

Comparações entre medidas de dissimilaridade e estatísticas multivariadas como critérios no direcionamento de hibridações em aveia

Comparisons among dissimilarity measures and multivariate statistics as criterions for directing hybridizations in oat

Giovani Benin¹ Fernando Irajá Félix de Carvalho² Antônio Costa de Oliveira² Volmir Sergio Marchioro³
Claudir Lorencetti³ Andreomar Jose Kurek⁴ José Antonio Gonzalez Silva³ Pedro Jacinto Cruz³
Irineu Hartwig⁵ Douglas André Mallmann Schmidt⁵

RESUMO

Medidas de dissimilaridade são artifícios que os melhoristas de plantas dispõem para identificarem genótipos distantes na obtenção de populações segregantes desejadas. Assim, o presente trabalho foi realizado para estabelecer relação entre as diferentes técnicas multivariadas utilizadas em estimativas de divergência genética. O experimento foi desenvolvido no ano agrícola de 2001, quando doze cultivares de aveia foram testados para sete caracteres agrônomicos, utilizando-se o delineamento experimental de blocos casualizados com quatro repetições. As distâncias Euclidiana e de Mahalanobis evidenciaram baixa correlação (0,529) e, quando utilizadas na confecção de dendrogramas, não possibilitaram a formação de agrupamentos similares. As análises gráficas por componentes principais e variáveis canônicas também apresentaram dispersões distintas. Entretanto, apesar das discrepâncias observadas entre as metodologias analisadas, foi possível a identificação de genótipos dissimilares e com média elevada, que poderão ser utilizados com grande probabilidade de sucesso em hibridações dirigidas em aveia.

Palavras-chave: escolha de genitores, componentes principais, variáveis canônicas, técnicas de agrupamento.

ABSTRACT

Genetic dissimilarity measures are commonly used by plant breeders to identify different genotypes to get

desired segregant populations. This study was proposed to establish the relationship between different multivariate techniques to estimate divergence. The experiment was performed during the growing season of 2001, where twelve oat cultivars were tested for seven agronomic traits, using the random blocks experimental design with four replications. Euclidian and Mahalanobis distances showed low correlation (0.529) and when used to build dendrograms did not show similar clustering. The graphic techniques analysis through principal components and canonical variables also showed distinct spreading patterns. However, in spite of the observed discrepancies among the methodology analyzed, it was possible to recognize dissimilar genotypes with high average that can be used with large success probability in selected artificial hybridizations in oats.

Key words: choice of parents, principal components, canonical variables, cluster analyses.

INTRODUÇÃO

A quantificação da dissimilaridade genética é um dos mais importantes parâmetros estimados pelos melhoristas de plantas, principalmente quando o objetivo for a obtenção de segregantes transgressivos e populações de ampla variabilidade genética.

¹Engenheiro Agrônomo, estudante do Curso de Doutorado em Agronomia (Fitomelhoramento) da Universidade Federal de Pelotas (UFPel), Pelotas-RS. Bolsista do CNPq. E-mail: bening@ufpel.tche.br - gibn@starmedia.com. Autor para correspondência.

²Engenheiro Agrônomo, PhD, Professor do Departamento de Fitotecnia, Faculdade de Agronomia Eliseu Maciel, UFPel, Campus Universitário, CP 354, 96001-970, Pelotas-RS. Email: carvalho@ufpel.tche.br

³Engenheiro Agrônomo, estudante do Curso de Pós-graduação em Agronomia (Fitomelhoramento), UFPel.

⁴Engenheiro Agrônomo, MSc, Professor do Curso de Agronomia, CEFET-PR/UNED-PB

⁵Estudante do Curso de Agronomia, UFPel, bolsista do CNPq.

Neste sentido, os métodos multivariados, em que diversos caracteres podem ser dimensionados simultaneamente, têm oferecido contribuições efetivas na identificação de genótipos para serem utilizados em programas de melhoramento genético de várias culturas (SANTOS et al., 2000), inclusive com indicação de caracteres mais representativos para a obtenção de populações geneticamente diferentes.

Dentre as várias técnicas multivariadas que podem ser aplicadas, CRUZ & REGAZZI (1997), citam as análises por componentes principais e variáveis canônicas e os métodos aglomerativos, que utilizam a distância Euclidiana ou a distância generalizada de Mahalanobis como medidas de dissimilaridade. Desta forma, é de fundamental importância escolher dentre estes, aqueles que permitam inferências mais efetivas das constituições genéticas avaliadas.

MALUF & FERREIRA (1983), comparando a distância Euclidiana e a distância generalizada de Mahalanobis na avaliação de dez cultivares de feijão de corda (*Phaseolus vulgaris* L.), encontraram baixa correlação entre as duas metodologias (0,273) e observaram similaridade na detecção de diferenças intergrupos e discrepância nas diferenças menores (intragrupos) pela análise de conglomeração.

Desta forma, o presente trabalho foi realizado com o objetivo de obter estimativas da dissimilaridade genética entre 12 cultivares de aveia cultivada hexaplóide, com vistas à hibridação; verificar a associação entre as técnicas de variáveis canônicas e componentes principais; estabelecer correlação entre a distância Euclidiana e de Mahalanobis, e verificar suas concordâncias em análises de conglomeração pela técnica vizinho mais próximo.

MATERIAL E MÉTODOS

O experimento foi instalado em julho de 2001 no Sistema de Plantio Convencional no campo experimental do Setor de Fitomelhoramento da Faculdade de Agronomia Eliseu Maciel - UFPel, município de Capão do Leão, em solo classificado como Podzólico Vermelho-amarelo da área de Mapeamento de Pelotas. O município está situado a 31° 52' 00'' de latitude Sul e 52° 21' 24'' de longitude W. Gr; a uma altitude de 13,24m. O clima, segundo a classificação de Koeppen, é do tipo Cfa, com uma precipitação pluvial média anual de 1280 mm (SILVA, 1999).

Os tratamentos compreenderam doze cultivares de aveia cultivada hexaplóide em uso no Brasil (Tabela 1), arranjados em delineamento experimental de blocos casualizados com quatro repetições, com parcelas de cinco linhas com cinco

metros de comprimento, espaçadas 0,2 metros entre si. O rendimento de grãos foi determinado nas três linhas centrais da parcela, com exclusão das duas laterais (bordaduras). Não foi realizado tratamento para o controle de moléstias, a fim de se verificar a ocorrência de patógenos nos diferentes genótipos.

Os caracteres avaliados foram: rendimento de grãos desaristados em kg.ha⁻¹ (RGD), corrigidos para 13% umidade, peso de mil grãos em gramas (PMG), peso do hectolitro em quilos/100 litros de grãos (PH), estatura de planta (EP) medida em cm da base do colmo ao topo da panícula sem a inclusão das aristas, dias da emergência à maturação (DEM), dias da floração a maturação (DFM) e reação à moléstia ferrugem da folha (FF) quantificada em porcentagem de pústula por unidade de área foliar.

Os dados foram submetidos à análise de variância univariada, pela qual foram obtidos o arquivo de médias e a matriz de variância e covariâncias residuais, sendo as médias comparadas pelo teste de SCOTT & KNOTT (1974). Este é um método utilizado como uma alternativa em procedimentos de comparação múltipla, com a característica de não apresentar ambigüidade nos resultados (SILVA et al., 1999). Foram aplicadas as análises de componentes principais e variáveis canônicas e o método aglomerativo vizinho mais próximo (CRUZ & REGAZZI, 1997), utilizando a distância Euclidiana e distância generalizada de Mahalanobis como medidas de dissimilaridade. A associação entre as medidas de dissimilaridade foi estimada pela correlação de Spearman (STEEL & TORRIE, 1980). As análises foram realizadas com emprego do programa computacional GENES (CRUZ, 2001).

RESULTADOS E DISCUSSÃO

Os resumos das análises de variância univariada inseridos na tabela 1 apresentaram diferenças significativas a 5% de significância em relação a todos os caracteres avaliados sendo, portanto, evidência da presença de dissimilaridade entre os genótipos testados.

Na tabela 1, também estão incluídas as médias dos caracteres analisados, comparadas pelo teste de Scott e Knott. As diferenças entre genótipos demonstram presença de variabilidade para os caracteres analisados. O caráter rendimento de grãos desaristados (RGD) variou de 619 a 2253kg.ha⁻¹ com os cultivares UFRGS 17 e URS 21, respectivamente. A elevada suscetibilidade à ferrugem da folha apresentada por alguns genótipos determinou maior número de diferenças significativas entre médias, fator

Tabela 1 – Comparação de médias e resumo das análises de variância univariada dos caracteres rendimento de grãos desaristados (RGD), peso de mil grãos (PMG), peso do hectolitro (PH), estatura de planta (EP), dias da emergência a maturação (DEM), dias da floração a maturação (DFM) e reação à ferrugem da folha (FF) visando estimar a dissimilaridade genética entre doze cultivares de aveia. Pelotas-RS, 2002.

Genótipos	RGD (kg.ha ⁻¹)	PMG (g)	PH (kg.hl ⁻¹)	EP (cm)	DEM (dias)	DFM (dias)	FF (%)
UPF 16	924 b*	16 b	26 d	85 a	98 b	31 a	68 b
UPF 17	1957 a	25 a	43 a	114 a	107 a	36 a	4 d
UPF 18	1957 a	25 a	43 a	117 a	107 a	36 a	4 d
UPF 19	642 c	18 b	27 d	101 a	98 b	30 a	70 b
UFRGS 14	626 c	17 b	30 c	95 a	104 a	28 a	66 b
UFRGS 15	644 c	19 b	34 b	66 a	103 a	25 a	68 b
UFRGS 17	619 c	16 b	32 b	98 a	104 a	32 a	56 c
UFRGS 18	1044 b	29 a	43 a	93 a	107 a	31 a	48 c
URS 20	890 b	24 a	40 a	103 a	105 a	35 a	18 d
URS 21	2253 a	28 a	46 a	106 a	97 b	31 a	5 d
OR 3	673 b	27 a	35 b	111 a	107 a	34 a	41 c
IAC 7	951 b	21 b	36 b	95 a	98 b	35 a	86 a
QM Genótipos	1.437.693,66**	90,75**	181,45**	775,41**	66,42**	471,93**	3.533,8**
QM Resíduo	45.440,14	5,52	3,43	53,70	4,86	6,51	90,18
Médias	1099	22	36	99	103	32	44
C V	19	10	5	7	2	7	21

*Médias seguidas de letras distintas na coluna diferem significativamente em 5% de significância pelo teste de Scott-Knott.

**Significativo em 5% de significância.

este que provavelmente afetou o caráter peso hectolítrico e peso médio de grãos, possibilitando a formação de grupos distintos.

Na tabela 2, estão inseridas as variâncias (autovalores), as variâncias percentuais e as variâncias acumuladas dos três primeiros componentes principais e das três primeiras variáveis canônicas. Conforme CRUZ & REGAZZI (1997), para uma interpretação satisfatória da variabilidade manifestada entre genótipos é necessário que os dois primeiros componentes principais e as duas primeiras variáveis canônicas permitam estimativas mínima de 80% da variação total contida no conjunto de caracteres analisados, as quais, neste trabalho, foram de apenas 76,7 e 75,5%, respectivamente (Tabela 2). Entretanto, pelo fato de as variâncias terem sido similares, fator de fundamental importância para a análise comparativa destas duas técnicas, foi possível utilizar seus escores em gráficos de dispersão. A dispersão gráfica dos escores pela técnica de componentes principais e variáveis canônicas (Figura 1) possibilitou a formação de três e quatro grupos distintos, respectivamente. No grupo I de ambas as técnicas empregadas, a única constituição genética não coincidente foi o cultivar UFRGS 15, que distanciou em relação aos demais pela

Tabela 2 – Variâncias, variância percentuais e variâncias acumuladas (%) em relação aos três primeiros componentes principais e às três primeiras variáveis canônicas visando estimar a dissimilaridade entre doze cultivares de aveia. Pelotas-RS, 2002.

Genótipos	Componentes Principais			Variáveis canônicas		
	1ª	2ª	3ª	1ª	2ª	3ª
UPF 16	13,2	17,5	11,1	17,6	-34,4	49,1
UPF 17	18,4	18,7	11,3	31,2	-36,3	53,2
UPF 18	18,5	18,6	11,4	31,1	-35,7	53,7
UPF 19	13,5	17,6	11,6	17,2	-31,6	52,4
UFRGS 14	13,5	19,2	11,2	18,8	-38,5	56,2
UFRGS 15	12,8	19,8	9,2	21,0	-47,3	51,4
UFRGS 17	14,3	18,9	11,8	19,8	-38,9	53,3
UFRGS 18	16,3	20,0	10,0	27,7	-42,9	52,0
URS 20	16,7	19,2	11,4	26,7	-38,8	52,0
URS 21	17,8	17,0	9,1	32,5	-35,1	47,7
OR 3	16,3	19,6	12,0	24,0	-34,3	56,1
IAC 7	14,6	17,5	11,2	20,7	-37,5	44,5
Variância	4,3	1,0	0,9	31,5	17,5	11,3
Variância (%)	62,2	14,4	5,5	48,5	27,0	17,3
Variância acum. (%)	62,2	76,7	89,7	48,5	75,5	92,8

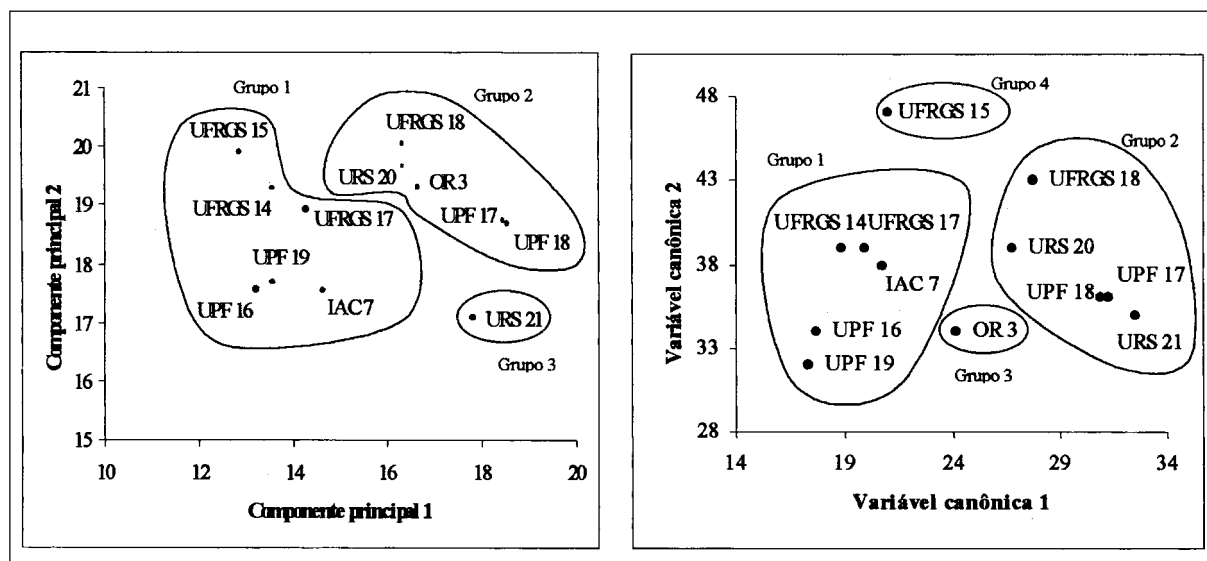


Figura 1 – Dispersão gráfica dos escores de doze cultivares de aveia em relação aos dois primeiros componentes principais e as duas primeiras variáveis canônicas. Pelotas, RS, 2002.

análise de variáveis canônicas. Da mesma forma, os cultivares URS 21 e OR 3 não coincidiram no grupo 2 das duas análises. A técnica de componentes principais identificou os cultivares UFRGS 15, UFRGS 18, IAC 7, UPF 16 e URS 21 como os mais dissimilares. Pela análise de variáveis canônicas, os mais dissimilares foram os cultivares UFRGS 15, UFRGS 18, URS 21, UPF 19 e OR 3. Entretanto, ambas as técnicas empregadas possibilitaram identificar os cultivares UFRGS 15, UFRGS 18 e URS 21 dentre os de maior dissimilaridade. Além de apresentarem agrupamentos não coincidentes, a técnica de variáveis canônicas permitiu uma maior distribuição espacial dos genótipos, formando grupos mais característicos, quando comparados com a técnica de componentes principais.

Tanto a distância Euclidiana quanto a distância de Mahalanobis (Tabela 3) identificaram as combinações UPF 17 - UPF 18, UFRGS 14 - UFRGS 17 e UPF 16 - UPF 19 como os pares menos dissimilares entre si. Entre os seis pares mais dissimilares, a análise das distâncias Euclidianas identificou as combinações UFRGS 15 - UPF 18, UFRGS 15 - UPF 17, UFRGS 15 - URS 21, UPF 16 - UPF 17, UPF 18 - UPF 16, UPF 18 - UPF 19, enquanto as distâncias de Mahalanobis identificaram as seis maiores distâncias entre as combinações UFRGS 15 - URS 21, UFRGS 14 - URS 21, UPF 19 - URS 21, UPF 19 - UFRGS 15, UPF 19 - UFRGS 18 e UPF 18 - UFRGS 15. Apesar de serem linhas irmãs do mesmo cruzamento, os cultivares UPF 16 e UPF 17 foram

identificados entre os seis pares mais dissimilares pela distância Euclidiana, provavelmente, devido à suscetibilidade à ferrugem da folha, que determinou menor RGD, PMG, PH, EP, DEM e DFM para o cultivar UPF 16. Apenas dois pares (UFRGS 15 - UPF 18 e UFRGS 15 - URS 21) coincidiram dentre os seis mais dissimilares em ambas as técnicas testadas, o que reflete a baixa correlação encontrada entre as distâncias Euclidianas e de Mahalanobis (coeficiente de correlação de Spearman = 0,529).

A discrepância entre a distância Euclidiana e distância generalizada de Mahalanobis pode ser mais bem visualizada através da análise de seus correspondentes dendrogramas, oriundos da análise conglomeração, inseridos na figura 2. O dendrograma A, que utilizou a distância Euclidiana como medida de dissimilaridade formou sete grupos: I - UPF 17 e UPF 18; II - UPF 18 e URS 20; III e IV - constituídos exclusivamente pelos cultivares URS 21 e OR 3, respectivamente; V - UFRGS 14 e UFRGS 17 (Subgrupo V¹), bem como UPF 16 e UPF 19 (Subgrupo V²) e; VI e VII - constituídos exclusivamente pelos cultivares IAC 7 e UFRGS 15, respectivamente. O dendrograma B, que utilizou a distância generalizada de Mahalanobis como medida de dissimilaridade, também formou sete grupos: I - UPF 17 e UPF 18; II - URS 20 e OR 3; III - UFRGS 18; IV - UFRGS 14 e UFRGS 17 (Subgrupo IV¹), bem como UPF 16 e UPF 19 (Subgrupo IV²) e; V, VI e VII - constituídos exclusivamente pelos cultivares IAC 7, UFRGS 15 e URS 21, respectivamente. A menor

Tabela 3 – Medidas de dissimilaridade entre doze cultivares de aveia com base na distância Euclidiana (lado superior da diagonal) e na estatística D² de Mahalanobis (lado inferior da diagonal). Pelotas-RS, 2002.

Genótipos	UPF 16	UPF 17	UPF 18	UPF 19	UFRGS 14	UFRGS 15	UFRGS 17	UFRGS 18	URS 20	URS 21	OR 3	IAC 7
UPF 16	-	5,4	5,4	1,3	2,1	3,0	2,0	4,3	3,9	5,1	4,1	2,2
UPF 17	208	-	0,2	5,1	4,9	6,0	4,2	3,1	2,1	2,9	2,8	4,4
UPF 18	211	0,7	-	5,2	5,0	6,2	4,3	3,2	2,2	3,0	2,9	4,4
UPF 19	24	220	215	-	1,8	3,5	1,8	4,1	3,6	4,9	3,4	2,1
UFRGS 14	75	170	169	66	-	2,3	1,2	3,4	3,4	5,2	3,2	2,9
UFRGS 15	182	230	244	261	107	-	3,1	3,8	4,5	5,7	4,6	3,9
UFRGS 17	48	137	139	64	13	78	-	3,2	2,5	4,8	2,6	2,4
UFRGS 18	200	69	77	241	127	79	90	-	2,0	3,6	2,1	3,3
URS 20	121	40	45	149	98	116	55	32	-	3,4	1,4	3,1
URS 21	228	40	45	268	276	295	212	120	81	-	4,1	4,2
OR 3	111	79	74	74	61	217	59	104	54	164	-	3,1
IAC 7	62	199	208	118	154	162	89	137	110	179	161	-

distância detectada por ambas as metodologias entre os cultivares UPF 17 e UPF 18 corroboram com as técnicas de componentes principais e variáveis canônicas (Figura 1) e com as menores distâncias das matrizes de dissimilaridade (Tabela 3). Entretanto, os dendrogramas A e B identificaram os cultivares UFRGS 15 e URS 21, respectivamente, como os mais dissimilares, evidenciando a inexistência de concordância entre a distância Euclidiana e de Mahalanobis na formação de agrupamentos, discordando de MALUF & FERREIRA (1983). Nos Sugrupos IV¹-IV² e V¹-V², que reuniram os mesmos cultivares (UFRGS 14 - UFRGS 17, UPF 16 - UPF 19) por ambas as metodologias, as maiores diferenças intragrupos foram observadas no dendrograma baseado na distância Euclidiana, estando de acordo com os autores anteriormente citados.

Analisando as divergências intra e intergrupos (Figura 2), foi verificado que as seis maiores distâncias Euclidianas e distâncias de Mahalanobis (Tabela 3) corresponderam a combinações intergrupais. Isto parece indicar uma concordância entre os métodos na detecção de grandes divergências (intergrupais). Contudo, nem sempre as duas técnicas empregadas estabeleceram a mesma formação de grupos e não concordaram na identificação de qual combinação intragrupal específica, dentre as muitas possíveis, é a mais divergente.

A utilização de uma ou outra técnica vai ao encontro dos objetivos do pesquisador e da situação em que os experimentos foram conduzidos. CRUZ (1990), recomenda a utilização

da distância Euclidiana e análise de componentes principais em experimentos que não contemplam repetição, por ser difícil a quantificação da influência do ambiente que atua sobre as constituições genéticas. Estas técnicas são recomendadas para a avaliação de acessos em bancos de germoplasma, onde o número considerável de genótipos inviabiliza a utilização de delineamento experimental. Entretanto, a distância generalizada de Mahalanobis e variáveis canônicas podem ser estimadas apenas quando o delineamento experimental contempla repetição, permitindo a quantificação dos efeitos do ambiente sobre as constituições genéticas. Desta forma, fica explícito que a distância Euclidiana e componentes principais possibilitaram inferências sobre a dissimilaridade fenotípica, enquanto a distância generalizada de Mahalanobis e variáveis canônicas quantificaram a dissimilaridade genética, fator este provavelmente responsável pelas diferenças observadas entre as técnicas avaliadas.

Em um programa de melhoramento é necessário que os genitores tenham ampla divergência genética entre si e médias elevadas para os caracteres alvo de melhoramento. Neste sentido, o conjunto das metodologias avaliadas credenciam os cultivares UFRGS 18, URS 20, URS 21, UPF 17 e UPF 18 como os genitores mais ajustados aos objetivos existentes nos programas de melhoramento genético. Assim, poderão ser consideradas as hibridações URS 21 x UPF 17, URS 21 x UPF 18, UPF 18 x UFRGS 18 e URS 21 x UFRGS 18 como as mais promissoras na obtenção de populações segregantes desejadas.

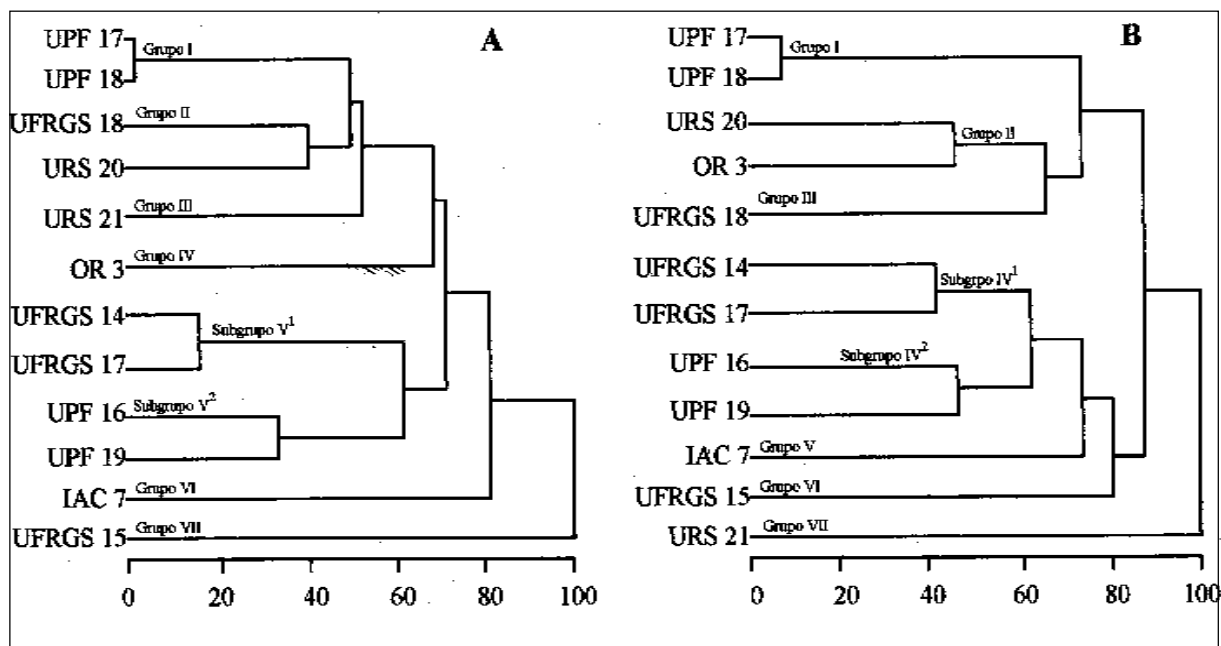


Figura 2 – Dendrogramas resultantes da análise de conglomeramento de doze cultivares de aveia, obtidos pelo método hierárquico vizinho mais próximo, utilizando a distância Euclidiana (A) e a distância de Mahalanobis (B) como medidas de dissimilaridade. Pelotas - RS, 2002.

CONCLUSÕES

A baixa correlação observada entre a distância Euclidiana e de Mahalanobis e as suas inconsistências quando utilizadas no estabelecimento de agrupamentos caracterizam estas duas estimativas como medidas de dissimilaridade diferentes. Da mesma forma, a técnica de componentes principais e variáveis canônicas evidenciaram dispersões gráficas não coincidentes, ficando evidente a diferença entre as duas metodologias.

REFERÊNCIAS BIBLIOGRÁFICAS

- CRUZ, C.D. **Programa genes: aplicativo computacional em genética e estatística**. Viçosa : UFV, 2001. 648p.
- CRUZ, C.D. **Aplicação de algumas técnicas multivariadas no melhoramento de plantas**. 1990. 188f. Tese (Doutorado em Genética e Melhoramento) - Escola Superior de Agronomia Luiz de Queiroz.
- CRUZ, C.D.; REGAZZI, A.J. **Métodos biométricos aplicados ao melhoramento genético**. Viçosa : UFV, 1997. 390p.

MALUF, W.R.; FERREIRA, P.E. Análise multivariada da divergência genética em feijão-vagem (*Phaseolus vulgaris* L.). **Horticultura Brasileira**, Brasília v.1, n.2, p.31-34, 1983.

SANTOS, R.C. et al. Classificação de genótipos de amendoim baseada nos descritores agromorfológicos e isoenzimáticos. **Ciência Rural**, Santa Maria, v.30, n.1, p.55-59, 2000.

SILVA, S.A. **Estimativa da herança do caráter "stay-green" em genótipos de trigo hexaplóide**. 1999. 44f. Dissertação (Mestrado em Agronomia) - Faculdade de Agronomia, UFPel.

SILVA, E.C. da; FERREIRA, D.F.; BEARZOTI, E. Avaliação do poder e taxas de erro tipo I do teste de Scott-Knott por meio do método de Monte Carlo. **Ciência Agrotécnica**, Lavras, v.23, n.3, p.687-696, 1999.

SCOTT, A. J.; KNOTT, M.A. Cluster Analysis methods for grouping means in the analysis of variance. **Biometrics**, Washington, v.30 p.507-512, 1974.

STEEL, R.G.D.; TORRIE, J.H. **Principles and procedures of statistics: a biometric approach**. 2.ed. New York : McGraw-Hill, 1980. 633p.