

# Consistência do padrão de agrupamento de cultivares de feijão conforme medidas de dissimilaridade e métodos de agrupamento

Alberto Cargnelutti Filho<sup>(1)</sup>, Nerinéia Dalfollo Ribeiro<sup>(1)</sup> e Cláudia Burin<sup>(1)</sup>

<sup>(1)</sup>Universidade Federal de Santa Maria, Departamento de Fitotecnia, Avenida Roraima, nº 1.000, Bairro Camobi, CEP 97105-900 Santa Maria, RS. E-mail: cargnelutti@pq.cnpq.br, nerineia@hotmail.com, clauburin@yahoo.com.br

Resumo – O objetivo deste trabalho foi avaliar a consistência do padrão de agrupamento de cultivares de feijão (*Phaseolus vulgaris* L.), obtido a partir da combinação de oito medidas de dissimilaridade e oito métodos de agrupamento. Treze cultivares de feijão foram avaliadas em nove experimentos conduzidos entre os anos agrícolas de 2000/2001 e 2004/2005, e agrupadas de acordo com caracteres de produção (produtividade de grãos, número de vagens por planta), de fenologia (número de dias da emergência ao florescimento e da emergência à colheita) e de morfologia (altura de inserção da primeira e da última vagem). Foram realizadas análises de variância, de correlação, de diagnóstico de multicolinearidade, de agrupamento e de comparação de médias. A consistência do padrão de agrupamento foi avaliada por meio do coeficiente de correlação cofenética. Há variabilidade na consistência do padrão de agrupamento das cultivares de feijão, obtido a partir da combinação de diferentes medidas de dissimilaridade e métodos de agrupamento. Maior consistência nos padrões de agrupamento de cultivares de feijão é verificada com o método da ligação média entre grupo, obtido a partir da matriz de distância euclidiana.

Termos para indexação: *Phaseolus vulgaris*, algoritmos de agrupamento, coeficiente de correlação cofenética, medidas de dissimilaridade.

## Clustering pattern consistency of common bean cultivars by dissimilarity measures and clustering methods

Abstract – The objective of this work was to evaluate the clustering pattern consistency in common bean cultivars (*Phaseolus vulgaris* L.) obtained from the combination of the eight dissimilarity measures and eight clustering methods. Thirteen common bean cultivars were evaluated in nine experiments conducted in the growing seasons between 2000/2001 and 2004/2005 and grouped according to productive (grain yield and number of pods per plant), phenological (number of days between seedling emergence and flowering, and number of days from emergence to harvest) and morphological (height of first pod insertion and height of the final pod insertion) characters. Analyses of variance, correlation, diagnoses of multicollinearity, cluster and comparison of means were carried out. Clustering pattern consistency was evaluated by the cophenetic correlation coefficient. There is variability in the clustering pattern consistency of common bean cultivars, derived from the combination of different dissimilarity measures and clustering methods. Greater clustering pattern consistency in common bean is found with the average linkage between groups method, based on the euclidian distance.

Index terms: *Phaseolus vulgaris*, clustering algorithms, cophenetic correlation coefficient, dissimilarity measures.

### Introdução

A identificação de cultivares divergentes, por meio de análise de agrupamento, tem sido empregada em programas de melhoramento de feijão. Para tanto, há diversas medidas de dissimilaridade entre pares de cultivares que são utilizadas nos diversos métodos de agrupamento. Os diferentes métodos de análise e combinações entre medidas levam a padrões de agrupamento

distintos (Cruz & Regazzi, 1997; Mingoti, 2005; Manly, 2008).

O coeficiente de correlação linear de Pearson entre os elementos da matriz de dissimilaridade (matriz de distâncias entre as cultivares, obtida a partir dos dados originais) e os elementos da matriz cofenética (matriz de distâncias entre as cultivares, obtida a partir do dendrograma) é denominado coeficiente de correlação cofenética.

Esse coeficiente pode ser utilizado para avaliar a consistência do padrão de agrupamento de métodos de agrupamento hierárquicos, sendo que valores próximos à unidade indicam melhor representação (Barroso & Artes, 2003; Cruz & Carneiro, 2003).

Comparações entre algumas medidas de dissimilaridade (Maluf & Ferreira, 1983) e entre alguns métodos de agrupamento (Bonett et al., 2006; Cargnelutti Filho et al., 2008) têm sido realizadas em estudos de divergência genética em feijão e revelado níveis de concordância variáveis. Entretanto, estudos comparativos entre os coeficientes de correlação cofenética obtidos por diferentes medidas de dissimilaridade e métodos de agrupamento, para a cultura do feijão, não foram encontrados na literatura.

O objetivo deste trabalho foi avaliar a consistência do padrão de agrupamento de cultivares de feijão (*Phaseolus vulgaris* L.), obtido a partir da combinação de oito medidas de dissimilaridade e oito métodos de agrupamento.

### Material e Métodos

Foram avaliadas 13 cultivares de feijão – Carioca, Diamante Negro, TPS Nobre, Guapo Brilhante, Guateian 6662, IAPAR 44, Macanudo, Macotaço, Minuano, Pérola, Rio Tibagi, TPS Bionobre e TPS Bonito – em nove experimentos, realizados em delineamento de blocos ao acaso, com três repetições, na Universidade Federal de Santa Maria, Santa Maria, Rio Grande do Sul (29°42'S, 53°49'W, a 95 m de altitude). As parcelas foram compostas de quatro fileiras de 4 m de comprimento, espaçadas de 0,50 m, e com área útil de 3 m<sup>2</sup>.

Os experimentos foram conduzidos nos anos agrícolas de 2000/2001 a 2004/2005, em duas épocas de cultivo: safra (semeadura em setembro ou outubro) e safrinha (semeadura em janeiro ou fevereiro). Em cada experimento, foram avaliados os caracteres de produção (produtividade de grãos a 13% de umidade e número de vagens por planta), de fenologia (número de dias da emergência da plântula ao florescimento e da emergência à colheita) e de morfologia (altura de inserção da primeira e última vagem).

Para avaliação dos efeitos das cultivares sobre esses caracteres, foi utilizada a análise de variância, com o teste F a 5% de probabilidade. Em seguida, em cada experimento, foi determinada a matriz de coeficientes de correlação linear de Pearson entre os caracteres (matriz fenotípica) e realizado o diagnóstico de multicolinearidade (Cruz, 2006), conforme critério de Montgomery & Peck (1982). Depois foram determinadas as seguintes matrizes de distância entre as cultivares: euclidiana (E), euclidiana padronizada (EP), euclidiana média (EM), euclidiana média padronizada (EMP), quadrado da distância euclidiana (QE), quadrado da distância euclidiana padronizada (QEP), Mahalanobis (M) e Mahalanobis padronizada (MP) (Cruz, 2006). Essas matrizes, em escala relativa, foram utilizadas como medida de dissimilaridade para a análise de agrupamento das cultivares pelos seguintes métodos hierárquicos: ligação simples (vizinho mais próximo), ligação completa (vizinho mais distante), Ward, mediana, ligação média dentro de grupo, ligação média entre grupo, Gower e Centroide (Cruz & Regazzi, 1997; Cruz & Carneiro, 2003; Cruz, 2006). Ao final, foram obtidos 576 dendrogramas a partir da combinação de nove experimentos, oito distâncias e oito métodos de agrupamento.

Para avaliar a consistência dos 576 agrupamentos, ou seja, verificar a capacidade do dendrograma em reproduzir as matrizes de dissimilaridade (E, EP, EM, EMP, QE, QEP, M e MP), calculou-se o coeficiente de correlação cofenética (Barroso & Artes, 2003; Cruz & Carneiro, 2003). A significância desse coeficiente foi verificada por meio do teste t de Student, a 5% de probabilidade, com  $n - 2 = 76$  graus de liberdade, em que  $n$  é igual a 78 distâncias resultantes da combinação de 13 cultivares, duas a duas. Compararam-se as médias dos coeficientes de correlação cofenética (média dos nove experimentos), das oito matrizes de distância dentro de cada método de agrupamento, e as médias dos oito métodos de agrupamento dentro de cada matriz de distância, por meio do teste t, a 5% de probabilidade. As análises estatísticas foram realizadas no programa Genes (Cruz, 2006).

## Resultados e Discussão

Em relação aos caracteres avaliados em nove experimentos, o teste F da análise de variância revelou que o efeito de bloco foi significativo ( $p \leq 0,05$ ) em 16 casos (29,63%) (Tabela 1). Esse resultado evidencia que os blocos foram heterogêneos e que o uso do delineamento blocos ao acaso foi adequado.

Entre os 54 casos avaliados, houve efeito significativo para cultivares em 47 casos (87,04%), evidenciando a possibilidade de utilização da técnica de agrupamento para a identificação de cultivares superiores em razão da heterogeneidade observada no germoplasma avaliado (Tabela 1). Nos 47 casos com efeito significativo de cultivares, a média da acurácia seletiva foi de 0,87, enquanto que, nos sete casos em que não foram detectadas diferenças significativas, a média da acurácia seletiva foi de 0,55. Esses resultados revelaram que a não identificação de diferenças entre as cultivares está associada a menor precisão experimental (Cargnelutti Filho & Storck, 2007, 2009; Resende & Duarte, 2007).

O diagnóstico de multicolinearidade, na matriz de coeficientes de correlação linear de Pearson entre os caracteres avaliados, revelou que o número de condição (NC) oscilou entre 15 (Safrinha 2004) a 102 (Safrinha 2002) (Tabela 1). De acordo com Montgomery & Peck (1982), a matriz pode apresentar multicolinearidade fraca ( $NC < 100$ ), moderada a forte ( $100 < NC < 1.000$ ) ou severa ( $NC > 1.000$ ). Em presença de multicolinearidade, o uso de todos os caracteres na análise de agrupamento não é um procedimento adequado, pois os caracteres multicolineares são implicitamente ponderados com maior peso (Barroso & Artes, 2003; Cruz & Carneiro, 2003; Hair et al., 2005; Corrar et al., 2007). Neste trabalho, em apenas um experimento o NC indicou multicolinearidade moderada a forte, porém próximo da multicolinearidade fraca encontrada nos outros oito experimentos. Dessa forma, o efeito da multicolinearidade não foi detectado e a análise de agrupamento com base nos seis caracteres foi considerada um procedimento adequado (Cargnelutti Filho et al., 2009b).

A identificação de cultivares divergentes, com base em apenas um experimento, comumente é pouco eficiente devido à variabilidade de resposta

das cultivares às condições ambientais (interação genótipo  $\times$  ambientes) (Bertan et al., 2006). O uso dos dados de sete e seis experimentos, respectivamente, para os métodos de Ward e de Tocher, foi considerado suficiente para o agrupamento de cultivares de feijão (Cargnelutti Filho et al., 2009a). Assim, se a finalidade principal desse estudo fosse identificar cultivares divergentes, a análise conjunta dos nove experimentos seria adequada. No entanto, esse estudo teve por objetivo avaliar a consistência do padrão de agrupamento de cultivares de feijão, obtidos a partir da combinação de diferentes medidas de dissimilaridade e métodos de agrupamento, e a análise individual é particularmente importante, por proporcionar uma quantidade maior de cenários.

O coeficiente de correlação cofenética (CCC) oscilou entre 0,2437 e 0,9221 (Tabela 2). A média entre os 576 valores de CCC formados pela combinação de nove experimentos, oito distâncias e oito métodos de agrupamento foi de 0,6733, revelando variabilidade na consistência do padrão de agrupamento entre os experimentos, matrizes de distância e métodos de agrupamento. Embora os 576 CCC tenham sido significativos ( $p \leq 0,05$ ), deve-se avaliar a magnitude dos valores de CCC tendo-se em mente que os valores próximos à unidade indicam maior consistência do padrão de agrupamento (Barroso & Artes, 2003; Cruz & Carneiro, 2003) e devem ser os preferidos.

Para um mesmo método de agrupamento, três pares de matrizes de distâncias entre as cultivares [par 1: euclidiana (E) e euclidiana média (EM); par 2: euclidiana padronizada (EP) e euclidiana média padronizada (EMP); e par 3: Mahalanobis (M) e Mahalanobis padronizada (MP)] deram origem ao mesmo padrão de agrupamento. O coeficiente de correlação linear de Pearson entre esses pares de matrizes de distância foi igual a 1,00 e explica a relação direta e linear existente entre elas. Assim, são esperados os mesmos valores de CCC, em cada método, a partir desses três pares de matrizes de distância entre as cultivares (E e EM, EP e EMP, M e MP) (Tabela 2). Portanto, ao investigar o método que proporciona maior consistência no padrão de agrupamento, apenas uma matriz de distância de cada um desses três pares pode ser usada.

Há relação direta e quadrática entre o par de matriz de distância euclidiana (E) e euclidiana média (EM) com a matriz quadrado da distância euclidiana (QE). A mesma relação existe entre o par de matriz de

**Tabela 1.** Análise de variância para os seis caracteres avaliados nas 13 cultivares de feijão, precisão experimental<sup>(1)</sup> de nove experimentos e número de condição (NC) da matriz de coeficientes de correlação linear de Pearson entre os caracteres.

FV <sup>(2)</sup>	GL	Experimento <sup>(3)</sup>								
		1	2	3	4	5	6	7	8	9
Produtividade de grãos										
Bloco	2	0,05 <sup>ns</sup>	0,07 <sup>ns</sup>	11,35*	0,09 <sup>ns</sup>	0,05 <sup>ns</sup>	0,04*	0,03 <sup>ns</sup>	0,11 <sup>ns</sup>	0,41*
Cultivar	12	0,54*	0,24*	0,23 <sup>ns</sup>	0,16*	1,31*	0,03*	0,11 <sup>ns</sup>	0,09*	0,41*
Erro	24	0,09	0,03	0,15	0,04	0,13	0,01	0,07	0,04	0,03
Média (Mg ha <sup>-1</sup> )	-	1,77	0,85	2,39	0,71	1,85	0,26	2,01	0,88	1,89
CV(%)	-	17,29	21,61	16,24	27,21	19,60	36,36	13,67	22,68	8,67
Fc	-	5,72	7,25	1,53	4,37	9,97	2,94	1,50	2,23	15,22
AS	-	0,91	0,93	0,59	0,88	0,95	0,81	0,58	0,74	0,97
Precisão	-	MA	MA	M	A	MA	A	M	A	MA
Número de vagens por planta										
Bloco	2	7,69 <sup>ns</sup>	51,18*	232,90*	13,35 <sup>ns</sup>	20,60 <sup>ns</sup>	5,85 <sup>ns</sup>	10,82 <sup>ns</sup>	3,99 <sup>ns</sup>	26,05*
Cultivar	12	20,38*	48,26*	29,84*	26,56*	29,08*	5,88*	12,38 <sup>ns</sup>	6,29 <sup>ns</sup>	22,21*
Erro	24	8,83	4,52	9,35	8,20	9,22	2,46	8,75	3,62	5,16
Média	-	10,31	10,55	17,91	9,61	11,63	6,42	14,19	9,79	14,92
CV(%)	-	28,82	20,15	17,07	29,81	26,11	24,40	20,85	19,44	15,23
Fc	-	2,31	10,68	3,19	3,24	3,15	2,39	1,41	1,73	4,30
AS	-	0,75	0,95	0,83	0,83	0,83	0,76	0,54	0,65	0,88
Precisão	-	A	MA	A	A	A	A	M	M	A
Número de dias da emergência ao florescimento										
Bloco	2	1,26 <sup>ns</sup>	0,95 <sup>ns</sup>	40,95*	13,64*	0,49 <sup>ns</sup>	1,46 <sup>ns</sup>	3,41 <sup>ns</sup>	0,95 <sup>ns</sup>	
Cultivar	12	16,09*	17,49*	3,18 <sup>ns</sup>	7,64*	10,08*	6,05*	6,68*	7,24*	7,75*
Erro	24	0,95	1,17	2,64	2,64	0,99	2,24	0,91	1,41	1,25
Média	-	38,21	33,95	38,26	34,21	37,10	42,38	43,08	36,56	34,03
CV(%)	-	2,55	3,19	4,25	4,75	2,68	3,53	2,21	3,25	3,29
Fc	-	16,92	14,94	1,20	2,89	10,21	2,70	7,37	5,14	6,18
AS	-	0,97	0,97	0,41	0,81	0,95	0,79	0,93	0,90	0,92
Precisão	-	MA	MA	B	A	MA	A	MA	MA	MA
Número de dias da emergência à colheita										
Bloco	2	2,39 <sup>ns</sup>	0,80 <sup>ns</sup>	248,18*	5,62 <sup>ns</sup>	9,56*	4,69 <sup>ns</sup>	4,00*	1,56 <sup>ns</sup>	9,54*
Cultivar	12	54,47*	5,44*	10,09*	47,36*	27,79*	4,42*	7,42*	50,64*	18,45*
Erro	24	2,22	0,71	1,79	3,62	1,62	1,97	0,69	2,23	1,01
Média	-	77,54	77,05	73,79	78,31	79,26	82,46	82,46	78,13	70,08
CV(%)	-	1,92	1,09	1,81	2,43	1,61	1,70	1,01	1,91	1,43
Fc	-	24,56	7,64	5,63	13,10	17,16	2,24	10,68	22,70	18,26
AS	-	0,98	0,93	0,91	0,96	0,97	0,74	0,95	0,98	0,97
Precisão	-	MA	MA	MA	MA	MA	A	MA	MA	MA
Altura de inserção da primeira vagem										
Bloco	2	18,29 <sup>ns</sup>	177,62*	48,67*	26,69 <sup>ns</sup>	6,76 <sup>ns</sup>	14,18*	4,54 <sup>ns</sup>	19,07 <sup>ns</sup>	30,17*
Cultivar	12	141,22*	52,74*	22,73*	20,04 <sup>ns</sup>	55,92*	6,56 <sup>ns</sup>	16,82*	14,68*	35,95*
Erro	24	15,32	23,70	9,35	17,25	12,32	3,34	3,93	6,54	5,39
Média (cm)	-	13,51	14,94	15,56	9,15	18,61	12,83	16,24	16,16	20,71
CV(%)	-	28,97	32,60	19,64	45,39	18,86	14,25	12,20	15,83	11,21
Fc	-	9,22	2,23	2,43	1,16	4,54	1,96	4,28	2,24	6,67
AS	-	0,94	0,74	0,77	0,37	0,88	0,70	0,88	0,74	0,92
Precisão	-	MA	A	A	B	A	A	A	A	MA
Altura de inserção da última vagem										
Bloco	2	126,84 <sup>ns</sup>	50,92 <sup>ns</sup>	12,73 <sup>ns</sup>	51,85 <sup>ns</sup>	1,40 <sup>ns</sup>	11,32 <sup>ns</sup>	38,23 <sup>ns</sup>	99,02 <sup>ns</sup>	14,61 <sup>ns</sup>
Cultivar	12	305,39*	97,90*	73,55*	126,67*	192,96*	26,88*	120,28*	224,63*	106,58*
Erro	24	84,20	38,93	30,39	54,03	40,73	10,21	37,75	36,76	28,34
Média (cm)	-	39,27	39,15	43,21	43,95	55,00	22,90	43,98	39,05	51,36
CV(%)	-	23,36	15,94	12,76	16,73	11,60	13,95	13,97	15,53	10,37
Fc	-	3,63	2,51	2,42	2,34	4,74	2,63	3,19	6,11	3,76
AS	-	0,85	0,78	0,77	0,76	0,89	0,79	0,83	0,91	0,86
Precisão	-	A	A	A	A	A	A	A	MA	A
NC	-	72	79	22	102	52	42	74	15	64

<sup>(1)</sup>Avaliada de acordo com Cargnelutti Filho & Storck (2009): MA, muito alta (AS $\geq$ 0,90); A, alta (0,70 $\leq$ AS $<$ 0,90); M, moderada (0,50 $\leq$ AS $<$ 0,70); e B, baixa (AS $<$ 0,50). <sup>(2)</sup>Os valores relatados para as fontes de variação (bloco, cultivar e erro) referem-se ao quadrado médio da análise de variância. <sup>(3)</sup>1, Safra 2000/2001; 2, Safra 2001/2002; 3, Safra 2001/2002; 4, Safra 2002; 5, Safra 2002/2003; 6, Safra 2003; 7, Safra 2003/2004; 8, Safra 2004; e 9, Safra 2004/2005. <sup>ns</sup>Não significativo. \*Significativo a 5% de probabilidade pelo teste F. GL, graus de liberdade; Fc, valor do teste F para cultivar; AS, acurácia seletiva.

distância euclidiana padronizada (EP) e euclidiana média padronizada (EMP) com a matriz quadrado da distância euclidiana padronizada (QEP). Esses resultados são esperados e explicados pelo

estimador de cada distância (Cruz, 2006). Portanto, nessas relações, o coeficiente de correlação linear de Pearson entre as matrizes de distância não é adequado para avaliar a concordância entre as matrizes. Essas

**Tabela 2.** Coeficiente de correlação cofenética<sup>(1)</sup> entre as matrizes de distância euclidiana, euclidiana padronizada, euclidiana média, euclidiana média padronizada, quadrado da distância euclidiana, quadrado da distância euclidiana padronizada, Mahalanobis e Mahalanobis padronizada e a matriz cofenética dos métodos de agrupamento hierárquicos da ligação simples (vizinho mais próximo, LS), da ligação completa (vizinho mais distante, LC), de Ward, da mediana, da ligação média dentro de grupos (LMDG), da ligação média entre grupos (LMEG), de Gower e do Centróide, calculados com base em seis caracteres avaliados em 13 cultivares de feijão, em nove experimentos.

Método	Experimento <sup>(2)</sup>									Média
	1	2	3	4	5	6	7	8	9	
Distância euclidiana e euclidiana média										
LS	0,8393	0,6563	0,7111	0,8888	0,6847	0,7214	0,8115	0,6985	0,8407	0,7614
LC	0,8731	0,6454	0,6434	0,7008	0,6424	0,6451	0,5685	0,6447	0,6987	0,6736
Ward	0,8621	0,6747	0,6307	0,7051	0,5773	0,5318	0,5937	0,5676	0,7525	0,6551
Mediana	0,8747	0,6351	0,5429	0,9142	0,7324	0,5099	0,8189	0,6497	0,7437	0,7135
LMDG	0,8599	0,6819	0,6689	0,8293	0,6151	0,5931	0,7696	0,5998	0,8186	0,7151
LMEG	0,8750	0,6712	0,7545	0,9187	0,7814	0,7207	0,8244	0,7406	0,8205	0,7897
Gower	0,8634	0,6982	0,4616	0,8842	0,6970	0,7564	0,7805	0,6262	0,7545	0,7247
Centróide	0,8650	0,6264	0,6183	0,7459	0,6159	0,6121	0,7532	0,6550	0,8213	0,7015
Média	0,8641	0,6612	0,6289	0,8234	0,6683	0,6363	0,7400	0,6478	0,7813	0,7168
Distância euclidiana padronizada e euclidiana média padronizada										
LS	0,7292	0,4353	0,6355	0,8778	0,6506	0,7607	0,4360	0,6410	0,6905	0,6507
LC	0,7974	0,6681	0,6137	0,8469	0,5759	0,5206	0,5996	0,6606	0,5921	0,6528
Ward	0,6557	0,6542	0,6567	0,7806	0,5818	0,5830	0,6039	0,6463	0,7495	0,6569
Mediana	0,7746	0,6805	0,6653	0,8657	0,7293	0,7486	0,6073	0,6769	0,6161	0,7071
LMDG	0,8025	0,6152	0,5707	0,8539	0,7256	0,6336	0,6256	0,6548	0,7557	0,6931
LMEG	0,7751	0,6976	0,6733	0,8669	0,7321	0,7899	0,6529	0,6819	0,7817	0,7390
Gower	0,5726	0,5944	0,5149	0,8576	0,5296	0,7317	0,4965	0,5076	0,7347	0,6155
Centróide	0,7581	0,6535	0,4168	0,8371	0,6259	0,6608	0,5903	0,4644	0,6674	0,6305
Média	0,7332	0,6249	0,5934	0,8483	0,6439	0,6786	0,5765	0,6167	0,6985	0,6682
Quadrado da distância euclidiana										
LS	0,7720	0,6221	0,6634	0,8347	0,6536	0,6564	0,7395	0,6472	0,7299	0,7021
LC	0,7998	0,5631	0,5461	0,5605	0,5456	0,5843	0,4784	0,5528	0,5648	0,5773
Ward	0,7980	0,5691	0,5563	0,5803	0,4615	0,4603	0,4920	0,5546	0,6017	0,5638
Mediana	0,8002	0,5295	0,4351	0,8729	0,6796	0,4061	0,7388	0,5548	0,5949	0,6235
LMDG	0,8009	0,5735	0,5658	0,7857	0,4944	0,5029	0,7072	0,4808	0,6945	0,6229
LMEG	0,8003	0,5706	0,6889	0,8784	0,7182	0,7455	0,7445	0,6735	0,7082	0,7253
Gower	0,8014	0,6313	0,4176	0,8768	0,6891	0,7208	0,7365	0,4441	0,6038	0,6579
Centróide	0,7934	0,5415	0,5265	0,5814	0,4942	0,5207	0,6931	0,5978	0,6995	0,6053
Média	0,7958	0,5751	0,5500	0,7463	0,5920	0,5746	0,6663	0,5632	0,6497	0,6348
Quadrado da distância euclidiana padronizada										
LS	0,6589	0,3271	0,5957	0,8408	0,5629	0,7212	0,3906	0,5577	0,6036	0,5843
LC	0,7383	0,6148	0,5618	0,8029	0,5251	0,4453	0,5257	0,5851	0,5065	0,5895
Ward	0,7350	0,6088	0,6131	0,7300	0,5298	0,4841	0,5438	0,5706	0,6895	0,6116
Mediana	0,6830	0,5482	0,6149	0,8145	0,6532	0,7107	0,5497	0,6004	0,5227	0,6330
LMDG	0,7269	0,5551	0,5806	0,7964	0,6514	0,6724	0,5698	0,5696	0,6874	0,6455
LMEG	0,6832	0,5568	0,5728	0,8208	0,6551	0,7405	0,5592	0,6033	0,7032	0,6550
Gower	0,6703	0,6248	0,5952	0,8101	0,6505	0,7353	0,5449	0,5779	0,6983	0,6564
Centróide	0,6807	0,6338	0,3150	0,7972	0,5607	0,5705	0,5354	0,3481	0,5817	0,5581
Média	0,6970	0,5587	0,5561	0,8016	0,5986	0,6350	0,5274	0,5516	0,6241	0,6167
Distâncias de Mahalanobis e de Mahalanobis padronizada										
LS	0,7442	0,4873	0,4051	0,8842	0,6431	0,7913	0,5535	0,6460	0,5063	0,6290
LC	0,5889	0,5250	0,5620	0,8132	0,5548	0,5978	0,5861	0,7332	0,6171	0,6198
Ward	0,5819	0,5125	0,5694	0,9179	0,6828	0,7075	0,5914	0,7347	0,6213	0,6577
Mediana	0,7755	0,6159	0,5999	0,9189	0,6912	0,7927	0,5997	0,7338	0,6196	0,7052
LMDG	0,5955	0,5958	0,5899	0,9142	0,6877	0,7134	0,5938	0,7330	0,6231	0,6718
LMEG	0,5975	0,6588	0,6080	0,9221	0,6919	0,7932	0,6012	0,7364	0,6234	0,6925
Gower	0,7732	0,7080	0,4406	0,9205	0,6831	0,7691	0,5897	0,7317	0,6161	0,6924
Centróide	0,6741	0,6339	0,2437	0,7368	0,6837	0,7863	0,5777	0,7240	0,5811	0,6268
Média	0,6664	0,5922	0,5023	0,8785	0,6648	0,7439	0,5866	0,7216	0,6010	0,6619

<sup>(1)</sup>Valor superior a |0,23| é significativo a 5% de probabilidade, pelo teste t, com 76 graus de liberdade. <sup>(2)</sup>1, Safra 2000/2001; 2, Safra 2001; 3, Safra 2001/2002; 4, Safra 2002; 5, Safra 2002/2003; 6, Safra 2003; 7, Safra 2003/2004; 8, Safra 2004; e 9, Safra 2004/2005.

relações levam a diferentes padrões de agrupamento e, por consequência, a diferentes valores de CCC.

Independentemente do método de agrupamento, os coeficientes de correlação cofenética obtidos a partir das matrizes de distâncias euclidiana e euclidiana média foram maiores ( $p \leq 0,05$ ) que os alcançados a partir da matriz quadrado da distância euclidiana (Tabela 3).

**Tabela 3.** Estimativa dos contrastes de médias do coeficiente de correlação cofenética, em valores absolutos, entre matrizes de distância (E, euclidiana; EP, euclidiana padronizada; EM, euclidiana média; EMP, euclidiana média padronizada; QE, quadrado da distância euclidiana; QEP, quadrado da distância euclidiana padronizada; M, Mahalanobis; e MP, Mahalanobis padronizada), em cada método de agrupamento, calculados com base em seis caracteres de 13 cultivares de feijão, avaliadas em nove experimentos.

Distância	EP e EMP	QE	QEP	M e MP
Ligação simples				
E e EM	0,111*	0,059*	0,177*	0,132*
EP e EMP		0,051 <sup>ns</sup>	0,066*	0,022 <sup>ns</sup>
QE			0,118*	0,073 <sup>ns</sup>
QEP				0,045 <sup>ns</sup>
Ligação completa				
E e EM	0,021 <sup>ns</sup>	0,096*	0,084*	0,054 <sup>ns</sup>
EP e EMP		0,076 <sup>ns</sup>	0,063*	0,033 <sup>ns</sup>
QE			0,012 <sup>ns</sup>	0,043 <sup>ns</sup>
QEP				0,030 <sup>ns</sup>
Ward				
E e EM	0,002 <sup>ns</sup>	0,091*	0,043*	0,003 <sup>ns</sup>
EP e EMP		0,093*	0,045*	0,001 <sup>ns</sup>
QE			0,048*	0,094 <sup>ns</sup>
QEP				0,046 <sup>ns</sup>
Mediana				
E e EM	0,006 <sup>ns</sup>	0,090*	0,080 <sup>ns</sup>	0,008 <sup>ns</sup>
EP e EMP		0,084 <sup>ns</sup>	0,074*	0,002 <sup>ns</sup>
QE			0,009 <sup>ns</sup>	0,082 <sup>ns</sup>
QEP				0,072*
Ligação média dentro de grupo				
E e EM	0,022 <sup>ns</sup>	0,092*	0,070*	0,043 <sup>ns</sup>
EP e EMP		0,070 <sup>ns</sup>	0,048*	0,021 <sup>ns</sup>
QE			0,023 <sup>ns</sup>	0,049 <sup>ns</sup>
QEP				0,026 <sup>ns</sup>
Ligação média entre grupo				
E e EM	0,051 <sup>ns</sup>	0,064*	0,135*	0,097*
EP e EMP		0,014 <sup>ns</sup>	0,084*	0,047 <sup>ns</sup>
QE			0,070*	0,033 <sup>ns</sup>
QEP				0,038 <sup>ns</sup>
Gower				
E e EM	0,109*	0,067*	0,068 <sup>ns</sup>	0,032 <sup>ns</sup>
EP e EMP		0,042 <sup>ns</sup>	0,041 <sup>ns</sup>	0,077 <sup>ns</sup>
QE			0,002 <sup>ns</sup>	0,035 <sup>ns</sup>
QEP				0,036 <sup>ns</sup>
Centroide				
E e EM	0,071 <sup>ns</sup>	0,096*	0,143*	0,075 <sup>ns</sup>
EP e EMP		0,025 <sup>ns</sup>	0,072*	0,004 <sup>ns</sup>
QE			0,047 <sup>ns</sup>	0,021 <sup>ns</sup>
QEP				0,069 <sup>ns</sup>

<sup>ns</sup>Não significativo. \*Contraste significativo a 5% de probabilidade, pelo teste t.

Comportamento similar foi observado entre as matrizes de distâncias euclidiana padronizada e euclidiana média padronizada em relação à matriz quadrado da distância euclidiana padronizada, exceto em relação ao método de Gower. As matrizes de distâncias de Mahalanobis e de Mahalanobis padronizada apresentaram valores intermediários de CCC (Tabelas 2 e 3). Esses resultados evidenciam, de maneira geral, que as matrizes de distâncias euclidiana e euclidiana média proporcionaram agrupamentos mais consistentes. Em análise de agrupamento, a padronização das variáveis tem sido recomendada para minimizar o efeito das diferentes escalas de medidas, fazendo com que todas as variáveis exerçam a mesma importância na definição dos grupos (Cruz & Regazzi, 1997; Barroso & Artes, 2003; Hair et al., 2005; Corrar et al., 2007). Assim, valores distintos de CCC entre medidas de dissimilaridade com e sem padronização são esperados. No entanto, não é o foco deste trabalho julgar a necessidade de padronização desse conjunto de caracteres.

Com raras exceções, o método da ligação média entre grupo a partir das matrizes de distâncias euclidiana, euclidiana média, euclidiana padronizada, euclidiana média padronizada e quadrado da distância euclidiana apresentou maiores coeficientes de correlação cofenética, quando comparado aos demais métodos (Tabelas 2 e 4). Portanto, quando a opção for por uma dessas medidas de dissimilaridade, esse método de agrupamento deve ser o preferido, pois proporcionou melhor consistência do padrão de agrupamento. Resultados semelhantes foram observados por Sokal & Rohlf (1962). Por sua vez, os valores do CCC revelaram consistências similares nos padrões de agrupamento entre os oito métodos, quando eles foram expressos pelas matrizes quadrado da distância euclidiana padronizada, Mahalanobis e Mahalanobis padronizada.

De maneira geral, o método da ligação média entre grupo, obtido a partir da matriz de dissimilaridade expressa pelas distâncias euclidiana e euclidiana média, proporcionou maior consistência nos padrões de agrupamento. Diferentes cenários formados pela combinação de distintos números de indivíduos (cultivares), números de variáveis, diferentes graus das escalas de medidas, além do agrupamento por diversos métodos, com base em diferentes medidas de dissimilaridade poderiam ser simulados e estudados antes dessas informações serem generalizadas.

**Tabela 4.** Estimativa dos contrastes de médias do coeficiente de correlação cofenética, em valores absolutos, entre os métodos de agrupamento hierárquico ligação simples (vizinho mais próximo, LS), da ligação completa (vizinho mais distante, LC), de Ward, da mediana, da ligação média dentro de grupos (LMDG), da ligação média entre grupos (LMEG), de Gower e do Centroide, em cada matriz de distância, calculados com base em seis caracteres de 13 cultivares de feijão, avaliadas em 9 experimentos.

Método	LC	Ward	Mediana	LMDG	LMEG	Gower	Centroide
Euclidiana ou euclidiana média							
LS	0,088*	0,106*	0,048 <sup>ns</sup>	0,046*	0,028*	0,037 <sup>ns</sup>	0,060*
LC		0,019 <sup>ns</sup>	0,040 <sup>ns</sup>	0,042 <sup>ns</sup>	0,116*	0,051 <sup>ns</sup>	0,028 <sup>ns</sup>
Ward			0,058 <sup>ns</sup>	0,060*	0,135*	0,070 <sup>ns</sup>	0,046 <sup>ns</sup>
Mediana				0,002 <sup>ns</sup>	0,076*	0,011 <sup>ns</sup>	0,012 <sup>ns</sup>
LMDG					0,075*	0,010 <sup>ns</sup>	0,014 <sup>ns</sup>
LMEG						0,065 <sup>ns</sup>	0,088*
Gower							0,023 <sup>ns</sup>
Euclidiana padronizada ou euclidiana média padronizada							
LS	0,002 <sup>ns</sup>	0,006 <sup>ns</sup>	0,056 <sup>ns</sup>	0,042 <sup>ns</sup>	0,088*	0,035 <sup>ns</sup>	0,020 <sup>ns</sup>
LC		0,004 <sup>ns</sup>	0,054 <sup>ns</sup>	0,040 <sup>ns</sup>	0,086*	0,037 <sup>ns</sup>	0,022 <sup>ns</sup>
Ward			0,050 <sup>ns</sup>	0,036 <sup>ns</sup>	0,082*	0,041 <sup>ns</sup>	0,026 <sup>ns</sup>
Mediana				0,014 <sup>ns</sup>	0,032 <sup>ns</sup>	0,092*	0,077*
LMDG					0,046*	0,078 <sup>ns</sup>	0,063*
LMEG						0,124*	0,109*
Gower							0,015 <sup>ns</sup>
Quadrado da distância euclidiana							
LS	0,125*	0,138*	0,079 <sup>ns</sup>	0,079*	0,023 <sup>ns</sup>	0,044 <sup>ns</sup>	0,097*
LC		0,014 <sup>ns</sup>	0,046 <sup>ns</sup>	0,046 <sup>ns</sup>	0,148*	0,081 <sup>ns</sup>	0,028 <sup>ns</sup>
Ward			0,060 <sup>ns</sup>	0,059 <sup>ns</sup>	0,162*	0,094 <sup>ns</sup>	0,042 <sup>ns</sup>
Mediana				0,001 <sup>ns</sup>	0,102*	0,034 <sup>ns</sup>	0,018 <sup>ns</sup>
LMDG					0,102*	0,035 <sup>ns</sup>	0,018 <sup>ns</sup>
LMEG						0,067 <sup>ns</sup>	0,120*
Gower							0,053 <sup>ns</sup>
Quadrado da distância euclidiana padronizada							
LS	0,005 <sup>ns</sup>	0,027 <sup>ns</sup>	0,049 <sup>ns</sup>	0,061 <sup>ns</sup>	0,071*	0,072 <sup>ns</sup>	0,026 <sup>ns</sup>
LC		0,022 <sup>ns</sup>	0,044 <sup>ns</sup>	0,056 <sup>ns</sup>	0,065 <sup>ns</sup>	0,067 <sup>ns</sup>	0,031 <sup>ns</sup>
Ward			0,021 <sup>ns</sup>	0,034 <sup>ns</sup>	0,043 <sup>ns</sup>	0,045 <sup>ns</sup>	0,054 <sup>ns</sup>
Mediana				0,012 <sup>ns</sup>	0,022 <sup>ns</sup>	0,023 <sup>ns</sup>	0,075 <sup>ns</sup>
LMDG					0,009 <sup>ns</sup>	0,011 <sup>ns</sup>	0,087*
LMEG						0,001 <sup>ns</sup>	0,097*
Gower							0,098*
Mahalanobis ou Mahalanobis padronizada							
LS	0,009 <sup>ns</sup>	0,029 <sup>ns</sup>	0,076*	0,043 <sup>ns</sup>	0,064 <sup>ns</sup>	0,063*	0,002 <sup>ns</sup>
LC		0,038 <sup>ns</sup>	0,085*	0,052*	0,073*	0,073 <sup>ns</sup>	0,007 <sup>ns</sup>
Ward			0,048 <sup>ns</sup>	0,014 <sup>ns</sup>	0,035 <sup>ns</sup>	0,035 <sup>ns</sup>	0,031 <sup>ns</sup>
Mediana				0,033 <sup>ns</sup>	0,013 <sup>ns</sup>	0,013 <sup>ns</sup>	0,078 <sup>ns</sup>
LMDG					0,021 <sup>ns</sup>	0,021 <sup>ns</sup>	0,045 <sup>ns</sup>
LMEG						0,000 <sup>ns</sup>	0,066 <sup>ns</sup>
Gower							0,066*

<sup>ns</sup>Não significativo. \*Contraste significativo a 5% de probabilidade, pelo teste t.

## Conclusões

1. Há variabilidade na consistência do padrão de agrupamento de cultivares de feijão (*Phaseolus vulgaris* L.), obtido a partir da combinação de diferentes medidas de dissimilaridade e métodos de agrupamento.

2. Maior consistência nos padrões de agrupamento de cultivares de feijão é verificada com o método da

ligação média entre grupo, obtido a partir da matriz de distância euclidiana.

## Agradecimentos

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico, pela concessão de bolsas de estudo aos autores.

## Referências

- BARROSO, L.P.; ARTES, R. **Análise multivariada**. Lavras: UFLA, 2003. 151p.
- BERTAN, I.; CARVALHO, F.I.F. de; OLIVEIRA, A.C. de; VIEIRA, E.A.; HARTWIG, I.; SILVA, J.A.G. da; SHIMIDT, D.A.M.; VALÉRIO, I.P.; BUSATO, C.C.; RIBEIRO, G. 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.; GONÇALVES-VIDIGAL, M.C.; SCHUELTER, A.R.; VIDIGAL FILHO, P.S.; GONELA, A.; LACANALLO, G.F. Divergência genética em germoplasma de feijoeiro comum coletado no estado do Paraná, Brasil. **Semina: Ciências Agrárias**, v.27, p.547-560, 2006.
- CARGNELUTTI FILHO, A.; RIBEIRO, N.D.; JOST, E. Número necessário de experimentos para a análise de agrupamento de cultivares de feijão. **Ciência Rural**, v.39, p.371-378, 2009a.
- CARGNELUTTI FILHO, A.; RIBEIRO, N.D.; REIS, R.C.P. dos; SOUZA, J.R. de; JOST, E. Comparação de métodos de agrupamento para o estudo da divergência genética em cultivares de feijão. **Ciência Rural**, v.38, p.2138-2145, 2008.
- CARGNELUTTI FILHO, A.; STORCK, L. Estatísticas de avaliação da precisão experimental em ensaios de cultivares de milho. **Pesquisa Agropecuária Brasileira**, v.42, p.17-24, 2007.
- CARGNELUTTI FILHO, A.; STORCK, L. Medidas do grau de precisão experimental em ensaios de competição de cultivares de milho. **Pesquisa Agropecuária Brasileira**, v.44, p.111-117, 2009.
- CARGNELUTTI FILHO, A.; STORCK, L.; RIBEIRO, N.D. Agrupamento de cultivares de feijão em presença e em ausência de multicolinearidade. **Ciência Rural**, v.39, p.2409-2418, 2009b.
- CORRAR, L.J.; PAULO, E.; DIAS FILHO, J.M. **Análise multivariada: para os cursos de administração, ciências contábeis e economia**. São Paulo: Atlas, 2007. 542p.
- CRUZ, C.D. **Programa GENES: análise multivariada e simulação**. Viçosa: UFV, 2006. 175p.
- 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.
- HAIR, J.F.; ANDERSON, R.E.; TATHAM, R.L.; BLACK, W.C. **Análise multivariada de dados**. 5.ed. Porto Alegre: Bookman, 2005. 593p.
- 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.
- MANLY, B.J.F. **Métodos estatísticos multivariados: uma introdução**. 3.ed. Porto Alegre: Bookman, 2008. 229p.
- MINGOTI, S.A. **Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada**. Belo Horizonte: UFMG, 2005. 297p.
- MONTGOMERY, D.C.; PECK, E.A. **Introduction to linear regression analysis**. New York: John Wiley & Sons, 1982. 504p.
- RESENDE, M.D.V. de; DUARTE, J.B. Precisão e controle de qualidade em experimentos de avaliação de cultivares. **Pesquisa Agropecuária Tropical**, v.37, p.182-194, 2007.
- SOKAL, R.R.; ROHLF, F.J. The comparison of dendrograms by objective methods. **Taxon**, v.11, p.33-40, 1962.

---

Recebido em 15 de outubro de 2009 e aprovado em 29 de janeiro de 2010