



Ciência Rural

ISSN: 0103-8478

cienciarural@mail.ufsm.br

Universidade Federal de Santa Maria  
Brasil

Cargnelutti Filho, Alberto; Dalfollo Ribeiro, Nerinéia; Padilha dos Reis, Rodrigo Citton; Rodrigues de Souza, Joseane; Jost, Evandro

Comparação de métodos de agrupamento para o estudo da divergência genética em cultivares de feijão

Ciência Rural, vol. 38, núm. 8, novembro, 2008, pp. 2138-2145

Universidade Federal de Santa Maria  
Santa Maria, Brasil

Disponível em: <http://www.redalyc.org/articulo.oa?id=33113633008>

- Como citar este artigo
- Número completo
- Mais artigos
- Home da revista no Redalyc

redalyc.org

Sistema de Informação Científica

Rede de Revistas Científicas da América Latina, Caribe, Espanha e Portugal

Projeto acadêmico sem fins lucrativos desenvolvido no âmbito da iniciativa Acesso Aberto

## Comparação de métodos de agrupamento para o estudo da divergência genética em cultivares de feijão

Comparison of cluster methods for the study of genetic diversity in common bean cultivars

Alberto Cargnelutti Filho<sup>®</sup> Nerinéia Dalfollo Ribeiro<sup>I</sup> Rodrigo Citton Padilha dos Reis<sup>II</sup>  
Joseane Rodrigues de Souza<sup>III</sup> Evandro Jost<sup>IV</sup>

### RESUMO

Os objetivos deste trabalho foram comparar métodos de agrupamento, com base nas medidas de dissimilaridade (euclidiana média padronizada e generalizada de Mahalanobis) e obter informações sobre a divergência genética em cultivares de feijão (*Phaseolus vulgaris* L.). Quatorze cultivares de feijão foram avaliadas em nove experimentos conduzidos em Santa Maria, Estado do Rio Grande do Sul (latitude 29°42'S, longitude 53°49'W e 95m de altitude), nos anos agrícolas de 2000/2001 a 2004/2005. Foi utilizado o delineamento aleatorizado em blocos, com três repetições, e foram avaliados os caracteres produtividade de grãos, número de vagens por planta e de sementes por vagem, massa de cem grãos, população final de plantas, número de dias da emergência ao florescimento e da emergência à colheita, altura de inserção de primeira e de última vagem e grau de acamamento. Agrupamentos com base na distância euclidiana média padronizada são distintos dos formados com base na distância generalizada de Mahalanobis. O método de Tocher e os métodos hierárquicos da ligação simples, de Ward, da ligação completa, da mediana, da ligação média dentro de grupo e da ligação média entre grupo, com base na distância generalizada de Mahalanobis formam grupos concordantes. A cultivar "Iraí" apresenta comportamento distinto das demais cultivares.

**Palavras-chave:** *Phaseolus vulgaris* L., medidas de dissimilaridade, algoritmos de agrupamento, coeficiente de correlação cofenética.

### ABSTRACT

The aim of this research was to compare cluster methods, on the basis of the dissimilarity (standardized average euclidian and Mahalanobis generalized) and obtain information on genetic diversity in common bean cultivars (*Phaseolus vulgaris*). Fourteen common beans cultivars were evaluated in nine experiments conducted in Santa Maria, Rio Grande do Sul State, Brazil (latitude 29°42' S, longitude 53°49' W, altitude 95m), in agricultural years from 2000/2001 to 2004/2005. Randomized blocks design with three repetitions was installed to evaluate the following characters: grain yield, number of pods per plant, number of seeds per pod, weight of 100 grains, final population of plants, number of days of the emergency to flowering, number of days of the emergency to harvest, height of first pod insertion and height of the final pod insertion. Clusters based on the standardized average euclidian distance are distinct from those formed on the basis of Mahalanobis generalized distance. The Tocher's method and hierarchical methods of the single linkage, Ward, complete linkage, median, the average linkage within the group and average linkage between groups, based on the Mahalanobis generalized distance form agreement cluster. The cultivar 'Iraí' presents distinct behavior from other cultivars.

**Key words:** *Phaseolus vulgaris* L., dissimilarity measures, cluster algorithms, cofenetic correlation coefficient.

<sup>I</sup>Departamento de Fitotecnia, Centro de Ciências Rurais (CCR), Universidade Federal de Santa Maria (UFSM), 97105-900, Santa Maria, RS, Brasil. E-mail: cargnelutti@pq.cnpq.br. \*Autor para correspondência.

<sup>II</sup>Curso de Estatística, UFRGS, Porto Alegre, RS, Brasil. E-mail: rodrigocpdosreis@gmail.com.

<sup>III</sup>Programa de Pós-graduação em Agronomia (Genética e Melhoramento de Plantas), Universidade Estadual Paulista (UNESP), Faculdade de Ciências Agrárias e Veterinárias (FCAV), Jaboticabal, SP, Brasil.

<sup>IV</sup>Programa de Pós-graduação em Agronomia, UFSM, Santa Maria, RS, Brasil.

## INTRODUÇÃO

A inferência sobre a divergência genética em um grupo de genitores pode ser de natureza quantitativa e preditiva. Na primeira, citam-se as análises dialélicas, que possibilitam determinar a capacidade geral e específica de combinação e a heterose manifestada nos híbridos. No entanto, a necessidade de avaliações de  $p$  genitores e de todas as suas combinações híbridas  $p(p-1)/2$ , aliada ao fato de que, em algumas culturas, a polinização manual é onerosa, difícil de ser executada e com pouca probabilidade de êxito na obtenção de semente híbrida, pode inviabilizar o estudo, principalmente quando o valor de  $p$  é elevado (CRUZ & REGAZZI, 1997; CRUZ & CARNEIRO, 2003). Nesse contexto, por dispensarem a obtenção prévia das combinações híbridas, destacam-se os métodos preditivos da divergência entre genitores, que se baseiam em diferenças morfológicas, fisiológicas e moleculares. Entre os métodos preditivos, a análise de agrupamento destaca-se, pois apresenta a finalidade de reunir, por algum critério de classificação, os genitores em grupos, de tal forma que exista homogeneidade dentro do grupo e heterogeneidade entre os grupos, sendo adequada para identificar os genótipos divergentes e com maior probabilidade de sucesso nos cruzamentos.

As distâncias euclidiana média padronizada e generalizada de Mahalanobis entre os pares de genótipos são amplamente utilizadas como medida de dissimilaridade nos métodos de agrupamento. A segunda oferece a vantagem de levar em consideração a existência de correlações entre os caracteres analisados por meio da matriz de variâncias e covariâncias residuais, porém, necessita de experimentos com repetições (CRUZ & REGAZZI, 1997; CRUZ & CARNEIRO, 2003).

A medida de dissimilaridade e o método utilizado devem garantir ao melhorista segurança na seleção de genitores para os cruzamentos. Em caso de concordância de agrupamentos, a escolha do método deve recair naquele de simples execução e de fácil interpretação. No entanto, caso haja discordância entre os métodos, a escolha de genitores passa a depender do método utilizado, havendo a necessidade de escolha do método mais eficiente.

Coefficiente de correlação linear de 0,97 entre as estimativas de distância euclidiana média e generalizada de Mahalanobis foi encontrado em estudo sobre divergência genética em milho (CRUZ, 1990). Por outro lado, MALUF & FERREIRA (1983), comparando a distância euclidiana e a distância generalizada de Mahalanobis na avaliação de 10 cultivares de feijão-

vagem (*Phaseolus vulgaris* L.) observaram baixa correlação entre as metodologias ( $r = 0,273$ ). BENIN et al. (2003) também compararam a distância euclidiana e de Mahalanobis e encontraram média correlação ( $r=0,529$ ) no estudo de 12 cultivares de aveia e, quando utilizadas na confecção de dendrogramas, não possibilitaram a formação de agrupamentos similares. Tal concordância ou discordância é dependente da magnitude das correlações residuais que possam existir entre os caracteres considerados (CRUZ, 1990).

Em relação à comparação de métodos, os agrupamentos por meio dos métodos de ligação simples, ligação completa e ligação média entre grupos, com base na distância euclidiana, em acessos das espécies *Paspalum guenoarum* e *Paspalum plicatum* foram estudados por TOTTI et al. (2001). Esses autores concluíram que o método da ligação média entre grupos foi o mais eficiente. Já KARASAWA et al. (2005) concluíram que o método de Tocher formou grupos, concordantes com os subgrupos obtidos pelo método da ligação simples, com base na distância generalizada de Mahalanobis, em 70 acessos de tomateiro. Concordância entre métodos de agrupamento da ligação simples e de Tocher, com base na distância generalizada de Mahalanobis, foi verificada em pimentas e pimentões por SUDRÉ et al. (2005).

Os métodos da ligação simples, da ligação completa, do centróide, da mediana, das médias das distâncias e de Ward, estudados por ALBUQUERQUE et al. (2006), em espécies florestais, revelaram coincidências nos agrupamentos, exceto os métodos do centróide e de Ward e os métodos do centróide e da mediana, em comparação com o de Ward, respectivamente, com base nas matrizes de Mahalanobis a partir dos dados originais e de "bootstrap". BERTAN et al. (2006) concluíram que os métodos da ligação média entre grupo e de Tocher produziram agrupamentos similares, utilizando a distância generalizada de Mahalanobis para o estudo de 19 genótipos de trigo. Os métodos de Tocher e ligação média entre grupos foram parcialmente concordantes no agrupamento de 63 cultivares crioulas de feijão (BONETT et al., 2006).

Estudos dessa natureza na cultura de feijão são escassos. Além disso, diversos métodos com base em diferentes medidas de dissimilaridade podem levar a distintos padrões de agrupamento (CRUZ & REGAZZI, 1997; MINGOTI, 2005). Assim, os objetivos deste trabalho foram comparar métodos de agrupamento, com base nas medidas de dissimilaridade (euclidiana média padronizada e generalizada de Mahalanobis), e obter informações sobre a divergência genética em cultivares de feijão (*Phaseolus vulgaris* L.).

## MATERIAL E MÉTODOS

Quatorze cultivares de feijão (“Carioca”, “Diamante Negro”, “TPS Nobre”, “Guapo Brilhante”, “Guatelian 6662”, “Iapar 44”, “Iraí”, “Macanudo”, “Macotaço”, “Minuano”, “Pérola”, “Rio Tibagi”, “TPS Bionobre” e “TPS Bonito”) foram avaliadas em nove experimentos conduzidos em área do Departamento de Fitotecnia, da Universidade Federal de Santa Maria, Santa Maria, Rio Grande do Sul (latitude 29°42S, longitude 53°49W e 95m de altitude), nos anos agrícolas 2000/2001, 2001/2002, 2002/2003, 2003/2004 e 2004/2005, em duas épocas de cultivo: safra - semeadura em setembro ou outubro e safrinha - semeadura em janeiro ou fevereiro.

Os nove experimentos foram conduzidos no delineamento aleatorizado em blocos, com três repetições. As parcelas foram compostas de quatro fileiras de 4m de comprimento, espaçadas de 0,50m, e a área útil de 3m<sup>2</sup>. A densidade de semeadura foi ajustada de acordo com o hábito de crescimento de cada cultivar (CEPEF, 2003).

Foram avaliados os caracteres produtividade de grãos a 13% de umidade, número de vagens por planta, número de sementes por vagem, massa de cem grãos, população final de plantas, número de dias da emergência ao florescimento, número de dias da emergência à colheita, altura de inserção da primeira vagem, altura de inserção da última vagem e grau de acamamento (escala de notas entre um a nove, sendo 1: planta ereta e 9: planta acamada).

Foram avaliados a análise de variância conjunta dos experimentos e o teste F a 5% de probabilidade, considerando o efeito de cultivar como fixo e os efeitos de bloco e de ambiente como aleatórios. Para os caracteres com relação entre os quadrados médios residuais dos ambientes maior que sete (GOMES, 1990; CRUZ & REGAZZI, 1997), fez-se a correção dos graus de liberdade (CRUZ, 2001).

Foram determinadas as matrizes de distância euclidiana média padronizada (D) e de distância generalizada de Mahalanobis (D<sup>2</sup>) entre as cultivares. Essas matrizes de distâncias, em escala relativa, foram utilizadas como medida de dissimilaridade para a análise de agrupamento das cultivares pelo método otimização de Tocher e pelos métodos hierárquicos da ligação simples (vizinho mais próximo), de Ward, da ligação completa (vizinho mais distante), da mediana, da ligação média dentro de grupo e da ligação média entre grupo (CRUZ & REGAZZI, 1997; CRUZ, 2001; CRUZ & CARNEIRO, 2003).

As duas matrizes (D e D<sup>2</sup>) foram comparadas por meio do coeficiente de correlação de Pearson. Para

validar os agrupamentos, ou seja, verificar a capacidade do dendrograma em reproduzir as matrizes de dissimilaridade (D e D<sup>2</sup>), calculou-se o coeficiente de correlação cofenética (CCC). O CCC é o coeficiente de correlação de Pearson entre as matrizes de distâncias (D e D<sup>2</sup>) e a matriz cofenética (C) (matriz de distâncias entre as cultivares, obtida a partir do dendrograma) e valores próximos à unidade indicam melhor representação (BUSSAB et al., 1990; BARROSO & ARTES, 2003; CRUZ & CARNEIRO, 2003). Verificou-se a concordância entre os métodos hierárquicos por meio do coeficiente de correlação de Pearson entre as matrizes cofenéticas (BUSSAB et al., 1990).

Utilizou-se a análise de variância multivariada (MANOVA) a 5% de significância para testar a hipótese da diferença entre vetores de média de cada grupo. Representou-se a média dos grupos em um gráfico de perfis de médias (BARROSO & ARTES, 2003). As análises estatísticas foram realizadas no programa GENES (CRUZ, 2001) e no aplicativo Office Excel.

## RESULTADOS E DISCUSSÃO

A relação entre o maior e o menor quadrado médio do erro (QMr<sup>+</sup>/QMr<sup>-</sup>), da análise de variância individual dos nove experimentos, oscilou entre 3,96 (número de dias da emergência ao florescimento) e 16,47 (produtividade de grãos) (Tabela 1). Testes estatísticos de Bartlett e F máximo (STEEL et al., 1997; GOMES, 1990; CRUZ & REGAZZI, 1997) são indicados para verificar a homogeneidade das variâncias residuais. No entanto, GOMES (1990) e CRUZ & REGAZZI (1997) consideram as variâncias residuais homogêneas e a análise conjunta adequada quando a relação QMr<sup>+</sup>/QMr<sup>-</sup> é menor que sete. Por outro lado, quando a relação excede esse valor, há necessidade de exclusão de experimento(s) com quadrados médios residuais discrepantes, formação de grupos de experimentos com variâncias residuais homogêneas ou ajuste dos graus de liberdade para a interpretação correta dos testes de hipóteses. Assim, para a realização adequada da análise de variância conjunta, foi realizado o ajuste dos graus de liberdade em relação aos caracteres produtividade de grãos (PROD), massa de cem grãos (MCG), população final de plantas (POP), altura de inserção da primeira vagem (APV) e altura de inserção da última vagem (AUV) (Tabela 1).

A análise de variância conjunta em relação a todos os caracteres, com exceção do número de sementes por vagem (NSV), mostrou presença de interação cultivar x ambiente (experimento) significativa, o que revela comportamento diferencial

Tabela 1 - Análise de variância conjunta e significância dos quadrados médios dos caracteres produtividade de grãos (PROD), número de vagens por planta (NVP), número de sementes por vagem (NSV), massa de cem grãos (MCG), população final de plantas (POP), número de dias da emergência ao florescimento (FLOR), número de dias da emergência à colheita (CICLO), altura de inserção da primeira vagem (APV), altura de inserção da última vagem (AUV) e grau de acamamento (ACA), média, coeficiente de variação experimental (CV) e relação entre o maior e o menor quadrado médio residual entre os ambientes (QMr<sup>+</sup>/QMr<sup>-</sup>) de 14 cultivares de feijão, avaliadas em nove experimentos. Santa Maria-RS, UFSM, 2007.

Fonte de variação	-----Quadrado Médio <sup>(1 e 2)</sup> -----				
	PROD (kg ha <sup>-1</sup> )	NVP	NSV	MCG (g)	POP (plantas 3m <sup>-2</sup> )
Blocos/Ambiente	1516431,13 <sup>*(18)</sup>	40,85 <sup>*(18)</sup>	0,39 <sup>ns(18)</sup>	17,14 <sup>*(18)</sup>	264,55 <sup>*(18)</sup>
Cultivares (C)	1253542,70 <sup>*(13)</sup>	82,42 <sup>*(13)</sup>	5,23 <sup>*(13)</sup>	550,16 <sup>*(13)</sup>	831,54 <sup>*(13)</sup>
Ambientes (A)	21674226,98 <sup>*(8)</sup>	460,58 <sup>*(8)</sup>	12,68 <sup>*(8)</sup>	546,72 <sup>*(8)</sup>	1288,87 <sup>*(8)</sup>
CxA	366510,90 <sup>*(75)</sup>	15,80 <sup>*(104)</sup>	0,31 <sup>ns(104)</sup>	12,82 <sup>*(74)</sup>	258,12 <sup>*(57)</sup>
Resíduo	95535,36 <sup>(162)</sup>	6,47 <sup>(234)</sup>	0,27 <sup>(234)</sup>	6,20 <sup>(160)</sup>	148,60 <sup>(119)</sup>
Média	1409,96	11,52	4,29	19,03	56,34
CV (%)	21,92	22,09	12,13	13,08	21,64
QMr <sup>+</sup> /QMr <sup>-</sup>	16,47	4,06	4,59	8,41	15,17

Fonte de variação	-----Quadrado Médio <sup>(1 e 2)</sup> -----				
	FLOR (dias)	CICLO (dias)	APV (cm)	AUV (cm)	ACA <sup>(3)</sup>
Blocos/Ambiente	5,87 <sup>*(18)</sup>	28,91 <sup>*(18)</sup>	35,07 <sup>*(18)</sup>	47,16 <sup>ns(18)</sup>	4,82 <sup>*(18)</sup>
Cultivares (C)	142,45 <sup>*(13)</sup>	184,34 <sup>*(13)</sup>	138,26 <sup>*(13)</sup>	921,91 <sup>*(13)</sup>	49,38 <sup>*(13)</sup>
Ambientes (A)	478,74 <sup>*(8)</sup>	613,25 <sup>*(8)</sup>	455,01 <sup>*(8)</sup>	3271,00 <sup>*(8)</sup>	62,65 <sup>*(8)</sup>
CxA	6,04 <sup>*(104)</sup>	17,02 <sup>*(104)</sup>	39,83 <sup>*(78)</sup>	103,06 <sup>*(88)</sup>	2,39 <sup>*(104)</sup>
Resíduo	1,74 <sup>(234)</sup>	2,07 <sup>(234)</sup>	14,77 <sup>(168)</sup>	48,65 <sup>(194)</sup>	1,40 <sup>(234)</sup>
Média	36,99	77,19	15,03	40,94	5,99
CV (%)	3,57	1,86	25,57	17,04	19,79
QMr <sup>+</sup> /QMr <sup>-</sup>	3,96	6,79	7,62	8,13	5,67

(1) ns Não-significativo. \* Significativo a 5% de probabilidade pelo teste F. (2) Valores subscritos entre parênteses referem-se aos graus de liberdade da fonte de variação. (3) Escala de notas entre um a nove (1 : planta ereta; 9 : planta acamada).

das cultivares nos ambientes. Esses resultados sugerem que agrupamentos de cultivares com base em apenas uma época de cultivo podem fornecer informações equivocadas, por não considerarem a variabilidade ambiental entre anos e épocas de cultivo dentro de um mesmo local. BERTAN et al. (2006) afirmam que a significância do efeito de cultivares x ambientes revela a necessidade da avaliação das cultivares em mais que um ambiente (ano), sendo tão mais segura quanto maior for o número de ambientes avaliados.

O teste F, da análise de variância conjunta, revelou efeito de cultivar significativo em relação a todos os caracteres, evidenciando a possibilidade de identificação de cultivares superiores devido à heterogeneidade do germoplasma (Tabela 1). Isso permite inferir presença de variabilidade genética e possibilidade de agrupar as cultivares.

O coeficiente de variação experimental da análise de variância conjunta oscilou de 1,86% (número de dias da emergência à colheita - CICLO) a 25,57% (altura de inserção da primeira vagem - APV). LÚCIO et

al. (1999) classificaram os experimentos com média precisão experimental (CV entre 15,5% e 27,0%), baixa precisão experimental (CV entre 27,0% e 35,5%) e muito baixa precisão experimental (CV > 35,5%) em relação à produtividade de grãos de feijão. Assim, de modo geral, esses resultados indicam uma boa precisão experimental em relação a todos os caracteres (Tabela 1).

Houve correlação linear significativa e de alta magnitude ( $r=0,92$ ) entre as estimativas das distâncias euclidiana média padronizada (D) e generalizada de Mahalanobis (D<sup>2</sup>), o que evidencia boa concordância entre as medidas. Assim, há possibilidade de ambas formarem agrupamentos semelhantes. No entanto, a distância euclidiana média padronizada é desaconselhada na quantificação da divergência genética quando se dispõe de caracteres correlacionados. Além disso, na sua estimativa, são ignorados as informações em nível de repetições, sendo usado apenas os valores padronizados das médias. Então, o procedimento recomendável é a utilização da estatística D<sup>2</sup>, que leva em consideração as associações

entre os caracteres por meio da matriz de variâncias e de covariâncias residuais (CRUZ & CARNEIRO, 2003). Correlação em magnitude semelhante foi observada por CRUZ (1990) ( $r=0,97$ ) e inferior por MALUF & FERREIRA (1983) ( $r=0,273$ ) e por BENIN et al. (2003) ( $r=0,529$ ). No entanto, nesses trabalhos não foi utilizada a distância euclidiana média padronizada.

O método de Tocher, com base nas distâncias euclidiana média padronizada ( $D$ ) e generalizada de Mahalanobis ( $D^2$ ), separou as cultivares em dois grupos, com comportamento similar dentro do grupo e divergente entre grupos. O grupo 1 foi composto pela cultivar “Iraí” e o grupo 2 pelas demais 13 cultivares (“Carioca”, “Diamante Negro”, “TPS Nobre”, “Guapo Brilhante”, “Guatêian 6662”, “Iapar 44”, “Macanudo”, “Macotaço”, “Minuano”, “Pérola”, “Rio Tibagi”, “TPS Bionobre” e “TPS Bonito”).

Em um dendrograma, grande mudança de nível indica a união de cultivares heterogêneas (BARROSO & ARTES, 2003). Assim, utilizando-se 50% de similaridade como critério para definição dos grupos, pelos métodos hierárquicos da ligação simples, de Ward, da ligação completa, da mediana, da ligação média dentro de grupo e da ligação média entre grupo, com base na distância generalizada de Mahalanobis ( $D^2$ ), houve formação de dois grupos de cultivares (Figura 1). Em todos os métodos, a cultivar “Iraí” formou o grupo 1 e as demais treze cultivares constituíram o grupo 2, ou seja, houve a separação de germoplasma de origem Andina (“Iraí”) e Mesoamericano. Portanto, pode-se inferir que o método otimização de Tocher e os métodos hierárquicos com base na  $D^2$  agrupam as cultivares de forma similar e coerente com a origem genética do germoplasma avaliado. Além disso, em todos os agrupamentos formados com base na distância generalizada de Mahalanobis, as cultivares “Rio Tibagi” e “Iapar 44” estão inseridas no mesmo grupo, lado a lado, nos dendrogramas, o que confirma a coerência dos grupos formados, pois a “Rio Tibagi” foi um dos parentais utilizados no desenvolvimento da cultivar “Iapar 44” (RIBEIRO, 2001). O mesmo ocorreu com as cultivares “Macanudo” e “Minuano”, provenientes do mesmo cruzamento.

Por outro lado, os grupos formados com base na distância euclidiana média padronizada ( $D$ ) e 50% de similaridade foram discordantes entre os métodos hierárquicos. Houve formação de três, quatro, quatro, três, nove e quatro grupos de cultivares, respectivamente, pelos métodos hierárquicos da ligação simples, de Ward, da ligação completa, da mediana, da ligação média dentro de grupo e da ligação média entre grupo (Figura 1). No entanto, a separação

das cultivares em dois grupos (grupo 1 = cultivar “Iraí” e grupo 2 = demais cultivares) é possível, mas somente a 85% de similaridade. Esses resultados reforçam a adequabilidade da utilização da distância generalizada de Mahalanobis para a formação dos grupos em dados provenientes de experimentos planejados com repetição (CRUZ & CARNEIRO, 2003). De maneira geral, esses resultados assemelham-se aos verificados por TOTTI et al. (2001), KARASAWA et al. (2005), SUDRÉ et al. (2005), ALBUQUERQUE et al. (2006), BERTAN et al. (2006) e BONETT et al. (2006).

O coeficiente de correlação cofenética (CCC) obtido entre a matriz de distância generalizada de Mahalanobis ( $D^2$ ) e a matriz de distância cofenética ( $C$ ), obtida a partir do dendrograma de cada método, foram de elevada magnitude ( $r=0,90$ ) e significativos, evidenciando consistência dos agrupamentos (Tabela 2). Já o CCC entre a matriz de distância euclidiana média padronizada ( $D$ ) e  $C$  foram significativos, porém, de magnitude inferior, oscilando entre 0,75 (Ward) e 0,92 (ligação média entre grupo). Esses resultados revelam a maior consistência dos agrupamentos formados com base na distância generalizada de Mahalanobis ( $D^2$ ) em relação à euclidiana média padronizada ( $D$ ) e reforçam a adequabilidade do uso da  $D^2$  em relação à  $D$ . Sugere, ainda, maior concordância entre os métodos de agrupamento com base em  $D^2$ .

O coeficiente de correlação entre os métodos de agrupamento, obtidos com base na matriz da distância generalizada de Mahalanobis ( $D^2$ ), oscilou entre 0,89 (Ward e ligação simples) e 1,00 (ligação média entre grupo e ligação simples e ligação média entre grupo e mediana). Já com base matriz de distância euclidiana média padronizada ( $D$ ), variou de 0,65 (ligação simples e Ward) e 0,98 (Ward e ligação média dentro de grupo). Os coeficientes de correlação entre os métodos de agrupamento obtidos com base em  $D^2$  apresentaram magnitudes superiores aos obtidos com base em  $D$ , o que comprova maior concordância entre os métodos de agrupamento com base em  $D^2$  (Tabela 2).

A análise de variância multivariada (MANOVA) revelou pelos critérios de Wilk's, de Lawley-Hotelling e de Pillai's, que os vetores de média dos dois grupos formados por meio do método de Tocher e dos métodos hierárquicos, com base em  $D^2$ , diferem. Esses resultados revelam que o perfil de médias dos grupos difere e que esses grupos são divergentes, possibilitando cruzamentos promissores (Figura 2).

Do ponto de vista do melhorista de plantas, o processamento dos dados por diversos métodos de agrupamento e com base em diversas medidas de dissimilaridade e a consideração das particularidades de cada um é adequada para uma melhor tomada de

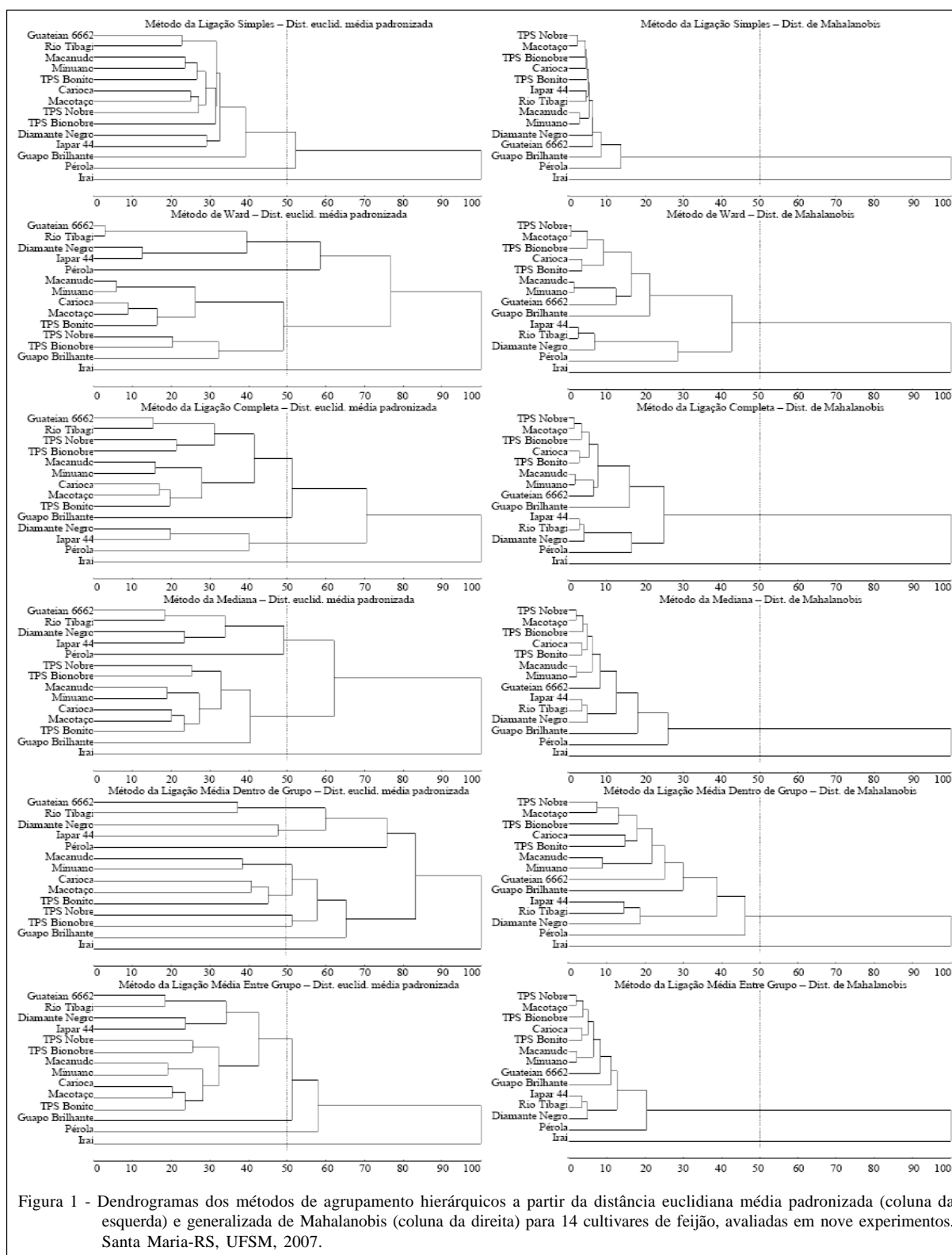


Figura 1 - Dendrogramas dos métodos de agrupamento hierárquicos a partir da distância euclidiana média padronizada (coluna da esquerda) e generalizada de Mahalanobis (coluna da direita) para 14 cultivares de feijão, avaliadas em nove experimentos. Santa Maria-RS, UFSM, 2007.

decisão em relação à escolha de cultivares para os cruzamentos. Os agrupamentos com base em  $D^2$  foram concordantes e qualquer método poderia ser utilizado.

Além disso, a  $D^2$  mostrou-se adequada em relação a  $D$ . No entanto, há necessidade de comparações adicionais antes dessas considerações serem generalizadas e

Tabela 2 - Coeficiente de correlação cofenética entre as matrizes de distância euclidiana média padronizada (D) e generalizada de Mahalanobis ( $D^2$ ) e os métodos de agrupamento hierárquicos e entre os métodos de agrupamento hierárquicos. Acima e abaixo da diagonal estão apresentados, respectivamente, os coeficientes a partir de D e  $D^2$ . Santa Maria-RS, UFSM, 2007.

	$D^2$	LS	WA	LC	ME	LMDG	LMEG
D		0,90*	0,75*	0,86*	0,86*	0,78*	0,92*
LS	0,97*		0,65*	0,80*	0,83*	0,69*	0,97*
WA	0,90*	0,89*		0,77*	0,94*	0,98*	0,77*
LC	0,96*	0,97*	0,97*		0,83*	0,79*	0,85*
ME	0,97*	0,99*	0,92*	0,98*		0,97*	0,90*
LMDG	0,94*	0,94*	0,98*	0,99*	0,97*		0,81*
LMEG	0,97*	1,00*	0,93*	0,98*	1,00*	0,97*	

<sup>ns</sup> Não-significativo. \* significativo a 5% de probabilidade, pelo teste t com 89 graus de liberdade. LS: Ligação Simples; WA: Ward; LC: Ligação Completa; ME: Mediana; LMDG: Ligação Média Dentro de Grupo; e LMEG: Ligação Média Entre Grupo.

ainda comparações com outras metodologias com base em outras medidas de dissimilaridade.

## CONCLUSÕES

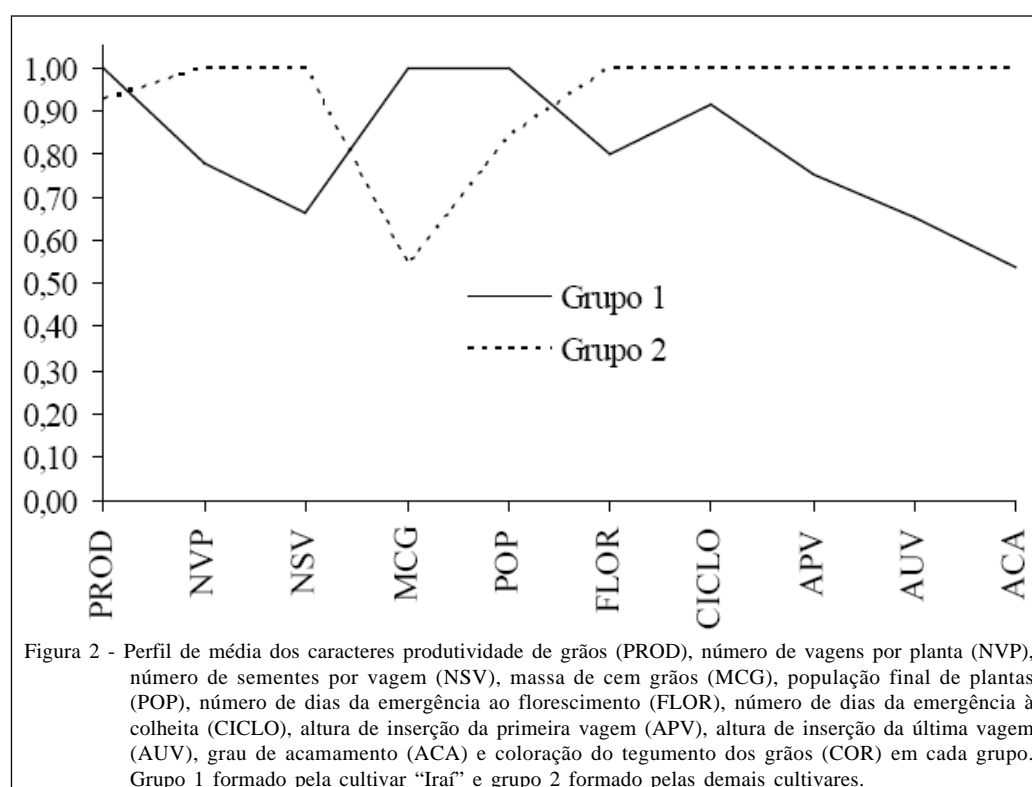
Agrupamentos com base na distância euclidiana média padronizada são distintos dos formados com base na distância generalizada de Mahalanobis, exceto o método de otimização de Tocher. O método de otimização de Tocher e os métodos hierárquicos da ligação simples, de Ward, da ligação completa, da mediana, da ligação média dentro de grupo

e da ligação média entre grupo, com base na distância generalizada de Mahalanobis, formam grupos concordantes e coerentes com a origem genética do germoplasma avaliado.

A cultivar “Iraí” possui comportamento distinto das demais cultivares e as técnicas de agrupamento, com base na distância generalizada de Mahalanobis, captaram essa divergência.

## AGRADECIMENTOS

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela concessão de bolsa de





produtividade em pesquisa à Alberto Cargnelutti Filho e Nerinéia Dalfollo Ribeiro. À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela concessão de bolsa de mestrado à Joseane Rodrigues de Souza e à Evandro Jost.

## REFERÊNCIAS

- ALBUQUERQUE, M.A. et al. Estabilidade em análise de agrupamento: estudo de caso em ciência florestal. **Revista Árvore**, v.30, p.257-265, 2006.
- BARROSO, L.P.; ARTES, R. **Análise multivariada**. Lavras: UFLA, 2003. 151p.
- BENIN, G. et al. Comparações entre medidas de dissimilaridade e estatísticas multivariadas como critérios no direcionamento de hibridações em aveia. **Ciência Rural**, v.33, p.657-662, 2003.
- BERTAN, I. et al. Comparação de métodos de agrupamento na representação da distância morfológica entre genótipos de trigo. **Revista Brasileira de Agrociência**, v.12, p.279-286, 2006.
- BONETT, L.P. et al. Divergência genética em germoplasma de feijoeiro comum coletado no estado do Paraná, Brasil. **Semina**, v.27, p.547-560, 2006.
- BUSSAB, W.O. et al. **Introdução à análise de agrupamentos**. São Paulo: ABE, 1990. 105p.
- COMISSÃO ESTADUAL DE PESQUISA DE FEIJÃO. CEPEF. **Indicações técnicas para a cultura do feijão no Rio Grande do Sul 2003/04**. Passo Fundo: UPF, 2003. 149p.
- CRUZ, C.D.; CARNEIRO, P.C.S. **Modelos biométricos aplicados ao melhoramento genético**. Viçosa: UFV, 2003. 585p.
- CRUZ, C.D.; REGAZZI, A.J. **Modelos biométricos aplicados ao melhoramento genético**. 2.ed. Viçosa: UFV, 1997. 390p.
- CRUZ, C.D. **Aplicação de algumas técnicas multivariadas no melhoramento de plantas**. 1990. 188f. Tese (Doutorado em Agronomia) - Programa de Pós-graduação em Genética e Melhoramento de Plantas, Escola Superior de Agricultura Luiz de Queiroz.
- CRUZ, C.D. **PROGRAMA GENES**: versão Windows: aplicativo computacional em genética e estatística. Viçosa: UFV, 2001. 648p.
- GOMES, F.P. **Curso de estatística experimental**. 13.ed. Piracicaba: São Paulo: Nobel 1990. 468p.
- KARASAWA, M. et al. Aplicação de métodos de agrupamento na quantificação da divergência genética entre acessos de tomateiro. **Horticultura Brasileira**, v.23, p.1000-1005, 2005.
- LÚCIO, A.D. et al. Classificação dos experimentos de competição de cultivares quanto a sua precisão. **Pesquisa Agropecuária Gaúcha**, v.5, p.99-103, 1999.
- MALUF, W.R.; FERREIRA, P.E. Análise multivariada da divergência genética em feijão-vagem (*Phaseolus vulgaris* L.). **Horticultura Brasileira**, v.1, p.31-34, 1983.
- MINGOTI, S.A. **Análise de dados através de métodos de estatística multivariada**. Belo Horizonte: UFMG, 2005. 297p.
- RIBEIRO, N.D. **Escolha de genitores de feijoeiro por meio da divergência genética**. 2001. 80f. Tese (Doutorado em Agronomia) – Programa de Pós-graduação em Agronomia, Universidade Federal de Santa Maria.
- STEEL, R.G.D. et al. **Principles and procedures of statistics a biometrical approach**. 3.ed. New York: McGraw Hill Book, 1997. 666p.
- SUDRÉ, C.P. et al. Divergência genética entre acessos de pimenta e pimentão utilizando técnicas multivariadas. **Horticultura Brasileira**, v.23, p.22-27, 2005.
- TOTTI, R. et al. Utilização de métodos de agrupamentos hierárquicos em acessos de Paspalum (*Graminea (Poaceae)*). **Semina**, v.22, p.25-35, 2001.