Characterization of Young Red Wines by Application of HJ Biplot Analysis to Anthocyanin Profiles

Celestino Santos,† Soledad S. Muñoz,† Yolanda Gutiérrez,† Eduardo Hebrero,† José L. Vicente,‡ Purificación Galindo,‡ and Julián C. Rivas*,†

Departamento de Química Analítica, Nutrición y Bromatología, and Unidad de Bioestadística, Facultad de Farmacia, Universidad de Salamanca, 37007 Salamanca, Spain

The anthocyanin contents of 48 young red wines with Certified Brands of Origin (CBO) "Ribera de Duero" and "Toro" were studied by using an HPLC method. The anthocyanin profiles were subjected to HJ biplot analysis to determine whether it would be possible to differentiate between the wines of different CBO. Anthocyanins in the form of acetates and the nonacylated anthocyanins/acylated anthocyanins ratio were the parameters with the best relative contribution to discrimination between the two CBO. A logistic model for estimating the probability of a sample belonging to a given CBO is proposed. According to the model, for cutoff point 0.65, 94.59% of the Ribera de Duero wines were well classified as were 81.82% of the Toro wines.

INTRODUCTION

Wines elaborated in specific areas and recognized as having Certified Brands of Origin (CBO) are of significant importance in the different wine-producing regions. The quality of the wines made is recognized and guaranteed. Accordingly, a series of specific parameters that will allow analysts to classify the different wines within their corresponding CBO is necessary. Among the parameters that can be used is the composition in certain metals, organic acids, certain polyphenolic compounds, etc. The values reached in these depend on a series of factors, among which are the varieties of grapes employed and the processes of elaboration and aging.

In the case of wines, and generally also of other food products, it is becoming increasingly important to have available methods for characterization, to prevent fraudulency and gain a better knowledge of their characteristics. The anthocyanin composition due to the nature of the grapes can be used as a criterion for their characterization (Roggero et al., 1986, 1988); this is also true of the total content in anthocyanins. These features later affect the characteristics displayed by the wines themselves such that it is reasonable to assume that anthocyanins can be used to characterize the wines.

The color of young red wines is mainly due to the anthocyanins extracted from the skins of the fruit during the maceration process. With time, the anthocyanins gradually disappear and the color becomes more and more due to polymeric pigments resulting from the condensation of the anthocyanins both among each other and with other components such as flavans. The vinification technique is of particular importance in the anthocyanin composition of wines; thus, wines from similar varieties of grapes may exhibit significant differences in their anthocyanin contents according to the technique used for their elaboration.

To classify the different wines within their corresponding CBO, it is possible to use techniques of statistical analysis. The techniques for data representation that show the results in the form of spatial plots across the coordinate axes are widespread. The analysis by principal components proposed by Hotelling (1933), whose theoretical basis was proposed by Pearson (1901), enables one to plot the operational taxonomic units (OTUs) with respect to new variables, a linear combination of the observable variables but with a maximum degree of variance. This technique is probably the one most used in the recognition of patterns. In this sense, Santamaria et al. (1986) applied linear discriminant analysis to different concentrations of different types of phenolic compounds to distinguish between red, claret, and rosé wines. Another common method involves plots and classification by tree diagrams in which the abscissa axis represents the unit under study and the Y axis represents distances; these plots are known as dendograms. This method has been applied by Vasconcelos and Chaves das Neves (1989) to classify wines by pattern analysis of free amino acid profiles. Callao et al. (1987) applied a discriminant analysis to 13 volatile compounds and 11 conventional enological parameters to characterize and differentiate wines from three Catalonian (Spain) regions. Mullet et al. (1987), applying an ADE statistical treatment and a clustering method to 17 wine samples from Majorca to differentiate between the high zone and the low zone, found six functions, among them the hue, that allowed them to classify all the samples correctly.

However, plots including the simultaneous analysis of OTUs and variables are of additional interest. It is extremely important to know the configuration of the OTUs and equally so to know which variables are responsible for such a configuration. In the present-day literature, and above all regarding applied works, almost the only technique for simultaneous representation is factor analysis of correspondences (FAC) (Benzecri, 1982); this kind of analysis is mainly designed for contingency tables but can also be applied to any data matrix for which it makes sense to work on profiles. Another less well-known way, although with important advantages over the former type, of representing the rows and columns of any matrix together is the HJ biplot method proposed by Galindo-Villardón (1986); this is an extension of the GH^T and HJ^T biplot method of Gabriel (1971).

In the present work three basic objectives were pursued; the first was to discover whether the anthocyanin composition can be used to differentiate wines according to their CBO. In the event of this being possible, the second aim was to study whether all the anthocyanins present in the wines are necessary or whether only one would be

† Nutrición y Bromatología.
‡ Unidad de Bioestadística.

0021-8561/91/1439-1086$02.50/0

sufficient for such purposes. Finally, given an individual sample, the third aim was to calculate the probability of that sample belonging to particular CBO. To do so, the variables chosen were used to establish the logistic prediction model. In this work, the interest in characterization is even greater if one takes into account the fact that the CBO studied belong to geographical zones close to one another and that these are within the same river basin (the River Duero); the grapes mainly employed in both zones are of the same variety, and hence the wines should, in principle, have similar characteristics.

EXPERIMENTAL PROCEDURES

Samples. The samples used in this work correspond to young red wines from the 1986 and 1987 crops and have the CBO of Toro (11 samples) and Ribera de Duero (37 samples). All the wines were from cellars affiliated with the respective Regulating Councils of both CBO. None of the wines were bought from stores; rather they were obtained directly from the wine-producing cellars themselves, thus ensuring both suitability and representativity of the samples analyzed.

Analysis of Anthocyanins. Anthocyanins were analyzed for individual compounds by HPLC (Hebrero et al., 1988). The chromatograph employed was a Varian 5000, connected to a Hewlett-Packard diode array detector (Model HP-1040M), in turn coupled to a HP-79994A data treatment station that permitted work both with the chromatograms and with the spectra obtained at any time during chromatographic analysis. Anthocyanin detection was carried out by joint use of HPLC and diode array spectroscopy according to the order of elution and the spectral properties of the compounds. To do so, the data obtained during analysis were compared with those previously collected at our laboratory (Hebrero et al., 1988).

Quantitative determination of each of the anthocyanins present in the wines was achieved from the areas of the chromatographic peaks, using a calibration table prepared from standards of malvidin 3-monoglucoside. The anthocyanin concentration in each sample is the mean of three replicates.

RESULTS AND DISCUSSION

The relative percent concentrations of the individual anthocyanins and anthocyanin fractions of the wines studied are shown in Table I. Different ratios among the former are presented too. Samples are numbered from 1 to 48; their origin is identified by RD (Ribera de Duero) or T (Toro) and crop year by 86 (1986) or 87 (1987). In the statistical analysis as INPUT a 48 × 23 matrix was used; this contained the values of the different parameters considered for the Ribera de Duero and Toro samples.

The markers for the columns in an HJ biplot coincide with the projections of the points of the scatter diagram N_j on the principal components of the OTU spaces. The markers for the rows in the HJ biplot coincide with the coordinates of the OTUs when these are referred to new variables, which are linear combinations of the original variables; these have maximum variance and are uncorrelated, and each is less important than the preceding one from the descriptive point of view. In other words, representing the OTUs by an HJ biplot is equivalent to decomposing the overall variability into its principal components. Taking into account that the factorization chosen for the HJ biplot corresponds to introducing a metric associated with the reciprocal of the matrix of covariances of the variables into the space of the rows, the distance between the markers, J, can be said to represent the Mahalanobis distance between the OTUs for plotting on the plane.

The markers for the columns of any matrix according to an HJ biplot and the markers for the rows can be represented on the same reference system (the factorial axis system) whose origin coincides with the equilibrium point of the data clouds; this enables one to interpret the proximity between the OTUs in terms of the similarity among them, and the proximity among variables in terms of covariation. As well as being able to compare the relative positions of different pairs of points–rows with respect to the set of variables and vice versa, it is also possible to interpret the distance from one observation to a variable in the sense that a variable that is close to an observation indicates that that variable has taken a high value in that OTU. The farther apart the points representing the characters of the center of gravity, the greater the variability shown by those characters in the study; the smaller the angle formed by two vectors that join the center of gravity to the points representing the variables, the more correlated the characters. Additionally, the larger the model of that vector, the greater the contribution of the element to inertia.

The HJ biplot on the maximum plane of inertia is shown in Figure 1. The inertia in the space is above 72%. The first eigenvalue is 49.57, the second 41.81, and the third 28.13, and hence the first and second axes could be interpreted as a single factorial plane, although it was necessary to interpret the third axis separately. Figure 1 shows that the first eigenvector essentially discriminates the sampling years; however, the most important variables regarding the differences in the CBO are characteristic variables of axis 3 (see Figure 2). It is evident that axis 2 separates observation 43 from the others and, thus, this axis has no relevant information. For wine 43 the relative contribution to axis 2 is very high, probably due to values obtained for the acetates of cyanidin 3-monoglucoside and peonidin 3-monoglucoside and for delphinidin 3-monoglucoside *p*-coumarate.

The greatest contributions relative to the first factorial axis were found for the variables shown in Table II. These variables are the most important in the differentiation between the years, above all in the case of the Ribera de Duero cellars, since for Toro the sample was smaller and hence the results were less reliable.

Upon analyzing the same information for axis 3 it is possible to discover the most important variables in the differentiation between the CBO (Table III). Multivariate analysis was then repeated for the above-specified samples and only the variables involved in the discrimination between both CBO. Figure 3 shows the projection of the samples onto the subspace of maximum inertia; the principal axis now separates the CBO, and the quality of representation is in this case much better. For the analysis with all the variables it was 78.22%, and now it is 97.32%, for both the samples and the variables. Figure 4 shows the samples and the variables plotted on the same reference system; this allows one to broaden the information in the sense that the variables petunidin 3-monoglucoside acetate and the anthocyanins in the form of acetates are strongly related (the angle that they form with the origin of the coordinates is very small); the sum of nonacylated anthocyanins/sum of anthocyanin acetates and malvidin 3-monoglucoside/malvidin 3-monoglucoside acetate are also related, and both these are related to petunidin 3-monoglucoside *p*-coumarate.

According to the criterion mentioned by Roggero et al. (1986), acylated anthocyanins vary considerably with different climatic conditions, which is quite possible between the two geographical zones involved in this study. However, the peonidin monoglucoside/malvidin monoglucoside ratio proposed by the above authors as a chemicotaxonomic parameters as an indicator of specific

Table I. Anthocyanin Composition (%) of Ribera de Duero and Toro Wine Samples and Different Ratios among Anthocyanin Fractions[a]

	percentages																		ratios				
sample	Dp	Cy	Pt	Pe	Mv	Dp-Ac	Cy-Ac	Pt-Ac	Pe-Ac	Mv-Ac	Dp-Cm	Cy-Cm	Pt-Cm	Pe-Cm	Mv-Cm	non-acyl	Ac	Cm	nonacyl/Ac	nonacyl/Cm	Mv/Mv-Ac	Mv/Mv-Cm	Pe/Mv
1-RD86	19.87	1.60	15.01	4.11	46.76	1.00	0.59	0.76	0.31	1.31	2.41	0.27	1.75	0.41	3.86	87.35	3.97	8.70	22.00	10.04	35.69	12.11	0.09
2-RD86	17.41	1.72	14.72	4.17	47.99	2.76	0.78	0.73	0.23	1.37	2.90	0.18	2.15	0.24	2.65	86.01	5.87	8.12	14.65	10.59	35.03	18.11	0.09
3-RD86	17.51	1.29	14.30	3.90	48.78	2.46	0.45	0.79	0.23	1.42	2.99	0.66	1.88	0.39	2.96	85.78	5.35	8.88	16.03	9.66	34.35	16.48	0.08
4-RD86	16.15	2.16	12.05	7.66	47.31	3.00	0.03	0.74	0.42	1.98	2.23	0.11	1.34	0.73	4.09	85.33	6.17	8.50	13.83	10.04	23.89	11.57	0.16
5-RD86	18.81	3.08	14.02	4.97	47.06	1.05	0.77	0.55	0.19	1.70	2.02	0.40	2.32	0.20	2.87	87.94	4.26	7.81	20.64	11.26	27.68	16.40	0.11
6-RD86	16.60	1.84	14.87	3.37	50.72	1.54	0.82	0.56	0.21	1.68	1.95	0.43	2.12	0.15	3.16	87.40	4.81	7.81	18.17	11.19	30.19	16.05	0.07
7-RD86	19.68	3.75	13.60	5.28	42.99	2.88	1.02	1.49	0.70	1.45	2.58	0.36	1.83	0.35	2.05	85.30	7.54	7.17	11.31	11.90	29.65	20.97	0.12
8-RD86	11.90	2.21	11.27	9.21	52.43	0.89	0.10	0.64	0.63	1.49	1.45	0.71	1.50	1.54	4.04	87.02	3.75	9.24	23.21	9.42	35.19	12.98	0.18
9-RD86	17.70	2.20	14.13	4.77	47.39	1.36	0.93	1.59	0.50	1.29	2.35	0.49	1.53	0.42	3.37	86.19	5.67	8.16	15.20	10.56	36.74	14.06	0.10
10-RD86	17.61	3.22	12.92	6.00	48.20	2.31	0.93	0.98	0.70	1.76	1.64	0.59	1.24	0.15	1.75	87.95	6.68	5.37	13.17	16.38	27.39	27.54	0.12
11-RD86	16.40	2.06	14.10	5.11	48.04	2.70	0.66	1.36	0.11	1.77	1.78	0.30	1.70	0.47	3.43	85.71	6.60	7.68	12.99	11.16	27.14	14.01	0.11
12-RD86	16.86	1.56	14.07	4.27	50.75	1.47	0.78	0.70	0.35	1.66	2.38	0.14	0.65	0.37	4.00	87.51	4.96	7.54	17.64	11.61	30.57	12.69	0.08
13-RD86	16.78	2.82	13.57	4.72	48.17	1.99	0.95	1.03	0.33	0.94	2.19	0.91	1.74	0.61	3.25	86.06	5.24	8.70	16.42	9.89	51.24	14.82	0.10
14-RD86	16.83	2.65	12.94	6.66	47.31	3.95	0.38	1.04	0.44	2.28	1.23	0.79	0.57	0.75	2.19	86.39	8.09	5.53	10.68	15.62	20.75	21.60	0.14
15-RD86	13.80	1.68	12.19	7.86	51.79	1.78	0.43	0.41	0.12	1.39	1.07	0.61	2.00	1.04	3.86	87.32	4.13	8.58	21.14	10.18	37.26	13.42	0.15
16-RD86	16.86	1.90	14.51	4.04	50.91	0.70	0.58	0.77	0.38	1.46	1.47	0.99	1.05	0.85	3.53	88.22	3.89	7.89	22.68	11.18	34.87	14.42	0.08
17-RD86	17.08	1.43	14.77	3.63	50.50	1.25	0.58	1.13	0.59	1.75	2.11	0.45	0.99	0.48	3.28	87.41	5.30	7.31	16.49	11.96	28.86	15.40	0.07
17-RD86	17.08	1.43	14.77	3.63	50.50	1.25	0.58	1.13	0.59	1.75	2.11	0.45	0.99	0.48	3.28	85.62	4.39	10.02	19.50	8.54	24.20	9.74	0.10
19-RD87	9.37	0.42	11.41	2.76	58.98	0.74	0.27	1.16	0.29	2.50	2.48	0.35	1.31	0.85	7.13	82.94	4.96	12.12	16.72	6.84	23.59	8.27	0.05
20-RD87	9.93	0.95	11.39	4.89	56.39	0.88	0.18	0.99	0.51	2.91	2.28	0.45	1.16	1.08	6.01	83.55	5.47	10.98	15.27	7.61	19.38	9.38	0.09
21-RD87	9.39	0.55	11.07	3.05	57.61	0.96	0.65	1.53	0.30	2.74	2.75	0.24	1.18	0.90	7.10	81.67	6.18	12.17	13.22	6.71	21.03	8.11	0.05
22-RD87	10.32	0.79	12.16	2.85	58.17	0.88	0.45	1.11	0.22	2.31	2.56	0.14	1.05	0.82	6.19	84.29	4.97	10.76	16.96	7.83	25.18	9.40	0.05
23-RD87	10.48	0.91	11.66	3.81	57.47	1.01	0.48	0.93	0.35	2.41	2.40	0.30	1.01	0.81	5.99	84.33	5.18	10.51	16.28	8.02	23.85	9.59	0.07
24-RD87	11.56	1.02	12.12	3.14	55.65	0.84	0.55	1.17	0.21	2.20	2.75	0.24	1.08	0.87	6.62	83.49	4.97	11.56	16.80	7.22	25.30	8.41	0.06
25-RD87	10.95	0.80	11.77	4.17	56.10	1.18	0.15	1.09	0.47	2.97	2.56	0.20	0.80	0.83	5.97	83.79	5.86	10.36	14.30	8.09	18.89	9.40	0.07
26-RD87	9.10	0.82	11.07	3.64	59.68	0.99	0.16	1.13	0.39	3.01	2.27	0.25	0.73	0.88	5.87	84.31	5.68	10.00	14.84	8.43	19.83	10.17	0.06
27-RD87	12.36	1.24	12.68	3.85	52.69	1.42	0.81	1.67	0.42	1.71	3.50	0.17	0.82	0.74	5.93	82.82	6.03	11.16	13.73	7.42	30.81	8.89	0.07
28-RD87	9.32	0.87	10.90	4.54	58.60	0.88	0.10	0.91	0.57	3.28	2.11	0.31	0.87	1.00	5.74	84.23	5.74	10.03	14.67	8.40	17.87	10.21	0.08
29-RD87	10.54	0.47	11.95	3.11	58.04	1.26	0.39	1.48	0.20	2.36	2.68	0.29	0.83	0.83	5.58	84.11	5.69	10.21	14.78	8.24	24.59	10.40	0.05
30-RD87	10.21	0.69	11.84	3.89	58.01	0.82	0.13	0.91	0.44	2.63	2.61	0.28	0.97	0.89	5.70	84.64	4.93	10.45	17.17	8.10	22.06	10.18	0.07
31-RD87	11.28	0.64	12.49	2.51	56.01	0.87	0.11	1.12	0.17	2.06	2.96	0.27	1.42	0.89	7.20	82.93	4.33	12.74	19.15	6.51	27.19	7.78	0.04
32-RD87	10.08	0.77	11.59	4.26	57.39	0.87	0.11	1.08	0.55	2.90	2.51	0.23	0.78	0.97	5.92	84.09	5.51	10.41	15.26	8.08	19.79	9.69	0.07
33-RD87	10.58	0.50	12.13	2.86	56.68	1.12	0.58	1.14	0.38	2.20	3.12	0.26	0.99	0.73	6.75	82.75	5.42	11.85	15.27	6.98	25.76	8.40	0.05
34-RD87	10.36	0.53	12.14	2.63	57.88	1.29	0.73	1.39	0.28	2.02	3.10	0.29	1.01	0.78	5.58	83.54	5.71	10.76	14.63	7.76	28.65	10.37	0.05
35-RD87	11.97	0.77	11.50	4.73	53.05	1.77	0.44	1.97	0.35	1.68	3.05	0.32	0.99	0.99	6.43	82.02	6.21	11.78	13.21	6.96	31.58	8.25	0.09
36-RD87	9.65	0.59	11.68	2.79	56.61	1.83	0.71	1.59	0.37	2.22	3.72	0.36	1.05	0.64	6.19	81.32	6.72	11.96	12.10	6.80	25.50	9.15	0.05
37-RD87	12.81	0.66	12.55	3.06	53.39	1.54	0.64	1.53	0.42	1.94	3.07	0.26	1.03	0.80	6.29	82.47	6.07	11.45	13.59	7.20	27.52	8.49	0.06
38-T86	14.98	1.14	11.88	5.79	50.60	4.64	1.18	1.62	0.79	1.97	1.94	0.23	1.32	0.18	1.75	84.39	10.20	5.42	8.27	15.57	25.69	28.91	0.11
39-T86	15.31	1.59	14.16	5.64	48.57	1.50	0.26	1.05	0.17	2.01	2.92	0.31	0.56	0.79	5.19	85.27	4.99	9.77	17.09	8.73	24.16	9.36	0.12
40-T86	10.97	1.59	12.72	5.61	52.96	2.96	1.45	0.97	0.22	2.62	3.15	0.41	0.35	0.37	3.67	83.85	8.22	7.95	10.20	10.55	20.21	14.43	0.11
41-T86	15.30	2.51	12.38	8.08	42.03	4.31	1.32	1.99	0.20	2.44	4.72	0.13	0.18	0.74	3.68	80.30	10.26	9.45	7.83	8.50	17.23	11.42	0.19
42-T86	13.65	1.43	13.32	4.00	52.70	1.43	0.62	1.42	0.52	2.86	3.23	0.11	0.26	0.50	3.95	85.10	6.85	8.05	12.42	10.57	18.43	13.34	0.08
43-T86	9.49	1.01	13.20	5.67	56.16	1.42	0.02	0.89	2.65	1.45	0.57	0.79	0.90	0.58	5.19	85.53	6.43	8.03	13.30	10.65	38.73	10.82	0.10
44-T87	11.56	0.66	13.03	4.53	52.56	1.03	0.61	1.45	0.10	2.70	3.05	0.36	1.03	0.97	6.35	82.34	5.89	11.76	13.98	7.00	19.47	8.28	0.09
45-T87	9.15	0.91	11.76	4.24	58.04	1.14	0.34	1.80	0.38	2.84	2.63	0.32	0.40	0.70	5.35	84.10	6.50	9.40	12.94	8.95	20.44	10.85	0.07
46-T87	13.59	1.81	12.66	6.51	47.87	2.53	0.66	1.78	0.34	2.36	3.40	0.30	0.61	0.75	4.83	82.44	7.67	9.89	10.75	8.34	20.28	9.91	0.14
47-T87	8.64	0.96	10.63	4.45	57.86	0.71	0.49	1.47	0.68	3.32	3.60	0.12	0.58	0.66	5.84	82.54	6.67	10.80	12.37	7.64	17.43	9.91	0.08
48-T87	9.10	1.24	11.84	5.74	52.90	1.37	0.72	1.63	0.51	2.82	3.82	0.24	0.70	1.05	6.32	80.82	7.05	12.13	11.46	6.66	18.76	8.37	0.11

[a] Dp, Cy, Pt, Pe, and Mv are the 3-monoglucosides of delphinidin, cyanidin, petunidin, peonidin, and malvidin, respectively. Dp-Ac, Cy-Ac, Pt-Ac, Pe-Ac, and Mv-Ac are the acetates of anthocyanin 3-monoglucosides. Nonacyl is the sum of nonacylated anthocyanins. Dp-Cm, Cy-Cm, Pt-Cm, Pe-Cm, and Mv-Cm are the *p*-coumarates of anthocyanin 3-monoglucosides. Cm is the sum of anthocyanins in the form of *p*-coumarates. Ac is the sum of anthocyanins in the form of acetates.

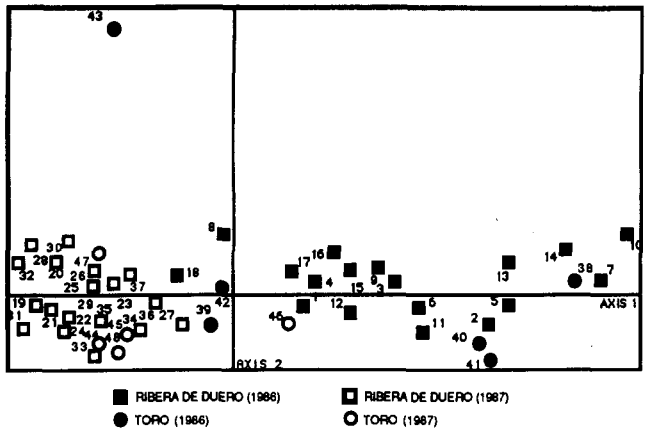

Figure 1. Projection of the samples onto the maximum plane of inertia.

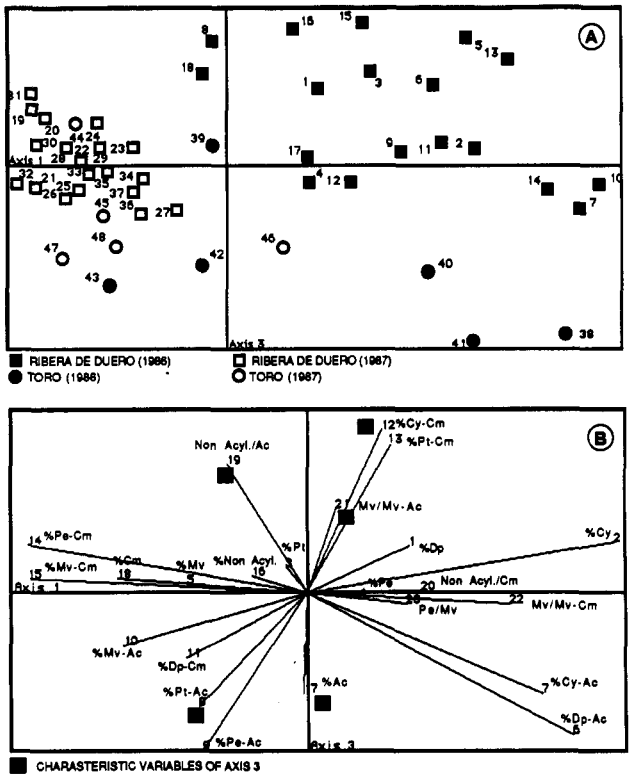

Figure 2. Projection of samples (A) and variables (B) onto planes 1 and 3.

Table II. Variables for Which the Highest Contributions Relative to Axis 1 Were Obtained

variable	rel contrib
malvidin 3-monoglucoside *p*-coumarate	910
sum % of anthocyanin *p*-coumarates	800
malvidin 3-monoglucoside	770
cyanidin 3-monoglucoside	660
peonidin 3-monoglucoside *p*-coumarate	660
malvidin 3-monoglucoside/malvidin 3-monoglucoside *p*-coumarate	660
malvidin 3-monoglucoside acetate	510
sum % of nonacylated anthocyanins	510
sum of nonacylated anthocyanins/sum of anthocyanin *p*-coumarates	460
delphinidin 3-monoglucoside acetate	450
peonidin 3-monoglucoside/malvidin 3-monoglucoside	180

enzymatic activity in grapes is not appropriate for the characterization of young red wines, in our case perhaps because we were dealing with the same variety of grapes.

In the Ribera de Duero CBO the values of petunidin

Table III. Variables for Which the Best Contributions Relative to Axis 3 Were Obtained

variable	rel contrib
anthocyanins in form of acetates	850
sum of nonacylated anthocyanins/sum of anthocyanin acetates	700
petunidin 3-monoglucoside *p*-coumarate	500
malvidin 3-monoglucoside/malvidin 3-monoglucoside acetate	370
petunidin 3-monoglucoside acetate	350
cyanidin 3-monoglucoside *p*-coumarate	320
petunidin 3-monoglucoside	150

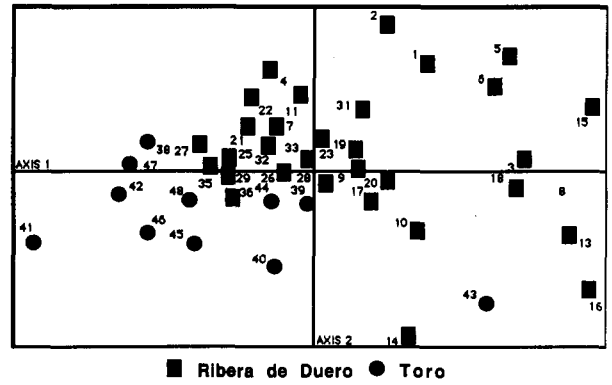

Figure 3. Projection of the samples onto the maximum plane of inertia using only the variables of maximum relative contribution shown in Table III.

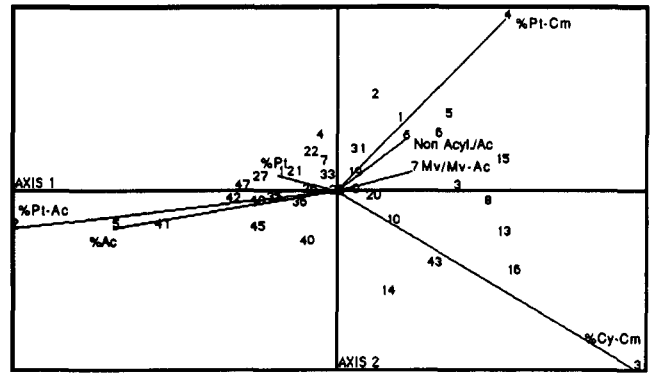

Figure 4. Samples and variables plotted on the same reference system.

3-monoglucoside *p*-coumarate and cyanidin 3-monoglucoside *p*-coumarate are higher; however, in the Toro CBO the petunidin 3-monoglucoside acetate and the anthocyanins in the form of acetate are higher.

The variables chosen in the biplot analysis were used to find the parameters of the logistic model that will serve to estimate the probability of a sample, for which the variables considered take particular values, belonging to a given CBO. The coefficients estimated and their corresponding standard errors together with the equation that serves as the prediction model are shown in Table IV. Given a wine sample, which may be from Ribera de Duero or from Toro, it is possible to establish whether it is from Ribera de Duero if the probability of its belonging to Toro is less than the probability of its belonging to Ribera de Duero. The error rate will always be higher when a prediction rule is used prospectively in a new group of wines (the test set) than when it is used in the group from which it was derived (the training set). There are several ways for estimating the misclassification when a prediction rule is applied. In our case, among the statistical cross-validation techniques, the jackknife method was selected. In the jackknife method, one wine is removed and the rule

Table IV. Parameters of the Logistic Regression Model[a]

variable	coefficient	SE
	4.153	12.789
X_1 petunidin 3-monoglucoside	−0.363	0.495
X_2 petunidin 3-monoglucoside acetate	−0.622	1.507
X_3 cyanidin 3-monoglucoside *p*-coumarate	0.426	2.740
X_4 petunidin 3-monoglucoside *p*-coumarate	2.712	1.389
X_5 anthocyanins in form of acetates	−0.384	0.973
X_6 sum of nonacylated/sum acylated anthocyanins	0.112	0.472
X_7 malvidin 3-monoglucoside/malvidin 3-monoglucoside acetate	0.002	0.094

[a] $P(Y = 1) = 1/1 + e^{-(4.153-0.363X_1-0.622X_2+0.426X_3+2.712X_4-0.384X_5+0.112X_6+0.002X_7)}$. $P(Y = 0) = 1 - P(Y = 1)$. $Y = 1$, if the observation is sampled from Ribera de Duero CBO. $Y = 0$, if the observation is sampled from Toro CBO.

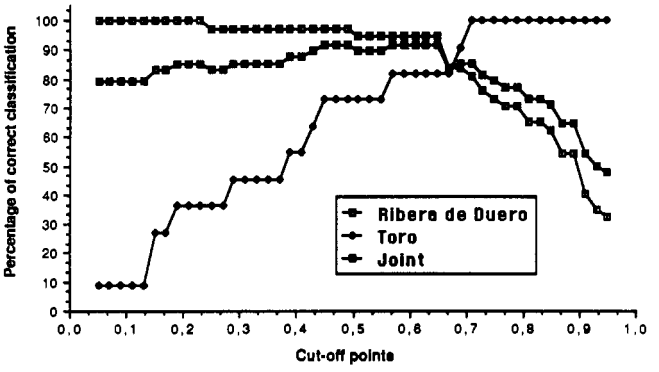

Figure 5. Plot of percentages of correct classification of samples according to the different cutoff points on applying the proposed logistic prediction model.

rederived and used to classify the excluded wine. The predictive finding is compared with the true state. Thus, the misclassification rate is determined by repeating the process for all wines. In the estimation of the model we assigned a score of 1 to the samples of Ribera de Duero and a score of 0 to those from Toro. However, when the probability with the model is estimated, such values would seldom be obtained; accordingly, one has to estimate the value, cutoff point, as from which it can be considered that the sample may belong to Ribera de Duero. Figure 5 shows different cutoff points ranging from 0.05 to 0.95 and correct success percentages ($S = 1$, Ribera de Duero) and failure percentages ($F = 0$, Toro), together with the total percentages.

The decision rule is as follows: if the probability estimated with the model surpasses the value of the cutoff point, the wine is classified as Ribera de Duero; if it is lower, the wine is classified as Toro. Taking as the cutoff point any value lower than 0.23, all the Ribera de Duero samples are classified well. Up to 0.49 only one is poorly classified; up to 0.65 two Ribera samples are poorly classified as belonging to Toro. That is, more than 94% are well classified. For a cutoff point of 0.65, 94.59% of the Ribera de Duero wines are well classified and 81.82% of the Toro wines are also well classified. In general terms and with no additional information, this would probably be a good cutoff point. Evidently, in view of the logistic plot it is possible to choose other cutoff points to be able to work with different degrees of reliability for both determinations.

ACKNOWLEDGMENT

We are grateful to enologist Dª. Lucía García de María and to cellars affiliated with the Denominaciones de Origen "Ribera de Duero" and "Toro" for providing sample facilities and to CICYT (No. PB86-201) and Junta de Castilla y León for financial support given.

LITERATURE CITED

Benzecri, J. P. *L'Analyse des Données. Tome II: L'Analyse des Correspondences*, 4th ed.; Dunod: Paris, 1982.

Callao, M. P.; Guasch, J.; Larrechi, M. S.; Rius, F. X. Aroma Components and Chemometric Analysis to Characterize Catalan Wine Producing Zones. *Proceedings*, II Congreso Internacional de Ciencia y Tecnología de Alimentos, March 3–6, 1987, Barcelona; Primo, E., Fito, P., Eds.; Univ. Politécnica: Valencia, Spain, 1987.

Gabriel, K. R. The Biplot Graphic Display of Matrices with Application to Principal Component Analysis. *Biometrika* **1971**, *58*, 453–467.

Galindo-Villardón, P. *Questio* **1986**, *10*, 13–23.

Hebrero, E.; Santos-Buelga, C.; Rivas-Gonzalo, J. C. High Performance Liquid Chromatography-Diode Array Spectroscopy Identification of Anthocyanins of *Vitis vinifera* Variety Tempranillo. *Am. J. Enol. Vitic.* **1988**, *39*, 227–233.

Hotteling, H. Analysis of a Complex of Statistical Variables into Principal Components. *J. Educ. Psychol.* **1933**, *24*, 417–441.

Mullet, A.; Berna, A.; Forcen, M.; Mora, J. Vitivinicultural Producing Zones in Majorca. *Proceedings*, II Congreso Internacional de Ciencia y Tecnología de Alimentos, March 3–6, 1987, Barcelona; Primo, E., Fito, P., Eds.; Univ. Politécnica: Valencia, Spain, 1987.

Pearson, K. On Lines and Planes of Closest Fit to Systems of Points in Space. *Philos. Mag.* **1901**, *2*, 559–572.

Roggero, J. P.; Coen, S.; Larice, J. L. *Bull. Liaison Groupe Polyphenols* **1986**, *13*, 380–388.

Roggero, J. P.; Larice, J. L.; Rocheville-Divorne, C.; Archier, P.; Coen, S. *Rev. Fr. Oenol.* **1988**, *112*, 41-48.

Santamaria, G.; Garrido, J. L.; Diez, C. The Use of Phenol Compounds as Parameters for Distinguished Red and Rose Wines from Pale Wine in Multivariate Analysis. *Z. Lebensm. Unters. Forsch.* **1986**, *182*, 112–114.

Vasconcelos, A. M. P.; Chaves das Neves, H. J. Characterization of Elementary Wines of *Vitis vinifera* Varieties by Pattern Recognition of Free Amino Acid Profiles. *J. Agric. Food Chem.* **1989**, *37*, 931–937.

Received for review April 3, 1990. Accepted December 3, 1990.

Registry No. Petunidin 3-monoglucoside-*p*-coumarate, 66594-07-8; cyanidin 3-monoglucoside-*p*-coumarate, 37306-13-1; petunidin 3-monoglucoside acetate, 66594-02-3