

**UNIVERSIDADE FEDERAL DO RECÔNCAVO DA BAHIA  
CENTRO DE CIÊNCIAS AGRÁRIAS, AMBIENTAIS E BIOLÓGICAS  
BACHARELADO EM ENGENHARIA FLORESTAL**

**DIVERSIDADE GENÉTICA A PARTIR DE DADOS COMPUTACIONAIS  
SIMULADOS PARA FINS DE CONSERVAÇÃO EM POPULAÇÕES  
NATURAIS SOB DERIVA GENÉTICA**

**DANIELA ACOSTA BRITO**

**Cruz das Almas-BA**

**Agosto/2017**

**DANIELA ACOSTA BRITO**

**DIVERSIDADE GENÉTICA A PARTIR DE DADOS COMPUTACIONAIS  
SIMULADOS PARA FINS DE CONSERVAÇÃO EM POPULAÇÕES  
NATURAIS SOB DERIVA GENÉTICA**

Trabalho de conclusão de curso apresentado ao Colegiado de Engenharia Florestal da Universidade Federal do Recôncavo da Bahia-UFRB pela estudante Daniela Acosta Brito como requisito parcial para obtenção do grau de Bacharel em Engenharia Florestal, sob orientação do Prof. Ricardo Franco Cunha Moreira.

**Orientador: Profº Drº Ricardo Franco Cunha Moreira**

**Co-orientador: Profº Drº Jair Wyzykowski**

**Cruz das Almas-BA**

**Agosto/2017**

DANIELA ACOSTA BRITO

**DIVERSIDADE GENÉTICA A PARTIR DE DADOS COMPUTACIONAIS  
SIMULADOS PARA FINS DE CONSERVAÇÃO EM POPULAÇÕES  
NATURAIS SOB DERIVA GENÉTICA**

Trabalho de conclusão de curso apresentado ao Colegiado de Engenharia Florestal da Universidade Federal do Recôncavo da Bahia-UFRB pela estudante Daniela Acosta Brito como requisito parcial para obtenção do grau de Bacharel em Engenharia Florestal, sob orientação do Prof. Ricardo Franco Cunha Moreira.

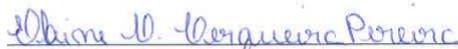
Aprovada em 21/05/17

Banca Examinadora



Prof<sup>o</sup> Ricardo Franco Cunha Moreira (Doutor Genética e Melhoramento de Plantas)  
- UFRB

Orientador



Prof<sup>a</sup> Elaine Costa Cerqueira Pereira (Doutora em Agronomia/Fitotecnia) - PNPB



Prof<sup>o</sup> Angelo Gallotti Prazeres (Doutor em Ciências Agrárias) - IFBA

## AGRADECIMENTOS

Agradeço a Deus e aos meus guias espirituais que iluminam e fortalecem minha caminhada.

À minha família, por sua capacidade de acreditar e investir em mim. Mãe, pelo cuidado e dedicação incondicional. Pai, por demonstrar o orgulho de me ver nesta caminhada. Tia Maryse, tia Marluce e vó Lusía por toda dedicação, incentivo, carinho, confiança, suporte e inspiração. A todos os primos, tias, tios e demais familiares da Bahia, São Paulo e Paraguai que me apoiaram.

A todos os professores que me acompanharam durante a graduação, em especial ao Prof. Ricardo Franco, responsável pela realização deste trabalho.

Ao Prof. Jesus Delgado Mendez que acompanha meu desenvolvimento acadêmico e humano junto ao PET Conexões de Saberes Socioambientais.

Aos colegas do Programa de Educação Tutorial pela amizade e experiências compartilhadas.

Ao curso de Engenharia Florestal da UFRB, local que pude compartilhar experiências e conhecer pessoas maravilhosas que irei levar sempre comigo. Em especial, Iracema por sempre me apoiar e pela amizade, Nai (mãe) pelo cuidado e zelo, Vel (pai) pelos “piriris” e alegria contagiante, Suélen pela companhia e inspiração nos estudos, Mari Bastos, Isabela, Flávia, Jônathan, Guto, Denise, Cacau e Taise.

A Rodrigo pelo incentivo, apoio e amizade.

A todos que direta ou indiretamente fizeram parte da minha formação, o meu muito obrigado.

## **Resumo**

A ruptura em uma unidade de paisagem que inicialmente era contínua por meio da ação antrópica de devastação define o processo de fragmentação florestal. Como consequências, as unidades remanescentes do processo ficam desconectadas, apresentando diferentes tamanhos e sujeitas a uma série de efeitos deletérios como é o caso da deriva genética. A deriva afeta de maneira significativa a frequência dos genes, ocasionando o afastamento dos mesmos dos que eram da população original, e de modo mais extremo contribui para a perda de alelos. Estudos genéticos de populações naturais buscam avaliar e quantificar como a variabilidade genética está distribuída no tempo e no espaço. Porém, no setor florestal existem poucas metodologias a serem seguidas e a implantação de estudos na área demanda altos custos. Por isso, o uso da simulação de dados na área da genética de populações proporciona o desenvolvimento de novos métodos estatísticos para auxiliar na análise de diversidade, define processos da história e estrutura da população, a fim de, prever frequências alélicas e genotípicas. O estudo objetiva testar diferentes procedimentos biométricos, através da simulação de dados provenientes de programas computacionais, a fim de auxiliar nas inferências em relação à conservação dos recursos genéticos de populações naturais sob condições de deriva genética. Observou-se que variáveis biométricas descritivas como heterozigosidade observada ( $H_o$ ) e heterozigosidade esperada ( $H_e$ ) são parâmetros eficientes para diferenciar aumento de heterozigotos em populações sob efeito de deriva genética e que as metodologias testadas são bons indicadores para diferenciar populações que sofrem processos que afetam as frequências gênicas.

**Palavras-chave: Evolução. Oscilações gênicas. Florestas.**

**Abstract**

The rupture of a landscape unit that was initially continuous through anthropic action of devastation defines the process of forest fragmentation. As a consequence, the remaining units of the process are disconnected, presenting different sizes and subject to a series of deleterious effects as is the case of genetic drift. The drift significantly affects the frequency of the genes, causing them to be removed from the original population, and in a more extreme way contributes to the loss of alleles. Genetic studies of natural populations seek to evaluate and quantify how genetic variability is distributed in time and space. However, in the forest sector there are few methodologies to be followed and the implementation of studies in the area demands high costs. Therefore, the use of data simulation in the area of population genetics provides the development of new statistical methods to aid in the analysis of diversity, defines processes of population history and structure, in order to predict allelic and genotype frequencies. The objective of this study is to test different biometric procedures by simulating data from computational programs in order to aid in inferences regarding the conservation of the genetic resources of natural populations under conditions of genetic drift. It was observed that descriptive biometric variables such as observed heterozygosity ( $H_o$ ) and expected heterozygosity ( $H_e$ ) are efficient parameters to differentiate increase of heterozygotes in populations with genetic drift effect and that the methodologies tested are good indicators to differentiate populations that undergo processes that affect the gene frequencies.

**Keywords: Evolution. Gene oscillations. Forests.**

# SUMÁRIO

<b>1. INTRODUÇÃO</b>	<b>8</b>
<b>2. OBJETIVOS</b>	<b>10</b>
<b>2.1. Objetivo geral</b>	<b>10</b>
<b>2.2. Objetivos específicos</b>	<b>10</b>
<b>3. REVISÃO DE LITERATURA</b>	<b>11</b>
<b>3.1. Importância da diversidade genética</b>	<b>11</b>
<b>3.2. Estrutura genética populacional</b>	<b>12</b>
<b>3.2.1. Parâmetros para caracterizar genética de populações</b>	<b>12</b>
3.2.1.1. Índice de fixação de Wright (Coeficiente de endogamia)	12
3.2.1.2. Heterozigosidade observada e esperada	13
3.2.1.3. Índice de Shanon-Wiever	14
3.2.1.4. Estatísticas F de Wright (FIT, FIS, FST, Nm)	14
3.2.1.5. Anova Molecular (AMOVA)	17
3.2.1.6. Conteúdo de população polimórfica (PIC)	18
3.2.1.7. Equilíbrio de Hardy-Weinberg (EWH)	18
3.2.1.8. Coeficiente de divergência genética entre populações (GST)	20
<b>3.3. Simulação de dados genéticos</b>	<b>20</b>
<b>4. MATERIAL E MÉTODOS</b>	<b>22</b>
<b>4.1. Simulação do conjunto de dados (Deriva genética)</b>	<b>22</b>
<b>4.2. Avaliação de diversidade genética</b>	<b>23</b>
<b>5. RESULTADOS E DISCUSSÕES</b>	<b>24</b>
<b>5.1. Deriva genética</b>	<b>24</b>
<b>5.2. Avaliação de diversidade genética</b>	<b>25</b>
5.2.1. Intrapopulacional	25
5.2.2. Interpopulacional	27
<b>5. CONCLUSÃO</b>	<b>31</b>
<b>REFERÊNCIAS</b>	<b>32</b>

## 1. INTRODUÇÃO

A genética de populações estuda o conjunto de forças que resultam nas mudanças evolutivas das espécies distribuídas no tempo e no espaço. Os indivíduos são diferenciados de acordo com sua constituição gênica e características que manifestam. Sendo assim, é estabelecida uma relação de genótipo e fenótipo, respectivamente, o que confere uma complexidade devido à dependência de diferentes genes e efeitos proporcionados pelo ambiente. Portanto, os princípios da genética de populações são essenciais para se obter uma perspectiva ampla da biologia e corrobora com a afirmação de Perez (2008) de que a variação genética é uma condição fundamental para que haja evolução (HARTL, 2010).

As florestas remanescentes foram em grande parte reduzidas a pequenos fragmentos isolados, ocasionado pela intensa ação antrópica de devastação. A formação desses mosaicos de vegetação é resultado da ruptura de uma unidade de paisagem que inicialmente era contínua. As unidades remanescentes do processo ficam desconectadas, apresentando diferentes tamanhos e ficam sujeitas a uma série de efeitos deletérios às populações que ali habitam. Como consequências, em curto prazo pode ocorrer deriva genética, que significa ter as frequências dos genes afastadas daquelas que eram da população original, a qual poderá ocorrer perda de alelos. Em longo prazo, pode haver aumento da endogamia devido à maior probabilidade de autofecundação e acasalamento entre indivíduos aparentados e, conseqüente, erosão genética (ZANELLA, 2011).

O estudo da estrutura genética permite conhecer modos de reprodução, estimar níveis de migração e dispersão, auxiliar no manejo de espécies ameaçadas de extinção, além de, proporcionar o diagnóstico do processo evolutivo de um grupo de táxons. A teoria da genética populacional envolve, principalmente, dois descritores básicos: frequência gênica e frequência genotípica. Modelos teóricos desenvolvidos nesta área contribuem com respostas sobre o comportamento das populações, a exemplo, do modelo reprodutivo de acasalamento ao acaso que influencia outros modelos da genética de populações, pois serve como ponto de partida ao considerar situações reais. Por isso, para o melhoramento genético, informações sobre a estrutura populacional permitem que o pesquisador realize predições em uma maior magnitude (CRUZ, 2008).

A simulação de sistemas genéticos consiste em imitar funcionalidades de modelos reais, de modo que avaliem o comportamento causado no sistema real, caso ocorram mudanças, ou seja, realiza predições futuras (ASSIS, 2005). Assim,

proporciona o desenvolvimento de novos métodos estatísticos para auxiliar na análise da diversidade, define processos da história e estrutura da população a fim de prever frequências alélicas (SANCHEZ, 2008). As populações simuladas podem servir de modelos que amparam estudos comparativos como uma alternativa viável na verificação de técnicas estatísticas mais eficientes (FERREIRA, 2007). Desta forma, estudos com utilização de modelos computacionais permitem avaliar o efeito de diversos fenômenos que incidem e atuam na diversidade genética em diferentes populações simuladas de tamanhos variados.

O objetivo deste trabalho é testar a eficiência de diferentes procedimentos biométricos através da simulação de dados e avaliar a diversidade genética entre e dentro de populações simuladas. Isso servirá para estimar parâmetros genéticos populacionais de utilidade no planejamento de programas de conservação e mitigação de fragmentos naturais.

## **2. OBJETIVOS**

### **2.1. Objetivo geral**

O estudo visa testar diferentes procedimentos biométricos, através da simulação de dados provenientes de programas computacionais, a fim de auxiliar nas inferências em relação à conservação dos recursos genéticos de populações naturais sob condições de deriva genética.

### **2.2. Objetivos específicos**

1. Comparar a diversidade genética entre e dentro de populações simuladas;
2. Observar eficiência de métodos biométricos;
3. Avaliar efeitos da deriva genética;
4. Estimar parâmetros genéticos populacionais de utilidade para o planejamento de programas de conservação e mitigação de fragmentos naturais.

### 3. REVISÃO DE LITERATURA

#### 3.1. Importância da diversidade genética

As florestas remanescentes foram em grande parte reduzidas a pequenos fragmentos isolados, ocasionado pela intensa ação antrópica de devastação. Como consequência da formação desses mosaicos de vegetação, problemas relacionados à biodiversidade local, como redução do tamanho de grandes populações, influenciam na alteração de processos ecológicos e genéticos de espécies naturais que ali ocorrem, tornando cada vez mais difícil a conservação e a manutenção ecossistêmica (ESTOPA et al., 2006).

Estudos genéticos de populações naturais, buscam avaliar e quantificar de que forma a variabilidade genética está distribuída no tempo e no espaço. A partir deste conhecimento, é possível compreender como a seleção influencia na adaptabilidade entre e dentro de populações, afinal, quanto maior a diversidade genética de uma população, maiores são as chances de perpetuação da espécie (ESTOPA et al., 2006). Desta forma, ao caracterizar os níveis de variabilidade e compreender a dinâmica de movimentação dos alelos nas populações naturais, adquirem-se subsídios para a maximização de estratégias de conservação (JUNIOR et al., 2004).

A junção de conceitos da ecologia de populações e da genética de populações é fundamental para o direcionamento de ações a serem realizadas e na definição de parâmetros adequados para o monitoramento destas (KAGEYAMA et al., 1998). Jaeger (2004) corrobora com esta afirmação e assegura que a união dos fundamentos teóricos e práticos da ecologia e da genética permite analisar, de acordo com o enfoque evolutivo, a dinâmica populacional, ou seja, é possível planejar o manejo de espécies e a recomposição de áreas preservando a variabilidade genética necessária para a manutenção de processos evolutivos.

Ao empregar marcadores polimórficos, a exemplo, dos microssatélites (SSR), observam-se parâmetros para estimativa da diversidade genética em populações naturais e as alterações decorrentes da ação antrópica. Os mais utilizados: são número médio de alelos por loco ( $A^{\wedge}$ ), heterozigosidade esperada ou diversidade gênica ( $H^{\wedge}e$ ); heterozigosidade observada ( $H^{\wedge}o$ ) e índice de fixação ( $f^{\wedge}$ ). Na avaliação, o número de alelos por loco e heterozigosidade esperada são altamente influenciados pela deriva genética. Já para inferir se a diversidade genética em uma população será mantida em

longo prazo é necessário realizar uma análise comparativa entre o número de alelos por loco e a heterozigosidade esperada (RAPOSO et al., 2007).

### **3.2. Estrutura genética populacional**

A estrutura populacional de uma espécie consiste em um conjunto de características genéticas e demográficas resultantes de ações e interações de uma série de mecanismos evolutivos e ecológicos. Nesta área, busca-se compreender os mecanismos micro evolutivos como: seleção, fluxo gênico e deriva genética. Assim, é possível diferenciar populações em ambientes heterogêneos e relacionar a distribuição dos indivíduos de acordo com o tempo e o espaço (MARTINS, 1987).

Desta forma, estudos que utilizam esta abordagem para comparar a estrutura genética populacional de espécies vegetais e graus de perturbação são fundamentais para o reconhecimento da perda genética real. Estes estudos direcionam ações que visam contribuir com o uso racional dos recursos naturais e permitem traçar estratégias de conservação para fins de sobrevivência das espécies (LIMA et al., 2015).

Ao aprofundar o estudo sobre a estrutura genética de populações, é possível detectar modos de reprodução e estrutura familiar, estimar níveis de migração e dispersão, inferir no melhor manejo de espécies ameaçadas de extinção, na manutenção de bancos de germoplasma e diagnosticar o histórico evolutivo de um conjunto de táxons. Especificamente para o geneticista ou melhorista, estas informações permitem prever mudanças em magnitude e sentido desejados (CRUZ et al., 2008).

#### **3.2.1. Parâmetros para caracterizar genética de populações**

##### **3.2.1.1. Índice de fixação de Wright (Coeficiente de endogamia)**

O coeficiente de endogamia ou índice de fixação foi introduzido por Wright em 1951. Consiste em um parâmetro que afeta diretamente a diversidade genética de uma população e define com propriedade a estrutura desta (MUNIZ et al., 2008).

A endogamia é o resultado do acasalamento de indivíduos aparentados de forma intencional ou não. Este método de acasalamento ocasiona a alteração da constituição genética da população devido ao aumento da homozigose e, conseqüente, redução da heterozigose. Portanto, há alteração da frequência fenotípica, mas não das frequências gênicas (QUEIROZ et al., 2000).

O efeito da endogamia sobre uma população é denominado “depressão por endogamia”. Assim, há uma perda do desempenho reprodutivo em uma população,

consequentemente, ocasiona alterações na fertilidade, na viabilidade de sementes, no vigor, na adaptação e na produtividade (CONTE, 2004)

A expressão de WRIGHT (1965) para medir a diversidade genética é a seguinte:

$$(1 - FIT) = (1 - FIS)(1 - FST) \text{ ou}$$

$$FIT = FST + (1 - FST).FIS$$

Onde:

FIT - índice de fixação para a população, ou seja, é a probabilidade total de identidade num determinado indivíduo I;

FIS - índice de fixação que ocorre dentro de população, ou seja, é a probabilidade de que 2 genes sejam homólogos no indivíduo I derivado do mesmo gene de um ancestral comum dentro da população;

FST - índice de fixação para alelos por locos, ou seja, é a probabilidade de que 2 genes sejam homólogos, combinados ao acaso na população, ambos originários de um gene na população.

### ***3.2.1.2. Heterozigosidade observada e esperada***

#### **a) Heterozigosidade observada ( $H_o$ )**

A quantidade de heterozigose de um determinado loco segundo os procedimentos de BROWN e WEIR (1983) são expressos pela seguinte fórmula:

$$H_o = 1 - \sum P_{ii}$$

Sendo:

$P_{ii}$  - Frequência observada de genótipos homozigotos do alelo i.

A heterozigosidade média observada é obtida pela soma dos valores de cada loco dividindo-se pelo número total de locos estudados.

#### **b) Heterozigosidade esperada ( $H_e$ )**

A quantidade de heterozigose é estimada de acordo com NEI (1978):

$$H_e = 1 - \sum P_i^2$$

Sendo:

$P_i$  - Frequência estimada do *i*ésimo alelo.

A heterozigosidade média esperada é obtida pela média aritmética de todos os locos estudados.

### 3.2.1.3. Índice de Shanon-Wiever

O índice de Shanon-Wiever ou riqueza fenotípica/genotípica é um parâmetro utilizado para mensurar a biodiversidade em dados categóricos. Leva-se em consideração o número de espécies e as espécies dominantes. Portanto, a função calcula a diversidade ou a riqueza de espécies em estudos de ecologia e a diversidade fenotípica ou genotípica dentro de uma população.

A estatística de Shannon-Wiever (1949) é definida como:

$$H = - \sum_{i=1}^c P_i \ln(P_i)$$

Onde:

c- número de classes genotípicas para um dado loco;

$P_i$ - frequência do *i*-ésimo genótipo.

Pode ser demonstrado que, para um dado número c de genótipos, o valor de H' poderá atingir o máximo dado por  $\ln(c)$ , que ocorre quando todos os genótipos estão presentes e com a mesma frequência.

### 3.2.1.4. Estatísticas F de Wright (FIT, FIS, FST, Nm)

#### a) Redução da heterozigosidade de um indivíduo com relação à metapopulação (FIT)

FIT é o índice de fixação para a população que representa a probabilidade total de identidade num determinado indivíduo I (FURLAN et al., 2007).

É expresso pela medida de heterozigosidade de um indivíduo em relação ao total da população:

$$FIT = \frac{Het - Ho}{Het}$$

Onde:

Ho = heterozigosidade observada (do indivíduo);

Het = heterozigosidade esperada da população total.

**b) Coeficiente de consanguinidade intrapopulacional (FIS)**

FIS é o índice de fixação que ocorre dentro de população. Representa a probabilidade de que 2 genes sejam homólogos no indivíduo I derivado do mesmo gene de um ancestral comum dentro da população (FURLAN et al., 2007). Expressa o desvio da casualidade de acasalamentos atuais pela seguinte fórmula matemática:

$$FIS = \frac{He - Ho}{He}$$

Onde:

H: é a heterozigosidade observada do indivíduo;

He: heterozigosidade esperada se a população estiver em Equilíbrio de Hardy-Weinberg.

Caso o valor de FIS seja maior que zero, a endogamia real supera o nível esperado sob acasalamentos ao acaso, implicando que os acasalamentos estão acontecendo entre indivíduos mais aparentados que a média, conseqüentemente, a população estará dividida em subpopulações (CARNEIRO, 2009).

**c) Coeficiente de consanguinidade entre subpopulações (FST)**

FST é o índice de fixação para alelos por locos que representa a probabilidade de que dois genes sejam homólogos, combinados ao acaso na população, ambos com origem de um gene na população (FURLAN et al., 2007).

Este índice é o coeficiente médio de endogamia esperado caso os reprodutores de cada período fossem acasalados de forma aleatória e foi estimado em populações hipotéticas produzidas pelo acasalamento ao acaso dos pais dentro de cada período (CARNEIRO, 2009). É obtido através da seguinte fórmula:

$$F_{ST} = \frac{Het - He}{Het}$$

Onde:

He: é a heterozigosidade esperada se a população está em Equilíbrio de Hardy-Weinberg (da subpopulação);

Het: é a heterozigosidade esperada da população total.

Valores de  $F_{ST}$  iguais a zero, são de populações que possuem frequência de alelos idênticas, já  $F_{ST}$  igual a 1, são populações que possuem alelos diferentes (McMANUS et al., 2011).

#### d) Fluxo gênico ( $Nm$ )

O fluxo gênico se caracteriza por um processo migratório que contribui intensamente na ampliação da variabilidade genética (CANTELMO, 2016). Ocorre a transferência de material genético de uma espécie para plantas de espécies diferentes ou iguais que determinam novas expressões fenotípicas.

Este escape gênico está associado também à poluição gênica, onde ocorre uma dispersão descontrolada de genes. Portanto, a migração pode resultar em grandes mudanças das frequências dos genes. Um fator significativo que afeta a taxa de fluxo gênico entre diferentes populações é a mobilidade, onde animais tendem a ser mais móveis que vegetais, conferindo-lhes maior potencial migratório. Nos vegetais, o fluxo gênico ocorre via pólen e depende de fatores como sincronismo de florescimento, elevada compatibilidade, abundância de vetores e métodos de difusão de pólen, distância de movimentação do pólen e condições ambientais apropriadas para polinização cruzada (CRUZ, 2008).

Para estimar fluxo gênico entre populações, se utiliza a seguinte equação proposta por Crow e Aoki (1984):

$$nm = \frac{1}{4k} \left( \frac{1}{F_{ST}} - 1 \right)$$

Onde:

$F_{ST}$ - coeficiente de consanguinidade entre subpopulações;

$k = [p / (p - 1)]^2$ ;

$nm$ - o número de migrantes;

p- o número de populações.

### 3.2.1.5. Anova Molecular (AMOVA)

A Análise Multivariada estuda a variedade de técnicas e algoritmos por meio de modelos multidimensionais. Tais técnicas conduzem à construção de modelos matemáticos que se ajustam aos dados obtidos experimentalmente e incluem os métodos de classificação, de otimização e de análise quantitativa e qualitativa, a exemplo, da Análise de Variância Molecular (AMOVA), que possibilita a identificação e classificação de materiais biológicos com estrutura hierárquica. Ao estimar variância em diferentes níveis hierárquicos, a AMOVA verifica a homogeneidade de um conjunto de dados, analisando se esta variabilidade está estruturada entre grupos, subgrupos ou está organizada dentro dos indivíduos (VASCONCELOS, 2006).

A Análise de Variância Molecular, é uma metodologia capaz de avaliar a diversidade entre populações a partir de dados moleculares e também de testar hipóteses a respeito de tal diferenciação. Uma variedade de dados moleculares, como informações de marcadores dominantes ou codominantes e dados de sequência, pode ser analisada usando este método (EXCOFFIER et al., 2001).

**Tabela 1:** Esquema de análise de variância gênica

FV	GL	SQ	QM	E(QM)
Entre populações	$g-1$	SQP	QMP	$\sigma_g^2 + 2\sigma_i^2 + 2k\sigma_p^2$
Indivíduos/populações	$\sum_{i=1}^g Ni - g$	SQI	QMI	$\sigma_g^2 + 2\sigma_i^2$
Alelos/indivíduos	$\sum_{i=1}^g Ni$	SQG	QMG	$\sigma_g^2$
Total	$2N-1$			

Fonte: Perez (2008)

Onde:

$$k = \frac{1}{g-1} \left( \sum_{i=1}^g Ni - \frac{\sum_{i=1}^g Ni^2}{\sum_{i=1}^g Ni} \right) \quad N = \sum_{i=1}^g Ni$$

As somas de quadrado são obtidas por meio das equações abaixo:

$$C = \frac{x^2}{2N}$$

$$SQP = \sum \frac{x_i^2}{2Ni} - C$$

$$SQI = \frac{1}{2} \sum_{i=1}^g \sum_{j=1}^{Ni} x_{ij}^2 - \sum \frac{x_{i..}^2}{2Ni}$$

$$SQG = \sum_{i=1}^g \sum_{j=1}^{Ni} \sum_{k=1}^2 x_{ijk}^2$$

Os estimadores dos componentes de variância são dados pelas equações abaixo:

$$\sigma_g^2 = QGM$$

$$\sigma_i^2 = \frac{QMI - QMG}{2}$$

$$\sigma_p^2 = \frac{QMP - QMG}{2k}$$

### 3.2.1.6. Conteúdo de população polimórfica (PIC)

O Conteúdo de Informação Polimórfica (PIC), é um parâmetro que indica qualidade de marcadores em estudos genéticos como: segregação, identificação de populações e controle de paternidade. Este índice quantifica o valor da informação do polimorfismo de um *locus* marcador e apresenta dependência do número de alelos e de suas frequências. É possível identificar qual homólogo de um pai foi transmitido à prole e de que forma o outro progenitor é genotipado.

O cálculo para obtenção do PIC de acordo com Bolstein (1980) é o seguinte:

$$PIC = 1 - \left( \sum_{i=1}^k x_i^2 \right) - \sum_{i=1}^{k-1} \sum_{j=i+1}^k 2x_i^2 x_j^2$$

Valores de PIC superiores a 0,5 são considerados muito informativos, valores entre 0,25 e 0,5 são mediamente informativos e valores inferiores a 0,25 são pouco informativos.

### 3.2.1.7. Equilíbrio de Hardy-Weinberg (EWH)

O Equilíbrio de Hardy-Weinberg (EHW) foi determinado em 1908 pelos pesquisadores Hardy e Weinberg. Os autores afirmaram que populações de tamanho

grande, em condições de acasalamento aleatório, sem seleção, mutação ou migração, permanecem com as frequências alélicas e genotípicas constantes de geração a geração, considerando-se uma população em completo equilíbrio (MENEZES, 2006).

É uma teoria importante, afinal, comprova o que ocorre com as frequências gênicas e genotípicas de populações. É possível observar as frequências de qualquer população que atravessa uma geração de cruzamentos ao acaso. Neste caso, as frequências genotípicas da prole são geradas exclusivamente pelas frequências gênicas dos progenitores, o que estabelece a relação de que frequência de homozigotos é igual ao quadrado da frequência de gene relevante e a frequência de heterozigoto é igual ao dobro do produto das frequências dos genes relevantes; e as frequências gênicas e genotípicas permanecem constantes de uma geração para outra (NICHOLAS, 2011).

McManus et al. (2011) esquematizam matematicamente o EHW, que pode ser demonstrado a partir de uma binominal simples ou multinominal, para dois alelos e alelos múltiplos, respectivamente, como é descrito a seguir:

$$p^2 + 2pq + q^2$$

Em um ambiente natural é difícil encontrar condições necessárias para alcançar o EHW, por isso a frequência alélica pode modificar de uma geração para outra. A partir do momento em uma população se distancia dos parâmetros do EHW por fatores externos, é possível avaliar os efeitos observados e esperados pela fórmula estatística qui-quadrado:

$$x^2 = \sum_{i=1}^k \frac{(O - E)^2}{E}$$

Onde:

**O:** é o número observado de indivíduos com genótipo específico;

**E:** é o número esperado de indivíduos com base no equilíbrio de Hardy-Weinberg (ex., para AA, o número esperado de indivíduos é:  $p^2 \cdot N$ );

**K:** é o número de genótipos.

### 3.2.1.8. Coeficiente de divergência genética entre populações (*GST*)

A medida *GST* determina a divergência genética relativa entre grupos. O índice varia entre 0 e 1 e expressa a proporção da diversidade total explicada por diferenças entre as variáveis. Matematicamente é equivalente à estatística *FST* de Wright. A magnitude relativa da diferenciação entre grupos é medida através da seguinte expressão (PEREZ, 2008):

$$GST = \frac{100DIT}{HIT}$$

Onde:

DIT: diversidade genética

HIT: heterozigosidade total

Segundo Hartl e Clark (1971) a equação de Nei pode ser aplicada a diferentes níveis da classificação hierárquica, como populações dentro de uma região geográfica, subpopulações dentro de populações, entre outros. Em cada caso se utiliza as frequências alélicas correspondentes da categoria em questão.

### 3.3. Simulação de dados genéticos

O uso da informática é imprescindível para o processamento e o gerenciamento de dados. E, juntamente, com o impulso da utilização desta ferramenta, a simulação tem ganhado espaço na busca de metodologias para avaliação de situações a serem estudadas no intuito de realizar previsões futuras. Na área da genética de populações, a simulação proporciona o desenvolvimento de novos métodos estatísticos para auxiliar na análise de diversidade, define processos da história e estrutura da população, a fim de, prever frequências alélicas e genotípicas (SÁNCHEZ, 2008).

A simulação de sistemas genéticos consiste em imitar funcionalidades de modelos reais, de modo que avaliem o comportamento causado no sistema real, caso ocorram mudanças (ASSIS, 2005). Por isso, Sánchez (2008) afirma que o uso dessa ferramenta é uma alternativa para que geneticistas e melhoristas possam obter resolução de problemas, aprimoramento de técnicas estatísticas e comparações entre metodologias.

Devido à dificuldade em realizar experimentos com número de observações e repetições necessárias para obtenção de estimativas precisas, a simulação de dados tem sido uma ferramenta de destaque entre pesquisadores, já que minimiza a utilização de recursos na instalação de experimentos em campo e permite estudos mais abrangentes (OLIVEIRA, 2015).

Além de poupar recursos, Ferreira (2007) ressalta que ao utilizar ferramentas de softwares de maneira inadequada, o geneticista ou melhorista pode escolher uma análise não condizente ao seu conjunto de dados. Portanto, a necessidade de informações e novos referenciais teóricos e práticos que orientem a utilização de aplicativos e recursos biométricos é de fundamental importância para o melhor aproveitamento dos dados e interpretação correta dos resultados.

#### 4. MATERIAL E MÉTODOS

O programa GENES vs 2015.5.0 (CRUZ, 2013) foi utilizado para simular o conjunto de dados sobre efeito de deriva genética e gerar arquivos de saída compatíveis com o programa POPGENE vs 1.3.2, com a finalidade de estimar parâmetros biométricos descritivos das populações simuladas.

##### 4.1. Simulação do conjunto de dados (Deriva genética)

Inicialmente foi realizada a simulação genotípica de populações multialélicas para marcadores codominantes do tipo microssatélites (SSR). Nesta simulação de populações e marcadores, foi determinada uma população contendo mil indivíduos, vinte alelos e dois alelos por locos, não havendo variação nos mesmos (Tabela 2).

**Tabela 2:** Simulação de populações e marcadores no programa GENES. Cruz das Almas, BA.

Número de populações	1
Número de indivíduos/população	1000
Varição indivíduo	0
Número de locos	20
Alelos/locos	2
Varição loco	0

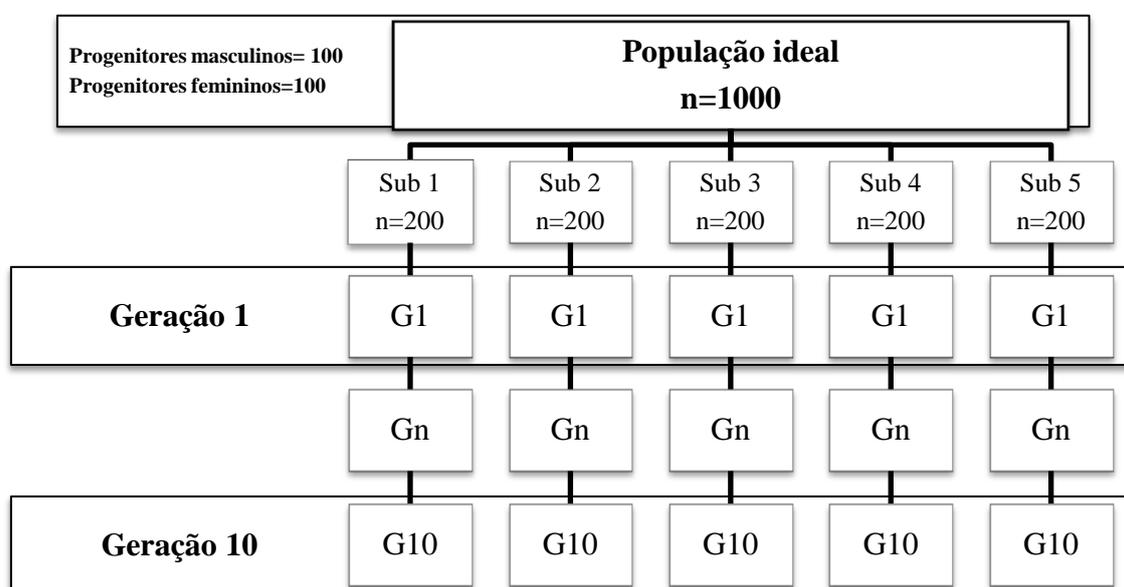
A partir do processamento destes dados, um arquivo (dat) foi gerado contendo códigos informativos dos alelos, onde genótipos homozigotos são descritos pelos caracteres 11 e 22 e heterozigotos são descritos pelo caractere 12. A Tabela 3 demonstra o modelo de arquivo gerado pelo programa GENES.

**Tabela 3:** Amostra de arquivos gerados pelo programa GENES Cruz das Almas, BA.

População	Indivíduos	Loco					
		1	2	...	...	19	20
1	1	11	12	...	...	22	11
1	2	12	11	...	...	12	11
1	...	...	...	...	...	...	...
1	999	11	12	...	...	12	11
1	1000	12	11	...	...	12	12

Obs.: 11 e 22 (indivíduos homozigotos)/12 (indivíduos heterozigotos).

A população gerada através de simulação genotípica multialélica serviu de modelo para a população ideal na simulação de gerações sob deriva genética. Portanto, leva-se em consideração que o tamanho da população ideal foi de 1000 indivíduos e que o marcador genético possui número de locos igual a 20. Na amostragem determinou-se o número de amostras igual a 200, ou seja, a população inicial foi subdividida em 5 subpopulações de forma igualitária. Além disso, foi determinado o tamanho de cada amostra igual a 20, o número de gerações igual a 10 e a quantidade de progenitores masculinos e femininos na mesma proporção, totalizando 100 indivíduos para cada, compondo assim, o número efetivo populacional ( $N_e$ ) que considera os indivíduos aptos sexualmente (Figura 1). Deste modo é possível avaliar a distribuição da frequência gênica e os efeitos causados pela deriva genética.



**Figura 1.** População ideal e gerações de subpopulações.

#### 4.2. Avaliação de diversidade genética

O cálculo da estrutura e da diversidade genética em níveis intra e interpulacionais foi avaliado com o auxílio dos programas computacionais GENES e POPGENE, onde foram testadas as seguintes técnicas biométricas descritivas:

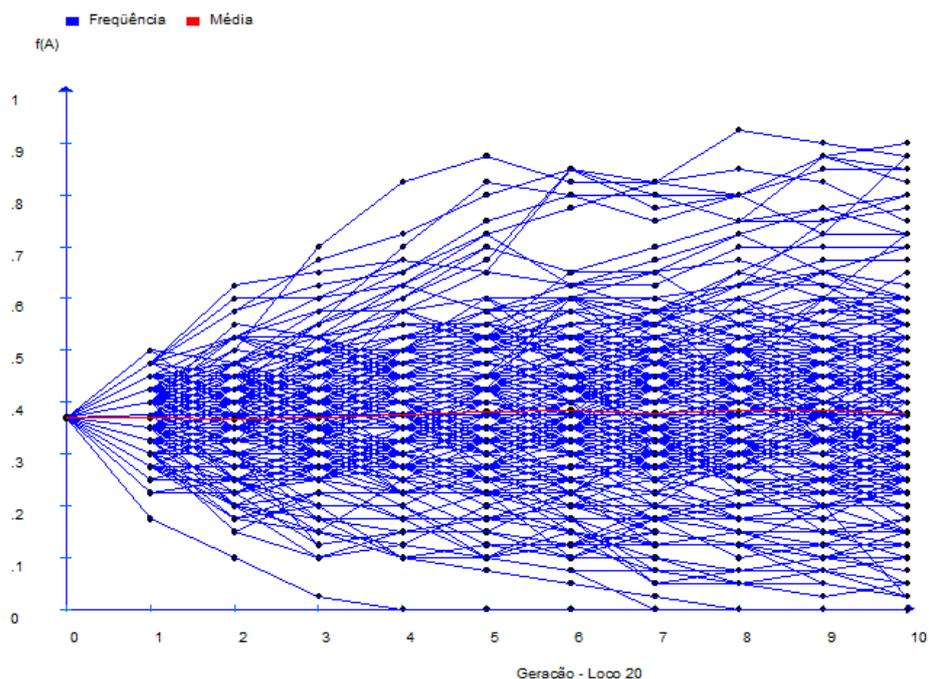
- I. Endogamia e heterozigose
- II. Índice de Shanon-Wiever
- III. Estatística F de Wright (1951)
- IV. Análise de variância molecular (AMOVA)
- V. Estatística de Nei e Conteúdo de população polimórfica (PIC)

## 5. RESULTADOS E DISCUSSÕES

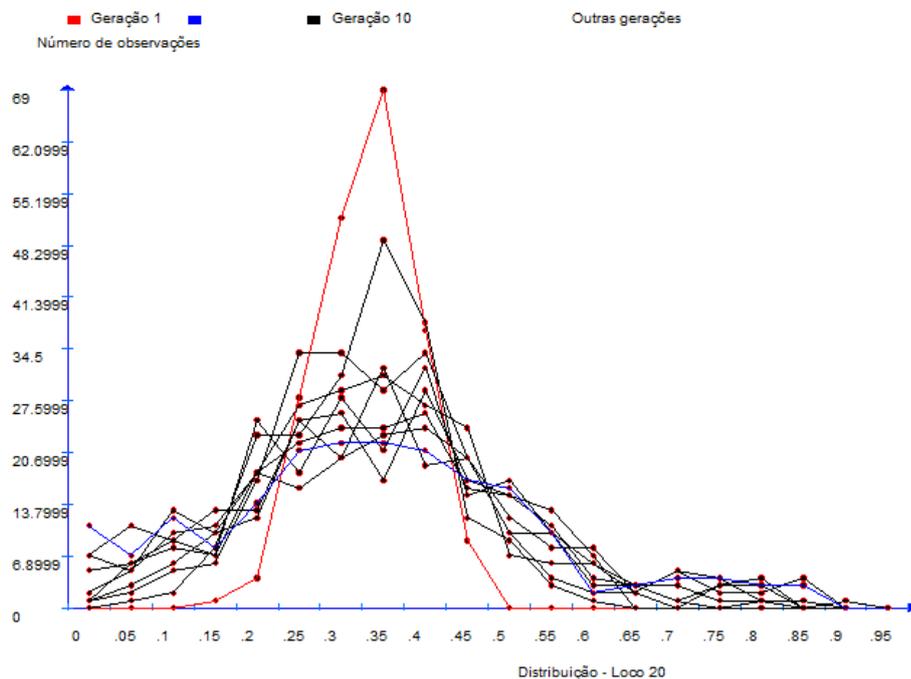
### 5.1. Deriva genética

As amostras seguem o Equilíbrio de Hardy-Weinberg, ou seja, as populações estão sob condições de acasalamento aleatório, sem seleção, mutação ou migração e tendem a permanecer com as frequências alélicas e genotípicas constantes de geração a geração, considerando-se uma população em completo equilíbrio. Porém, na análise do gráfico de oscilações das frequências alélicas observa-se que sob deriva genética, as frequências se comportam de maneira irregular com flutuações em todas as direções, comprovando que o índice é instável sob deriva e impossibilita a realização de futuras predições (Figura 2).

Observa-se no pico a primeira geração de acasalamento ao acaso com uma amostra no índice 0, nenhuma no índice 1 e as demais variando entre 0,23 e 0,52. Já na última geração de acasalamento ao acaso existem seis amostras no índice 0, nenhuma no índice 1 e as demais variando entre 0,05 e 0,73 (Figura 3). Quando um alelo atinge a frequência 1, diz-se que este fica fixado naquela linha e todos os indivíduos da linha são geneticamente idênticos, quando chega até a frequência 0 haverá perda de alelos. Portanto, pode-se inferir que na geração 10, houve maiores perdas alélicas em relação à geração 1 e em ambas houve baixa probabilidade das frequências representarem indivíduos geneticamente iguais.



**Figura 2.** Oscilação das frequências alélicas em 5 subpopulações, contendo 200 indivíduos simulados cada uma, 10 gerações de acasalamento ao acaso.



**Figura 3.** Frequência alélica ao longo de 10 gerações de acasalamento ao acaso.

## 5.2. Avaliação de diversidade genética

### 5.2.1. Intrapopulacional

#### I. Endogamia e heterozigose

A deriva genética causa perda de diversidade, o que significa ter as frequências dos genes afastadas daquelas que eram da população original (ZANELLA, 2011). Portanto, como esperado, as subpopulações simuladas sofreram perda significativa da variabilidade genética. Observa-se que nas subpopulações 1, 3, 4 e 5 (Tabela 4), as estimativas de heterozigosidade observada ( $H_o$ ) são maiores do que as estimativas de heterozigosidade esperada ( $H_e$ ), o que sugere um excesso de genótipos homozigotos.

Logo, com o aumento da homozigose verifica-se uma tendência crescente do coeficiente de endogamia, assim como observado por Sánchez (2008) e Perez (2008). Os autores atribuem a oscilação da frequência dos genes ao aumento da taxa de autofecundação e cruzamento entre indivíduos aparentados, que os tornam endogâmicos e susceptíveis à erosão genética.

**Tabela 4.** Médias de heterozigosidade observada (Ho) e heterozigosidade esperada (He), de 5 subpopulações com 10 gerações de acasalamento ao acaso. Cruz das Almas, BA.

<b>Subpopulação</b>	<b>Ho</b>	<b>He</b>
1	0,641	0,354
2	0,332	0,355
3	0,281	0,280
4	0,344	0,329
5	0,347	0,339

## **II. Índice de Shanon-Wiever**

A equação do índice de Shannon-Wiever calcula a diversidade ou a riqueza de espécies em estudos de ecologia e a diversidade fenotípica ou genotípica dentro de uma população. É possível, através deste parâmetro, verificar a proporção da diversidade total que se concentra dentro de grupos e entre grupos (BERGOCH, 2000). Assim, quanto maior o valor do índice, maior é o número de espécies e menor é o domínio da comunidade por uma ou algumas espécies (ODUM, 1975).

Observa-se que as estimativas do Índice de Diversidade de Shannon-Weaver entre as 5 subpopulações simuladas de acasalamento ao acaso variaram no valor mínimo de 0,430 na subpopulação 3 e no valor máximo de 0,525 na subpopulação 1 (Tabela 5). Ambas as variáveis apresentaram baixos resultados de acordo com a classificação de Amaral (2010). Sugere-se que há um maior domínio por parte de determinadas espécies dentro das subpopulação simuladas, o que ocasiona perda da biodiversidade de espécies e o estabelecimento e domínio por parte de uma ou poucas espécies na comunidade.

Apesar das médias do parâmetro em questão apresentarem baixos resultados do ponto de vista de variedade de espécies, as subpopulações 1 e 2 apresentaram os maiores índices, 0,525 e 0,523, respectivamente. Implica em associar que estas subpopulações possuem um elevado número de espécies maduras e menor número de espécies em processo de regeneração natural. Higuchi (2006), constata que este processo ocorre devido à predominância de árvores de vegetação adulta que possibilitam o desenvolvimento de novas espécies em estádios iniciais de sucessão devido às condições microclimáticas propiciadas pela formação de dossel.

**Tabela 5.** Médias de índice de diversidade de Shannon-Wiever, de 5 subpopulações com 10 gerações de acasalamento ao acaso. Cruz das Almas, BA.

<b>Subpopulação</b>	<b>I</b>
1	0,525
2	0,523
3	0,430
4	0,495
5	0,502

### **5.2.2. Interpopulacional**

#### **I. Estatística F de Wright (1951)**

A estatística de Wright (1965) é comumente utilizada para avaliar a estrutura genética de populações. Na Tabela 6 é possível observar parâmetros mínimo, médio e máximo de índices de fixação dentro de subpopulações (FIS) que varia de 0 a 1 na avaliação do grau de endogamia, diversidade genética entre subpopulações (FIT) que varia de 0 a 1 na avaliação de heterozigose e de fixação para o conjunto de populações (FST) que varia de 0 a 1 na avaliação do maior grau de diferenciação entre populações.

**Tabela 6:** Valores mínimo, médio e máximo de índices de fixação dentro de subpopulações (FIS), diversidade genética entre subpopulações (FIT) e de fixação para o conjunto de populações (FST) de acordo com Wright (1965), das subpopulações com tamanho de amostra igual a 200. Cruz das Almas, BA.

	<b>FIS</b>	<b>FIT</b>	<b>FST</b>
Mínimo	-0,1328	-0,0484	0,0562
Média	-0,0059	0,2915	0,2956
Máximo	0,1247	0,6060	0,6178

O índice de fixação entre subpopulações (FIS) considera valores próximos a 1 com maior grau de endogamia entre subpopulações. A média deste parâmetro apresenta o valor de -0,0059 o que significa que há menor grau de endogamia entre subpopulações e maior grau de endogamia dentro de subpopulações. É possível justificar o evento pelo aumento da taxa de autofecundação e cruzamento entre indivíduos aparentados dentro de subpopulações, e maiores trocas de frequências alélicas entre os indivíduos de diferentes populações, que possibilita uma maior variabilidade genética.

O índice de diversidade genética entre subpopulações (FIT), apresenta valores próximos a 1 com maior quantidade de indivíduos homozigotos. A média deste parâmetro apresenta o valor de 0,2915, o que significa que há maior quantidade de indivíduos com genótipo heterozigotos entre subpopulações e maior quantidade de indivíduos homozigotos dentro de subpopulações. Portanto, corrobora com o resultado e avaliação do índice de fixação entre populações (FIS) descritas anteriormente.

O índice de fixação para o conjunto de populações (FST), conceitua valores próximos a 1 com maior grau de diferença entre populações. A média deste parâmetro apresenta o valor de 0,5956, o que significa que a população ideal simulada apresenta frequências alélicas semelhantes, provavelmente, ocasionados pela deriva genética.

## II. Análise de variância molecular (AMOVA)

As estimativas geradas pelos marcadores microssatélites (SSR) demonstram que há maior proporção da variabilidade genética entre populações do que dentro de populações, com estimativa de 51,1% (Tabela 7).

A taxa de diferenciação genética entre populações pode ser explicada pela ocorrência de acasalamento ao acaso entre as subpopulações corroborando com o estudo de Oliveira (2015) que ao realizar distribuição da variabilidade entre e dentro de populações de açazeiro concluiu que devido à influência do fluxo gênico via pólen, há uma maior distribuição das frequências alélicas entre populações.

**Tabela 7:** Análise de variância molecular (AMOVA), com referência à 10<sup>a</sup> geração ao acaso, das subpopulações de 200 indivíduos.

<b>FV</b>	<b>GL</b>	<b>SQ</b>	<b>QM</b>
Entre	4	277,656	69,414
Dentro	995	328,789	0,330
Total	999	606,445	0,607

<b>Componente da variância</b>	<b>Estimativa</b>	<b>Porcentagem</b>
Entre	0,345	51,1
Dentro	0,330	48,9
Total	0,675	100
Estatística ØST	0,511	-

### III. Estatística de Nei

A Estatística de Nei calcula a diferenciação genética entre as populações. Para isso, utiliza-se de parâmetros como diversidade gênica total (HIT), diversidade gênica média dentro de populações (HIS) e diversidade gênica média entre as populações (DST). Ressalta-se que para o cálculo do coeficiente de divergência genética entre populações (GST) se utiliza todas as médias de HIT e HIS dos locus (Tabela 8).

**Tabela 8:** Estimativas de diversidade gênica total (HIT), diversidade gênica dentro das amostras (HIS), diversidade gênica entre as amostras (DST) e divergência genética entre populações (GST). Cruz das Almas, BA.

Descrição	Total	Média
HIT	234,786	9,391
HIS	33,074	6,614
DST	34,707	3,470
GST	2,070	0,207

Nei (1978) cita que o coeficiente de divergência genética entre populações (GST) varia entre 0 e 1 e expressa a proporção da diversidade total explicada por diferenças entre os grupos. Observa-se que a média do GST entre as gerações das 5 subpopulações simuladas é bem baixo, com o valor de 0,207. Portanto, confirma como o resultado da estatística de Wright, o FST, o qual é um parâmetro equivalente e evidencia que há menor diversidade genética dentro de subpopulações.

### IV. Conteúdo de população polimórfica (PIC)

O conteúdo de informação polimórfica (PIC) quantifica o valor da informação do polimorfismo de um locus marcador, portanto, possui dependência do número de alelos e de suas frequências. A classificação de PIC é determinada como muito informativo, medianamente informativo e pouco informativo. Com valores superiores a 0,5, entre 0,25 e 0,50 e, inferiores a 0,25, respectivamente (BOLSTEIN et al., 1980).

Observa-se na Tabela 9 que os valores médios de PIC nas 5 subpopulações encontram-se entre 0,25 e 0,5, exceto a subpopulação 3 que possui o parâmetro no valor de 0,225. Portanto, o polimorfismo do locus marcador codominante da população simulada é medianamente informativo.

**Tabela 9:** Estimativas de conteúdo de população polimórfica (PIC), de 5 subpopulações de acasalamento ao acaso. Cruz das Almas, BA.

<b>Subpopulação</b>	<b>H médio</b>	<b>PIC médio</b>
Sub 1	0,3534	0,282
Sub 2	0,3542	0,281
Sub 3	0,2794	0,225
Sub 4	0,3282	0,263
Sub 5	0,3385	0,268

## 5. CONCLUSÃO

As metodologias testadas no estudo foram eficientes em detectar diferenciação entre populações simuladas ao passar por processos que afetam as frequências gênicas, como foi o caso da deriva genética.

A comparação entre a heterozigosidade observada ( $H_o$ ) e heterozigosidade esperada ( $H_e$ ) são parâmetros eficientes para diferenciar aumento de heterozigotos em populações sob efeito de deriva genética.

Os índices de fixação de Wright são bons indicadores de diferenciação genética, principalmente, nas inferências sobre o efeito de populações fragmentadas em subpopulações.

## REFERÊNCIAS

AMARAL, L. P. **Geoestatística na caracterização do solo e da vegetação em Floresta Ombrófila Mista**. Dissertação de Mestrado em Agronomia, UNICENTRO, Guarapuava, 2010.

ASSIS, G.M.L. **Efeito do número de genes na avaliação genética utilizando dados simulados**. Dissertação de Mestrado em Genética e Melhoramento). Universidade Federal de Viçosa, Viçosa, 2005.

BOLSTEIN, D.; WHITE, R.L.; SKOLNICK, M.H; DAVIES, R.W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. **American Journal of Human Genetics**, v.32, p.314-31,1980.

BERGOCH, J.; MARTINS, P.S.; ANDO, A.; Distribuição da diversidade isoenzimática e morfológica da mandioca na agricultura autóctone de Ubatuba Sci. agric. vol.57 n.1 Piracicaba Jan./Mar. 2000.

BROWN, A.D.H.; WEIR, B.S. **Measuring genetic variability in plant population**. In: TANKESLEY, S. D., ORTON, T. J. (Eds.). Isoenzymes in plant genetics and breeding. Part A. Amsterdam: Elsevier Science,. p. 219-239,1983.

CANTELMO, N. F.; PINHO, R. G.; PINHO, E. V. R.; CAMARGOS, R. B.; PIRES, L. P. M.; PINHO, I. V. Fluxo gênico em milho sob diferentes tamanhos de amostras e distâncias de amostragem. **Revista Brasileira de Milho e Sorgo**, v.15, n.2, p. 328-334, 2016.

CARNEIRO, A. P. S.; MALHADO, C. H. M.; FILHO, R. M.; CARNEIRO, A. P. S.; SILVA, F. F.; TORRES, R. A. A raça Indubrasil no Nordeste brasileiro: melhoramento e estrutura populacional. **Revista Brasileira de Zootecnia**, v.38, n.12, p.2327-2334. 2009.

CONTE, Rudimar. **Estrutura genética de populações de Euterpe edules MART. submetidas à ação antrópica utilizando marcadores alozímicos e microssatélites**. . Tese de Doutorado em Agronomia). Universidade de São Paulo, Piracicaba,2004.

CRUZ, C. D.; FERREIRA, F.M.; PESSONI, L. A. **Biometria aplicada ao estudo da diversidade genética**. Edição 1. Viçosa. Editora UFV. 528 p, 2008.

Cruz, C.D. GENES - a software package for analysis in experimental statistics and quantitative genetics. Acta Scientiarum. v.35, n.3, p.271-276, 2013

ESTOPA, R. A.; SOUZA, A. M. S.; MOURA, M.C. O; BOTREL, C. G.; MENDONÇA, E. G.; CARVALHO, D. Diversidade genética em populações naturais de candeia (*Eremanthus erythropappus* (DC.) MacLeish). **Scientia Florestalis**, Lavras. v..70, p. 97-106, 2006.

EXCOFFIER, L. **Analysis of population subdivision**. IN: BALDING, D.J., BISHOP, M., CANNINGS, C. (eds). Handbook of statistical genetics. Chichester (UK): John Wiley & Sons. 2001, p. 271-307.

FERREIRA, Fabio. Medeiros. **Diversidade em populações simuladas com base em locos multialélicos**. Tese de Doutorado em Genética e Melhoramento. UFV, Viçosa, , 2007.

FURLAN, R. A.; MORI, E. S.; TAMBARUSSI, E. V.; MORAES, C. B.; JESUS, F. A.; ZIMBACK, L. Estrutura genética de populações de melhoramento de *Pinus caribaea* var *hondurensis* por meio de marcadores microssatélites. **Bragantia**, Campinas, v.66, n.4, p.553-563, 2007.

GOUDET, J. FSTAT (Version 1.2): a computer program to calculate F-Statistics. **The Journal of Heredity**, v.86, n.6) p.485-486, 1995.

HARTL, D. L.; CLARK, A.G. Principles of population genetics. **Evolution**, v.25, p. 276-280, 1971.

HARTL, D. L.; CLARK, A. G. Princípios de Genética de Populações. 4 ed. Porto Alegre: Artmed, 2010. 660 p.

HIGUCH, P.; REIS, M.G.F.; PINHEIRO, A.L.; SILVA, C.T.; OLIVEIRA, C.H.R. Composição florística da regeneração natural de espécies arbóreas ao longo de oito anos sem um fragmento de floresta estacional semidecidual, em Viçosa, MG. *Rev. Árvore* vol.30 no.6 Viçosa Nov./Dec. 2006

JAEGER, Peterson. **Caracterização genética e demográfica de populações *Xylopia emarginata* Mart. (Anonaceae)**. Dissertação de Mestrado em Engenharia Florestal. UFLA, Lavras, 2004.

JUNIOR, A. F. M.; CARVALHO, D.; PÓVOA, J. S. R.; BEARZOTI, E. Estrutura genética de populações naturais de pequizeiro (*Caryocar brasiliense* Camb.). **Scientia Florestalis**, n.66. p. 56-65, 2004.

KAGEYAMA, P.Y.; GANDARA, F. B. Consequências genéticas da fragmentação sobre populações de espécies arbóreas. **Série técnica ipef**, v. 12, n. 32, p. 65-70, 1998.

LIMA, R. A.; LOPES, M. T. G.; BENTES, J. L. S.; VALENTE, M. S. F.; PEREIRA, J. O.; MUNIZ, G. I. B. Diversidade e estrutura genética de *Senna reticulata*. **FLORESTA**, Curitiba, PR, v. 45, n. 3, p. 507-514, 2015.

MACIEL, Kelly de Jesus Silva. **Análise da diversidade e divergência genética em clones de *Eucalyptus* spp. potencialmente importantes para Goiás**. Dissertação de Mestrado em Genética e Melhoramento de plantas, Universidade Federal de Goiás, Goiânia, 2014.

MANLY, B. F. J. Randomization and regression methods for testing for associations with geographical, environmental and biological distances between populations. **Researches on Population Ecology**, v.28, n.2, p.201-218, 1986.

MANTEL, N. The detection of disease clustering and a generalized regression approach. **Cancer Research**, v.27, p.209-220, 1967.

MARTINS, P. S. Estrutura populacional, fluxo gênico e conservação “in situ”. **IPEF**, n.35, p.71-78, 1987.

McMANUS, C.; PAIVA, S.; CORRÊA, P. S.; SEIXAS, L.; MELO, C. B. **Estatísticas para descrever Genética de Populações**. INCT: Informação Genético-Sanitária da Pecuária Brasileira. (Série Técnica: Genética), 2011

MENEZES, M. P. C.; MARTINEZ, A. M.; RIBEIRO, M. N.; FILHO, E. C. P.; BERMEJO, J. V. D. Caracterização genética de raças caprinas nativas brasileiras utilizando-se 27 marcadores microssatélites. **Revista Brasileira de Zootecnia**, v.35, n.4, p.1336-1341, 2006.

MUNIZ, J. A.; CAMARGO, M. S.; FERREIRA, D. F.; VEIGA, R. D. Métodos de estimação do coeficiente de endogamia em uma população diplóide com alelos múltiplos. **Ciênc. agrotec.**, Lavras, v. 32, n. 1, p. 93-102, 2008.

NEI, M. Genetic distance between populations. **American Naturalist**, Chicago, v. 106. p. 283-292, 1972.

NEI, M. 1978. *Molecular Evolutionary Genetics*. Columbia University Press, New York, 512 pp.

NICHOLAS, F. W. **Introdução à genética veterinária**..3.ed. Porto Alegre: Artmed, 347 p, 2011.

ODUM, E.P. **Ecology**. London: Holt, Rinehart and Winston, 1975. 244p.

OLIVEIRA, Ana Maria Cruz **Associação das abordagens de genética quantitativa e de genética de populações no melhoramento de plantas**. Tese de Doutorado em Genética e Melhoramento). UFV, Viçosa, 2015.

PEREZ, Cristhian Carmelo Mendoza. Medidas de diferenciação em populações simuladas sob endogamia e seleção divergente. Dissertação de Mestrado em Biologia Celular e Estrutural). UFV, Viçosa, 2008.

QUEIROZ, S. Q.; ALBUQUERQUE, L. G.; LAZONI, N. A. Efeito da Endogamia sobre Características de Crescimento de Bovinos da Raça Gir no Brasil. **Rev. bras. zootec.**, v.29, n.4, p.1014-1019, 2000.

RAPOSO, A.; MARTINS, K.; CIAMPI, A. Y.; OLIVEIRA, W.; VEASEY, E. A. Diversidade genética de populações de andiroba no Baixo Acre. **Pesq. agropec. bras.**, Brasília, v.42, n.9, p.1291-1298, 2007.

SÁNCHEZ, Carlos Filipe Barreira. **Diversidade entre e dentro de populações simuladas sob deriva genética**. 2008. 85 f. Dissertação de Mestrado em Genética e Melhoramento, UFV, Viçosa, 2008.

SHANNON, C. E.; WEAVER, W..**The Mathematical Theory of Communication**. Urbana, University of Illinois Press, 117 p, 1949.

TELES, M. P. C.; VALVA, F. D.; BANDEIRA, L. F.; COELHO, A. S. G. Caracterização genética de populações naturais de araticunzeiro (*Annona crassiflora* Mart. - Annonaceae) no Estado de Goiás. **Revista Brasileira de Botânica**, v.26, n.1, p.123-129,2003.

VASCONCELOS, M. E. C.; RODA, E. A. Análise multivariada de dados moleculares de Rizóbio-Phaseolus isolados de nódulos de feijoeiro. **Ciências Exatas e Tecnológicas**, v.5, p. 53-57, 2006.

WRIGHT, S. The interpretation of population structure by F-statistics with special regard to system of mating. **Evolution**, Lawrence, v. 19, p.395-342, 1965.

ZANELLA, Lisiane. **Análise da interferência antrópica na fragmentação da mata atlântica e modelos de simulação da paisagem na microrregião da Serra da Mantiqueira do estado de Minas Gerais.** Dissertação de mestrado, Universidade Federal de Lavras, Lavras, 2011.