Linear relationships between biometric variables in lettuce seedlings[1]

Gabriella Rodrigues Gonçalves[2], *Cláudia Lopes Prins*[2], *Maria Inês Diel*[3], *Alessandro Dal'Col Lúcio*[3]

10.1590/0034-737X202269060004

ABSTRACT

Seedling production is an important step in growing vegetables. The study of relationships between variables has applications in practically all areas of study and is useful in identifying relationships among quality variables in vegetable seedlings. Thus, this work aims to estimate linear relationships between biometric variables in lettuce seedlings. The seedlings were grown in expanded polystyrene trays (n = 4) with 128 and 200 cells each. Leaf area, stem diameter, shoot height, length, area, volume of the root, dry and fresh shoot mass variables were measured and evaluated with Pearson's correlation analysis, multicollinearity diagnosis, condition number and the variance inflation factor, path analysis and canonical correlations. The results show linear relationships between the variables. For the 128 cell trays, there was a relationship between the total fresh mass and fresh shoot mass variables and between the total fresh mass and dry shoot mass variables. For the 200-cell trays, significant correlation was observed between the characters total fresh mass and fresh shoot mass variables and between root area and root volume variables, indicating that the greater the total fresh mass, the greater the fresh shoot mass and the greater the root area, the greater the root volume.

Keywords: *Lactuca sativa* L.; correlation; path analysis.

INTRODUCTION

Lettuce (*Lactuca sativa L.*) is the most cultivated leafy vegetable in Brazil, with a growing area of approximately 30,000 ha per year (Sala & Costa, 2012) and national production of approximately 671,509 metric tons (IBGE, 2017). Besides being one of the most consumed vegetables in the Brazilian diet, lettuce is of great socio-economic importance for the country due to its role in job creation, income, and tax revenues, in addition to helping maintain large sections of the population in rural areas (Silva *et al.*, 2017a).

The production of seedlings is one of the most important steps in the cultivation of vegetables that strongly influences the productive performance of plants and quality of the produce sent to consumer markets (Souza *et al.*, 2008). The production of healthy lettuce requires high quality seedlings. Quality seedlings are those that are well formed, healthy, and free of weeds, as there is a direct relation between healthy seedlings and productive plants in the field (Campanharo *et al.*, 2006). In contrast, producers avoid seedlings possessing compromised development, believing they likely will result in an underdeveloped crop, a downward spiral of production losses, and reduced profits (Guimarães *et al.*, 2002).

Seedling quality could be altered in the nursery period

Submitted on April 19th, 2021 and accepted on February 01rst, 2022.

[1] This work was extracted from a master's dissertation by Gabriella Rodrigues Gonçalves, titled "Plot and sample sizes and linear relationships between biometric variables in lettuce transplants" and received financial support from the Brazilian federal agency *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* (CAPES), Brazil (Financial Code 001).

[2] Universidade Estadual do Norte Fluminense Darcy Ribeiro, Laboratório de Fitotecnia, Campos dos Goytacazes, Rio de Janeiro, Brazil. rdgabriella@gmail.com, prins@uenf.br
[3] Universidade Federal de Santa Maria, Departamento de Fitotecnia, Santa Maria, Rio Grande do Sul, Brazil. mariaines.diel@hotmail.com, adlucio@ufsm.br
*Corresponding author: rdgabriella@gmail.com

through management, however it is necessary knowledge about the above and belowground growth relationship over time for effective results (Johnson & Cline, 1991). Morphological traits are the most used to evaluate seedling quality and to predict its survival in the field (Tsakaldimi *et al.*, 2013). There is a range of studies concerning root and shoot growth relationship in seedlings of forest (Johnson & Cline, 1991), maize (Ali *et al.*, 2013), Arabidopsis (Bouteillé *et al.*, 2012), rice (Zhang *et al.*, 2009) and cotton (Ayubov *et al.*, 2021), but little is known about vegetable crops. According to Ayubov *et al.*, (2021) identification of easier to screen traits is important to improve research efficiency.

Two or more associated variables may show a correlation such that modifications in one variable result in alterations to the other. Such correlations can help determine quality or reveal other variables whose measurement is complex or destructive (Volpato & Barreto, 2016). In addition, the improved understanding of correlations can help shape specific management techniques to promote processes that enhance quality or allow the selection of a variable that adequately indicates a general assessment of quality in a seedling.

The study of correlations between variables has applications in practically all fields of research. Simple linear correlations allow researchers to assess the magnitude and direction of association between two variables without providing the information necessary related to the direct and indirect effects of a variable group with respect to a dependent variable of greater importance (Cruz *et al*., 2012). Linear analyses, however, are unable to explain such responses because they may present problems of multicollinearity, which obfuscate real results. In order to better understand the causes involved in such associations, Wright (1921), developed the method of path analysis, which reveals the direct and indirect effects of correlations between independent variables and main dependent variables.

Besides path analysis, canonical correlation analysis can verify the correlations between two groups of variables (Cruz *et al.*, 2012). Biometric analyses via canonical correlations make it possible to select groups of variables more effectively than with only two variables (Bezerra Neto *et al.*, 2006).

There are few studies in the literature of the linear relations between variables with seedlings and, even among those that do exist, none dealing with lettuce cultivation were encountered. Thus, the present work sought to estimate the linear relations that exist among the biometric variables in lettuce seedlings.

MATERIAL AND METHODS

Initial procedures

To carry out the experiment, four 128-cell expanded polystyrene trays and four 200-cell expanded polystyrene trays were filled with commercial substrate mixed with NPK 4-14-8 (80g/25kg), each tray was considered as repetition. Each cell received three lettuce seeds of the 'Cinderella' cultivar. The trays were placed on 1-meter high countertops, in a greenhouse, where they remained until assessment. Automatic irrigation was applied 3-4 times per day with a duration of 2.5 minutes for each watering. Thinning was performed 9 and 10 days after the emergence of seedlings in order to leave only one seedling in each tray cell.

The experiment was carried out in the autumn of 2019. Maximum and minimum temperatures were 28 °C and 18.1 °C. Average relative humidity, evaporation and insolation were, respectively, 78.3%, 69.4 mm and 174.2 h.

Harvest and variables analyzed

The harvest and assessment were performed 21 days after emergence of the seedlings (DAE). All variables were analyzed in all the seedlings (total = 1,312 seedlings). After harvesting, the seedling roots were washed in running water with the aid of a sieve to prevent loss of root material. Next, the seedlings were placed on absorbent paper to remove excess water. Then the following measurements were recorded (in centimeters): total length (TL) (from the tip of the root to the tip of the highest leaf), shoot length (SL), root length (RL) (the difference between the total length and the shoot length), and the stem diameter (SD) (measured close to the root). After that, the following masses were recorded (in grams): the total fresh mass (TFM), the fresh shoot mass (FSM), and the fresh root mass (FRM). The leaves and roots were then placed on transparent acetate sheets and scanned to obtain digital images. The leaf area (LA) (cm^2) was measured with the Digimizer® version 5.3.5 application (Medical Software Broekstraat), while the root volume (RV) (mm^2) and root area (RA) (mm^2) were measured with the Safira® application (Embrapa, 2010). Next the material was dried in a circulating air oven at 65°C until a constant mass was obtained to determine the dry root mass (DRM) (g) and the dry shoot mass (DSM) (g).

Statistical analyses

Relationships between the variables were estimated by Pearson linear correlation coefficients. Direct and indirect effects were estimated using path analysis performed in two stages. The first stage defined the leaf area as the main dependent variable and the second defined the fresh shoot mass as the main dependent variable. The other variables were submitted to stepwise analysis in order to select the explanatory variables that did not cause multicollinearity in the correlation matrix.

After selecting the variables through stepwise analysis, multicollinearity assessment between the explanatory variables was performed with condition number analysis using Expression 1 as follows:

$$NC = \frac{\lambda\, max}{\lambda\, min} \quad (1)$$

NC represents the ratio between the highest and lowest eigenvalue in the correlation matrix. Thus, NC < 100 indicates no serious problems related to multicollinearity (weak), 100 < NC < 1000 indicates moderate problems related to multicollinearity, and NC > 1000 indicates strong evidence of multicollinearity.

With respect to the variance inflation factor (VIF), a VIF < 10 indicates weak multicollinearity, while VIF > 10 indicates elevated multicollinearity. Assessment was performed as follows:

$$VIF = \frac{1}{1-R_j^2} \quad (2)$$

Where, R_j^2 = coefficient of determination.

Coefficient of determination is the ratio of the model's sum of squares (SQM) to total sum of squares.

Multicollinearity diagnosis was performed within each group of variables by the variance inflation factor (VIF) below 10 (Hair *et al.*, 2009) and the condition number (CN) below 100 (Montgomery *et al.*, 1982) using only the variables that did not cause multicollinearity.

Path analysis for each treatment tested (trays with 128 cells and trays with 200 cells) was performed and calculated from the Pearson correlation matrix. The path coefficients were obtained using the methodology proposed by Cruz *et al.* (2012).

Canonical correlation analysis was also carried out for each treatment. The groups formed for this purpose were comprised of shoot variables versus root variables related to the lettuce seedlings produced in 128 and 200-cell trays.

Statistical analyses were performed at 5% significance with the aid of MASS packages for the stepwise analysis (Venables & Ripley, 2002), biotools for the path analysis (Silva *et al.*, 2017b), and yacca for the canonic correlation analyses available on the R program (R Development Core Team, 2016).

RESULTS AND DISCUSSION

The Pearson correlation, which measures degrees of correlation between two variables, presented magnitudes varying between 0.01 and 0.92 for the tray with 128 cells and 0.01 and -0.03 for the 200-cell tray (Table 1).

Table 1: Pearson correlation coefficient between the variables: total length (TL), root length (RL), shoot length (SL), stem diameter (SD), total fresh mass (TFM), fresh root mass (FRM), fresh shoot mass (FSM), dry root mass (DRM), dry shoot mass (DSM), root volume (RV), root area (RA), and leaf area (LA) in seedlings of the Cinderella lettuce produced in 128-cell trays (below the blank diagonal pathway) and in 200-cell trays (above the blank diagonal pathway).

	TL	RL	SL	SD	TFM	FRM	FSM	DRM	DSM	RV	RA	LA
TL	-	0.57*	**0.79***	0.22*	0.39*	0.23*	0.35*	0.14*	0.41*	0.16*	0.16*	0.59*
RL	0.63*	-	0.05	0.20*	0.13*	0.36*	0.07	0.33*	0.12*	0.25*	0.25*	0.10*
SL	**0.76***	0.03	-	0.13*	0.37*	0.01	0.38*	0.08*	0.41*	0.01	0.01	0.64*
SD	0.18*	0.17*	0.08	-	0.22*	0.36*	0.16*	0.41*	0.38*	0.10*	0.12*	0.33*
TFM	0.55*	0.20*	0.54*	0.43*	-	0.16*	**0.98***	0.22*	0.45*	0.14*	0.13*	0.49*
FRM	0.23*	0.27*	0.06	0.24*	0.53*	-	-0.03	0.66*	0.37*	0.41*	0.44*	0.27*
FSM	0.53*	0.10*	0.60*	0.39*	**0.92***	0.14*	-	0.10*	0.38*	0.06	0.05	0.44*
DRM	0.35*	0.48*	0,05	0.33*	0.47*	0.49*	0.31*	-	0.45*	0.34*	0.39*	0.22*
DSM	0.45*	0.25*	0.37*	0.51*	**0.86***	0.39*	**0.81***	0.59*	-	0.21*	0.21*	**0.79***
RV	0.02	0.09*	0.11*	0.02	0.09*	0.11*	0.05	0.01	0.11*	-	**0.86***	0.12*
RA	0.19*	0.24*	0.04	0.23*	0.32*	0.29*	0.23*	0.41*	0.35*	0.57*	-	0.11*
LA	0.53*	0.11*	0.59*	0.43*	**0.89***	0.28*	**0.90***	0.32*	**0.85***	0.08	0.28*	-

* Significant coefficient at 5% probability by *Student's t* test.

For the 128-cell tray a positive and significant correlation was observed between the TFM and FSM variables; i.e., when the TFM increased, the FSM increased, and vice versa. Significant correlation was also found between the TFM and DSM, suggesting that the higher the total fresh mass, the higher the dry mass from the shoot. The LA variable presented positive correlation with the TFM, FSM, and DSM variables. Other significant correlations were found between the DSM and FSM variables and between the SL and TL variables (Table 1).

For the 200-cell tray positive and significant correlation was observed between the TFM and FSM variables such that when TFM increased, so did FSM, and vice versa. Positive and significant correlation was also found between the RA and RV variables, indicating that the higher the root volume, the higher the root area. While there was negative correlation between the FSM and FRM, it was of very low magnitude and not significant. The SLand TL variables and the LA and DSM variables also presented high and significant correlation.

When LA was used as the main dependent variable for path analyses among the seedlings cultivated in the 128-cell tray, the selection of explanatory variables via stepwise resulted in a low condition number (CN) and low variance inflation factor (VIF) (Table 2). Likewise, the decomposition of the linear correlations in direct and indirect effects presented a coefficient of determination of 89% and very low residual effect (0.34), proving that the variables selected by the model by stepwise explain a large portion of the variance observed.

Table 2: Path analysis of the direct (bold text diagonal pathway) and indirect effects of the variables total length (TL), root length (RL), stem diameter (SD), total fresh mass (TFM), fresh root mass (FRM), dry root mass (DRM), dry shoot mass (DSM), root volume (RV), and root area (RA) on the variable leaf area (LA) of a lettuce cultivar cultivated in 128-cell trays

Variables	TL	RL	SD	TFM	FRM	DRM	DSM	RV	RA	r
TL	**0.171**	-0.075	0.005	0.312	-0.030	-0.059	0.192	0.001	0.013	0.529
RL	0.108	**-0.120**	0.005	0.112	-0.036	-0.081	0.106	-0.003	0.017	0.107
SD	0.030	-0.020	**0.029**	0.244	-0.031	-0.056	0.219	-0.001	0.016	0.430
TFM	0.094	-0.024	0.012	**0.568**	-0.069	-0.078	0.368	-0.003	0.022	0.891
FRM	0.039	-0.033	0.007	0.299	**-0.132**	-0.082	0.168	-0.003	0.020	0.282
DRM	0.061	-0.058	0.010	0.264	-0.065	**-0.167**	0.253	0.000	0.028	0.325
DSM	0.076	-0.030	0.015	0.486	-0.051	-0.098	**0.430**	-0.003	0.024	0.848
RV	-0.004	-0.011	0.001	0.049	-0.014	-0.001	0.047	**-0.031**	0.039	0.075
RA	0.032	-0.029	0.007	0.180	-0.039	-0.068	0.149	-0.018	**0.069**	0.284
R^2	0.89									
Residual	0.34									
VIF*	2.94	2.53	1.40	6.78	1.92	2.55	5.65	1.74	2.02	
CN	47.78									

*VIF: Variance inflation factor; CN: Condition number

The variables RL, FRM, DRM, RV, and RA do not possess a cause and effect relationship with the LA variable, as those correlations are negligible. The variables TFM and DSM presented a positive and high, and high direct effect Pearson coefficient (r), indicating robust explanation of LA by the variables TFM and DSM; i.e., there is a direct effect between the variables. The variable TL presented a high r and low direct effect due to indirect effects.

To analyze the path analysis among the seedlings cultivated in the 128-cell tray, using FSM as the main dependent variable, the selection of the explanatory variables via stepwise resulted in a low condition number (CN) and low variance inflation factor (VIF) (Table 3). In the same way, the decomposition of the linear correlations in direct and indirect effects presented a coefficient of determination of 84% and very low residual effect (0.40), proving that the variables selected through the stepwise model explain a great deal of the variation observed.

As their correlations are negligible, the variables RL, SD, FRM, RV, and RA do not possess a cause and effect relationship with the variable FSM. The variable LA presented positive, high, and high direct effect of the Pearson correlation, confirming that FSM is completely explained by the variable LA; i.e., there is a cause and effect relationship between the variables. The variable SL presented high r and negligible direct effect, as this was caused by the indirect effects (Table 3).

Table 3: Path analysis of the direct (bold text diagonal pathway) and indirect effects of the variables root length (RL), shoot length (SL), stem diameter (SD), fresh root mass (FRM), root volume (RV), root area (RA), and leaf area (LA) on the variable fresh shoot mass (FSM) of a lettuce cultivar cultivated in 128-cell trays

Varibles	RL	SL	SD	FRM	RV	RA	LA	R
RL	**0.0405**	-0.0032	0.0052	-0.0340	0.0005	-0.0001	0.0928	0.1016
SL	-0.0013	**0.0968**	0.0026	-0.0073	-0.0005	0.0000	0.5098	0.6000
SD	0.0068	0.0081	**0.0312**	-0.0295	0.0001	-0.0001	0.3713	0.3879
FRM	0.0111	0.0057	0.0075	**-0.1238**	0.0006	-0.0002	0.2438	0.1447
RV	0.0038	-0.0102	0.0007	-0.0131	**0.0052**	-0.0003	0.0650	0.0510
RA	0.0098	0.0037	0.0073	-0.0362	0.0030	**-0.0006**	0.2458	0.2328
LA	0.0043	0.0571	0.0134	-0.0349	0.0004	-0.0002	**0.8643**	0.9045
R^2	0.84							
Residual	0.40							
VIF*	1.13	1.71	1.35	1.22	1.55	1.78	2.18	
CN	7.90							

*VIF: Variance inflation factor; NC: Condition number

For the path analysis, with LA as the dependent variable in seedlings cultivated in the 200-cell tray, the multicollinearity criteria were satisfied (Table 4). The direct and indirect effects presented a coefficient of determination of 76%. The variables RL, SD, TFM, FRM, RV, and RA did not possess a cause and effect relationship with the variable LA, as the correlations are negligible. The variables TL and DSM presented positive and high Pearson correlation, along with high direct effect, confirming that LA is completely explained by the variables TL and DSM, meaning there is a cause and effect relationship between the variables (Table 4).

Table 4: Path analysis of the direct (bold text diagonal pathway) and indirect effects of the variables total length (TL), root length (RL), stem diameter (SD), total fresh mass (TFM), fresh root mass (FRM), dry root mass (DRM), dry shoot mass (DSM), root volume (RV), and root area (RA) on the variable leaf area (LA) of a lettuce cultivar cultivated in 200-cell trays

Variables	TL	RL	SD	TFM	FRM	DRM	DSM	RV	RA	r
TL	**0.408**	-0.119	0.010	0.033	0.018	-0.015	0.257	-0.001	-0.006	0.586
RL	0.233	**-0.208**	0.009	0.011	0.027	-0.036	0.073	-0.001	-0.009	0.099
SD	0.091	-0.041	**0.045**	0.019	0.027	-0.044	0.239	0.000	-0.004	0.330
TFM	0.159	-0.028	0.010	**0.085**	0.012	-0.024	0.280	-0.001	-0.005	0.489
FRM	0.096	-0.076	0.016	0.013	**0.075**	-0.072	0.234	-0.002	-0.016	0.268
DRM	0.055	-0.069	0.018	0.019	0.050	**-0.108**	0.284	-0.002	-0.014	0.233
DSM	0.168	-0.024	0.017	0.038	0.028	-0.049	**0.625**	-0.001	-0.008	0.794
RV	0.066	-0.051	0.005	0.012	0.030	-0.037	0.130	**-0.005**	-0.031	0.118
RA	0.067	-0.052	0.005	0.011	0.033	-0.042	0.132	-0.004	**-0.037**	0.113
R^2	0.762									
Residual	0.488									
VIF*	2.163	1.936	1.316	1.363	2.057	2.326	1.885	3.880	4.019	
CN	25.359									

*VIF: Variance inflation factor; CN: Condition number

Multicollinearity was low when FSM was the main dependent variable (Table 5). The coefficient of determination was 76%. The variables RL, SD, TFM, FRM, DRM, RV, and RA do not possess a cause and effect relationship with the variable FSM, as the correlations are negligible. The variables TL and DSM have a cause and effect relationship with FSM, confirming that FSM is completely explained by the variables TL and DSM (Table 5).

Table 5: Path analysis of the direct (bold text diagonal pathway) and indirect effects of the variables total length (TL), root length (RL), stem diameter (SD), total fresh mass (TFM), fresh root mass (FRM), dry root mass (DRM), dry shoot mass (DSM), root volume (RV), and root area (RA) on the variable fresh shoot mass (FSM) of a lettuce cultivar cultivated in 200-cell trays

Variables	TL	RL	SD	TFM	FRM	DRM	DSM	RV	RA	r
TL	**0.41**	-0.12	0.01	0.03	0.02	-0.01	0.26	-0.001	-0.006	0.586
RL	0.23	**-0.21**	0.01	0.01	0.03	-0.04	0.07	-0.001	-0.009	0.099
SD	0.09	-0.04	**0.04**	0.02	0.03	-0.04	0.24	-0.001	-0.004	0.329
TFM	0.16	-0.03	0.01	**0.08**	0.01	-0.02	0.28	-0.001	-0.005	0.489
FRM	0.10	-0.08	0.02	0.01	**0.07**	-0.07	0.23	-0.002	-0.016	0.268
DRM	0.06	-0.07	0.02	0.02	0.05	**-0.11**	0.28	-0.002	-0.014	0.233
DSM	0.17	-0.02	0.02	0.04	0.03	-0.05	**0.63**	-0.001	-0.008	0.794
RV	0.07	-0.05	0.00	0.01	0.03	-0.04	0.13	**-0.005**	-0.031	0.118
RA	0.07	-0.05	0.01	0.01	0.03	-0.04	0.13	-0.004	**-0.037**	0.113
R^2	0.76									
Residual	0.49									
VIF *	2.16	1.94	1.32	1.36	2.06	2.33	1.88	3.88	4.02	
CN	25.36									

* VIF: Variance inflation factor; CN: Condition number

The canonical correlation analyses were carried out after diagnoses of multicollinearity within each group of variables and they presented relatively low values. The canonical correlation interpretations were performed considering the significance of the F test and the magnitude of correlation of the variable groups. While the canonical correlations among the shoot and root variables of lettuce cultivated in 128-cell trays presented the first three significant pairs, only the first pair presented elevated magnitude, showing the dependence of the groups and the ability to use their coefficients to study the variables of these groups. The first canonical pair presented a correlation of r = 0.697, with the canonical cross load of the group of DSM shoot variables possessing high magnitude and negative values. The fact that the DRM and FRM variables bear the same magnitude and negativity indicates that while a reduction in DRM and FRM would reduce DSM, increases in these variables would increase the dry shoot variable as well (Table 6).

For the lettuce seedlings cultivated in 200-cell trays, the canonical correlations presented the first three significant pairs; however, only the first pair presents elevated magnitude and can be used to explain the dependencies existing among the groups of variables. The first canonical pair presented a correlation of r = 0.551 and, as in the case with the lettuce seedlings cultivated in 128-cell trays, the canonical cross load results of the DSM shoot variable group are influenced by variables of the DRM and FRM root groups. These variables cause a reduction in DSM when compromised and an increase in DSM when the variables increase (Table 7). With respect to the 128-cell tray, the significant and positive relationships between the TFM and FSM variables and between the TFM and DSM variables were expected, as the largest portion of total fresh mass is composed of fresh shoot mass. The same association exists for TFM and DSM.

The correlations between LA and the variables TFM, FSM, and DSM were high, positive, and significant. This indicates a selection based on any one of these three variables will positively influence LA or that an increase in LA will help cause increases in the other three. The advantage here is that by selecting an easily measured characteristic (such as fresh shoot mass), one can obtain important results for characteristics that are not easily measured (such as leaf area). While there were correlations between the variables DSM, FSM, TL, and SL, they were not obvious, as they are interdependent variables.

Table 6: Canonical correlations between the shoot and root variable groups of lettuce seedlings cultivated in 128-cell trays

Variables	G1	G2	G3	G4
Shoot variables				
SL	-0.039	0.077	-0.047	-0.146
FSM	-0.306	-0.018	0.080	-0.120
DSM	-0.606	0.031	0.047	-0.063
LA	-0.349	0.130	0.076	-0.101
Root variables				
RL	-0.310	-0.045	-0.005	0.045
FRM	-0.449	0.253	-0.034	0.031
DRM	-0.676	-0.060	-0.013	-0.022
RV	-0.118	0.025	0.139	0.084
RA	-0.330	0.094	0.133	-0.039
Canonical correlation	0.697	0.356	0.171	0.156
Degree of freedom	20	12	6	2
p- value	0.00	0.00	0.0001143	NS

Table 7: Canonical correlations between the shoot and root variable groups of lettuce seedlings cultivated in 200-cell trays

Variables	G1	G2	G3	G4
Shoot variables				
SL	0.068	0.099	0.043	0.038
FSM	-0.086	-0.157	-0.030	0.034
DSM	-0.467	0.053	0.011	0.022
LA	-0.247	0.123	-0.037	0.032
Root variables				
RL	-0.144	-0.016	-0.110	0.014
FRM	-0.400	0.178	-0.018	0.007
DRM	-0.545	-0.027	-0.009	-0.002
RV	-0.221	-0.005	0.026	0.040
RA	-0.226	0.004	0.042	0.030
Canonical correlation	0.551	0.277	0.133	0.044
Degree of freedom	20	12	6	2
p-value	0.00	0.00	0.01564	NS

There was a relationship in the 200-cell trays between the root variables RA and RV such that the higher the root volume, the greater the area and vice versa. This relationship is beneficial since a more developed root system, in terms of quantity, especially with regard to the thin roots, can result in a more vigorous seedling with enhanced survival probability in the field under poor nutritional and water supply conditions (Oliveira *et al.*, 2006). In addition, root volume measuring techniques are easy to perform and, in light of the correlation, may be utilized to infer root area.

Although correlation studies can be useful for determining the linear relationships that influence seedling quality, they do not prove the direct and indirect effects of the independent variables (Cruz *et al.*, 2012). Thus, path analyses can realistically direct us to the contribution of each variable. According to Silva *et al.* (2010), characters with high positive correlation to the main variable and direct favorable effect indicate the presence of cause and effect. Thus, the path analyses revealed the cause and effect relationship of the dependent variable LA with the independent characters TFM and FSM in the 128-cell tray (Table 2) and for the dependent variable FSM with the independent variable LA (Table 3), also in the 128-cell tray. Such results were expected, as the leaf area will influence both the total fresh mass as well as the fresh shoot mass and vice versa. Olivoto *et al.* (2015), assessed the relationships between biometric parameters in cucumber seedlings in a natural setting and a protected setting and in the latter case obtained results similar to those reported in the present work. Those authors determined that the dependent variable LA had a direct effect on the FSM variable, providing evidence the accumulation of fresh shoot mass contributes to an increase in leaf area. The same authors observed the direct effect of DSM on the dependent variable, affirming its role in increasing the variable. The indirect effects were observed with the dry shoot mass (DSM), plant height (PH), and root length (RL).

Path analysis with the 200-cell tray revealed a cause and effect relationship of the dependent variable LA with the independent variables TL and DSM (Table 4). For the dependent variable FSM in the 200-cell tray there was also a cause and effect relationship with the independent variables TL and DSM (Table 5). Using dry root mass (DRM) as the dependent variable with guarana seedlings, Nascimento Filho *et al.* (1993), found strong negative correlation with the dry shoot mass (DSM), indicating a lack of cause and effect, and explained this finding through indirect influence via the branch basal diameter and main leaf area. As in the present work, those authors confirmed that the dry shoot mass had large influence on the leaf area and vice versa. With respect to TL, although Gomes & Paiva (2004), asserted the use of seedling length was inadequate as the sole criterion for assessing quality standards, this variable has potential as a quality indicator due to its correlation with fresh mass and ease of measurement.

Dardengo *et al.* (2013), observed with coffee seedlings that although the total dry matter and stem diameter exerted the greatest direct effect on DQI, the dry matter of the shoot and root also presented high positive correlation with DQI. These results reaffirm the influence of dry shoot mass on features indicating quality in seedlings.

While variable association studies with lettuce seedlings are scarce in the literature, they are more plentiful with seedlings of other species. With guarana, for example, Nascimento Filho *et al.* (2012), found strong correlation between the root and shoot groups, indicating canonical correlation and that the groups are dependent on each other. The variables with the greatest relationship were the branch length (BL) with the dry root matter weight (DRW), suggesting the ability to improve root systems through increasing branch length. Thus, branch lengths can be measured in order to indicate root development without the need to destroy seedlings.

Bouteillé *et al.*, (2012) report common QTL for root and shoot growth in Arabidopsis highlighting the relationship between both plant parts. The authors also suggest carbon partitioning as responsible for root-shoot growth balance. Shipley & Meziane (2002) tested the balance growth hypothesis and allometry of leaf and root biomass allocation. They found that the balance growth hypothesis is suitable for root-shoot growth dynamics. According to this hypothesis biomass allocation will occur preferentially towards those organs under limited resource growth. Thus, the balance could shift according to the environment and under a non limiting environment seedlings show a near equal allocation. This could explain the findings in the present study since seedlings production was not affected by stresses and correlations between growth traits are observed for almost all variables.

The correlations reported in this work are relevant and can be used in future studies. These results indicate that some variables exert influence over each other, other variables possess a cause and effect relationship, and the root and shoot groups have a relationship with each other, suggesting that gains or losses by one group will result in gains or losses of the other.

CONCLUSION

In conclusion, there is a linear relationship among biometric variables in lettuce seedlings. Thus, it is possible to select simple features of analysis, or ones that require less complex equipment, to make inferences about aspects of another associated variable related to adequate seedling development and, consequently, indicate quality level. The fresh shoot mass can be used as an indicator of quality, as it is positively related to other aspects such as leaf area and dry mass, as well as root volume as indicative of root area with the advantage that the former could be obtained more simply. In addition to allowing greater efficiency in evaluations in scientific research, this result also represents an alternative for commercial producers who can carry out quality assessments of lettuce seedlings through simple measurement variables.

ACKNOWLEDGEMENTS, FINANCIAL SUPPORT AND FULL DISCLOSURE

This work received financial support from the Brazilian federal agency *Coordenação de Aperfeiçoamento de Pes-*

soal de Nível Superior (CAPES), Brazil (Financial Code 001).

We thank our colleagues from the laboratory: Leandro, Enzo, Letícia, Lia, Israel, Patricia, André and Francielle for their help in collecting the plant data and analyzing. We would also to thank Pesagro: José Márcio, Fátima and Amaro for assistance with seedling production.

The authors declare they have no conflicts of interest. The founding sponsors had no role in the study design, analysis or interpretation of data, writing of the manuscript, or decision to publish the results.

REFERENCES

Ali Q, Ahsan M & Bin Mustafa HS (2013) Studies of genetic variability and correlation among morphological traits of maize (Zea mays L.) at seedling stage. Albanian Journal of Agricultural Sciences, 12:405-410.

Ayubov MS, Norov TM, Saha S, Tseng TM, Reddy KR, Jenkins JN, Abdurakhmonov IY & Stelly DM (2021) Alteration of root and shoot morphologies by interspecific replacement of individual Upland cotton chromosome or chromosome segment pairs. Euphytica, 217:01-23.

Bezerra Neto FV, Leal NR, Costa FR, Gonçalves GM, Do Amaral Júnior AT, Vasconcellos HO & Mello M (2006) Análise biométrica de linhagens de abóbora. Horticultura Brasileira, 24:378-380.

Bouteillé M, Rolland G, Balsera C, Loudet O & Muller B (2012) Disentangling the intertwined genetic bases of root and shoot growth in Arabidopsis. PLoS One, 7:e32319.

Campanharo M, Rodrigues JJV, Lira Junior MA, Espindula MC & Costa JVT (2006) Características físicas de diferentes substratos para produção de mudas de tomateiro. Revista Caatinga, 19:140-145.

Cruz CD, Regazzi AJ & Carneiro PCS (2012) Modelos biométricos aplicados ao melhoramento genético. 4ª ed. Viçosa, UFV. 514p.

Dardengo MCJD, Sousa EF, Reis EF & Gravina GA (2013) Crescimento e qualidade de mudas de café conilon produzidas em diferentes recipientes e níveis de sombreamento. Coffee Science, 8:500-509.

Embrapa - Empresa Brasileira de Pesquisa Agropecuária (2010) Safira – Software para análise de fibras e raízes por imagem. Available at: https://www.embrapa.br/instrumentacao/busca-de-solucoes-tecnologicas/-/produto-servico/2016/safira---software-para-analise-de-fibras-e-raizes-por-imagem. Accessed on: April 17th, 2019.

Gomes JM & Paiva HN (2004) Viveiro florestais – propagação sexuada. 3ª ed. Viçosa, UFV. 116p.

Guimarães VF, Echer MM & Minami K (2002) Métodos de produção de mudas, distribuição de matéria seca e produtividade de plantas beterraba. Horticultura Brasileira, 20:505-509.

Hair JF, Black WC, Babin BJ, Anderson RE & Tatham RL (2009) Análise multivariada de dados. 6ª ed. Brasil, Bookman Editora. 688p.

IBGE – Instituto Brasileiro de Geografia e Estatística (2019) Censo Agropecuário 2017. Available at: <https://censoagro2017.ibge.gov.br/templates/censo_agro/resultadosagro/indexhtml>. Accessed on: April 17th, 2019.

Johnson JD & Cline ML (1991) Seedling quality of southern pines. Forest regeneration manual, 36:143-159.

Medical Software Brolkstraat (2019) Digimizer image analysis software. Version 5.3.5. Belgium, MedCalc. Available at: https://www.digimizer.com/. Accessed on: April 17th, 2019.

Montgomery DC, Peck EA & Vining GG (2012) Introduction to linear regression analysis. 5º ed. United States of America, John Wiley & Sons. 672p.

Nascimento Filho FJ, Ando A, Cruz CD & Garcia TB (1993) Análise de caminhamento em mudas de guaraná. Pesquisa Agropecuária Brasileira, 28:447-452.

Nascimento Filho FJ, Atroch AL & Cecon PR (2012) Associação entre características da parte aérea e do sistema radicular em mudas de guaranazeiro por correlações canônicas. Acta Amazonica, 42:241-244.

Oliveira AN, Da Silva AC, Da Silva SCR & Rodrigues ÉAC (2006) Variações genéticas para características do sistema radicular de mudas de baru (Dipteryx alata Vog.). Revista Árvore, 30:905-909.

Olivoto T, De Souza VQ, Carvalho IRC, Nardino M & Follmann DN (2015) Análise de trilha para caracteres relacionados ao crescimento de mudas de pepineiro. Enciclopédia Biosfera, 11:69-80.

R Development Core Team (2016) R: a language and environment for statistical computing. Vienna, R Foundation for Statistical Computing. Available at: https://www.r-project.org/. Accessed on: September 17th, 2019.

Sala FC & Costa CP (2012) Retrospectiva e tendência da alfacicultura brasileira. Horticultura Brasileira, 30:187-194.

Shipley B & Meziane D (2002) The balanced-growth hypothesis and the allometry of leaf and root biomass allocation. Functional Ecology, 16:326-331.

Silva AC, Silva VSG, Mantovanelli BC & Santos GM (2017a) Formação de mudas de alface em diferentes bandejas e substratos. Revista da Universidade Vale do Rio Verde, 15:465-471.

Silva AR, Malafaia G & Menezes IPP (2017b) Biotools: an R function to predict spatial gene diversity via an individual-based approach. Genetics and Molecular Research, 16:gmr16029655.

Silva JI, Vieira HD, Viana AP & Barroso DG (2010) Desenvolvimento de mudas de *Coffea canephora* Pierre ex A. froehner em diferentes combinações de substrato e recipiente. Coffee Science, 5:38-48.

Souza SR, Fontinele YR, Saldanha CS, Araújo Neto SE & Kusdra JF (2008) Produção de mudas de alface com o uso de substrato preparado com coprólitos de minhoca. Ciência e Agrotecnologia, 32:115-121.

Tsakaldimi M, Ganatsas P & Jacobs DF (2013) Prediction of planted seedling survival of five Mediterranean species based on initial seedling morphology. New forests, 44:327-339.

Venables WN & Ripley BD (2002) Modern applied statistics with S. 4º ed. New York, Springer. 501p.

Volpato & Barreto (2016) Estatística sem dor!!!. 2ª ed. Botucatu, Best Writing. 160p.

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

Zhang H, Xue Y, Wang Z, Yang J & Zhang J (2009) Morphological and physiological traits of roots and their relationships with shoot growth in "super" rice. Field Crops Research, 113:31-40.