Prediction of Aged Red Wine Aroma Properties from Aroma Chemical Composition. Partial Least Squares Regression Models

ARTICLE *in* JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY · MAY 2003

Impact Factor: 2.91 · DOI: 10.1021/jf026115z · Source: PubMed

4 AUTHORS:

Margarita Aznar
University of Zaragoza
30 PUBLICATIONS **1,147** CITATIONS
SEE PROFILE

Ricardo Lopez
University of Zaragoza
48 PUBLICATIONS **2,031** CITATIONS
SEE PROFILE

Juan Cacho
University of Zaragoza
230 PUBLICATIONS **6,376** CITATIONS
SEE PROFILE

Vicente Ferreira
University of Zaragoza
212 PUBLICATIONS **6,183** CITATIONS
SEE PROFILE

JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Prediction of Aged Red Wine Aroma Properties from Aroma Chemical Composition. Partial Least Squares Regression Models

MARGARITA AZNAR, RICARDO LÓPEZ, JUAN CACHO, AND VICENTE FERREIRA*

Department of Analytical Chemistry, Faculty of Sciences, University of Zaragoza, 50009 Zaragoza, Spain

Partial least squares regression (PLSR) models able to predict some of the wine aroma nuances from its chemical composition have been developed. The aromatic sensory characteristics of 57 Spanish aged red wines were determined by 51 experts from the wine industry. The individual descriptions given by the experts were recorded, and the frequency with which a sensory term was used to define a given wine was taken as a measurement of its intensity. The aromatic chemical composition of the wines was determined by already published gas chromatography (GC)-flame ionization detector and GC-mass spectrometry methods. In the whole, 69 odorants were analyzed. Both matrixes, the sensory and chemical data, were simplified by grouping and rearranging correlated sensory terms or chemical compounds and by the exclusion of secondary aroma terms or of weak aroma chemicals. Finally, models were developed for 18 sensory terms and 27 chemicals or groups of chemicals. Satisfactory models, explaining more than 45% of the original variance, could be found for nine of the most important sensory terms (*wood-vanillin-cinnamon, animal-leather-phenolic, toasted-coffee, old wood-reduction, vegetal-pepper, raisin-flowery, sweet-candy-cacao, fruity*, and *berry fruit*). For this set of terms, the correlation coefficients between the measured and predicted *Y* (determined by cross-validation) ranged from 0.62 to 0.81. Models confirmed the existence of complex multivariate relationships between chemicals and odors. In general, pleasant descriptors were positively correlated to chemicals with pleasant aroma, such as vanillin, β damascenone, or (*E*)-β-methyl-γ-octalactone, and negatively correlated to compounds showing less favorable odor properties, such as 4-ethyl and vinyl phenols, 3-(methylthio)-1-propanol, or phenylacetaldehyde.

KEYWORDS: Wine; aroma; flavor; models; PLS; sensory description

INTRODUCTION

One of the final aims of flavor chemistry is to create mathematical models that allow the establishment of the existing relationship between the content of the product in aroma-active compounds and the sensory properties, mainly aromatic, of the aforementioned product. In the case of wine, this aim is difficult to achieve for several reasons. First, the number of odorants that can be found in a wine in concentrations above the threshold not only is relatively big (up to 50) (*1*−*4*) but also the quantitative analysis of some of these components is extremely complicated and expensive (*1*, *5*), which impedes the analysis of a large number of samples. Second, in many complex wines, there are no impact compounds that determine the aroma. On the contrary, the aroma is due to the mixture of different odorants (*6*−*9*). The study of these systems requires a statistical approximation. Finally, the measurement of the sensory attributes of wines is not exempt of difficulties either, especially when working with wide groups of wines with very complex aromas (*10*), which is the case of the aged red wines approached in the present study. These reasons explain why it is possible to find, in the literature, studies that penetrate the chemical base of aroma, abandoning any statistical approximation and using a limited sensory analysis (*1*, *4*), works that concentrate on the study of the role of just one or a small group of odorants (*11*−*16*), or statistical studies in which the sensory descriptors are explained by means of data on volatile components whose aromatic contribution has not been previously verified (*9*, *17*−*19*).

The strategy followed in the present study lies halfway between the limits defined above. Our aim is to build up models correlating a matrix of sensory data, *Y*, with one of chemical data, *X*, guaranteeing that the number of cases studied is sufficient to produce robust and reliable models. Furthermore, the matrix *X* will only contain data on chemical components that are, or can be, aromatically active in the wines. The modeling technique used here has been partial least squares regression (PLSR), which offers several advantages over the classic regression techniques, as well as over other multivariate

* Author to whom correspondence should be addressed. Phone: 34 976762067. Fax: 34 976761292. E-mail: vferre@posta.unizar.es.

10.1021/jf026115z CCC: $25.00
Published on Web 03/27/2003

techniques used to correlate sets of sensory and chemical data (*20*). This technique has been previously used in the modeling of the aroma of wines of the variety Traminer (*18*), in the measurement of the influence of the de-alcoholization in the aroma of the wine (*21*), in the study of the relationship between the sensory descriptors and hydrolyzed compounds of extracts of precursors (*22*), and in a study to determine volatile components related to the sensory changes observed in Chardonnay wines stored at high temperatures (*9*).

The sensory analysis performed here has been partially inspired by the methodology developed by McCloskey et al. (*17*, *23*), having used panels of judges formed entirely by enologists involved in the production of the wines to study. This election facilitates the later communication of the results to this group of professionals and allows, a priori, the difficulties derived from a sensory evaluation of samples of very complex aroma to be overcome.

The group of wines to study constitutes a more or less homogeneous category of high quality products. These are red wines aged between one and five (or more) years in oak cask, and at least two more in the bottle, produced in prestigious wine-making areas of the North and North-East of Spain. The aroma of these wines has a common component due to the cask aging and a much more diverse part in which different aromatic notes can be recognized. The olfactometric profile of four of these wines has been recently studied (*7*, *8*), so that most of the odorants are known. The experimental design used here assumes several simplifications—not to take into account the interaction of the aromas with the elements of the matrix—and limitations—not all the aromatically active components have been quantified—but we think that some of the conclusions of this study will be of interest to enologists and to flavor chemists.

MATERIALS AND METHODS

Wines. 57 Spanish red aged wines from 7 different Spanish Denominations of Origin: Campo de Borja (7 samples), Cariñena (7 samples), Navarra (9 samples), Penedés (10 samples), Ribera del Duero (9 samples), Rioja (10 samples) and Somontano (5 samples). The wines were selected by experts from each of the Denominations as the most valuable and representative of the region (**Table 5**).

Reagents. *Solvents*. Dichloromethane was purchased from Fischer (Leicester, UK), absolute ethanol was from Panreac (Barcelona, Spain), methanol was from Merck (Darmstadt, Germany), and water was purified in a milliQ system from Millipore (Bedford, MA).

Resins. Lichrolut EN cartridges were supplied by Merck (Darmstadt, Germany).

Standards. Compounds numbered in **Table 1** as **1**−**3**, **6**, **7**, **10**, **12**, **15**, **17**−**19**, **21**, **27**, **28**, **30**, **34**, **36**, **38**, **45**, **48**, **49**, **52**, **55**−**57**, **59**, **60**, **62**, **63**, and **65**−**69** were supplied by Aldrich (Steinheim, Germany); **4** was supplied by Merck (Darmstadt, Germany); **5**, **9**, **26**, **40**, and **47** were supplied by ChemService (West Chester, USA); **8**, **13**, **22**, **23**, **25**, **29**, **32**, **35**, **41**−**44**, **46**, **51**, and **58** were supplied by Fluka (Buchs, Switzerland); **11**, **50**, and **53** were supplied by Sigma (St. Louis, USA); **14**, **16**, **20**, and **24** were supplied by PolyScience (Niles, USA); **31**, **37**, **39**, **54**, **61**, and **64** were supplied by Lancaster (Strasbourg, France); and **33** was supplied by Firmenich (Geneva, Switzerland).

Quantitative Analysis and Selection of the Main Odorants. *Quantitative Analysis of Major Compounds*. The analysis was carried out by the method published by Ortega et al. (*24*). In this analysis, 4.5 g of $(NH_4)_2SO_4$, 3 mL of wine, 7 mL of water, 0.2 mL of dichloromethane, and an internal standard were added to a 15 mL screw-capped centrifuge tube; the tube was first shaken and then centrifuged at 2500 rpm for 10 min. Once the phases were separated, the dichloromethane phase was recovered and injected into a Hewlett-Packard 5890 series II gas chromatograph under the following conditions: The initial temperature was 40 °C, held for 5 min and then raised at 3 °C/min to 200 °C. The carrier gas was H_2 at 3 mL/min. Three microliters was injected in split mode. Split flow was 30 mL/min. The column (50 m × 0.32 mm and 0.5 μm film thickness) was a DB-20 from J&W Scientific (Folsom, CA) and detection was by flame ionization detector (FID).

Quantitative Analysis of Minor and Trace Compounds. The analysis was carried out by the method published by Lopez et al. (*25*). The cartridges were previously conditioned with dichloromethane, methanol, and water/ethanol (12%). After this, 50 mL of wine were passed through a 200 mg SPE cartridge of Lichrolut EN resins at about 2 mL/min. The sorbent was then dried by letting air pass through it. Analytes were recovered by elution with 1.3 mL of dichloromethane. The extract was spiked with the internal standard solution and injected into a 3400CX gas chromatograph fitted to a Saturn 4 electronic impact ion trap mass spectrometer from Varian. The oven initial temperature was set at 40 °C, held for 5 min, and then raised at 2 °C/min to 230 °C. The carrier gas was helium at 1 mL/min. Three μL of extract were injected into a SPI injector from Varian, whose temperature was set initially at 30 °C for 0.6 min and then was raised to 230 °C at 200 °C/min. The column (60 m × 0.25 mm and 0.5 μm film thickness) was a DB-WAXetr from J&W Scientific (Folsom, CA). Calibration graphs were prepared for 69 compounds by the analysis of synthetic samples containing known amounts of odorants.

Sensory Analysis and Selection of the Wine Descriptors. *Sensory Analysis.* The sensory panel, formed by 51 judges, was actually divided into 5 local committees (Ribera Duero, Rioja, Navarra, Aragón, and Cataluña). Each committee was formed by 8 to 12 professional enologists, all of them involved in the production of wines similar to those of the experiment and with a long experience as wine tasters. Each committee analyzed a complete set of the 57 wine samples, having at hand a second bottle to replace spoiled samples. The tasting took place in four sessions of about 1 h in two different days. The tasting sessions were conducted in tasting rooms with isolated booths, following the usual practices in wine tasting, and were conducted by a member of the research group. An incomplete block design was used, with each judge given a different subset of the 57 wines. During each session, 9 different wines and a replicate were given to the judges. Therefore, each judge tasted up to 36 different samples and 4 replicates. The whole experiment was carefully randomized to ensure that all wines were tasted by approximately the same number of judges. Judges were previously informed of the goals of the experiment and were asked to note down the main aromatic descriptors that best defined in their opinion the aroma of the wine. The tasting was not limited to orthonasal aroma evaluation. Judges were asked to proceed following the standard wine tasting procedures (visual inspection, aroma evaluation, in-mouth evaluation, retro-nasal aroma evaluation, and evaluation of aftertaste). To keep to minimum bottle variation, all the bottles used in the study belonged, strictly, to the same batch and were stored in the same place under the same conditions. Before analysis, (chemical or sensorial) the bottle was tasted to search for possible off-flavors and was rejected if some abnormal odor appeared.

Selection of the Wine Descriptors. All the descriptions given by the judges were pooled down into a database. Nondescriptive or abstract terms were directly eliminated, and obvious synonyms were grouped together. After this, terms used less than 30 times were eliminated, except if they defined, specifically, a given wine (minimum frequency arbitrarily set at 4). The performance of the judges was evaluated by the study of the replicate samples. Judges with a mismatch in the terms used higher than 50% were eliminated. The similarity of the terms was studied by cluster analysis using different correlation coefficients (Pearson's, Kendall's, and Spearman's) as similarity measures. In all cases, a sequential agglomerative hierarchical nested cluster analysis (SAHN) was carried out. Complete linkage and un-weighted pair-group (UPGMA) clustering methods were studied using SPSS (SPSS Inc., Chicago, IL) and NTSYS (Exeter Software, Setauket, NY). Correlated terms defining similar, or at least not very dissimilar, aromas were clustered together if the regression model for the global term was better than any of the models for the single terms.

Modeling Sensorial Descriptors from Analytical Composition by PLSR Analysis. PLSR analysis was performed with UNSCRAMBLER 7.5 (Camo, Asa, Norway). The following strategy was used in the model building: A first initial model was built for a given descriptor using

Table 1. Concentrations[a] (μg/L), Olfaction Thresholds (μg/L), and Odor Activity Values (OAV)

	compounds	threshold	avg concd (RSD)	avg OAV	max OAV
1	**acetaldehyde**	**500**[b]	**45 900 (23.4)**	**91.8**	**153**
2	**2.3-butanodione**	**100**[b]	**1250.(63.3)**	**12.5**	**33.9**
3	**ethyl butyrate**	**20.0**[b]	**386 (36.9)**	**19.3**	**55.9**
4	**2-methylpropanol**	**40 000**[b]	**84 800 (40.4)**	**2.12**	**5.75**
5	**3-methylbutyl acetate**	**30.0**[b]	**437 (32.3)**	**14.6**	**28.3**
6	1-butanol	150 000[c]	1730 (62.9)	<1	<1
7	**3-methylbutanol**	**40 000**[b]	**248 000 (32.3)**	**6.20**	**11.8**
8	**ethyl hexanoate**	**14.0**[b]	**425 (30.2)**	**30.4**	**59.3**
9	hexyl acetate	1500[c]	97.1 (65.6)	<1	<1
10	**3-hydroxy-2-butanone**	**150 000**[b]	**54 900 (87.5)**	**<1**	**1.69**
11	1-hexanol	8000[b]	3500 (24.4)	<1	<1
12	**(Z)-3-hexenol**	**400**[b]	**347 (58.9)**	**0.87**	**2.60**
13	**ethyl octanoate**	**5.00**[b]	**196 (25.3)**	**39.2**	**69.7**
14	**propanoic acid**	**8100**[c]	**4160 (71.3)**	**<1**	**1.47**
15	**2-methylpropanoic acid**	**2300**[b]	**3510 (34.1)**	**1.53**	**3.34**
16	**butyric acid**	**173**[b]	**2020 (37.7)**	**11.7**	**25.9**
17	γ-butyrolactone	50 000[g]	36 100 (54.6)	<1	<1
18	**3-methylbutyric acid**	**33.4**[b]	**1810 (28.9)**	**54.2**	**105**
19	**3-(methylthio)-1-propanol**	**1000**[b]	**3660 (41.1)**	**3.66**	**10.8**
20	**hexanoic acid**	**420**[b]	**2820 (27.7)**	**6.71**	**13.9**
21	benzyl alcohol	200 000[f]	654 (165)	<1	<1
22	**2-phenylethanol**	**14 000**[b]	**58 800 (40.7)**	**4.20**	**11.9**
23	**octanoic acid**	**500**[b]	**2500 (26.1)**	**5.00**	**9.94**
24	**decanoic acid**	**1000**[b]	**671 (52.4)**	**<1**	**2.00**
25	ethyl decanoate	200[b]	18.7 (20.4)	<1	<1
26	phenylethyl acetate	250[d]	21.9 (51.1)	<1	<1
27	**2-methoxyphenol**	**9.50**[d]	**5.80 (37.7)**	**<1**	**1.40**
28	**(*E*)-β-methyl-γ-octalactone**	**67.0**[d]	**211 (52.9)**	**3.15**	**7.73**
29	furfuryl alcohol	1410[g]	57.3 (122)	<1	<1
30	γ-nonalactone	30.0[d]	11.4 (33.4)	<1	<1
31	2-methoxy-4-vinylphenol	1100[d]	66.7 (77.9)	<1	<1
32	furfural	14 100[b]	44.2 (56.3)	<1	<1
33	**β-damascenone**	**0.05**[d]	**1.50 (40.2)**	**29.6**	**67.7**
34	**ethyl isobutyrate**	**15.0**[b]	**211 (47.5)**	**14.1**	**41.1**
35	**ethyl 2-methylbutyrate**	**18.0**[d]	**14.9 (50.9)**	**<1**	**2.08**
36	1-(4-hydroxy-3-methoxyphenyl)ethanone	1000[d]	66.8 (44.2)	<1	<1
37	ethyl vanillate	990[d]	158 (40.9)	<1	<1
38	vanillin	200[d]	58.7 (55.3)	<1	<1
39	methyl vanillate	3000[d]	17.5 (59.5)	<1	<1
40	isobutyl acetate	1600[b]	35.7 (42.2)	<1	<1
41	**ethyl isovalerate**	**3.00**[b]	**43.1 (51.8)**	**14.4**	**45.9**
42	butyl acetate	1880[d]	3.10 (42.0)	<1	<1
43	5-methylfurfural	20 000[d]	14.8 (124)	<1	<1
44	ethyl furoate	16 000[b]	12.1 (45.9)	<1	<1
45	methyl benzoate	30.0[e]	0.40(30.1)	<1	<1
46	2-methylbutyric acid	2520[g]	204 (45.4)	<1	<1
47	α-terpineol	250[d]	11.7 (51.8)	<1	<1
48	linalool	25.0[d]	3.70 (51.6)	<1	<1
49	citronelool	100[c]	2.10 (62.3)	<1	<1
50	α-ionone	2.60[b]	0.10 (120)	<1	<1
51	**ethyl dihydrocinnamate**	**1.60**[d]	**0.80 (46.2)**	**<1**	**1.69**
52	(Z)-β-methyl-γ-octalactone	790[b]	53.7 (49.9)	<1	<1
53	**β-ionone**	**0.09**[d]	**0.40 (129)**	**4.44**	**21.7**
54	**4-ethylguaiacol**	**33.0**[d]	**76.5 (105)**	**2.30**	**12.7**
55	*m*-cresol	68.0[b]	1.90 (53.0)	<1	<1
56	2-methoxy-4-propylphenol	10.0[f]	2.60 (99.0)	<1	1.31
57	**(*E*)-ethyl cinnamate**	**1.10**[d]	**1.30 (65.6)**	**1.10**	**5.60**
58	γ-decalactone	88.0[b]	6.00 (46.4)	<1	<1
59	**4-allyl-2-methoxyphenol**	**6.00**[d]	**28.5 (54.8)**	**4.80**	**12.2**
60	**4-ethylphenol**	**440**[d]	**436 (131)**	**1.00**	**8.40**
61	δ-decalactone	386[b]	12.5 (41.5)	<1	<1
62	2,6-dimethoxyphenol	570[d]	31.0 (31.4)	<1	<1
63	2-methoxy-4-propenylphenol	6.00[f]	3.40 (60.6)	<1	2.00
64	4-vinylphenol	180[d]	35.2 (65.5)	<1	<1
65	4-allyl-2,6-dimethoxyphenol	1200[g]	12.3 (41.2)	<1	<1
66	phenylacetic acid	1000[g]	59.4 (54.6)	<1	<1
67	4-hydroxy-3,5-dimethoxybenzaldehyde	>50 000[g]	10.5 (137)	<1	<1
68	*o*-cresol	31.0[d]	2.20 (35.4)	<1	<1
69	**phenylacetaldehyde**	**5.00**[b]	**4.70 (68.9)**	**<1**	**2.60**

[a] Standard deviation below 10% for all cases with the exception of 4-vinylphenol (17%), furfuryl alcohol (30%), and vanillin (11%). [b] Thresholds from ref *2*. [c] Thresholds from ref *34*. [d] Thresholds from ref *25*. [e] Thresholds from ref *7*. [f] Calculated in the laboratory. [g] Database of L. J. van Gemert.

all *X* variables. After this, the existence of outliers was checked, and samples with a clear deviation from the model were eliminated and kept out from the calibration process (up to a maximum of 5). The model was then recalculated. An iterative process was then begun, to

Table 2. Basic Statistic Parameters of the 18 Sensory Descriptors (Data are Given as %)

	mean	max	min	SD	I_{50}[a]	asymmetry	kurtosis	no. of zeros
wood-vanillin-cinnamon (1)	46	82	15	17	25	0.3	–0.7	0
total fruits (2)	37	67	7	16	23	–0.3	–0.9	0
animal-leather-phenolic (3)	30	87	0	21	35	0.7	0.0	4
toasted-coffee (4)	24	69	0	15	15	1.0*	1.4	3
balsamic-licorice (5)	22	50	0	12	17	0.2	–0.2	3
spicy (6)	21	47	0	13	21	0.3	–1.1	3
berry-fruit (7)	21	47	0	10	14	0.2	–0.5	1
ripe-fruit-lactone-jam (8)	20	45	0	12	17	0.0	–0.8	5
old-wood-reduction (9)	19	67	0	17	20	1.3[b]	1.4	6
vegetal-pepper (10)	14	44	0	11	16	0.6	–0.3	11
raisin-flowery (11)	14	53	0	12	13	1.6[b]	2.7[b]	4
chemical (12)	13	60	0	11	9	2.1[b]	5.9[b]	6
sweet-candy-cacao (13)	13	38	0	10	13	0.9	0.4	8
lactic (14)	11	40	0	9	12	0.9	0.9	9
alcoholic (15)	10	50	0	9	7	1.9[b]	6.2[b]	13
tobacco-herbaceous (16)	7	35	0	7	12	1.4[b]	3.0[b]	17
smoky (17)	6	20	0	6	7	0.9	0.4	21
dried fruits (18)	5	30	0	7	7	1.8[b]	3.7[b]	27

[a] Range interquartiles. [b] Significantly different from the normal distribution.

Table 3. Quality Parameters of the PLSR Models Built for the 18 Wine Descriptors

descriptors	%EV[a]	RMSEP[b]	m[c]	offset[d]	CC[e]	no. X[f]	no. PC[g]
wood-vanillin-cinnamon (1)	53.2	2.28	0.55	4.04	0.72	10	2
total fruits (2)	46.4	2.40	0.45	4.41	0.67	10	1
animal-leather-phenolic (3)	62.1	2.72	0.63	2.24	0.79	10	2
toasted-coffee (4)	70.7	1.73	0.68	1.50	0.77	14	5
balsamic-licorice (5)	40.1	1.73	0.40	2.61	0.62	8	2
spicy (6)	29.2	1.83	0.36	2.95	0.54	8	3
berry-fruit (7)	48.2	1.95	0.52	2.21	0.68	7	2
ripe fruit-lactone-jam (8)	38.4	1.88	0.37	2.79	0.60	8	1
old-wood-reduction (9)	55.6	2.40	0.57	1.82	0.73	6	2
vegetal-pepper (10)	59.6	1.44	0.57	1.20	0.76	4	2
raisin-flowery (11)	66.3	1.20	0.70	0.79	0.81	8	3
chemical (12)	5.2	1.27	0.07	1.99	0.20	6	1
sweet-candy-cacao (13)	54.6	1.34	0.56	1.09	0.73	13	3
lactic (14)	16.0	1.67	0.17	2.07	0.37	6	1
alcoholic (15)	18.7	1.40	0.22	1.40	0.40	6	2
tobacco-herbaceous (16)	30.9	1.10	0.30	0.87	0.53	6	1
smoky (17)	33.5	0.97	0.30	0.77	0.52	6	1
dried fruits (18)	11.4	0.81	0.12	0.64	0.31	6	1

[a] Percentage of variance explained by the model. [b] Root-mean-square prediction error. [c] Slope of the regression curve between real and predicted Y variables. [d] Offset of the regression curve between real and predicted Y variables. [e] Correlation coefficient between real and predicted Y variables. [f] Number of X variables in the model. [g] Number of principal components in the model.

reduce the number of X variables in the model, searching for the simplest model with the best prediction ability. A *full cross validation* was carried out to estimate the prediction ability of the models for new sets of samples. The parameters studied to evaluate the prediction ability were the RMSEP (root mean square prediction error); the percentage of variance explained by the model, %EV; the correlation coefficient between real and predicted Y variables, CC; the slope of the regression curve between real and predicted Y variables, m; and the offset of the regression curve between real and predicted Y variables.

RESULTS AND DISCUSSION

Quantitative Analysis and Reduction of the X Variables. A summary of the results from the quantitative analysis of the 57 wines can be seen in **Table 1**. The table gives the mean concentration and its relative standard deviation (RSD), and the mean and maximum number of odor units reached in the set of wines for each of the 69 quantified odorants. Thirty-four compounds out of the 69 were found at concentrations above its threshold in at least one wine. Most of the other 35 compounds not reaching the threshold were not considered in further stages of the analysis. There are several exceptions to this rule, however. Vanillin and several related compounds were kept for the model-building steps, because despite not reaching the calculated threshold, there are some reasons to think that these compounds can play some role. First, there are at least four compounds with similar aromas that may have some additive and/or synergic effect (vanillin, ethyl vanillate, methyl vanillate, and 1-(4-hydroxy-3-methoxyphenyl)ethanone). Second, some odor thresholds reported in the literature are far lower than those reported in **Table 1** (*26*).

To reduce further the number of components that must take part in the models, a cluster study was performed using different correlation coefficients in order to locate those variables that maintain a high degree of correlation in the data set. In some cases, it was possible to identify the existence of groups of highly correlated components, because they share the same biochemical origin, and moreover, they show sensory properties relatively similar as a consequence of their similar chemical structures. In these cases, a new variable was built by linear combination of these correlated components. The weight of each component in the new variable was obtained from its corresponding threshold value (combined variable = compound_1/odor threshold of C_1 + compound_n/odor threshold of C_n). The new combined variables were the following:

Table 4. Regression Coefficients of the Odorants Included in the Study with the Models Explaining More than 45% of the Original Variance

compounds	D1[a]	D2	D3	D4	D7	D9	D10	D11	D13
acetaldehyde	–0.672			0.426					
2,3-butanodione					1.101				0.386
3-methylbutyl acetate	–0.481			1.016					
3-hydroxy-2-butanone		0.388	–0.601					0.660	0.427
(*Z*)-3-hexenol		–0.723	0.839	0.630	–1.10				
butyric acid				–0.668					
3-(methylthio)-1-propanol			0.798	0.851				–1.357	–0.388
2-phenylethanol									–0.246
2-methoxyphenol			0.576	2.501				–0.861	
(*Z*)-β-methyl-γ-octalactone	0.989			–0.474					0.341
β-damascenone		0.594	–0.760	1.008					0.345
vanillin	0.889	0.257	–1.032	–1.139		–1.682			
methyl vanillate		0.419		–0.300	0.827	–0.895			0.296
β-ionone	0.394				0.684		0.366	0.692	0.414
4-ethylguaiacol	–0.633	–0.457	1.229		–0.473				
2-methoxy-4-propylphenol			0.768					–0.522	
4-allyl-2-methoxyphenol	0.782								
4-ethylphenol	–0.744	–0.533	1.519		–0.536				–0.584
2-methoxy-4-propenylphenol		0.173	0.880	0.534		–1.150		–0.184	0.418
phenylacetaldehyde	–0.537	–0.544		–0.480	–0.510	3.119			–0.705
group 1									
group 2		–0.701							
group 3									
group 4				–1.231			0.755	0.918	0.331
group 5				–1.282		–0.592	0.766		
group 6							0.660		
ethyl vanillate	0.567					–1.005		0.257	

[a] The number is the code for the descriptor as shown in **Table 3**.

Group 1 (cinnamates): (*E*)-ethyl cinnamate + ethyl dihydrocinnamate

Group 2 (fatty acids): hexanoic acid + octanoic acid

Group 3 (ethyl esters of fatty acids): ethyl butyrate + ethyl hexanoate + ethyl octanoate

Group 4 (ethyl esters of isoacids): ethyl isovalerate + ethyl 2-methylbutyrate + ethyl isobutyrate

Group 5 (fusel alcohols): 3-methylbutanol + 2-methylpropanol

Group 6 (isoacids): 2-methylpropanoic acid + 3-methylbutyric acid

After these operations, the number of components or groups of components considered in the building of the models was 27, listed in **Table 4**.

Sensory Analysis and Reduction of the Wine Descriptors—*Y* Variables. The number of terms used by a judge to define a wine was between 1 and 12, with 3 being the most frequent. In 80% of the cases, between 2 and 5 terms were used. This number of descriptors can be considered to be normal in this type of study (*27*).

The total number of citations (excluding the hedonist and nondescriptive terms) compiled in the sensory experiment was of 3777, and the total number of descriptors used by the judges was above one hundred. As it was done in the case of the chemical components, this number was reduced, seeking to eliminate hardly relevant or badly defined terms and combining highly correlated terms or those with very nearby meanings. In a first stage, those terms which were obvious synonymous were grouped. After this, all the terms mentioned in less than 30 occasions, and moreover which had not been used in a particular wine more than 3 times, were eliminated.

Once the descriptors had been defined, the reproducibility of the judges was studied by means of the replicates performed in the tasting. Of the 51 judges, 5 were eliminated. After this operation, the number of terms considered in the study was of 3091 distributed among 33 descriptors. This number of descriptors was still considered excessively high for statistical purposes; therefore, we studied the existence of correlations among them and the possibility of using terms grouping two or three descriptors.

The cluster study using different correlation coefficients as measurement of similarity emphasized the existence of highly correlated terms. In most cases, this correlation was supported by an obvious aromatic proximity, as, for example, the terms animal and leather or (oak) wood, vanilla, and cinnamon. In other cases, the aromatic proximity is more questionable, as that between flowery and raisin, or that between old wood and reduction. Nevertheless, it was decided to group such terms in the statistical processing, because a clear improvement is obtained in the results when proceeding in this manner. This is due to the fact that the most general terms achieve a higher number of citations, so that the frequency distributions become more similar to the normal distribution. This is coherent with observations by other authors about the utility of using nonexcessively specific terms (*10, 23, 28*). The descriptors finally considered are the 18 shown in **Table 2**.

The terms in this table are arranged by the average frequency of citation per wine. The table can be seen to be divided in two halves. The first nine terms are those most frequently used to define these wines, with average frequencies of citation higher than 19%. Conversely, all nine terms of the lowest part of the table are much more rarely used, with average frequencies of citation below 15%. The first terms are therefore essential in the definition of such wines, whereas the last ones are secondary, even though they can define some subset of the original group. Data from this table also indicate that the scores of most of the descriptors are distributed in an approximately normal form, though in the cases 4, 9, 11, 12, 15, 16, and 18, the distribution is distinctly slanted, being rather normal-log.

Modeling Sensorial Descriptors from Analytical Composition. Models were built based on the algorithm PLS1 for all 18 descriptors indicated in **Table 2**. PLS1 algorithm was used

Table 5. Wines Analyzed in the Experiment

wine	vintage year	denomination of origin	grape variety*	alcohol %(v/v)
Federico	1992	Ribera Duero	na	13.0
Hacienda Monasterio	1992	Ribera Duero	na	12.5
Monte Ducay	1993	Cariñena	T, G, CS	12.5
Mas Comtal	1993	Penedés	Me, CF	13.0
Viña Ainzon	1994	Borja	G, T	13.0
Marques de Tosos	1994	Cariñena	T, G, Ca	12.5
Señorio del Aguila	1994	Cariñena	G, T, CS	12.5
Jean Leon	1994	Penedés	CS	13.5
Borsao	1995	Borja	G, T, CS	13.0
Coto hayas	1995	Borja	G, T, CS	13.5
Señor Atares	1995	Borja	G, T, CS	13.0
Torrelongares	1995	Cariñena	G, T	13.0
Lan	1995	Rioja	na	13.0
Borsao	1996	Borja	G, T, CS	13.0
Coto Hayas	1996	Borja	G, T	13.0
Gran Campelles	1996	Borja	G, T, CS	13.0
Monte Ducay	1996	Cariñena	G, T, Ca	12.5
Señorio del Aguila	1996	Cariñena	G, T, CS	12.5
Castillo de Monjardin	1996	Navarra	T, CS, Me	13.0
Evena	1996	Navarra	T, CS	13.0
Gran Feudo	1996	Navarra	G, T, CS	12.5
Gran Feudo	1996	Navarra	T, CS	12.5
Gran Irache	1996	Navarra	na	12.5
Montecristo	1996	Navarra	T, CS, Ma	12.5
Nekeas	1996	Navarra	T, CS, Me	13.0
Palacio de la Vega	1996	Navarra	T, CS	13.0
Palacio de Otazu	1996	Navarra	T, Me, CS	13.0
Cavas Hill	1996	Penedés	CS	12.5
Jané Ventura	1996	Penedés	CS	13.0
Mas d'Aranyó	1996	Penedés	T	12.5
Naveran	1996	Penedés	CS	12.5
Emilio Moro	1996	Ribera Duero	T	13.0
Pago Capellanes	1996	Ribera Duero	T, CS	13.5
Viña Mayor	1996	Ribera Duero	T	13.0
Alta Rio	1996	Rioja	T	12.5
Coto de Imaz	1996	Rioja	T	12.5
Longrande	1996	Rioja	na	13.0
Marqués de Vargas	1996	Rioja	na	14.0
Muga	1996	Rioja	na	13.0
Murua	1996	Rioja	na	13.5
Puerta Vieja	1996	Rioja	na	12.5
Viña Albina	1996	Rioja	na	13.0
Gran Vos	1996	Somontano	na	13.5
Can Feixes	1997	Penedés	na	13.0
Dehesa de los Canonigos	1997	Ribera Duero	na	13.0
Emilio Moro	1997	Ribera Duero	T	13.0
Pago de los Capellanes	1997	Ribera Duero	na	13.0
Viña Pedrosa	1997	Ribera Duero	na	13.0
Viñas de Gain	1997	Rioja	T	13.0
Enate	1997	Somontano	T, CS	13.0
Señorio de Lazán	1997	Somontano	T, CS, Mo	13.0
Marques de Ballestar	1998	Cariñena	G, T	13.0
Albet i Noya	1998	Penedés	S	13.0
Augustus	1998	Penedés	CF	12.5
Coronas	1998	Penedés	na	13.0
Duque de Azara	1998	Somontano	T, CS, Me	13.5
Montesierra	1998	Somontano	T, CS, Mo	13.0

[a] Ca, Carignan; CF, Cabernet Franc; CS, Cabernet Sauvignon; G, Grenache; Ma, Mazuelo; Me, Merlot; Mo, Moristel; S, Syrah; T, Tempranillo; na, not available.

because PLS2 algorithm results were not satisfactory and difficult to interpret. Different transformations of the chemical and sensory data were tested to avoid the problems derived from the lack of normality observed in some cases. Nevertheless, the best results were obtained by direct correlation between the frequency of citation of an odorant and the chemical data without any transformation further than a centering and an auto scaling. The best models obtained following the strategy delineated in the methods section are shown in **Tables 3** and **4**. **Table 3** shows the basic statistical data of the models, and **Table 4** shows the correlation coefficients of the different variables.

Data in **Table 3** show that some models are capable of explaining a high percentage of the original variance. In the cases *wood-vanillin-cinnamon*, *animal-leather-phenolic*, *toasted-coffee*, *old-wood-reduction*, *vegetal-pepper*, *raisin-flowery*, and *sweet-candy-cacao*, the percentage of initial variance explained by the validated model was higher than 50%. In the opposite case stood the models obtained for the descriptors *chemical*, *lactic*, *alcoholic*, and *dried fruits*, which did not manage to explain more than 20% of the initial variance. Fortunately, these descriptors belong to the group of secondary descriptors for this type of wines, as was stated previously. The low efficiency of these models can be attributed as much to the difficulty of modeling variables containing a high number of zeros as to the

possibility that these might be terms with a non-univocal definition. On the contrary, the relatively poor results of the models built for the descriptors 5, 6, and 8, which are well defined descriptors and with a high frequency of use, must be ascribed to the nonexistence of data on aromatic compounds with more than probable effect in these descriptors, such as 3-hydroxy-4,5-dimethyl-2(5*H*)-furanone, 4-hydroxy-2,5-dimethyl-3(2*H*)-furanone, and the methoxypyrazines. Another result of undoubted importance derived from the information on **Table 3** is that no model is univariate. The simplest is integrated by 4 variables, but most models required between 6 and 10 variables. This result is in agreement with the recognized complexity of the aroma of wine and partly explains the difficulty of its study.

Models. The data presented in **Table 4** confirm this complexity, because the 27 odorants or groups of odorants selected for the modeling intervened in some of the 9 models shown in the table. The correlations are positive as well as negative, which suggests that the perception of an aromatic note is influenced not only by the presence of a few components whose aroma form the note, but also by the presence of other odorants that affect negatively in the perception of such aromatic note. This is the case of the most important aromatic note of these wines, the descriptor *wood-vanillin-cinnamon* (1). This aromatic note is due mainly to the three most important aromas of wood, (*Z*)-β-methyl-γ-octalactone, vanillin, and 4-allyl-2-methoxyphenol, but its perception comes limited by the presence of 4-ethylphenol, 4-ethylguaiacol, acetaldehyde, and phenylacetaldehyde. It should be noted that the contribution of vanilla in this model cannot be explained attending to the aroma values calculated in **Table 1**. The following descriptor in importance is the fruity term (*2*) whose model shows some interesting relationships. The fruity tone depends primarily on the content of the wine in β-damascenone, which is coherent both with the aroma values of this component and with its aroma. Nevertheless, in the model, there is no longer any other component with fruity aroma that may positively affect the perception of this descriptor. On the other hand, it is again observed that the perception of this note comes impeded by the presence of four components. The same behavior is observed in the case of another two descriptors related to fruity characteristics, berry-fruit (7) and raisin-flowery (11). Berry-fruit is a note formed of 2,3-butanodione, methyl vanillate, and β-ionone, and its perception seems to be limited by the same components that are opposed to the fruity perception. The ethyl esters of the acids 3-methylbutyric, 2-methylbutyric, and 2-methylpropanoic (group 4) β-ionone and 3-hydroxy-2-butanone are responsible for the raisin-flowery note. In this case, the perception of this note seems to be affected by the levels of 3-(methylthio)-1-propanol and 2-methoxyphenol, which are odorants with pungent aroma.

The model built for the descriptor *toasted-coffee* (4) is the most complex, and despite being the model that manages to explain the highest percentage of variance, its interpretation is not easy. The odorant with the strongest weight in the model is 2-methoxyphenol, which is coherent with its aromatic description. Nevertheless, none of the other odorants with positive weight in the model possesses notes of this type. This complexity might be due to the absence of quantitative data on 2-furfurylthiol, an odorant whose implication in the coffee notes of some wines has already been proved (*29*). A similar situation is observed in the case of the model *sweet-candy-cacao* (13). Though the predictive capacity of the model is satisfactory, its excessive complexity might be due to the absence of data on 4-hydroxy-2,5-dimethyl-3(2*H*)-furanone. Nevertheless, the model shows that this descriptor is related to aromas of sweet and fruity tones and comes limited by the presence of 4-ethylphenol, phenylacetaldehyde, and 3-(methylthio)-1-propanol.

Conversely, the interpretation of the model built for the descriptor *animal-leather-phenolic* (3) turns out to be very easy. The two components with the strongest weight in the model are 4-ethylphenol and 4-ethylguaiacol, odorants whose contribution to these less favorable notes in the quality of a wine has already been shown (*30*−*32*). The negative contribution of odorants of agreeable aroma such as vanillin or β-damascenone is remarkable, and therefore it can be said that the descriptor *animal-leather-phenolic* has a structure opposed to that of descriptors 1 and 2. Similar behavior is observed in the case of the descriptor *old-wood-reduction* (9), which is due exclusively to the presence of phenylacetaldehyde in the wine but whose intensity comes limited by the presence of odorants related to fruity notes or to the wood term.

The model obtained for the descriptor *vegetal-pepper* (10) is very satisfactory, despite not having available data on methoxypyrazines. These components have not been considered as potentially important aromas in any of the olfactometries performed on Spanish aged wines (*7*, *8*), though they have been reported by other authors as contributors to the vegetal and pepper notes of Cabernet Sauvignon wines (*33*). The model suggests, nevertheless, that the vegetal and pepper character of the wines in this study is due to the joint action of the ethyl esters of the isoacids, to the fusel alcohols, and to the isoacids themselves.

LITERATURE CITED

(1) Guth, H. Quantitation and sensory studies of character impact odorants of different white wine varieties. *J. Agric. Food Chem.* **1997**, *45*, 3027−3032.

(2) Ferreira, V.; Lopez, R.; Cacho, J. F. Quantitative determination of the odorants of young red wines from different grape varieties. *J. Sci. Food Agric.* **2000**, *80*, 1659−1667.

(3) Ferreira, V.; López, R.; Aznar, M. Olfactometry and Aroma Extract Dilution Analysis of Wines. In *Analysis of Taste and Aroma*; J. Jackson, Ed.; Springer-Verlag: Berlin, 2001; pp 89−122.

(4) Ferreira, V.; Ortín, N.; Escudero, A.; López, R.; Cacho, J. Chemical characterization of the aroma of grenache rosé wines. Aroma Extract Dilution Analysis, quantitative determination and sensory reconstitution studies. *J. Agric. Food Chem.* **2002**, *50*, 4048−4054.

(5) Tominaga, T.; Murat, M. L.; Dubourdieu, D. Development of a method for analyzing the volatile thiols involved in the characteristic aroma of wines made from Vits vinifera L. Cv Sauvignon Blanc. *J. Agric. Food Chem.* **1998**, *46*, 1044−1048.

(6) López, R.; Ferreira, V.; Hernández, P.; Cacho, J. F. Identification of impact odorants of young red wines made with Merlot, Cabernet Sauvignon, and Grenache grape varieties: a comparative study. *J. Sci. Food Agric.* **1999**, *79*, 1461−1467.

(7) Aznar, M.; López, R.; Cacho, J. F.; 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.* **2001**, *48*, 2924−2929.

(8) Ferreira, V.; Aznar, M.; López, R.; Cacho, J. Quantitative gas chromatography−Olfactometry carried out at different dilutions of an extract. Key differences in the odor profiles of four high-quality Spanish aged red wines. *J. Agric. Food Chem.* **2001**, *49*, 4818−4824.

(9) Noble, A.; Ebeler, S. E. Use of multivariate statistics in understanding wine flavor. *Food Rev. Int.* **2002**, *18*, 1−20.

(10) Lawless, H. T. Descriptive analysis of complex odors: reality, model, or illusion? *Food Qual. Pref.* **1999**, *10*, 325−332.

(11) Etiévant, P. X.; Issanchou, S. N.; Marie, S.; Ducruet, V.; Flanzy, V. Sensory impact of volatile phenols on red wine aroma: influence of carbonic maceration and time of storage. *Sci. Aliments* **1989**, *9*, 19−33.
(12) Allen, M. S.; Lacey, M. J.; Harris, R. L. N.; Brown, W. V. Contribution of methoxypyrazines to Sauvignon blanc wine aroma. *Am. J. Enol. Vitic.* **1991**, *42*, 109−112.
(13) Martin, B.; Etievant, P. X.; Quere, J.; Schlich, P. More clues about sensory impact of sotolon in some flor sherry wines. *J. Agric. Food Chem.* **1992**, *40*, 475−478.
(14) Ferreira, V.; Peòa, C.; Escudero, A.; Fernandez, P.; Cacho, J. Investigation on the role played by fermentation esters in the aroma of young Spanish wines by multivariate analysis. *J. Sci. Food Agric.* **1995**, *67*, 381−392.
(15) Escudero, A.; HernandezOrte, P.; Cacho, J.; Ferreira, V. Clues about the role of methional as character impact odorant of some oxidized wines. *J. Agric. Food Chem.* **2000**, *48*, 4268−4272.
(16) Murat, M.; Tominaga, T.; Dubourdieu, D. Mise en évidence de composés clefs dans l′arôme des vins rosés et clairets de Bordeaux. *J. Int. Sci. Vigne Vin* **2001**, *35*, 99−105.
(17) Arrhenius, S. P.; Mccloskey, L. P.; Sylvan, M. Chemical Markers for Aroma of Vitis-Vinifera Var Chardonnay Regional Wines. *J. Agric. Food Chem.* **1996**, *44*, 1085−1090.
(18) Bertuccioli, M.; Rosi, I.; Picchi, M.; Viviani, C. Prediction of wine sensory aroma attributes by wine headspace volatile compounds. *Fourth International Symposium on Cool Climate Viticulture and Enology, Rochester, NY.* 1997, Vol II, pp 33−37.
(19) Francis, I. L.; Kassara, S.; Noble, A. C.; Williams, P. J. The contribution of glycoside precursors to cabernet sauvignon and merlot aroma. In *Chemistry of wine flavor*; A. L. Waterhouse, S. E. Ebeler, Eds.; American Chemical Society: Washington, DC, 1998; pp 13−30.
(20) Martens, H.; Martens, M. *Multivariate analysis of quality. An introduction*; Wiley: Chichester, England, 2001.
(21) Fischer, U.; Berger, R. G. The impact of dealcoholization on the flavor of wine: Correlating sensory and instrumental data by PLS analysis using different transformation techniques. *4th international symposium on cool climate viticulture and enology, Rochester, NY*; 1997; pp Vol I, pp 29−36, Vol II, pp 13−17.
(22) Francis, I.; Kassara, S.; Noble, A.; Williams, P. The contribution of glycoside precursors to Cabernet Sauvignon and Merlot Aroma. In *Chemistry of wine flavor*; A. Waterhouse, S. Ebeler, Eds.; American Chemical Society: Washington, DC, 1998, pp 13−30.
(23) McCloskey, L. P.; Sylvan, M.; Arrhenius, S. P. Descriptive analysis for wine quality experts determining appellations by Chardonnay wine aroma. *J. Sens. Stud.* **1996**, *11*, 49−67.
(24) Ortega, C.; López, R.; Cacho, J.; Ferreira, V. Fast analysis of important wine volatile compounds. Development and validation of a new method based on gas chromatographic-flame ionisation detection analysis of dichloromethane microextracts. *J. Chromatogr. A* **2001**, *923*, 205−214.
(25) López, R.; Aznar, M.; Cacho, J.; Ferreira, V. Quantitative determination of minor and trace volatile compounds in wine by Solid-phase Extraction and Gas Chromatography with Mass Spectrometric detection. *J. Chromatogr. A* **2002**, *966*, 166−177.
(26) Chatonnet, P.; Dubourdieu, D.; Boidron, J. N. Incidence des conditions de fermentation et d′élevage des vins blancs secs en barriques sur leur composition en substances cédées par le bois de chene. *Sci. Aliments* **1992**, *12*, 665−685.
(27) Brochet, F.; Dubourdieu, D. Wine descriptive language supports cognitive specificity of chemical senses. *Brain Lang.* **2001**, *77*, 187−196.
(28) Civille, G. V.; Lawless, H. T. The importance of language in describing perceptions. *J. Sens. Stud.* **1986**, *1*, 203-215.
(29) Tominaga, T.; Blanchard, L.; Darriet, P.; Dubourdieu, D. A powerful aromatic volatile thiol, 2-furanmethanethiol, exhibiting roast coffee aroma in wines made from several Vitis vinifera grape varieties. *J. Agric. Food Chem.* **2000**, *48*, 1799−1802.
(30) Chatonnet, P.; Dubourdieu, D.; Boidron, J. N.; Pons, M. The origin of ethylphenols in wines. *J. Sci. Food Agric.* **1992**, *60*, 165−178.
(31) Chatonnet, P.; Dubourdieu, D.; Boidron, J. N. The influence of *Brettanomyces dekkera* sp. yeasts and lactic acid bacteria on the ethylphenol content of red wines. *Am. J. Enol. Vitic.* **1995**, *46*, 463−468.
(32) Chatonnet, P.; Viala, C.; Dubourdieu, D. Influence of polyphenolic components of red wines on the microbial synthesis of volatile phenols. *Am. J. Enol. Vitic.* **1997**, *48*, 443−447.
(33) Noble, A. C. Wine flavour. In *Understanding natural flavors*; Piggott, J. R., Patterson, A., Eds.; Blackie Academic and Professional: London, 1994; pp 228−242.
(34) Etievant, P. X. Wine. In *Volatile Compounds of Food and Beverages*; H. Maarse, Ed.; Marcel Dekker: New York, 1991; 483−546.

Received for review November 11, 2002. Revised manuscript received February 7, 2003. Accepted February 16, 2003.

JF026115Z