NATURA MONTENEGRINA, Podgorica, 2013, 12(1): 205-216

ORIGINAL RESEARCH PAPER

SEQUENTIAL PATH ANALYSIS OF SPINACH YIELD USING SEVERAL QUANTITATIVE AND QUALITATIVE TRAITS

Naser SABAGHNIA[1] Hossein Ali ASADI-GHARNEH[2] & Mohsen JANMOHAMMADI[1]

1 Dep. of Agronomy and Plant Breeding, Faculty of Agriculture, University of Maragheh, Maragheh, Iran.
Email: sabaghnia@maragheh.ac.ir

2 Department of Horticulture Science, Islamic Azad University, Khorasgan Branch, Isfahan, Iran.

Key words:
bootstrap analysis,
conventional path analysis,
multi-collinearity.

SYNOPSIS

Development of new spinach (*Spinacia oleracea* L.) verities needs efficient tools to mange character relationships in a breeding program. This research uses correlation coefficients as well as sequential path analysis to determine the interrelationships among dry yield and several quantitative and qualitative related traits. Fifty-four spinach landraces, which collected from different geographical regions of Iran, were grown to determine the important components of dry yield. Sequential path analysis identified the petiole diameter, petiole length and days to flowering as important first order traits that influenced dry yield. Leaf length, leaf width, leaf numbers in flowering and stem anthocyanin content were indentified as important second order traits that influenced dry yield. All direct effects were significant, as indicated by bootstrap analysis. The results suggest that petiole and leaf properties could be used as a selection criterion in selecting for increased yield in spinach. Such similar outcome could be used in the future to delineate predictive, more rigorous secetion strategies as well as to help define indicator traits for recommendations for spinach and other vegetable crops.

INTRODUCTION

Spinach (*Spinacia oleracea* L.) is one of the important salad vegetable crops providing remarkable amounts of vitamin A, vitamin C, calcium and iron. Its increasing consumption during the two past decades can be related to growing public awareness of the health benefits (Brandenberger et al., 2007). According to Goldman (2003) functional foods have health benefits as well as fulfill the human's nutritional needs and spinach could be regarded as functional foods which is increasing antioxidant intake in diets (Brandenberger et al., 2004). Spinach as cool season crop is grown for fresh or processing market and is resistant to low temperatures and drought reduce yield and deteriorate the quality in spinach (Salk et al., 2008). Different spinach breeding programs aim to incorporate important traits into improved spinach cultivars including high yield, adaptation, resistances diseases, pests, and herbicides, and nutritional quality (Murphy & Morelock, 2000).

Awareness from traits relationships with yield analysis is important in breeding programs that can enhance some yield components. Therefore, the study of the correlation among yield-related traits can help the plant breeders in selection for higher yields. Different statistical methods have been introduced in modeling yield characteristic, including simple correlation coefficient, linear regression model, path analysis, factor analysis and clustering procedure. Path analysis is a statistical method for partitioning the correlation coefficients into direct and indirect effects, so that the contribution of each characteristic to yield could be estimated (Sabaghnia et al., 2010). The coefficients of path analysis are standardized partial regression coefficients which are scaled in terms of standard deviations so that the direct and indirect effects related with different yield components.

Most of path analysis investigations ignored the importance of the causal relationship and usually used a model, in which bidirectional causation among yield components is assumed (Mohammadi et al., 2003). Path analysis needs the determination of causal relationships among different characteristics, based on either a priori evidence or a postulated hypothesis. Yield components develop sequentially with later-developing characteristics under the control of earlier-developing. The analysis of the values of correlation coefficients of the different characteristics with yield helps in deciding their importance as selection criteria for yielding ability.

This conventional path analysis might result in multi-collinearity for involved traits, particularly when correlations among traits were high. The multi-collinearity maybe causes difficulties in interpretation of the actual contribution of each trait. Feyzian et al. (2009) used a sequential path analysis for determining the relationships among yield and its components in rice by analyzing several predictor traits in first, second and third order paths. Recently sequential path analysis is used in different crops including maize (Mohammadi et al., 2003), potato (Asghari-Zakaria et al., 2007) and rapeseed (Sabaghnia et al., 2010). Although, there are several investigations about correlation and path coefficient analysis of several agronomic crops, but such detailed cause and effect relationships using sequential path analysis have not been examined in horticultural crops. The main objectives of this investigation were to analyze the correlation between yield and related traits in spinach by applying the sequential path analysis and identify traits of genotypes which may prove useful in breeding higher-yielding genotypes.

MATERIALS AND METHODS

For the purpose of this investigation, fifty-four native Iranian spinach germplasms were collected from different spinach producing areas of Iran. The landraces were evaluated in the field using a randomized complete block design (RCBD) layout with four replicates. The geographical characteristics of the 54 regions of the spinach landraces were given in Table 1. Field soil was calcareous, loamy structure, low organic matter, low salt content and had poor nitrogen and phosphorous. Fertilization was done by spreading 80 kg N ha^{-1} (half at sowing and half at seedling emergence) and sowing was performed at the rate of 50 kg seed ha^{-1}. Weed control was performed twice when the weed density was high, in the pre-flowering and post-flowering stages. Each field plot contained six rows (3 m long with 25 cm space between rows) and 4.5 m^2 plot size. The harvested plot size was 2.5 m^2 (four 2.5 rows at the center of each plot).

Several quantitative traits consist on leaf length (LL), leaf width (LW), petiole length (PL), petiole diameter (PD), leaf area (LA), leaf numbers in flowering (LN), days to flowering (DF), female plants percent (FP), fresh yield (FY) and dry yield (DY) were measured. Also, various qualitative traits consist on leaf texture (LT, 1=smooth, 2=slight crinkled, 3= crinkled), seed type (ST, 1=smooth, 2=prickly), stem anthocyanin (SA, 1=very low, 3=low, 5=intermediate, 7=high, 9=very high), petiole attitude (PA, 1=erect, 2=semi-spared, 3= spared), vegetative leaf shape (VL, 1=elliptic, 2=broad elliptic, 3=circular, 4=ovate, 5=broad ovate, 6=triangular), reproductive leaf shape (RL, 1=smooth, 2=Pointy); leaf edge (LE, 1=smooth, 2= Rippler); leaf color (LC, 1=yellow-green, 2=grey-green, 3=blue-green); seed color (SC, 1=yellow-green, 2=grey-green, 3=blue-green) were measured.

The datasets were first tested for normality by Anderson and Darling normality test using MINITAB version 16 (2010) software. The correlation coefficients were calculated for all the possible comparisons using the Pearson (quantitative traits) and Spearman (qualitative traits) correlation coefficients. The correlation coefficients were partitioned into direct and indirect effects using the conventional path coefficient analysis. Sequential stepwise regression was performed to organize the predictor traits into first, second and third order paths on the basis of their respective contributions to the total variation of dry yield and minimal collinearity by SPSS 16 (SPSS, 2008).

The sequential path model consisted of predictor and response variables. The level of multi-collinearity in each component path was measured from two common measures, namely the Tolerance value and its inverse, the Variance Inflation Factor (VIF). Tolerance value is the amount of variability of the selected independent variable not explained by other independent variables ($1-R^2$), where R^2 is the coefficient of determination for the prediction variable by the predictor variables). Variance inflation factor (VIF) indicates the extent of effects of other independent variables on the variance of the selected independent variable [VIF $=1/(1/\ R^2)$]. Thus, very small tolerance values (much below 0.1) or large variance inflation factor values (above 10) indicate high collinearity (Mohammadi et al., 2003). Partial coefficients of determination (analogous of linear regression) were calculated from the path coefficients for all predictor variables. To estimate the standard error path coefficients, bootstrap analysis was performed by S-PLUS version 8 (Springr-Verlag Inc., 2007) statistical package.

RESULTS AND DISCUSSION

The results of ANOVA indicated highly significant differences among spinach landraces for all the measured traits under investigation (data are not shown). In order to determine in the most precise manner the interrelation of direct and indirect dry yield components, simple correlation coefficients were established and then path coefficient analysis was performed. The Pearson simple correlation coefficients among quantitative traits (Table 1) showed that there were significant positive correlations between leaf length (LL) and all of the measured traits except days to flowering (DF) and female plants percent (FP). Eftekhari et al. (2010) reported relatively similar results for association between LL and most of traits which are mentioned here, but they found negative significant correlation between LL and FP. The petiole length showed significant positive correlation with leaf area (LA), leaf numbers in flowering (LN), and dry yield (DY). Asadi & Hasandokht (2007) found similar results for interrelationships among DY, PL and PD.

The leaf width had significant positive correlation with other traits except petiole diameter (PD) and female plants percent (FP). The petiole diameter indicated positive significant correlation with leaf area (LA), leaf numbers in flowering (LN), days to flowering (DF), and dry yield (DY). The leaf area (LA) had significant positive correlation with leaf numbers in flowering (LN), days to flowering (DF), and dry yield (DY). The leaf numbers in flowering (LN) indicated significant positive correlation with to flowering (DF), and dry yield (DY). It is clear that all quantitative traits except female plants percent (FP) had significant positive correlation with dry yield and so it is better to extract the direct and indirect of these traits on spinach dry yield via path analysis. The obtained results verified the investigations of Asadi & Hasandokht (2007) and

Eftekhari et al. (2010) on some Iranian spinach landraces and showed the importance of most spinach quantitative traits on dry yield.

Table 1: Pearson's correlation coefficients between nine quantitative traits of 54 Iranian spinach landraces.

	LL	LW	PL	PD	LA	LN	DF	FP
LW	0.53**							
PL	0.50**	0.77**						
PD	0.64**	0.18ns	0.25ns					
LA	0.78**	0.87**	0.68**	0.46**				
LN	0.46**	0.53**	0.55**	0.41**	0.61**			
DF	0.10ns	0.31*	0.19ns	0.28*	0.27*	0.33**		
FP	0.20ns	0.09ns	0.02ns	0.06ns	0.12ns	0.00ns	0.16ns	
DY	0.49**	0.36**	0.40**	0.70**	0.48**	0.50**	0.57**	-0.04ns

†Critical values of correlation *P* < 0.05 and *P* < 0.01 (df= 52) are 0.35 and 0.27, respectively.
** Significant at 1% level; * Significant at 5% level; ns Non-significant.

Table 2: Spearman's correlation coefficients between eight qualitative traits of 54 Iranian spinach landraces.

	LT	ST	SA	PA	VL	RL	LE
ST	0.10ns						
SA	-0.06ns	-0.15ns					
PA	-0.17ns	-0.19ns	0.23ns				
VL	-0.16ns	-0.27*	0.18ns	0.27ns			
RL	-0.08ns	-0.16ns	0.16ns	0.17ns	0.41**		
LE	0.19ns	0.05ns	0.05ns	0.00ns	0.31*	0.20ns	
LC	0.18ns	-0.29*	0.34*	0.04ns	0.23ns	0.01ns	0.31*

LT, Leaf texture; Seed type; SA, Stem anthocyanin; PA, Petiole attitude; VL, Vegetative leaf shape; RL, Reproductive leaf shape; LE, Leaf edge; LC, Leaf color; SC, Seed color.
†Critical values of correlation *P* < 0.05 and *P* < 0.01 (df= 52) are 0.35 and 0.27, respectively.
** Significant at 1% level; * Significant at 5% level; ns Non-significant.

The Spearman's rank correlation coefficients between pairs of qualitative traits of spinach (Table 2) indicated that most of these traits were not correlated with each other. Only, vegetative leaf shape (VL) had significant positive correlation with reproductive leaf shape (RL) and leaf edge (LE). Also, leaf color showed significant positive correlation with content of stem anthocyanin (SA) and leaf edge (LE); and indicated significant negative correlation with seed type (SA). Our results are in good agreement with the findings of Asadi & Hasandokht (2007) in investigation of some qualitative traits in Iranian spinach landraces. Therefore it seems that, the most of spinach qualitative traits are not good indicators for assessing yield performance quantity, but they maybe proper for increasing quality of spinach.

Table 3: Direct effects of first-order predictor variables on the dry yield (DY) of 54 spinach landraces and two common measures of collinearity indices in path analysis.

	Direct effect	Tolerance	VIF†
LL	0.30	0.21	4.75
LW	0.20	0.09	10.78
PL	0.05	0.29	3.46
PD	0.54	0.33	3.00
LA	-0.44	0.08	13.29
LN	0.24	0.39	2.57
DF	0.41	0.60	1.65
FP	-0.06	0.60	1.65
LTT	-0.05	0.72	1.40
ST	0.03	0.58	1.73
SA	-0.26	0.42	2.36
PA	0.12	0.55	1.80
VL	-0.07	0.58	1.71
RL	0.25	0.64	1.56
lee	-0.03	0.66	1.52
lc	0.05	0.57	1.75

†VIF: variance inflation factor.

To determine the relative importance of the measured traits on spinach dry yield, the data were subjected to conventional path analysis which permits the separation of simple correlation coefficients with components of direct and indirect effects. The results pertaining to direct effects of components traits, where the traits were considered as first order variables with spinach dry yield as the response variable, are presented in Table 3. Also, the results of two measures of multicollinearity analysis (Tolerance and Variance Inflation Factor) for this path analysis are given in Table 3. According to the results of conventional path analysis and multicollinearity analysis, there are inconsistent patterns of relationships among the studied variables. Results of this analysis, where the all traits were considered as the first-order variables (Model I) with dry yield as the response variable, indicated high multicollinearity for some traits, particularly for those showing high direct effects such as leaf width (LW) and leaf area (LA) and thus these traits were removed as the first-order variables from the analysis. Such similar strategy in evaluation of different traits correlations in path analysis was used by Feyzian et al. (2009) in rice, Mohammadi et al. (2003) in maize and Sabaghnia et al. (2010) in rapeseed.

Estimation of direct effects by the second strategy (sequential path analysis), where the yield-related traits were considered as grouped into first, second, and third-order variables with dry yield (Model II), and analysis of multicollinearity indicated a better understanding of the interrelationships among various traits and their relative contribution to spinach dry yield (Table 4). Results of Tolerance and variance inflation factor (VIF) values for the predictor variables indicated the remarkable reduction of VIF values in Model I in comparison to Model II. It is clear that stepwise regression in the present study minimized the collinearity measures (Tolerance and VIF) of all variables, thereby facilitating detection of actual contributions of each predictor variables in different path components, with negligible confounding effects and interference. The advantage of sequential path procedure over conventional path analysis in minimizing collinearity problems and identifying actual contributions of each component in different path components are similar to those reported in the other investigations (Asghari-Zakaria et al., 2007; Sabaghnia et al., 2010; Ibrahim & Ramadan, 2013), indicating that it would be very effective to achieve a favorable results.

Table 4: Measures of collinearity values (tolerance and variance inflation factor, VIF) for predictor variables in conventional path analysis (CPA, all predictor variables as first-order variables) and sequential path analysis (SPA, predictors grouped into first-, second-, and third-order variables).

Variable		Tolerance		VIF	
Predictor	Response	CPA	SPA	CPA	SPA
DY	PD	0.33	0.88	3.00	1.13
	DF	0.60	0.91	1.65	1.10
	PL	0.29	0.92	3.46	1.09
PD	LL	0.21	0.96	4.75	1.04
	SA	0.42	0.96	2.36	1.04
DF	LN	0.39	1	2.57	1
PL	LW	0.09	1	10.78	1

In most statistical analysis procedures, the plant breeder is usually interested in obtaining not only a point estimate of a statistic but also an estimate of the variation in this point estimate and a confidence interval for the true value of the parameter. Resampling techniques, such as the bootstrap, provide good estimates of the standard error, confidence intervals, and distributions for any statistics. To apply above procedures on path analysis, the mean direct effects estimated from a set of 1000 bootstrap samples were in close agreement with the observed direct effects of different measured traits (Table 5). The low standard error for all the direct effects and low bias also indicated the robustness of the sequential path analysis. The T-test of significance, using standard error values obtained through bootstrap resampling, indicated that all the direct effects were significant (data are not shown).

The adjusted coefficient of determination (R^2 Adj. = 0.67) represents the influence of the petiole diameter (PD), days to flowering (DF), and petiole length (PL) traits as first-order variables involved in the research on total variability of dry yield (Table 5). Among these three traits, PD, DF and PL traits, PD had the greater direct effect (0.55) than two other traits on spinach dry yield (Table 6). Indirect effect to the DY was low and positive (0.11) via DF but indirect effect to the DY was very low and positive (0.05) via PL. The DF had the moderate direct effect (0.38) on spinach dry yield and indirect effect to the DY was low and positive (0.15) via PD but indirect effect to the DY was very low and positive (0.04) via PL (Table 6). The PL had the low direct effect (0.19) on dry yield and indirect effect to the DY was low and positive (0.14) via PD but indirect effect to the DY was very low and positive (0.07) via DF (Table 6).

Table 5: Estimation of standard error values of path coefficients (direct effects) using bootstrap analysis.

Variable		Direct effect	Adj. R^2	Bootstrap		
Predictor	Predictor			Mean	Bias	SE
DY	PD	0.55	0.67	0.5643	0.0141	0.1003
	DF	0.38		0.3951	0.0192	0.1174
	PL	0.19		0.1676	-0.0189	0.1321
PD	LL	0.59	0.48	0.5938	-0.0170	0.1405
	SA	0.27		0.2689	0.0168	0.1430
DF	LN	0.33	0.33	0.3264	-0.0036	0.1222
PL	LW	0.77	0.60	0.7731	0.0069	0.1421

Table 6: Direct and indirect effects for the predictor variables in sequential path analysis (grouped into first, second and third order variables).

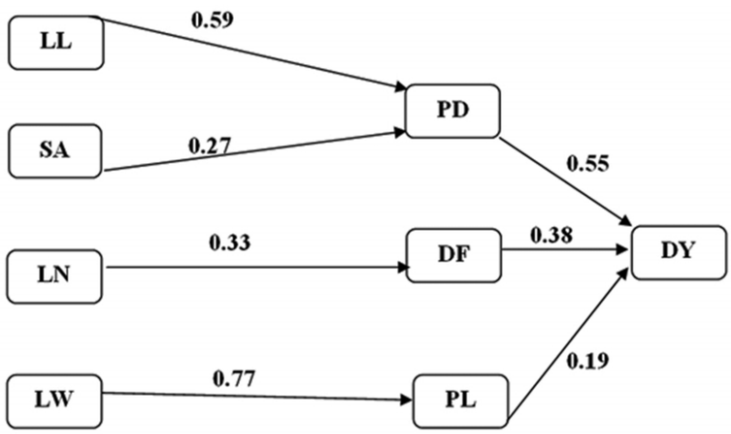
DY			
	PD	DF	PL
PD	0.55	0.11	0.05
DF	0.15	0.38	0.04
PL	0.14	0.07	0.19
PD			
	LL	SA	
LL	0.05	0.27	
SA	0.64	0.38	

According to this famous belief that “a picture is worth a thousand words”, large attempts have been performed to graphically present statistical outputs in plant breeding areas. The diagrams of sequential path analysis (Graph 1) provided a better understanding of the interrelationships among various traits and their relative contribution to spinach dry yield. Based on this diagram, when the second-order traits were used as predictors and first-order variables as response variables indicated that LL and SA positively influenced DY and accounted for more than 48% of observed variation (Graph 1). Also, LN positively influenced DF and accounted for more than 33% of observed variation while LW positively influenced PL and accounted for more than 60% of observed variation (Table 5). Yield component study indicated, petiole properties (length and diameter) were positively associated with dry yield and had the important direct effect on dry yield. Conventional path analysis and yield component studies provide the usual conflicting results.

For future spinach breeding and selection of the most favorable genotypes, it is important to ascertain the variation available for plant structure and yield components in these landraces. Better understanding of how yield components influence the yield formation process can be obtained by applying path analysis to determine the direct and indirect effects of primary, secondary and tertiary traits on the yield formation process. The main advantage is that path analysis not only identifies the most important factor directly affecting a trait, but also indicates how factors affect the trait indirectly through other factors. Previous investigations have indicated that the path analysis provides more information on the interrelationships between the yield components and yield than do the correlation coefficients (Asghari-Zakaria et al., 2007; Sabaghnia et al., 2010). Path analysis helps determine whether yield component compensation is occurring.

Yield component compensation is when two or more yield components affecting yield or any other yield component act inversely in their effects.

Our study demonstrates the utility of the sequential path analysis over conventional path analysis in discerning direct and indirect effects of different yield-related traits. Various characters consist on PD, DF, PL, LL, SA, LN and LW traits were identified as the first, second and third order variables (Graph 1). The role of petiole (length and diameter) regarding direct and indirect effects on spinach dry yield was important. These findings could be verified regarding the investigations of Asadi & Hasandokht (2007) and Eftekhari et al. (2010), who worked on different sets of Iranian spinach landraces. Also, the leaf properties (LL and LW) were identified as the other influencing factors on dry yield.



Graph 1: Sequential path analysis diagram illustrating the interrelationships among various traits contributing to dry yield of spinach.

In general, based on the findings of correlation and sequential path analyses in this research very close interrelationships were recorded among dry yield with some measured traits. On the other hand, the most important yield components which do not differ in different studies were leaf and petiole properties which could be used in conducted breeding research to increase yield. As a result, it may be recommended that genotypes having good manner of these mentioned traits should be criteria preferred specifically in both cultivation and selection in spinach. It can be concluded the great importance of above traits for selection in breeding programs which is meaning with the improvement of spinach yield under the native condition.

REFERENCES:

ASADI, H. A. & HASANDOKHT, M. R. 2007: An evaluation of genetic diversity of Iranian spinach landraces. – *Iranian Journal of Horticulture Science*, 38: 257–265.

ASGHARI–ZAKARIA, R., FATHI, M. & HASAN–PANAH, D. 2007: Sequential path analysis of yield components in potato. – *Potato Research*, 49: 273–279.

BRANDENBERGER, L. P., WELLS, L. K. & HAIGH, M. M. 2004: Yield and quality characteristics of spring spinach grown in Oklahoma. – *HortTechnology*, 14: 602–605.

BRANDENBERGER, L., CAVINS, T., PAYTON, M., WELLS, L. & JOHNSON, T. 2007: Yield and quality of spinach cultivars for greenhouse production in Oklahoma. – *HortTechnology*, 17: 269–272.

EFTEKHARI, S. A. A., HASANDOKHT, M. R., FATAHI-MOGHADAM, M. R. & KASHI, A. A. K. 2010: Genetic diversity of some Iranian spinach (*Spinacia oleracea* L.) landraces using morphological traits. – *Iranian Journal of Horticulture Science*, 41: 83–93.

FEYZIAN, E., DEHGHANI, H., REZAI, A. M. & JALALI, M. 2009: Correlation and sequential path model for some yield-related traits in melon (*Cucumis melo* L.). – *Journal of Agriculture Science and Technology*, 11: 341–353.

GOLDMAN I.L., 2003. Recognition of fruit and vegetables as healthful: Vitamins and phytonutrients. – *HortTechnology*, 13: 252–258.

IBRAHIM, E. A. & RAMADAN A. Y. 2013: Correlation and path coefficient analyses in sweet melon (*Cucumis melo* Var. *Aegyptiacus* L.) under irrigated and drought conditions. – *Pakistan Journal of Biological Sciences*, 16: 610–616.

MINITAB INC. 2010: MINITAB 16. MINITAB User's guide. - *Minitab Inc*, Harrisburg, Pennsylvania. USA, 345 p.

MOHAMMADI S. A., PRASANNA B. M., & SINGH N. N. 2003: Sequential path model for determining interrelationships among grain yield and related characters in maize. – *Crop Science*, 43: 1690–1697.

MURPHY, J. B. & MORELOCK, T. E. 2000: Spinach breeding program yields lines containing high levels of carotenoid antioxidants. – *Arkansas Agricultural Experiment Station Research Series*, 475: 36–39.

SABAGHNIA, N., DEHGHANI, H., ALIZADEH, B. & MOHGHADDAM, M. 2010: Interrelationships between seed yield and 20 related traits of 49 canola (*Brassica napus* L.) genotypes in non-stressed and waterstressed environments. – *Spanish Journal of Agriculture Research*, 8: 356–370.

SALK A., ARIN L., DEVECI M. & POLAT S. 2008: Special Vegetable Production. - *Onur Press*, Tekirdag, 488 p.

SPRINGR-VERLAG INC 2007: User's guide S-PLUS® 8 for Windows. - *Springr-Verlag Inc*, Berlin, Germany.

SPSS INC., 2008: SPSS 16 for Windows, SPSS User's guide. - *SPSS Inc*, Chicago, IL. USA. (http://www.spss.com).

Receive: 2 August 2013.