

ANÁLISE DE AGRUPAMENTO UTILIZANDO VARIÁVEIS QUANTITATIVAS E QUALITATIVAS PARA O ESTUDO DA DIVERSIDADE GENÉTICA EM ACESSOS DE ABACAXIZEIRO

Carlos Alberto da Silva Ledo¹, Leandro Simões Azeredo Gonçalves², Everton Hilo Souza³, Fernanda Vidigal Duarte Souza⁴

Resumo

A análise conjunta de variáveis qualitativas e quantitativas tem sido apontada como uma ferramenta útil na estimativa da diversidade genética entre os acessos de uma coleção de germoplasma. O objetivo deste trabalho foi promover a análise simultânea de variáveis qualitativas e quantitativas e posterior agrupamento para definição da diversidade genética entre acessos de abacaxizeiro por meio do algoritmo de Gower. Foram avaliadas 11 características quantitativas e 12 qualitativas em 18 acessos do banco ativo de germoplasma de abacaxizeiro da Embrapa Mandioca e Fruticultura Tropical. A análise simultânea de variáveis qualitativas e quantitativas foi eficiente em expressar o grau de diversidade genética entre os acessos de abacaxizeiro quando comparadas com a análise individual de cada variável.

Introdução

A Embrapa Mandioca e Fruticultura Tropical, situada no município de Cruz das Almas-BA, possui um Banco Ativo de Germoplasma (BAG) de abacaxizeiro que reúne um total de 628 acessos no campo, do gênero *Ananas* e outras bromeliáceas, sendo uma das maiores coleções de germoplasma de abacaxi do mundo (CABRAL, 2000). Estima-se que a maior parte da variabilidade genética intra e interespecífica do abacaxi esteja aí representada. A variabilidade existente no BAG agrega uma extensa fonte de cores, formas, tamanhos de frutos e coroas, assim como diferentes arquiteturas de plantas a serem exploradas. Portanto, faz-se necessário quantificar essa variabilidade pela avaliação de variáveis qualitativas (morfológicas e moleculares) e variáveis quantitativas (agronômicas).

As análises de agrupamento dessas variáveis são feitas individualmente, visto que as distâncias genéticas são calculadas em função do tipo de variável utilizada. Cruz (2008) apresenta os procedimentos para estimar medidas de dissimilaridade com base em variáveis quantitativas (distâncias euclidianas ou de Mahalanobis), binárias (índice de Jaccard, Nei e Li etc.) e multicategóricas (Cole-Rodgers). Nesse sentido, observam-se várias discrepâncias em relação aos agrupamentos e às inferências em relação à quantificação da variabilidade entre acessos de um banco de germoplasma. Uma técnica que permite a análise simultânea de dados quantitativos e qualitativos foi proposta por Gower (1971). Este método permite que valores da matriz de distância fiquem compreendidos entre 0 e 1, sendo necessário a padronização das variáveis quantitativas e qualitativas. Este trabalho teve por objetivo avaliar a variabilidade genética entre acessos de abacaxizeiro por meio da análise simultânea de variáveis qualitativas e quantitativas em acessos de abacaxizeiro.

Material e Métodos

Foram caracterizados 18 acessos do banco ativo de germoplasma de abacaxizeiro da Embrapa Mandioca e Fruticultura Tropical em que se avaliaram 23 descritores morfológicos (IBPGR, 1991), sendo onze descritores quantitativos: altura da planta; diâmetro da copa; comprimento e largura da folha; comprimento e diâmetro da haste; comprimento e diâmetro da inflorescência; comprimento e diâmetro do fruto após fechamento da última flor; comprimento da coroa após fechamento da última

¹ Primeiro autor é pesquisador da Embrapa Mandioca e Fruticultura Tropical, CP 07, Cruz das Almas, BA, CEP 44.380-000. E-mail: ledoc@cpmfmf.embrapa.br

² Segundo autor é doutorando do Laboratório de Melhoramento Genético Vegetal, Centro de Ciências e Tecnologias Agropecuárias, Universidade Estadual do Norte Fluminense Darcy Ribeiro, Campos dos Goytacazes, RJ, CEP 28.013-600. E-mail: lsagrural@yahoo.com.br

³ Terceiro autor é mestrando do curso de Pós-Graduação em Ciências Agrárias, Universidade Federal do Recôncavo da Bahia, Cruz das Almas, BA, CEP 44.380-000. E-mail: hilosouza@gmail.com

⁴ Quarta autora é pesquisadora da Embrapa Mandioca e Fruticultura Tropical, CP 07, Cruz das Almas, BA, CEP 44.380-000. E-mail: fernanda@cpmfmf.embrapa.br

flor, e 14 descritores qualitativos: hábito da planta; presença de estolão; variação das folhas; cor das folhas; presença de antocianina nas folhas; presença de espinho; cor dos espinhos; forma da haste; forma do fruto; cor do fruto após fechamento da última flor; cor do fruto aos 30 dias após fechamento da última flor; cor da bráctea. Os seguintes acessos que apresentaram potencial ornamental foram avaliados: três acessos de *Ananas comosus* var. *bracteatus* (BGA-20, BGA-126, BGA-128), três de *Ananas comosus* var. *erectifolius* (BGA-739, BGA-750, BGA-804), três de *Ananas comosus* var. *ananassoides* (BGA-208, BGA-464, BGA-475), três de *Ananas comosus* var. *comosus* (BGA-16, BGA-380, BGA-511), três de *Ananas* sp. (BAG-324, BAG-325, BAG-388) e três de *Ananas macrodontes* (BGA-81, BGA-83, BGA-719).

Uma análise individual para as variáveis quantitativas e qualitativas foi realizada. Para as variáveis quantitativas calculou-se a distância euclidiana média e para as variáveis qualitativas utilizou-se a distância de Cole-Rodgers (COLE-RODGERS *et al.*, 1997). Para a análise simultânea das variáveis quantitativas e qualitativas utilizou-se o algoritmo de Gower (GOWER, 1971). Os agrupamentos hierárquicos das análises individuais e simultâneas a partir das matrizes de distância genética foram obtidos pelos métodos de UPGMA - *Unweighted Pair Group Method with Arithmetic Mean* (SNEATH; SOKAL 1973). A validação dos agrupamentos foi determinada pelo coeficiente de correlação cofenético (SOKAL; ROHLF, 1962). A significância dos coeficientes de correlação cofenético foi calculada pelo teste de Mantel com 1000 permutações (MANTEL, 1967).

Para a obtenção das matrizes de distância genética das análises individuais e cálculo dos coeficientes de correlação cofenético foi utilizado o programa Genes (Cruz, 2008). A matriz de distância genética utilizando o algoritmo de Gower foi obtido pelo programa SAS (SAS INSTITUTE, 2006). O dendrograma foi obtido pelo programa Statistica 7.1 (STATSOFT, 2005).

Resultados e Discussão

A análise simultânea das variáveis quantitativas e qualitativas pelo algoritmo de Gower apresentou o maior coeficiente de correlação cofenético entre a matriz de distância genética e a matriz de agrupamento, com valor de correlação de 0,89**. As análises individuais para as variáveis quantitativas e qualitativas apresentaram valores de 0,82** e 0,84**, respectivamente (Tabela 1). Observa-se que a matriz de distância conjunta apresentou valores de correlações altos e significativos com as matrizes de distância obtidas nas análises individuais, 0,77** e 0,84**, para as variáveis quantitativas e qualitativas, respectivamente (Tabela 2).

Portanto, a análise simultânea pelo algoritmo de Gower foi utilizada para representar graficamente a variabilidade dos acessos de abacaxizeiro, pela sua maior eficiência em quantificar as diferenças entre os acessos. (Figura 1). Observa-se que o ponto de corte, média das distâncias da matriz de agrupamento, definiu a formação de 3 grupos: o grupo I formado pelos acessos de *Ananas comosus* var. *bracteatus*, *Ananas comosus* var. *comosus* e *Ananas macrodontes*; o grupo II formado pelos acessos de *Ananas comosus* var. *ananassoides* e *Ananas* sp.; e o grupo III formado pelos acessos de *Ananas comosus* var. *erectifolius*.

A análise simultânea permite que, além de variáveis agronômicas e morfológicas, sejam utilizadas variáveis provenientes de marcadores moleculares. Segundo Gonçalves *et al.* (2009), a escolha das variáveis e o número de variáveis a serem usadas podem comprometer a eficiência da análise simultânea, principalmente no caso de se utilizar um grande número de variáveis binárias, provenientes de marcadores moleculares, na quantificação da diversidade genética em acessos de bancos de germoplasma.

Conclusões

A análise pelo algoritmo de Gower foi eficiente em expressar o grau de diversidade genética entre os acessos de abacaxizeiro avaliados, demonstrando que a análise simultânea proporciona uma maior eficiência no conhecimento da diversidade genética entre acessos de bancos de germoplasma.

Referências

- CABRAL, J. R. S. *Banco ativo de germoplasma de abacaxi*. Cruz das Almas: EMBRAPA-CNPMPF, 2000. Folder.
- COLE-RODGERS, P.; SMITH, D. W.; BOSLAND, P. W. A novel statistical approach to analyze genetic resource evaluations using *Capsicum* as an example. *Crop Science*, v. 37, p. 1000-1002. 1997.
- CRUZ, C. D. *Programa genes* (versão Windows): aplicativo computacional em genética e estatística. Viçosa: UFV, 2008.
- GONÇALVES, L. S.; RODRIGUES, R.; AMARAL, A. T. JR.; KARASAWA, M; SUDRE, C. P. Heirloom tomato gene bank: assessing genetic divergence based on morphological, agronomic and molecular data using a Ward-modified location model. *Genet. Mol. Res.*, v. 8, p. 364-374. 2009.
- GOWER, J. C. A general coefficient of similarity and some of its properties. *Biometrics*, Arlington, v. 27, n. 4, p. 857-874. 1971.
- IBPGR – International Board for Plant Genetic Resources. *Descriptors for pineapple*. Rome, 1991. 41p.
- MANTEL, N. The detection of disease clustering and generalized regression approach. *Cancer Research*, Birmingham, v.27, n.2, p.209-220, 1967.
- SAS INSTITUTE. SAS Technical Report. *SAS/STAT software: Changes and Enhancement*, Release 9.1. 3, Cary NC: SAS Institute. 2006.
- SNEATH, P. H.; SOKAL, R. R. *Numerical taxonomy*: The principles and practice of numerical classification. San Francisco: W.H. Freeman, 1973. 573p.
- SOKAL, R. R. and ROHLF, F. J. The comparison of dendrograms by objective methods. *Taxon*, v.11 p.33-40. 1962.
- STATSOFT, Inc. *Statistica for Windows (data analysis software system)*, version 7.1. Statsoft, Tulsa, Oklahoma (USA), 2005.

Tabela 1. Coeficiente de correlação cofenético e número de grupos formados em função das matrizes de distâncias genéticas estimada para os dados quantitativos, qualitativos e análise simultânea para 18 acessos de abacaxizeiro.

Matriz de distância	Coeficiente de correlação cofenético	Número de grupos formados
Quantitativo	0,82**	2 ¹
Qualitativo	0,84**	4
Conjunta	0,89**	3

** Significativo a 1% de probabilidade pelo teste de Mantel com 1000 permutações. ¹ Baseado na média da matriz de agrupamento.

Tabela 2. Correlação entre matrizes de distância genética estimada para os dados quantitativos, qualitativos e análise simultânea para 18 acessos de abacaxizeiro.

Matriz de distância	Quantitativo	Qualitativo	Conjunta
Quantitativo	1		
Qualitativo	0,38*	1	
Conjunta	0,77**	0,84**	1

** e * significativo a 1 e 5%, respectivamente, pelo teste de Mantel com 1000 permutações.

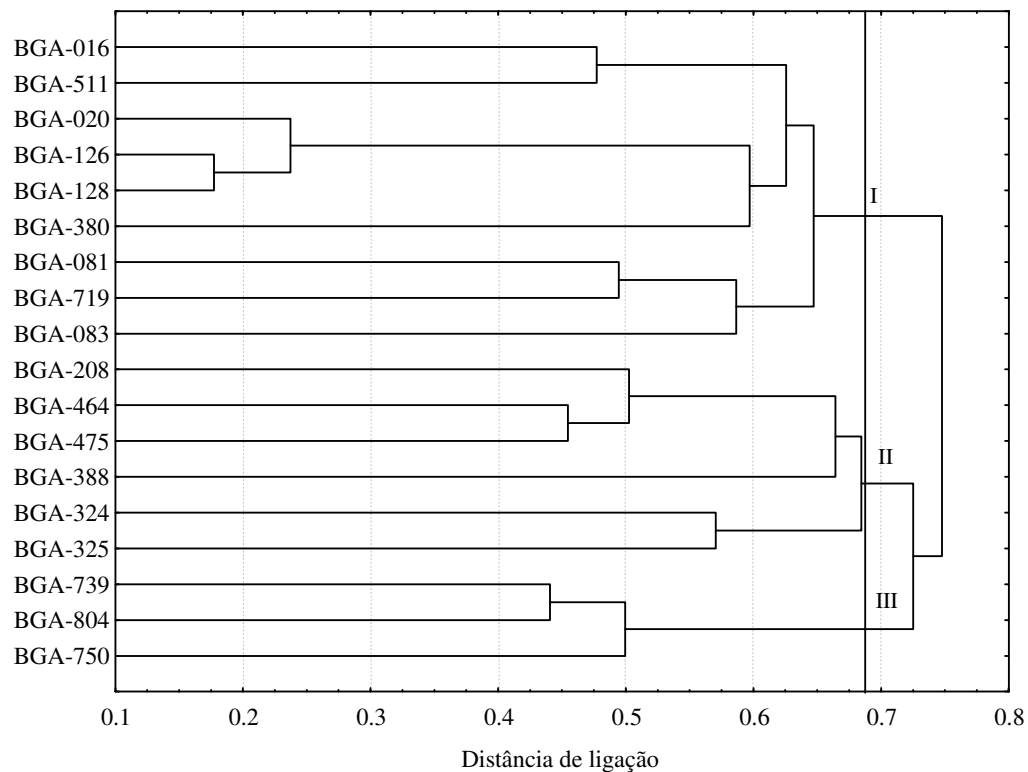


Figura 1. Dendrograma baseado na distância de Gower e método de agrupamento UPGMA na análise de 11 variáveis quantitativas e 12 variáveis qualitativas em 18 acessos de abacaxizeiro.