Orthonasal aroma characteristics of Spanish red wines from different price categories and their relationship to expert quality judgements

M.-P. SÁENZ-NAVAJAS[1,2], M. GONZÁLEZ-HERNÁNDEZ[3], E. CAMPO[4], P. FERNÁNDEZ-ZURBANO[1] and V. FERREIRA[4]

[1] Department of Chemistry, University of La Rioja, Institute of Vine and Wine Sciences (UR-CSIC-GR), Madre de Dios 51, E-26006 Logroño, La Rioja, Spain

[2] Centre des Sciences du Goût et de l'Alimentation, UMR5170 CNRS – University of Burgundy – INRA, 9E Boulevard Jeanne d'Arc, 21000 Dijon, France

[3] Department of Food Science, University of La Rioja, Institute of Vine and Wine Sciences (UR-CSIC-GR), Madre de Dios 51, E-26006 Logroño, La Rioja, Spain

[4] Laboratory for Aroma Analysis and Enology, Aragón Institute of Engineering Research (I3A), Department of Analytical Chemistry, Faculty of Sciences, University of Zaragoza, 50009 Zaragoza, Spain

Corresponding author: Professor Vicente Ferreira, fax +34 9 7676 1292, email vferre@unizar.es

Abstract

Background and Aims: Research aimed to define the olfactory sensory spaces of Spanish red wines from different price segments and to determine relationships between sensory descriptors and quality as evaluated by experts.

Methods and Results: A trained panel using a frequency of citation method assessed the orthonasal sensory characteristics of 96 Spanish wines belonging to three price segments (premium; high standard; low standard). Eleven robust sensory terms were built by combining similar terms using statistical criteria. A panel of experts consistently assessed quality. Negative aroma profiles, some of them very skewed, were identified in the different sets with 'animal', 'undergrowth', 'vegetal', 'dried fruit' or 'evolved/oxidized' notes. While skewed aroma profiles were not necessary for achieving low quality, the role of some terms was highly dependent on the category. 'Dried fruit' was positive in premium wines and negative in the other categories, in which its role was played by 'berry'. The low standard wines had to be segmented into wooded and unwooded samples to find clear relationships with quality. Satisfactory models for quality could be built in all segments.

Conclusions: Orthonasal aroma can explain wine quality. The sensory pair 'woody/animal' was confirmed as the most relevant and influential for wine quality. The role of the other terms depended on the category. Oxidation-related attributes are particularly detrimental to younger wines.

Significance of the Study: The results provide a reliable and complete sensory database. They identified aroma profiles related to quality, provided an insight into the hierarchies of terms and provided information that will assist in the identification of the aroma chemicals potentially involved.

Keywords: *berry, oxidation, quality, red wine, sensory analysis, wood*

Introduction

Within the field of food science, the concept of perceived quality has aroused interest for decades. The quality of a product is a parameter that can be complex to define, and different definitions can be found in the literature. A general definition from the International Standards Organization (ISO NORM 9000 2000) is the 'degree to which the inherent characteristics of a product fulfill requirements'. The level of perceived quality has also been defined as the 'judgment about a product's overall excellence or superiority' (Zeithaml 1988).

Wine is a particular case study within the general food and beverages domain. Compared with other beverages, wine is a product that can evoke a wide range of aroma and taste attributes as a result of several variables: grape varieties employed as raw material, winemaking methods, viticultural practices, geographical origins or vintage. However, the intrinsic sensory properties of wine flavour are only one component in the modern consumer definition of quality (Bisson et al. 2002). Extrinsic factors such as bottle and label design are also important drivers of preference in wine selection (Charters and Pettigrew 2007). Notwithstanding, several studies have shown that flavour is a dominating factor for wine choice and quality perception (Keown and Casey 1995, Thompson and Vourvachi 1995), and when purchasing a wine, the flavour is the risk that concerns consumers the most (Mitchell and Greatorex 1988, Spawton 1991).

Understanding wine quality has been repeatedly described to be strongly dependent on the level of consumer involvement (Lawless et al. 1997, Charters and Pettigrew 2007). As stated by Charters and Pettigrew (2007), the importance of the different quality dimensions seems to be related to the informant's level of involvement with wine, with the most-highly involved focusing on more cognitive dimensions and the less-highly involved

doi: 10.1111/j.1755-0238.2012.00195.x

on the more sensory dimensions. For low-involved consumers, the term 'quality' seems to be rather subjective, and what one consumer considers attractive may be perceived as unattractive by another, thus leading to a huge diversity of preferences. On the contrary, highly involved consumers seem to have much in common with wine experts as they assert to divorce the evaluation of quality from pleasure in wine and are inclined to conceptualise wine quality as objective (Charters and Pettigrew 2007). Although some subjective criteria may always exist for highly involved consumers, as for experts, the judgement of product quality is usually based on their previous wine tasting experience and on the alignment of concepts of what is a good example of a grape variety, wine style or quality. It has been reported that experts are able to measure the fit between the different wine exemplars and their memorised models or prototypes (Brochet and Dubourdieu 2001). Thus, overall sensory properties (colour, taste, aroma and mouthfeel) allow experts to elaborate a quality judgement for wine and thus to decide the price range where the product should be located in the market. Obviously, quality concept changes with time and has strong regional differences, although the power of the wine 'gurus' implies the existence of a certain convergence and the dominance of certain quality prototypes (Machado et al. 2011).

Given the importance wine experts exert on the wine market, it is important to study their wine quality perception in order to establish any program for standardisation and improvement. To the best of our knowledge, little is known about the link between quality perceived by wine experts and the aroma attributes shared by products identified as quality exemplars. An interesting approach to deal with this subject is relating the quality scores given by experts with the sensory properties measured by panels of trained observers by descriptive analysis. Similar strategies have been successfully employed in recent studies regarding Uruguayan Tannat wines (Varela and Gambaro 2006), Cabernet Sauvignon and Shiraz wines belonging to the Australian wine market (Lattey et al. 2010) as well as in top-rated dry red wines from diverse origins (Machado et al. 2011).

The present paper is the result of a study aimed at providing some insights into this field by taking a relatively large number of commercial Spanish red wines as a case study. Therefore, a group of trained panellists defined the aroma frame of a large set of Spanish red wines belonging to three price segments selected to encompass a wide range of sensory characteristics. In parallel, a group of experts with long experience in the wine industry evaluated quality with the aim of establishing the most important criteria Spanish experts use to judge intrinsic wine quality linked to orthonasal aroma perception.

Materials and methods

Samples

Three commercially available sets of Spanish red wines were selected based on sales records to obtain a random sample representative of the Spanish red wine market in the year 2009. The first group of wines, called premium, consisted of 24 ranging in retail price from 25€ to 15€, the second group (high standard) consisted of 34 wines ranging from 14€ to 6€ and the third group, called low standard, included 36 wines with price from 5€ to 1€. More details of wine characteristics are shown elsewhere (San Juan et al. 2012).

Descriptive analysis by a trained panel

Panel training. A total of 37 students and staff members from the University of La Rioja (Spain) were recruited on the basis of their interest (they were not paid for their participation) and availability over 25 weeks (1-h session per week). Panellists attended eight descriptive sensory training sessions over a period of 2 months during which they worked in subgroups of five to seven people. They were provided with a list of 110 terms arranged by odour families (fruity, floral, spicy, vegetal, roasted/woody, animal, undergrowth and others) obtained from the literature (Campo et al. 2008). During training, different reference standards representative of aroma were presented and discussed with the panellists. Standards were either commercially available odorants taken from International Flavour and Fragrances (Dijon, France), Sentosphère (Paris, France), Le Nez du Vin (Jean Lenoir, Provence, France), Firmenich (Geneva, Switzerland) or fresh products (fruits, juices, spices, vegetables, etc.) prepared at the beginning of each session as described elsewhere (Sáenz-Navajas et al. 2011b). The training period consisted of two main phases: general (four 1-h sessions) and product-specific training (four 1-h sessions) following the described methodology (Sáenz-Navajas et al. 2011b).

Wine characterisation. For the analysis of the 24 premium wines in duplicate, trained panellist completed a total of five sessions (c. 45 min each) over a period of 5 weeks (24 samples × two replicates). Four formal sessions were completed for the analysis of the 38 high-standard wines (34 samples + 4 replicates: one replicate per session) and another four sessions were carried out for the 40 low-standard wines (36 samples + 4 replicates: one replicate per session). The number of samples presented in each session ranged between nine and ten.

Panellists described the orthonasal odour of each sample according to a citation frequency method (Campo et al. 2010, Sáenz-Navajas et al. 2011b) by choosing a maximum of five attributes from the proposed list of 110 terms. Ten-millilitre wine samples were presented in dark (ISO NORM 3591 1977) wine glasses labelled with three-digit random codes and covered by plastic Petri dishes according to a random arrangement. Formal sessions took place from May to December 2010. All wines were served at room temperature (20–22°C) and evaluated in individual booths. Panellists were not informed about the nature of the samples.

Quality assessment by wine experts

The panel of experts was composed of eight females and fourteen males (30–60 years of age), all of them with a long experience as wine tasters but with different professional backgrounds: seven were aroma researchers (AR), seven were winemakers (WM) and eight were sommeliers (S). Each panellist participated individually in three formal sessions (one for each market segment). Participants had to sort wine samples into five quality categories: exceptional (scored as 5 during data recording), good or very good (scored as 4), acceptable or approved (scored as 3), poor or disappointing (scored as 2) and defective (scored as 1). Panellists were required to smell and taste each sample in the proposed order to minimise any bias introduced by the order of presentation, then experts could taste and smell samples as many times as they considered in the order they preferred. Once the participants had finished the task, they were asked to explain the criteria they used for evaluating wine quality. Experts were informed about the general retail price of each group sample (premium, high- or low-standard wines) before the tasting session, but no more data were disclosed.

Thirty-millilitre samples were presented as for the wine characterisation. All wines were checked for the presence of cork-related taints before being presented for assessment, and if evident, a new bottle was opened. Where the second bottle was free from cork taint, the first bottle was discarded, but if the problem persisted, the sample was included in the experiment because samples representative for the real wine market were required for the experiment.

Data analysis

Quality judgements by experts. A quality index for each wine was computed by averaging all the individual scores obtained for each wine. To check for significant differences in scoring quality between groups of professionals, a two-way analysis of variance (ANOVA) with the kind of expert (AR, WP and S) and wine as fixed factors was performed. Data from one of the experts had to be removed because of outlier behaviour. A two-way ANOVA calculated on the remaining 21 experts revealed no significant differences among the scores given by the different groups of experts ($F = 1.89$; $P = 0.151$). Further, a Principal Component Analysis (PCA) was run for each wine segment on quality perception (judges in columns and wines in rows). Results showed that judges' projections were grouped together in each of the loading plots yielded for the three quality segments (data not shown), thereby confirming consensus among experts in the quality concept.

A one-way ANOVA with repeated measurements was performed on quality scores in order to check differences in quality among samples. The wine effect was significant for quality evaluated in all categories: premium ($F = 5.89$; $P < 0.001$), high standard ($F = 4.90$; $P < 0.001$) and low standard ($F = 8.94$; $P < 0.001$). This indicates that experts are able to differentiate among wines in terms of quality.

Analyses were carried out using SPSS (SPSS Inc., Chicago, Illinois, USA) for Windows, version 15.

Sensory descriptive analysis by a trained panel. To assess the individual performance of panellists, an average reproducibility index (Ri) was calculated for each of the panellists from duplicate assessments of 32 wines (24 premium, 4 high standard and 4 low standard). This parameter, ranging from 0 to 1, has been used in previous works (Campo et al. 2008). The minimum Ri required to keep a judge response was set at 0.20. According to this criterion and attendance, data from 32 panellists (12 males and 20 females, average age 30) were processed in the premium wine set, from 28 participants (9 males and 19 females, average age 30) in the high-standard wine set and from 34 panellists (14 males and 20 females, average age 29) in the low-standard set out of the 37 initial attendants.

Multivariate analysis. Correspondence Analysis (CA) was performed on the contingency table encoded in a wine × aroma attributes table, in which each cell represents the average frequency of citation (FC) for a term in a wine. In order to choose the number of factors that should be retained, dimensions with an eigenvalue higher than the mean eigenvalue (Kaiser law) were calculated for CA spaces. The average (among wine experts) quality score for each wine sample was projected as an illustrative variable on the CA bi-plot.

Hierarchical cluster analysis (HCA) with the Ward criteria was finally applied to the factorial coordinates. The clusters identified by truncating the tree diagram were consolidated by aggregation around mobile centres. The statistical software package used for these analyses was SPAD (version 5.5, CISIA-CESRESTA, Montreuil, France).

Partial Least Square Regression (PLSR1) was used to calculate predictive models for quality from orthonasal aroma descriptors. Therefore, the matrix of descriptive aroma data was firstly centred and scaled in order to obtain normalised data for each of the three groups of wines. All variables were weighted by dividing them by their own standard deviation, thus allowing all variables to contribute equally to the model. The strategy followed for building the models was: an initial model was built for quality using all X variables (average FC given by the trained panel for orthonasal aroma description), after this, an iterative process was begun to reduce the number of X variables in the model, searching for the simplest model with the best prediction ability (Sáenz-Navajas et al. 2010). Therefore, Martens' uncertainty test was used to identify and keep only significant variables. This test considers significant variables for which uncertainty limits do not cross the zero line. Then, the existence of outliers was checked, and samples with a clear deviation from the model were eliminated and kept from the calibration process. The model was then recalculated. A full-cross validation was carried out to estimate the prediction ability of the models for the new set of samples. PLS analysis was performed by using the Unscrambler software (version 9.7, CAMO, Trondheim, Norway).

Results and discussion

Selection of significant terms

The orthonasal aroma descriptors of 94 different wines classified into three different price categories have been measured by using an FC method in which a relatively large number of trained panellists were asked to choose the descriptors that best fit each one of the samples from a previously developed list of 110 aroma terms (Campo et al. 2008). According to judges' consensual decisions, these individual terms were included into seven family categories (fruity, floral, vegetal, spicy, animal, undergrowth, roasted/woody and 'others'). On average, each judge selected 3.6, 4.1 or 4.2 terms out of the list for the description of wines from the premium, high-standard or low-standard sets, respectively. With these data, and using basic combinatory statistics, the expected distributions of frequencies for each term in each one of the wine sets were estimated assuming a completely randomised selection of the terms. Those expected distributions of frequencies were compared with those obtained in each experiment, and calculating the χ^2 distribution, it was possible to determine which terms were used with frequencies higher than those expected by chance. With this criterion, it was possible to determine that 29, 22 or 35 terms were relevant for the definition of the sensory spaces of the three respective sets. These lists of terms (Supporting information: Tables S1–S3) made it possible to obtain fine descriptions of the wines in the sets, but the comparison between wines in the different sets or even the graphical representation was not easy. Because of this, some of the terms were combined with other terms belonging to the same sensory family category, in order to obtain a reduced number of more general terms reaching higher FCs and higher ranges. Such combined terms were also statistically tested for the two aforementioned criteria. This final list of terms is shown in Table 1. A total of 11 terms was defined. Seven of these, 'berry', 'dried fruit', 'roasted', 'woody', 'spicy', 'alcohol' and 'animal', were common to the three categories, although the exact definition of the combined term may differ in some cases. For instance, the term 'dried fruit' does not include the 'date' term in the two standard categories, the 'berry' term includes specifically 'raspberry' and 'strawberry' only in the low-standard category, in which the term 'spicy' also becomes more complex (Table 1). The term 'floral' was found

Table 1. Combined terms (Cx) formed by x individual attributes (significance in brackets) with their significance according to the χ^2 distribution, their range of frequency of citation (FC) (max–min, expressed as %) and average of FC (expressed as %).

	PREMIUM	HIGH STANDARD	LOW STANDARD
	Berry (C3)	**Berry (C3)**	**Berry (C5)**
	Red fruits ($P < 0.001$)	Red fruits ($P < 0.001$)	Red fruits ($P < 0.001$)
	—	—	Raspberry ($P < 0.001$)
	—	—	Strawberry ($P < 0.001$)
	Black fruits ($P < 0.001$)	Black fruits ($P < 0.001$)	Black fruits ($P < 0.001$)
	Blackberry ($P < 0.001$)	Blackberry ($P < 0.001$)	Blackberry ($P < 0.001$)
Significance	0.034	<0.001	<0.001
Max–Min (%FC)	20.3	35.7	38
Average (%FC)	17.6	22.8	38.2
	Dried fruit (C4)	**Dried fruit (C3)**	**Dried fruit (C3)**
	Dried fruits ($P < 0.001$)	Dried fruits ($P < 0.0001$)	Dried fruits ($P < 0.001$)
	Date ($P = 0.048$)	—	—
	Prune ($P = < 0.001$)	Prune ($P < 0.001$)	Prune ($P < 0.001$)
	Dried fig ($P < 0.001$)	Dried fig ($P = 0.063$)	Dried fig ($P < 0.001$)
Significance	<0.001	<0.001	<0.001
Max–Min (%FC)	37.5	39.3	41
Average (%FC)	23.8	28.8	26.1
	Roasted (C5)	**Roasted (C5)**	**Roasted (C4)**
	Roasted ($P = 0.846$)	Roasted ($P < 0.001$)	—
	Caramel ($P < 0.001$)	Caramel ($P < 0.001$)	Caramel ($P < 0.001$)
	Coffee ($P < 0.001$)	Coffee ($P < 0.001$)	Coffee ($P < 0.001$)
	Toasted bread ($P < 0.001$)	Toasted bread ($P < 0.001$)	Toasted bread ($P < 0.001$)
	Smoky ($P < 0.0001$)	Smoky ($P < 0.001$)	Smoky ($P < 0.001$)
Significance	<0.001	<0.001	<0.001
Max–Min (%FC)	26.6	46.4	41
Average (%FC)	33.1	40.2	28.3
	Woody (C2)	**Woody (C2)**	**Woody (C2)**
	Woody ($P < 0.001$)	Woody ($P < 0.001$)	Woody ($P < 0.001$)
	New wood ($P < 0.001$)	New wood ($P < 0.001$)	New wood ($P < 0.001$)
Significance	<0.001	<0.001	<0.001
Max–Min (%FC)	23.4	39.3	31
Average (%FC)	22.7	27.7	14.6
	Spicy (C4)	**Spicy (C5)**	**Spicy (C7)**
	Spicy ($P < 0.001$)	Spicy ($P < 0.001$)	Spicy ($P < 0.001$)
	Vanilla ($P < 0.001$)	Vanilla ($P < 0.001$)	Vanilla ($P < 0.001$)
	Liquorice ($P < 0.001$)	Liquorice ($P < 0.001$)	Liquorice ($P < 0.001$)
	Menthol/fresh ($P < 0.001$)	Menthol ($P < 0.001$)	Menthol/fresh ($P < 0.001$)
	—	Black pepper ($P < 0.001$)	Black pepper ($P < 0.001$)
	—	—	Clove ($P = 0.172$)
	—	—	Nutmeg ($P < 0.001$)
Significance	<0.001	<0.001	<0.001
Max–Min (%FC)	25	35.7	50
Average (%FC)	27.7	39.8	41.1
	Alcohol (C1)	**Alcohol (C1)**	**Alcohol (C1)**
Significance	<0.001	<0.001	<0.001
Max–Min (%FC)	17.2	39.3	26
Average (%FC)	16.7	19.9	11.8

Table 1. *Continued*

	PREMIUM	HIGH STANDARD	LOW STANDARD
	Animal (C2)	**Animal (C1)**	**Animal (C3)**
	Animal ($P = 0.203$)	—	Animal ($P = 0.804$)
	Leather ($P < 0.001$)	Leather($P < 0.001$)	Leather ($P < 0.001$)
	—	—	Cat urine ($P < 0.001$)
Significance	<0.001	<0.001	<0.001
Max–Min (%FC)	43.6	32.1	34
Average (%FC)	14.3	13.3	14.1
	Vegetal (C4)		**Vegetal (C4)**
	Vegetables ($P < 0.001$)		Vegetables ($P = 0.0023$)
	Bell pepper ($P = 0.025$)		Olive ($P = 0.143$)
	Asparagus ($P = 0.006$)		Backed potato ($P = 0.10$)
	Green beans ($P = 0.067$)		Green beans ($P = 0.003$)
Significance	<0.001		0.0258
Max–Min (%FC)	25		38
Average (%FC)	15		17.1
	Floral (C1)	**Floral (C1)**	**—**
Significance	<0.001	<0.001	—
Max–Min (%FC)	12.5	25	—
Average (%FC)	8.72	12.2	—
	Undergrowth (C3)		**Evolved/oxidised (C6)**
	Humus/earthy ($P < 0.001$)		Apple ($P = 0.0255$)
	Mouldy ($P = 0.396$)		Quince ($P = 0.0181$)
	Mushroom ($P = 0.112$)		Honey ($P = 3.98E\text{-}04$)
			Bitter almond ($P = 0.0266$)
			Candied fruits ($P < 0.001$)
			Vinegar ($P = 0.053$)
			Cauliflower ($P < 0.001$)
Significance	<0.0001	—	<0.0001
Max–Min (%FC)	31.3	—	68
Average (%FC)	11	—	20.4

—, terms with no significant distribution of frequencies.

significant only in the premium and high-standard categories, the 'vegetal' only in the premium and low standard, and two other terms, 'undergrowth' and 'evolved/oxidized' were used to describe only premium or low-standard wines, respectively.

Sensory space defined by orthonasal aroma description

A closer look at data in Table 1 gives some clues about differences between categories. It is noteworthy that the premium category seems to be more homogeneous and less aroma explicit because it had the smallest FC ranges (max–min) in all terms except for 'animal' and reached smaller average FC in nearly all cases. On the other hand, the low-standard category seems to be the most heterogeneous and aroma explicit, as suggested by both the highest number of significant individual terms (35 terms) used for its definition and by the widest ranges in five of its combined descriptors. This could be partly explained by the fact that the low-standard category has the greatest heterogeneity in terms of wine production regions (24 denominations of origin) followed by the high-standard group (21) and the low-standard category (13), which could lead to a wider range of aroma profiles as has been published by other authors (Parr 2009, Green et al. 2011).

The ten, eight and nine final significant terms were taken into account for CA in three respective sets. The projection of wines and terms into their corresponding bidimensional CA maps can be seen in Figure 1. HCA was calculated on all CA factors to identify the potential existence of clusters within each set. The sensory properties of each cluster are summarised in Tables 2, 3 and 4 and the corresponding spider web diagrams are given in Figure 2.

Premium wine set. The first three factors of the CA are significant and explain 75% of the original variance. The first factor is mainly defined by the attribute 'animal' (contributes 48%), the second factor, explaining more than 21% of the total variance, is mainly driven by the terms 'vegetal' (30%) and 'animal' (24%), while the third factor, representing 12% of variance, is primarily defined by the attribute 'undergrowth' (50%). HCA showed that this set contains five stable clusters, two of them containing a single sample (Table 2 and Figure 2a). As can be seen, the most important differences lie on the attributes 'animal', 'vegetal' and 'undergrowth' which in all cases are highest in the samples 823 (cluster 4, highest in 'vegetal' and 'undergrowth') and 984 (cluster 5, highest in 'animal'), which are clearly the most different samples in the dataset. Both samples

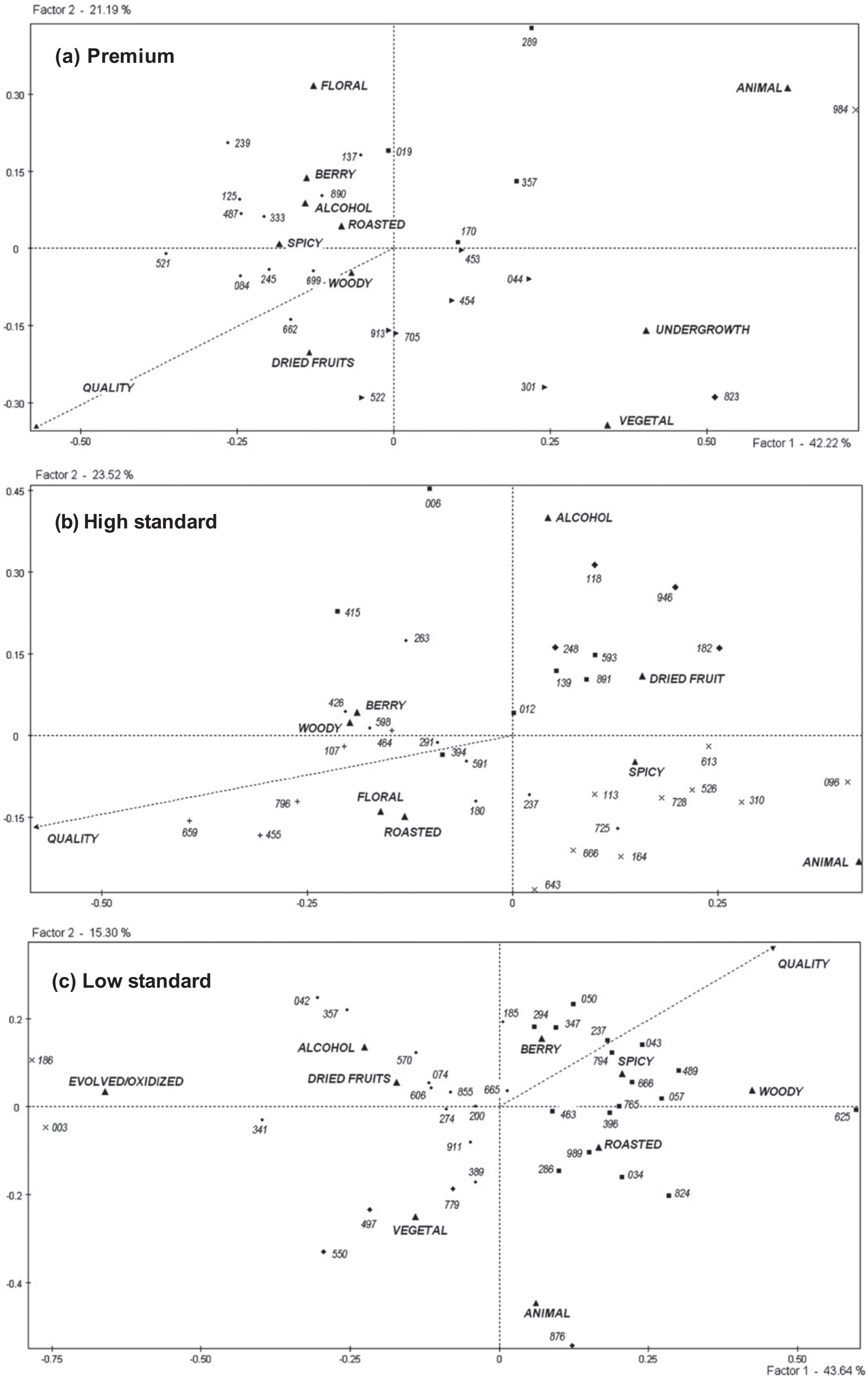

Figure 1. Projection of aroma descriptors and (a) premium wines in the Correspondence Analysis (CA) space (dimensions 1 and 2). Cluster 1 (●), 2 (▶), 3 (■), 4 (◆) and 5 (✕) yielded by the HCA, (b) high-standard wines in the CA space (dimensions 1 and 2). Cluster 1 (■), 2 (◆), 3 (●), 4 (+) and 5 (✕) yielded by the hierarchical cluster analysis (HCA), and (c) low-standard wines in the CA space (dimensions 1 and 2). Cluster 1 (✕), 2 (◆), 3 (●) and 4 (■) yielded by the HCA. The arrow (illustrative variable) shows quality projection according to the categorisation task carried out by the panel of wine experts.

Table 2. Average frequency of citation expressed in % (error deviations are calculated as $s/(n)^{1/2}$; s, standard deviation; *n*, number of panellists) of the significant combined terms describing each cluster of the premium wines and their significance (*P*).

Cluster	1	2	3	4	5	*P**
Samples	333†, 239, 487, 084, 699, 125, 137, 890, 245, 662, 521	522, 454, 453, 44, 913, 705†, 301	170, 019, 289, 357†	823	984	
Berry	20.0 ± 1.3[a]	14.4 ± 0.3[ab]	19.7 ± 0.9[a]	9.4 ± 0.8[b]	12.5 ± 0.8[b]	0.097
Dried fruit	25.3 ± 0.9[a]	27.8 ± 2.2[a]	15.6 ± 1.6[bc]	21.9 ± 1.6[ab]	12.5 ± 1.6[c]	0.083
Roasted	34.4 ± 1.3[a]	34.4 ± 0.9[a]	33.1 ± 2.5[a]	23.4 ± 1.2[ab]	17.2 ± 1.2[b]	0.101
Woody	23.4 ± 0.9	22.2 ± 0.9	22.1 ± 1.9	26.6 ± 1.2	10.9 ± 1.2	0.294
Spicy	31.6 ± 0.9[a]	26.9 ± 1.3[a]	21.3 ± 1.9[ab]	14.1 ± 1.6[b]	15.6 ± 1.6[b]	0.091
Alcohol	20.0 ± 0.6[a]	13.8 ± 0.9[ab]	12.2 ± 1.3[b]	12.5 ± 0.4[b]	14.1 ± 0.4[ab]	0.081
Animal	7.8 ± 0.6[d]	14.4 ± 0.9[c]	17.2 ± 1.9[bc]	21.9 ± 0.8[b]	46.9 ± 0.8[a]	**<0.001**
Vegetal	9.7 ± 0.6[c]	24.4 ± 0.6[ab]	10.0 ± 1.3[c]	29.7 ± 1.2[a]	20.3 ± 1.2[b]	**<0.001**
Undergrowth	7.2 ± 0.6[b]	10.6 ± 0.6[b]	10.0 ± 1.3[b]	34.4 ± 1.2[a]	10.9 ± 1.2[b]	**<0.001**
Floral	10.0 ± 0.6	6.6 ± 0.6	12.5 ± 1.3	3.1 ± 0.8	3.1 ± 0.8	0.121
Quality	3.3 ± 0.1	3.0 ± 0.1	2.6 ± 0.2	2.6 ± 0.2	1.9 ± 0.2	0.186

† Wines closest to the centre of gravity of the cluster. **P*-values in bold indicate the existence of significant differences between clusters ($\alpha < 0.05$) (Student–Newmans–Keuls test). Different letters mean significant differences in the frequency of citation for each term among clusters.

Table 3. Average frequency of citation expressed in % (error deviations are calculated as $s/(n)^{1/2}$; s, standard deviation; *n*, number of panellists) of the significant combined terms describing each cluster of the high-standard wines and their significance (*P*).

Cluster	1	2	3	4	5	*P**
Samples	012†, 139, 891, 593, 415, 394, 426, 006	182†, 248, 118, 946	591†, 291, 180, 598, 237, 263, 725	796†, 107, 464, 455, 659	526†, 164, 728, 310, 613, 666, 113, 096, 643	
Berry	24.1 ± 1.4[ab]	17.0 ± 1.2[b]	28.1 ± 1.5[a]	28.6 ± 1.4[a]	17.1 ± 1.2[b]	**0.013**
Dried fruit	25.4 ± 1.2[b]	42.9 ± 1.0[a]	31.1 ± 1.5[b]	25.7 ± 1.7[b]	25.4 ± 1.4[b]	**0.005**
Roasted	30.4 ± 1.5[b]	32.1 ± 1.5[b]	51.0 ± 1.0[a]	55.0 ± 1.8[a]	35.7 ± 2.1[b]	**<0.001**
Woody	29.5 ± 1.3[ab]	22.3 ± 0.8[b]	28.1 ± 1.7[ab]	38.6 ± 1.7[a]	22.2 ± 1.4[b]	**0.008**
Spicy	39.3 ± 1.3	36.6 ± 1.5	37.2 ± 1.2	36.4 ± 1.7	45.6 ± 2.0	0.207
Alcohol	27.7 ± 1.4[a]	27.7 ± 0.6[a]	17.3 ± 1.6[b]	14.3 ± 1.3[b]	14.7 ± 1.1[b]	**0.001**
Animal	9.8 ± 0.9[bc]	10.7 ± 1.0[bc]	12.8 ± 0.5[b]	5.7 ± 0.6[c]	22.2 ± 0.9[a]	**<0.001**
Floral	16.1 ± 1.1[a]	5.4 ± 0.7[c]	7.7 ± 0.8[bc]	17.9 ± 0.8[a]	12.3 ± 1.0[ab]	**0.001**
Quality	3.0 ± 0.1[ab]	2.6 ± 0.1[b]	3.2 ± 0.1[a]	3.1 ± 0.1[ab]	2.7 ± 0.1[ab]	**0.043**

† Wines closest to the centre of gravity of the cluster. **P*-values in bold indicate the existence of significant differences between clusters ($\alpha < 0.05$) (Student–Newmans–Keuls test). Different letters mean significant differences in the frequency of citation for each term among clusters.

have also low scores in some attributes with positive character, such as 'berry' or 'spicy' and have also low scores in 'quality'. The olfactometric and chemical analysis of these samples revealed that sample 984 had a large content of 4-ethylphenol (1214 µg/L) and that sample 823 contained 18 µg/L of methanethiol (San Juan et al. 2011), which may explain the skewed aroma profiles of these samples that can be considered as prototypes of atypical defective profiles. It should be noted, however, that they are not the samples with lowest scores in 'quality', which certainly tells us that it is not necessary to have a skewed aroma profile to achieve low 'quality' scores.

The differences between clusters 1, 2 and 3 are smaller in magnitude, but still are larger enough to indicate three different sensory profiles, as suggested by the different positions of the samples in the CA plane (Figure 1). Differences again affect the attributes 'animal' ($P < 0.001$), 'vegetal' ($P < 0.002$) and 'undergrowth' ($P < 0.0001$) and also 'dried fruit' ($P < 0.05$). Cluster 1 is the most homogeneous and it is characterised by the lowest scores in 'animal', 'vegetal' and 'undergrowth' and by high scores in 'dried fruit'. Cluster 2 has maxima scores in 'vegetal' and 'dried fruit', and intermediate in 'undergrowth' and 'animal'. Cluster 3 is the least homogeneous and has maxima scores in 'animal' and 'undergrowth' and minima in 'dried fruit' and 'vegetal', and has also the lowest 'quality' scores, although differences in this parameter are not significant. The lack of significance must be attributed to the large variability in the 'quality' scores of samples in this cluster.

High-standard wine set. The CA yielded three significant factors explaining 70% of the original variance. Figure 1 shows the bi-plot with the first two factors. The first dimension (c. 30% of variance) is mainly driven by the term 'animal', while the

Table 4. Average frequency of citation expressed in % (error deviations are calculated as $s/(n)^{1/2}$; s, standard deviation; n, number of panellists) of the significant combined terms describing each cluster of the low-standard wines and their significance (P).

Cluster	1	2	3	4	5
Samples	003†, 186	497†, 779, 550, 876	274†, 074, 855, 606, 665, 570, 911, 185, 200, 341, 042, 389, 357	666†, 057, 043, 463, 989, 489, 050, 396, 237, 294, 794, 034, 347, 765, 286, 824, 625	P*
Berry	22.8 ± 1.2[b]	26.5 ± 1.5[b]	36.5 ± 1.2[a]	39.7 ± 1.5[a]	**0.010**
Dried fruit	43.4 ± 0.9[a]	26.5 ± 1.1[b]	23.8 ± 1.7[b]	25.9 ± 1.4[b]	**0.044**
Roasted	23.5 ± 0.7	27.2 ± 2.6	23.6 ± 1.2	32.7 ± 1.7	0.078
Woody	2.2 ± 0.2[c]	7.4 ± 0.7[bc]	11.1 ± 0.8[b]	20.4 ± 1.2[a]	**<0.001**
Spicy	16.2 ± 0.4[c]	32.4 ± 1.5[b]	36.5 ± 1.5[b]	49.6 ± 1.4[a]	**<0.001**
Alcohol	16.9 ± 1.6[a]	7.4 ± 0.7[ab]	14.8 ± 1.1[ab]	9.9 ± 0.8[b]	**0.031**
Animal	9.6 ± 2.0[b]	26.5 ± 1.0[a]	11.7 ± 0.9[b]	13.7 ± 1.0[b]	**0.001**
Vegetal	16.9 ± 0.5	25.0 ± 2.6	18.9 ± 1.2	13.9 ± 1.0	0.059
Evolved/oxidised	64.7 ± 0.7[a]	23.5 ± 2.1[b]	25.3 ± 1.3[b]	10.7 ± 0.9[c]	**<0.001**
Quality	1.6 ± 0.1[b]	2.2 ± 0.1[b]	2.9 ± 0.1[a]	2.8 ± 0.1[a]	**0.005**

† Wines closest to the centre of gravity of the cluster. *P-values in bold indicate the existence of significant differences between clusters ($\alpha < 0.05$) (Student–Newmans–Keuls test). Different letters mean significant differences in the frequency of citation for each term among clusters.

second factor (24% of variance) is primarily constituted (58%) by the individual term 'alcohol'. The third factor or dimension is mainly built by the terms 'floral' (35%), 'roasted' (26%) and 'dried fruit' (25%). HCA showed again that this set can be classified into five groups, defined in Table 3 and Figure 2b. Differences between clusters in this category are rather clear because all terms except for 'spicy', scored significantly different between clusters. As Figures 1, 2b and Table 3 show, clusters 2, 4 and 5 are the most different. Clusters 2 and 5 have in common low scores in 'berry' and 'woody', low in 'roasted' and also the fact that both clusters score highest in a unique sensory note: 'dried fruit' in the case of cluster 2 and 'animal' in the case of cluster 5. The higher level of 'dried fruit' in the wines in cluster 2 is consistent with the fact that they contain the highest level of methional found in this set, while four of the samples in cluster 5 with higher animal note contained a high level of 4-ethylphenol and 4-ethylguaicol. Cluster 2 is additionally characterised by high scores in 'alcohol' and very low in 'floral', while cluster 5 is very low in 'dried fruit'. Both clusters are also significantly rated low in 'quality'. On the other hand, cluster 4 is characterised by highest scores in 'berry', 'roasted', 'woody' and 'floral', and the lowest scores in 'animal' and 'alcohol'. Cluster 3 is relatively close to this cluster 4, although scores in 'animal' are slightly higher and in 'floral' slightly lower. These two clusters are scored very high in 'quality'. Cluster 1, finally, has intermediate scores in 'berry', 'woody' and 'animal', highest in 'alcohol' and lowest in 'roasted' and 'dried fruit'.

Low-standard wine set. This set has as a major difference the prominence in some samples of the term, not used in the two other categories, 'evolved/oxidised'. This term is formed by combining the terms apple, quince, honey, bitter almond, candied fruits, vinegar and cauliflower, all of them related to oxidation or evolution processes often linked to the formation of aldehydes such as acetaldehyde which supplies ripe apple notes (Azzara and Campbell 1992), methional, which is related to the raisin (San Juan et al. 2011) and cauliflower aromas (Breme et al. 2009), phenylacetaldehyde, responsible for honey-like notes (Ferreira et al. 2002) or old wood nuances of aged red wines (Aznar et al. 2003).

For this set of wines, the CA reveals that the two first factors are significant and explain 60% of the original variance as shown in Figure 1. The first factor (44% of variance) is mainly defined by the term 'evolved/oxidised' while the second factor (15% of variance) is primarily driven by the animal attribute (50%). HCA calculated with all CA factors yielded four stable clusters whose properties are summarised in Figure 2c and Table 4. Differences between clusters are again quite remarkable, because the scores of nearly all attributes significantly differ between them. Cluster 1 groups the two most different samples, scoring highest in 'dried fruit', 'alcohol' and 'evolved/oxidized', and low in 'berry', 'woody', 'spicy' and 'quality'. These two samples can be considered characteristic of a clear oxidative off-odour, and their chemical analysis revealed the largest amount of acetaldehyde (60 and 105 mg/L) and of other aldehydes, such as E-2-nonenal, phenylacetaldehyde, methional and isovaleraldehyde (San Juan et al. 2012). Cluster 2 contains samples also rated low in 'quality' and scoring highest in 'animal' and 'vegetal' and low in 'berry'. The chemical composition did not reveal the existence of a consistent pattern, although two samples in the cluster had around 6 ng/L of isobutyl-2-methoxypyrazine and a third one a relatively high amount of 4-ethylphenol, 4-ethylguaiacol and of isovaleraldehyde which may be related to the observed aroma profiles. Both clusters 3 and 4 contain samples scoring well in 'quality' and characterised by high 'berry', low 'dried fruit', 'animal', 'vegetal' and 'evolved/oxidised' nuances. Cluster 4 is further characterised by the highest 'woody', 'spicy' and lowest 'evolved/oxidised' notes.

Comparison between categories. The aroma profiles of the 58 samples with 'quality' scores higher than 2.7 were compared to assess the influence of the price category on the seven common aroma terms. The plot of these results (Figure 3) reveals that, as expected, cheap (and young) wines have a different aroma profile characterised by higher levels of 'berry' and smaller scores for 'alcohol', 'roasted' and 'woody' notes. What is more interesting is the fact that young wines have significantly higher scores for 'spicy' notes than those of the premium category. This term was also more complex in the low-standard category, because it included citations from six

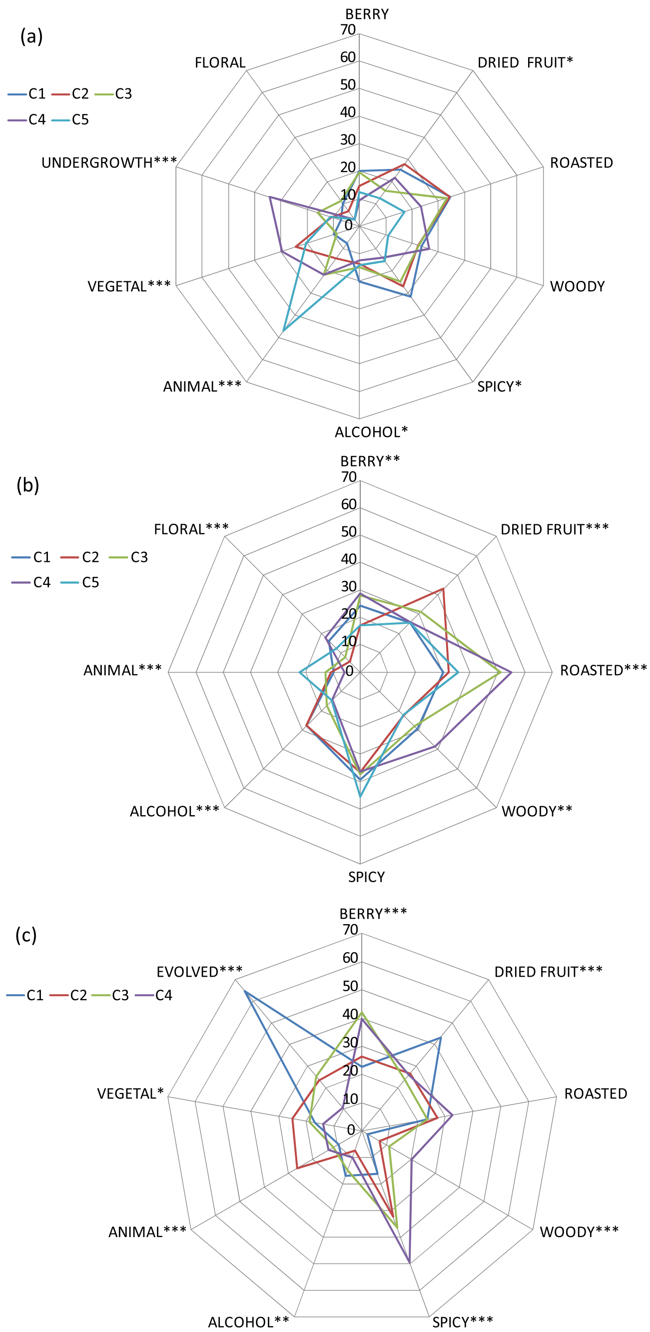

Figure 2. Sensory description expressed as the average frequency of citation (%) of attributes characterising samples belonging to each cluster (C1–C5) in the (a) premium, (b) high-standard and (c) low-standard wine set with their significance (P): $*P < 0.1$; $**P < 0.01$; $***P < 0.001$. (—) C1, (—) C2, (—) C3, (—) C4, (—) C5.

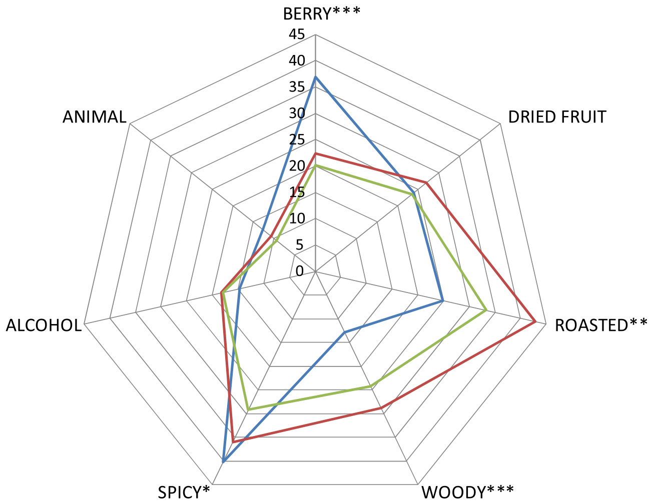

Figure 3. Sensory description expressed as the average frequency of citation (%) of attributes characterising samples with quality scores higher than 2.7 for low-standard (blue, —), high-standard (red, —) and premium (green, —) sets with their significance (P): $*P < 0.05$; $**P < 0.01$; $***P < 0.001$.

different individual aroma terms (Table 1). Also, the level of 'dried fruit' and 'animal' did not differ between categories. A second relevant observation is that the average aroma profile of a high-standard wine is quite similar to that of an average premium wine, although the former has higher scores for 'roasted', 'woody' and 'spicy' notes than the latter. The highest scores for berry notes in young wines can be easily interpreted in terms of their higher content in β-damascenone, fatty acid ethyl esters and fusel alcohol acetates and lower concentration of ethanol. Their higher scores for 'spicy' notes, however, did not have an easy chemical interpretation. Similarly, the higher 'roasted', 'woody' and 'spicy' scores observed in high-standard wines cannot be explained in terms of a higher content of wood extractables, because premium wines have a higher level of these compounds (San Juan et al. 2012). This effect, congruent with the practical observation that 'woody' notes tend to decrease during bottle ageing, must then be attributed to the existence of different perceptual and chemophysical interactions linked to a longer ageing process. We have recently observed (unpublished) that 'woody' notes are strongly linked not only to the presence of wood extractables, but to the presence of β-damascenone and of branched acids (isobutyric, 2 and 3-methylbutyric), compounds that are more concentrated in the high-standard category (San Juan et al. 2012).

It is also noteworthy that the average FC scores for the sensory attributes describing premium wines are in all cases lower than those in the two other categories (Figure 3) in apparent contrast with the fact that 'quality' scores of premium wines are comparatively higher. This may suggest that expert 'quality' scores are more related to the aroma harmony reached by the mixture of different sensory attributes than by the presence of a predominant aroma note in wine.

Correlation between orthonasal aroma descriptors and quality perception

The relationship between the aroma attributes describing wine samples and the quality perception evaluated by wine experts was investigated by simple correlation analysis (Table 5). Results show that in the premium group, the attributes 'spicy', 'dried fruit', 'woody' and, to a lesser extent, 'alcohol' were positive and significantly correlated to quality perception, while the term 'animal' was negatively correlated. Similarly, in the high-standard segment, 'quality' was negatively correlated to the term 'animal', while the attributes positively correlated to 'quality' were 'woody', 'roasted' and 'berry'. Again, in the low-standard samples, the 'animal' descriptor was negatively correlated to 'quality', as well as the term 'evolved/oxidised' that presents a similar trend. On the contrary, attributes such as 'woody', 'roasted' and 'berry' were positively correlated to 'quality'.

It should be remarked that 'quality' seems to be more related to aroma descriptors in the premium and high-standard wine sets than in the low standard, in which correlation

Table 5. Correlation (Pearson coefficient, *r*, and significance, *P*) between attributes assessed by the trained panel and quality perception of experts in the three wine sets.

	Premium		High standard		Low standard	
	r	*P*	*r*	*P*	*r*	*P*
Berry	0.239	0.260	**0.308**	**0.081***	**0.379**	**0.022****
Dried fruits	**0.581**	**0.003*****	−0.273	0.124	−0.265	0.118
Roasted	0.165	0.440	**0.476**	**0.005*****	0.199	0.244
Woody	**0.525**	**0.008*****	**0.619**	**0.0001*******	**0.358**	**0.032****
Spicy	**0.691**	**0.0002*******	0.050	0.782	0.243	0.153
Alcohol	**0.374**	**0.072***	−0.088	0.627	0.070	0.686
Animal	**−0.584**	**0.003*****	**−0.344**	**0.0498****	**−0.302**	**0.077***
Vegetal	−0.139	0.561	—	—	−0.229	0.179
Undergrowth	−0.169	0.430	—	—	—	—
Floral	−0.085	0.692	—	—	—	—
Evolved/oxidised	—	—	—	—	**−0.456**	**0.005*****

*$P < 0.1$; **$P < 0.05$; ***$P < 0.01$; ****$P < 0.001$. Significant P values ($\alpha < 0.05$) are highlighted in bold. —, no significant combined term.

coefficients were much smaller. Furthermore, in this case, correlations were artificially inflated by the six samples in the two clusters of lower quality (Table 4) and, in fact, if these six samples are removed, the correlations with 'quality' completely disappear. Things completely changed in this set when the quality study was carried out separately on wooded and unwooded samples. In the subset containing the 24 samples with low scores for 'woody' (FC < 6), 'quality' was found to be significantly correlated with 'berry' ($r = 0.50$, $P = 0.005$), 'evolved/oxidized' ($r = -0.49$, $P = 0.006$) and 'dried fruit' ($r = -0.35$, $P = 0.07$). Similar results were obtained when the group of samples included in cluster 4 (rich in 'wood') was excluded. In this 18 sample subset, 'quality' was significantly correlated to 'berry' ($r = 0.63$; $P = 0.004$), 'evolved/oxidized' ($r = -0.64$; $P = 0.003$), 'vegetal' ($r = -0.41$, $P = 0.08$) and 'dried fruit' ($r = -0.39$, $P = 0.09$). In the set of 12 wooded samples (arbitrarily selected as those with scores for 'woody' equal or higher than 6), 'quality' was found to be significantly correlated to the terms 'woody' ($r = 0.53$, $P = 0.07$) and 'animal' ($r = -0.67$, $P = 0.01$). What this suggests is that the assessment of quality given by experts is strongly influenced by the presence of 'woody' notes. Moreover, it seems that experts handle two completely different quality prototypes in this set as can be seen in Figure 4.

In the absence of 'woody' character, 'berry' and 'dried fruit' are determinant for quality, but they lose such prominence in favour of 'woody' and 'animal', as soon as wood becomes detectable.

The differences in the observed relationship of a given descriptor with 'quality' between sets deserve a further comment. On the one hand, some descriptors, such as 'berry', 'roasted', 'woody' and 'animal', hold a relationship with 'quality' relatively similar in the three different sets and, in general, the higher the range of the attribute in the set (Table 1), the higher the correlation coefficient. This rule is not fulfilled in the case of 'dried fruit' in premium wines, because a high correlation is observed with a small range. On the other hand, the attributes 'spicy' and 'alcohol' were only significantly correlated to 'quality' in the premium wine set, in spite of the fact that it is in this set in which they reach the shortest ranges (Table 1). The pattern found in the attribute 'dried fruit' is still more striking, while ranges and averages are similar in the three

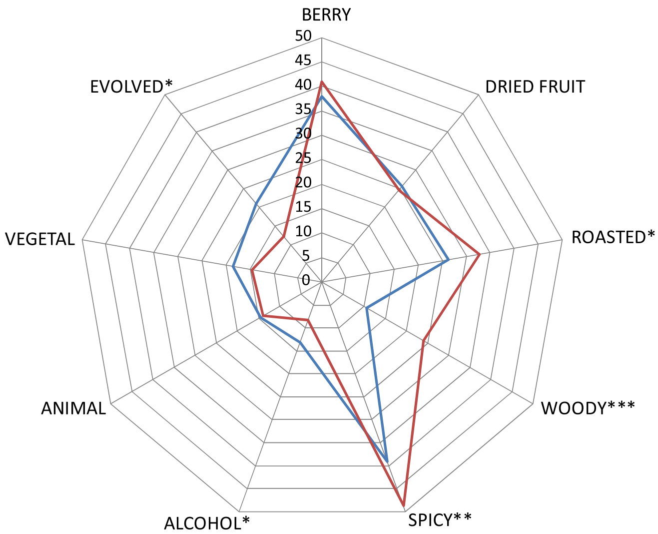

Figure 4. Sensory description expressed as the average frequency of citation (%) of both wooded (blue, —) and unwooded (red, —) wines in the low-standard set with their significance (*P*): *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

categories; this attribute correlates with 'quality' in the premium wine set, while in the two other sets, it holds a negative correlation (Table 5). This seems to further support the existence of different quality prototypes; 'dried fruit' character, which is related in part to oxidation, would be more than tolerated by experts in premium wines while constitutes a rather negative feature in the other sets.

Quality prediction from aroma attributes by PLS regression

PLSR1 regression models have been used for predicting the general quality scores from orthonasal aroma sensory data. In the premium wine set, the best final model is able to explain 67% of the original variance (63% by cross-validation) and has a Root Mean Square Error (RMSE) of 0.36. Four sensory variables are statistically significant in predicting quality according to Marten's criterion: 'spicy', 'woody', 'dried fruit' and 'animal'. The model for quality is:

$$\text{Quality} = 1.302 + 0.059 * \text{Dried fruit} + 0.084 * \text{Spicy} + 0.096 * \text{Woody} - 0.050 * \text{Animal} \quad (1)$$

It is worth mentioning that the model strongly underestimated the quality of sample 239, which had the highest quality score of the set. Leaving aside the difficulty in modelling an extreme sample by cross-validation, the fact is that this sample presented the highest scores for astringency, global intensity and persistence evaluated in-mouth, terms highly correlated to quality perception (Sáenz-Navajas et al. 2011b) and the lowest L_{10}* score (e.g. it is the darkest sample), parameter negatively correlated to quality perception (Sáenz-Navajas et al. 2011a).

PLSR1 calculated on the sensory attributes of high-standard wines provided also a quite good model able to explain 65% of the original variance (64.6% by cross-validation) and with a quite low average prediction error, being the RMSE = 0.24 (0.26 by cross-validation). The model, with two PCs, included also four significant variables:

$$\text{Quality} = 1.865 + 0.035 * \text{Roasted} + 0.056 * \text{Berry} + 0.066 * \text{Woody} - 0.048 * \text{Animal} \quad (2)$$

The model strongly overestimates the quality of samples 107, 946 and 659 and underestimates that of sample 666. These limitations may be attributed to a possible higher weight of colour and in-mouth properties than aroma in the quality assessment of these samples. In fact, wines 107, 946 and 659 present the lowest scores for sourness, global intensity or persistence (Sáenz-Navajas et al. 2011b), moreover, sample 107 is the least dark sample (highest L_{10}*), while sample 666 results to be the most persistent wine in-mouth and the darkest sample (lowest L_{10}*) (Sáenz-Navajas et al. 2011a).

Finally, the prediction of quality from aroma attributes in the cheapest set required, as explained before, the division of the set in two subsets attending to the presence or not of woody notes. In the subset containing the 24 samples with low scores for woody (FC < 6), a significant model able to explain 74% of the original variance (66% by cross-validation) was developed, as can be seen in the following equation 3:

$$\text{Quality} = 1.96 + 0.078 * \text{Berry} + 0.063 * \text{Alcohol} - 0.057 * \text{Evolved/oxidised} \quad (3)$$

In the subset of the 18 samples included in the clusters 1, 2 and 3 (Table 4), a similar model explaining 82% of the original variance (64% by cross-validation) could be found:

$$\text{Quality} = 2.42 + 0.072 * \text{Berry} + 0.068 * \text{Alcohol} - 0.057 * \text{Evolved/oxidised} - 0.063 * \text{Vegetal} \quad (4)$$

In the case of wooded samples, a satisfactory model for quality could be built, but it could not be properly validated because of the low number of samples. The model explains 60% of the original variance and is expressed by the following equation:

$$\text{Quality} = 2.80 + 0.105 * \text{Woody} - 0.160 * \text{Animal} \quad (5)$$

The models confirm previous observations and are also consistent with previous models developed for quality and aroma chemical composition (Ferreira et al. 2009, San Juan et al. 2012).

Conclusions

The present large-scale experiment confirms that orthonasal aroma perception has a strong influence on wine quality evaluated by experts. The comparison between categories and the need to split the young wine set into wooded and unwooded samples reveal that within a homogeneous set, quality can be satisfactorily explained by the sensory quantitative scores, but that such quantitative relationships cannot be generalised among sets. This suggests the existence of different quality prototypes, each one characterised by a given sensory profile. Remarkably, premium wines have on average, lower sensory scores but higher quality scores, suggesting that harmony is essential. While the sensory pair 'woody/animal' is confirmed as the most general influence for quality whenever contact with wood has taken place, other similar descriptors have been found to play different roles depending on the wine set. 'Dried fruit' emerges as an important contributor of premium and aged Spanish red wines, while it is perceived as negative in younger wines. The relevance of 'berry' decreases with ageing, becoming essential in unwooded young wines. In these, oxidation and vegetal notes play prominent negative roles.

Acknowledgements

The research reported in this article was financially supported by *Instituto de Estudios Riojanos* and the Spanish Education Ministery (AGL2007-65139 and AGL2010-22355-C02) projects. M.-P.S.-N. and E.C. (Juan de la Cierva) acknowledge the Spanish Education Ministery (M.E.C.) for their postdoctoral fellowships.

References

Aznar, M., López, R., Cacho, J. and Ferreira, V. (2003) Prediction of aged red wine aroma properties from aroma chemical composition. Partial least squares regression models. Journal of Agricultural and Food Chemistry **51**, 2700–2707.

Azzara, C.D. and Campbell, L.B. (1992) Off-flavors of dairy products. In: Off-flavors in foods and beverages. Ed. G. Charalambous (Elsevier: Amsterdam) pp. 329–374.

Bisson, L.F., Waterhouse, A.L., Ebeler, S.E., Walker, M.A. and Lapsley, J.T. (2002) The present and future of the international wine industry. Nature **418**, 696–699.

Breme, K., Langle, S., Fernandez, X., Meierhenrich, U.J., Brevard, H. and Joulain, D. (2009) Character impact odorants from Brassicaceae by aroma extract dilution analysis (AEDA): Brassica cretica and Brassica insularis. Flavour and Fragrance Journal **24**, 88–93.

Brochet, F. and Dubourdieu, D. (2001) Wine descriptive language supports cognitive specificity of chemical senses. Brain and Language **77**, 187–196.

Campo, E., Do, B.V., Ferreira, V. and Valentin, D. (2008) Aroma properties of young Spanish monovarietal white wines: a study using sorting task, list of terms and frequency of citation. Australian Journal of Grape and Wine Research **14**, 104–115.

Campo, E., Ballester, J., Langlois, J., Dacremont, C. and Valentin, D. (2010) Comparison of conventional descriptive analysis and a citation frequency-based descriptive method for odor profiling: an application to Burgundy Pinot noir wines. Food Quality and Preference **21**, 44–55.

Charters, S. and Pettigrew, S. (2007) The dimensions of wine quality. Food Quality and Preference **18**, 997–1007.

Ferreira, A.C.S., de Pinho, P.G., Rodrigues, P. and Hogg, T. (2002) Kinetics of oxidative degradation of white wines and how they are affected by selected technological parameters. Journal of Agricultural and Food Chemistry **50**, 5919–5924.

Ferreira, V., San Juan, F., Escudero, A., Cullere, L., Fernandez-Zurbano, P., Saenz-Navajas, M.P. and Cacho, J. (2009) Modeling quality of premium Spanish red wines from gas chromatography-olfactometry data. Journal of Agricultural and Food Chemistry **57**, 7490–7498.

Green, J.A., Parr, W.V., Breitmeyer, J., Valentin, D. and Sherlock, R. (2011) Sensory and chemical characterisation of Sauvignon blanc wine: influence of source of origin. Food Research International **44**, 2788–2797.

ISO NORM 3591 (1977) Sensory analysis: Apparatus wine tasting glass. (International Organization for Standardization: Geneva, Switzerland).

ISO NORM 9000 (2000) Quality management systems-fundamentals and vocabulary. (International Organization for Standardization: Geneva, Switzerland).

Keown, C. and Casey, M. (1995) Purchasing behaviour in the Northern Ireland wine market. British Food Journal **97**, 17–20.

Lattey, K.A., Bramley, B.R. and Francis, I.L. (2010) Consumer acceptability, sensory properties and expert quality judgements of Australian Cabernet Sauvignon and Shiraz wines. Australian Journal of Grape and Wine Research **16**, 189–202.

Lawless, H., Liu, Y.F. and Goldwyn, C. (1997) Evaluation of wine quality using a small-panel hedonic scaling method. Journal of Sensory Studies **12**, 317–332.

Machado, B., Graça, A., Hirson, G. and Heymann, H. (2011) Revealing the secret preferences for top-rated dry red wines through sensometrics. 16th International Oenology Symposium, Bolzano, Italy (International Association of Enology, Management and Wine Marketing: Germany) pp. 145–149.

Mitchell, V.W. and Greatorex, M. (1988) Consumer risk perception in the UK wine market. European Journal of Marketing **22**, 5–15.

Parr, W. (2009) Geographical indication and the concept of wine typicality. ChemoSense **11**, 16–17.

Sáenz-Navajas, M.P., Fernandez-Zurbano, P., Tao, Y.S., Dizy, M. and Ferreira, V. (2010) Relationship between nonvolatile composition and sensory properties of premium Spanish red wines and their correlation to quality perception. Journal of Agricultural and Food Chemistry **58**, 12407–12416.

Sáenz-Navajas, M.P., Echavarri, F., Ferreira, V. and Fernández-Zurbano, P. (2011a) Pigment composition and color parameters of commercial Spanish red wine samples: linkage to quality perception. European Food Research and Technology **232**, 877–887.

Sáenz-Navajas, M.P., Martín-López, C., Ferreira, V. and Fernández-Zurbano, P. (2011b) Sensory properties of Premium Spanish red wines and their implication in wine quality perception. Australian Journal of Grape and Wine Research **17**, 9–12.

San Juan, F., Ferreira, V., Cacho, J. and Escudero, A. (2011) Quality and aromatic sensory descriptors (mainly fresh and dried fruit character) of Spanish red wines can be predicted from their aroma-active chemical composition. Journal of Agricultural and Food Chemistry **59**, 7916–7924.

San Juan, F., Ferreira, V., Cacho, J. and Escudero, A. (2012) The aroma chemical composition of red wines from different price categories and its relationship with quality. Journal of Agricultural and Food Chemistry **60**, 5045–5056.

Spawton, T. (1991) The wine-marketing environment. European Journal of Marketing **25**, 12–18.

Thompson, K.E. and Vourvachi, A. (1995) Social and attitudinal influences on the intention to drink wine. International Wine Marketing Journal **7**, 35–44.

Varela, P. and Gambaro, A. (2006) Sensory descriptive analysis of Uruguayan Tannat wine: correlation to quality assessment. Journal of Sensory Studies **21**, 203–217.

Zeithaml, V.A. (1988) Consumer perceptions of price, quality and value – a means-end model and synthesis of evidence. Journal of Marketing **52**, 2–22.

Manuscript received: 1 February 2012

Revised manuscript received: 21 May 2012

Accepted: 28 May 2012

Supporting information

Additional Supporting Information may be found in the online version of this article: http://onlinelibrary.wiley.com/doi/10.1111/j.1755-0238.2012.00195.x/abstract

Table S1. Relevant individual terms for the definition of the sensory space of the premium wine set. Maximum citation frequency (Cfmax) for the term among the wines in the set and significance according to the χ^2 criterion.

Table S2. Relevant individual terms for the definition of the sensory space of the high-standard wine set. Maximum citation frequency (Cfmax) for the term among the wines in the set and significance according to the χ^2 criterion.

Table S3. Relevant individual terms for the definition of the sensory space of the low-standard wine set. Maximum citation frequency (Cfmax) for the term among the wines in the set and significance according to the χ^2 criterion.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.