Best linear unbiased prediction in combination with path analysis in processing grapes

Cinthia Souza Rodrigues[2]*[iD], Mara Fernandes Moura[1][iD], Geovani Luciano de Oliveira[3][iD], Marlon Jocimar Rodrigues da Silva[4][iD], Marco Antonio Tecchio[4][iD]

[1]Instituto Agronômico/Centro Avançado de Pesquisa e Desenvolvimento de Frutas, Av. Luiz Pereira dos Santos, 1500 – 13214-820 – Jundiaí, SP – Brasil.
[2]Instituto Agronômico/Centro de Fibras e Grãos, Av. Theodureto de Almeida Camargo, 1500 – 13075-630 – Campinas, SP – Brasil.
[3]Universidade Estadual de Campinas/Centro de Biologia Molecular e Engenharia Genética, Av. Cândido Rondom, 400 – 13418-900 – Campinas, SP – Brasil.
[4]Universidade Estadual Paulista "Júlio de Mesquita Filho"/Faculdade de Ciências Agronômicas, Av. Universitária, 3780 – 18610-034 – Botucatu, SP – Brasil.
*Corresponding author <cinthia.rodrigues@iac.sp.gov.br>

ABSTRACT: The knowledge of correlations between multiple characteristics in plant breeding leads to more effective selection strategies. The path analysis allows refining these correlations and partitioning them into direct and indirect effects on the main variable. The path analysis becomes more effective when based on predicted genotypic values rather than phenotypic values. The objective was to evaluate correlations between the main agronomic characteristics of grapevine cultivation and their direct and indirect effects on yield per plant to improve selection strategies to reach superior progenies. A randomized complete block design was installed using four cultivars and two rootstocks, five repetitions, and plots of four plants. Data from three crop seasons were analyzed from a mixed model and genetic correlations were subject to the path analysis. A high and positive significant correlation was found between average fruit production and the number of clusters per plant. On the other hand, the average production per plant showed a low correlation to cluster width and height per grapevine. Wider and higher berries tend to increase berry fresh mass and therefore increase the contents of soluble solids and reducing sugars. Among the features, the number of clusters per plant has the strongest direct effect on fruit production in grape cultivars. Berry fresh mass, berry length, and berry width were indirectly influenced by the number of clusters and showed high heritability compared to yield and number of clusters. These characteristics could be used in indirect selection.

Keywords: *Vitis* spp., correlation, genotypic values, mixed models, tropical vitiviniculture

Introduction

Plant breeding programs seek improvements in agronomic features of interest. In this regard, multiple characteristics are simultaneously considered during the evaluation. Correlation studies aim to understand the relationship between such characteristics, which may result in genetic enhancement in future generations or the selection of a genotype of interest (Cargnin, 2019; Moreira et al., 2019). Thus, the selection for a main feature with low heritability and/or measuring difficulties can be performed based on one or more characteristics of moderate to high heritability, allowing the breeder to advance the use of indirect selection, saving time and effort (Diniz and Oliveira, 2019).

The correlation estimates a connection between pair of variables, regardless of their causes. However, the correlation allows estimating only the magnitude and the sense of linear association between two features, which may not represent their actual association, since a high or low relation may result either from the biased of a third characteristic or a set of characteristics (Cruz et al., 2012). For that reason, more detailed studies about relationships between characteristics become crucial, such as the path analysis developed by Wright (1921), which allows a better analysis of the coefficients that determine the influence that one characteristic has on another. The path analysis allows to unfold biased correlations into direct and indirect effects of features on

the main variable, generating more accurate estimates of cause and effect (Lombardi et al., 2015).

The results of the path analysis become more effective when based on predicted genotypic values rather than phenotypic ones. The use of components of variance estimated by maximum restricted likelihood (REML) and by predicted genotypic values by the best linear unbiased prediction (BLUP) leads to more precise and accurate inferences, increasing accuracy of the analysis and efficiency of breeding programs (Olivoto et al., 2017), since the phenotypic values are corrected to the environmental effects and weighed by the character heritability through BLUP (Butler et al., 2017).

This study aimed to evaluate the correlations between the main agronomic characteristics of vine cultures and their direct and indirect effects on average production by the plant of promising grapes varieties for processing using the path analysis based on genotypic values via BLUP.

Materials and Methods

Study site

The experiment was conducted at the Centro de Seringueira e Sistemas Agroflorestais of the Instituto Agronômico (IAC) in Votuporanga, São Paulo State, Brazil (20°15′ S, 50°30′ W, 483 m altitude). According to the Köppen classification, the climate in the region is humid tropical (*Aw*) (Beck et al., 2018).

The vineyard was implanted in Aug 2013. The plants were spaced in 2.0 m × 1.1 m and sustained in an espalier system with a unilateral cordon. Cultural practices followed the recommendations for regional viticulture, with daily micro-sprinklers irrigation, and installation of polyethylene screens. For all seasons, vines were cane pruned to leave one to two nodes, as well as the usual cultivation practices in the region. Subsequently, 5 % hydrogen cyanamide was applied to the buds to induce and standardize the sprouting.

Treatments and experimental design

The experimental design was randomized blocks in plots with four cultivars grafted onto two rootstocks, totaling eight cultivar/rootstock combinations. Five replicates with four plants per plot were used. The treatments consisted of the cultivars Isabel Precoce (Isabella mutation), BRS Carmem (Muscat Belly A × H 65.9.14), BRS Cora (Muscat Belly A × H. 65.9.14), and IAC 138-22 Máximo (Seibel 11342 × Syrah) grafted on the most commonly used rootstocks IAC 766 'Campinas' (106-8 Mgt × *Vitis caribaea*) and IAC 572 'Jales' (*V. caribaea* × 101-14 Mgt). All the cultivars are grapes intended for juice production, with IAC 138-22 Máximo used for juice and wine productions.

Three production cycles were evaluated in the first halves of three years: cycle 1 - from Mar 2017 (pruning) to June 2017 (harvest); cycle 2 - from Mar 2018 (pruning) to June 2018 (harvest); cycle 3 - from Mar 2019 (pruning) to July 2019. In all pruning events, four to six buds per branch were kept.

Evaluated variants

The components evaluated were as follows: cluster fresh mass (CM), berry fresh mass (BM), rachis fresh mass (RM), cluster length (CL) and width (CW), berry length (BL) and width (BW), total soluble solids (SS) expressed in °Brix, titratable acidity (AT) expressed in tartaric acid percentage, pH, maturation index (MI), and reducing sugars content (RS) in glucose percentage. At harvesting, the number of clusters per vine (NCV) was measured and the production (kg) per vine was obtained (Yield) through plant mass.

For the evaluation of the physical characteristics of clusters, rachis, and berries, ten representative clusters were used per experimental plot, and in each cluster, ten berries were collected from the upper, middle, and lower parts of the clusters (3:4:3), totaling 100 berries per plots size. Following the physical analysis, the same berries from each cluster were smashed to obtain grape must determine SS, AT, pH, MI, and RS in triplicates.

An analytically accurate scale was used to measure CM, BM, and RM (in g). A graduated measuring ruler with a length of 30 cm was used to quantify CL, CW, BL, and BW (in cm). SS was determined by an Atago® digital refractometer using the juice extracted from the grape pulp. RS was determined according to the colorimetric method proposed by Somogyi-Nelson (Nelson, 1944), which is based on an analytical glucose curve at the absorbance of 510 nm. AT was obtained by titration of 0.1 N NaOH to the equivalence point of pH = 8.2 as indicated by a color change, the MI corresponds to SS/AT ratio, and the pH was determined by a potentiometer Micronal brand, model B274.

Statistical analysis

The joint analysis of crop seasons from 2017, 2018, and 2019 was analyzed according to the following mixed linear model described:

$$Y = Xr + Zg + Wi + e$$

where: Y is the data vector; r is the vector of plot effects within different seasons (fixed); g is the vector of genetic values effects (random), where σ_g^2 corresponds to genetic variance; i is the vector effects of cultivars × seasons interaction (random), where σ_i^2 is the variance of cultivars × seasons interaction; e is the vector of random errors, where σ_e^2 the variance of errors. X, Z, and W represent the incidence matrices that fit r, g, and i to the Y data vector.

The estimation of the effects of random effects (best unbiased linear prediction - BLUP) of the model was carried out by the equation system of Henderson (1975). The residual maximum likelihood (REML) method was used to estimate the variance components (σ_g^2, σ_i^2 and σ_e^2), as described by Butler et al. (2017).

The significance of variance components was verified by the likelihood ratio test (LRT), according to Resende (2007). The genetic correlations between the characteristics studied were estimated based on the predictions of genetic values, and each characteristic was used to calculate broad sense heritability values and selective accuracy. The statistical analysis was carried out applying R version 3.6.3 using lmer functions and ranef from lme4 library to estimate and predict the mixed model.

For the path analysis, the genetic correlation matrix of the explanatory variables was verified in relation to multicollinearity. The test used was the assessment of the condition number (CN) as proposed by Montgomery and Peck (1981), which examines the ratio between the highest and lowest eigenvalues of the correlation matrix. Multicollinearity is considered low when CN < 100, moderate to high if 100 < CN < 1000, and severe if CN > 100. When severe multicollinearity was identified, a constant k was added to the main diagonal of the correlation matrix, similar to the ridge regression method (Carvalho and Cruz, 1996).

The adequate value regarding constant K was determined by examining the ridge (Hoerl and Kennard, 1970), which was obtained by plotting the estimated parameters (path coefficients) as a function of K values

within the range of $0 < K < 1$. The smallest K value capable of stabilizing most estimates of path coefficient was used. Therefore, path analysis was used to analyze all variables, considering a characteristic yield per vine as a basic variable and the other variables as independent or explanatory. GENES computational program (Cruz, 2013) was applied for the path analysis.

Results and Discussion

The treatments, constituted of different grape cultivar/ rootstock combinations for processing, showed significant differences by the likelihood ratio test (LRT) for CM, CL, CW, BM, BL, BW, RM, RM/CM, SS, AT, and RS (Table 1). Essential characteristics, such as the number of clusters per vine (NCV) and average fruit production per vine, did not differ genetically among the grapevine cultivars evaluated, which proceeded in low estimates of heritability. These results point to the need to investigate and select based on the features that are less influenced by the environment and that affect fruit production, allowing indirect genetic progress.

    The interaction of the cultivars × seasons affected almost all characteristics evaluated, except for BL. In other words, the cultivars did not show similar behavior in different seasons for this feature (Table 1). The

presence of the genotype × environment interaction makes the selection process difficult and may alter the estimates of several parameters, such as heritability, genetic variance, and even correlation. Correlation is a changeable parameter; thus, it might be underestimated if one of the characteristics offers slight variation due to low environmental control or a strong interaction between genotypes × seasons (Lira et al., 2017). Therefore, evaluating the genotypes in different seasons is essential for a more liable selection. When genotypes are evaluated in a more significant number of places or seasons, the effect of environments and genotypes × environment is estimated with greater precision, and, consequently, the contribution of the genetic effect on phenotypic variation is more accurate (Zambiazzi et al., 2017).

    The average yield per vine showed a high and positive genetic correlation with the number of clusters per vine (0.80) (Figure 1). This result was expected since the number of clusters and the cluster mass are directly related to the average production of grapevines. Gupta et al. (2015) also found a similar result while evaluating 20 cultivars of table grapes. The authors also observed that the average production per vine presented a low correlation with the cluster length and width and a high correlation with the number of clusters per vine.

Table 1 – Genetic parameters for the characteristics: NCV = number of clusters per vine; Yield; CM = cluster mass; CL = cluster length; CW = cluster width; BM = berry mass; BL = berry length; BW = berry width; RM = rachis mass; RM/CM; SS = total soluble solids; pH; AT = titratable acidity; MI = maturation index; RS = reducing sugar – estimated by the analysis of cultivars assessed in the harvests of 2017, 2018, and 2019, in Votuporanga, São Paulo State, Brazil.

Parameters	NCV	YIELD	CM	CL	CW
σ_g^2	27.98	5.80E-05	570.35**	0.67**	0.52**
σ_i^2	110.81**	1.50**	374.39**	0.98**	0.49**
σ_e^2	33.08	4.26E-01	268.42	0.78	2.90E-01
h^2	31.67	0.0034	78.19	62.14	73.12
rgg	56.23	0.6	88.42	78.82	85.51
Cve	15.23	17.8	13.77	7.23	7.67E+00
Average	37.76	3.67	119.11	12.2	7.02E+00
	BM	BL	BW	RM	RM/CM
σ_g^2	5.348e-01**	0.045**	0.023**	1.36**	0.653**
σ_i^2	0.02685**	0.0061	0.0012**	0.26**	0.060 **
σ_e^2	3.21E-02	0.0316	1.69E-03	0.43	0.149
h^2	97	90.5	97.47	92	94.5
rgg	98	95	98.7	95.9	97.2
Cve	7.9	9.98	2.72	19.4	13.5
Average	2.53	1.78	1.513	3.389	2.87
	SS	pH	AT	MI	RS
σ_g^2	0.36**	2.93E-04	3.900e-02**	13.7458	3.189**
σ_i^2	0.83**	0.008**	0.03971**	20.052**	0.3672**
σ_e^2	6.49E-01	3.53E-03	5.87E-03	7.542	9.09E-01
h^2	52	9	74	65.39	92
rgg	72.2	30	86	80.8	95.9
Cve	4.72	1.87	7.08	15.6	6.98
Average	3.17	3.17	1.08	17.56	13.67

** and * Significant at 1 and 5 %, respectively, by likelihood ratio test (LRT); σ_g^2 = genotypic variance between cultivar/rootstock combinations; σ_i^2 = variance of cultivars × season interaction; σ_e^2 = residual variance; h^2 = heritability; rgg = selective accuracy; Cve = coefficient of variation.
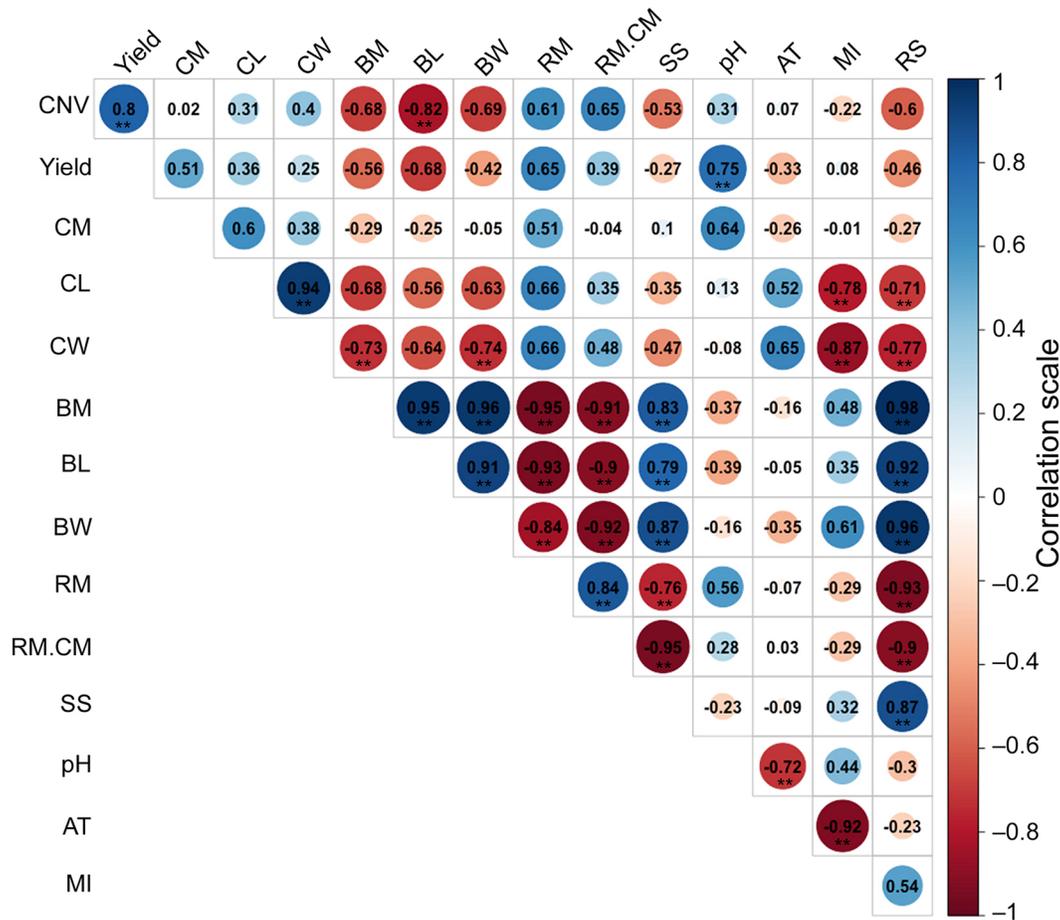
Figure 1 – Estimates of the correlations between vectors of genotypic average predicted by BLUP. NCV = n° of clusters per vine; Yield; CM = cluster mass; CL = cluster length; CW = cluster width; BM = berry mass; BL = berry length; BW = berry width; RM = rachis mass; RM/CM; SS = total soluble solids; pH; AT = Titratable acidity; MI = Maturation index; RS = reducing sugar. **Significant at 5 % of probability and other nonsignificant correlations by the t test.

The correlation between CL with CW (0.94) and BL with BW (0.91) was highly positive. This means that large berries, according to measurements such as length, tend to be bigger in other measurements, such as width; thus, longer grape clusters tend to be wider (Abiri et al., 2020).

The significant and positive correlation reported between SS and RS was 0.87 (Figure 1). The fact that these features are associated suggests that high SS may be used as an indirect measurement to select genotypes with high amounts of reducing sugars. Approximately 95 % of grapes' total soluble solids content are free sugars, which explains the highly significant correlation between SS and RS. However, the sugar content is a crucial characteristic of grapes for juice production (Borghezan, 2017).

There was also a high and positive correlation between BL and BM (0.95), BW and BM (0.96), BM and SS (0.83), BM and RS (0.98), BL and SS (0.79), BW and SS (0.87), BL and RS (0.92), and BW and RS (0.96) (Figure 1). This shows that the larger the berries, the larger their

fresh mass and the higher the content of soluble solids and reducing sugars.

Sugars in ripe berries are present at high contents in the flesh, not in the skin (Coombe et al., 1987; Possner and Kliewer, 1985). Since most berry weight is located in the flesh, the high and positive correlation between reducing sugars and berry fresh mass is easily explained. Moreover, studies have indicated that the amount of sugar per berry was not constant but increased linearly with the berry size (Ferrer et al., 2014; Matthews and Nuzzo, 2007; Roby et al., 2004). These results were similar to our findings and confirmed that the sugar content is proportional to the berry size.

Nevertheless, increases in the sugar content with fresh mass may not be enough to avoid a lower SS in larger berries. Other studies have already observed an inverse correlation between SS and berry mass (Cawthon and Morris, 1982; Roby et al., 2004). The sources of variation in berry sizes, such as viticultural practices, environmental conditions, and grape cultivars, are more

critical in determining composition than berry size *per se* (Matthews and Nuzzo, 2007; Ferrer et al., 2014).

On the other hand, there was a high and negative correlation between CL and CW with RS, providing evidence that the sugar content decreases as cluster size increases. Sugar concentration usually increases with decreasing fruit load (measured as leaf-to-fruit ratio) in many fruits, including grape (Dai et al., 2016). In grapevine, Nuzzo and Matthews (2006) showed that the rate of sugar accumulation occurred more rapidly in berries under a lower fruit load than in those under a higher fruit load.

The NCV and BL were correlated, although in opposite directions, suggesting that the more grape clusters a plant produces, the smaller the size of cluster berries. A negative correlation may be found between the production components, mainly because of competition between these components during plant development in each crop cycle. As stated by Leão et al. (2020), the negative correlation between the number of clusters and berries length arises from the competition for photoassimilates and unbalance in the source/drain.

The maturation index (MI), essential to define fruit flavor, presented a higher correlation to acidity (–0.92) than soluble solids content. These results indicate that fruits with the best flavor are readily selected based on acidity alteration. Acidity, however, is crucial for the industrial sector since it prevents deterioration by microorganisms and allows more flexibility to add sugar (Morgado et al., 2010). On the other hand, the selection to increase the SS/AT ratio is desirable to obtain fruits with good acceptance for table grapes (Ribeiro and Freitas, 2020).

There was also a negative correlation between BM with RM (–0.95), BW with RM (–0.93), BL with RM (–0.84), RM with SS (–0.76), and RM with RS (–0.93) (Figure 1), indicating that both variables move in opposite directions. A close-ratio between these characteristics could either facilitate gene introgression or make it more difficult, since a strong selection for a desirable characteristic may favor the presence of another undesirable germplasm characteristic (Dicenta and Garcia, 1992).

For the path analysis, the multicollinearity diagnosis showed NC = 4037.6, leading to severe multicollinearity (NC > 100). To minimize the adverse effects of multicollinearity, the ridge regression method was applied in which a constant (k) is added to the diagonal elements of the X′X matrix. The k value is the lowest value capable of stabilizing most estimators of the path coefficients (Viotto Del Conte et al., 2020). From there, a value k = 0.1269 is adopted, and all variables were used. In this analysis, the determination coefficient (R^2) was 0.91, and the residual variable effect was 0.298, indicating a suitable model adjustment to explain the genetic effects related to the variables under analysis (Figure 2A).

Some methodologies can be used to mitigate problems caused by multicollinearity, such as the elimination of variables with interrelationships and the ridge regression method. There is difficulty in discarding variables because a variable that is not necessarily more economically important may better explain the primary variable (Bizeti et al., 2004; Viotto Del Conte et al., 2020). Additionally, it may be required to discard many variables. Therefore, the ridge regression method becomes an efficient method to avoid the multicollinearity effect and to include and analyze all variables.

The number of clusters per vine variable (NCV) provided the most significant maximum direct effect on grape production (0.599) and low indirect effects regarding the average production of grapes (Figure 2B). In other words, NCV presented a more relevant direct effect than the residual one and may be considered the most determining variation of grapevine production. In plant breeding, finding correlated variables with a high direct and favorable effect on the primary variable is desirable. In this study, the number of clusters per vine was the feature most correlated to grape production, with the most strongly direct effect on this feature (Figure 2A), confirming the relevance of this variable in the selection process to increase grape production.

Direct positive effects of low magnitude were observed for CM, RM, CL, SS, RS, and maturation indexes. These values were below the residual (Figure 2A). In comparison, direct negative smaller effects were found in CW, BM, BW, BL, RM/CM, and AT. This low magnitude indicates a small contribution of these variables to grape production.

It is known that yield per grapevine is a polygenic characteristic greatly influenced by the environment, which makes the direct increase of grape production a real challenge (Fanizza et al., 2005; Houel et al., 2015; Muñoz-Espinoza et al., 2016). Thus, it is essential to consider the variables that indirectly influence the increase in fruit production, mainly the variables with more significant heritability and easiness of evaluation. Berry mass, berry length, berry width, rachis mass, RM/CM, SS, and reducing sugar indirectly influenced fruit production through NCV. (Figure 2B). These characteristics exhibited high heritability in comparison to yield and NCV and could be used in the indirect selection in breeding programs aiming at increasing fruit production per plant.

Conclusion

The number of clusters per vine demonstrated a higher correlation and direct influence on fruit production in the grapevine. On the other hand, berry fresh mass and length and width were indirectly influenced by the number of clusters for the production variable of grapevine. However, these variables presented higher heritability and may be an excellent selective strategy to obtain indirect gains in the production of grapes for juice production.
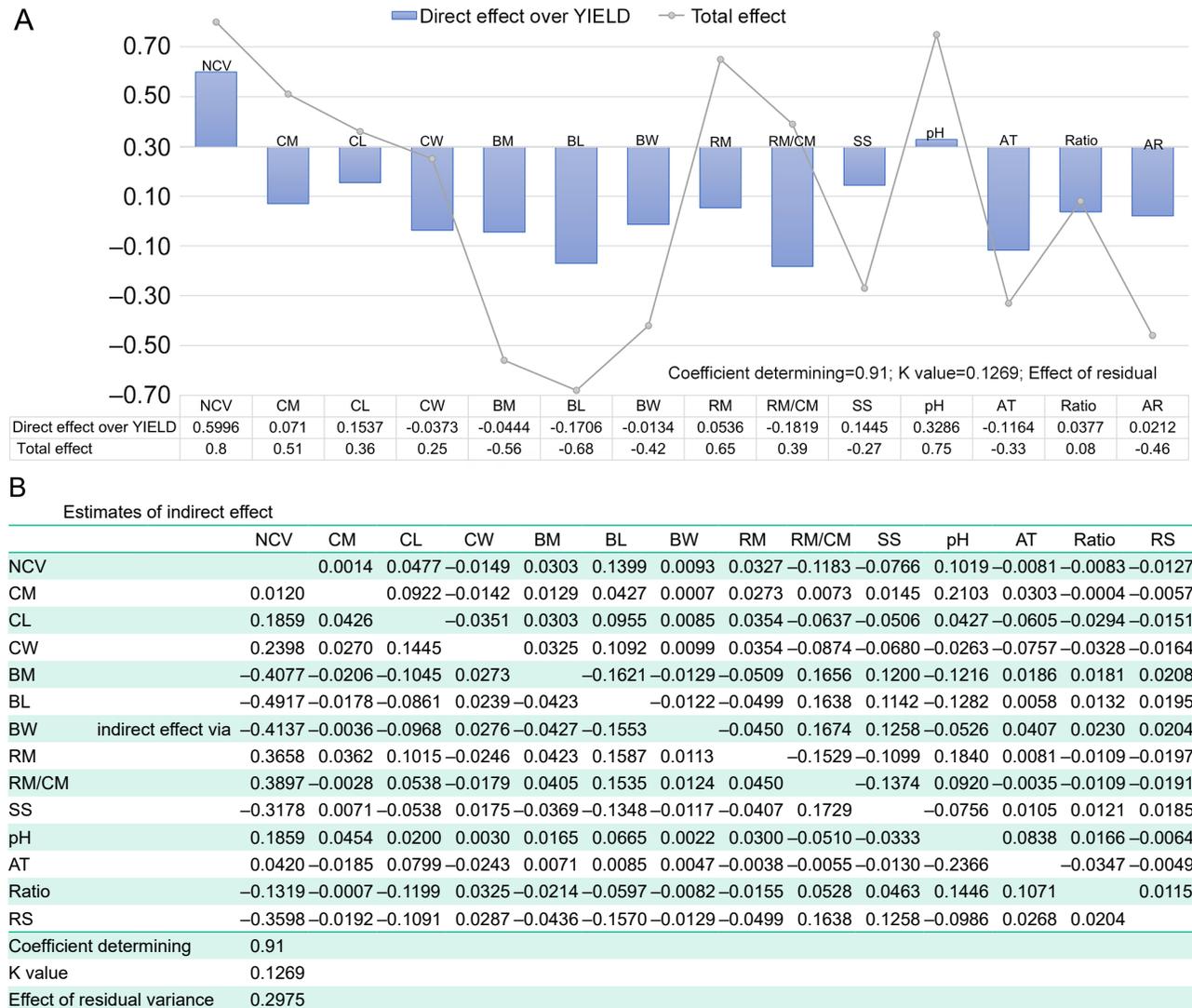
A

	NCV	CM	CL	CW	BM	BL	BW	RM	RM/CM	SS	pH	AT	Ratio	AR
Direct effect over YIELD	0.5996	0.071	0.1537	-0.0373	-0.0444	-0.1706	-0.0134	0.0536	-0.1819	0.1445	0.3286	-0.1164	0.0377	0.0212
Total effect	0.8	0.51	0.36	0.25	-0.56	-0.68	-0.42	0.65	0.39	-0.27	0.75	-0.33	0.08	-0.46

B

Estimates of indirect effect		NCV	CM	CL	CW	BM	BL	BW	RM	RM/CM	SS	pH	AT	Ratio	RS
NCV			0.0014	0.0477	–0.0149	0.0303	0.1399	0.0093	0.0327	–0.1183	–0.0766	0.1019	–0.0081	–0.0083	–0.0127
CM		0.0120		0.0922	–0.0142	0.0129	0.0427	0.0007	0.0273	0.0073	0.0145	0.2103	0.0303	–0.0004	–0.0057
CL		0.1859	0.0426		–0.0351	0.0303	0.0955	0.0085	0.0354	–0.0637	–0.0506	0.0427	–0.0605	–0.0294	–0.0151
CW		0.2398	0.0270	0.1445		0.0325	0.1092	0.0099	0.0354	–0.0874	–0.0680	–0.0263	–0.0757	–0.0328	–0.0164
BM		–0.4077	–0.0206	–0.1045	0.0273		–0.1621	–0.0129	–0.0509	0.1656	0.1200	–0.1216	0.0186	0.0181	0.0208
BL		–0.4917	–0.0178	–0.0861	0.0239	–0.0423		–0.0122	–0.0499	0.1638	0.1142	–0.1282	0.0058	0.0132	0.0195
BW	indirect effect via	–0.4137	–0.0036	–0.0968	0.0276	–0.0427	–0.1553		–0.0450	0.1674	0.1258	–0.0526	0.0407	0.0230	0.0204
RM		0.3658	0.0362	0.1015	–0.0246	0.0423	0.1587	0.0113		–0.1529	–0.1099	0.1840	0.0081	–0.0109	–0.0197
RM/CM		0.3897	–0.0028	0.0538	–0.0179	0.0405	0.1535	0.0124	0.0450		–0.1374	0.0920	–0.0035	–0.0109	–0.0191
SS		–0.3178	0.0071	–0.0538	0.0175	–0.0369	–0.1348	–0.0117	–0.0407	0.1729		–0.0756	0.0105	0.0121	0.0185
pH		0.1859	0.0454	0.0200	0.0030	0.0165	0.0665	0.0022	0.0300	–0.0510	–0.0333		0.0838	0.0166	–0.0064
AT		0.0420	–0.0185	0.0799	–0.0243	0.0071	0.0085	0.0047	–0.0038	–0.0055	–0.0130	–0.2366		–0.0347	–0.0049
Ratio		–0.1319	–0.0007	–0.1199	0.0325	–0.0214	–0.0597	–0.0082	–0.0155	0.0528	0.0463	0.1446	0.1071		0.0115
RS		–0.3598	–0.0192	–0.1091	0.0287	–0.0436	–0.1570	–0.0129	–0.0499	0.1638	0.1258	–0.0986	0.0268	0.0204	
Coefficient determining	0.91														
K value	0.1269														
Effect of residual variance	0.2975														

Figure 2 – Path analysis on the direct (A) and indirect effects (B) of NCV = n° of clusters per vine; CM = cluster mass; CL = cluster length; CW = cluster width; BM = berry mass; BL = berry length; BW = berry width; RM = rachis mass; RM/CM; SS = total soluble solids; pH; AT = Titratable acidity; MI = Maturation index; RS = reducing sugar over the yield performed with average data of vine cultivars of grapevines in Votuporanga, São Paulo State, Brazil, in the harvests of 2017, 2018, and 2019.

Acknowledgments

Authors' Contributions

Conceptualization: Rodrigues CS, Moura MF. **Data curation:** Rodrigues CS, Moura MF, Oliveira GL, Silva MJR, Tecchio MA. **Formal analysis:** Rodrigues CS. **Funding acquisition:** Silva MJR, Tecchio MA, Moura MF. **Investigation:** Rodrigues CS, Moura MF, Oliveira GL. **Methodology:** Rodrigues CS, Moura MF, Oliveira GL, **Project administration:** Rodrigues CS, Moura MF. **Resources:** Rodrigues CS, Moura MF, Oliveira GL, Silva MJR, Tecchio MA. **Supervision:** Rodrigues CS, Moura MF. **Writing-original draft:** Rodrigues CS, Moura MF, Oliveira GL. **Writing-review & editing:** Rodrigues CS, Moura MF, Oliveira GL, Silva MJR, Tecchio MA.

References

Abiri K, Rezaei M, Tahanian H, Heidari P, Khadivi A. 2020. Morphological and pomological variability of a grape (*Vitis vinifera* L.) germplasm collection. Scientia Horticulturae 266: 109285. https://doi.org/10.1016/j.scienta.2020.109285

Beck HE, Zimmermann NE, McVicar TR, Vergopolan N, Berg A, Wood EF, 2018. Present and future Köppen-Geiger climate classification maps at 1-km resolution. Scientific Data 5: 180214. https://doi.org/10.1038/sdata.2018.214

Bizeti HS, Carvalho CGP, Souza JRP, Destro D. 2004. Path analysis under multicollinearity in soybean. Brazilian Archives of Biology and Technology 47: 669-676. https://doi.org/10.1590/S1516-89132004000500001

Borghezan M. 2017. Formation and ripening of grape and effects on the wines: Review. Ciência e Técnica Vitivinícola 2: 126-141. (in Portuguese, with abstract in English) https://doi.org/10.1051/ctv/20173202126.

Butler DG, Cullis BR, Gilmour AR, Gogel BG, Thompson R. 2017. ASReml-R Reference Manual Version 4. VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.

Cargnin A. 2019. Canonical correlations among grapevine agronomic and processing characteristics. Acta Scientiarum. Agronomy 41: 42619. https://doi.org/10.4025/actasciagron.v41i1.42

Carvalho SP, Cruz CD. 1996. Diagnosis of multicollinearity: assessment of the condition of correlation matrices used in genetic studies. Brazilian Journal of Genetics 19: 479-484.

Cawthon DL, Morris JR. 1982. Relationship of seed number and maturity to berry development, fruit maturation, hormonal changes, and uneven ripening of Concord (*Vitis labrusca* L.) grapes. Journal of the American Society for Horticultural Science 107: 1097-1104. https://doi.org/10.21273/JASHS.107.6.1097

Coombe BG, Bovio M, Schneider A. 1987. Solute accumulation by grape pericarp cells V. relationship to berry size and the effects of defoliation. Journal of Experimental Botany 38: 1789-1798. https://doi.org/10.1093/jxb/38.11.1789

Cruz CD, Regazzi AJ, Carneiro PCS. 2012. Biometric Models Applied to Genetic Breeding = Modelos Biométricos Aplicados ao Melhoramento Genético. Editora UFV, Viçosa, MG, Brazil (in Portuguese).

Cruz CD. 2013. GENES: a software package for analysis in experimental statistics and quantitative genetics. Acta Scientiarum. Agronomy 3: 271-276. https://doi.org/10.4025/actasciagron.v35i3.21251

Dai Z, Wu H, Baldazzi V, Van Leeuwen C, Bertin N, Gautier H, et al. 2016. Inter-species comparative analysis of components of soluble sugar concentration in fleshy fruits. Frontiers in Plant Science 7: 649. https://doi.org/10.3389/fpls.2016.00649

Dicenta F, Garcia JE. 1992. Phenotypical correlations among some traits in almond. Journal of Genetics and Breeding 46: 241-246.

Diniz RP, Oliveira EJ. 2019. Genetic parameters, path analysis and indirect selection of agronomic traits of cassava germplasm. Anais da Academia Brasileira de Ciências 3: 1-11. https://doi.org/10.1590/0001-3765201920180387

Fanizza G, Lamaj F, Costantini L, Chaabane R, Grando MS. 2005. QTL analysis for fruit yield components in table grapes (*Vitis vinifera*). Theoretical and Applied Genetics 4: 658-664. https://doi.org/10.1007/s00122-005-2016-6

Ferrer M, Echeverría G, Carbonneau A. 2014. Effect of berry weight and its components on the contents of sugars and anthocyanins of three varieties of *Vitis vinifera* L. under different water supply conditions. South African Journal of Enology and Viticulture 35: 103-113.

Gupta N, Brar K, Gill M, Arora N. 2015. Studies on variability, correlation and path analysis of traits contributing to fruit yield in grapes. Indian Journal of Plant Genetic Resources 28: 317-320. https://doi.org/10.5958/0976-1926.2015.00042.X

Henderson CR. 1975. Best linear unbiased estimation and prediction under a selection model. Biometrics 31: 423-447. https://doi.org/10.2307/2529430

Houel C, Chatbanyong R, Doligez A, Rienth M, Foria S, Luchaire N, et al. 2015. Identification of stable QTLs for vegetative and reproductive traits in the microvine (*Vitis vinifera* L.) using the 18K Infinium chip. BMC Plant Biology 15: 205. https://doi.org/10.1186/s12870-015-0588-0

Hoerl AE, Kennard RW. 1970. Ridge regression: biased estimation for nonorthogonal problems. Technometrics 12: 55-67. https://doi.org/10.2307/1267351

Leão PCS, Nascimento JHB, Moraes DS, Souza ER. 2020. Yield components of the new seedless table grape 'BRS Ísis' as affected by the rootstock under semi-arid tropical conditions. Scientia Horticulturae 263: 109114. https://doi.org/10.1016/j.scienta.2019.109114

Lira EG, Amabile RF, Fagioli M, Montalvão APL. 2017. Genetic parameters, phenotypic, genotypic and environmental correlations and genetic variability on sunflower in the Brazilian Savannah. Ciência Rural 47: e20160719. https://doi.org/10.1590/0103-8478cr20160719

Lombardi GMR, Nunes JAR, Parrella RAC, Teixeira DHL, Bruzi AT, Durães NNL, et al. 2015. Path analysis of agro-industrial traits in sweet sorghum. Genetics and Molecular Research 14: 16392-16402. https://doi.org/10.4238/2015.December.9.8

Matthews M, Nuzzo V. 2007. Berry size and yield paradigms on grapes and wines quality. Acta Horticulturae 754: 423. https://doi.org/10.17660/ActaHortic.2007.754.56

Montgomery DC, Peck EA. 1981. Introduction to Linear Regression Analysis. John Wiley, New York, NY, USA.

Moreira FF, Hearst AA, Cherkauer KA, Rainey KM. 2019. Improving the efficiency of soybean breeding with high-throughput canopy phenotyping. Plant Methods 15: 139. https://doi.org/10.1186/s13007-019-0519-4

Morgado MAD, Santos CEM, Linhales H, Bruckner CH. 2010. Phenotypic correlations in physicochemical characteristics of Passion fruits. Acta Agronómica 59: 457-461 (in Portuguese, with abstract in English).

Muñoz-Espinoza C, Di Genova A, Correa J, Silva R, Maass A, González-Agüero M, et al. 2016. Transcriptome profiling of grapevine seedless segregants during berry development reveals candidate genes associated with berry weight. BMC Plant Biology 16: 104. https://doi.org/10.1186/s12870-016-0789-1

Nelson N. 1944. A photometric adaptation of the somogyi method for the determination of glucose. Journal of Biological Chemistry 153: 375-380. https://doi.org/10.1016/S0021-9258(18)71980-7

Nuzzo V, Matthews MA. 2006. Response of fruit growth and ripening to crop level in dry-farmed Cabernet Sauvignon on four rootstocks. American Journal of Enology and Viticulture 57: 314-324. https://doi.org/10.5344/ajev.2006.57.3.314

Olivoto T, Nardino M, Carvalho IR, Follmann DN, Ferrari M, Szareski VJ, et al. 2017. REML/BLUP and sequential path analysis in estimating genotypic values and interrelationships among simple maize grain yield-related traits. Genetics and Molecular Research 16: gmr16019525. https://doi.org/10.4238/gmr16019525

Possner DRE, Kliewer WM. 1985. The localization of acids, sugars, potassium, and calcium into developing grape berries. Vitis 24: 229-240. https://doi.org/10.5073/vitis.1985.24.229-240

Resende MDV. 2007. Software Selegen-REML/BLUP: Statistical System and Computerized Genetic Selection via Mixed Linear Models. Embrapa Forests = Sistema Estatístico e Seleção Genética Computadorizada via Modelos Lineares Mistos. Embrapa Florestas, Colombo, PR, Brazil (in Portuguese).

Ribeiro BS, Freitas ST. 2020. Maturity stage at harvest and storage temperature to maintain postharvest quality of acerola fruit. Scientia Horticulturae 260: 108901. https://doi.org/10.1016/j.scienta.2019.108901

Roby G, Harbertson JF, Adams DA, Matthews MA. 2004. Berry size and vine water deficits as factors in winegrape composition: anthocyanins and tannins. Australian Journal of Grape and Wine Research 10: 100-107. https://doi.org/10.1111/j.1755-0238.2004.tb00012.x

Viotto Del Conte M, Carneiro PCS, Resende MDV, Silva FL, Peternelli LA. 2020. Overcoming collinearity in path analysis of soybean [*Glycine max* (L.) Merr.] grain oil content. Plos One 15: e0233290. https://doi.org/ 10.1371/journal.pone.0233290

Wright S.1921. Correlation and causation. Journal of Agricultural Research 20: 557-585.

Zambiazzi EV, Bruzi AT, Guilherme SR, Pereira DR, Lima JG, Zuffo AM, et al. 2017. Estimates of genetics and phenotypics parameters for the yield and quality of soybean seeds. Genetics and Molecular Research 16: 1-12. https://doi.org/10.4238/gmr16039801