURYLETONE	RS; MS	560.99	4.38×10^{-12}	5.13×10^{11}	0.0011	0.0014	0.0016
1-methyl-4-piperidinone	M1PIP4DIONE	MS	558.99	removed from analysis				
3-methyl-1,2-cyclopentanedione	M3CPT12DONE	MS	544.81	removed from analysis				
4-methylimidazole	M4IMDZOLE	MS	543.79	removed from analysis				
nonanoic acid	NONOIC	RS; MS	534.87	removed from analysis				
2-methoxy-4-vinyl-phenol	MXY2VYL4PHEOL	MS	531.37	removed from analysis				
β-myrcene	BMYRCNE	MS	526.99	1.22×10^{-11}	1.84×10^{11}	−0.0012	−0.0014	0.0015
3-methyl-2-pentanol	M2PT2OL	MS	504.97	7.52×10^{-12}	2.99×10^{11}	−0.0007	−0.0004	0.0015
3-hydroxy-2,3-dihydro-Maltol	HX3DH23MALTOL	MS	492.74	removed from analysis				
2-cyclohexenol	CHEX2E1OL	MS	443.59	removed from analysis				
phoracanthol	PHOCATHOL	MS	432.85	6.74×10^{-13}	3.34×10^{12}	0.0022	0.0019	0.0003
linalool	LINOL	RS; MS	384.56	1.25×10^{-11}	1.80×10^{11}	0.0006	0.0012	0.0031
4-cyclopentene-1,3-dione	CYPT4E13DONE	MS	365.04	1.26×10^{-11}	1.79×10^{11}	0.0008	0.0015	0.0021
1-(2-furanyl)-2-hydroxy-ethanone	FURYLHXYEONE	MS	356.33	removed from analysis				
2-methyl-crotonal	M2CROTNAL	MS	350.13	1.80×10^{-11}	1.25×10^{11}	−0.0004	0.0008	−0.0018
ethylene glycol butyl ether	ENGOLBETHR	MS	349.45	1.62×10^{-11}	1.39×10^{11}	0.0004	0.0013	0.0006
2-methy-2-butanol	M2BT2OL	MS	342.89	9.94×10^{-12}	2.26×10^{11}	−0.0011	−0.0009	0.0009
cis-4-decenol	CDEC4E1OL	MS	327.97	removed from analysis				
2,4-dimethyl-1-heptene	DM24HPT1ENE	MS	324.27	removed from analysis				
triethylene glycol	TETYNEGLOL	MS	304.57	removed from analysis				
1-nonanol	NON1OL	MS	298.22	4.43×10^{-12}	5.08×10^{11}	0.0009	−0.0011	−0.0016
trans-2-nonene	TNONENE	MS	293.65	removed from analysis				
heptanoic acid	HEPTOIC	RS; MS	289.77	removed from analysis				
octanal	OCTAL	RS; MS	289.51	removed from analysis				
pentanal	PTNAL	RS; MS	288.09	2.42×10^{-11}	9.30×10^{10}	−0.0011	−0.0007	0.0008
p-methoxy-styrene	PMTXYESTYNE	MS	285.67	removed from analysis				
4.6-dimethyl-dodecane	DM45DODCANE	MS	274.61	removed from analysis				
5-methyl-2-hexanone	M5HX2AONE	MS	274.44	1.90×10^{-11}	1.19×10^{11}	−0.0007	0.0000	0.0017
trans-2-nonen-1-ol	TNON1E2OL	MS	270.13	6.21×10^{-12}	3.62×10^{11}	−0.0011	−0.0007	0.0007
5-hydroxy-maltol	HX5MALTOL	MS	269.56	removed from analysis				
2-heptanone	HPT2ONE	MS	266.55	removed from analysis				
furfuryl formate	FURYLFMTE	MS	266.49	removed from analysis				
2-ethyl hexanoic acid	E2HEXOIC	MS	264.43	removed from analysis				
trans-2-octen-1-ol	TOCT1E2OL	MS	255.47	9.66×10^{-12}	2.33×10^{11}	0.0001	0.0007	−0.0006
2-methyl-benzaldehyde	M2BNZAL	MS	255.45	removed from analysis				

[a]ID—identification-type method used: RS (identified by reference standard) and MS (tentatively identified by NIST14 mass spectral library). [b]F—Fischer value from one-way ANOVA test. [c]W—Wilks value from LDA. [d]F—Fischer value from LDA. [e]CDF—canonical discriminant function coefficients.

https://dx.doi.org/10.1021/acs.jafc.0c07554
J. Agric. Food Chem. 2021, 69, 3548−3558

in Table S12. The LDA line plots of all cultivars classified according to CDF1, CDF2, and CDF3 are shown in Figure 3A−C, respectively. The LDA line plot of selected variables for CDF1, CDF2, and CDF3 are shown in Figure 3D−F, respectively. The CDF coefficient 3D plots of the six cultivar centroids and selected variables for CDF1, CDF2, and CDF3 are presented in Figure 5A,B, respectively.

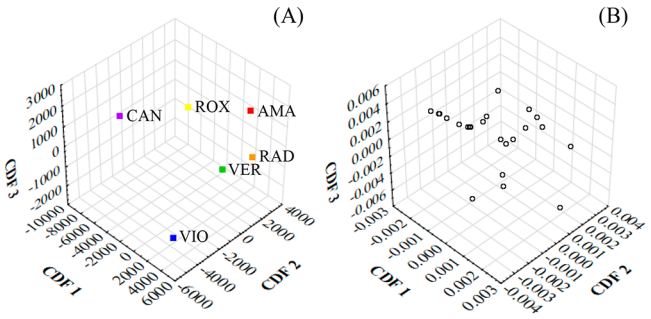

Figure 5. The LDA loading 3D plot of all cultivar centroids (A) and the 23 most predictive VOCs (B) for the three main components.

Throughout all the 30 steps of the backward selection method ($p < 0.05$), 29 VOCs were removed. Consequently, LDA analysis results were based on 23 VOCs, namely, MNTHOL, DM44PT2ONE, 2,3-dimethyl-2-hexanol (DM23HX2OL), E3M2HX13DENE, M4HPT2ONE, BENZACETAL, 2-ethylhexyl octanoate (EHESTOA), 2-nonanol (NON2OL), propanoic acid (PPANOIC), 1-(2-furanyl)-ethanone (FURYLETONE), BMYRCNE, 3-methyl-2-pentanol (M2PT2OL), phoracanthol (PHOCATHOL), linalool (LINOL), 4-cyclopentene-1,3-dione (CYPT4E13DONE), 2-methyl-crotonal (M2CROTNAL), ethylene glycol butyl ether (ENGOLBETHR), M2BT2OL, 1-nonanol (NON1OL), pentanal (PTNAL), 5-methyl-2-hexanone (M5HX2AONE), *trans*-2-nonen-1-ol (TNON1E2OL), and TOCT1E2OL. All the 36 SC samples were classified at a 100% correct rate. Thus, the dimension of the matrix was reduced to 23 VOCs, corresponding to 8.8% of the original matrix of 260 VOCs.

The LDA 3D plot in Figure 5A presented a high level of discrimination along the three CDFs between all cultivars. In CDF 1, a prominent discrimination was obtained between the CAN cultivar and the other southern cultivars. Also, in CDF1, an interesting discrimination was observed between two clusters, one formed by ROX and VER cultivars and the other formed by AMA, RAD, and VIO cultivars. In CDF 2, a higher discrimination was observed to CAN and VIO cultivar from the other cultivars, where substantial discrimination was also verified between VER cultivar and other cultivars. Notably, in CDF3, an equilibrate and a reasonable discrimination between all cultivars under analysis were observed. According to the LDA results, discriminate was possible and all cultivar samples were classified correctly based only on selected 23 of 260 VOCs.

PLSs and HCA. A second PLS analysis was performed to verify the differentiation/correlation structure between all SC cultivars when only the 23 most predictive VOCs were used. The PLS information is summarized in Table S13. The loading results and VIP scores of PLS analysis for each one of 23 VOCs are summarized in Table 1. The PLS loading results of six cultivar centroids and the score results of all samples under analysis are described in Table S12. The PLS loadings line plots of all cultivars classified according to PLS1, PLS2, and PLS3 are shown in Figure S4A−C, respectively. The PLS loading line plot of all 23 VOCs for PLS1, PLS2, and PLS3 are shown in Figure S4D−F, respectively. The PLS loading 3D plots of the six cultivar centroids and 23 VOCs for PLS1, PLS2, and PLS3 are shown in Figure 6A,B, respectively.

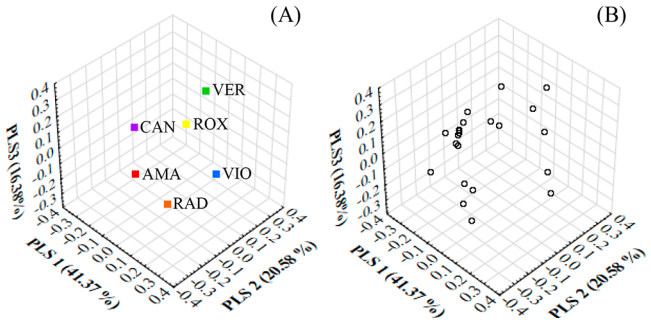

Figure 6. The PLS loading 3D plot of all cultivar centroids (A) and the 23 most predictive VOCs (B) for the three main components.

The PLS analysis based on PLS1, PLS2, and PLS3 explained 78.3% of TVA, where the sum of all 14 components explained 99.9%. As expected, the results presented in Figure 6A demonstrate that the six cultivar centroids were undoubtedly separated from each other, presenting a high inter-cultivar variance for all cultivars. Moreover, the difference level of variance between cultivars was clearly higher than that obtained in the previous PLS (complete matrix). Throughout the individual analysis of PLS1 projection (41.4%), it was possible to visualize a higher variance between CAN cultivar and the southern cultivars. In PLS2 projection (20.6%), a considerable variance among the five was observed as follows: AMA−RAD−ROX−VER−VIO. For PLS3 projection (16.4%), a high variance was observed between VIO and RAD cultivars from the other cultivars, principally from ROX and VER cultivars.

HCA was performed based on the previously selected 23 VOCs in order to determine the Euclidean linkage distances between all samples and complete an appropriate measure of distance and linkage criterion between the SC cultivars. The HCA dendrogram is presented in Figure 7.

Once again, the higher distance was obtained between the CAN cultivar and southern cultivars. Among the southern cultivars, the linkage level decreases in the following order: VIO > ROX > AMA > VER > RAD. The samples from CAN,

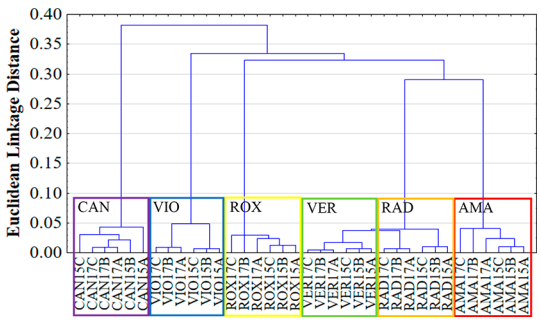

Figure 7. The HCA dendrogram based on Euclidean linkage distances for all replicates (A−C) from six SC cultivars during 2015 and 2017 harvest years.

https://dx.doi.org/10.1021/acs.jafc.0c07554
J. Agric. Food Chem. 2021, 69, 3548−3558

VIO, ROX, and AMA cultivars presented a high level of differentiation, being easily distinguished from all other regional cultivars. However, low distances were obtained for samples from RAD and VER cultivars, indicating the probable existence of a common pre-antecedent SC cultivar. The PLS projection, a supervised pattern recognition method, and HCA, an unsupervised pattern recognition method, presented similar results, allowing to infer that all cultivars can be well differentiated based only on the 23 most predictive VOCs.

Remarks on Differentiation of SC Cultivars. The present study represents the first application to differentiate molecularly the SC cultivars from Madeira Island, Portugal. The establishment of the volatile profile from all cultivars was performed by the HS-SPME/GC−MS method, recognizing a total of 260 different VOCs belonging to 15 different chemical classes, namely, ALC, ALD, BNZ, CAC, EST, ETH, FUR, HYD, KET, NPH, NIT, PHE, PYR, SUL, and TER. Based on the chemometric analysis, we concluded that it was possible to establish an exclusive volatile profile for each one of the six different SC cultivars and consequently achieve a high level of differentiation between all cultivars. The CAN cultivar appears to be more molecularly different from the other regional cultivars. Interestingly, the CAN cultivar is empirically indicated by farmers as the oldest cultivar in Madeira Island, which is a potential cultivar originated from the group formed by the first cultivars introduced in the 14th century due to its agronomic isolation into a small cluster localized in the northern part of the island. Nevertheless, its volatile composition is also the result of its adaptation to the unique agronomic and climatic conditions of the north of the island. Among the southern cultivars, ROX and VIO cultivars presented a higher level of differentiation, followed by AMA cultivar, and in minor extension by VER and RAD cultivars. ROX cultivar is the most appreciated by regional farmers due to its high productivity when grown in areas with strong sunlight exposure and high water availability, which is responsible for approximately 75% of the SC total production. Also, AMA and VER cultivars are commonly produced in areas with lower water availability, but with lower productivity. VIO and RAD cultivars were recently introduced in the regional SC production system due to its apparently high pest robustness.

In addition, some VOCs appear to have a specific correlation for a particular SC cultivar and can be recognized as a potential marker for those cultivars. The CAN cultivar was characterized by VOCs from KET, HYD, and TER chemical classes, such as M4PT2ONE, DM46HPT2ONE, M2HPTENE, DM24HPT1ENE, APHELDNE, and BPHELDNE. The ROX cultivar was influenced by VOCs from EST chemical class, such as EESTPA, M2EESTPA, PESTAA, M2PESTAA, and E2MESTPTA; the VIO cultivar was influenced by VOCs from KET and ALC, such as HPT2ONE, M5HX3E4ONE, OCT2OL, NON2OL, and CDEC5E1OL. The remaining regional cultivars were directly influenced by a few VOCs, namely, PTDONE and MALTOL for the AMA cultivar, M2FUR for the RAD cultivar, and DEC1OL for the VER cultivar. Furthermore, LDA, PLS, and HCA demonstrated that it is possible to differentiate all regional SC cultivars only based on the 23 most predictive VOC dataset, namely, MNTHOL, DM44PT2ONE, DM23HX2OL, E3M2HX13DENE, M4HPT2ONE, BENZACETAL, EHESTOA, NON2OL, PPANOIC, FURYLETONE, BMYRCNE, M2PT2OL, PHOCATHOL, LINOL, CYPT4E13DONE, M2CROTNAL, ENGOLBETHR, M2BT2OL, NON1OL, PTNAL, M5HX2AONE, TNON1E2OL, and TOCT1E2OL.

Finally, the information obtained in this study about the regional SC cultivars represents an important contribution for the maintenance of biodiversity and subsistence of the sugar industry not only in Madeira Island, Portugal, but also in Europe. Moreover, this study is also a fundamental input to establish the typicality of traditional SC-based products, such as SCH, and its submission to an EU certification.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jafc.0c07554.

Typical GC−MS chromatograms for *Amarela*, *Radiada*, *Roxa*, *Verde*, *Violeta*, and *Canica* (F) SC cultivars; PCA loading line plots of all cultivar samples based on the selected 259 VOCs for PC1, PC2, and PC3 and the PLS loading line plots of all cultivar samples for PLS1, PLS2, and PLS3; LDA line plots of all cultivars based on the 23 most predictive VOCs for CDF1, CDF2, and CDF3 and LDA line plots of the 23 most predictive VOCs for CDF1, CDF2, and CDF3; PLS loading line plots of all cultivar samples based on the 23 most predictive VOCs for PLS1, PLS2, and PLS3 and the PLS loading line plots of the 23 most predictive VOCs for PLS1, PLS2, and PLS3; PLS 3D plot and HCA dendrogram for results from the matrix reduction procedure based on the VIP scores; ID replicate number, ID replicate code, ID sample code, ID variety code, variety empirical name, harvest year, and geographical area of SC samples; ID number, retention time, main ions (m/z), target ion, match percent, identification type, IUPAC name, NIST database, abbreviation, CAS number, molecular formula, and main chemical class of identified VOCs in SC samples; mean and RSD values of identified VOCs in SC samples from "*amarela*", "*radiada*", and "*roxa*" cultivars; mean and RSD values of identified VOCs in SC samples from "*Verde*", "*Violeta*", and "*Canica*" cultivars; mean, minimum, and maximum peak area values of identified VOCs in SC samples; VOCs from regional cultivar samples previously identified in other SC-based products; summary of number of VOCs identified, RPA, and total RPA (%) values of main chemical classes identified in SC samples; one-way ANOVA with post-hoc Tukey's test results based on the RPAs of the identified VOCs in SC samples; information summary of PCA and PLS analysis based on the RPAs of the identified VOCs in SC samples; loading results and VIP scores of variables from PCA and PLS analysis based on the RPAs of the identified VOCs in SC samples; loading results of samples and variables from PCA and PLSs analysis based on the RPAs of the identified VOCs in SC samples; score results of all cases from PCA and PLS analysis based on the RPAs of the identified VOCs in SC samples; canonical discriminant function coefficients and highest probability classification results of samples from LDA after matrix reduction method obtained from one-way ANOVA test based on the RPAs of identified VOCs in SC samples; information summary of PLS analysis based only on the RPAs of the 23 most predictive VOCs identified in SC samples; and results of variables from

https://dx.doi.org/10.1021/acs.jafc.0c07554
J. Agric. Food Chem. 2021, 69, 3548−3558

LDA after matrix reduction method to 20% of original dimension according to higher VIP score values from the PLS analysis based on the RPAs of the identified VOCs in SC samples (PDF)

AUTHOR INFORMATION

Corresponding Author

José S. Câmara − *CQM, Centro de Química da Madeira, Universidade da Madeira, 9020-105 Funchal, Portugal; Departamento de Química, Faculdade de Ciências Exactas e Engenharia, Universidade da Madeira, 9020-105 Funchal, Portugal;* orcid.org/0000-0003-1965-3151; Phone: +351 291705112; Email: jsc@staff.uma.pt; Fax: +351 291705149

Authors

Pedro Silva − *CQM, Centro de Química da Madeira, Universidade da Madeira, 9020-105 Funchal, Portugal*

Jorge Freitas − *CQM, Centro de Química da Madeira, Universidade da Madeira, 9020-105 Funchal, Portugal*

Fernando M. Nunes − *CQ-VR, Centro de Química—Vila Real, Departamento de Química, Universidade de Trás-os-Montes e Alto Douro, 5001-801 Vila Real, Portugal;* orcid.org/0000-0001-5540-318X

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jafc.0c07554

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors acknowledge the Fábrica Mel-de-Cana Ribeiro Sêco de V. Melim, Lda, (FMCRS), Funchal, Portugal, and the Agência Regional para o Desenvolvimento da Investigação, Tecnologia e Inovação (ARDITI), through the support granted under the M1420 Project—09-5369-FSE-000001—PhD Scholarship in companies for PhD grant of the author P.S.. This work was supported by FCT (project PEst-OE/QUI/UI0674/2019, CQM, Portuguese Government funds) and through Madeira 14-20 Program, project PROEQUIPRAM—Reforço do Investimento em Equipamentos e Infraestruturas Científicas na RAM (M1420-01-0145-FEDER-000008) and by ARDITI through the project M1420-01-0145-FEDER-000005—Centro de Química da Madeira—CQM+ (Madeira 14-20).

REFERENCES

(1) Dias, H. B.; Sentelhas, P. C. Evaluation of Three Sugarcane Simulation Models and Their Ensemble for Yield Estimation in Commercially Managed Fields. *Field Crops Res.* **2017**, *213*, 174−185.

(2) Singh, R. B.; Mahenderakar, M. D.; Jugran, A. K.; Singh, R. K.; Srivastava, R. K. Assessing Genetic Diversity and Population Structure of Sugarcane Cultivars, Progenitor Species and Genera Using Microsatellite (SSR) Markers. *Gene* **2020**, *753*, 144800.

(3) Piperidis, N.; D'Hont, A. Sugarcane genome architecture decrypted with chromosome-specific oligo probes. *Plant J.* **2020**, *103*, 2039−2051.

(4) Cheavegatti-Gianotto, A.; de Abreu, H. M. C.; Arruda, P.; Bespalhok Filho, J. C.; Burnquist, W. L.; Creste, S.; di Ciero, L.; Ferro, J. A.; de Oliveira Figueira, A. V.; de Sousa Filgueiras, T.; Grossi-de-Sá, M. d. F.; Guzzo, E. C.; Hoffmann, H. P.; de Andrade Landell, M. G.; Macedo, N.; Matsuoka, S.; de Castro Reinach, F.; Romano, E.; da Silva, W. J.; de Castro Silva Filho, M.; César Ulian, E. Sugarcane (Saccharum X Officinarum): A Reference Study for the Regulation of Genetically Modified Cultivars in Brazil. *Trop. Plant Biol.* **2011**, *4*, 62−89.

(5) Vieira, M. L. C.; Almeida, C. B.; Oliveira, C. A.; Tacuatiá, L. O.; Munhoz, C. F.; Cauz-Santos, L. A.; Pinto, L. R.; Monteiro-Vitorello, C. B.; Xavier, M. A.; Forni-Martins, E. R. Revisiting Meiosis in Sugarcane: Chromosomal Irregularities and the Prevalence of Bivalent Configurations. *Front. Genet.* **2018**, *9*, 213.

(6) Godshall, M. A.; Legendre, B. L. SUGAR|Sugarcane. *Encyclopedia of Food Sciences and Nutrition*; Elsevier, 2003; pp 5645−5651.

(7) de Matos, M.; Santos, F.; Eichler, P. Sugarcane World Scenario. *Sugarcane Biorefinery, Technology and Perspectives*; Elsevier, 2020; pp 1−19.

(8) Silva, P.; Freitas, J.; Silva, C. L.; Perestrelo, R.; Nunes, F. M.; Câmara, J. S. Establishment of Authenticity and Typicality of Sugarcane Honey Based on Volatile Profile and Multivariate Analysis. *Food Control* **2017**, *73*, 1176−1188.

(9) Silva, P.; Silva, C. L.; Perestrelo, R.; Nunes, F. M.; Câmara, J. S. Fingerprint Targeted Compounds in Authenticity of Sugarcane Honey - An Approach Based on Chromatographic and Statistical Data. *LWT–Food Sci. Technol.* **2018**, *96*, 82−89.

(10) Cramer, A.-C. J.; Mattinson, D. S.; Fellman, J. K.; Baik, B.-K. Analysis of Volatile Compounds from Various Types of Barley Cultivars. *J. Agric. Food Chem.* **2005**, *53*, 7526−7531.

(11) Wadood, S. A.; Boli, G.; Xiaowen, Z.; Raza, A.; Yimin, W. Geographical Discrimination of Chinese Winter Wheat Using Volatile Compound Analysis by HS-SPME/GC-MS Coupled with Multivariate Statistical Analysis. *J. Mass Spectrom.* **2020**, *55*, No. e4453.

(12) Bryant, R. J.; McClung, A. M. Volatile Profiles of Aromatic and Non-Aromatic Rice Cultivars Using SPME/GC-MS. *Food Chem.* **2011**, *124*, 501−513.

(13) Esteki, M.; Simal-Gandara, J.; Shahsavari, Z.; Zandbaaf, S.; Dashtaki, E.; Vander Heyden, Y. A Review on the Application of Chromatographic Methods, Coupled to Chemometrics, for Food Authentication. *Food Control* **2018**, *93*, 165−182.

(14) Lubes, G.; Goodarzi, M. Analysis of Volatile Compounds by Advanced Analytical Techniques and Multivariate Chemometrics. *Chem. Rev.* **2017**, *117*, 6399−6422.

(15) Sengun, I. Y.; Karabiyikli, S. Importance of Acetic Acid Bacteria in Food Industry. *Food Control* **2011**, *22*, 647−656.

(16) Buśko, M.; Jeleń, H.; Góral, T.; Chmielewski, J.; Stuper, K.; Szwajkowska-Michałek, L.; Tyrakowska, B.; Perkowski, J. Volatile Metabolites in Various Cereal Grains. *Food Addit. Contam., Part A: Chem., Anal., Control, Exposure Risk Assess.* **2010**, *27*, 1574−1581.

(17) Tian, J. P.; Ma, Z. Y.; Zhao, K. G.; Zhang, J.; Xiang, L.; Chen, L. Q. Transcriptomic and Proteomic Approaches to Explore the Differences in Monoterpene and Benzenoid Biosynthesis between Scented and Unscented Genotypes of Wintersweet. *Physiol. Plant.* **2019**, *166*, 478−493.

(18) Niederbacher, B.; Winkler, J. B.; Schnitzler, J. P. Volatile Organic Compounds as Non-Invasive Markers for Plant Phenotyping. *J. Exp. Bot.* **2015**, *66*, 5403−5416.

(19) Maffei, M. E.; Gertsch, J.; Appendino, G. Plant Volatiles: Production, Function and Pharmacology. *Nat. Prod. Rep.* **2011**, *28*, 1359−1380.

(20) Hammerbacher, A.; Coutinho, T. A.; Gershenzon, J. Roles of Plant Volatiles in Defence against Microbial Pathogens and Microbial Exploitation of Volatiles. *Plant, Cell Environ.* **2019**, *42*, 2827−2843.

(21) McGorrin, R. J. The Significance of Volatile Sulfur Compounds in Food Flavors. *Volatile Sulfur Compounds in Food*; ACS Symposium Series; American Chemical Society, 2011; Vol. *1068*, pp 3−31.

(22) Cerny, C. Thermal Generation of Aroma-Active Volatiles in Food. *The Chemistry and Biology of Volatiles*; John Wiley & Sons, Ltd: Chichester, UK, 2010; pp 231−252.

(23) Zhang, B.; Zhong, Z.; Min, M.; Ding, K.; Xie, Q.; Ruan, R. Catalytic Fast Co-Pyrolysis of Biomass and Food Waste to Produce Aromatics: Analytical Py-GC/MS Study. *Bioresour. Technol.* **2015**, *189*, 30−35.

(24) Obulesu, M.; Bhattacharya, S. Color Changes of Tamarind (Tamarindus indicaL.) Pulp During Fruit Development, Ripening, and Storage. *Int. J. Food Prop.* **2011**, *14*, 538−549.
(25) van Hylckama Vlieg, J. E.; Veiga, P.; Zhang, C.; Derrien, M.; Zhao, L. Impact of Microbial Transformation of Food on Health—from Fermented Foods to Fermentation in the Gastro-Intestinal Tract. *Curr. Opin. Biotechnol.* **2011**, *22*, 211−219.
(26) Porker, K.; Zerner, M.; Cozzolino, D. Classification and Authentication of Barley (Hordeum Vulgare) Malt Varieties: Combining Attenuated Total Reflectance Mid-Infrared Spectroscopy with Chemometrics. *Food Anal. Methods* **2017**, *10*, 675−682.
(27) Khalil, M. N. A.; Fekry, M. I.; Farag, M. A. Metabolome Based Volatiles Profiling in 13 Date Palm Fruit Varieties from Egypt via SPME GC−MS and Chemometrics. *Food Chem.* **2017**, *217*, 171−181.
(28) Giannetti, V.; Boccacci Mariani, M.; Mannino, P.; Marini, F. Volatile Fraction Analysis by HS-SPME/GC-MS and Chemometric Modeling for Traceability of Apples Cultivated in the Northeast Italy. *Food Control* **2017**, *78*, 215−221.

https://dx.doi.org/10.1021/acs.jafc.0c07554
J. Agric. Food Chem. 2021, 69, 3548−3558