



Comparação entre dois programas na análise de dados obtidos por meio de marcadores dominantes ISSR*

Jarbson Henrique Oliveira Silva¹; Raiane de Sousa Oliveira²; Maria Fernanda da Costa Gomes³; Sérgio Emílio dos Santos Valente⁴; Raimundo Bezerra de Araújo Neto⁵ Paulo Sarmanho da Costa Lima⁵

¹Estudante de Ciências Biológicas/UFPI, estagiário da Embrapa Meio-Norte, oliveirahenrique2@outlook.com ²Estudante de Ciências Biológicas/UESPI, estagiária da Embrapa Meio-Norte; ³Doutoranda em Ciências Biológicas/UFPE. ⁴Professor Associado II na UFPI. ⁵Pesquisador da Embrapa Meio-Norte, paulo.costa-lima@embrapa.br

Existem diversas plataformas e programas que possibilitam realizar análises de dados moleculares. O presente trabalho teve como objetivo avaliar as divergências entre dois programas quanto aos resultados das análises de uma matriz binária de dados, baseada em presença ou ausência de bandas amplificadas por marcadores dominantes do tipo ISSR. Os dois programas utilizados foram o PAST e o R. Para o estudo utilizou-se uma matriz com dados moleculares de 9 indivíduos de pau-ferro (*Caesalpinia ferrea* Mart. ex Tul. Var. *ferrea*) do Banco de Forrageiras Nativas da Embrapa Meio-Norte. No PAST, realizaram-se todos os procedimentos para se obter uma matriz de distâncias e um dendrograma, ambos baseados no índice de dissimilaridade de Jaccard. Para a formação do dendrograma, utilizou-se o agrupamento hierárquico *average*. No R, as análises foram feitas após a instalação e leitura do pacote “vegan”, que permite a execução de comandos de agrupamento e de formação de um dendrograma; e assim como no PAST, foram utilizados o índice de dissimilaridade de Jaccard e o agrupamento hierárquico *average*. Nos dois programas, as análises resultaram em um dendrograma que dividiu a população estudada em dois grupos: um grupo menor com três indivíduos (PF-01, PF-02 e PF-03) e outro grupo maior com seis indivíduos (PF-04, PF-05, PF-06, PF-07, PF-08 e PF-09). As correlações cofenéticas apresentadas pelos dois programas foram similares: no R, a correlação cofenética foi de 0,8242, enquanto no PAST, foi de 0,8245. A diferença entre os dois programas é observada na matriz de distância entre os indivíduos, pois, apesar de terem sido baseadas no mesmo índice de dissimilaridade, as matrizes apresentam diferenças nos valores das distâncias. Os dois programas são eficientes na análise da matriz de dados moleculares e se mostram bastante similares nos resultados, com exceção da matriz de distâncias.

Palavras-chave: PAST, Plataforma R, cluster.

Agradecimentos: À Embrapa Meio-Norte, pelo apoio técnico e científico.

*Trabalho financiado pelo Macroprograma 1.