

## **ESTUDOS SOBRE DIVERGÊNCIA GENÉTICA. III. COMPARAÇÃO DE TÉCNICAS MULTIVARIADAS <sup>1/</sup>**

Cosme Damião Cruz <sup>2/</sup>  
Roland Vencovsky <sup>3/</sup>  
Samuel Pereira de Carvalho <sup>4/</sup>

### **1. INTRODUÇÃO**

A avaliação da divergência com base em evidências científicas é de grande importância em programas de melhoramento, principalmente na escolha de progenitores para hibridações. Também, é de grande utilidade em estudos evolutivos por prover informações sobre os recursos genéticos disponíveis e auxiliar na localização e no intercâmbio dos membros.

No estudo da divergência genética, vários métodos multivariados podem ser aplicados. Dentre eles, citam-se a análise por componentes principais e por variáveis canônicas e os métodos aglomerativos. A escolha do método mais adequado tem sido determinada pela precisão desejada pelo pesquisador, pela facilidade da análise e pela forma como os dados foram obtidos.

Os métodos aglomerativos diferem dos demais, em razão de dependerem fundamentalmente de medidas de similaridade estimadas previamente, como as distâncias eu-

---

<sup>1/</sup> Parte da tese apresentada pelo primeiro autor à ESALQ/USP, como uma das exigências para a obtenção do título de Doutor em Agronomia.

Aceito para publicação em 17.09.1993.

<sup>2/</sup> Departamento de Biologia Geral, Universidade Federal de Viçosa, 36570-000 Viçosa, MG. Enviar correspondência para C.D.C.

<sup>3/</sup> Departamento de Genética, ESALQ/USP, 13400-000 Piracicaba, SP.

<sup>4/</sup> Aluno do curso de pós-graduação em Genética e Melhoramento, em nível de doutorado. Universidade Federal de Viçosa, Viçosa-MG.

DF, em relação à análise por componentes principais, foram os que menos contribuíram para a divergência dos progenitores.

Essas técnicas multivariadas são sensíveis em discriminar as variáveis menos importantes que se caracterizam pela redundância, por já estarem correlacionadas com outras variáveis presentes na análise ou pela sua invariância. Assim, na análise por variáveis canônicas, os caracteres PROD e FAE possivelmente tiveram menor importância para a diversidade entre os cultivares, por terem apresentado o menor quociente entre os quadrados médios de tratamentos e de resíduos. Os caracteres AE e MS foram também considerados de menor importância, possivelmente em razão do aspecto de redundância, uma vez que AE é indiretamente representado por AP (correlações fenotípicas e genotípicas iguais a 0,9755 e 0,9832, respectivamente) e MS por AF (correlações fenotípicas e genotípicas iguais a 0,9497 e 0,9820, respectivamente).

Considerando as correlações entre os caracteres e as coincidências observadas, verifica-se que as técnicas multivariadas dos componentes principais e das variáveis canônicas tiveram concordância satisfatória no contexto da identificação de caracteres mais importantes para a determinação da diversidade genética entre os cultivares.

Conclui-se então que há grande concordância na quantificação da divergência genética tanto por meio da distância generalizada de Mahalanobis quanto pelas distâncias euclidianas médias, obtidas a partir das variáveis padronizadas, de escores dos dois primeiros componentes principais e de escores das duas primeiras variáveis canônicas. As correlações simples entre estas quatro medidas da divergência foram superiores a 0,93.

As altas correlações entre as medidas de dissimilaridade revelam considerável robustez da distância euclidiana e dos componentes principais, uma vez que, no presente estudo, a matriz de correlações residuais foi estatisticamente diferente da matriz identidade.

Os caracteres que menos contribuíram para a divergência entre os cultivares, identificados pelos coeficientes (autovetores) das últimas variáveis canônicas, foram altura de espiga, produção de grãos por planta, produção de matéria seca e porcentagem de folhas acima da espiga. Considerando a redundância de certas variáveis, concluiu-se que as técnicas dos componentes principais e das variáveis canônicas foram também concordantes na discriminação da importância de caracteres para a determinação da divergência genética do material avaliado.

#### 4. RESUMO

Foi verificado o grau de concordância das estimativas da divergência genética entre cultivares de milho, obtidas por meio de quatro técnicas multivariadas: Distância Generalizada de Mahalanobis ( $D^2$ ); Distância Euclidiana Média, obtida a partir dos dados padronizados; Distância Euclidiana Média, obtida a partir dos escores das duas primeiras variáveis canônicas; e Distância Euclidiana Média, obtida a partir dos dois primeiros componentes principais, que representavam mais de 94% da variabilidade total disponível.

Foi demonstrada a existência de alto grau de concordância entre as medidas de dissimilaridade utilizadas, por meio de coeficientes de correlações, cujas estimativas foram superiores a 0,93. Foi também enfatizada a razoável concordância na identificação dos caracteres que menos contribuíram para a diversidade genética. Constatou-se a robustez das técnicas euclidianas e de componentes principais, a despeito da existência de correlação residual entre caracteres.

## 5. SUMMARY

### (STUDIES ON GENETIC DIVERGENCE. III. COMPARISON OF MULTIVARIATE TECHNIQUES)

The degree of concordance of four multivariate techniques applied to estimate the genetic divergence among corn cultivars was evaluated. The Generalized Distance of Mahalanobis ( $D^2$ ), the average Euclidean distance based on standardized variables and on scores of the two first canonical variables and principal components were considered.

The first two scores explained more than 94% of the total variance present. Estimates of correlation higher than 0.93 between measures of divergence indicated a high degree of concordance between applied techniques. A reasonable correspondence was found in the identification of traits with smallest contribution to diversity. Euclidean distances and principal components showed robustness despite the existence of residual correlations between traits.

## 6. LITERATURA CITADA

1. DAS, P.K. & GUPTA, T.D. Multivariate analysis in black gram (*Vigna mungo*, L. Hepper). *The Indian Journal of Genetics & Plant Breeding*, 44(2): 243-247, 1984.
2. GODOI, C.R. de M. *Análise estatística multidimensional*. Piracicaba, ESALQ/USP, 1985. 187 p.
3. HUSSAINI, S.H.; GOODMAN, M.M. & TIMOTHY, D.H. Multivariate analysis and the geographical distribution of the collection of finger millet. *Crop Sci.*, 17(2):257-263, 1977.
4. JAIN, K.C.; PANDYA, B.P. & PANDE, K. Genetic divergence in chickpea. *The Indian Journal of Genetics & Plant breeding*, 41(2):220-225, 1981.
5. MARDIA, K.V.; KENT, J.T. & BIBBY, J.M. *Multivariate analysis*. London, Academic Press, 1979. 521 p.
6. PETER, K.V. & RAI, B. Genetic divergence in tomato. *The Indian Journal of genetics & Plant Breeding*, 36(3):379-383, 1976.
7. TATSUOKA, M.M. *Multivariate analysis*. New York, Wiley, 1971. 451 p.
8. VARMA, N.S. & GULATI, S.C. Genetic divergence in 2-rowed and 6-rowed barley. *The Indian Journal of Genetics & Plant Breeding*, 42(3):314-318, 1982.