Acta Scientiarum
http://periodicos.uem.br/ojs/acta
ISSN on-line: 1807-8672
Doi: 10.4025/actasciagron.v41i1.42619

Canonical correlations among grapevine agronomic and processing characteristics

Adeliano Cargnin

Empresa Brasileira de Pesquisa Agropecuária , Embrapa Uva e Vinho, Rua Livramento, 515, Cx. postal 130, 95701-008, Bento Gonçalves, Rio Grande do Sul, Brazil. E-mail: adeliano.cargnin@embrapa.br

ABSTRACT. Canonical correlation analysis allows conclusions to be drawn about the occurrence and magnitude of associations between two groups of characteristics. This study estimated the magnitude of association and interdependence between two trait groups of clones of two varieties of *Vitis vinifera* grapes. The study was based on the mean data of eight characteristics from two experiments to test the performance of these clones: the first experiment provided data from seven clones of Cabernet Sauvignon, which evaluated in the northern San Joaquin Valley, California, United States, and the second data set was from six clones of the Chardonnay variety, which were evaluated in the Napa Valley, California, United States. The canonical correlations among the two varieties were significant, indicating that the two groups of characteristics were related or not independent. The quality and quantity of the raw material required for fine-wine production were inversely proportional and specific to each variety. For Cabernet Sauvignon, the high number of berries per bunch and high number of bunches per plant resulted in fruits with a high °Brix level and increased pH. For Chardonnay, the increase in the bunch weight and number of berries decreased the fruit pH.

Keywords: *Vitis vinifera*; multivariate analysis; variety; selection; clones.

Received on October 24, 2017.
Accepted on March 26, 2018.

Introduction

The grape is one of the most important plant species, and it ranks third in economic importance among fruit trees. In Brazil, the acreage of grapes was approximately 82,500 ha in 2012, and the social and economic importance of the wine industry is high in the country due to its greater impact on income and employment generation compared to other activities typical of family agriculture. The grape is a key factor in the economy of several municipalities of at least 12 Brazilian states (approximately 30,000 families of small farmers mainly depend on wine production). This crop promotes remarkable benefits in terms of indices of quality of life, employment, income and the maintenance of producers in their primary activity (Mello, 2013).

Table grapes (*Vitis vinifera*) are used around the world for both fresh consumption and processing (fine wines). In Brazil, the use of American grapes (*Vitis labrusca* and hybrids) is also common, so the Brazilian wine industry is predominantly characterized by diversity. In this sense, breeding has contributed to the progress and development of the domestic wine industry by establishing grape cultivars and clones for different purposes that are adapted to the soil and climate conditions of Brazil.

Despite the gains obtained by clonal selection, mainly in yield, there are difficulties related to selecting and identifying superior genotypes, as with any other breeding method. Therefore, according to Sudre et al. (2010), multivariate methods, by which several characteristics can be simultaneously evaluated, effectively contribute to the identification of superior genotypes in breeding programs for different crops. The study of trait correlations became fundamental because it enabled the changes that occur in a particular characteristic that lead to selection losses in another to be visualized (Cruz & Carneiro, 2004). According to Tavares, Mello, and & Scivittaro (1999), path analysis can be used to improve the understanding of genetic correlation estimates as well as the complexity of these factors, thus increasing selection efficiency. However, the relationship between two groups, especially in terms of agronomically important traits, cannot be weighted by this technique.

In this context, canonical correlation analysis is the most suitable technique for measuring relationships between two trait sets, both between groups of primary and secondary yield characteristics as well as groups of physiological and agronomic characteristics. In both cases, this analytical approach is useful for the identification and selection of the most appropriate genotype in response to current agricultural demands (Rigão, Storck, Bisognin, & Lopes, 2009). Perhaps the most relevant advantage of this technique is for breeding work involving more than one dependent characteristic because it allows effort to be focused on readily measurable traits with high heritability and minimal complexity. As stated by Trugilho, Lima, and Mori (2003) and Coimbra et al. (2004), the canonical correlation technique has been used to estimate the relationship between different groups of plant characteristics in several crops, e.g., shoot and root system characteristics, primary and secondary yield characteristics, and physiological and agronomic characteristics. According to Silva, Soares, Ferreira, Silva, and Silva (2007), another essential application of the technique is in exploratory research, in which the number of characteristics is very high, to analyse trait combinations with higher correlations.

The objective of this study was to estimate the magnitude of association and interdependence between two trait groups of clones of two varieties of *Vitis vinifera*.

Material and methods

Experimental data of the Cabernet Sauvignon and Chardonnay varieties were used that came from agronomic and processing performance evaluations of seven and six wine grape clones, respectively, published by Wolpert, Kasimatis, and Weber (1994) and Wolpert, Kasimatis, and Verdegaal (1995). The experiment was arranged in a randomized block design with three replications. In 1984, Cabernet Sauvignon clones were grafted on a Harmony rootstock and planted in the field at a density of 1397 plants ha^{-1} in an espalier system in the northern San Joaquin Valley, California, United States. In the same year, Chardonnay clones were grafted on an AXR#1 rootstock and planted in the field at a density of 1346 plants ha^{-1} and trained in an espalier system in the Napa Valley, California, United States.

The clones were evaluated over three growing seasons (years) for the following characteristics: fruit yield (kg plant^{-1}), number of bunches (bunches plant^{-1}), bunch weight (g), number of berries (berries bunch $^{-1}$), berry weight (g), °Brix, pH, and titratable acidity (g L^{-1}). For the canonical correlation analysis, the characteristics were assigned to two trait groups, of which group I consisted of agronomic traits (fruit yield, number of bunches, bunch weight, number of fruits and berry weight) and group II included industrial characteristics (°Brix, pH, and titratable acidity).

Canonical correlation analysis was performed as a measure of association for each wine grape variety according to the methodology Cruz and Carneiro (2004), described below, with two trait groups, X and Y, defined as follows.

$X' = [x_1\ x_2\ ...\ x_p]$ is the vector of observations of p characteristics, constituting group I, and

$Y' = [y_1\ y_2\ ...\ y_q]$ is the vector of observations of q characteristics, constituting group II.

The problem was to estimate the maximum correlation between linear combinations of the group I and II characteristics and to estimate the respective trait-weighting coefficients for each linear combination. If X_1 and Y_1 each represent one of the linear combinations of the characteristics of groups I and II, it follows that:

$X_1 = a_1x_1 + a_2x_2 + \cdots + a_px_p$, and

$Y_1 = b_1y_1 + b_2y_2 + \cdots + b_qy_q$, where:

$a' = [a_1\ a_2\ ...\ a_p]$ is vector 1 x p of the weights of traits of group I, and

$b' = [b_1\ b_2\ ...\ b_q]$ is vector 1 x q of the weights of traits of group II.

In this way, the first canonical correlation was defined as the one that maximizes the relationship between X_1 and Y_1. Functions X_1 and Y_1 are the first canonical pair associated with the canonical correlation:

$$r = \frac{C\hat{o}v(X_1, Y_1)}{\sqrt{\hat{V}(X_1).\hat{V}(Y_1)}}$$

Thus, it follows that:

$C\hat{o}v(X_1, Y_1) = a'S_{12}b$;

$\hat{V}(X_1) = a'S_{11}a$; and

$\hat{V}(Y_1) = b'S_{22}b$.

In this way:

S_{11} is the p x p matrix of covariances between the characteristics of group I;

S_{22} is the q x q matrix of covariances between the characteristics of group II;

S_{12} is the p x q matrix of covariances between the characteristics of groups I and II.

In cases of standardized variables, $S_{11} = R_{11}, S_{22} = R_{22}$ and $S_{12} = R_{12}$, where R represents a correlation matrix. Where R is the correlation matrix for the two groups of variables, it corresponds to:

$R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$ and $R_{21} = R'_{12}$.

Vectors a and b are estimated by maximizing function r under the restriction that $a'R_{11}a = b'R_{22}b = 1$. This restriction is necessary to provide estimators of only a and b and indicate that the variance of each linear combination is 1 (Cruz & Carneiro, 2004).

In a first step, the eigenvalues (λ) of the specific equations were determined:

$|{R_{11}}^{-1}R_{12}{R_{22}}^{-1}R_{21} - \lambda I| = 0$; and

$|{R_{22}}^{-1}R_{21}{R_{11}}^{-1}R_{12} - \lambda I| = 0$.

Then, the respective associated eigenvectors (a) were calculated. The eigenvalues can be estimated by two different specific equations from two different matrices, one of the order p and the other of order q. When $p = q$ and the variables $X_1, X_2, \dots X_P$ as well as the variables $Y_1, Y_2, \dots Y_q$ are linearly independent, there will be p = q nonzero eigenvalues and $p = q$ canonical variable pairs. However, if, for example, $p < q$, there will be q - p null eigenvalues of the matrix $R_{22}{R_{22}}^{-1}{R_{11}}^{-1}R_{12}$ and only p canonical variable pairs.

Thus, the system of linear equations is given by:

$({R_{11}}^{-1}R_{12}{R_{22}}^{-1}R_{21} - \lambda I)\ a = \theta$, and

$({R_{22}}^{-1}R_{21}{R_{11}}^{-1}R_{12} - \lambda I)\ a = \theta$.

Thus, (1) the first canonical correlation (r_1), between the linear combination of the characteristics of groups I and II is given by $r = (^1)^{0.5}$, where λ^1 is the largest eigenvalue of matrix z ${R_{11}}^{-1}R_{12}{R_{22}}^{-1}R_{21}$, which is square and generally has a non-symmetrical order p. (2) The first canonical factor is given by $X_1 = a'X$ e $Y_1 = b'Y$, where a is the eigenvector associated with the first eigenvalue ${R_{11}}^{-1}R_{12}{R_{22}}^{-1}R_{21}$, and eigenvector b is associated with the first eigenvalue ${R_{22}}^{-1}R_{21}{R_{11}}^{-1}R_{12}$. (3) The other correlations and canonical factors are estimated with the eigenvalues and eigenvectors of the above expressions, corresponding to the order of the pth or qth estimated correlation.

The significance of the null hypothesis that all possible canonical correlations are zero was assessed by the chi-square test. Canonical correlation analysis was performed to detect relationships between the groups of agronomic (group I) and processing traits (group II). For all analyses, Genes software was used (Cruz, 2013).

Results and discussion

The results of the data variance analysis indicated good experimental precision with a satisfactory range of coefficients of variation (CV) for this type of experiment, i.e., < 20%. The F test results indicated genetic variability ($p < 0.01$) for all the parameters evaluated in the clones of the two varieties, indicating the possibility of selecting and recommending promising Cabernet Sauvignon and Chardonnay clones.

The phenotypic correlations between the eight traits evaluated in the Cabernet Sauvignon and Chardonnay clones are shown in Table 1. The correlation coefficients between the characteristics of the clones of the two grape varieties varied from -0.9850 to 0.9831.

For the Cabernet Sauvignon variety in the study environment, bunch weight (0.9831; $p < 0.01$) and berry weight had positive and significant phenotypic correlations (0.9831; $p < 0.01$ and 0.8664; $p < 0.05$, respectively) with fruit yield (weight) per plant, indicating that clones with higher bunch and berry weights (characteristics correlated with each other, r = 0.8927; $p < 0.01$) are more indicated for increased fruit yield. Similarly, there was a significant positive correlation between the phenotypic characteristics of the number of bunches (0.7753; $p < 0.05$) and pH (0.8910; $p < 0.01$) with °Brix, suggesting that clones with a greater number of bunches per plant and high pH can have high °Brix values. However, the phenotypic correlations of bunch weight (0.8281; $p < 0.05$) and berry weight (-0.8214; $p < 0.05$) with the number of bunches were negative and significant, suggesting that clones with a higher number of bunches per plant may have a lower potential for fruit yield. Thus, more bunches per plant will result in lower bunch weight and lower berry weight and consequently lower fruit yield. On the other hand, a higher number of bunches may lead to an increase in °Brix and pH.

However, for the Chardonnay variety in the environment under study, the phenotypic correlation between bunch weight (0.9095; $p < 0.05$) and the number of berries per bunch (0.8791; $p < 0.05$) with fruit yield per plant was positive and significant, indicating that clones with these characteristics (correlation $r = 0.9728$; $p < 0.01$) have greater potential for increased fruit yield. On the other hand, there was a significant negative correlation between the phenotypic characteristics of fruit yield per plant (-0.9449; $p < 0.01$), bunch weight (-0.9850; $p < 0.01$) and the number of berries per bunch (-0.9841; $p < 0.01$) with pH, suggesting that the pH values may be lower for clones with higher yield per plant (higher fruit yield per plant, increased bunch weight and greater number of berries per bunch). Thus, more berries per bunch will result in higher bunch weight and consequently increased fruit yield per plant. On the other hand, increases in these traits can lead to reduced pH.

Table 1. Matrix of the correlations between eight agroindustrial traits of clones of the *Vitis vinifera* varieties Cabernet Sauvignon (above the diagonal) and Chardonnay (below the diagonal).

Characteristics	Fruit yield	Number of bunches	Bunch weight	Number of berries	Berry weight	°Brix	pH	Titratable acidity
Fruit yield	-	-0.7183	0.9831**	0.1615	0.8664*	-0.6001	-0.3616	0.1820
Number of bunches	0.5583	-	-0.8281*	0.3112	-0.8214*	0.7753*	0.4429	0.2086
Bunch weight	0.9095*	0.1668	-	0.550	0.8927**	-0.7040	-0.4389	0.0767
Number of berries	0.8791*	0.1717	0.9728**	-	-0.1862	0.2753	0.3595	0.3850
Berry weight	-0.6817	-0.4358	-0.6315	-0.7784	-	-0.6457	-0.4075	-0.0574
°Brix	-0.1053	0.5304	-0.3656	-0.2919	-0.1852	-	0.8910**	0.1359
pH	-0.9449**	-0.2939	-0.9850**	-0.9841**	0.7427	0.2339	-	0.1493
Titratable acidity	0.8132	0.2338	0.8388	0.7447	-0.3630	-0.0553	-0.8163	-

** and * indicate significance at 1 and 5% probability, respectively, according to a t test. Number of observations underlying the correlations: 18.

A negative correlation can occur between yield components mainly due to competition among them during plant development in each crop cycle (Silva et al., 2009). It is noteworthy that the phenotypic correlation between the observed means of the clones indicates little about the real correlations between means. This correlation is based on genetic and environmental factors, but only genetic correlations consider inheritable associations.

The first two canonical correlations (Table 2) for Cabernet Sauvignon clones were significant by the chi-square test, demonstrating dependent relationships between the two trait groups (Cruz & Carneiro, 2004). Consequently, the first two canonical pairs are of interest for this study.

Table 2. Estimated canonical correlations between agronomic (group I) and processing traits (group II) in clones of the *Vitis vinifera* varieties Cabernet Sauvignon and Chardonnay.

	Cabernet Sauvignon			Chardonnay		
	Canonical pairs			Canonical pairs		
Characteristics	1°	2°	3°	1°	2°	3°
Fruit yield	-0.0191	-0.6496	0.6978	0.2267	0.9238	0.0610
Number of bunches	0.1573	0.9430	-0.1193	0.6634	0.2021	0.0731
Bunch weight	-0.0952	-0.7591	0.5980	-0.0202	0.9991	0.0110
Number of berries	0.4641	0.0617	0.3014	0.1142	0.9741	-0.1436
Berry weight	-0.1438	0.7071	0.3345	-0.6330	-0.6472	0.3930
°Brix	0.6212	0.6922	-0.3671	0.8606	-0.3502	0.3697
pH	0.8651	0.2959	-0.4048	-0.1436	-0.9892	0.0297
Titratable acidity	0.5008	0.1942	0.8434	0.0407	0.8358	0.5475
R	1.00**	0.99**	0.54^{ns}	1.00**	1.00**	0.95**

^{ns} not significant and ** significant at 1% probability by the chi-square test. Number of observations underlying the correlations: 18.Thus, inter-group associations were established, especially due to the following influences.

i) The first canonical pair (correlation 1.00; $p < 0.01$) associates plants with more berries per bunch and more bunches per plant with plants with fruits with high °Brix and higher pH.

ii) The second canonical pair (correlation 0.99; $p < 0.01$) associates plants with fruits with high Brix with those with a larger number of bunches per plant and higher berry weight but lower bunch weight and fruit yield.

Therefore, to select plants with high fruit °Brix, high yield per plant is not advisable. Plants with more bunches and more berries per bunch tend to produce fruits with high fruit °Brix. In other words, the quality and quantity of the raw material used to produce fine wines are inversely proportional, but from a physiological point of view, the opposite would be expected. This is the reason that fruits are thinned to reduce their quantity and raise their sugar content.

For clones of the Chardonnay variety, the first three canonical correlations (Table 2) were significant by the chi-square test, which also shows the dependence of the two trait groups (Cruz & Carneiro, 2004). Consequently, the first three canonical pairs are of interest for this study.

Thus, it appears that intergroup associations are mainly established due to the influence of:

i) the first canonical pair (correlation 1.00; $p < 0.01$), which associates plants with more bunches and lower berry weight with plants with high fruit °Brix;

ii) the second canonical pair (correlation 1.00; $p < 0.01$), which associates plants with greater bunch weight, more bunches per plant and increased fruit yield per plant with plants with lower pH and higher titratable acidity;

iii) the third canonical pair (correlation 0.95; $p < 0.01$), which relates plants with higher berry weight and fewer berries with those with fruits with higher acidity.

Therefore, to select plants with high fruit °Brix, the target must not be higher yield per plant. Plants with a higher number of bunches per plant and lower berry weight tend to produce fruits with higher °Brix content, whereas plants with higher bunch weight and more berries per bunch may have lower pH. In this way, the quality and quantity of the raw material used to produce fine wines are inversely related.

Conclusion

The canonical correlations were significant, indicating that the analysed trait groups are not independent in either variety.

The quality and quantity of the raw material used to produce fine wines are inversely related and specific to each variety (type of target product).

A high °Brix content in fruits is associated with more bunches per plant, but it tends to decrease with increasing fruit yield.

For the Cabernet Sauvignon variety, more berries per bunch and more bunches per plant resulted in fruits with a high °Brix content and higher pH.

For the Chardonnay variety, a higher bunch weight and number of berries reduced fruit pH.

References

Coimbra, J. L. M., Guidolin, A. F., Almeida, M. L., Sangoi, L., Ender, M., & Junior, A. M. (2004). Análise de trilha dos componentes do rendimento de grãos em genótipos de canola. *Ciência Rural*, *34*(5), 1421-1428. DOI: http://dx.doi.org/10.1590/S0103-84782004000500015.

Cruz, C. D. (2013). Genes - a software package for analysis in experimental statistics and quantitative genetics. *Acta Scientiarum. Agronomy*, *35*(3), 271-276. DOI: http://dx.doi.org/10.4025/actasciagron.v35i3.21251

Cruz, C. D., & Carneiro, P. C. S. (2004). *Modelos biométricos aplicados ao melhoramento genético*. Viçosa, MG: UFV.

Mello, L. M. R. (2013) *Vitivinicultura Brasileira*: Panorama 2012. Bento Gonçalves, RS: Embrapa Uva e Vinho.

Rigão, M. H., Storck, L., Bisognin, D. A., & Lopes, S. J. (2009). Correlação canônica entre caracteres de tubérculos para seleção precoce de clones de batata. *Ciência Rural*, *39*(8), 2347-2353. DOI: http://dx.doi.org/10.1590/S0103-84782009005000190.

Silva, F. L., Pedrozo, C. A., Barbosa, M. H. P., Resende, M. D. V., Peternelli, L. A., Costa, P. M. A., & Vieira, M. S. (2009). Análise de trilha para os componentes de produção de cana-de-açúcar via blup. *Revista Ceres*, *56*(3), 308-314.

Silva, J. W., Soares, L., Ferreira, P. V., Silva, P. P., & Silva, M. J. C. (2007). Correlações canônicas de características agroindustriais em cana-de-açúcar. *Acta Scientiarum. Agronomy*, *29*(3), 345-349.

Sudre, C. P. Gonçalves, L. S. A., Rodrigues, R., Amaral Junior, A. T., Riva-Souza, E. M., & Bento, C. S. (2010). Genetic variability in domesticated *Capsicum* spp. as assessed by morphological and agronomic data in mixed statistical analysis. *Genetics and Molecular Research*, *9*(1), 283-294. DOI: 10.4238/vol9-1gmr698

Tavares, M., Mello, A. M. T., & Scivittaro, W. B. (1999). Efeitos diretos e indiretos e correlações canônicas para caracteres relacionados com a produção de pimentão. *Bragantia*, *58*(1), 41-47. DOI: http://dx.doi.org/10.1590/S0006-87051999000100006

Trugilho, P. F., Lima, J. T., & Mori, F. A. (2003). Correlação canônica das características químicas e físicas da madeira de clones de *Eucalyptus grandis* e *Eucalyptus saligna*. *Cerne*, *9*(1), 84-94.

Wolpert, J. A., Kasimatis, A. N., & Weber, E. (1994). Filed performance of six Chardonnay clones in the Napa Valley. *American Journal of Enology and Viticulture*, *45*(1), 393-400.

Wolpert, J. A., Kasimatis, A. N., & Verdegaal, P. S. (1995). Viticultural performance of seven Cabernet Sauvignon clones in the northern San Joaquin Valley, California. *American Journal of Enology and Viticulture*, *46*(1), 437-441.