



## INTEGRAÇÃO DE DESCRITORES DE FOLHA QUANTITATIVOS E MULTICATEGÓRICOS NA DIVERGÊNCIA GENÉTICA EM CAFÉ ARÁBICA.

**Sabrina Batista de Oliveira, Tiago de Oliveira Godinho, Sarah Ola Moreira.**

Instituto Capixaba de Pesquisa, Assistência Técnica e Extensão Rural – Incaper, Centro Regional de Desenvolvimento Rural Centro Serrano. Rodovia BR 262, Km 94, Fazenda do Estado, 29375-000, Venda Nova do Imigrante-ES, Brasil, [sabrina.bo@live.com](mailto:sabrina.bo@live.com), [godinhoto@hotmail.com](mailto:godinhoto@hotmail.com), [sarah.ola@gmail.com](mailto:sarah.ola@gmail.com).

**Resumo** – A estatística multivariada e a análise conjunta de descritores possibilitam quantificar de forma mais precisa a diversidade genética, auxiliando na definição das estratégias a serem usadas em programas de melhoramento. O objetivo desse trabalho foi integrar variáveis quantitativas e multicategóricas na avaliação da divergência genética de café arábica, utilizando diferentes métodos de agrupamento. Foram utilizados 18 genótipos de café arábica e dez descritores multicategóricos e três quantitativos. Os descritores quantitativos foram convertidos em multicategóricos pela divisão equitativa da amplitude. Foi obtida a matriz de dissimilaridade e os agrupamentos pelos métodos hierárquico e de Tocher. A correlação cofenética (CCC) foi usada para definição de melhor método de agrupamento hierárquico. O agrupamento pelo método da ligação média entre os grupos (UPGMA) foi o de maior CCC. A diversidade genética para caracteres de folha de café arábica é baixa devido ao uso dos mesmos genitores para a obtenção de cultivares. Os descritores quantitativos convertidos em multicategóricos foram eficientes para a ampliar a dissimilaridade entre os genótipos avaliados.

**Palavras-chave:** *Coffea arabica*; análise multivariada; métodos hierárquicos; coeficiente de correlação cofenético.

**Área do Conhecimento:** Engenharia Agrônoma.

### Introdução

O sucesso da cafeicultura brasileira deve-se, em grande parte, ao desenvolvimento de cultivares mais adaptadas às diferentes condições de cultivo, de alto padrão genético e elevado potencial de produção. Para que maiores avanços nessa área sejam atingidos é necessário ampliar os estudos de diversidade genética, que são essenciais no planejamento de programas e na definição de estratégias de trabalhos (FREITAS et al., 2007; GUEDES et al., 2013). Assim, a caracterização e o estudo da divergência genética entre os acessos de cafeeiro contribuem para o aumento da eficiência na seleção de genitores a serem utilizados em programas de melhoramento genético (SILVA et al., 2013).

A divergência genética, simplificada, pode ser definida como a distância genética entre as populações ou indivíduos, avaliada por meio de caracteres quantitativos, multicategóricos ou binários. No entanto, a geração de um grande número de dados de diferentes categorias (qualitativos e quantitativos) pode ser um fator que dificulta a análise e a interpretação dos resultados, muitas vezes não permitindo a completa distinção entre os acessos. Ao longo dos anos, houve um aprimoramento das técnicas multivariadas para quantificação da divergência genética, permitindo considerar simultaneamente inúmeras características (MOURA et al., 2010; SUDRÉ et al., 2010). Para o cafeeiro, a conversão de todas as variáveis em um único padrão, por exemplo, transformação das variáveis quantitativas em multicategóricas, e a integração delas em uma única análise, favoreceu a melhor visualização da divergência genética, com maior número de grupos formados (SILVA et al., 2013).

As estimativas de dissimilaridade, apesar de atenderem aos objetivos do melhorista, por quantificarem e informarem o grau diferença entre dois genótipos quaisquer, não são viáveis para o reconhecimento dos grupos de indivíduos semelhantes, especialmente quando se avalia muitos genótipos. Assim, com uso de técnicas de agrupamento, pode-se classificar os genótipos em grupos de forma que exista homogeneidade dentro e heterogeneidade entre os grupos (CRUZ; CARNEIRO; REGAZZI, 2014). Dentre os métodos de agrupamento existentes, a escolha pelo melhor deve-se basear na melhor correlação cofenética, que é a relação entre a matriz de dissimilaridade e o dendrograma gerado, ou seja, entre as distâncias reais entre os acessos e as distâncias representadas

graficamente, sujeitas ao acúmulo de erro (SOKAL; ROHLF, 1962). Diante do exposto, o objetivo desse trabalho foi avaliar a diversidade genética para caracteres de folha de café arábica diferentes métodos de agrupamento e integrando dados quantitativos em multicategóricos.

## Metodologia

O experimento foi instalado na Fazenda Experimental Mendes da Fonseca do Instituto Capixaba de Pesquisa, Assistência Técnica e Extensão Rural (Incaper), localizado no município de Domingos Martins, ES (20,383° S; 41,050° W), a 950 m de altitude. Segundo a classificação de Köppen, o clima da região é temperado com verão ameno (Cfb), com temperatura média do ar mensal de 18,2°C e precipitação anual de 1.263 mm (ALVARES et al., 2013).

Foram avaliadas 18 cultivares de café arábica (*Coffea arabica*): 1 = Catucaí Amarelo 2 SL; 2 = Catucaí Amarelo 24/137; 3 = Catucaí Amarelo 20/15 CV 479; 4 = Catucaí Vermelho 20/15 CV 479; 5 = Sabiá 398; 6 = Pauma II; 7 = Acauã; 8 = Oeiras MG 6851; 9 = Catiguá MG 01; 10 = Sacramento MG 1; 11 = Catiguá MG 02; 12 = Araponga MG 1; 13 = Paraíso H 419-3-3-7-16-4-1; 14 = Pau Brasil MG; 15 = material crioulo (Cassine); 16 = Obatã; 17 = Iapar 59; e, 18 = IPR 105. O plantio foi efetuado em abril de 2006, em delineamento de blocos ao acaso com quatro repetições e 10 plantas por parcela. O espaçamento utilizado foi de 2,0 m entre linhas por 0,9 m entre plantas. A adubação, manejo e demais recomendações para a cultura foram realizadas conforme Sakiyama et al. (2015).

Para as análises de diversidade genética foram utilizados 13 descritores de folha estabelecidos para *Coffea* spp. e *Psilanthus* spp. pelo *International Plant Genetic Resources Institute* - IPGRI (1996):

- Comprimento do limbo: definido com a distância entre o ponto de inserção do pecíolo e o ápice da folha no limbo foliar, medido em cm;
- Largura do limbo: determinado na porção mais extensa da folha, medido em cm;
- Comprimento do pecíolo: medido da bainha à lâmina foliar, em cm;
- Forma do limbo: 1 = obovada; 2 = ovada; 3 = elíptica; 4 = lanceolada;
- Forma do ápice da folha: 1 = arredondado; 2 = obtuso; 3 = agudo; 4 = acuminado; 5 = aculeado; 6 = espatulada;
- Cor do pecíolo: 1 = verde; 2 = marrom escuro; 3 = amarelo;
- Cor da folha jovem: 1 = verde; 2 = bronze; 3 = verde e bronze;
- Cor da folha adulta: 1 = verde claro; 2 = verde escuro;
- Intensidade da nervura secundária: 1 = fraca; 3 = média; 5 = forte;
- Ondulação das bordas das folhas: 0 = ausente; 1 = presente;
- Intensidade da ondulação da borda da folha: 1 = fraca; 3 = média; 5 = forte;
- Domácia: 0 = ausente; 1 = parcialmente desenvolvida; 2 = bem desenvolvida; e,
- Pubescência na domácia: 0 = ausente; 1 = presente.

Para a avaliação foram coletadas amostras de três folhas completamente expandida em seis plantas de cada cultivar. As folhas foram coletadas no quarto par de folhas, a partir da porção distal do ramo, a uma altura aproximada de 1m do solo.

Os descritores quantitativos (comprimento e largura do limbo e comprimento do pecíolo) foram transformados em multicategóricos por meio da codificação dos dados quantitativos em um número pré-definido de classes, utilizando a divisão equitativa da amplitude (DEA). Com isso, foi possível realizar a análise da divergência genética envolvendo todas as características avaliadas, independentemente de sua natureza (SILVA et al., 2013). Para essas três variáveis, foram formadas quatro classes.

Após a conversão dos dados, para a obtenção da matriz de dissimilaridade foi utilizado o complemento aritmético do índice de coincidência simples, com base na moda de cada variável por acesso. A partir dessa matriz foi realizado o agrupamento dos indivíduos pelos métodos hierárquicos e de otimização (Tocher). Para definição de melhor método de agrupamento hierárquico foi utilizado o coeficiente de correlação cofenético e a sua significância pelo teste de Mantel. A análise dos dados foi realizada com auxílio do Programa Genes (CRUZ, 2013).

## Resultados

Os agrupamentos obtidos tiveram coeficientes de correlação cofenéticos (CCC) variando entre 53,80% e 73,76% para os métodos da Ligação completa - vizinho mais distante e o de ligação média entre os grupos (UPGMA). No entanto, todos os valores de CCC foram significativos pelo teste de Mantel. Os métodos Ligação simples - Vizinho mais próximo, Ligação Média dentro de Grupo, WPGMC e UPGMC tiveram CCC bem próximos, entre 60 e 63%, indicando a presença de similaridade entre eles (Tabela 1).

Tabela 1 – Coeficientes de correlação cofenético de diferentes métodos de agrupamento hierárquicos obtidos a partir de 13 descritores multicategóricos em 18 cultivares de café arábica.

Método de agrupamento	Coeficiente de correlação cofenético (%)
Ligação simples - Vizinho mais próximo	62,71**
Ligação completa - Vizinho mais distante	53,80**
Método de Ward	58,77**
Método da Mediana (WPGMA)	72,66**
Ligação Média dentro de Grupo	61,85**
Ligação Média entre Grupo (UPGMA)	73,76**
Método de Gower (WPGMC)	60,11**
Método do Centróide (UPGMC)	60,45**

\*\* Significativo pelo teste de Mantel a 1% de probabilidade.

Fonte: os autores

Por ter maior CCC entre os métodos analisados, os genótipos de café arábica foram agrupados pelo método UPGMA. Observou-se a formação de apenas três grupos, sendo um deles (Grupo I) formado por 16 das 18 cultivares de café avaliadas (88,89%). O Grupo II foi formado apenas pelo 'Paraíso H 419-3-3-7-16-4-1' e o Grupo III pela cultivar Sacramento MG1 (Figura 1). O material genético Paraíso H 419-3-3-7-16-4-1 foi o único genótipo a ter intensidade da nervura secundária fraca associado a cor verde e bronze na folha jovem. A 'Sacramento MG1' foi o único material genético avaliado com intensidade de ondulação da borda da folha fraca e intensidade da nervura secundária forte.

Dentre os descritores utilizados, a ondulação da borda da folha, a domácia, a pubescência da domácia e a cor do pecíolo não tiveram diferenças entre os materiais avaliados. As maiores variabilidades foram obtidas para os descritores quantitativos convertidos em multicategóricos - comprimento e largura do limbo e comprimento do pecíolo.

O agrupamento dos genótipos pelo método de Tocher se diferenciou do agrupamento pelo método hierárquico, havendo a formação de quatro grupos, permitindo uma maior diferenciação entre os genótipos: o grupo I com 13 genótipos, o Grupo II e III com dois materiais genéticos e o Grupo IV formado apenas pelo 'Paraíso H 419-3-3-7-16-4-1', assim como no UPGMA (Tabela 2).

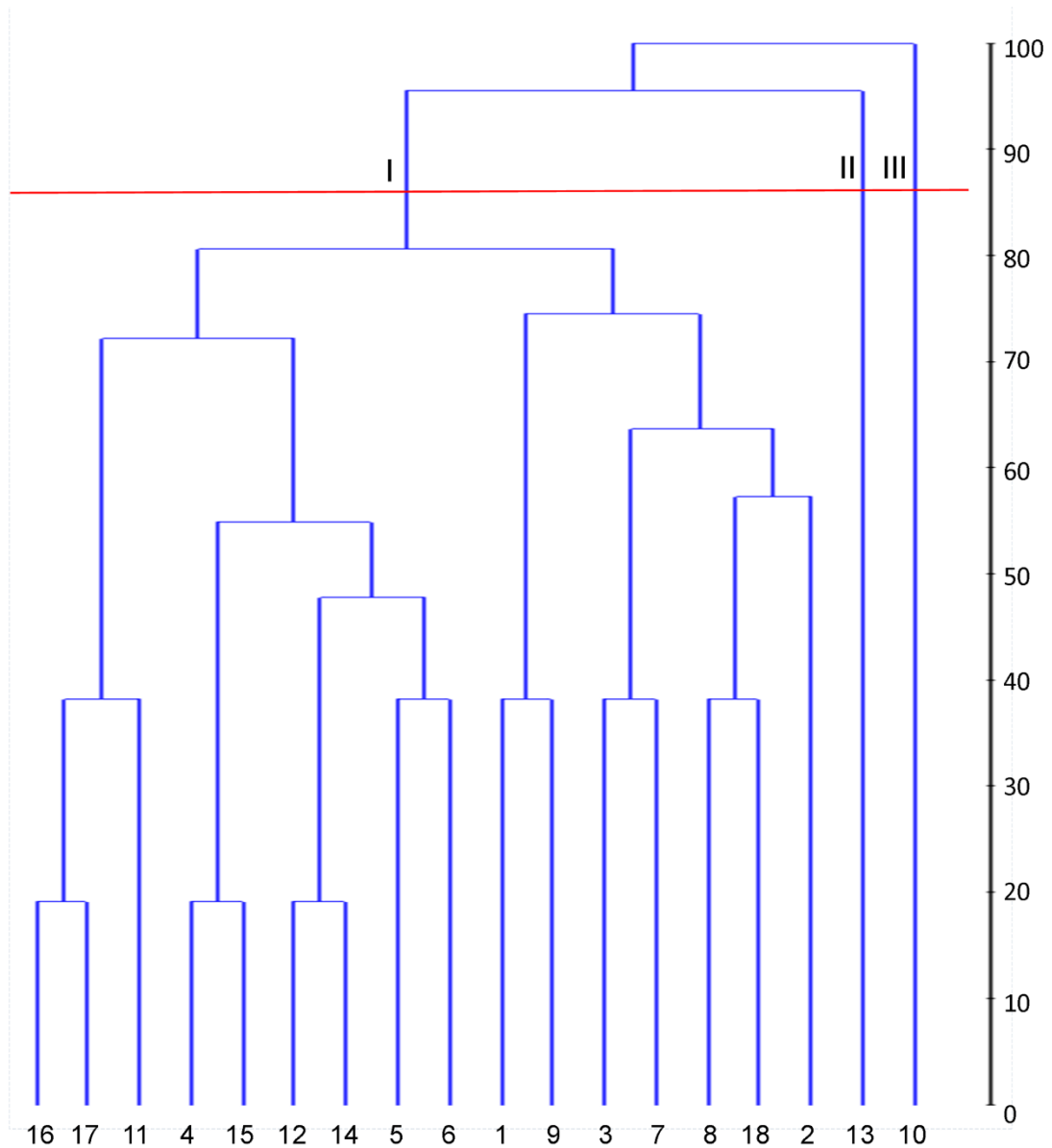
## Discussão

A correlação cofenética é uma análise que estabelece uma relação entre a matriz de dissimilaridade e o dendrograma gerado a partir dessa matriz. Assim, na concepção de Sokal; Rohlf (1962), valores de correlação cofenética superiores a 0,80 indicam bom ajuste entre as matrizes originais de distância e as derivadas das distâncias gráficas. Os CCC obtidos para os diferentes métodos de agrupamento não atingiram os 80% indicados por Sokal; Rohlf (1962), no entanto, todos eles foram significativos pelo teste de Mantel. Assim, é possível realizar inferências sobre a divergência genética entre os materiais avaliados por meio da avaliação do dendrograma.

Dentre os métodos de agrupamento hierárquicos avaliados, o método da Ligação Média entre Grupo (UPGMA) teve melhor consistência (CCC = 73,76%). Segundo Cargnelutti Filho et al. (2010), dentre os diferentes métodos de agrupamento, o UPGMA permite a formação de grupos com maior aderência à matriz de dissimilaridade. Segundo Cruz; Carneiro; Regazzi (2014), o UPGMA considera as médias aritméticas das medidas de dissimilaridade, o que evita quantificar a divergência genética por valores

extremos entre os genótipos. Por isso, dentre os métodos hierárquicos, o UPGMA é o mais utilizado em trabalhos de avaliação da diversidade genética.

Figura 1 - Dendrograma de dissimilaridade genética entre 18 cultivares de café arábica obtido pelo método hierárquico de ligação média entre grupos (UPGMA), com base em 13 descritores multicategóricos.



Fonte: os autores

Tabela 2 – Agrupamento de 18 cultivares de café arábica obtido pelo método de Tocher, com base em 13 descritores multicategóricos.

Grupos	Genótipos
Grupo I	Catucaí Amarelo 2 SL, Catucaí Amerelo 20/15 CV 479, Catucaí Vermelho 20/15 CV 479, Sabiá 398, Pauma II, Oeiras MG 6851, Catiguá MG 02, Araponga MG 1, Pau Brasil MG, Material crioulo (Cassine), Obatã; Iapar 59, IPR 105
Grupo II	Acauã e Catucaí Amarelo 24/137
Grupo III	Catiguá MG 01 e Sacramento MG 1
Grupo IV	Paraíso H 419-3-3-7-16-4-1

Fonte: os autores.

Os materiais genéticos avaliados pelo método UPGMA ficaram divididos em três grupos, indicando a baixa variabilidade genética para descritores de folhas entre os materiais genéticos observados. Em trabalho realizado por Silva et al. (2013), foi observado que o agrupamento de 88 acessos de café arábica formaram 12 grupos, sendo que 48 deles (54,5%) foram alocados em um só grupo. Diante disso, concluíram que a população estudada tinha baixa variabilidade genética entre os acessos para as características avaliadas (SILVA et al., 2013).

O número de descritores utilizados para essa análise pode ter contribuído para baixa variabilidade genética identificada. Além disso, dos 13 descritores avaliados, três não tiveram diferenças entre os genótipos avaliados, o que corrobora para esse resultado. No entanto, descritores multicategóricos são mais efetivos na diferenciação de genótipos quando comparados aos quantitativos ou binários, e a inclusão de novas características tem de ser avaliada com critério, uma vez que isso pode não proporcionar alterações nos agrupamentos (MOURA et al., 2010). Por outro lado, os descritores quantitativos que foram convertidos em multicategóricos foram os discrepantes entre as características avaliadas, indicando que essa técnica pode ser usada com sucesso para a análise conjunta de variáveis. Segundo Silva et al. (2013), a estratégia de conversão dos dados quantitativos em multicategóricos e a integração dos dados de diferente natureza, favoreceu a melhor visualização da divergência genética, devido ao maior número de grupos formados em comparação ao agrupamento para cada conjunto de características (quantitativas ou multicategóricas).

O agrupamento pelo método de Tocher proporcionou a formação de quatro grupos, diferentemente do obtido pelo método UPGMA. Número divergente de grupos por esse dois métodos de agrupamento foram obtidos por Guedes et al. (2013), avaliando 12 acessos de café arábica Maragogipe. Os métodos de agrupamento hierárquicos e de otimização se distinguem pelo tipo de resultado a ser fornecido e pelas formas de se definir a proximidade entre um indivíduo e um grupo já formado, ou entre dois grupos quaisquer (CRUZ; CARNEIRO; REGAZZI, 2014). Com isso, divergências entre os dois agrupamentos são esperadas. No entanto, os genótipos Sacramento MG1 e Paraíso H 419-3-3-7-16-4-1 foram os mais divergentes dentre os materiais genéticos estudados nos dois agrupamentos utilizados.

## Conclusão

Os dados quantitativos convertidos em multicategóricos foram os descritores mais eficientes para a avaliação da diversidade genética.

O método de agrupamento da Ligação Média entre Grupo (UPGMA) foi o mais aderente à matriz de dissimilaridade obtida.

A diversidade genética para caracteres de folha de café arábica é baixa e estratégias para aumentar a variabilidade devem ser empregadas para potencializar os resultados dos programas de melhoramento genético.

## Agradecimentos

Ao Consórcio Pesquisa Café pela concessão de bolsa ao primeiro autor e pelo apoio financeiro ao projeto (Convênio 793706/2013; Meta 2.3).

## Referências

- ALVARES, C. A. et al. Köppen's climate classification map for Brazil. **Meteorol. Z**, v. 22, n. 6, p. 711-728, 2013.
- CARGNELUTTI FILHO, A. et al. Consistência do padrão de agrupamento de cultivares de feijão conforme medidas de dissimilaridade e métodos de agrupamento. **Pesq. Agropec. Bras.**, v. 45, n. 3, p. 236-243, 2010.
- CRUZ, C. D.; CARNEIRO, P. C. S.; REGAZZI, A. J. **Modelos biométricos aplicados ao melhoramento genético**: volume 2. 3 ed. Viçosa: UFV, 2014. 668p.
- CRUZ, C. D. Genes - a software package for analysis in experimental statistics and quantitative genetics. **Acta Sci. Agron.**, v. 35, n. 3, p. 271-276, 2013.
- FREITAS, Z. M. T. S. et al. Avaliação de caracteres quantitativos relacionados com o crescimento vegetativo entre cultivares de café arábica de porte baixo. **Bragantia**, v. 66, n. 2, p. 267-275, 2007.
- GUEDES, J. M. et al. Divergência genética entre cafeeiros do germoplasma Maragogipe. **Bragantia**, v. 72, n. 2, p. 127-132, 2013.
- IPGRI. **Descriptors for coffee (Coffea spp. and Psilanthus spp.)**. Roma, International Plant Genetic Resource Institute, 1996. 36p.
- MOURA, M. C. C. L. et al. Algoritmo de Gower na estimativa da divergência genética em germoplasma de pimenta. **Hortic. bras.**, v. 28, n. 2, p. 155-161, 2010.
- SAKIYAMA, N. et al. **Café arábica do plantio a colheita**. Viçosa: UFV, 2015, 316p.
- SUDRE, C. P. et al. Genetic variability in domesticated *Capsicum* spp as assessed by morphological and agronomic data in mixed statistical analysis. **Genet Mol Res.**, v. 9, n. 1, p. 283-294, 2010.
- SILVA, F. L. et al. Integração de dados quantitativos e multicategóricos na determinação da divergência genética entre acessos de cafeeiro. **Bragantia**, v. 72, n. 3, p. 224-229, 2013.
- SOKAL, R. R.; ROHLF, F. J. The comparison of dendrograms by objective methods. **Taxon**, v. 11, p. 30-40. 1962.