LWT - Food Science and Technology 203 (2024) 116337

Contents lists available at ScienceDirect

LWT

journal homepage: www.elsevier.com/locate/lwt

Indirect evaluation of watermelon volatile profile: Detection of subtle changes with e-nose systems

Alejandro Fredes [a], Jaime Cebolla-Cornejo [b], Joaquín Beltrán [c], Carmina Gisbert [d], Belén Picó [d], Mercedes Valcárcel [a], Salvador Roselló [a,*]

[a] Joint Research Unit UJI-UPV - Improvement of Agri-food Quality, Biology, Biochemistry and Natural Sciences Department, Universitat Jaume I, Avda. Sos Baynat s/n, 12071, Castelló de la Plana, Spain
[b] Joint Research Unit UJI-UPV - Improvement of Agri-food Quality, COMAV, Universitat Politècnica de València, Cno. de Vera s/n, 46022, València, Spain
[c] Environmental and Public Health Analytical Chemistry, Research Institute for Pesticides and Water (IUPA), Universitat Jaume I, Av. Sos Baynat s/n, 12071, Castellón de la Plana, Spain
[d] Instituto de Conservación y Mejora de la Agrodiversidad Valenciana, Universitat Politècnica de València, Cno de Vera s.n, 46022, Spain

ARTICLE INFO

Keywords:
Citrullus lanatus
Volatiles
Aroma
Breeding
Grafting

ABSTRACT

The effectiveness of e-nose systems as high-throughput tools for volatile profiling in watermelon was investigated focusing on discerning subtle changes induced by the use of different rootstocks. Partial Least Square Discriminant Analysis (PLS-DA) models, both GC-MS and e-nose data, demonstrated moderate performance in classification due to nuanced differences among groups (the same F1 hybrid was used as scion). However, PLS-DA biplots revealed a clear correlation between GC-MS and e-nose data. This methodology enabled the e-nose system to identify the effects of specific root-scion combinations compared to non-grafted controls and detect combinations with more variable volatile profiles. Remarkably, the e-nose system identified samples with anomalous volatile profiles, mirroring the capabilities of GC-MS data. Additionally, PLS models were developed to provide reasonably accurate predictions of key compound contents like geranylacetone, (Z)-6-nonen-1-ol, or (Z)-6-nonenal, crucial for watermelon flavor and taste perception. Overall, this study highlights the potential of e-nose systems in discerning nuanced variations in watermelon volatile profiles affecting aroma. Incorporating volatile profile evaluation capabilities using such systems will significantly optimize quality control processes and plant breeding programs.

1. Introduction

Watermelon is a highly appreciated fruit. In 2020, 101.9 million t were produced worldwide, reaching a consumption of 29.36 g per capita and day, which is exceptionally high in Asia, topping 39.07 g per capita and day (https://www.fao.org/faostat). Apart from being refreshing, watermelon is also appreciated for its appealing flavour. This flavour is influenced by the accumulation of soluble solids (taste perception) and volatile organic compounds (VOCs) contributing to the aroma. Among these compounds, a major emphasis has been placed on the study of sugar accumulation conditioning sweetness perception. Organic acids have been less studied, though they can tinge sweetness perception or even create new flavours (Gao, Zhao, Lu, He, & Liu, 2018).

Regarding aroma, around 80 VOCs have been described in watermelon, with C6 and C9 aldehydes and alcohols being the most abundant. Other compounds, such as esters, ketones, lactones, furans, and apocarotenoids are also present (Beaulieu & Lea, 2006; Lewinsohn et al., 2005; Pino, Marbot, & Aguero, 2003; Yajima, Sakakibara, Ide, Yanai, & Hayashi, 1985).

The large number of compounds involved in the taste of watermelons, especially VOCs, presents a challenge in evaluation sensory quality This complexity is particularly pronounced when evaluating numerous samples with different profiles, as those used in breeding programs and quality control processes such scenarios, assessment by sensory panels or by precise analytical methodologies, such as GC-MS (Alejandro Fredes et al., 2016; Beaulieu & Lea, 2006; Verzera, Dima, Tripodi, & Ziino, 2011) proves impractical as they have not been designed for high-throughput evaluation.

Alternatively, the volatile profile of fruits can be analyzed following an indirect approach using electronic nose (e-nose) devices. Electronic

* Corresponding author.
E-mail address: rosello@uji.es (S. Roselló).

https://doi.org/10.1016/j.lwt.2024.116337
Received 7 March 2024; Received in revised form 13 May 2024; Accepted 11 June 2024
Available online 13 June 2024

noses utilize arrays of electronic sensors, mostly metal oxide sensors (MOS), which change their electrical resistance based on the type and concentration of the stimulating volatile compounds. After chemometric analyses qualitative and semi-quantitative evaluations can be performed by comparing response patterns with reference samples previously evaluated (Gardner & Bartlett, 1999). Such systems offer several advantages: greater objectivity, the ability to evaluate a large number of samples, higher discrimination sensitivity, and the potential to build scalable long-term maps for assessing and comparing volatile profiles of numerous samples over extended periods. However, drift issues caused by environmental variations in temperature or humidity (Gardner and Persaud, 2001) or sensor characteristics (Holmberg & Artursson, 2003) also hinder their application. Nonetheless, recent progress in drift correction strategies opens possibilities to overcome these limitations (Valcárcel et al., 2021).

Although in crops such as tomatoes, the use of e-nose systems has been extensively analyzed (Hernández, Wang, Hu, & Pereira, 2008; Sinesio et al., 2000; Valcárcel et al., 2021), in the case of cucurbits, a lower effort has been placed, usually analyzing evident differences. For example, e-nose systems have been used in melon for differentiating the ripening stage (Benady, Simon, Charles, & Miles, 1995), climacteric ripening (Chaparro-Torres, Bueso, & Fernández-Trujillo, 2016) or the effects of crop management (Wang et al., 2023); in *Cucurbita*, for the differentiation of specific aromatic tinges (Junxing et al., 2022) and differentiating species (C.-L. Zhou, Mi, Hu, & Zhu, 2017); and in cucumber for analyzing the effect of transgenic lines (Zawirska-Wojtasiak, Gośliński, Szwacka, Gajc-Wolska, & Mildner-Szkudlarz, 2009) and postharvest evolution (Feng, Zhang, Bhandari, & Guo, 2018). To our knowledge, in watermelon e-nose has only been applied to the differentiation of two cultivars at different ripening stages (Bianchi, Provenzi, & Rizzolo, 2020).

In this context, the objective of this study was to validate the potential use of an electronic nose for rapid phenotypic evaluation of volatile profiles in watermelon fruits, which could be very useful in scenarios where a high number of samples must be evaluated as different genotype selection in breeding programs or quality control tasks. The electronic fingerprints of these samples would be used not to classify them into groups but to develop similarity maps of volatile profiles, allowing for the sample comparison with elite references for selection or control purposes. As a hard case study to analyze e-nose phenotyping capabilities for selection compared with GC-MS evaluation, both systems were used to evaluate the modulation of VOCs profile from one watermelon commercial hybrid when grafted on a set of commercial and experimental rootstocks. This set of scion-rootstocks combinations was selected considering the differences in the volatile profile estimated using GC-MS that had previously been reported (Fredes et al., 2017).

2. Materials and methods

2.1. Experimental design and growing conditions

The commercial watermelon cultivar Oneida F1 (Rijk Zwaan, Almería, Spain) was used as scion with four rootstocks combinations. The rootstocks used included the accession BGV0005167 *Citrullus lanatus* var. citroides (also known as *Citrullus amarus*), GC, with high nematode resistance (Gisbert et al., 2017); an accession of Cucurbita pepo (GPepo), and two interspecific hybrids of *Cucurbita maxima* x *Cucurbita moschata* (GMM1 and GMM2). GMM1 is an experimental line derived from the cross between *C. maxima* VAV 1860 (Large Warted Hubbard, Australia) and *C. moschata* PI 550689 (Canada Crookneck Squash). GMM2 is the commercial rootstock F1 Cobalt (Rijk Zwaan, Almería, Spain). The first two accessions were obtained from the Germplasm Bank of the Institute for the Conservation and Improvement of Valencian Agrodiversity (COMAV, Valencia, Spain). Non-grafted (NG) and self-grafted plants of Oneida F1 were used as controls.

A completely randomized block design was employed. Samples were collected in triplicate for each scion-rootstock combination, with each replicate consisting of six plants. Grafting was performed using the approach grafting method, and one month later, all the plants were transplanted to an experimental field at Rijk Zwaan's facility in Picassent, Valencia, Spain. The planting scheme used was 2.0 m × 1.0 m, and irrigation and fertilization were managed according to standard cultural practices for the crop in that region.

Table 1
Composition (ng mL^{-1}) of the synthetic standard used for drift correction.

Volatile	Concentration	Volatile	Concentration
(Z)-3-Nonen-1-ol	56.08	Propyl butyrate	2.75
1-Hexanol	49.44	Phenol	2.69
(E,Z)-2,6-Nonadienal	47.30	(E,E)-2,4-Heptadienal	2.61
Hexanal	31.38	Eucalyptol	2.08
(Z)-3-hexen-1-ol	24.94	Butyl acetate	1.95
Methyl-2-methyl butyrate	24.06	Eugenol	1.71
Ethy 1-butanoate (37)	19.76	Beta-Ionone	1.71
Nonanal	17.13	1-Pentanol	1.60
(Z)-6-nonenal	15.33	2-methylpropyl butyrate	1.48
1-Octanol	13.97	Amyl acetate	1.45
Ethyl-2-methyl butyrate	12.64	6-methyl-5-Hepten-2-one	1.37
(Z)-6-Nonen-1-ol	11.85	Heptanal	1.36
1-octen-3-ol	9.89	(E,Z)-2,6-Nonadien-1-ol	1.33
Octyl acetate	5.83	(E,E)-2,4-Decadienal	1.33
1-Nonanol	5.37	methyl hexanoate	1.32
(E)-2-Heptenal	5.00	Beta-ciclocytral	1.32
3-methylbutyl acetate	4.97	Ethyl hexanoate	1.29
(E)-2-Nonenal	4.89	Butyl isobutyrate	1.16
Octanal	3.70	Decanal	0.92
2-methylbutylacetate	3.56	Phenethyl acetate	0.92
Benzaldehyde	3.46	(E)-2-Octenal	0.75
Phenylacetaldehyde	3.43	Butyl butyrate	0.74
1-Decanol	3.23	(E,E)-2,4-Nonadienal	0.67
Benzyl Alcohol	3.23	(E)-2-hexenal	0.64
Hexyl acetate	3.04	ethyl heptanoate	0.60
2-Methyl propyl acetate	2.77	Benzyl acetate	0.29
Geranylacetone	2.76	(Z)-3-Hexen-1-ol, acetate	0.27

2.2. Sampling

Fruits were sampled from all the plants in each replicate when characteristic external signs of maturity were observed. A 5 cm portion from the equatorial section of each fruit was obtained. For each plant material, samples from fruits of the same replicate were processed together to obtain a biological mean for that replicate. Before sample homogenization, the pericarp, approximately 2 mm of adjacent flesh, and the seeds were removed. Homogenization was performed using a blender (Krups KB720, Groupe Seb Iberica, Barcelona, Spain), and the samples were kept frozen at −80 °C until analysis.

2.3. Reagents and volatiles extraction

The reference standards for each of the studied volatile compounds (purity grade 90–99.55%) were obtained from Supelco (Sigma-Aldrich and Fluka, Barcelona, Spain). These standard solutions were prepared, handled, and stored as described in Fredes et al. (2016).

For the quantification of volatiles by GC-MS, solvents used for desorption and dilution of analytes retained in Supelclean™ ENVI-Carb™ 120–400 mesh cartridges (6 mL SPE tubes) were all of GC-grade quality.

A volatile standard mixture (Table 1) was used for drift correction and inter-sequence standardization following the strategy proposed by

Valcárcel et al. (2021). This volatile standard mixture was developed based on the average concentrations of a group of samples previously analyzed by GC-MS Fredes et al. (2016) and adjusted to the sensitivity levels of the electronic nose.

2.4. GC-MS volatile analysis

The determination of volatile compounds was performed as described by Fredes et al. (2016). In this case, 61 compounds were analyzed using purge and trap extraction with 500 mg Carbowax cartridges, which were preconditioned with 5 mL of diethyl ether and 5 mL of n-hexane. For the extraction, 30 g of sample were placed in a 150 mL flask, and the system was sealed with a ground glass stopper equipped with a nitrogen gas inlet and an outlet connected to the volatile trap. During extraction, the samples were stirred at 300 rpm and heated to a temperature of 40 °C, with a nitrogen gas flow rate of 1.6 L min−1. After 49 min of extraction, the retained compounds were eluted with 5 mL of a diethyl ether:hexane mixture (1:1) and 5 mL of diethyl ether. Subsequently, the extract was concentrated under a nitrogen gas flow in a water bath at 35 °C to a volume of 0.5 mL.

The analysis of volatile compounds was performed using a gas chromatograph (Varian CP-3800) coupled to a mass spectrometer (ion trap, Saturn 4000, Varian). A 1 μL sample was injected in splitless mode (injector temperature 220 °C). Separation was achieved using a capillary column (Supelco, Bellefonte, PA) of dimensions 30 m × 0.25 mm, Supelcowax 10 (0.25 μm internal diameter). Helium was used as the carrier gas at a flow rate of 1 mL min-1. The following temperature gradient was applied: 40 °C for 5 min, a ramp of 4 °C min-1 to 160 °C, and finally a ramp of 30 °C min-1 to 250 °C, which was held for 1 min, resulting in a total analysis time of 39 min. Full-scan spectra (m/z 50–200 Da) were acquired using electron impact ionization (70 eV) in positive mode with an external ionization configuration. The GC-MS interface temperature was set at 275 °C, the ion trap temperature at 190 °C, and the manifold temperature at 60 °C. Compound identification was performed by comparing the mass spectra, retention times of peaks, and retention indices calculated according to the formula used by Kovats (1958) based on a mixture of n-alkanes (C7–C30) under the same analysis conditions. External calibration curves in solvent were used for the quantification of all studied compounds.

2.5. E-nose volatile evaluation

For e-nose phenotyping, a commercial system Fox 4000 (Alpha M.O. S., Toulouse, France) was used, which consists of 18 sensors based on metal oxide semiconductors (MOS) arranged in three chambers connected in series. The system also includes an automated injection system (CombiPAL HS100, CTC Analytics, Zwingen, Switzerland) and computer software (AlphaSoft v9) for control and data acquisition. The analysis was performed as described by Valcárcel et al. (2021) with minor modifications (10 min of vial incubation for 10 min at 45 °C with 500 rpm of agitation). Each sample was analyzed in duplicate, with sequences of 16 vials (6 samples per sequence) and four synthetic standards placed randomly for subsequent drift correction.

The response of the sensors was measured as the variation in electrical resistance due to the reaction of volatile compounds in the sample on the sensor's active surface. Two types of parameters were obtained from the graphical representation of each sensor's response to volatile stimuli: the normalized maximum intensity of each sensor (INmáx) calculated as (Ri-Rmax)/Ri (where Ri is the electrical resistance of the sensors when detecting clean air and Rmax is the electrical resistance during compound detection) and the slope (K) calculated using the formula (Lin & Zhang, 2016): $K = \frac{t_{máx} - t_0}{IN_{máx}}$, (where t_{max} is the time of maximum signal intensity, INmáx, and t_o is the delay time in signal start). This slope is also related to the process of volatile absorption on the surface of each sensor (García-González & Aparicio, 2002).

Drift correction and inter-sequence standardization were performed following the procedure described by Valcárcel et al. (2021). Inter-sequence standardization was calculated using a standardization coefficient obtained from the difference of signals of the same reference standard (Table 1) measured in two different sequences.

2.6. Data analysis

Graphical MANOVA Biplot representations were carried out to study the scion-rootstock effect considering both the volatiles' contents analyzed via GC-MS and the e-nose parameters measured. Bonferroni circles were used to represent the confidence intervals ($\alpha = 0.05$), and their projection on each variable enable the identification of significant differences between groups. Multibiplot free software was used to perform the Biplot analysis (Vicente-Villardon, 2015).

Discriminant analysis based on partial least squares regression (PLS-DA) was chosen due to the orthogonality of the calculated latent discriminant variables. This helps to avoid collinearity issues that may arise when using other types of analyses, such as linear discriminant analysis (LDA), because of the nature of the original variables (groups of volatile compounds sharing chemical nature and biosynthetic pathways) and the similarity of sensor groups in terms of material composition (Hubert & Branden, 2003).

To develop the models effectively, the raw data were pre-processed using autoscaling to prevent scale differences in certain variables from biasing the modelling process. The obtained dimensionality reduction and classification models were validated using the Venetian blinds cross-validation method. The goal was to minimize the root mean squared error of cross-validation (RMSECV) and identify the optimal number of latent variables, avoiding overfitting issues (Brereton & Lloyd, 2014). The criterion followed was to stop including latent variables when the reduction in RMSECV was less than 2%, always aiming to use the fewest possible number of latent variables. Since the presence of outliers can heavily influence these models, they were eliminated if they exceeded the confidence thresholds of Hotelling's T2 statistics and Q residuals (Jackson & Mudholkar, 1979). %RMSECV was calculated as the ratio between RMSECV and the maximum value of the variable and expressed as a percentage.

The quality of classification will be determined based on the accuracy percentage for each class, relying on specificity (the number of samples predicted outside the class divided by the number of samples that should be outside the class) and sensitivity (the number of samples predicted within the class divided by the number of samples in the class). Sensitivity and specificity are obtained for each class, both for the calibration (self-prediction) results and cross-validation. The percentage of correct predictions for each class is determined as a weighted average of their specificity and sensitivity values (Wise et al., 2006).

Predictive regression models for volatile compounds' content using electronic signal parameters have been developed using the partial least squares regression (PLS) technique to better relate electronic sensors parameters with VOCs quantified. The pre-processing criteria, model validation, selection of the number of latent variables, and elimination of anomalous data have been similar to those used in PLS-DA analyses. To refine the initial models, a hierarchical methodology of variable selection (interval PLS variable selection, iPLS) was employed to reduce the number of explaining variables used in the models and improve their goodness of fit. This includes both the incremental option (sequential inclusion of variables, forward selection) and the detrimental option (sequential exclusion of variables, reverse selection). The final models for each compound were selected considering the RMSECV value (Wise et al., 2006).

All the described PLS related calculations have been performed using the PLS Toolbox version 8.2.1 (Eigenvector Research Inc, Wenatchee, WA, USA) within the MATLAB environment (version R2021a, 8.3.0.532, MathWorks, Inc., Natick, MA, USA).

To better visualize the importance of the electronic sensor

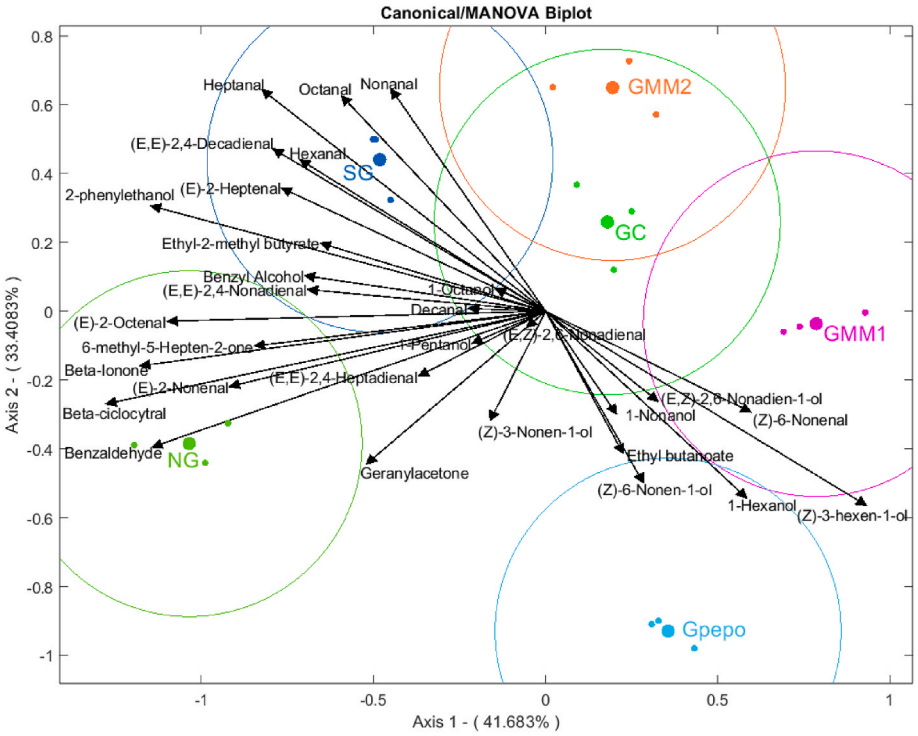

Fig. 1. MANOVA biplot of the accumulation of volatiles in watermelon fruits (*Citrullus lanatus* F1 Oneida) grown ungrafted (NG), self-grafted (SG) and grafted on: experimental rootstock of *Cucurbita* (*C. maxima* x *C. moschata*) F1 hybrid (GMM1), *Cucurbita* (*C. maxima* x *C. moschata*) Cobalt commercial F1 hybrid (GMM2), Cucurbita (*C. pepo*) rootstock (GPepo) and citroides (*Citrullus lanatus* var. citroides) experimental rootstock (GC). Circles represent Bonferroni confidence intervals. The significance of differences among treatments is inferred when the projections of confidence intervals on each vector do not overlap. Variance (%) explained by each latent variable indicated in parenthesis in each axis.

parameters in each VOC model, an expression heatmap was calculated using the loadings for each e-nose sensor obtained in the PLS model developed to predict volatile accumulation. Euclidean distance and UPGMA were selected for clustering sensor loadings. The software heatmapper (http://www.heatmapper.ca) was used for this purpose.

3. Results and discussion

3.1. CG-MS volatile profile

Of the 61 analyzed volatiles, 30 had contents above the quantification limit. Major components included alcohols and aldehydes, with a lower accumulation of other classes, such as apocarotenoids derived from carotenoid degradation (Supplementary Fig. 1). This profile of VOC accumulation was similar to that described by other authors in watermelon (Beaulieu & Lea, 2006; Lewinsohn et al., 2005; Liu et al., 2012; Petropoulos et al., 2014).

The specific effect of grafting was analyzed with a MANOVA biplot in order to gain a general view of the complex profile of watermelon volatiles. Grafting *per se* did not alter the volatile profile, as fruits from non-grafted and self-grafted plants exhibited a similar profile (left side of Fig. 1). Nonetheless, using different rootstocks affected the volatile profile to a greater or lesser degree. In general, the effect of most of these rootstocks on the volatile profile was coherent with those previously described in Fredes et al. (2017).

Citroides rootstock (GC) outstood for providing a similar profile to the self-grafted control. On the other hand, *Cucurbita* rootstocks (GMM1, GMM2 and Gpepo) tended to increase C9 aldehydes contents but to a different degree, being most notable in the experimental line GMM1. At the same time, GMM2 offered a profile more similar to the self-grafted control (Fig. 1). Petropoulos et al. (2014) also reported this effect using *Cucurbita* rootstocks.

Cucurbita rootstocks also tended to increase some alcohols, among them (Z)-6-nonen-1-ol. An increase of (Z)-6-nonen-1-ol levels had already been described by Fredes et al. (2016), but this time it could be checked that this effect also applied when watermelons are grown on *Cucurbita pepo*. This compound is especially important as it confers pumpkin-like odors (Leffingwell & Associates, 2023). In fact, it is considered detrimental to fruit quality and frequently associated with overripe watermelons (Saftner, Luo, McEvoy, Abbott, & Vinyard, 2007).

Tripodi, Condurso, Cincotta, Merlino, and Verzera (2020) also found that different hybrid genotypes of *C. maxima* x *C moschata* affected the volatile profile of mini watermelons differently, and it is possible to identify materials with a lower impact. In that case, the use of *Lagenaria siceraria* rootstock had a major impact on the volatile profile with increased levels of (Z)-6-nonenal and lower amounts of hexanal, (E,Z)-2, 6-nonadienal and (E,Z)-2,4-nonadienal. Interestingly, the Cucurbita hybrid Shintosa and the *Lagenaria* rootstock resulted in higher levels of pumpkin-like flavors. Nonetheless, it should be considered that Guler, Candir, Yetisir, Karaca, and Solmaz (2014) showed that a high variation exists in the impact of different rootstocks on the volatile profile.

3.2. E-nose volatile profile evaluation

Independently of the specific effect of rootstock on the volatile profile, it was clear that differences in the volatile profile were detected among the studied rootstocks. Thus, the sample set represented an ideal opportunity to test the ability of an e-nose system to discriminate between treatments. A MANOVA biplot integrating two types of sensor responses showed that, in general, normalized maximum intensity and sensor response slope of each sensor had a negative correlation, basically confirming that samples could be separated attending any of the two groups of sensor responses. Most variables contributed to the separation of non-grafted and self-grafted controls and the use of citron rootstock from *Cucurbita* rootstocks, while the variables related to the sensors T40/1, T30/1, and LY2/LG tended to separate the different

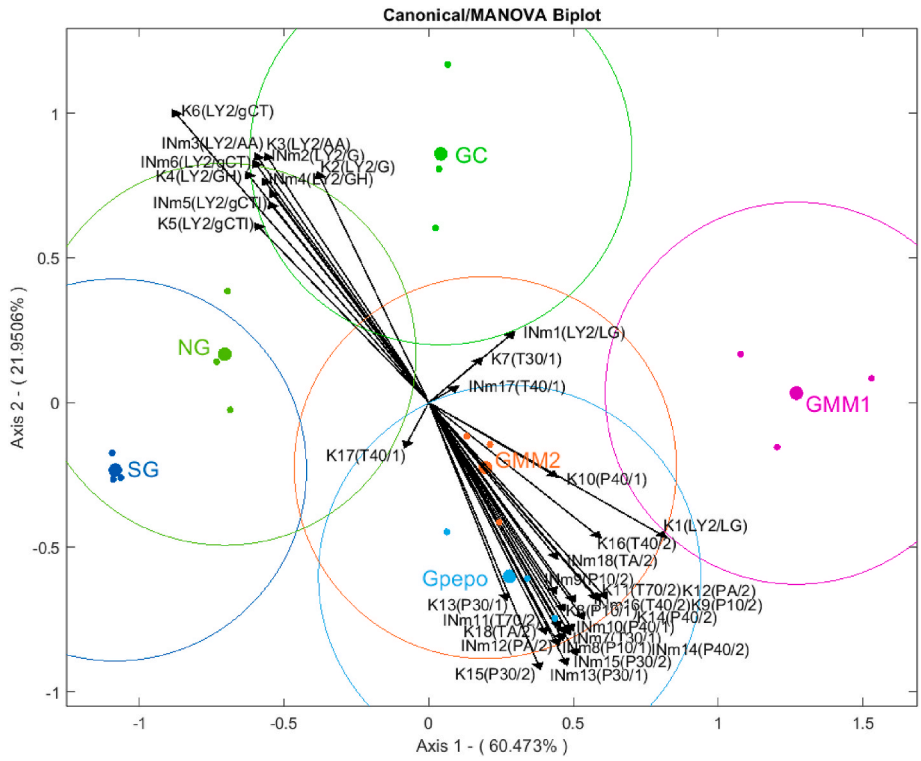

Fig. 2. MANOVA biplot of the sensor responses (INm: normalized maximum intensity; K: slope of sensor response) of the e-nose system analyzing watermelon volatile profiles of *Citrullus lanatus* F1 Oneida grown ungrafted (NG), self-grafted (SG) and grafted on: experimental pattern of *Cucurbita* (*C. maxima* x *C. moschata*) F1 hybrid (GMM1), *Cucurbita* (*C. maxima* x *C. moschata*) Cobalt commercial F1 hybrid (GMM2), *Cucurbita pepo* rootstock (GPepo) and citroides (*Citrullus lanatus* var. *citroides*) experimental rootstock (GC). Circles represent Bonferroni confidence intervals. The significance of differences among treatments is inferred when the projections of confidence intervals on each vector do not overlap. Variance (%) explained by each latent variable indicated in parenthesis in each axis.

Table 2
Performance of PLS-DA classification models in calibration and cross-validation of samples analyzed by gas chromatography and e-Nose. NG: Non-grafted control *Citrullus lanatus* F1 Oneida, SG: self-grafted control (SG), GMM1 grafted on: experimental pattern of Cucurbita (*C. maxima* x *C. moschata*) F1 hybrid, GMM2 Cucurbita (*C. maxima* x *C. moschata*) commercial F1 hybrid Cobalt (GMM2), Gpepo: *C. pepo* rootstock, GC: citroides (*Citrullus lanatus* var. citroides) experimental rootstock.

		GMM1	GMM2	GPepo	GC	SG	NG
Gas chromatography data	**Sensitivity**						
	Calibration	1.00	1.00	0.67	1.00	1.00	1.00
	Cross-validation	0.67	0.00	0.33	0.67	0.67	0.33
	Specificity						
	Calibration	0.87	0.33	0.87	0.53	0.60	0.87
	Cross-validation	0.87	0.60	0.87	0.60	0.60	0.73
	Accuracy						
	Calibration	0.89	0.22	0.56	0.72	0.5	0.89
	Cross-validation	0.83	0.28	0.61	0.78	0.67	0.78
	Precision						
	Calibration	0.6	0.13	0.22	0.25	0.2	0.6
	Cross-validation	0.5	0.08	0.25	0.33	0.33	0.4
E-nose data	**Sensitivity**						
	Calibration	1.00	0.67	0.67	1.00	1.00	1.00
	Cross-validation	0.33	0.00	0.67	0.67	0.67	0.33
	Specificity						
	Calibration	0.87	0.73	0.73	0.47	0.80	0.73
	Cross-validation	0.80	0.80	0.73	0.47	0.73	0.80
	Accuracy						
	Calibration	0.89	0.71	0.5	0.72	0.56	0.78
	Cross-validation	0.72	0.75	0.61	0.83	0.61	0.72
	Precision						
	Calibration	0.6	0.33	0	0.33	0.3	0.43
	Cross-validation	0.25	0.4	0.17	0.5	0.33	0.33

treatments within each of these two groups (Fig. 2). In essence, the MANOVA biplot with e-nose data was coherent with the analysis performed with chromatographic data. Accordingly, little differences were found between the sensor profile of the non-grafted and self-grafted controls. The use of rootstocks citroides (GC) and GMM2 minimized the influence of grafting (groups close to NG group), while GMM1 and Gpepo had a major impact on the volatile profile.

In order to check the performance in sample classification, a discriminant classification analysis based on partial least squares-discriminant analysis (PLS-DA) was conducted (Table 2). In our case,

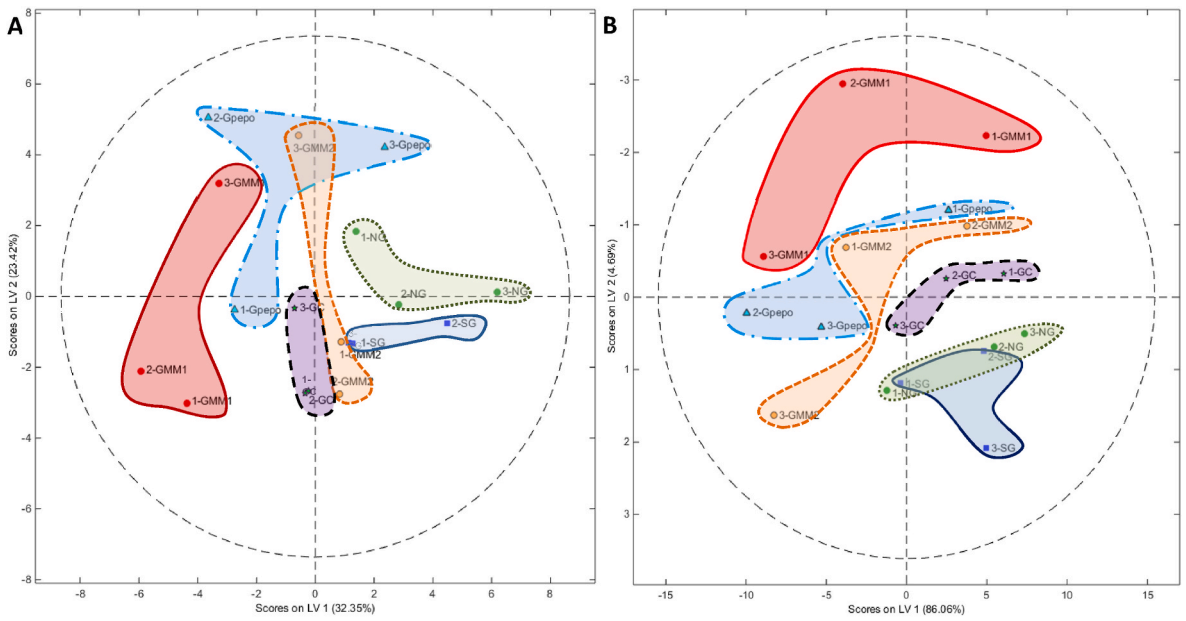

Fig. 3. PLS-DA plots of the volatile profile obtained with gas chromatography data (**A**) and e-nose data (**B**). Scion: *Citrullus lanatus* F1 Oneida. NG: Non-grafted control, SG: self-grafted control (SG), GMM1 grafted on: experimental pattern of Cucurbita (*C. maxima* x *C. moschata*) F1 hybrid, GMM2 Cucurbita (*C. maxima* x *C. moschata*) commercial F1 hybrid Cobalt (GMM2), Gpepo: *C. pepo* rootstock, GC: citroides (*Citrullus lanatus* var. citroides) experimental rootstock.

the 30 original variables (quantified volatiles) were reduced to two underlying discriminant variables that explained 55.8% of the variability in volatile content.

The system's performance for classification was moderate, regardless of the type of data used. Using gas chromatography data, the sensitivity obtained during the calibration of the model was perfect except for GMM2 (Table 2). The specificity was also good for all the treatments but for GMM2 and GC. Accordingly, good accuracy levels were obtained except for these rootstocks. In the case of e-nose data, sensitivity reached 1.0 except for GMM2 and Gpepo, while specificity was rather good except for GC. Accuracy reached reasonable values for GMM1, GMM2, GC and NG, being lower in the rest of the conditions. On the other hand, the precision of the model was low for most of the scion-rootstock combinations. When the model was cross-validated, the performance was, in general, similar to the one obtained during the calibration of the model, with isolated drops in specific cases.

One of the factors that justify the moderate performance obtained in classification is the high variability observed within each scion-rootstock combination for both gas chromatography and e-nose data even after PLS-DA (Fig. 3A and 3B, respectively). Regarding chromatographic data, non-grafted and self-grafted samples plotted in the PLS-DA biplot together and close to GMM2 and GC. GMM1 and Gpepo plotted at a certain distance, but one of the samples of GMM1 and GMM2 potted close to the Gpepo group, while one of the samples of Gpepo plotted between the GMM1 and the GC and GMM2 groups. A similar clustering was obtained with the e-nose data. The same samples of GMM1 and GMM2 plotted close to the group of the Gpepo, while one of the samples of Gpepo plotted between the GMM1 and the GMM2 and GC group. The same samples of GMM1 and GMM2 that plotted close to the Gpepo group with chromatographic data also plotted in the same way with the e-nose data. In essence, the PLS-DA plot representation of e-nose data proves to be a clear and effective method for identifying subtle changes in volatile profiles, similar to chromatographic data.

This variability cannot be attributed to genotypic effects, as a commercial F1 hybrid cultivar (Oneida) thus assuring genetic uniformity. Therefore, it may be attributable to micro-environmental effects or differences in each plant's development stage. In the case of watermelon, it is rather difficult to establish the best date for harvesting. In the case of grafted plants, it is even more challenging, as maturity can be delayed by grating (Fallik & Ziv, 2020). This does not seem to be the case, as previously published results confirmed that no differences were observed in fruit firmness with these rootstocks, a trait usually linked to the maturity stage. Nonetheless, finding differences in maturity between plants within the same treatments is quite frequent. This effect is evident in different works dealing with watermelon. For example, in the principal component analysis performed by Kyriacou, Soteriou, Rouphael, Siomos, and Gerasopoulos (2016) with fruits from self-grafted plants and plants grafted on a Cucurbita hybrid harvested at different dates, a high dispersion can be found between samples of the same condition.

Indeed, this is probably one of the main reasons for the variability of the volatile profile within each scion-rootstock combination. This effect did not apply equally to all the conditions analyzed, as non-grafted, self-grafted, and grafting on GC resulted in a much higher uniformity of the volatile profile (Fig. 3A and B).

Although it has been reported that extraction procedures can introduce variability in the volatile profiles of the same samples (Rambla et al., 2015), it does not seem to be the case, as a similar variability is observed in data from GC-MS and e-nose, while the extraction procedure in the second case is straightforward and reproducible.

It seems clear that, although grafting exerts an evident influence on the volatile profile, the high variability found among plants hinders the identification of apparent differences between different scion-rootstock combinations. The moderate results obtained in classification methods using e-nose data are not a result of a low discriminant power of the system, as similar results were obtained with chromatographic data. This situation differs from classification attempts made in melon, where chromatographic data clearly separated melon genotypes while the e-nose system provided lower discrimination power (Chaparro-Torres et al., 2016).

To our knowledge, few groups have applied e-nose to evaluate the volatile profile of watermelon. Bianchi et al. (2020) characterized the volatile profile of two watermelon cultivars harvested at different developmental stages (unripe, ripe and overripe). Both cultivars had a completely different volatile profile, with 'Rugby' presenting fivefold higher amounts of volatiles than 'Cuoredolce'®. Additionally, 'Ruby' had a high level of apocarotenoids. With such differences, the e-nose system could distinguish both cultivars using PCA analysis, but it was unable to differentiate samples of 'Cuoredolce'® at different

Table 3
Coefficients (a_i) of the PLS prediction models[a] for volatile contents (ng g^{-1}) obtained with e-nose sensors response showing a moderate to high performance.

E-nose sensor	(Z)-6-Nonen-1-ol	Geranylacetone	(Z)-6-Nonenal	1-Hexanol	Heptanal	(Z)-3-Nonen-1-ol	Ethyl butanoate	(Z)-3-hexen-1-ol	1-Nonanol	Hexanal	6-methyl-5-Hepten-2-one	(E,E)-2,4-Heptadienal	Decanal	(E,Z)-2,6-Nonadienal
Bo	632.9	−18.2	−3069.6	−227.4	−75.1	−13854.0	−493.5	710.5	−1654.7	−1297.4	−1583.8	−90.8	36.6	−2515.6
INm1(LY2/LG)	275.3	−3701.9	2644.5	−293.1	−46.2	−12341.7	−459.5			−299.5	−602.8			3692.3
INm2(LY2/G)			170.8		4.2								−0.4	
INm3(LY2/AA)	−34.0		141.0											
INm4(LY2/GH)					2.7			−68.1					0.2	
INm5(LY2/gCTl)	−63.5						−5.9					−2.3	0.9	
INm6(LY2/gCT)	−258.1				8.9		−158.2			180.1			−0.9	8071.7
INm7(T30/1)	−21.7													
INm8(P10/1)	−36.6												−8.8	
INm9(P10/2)							−151.1							
INm10(P40/1)	−36.6													
INm11(T70/2)	38.6											21.3		
INm12(PA/2)	−27.0												−5.2	
INm13(P30/1)	−199.5													
INm14(P40/2)	−34.5		1550.7										−9.5	
INm15(P30/2)	−196.8		1701.3				−232.3			−151.0			−20.2	
INm16(T40/2)	55.1		2139.3					289.8					−15.9	
INm17(T40/1)	−2718.9	8011.0								3535.0	5444.6	98.7		
INm18(TA/2)		−6589.4			142.5	23988.5	1125.9	−1268.7			−1742.0			
K1(LY2/LG)		141759.0			2764.2	−39656.2	14091.7							
K2(LY2/G)											−1250.9			11577.4
K3(LY2/AA)	−36.2	−696.0	1324.4							154.4	−1305.2			
K4(LY2/GH)			2383.3								1963.0	34.5		
K5(LY2/gCTl)	−180.8		−1036.4		−48.7					232.1	1097.0		−42.2	
K6(LY2/gCT)	−2205.7	10300.1	9922.7		123.2		−1827.9					310.7	−7.1	
K7(T30/1)			−6947.7							−1973.9			−288.5	
K8(P10/1)	−3080.4		−7633.6					−699.1		−75.0			−260.7	
K9(P10/2)		−14813.3	−15146.7						14761.6	−2539.0	−8919.4	−493.8		
K10(P40/1)	−4381.4						−1670.7	−2125.1		124.5	−567.1	209.1		−7892.9
K11(T70/2)	2323.7				−231.2			1879.5		−1089.9			139.0	
K12(PA/2)		−12780.2	9229.8		−68.3			1444.8						
K13(P30/1)	4010.6						−2469.8	−1320.3	7745.2			−206.2		
K14(P40/2)	4204.5	38221.1	18198.3		150.4			4315.6						68299.8
K15(P30/2)			−4885.9	2817.4		37590.1					2284.4	256.7	−279.3	
K16(T40/2)	4094.7		7185.5				−2419.8						690.2	79394.7
K17(T40/1)	24678.1						16338.7			−4310.9			−106.6	−135214.0
K18(TA/2)	13607.3	13667.7		6774.2		89618.5	1157.8		19029.0			790.0	900.9	70159.3
R^2 Cal [b]	0.980	0.974	0.953	0.926	0.899	0.886	0.881	0.875	0.875	0.872	0.856	0.855	0.844	0.796
R^2 CV [c]	0.937	0.827	0.769	0.864	0.721	0.787	0.625	0.735	0.852	0.786	0.724	0.412	0.665	0.646
RMSECV [d]	16.79	24.49	28.17	9.62	0.37	264.00	8.94	7.01	40.40	14.32	11.28	0.80	0.74	79.80
%RMSECV[e]	13%	17%	16%	11%	10%	14%	16%	15%	14%	12%	20%	22%	14%	21%

[a] Prediction equations: $Y_{compound} = B_0 + a_1IN_{máx}1 + a_2IN_{máx}2 + \ldots + a_{18}IN_{máx}18 + b_1K1 + b_2K2 + \ldots + b_{18}K_{18}$, Where $IN_{max}i$ is the maximum relative intensity, Ki is the slope for each sensor, a_1-a_{18} coefficients of the equation showing the best fit.
[b] R^2_Cal: regression coefficients of calibration models.
[c] R^2_CV: regression coefficients of cross-validation models.
[d] RMSECV: Root Mean Squared Standard Error for Cross-Validation (ng g^{-1}).
[e] %RMSECV: RMSECV scaled to the maximum value of the compound being considered (%).

Table 4
Coefficients (a_i) of the PLS prediction models[a] for volatile contents (ng g^{-1}) obtained with e-nose sensors response showing a moderate to low performance.

E-nose sensor	(E,Z)-2,6-Nonadien-1-ol	(E)-2-Octenal	Beta-ciclocytral	Beta-Ionone	Nonanal	1-Octanol	2-phenylethanol	Octanal	Benzyl Alcohol	1-Pentanol	(E,E)-2,4-Decadienal	(E)-2-Nonenal	(E)-2-Heptenal	(E,E)-2,4-Nonadienal	Benzaldehyde	Ethyl-2-methyl butyrate
Bo	2055.4	−105.6	−6.6	283.5	3110.8	−170.6	9.0	80.9	69.6	−229.8	28.7	−168.3	1.3	8.6	20.2	580.1
INm1(LY2/LG)		−72.6	−14.6	−509.5		−59.1			−123.9	−141.5	−64.1			−32.0	−38.0	−677.4
INm2(LY2/G)			0.3								1.4			0.3		
INm3(LY2/AA)			0.3								1.1			0.3		
INm4(LY2/GH)			0.3	148.7							1.1			0.4		
INm5(LY2/gCTl)				200.0	220.0								3.5	0.5		
INm6(LY2/gCT)			1.2										11.5	1.4		
INm7(T30/1)											1.8					
INm8(P10/1)			0.2	−0.3							1.0			2.4		
INm9(P10/2)			−1.2					−69.2			−8.6			3.4		
INm10(P40/1)			0.5								2.8			4.6		
INm11(T70/2)			0.0	−0.8							0.0			1.0		
INm12(PA/2)			0.1	−0.3							0.7			1.5		
INm13(P30/1)		−11.3	0.1								0.9			2.6		
INm14(P40/2)		−12.6	−0.5	−264.3							−1.5			0.7		
INm15(P30/2)		−13.2	−0.3	−201.7							0.1			0.8		−169.6
INm16(T40/2)	334.3		−1.7	−827.3							−6.3					
INm17(T40/1)		297.1	23.9			271.4	−1.2			536.9			13.7			
INm18(TA/2)	−3032.7				−2630.4											
K1(LY2/LG)	−17997.3						−560.4									
K2(LY2/G)											5.4			0.4		
K3(LY2/AA)											3.0					
K4(LY2/GH)		30.7	3.0	154.6							11.4			9.4		
K5(LY2/gCTl)	−961.8		2.0	879.6							0.2		18.3	2.8		
K6(LY2/gCT)								119.0					57.4			
K7(T30/1)					−4237.0		−55.0	−164.6								
K8(P10/1)														22.3		
K9(P10/2)																
K10(P40/1)	−3850.1	−40.8									77.4	−29051.0				
K11(T70/2)	5523.8	−50.0									0.0			−9.3		
K12(PA/2)											40.7	6682.6				
K13(P30/1)			11.3	382.3		318.1		−154.5								
K14(P40/2)			8.0	488.9							−20.8	−22196.6		−59.1		
K15(P30/2)	−6924.0				−4529.2			−175.3			−9.5					−183.5
K16(T40/2)	11207.1	−105.1									19.2			−50.8		
K17(T40/1)					−21858.2											
K18(TA/2)		112.0	32.6	114.9		784.3					103.7	79416.6				119.1
R^2 Cal [b]	0.725	0.702	0.642	0.527	0.510	0.496	0.448	0.444	0.402	0.382	0.375	0.360	0.294	0.271	0.254	0.225
R^2 CV [c]	0.422	0.466	0.330	0.335	0.379	0.370	0.247	0.349	0.246	0.161	0.111	0.119	0.065	0.019	0.141	0.010
RMSECV [d]	10.18	1.79	0.32	1.18	84.23	4.16	0.44	3.81	2.98	5.20	2.05	170.56	2.19	1.25	1.21	3.24
%RMSECV [e]	28%	25%	27%	29%	30%	21%	32%	41%	18%	16%	48%	30%	37%	42%	35%	43%

[a] Prediction equations: $Y_{compound} = B_0 + a_1IN_{máx}1 + a_2IN_{máx}2 + \ldots + a_{18}IN_{máx}18 + b_1K1 + b_2K2 + \ldots + b_{18}K_{18}$, Where $IN_{max}i$ is the maximum relative intensity, Ki is the slope for each sensor, a_1-a_{18} coefficients of the equation showing the best fit.
[b] R^2_Cal: regression coefficients of calibration models.
[c] R^2_CV: regression coefficients of cross-validation models.
[d] RMSECV: Root Mean Squared Standard Error for Cross-Validation (ng g^{-1}).
[e] %RMSECV: RMSECV scaled to the maximum value of the compound being considered (%).

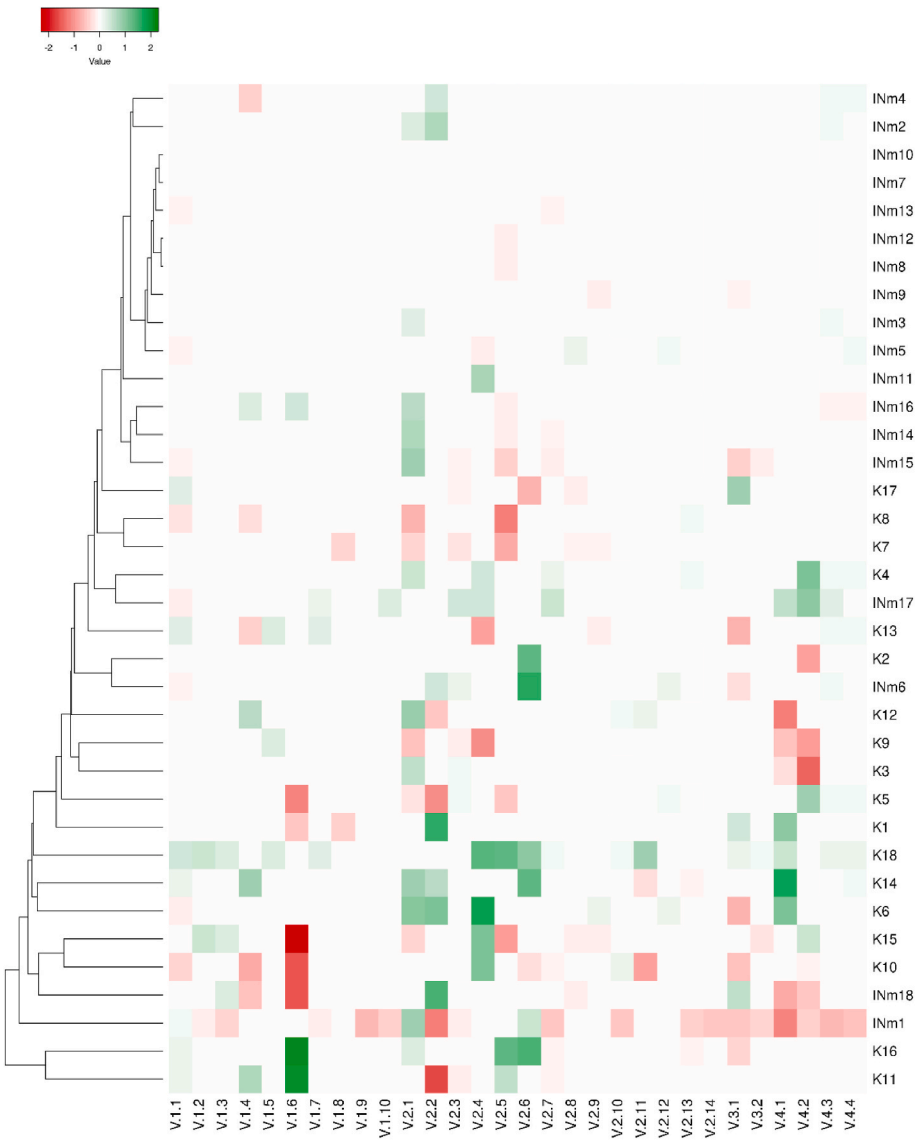

Fig. 4. Heatmap of loadings values of PLS models for the prediction of the accumulation of alcohols (V.1), aldehydes (V.2), esters (V.3) and apocarotenoids (V.4) using e-nose sensors datasensor (INm: normalized maximum intensity; K: slope of sensor response). Higher absolute values denote a higher weight of the specific sensor in the prediction model. V.1-1: (Z)-6-Nonen-1-ol: V.1-2: 1-Hexanol: V.1-3: (Z)-3-Nonen-1-ol: V.1-4: (Z)-3-hexen-1-ol: V.1-5: 1-Nonanol: V.1-6: (E,Z)-2,6-Nonadien-1-ol: V.1-7: 1-Octanol: V.1-8: 2-phenylethanol: V.1-9: Benzyl Alcohol: V.1-10: 1-Pentanol: V.4–1: Geranylacetone: V.4–2: 6-methyl-5-Hepten-2-one: V.4–3: Beta-ciclocytral: V.4-4: Beta-Ionone: V.2–1: (Z)-6-Nonenal: V.2-2: Heptanal: V.2-3: Hexanal: V.2-4: (E,E)-2,4-Heptadienal: V.2-5: Decanal: V.2-6: (E,Z)-2,6-Nonadienal: V.2-7: (E)-2-Octenal: V.2-8: Nonanal: V.2-9: Octanal: V.2-10: (E,E)-2,4-Decadienal: V.2-11: (E)-2-Nonenal: V.2-12: (E)-2-Heptenal: V.2-13: (E,E)-2,4-Nonadienal: V.2-14: Benzaldehyde: V.3–1: Ethyl butanoate: V.3–2: Ethyl-2-methyl butyrate. Sensors: 1: LY2/LG; 2: LY2/G; 3: LY2/AA; 4: LY2/GH; 5: 2/gCTl; 6: Y2/gCT; 7: T30/1; 8: P10/1; 9: P10/2; 10: P40/1; 11: T70/2; 12: PA/2; 13: P30/1; 14: P40/2; 15: P30/2; 16: T40/2; 17: T40/1; 18: TA/2.

developmental stages. In the case of 'Rugby', though, it could differentiate unripe simples from ripe and over-ripe, but the last two plotted close together.

The evident correlation between PLS-DA plots obtained with chromatographic and e-nose data opens a high value of e-nose as high-throughput volatile profile evaluation systems. An example of complex task which can be faced with this system is the germplasm evaluation. Indeed, one of the challenges in germplasm evaluation is to get an idea of the volatile profile with a minimum investment, operational cost and time requirements. For example, one of the obvious applications would be the development of breeding programs. In such programs, evaluating a high number of genotypes in a short time and at reasonable costs is necessary. In these cases, sensory analysis with panelists is only feasible for selecting a limited number of finalist lines. GC-MS opens the possibility of overcoming the limitations of sensory evaluation in terms of cost and time. However, these analytic procedures are not affordable in terms of cost and time, with a high number of samples. As an affordable alternative, using e-nose systems with simple and rapid extraction procedures opens new possibilities for aroma evaluation.

It must be remembered that the lack of evaluation of volatiles in breeding programs has resulted in dramatic quality losses in different crops. Tomato is one of the most evident case. It has been proved that the lack of evaluation of the aroma profile in tomato breeding programs and the excessive focus placed on yield led to the loss of alleles related to the accumulation of volatiles, which not only offered positive odour notes but also potentiated the perception of sweetness (Tieman et al., 2017). Consequently, the necessity to evaluate the volatile profile and the presence of critical alleles in the development of breeding programs in order to develop new flavourful varieties has been stressed (Kaur, Abugu, & Tieman, 2023). This problem might be extrapolated to

Cucurbits. For example, in the case of melon, the evaluation of introgression lines proved that the lightest alteration of the genetic background entails a dramatic effect on the accumulation of key volatiles (Perpiñá et al., 2021).

Two-dimensional PLS-DA representation as a measure of similarity of the electronic fingerprints of the volatile profile has already been proposed to evaluate tomato germplasm (Valcárcel et al., 2021). In the case of watermelon, it seems to provide an efficient evaluation system, as the evaluation of e-nose data enables the detection of similarities and differences of sample groups. In our case, it clearly evaluated the possible effects of grafting and scion-rootstock interaction in similar materials. Despite using a single hybrid F1 commercial cultivar as scion, the system was capable to detect subtle differences, offering the same conclusions obtained with GC-MS. The system's efficiency surpassed our expectations, as the PLS-DA representation of e-nose data successfully identified anomalous sample profiles within the same scion-rootstock combination as detected with GC-MS data. In fact, it was even capable of detecting which scion-rootstock combinations presented higher within-group variability (e.g. GMM1 vs SG).

3.3. E-nose prediction models for specific volatile accumulation

The consistent results obtained with GC-MS and e-nose data opened the possibility of predicting specific compound accumulation using e-nose fingerprints. For that purpose, partial least squares regressions (PLS) followed by hierarchical variable selections (i-PLS) were used to develop models for the indirect prediction of compound accumulation.

The prediction models obtained for (Z)-6-nonen-1-ol, geranyl acetone, (Z)-6-nonenal, 1-hexanol, and heptanal were particularly good (Table 3), with regression coefficients for linear calibration models above 0.89. Even though a low number of samples were used, the cross-validated models also showed a good fit, with linear regression coefficients above 0.72 and low cross-validated errors (RMSECV). In general, RMSEC values for these models ranged between 10% and 17% of the maximum content of the compound being considered. That would imply that e-nose would be a reliable system for the indirect preliminary evaluation of compounds, such as (Z)-6-nonen-1-ol, with a major impact on watermelon aroma perception. Even more, expanding the model with more samples would probably lead to better model performance. Undoubtedly, broadening the scope of evaluation to include a greater diversity of scions, rootstocks, environments, and cultivation methods will significantly bolster the development of more robust and comprehensive models, thereby augmenting their overall generalizability.

Other compounds such as (Z)-3-nonen-1-ol, ethyl butanoate, (Z)-3-hexen-1-ol, 1-nonanol, hexanal, 6-methyl-5-hepten-2-one, (E,E)-2,4-heptadienal, decanal, and (E,Z)-2,6-nonadienal showed R^2 values ranging between 0.70 and 0.88 in calibration and 0.41 to 0.85 in cross-validation (Table 4). These models, although less precise than the previous ones, could be used as exploratory semi-quantitative models for these compounds. In the remaining cases, the calibration R^2 values were below 0.75, and the cross-validation R^2 values were below 0.47 (Table 4). The lower fit of these prediction models may be due to the fact that the electronic sensors used are less sensitive to some compounds (Song et al., 2010). Therefore, using them for prediction purposes would not be prudent because the prediction errors would be too high. In fact, %RMSECV values ranged from 21% to 42%.

In general, the use of slope (k values) for the development of prediction models was more informative than the normalized maximum intensity (INm), as the loadings obtained in the development of the model presented high absolute values (Fig. 4). In fact, most INm sensor loadings cluster together except for LY2/LG, TA/2, Y2/gCT and T40/1. Interestingly, as previously stated, the sensors T40/1, and LY2/LG tended to separate the different treatments within each of these two groups in the MANOVA biplot (Fig. 2).

E-nose systems had been previously used for fast indirect evaluation of the accumulation of specific volatiles. It would be the case of specific alkaloids such as nicotine in tobacco (Lin & Zhang, 2016), ethanol in beers (Voss, Mendes Júnior, Farinelli, & Stevan, 2019) or linalool in *Osmanthus fragrans* (C. Zhou et al., 2022). In our case, the direct evaluation of the volatile profile of watermelon would not only offer insights into proximity to ideal volatile profiles, but it would also be reasonably efficient in the selection for increased levels of apocarotenoids, such as geranylacetone, exerting positive aroma and taste influences, or the negative selection for detrimental compounds such as (Z)-6-nonen-1-ol, offering negative pumpkin tinges (Kyriacou, Leskovar, Colla, & Rouphael, 2018), or Z-6-nonenal and E,Z-2,6-nonadienal, associated with greenish notes (Beaulieu & Lea, 2006).

4. Conclusion

The evaluation of the volatile profile using e-nose systems avoids complex volatile extraction procedures and reduces time analysis and operational costs compared to GC-MS alternatives. At the same time, PLS-DA analysis of e-nose data efficiently identifies subtle volatile profile changes as those induced by the different rootstock combinations with the same scion. It not only differentiates treatments in comparison with the ungrafted control, but it also addresses the level of variability within group and detects specific sample profiles obtained within the same rootstock-scion combination. Furthermore, PLS models have been developed for the indirect quantification of major volatiles with a known impact on the aroma of watermelon, with reasonable efficiency. Indirect analysis with e-nose systems will never be as efficient as a sensory panel evaluation of GC-MS quantification, but it can be operated as a high-throughput system, enabling the evaluation of the volatile profile in large scale sample management tasks as needed in breeding programs for new cultivars and rootstocks, in the assessment of the effects of different crop management procedures and in commercial watermelon quality control chains.

CRediT authorship contribution statement

Alejandro Fredes: Writing – review & editing, Visualization, Investigation, Formal analysis. **Jaime Cebolla-Cornejo:** Writing – review & editing, Writing – original draft, Visualization, Formal analysis. **Joaquín Beltrán:** Writing – review & editing, Investigation. **Carmina Gisbert:** Writing – review & editing, Resources. **Belén Picó:** Writing – review & editing, Supervision, Project administration, Investigation. **Mercedes Valcárcel:** Writing – review & editing, Investigation, Formal analysis, Conceptualization. **Salvador Roselló:** Writing – review & editing, Visualization, Supervision, Project administration, Formal analysis, Conceptualization.

Declaration of competing 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.

Data availability

10.5281/zenodo.8191669

Acknowledgments

This research has been funded by MCIN/AEI/10.13039/501100011033/grant AGL2014-53398-C2-2-R, by MCIN/AEI/10.13039/501100011033/and by "ERDF: A way of making Europe" grant AGL2017-85563-C2-1-R-AR, by MCIN/AEI/10.13039/501100011033/grant PID2020-116055RB-C21, and by grant PROMETEO/2021/072 (to promote excellence groups) funded by Conselleria de Educación, Universidades y Empleo (Generalitat Valenciana, Spain). A. Fredes acknowledges a scholarship from the "Santiago Grisolía/

2013/032″ program funded by Generalitat Valenciana, as well as the scholarship funded by Banco Santander through the "Scholarships for Doctoral Studies for Students from Latin America" program.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.lwt.2024.116337.

References

Beaulieu, J. C., & Lea, J. M. (2006). Characterization and semiquantitative analysis of volatiles in seedless watermelon varieties using solid-phase microextraction. *Journal of Agricultural and Food Chemistry, 54*(20), 7789–7793. https://doi.org/10.1021/jf060663l

Benady, M., Simon, J. E., Charles, D. J., & Miles, G. E. (1995). Fruit ripeness determination by electronic sensing of aromatic volatiles. *Transactions of the ASAE, 38*(1), 251–257. https://doi.org/10.13031/2013.27837

Bianchi, G., Provenzi, L., & Rizzolo, A. (2020). Evolution of volatile compounds in 'Cuoredolce®' and 'Rugby' mini- watermelons (Citrullus lanatus (Thunb.) Matsumura and Nakai) in relation to ripening at harvest. *Journal of the Science of Food and Agriculture, 100*(3), 945–952. https://doi.org/10.1002/jsfa.10023

Brereton, R. G., & Lloyd, G. R. (2014). Partial least squares discriminant analysis: Taking the magic away. *Journal of Chemometrics, 28*(4), 213–225. https://doi.org/10.1002/cem.2609

Chaparro-Torres, L. A., Bueso, M. C., & Fernández-Trujillo, J. P. (2016). Aroma volatiles obtained at harvest by HS-SPME/GC-MS and INDEX/MS-E-nose fingerprint discriminate climacteric behaviour in melon fruit. *Journal of the Science of Food and Agriculture, 96*(7), 2352–2365. https://doi.org/10.1002/jsfa.7350

Fallik, E., & Ziv, C. (2020). How rootstock/scion combinations affect watermelon fruit quality after harvest? *Journal of the Science of Food and Agriculture, 100*(8), 3275–3282. https://doi.org/10.1002/jsfa.10325

Feng, L., Zhang, M., Bhandari, B., & Guo, Z. (2018). Determination of postharvest quality of cucumbers using nuclear magnetic resonance and electronic nose combined with chemometric methods. *Food and Bioprocess Technology, 11*(12), 2142–2152. https://doi.org/10.1007/s11947-018-2171-9

Fredes, A., Roselló, S., Beltrán, J., Cebolla-Cornejo, J., Pérez-de-Castro, A., Gisbert, C., et al. (2017). Fruit quality assessment of watermelons grafted onto citron melon rootstock. *Journal of the Science of Food and Agriculture, 97*(5). https://doi.org/10.1002/jsfa.7915

Fredes, A., Sales, C., Barreda, M., Valcárcel, M., Roselló, S., & Beltrán, J. (2016). Quantification of prominent volatile compounds responsible for muskmelon and watermelon aroma by purge and trap extraction followed by gas chromatography-mass spectrometry determination. *Food Chemistry, 190*. https://doi.org/10.1016/j.foodchem.2015.06.011

Gao, L., Zhao, S., Lu, X., He, N., & Liu, W. (2018). 'SW', a new watermelon cultivar with a sweet and sour flavor. *HortScience, 53*(6), 895–896. https://doi.org/10.21273/HORTSCI12857-18

García-González, D. L., & Aparicio, R. (2002). Sensors: From biosensors to the electronic nose. *Grasas y Aceites Sevilla, 53*(1), 96–114.

Gardner, J. W., & Bartlett, P. N. (1999). Electronic noses. Principles and applications. *Measurement Science and Technology, 11*(7), 1087. Retrieved from http://stacks.iop.org/0957-0233/11/i=7/a=702.

Gardner, J. W., & Persaud, K. C. (Eds.). (2001). Electronic noses and olfaction 2000. *Measurement Science and Technology, 12*(11), 2019. https://doi.org/10.1088/0957-0233/12/11/702

Gisbert, C., Gammoudi, N., Munera, M., GinÃ©, A., Pocurull, M., Sorribas, F. J., et al. (2017). Evaluation of two potential Cucumis spp. resources for grafting melons. In *Acta horticulturae* (pp. 157–162). Leuven, Belgium: International Society for Horticultural Science (ISHS). https://doi.org/10.17660/ActaHortic.2017.1151.25.

Guler, Z., Candir, E., Yetisir, H., Karaca, F., & Solmaz, I. (2014). Volatile organic compounds in watermelon (Citrullus lanatus) grafted onto 21 local and two commercial bottle gourd (Lagenaria siceraria) rootstocks. *The Journal of Horticultural Science and Biotechnology, 89*(4), 448–452. https://doi.org/10.1080/14620316.2014.11513105

Hernández, A., Wang, J., Hu, G., & Pereira, A. G. (2008). Monitoring storage shelf life of tomato using electronic nose technique. *Journal of Food Engineering, 85*(4), 625–631. https://doi.org/10.1016/j.jfoodeng.2007.06.039

Holmberg, M., & Artursson, T. (2003). Drift compensation, standards, and calibration methods. In T. C. Pearce, S. S. Schiffman, H. T. Nagle, & J. W. Gardner (Eds.), *Handbook of machine olfaction* (pp. 325–346). Wiley-VCH Verlag GmbH & Co. KGaA. https://doi.org/10.1002/3527601597.ch13.

Hubert, M., & Branden, K. V. (2003). Robust methods for partial least squares regression. *Journal of Chemometrics, 17*(10), 537–549. https://doi.org/10.1002/cem.822

Jackson, J. E., & Mudholkar, G. S. (1979). Control procedures for residuals associated with principal component analysis. *Technometrics, 47*, 64–79.

Junxing, L. I., Aiqing, M., Gangjun, Z. H. A. O., Xiaoxi, L., Haibin, W., Jianning, L., et al. (2022). Assessment of the 'taro-like' aroma of pumpkin fruit (Cucurbita moschata D.) via E-nose, GC–MS and GC-O analysis. *Food Chemistry X, 15*(May), Article 100435. https://doi.org/10.1016/j.fochx.2022.100435

Kaur, G., Abugu, M., & Tieman, D. (2023). The dissection of tomato flavor: Biochemistry, genetics, and omics. *Frontiers in Plant Science, 14*(June), 1–18. https://doi.org/10.3389/fpls.2023.1144113

Kovats, E. (1958). Gas-chromatographische charakterisierung organischer verbindungen. Teil 1: retentionsindices aliphatischer halogenide, alkohole, aldehyde und ketone. *Helvetica Chimica Acta, 41*, 1915–1932. https://doi.org/10.1002/hlca.19580410703

Kyriacou, M. C., Leskovar, D. I., Colla, G., & Rouphael, Y. (2018). Watermelon and melon fruit quality: The genotypic and agro-environmental factors implicated. *Scientia Horticulturae, 234*(June 2017), 393–408. https://doi.org/10.1016/j.scienta.2018.01.032

Kyriacou, M. C., Soteriou, G. A., Rouphael, Y., Siomos, A. S., & Gerasopoulos, D. (2016). Configuration of watermelon fruit quality in response to rootstock-mediated harvest maturity and postharvest storage. *Journal of the Science of Food and Agriculture, 96*(7), 2400–2409. https://doi.org/10.1002/jsfa.7356

Leffingwell & Associates. (2023). Alkenols and molecular structures [Online]. *Odor properties & molecular visualization*. from http://www.leffingwell.com/alkenol.htm. (Accessed 15 November 2023).

Lewinsohn, E., Sitrit, Y., Bar, E., Azulay, Y., Meir, A., Zamir, D., et al. (2005). Carotenoid pigmentation affects the volatile composition of tomato and watermelon fruits, as revealed by comparative genetic analyses. *Journal of Agricultural and Food Chemistry, 53*(8), 3142–3148. https://doi.org/10.1021/jf047927t

Lin, S., & Zhang, X. (2016). A rapid and novel method for predicting nicotine alkaloids in tobacco through electronic nose and partial least-squares regression analysis. *Analytical Methods, 8*(7), 1609–1617. https://doi.org/10.1039/C5AY02257F

Liu, C., Zhang, H., Dai, Z., Liu, X., Liu, Y., Deng, X., et al. (2012). Volatile chemical and carotenoid profiles in watermelons [Citrullus vulgaris (Thunb.) Schrad (Cucurbitaceae)] with different flesh colors. *Food Science and Biotechnology, 21*(2), 531–541. https://doi.org/10.1007/s10068-012-0068-3

Perpiñá, G., Roselló, S., Esteras, C., Beltrán, J., Monforte, A. J., Cebolla-Cornejo, J., et al. (2021). Analysis of aroma-related volatile compounds affected by 'Ginsen Makuwa' genomic regions introgressed in 'Vedrantais' melon background. *Scientia Horticulturae, 276*. https://doi.org/10.1016/j.scienta.2020.109664 (October 2020).

Petropoulos, S. A., Olympios, C., Ropokis, A., Vlachou, G., Ntatsi, G., Paraskevopoulos, A., et al. (2014). Fruit volatiles, quality, and yield of watermelon as affected by grafting. *Journal of Agricultural Science and Technology A, 16*(4), 873–885.

Pino, J. A., Marbot, R., & Aguero, J. (2003). Volatile components of watermelon (Citrullus lanatus [Thunb.] Matsum. et Nakai) fruit. *Journal of Essential Oil Research, 15*(6), 379–380. https://doi.org/10.1080/10412905.2003.9698616

Rambla, J. L., Alfaro, C., Medina, A., Zarzo, M., Primo, J., & Granell, A. (2015). Tomato fruit volatile profiles are highly dependent on sample processing and capturing methods. *Metabolomics, 11*(6), 1708–1720. https://doi.org/10.1007/s11306-015-0824-5

Saftner, R., Luo, Y., McEvoy, J., Abbott, J. A., & Vinyard, B. (2007). Quality characteristics of fresh-cut watermelon slices from non-treated and 1-methylcyclopropene- and/or ethylene-treated whole fruit. *Postharvest Biology and Technology, 44*(1), 71–79. https://doi.org/10.1016/j.postharvbio.2006.11.002

Sinesio, F., Di Natale, C., Quaglia, G. B., Bucarelli, F. M., Moneta, E., Macagnano, A., et al. (2000). Use of electronic nose and trained sensory panel in the evaluation of tomato quality. *Journal of the Science of Food and Agriculture, 80*(1), 63–71. https://doi.org/10.1002/(SICI)1097-0010(20000101)80:1<63::AID-JSFA479>3.0.CO;2-8

Song, S., Zhang, X., Hayat, K., Jia, C., Xia, S., Zhong, F., … Niu, Y. (2010). Correlating chemical parameters of controlled oxidation tallow to gas chromatography-mass spectrometry profiles and e-nose responses using partial least squares regression analysis. *Sensors and Actuators, B: Chemical, 147*, 660–668. https://doi.org/10.1016/j.snb.2010.03.055

Tieman, D., Zhu, G., Resende, M. F. R., Lin, T., Nguyen, C., Bies, D., et al. (2017). A chemical genetic roadmap to improved tomato flavor. *Science (New York, N.Y.), 355*(6323), 391–394. https://doi.org/10.1126/science.aal1556

Tripodi, G., Condurso, C., Cincotta, F., Merlino, M., & Verzera, A. (2020). Aroma compounds in mini-watermelon fruits from different grafting combinations. *Journal of the Science of Food and Agriculture, 100*(3), 1328–1335. https://doi.org/10.1002/jsfa.10149

Valcárcel, M., Ibáñez, G., Martí, R., Beltrán, J., Cebolla-Cornejo, J., & Roselló, S. (2021). Optimization of electronic nose drift correction applied to tomato volatile profiling. *Analytical and Bioanalytical Chemistry, 413*(15), 3893–3907. https://doi.org/10.1007/s00216-021-03340-5

Verzera, A., Dima, G., Tripodi, G., & Ziino, M. (2011). Fast quantitative determination of aroma volatile constituents in melon fruits by headspace–solid-phase microextraction and gas chromatography–mass spectrometry. *Food Analytical Methods, 4*, 141–149. https://doi.org/10.1007/s12161-010-9159-z.

Vicente-Villardon, J. L. (2015). MULTBIPLOT: A package for multivariate analysis using biplots. *Departamento de Estadística*. Universidad de Salamanca. Retrieved from http://biplot.dep.usal.es/multbiplot/introduction.html.

Voss, H. G. J., Mendes Júnior, J. J. A., Farinelli, M. E., & Stevan, S. L. (2019). A prototype to detect the alcohol content of beers based on an electronic nose. *Sensors, 19*(11), 2646. https://doi.org/10.3390/s19112646

Wang, Q., Chen, X., Zhang, C., Li, X., Yue, N., Shao, H., et al. (2023). Discrimination and characterization of volatile flavor compounds in fresh oriental melon after forchlorfenuron application using electronic nose (E-Nose) and headspace-gas chromatography-ion mobility spectrometry (HS-GC-IMS). *Foods, 12*(6). https://doi.org/10.3390/foods12061272

Wise, B. M., Gallagher, N. B., Bro, R., Shaver, J. M., Windig, W., & Koch, R. S. (2006). *Chemometrics tutorial for PLS_Toolbox and solo*. Wenatchee: Eigenvector Research Inc.

Yajima, I., Sakakibara, H., Ide, J., Yanai, T., & Hayashi, K. (1985). Volatile flavor components of watermelon (Citrullus vulgaris). *Agricultural and Biological Chemistry, 49*(11), 3145–3150. https://doi.org/10.1271/bbb1961.49.3145

Zawirska-Wojtasiak, R., Gośliński, M., Szwacka, M., Gajc-Wolska, J., & Mildner-Szkudlarz, S. (2009). Aroma evaluation of transgenic, thaumatin II-producing

cucumber fruits. *Journal of Food Science, 74*(3), 204–210. https://doi.org/10.1111/j.1750-3841.2009.01082.x

Zhou, C., Fan, J., Tan, R., Peng, Q., Cai, J., & Zhang, W. (2022). Prediction of linalool content in Osmanthus fragrans using E-nose technology. *Journal of Sensors*. https://doi.org/10.1155/2022/7349030, 2022.

Zhou, C.-L., Mi, L., Hu, X.-Y., & Zhu, B.-H. (2017). Evaluation of three pumpkin species: Correlation with physicochemical, antioxidant properties and classification using SPME-GC–MS and E-nose methods. *Journal of Food Science and Technology, 54*(10), 3118–3131. https://doi.org/10.1007/s13197-017-2748-8