Multi-Science Journal *2015; 1(3):9-12*

COMUNICAÇÃO BREVE

Phenotypic correlation network analysis of garlic variables

Anderson Rodrigo da Silva[1], Paulo R. Cecon[2], Mário Puiatti[2]

ABSTRACT

In this paper we applied weighted correlation networks in order to discover correlation structures and link patterns of sixteen garlic variables related to leaf, bulb and other vegetative and growth variables. By using the Fruchterman-Reingold algorithm, correlation clusters and other structures could be easily identified. Overall, we detected a link between clusters of leaf and bulb variables. The harvest index was negatively associated with vegetative variables, as expected. In addition, bulb growth rate was positively associated with leaf area rate, root growth rate and plant liquid assimilation rate.

Keywords: *Allium sativum* L., correlation matrix, covariance, plant breeding

Análise de redes ponderadas de correlação de variáveis de alho

ABSTRACT

A análise de redes ponderadas de correlação foi aplicada com o objetivo de descobrir estruturas de correlação e padrões de ligação entre dezesseis variáveis de alho relacionadas à folha, bulbo e outras variáveis vegetativas. Por meio do algoritmo Fruchterman-Reingold, grupos de correlação e outras estruturas puderam ser facilmente identificados. Um padrão geral de ligação entre grupos de variáveis de folha e de bulbo foi detectado. O índice de colheita foi negativamente correlacionado com variáveis vegetativas, como esperado. Além disso, a taxa de crescimento do bulbo foi positivamente associada com o índice de área foliar, taxa de crescimento radicular e taxa de assimilação líquida da planta.

Keywords: *Allium sativum* L., matriz de correlação, covariância, melhoramento de plantas.

Autor para correspondência: Anderson Rodrigo da Silva
Rodovia Geraldo Silva Nascimento, km 2,5, s/n, Zona Rural, Urutaí, GO, Brasil.
E-mail: anderson.silva@ifgoiano.edu.br

Recebido em: 23 mai. 2015
Aceito em: 15 jun. 2015
Editor responsável: Prof. Dr. Guilherme Malafaia

[1] Instuto Federal Goiano – Câmpus Urutaí, GO, Brasil.

[2] Universidade Federal de Viçosa (UFV), MG, Brasil.

INTRODUCTION

Clonal selection is a major breeding method for garlic since plant sterility usually precludes crop improvement through cross hybridization. Nevertheless, garlic shows wide morphological and agronomic variations in colour, size of bulb, plant height, flowering, number and size of cloves, days to harvesting, resistance to storage capacity, dormancy and adaptation to agro-climatic situations (Mario et al., 2008, Singh et al., 2013). According to Honorato (2012), there is a large diversity of garlic cultivars in Brazil, and this is a result of selection and somatic mutations. Reis et al. (2014) mention the necessity of exploring the genetic variability in germoplasm banks. When evaluating accessions of these banks, several characters should be considered.

According to Olawuyi et al. (2013), the correlation coefficient measures the mutual relationship among several characters and determines the component on which selection can be based upon. Barad et al. (2012) highlighted the value of studying character association patterns for effective selection. Singh et al. (2013) pointed out the importance of considering several components affecting garlic yield, directly or indirectly. However, one may find out some difficulties when analysing large correlation matrices.

According to Epskamp et al. (2012), a correlation network analysis can help one to identify suggestive correlation patterns in a correlation matrix that are hard to extract by analysing it numerically. A network graph is created to represent variables as nodes. These nodes are connected by edges, whose width is proportional to the strength of the correlation. Correlation network plots are biologically intuitive (Langfelder & Horvath, 2008) and useful to identify clusters of variables and how these clusters are connected to each other (Ursem et al., 2008).

Correlation networks of garlic variables may increase the selection effectiveness of of garlic cultivars, since not only correlated variables but clusters of correlated variables can be identified. And, although there are studies concerning correlation analysis of garlic variables (Singh et al., 2011, Barad et al., 2012, Singh et al., 2013), we could not find any published work on correlation network with this important crop. In this context, the objective of this work was to identify patterns or structures in the matrix of phenotypic correlations among agronomic and growth-related garlic variables based upon weighted correlation network analysis.

MATERIAL AND METHODS

Our results are based on data from an experiment installed under the randomized block design, with four replications, at the municipality of Viçosa, MG, Brazil, coordinates 20°45′ S and 42°51′ W. We evaluated 89 garlic accessions registered in the Germoplasm Bank of the Federal University of Viçosa (BGH/UFV). Experimental units consisted of four transversal 1.0 m long lines, with plants on spacing 0.25 × 0.10 cm, which results in 40 plants per experimental unit.

At 60, 90, 120, 150 days after planting and at harvest (~ 170 days), we evaluated the following variables related to leaf and leaf growth: leaf dry matter (LDM), leaf area (LA), number of leaves per plant (NLP), leaf area ratio (LAR) and leaf area endurance (LAE); and the following variables related to bulb and bulb growth: bulb dry matter (BDM), bulb diameter (BD), bulb length (BL), number of cloves per bulb (NCB), bulb weight (BW), bulb growth rate (BGR); and some other vegetative and growth related variables: pseudo-stem dry matter (PDM), root dry matter (RDM), root growth rate (RGR), plant liquid assimilation rate (PAR) and harvest index (HI). Growth related variables were determined as suggested by Evans (1982).

Data were subjected to multivariate analysis of variance. For estimating the phenotypic correlation matrix, RP, we took the matrix of mean squares and cross-products of genotype, MG, as the estimate of phenotypic covariance matrix. Then, considering m_{ij} as the element (covariance) at the i-th (i = 1, 2, ..., 16) row and j-th (j = 1, 2, ..., 16) column, the corresponding correlation is given by: $r_{ij} = m_{ij} / \sqrt{m_{ii} m_{jj}}$. This procedure is equivalent to computing the correlation matrix from the accession means. We then analysed the correlation matrix by creating weighted correlation networks, at which connections among variables are determined by the adjacency matrix A = h(RP), with function $h(r_{ij}) = \frac{1}{2}\{\mathrm{sgn}(|r_{ij}| - \rho) + 1\}$, where ρ is the hard threshold parameter, which was set to be null so that all the connections among variables can be seen. However, we have controlled the thickness of edges by applying a cut-off value equal to 0.5, meaning that only $|r_{ij}| \geq 0.5$ have their edges highlighted. The Fruchterman-Reingold algorithm (Fruchterman & Reingold, 1991) was used to create a force-directed layout for the network, in which the proximity between nodes (variables) is proportional to the absolute value of the correlation between those nodes. Finally, positive correlations are coloured in dark green, whereas negatives are shown in red. All the analyses were performed using the software R version 3.1.2 (R Core Team, 2014).

The correlation network procedure was done using the package qgraph (Epskamp et al., 2012).

RESULTS AND DISCUSSION

The multivariate analysis of variance has revealed significant ($p < 0.01$) effect of accessions, meaning that at least one covariance (and correlation) component is non-null.

Figure 1 shows the network that was constructed from all pairwise phenotypic correlations between variables of leaf, bulb and other related to vegetative growth. We set three different colours to indicate the type of variable or part of the plant. It is noteworthy the central clustering involving all the five leaf variables and four bulb variables related to bulb yield: BW, BL, BD and NCB, all positively correlated. The association between these two clusters seems to be quite reasonable since larger leaf development usually indicates higher photosynthetic rate and, hence, bulb yield. Singh et al. (2013) found positive significant ($p < 0.01$) phenotypic correlations between NLP and BW, and strong (>0.9) correlation between BD and BW. Significant ($p < 0.05$) phenotypic correlations between NLP and BD and NCB were found by Barad et al. (2012).

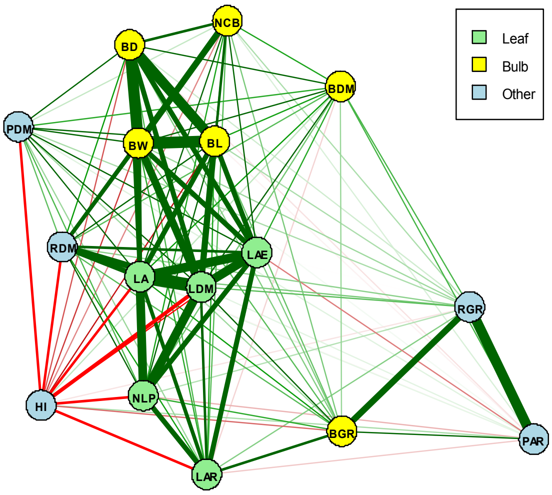

Figure 1. Phenotypic correlation network of garlic variables. Red and green lines represent negative and positive correlations, respectively. Line width proportional to the strength of the correlation.

Another correlation structure involves HI being negatively correlated with PDM, RDM, BD, BW and all leaf variables. This result is in agreement with those found by Kassahun et al. (2010) when evaluated Ethiopian garlic accessions. In fact, because HI is basically the ratio between BDM and the plant total dry matter, those associations are expected.

PAR and RGR presented strong (> 0.9) correlation. In addition, bulb growth rate (BGR) is associated with the following growth variables: LAR, RGR and PAR.

Because bulb and leaf variables form a correlated group, it means that these accessions present aptitude for selection. In this paper, we used network plots to simultaneously visualize all correlations between garlic variables. By using the Fruchterman-Reingold algorithm, correlation structures and link patterns could be easily identified. Overall, we detected a link between clusters of leaf and bulb variables. The relationships we observed through our correlation network may also serve as general guidelines for breeders.

ACKNOWLEDGEMENTS

Authors would like to thank the Coordination for the Improvement of Higher Level Personnel (CAPES) and the National Council for Scientific and Technological Development (CNPq) for granting the scholarships.

REFERENCES

Epskamp, A., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D. & Borsboom, D. (2012). qgraph: Network Visualizations of Relationships in Psychometric Data. *Journal of Statistical Software*, *48*(4), 1-18.

Evans, G. C. (1982). *The quantitative analysis of plant growth*. London: Blackwell Scientific Publications.

Fruchterman, T. & Reingold, E. (1991). Graph drawing by force-directed placement. *Software – Practice & Experience*, *21*(11), 1129-1164.

Honorato, A. R. F. (2012). *Avaliação de cultivares de alho na região de Mossoró-RN*. (Dissertação de mestrado). Universidade Federal Rural do Semi-Árido, Brasil.

Kassahun, T., Akhilesh, T. & Kebede, W. (2010). Genetic variability, correlation and path coefficient among bulb yield and yield traits in Ethiopian garlic germoplasm. *Indian Journal of Horticulture*, *64*(4), 489-499.

Langfelder, P. & Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, *9*(559), 1-13.

Mario, P. C., Viviana, B. V. & Marya, I. A. (2008). Low genetic diversity among garlic accessions detected using RAPD. *Chilean Journal of Agricultural Research, 68*(1), 3-12.

Olawuyi, O. et al. (2014). Accession × Treatment Interaction, Variability and Correlation Studies of Pepper (*Capsicum* spp.) under the Influence of Arbuscular Mycorrhiza Fungus (*Glomus clarum*) and Cow Dung. *American Journal of Plant Sciences*, *5*(5), 683-690.

Puiatti, G. A. et al. (2013). Cluster analysis applied to nonlinear regression models selection for the description of dry matter accumulation of garlic plants. *Revista Brasileira de Biometria*, *31*(3), 337-351.

R Core Team (2014). *R*: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. Available from: <http://www.R-project.org/>. Accessed on 01/11/2014.

Reis, R. M. et al. (2014). Nonlinear regression models applied to clusters of garlic accessions. *Horticultura Brasileira*, *32*(2), 178-183.

Singh, R. K., Dubey, B. K., Bhonde, S. R. & Gupta, R. P. (2011). Correlation and path coefficient studies in garlic (*Allium sativum* L.). *Journal of Spices and Aromatic Crops*, *20*(2), 81–85.

Singh, S. R. et al. (2013). Character association and path analysis in garlic (*Allium sativum* L) for yield and its attributes. *SAARC Journal of Agriculture, 11*(1): 45-52.

Ursem, R., Tikunov, Y., Bovy, A., Berloo, R. & Eeuwijk, F. (2008). A correlation network approach to metabolic data analysis for tomato fruits. *Euphytica*, *161*(1-2), 181–193.