Characterization of the aromatic profile of Ruché wine from Piedmont (Italy) with gas chromatography–mass spectrometry and unsupervised machine learning techniques

Published version:

DOI:10.1002/jsfa.14083

Characterization of the aromatic profile of Ruchè wine from Piedmont (Italy) with gas - chromatography coupled to mass spectrometry and unsupervised Machine Learning techniques

Ciro Orecchio[a], Roberto Rabezzana[a*], Marco Vincenti[a], Elisabetta Bonometti[a], Enzo Laurenti[a], Lorenza Operti[a], Monica Rigoletto[a]

[a] Department of Chemistry, University of Turin, Via Giuria 7, 10125 Turin, Italy

*Correspondence: roberto.rabezzana@unito.it (R.R.)

Abstract

Background: The volatile composition of Ruchè, a red wine produced from a native grape in Piedmont (Italy), was investigated by means of headspace-solid phase microextraction (SPME) coupled to gas chromatography‐mass spectrometry (GC/MS). The main volatile compounds of the wine were identified and quantified. Chemometric techniques were applied to identify features and possible clusters among the different samples.

Results: Forty volatile compounds were unambiguously identified in the 36 wine samples from different producers. The aroma profile is mainly composed by different alcohols and esters, but also features appreciable concentrations of terpenes, aldehydes and octanoic acid. 2-methyl benzaldehyde was identified for the first time. High concentration of isoamyl alcohol significantly contributes to the aroma complexity. Differences between producers are highlighted.

Conclusion: The present work is the first report about the volatile profile of Ruchè wines investigated by chemometric methods on quantitative results. Multivariate exploratory approaches allowed to highlight little but peculiar differences among the wines studied, conceivably ascribed to different winemaking procedures. This study may be developed in future by investigating possible differences between wines according to different vintages of production.

Keywords: Ruchè; Volatile Organic Compounds; HS-SPME; GC-MS; Chemometrics

INTRODUCTION

Ruchè is a red wine produced from a native grape variety in a very restricted area of Piedmont, a region in north-west Italy. According to the disciplinary of production, the Ruchè vine can be cultivated in the territory of seven villages only (Castagnole Monferrato, Montemagno, Portacomaro, Refrancore, Scurzolengo, Viarigi, Grana): among them, the most important is by far Castagnole Monferrato, where most of the wine makers (≈40) are located. This is the reason why the complete and official name of the DOCG (denomination of controlled and guaranteed origin – official recognition in 2010 - corresponding to PDO - protected designation of origin, according to European regulations) is "Ruchè of Castagnole Monferrato". Despite the very limited territory of production, the progressive popularity of Ruchè wine resulted in a rapid increase of the vineyard area, which expanded from the original 10^2 km^2 (≈25 acres) in 1988, to about 10^3 km^2 (≈250 acres) in 2010, and further doubling to $2{,}04 \cdot 10^3$ km^2 in 2022 (1.1 million bottles produced).

Piedmont is one of the most important wine producing regions in Italy, and Ruchè is one of the seventeen wines from Piedmont to hold the DOCG denomination[1]. Ruchè vine had been known since ancient times, indeed various folkloristic stories arose around its denomination: some studies attribute its name to the vineyards proximity to a Benedictine convent dedicated to Saint "Rocco", subsequently destroyed, while others, relate its name to the vine's predilection for the steepest and sunniest "rocks". Despite its ancient origin, it was only in the early 80's of the 20th century that the wine was acknowledged with the denomination of controlled origin (DOC). Lastly, thanks to the production regulation of 2010, Ruchè effectively became the smallest DOCG in the Monferrato area.

Ruchè wine features a specific aromatic flavour, where floral nuances are predominant. Bonino et al.[2] investigated its aromatic composition, focussing on the primary aromatic substances, i.e. those directly deriving from the grape. Their analyses were conducted by headspace solid phase micro extraction (HS-SPME) coupled to gas-chromatography and mass spectrometry (GC-MS). According to their results, the primary aroma profile of Ruchè shows some similarities with Malvasia, a wine produced from an aromatic vine. Subsequently, Genovese et al.[3] employed a solvent extraction procedure to perform GC-olfactometry and GC-MS analysis on a Ruchè wine extract, with the aim to identify the main analytes responsible for the perceived fragrances. Their results agreed with those of Bonino et al.[2] in highlighting the role played by varietal aromas.

Purpose of the present paper was to study the whole aromatic composition of Ruchè wine and compare the composition of Ruchè wines arising from different producers, in order to verify the variability of Ruchè components. It is well known that the aromatic profile of a wine may be adopted as a scientific basis for wine classification and for characterization of typical wine varietals, their provenance, soil composition, and aging conditions, since volatile organic compounds (VOCs) are closely related to the *bouquet*, that is the characteristic and peculiar fragrance of wine[4,5]. Moreover, the study of aromatic compounds may help to improve the wine quality during the various steps of vine cultivation and wine processing. The analytical measurements were performed by HS-SPME coupled to GC-MS. SPME is a fast, robust and unexpensive technique[6] which has been consistently used for the characterization of several wine varieties and analysis of their flavour components[7–9], even if the SPME fiber can capture only a small amount of volatile substances, resulting in a limited sensitivity especially toward the most polar flavour components, such as some terpene derivatives. The experimental data gathered in this work were processed by traditional and advanced exploratory chemometric techniques in order to hypothesize possible clusters and peculiar features occurring among the investigated samples.

MATERIALS AND METHODS

Wine samples

Thirty-eight wine samples from different producers were purchased directly from the wine maker or in different shops and markets. Thirty-one wine samples were from 2021 vintage, while the 2021 vintage was no more available from some producers. Consequently, different vintages were sampled: 2019 (two wines), 2020 (three) and 2022 (two). As previously stated, most producers are located in the Castagnole Monferrato territory, yielding 22 out of 36 wines (labelled CM, n. 22), while the remaining 14 came from the other four villages where the Ruchè vine cultivation is allowed, according to its disciplinary of production: Grana (GR, n. 4), Montemagno (MO, n. 4), Portacomaro (PO, n. 3), Viarigi (VI, n. 3). Alle these villages are located in the territory of Asti (44°N, 8°E), where the yearly average rainfall is 900 mm. The average humidity is 74%, and the average annual radiation is 5.5 MJ/m^2. Forests represent ca. 26% of the territory, which displays an average altitude of 240 m a.s.l.

Table 1 reports the list of the wines studied with the labels employed throughout the paper to identify them, the year and the site of production. All the wine bottles displayed the DOCG certification mark, as a warranty of quality and geographical origin. Winery identities were not reported in order to protect proprietary interests. The average alcohol content of wines was 15% v/v. For all wines, fermentation and ageing were conducted in stainless steel. All samples were stored at 4°C until analysis.

Reagents and standards

All analytical standards, 3,4-dimethylbenzaldehyde, used as the internal standard, and *n*-alkane mixture (C_8–C_{20}) used for determining the retention indexes, were purchased at analytical purity degree from Sigma Aldrich (St Louis, MO, USA) and used as received. For headspace sampling, 5 mL wine aliquots were placed in a 20 mL glass vial (headspace volume 15 mL) with 1.0 g of sodium chloride and 5 µL of a 3,4-dimethylbenzaldehyde ethanolic solution (0.2 g kg^{-1}, $2 \cdot 10^{-4}$ g kg^{-1} in the sample). The vial was sealed with an aluminium-coated silicone rubber septum and conditioned at 40°C for 45 minutes before sampling. Headspace extraction was performed with a divinylbenzene/carboxen/polydimethylsiloxane (CAR/PDMS/DVB) 50/30 µm fiber from Supelco (Bellefonte, PA, USA) for 60 minutes at 40°C under continuous stirring. Then, the fiber was transferred into the GC inlet. Extraction time and temperature were selected after extensive testing until optimal reproducibility was reached. Each experiment and analysis was conducted in triplicate.

GC-MS

A Focus GC Thermo Scientific (Waltham, MA, USA) gas chromatograph with quadrupole mass spectrometer operated in electron ionization (EI – 70 eV) mode was used for measurements. The GC system was equipped with a fused silica capillary column (Supelcowax® 10, 30 m × 0.25 mm ID; 0.25 µm film thickness) from Supelco. GC conditions were as follows: inlet at 200 °C; splitless time: 3 min; oven initial temperature: 40 °C (3 min); then, from 40 °C to 160 °C at a rate of 2.0 °C/min (1 min hold), then to 200 °C at a rate of 10°C/min (2 min hold). GC-MS transfer line at 250 °C; carrier gas He (5.5 grade of purity) at 1.2 mL/min. The MS ion source temperature was kept at 250 °C. Mass spectra and reconstructed total ion chromatograms (TIC) were obtained by automatic scanning in the *m/z* 35–600 mass range. GC–MS data were processed with the Excalibur 1.4 software. Every 10 analyses, a run of blank fiber was performed in order to check for memory effects. The VOCs identification was achieved by comparing mass spectra with those of the data system library (NIST 98, P > 90%) and comparing the experimental vs. theoretical retention indexes. Experimental indexes were determined according to Golovnya et al.[10], by injecting a *n*-alkane mixture (C_8–C_{20}) into the chromatograph. Theoretical indexes were derived from the NIST 98 library. In doubtful cases, injection of pure standards was performed. Quantification was carried out by the internal standard method. To this purpose, different standard solutions (n. 8) were prepared by dissolving exact volumes of the analytical standard compounds, listed in Table 2, in a 12.0% v/v ethanol solutions at 0.1 g kg^{-1} concentration. Next, a set of six standard mixtures containing known concentrations of the chemical standards were prepared by dilution. To every synthetic standard solution $6 \cdot 10^{-3}$ g cm^{-3} of tartaric acid and the internal standard (IS, same concentration as the samples) were added. The standard mixtures were analysed and a six-points calibration graph of relative component area ($A_{analyte}/A_{IS}$) versus analyte concentration ($C_{analyte}$) was drawn to confirm a linear detector response and from which the amount of the analyte can be determined. Each chemical standard was used to quantify several compounds with similar chemical structure (Table 2).

Chemometrics

The data analysis was performed with Python Programming Language[11]. The adopted version (3.11.9) provides a plenty of essential packages and libraries, including Numpy[12] for scientific computing, Pandas[13] for data structuring and manipulation, Matplotlib[14] and Plotly for graphical outcomes, and Scikit-learn[15] for Machine Learning implementations. By the last library, different unsupervised techniques were applied to the data, including Principal Component Analysis (PCA), Hierarchical Cluster Analysis (HCA), and t-distributed Stochastic Neighbour Embedding (t-SNE). All the data were auto-scaled before performing any type of unsupervised analysis.

RESULTS AND DISCUSSION

Volatile Compounds

A total of 40 compounds were unambiguously identified in the volatile fraction of Ruchè wines. However, 13 of these compounds showed very low concentrations and were present in few samples only; therefore, they were not included in the following discussion. It is worth noting that, among the analytes discarded, three varietal compounds are included, namely nerol oxide, terpinene-4-ol and hotrienol. They all belong to the class of terpene derivatives, which generally display very low olfactory perception thresholds and feature fruity and floral notes. Therefore, it is conceivable that they contribute to the Ruchè *bouquet* to an extent comparable to those of analytes detected in higher abundance, but their detected concentration is possibly underestimated, because the extraction of such polar compounds is likely to be incomplete on the apolar fiber coating we adopted. Table 3 reports the 27 analytes considered in the subsequent elaborations with their mean concentrations (g kg^{-1}) and relative standard deviations. The compounds are listed according to their chemical class. The odour thresholds (OTH) and odour activity values (OAV) are also reported in Table 3.

In general, the triphasic fiber coating CAR/DVB/PDMS is largely adopted in food matrix sampling, including wines[2,7–9], because of its versatility, since it may extract a broad range of analytes with different chemical structures. However, the extraction efficiency of CAR/DVB/PDMS fibers toward the most polar wine volatile components, such as alcoholic terpene derivatives, is modest, and limited their random detection to a small number of our samples, without providing statistical significance. Therefore, the statistical analysis of our study has been mainly focused on the fermentative and aging varietal compounds, mostly composed by slightly polar esters and higher alcohols, which were detected in appreciable abundance.

Higher alcohols represent the most abundant components, since they amount to about 80% of the total concentration of volatile organic compounds. Among them, 3-methyl-1-butanol (isoamyl alcohol) represents by far the most abundant analyte with an average concentration above 0.4 g kg^{-1}. Isoamyl alcohol is the result of leucine and valine metabolism, but can also be formed from pyruvate[16]. The second most abundant alcohol is 2-phenylethanol, which is produced via the Ehrilch pathway starting from 2-phenylalanine[17]. It contributes to the floral nuance of Ruchè wine, but is also present in most wines varieties and could not be considered peculiar to it. Nevertheless, the detected 2-phenylethanol concentration largely exceeded the OAV value of unity, and therefore its contribution to the Ruché aroma cannot be neglected. Generally, moderate concentrations of some volatiles showing high fragrance intensity, such as isoamyl alcohol and 2-phenylethanol, provide positive sensory features to the wine, conferring flower, honey, and fruit aroma notes. However, this typically occurs when their concentrations do not exceed 0.3 – 0.35 g kg^{-1}, while, according to some authors[18,19], higher alcohols concentration becomes detrimental for the wine aroma when they exceed 0.4 g kg^{-1}, since they add pungent and unpleasant hints, making the ‘spirit’ nuance too strong and covering other aromas. The third most abundant alcohol is 1-hexanol, which is reported to produce an herbaceous scent. In Ruchè, its OAV does not exceed unity, making it perceptible but not be detrimental to the overall wine aroma. The C_7-C_9 linear alcohols are also present in low concentrations. Like hexanol, they are formed by decarboxylation of fatty acids[20] and display fruity aromas. Their concentration in wines normally decrease with aging due to esterification reactions.

The second most abundant class of Ruchè components is represented by esters. They are scarcely present in grapes, but are generated both during fermentation - by enzymatic activity – and with aging. In particular, esterification reactions take place during storage with formation of a variety of esters that contribute to the wine aroma. The most abundant ester in Ruchè wine is diethyl succinate; it features fruity odour (watermelon), but combined with a high perception threshold, resulting in an OAV lower than 1. It is considered an aging ester[17], *i.e.* formed during the second fermentation of wine. Ethyl lactate is also formed during malo-lactic fermentation; despite being quite abundant, it also displays a high odour threshold that makes its OAV lower than 1. Ethyl octanoate is also quite abundant in Ruchè wine and imparts fruity and ethereal odours[20]. Acetic esters are formed by the reaction of coenzyme acyl-S-CoA with higher alcohols[16]. The most abundant in Ruchè are isoamyl acetate (banana scent) and phenylethyl acetate (floral scent), both arising from the alcohols with the highest concentration. It is worth noting that the esters mentioned in this paragraph are quite common in red wines and, therefore, cannot be considered as peculiarly responsible for the uniqueness of Ruché *bouquet*. Indeed, this is largely ascribed to varietal compounds, and only the esters featuring OAVs higher than 1 play a significant role in the overall wine aroma.

Aldehydes only account for 0.28% of total aromatic composition. This class of compounds is not very abundant in grapes but are formed in wine through different pathways[16]. Benzaldehyde is a product of fermentation, but it could also originate by *Botrytis cinerea* spoilage[21]. At low concentrations, it imparts a pleasant oak flavour, but when its concentration is too high, an unpleasant 'jam' aroma is perceived, thus representing a negative factor. Also the presence of furfural in wine may be ascribed to *Botrytis cinerea* spoilage[16]. However, its average concentration is quite low and its OAV far lower than unity. 2-methylbenzaldehyde is another component of Ruchè, which, to the best of our knowledge, has never been detected before in wine. This compound was reported in lamb meat extracts[22] and it is also a component of Swiss cheese flavour[23]. Herein, the authors propose a mechanism of formation starting from the complexation of proline with ethanal. Proline is one of the most abundant amino acids present in must[16] and ethanal is formed by ethanol oxidation, making this hypothesis a conceivable formation pathway for 2-methylbenzaldehyde in wine. Nevertheless, it cannot be considered as a characteristic marker of Ruchè wine, since it is not present in all the examined samples.

Terpenes are varietal aromas present in the skin of aromatic vines and are transferred to wine during various production steps. Bonino et al.[2] stated that Ruchè shares several features with aromatic wines, despite being produced from a non-aromatic vine, because numerous terpenes and other varietal aromas were detected in its volatile composition. However, it should be noted that the wines under examination were produced by micro-vinification by the authors, and were not commercial products. As mentioned above, the intrinsic limitations of the adopted CAR/PDMS/DVB fibers in extracting polar substances allowed us to detect only three terpene derivatives, present in most samples at appreciable concentrations, namely citronellol, linalool, and terpineol. As outlined at the beginning of the Discussion section, three further terpene derivatives were detected, but only in few samples. Overall, the presence of varietal compounds such as terpenes is crucial in defining the distinctive aromatic profile of a wine, even if they are present in low concentrations, since they are mostly characterized by low detection thresholds. Among the three most abundant monoterpenes, citronellol and terpineol displayed OAV values above unity. These are among the most odoriferous compounds and impart lemon and rose scent, respectively. However, the role played by the aromatic compounds is synergistic and it is therefore conceivable that also the less represented varietal aromas effectively contribute to the complexity of the Ruchè *bouquet*.

Only one fatty acid was detected in relatively high concentration, namely caprylic (octanoic) acid. Fatty acids in wine are produced by yeast. Fatty acids hinder fermentation when their concentration is too high (above $1 \cdot 10^{-3}$ g kg^{-1})[16] and actually they may be deliberately added to the must with the purpose of producing sweet wines[24]. This addition has also the effect of reducing the amount of SO_2 needed. However, the average concentration found in Ruchè wines was below $1 \cdot 10^{-3}$ g kg^{-1}.

Overall, the present results generally agree with those of Genovese et al.[3] in the identification of the most abundant volatile substances. However, the solvent-extraction procedure employed by Genovese allowed the determination of a higher number of analytes, especially among the class of varietal compounds, which play a major role in the definition of the wine aroma profile, such as geraniol, geranic acid, and β-damascenone. Nevertheless, SPME is nowadays largely employed in food aroma determination because of its advantages, including the avoidance of halogenated solvents and ease of use. Detection of polar analytes such as terpenols may be improved by employing different fiber coatings, which exhibit different selectivity profiles.

Unsupervised methods

In the studies involving the comparison of several samples, large amount of data is generated by the modern analytical techniques and the adoption of multivariate statistical approaches is mandatory to exploit the useful information potential present in the chemical data and gain a deeper understanding of complex chemical systems[25,26]. In the present study, all wine samples were produced in a restricted area from the same native grape variety, suggesting a relative uniformity of their composition and no predetermined class separation. Therefore, the data were studied by several multivariate data exploration techniques, including the well-known PCA[27] and HCA[28], together with a more recent unsupervised method, namely t-SNE[29], in order to investigate in detail the differences and similarities occurring among the assorted wines productions and the possible correlations among the substances that define their sensory properties.

The less-known t-SNE technique may deserve some description. It is a nonlinear and manifold based machine learning approach particularly suitable for embedding *p*-dimensional data in a two- or three-dimensional space while preserving the significant structure of the original data[30]. This SNE-based algorithm is able to convert the high-dimensional Euclidean distance between samples into conditional probabilities, that can be interpreted as values of similarity[29].

Multivariate data analysis

Starting from an initial dataset characterized by 36 samples and 40 variables, a first data filter was applied to make the final data matrix more manageable. 13 out of 40 variables were removed because they were not consistently detected in all the samples. The final data matrix included 36 samples and 27 variables (VOCs). A further column indicating the origin (village) was introduced "*a posteriori"* to make the graphic interpretation based on territory clearer. An overview of the concentration (ppm) distribution of all detected analytes based on the wine origin is provided in Figure 1.

PCA was performed to the final data matrix. The variance explained by the first three principal components is 36%, 14% and 12% respectively, for a total of 62%. The corresponding scores plots are reported in Figures 2a (PC1 vs. PC2) and 2b (PC1 Vs. PC3). From Figure 2a it is possible to detect a partial separation between the joined CM and GR samples and the others, in particular VI and PO, along the bisector between the first and third quadrant. The same trend is enhanced in Figure 2b: all the samples belonging to VI and PO stand above this bisector while those of group GR are located below it together with the majority of CM samples. Preliminarily, wine samples from CM and GR production are apparently similar, likewise VI and PO productions, since their data-points are roughly concentrated in the same PC space portion. In contrast, the two sub-groups show differences that should be referred to their composition. In the case of MO samples which are equally divided into the two sub-groups, the geographic location appears not to play a role. Instead, the production process adopted by the different wineries and the vintage may rationally represent other data influencing factor that should be taken into account for all samples sub-sets, remembering the restricted geographical area of Ruchè wine production (around 80 km^2). In particular, the possible effect of the vintage was attempted, even if a clear predominance of a specific vintage (2021) was present: no difference was highlighted between 2021 and "non-2021" wines, as the latter are scattered throughout the PCA space (Supporting Information, Figure 1S).

The interpretation of the partial separation observed in Figure 2 relies on the loadings plots depicted in Figure 3a and 3b. In particular, the variables with greater influence (relative abundance) in describing the wines represented in the score plots (Figure 2) have the same orientation in the loadings plots (Figure 3). Therefore, the samples located above the bisector previously mentioned (for example, VI and PO) are characterized by a high content of benzaldehyde, 2-methyl-benzaldehyde, methyl-benzoate acid and 2-methyl-1-propanol (isobutanol), all situated in the second quadrant of Figure 3. The high abundance of the two aldehydes may be related to the presence of *Botrytis cinerea* spoilage, that is able to transform benzyl alcohol in benzaldehyde[31]. Notably, both benzaldehyde and methyl benzoate confer to the wine a marked bitter/green almond hint. Methyl benzoate is not frequently found in wines; in a 2001 paper, Aznar et al.[32] reported the first identification of this component in Spanish red wines from Rioja. Together with benzaldehyde and its methyl derivative, methyl benzoate belongs to the class of benzenoids, reported to provide lesser differentiation among wines produced in the same area and similar procedures[33]. Asproudi et al.[34] suggested that environmental stress may activate the chemical pathways leading to volatile benzenoids formation in grapes, which also show higher concentrations in grapes from old-vine vineyards. Accordingly, the websites of some producers located in the PO territory report that their Ruchè stems from forty-years old vines, in agreement with the high benzenoids concentration found in them. Isobutanol characterize VI samples (not PO) with a slightly higher concentration. At high concentration level, isobutanol and isoamyl alcohol are essentially detrimental to wine quality[35]; an anomalous amount (Fig. 1) of these components was only found in a single CM wine (ru37), suggesting the occurrence of peculiar production conditions (*i.e.*, yeast selection, fermentation conditions, wine blending). While the producer's website does not report any information regarding the winemaking procedure, its activity only started in 2020, notably making it the youngest winemaker in the consortium.

In the opposite direction (forth quadrant) are mostly found the GR and CM samples, characterized by higher concentration of several ethyl esters, including octanoic, nonanoic, decanoic, and dodecanoic, together with citronellol and acetic acid, phenylethyl ester, some of which are particularly low in VI and PO samples. The components depicted in the first quadrant of Figure 3, basically higher alcohols such as isoamyl alcohol, 2-phenylethanol, 1-hexanol, and 1-heptanol, do not characterize specific production territories, but rather single samples. These compounds bring to the wine flavours, either pleasing or distasteful, that strongly depends on their concentration.

Further insight into the data unsupervised exploration is provided by HCA, from which the dendrogram shown in Figure 4 is derived. In this plot, the Euclidean distance between couples of samples is reported on the y-axis and the Ward linkage approach is used for their aggregation: the resulting clusters are in full agreement with the sample distribution observed in Fig.2. It is possible to highlight (i) the presence of the cluster related to the first samples (ru1-ru14, without ru07) corresponding to the samples above the bisector between first and third quadrant, (ii) an outlier evidenced as a singleton (ru37), (iii) the turquoise cluster that includes all samples characterized by high alcohols and esters concentrations. The heatmap depicted in Figure 5 complete the information with the VOCs clustering and provides an immediate overview on the volatile composition of all Ruchè samples.

Finally, the dimensionality reduction performed by t-SNE is showed in Figure 6: once again this algorithm underlines the separation already mentioned, endorsing the results provided by the previously unsupervised methods.

CONCLUSIONS

Aim of this study was to characterize the volatile profile of Ruchè, a typical wine of Piedmont (Italy) and highlight the minute differences of Ruchè wine specimen produced in the restricted cultivation area of this native vine variety. This wine features a high concentration of higher alcohols and various unusual components, which make Ruchè wine largely appreciated thanks to the presence of different aromatic combinations. In the present study, the sampling technique employed allowed to efficiently extract and, therefore, to highlight the contribution of the most abundant semi-polar aromatic species, namely 2-phenylethanol (floral nuances), diethyl succinate (fruity), and isoamyl acetate (banana). Conversely, the extraction of polar compounds such as varietal aromas, which are normally less abundant but are peculiar in defining the uniqueness of a wine, was less efficient. Most of the detected aromatic compounds show an OAV above 1, thus contributing to build up the complex wine *bouquet*.

Multivariate analysis allowed to identify small but peculiar differences among the studied samples flavours, despite their homogeneity in terms of vintage and production territory. The outcome of the present study is likely to represent a useful indication for producers in order to better characterize their production and improve the winemaking procedures.

FUNDING

This study was carried out within the Agritech National Research Center and received funding from the European Union Next-GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) – MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4 – D.D. 1032 17/06/2022, CN00000022).

ACKNOWLEDGEMENTS

Authors acknowledge support from the Project CH4.0 under the MUR program "Dipartimenti di Eccellenza 2023-2027" (CUP: D13C22003520001).

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the outcomes of this study are available from the corresponding author upon reasonable request.

AUTHORS CONTRIBUTION

Conceptualization, R.R., C.O., M.V.; methodology, R.R., C.O., M.V.; software, C.O.; formal analysis, E.B., E.L., M.R.; investigation, R.R., C.O., E.B., E.L.; resources, L.O., C.O. M.V.; data curation, R.R., C.O., M.V.; writing-original draft preparation, R.R., C.O.; writing-review and editing, M.V., C.O., LO.; funding acquisition, R.R., M.V. All authors have read and agreed to the published version of the manuscript.

REFERENCES

1 Valoritalia. Annual-Report Valoritalia. (2022).

2 Bonino M, Schellino R, Rizzi C, Aigotti R, Delfini C, and Baiocchi C. Aroma compounds of an Italian wine (Ruche´) by HS-SPME analysis coupled with GC-ITMS. *Food Chem* (2003).

3 Genovese A and Gambuti A. Analisi di componenti odorosi mediante tecniche poli-strumentali. *Ingredienti Alimentari* **39** (2008).

4 Dall'Asta C, Cirlini M, Morini E, and Galaverna G. Brand-dependent volatile fingerprinting of Italian wines from Valpolicella. *J Chromatogr A* **1218**:7557–7565 (2011).

5 Sagratini G, Maggi F, Caprioli G, Cristalli G, Ricciutelli M, Torregiani E, and Vittori S. Comparative study of aroma profile and phenolic content of Montepulciano monovarietal red wines from the Marches and Abruzzo regions of Italy using HS-SPME-GC-MS and HPLC-MS. *Food Chem* **132**:1592–1599 (2012).

6 Pawliszyn J. Applications of Solid Phase Microextraction. Pawliszyn J, ed. The Royal Society of Chemistry; 1999.

7 Avellone G, Salvo A, Costa R, Saija E, Bongiorno D, Stefano V Di, Calabrese G, and Dugo G. Investigation on the influence of spray-drying technology on the quality of Sicilian Nero d'Avola wines. *Food Chem* Elsevier Ltd; **240**:222–230 (2018).

8 Alves RF, Nascimento AMD, and Nogueira JMF. Characterization of the aroma profile of Madeira wine by sorptive extraction techniques. *Anal Chim Acta* Elsevier; **546**:11–21 (2005).

9 Condurso C, Cincotta F, Tripodi G, Sparacio A, Giglio DML, Sparla S, and Verzera A. Effects of cluster thinning on wine quality of Syrah cultivar (Vitis vinifera L.). *European Food Research and Technology* Springer Verlag; **242**:1719–1726 (2016).

10 Golovnya R V and Arsen'ev YN. System of Retention Indices and Its Physicochemical Application. Russian Chemical Reviews.

11 Rossum G Van. Centrum voor Wiskunde en Informatica Python reference manual. 1995.

12 Harris CR, Millman KJ, Walt SJ van der, Gommers R, Virtanen P, Cournapeau D, Wieser E, Taylor J, Berg S, Smith NJ, Kern R, Picus M, Hoyer S, Kerkwijk MH van, Brett M, Haldane A, Río JF del, Wiebe M, Peterson P, Gérard-Marchant P, Sheppard K, Reddy T, Weckesser W, Abbasi H, Gohlke C, and Oliphant TE. Array programming with NumPy. Nature. Nature Research; p. 357–362 2020.

13 Mckinney W. Data Structures for Statistical Computing in Python. 2010.

14 Hunter JD. Matplotlib: a 2D Graphics Environment. *Comput Sci Eng* (2007).

15 Pedregosa F, Michel V, Grisel O, Blondel M, Prettenhofer P, Weiss R, Vanderplas J, Cournapeau D, Pedregosa F, Varoquaux G, Gramfort A, Thirion B, Grisel O, Dubourg V, Passos A, Brucher M, Perrot

M, and Duchesnay E. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **12**:2825–2830 (2011).

16 Ribéreau-Gayon P, Glories Y, Maujean A, and Dubourdieu D. Handbook of Enology - The Chemistry of Wine Stabilization and Treatments. 2nd ed. John Wiley & Sons Ltd, ed. Chichester, UK; 2006.

17 Korenika AMJ, Preiner D, Tomaz I, and Jeromel A. Volatile profile characterization of croatian commercial sparkling wines. *Molecules* MDPI AG; **25** (2020).

18 de-la-Fuente-Blanco A, Sáenz-Navajas MP, and Ferreira V. Levels of higher alcohols inducing aroma changes and modulating experts' preferences in wine model solutions. *Aust J Grape Wine Res* Blackwell Publishing Ltd; **23**:162–169 (2017).

19 Carpena M, Fraga-Corral M, Otero P, Nogueira RA, Garcia-Oliveira P, Prieto MA, and Simal-Gandara J. Secondary aroma: Influence of wine microorganisms in their aroma profile. Foods. MDPI AG; 2021.

20 Panighel A and Flamini R. Applications of solid-phase microextraction and gas chromatography/mass spectrometry (SPME-GC/MS) in the study of grape and wine volatile compounds. Molecules. MDPI AG; p. 21291–21309 2014.

21 Tomasino E, Harrison R, Breitmeyer J, Sedcole R, Sherlock R, and Frost A. Aroma composition of 2-year-old New Zealand Pinot Noir wine and its relationship to sensory characteristics using canonical correlation analysis and addition/omission tests. *Aust J Grape Wine Res* **21**:376–388 (2015).

22 Bueno M, Resconi VC, Campo MM, Cacho J, Ferreira V, and Escudero A. Gas chromatographic-olfactometric characterisation of headspace and mouthspace key aroma compounds in fresh and frozen lamb meat. *Food Chem* **129**:1909–1918 (2011).

23 Griffith R and Hammond EG. Generation of Swiss Cheese Flavor Components by the Reaction of Amino Acids with Carbonyl Compounds. *J Dairy Sci* **72**:604–613 (1989).

24 Baniţă C, Antoce OA, and Cojocaru GA. Evaluation by a GC Electronic Nose of the Differences in Volatile Profile Induced by Stopping Fermentation with Octanoic and Decanoic Acid to Produce Sweet Wines. *Chemosensors* MDPI; **11** (2023).

25 González-Domínguez R, Sayago A, and Fernández-Recamales Á. An Overview on the Application of Chemometrics Tools in Food Authenticity and Traceability. Foods. MDPI; 2022.

26 Consonni V, Todeschini R, Gosetti F, and Ballabio D. Authenticity and chemometrics basics. *Chemometrics and Authenticity of Foods of Plant Origin* CRC Press; p. 9–35 2022.

27 Bro R and Smilde AK. Principal component analysis. Analytical Methods. Royal Society of Chemistry; p. 2812–2831 2014.

28 Massart DL, Vandeginste BGM, Buydens LMC, Jong S De, Lewi PJ, and Smeyers-Verbreke J. Handbook of Chemometrics and Qualimetrics Part A. Amsterdam: Elsevier; 1997.

29 Maaten L Van Der and Hinton G. Visualizing Data using t-SNE. Journal of Machine Learning Research. 2008.

30 Melit Devassy B and George S. Dimensionality reduction and visualisation of hyperspectral ink data using t-SNE. *Forensic Sci Int* Elsevier Ireland Ltd; **311** (2020).

31 Delfini C, Gaia P, Bardi L, Mariscalco G, Contiero M, and Pagliara A. Production of benzaldehyde, benzyl alcohol and benzoic acid by yeasts and Botrytis cinerea isolated from grape musts and wines. Vitis. 1991.

32 Aznar M, López R, Cacho JF, and Ferreira V. Identification and quantification of impact odorants of aged red wines from Rioja, GC-olfactometry, quantitative GC-MS, and odor evaluation of HPLC fractions. *J Agric Food Chem* American Chemical Society; **49**:2924–2929 (2001).

33 Slaghenaufi D, Guardini S, Tedeschi R, and Ugliano M. Volatile terpenoids, norisoprenoids and benzenoids as markers of fine scale vineyard segmentation for Corvina grapes and wines. *Food Research International* Elsevier Ltd; **125** (2019).

34 Asproudi A, Bonello F, Ragkousi V, Gianotti S, and Petrozziello M. Aroma precursors of Grignolino grapes (Vitis vinifera L.) and their modulation by vintage in a climate change scenario. *Front Plant Sci* Frontiers Media SA; **14** (2023).

35 De-La-Fuente-Blanco A, Sáenz-Navajas MP, and Ferreira V. On the effects of higher alcohols on red wine aroma. *Food Chem* Elsevier Ltd; **210**:107–114 (2016).

36 Tat L, Comuzzo P, Battistutta F, and Zironi R. Sweet-like off-flavor in aglianico del vulture wine: Ethyl phenylacetate as the mainly involved compound. *J Agric Food Chem* **55**:5205–5212 (2007).

37 Arcari SG, Caliari V, Sganzerla M, and Godoy HT. Volatile composition of Merlot red wine and its contribution to the aroma: optimization and validation of analytical method. *Talanta* Elsevier B.V.; **174**:752–766 (2017).

38 Verzera A, Tripodi G, Dima G, Condurso C, Scacco A, Cincotta F, Giglio DML, Santangelo T, and Sparacio A. Leaf removal and wine composition of Vitis vinifera L. cv. Nero d'Avola: The volatile aroma constituents. *J Sci Food Agric* John Wiley and Sons Ltd; **96**:150–159 (2016).

39 Guth H. Quantitation and Sensory Studies of Character Impact Odorants of Different White Wine Varieties. 1997.

40 Tufariello M, Capone S, and Siciliano P. Volatile components of Negroamaro red wines produced in Apulian Salento area. *Food Chem* Elsevier Ltd; **132**:2155–2164 (2012).

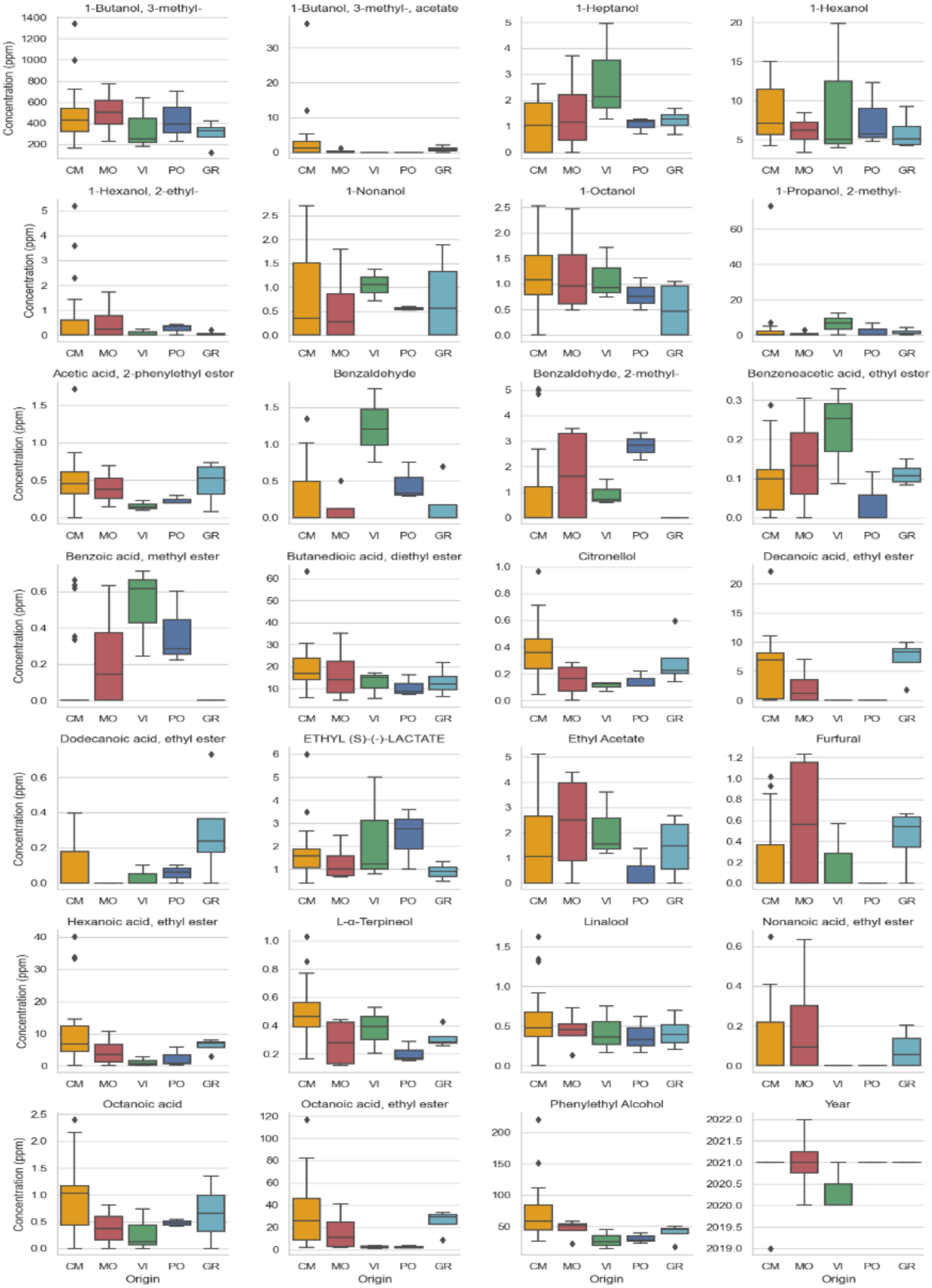

Figure 1. Boxplots of all volatile compounds detected in Ruchè wine samples. (CM = Castagnole Monferrato MO = Montemagno VI = Viarigi PO = Portacomaro GR = Grana)

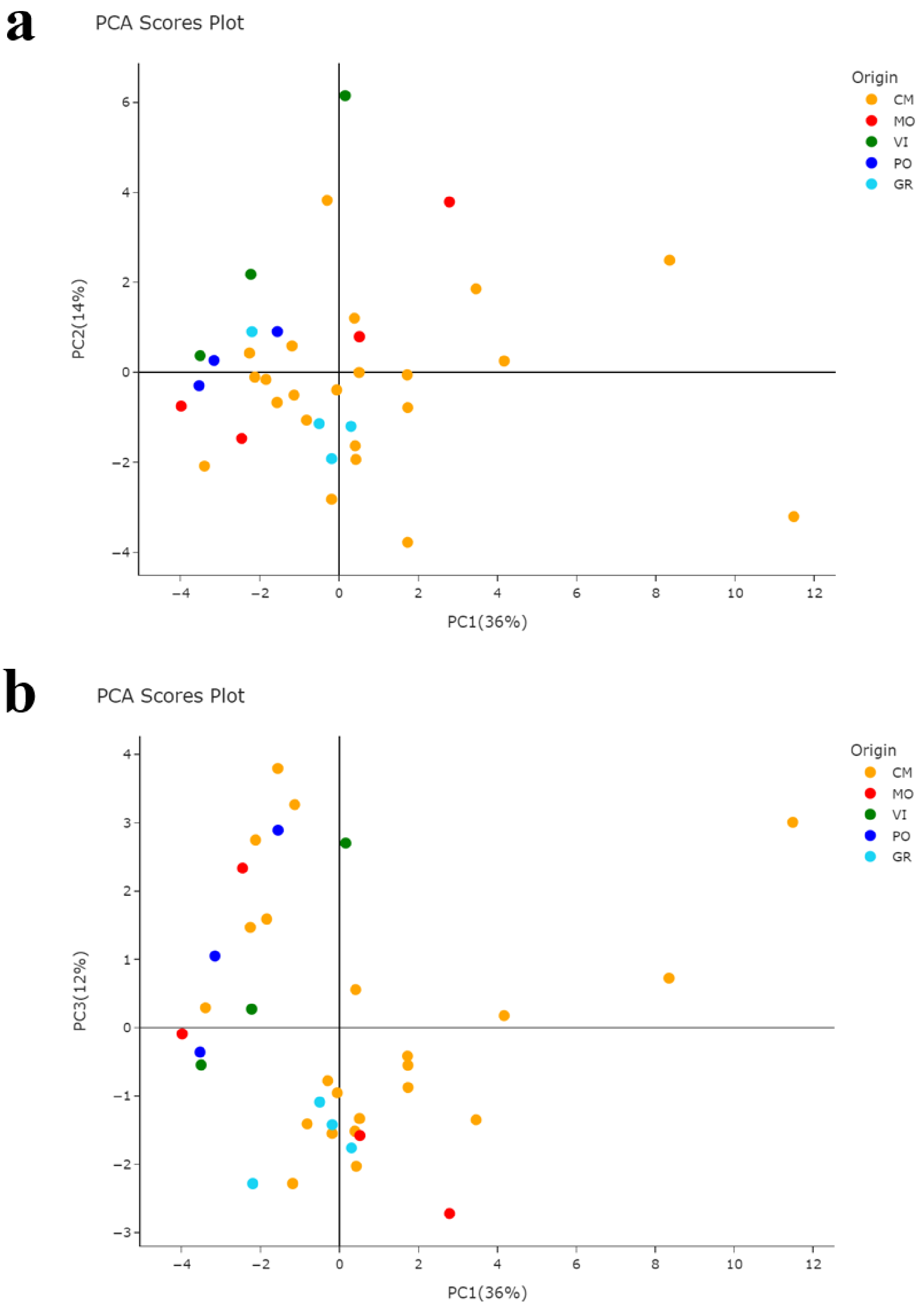

Figure 2. (a) Scores plot showing the projection of the samples on the first two principal components (CM = Castagnole Monferrato, MO = Montemagno, VI = Viarigi, PO = Portacomaro, GR = Grana). **(b)** Scores plot showing the projection of the samples on the first and third principal components.

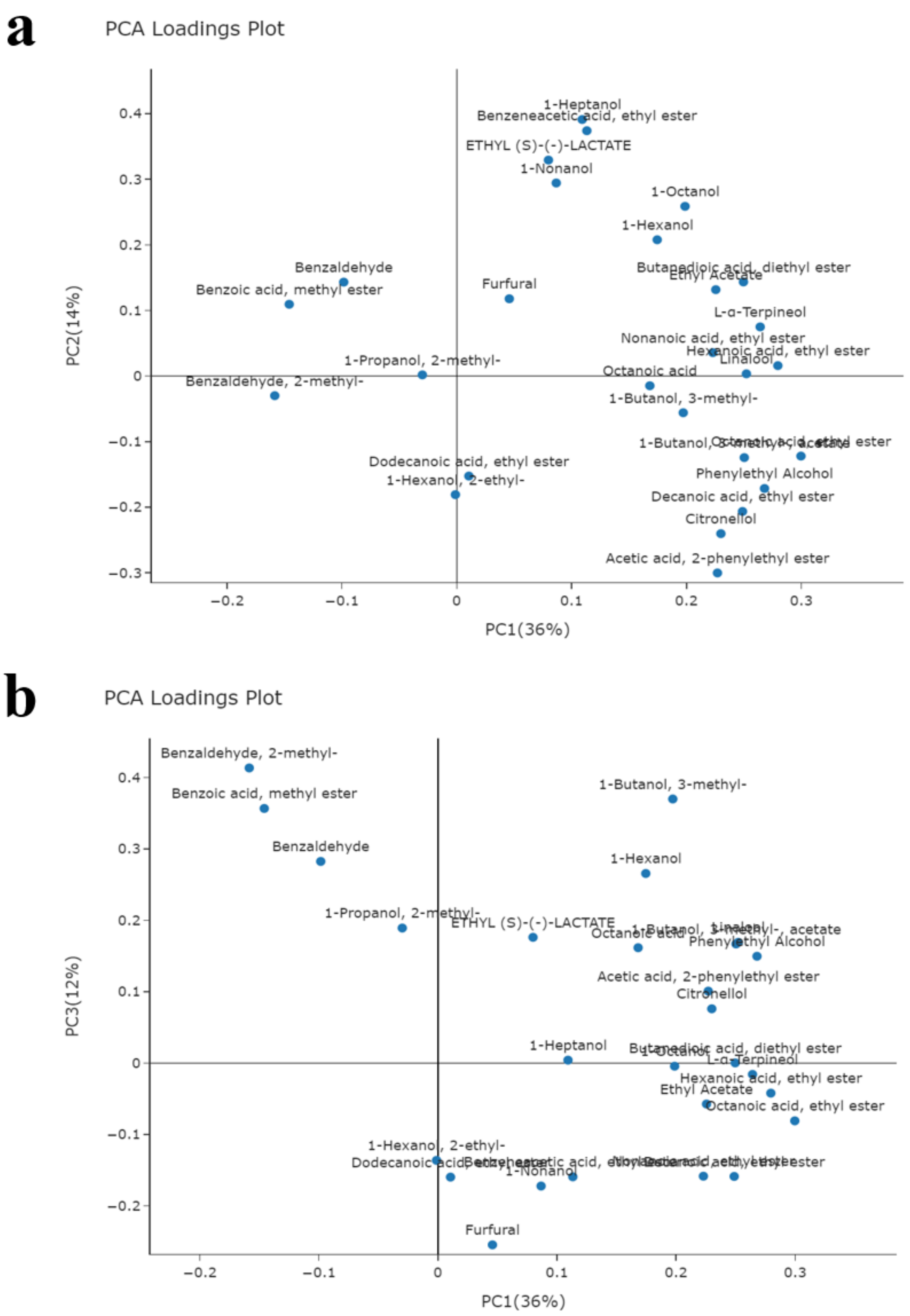

Figure 3. (a) Loadings plot showing the projection of the volatile compounds on the first two principal components. **(b)** Loadings plot showing the projection of the volatile compounds on the first and the third principal components.

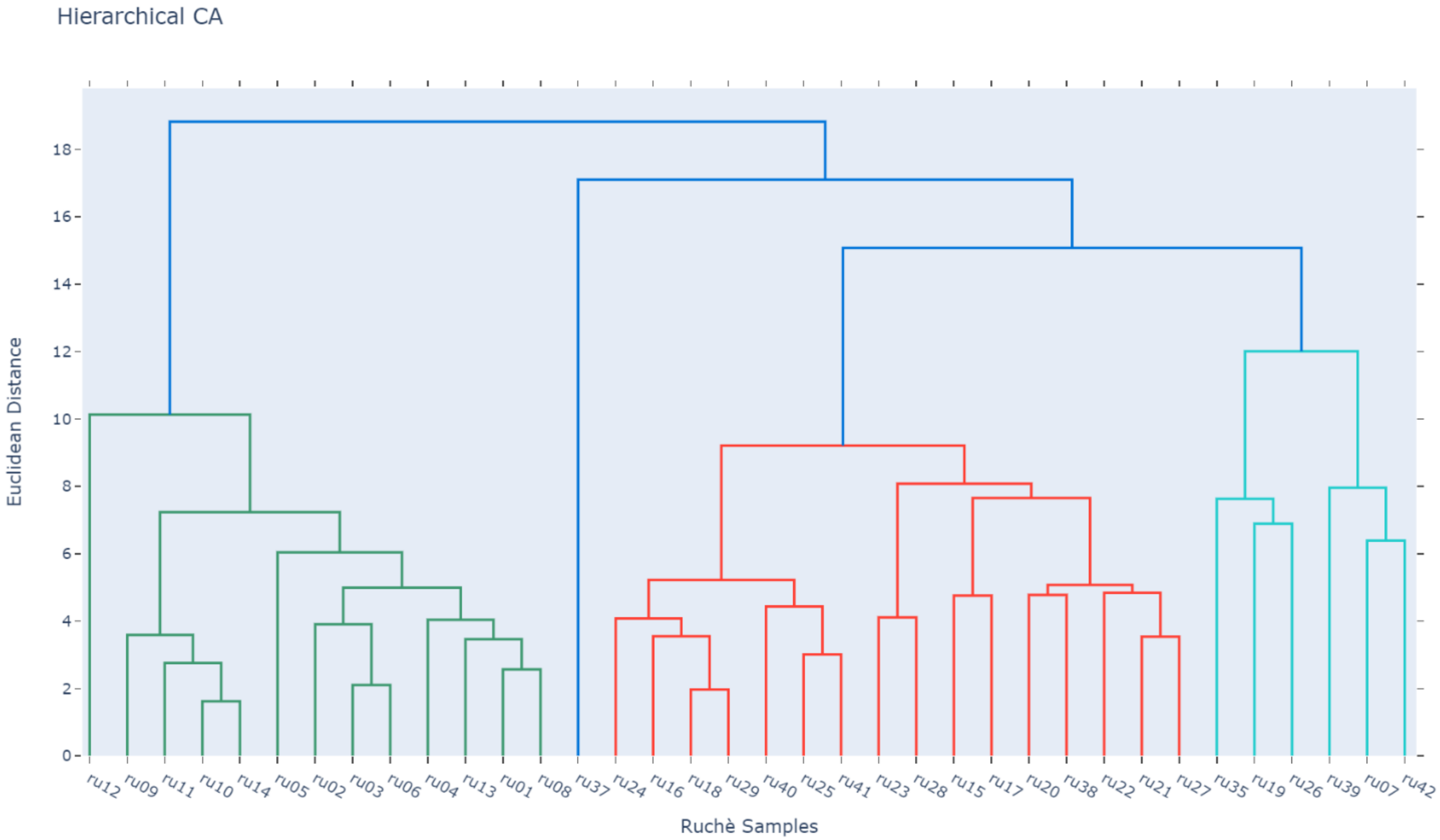

Figure 4. HCA performed on the full dataset, using Ward’s method and Euclidean distance.

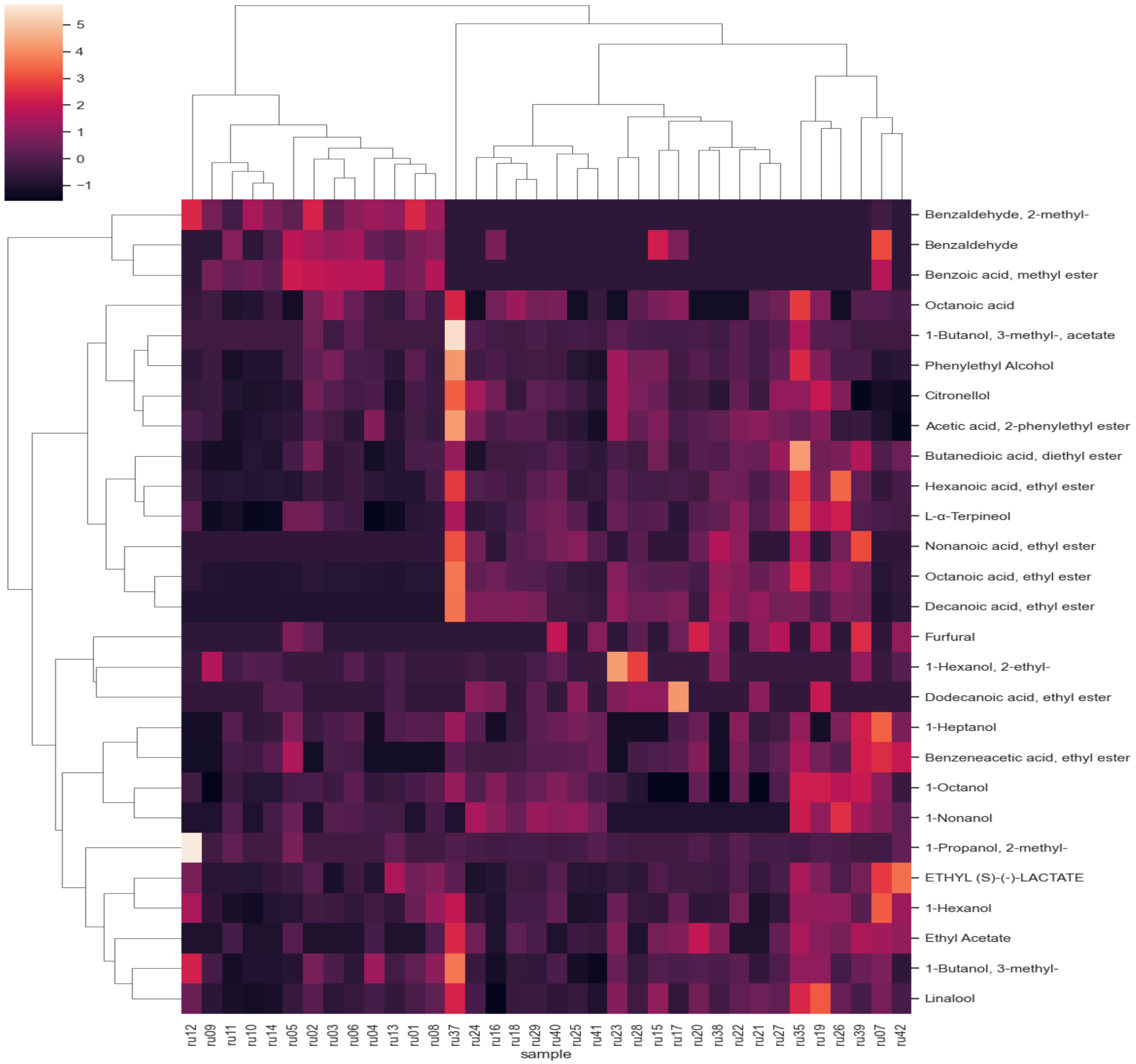

Figure 5. Cluster heat map analysis of VOCs in Ruchè wine samples. The chromatic scale (from low values, dark violet, to high values, skin color) indicates the correlation between the variable concentration and the wine sample.

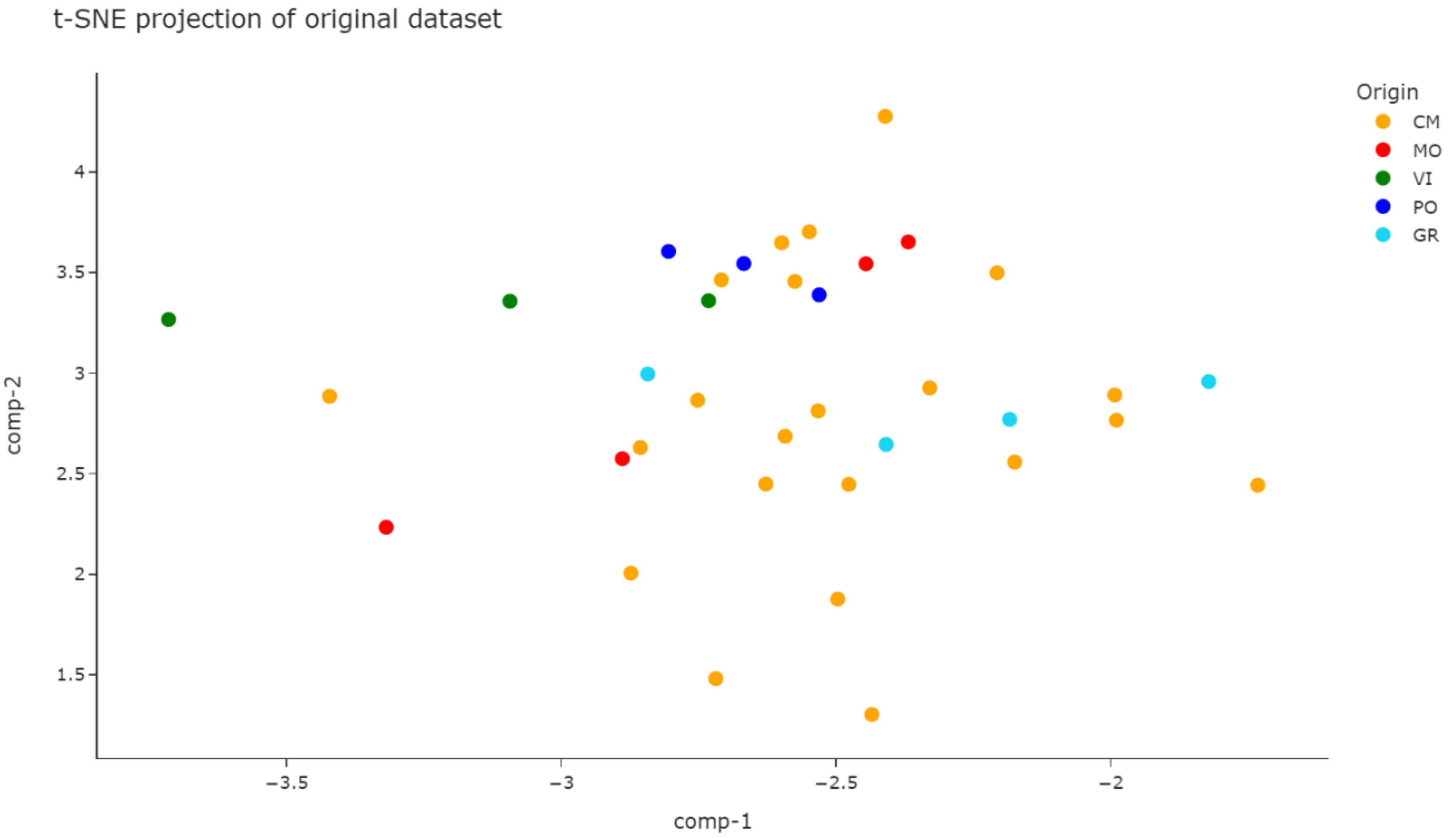

Figure 6. Dimensionality reduction performed by t-SNE approach (CM = Castagnole Monferrato MO = Montemagno VI = Viarigi PO = Portacomaro GR = Grana).

Table 1. Wine samples

Year	CM	GR	MO	PO	VI
2019	ru25				
2020			ru20		ru07, ru11
2021	ru01, ru02, ru03, ru06, ru09, ru12, ru15, ru16, ru18, ru19, ru22, ru23, ru26, ru27, ru28, ru29, ru35, ru37, ru38, ru40, ru42	ru17, ru21, ru24, ru41	ru10, ru39	ru08, ru13, ru14	ru05
2022			ru04		

CM = Castagnole Monferrato; GR = Grana; MO = Montemagno; PO = Portacomaro; VI = Viarigi.

Table 2. List of analytical standards used and molecules quantified by each standard.

Analytical standards	Components considered in quantitative determinations
1-butanol, 3-methyl	aliphatic alcohols
1-butanol, 3-methyl acetate	acetates
benzaldehyde	aldehydes
diethyl succinate	diethyl succinate
ethyl octanoate	esters of fatty acids
octanoic acid	octanoic acid
phenyl ethyl alcohol	phenyl ethyl alcohol
terpineol	terpenes

Table 3. Volatile composition of Ruchè wines: concentration (g kg^{-1}, mean ± standard deviation), calculated and theoretical retention indexes, odor threshold (OTH), and odor activity (OAV) values.

Compound	Mean concentration (g kg^{-1})	Standard deviation	Calculated retention index	Theoretical retention index	OTH (g kg^{-1})	OAV[a]
Esters						
1-butanol, 3-methyl-, acetate	$2.31 \cdot 10^{-3}$	6.4	1123	1122	$3 \cdot 10^{-5}$[b]	77
acetic acid, 2-phenylethyl ester	$4.3 \cdot 10^{-4}$	0.3	1812	1813	$2.5 \cdot 10^{-4}$[b]	1.7
benzeneacetic acid, ethyl ester	$1.1 \cdot 10^{-4}$	0.1	1783	1783	$7.3 \cdot 10^{-5}$[c]	1.5
benzoic acid, methyl ester	$1.7 \cdot 10^{-4}$	0.3	1615	1612	$3 \cdot 10^{-5}$[d]	5.7
butanedioic acid, diethyl ester	$1.761 \cdot 10^{-2}$	10.8	1681	1680	$2.00 \cdot 10^{-1}$[b]	0.088
decanoic acid, ethyl ester	$4.48 \cdot 10^{-3}$	5.0	1639	1638	$2 \cdot 10^{-4}$[b]	22.4
dodecanoic acid, ethyl ester	$9 \cdot 10^{-5}$	0.2	1853	1841	$1.5 \cdot 10^{-3}$[b]	0.06
ethyl acetate	$1.50 \cdot 10^{-3}$	1.5			$7.5 \cdot 10^{-3}$[b]	0.2
ethyl lactate	$1.71 \cdot 10^{-3}$	1.2	1348	1347	$1.4 \cdot 10^{-2}$[b]	0.12
hexanoic acid, ethyl ester	$8.13 \cdot 10^{-3}$	9.5	1236	1233	$1.4 \cdot 10^{-5}$[b]	581
nonanoic acid, ethyl ester	$1.1 \cdot 10^{-4}$	0.2	1539	1531	$1.3 \cdot 10^{-3}$[e]	0.085
octanoic acid, ethyl ester	$2.426 \cdot 10^{-2}$	25.5	1436	1435	$5 \cdot 10^{-6}$[b]	4852

Subtotal	$6.091 \cdot 10^{-2}$					
%	$1.028 \cdot 10^{-2}$					
Alcohols						
1-butanol, 3-methyl-	$4.5505 \cdot 10^{-1}$	246.2	1216	1209	$3.0 \cdot 10^{-2}$[b]	15.2
1-heptanol	$1.27 \cdot 10^{-3}$	1.1	1458	1453	$3 \cdot 10^{-4}$[b]	4.2
1-hexanol	$7.91 \cdot 10^{-3}$	3.7	1355	1355	$8 \cdot 10^{-3}$[b]	0.99
1-hexanol, 2-ethyl-	$5.1 \cdot 10^{-4}$	1.1	1488	1491	N/A[f]	
1-nonanol	$7.4 \cdot 10^{-4}$	0.8	1662	1660	N/A[f]	14.8
1-octanol	$1.05 \cdot 10^{-3}$	0.7	1560	1557	$9 \cdot 10^{-4}$[g]	1.2
1-propanol, 2-methyl-	$3.74 \cdot 10^{-3}$	12.2	1115	1092	4.0×10^{-2}[b]	0.094
phenylethyl alcohol	$5.778 \cdot 10^{-1}$	39.3	1907	1906	$1.4 \cdot 10^{-2}$[b]	4.12
Subtotal	$5.2805 \cdot 10^{-1}$					
%	$8.910 \cdot 10^{-1}$					
Terpenes						
citronellol	$3.0 \cdot 10^{-4}$	0.2	1769	1765	$1 \cdot 10^{-4}$[h]	3
linalool	$5.3 \cdot 10^{-4}$	0.3	1552	1547	$2.5 \cdot 10^{-5}$[b]	21.2
α-terpineol	$4.3 \cdot 10^{-4}$	0.2	1694	1697	$3.3 \cdot 10^{-4}$[b]	1.3
Subtotal	$1.26 \cdot 10^{-3}$					
%	$2.1 \cdot 10^{-4}$					
Aldehydes						
benzaldehyde	$3.3 \cdot 10^{-4}$	0.5	1492	1520	$3.5 \cdot 10^{-4}$[i]	0.94
benzaldehyde, 2-methyl-	$1.09 \cdot 10^{-3}$	1.6	1638	1632	N/A[f]	
furfural	$2.7 \cdot 10^{-4}$	0.4	1466	1461	$1.41 \cdot 10^{-2}$[g]	0.019
Subtotal	$1.68 \cdot 10^{-3}$					
%	$2.8 \cdot 10^{-4}$					
Carboxylic acids						
octanoic acid	$7.4 \cdot 10^{-4}$	0.6			$5 \cdot 10^{-4}$[g]	1.5
%	$1.2 \cdot 10^{-4}$					

[a] Calculated as the ratio concentration/odour threshold value. [b] Retrieved from Ref[7]. [c] Retrieved from Ref.[36] [d] Retrieved from Ref [32]. [e] Retrieved from Ref [37]. [f] N/A = not available. [g] Retrieved from Ref [38] [h] Retrieved from Ref [39]. [i] Retrieved from Ref [40] .