

Técnicas de agrupamento aplicadas no mapeamento da divergência genética de subpopulações de *Araucaria angustifolia* (Bert.) O. Ktze em Irati, Pr e Caçador, SC por marcadores isoenzimáticos

Ricardo Alexandre Valgas¹, Anselmo Chaves Neto¹,
Osmir José Lavoranti², Valderês Aparecida de Sousa²

¹Programa de Pós Graduação em Métodos Numéricos em Engenharia
Universidade Federal do Paraná, Curitiba, PR

²Embrapa Florestas, Colombo, PR
e-mail: ricardo.valgas@ufpr.br

Resumo

Este trabalho teve por objetivo avaliar a divergência genética entre subpopulações de *Araucaria angustifolia* por marcadores isoenzimáticos. Foram amostradas árvores em Irati, PR e Caçador, SC e, para cada árvore, foram avaliados seis *locus* e observados os alelos do endosperma e dos embriões (homozigotos e heterozigotos). Os resultados foram explorados por técnicas multivariadas de Análise de Agrupamento em que as semelhanças entre as unidades experimentais foram obtidas pelas medidas Euclidiana e Manhattan. Os métodos de ligação usados na construção dos grupos foram: Ward, Vizinho Mais Próximo e Ligação Média. Observou-se a menor divergência genética entre as amostras das subpopulações de Floresta Intocada e Floresta Explorada em Irati, e Capão 3 e Capão 4 em Caçador. A Floresta Plantada em Irati e as subpopulações de Capão 1 e Capão 2 em Caçador foram as regiões mais críticas para conservação.

Palavras-chave: mapeamento genético, marcadores isoenzimáticos, análise de agrupamento, *Araucaria angustifolia*.

Abstract

This work has as objective to determine the genetic divergence between local populations of *Araucaria angustifolia* by isozyme markers. For this, trees in Irati, PR and Caçador, SC were sampled, for each tree six *locus* were evaluated and observed the alleles of endosperm and embryos (homozygous and heterozygous). The results were explored by multivariate techniques as Analysis of Group where the similarities between the experimental units were obtained by the measures Euclidian and Manhattan. The

linkage methods used to perform the groups were: Ward, Next Neighborhood and Average Linkage. It was observed the smallest genetic divergence between local population samples of Pristine Forest and Logged Forest in Irati, even as Capão 3 and Capão 4 in Caçador. The Planted Forest in Irati and the local populations of Capão 1 and Capão 2 in Caçador were the most critical regions observed for conservation.

Keywords: genetic mapping, isozyme markers, analysis of grouping, *Araucaria angustifolia*.

1. Introdução

Muitas comunidades biológicas, que se desenvolveram ao longo de milhares de anos, agora estão sendo devastadas pela ação antrópica. Com isso, muitas espécies estão sob o risco de extinção. Nos *habitats* naturais, preservados como reservas naturais ou parques nacionais, são necessários cuidados especiais para prevenir a extinção das espécies remanescentes, pois a conservação “in situ” permite a continuidade da evolução natural das espécies.

Diante desse contexto, surgiu a biologia da conservação, a qual investiga os impactos negativos das atividades humanas sobre os ecossistemas e desenvolve práticas que previnem a extinção de espécies reintegrando-as, quando possível, a ecossistemas funcionais. No Brasil, as florestas naturais têm sido foco da biologia da conservação devido a sua riqueza de espécies, complexas estruturas de organização em comunidades e ameaças constantes aos seus ecossistemas.

A necessidade de se conservar os remanescentes florestais é urgente. Para que esse procedimento seja corretamente conduzido, são necessários conhecimentos biológicos e ecológicos relacionados às populações a serem conservadas. Parâmetros como a quantidade e distribuição da variabilidade genética, biologia reprodutiva das espécies, fluxo gênico, regeneração, são, dentre outras, informações essenciais nesse processo (TRINDADE, 2001).

1.1 A espécie *Araucaria angustifolia*

Durante séculos, as florestas tropicais brasileiras foram consideradas fontes inesgotáveis de matéria-prima para os mais diversos fins, tais como: madeira, frutos, fibras, entre outros produtos. Elas ocupavam, aproximadamente, 20% da área terrestre (cerca de 17 milhões de km²) e abrigavam mais de um terço das espécies do planeta, entretanto esses recursos foram explorados de forma predatória e muitos deles encontram-se ameaçados de extinção.

Um exemplo dessa exploração está na Floresta Ombrófila e, prin-

principalmente, nas Florestas com Araucária (Floresta Ombrófila Mista), ecossistema encontrado no Brasil em altitudes elevadas, frequente nos estados do Paraná, Santa Catarina e Rio Grande do Sul.

No Brasil, a espécie *Araucaria angustifolia* corresponde ao principal componente das Matas de Araucária, estendendo-se desde o oeste de Minas Gerais até o Rio Grande do Sul, principalmente no Planalto Sul Brasileiro (MANTOVANI *et al*, 2004).

A araucária é altamente valorizada devido ao consumo de suas sementes e a utilização da madeira para a confecção de artefatos em marcenaria. O pinhão é rico em amido e pode ser utilizado tanto na alimentação de animais quanto na humana, integrando por um longo período a alimentação de alguns grupos indígenas, dos primeiros colonos e dos caboclos que habitaram o interior dos estados do Sul (SOUZA, 2006).

Anselmini (2005) relata que a intensa exploração da floresta de araucária foi responsável pelo primeiro ciclo econômico no sul do Brasil, resultando em um enorme desmatamento nesta região. No Paraná, a criação da Estrada da Graciosa, em 1873, serviu para facilitar o transporte de toras e tábuas da cidade de Curitiba até Antonina, assim como a Ferrovia Paranaguá-Curitiba em 1885 (KOCH e CORRÊA, 2002).

Anos depois, com o início da Primeira Guerra Mundial, a madeira do pinheiro foi destinada ao mercado interno, principalmente para São Paulo e Rio de Janeiro. A partir de 1934, a exploração foi intensificada de tal forma que aproximadamente 18 bilhões de m³ de madeira foram exportados, quase tudo extraído das Matas de Araucária. Desse modo, já em 1978, a mata foi reduzida para apenas 8% da sua distribuição original (FUPEF, 1978).

Laboreau e Matos Filho (1948) citam a exploração das Matas de Araucária em três fases: a primeira fase explorou as árvores de maior valor comercial; a segunda explorou as árvores remanescentes da primeira fase e a terceira proporcionou a queimada da mata para a criação de áreas de pastoreio e agricultura.

1.2 Diversidade genética em espécies arbóreas

Geralmente, para se expressar a diversidade entre espécies naturais, é necessário apresentar medidas de diversidade que podem ser classificadas em três grupos: diversidade taxonômica, diversidade filogenética e diversidade genética.

A diversidade genética (ou variabilidade genética) mensura as diferenças existentes tanto entre indivíduos de uma mesma espécie quanto entre populações, as quais podem estar separadas entre si pela fragmentação dos *habitats* naturais.

De acordo com Francisco (2005), a diversidade genética é criada

pelo processo de mutação, a qual gera diferentes alelos. Entende-se por alelo como cada uma das várias formas alternativas do mesmo gene, ocupando um *locus* (posição) em um cromossomo.

Tanto na conservação como no melhoramento de recursos genéticos, é muito importante o conhecimento da estrutura genética das espécies naturais, bem como de seus sistemas reprodutivos, a fim de se determinar as áreas mais importantes para conservação.

A introdução da técnica de eletroforese de isoenzimas, no início da década de 60, além de iniciar a era dos marcadores moleculares, ampliou o número de marcadores que poderiam ser utilizados. Além disso, a utilização de isoenzimas permitiu a identificação de polimorfismo entre plantas que não apresentavam diferenças morfológicas (CONTE, 2004).

A atividade de uma enzima específica pode ser detectada através da eletroforese: quanto maior a carga elétrica de uma proteína, mais rápida a sua migração no gel. A consequência é o aparecimento de uma ou mais bandas que fornecem as distintas formas de uma mesma enzima, as isoenzimas, sendo identificadas em diferentes regiões do gel caso apresentem diferentes mobilidades eletroforéticas.

Após a introdução da eletroforese em gel de amido, as isoenzimas têm sido muito utilizadas em pesquisas na área biológica em investigações sobre variação intra-específica, genética de populações, evolução e mapeamento genético. Em plantas, elas têm sido utilizadas principalmente em genética de populações, evolução e caracterização de germoplasma.

2. Técnicas de agrupamento

Neste trabalho, foram utilizadas técnicas estatísticas multivariadas exploratórias, conhecidas por Análise de Agrupamento por método hierárquico para mensurar a divergência genética entre subpopulações de *Araucaria angustifolia*.

Os métodos hierárquicos são classificados em aglomerativos e divisivos, geralmente utilizados em análises exploratórias dos dados para identificar possíveis agrupamentos (MINGOTI, 2007).

Para se estabelecer até que ponto dois elementos do conjunto de dados podem ser considerados semelhantes ou não, é necessário considerar medidas que descrevam a similaridade entre eles, de acordo com as características que foram observadas. Mingoti (2007) explica que se as observações amostrais referentes à p variáveis estão armazenadas em um vetor, os elementos amostrais podem ser comparados através de medidas matemáticas (métricas), como as medidas de distâncias. Assim podem-se calcular as distâncias entre os vetores amostrais de observações e agrupar aqueles de menor distância.

Conforme Johnson e Wichern (1998), os elementos agrupados em *clusters* representam uma medida de “proximidade” através de algum tipo de distância, na maioria dos casos.

A distância Euclidiana é uma métrica poderosa e certamente a mais conhecida e mais utilizada para calcular a distância entre dois elementos no espaço p -dimensional. Considerando dois vetores, x e y , da matriz de observações, a distância Euclidiana entre eles é definida por:

$$d(\underline{x}, \underline{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_p - y_p)^2} = \sqrt{\sum_{i=1}^p (x_i - y_i)^2} \quad (1)$$

A distância de Minkowski é dada por:

$$d(\underline{x}, \underline{y}) = \sqrt[n]{|x_1 - y_1|^n + |x_2 - y_2|^n + \dots + |x_p - y_p|^n} = \sqrt[n]{\sum_{i=1}^p |x_i - y_i|^n} \quad (2)$$

que é uma generalização da fórmula da distância Euclidiana. Assim, quando $n = 1$, a equação 2 é conhecida por distância de Manhattan ou “métrica de *city-block*” e para $n = 2$ é a própria distância Euclidiana, mas variações nos valores de n podem causar trocas nos pesos dados (GROBE, 2005).

De acordo com Braulio (2005), grupos homogêneos de elementos podem ser formados a partir do cálculo das distâncias entre eles. Essas distâncias são calculadas com base nos vetores de características desses elementos e formam uma matriz de distâncias a partir da qual os elementos serão agrupados segundo algum método de ligação.

Cada distância é calculada de uma maneira própria, portanto é possível que sejam formados grupos de elementos diferentes para cada distância utilizada. Nesse caso, torna-se interessante verificar a consistência dos grupos formados.

Validar os agrupamentos encontrados significa garantir que realmente os grupos diferem entre si. Tal verificação pode ser realizada empregando-se desde testes estatísticos univariados para comparação de médias (t de Student e testes de comparação múltipla como Tukey, Duncan, HSD – *Honestly Significant Difference*, Dunnet e Scheffé), até testes multivariados como, por exemplo, a MANOVA (Análise de Variância Multivariada) e o teste T^2 de Hotteling.

Através da MANOVA (Quadro 1), verifica-se se há diferença estatisticamente significativa entre os vetores de médias amostrais de cada grupo, isto é, tem-se, como hipótese nula (H_0), que os vetores de médias amostrais não diferem entre si: $H_0 : \underline{\mu}_1 = \underline{\mu}_2 = \dots = \underline{\mu}_k$.

Quadro 1. Descrição da MANOVA.

Fonte de Variação	Matriz Soma de Quadrados	Graus de Liberdade
Tratamentos	$B = \sum_{j=1}^k n_j (\bar{x}_j - \bar{x})(\bar{x}_j - \bar{x})'$	$k - 1$
Resíduos	$W = \sum_{j=1}^k \sum_{i=1}^{n_j} (\bar{x}_{ij} - \bar{x}_j)(\bar{x}_{ij} - \bar{x}_j)'$	$n - k$
Total	$B + W$	$n - 1$

Ao se avaliar a hipótese nula (H_0), pelo teste de Lâmbda de Wilks, utiliza-se a estatística Λ , dada pela expressão:

$$\Lambda = \frac{|W|}{|B+W|} \quad (3)$$

em que:

B : é a matriz que resume a variação entre os efeitos dos tratamentos;
 W : é a matriz que resume a variação dentro dos grupos.

Ao se avaliar a hipótese nula (H_0), pelo teste de Hotteling, utiliza-se a estatística T^2 , dada pela expressão:

$$T^2 = [(\bar{X}_1 - \bar{X}_2) - (\underline{\mu}_1 - \underline{\mu}_2)]' \left[\left(\frac{1}{n_1} + \frac{1}{n_2} \right) S_p \right]^{-1} [(\bar{X}_1 - \bar{X}_2) - (\underline{\mu}_1 - \underline{\mu}_2)] \quad (4)$$

em que S_p é a estimativa da matriz de covariância conjunta. A estatística T^2 possui distribuição F com p graus de liberdade no numerador e $n_1 + n_2 - p - 1$ graus de liberdade no denominador.

No emprego dos testes multivariados para a comparação de médias, é necessário que as amostras aleatórias sejam independentes e oriundas de populações Gaussianas, ou seja, $\underline{X}_i \sim N_p(\underline{\mu}_i, \Sigma)$, $i=1, \dots, n$.

O pressuposto da Gaussianidade para o vetor \underline{X} pode ser verificado por um teste estatístico. No caso da distribuição normal bivariada com $p = 2$ variáveis e nível de significância α fixado, têm-se:

$$P\left[(\underline{X} - \underline{\mu})' \Sigma^{-1} (\underline{X} - \underline{\mu}) \leq \chi^2_2(1-\alpha) \right] = 1 - \alpha \quad (5)$$

em que $\underline{\mu}$ é substituído pela estimativa \bar{X} e a matriz de covariância Σ pela estimativa S .

2.1 Análise de Agrupamento por método hierárquico

O Método Hierárquico Aglomerativo é descrito por Johnson e Wichern (1998) e consiste de um estudo sistemático em etapas:

Etapa 1) Parte-se dos n elementos ou amostras para originar a matriz de semelhança para cada par de amostras ou elementos por alguma medida. Neste estudo, utilizou-se a matriz de distâncias $D = (d_{ij})$ de dimensão $n \times n$, na qual d_{ij} representa a distância do objeto i ao objeto j :

$$D = \begin{bmatrix} d_{11} & d_{12} & \dots & d_{1n} \\ d_{21} & d_{22} & \dots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n2} & \dots & d_{nn} \end{bmatrix} \quad (6)$$

em que $d_{11} = d_{22} = \dots = d_{nn} = 0$.

Etapa 2) O processo se inicia a partir das duas amostras A e B mais semelhantes (menor distância).

Etapa 3) A seguir, redimensiona-se a matriz de distâncias, substituindo as linhas e colunas que contém as amostras A e B por uma única coluna obtida dos valores de A e B , calculados por algum método de ligação, recalculando-se as distâncias a esse novo grupo AB formado.

Etapa 4) Os passos 2 e 3 devem ser repetidos $g - 1$ vezes (até que todos os elementos sejam incluídos em um único grupo) observando-se as identidades dos grupos que vão sendo formados e os níveis em que eles se fundem.

Ligações simples (vizinho mais próximo)

Neste método, a distância entre dois grupos é representada pela menor distância entre todas as combinações possíveis de dois pares de elementos. Assim, forma-se o primeiro grupo, por exemplo, grupo AB , onde a distância entre esse grupo a algum outro elemento C é calculada pela expressão:

$$d_{(AB),C} = \min \{ d_{AC}, d_{BC} \} \quad (7)$$

Ligação média

Neste procedimento, a distância considerada é a distância média

entre todos os pares possíveis de elementos. Então, a partir da matriz de distâncias $D = (d_{ij})$, agrupam-se os elementos com menor distância, por exemplo, A e B formam um grupo se $d_{AB} = \min\{d_{ij}\}, \forall i, j$ e a distância de (AB) a C é expressa por:

$$d_{(AB),C} = \frac{\sum_i \sum_j d_{ij}}{n_{(AB)}n_C} \quad (8)$$

em que $n_{(AB)}$ e n_C representam as quantidades de elementos presentes em cada grupo.

Método de Ward

Ward, em 1963, propôs uma técnica que busca minimizar a soma das diferenças entre os elementos de cada grupo e o valor médio do grupo. Isso equivale a minimizar o desvio padrão entre os elementos de cada grupo.

Johnson e Wichern (1998) descrevem que este método baseia-se na “perda mínima das informações” ao se formarem novos grupos, utilizando o critério da soma ao quadrado dos erros entre dois agrupamentos, para todas as amostras.

Inicialmente, existe um grupo para cada vetor componente do conjunto de dados, fornecendo um erro nulo para cada grupo, pois o vetor que compõe cada um representa o vetor médio do grupo. Posteriormente, os grupos são aglutinados 2 a 2, conforme o menor erro do grupo. É necessário um grande esforço computacional, caso o tamanho da amostra seja elevado e a cada iteração o número de elementos pertencentes a cada grupo aumenta, elevando o número de cálculos para o erro de cada um.

3. Material e método

As árvores amostradas para o presente estudo foram obtidas em áreas de proteção ambiental situadas nos municípios de Irati, PR e Caçador, SC.

A Floresta Nacional de Irati está localizada no estado do Paraná a 880m de altitude. Possui uma precipitação anual média de 1.442mm, ocupando uma área de 3.495ha, sendo 36% de floresta natural e apenas 2,4% de área de reflorestamento.

A Floresta Nacional de Caçador está localizada no estado de Santa Catarina a 1.100m de altitude numa área de aproximadamente 1.600ha. O clima é considerado Cfb – clima temperado úmido com verão temperado (de acordo com Köppen), com dias de frio intenso durante o inverno.

Nesta região, a vegetação é chamada “Floresta de Pinheiro Brasileiro”, sendo a *Araucaria angustifolia* a espécie dominante.

Em cada uma dessas regiões, foram coletadas amostras em grandes áreas florestais e em locais de menor incidência (ilhas de floresta ou capões) que variam de 12 a 60ha. Foram amostradas 10 árvores para cada subpopulação, totalizando 50 árvores em Irati e 60 em Caçador. O Quadro 2 fornece mais algumas informações sobre as amostras.

O procedimento de eletroforese aplicado partiu de um conjunto de 8 sementes de cada uma das árvores, pois esse é o número que seguramente estima o genótipo nuclear (SOUSA *et al.*, 2002).

As sementes foram armazenadas a -30°C até a extração da enzima para a eletroforese. Embriões e gametófitos femininos foram homogeneizados separadamente em uma extração tampão, contendo 5g de Sacarose, 150mg de DTT (Dithiothreitol) e 3g de PVP (Polivinilpirrolidona) dissolvido em 100ml de 0,1M Tris-HCl tampão, num pH 7,5. Extratos da enzima obtida do gametófito feminino e do embrião foram posicionados de forma adjacente no gel. O gel de amido horizontal (10,5% de amido e 2,5 – 3,5% sacarose) foi preparado seguindo Feret e Bergmann (1976), Conkle *et al.* (1982) e Liengsiri *et al.* (1990).

Quadro 2. Descrição das subpopulações

População/Município	Latitude/Longitude	Subpopulação
Irati, PR	25°30'/50°36'	Floresta Intocada
		Capão 1
		Capão 2
		Floresta Explorada
		Floresta Plantada
Caçador, SC	26°50'/51°00'	Floresta Intocada
		Capão 1
		Capão 2
		Capão 3
		Capão 4
		Floresta Explorada

O sistema tampão de eletroforese utilizado foi o ASHTON, ajustado para pH 8,6 para GOT e PGM, e o sistema TRIS-Citrato, ajustado para pH 7,3 para as outras enzimas. Os géis correram a 80 mA para o sistema ASHTON e a 180 mA para o sistema TRIS-Citrato. Foram examinados 6 *locus*: GOT-B, GOT-C, PGM-A, SKDH-B, MDH-B e 6-PGDH-B. Os detalhes sobre a metodologia da eletroforese de embriões e gametófitos e a interpretação genética dos zimogramas estão presentes em Sousa *et al.* (2002).

Embora exista uma grande quantidade de marcadores genéticos, o princípio de análise desses marcadores é simples: marcadores comuns significam semelhança genética e marcadores não comuns significam diferenças genéticas. As semelhanças e diferenças genéticas avaliadas em um banco de germoplasma geram grande quantidade de informações sobre a diversidade genética.

A primeira etapa do processo de análise de dados de marcadores genéticos é a codificação em dados de coincidência alélica (no caso de marcadores co-dominantes), ou seja, 1 para alelos comuns no *locus* e 0 para a ausência de alelos comuns no *locus*. A adequada aplicação dos métodos estatísticos multivariados exigiu tal codificação. Cada *locus* analisado é formado por dois alelos, pois, nesse caso, trata-se de genótipos ordenados e estes podem assumir os valores 1, 2 ou 3, de acordo com a interpretação obtida no gel de amido.

A segunda etapa na análise dos dados consiste em utilizar as informações codificadas para estimar os índices de similaridade ou de distância genética entre pares de observações. Para isso, é preciso decidir *a priori* quais serão utilizadas, pois existem várias medidas diferentes e cada uma delas produz um determinado tipo de agrupamento.

De acordo com a literatura existente sobre mapeamento genético, as medidas de semelhança mais utilizadas e que foram analisadas neste trabalho são as medidas de dissimilaridade Euclidiana, devido a sua ampla aplicação, e de Manhattan, pois é menos afetada pela presença de valores discrepantes na amostra que a distância Euclidiana.

A partir daí, calculam-se as respectivas matrizes de dissimilaridade entre as unidades experimentais, as quais formam a base das análises de agrupamento. A estrutura de grupos contida nos dados é visualizada em gráficos denominados dendrogramas (JOHNSON E WICHERN, 1998).

Neste trabalho, examinou-se o método do Vizinho mais próximo, o método da Ligação média e o método de Ward, justamente pelas suas características quanto à maneira de formação dos grupos. No primeiro método, a similaridade é definida pelos elementos com menor distância; no segundo, a distância entre dois grupos é tratada como a média das distâncias entre todos os pares de elementos que podem ser formados com os

elementos dos dois grupos e, finalmente, o terceiro método faz com que a variação entre grupos seja mínima e entre grupos seja notável. Dias (1998) descreve a aplicação dessas metodologias relatando as principais vantagens e desvantagens de cada uma nesse contexto.

4. Resultados

As análises estatísticas foram realizadas a partir dos vetores de médias amostrais de cada uma das subpopulações de Irati e Caçador. Estes vetores amostrais representam a média de árvores que apresentam alelos comuns no mesmo *locus*, ou seja, são homozigotos, de acordo com a codificação descrita anteriormente.

4.1 Irati

O processamento das análises estatísticas foi realizado no *software* R. Observa-se, na Figura 1, que as subpopulações de Floresta Intocada e Floresta Explorada são as mais semelhantes e formam um agrupamento, uma vez que apresentam a menor distância (0,2236). Capão 1 e Capão 2 formam o segundo grupo com menor distância (0,2290). O dendrograma gerado com o método de Ligação média apresenta uma pequena diferença dos demais: os dois maiores grupos unem-se para somente depois se unirem à subpopulação de Floresta Plantada, enquanto que nos outros dendrogramas a Floresta Plantada primeiro se une ao grupo de Capão 1 e Capão 2 para depois se unir ao grupo de Floresta Intocada e Floresta Explorada.

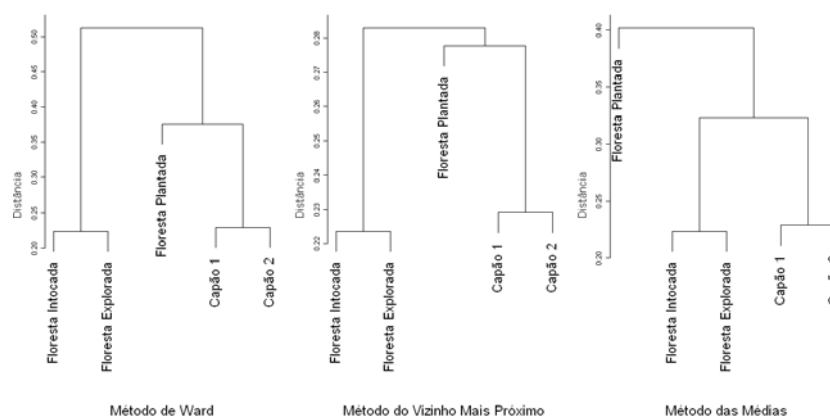


Figura 1. Dendrogramas das subpopulações de Irati, considerando a distância Euclidiana.

Os dois agrupamentos formados na Figura 2 são os mesmos observados na Figura 1, apenas com a inclusão da subpopulação de Floresta Plantada no grupo dos Capões 1 e 2 no método do Vizinheiro mais próximo, além de uma pequena diferença na escala.

O vetor médio das subpopulações de Floresta Intocada e Floresta Explorada foi comparado ao vetor médio das subpopulações de Capão 1, Capão 2 e Floresta Plantada, através do teste T^2 de Hotelling, com 95% de confiança. O p -valor do teste encontrado é inferior a 0,05, indicando uma diferença significativa entre eles. Do mesmo modo, foi comparado o vetor médio da subpopulação de Floresta Plantada com o vetor médio das subpopulações de Capão 1 e Capão 2, no qual o teste T^2 de Hotelling também forneceu um p -valor inferior a 0,05, acusando uma diferença significativa entre eles.

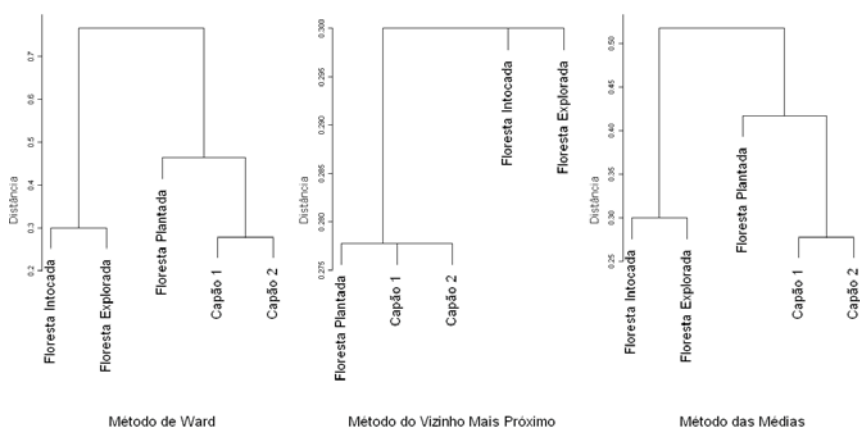


Figura 2. Dendrogramas das subpopulações de Irati, considerando a distância de Manhattan.

Na comparação das subpopulações de Capão 1 e Capão 2, o p -valor do teste é superior a 0,05. Logo, não existe diferença significativa entre os vetores médios dessas subpopulações. Na comparação do outro grupo formado pelas subpopulações de Floresta Intocada e Floresta Explorada, o teste T^2 de Hotelling também forneceu um p -valor superior a 0,05 e, mais uma vez, a hipótese nula não foi rejeitada.

Em todos esses testes foi utilizada a pseudo-inversa da matriz de covariância conjunta através da transformação de Moore-Penrose e a Gaussianidade dos vetores aleatórios observados foi verificada aplicando-se o teste baseado em uma estatística qui-quadrado descrito anteriormente.

4.2 Caçador

De acordo com os dendrogramas das Figuras 3 e 4, verifica-se que a subpopulação de Capão 1 se assemelha ao grupo de menor distância formado por Capão 3 e Capão 4. O segundo grupo mais homogêneo é formado pelas subpopulações de Floresta Intocada e Floresta Explorada.

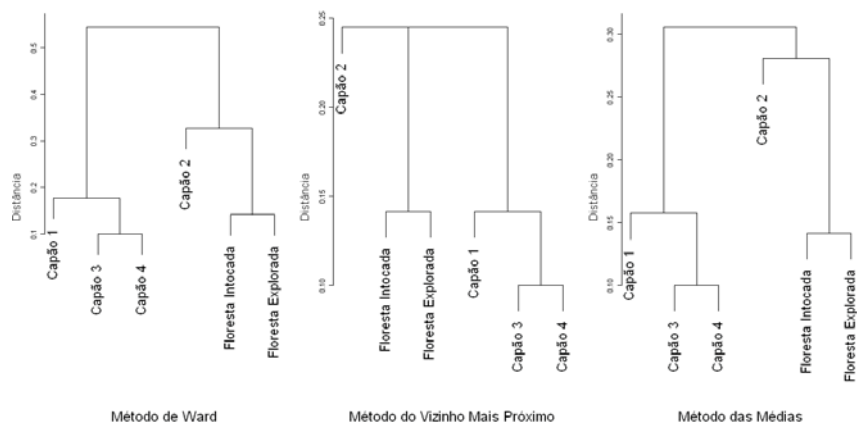


Figura 3. Dendrogramas das subpopulações de Caçador, considerando a distância Euclidiana.

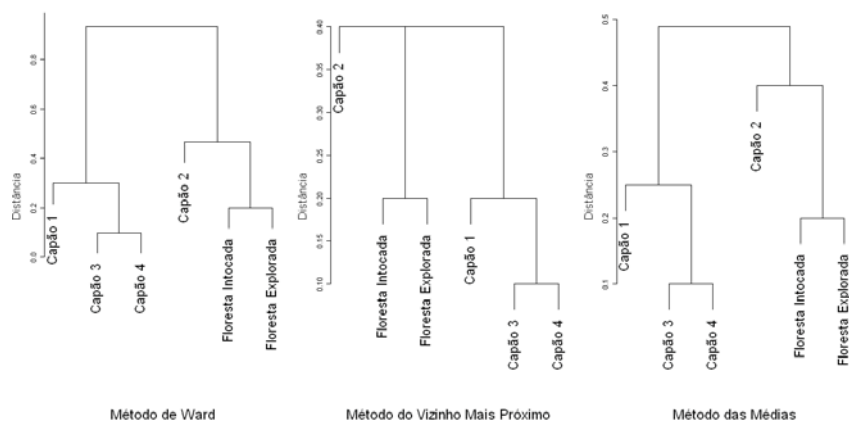


Figura 4. Dendrogramas das subpopulações de Caçador, considerando a distância de Manhattan.

Mais uma vez a validação dos agrupamentos encontrados deu-se através do teste T^2 de Hotelling, utilizando a pseudo-inversa da matriz de covariância conjunta, com 95% de confiança. O vetor médio das subpopulações de Capão 1, Capão 3 e Capão 4 foi comparado ao vetor médio das subpopulações de Capão 2, Floresta Intocada e Floresta Explorada. O p -valor do teste encontrado é inferior a 0,05, indicando uma diferença significativa entre eles.

Outras duas comparações foram feitas: uma entre o vetor médio da subpopulação de Capão 1 e o vetor médio do grupo composto pelas subpopulações Capão 3 e Capão 4, e outra entre o vetor médio de Capão 2 e o vetor médio do grupo formado pela Floresta Intocada e Floresta Explorada. Pelo teste T^2 de Hotelling e com 95% de confiança, nos dois casos o p -valor é inferior a 0,05, diferenciando significativamente a subpopulação de Capão 1 e a subpopulação de Capão 2 dos respectivos grupos as quais foram comparadas.

Na comparação dentro dos grupos, tanto o teste realizado para comparar as subpopulações de Capão 3 e Capão 4 quanto o realizado para comparar as subpopulações de Floresta Intocada e Floresta Explorada, forneceram um p -valor maior que 0,05. Assim, as hipóteses nulas em cada teste não foram rejeitadas e, portanto, os vetores de médias das subpopulações de cada grupo não diferem entre si.

5. Conclusões

Este artigo fornece uma parcela de contribuição ao conhecimento das semelhanças existentes entre as subpopulações de *Araucaria angustifolia* (Bert.) O. Ktze, com base na estrutura genética de cada uma delas. Além disso, gera informações técnico-científicas que possibilitam o delineamento de estratégias de conservação e aproveitamento racional para o manejo extrativista sustentável pois, através do manejo, é possível extrair da natureza os recursos florestais utilizados pelo homem para o seu crescimento sem devastá-la, conservando suas características genéticas e evitando a extinção das espécies.

Mostra, também, o poder da análise de agrupamentos por método hierárquico na fase exploratória quanto ao mapeamento da divergência genética em populações dessa natureza, identificando aquelas que possuem características genéticas mais semelhantes através dos agrupamentos formados.

Os esforços necessários para a preservação das araucárias deve se concentrar principalmente na subpopulação de Floresta Plantada em Irati e nas subpopulações de Capão 1 e Capão 2 em Caçador, pois foram estas que apresentaram as menores semelhanças com as demais subpopulações.

6. Bibliografia

- ANSELMINI, J. I. **Fenologia reprodutiva da Araucaria angustifolia (Bert.) O. Ktze, na região de Curitiba – PR.** 52 p. Dissertação (Mestrado em Agronomia) – Setor de Ciências Agrárias, Universidade Federal do Paraná, Curitiba, 2005.
- BRAULIO, S. N. **Proposta de uma metodologia para a avaliação de imóveis urbanos baseado em métodos estatísticos multivariados.** 131 p. Dissertação (Mestrado em Métodos Numéricos em Engenharia) – Setores de Tecnologia e de Ciências Exatas, Universidade Federal do Paraná, Curitiba, 2005.
- CONKLE, M. T. *et al.* Starch Gel Electrophoresis of Conifer Seeds: a Laboratory Manual. **Pacific Southwest Forest and Range Experimental Station.** 18 p. USDA, 1982.
- CONTE, R. **Estrutura genética de populações de Euterpe edulis MART. submetidas à ação antrópica utilizando marcadores alozímicos e microssatélites.** 124 p. Tese (Doutorado em Agronomia) – Setor de Ciências Agrárias, Escola Superior de Agricultura “Luiz de Queiroz”, Universidade Federal de São Paulo, Piracicaba, 2004.
- DIAS, L. A. S. Análises multidimensionais. In: ALFENAS, A. C. (Ed.). **Eletroforese de isoenzimas e proteínas afins.** Viçosa: Editora UFV, 1998. p. 405 – 475.
- FERET, P. P.; BERGMANN, F. Gel electrophoresis of proteins and enzymes. **Modern Methods in Forest Genetics**, Springer, Berlin, Heidelberg, p. 47 – 77, 1976.
- FRANCISCO, M. R. **Estruturação genética em populações do tangará-dançarino *Chiroxiphia caudata* (Aves, Pipridae) no corredor costeiro da mata atlântica (SP) e sua importância para a conservação.** 103 p. Tese (Doutorado em Ecologia e Recursos Naturais) – Setor de Ciências Biológicas, Universidade Federal de São Carlos, São Carlos, 2005.
- FUPEF (Fundação de Pesquisas Florestais do Paraná). **Inventário florestal do pinheiro no sul do Brasil.** Relatório Final. Curitiba, Brasil, 1978.
- GROBE, J. R. **Aplicações da estatística multivariada na análise de resultados em experimentos com solos e animais.** 134 p. Dissertação (Mestrado em Métodos Numéricos em Engenharia) – Setores de Tecnologia e de Ciências Exatas, Universidade Federal do Paraná, Curitiba, 2005.

JOHNSON, R.A; WICHERN, D.W. **Applied multivariate statistical analysis**. 2. ed. New Jersey: Prentice Hall, Inc., 1998.

KOCH, Z.; CORRÊA, M. C. **Araucária: a floresta do Brasil meridional**. Curitiba: Olhar Brasileiro, 2002.

LABOREAU, L. F. G. & MATOS FILHO, A. **Notas Preliminares sobre a região de Araucária**. Em: Anuário Brasileiro de Economia Florestal, 1. Rio de Janeiro, Brasil, 1948.

LIENGSI, C.; PIEWLUANG, C.; BOYTLE, T. J. B. Starch gel electrophoresis of tropical trees. A Manual. The ASEAN – Canada Forest Tree Seed Centre, 51 p. Saraburi, 1990.

MANTOVANI, A.; MORELLATO, P. C.; REIS, M. S. DOS. Fenologia reprodutiva e produção de semente em *Araucaria angustifolia* (Bert.) O. Kuntze. **Revista Brasileira de Botânica**, v. 27, n. 4, p. 787 – 796, 2004.

MINGOTI, S. A. **Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada**. 1ª reimpressão. Belo Horizonte: Editora UFMG, 2007.

SOUSA, V. A., HATTEMER, H. H.; ROBINSON, I. P. Inheritance and linkage relationships of isozyme variants of *Araucaria angustifolia* (Bert.) O. Ktze. **Silvae Genetica** 51. Frankfurt, v. 5 – 6, p. 191-196, 2002.

SOUZA, M. I. F. de. **Análise da diversidade genética de populações de *Araucaria angustifolia* (Bertol.) Kuntze utilizando marcador AFLP**. 97 p. Dissertação (Mestrado em Genética) – Setor de Ciências Biológicas, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2006.

TRINDADE, M. da G. **Estrutura Genética de Populações Naturais de Cagaiteira (*Eugenia dysenterica*) do Nordeste de Goiás, Brasil**. 104 p. Dissertação (Mestrado em Agronomia) – Setor de Ciências Agrárias, Universidade Federal de Goiás, Goiânia, 2001.

Submetido em: 10/setembro/2009

Aceito em: 23/08/2010