

**CARLOS FELIPE BARRERA SÁNCHEZ**

**DIVERSIDADE ENTRE E DENTRO DE POPULAÇÕES  
SIMULADAS SOB DERIVA GENÉTICA**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Magister Scientiae*.

**VIÇOSA  
MINAS GERAIS – BRASIL  
2008**

**CARLOS FELIPE BARRERA SÁNCHEZ**

**DIVERSIDADE ENTRE E DENTRO DE POPULAÇÕES  
SIMULADAS SOB DERIVA GENÉTICA**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Magister Scientiae*.

APROVADA: 15 de Dezembro de 2008.

---

Prof. Marcos Ribeiro Furtado  
(Co-orientador)

---

Prof. Pedro Crescêncio Souza Carneiro  
(Co-orientador)

---

Prof. Paulo Roberto Cecon

---

Prof. Adesio Ferreira

---

Prof. Cosme Damião Cruz  
(Orientador)

## AGRADECIMENTOS

A Deus, por sempre se fazer presente me iluminando todos os dias.

Aos meus pais, Carlos Barrera e Magnólia Sánchez por todos os ensinamentos e conselhos, e os meus irmãos Alex e Luis, pelo apoio, amizade e compreensão em todos os momentos.

Aos meus tios Fernando, Cecília e William pelo apoio a todas as minhas idéias.

Ao professor Cosme Damião Cruz, pela amizade, paciência, confiança e incentivo.

À Universidade Federal de Viçosa, pela oportunidade de realização deste curso.

Aos amigos do laboratório de Bioinformática, Moisés, Livia, Danielle, Márcio, Leonardo, Carol e Rafael pelo convívio agradável durante a realização deste curso.

À Maria, uma grande amiga, minha irmã de coração.

À Blanca, minha namorada por todos os momentos felizes e o apoio nos dias difíceis.

Aos grandes amigos feitos em viçosa, Caio, Geovanny, Fábio, Diogo, Gustavo, Sandra e Eder, pela convivência, paciência, cumplicidade e companheirismo.

Aos amigos colombianos Natalia, Juliana, Carolina, Jorge, Gabriel, Enrique e Alex que longe de nosso país foram um grande apoio e companhia,

Aos professores de graduação e de pós-graduação, pela atenção, pela disponibilidade e pelos ensinamentos.

Às secretárias do curso de pós-graduação em genética e melhoramento, Rose e Rita, pelo apoio, dedicação, atenção e amizade.

## SUMÁRIO

	Página
RESUMO.....	v
ABSTRACT.....	vii
1. INTRODUÇÃO.....	1
2. REVISÃO DE LITERATURA.....	3
2.1. Efeito Deriva genética .....	3
2.2 Propriedades de populações finitas.....	8
2.3. Simulação de dados genéticos.....	10
2.4. Medidas de Diversidades entre e dentro populações.....	11
2.4.1 Índice de diversidade - Heterozigose e Endogamia .....	12
2.4.2 Índice de diversidade - Shannon-Wiener.....	13
2.4.3 Estatística <i>H</i> de Nei.....	14
2.4.3.1 Medidas de Identidade genética.....	15
2.4.3.2 Medida de Heterozigosidade.....	16
2.4.3.3 Medida de diversidade.....	18
2.4.3.4 Coeficiente de diversidade relativa entre grupos.....	19
2.4.4 Índices de Fixação – Estatística <i>F</i> de Wright.....	19
2.4.5 Heterozigosidade – Weir (1996).....	24
2.4.6 Análise molecular de variância – AMOVA.....	26
2.5 Diversidade genética.....	32
3. MATERIAL E MÉTODOS.....	37
3.1 Material.....	37
3.2 Informações do conjunto de dados.....	41
3.3 Análises de diversidade genética.....	42
4. RESULTADOS E DISCUSSÃO.....	44
4.1 Distribuição das frequências gênicas.....	44
4.2 Diversidade Genética Em Nível Intrapopulacional.....	51
4.2.1 Índice de diversidade - Endogamia e Heterozigose.....	51

4.2.2 Índice de diversidade - Shannon-Wiener.....	54
4.3 Diversidade Genética Em Nível Interpopulacional.....	57
4.3.1 Estatística F de Wright (1951).....	57
4.3.2 Heterozigosidade de Weir (1996).....	61
4.3.3 Análise de Variância Molecular (AMOVA).....	65
4.3.4 Estatística de Nei (1973).....	69
5. CONCLUSÕES.....	73
6. REFERÊNCIAS BIBLIOGRÁFICAS.....	75

## RESUMO

BARRERA SÁNCHEZ, Carlos Felipe, M.Sc., Universidade Federal de Viçosa, dezembro de 2008. **Diversidade entre e dentro de populações simuladas sob deriva genética.** Orientador: Cosme Damião Cruz. Co-orientadores: Pedro Crescêncio Souza Carneiro e Marcos Ribeiro Furtado.

O efeito da deriva genética sobre as populações de pequenos tamanhos tem sido objeto de diversos estudos, apresentando efeito sobre a composição genética das populações. A amostragem dos gametas ao acaso dentro de pequenas populações com perda de diversidade genética e fixação de alelos, tem conseqüências de grande significado para a conservação da diversidade genética e das espécies. A diversidade genética é fundamental para que ocorra a evolução. A seleção natural atua entre as variantes dentro das populações em relação à adaptação ao ambiente, proporcionando variação entre populações e, por fim, variação entre espécies. A manutenção da variabilidade genética em populações é a base da conservação de espécies e, portanto, a descrição e o estudo do efeito da deriva genética sobre as populações se tornam de grande importância. Deste modo, o objetivo deste trabalho foi analisar o efeito da deriva genética sobre a diversidade genética de diferentes populações simuladas de diferentes tamanhos. Para o estudo da diferenciação genética dentro e entre populações sujeitas ao efeito da deriva genética, foram simuladas 4000 populações originadas de marcas moleculares co-dominantes. As amostras geradas foram de 20, 50, 100 e 200 indivíduos com 100 subpopulações submetidas a 10 gerações de acasalamento ao acaso com base em 10 locos. As populações foram avaliadas por meio dos seguintes índices de diversidade genética entre e dentro: Endogamia e Heterozigosidade, Distância Shannon – Wiener (1978), Índice de diversidade de Nei (1973), Índice de Fixação de Wright (1951, 1978), Heterozigosidade de Weir (1996), Análise de variância molecular (AMOVA) de Excoffier, et al., (1992). Conclui-se que o efeito da deriva genética alterando a frequência gênica depende do tamanho da população. Quanto menor a população, maiores são os efeitos da amostragem, evidenciados pelas estatísticas  $G_{ST}$ ,

$F_{ST}$  e  $\phi_{ST}$ , que aumentam seu valor com o avanço das gerações, indicando perda de variação genética dentro e aumento da divergência genética entre as populações.

## ABSTRACT

BARRERA SÁNCHEZ, Carlos Felipe, M.Sc., Universidade Federal de Viçosa, December, 2008. **Diversity among and within populations under simulated genetic drift.** Orientador: Cosme Damião Cruz. Co-orientadores: Pedro Crescêncio Souza Carneiro and Marcos Ribeiro Furtado.

The effect of genetic drift on populations has been the subject of many studies due to a reduction in the effective population size. It may seem that the case has only effect on the genetic composition of populations. Meanwhile, a random sampling of gametes in small populations has significant consequences for the conservation of genetic diversity, as the loss of genetic diversity and fixation of alleles within the population. Genetic diversity is fundamental to the evolution. The natural selection operates between the variations within populations to adjust to the environment, providing variation among populations and, finally, variation among species. Thus, the higher the genetic variability exists in the population, the greater the chances of exploitation by breeders in their work. The maintenance of genetic variability in populations is the basis for the conservation of species and hence the description and study of the effect of genetic drift which appears on it, is reason for the search. Thus, the research aimed to analyze the effect of genetic drift on the genetic diversity of different populations. For that were generated and analyzed data from the simulation of populations. One of the key factors to evaluate the effect of genetic drift is the size of populations to be worked. To study and acquisition of genetic differentiation within and between populations under the effect of genetic drift, 4000 were simulated populations originated from molecular markers co-dominant. The samples were generated, 20, 50, 100 and 200 individuals with 100 sub-populations subjected to 10 generations of random mating with emphasis on 10 loci. The populations were evaluated with different indices of genetic diversity between and within: heterozygosity and inbreeding, Distance Shannon - Wiener (1978), Index of diversity of Nei (1973), Index of Fixing Wright (1951, 1978), heterozygosity for Weir (1996), Analysis of molecular variance (AMOVA) of Excoffier, et al., (1992). We conclude that the effect of genetic drift by changing the gene frequency depends on the size of the population. The lower the population, most obvious are the effects of random sampling, the statistics  $G_{ST}$ ,  $F_{ST}$  and  $\phi_{ST}$  increase their value to the advancement of generations, a result of loss of genetic variation within and genetic divergence among populations.





## 1. INTRODUÇÃO

A variação genética das espécies cultivadas é fortemente reduzida pelos programas de melhoramento. Para a conservação e manutenção da biodiversidade, uma opção é contar com coleções de germoplasmas representativas de cada espécie. Entretanto, nessas coleções, a variação genética é frequentemente reduzida em função do tamanho das amostras mantidas, geralmente reduzido em função de custos, escassez de área e mão de obra.

Em muitas espécies cultivadas, a variação genética da população base é muito limitada pelo pequeno tamanho populacional ou pela seleção. Portanto, as populações apresentam pequeno número de indivíduos ou pequeno tamanho efetivo, mostrando níveis reduzidos de diversidade e levando a acontecimento da deriva genética (Moreira, 1999).

Em populações naturais, a distribuição da variação genética é influenciada pelo modo de reprodução, sistema de acasalamento, tamanho da população, distribuição geográfica e forças evolutivas. A mutação, a seleção natural, a migração e a deriva genética são as forças evolutivas que alteram o conjunto gênico da população. Elas, em geral, isoladamente, não alteram a estrutura genética, mas em conjunto atuam compensando ou complementando uma a outra (Hamrick & Godt, 1990).

A deriva genética refere-se às alterações aleatórias das frequências alélicas de uma população devido ao tamanho da amostra, levando eventualmente a fixação ou a perda do alelo. À medida que as subpopulações se distanciam em frequências alélicas, tomam-se também diferentes pelas frequências genotípicas, tendo em geral importantes conseqüências. Ocorre redução da variação genética dentro de uma pequena população e diferenciação entre as subpopulações. A amostragem pode levar ao aumento da frequência de alelos deletérios, recessivos ou dominantes, trazendo graus prejuízos às

espécies. De maneira que uma amostra em particular pode não ser representativa da população de origem (Moreira, 1999).

Diferentes medidas biométricas têm sido propostas e utilizadas para o estudo da diversidade genética, visando elucidar os efeitos da deriva genética, quantificar ou prever o nível de variação genética existente e verificar sua distribuição entre e ou, dentro de populações ou espécies. O estudo da eficácia destas metodologias é fundamental para que se possa, em estudos futuros, adotar a mais conveniente.

O conhecimento do padrão de variação genética entre e dentro de populações introduz o conceito de estrutura genética de populações. Essa pode ser caracterizada a partir do estudo de amostras de diferentes populações ou mesmo de diferentes subpopulações da mesma população.

Técnicas de simulação desempenham papel importante na aplicação genética e no desenvolvimento de novos métodos estatísticos para análise da diversidade genética, definindo processos da história, comportamento e estrutura populacional e, ainda, tornando possível a predição de frequências alélicas. Este trabalho teve como objetivo avaliar a variação genética dentro e entre populações simuladas sob deriva genética e avaliar a eficiência de métodos biométricos na quantificação da diversidade genética.

## 2. REVISÃO DE LITERATURA

### 2.1 Efeito Deriva Genética

A deriva genética refere-se às alterações nas frequências alélicas de uma população devido ao tamanho da amostra. Desde que essas alterações da frequência alélica sejam cumulativas e relativamente permanentes, a deriva pode ocorrer não somente em populações permanentemente pequenas, mas também naquelas que ficam pequenas só periodicamente. Também pode ocorrer nas que são pequenas desde sua fundação, isto é, as constituídas por poucos imigrantes, ou fundadores, e que carregam uma pequena porção de variação genética de uma população maior (Madsen et al., 1999).

O efeito da deriva genética sobre as populações tem sido objeto de diversos estudos devido à diminuição do tamanho efetivo populacional, tendo como consequência a perda da variação genética e a redução da heterozigosidade. A perda de variação genética resulta na redução da habilidade das populações de se adaptarem a mudanças ambientais, levando muitas espécies a atingirem uma limitação evolutiva, tornando este efeito muito importante em estudos de composição genética de populações (Barret & Kohn, 1991; Heldrick & Miller, 1992; Ellstrand & Elam, 1993; Frankhan, 1996).

Quando uma amostra é adequada espera-se que a frequência alélica estimada ( $\hat{p}$ ), tendo-se por base o número de ocorrência de cada classe genotípica na amostra, reflète a frequência alélica da população original ( $p$ ), de forma que se tenha:

$$E(\hat{p}) = p$$

$$E(\hat{q}) = E(1 - \hat{p}) = 1 - E(\hat{p}) = 1 - p = q$$

Entretanto, as várias amostras de uma população apresentarão diferentes valores de  $\hat{p}$ , dado que a ocorrência do alelo A poderá ser de diferente magnitude nas várias amostras.

Admitindo que numa população original, de tamanho N, encontram-se  $n_1$  genótipos (AA),  $n_2$  (Aa) e  $n_3$  (aa), tem-se que:

$$f(A) = p = \frac{2n_1 + n_2}{2n}$$

e

$$f(a) = q = \frac{2n_3 + n_2}{2n}$$

Ao ser tomado, desta população original, uma amostra de tamanho n ( $n < N$ ) e admitido que ela seja originada do encontro entre 2n gametas, dos quais n originaram dos genitores femininos e os outros n, dos genitores masculinos, tem-se que a relação genotípica esperada (RGe) na amostra será:

$$RGe = [p + q]^{2n}$$

Na amostra, que tem n indivíduos e, portanto, 2n alelos, a ocorrência do alelo A poderá ser quantificada pela variável x, de distribuição binomial, que assume os seguintes valores: 0,1,2...,2n. Tendo x distribuição binomial, verificam-se as propriedades:

$$\bar{x} = E(x) = 2np$$

$$V(x) = 2npq$$

Para uma particular amostra, a frequência do alelo A será quantificada por meio de:

$$\hat{p} = \frac{x}{2n} \quad (0 \leq \hat{p} \leq 1)$$

portanto,

$$E(\hat{p}) = E\left(\frac{x}{2n}\right) = \frac{2np}{2n} = p$$

e

$$V(\hat{p}) = V\left(\frac{x}{2n}\right) = \frac{V(x)}{4n^2} = \frac{pq}{2n}$$

A probabilidade de que  $\hat{p}$  assumira um particular valor, por exemplo, igual a  $k/2n$ , poderá ser calculada por meio de:

$$P\left(\hat{p} = \frac{k}{2n}\right) = \frac{2n!}{k!(2n-k)!} \hat{p}^k (1-\hat{p})^{2n-k}$$

A variação esperada na frequência gênica da população pelo processo de amostragem pode, então, ser quantificado por:

$$\Delta p = p - \hat{p}$$

e

$$V(\Delta p) = V(p - \hat{p}) = \frac{pq}{2n}$$

Deve ser lembrado que  $n$  é o número de indivíduos da população original que efetivamente contribuíram para a formação da amostra. Muitas vezes o que se procura é exatamente conhecer o valor do tamanho efetivo da amostra.

A subdivisão de uma população pode ser provocada, por causas geográficas ou ecológicas, sob condições naturais, ou condições experimentais utilizando-se acasalamentos controlados. Todas as subpopulações, em conjunto, constituem a população completa e cada subpopulação é uma “pequena população”, na qual as frequências alélicas estão sujeitas a deriva genética.

O efeito da deriva genética é inversamente proporcional ao tamanho da subpopulação ( $N_e$ ). Isso significa que quanto maior o tamanho da subpopulação, maior é a probabilidade das frequências gênicas da amostra estar próximas daquelas da

população de origem. À medida que a deriva genética ocorre, a variação entre as populações aumenta de geração a geração. Suponha, por exemplo, que algumas colônias tenham derivado de  $p=0,50$  até  $p=0,55$  (aproximadamente o mesmo número terá derivado de 0,50 a 0,45). Entre estas, algumas irão ter sua frequência gênica aumentada e, outras, diminuída com uma variância de  $(0,55)(0,45)/2n$ , havendo, portanto, uma expansão na gama de frequências gênicas. Após algum tempo, a distribuição das frequências gênicas será tão ampla que todas as frequências gênicas possíveis entre 0 e 1 estarão igualmente representadas entre as populações e algumas destas atingirão as frequências 0 e 1. Quando  $p = 1$  diz-se que o alelo *A* está fixado e a população é monomórfica, *AA*. Quando  $p = 0$ , o alelo *a* é que foi fixado. Entretanto, se um alelo foi fixado, o outro não poderá reaparecer, exceto por uma nova mutação ou através do fluxo gênico de outras populações (Futuyma, 2003).

O processo da deriva genética pode ser encarado como uma flutuação aleatória na frequência alélica, levando, eventualmente, à fixação ou à perda do alelo. O processo ocorre em toda população finita, mas é tanto mais rápido quando menor a população. A deriva genética apresenta algumas consequências evolutivas, duas das quais merecem ênfase especial: a deriva genética resulta na perda de variação genética dentro das populações e na divergência genética entre elas (Futuyma, 2003). A fragmentação de áreas de floresta, pelo avanço da agricultura e de outras atividades antrópicas, leva ao isolamento de subpopulações e, conseqüentemente, à deriva genética e à endogamia (Robinson, 1998).

Três importantes estudos, todos com *Drosophila melanogaster*, que comprovam os efeitos da deriva em pequenas populações, foram realizados por Kerr & Wright (1954). No primeiro, avaliaram 96 subpopulações constituídas de 8 indivíduos tomados ao acaso (quatro machos e quatro fêmeas). Cada subpopulação possuía um gene para cerdas 'forked' e o seu alelo selvagem (normal), ambos com frequência 0,5. Após 16 gerações de cruzamentos aleatórios os alelos estavam segregando apenas em 26 das subpopulações, sendo que o alelo 'forked' havia se fixado em 29 e o selvagem em 41 das subpopulações. A taxa de fixação de ambos alelos alcançou equilíbrio a partir da quarta geração em torno de 8,9%.

No segundo, Wright & Kerr (1954) estudaram a oscilação do gene *bar* em 108 subpopulações de *D. melanogaster*, cada uma de tamanho 8 (quatro machos e quatro fêmeas), durante dez gerações. Iniciando com uma frequência gênica de 0,50, uma

aparente estabilidade na dispersão das frequências gênicas ocorreu na quarta geração, com uma taxa de fixação de 22% para o tipo selvagem e 0,7% para o tipo *bar*. Após dez gerações, 95 subpopulações haviam fixado alelo selvagem e 3 o tipo *bar*, o que demonstra que houve uma forte seleção contra o gene *bar*. No terceiro, estudaram a oscilação dos genes *aristopedia* e *spineless* em 113 subpopulações dessa mesma espécie, sendo iguais ao estudo anterior o número de gerações e o tamanho efetivo das subpopulações. A frequência média do gene *aristopedia* oscilou de 0,50 na geração 1 a 0,3544 na geração 5, e nas últimas três gerações a frequência média foi de 0,4032. A variância alcançou 0,0235 na segunda geração e flutuou em torno deste valor nas gerações seguintes. Além da rápida estabilização na distribuição das frequências gênicas, os autores concluem que houve uma forte pressão de seleção contra os homozigotos.

Para Shah et al., (2008), uma possível explicação para a obtenção de altos índices de diferenciação genética entre populações de *T. fuana* seja o fluxo gênico e a deriva genética, uma vez que a dispersão de sementes da espécie é via eólica e por vertebrados. A distância genética de Nei ( $D$ ) variou de 0,018 a 0,3514 o que sugeriu um alto nível de diferenciação genética entre pares de populações. Uma significativa diferença genética entre populações também foi detectada pela análise de AMOVA. Pela análise de variância molecular obteve-se estimativa de  $\phi_{ST} = 0,5685$ , revelando que 56,85% da variação foi atribuída à diversidade entre populações e o restante, 43,15%, à diversidade dentro de populações. Entretanto, o fluxo gênico estimado ( $N_m$ ) foi de 0,3558 indivíduos por geração entre populações, considerado, segundo os autores, baixo fluxo gênico. Portanto, pela limitada dispersão de polens e sementes entre populações, unidos aos altos níveis de diferenciação genética entre populações, indica que o mais provável é que a deriva genética esteja afetando a estrutura genética de populações e ocasionando o aumento da diferenciação entre populações da espécie. Com evidências de populações pequenas e isoladas, a fragmentação do habitat, da população original, levou os autores a acreditar que a formação de entidades isoladas, pode estar dificultando o fluxo gênico entre populações. Conseqüentemente, tais fatos os levaram também a concluir que o principal fator causador da perda da diversidade genética dentro de entidades e do aumento da diferenciação genética entre as entidades populacionais seja a deriva genética.

Moraes & Carvalho (2003) estudaram a diferenciação genética de populações de *Cryptocarya aschersoniana* Mez, por meio da abordagem introduzida por Nei (1973) da análise da diversidade gênica em populações subdivididas. Observou-se que a diversidade gênica total ( $H_{IT}$ ), que é essencialmente a heterozigosidade esperada, quando as populações são agrupadas e de cruzamentos aleatórios, foi 0,552 e o coeficiente de diferenciação gênica ( $G_{ST}$ ) foi 0,340. Por sua vez, a diversidade gênica dentro das populações ( $H_{IS} = 0,365$ ) foi responsável por 66,12% da diversidade gênica total, indicando a existência de uma maior variabilidade ocorrendo dentro das populações do que entre as mesmas.

A divergência obtida através das estimativas de  $G_{ST}$  para as populações de *Cryptocarya aschersoniana* sugere a existência de deriva genética significativa e de efeitos de seleção natural entre populações devido ao fato de ser bastante superior ao valor médio encontrado para espécies tropicais com sistema reprodutivo alogâmico, 11,8%. (Loveless & Hamrick, 1987), o que pode ocorrer devido às baixas taxas de fluxo gênico entre as populações (Wright, 1965), conforme indicação feita por Moraes & Carvalho (2003) para as mesmas populações analisadas. Conseqüentemente, estas populações podem se enquadrar na situação em que a deriva genética sobrepujaria os efeitos do fluxo gênico (Moraes & Carvalho, 2003).

## **2.2 Propriedades de Populações Finitas**

A estrutura populacional de uma espécie é definida como o conjunto de características genéticas e demográficas resultantes da ação e das interações de vários mecanismos evolutivos e ecológicos. Do ponto de vista ecológico, a estrutura genética é caracterizada pela densidade populacional, natureza das relações entre os indivíduos e o ambiente e as interações entre indivíduos e populações locais. Já do ponto de vista genético evolutivo, a estrutura genética é caracterizada pela variabilidade morfológica e quantitativa existentes entre indivíduos, sistema reprodutivo, padrões de fluxo gênico e estratégias adaptativas aos ambientes locais (Martins, 1987).

O teorema de Hardy-Weinberg, a base da teoria genética da evolução, presume que uma população tem tamanho infinito e que todos os indivíduos na população se acasalam aleatoriamente. Quando ocorrem desvios das proporções esperadas pelo

Equilíbrio de Hardy-Weinberg significa que alguma das pressuposições foi violada e, então, são propostos modelos alternativos para a verificação das possíveis causas do desvio. As principais alterações nas frequências alélicas nas populações naturais podem ser devidas a processos sistemáticos como mutação, fluxo gênico (migração) e seleção natural ou processo dispersivo, como a deriva genética (Falconer & Mackay, 1997).

Toda população é uma entidade finita, por maior que seja; dessa forma, portanto, o máximo que se pode ter, em condições naturais, são estados próximos ao equilíbrio. Ademais os indivíduos de uma dada geração são formados de  $2n$  gametas da geração precedente, e esses somente representam exatamente o reservatório gênico da população quando  $N$  é infinito. Conclui-se, pois, que os gametas carregam apenas uma amostra do reservatório gênico de uma população e, por isso, estão sujeitos a erros de amostragem quando  $n$  é finito. Os gametas, que transmitem genes para a geração seguinte, carregam a amostra de genes da geração parental e se a amostra não for suficientemente grande, as frequências alélicas estarão sujeitas a modificações entre esta geração e a seguinte. Quanto menor o tamanho da população e, portanto, o tamanho da amostra de gametas, maior fica sendo a influência da variação devida à amostragem.

O acasalamento ao acaso pode ser aplicado a sistemas de acasalamento onde cada membro da população tem a mesma probabilidade de produzir descendentes e que cada gameta feminino deve ter igual probabilidade de ser fertilizado por qualquer gameta masculino (Alter et al., 2007). Entretanto, dificilmente isto deve ocorrer em condições naturais, pois sempre existe alguma forma de seleção ocorrendo. Preferências de acasalamento tende a aumentar dramaticamente a proporção de homozigose (baixa diversidade) na população. Nos grupos em que predominam o comportamento de harém, a maioria das fêmeas são fecundadas por um mesmo indivíduo, e nas espécies territoriais alguns machos contribuem com mais genes que os demais. Acasalamento randômico é aquele que segue proporcionalmente as mesmas frequências de genes na população.

O conhecimento da estrutura genética de populações pode ajudar na explicação da evolução da alta diversidade encontrada nos trópicos (Eguiarte et al., 1992). Vários autores formularam a hipótese de que a alta diversidade de espécies tropicais seria o produto de especiação não adaptativa, resultante de endogamia e deriva genética, provocadas por baixas densidades de indivíduos coespecíficos e autofecundação predominante (Fedorov, 1966). Essa hipótese, mais conhecida como hipótese de Fedorov, tem uma clara relação com a teoria de equilíbrio instável de Wright (1932), em

que a subdivisão de uma população em demes pequenos, havendo baixo fluxo gênico entre os mesmos, pode resultar em uma rápida diferenciação genética devido à deriva, levando a uma eventual especiação simpátrica (Bawa, 1992).

### **2.3 Simulação de Dados Genéticos**

Na genética são enfatizados temas relativos à evolução, à heredabilidade e ao melhoramento das espécies. Estudos genéticos em geral envolvem grande conjunto de dados, compostos por variáveis (características fenotípicas, multicategóricas, marcadores moleculares e, ou, geográficas) e unidades taxonômicas (espécies, populações, cultivares, famílias, indivíduos etc.). A complexidade de um estudo genético aumenta quando informações adicionais sobre a espécie são requeridas, além de o pesquisador ter que saber quais metodologias de análise genética estão disponíveis e se mostram mais adequadas. Também é comum a adoção de modelos matemáticos capazes de caracterizar o fenômeno biológico e, assim, prover estimativas de vários parâmetros genéticos. Diante deste cenário, o uso da informática passou a ser ferramenta imprescindível ao processamento e gerenciamento de dados.

A importância da simulação ganhou grande espaço no âmbito científico. Impulsionada pela informática, a simulação não se limita apenas a modelos que representam a entidade a ser investigada e sim a uma metodologia para avaliação destes. Técnicas de reamostragem e permutação e alguns algoritmos são em essência processos de simulação que em geral visam a estimação de parâmetros e definições de distribuições. Em estudos de genética de populações com o uso de marcadores, freqüentemente utiliza-se de métodos de reamostragem para a estimação de parâmetros genéticos populacionais e seus respectivos desvios-padrão (Carlini et al., 2003).

Obviamente que o estudo de simulação somente se justifica se soluções analíticas não existem ou o grau de dificuldades e o número de variáveis envolvidas não permitem a realização de inferências adequadas sobre o problema. O uso da simulação deve ser encarado pelos geneticistas e melhoristas como uma alternativa para a resolução de seus problemas, detecção e comprovação de técnicas estatísticas mais eficientes, comparações metodológicas entre outras (Ferreira, 2005).

A área de simulação em genética permitiu a criação de novos métodos e, conseqüentemente, *softwares*, com simuladores específicos e capazes de realizar estudos didáticos e pesquisas, a exemplo do programa GENES (Cruz, 2006a), PowerMarker (Liu & Muse, 2005) e GDA (Lewis & Zaykin, 2002).

## 2.4 Medidas de Diversidades Entre e Dentro de Populações

A genética de populações descreve e analisa a variação genética dentro e entre populações a partir das frequências alélicas e genotípicas. A interpretação desta variação nos organismos tem como princípio o Equilíbrio de Hardy-Weinberg, o qual estabelece que nas populações panmíticas e na ausência de forças evolutivas como migração, seleção, deriva genética e mutação, as frequências alélicas permanecem constantes de geração em geração, o que permite prever as frequências genotípicas a partir das alélicas. Descrever a estrutura genética das populações de organismos é um passo prévio para entender quais forças estão envolvidas na evolução das mesmas.

Para que os geneticistas, ou os melhoristas, possam inferir sobre a variação genética das populações, diversas metodologias de análise de dados genéticos foram desenvolvidas. Trata-se de um processo contínuo e que caminha juntamente com os avanços na área da biologia molecular e da informática. Elas podem ser caracterizadas pelo nível de inferência a que se destinam, ou seja, dentro e, ou, entre as unidades taxonômicas.

É comum a utilização de estatísticas descritivas nos estudos de populações, pois elas dão uma idéia inicial do polimorfismo genético de uma população. Parâmetros como: número de alelos por loco ou total, número de alelos efetivos, raros, exclusivos, conteúdo médio de informação polimórfica (PIC), heterozigidade observada e esperada, coeficiente de fixação/endogamia ( $f$ ), riqueza genotípica (índice Shannon-Wiener), entre outras, são alguns dos descritores que ajudam a caracterizar uma população.

Uma das medidas mais importantes para determinar a diversidade genética é a heterozigidade, que é uma medida geral da variação por loco em uma população. Imagine um loco no qual dois alelos ( $A$  e  $a$ ) estão presentes na população em que a frequência de  $A$  é  $p$ , a frequência de  $a$  é  $q$ , a heterozigidade é definida como a chance

de ser retirados dois alelos diferentes, se genes aleatórios amostrados da população (para um loco). A chance de serem retiradas duas cópias de  $A$  é  $p^2$ , e a chance de se retirar duas cópias de  $a$  é  $q^2$ . A chance total de se retirar duas cópias gênicas idênticas é  $p^2 + q^2$ . Assim, a chance de se retirar dois genes diferentes é 1 menos a chance de retirar dois alelos idênticos (Ridley, 2006), ou seja:

$$H = 1 - (p^2 + q^2)$$

Em geral, uma população pode conter vários alelos em um loco. Os diferentes alelos podem ser distinguidos por números subscritos. Por exemplo, se uma população possui três alelos, suas frequências podem ser escritas por  $p_1$ ,  $p_2$ , e  $p_3$  pode-se simbolizar a frequência do  $i$ -ésimo alelo por  $p_i$  (onde  $i$  possui tantos valores quantos alelos existirem na população) (Ridley, 2006). Nesta situação, tem-se:

$$H = 1 - \sum p_i^2$$

#### **2.4.1 Índice de Diversidade - Heterozigose e Endogamia**

Procura-se, de modo geral quantificar a diversidade genética e o grau de fixação dentro de uma população por meio dos coeficientes de endogamia, heterozigose esperada e observada e do conteúdo médio de informação polimórfica (PIC). Vários autores têm sugerido que a retenção de heterose em gerações sucessivas tem sido proporcional à retenção de heterozigose para a maioria das características de importância econômica. Portanto, conhecendo a heterozigose, tem-se uma estimativa da heterose (Cruz, 2008).

Com base na genética, pode-se inferir que haverá heterose quando houver diferença em frequência gênica entre as populações envolvidas no cruzamento e o efeito de dominância entre alelos não for zero. Se qualquer dessas situações deixarem de existir, a heterose será nula. Isso pode ser mais bem entendido se for considerado que as populações, durante o processo de formação, permanecem geneticamente isoladas e são submetidas a pressões de seleção variáveis, tanto artificial quanto natural. Esse processo resulta em alguma endogamia que, juntamente com a flutuação aleatória na frequência

gênica, contribui para fixação de alguns homozigotos. Esses homozigotos produzidos tanto podem ser de genes com efeitos indesejáveis quanto de genes cuja combinação heterozigótica produz resultados favoráveis (Cruz, 2008).

Assim, parece muito pouco provável que as diferentes populações tenham os mesmos alelos indesejáveis fixados na forma homozigótica. Isso será tanto mais verdade quanto mais distantes na origem e mais separadas espacialmente forem as populações. Dessa forma, ao se cruzarem populações diferentes, as progênies terão os efeitos deletérios dos genes recessivos encobertos pelos genes dominantes e maior taxa de heterozigose (Cruz, 2008).

#### 2.4.2 Índice de Diversidade - Shannon-Wiener

A função de Shannon-Wiener foi proposta para medir a diversidade ou riqueza de espécies em estudos de ecologia, sendo adequada para análise de amostras coletadas ao acaso em grandes comunidades (Brown, 1983). Este método considera a riqueza das espécies, calculando o índice de diversidade, representado pelo número de espécies e a equitabilidade na distribuição de indivíduos entre as espécies. A partir do cálculo de diversidade, pode-se isolar o componente equitabilidade e determinar o quanto a diversidade encontrada para um determinado local difere da diversidade hipotética máxima possível para o mesmo (Jamhour & Laroca, 2004). Também pode ser empregada para medir a diversidade fenotípica ou genotípica de uma população. Define-se a estatística de Shannon-Wiener como:

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

É possível calcular a variância da estatística H' pela seguinte fórmula:

$$V_{H'} = \frac{\sum_{i=1}^c P_i \ln^2(P_i) - \left[ \sum_{i=1}^c P_i \ln(P_i) \right]^2}{n} + \frac{c-1}{2n^2}$$

em que:

c: é o número de classes genotípicas para um dado loco;

$P_i$  é a frequência do i-ésimo genótipo.

n o tamanho da amostra ou o número de indivíduos da população i

### 2.4.3 Estatística *H* de Nei

Nei (1973) propôs um método que possibilita obter a medida do grau de diferenciação genética entre pares de populações. Este método é baseado nas identidades de dois genes escolhidos ao acaso dentro e entre populações e independente do número de alelos. Este autor usou o conceito de heterozigosidade (*H*) gênica ou alélica para quantificar a diversidade genética entre e dentro das unidades experimentais (populações, subpopulações, etc.). Este método independe do padrão de forças evolutivas tais como: mutação, seleção e migração. Ele é aplicável a qualquer organismo, a qualquer sistema reprodutivo e a qualquer nível de ploidia, desde que a frequência gênica possa ser determinada. Nei (1973) afirma que tal método parece ser necessário para uma análise rápida de uma grande quantidade de informações da frequência gênica.

Considerando uma população subdividida em *g* subpopulações, com  $P_{ijk}$  sendo a frequência do *k*-ésimo alelo, do *j*-ésimo loco, na *i*-ésima população pode-se obter, para estas subpopulações medidas de identidade gênica (igual ao complemento aritmético da diversidade gênica), da heterozigosidade e da diversidade genética. Segundo Robinson (1998) a equação de Nei pode ser aplicada a diferentes níveis da classificação hierárquica, como populações dentro de uma região geográfica, subpopulações dentro de populações, indivíduos dentro de subpopulações, etc. utilizando-se em cada caso, as frequências alélicas.

Nei (1973, 1978) determinou que probabilidade de identidade e não identidade gênica de dois genes escolhidos ao acaso é dada por:  $J_i = \sum_k p_{ijk}^2$  e  $H_i = 1 - J_i$ , respectivamente, sendo esta a identidade genética dentro de subpopulação. A probabilidade de não identidade, *H*, portanto, é igual ao complemento aritmético da identidade gênica, o que é usualmente denominado de Heterozigosidade. Nei (1973) ressalta que este conceito, entretanto, não é apropriado para populações sob acasalamentos direcionados.

### 2.4.3.1 Medidas de Identidade Genética

A identidade genética na  $i$ -ésima subpopulação, para um particular loco, é dada por:

$$J_i = \sum_{k=1}^{a_j} p_{ijk}^2$$

Sendo  $a_j$  o número de alelos no  $j$ -ésimo loco.

A identidade genética entre a  $i$ -ésima e a  $i'$ -ésima subpopulações é dada por:

$$J_{ii'} = \sum_{k=1}^{a_j} p_{ijk} p_{i'jk}$$

A partir destes valores podem ser estabelecidas as seguintes estatísticas:

#### Identidade dentro de subpopulações

A identidade genética total dentro das subpopulações, para um loco, é dada por:

$$J_{IS} = \sum_{i=1}^g J_i = \sum_{i=1}^g \sum_{k=1}^{a_j} p_{ijk}^2$$

e o valor médio é denotado por:

$$\bar{J}_{IS} = \frac{J_{IS}}{g}$$

Sendo  $g$  o número de populações

#### Identidade entre subpopulações

A identidade genética total entre os pares de subpopulações, para um loco, é dada por:

$$J_{ST} = \sum_{i=1}^g \sum_{i' < i}^g J_{ii'} = \frac{1}{2} \sum_{i=1}^g \sum_{i' \neq i}^g J_{ii'}$$

e o valor médio é dado por:

$$\bar{J}_{ST} = \frac{2J_{ST}}{g(g-1)}$$

Identidade total das subpopulações

$$J_{IT} = \sum_i J_i + \sum_{i \neq 1}^g \sum_{i'}^g J_{ii'} = J_{IS} + 2J_{ST}$$

e o valor médio é denotado por:

$$\bar{J}_{IT} = \frac{J_{IT}}{g^2} = \frac{1}{g} \bar{J}_{IS} + \frac{g-1}{g} \bar{J}_{ST}$$

### 2.4.3.2 Medida de Heterozigosidade

Nei (1978) apresenta as expressões para o cálculo da diversidade (comumente denominada de heterozigosidade esperada) de uma população. Para um único loco, de uma população de acasalamento ao acaso, a heterozigosidade é definida como:

$$h_{ij} = 1 - \left( \sum_{k=1}^{a_j} P_{ijk}^2 \right)$$

Em que  $P_{ijk}$  é frequência do k-ésimo alelo, para o j-ésimo loco que apresenta  $a_j$  alelos.

Para múltiplos locos, calcula-se a heterozigosidade média de uma população  $i$  ( $H$ ) que é a média aritmética dos valores de  $h$  sobre todos os locos, ou seja:

$$H_i = \frac{1}{L} \sum_{j=1}^L h_{ij}$$

em que  $h_{ij}$  é estimativa de heterozigosidade no j-ésimo loco e  $L$  é o numero de locos amostrados.

Em uma população em equilíbrio de Hardy-Weinberg, a estimativa  $h$  é igual à proporção esperada de heterozigotos. Porém nos organismos e, ou, situações em que a proporção de heterozigotos não possa ser definida conforme as suposições de equilíbrio, a medida não deve ser chamada de heterozigosidade, devendo ser compreendida apenas como uma medida de variabilidade genética e chamada de índice de heterogeneidade ou diversidade genética (Torggler et al., 1995).

Para uma subpopulação, tem-se:

$$H_i = 1 - J_i \qquad H_{ii} = 1 - J_{ii}$$

A partir destes valores podem ser estabelecidas as seguintes estatísticas:

### **Heterozigosidade dentro de subpopulações**

A heterozigose genética total dentro das subpopulações, para um loco, é dada por:

$$H_{IS} = \sum_{i=1}^g H_i = g - J_{IS}$$

e o valor médio é denotado por:

$$\bar{H}_{IS} = \frac{H_{IS}}{g} = 1 - \bar{J}_{IS}$$

### **Heterozigosidade entre subpopulações**

A heterozigosidade genética total entre todos os pares de subpopulações, para um loco, é dada por:

$$H_{ST} = \sum_{i=1}^g \sum_{i' < i}^g H_{ii'} = \frac{g(g-1)}{2} - J_{ST}$$

e o valor médio é denotado por:

$$\bar{H}_{ST} = \frac{2H_{ST}}{g(g-1)} = 1 - \bar{J}_{ST}$$

### **Heterozigosidade total das subpopulações**

A heterozigosidade media total é dada por:

$$H_{IT} = \sum_{i=1}^g H_i + \sum_{i=1}^g \sum_{i' \neq i}^g H_{ii'} = H_{IS} + 2H_{ST}$$

Define-se a heterozigose média das subpopulações pela expressão:

$$\bar{H}_{IT} = \frac{1}{g}(1 - \bar{J}_{IS}) + \frac{g-1}{g}(1 - \bar{J}_{ST}) = 1 - \bar{J}_{IT}$$

### 2.4.3.3 Medida de Diversidade

A diversidade genética ( $D_{ii'}$ ) entre a  $i$ -ésima e  $i'$ -ésima subpopulações é estimada por:

$$D_{ii'} = H_{ii'} - \frac{H_i + H_{i'}}{2}$$

Como se trata de diversidade, tem-se:

$$D_{ii} = 0 \quad \text{e} \quad D_{ii'} = D_{i'i}$$

### Diversidade dentro de subpopulações

Como  $D_{ii}=0$ , tem-se

$$D_{IS} = \sum_i D_{ii'} = 0$$

### Diversidade ente subpopulações

$$D_{ST} = \frac{g-1}{2} J_{IS} - J_{ST} \quad \text{e} \quad \bar{D}_{ST} = \frac{2}{g(g-1)} D_{ST} = \bar{J}_{IS} - \bar{J}_{ST}$$

### Diversidade total das subpopulações

A diversidade total computa todas as populações, é denotada por  $D_{IT}$  e calculada por meio de:

$$\bar{D}_{IT} = \frac{2D_{ST}}{g^2} = \frac{g-1}{g} \bar{D}_{ST}$$

#### 2.4.3.4 Coeficiente de Diversidade Relativa Entre Grupos

A magnitude relativa da diferenciação entre subpopulações pode ser medida por:

$$G = \frac{100\overline{D}_{ST}}{\overline{H}_{ST}} \quad \text{ou} \quad G_{ST} = \frac{100\overline{D}_{IT}}{\overline{H}_{IT}}$$

A medida  $G_{ST}$  é denominada de coeficiente de diversidade relativa entre grupos, varia entre 0 e 1 e expressa a proporção da diversidade total explicada por diferenças entre os grupos. Matematicamente, esta medida equivale à estatística  $F_{ST}$  de Wright. Segundo Hartl & Clark (1997) a equação de Nei pode ser aplicada a diferentes níveis da classificação hierárquica, como populações dentro de uma região geográfica, subpopulações dentro de populações, indivíduos dentro de subpopulações, etc., utilizando-se em cada caso as frequências alélicas correspondentes da categoria em questão.

#### 2.4.4 Índices de Fixação – Estatística F de Wright

Nos estudos genéticos de populações estruturadas, Wright (1951, 1978) mostrou que a variação na frequência gênica entre subpopulações pode ser analisada pelos índices de fixação ou pela estatística F. Ele derivou a fórmula:

$$F=1-F_{IT} = (1- F_{IS}) (1- F_{ST}) \quad (1)$$

Onde  $F_{IT}$  e  $F_{IS}$  são correlações entre duas unidades gaméticas para a formação de indivíduos relativos à população total e relativos à subpopulações, respectivamente. O  $F_{ST}$  é a correlação da união aleatória de dois gametas de cada subpopulação. Segundo Robinson (1998), Wright descreveu três coeficientes de fixação aplicáveis a uma população com um nível hierárquico de subdivisão:

$F_{IT}$  (ou  $F$ ) – que mede o desvio das frequências genotípicas da população em relação ao equilíbrio de Hardy-Weinberg. Esses desvios resultam de cruzamentos não ao acaso dentro da população (incluindo a endogamia em todos os níveis).  $F$  também é uma

medida de correlação entre duas unidades gaméticas que formam um zigoto na população (ou do conjunto de populações amostradas);

$F_{ST}$  (ou  $\theta$ ) - é o coeficiente de ancestralia, representando a probabilidade de que dois indivíduos, pertencentes à subpopulações diferentes, possuam uma alelo idêntico por descendência. Assim,  $\theta$  é uma medida da correlação de gametas entre as subpopulações. Neigel (2002) comenta que as estimativas de  $F_{ST}$  são aplicáveis aos modelos de ilhas, assumindo que uma população possa ter infinitas subpopulações, no entanto, não há seleção, mutação, migração ou fluxo gênico entre as mesmas. E tais subpopulações não são afetadas pela distância geográfica existente entre elas;

$F_{IS}$  (ou  $f$ ) - mede a fixação ao nível de indivíduos, ou seja, mede a probabilidade de que os dois alelos de um loco presentes no mesmo indivíduo sejam idênticos por descendência. Assim,  $f$  é também uma medida de correlação de gametas devido à endogamia dentro das subpopulações.

### **Relação entre os índices de fixação**

As estatísticas  $F$ , que expressam índices de fixação, foram redefinidas por Cockerham (1973) e relacionadas com o coeficiente de endogamia e de coancestralidade, como descrito a seguir:

$$F_{IS} = \frac{F_{IT} - F_{ST}}{1 - F_{ST}} = \frac{F - \theta}{1 - \theta} \qquad F_{ST} = \frac{F_{ST} - F_{IS}}{1 - F_{IS}} = \frac{\theta - f}{1 - f}$$

$$F_{IT} = \frac{F_{IT} - F_{IS}}{1 - F_{IS}} = \frac{F - f}{1 - f}$$

Onde:

$$F = F_{IT}$$

$$f = F_{IS}$$

$$\theta = F_{ST}$$

em que os indexadores I, S e T representam indivíduos, subpopulações e população total, respectivamente.

Wright (1951) introduziu a seguinte equação, como medida do grau de diferenciação genética entre populações:

$$F_{ST} = \frac{\sigma^2}{\bar{f}(1-\bar{f})} \quad (2)$$

Onde  $\bar{f}$  é a frequência média de um dos alelos na população, e  $\sigma^2$  é a variância na frequência de um dos alelos entre as populações.

Nei (1973) estendeu esta equação para o caso de múltiplos alelos e chamou de coeficiente de diferenciação gênica. Este coeficiente está definido como:

$$G_{ST} = (H_T - H_S) / H_T \quad (3)$$

Em que  $H_S$  é a média da diversidade genética dentro de populações e  $H_T$  é a média da diversidade genética na população total. A frase diversidade genética se refere à heterozigiosidade esperada sob o Equilíbrio de Hardy-Weinberg. A diferença  $D_{ST} = H_T - H_S$  é chamada de diversidade genética interpopulacional. Portanto,  $G_{ST} = D_{ST} / H_T$ . Se na população existem apenas dois alelos por loco, a equação (3) se reduz a equação (2). Quando existem múltiplos alelos,  $G_{ST}$  tem o mesmo significado que  $F_{ST}$ . Originalmente,  $G_{ST}$  foi interpretado como uma medida de diversidade genética para um grande número de loci, mas esta também pode ser computada para cada loco separadamente (Nei, 1977).

Nei (1977) mostra que as estatísticas  $F$  de Wright podem ser definidas como razões entre estatísticas  $H$ , ao invés de correlações entre unidades gaméticas. Os índices de fixação de Wright (1978) medem a relação entre a Heterozigiosidade observada e a esperada em uma população de cruzamentos aleatórios, conforme as expectativas do Equilíbrio de Hardy-Weinberg.

Estas definições, segundo Nei (1977), são independentes do número de alelos envolvidos e da atuação da seleção natural. Neste caso tem-se:

$F_{IS} = \frac{H_S - H}{H_S}$  → Diversidade genética (ou Heterozigosidade) dentro de subpopulações em relação à diversidade genética esperada entre subpopulações.

$F_{ST} = \frac{H_T - H_S}{H_T}$  → Diversidade genética (ou Heterozigosidade) entre subpopulações em relação à diversidade genética esperada na população total.

$F_{IT} = \frac{H_T - H}{H_T}$  → Diversidade genética (ou Heterozigosidade) dentro de subpopulações em relação à diversidade genética esperada na população total.

em que:

$H_T$ : correspondem à diversidade genética (ou heterozigosidade) na população total, dada por:

$$H_T = 1 - \sum_{k=1}^{a_j} \bar{P}_{.jk}^2$$

e 
$$\bar{P}_{.jk} = \sum_{k=1}^s W_i P_{ijk}$$

Sendo

$P_{ijk}$  a frequência do alelo k, do loco j, na subpopulação i,  $W_i$  o tamanho relativo da i-ésima população. Quando se compara populações com o mesmo número de indivíduos, ou se desconhece a proporcionalidade de indivíduos entre as populações, considera-se  $W_i=1/g$  para todo i.

$H_S$ : correspondem à diversidade genética (ou heterozigosidade) esperada entre subpopulações, dada por:

$$H_S = 1 - \sum_{i=1}^g W_i \left( \sum_{k=1}^{a_j} P_{ijk}^2 \right)$$

$H_o$ : correspondem à diversidade genética (ou heterozigosidade) observada dentro de subpopulações.

$$H_o = \sum_{j=1}^g W_i P_{ijkk'}^2 = \sum_{i=1}^g W_i H_{oi}$$

Em que

$P_{ijkk'}$  é a frequência heterozigoto  $A_k A_{k'}$  em relação ao loco  $j$  na população  $i$ .

A teoria da estatística  $F$  de Wright é baseada na pressuposição de que as populações são infinitamente grandes. Se não existe fluxo gênico entre as populações e a frequência gênica inicial é a mesma para todas as populações, o valor de  $F_{ST}$  na  $t$ -ésima geração é dada por:

$$F_{ST} = 1 - \left(1 - \frac{1}{2N}\right)^t \approx 1 - e^{-t/2N}$$

em que:  $n$  é o tamanho da subpopulação e  $t$  é o número de gerações.

Os índices de fixação são, portanto, índices de diferenciação genética porque eles permitem uma comparação objetiva de todos os efeitos de uma população estruturada em subpopulações (Hartl & Clark, 1997).

Uma importante consequência em populações estruturadas em subpopulações é a redução na proporção média de genótipos heterozigotos relativos ao esperado em uma população sob acasalamento ao acaso (Hartl & Clark, 1997). Este fato está relacionado com o estabelecimento de uma nova população por uns poucos fundadores originais que contêm somente uma pequena fração da variação genética total da população parental. O efeito fundador é o estabelecimento de uma nova população por um pequeno número de fundadores. Eventos fundadores ocorrem, e são caracterizados como eventos de amostragem populacional. Uma população pode ser descendente de um pequeno número de indivíduos ancestrais por duas possíveis razões: i) um pequeno número de indivíduos pode colonizar um local previamente desabitado pelas suas espécies e estes fundadores possuem uma amostra limitada da variação genética; ii) uma população que está estabelecida em uma área pode variar em tamanho; o efeito do fundador ocorre, então, quando a população passa por um “gargalo de garrafa”, no qual apenas uns poucos indivíduos sobrevivem, e, mais tarde, ela se expande novamente, quando o período favorável retorna. A chance de dois indivíduos retirados aleatoriamente da população,

serem ambos AA é  $(p^2)^2$ ; em geral, a chance de se retirar  $N$  homozigotos idênticos é  $(p^2)^N$ . A população fundadora seria homozigota também, porque ela é constituída de  $N$  homozigotos AA ou  $N$  homozigotos aa, e a chance total de homozigosidade é, portanto:

$$\text{Chance de homozigosidade} = [(p^2)^N + (q^2)^N]$$

Entretanto, eventos fundadores não são eficientes na produção de uma população geneticamente monomórfica. A chance de uma população fundadora ser homozigota depende do número de fundadores e das frequências gênicas. Se há menor variação e poucos fundadores, a chance de homozigosidade é maior (Ridley, 2006).

#### **2.4.5 Heterozigosidade – Weir (1996)**

Uma medida simples da variação em uma população é a quantidade observada de heterozigotos para um único loco e para todos os locos, considerada em média. Portanto, quando se estuda várias populações em relação às informações genéticas de vários locos pode-se formular varias questões a respeito da diversidade entre estas populações e, em caso afirmativo, haverá indicativo de que elas estão passando por algum tipo de diferenciação. Também se pode questionar se esta heterozigosidade varia de loco para loco e, em caso afirmativo, entende-se que, em certos locos, os alelos estão sujeitos a pressão de seleção resultando em variações alélicas ou genotípicas. Adicionalmente, a análise simultânea destes dois fatores, locos e populações, possibilitam estudos de interação de grande interesse no contexto genético.

Para fazer análise da heterozigose, adotam-se duas estratégias básicas. A primeira relaciona-se à obtenção dos valores médios da heterozigose para locos e populações, suas respectivas variâncias e testes de hipóteses ou intervalos de confiança. A segunda se fundamenta na análise de variância com obtenção de componentes de variância com significado biológico que quantificam a variação da heterozigose entre populações, entre locos e da interação locos x populações.

Para análise de variância dos valores que representam a heterozigose da população são obtidas as seguintes somas de quadrados:

- Soma de quadrados de populações

$$SQP = \frac{1}{LN} \sum_{i=1}^g X_{i..}^2 - C$$

Sendo,

$$C = \frac{X_{...}^2}{NLg}$$

- Soma de quadrados de indivíduos/populações

$$SQI/P = \frac{1}{L} \sum_{i=1}^g \sum_{j=1}^N X_{ij.}^2 - \frac{1}{NL} \sum_{i=1}^g X_{i..}^2$$

- Soma de quadrados de locos

$$SQL = \frac{1}{Ng} \sum_{i=1}^L X_{..i}^2 - C$$

- Soma de quadrados de locos x população

$$SQLP = \frac{1}{N} \sum_{i=1}^g \sum_{l=1}^L X_{i.l}^2 - \frac{1}{LN} \sum_{i=1}^g X_{i..}^2 - \frac{1}{Ng} \sum_{l=1}^L X_{..l}^2 + C$$

- Soma de quadrados de locos x indivíduos/populações

$$SQL(I/P) = \sum_{i=1}^g \sum_{j=1}^N \sum_{l=1}^L X_{ijl}^2 - \frac{1}{L} \sum_{i=1}^g \sum_{j=1}^N X_{ij.}^2 - \frac{1}{N} \sum_{i=1}^g \sum_{l=1}^L X_{i.l}^2 - \frac{1}{LN} \sum_{i=1}^g X_{i..}^2$$

O quadro de análise de variância é apresentado a seguir:

<b>FV</b>	<b>GL</b>	<b>QM</b>	<b>E(QM)</b>
Populações (Pop.)	g-1	QMP	$\sigma_{li/p}^2 + L\sigma_{i/p}^2 + N\sigma_{pl}^2 + LN\sigma_p^2$
Ind./Pop.	g(N-1)	QMI	$\sigma_{li/p}^2 + L\sigma_{i/p}^2$
Locos	L-1	QML	$\sigma_{li/p}^2 + N\sigma_{pl}^2 + \psi$
Locos x Pop.	(g-1)(L-1)	QMLP	$\sigma_{li/p}^2 + N\sigma_{pl}^2$
Locos x Ind./Pop.	g(N-1)(L-1)	QMR	$\sigma_{li/p}^2$

$$\psi = \frac{1}{L-1} \sum_{l=1}^L (H_l - \bar{H})^2$$

#### 2.4.6 Análise Molecular de Variância - AMOVA

Excoffier et al., (1992) formularam uma nova metodologia de análise da distribuição da variabilidade genética entre e dentro das populações pela análise de variância molecular (AMOVA). Neste método, a matriz de distância entre todos os pares de haplótipos (genótipos) é utilizada em um esquema de análise de variância hierarquizada produzindo estimativas de componentes de variância análogas às estatísticas F de Wright.

Alguns autores já haviam tentado traduzir as informações do DNA, haplótipos, utilizando endonucleases de restrição em estimativas da magnitude de uma subdivisão populacional intraespecífica. Lynch & Milligan (1994) estudando filogenia de alguns haplótipos, estimaram a diversidade por variância nucleotídica por diferentes processos amostrais. Takahata & Palumbi (1985) computaram a fração da diversidade nucleotídica devido às diferenças genéticas interpopulacionais e fizeram analogia com as estimativas de Nei (1973) e com o coeficiente de diferenciação genética ( $G_{ST}$ ). Ambos os métodos envolveram transformações não lineares da informação original em estimativas da diversidade genética (Excoffier et al., 1992).

Excoffier et al., (1992), na tentativa de obter uma metodologia mais generalizada, que não dependesse de pressuposições específicas, objetivaram traduzir as informações moleculares, como a identificação de haplótipos, pelo uso de enzimas de restrição, para a estimação da magnitude da subdivisão populacional intraespecífica. Estes autores

propuseram uma extensão dos trabalhos de Cockerham (1973) em que uma correlação alélica entre demes é usada para comparar uma análise de diversidade haplotípica. Usando o fato de que uma convencional soma de quadrados de uma análise de variância pode ser escrita como uma soma de quadrados das diferenças entre pares de observações, Excoffier et al., (1992) construíram uma análise de variância molecular hierárquica diretamente a partir de uma matriz dos quadrados das distâncias entre todos os pares de haplótipos (genótipos). Esta análise de variância hierarquizada produz estimativas de componentes de variância análogas às estatísticas F de Wright.

Para os autores, estava bem claro que o novo método tinha uma vantagem adicional, várias pressuposições diferentes poderiam ser impostas no processo de diferenciação haplotípica e cada uma seria traduzida em uma matriz de diferença de distâncias, e que não mudaria a estrutura da análise. O termo haplótipo, empregado neste contexto, refere-se à combinação de sítios de amplificação presente em alguma área definida do genoma. No caso de marcadores moleculares dominantes, o haplótipo de um indivíduo é geralmente representado por um vetor com valor 1, se a banda homóloga estiver presente e 0, caso contrário.

Portanto, a análise de variância molecular (AMOVA) é uma metodologia capaz de estudar a diversidade entre populações a partir de dados moleculares e também de testar hipóteses a respeito de tal diferenciação. Uma variedade de dados moleculares tais como informações de marcadores dominantes ou co-dominantes e dados de seqüência, pode ser analisada usando este método (Excoffier et al., 1992).

AMOVA estuda qualquer tipo de dados moleculares estruturados em um vetor  $p_j$  de variável Booliana, ou seja, uma matriz  $1 \times n$  constituída por uns ou zeros, sendo que 1 que indica a presença de um marcador e 0 sua ausência. Um marcador poderia ser uma base de nucleotídeo, uma seqüência básica, um fragmento de restrição ou um evento de mutacional (Excoffier et al., 1992).

Para dois haplótipos  $j$  e  $k$ , os vetores contendo as informações de 4 marcadores moleculares, poderiam ser representados por:

$$p_j = \begin{bmatrix} m1 & m2 & m3 & m4 \\ 1 & 1 & 0 & 1 \end{bmatrix}$$

$$p_k = \begin{bmatrix} 0 & 1 & 1 & 0 \end{bmatrix}$$

Para realizar a AMOVA é necessário, inicialmente, calcular a distância euclidiana entre pares de vetores (ou pares de indivíduos) subtraindo o vetor Booliano de um haplótipo de outro, de acordo com a fórmula  $(p_j - p_k)$ . A distância euclidiana é um escalar que expressa a menor distância entre esses dois indivíduos. Normalmente, calculam-se os quadrados das distâncias euclidianas por meio da expressão:

$$d_{jk}^2 = (p_j - p_k)' W (p_j - p_k)$$

em que W é uma matriz de pesos atribuídos a cada informação molecular. Geralmente estes pesos são estabelecidos a partir de informações sobre mudanças moleculares localizadas em diferentes posições numa seqüência ou numa árvore filogenética (Excoffier, et al., 1992). Na falta de informações sobre estes pesos, utiliza-se uma matriz de identidade.

Para análise da diversidade entre e dentro de populações pela metodologia proposta por Excoffier, et al., (1992) deve-se, inicialmente, calcular a matriz de dissimilaridade, cujos elementos são os quadrados das distâncias euclidianas entre pares de acessos, denotada por  $D^2$ . Esta matriz é, então, particionada agrupando as informações de pares de haplótipos dentro de cada subpopulação. Assim, a matriz origina  $D^2$  pode ser apresentada das seguintes formas:

$$D^2 = [\tilde{T}] \quad \text{ou} \quad D^2 = \begin{bmatrix} [\tilde{D}_1] & & & \\ & [\tilde{D}_2] & & \\ & & \dots & \\ & & & [\tilde{D}_p] \end{bmatrix}$$

### **Obtenção de somas de quadrados**

*Soma de quadrados das distâncias – total*

$$SQT = \frac{1}{2N} T$$

*Soma de quadrados das distâncias – dentro de subpopulações*

$$SQD = \sum_{i=1}^g \frac{D_i}{2N_i}$$

*Soma de quadrado das distâncias – entre subpopulações*

$$SQE = SQT - SQD$$

### **Esquema da análise de variância**

De posse dos valores de somas de quadrados é estabelecido o quadro de análise de variância molecular (AMOVA) conforme descrito a seguir. Os valores das esperanças dos quadrados médios são obtidos considerando um modelo de dois fatores de classificação em que as medidas de dissimilaridade são estabelecidas a partir das informações das variações entre e dentro de subpopulações, ou seja:

$$Y_{ij} = \mu + E_i + D_{ij}$$

em que:

$Y_{ij}$ : medida da dissimilaridade entre pares de indivíduos (j) numa subpopulação i.

$\mu$ : constante

$E_i$ : efeito da subpopulação i

$D_{ij}$ : efeito da dissimilaridade entre o par de indivíduos j na subpopulação i.

Esquema da análise de variância molecular (AMOVA), com dados agrupados em dois níveis hierárquicos

<b>Fonte de Variação</b>	<b>GL</b>	<b>SQ</b>	<b>QM</b>	<b>E(QM)</b>
Entre Populações	g-1	SQE	QME	$\sigma_d^2 + n\sigma_e^2$
Dentro Populações	N-g	SQD	QMD	$\sigma_d^2$
Total	N-1	SQT	-	$\sigma_T^2$

em que  $n$  representa o número médio de haplótipos amostrados por população. Com amostras de tamanhos desiguais,  $n$  é obtido por:

$$n = \frac{N - \sum_i \frac{N_i^2}{N}}{g - 1}$$

onde  $N_i$  é o número de indivíduos ou haplótipos da  $i$ -ésima subpopulação.

### **Estimação de componentes de variância**

Os estimadores dos componentes de variância que expressam as diferenciações entre e dentro de subpopulações são obtidos por meio de:

$$\hat{\sigma}_d^2 = QMD$$

e

$$\hat{\sigma}_e^2 = \frac{QME - QMD}{n}$$

### **Associação entre componente de variância e estatísticas $\Phi$**

As estatísticas obtidas pela AMOVA, denominadas de “estatísticas  $\Phi$ ”, refletem a correlação da diversidade de haplótipos em diferentes níveis da subdivisão hierárquica. Para o caso em que se consideram dois níveis hierárquicos, (populações e indivíduos ou haplótipos avaliados).

Assim, o coeficiente de correlação entre haplótipos, amostrados aleatoriamente entre populações, é dado por:

$$r = \Phi_{ST} = \frac{\text{Cov}(Y_{ij}, Y_{ij'})}{\sqrt{V(Y_{ij})V(Y_{ij'})}}$$

logo,

$$\Phi_{ST} = \frac{\hat{\sigma}_e^2}{\hat{\sigma}_T^2}, \text{ que é uma medida da diversidade relativa entre as populações}$$

avaliadas.

## Testes estatísticos dos componentes de variância e das estatísticas $\Phi$

A significância dos componentes  $\sigma_e^2$  e  $\Phi_{ST}$  pode ser testada pela permutação de haplótipos entre populações, através de procedimentos fundamentados em permutação dos dados. Dois tipos de permutações são realizados. No primeiro, para estabelecimento da nulidade dos componentes  $\sigma_d^2$  e  $\sigma_e^2$ , todos os dados de cada informação molecular são permutados, independentemente das populações a que pertencem. No segundo teste, admite-se ser verdadeira a divisão da população em subpopulações, e testa-se apenas a existência de variabilidade dentro das subpopulações. Neste caso, a permutação dos dados deve ocorrer apenas dentro de subpopulações.

### Teste de Hipótese

Uma estatística  $\Phi$  pode ser tratada como instrumento de avaliação de uma hipótese sobre o nível de diferenciação de uma população; por exemplo, estatística  $\Phi$  associada a  $\Phi_{ST}$  pode ser tratada como ferramenta para testar a hipótese sobre diferenciação da população em seus componentes de subpopulações. As hipóteses que usam a distribuição nula dos componentes de variação podem ser testadas; se a variação das subpopulações não difere significativamente da distribuição nula da variação da população, a hipótese que essas subpopulações são diferenciadas da população original seria rejeitada.

A permutação é feita diretamente na matriz de distância ( $D^2$ ), alterando-se as posições dos elementos de linhas e das colunas correspondentes. A permutação pode se dar em toda a matriz, sem distinção de subespaço de regiões ou de subpopulações, ou admitir que a subdivisão em regiões e, ou, em subpopulações é verdadeira. Nestes casos, a permutação deve ocorrer dentro dos subespaços apropriados da matriz.

## 2.5 Diversidade Genética

A existência de variação genética é fator fundamental para que ocorra a evolução. A seleção natural atua entre as variantes dentro das populações em relação à adaptação ao ambiente, proporcionando variação entre populações e, por fim, variação entre espécies (Torggler et al., 1995). Desta forma, quanto maior a variabilidade genética existente na população, maiores são suas chances de perpetuação. A manutenção da variabilidade genética em populações é a base da conservação de espécies (Yeoh et al., 1996) e, portanto, sua descrição e distribuição são fundamentais para o estabelecimento de práticas conservacionistas realmente efetivas.

O padrão espacial ou a estrutura da diversidade genética dentro de populações é um componente importante dos processos genético-ecológicos e evolucionários de populações naturais de plantas (Epperson, 1990). A destruição de habitats e a conseqüente fragmentação de populações naturais podem acarretar uma limitação evolutiva para as espécies que os compõem, devido à perda de variabilidade genética e a conseqüente redução na capacidade de adaptação às mudanças ambientais (Barret & Kohn, 1991). A fragmentação de habitats pode levar à perda de variação genética por duas vias: i) a redução do tamanho populacional que cria gargalos genéticos, dado que os indivíduos remanescentes contêm apenas uma pequena amostra do conjunto gênico original e; ii) como conseqüência, a pequena população remanescente, caso permaneça isolada por muitas gerações, terá contínua perda de alelos devido à deriva genética. O gargalo contribui para a perda de alelos, especialmente dos raros e isto é mais efetivo do que a perda de heterozigidade (Barret & Kohn, 1991; Charlesworth & Charlesworth, 1987).

A variabilidade é introduzida continuamente nas populações por mutação ou migração de indivíduos de outras populações e é perdida por deriva genética, por endocruzamentos e, no caso de genes neutros, pela maior parte dos tipos de seleção natural (Cole, 2003). O padrão espacial ou a estrutura da diversidade genética dentro de populações é um componente importante dos processos genético-ecológicos e evolucionários de populações naturais de plantas. O conhecimento da estrutura genética

espacial pode melhorar a eficiência da amostragem para maximizar a diversidade gênica ou minimizar os cruzamentos endogâmicos (Epperson, 1990).

Nei (1977, 1987) mostrou que a diversidade genética da população total poderia ser partida em seus componentes de diversidade genética intra e inter populacional, quando a mesma fosse definida como a heterozigosidade esperada sob Equilíbrio de Hardy-Weinberg (EHW). Hamrick & Murawski (1991) afirmam que as espécies que mantêm populações com alta densidade demográfica apresentam maiores níveis de diversidade genética do que as de baixa densidade, aumentando a probabilidade de ocorrerem mutações e a incorporação de novos alelos nas populações. Estudos efetuados por Cornuet & Luikart (1996) e Piry et al., (1999) baseiam-se no princípio de que populações que experimentaram uma recente redução de seu tamanho efetivo populacional exibem redução do número de alelos e da diversidade gênica (heterozigosidade esperada pelas proporções de EHW) nos locos polimórficos.

A estrutura genética de populações vegetais é refletida de interações de vários processos evolutivos bem como fragmentação de habitats, isolamento populacional, mutação, deriva genética, sistema de acasalamento, fluxo gênico, seleção, e ainda, incluindo um longo tempo de história evolutiva (Shah et al., 2008)

Farina et al., (2004) avaliaram os efeitos da fragmentação sobre a estrutura e a diversidade genética de populações de *Chorisia speciosa*. Populações naturais de diferentes tamanhos, situadas em quatro fragmentos florestais na região de Bauru, SP, foram estudadas por oito locos isoenzimáticos. Detectou-se que a diversidade genética presente nas populações de *C. speciosa* é superior à média estimada para as espécies arbóreas. A heterozigosidade observada ( $\hat{H}_o$ ), a diversidade gênica ( $\hat{H}_e$ ), a porcentagem de locos polimórficos ( $PIC\ 95\%$ ), o número médio de alelos por locos ( $\hat{A}$ ) e o número efetivo de alelos por loco ( $\hat{A}_e$ ) foram de 0,252, 0,344, 2,38 e 1,97, respectivamente. As conseqüências genéticas da fragmentação observadas nas populações foram: a perda de alelos raros, a fixação de alelos e as oscilações aleatórias das frequências alélicas. A diferenciação genética entre populações foi também alta ( $F_{IS} = 0,273$ ), refletindo que as populações estão isoladas nos fragmentos. O índice médio de fixação dentro das populações ( $F_{ST} = 0,140$ ) foi menor que o índice médio de fixação estimado para o conjunto das populações ( $F = 0,375$ ), indicando desvios das proporções do Equilíbrio de Hardy-Weinberg, combinados com efeitos da deriva genética.

Gusson et al., (2005), usando marcadores isoenzimáticos, caracterizaram os níveis de diversidade e estrutura genética espacial de uma espécie arbórea de alta densidade populacional, *Eschweilera ovata*, em duas populações naturais, Camarugipe e Itaparica, trabalhando com 60 indivíduos adultos em cada população, na Floresta Tropical Atlântica do Estado da Bahia-Brasil. A diversidade genética intrapopulacional foi caracterizada pelas frequências alélicas, número médio de alelos por loco ( $\hat{A}$ ), porcentagem de locos polimórficos (PIC), heterozigosidade observada ( $\hat{H}_o$ ), diversidade gênica esperada segundo o modelo de equilíbrio de Hardy-Weinberg ( $\hat{H}_e$ ) e índices de fixação de Wright (F). Os resultados com os parâmetros avaliados indicaram que a maior parte da diversidade genética encontra-se distribuída dentro das populações.

Espécies que mantêm populações com alta densidade demográfica apresentam maiores níveis de diversidade genética do que as de baixa densidade, aumentando a probabilidade de ocorrerem mutações e a incorporação de novos alelos nas populações. Os gradientes de variação das frequências alélicas ao longo de uma distância geográfica ocorrem frequentemente em resposta a gradientes ambientais, podendo a ação aleatória da deriva provocar o isolamento por distância, sem que haja envolvimento de qualquer fator seletivo no processo (Gusson et al., 2005).

Trabalhos com *Eremanthus erythropappus* MacLeish, uma árvore de grande importância ecológica e econômica, de ocorrência no sul do Estado de Minas Gerais foram realizados por Abjoud et al., (2006), com o objetivo de caracterizar a variabilidade genética intra e interpopulacional, de duas populações naturais localizadas no sul de Minas Gerais, por meio do marcador molecular RAPD. Na população I (Aiuruoca) 50 indivíduos foram coletados, porém 25 indivíduos foram amostrados num gradiente de 1289 a 1385 m, sendo considerada subpopulação I (Aiuruoca I) e, os demais indivíduos de 1415 a 1578 m, subpopulação II (Aiuruoca II), ambas localizadas a uma distância de 1260 m. Esta amostragem teve como objetivo avaliar a influência da altitude sobre a variabilidade genética. Na população II (São Tomé das Letras), distanciada a 37 km de Aiuruoca, 25 indivíduos foram amostrados. Para todas as áreas de estudo o principal fator considerado durante a amostragem foi à livre polinização entre as árvores e uma distância mínima de 100m entre as mesmas, visando garantir uma melhor representatividade da variabilidade genética (Abjoud et al., 2006).

Abjoud et al., (2006) encontraram, para populações de *E. erythropappus*, valores de variação genética segundo o índice de Shannon e apresentaram que as estimativas de

fluxo gênico não são suficientes para contrapor os efeitos da deriva genética, que muitas vezes é gerada pelo isolamento das populações naturais devido à fragmentação. Segundo Hartl e Clark (1997), quando o fluxo gênico entre populações excede quatro migrantes por geração, ocorre a homogeneização dos alelos entre estas, que funcionam como populações panmíticas. Para as subpopulações I e II de Aiuruoca a estimativa do fluxo foi de 3,38; já entre as populações de Aiuruoca e São Tomé das Letras esta estimativa foi de 2,79. Estes resultados indicam uma relação entre o fluxo gênico e a distância geográfica das populações. As subpopulações I e II de Aiuruoca por estarem distanciadas uma da outra por apenas 1260 m, apresentaram uma maior taxa de fluxo gênico, o que era de se esperar. Porém, avaliando-se esta taxa entre as subpopulações de Aiuruoca e a população de São Tomé das Letras, não foi observada uma diferença considerável entre os valores obtidos, embora, estas populações estejam distanciadas aproximadamente em 37 km.

Oliveira et al., (2006) estudaram a diversidade genética de duas populações naturais e de um reflorestamento de *Caesalpinia echinata* (Lam.), Pau-brasil, na Estação Ecológica do Tapacurá, PE, mediante a análise eletroforética de isoenzimas. Segundo estes autores, a importância das informações sobre a estrutura genética populacional desta espécie, remanescente protegida, está consoante com a estratégia de conservação da mesma. Isto porque, atualmente, existem poucas reservas, com pequeno número de indivíduos e em áreas limitadas. As populações foram identificadas como: Mata do Camucim (CAM), Mata do Toró (TOR) e reflorestamento (REF). Avaliaram-se os seguintes parâmetros genéticos populacionais: frequência alélica, heterozigidade, índice de fixação de Wright e divergência genética dentro e entre as populações. As frequências alélicas variaram de 0,05 a 1,00 entre as populações, ou seja, variou desde a completa fixação, para o alelo A em um sistema enzimático nas populações REF e TOR, até frequências muito baixas (0,05) para o alelo B em outro sistema enzimático na população TOR. O número médio de alelos por loco variou de 2,0 a 2,6 entre as três populações.

Os valores de heterozigidade observada ( $H_o$ ) e esperada ( $H_e$ ) para as populações de *C. echinata* variaram de 0,073 a 0,475 e 0,286 a 0,468, respectivamente. A população CAM apresentou maior heterozigidade ( $H_o=0,475$ ) e a população REF ( $H_o=0,073$ ) apresentou a maior homozigidade. A população da Mata do Camucim apresentou diversidade genética (0,475) superior às médias encontradas em outras

espécies arbóreas (0,149) (Hamrick & Godt, 1990) o que indica que esta população apresenta alto nível de diversidade genética. Valores próximos a este têm sido encontrados em populações naturais de outras espécies. Santos (1994) encontrou o valor de 0,451 para heterozigosidade em população natural de *Bauhinia forficata* Link. Bawa (1992) encontraram heterozigosidade média de 0,282 por população de *Stemonoporus oblongifolius*.

Oliveira et al., (2006) comentam que os altos níveis de diversidade genética observados em *C. echinata*, para as populações de CAM e TOR, são explicados pela ampla distribuição geográfica da espécie. E, a baixa heterozigosidade na população de reflorestamento pode ser explicada pelo fato de que na ocasião da coleta de sementes para formação de mudas que deram origem a esta população, a amostragem tenha sido feita em poucas árvores, ocasionando baixa representatividade da variabilidade total da população original. Estes autores comentam, ainda, que é possível que muitas das sementes tenham sido originadas de autofecundação ou cruzamento entre parentes. Estes dois argumentos levariam, portanto, a redução da variabilidade genética na população de reflorestamento, ou por efeito de amostragem ou pelo sistema de acasalamento da população inicial.

### 3. MATERIAL E MÉTODOS

#### 3.1 Material

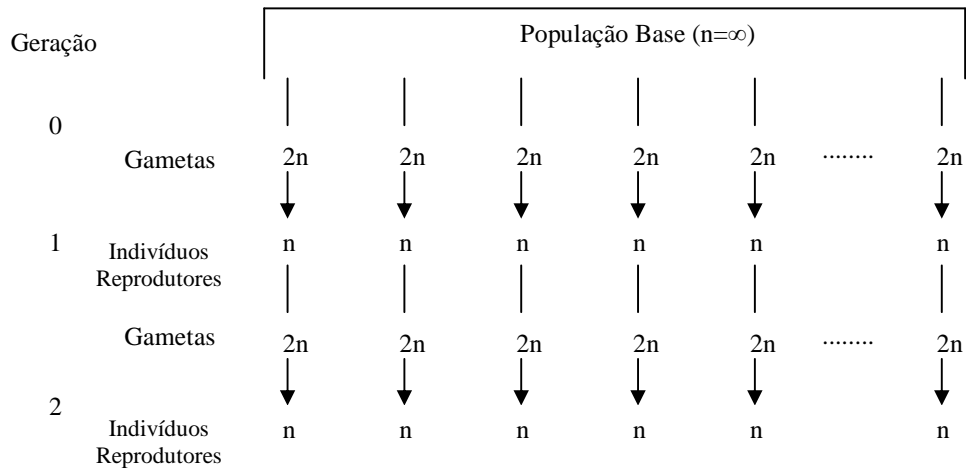
Para quantificar os efeitos da deriva genética proporcionada pela subdivisão de uma população, foram simuladas 4000 subpopulações de uma espécie diplóide, originadas a partir de uma população base simulada de tamanho grande, gerada por acasalamentos ao acaso, com frequência gênica de 0,5.

A partir da população base, simularam-se 10 gerações de acasalamentos ao acaso, com 100 subpopulações por geração. Essa simulação foi feita com quatro tamanhos populacionais diferentes os quais sejam: 20, 50, 100 a 200 indivíduos por subpopulação. Em todas as subpopulações foram analisados 10 locos. Todas as simulações foram feitas por meio do programa Genes-vs 2008.6.0.

- a. O acasalamento é restrito a membros da mesma subpopulação.
- b. As gerações são distintas e não se sobrepõem.
- c. O número de indivíduos que se acasalam em cada subpopulação é o mesmo em todas as subpopulações e gerações.
- d. Dentro de cada subpopulação, os acasalamentos são ao acaso, incluindo autofecundação em quantidade aleatória.
- e. Não há seleção em nenhum estágio.
- f. Não há mutação.

Para entender melhor as conseqüências do processo dispersivo e divisão da população, considere a população idealizada (Figura 1) como uma população original (base) de tamanho grande (infinito) na qual o acasalamento é casualizado e que é subdividida em um número grande de subpopulações ou linhagens, cada uma de

tamanho  $n$ , que permanece constante de geração para geração. A subdivisão pode surgir de causas geográficas ou ecológicas em populações naturais ou de um programa de acasalamento controlado em cativeiro ou laboratório. Todas as subpopulações em conjunto constituem a população base e cada linhagem é uma população pequena na qual as frequências alélicas estão sujeitas a dispersão.



**Figura 1-** População idealizada.

No diagrama considera-se que a população base é dividida em várias amostras (ou linhas) ao longo do tempo. Em cada geração tem-se que  $2n$  gametas de uma amostra se unem para gerar os  $n$  indivíduos da próxima subpopulação. Admite-se que a frequência alélica na população original seja  $p_0$  e  $q_0$  e que a frequência do alelo  $A$  na amostra (ou linha)  $j$  na subpopulação  $t$  seja dada por  $\hat{p}_{jt}$  ou simplesmente  $\hat{p}_t$ .

As seguintes análises poderão ser feitas:

#### *Análise da subpopulação 1*

Com já visto anteriormente tem-se:

$$E(\hat{p}_{j1}) = E(\hat{p}_1) = p_0 \quad V(\hat{p}_{j1}) = V(\hat{p}_1) = \frac{p_0 q_0}{2n}$$

#### *Análise da subpopulação 2*

A esperança matemática da frequência gênica da subpopulação 2, em relação à subpopulação 1 é dada por:

$$E_2(\hat{p}_{j_2}) = E_2(\hat{p}_2) = \hat{p}_1$$

Entretanto é interessante que a referência seja a população original e não a antecessora. Assim, deve-se quantificar a esperança matemática da frequência gênica da subpopulação 2, em relação à população original, denotada por:

$$E_1 E_2(\hat{p}_{j_2}) = E_1 E_2(\hat{p}_2) = E_1(\hat{p}_1) = p_0$$

A variância da frequência gênica na subpopulação 2 é dada por:

$$V(\hat{p}_{j_2}) = V(\hat{p}_2) = E_2[\hat{p}_2 - E_2(\hat{p}_2)]^2 = E_2(\hat{p}_2 - p_0)^2$$

Em relação à população original é esperado que:

$$\begin{aligned} V(\hat{p}_2) &= E_1 E_2(\hat{p}_2 - p_0)^2 = E_1 E_2[(\hat{p}_2 - \hat{p}_1) + (\hat{p}_1 - p_0)]^2 \\ &= E_1 E_2[(\hat{p}_2 - \hat{p}_1)]^2 + 2E_1 E_2[(\hat{p}_2 - \hat{p}_1)(\hat{p}_1 - p_0)] + E_1 E_2[(\hat{p}_1 - p_0)]^2 \end{aligned}$$

Para obter  $V(\hat{p}_2)$  deve-se investigar cada termo da expressão anterior. Assim, tem-se:

- $E_1 E_2[(\hat{p}_2 - \hat{p}_1)]^2 = \frac{1}{2n} [p_0 - p_0^2 - V(\hat{p}_1)] = \frac{p_0(1-p_0)}{2n} \left(1 - \frac{1}{2n}\right)$
- $2E_1 E_2[(\hat{p}_2 - \hat{p}_1)(\hat{p}_1 - p_0)] = 0$
- $E_1 E_2[(\hat{p}_1 - p_0)]^2 = E_1[(\hat{p}_1 - p_0)]^2 = V(\hat{p}_1) = \frac{p_0(1-p_0)}{2n}$

Voltando à expressão original, pode-se concluir que:

$$V(\hat{p}_2) = \frac{p_0(1-p_0)}{2n} \left[ \left(1 - \frac{1}{2n}\right) + 1 \right]$$

#### *Análise da subpopulação t*

A esperança matemática da frequência gênica da subpopulação t, em relação à população original é dada por:

$$E_1 \dots E_t(\hat{p}_{j_t}) = p_0$$

Pode-se deduzir que:

$$V(\hat{p}_t) = \frac{p_0(1-p_0)}{2n} \left[ \left(1 - \frac{1}{2n}\right)^{t-1} + \left(1 - \frac{1}{2n}\right)^{t-2} + \dots + \left(1 - \frac{1}{2n}\right)^1 + 1 \right]$$

logo

$$V(\hat{p}_t) = \frac{p_0(1-p_0)}{2n} \left[ \frac{1 - (1/2n)^t}{1/2n} \right] = p_0(1-p_0) \left[ 1 - \left(1 - \frac{1}{2n}\right)^t \right]$$

Sabendo que

$$\lim_{t \rightarrow \infty} \left(1 - \frac{1}{2n}\right)^t = 0$$

tem-se:

$$V(\hat{p}_{t \rightarrow \infty}) = p_0(1-p_0)$$

Observe que, a variância da frequência gênica aumenta  $2n$  vezes em relação à existente na subpopulação inicial. Também se constata que o valor  $p_0(1-p_0)$  corresponde a exatamente à variância da frequência gênica em um conjunto de amostras em que o alelo  $A$  foi fixado em  $p_0$  linhas (ou amostras) e o alelo  $a$  foi fixado em  $q_0$  linhas.

Nas populações simuladas o número de indivíduos que se acasalam para dar origem à próxima geração é constante e igual a  $n$  em todas as amostras e em todas as subpopulações. Portanto, considerou-se que em cada subpopulação as frequências gênicas estavam sujeitas ao processo de deriva genética. As frequências gênicas nas amostras terão um valor médio igual ao encontrado na população-base  $q_0$  e estarão distribuídas em torno dessa média, com a variância  $p_0q_0/2N$ . Portanto, a magnitude da mudança na frequência gênica resultante do processo dispersivo, devido à re-amostragem em uma geração, pode ser expressa, em termos de sua variância:  $\sigma^2 = p_0q_0/2N$ . Esta variância expressa à mudança esperada, em qualquer subpopulação. Seu efeito é uma dispersão das frequências gênicas entre as subpopulações. As subpopulações tornam-se diferentes em frequência gênica, embora a média na população como um todo permaneça constante.

### 3.2 Informações do Conjunto de Dados

A caracterização dos indivíduos ocorreu de forma que a descrição genotípica de cada um foi representada por códigos numéricos informativos dos alelos que eles possuem. Na existência de dois alelos  $A_1$  e  $A_2$  em um loco qualquer, os genótipos homocigotos foram descritos por 11, e 22 e os genótipos heterocigotos descritos por 12. Uma amostra do arquivo de dados gerado está apresentada na Tabela 1. Para cada tamanho populacional e cada uma das dez gerações, foi gerado um arquivo de dados.

Todo o processo de simulação e codificação inicial dos dados foi realizado pelo programa GENES versão 2008.6.0.

**Tabela 1-** Amostra do arquivo de dados gerado pelo programa GENES vs 2008.6.0.

Subpopulação	Ind.	Loco (j)				
		1	2	...	9	10
1	1	11	12	...	11	22
...	2	....	...	...	...	...
...	....	...	...	...	...	...
1	20	11	22	....	11	12
2	1	12	12	....	12	22
...	2	...	...	...	...	...
...	....	...	...	...	...	...
2	20	12	12	....	22	11
...	....	...	...	...	...	...
...	....	...	...	...	...	...
100	1	12	12	....	11	22
...	2	...	...	...	...	...
...	....	...	...	...	...	...
100	20	12	11	....	22	12

### 3.3 Análises de Diversidade Genética

Inferiu-se sobre a estrutura e diversidade genética ao nível intra e inter-populacional por meio das seguintes técnicas biométricas.

I. Endogamia e Heterozigosidade.

II. Distância Shannon – Wiener (1978).

III. Índice de diversidade de Nei (1973).

IV. Índice de Fixação de Wright (1951, 1978).

V. Heterozigosidade de Weir (1996).

VI. Análise de variância molecular (AMOVA) de Excoffier et al., (1992).

Diferentes parâmetros foram empregados para quantificar o conteúdo de informação de cada população e, assim, caracterizar a diversidade genética em relação ao efeito da deriva genética, conforme as seguintes medidas:

#### a. Frequência alélica:

A frequência alélica foi estimada a partir de ocorrência das diferentes classes genotípicas. Assim, denominando a ocorrência de homozigotos na população de  $n_{ii}$  e de heterozigotos de  $n_{ij}$ , tem-se:

$$p_i = f(A_i) = \frac{2n_{ii} + \sum_{j=1, j \neq i}^a n_{ij}}{2n}$$

Sendo  $a$  o número de alelos apresentados pelo loco estudado e  $n$  o número total de indivíduos da população.

**b. Heterozigose observada**

$$H_o = \frac{\sum_{j=1, j \neq 1}^a n_{ij}}{n}$$

**c. Heterozigose esperada**

$$H_e = 1 - \sum_{j=1}^a p_i^2$$

**d. Conteúdo médio de informação polimórfica (PIC)**

$$PIC = 1 - \sum_{i=1}^a p_i^2 - \sum_{i, j=1(i \neq j)}^a \sum_{j=1}^a p_i^2 p_j^2$$

O valor do PIC fornece uma estimativa do poder discriminatório do marcador por considerar não somente o número de alelos por loco, mas também a frequência relativa desses alelos. De maneira geral, os valores de PIC são inferiores aos obtidos para a heterozigosidade.

**e. Coeficiente de endogamia**

Estima-se o coeficiente médio de endogamia da população por meio de:

$$F = 1 - \frac{H_e}{H_o}$$

*H<sub>o</sub>*: frequência de heterozigotos numa população supostamente em equilíbrio de Hardy-Weinberg.

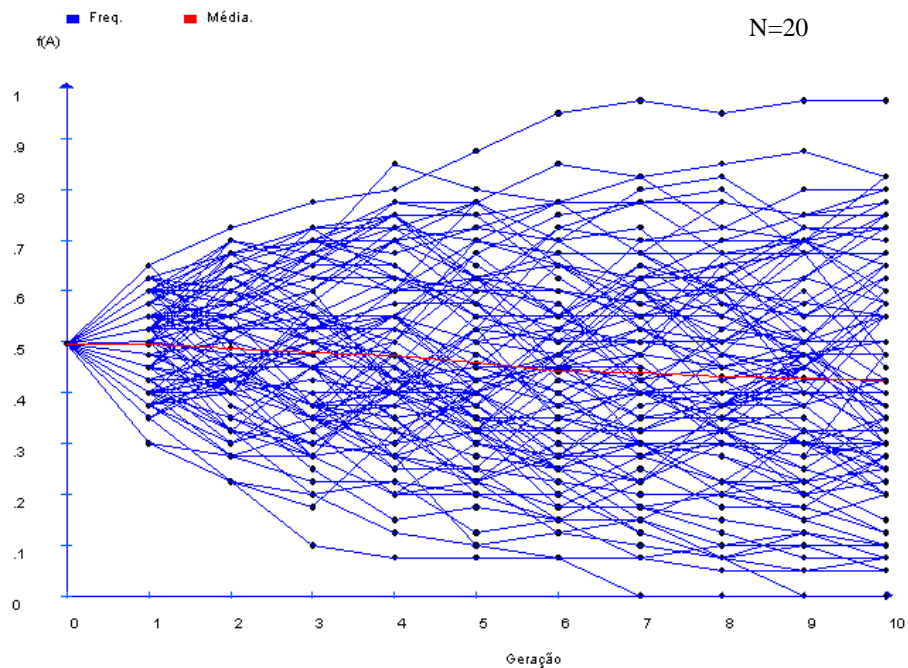
*H<sub>e</sub>*: frequência de heterozigotos numa população sujeita a acasalamento entre aparentados (dada por  $2pq-2^e$ , quando se consideram apenas dois alelos por loco e  $e = pqF$ ).

## **4. RESULTADOS E DISCUSSÃO**

### **4.1 Distribuição das Frequências Gênicas**

Os resultados demonstram uma redução na amplitude da variação nas frequências gênicas com o incremento de número de indivíduos por população, o qual é mais representativo na subpopulação com tamanho  $N=20$  (Figura 2). Todas as subpopulações iniciam com frequência de 0,5. Apresentam-se flutuações muito irregulares variando de 0 a 1, enquanto as populações constituídas por mais indivíduos apresentam uma flutuação perto da meia. Fica claro que as frequências gênicas com influência da deriva genética se comportaram de modo irregular em qualquer população fazendo sua previsão praticamente impossível.

Nas subpopulações constituídas por  $N=20$  indivíduos observou-se que 30% dos locos (2, 3, 4) apresentam fixação dos alelos com o transcorrer das dez gerações de acasalamento ao acaso. Enquanto 70% dos locos atingiram frequências alélicas muito próximas de 0 ou 1, com padrão de distribuição característicos das figuras 2 e 3. Isso indica que populações naturais fragmentadas irão experimentar estes efeitos em todos os seus locos genéticos, com fragmentos pequenos sofrendo deriva genética mais intensa do que fragmentos maiores (Frankham et al., 2008).

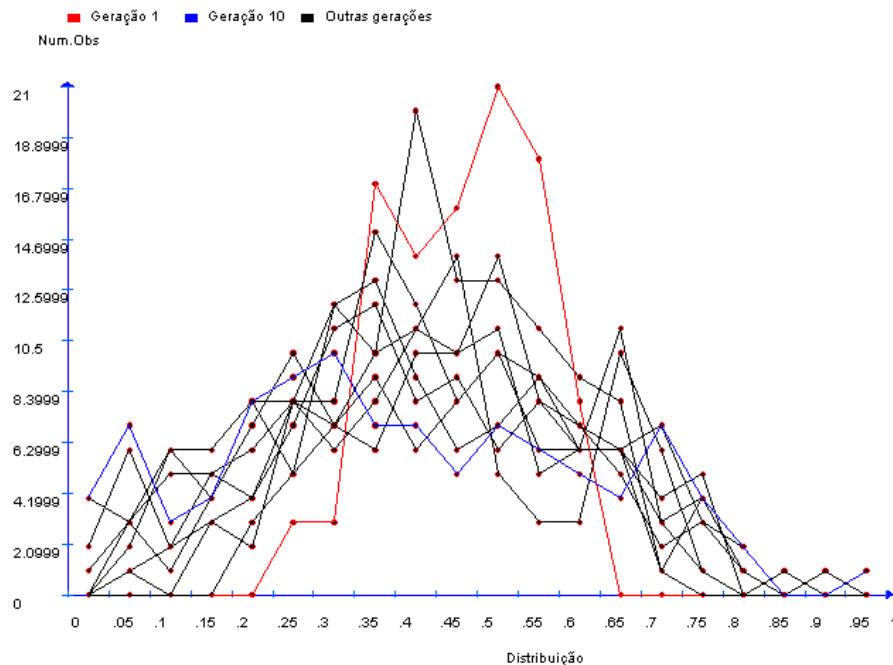


**Figura 2-** Oscilação ao acaso das frequências alélicas em 100 subpopulações, cada uma com 20 indivíduos simulados por 10 gerações de acasalamento ao acaso.

A Figura 2 caracteriza o fato de que a fixação não começa imediatamente e alguma dispersão das frequências gênicas ocorre antes, que qualquer linha tenha probabilidade de atingir a fixação. Falconer & Mackay (1997) citam que, uma vez que as amostras populacionais estejam sob acasalamentos ao acaso (população panmítica) e em Equilíbrio de Hardy-Weinberg, as frequências gênicas médias em uma geração serão as mesmas que na geração anterior, ao longo prazo, as frequências gênicas e genotípicas médias serão as do equilíbrio, mas com um pouco de distorção ao redor delas. No entanto, não é isso o que ocorre. O equilíbrio polimórfico de Hardy-Weinberg pode ser instável uma vez que a deriva genética ocorre, os efeitos da deriva fazem com que as frequências gênicas flutuem em todas as direções. Os genes são aleatoriamente amostrados para produzir a próxima geração, fazendo com que, em longo prazo, um dos alelos eventualmente, seja fixado ou perdido.

Em alguns locos, as frequências gênicas alcançaram limites de 0 ou 1, em algumas gerações (Figura 3). Nesta figura, fica em maior evidência a primeira e a última geração de acasalamento ao acaso onde, no pico da primeira geração (vermelho), cerca de 5 amostras (21-16) possuem frequências entre 0,53 e 0,47. Na última geração de

acasalamento ao acaso (azul), cerca de 2 amostras populacionais possuem freqüências de 0.95.



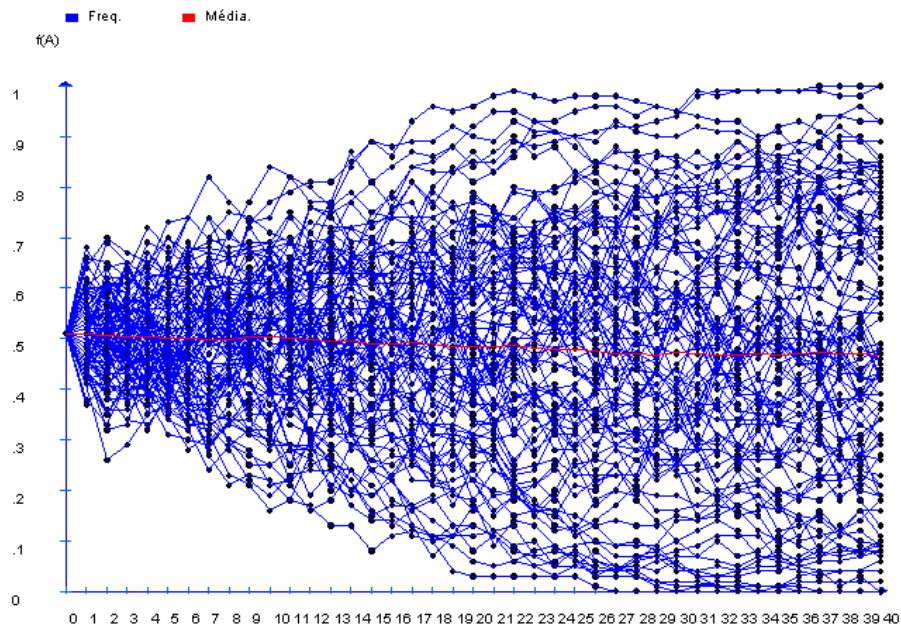
**Figura 3-** Freqüência do alelo (A), com tamanho  $N=20$ , ao longo de dez gerações de acasalamentos ao acaso.

Existem limites para a diferenciação entre as linhas, ocasionada pelo processo dispersivo. A freqüência gênica não pode ser mudada além dos limites de 0 a 1 e, eventualmente, cada linha irá atingir um desses limites. Quando um alelo particular atingir a freqüência de 1 diz-se que ele foi fixado naquela linha, quando atingir a freqüência de 0, será perdido. Uma vez fixado o alelo em uma linha, neste ponto, os indivíduos de uma linha são, então, geneticamente idênticos, com respeito àquele loco. (Falconer & Mackay, 1997)

Nas subpopulações com cinquenta indivíduos, não foi encontrada nenhuma freqüência alélica próxima de 0 ou 1. Entretanto, observou-se que os locos apresentaram amplitude das freqüências gênicas de 0,15 a 0,9, ao longo de todas as gerações (Figuras 4 e 5). Observou-se que não há oscilações bruscas das freqüências gênicas, ao ponto de se distanciarem tanto da média e acontecer fixação dos alelos, como o observado nas subpopulações de 20 indivíduos.



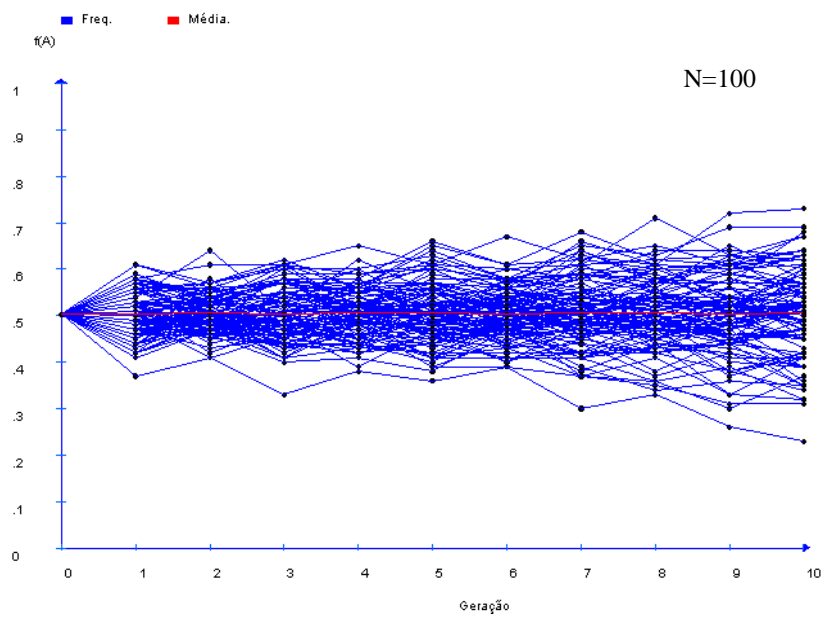
Para estabelecer com quantas gerações começa a acontecer fixação dos alelos em populações constituídas por 50 indivíduos, se efetuar uma simulação durante quarenta gerações de acasalamento ao acaso. O primeiro caso de fixação é que ocorreu com 26 gerações (Figura 6), ilustrando claramente que a deriva genética é maior em pequenas populações.



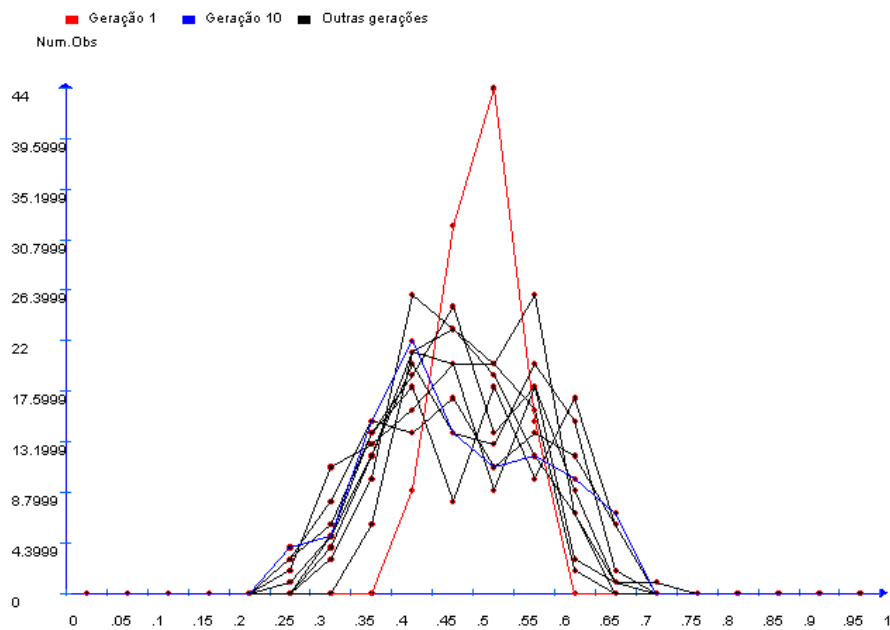
**Figura 6-** Deriva genética ao acaso de dois alelos. Tamanho populacional  $N=50$  em quarenta gerações de acasalamento ao acaso.

Ridley (2006) cita que existem duas fases do processo dispersivo. Durante a fase inicial, as frequências gênicas estão se distanciando do valor inicial (Figura 6). Essa conduz a uma segunda fase, estacionária, que ocorre quando as frequências gênicas estão uniformemente espalhadas sobre a extensão entre os dois limites, e todas as frequências gênicas, exceto os dois limites, são igualmente prováveis.

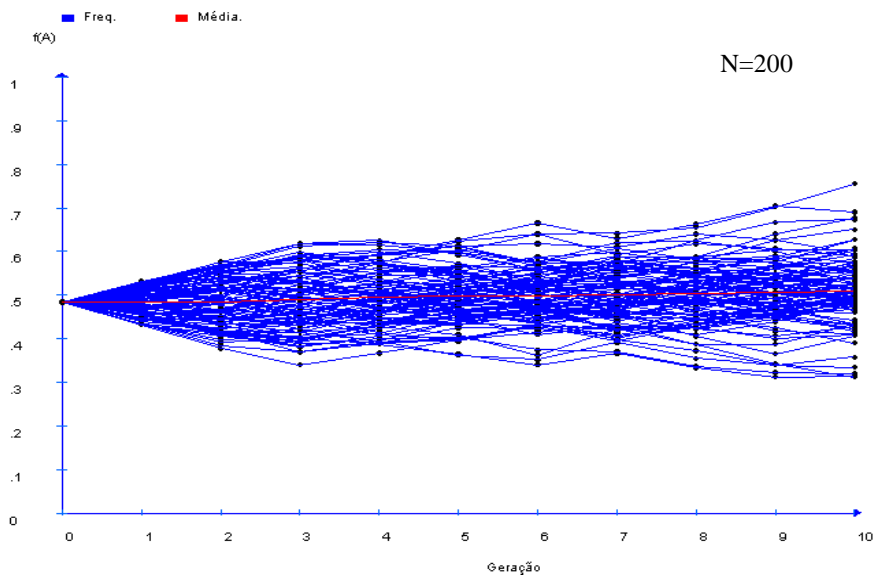
As populações com 100 e 200 indivíduos amostrados não apresentaram frequências alélicas próximas de 0 ou 1. Observou-se que as populações apresentaram frequências alélicas muito próximas da média 0,5 ao longo de todas as gerações (Figura 7, 8, 9 e 10).



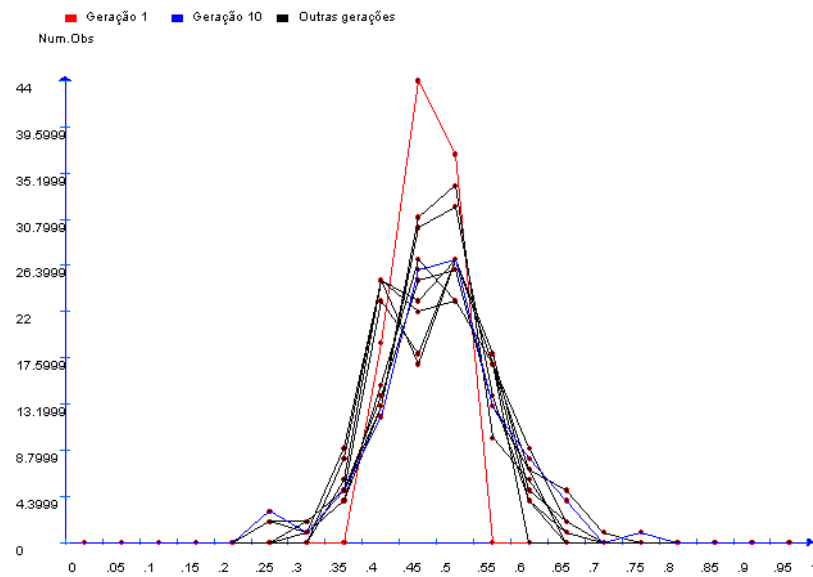
**Figura 7-** Deriva genética ao acaso de dois alelos. Tamanho populacional N=100 em dez gerações de acasalamento ao acaso.



**Figura 8-** Freqüência do alelo (A), com tamanho N=100, ao longo de dez gerações de acasalamentos ao acaso.



**Figura 9-** Deriva genética ao acaso de dois alelos. Tamanho populacional  $N=200$  em dez gerações de acasalamento ao acaso.



**Figura 10-** Frequência do alelo (A), com tamanho  $N=200$ , ao longo de dez gerações de acasalamentos ao acaso.

Comparando-se a população constituída de vinte indivíduos na segunda geração (Figura 2) e a população de duzentos indivíduos, na décima geração (Figura 9), pode-se determinar que a amplitude da frequência apresentada pelas populações se equivale. A deriva para a homozigotidade é mais rápida em uma população pequena, porém, em

qualquer população de tamanho finito, na ausência de mutação, a homoziguidade será o resultado final. Ridley (2006) comenta que se a população é grande, as frequências gênicas ficarão constantes de geração a geração e as frequências genotípicas também serão constantes, nas proporções de Hardy-Weinberg. No caso de um único loco gênico, com dois alelos seletivamente neutros A e a, o Equilíbrio de Hardy-Weinberg supõe a ausência de deriva genética em  $p^2$ ,  $2pq$  e  $q^2$ , mas em populações menores como no caso das populações de 20 indivíduos, as frequências gênicas podem flutuar em todas as direções.

Portanto, com o passar do tempo, a deriva genética faz com que a estrutura genética da população seja direcionada para a homoziguidade em um loco uma vez que as frequências alélicas tendem a valores extremos. A amostragem aleatória pode fazer com que mudem as frequências gênicas do alelo A, de um determinado gene A/a, com mesma probabilidade de diminuir ou de aumentar, entre gerações (Ridley, 2006).

## **4.2 Diversidade Genética em Nível Intrapopulacional**

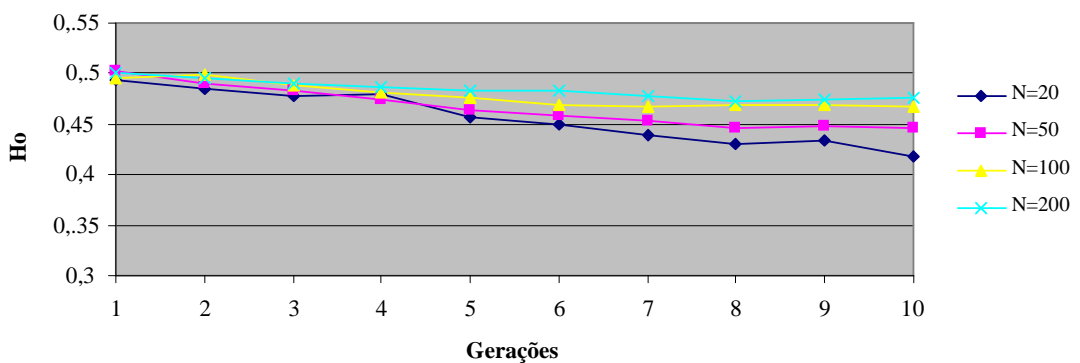
### **4.2.1 Índice de Diversidade - Endogamia e Heterozigose**

Como esperado, a deriva genética causou perdas de diversidade genética nas populações simuladas (Figura 11). A perda da diversidade genética ou da heteroziguidade foi mais rápida nas populações menores, concordando com Wright (1932 e 1951), quem afirmou que a deriva genética acontece de forma mais rápida dentro de pequenas populações. Os resultados mostraram pequena redução dos valores de heterozigose observada ( $H_o$ ) após 10 gerações, indicando aumento da frequência de homozigotos a expensas dos heterozigotos (Falconer & Mackay, 1997). Nas populações de 20 indivíduos as amplitudes de  $H_o$  foram (0,493 – 0,417) enquanto as populações de 200 indivíduos (0,500 - 0,475) no transcurso das dez gerações (Tabela 2), ratificando que a diversidade genética dentro de uma população pode ser quantificada por meio dos parâmetros de heterozigose esperada ( $H_e$ ) e observada ( $H_o$ ) graças à oscilação genética resultante do processo de acasalamento associado ao tamanho da população.

**Tabela 2-** Heterozigose observada, das médias de dez locos em dez gerações.

Gerações	N=20	N=50	N=100	N=200
1	0,49345	0,50212	0,4958	0,5006
2	0,48535	0,49051	0,49884	0,49476
3	0,477	0,48309	0,4878	0,490905
4	0,4798	0,47479	0,4812	0,486335
5	0,45705	0,46455	0,47528	0,48304
6	0,4495	0,45802	0,46976	0,482305
7	0,4393	0,45293	0,46652	0,477605
8	0,43025	0,44635	0,46836	0,472005
9	0,43295	0,4474	0,46866	0,47407
10	0,41765	0,447	0,46712	0,4756

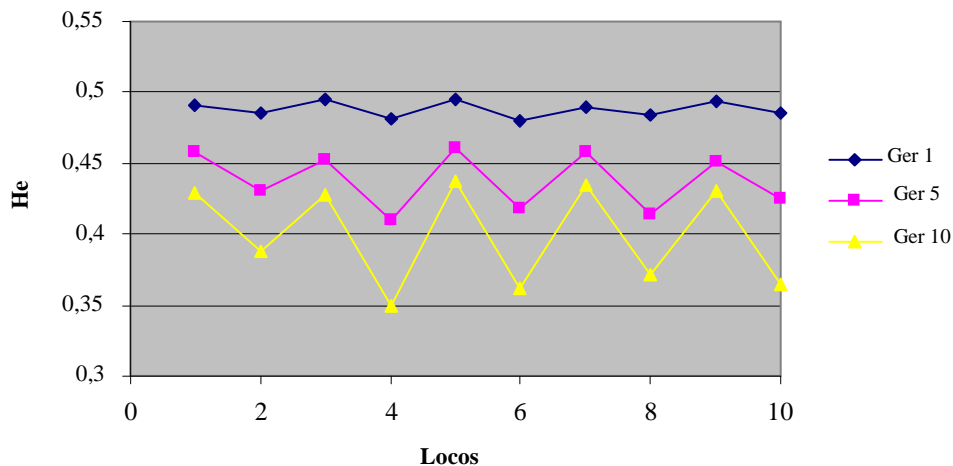
Teixeira et al., (2007) observaram, em seus estudos, que os valores da heterozigosidade observada variaram de 0,133 a 0,556, e a heterozigosidade esperada variou de 0,170 a 0,810 em seis locos de duas pequenas populações constituídas por 6 e 11 árvores de *Hymenaea stigonocarpa*. Cinco locos apresentaram excesso significativo de homozigotos em comparação ao esperado pelo modelo de Equilíbrio de Hardy-Weinberg. Estes resultados decorem, possivelmente, do efeito da deriva genética sobre as populações em estudo.



**Figura 11-** Heterozigose observada, das médias de dez locos em dez gerações.

A Heterozigose esperada em todas as populações (20, 50, 100, 200) apresenta desvios do equilíbrio de HW em *loco*, suas frequências genóticas não são mais as frequências de Hardy-Weinberg (HW) apropriadas à frequência gênica original ou média (Figura 12). Resultados similares foram encontrados por Luikart et al., 1999 e Barker et al., 2001, trabalhando com marcadores co-dominantes, onde detectarem déficit global de heterozigotos em todos os loci analisados (efeito Wahlund,  $P < 0,001$ ) como resultado de congregar as mostras de indivíduos de diferentes grupos com déficit de heterozigotos. Tal desvio pode ser devido à ocorrência de alelos nulos ou não amplificado, vícios de grupamento de alelos, efeito da estrutura de amostragem (efeito Wahlund).

Como se trata de populações simuladas sabe-se que a causa do desvio do HW seja originado por efeito Wahlund onde a relação de HW entre frequências alélicas e genóticas, embora se mantenham bem dentro de cada subpopulação, separadamente, não se mantém caso estas sejam consideradas em conjunto como uma única população. Neste caso, haverá aumento da frequência de homozigotos na população. Em outras palavras, quando a população é dividida em grandes isolados, o isolamento não terá efeito evolutivo, pois somente haverá alteração das frequências genóticas, mas não das frequências gênicas. (Wahlund, 1928).



**Figura 12-** Heterozigose esperada, de três gerações em dez locos em populações de tamanho 20.

Pode-se concluir com base nos resultados, que os parâmetros de heterozigosidade observada e heterozigosidade esperada são adequados para quantificar a diversidade

genética dentro de uma população, mas para apresentar o efeito da deriva genética seria preciso muitas gerações de acasalamento ao acaso, em populações de tamanho acima de 50 indivíduos.

#### **4.2.2 Índice de Diversidade - Shannon-Wiener**

O índice de Shannon-Wiener é a medida do nível de incerteza ao predizer corretamente a espécie a que pertence o próximo indivíduo coletado em uma amostra. Mais precisamente, o índice de diversidade de Shannon-Wiener ( $H'$ ) tem sido utilizado para expressar a biocenose numa expressão numérica e estabelecer uma interdependência entre a uniformidade de um sistema e seu número de espécies (riqueza). Também, tem sido empregado especificamente em estudos genéticos (Holcomb et al., 1977 e Picoli et al., 2004) como medida de diversidade dentro da população e se assemelha a um índice de riqueza genotípica. O valor máximo de  $H'$  ( $H'_{\text{máx}}$ ) ocorre quando se tem frequências genotípicas iguais no loco e pode ser definido como  $H'_{\text{máx}} = \log(nc)$ , em que  $nc$  é o número de classes genotípicas observadas. Assim, os valores de  $H'$  se distribuem no intervalo de 0 a  $\log(nc)$ . Quando se deseja que  $H'_{\text{máx}}$  seja igual a unidade (1), normaliza-se seu valor dividindo-o por  $\log(nc)$ .

Vale ressaltar a importância de índices genotípicos, principalmente em estudos de melhoramento de plantas. A seleção artificial atua diretamente no organismo ao nível genotípico, portanto, medidas de diversidade gênica nem sempre são capazes de contabilizar e discernir a variabilidade genotípica entre ou dentro de populações. Para um determinado loco, duas populações podem ter as mesmas frequências alélicas, mas possuírem diferentes arranjos genotípicos (Hedrick, 1971).

A estatística  $H'$  Shannon-Wiener, em todas as interações avaliadas foi ineficiente em detectar a diferenciação genética dentro das populações proporcionada pela deriva genética em 10 gerações. Para tamanho igual a 20, o valor médio das 100 populações de  $H'$  na primeira geração foi 0,968 sendo reduzido para 0,853 depois das sucessivas gerações. Resultados similares foram determinados para os diferentes tamanhos populacionais (Tabela 3).

**Tabela 3-** Valores médios da estatística Shannon-Wiener ( $H'$ ) de 100 populações para 10 locos, obtidas em diferentes tamanhos de populações na geração 1, 5 e 10.

<b>Ind./pop.</b>	<b>Geração 1</b>	<b>Geração 5</b>	<b>Geração 10</b>
N=20	0,968	0,916	0,853
N=50	1,028	0,990	0,913
N=100	1,011	0,992	0,937
N=200	1,034	1,001	0,992

A estatística  $H'$  não apresenta diferenças entre os valores ao se avaliar isoladamente cada loco nas 100 populações, isto se deve ao fato que um loco separado não consegue expressar o efeito da deriva genética sobre as frequências genotípicas. Isto se pode observar nos valores de dez populações para os quatro tamanhos populacionais (Tabela 4).

**Tabela 4-** Estatística ( $H'$ ) Shannon-Wiener para um loco obtida em 10 populações sob diferentes tamanhos populacionais nas gerações 1 e 10.

<b>Populações</b>	<b>N=20</b>		<b>N=50</b>		<b>N=100</b>		<b>N=200</b>	
	<b>Gerações</b>							
	<b>1</b>	<b>10</b>	<b>1</b>	<b>10</b>	<b>1</b>	<b>10</b>	<b>1</b>	<b>10</b>
1	1,029	0,823	1,019	1,018	1,032	0,943	1,026	0,821
2	1,029	0,974	0,961	0,915	1,046	1,039	1,044	0,998
3	1,048	0,673	1,072	1,013	1,029	1,004	1,048	1,023
4	1,039	0,791	1,029	0,974	1,020	1,027	1,038	0,997
5	0,823	0,948	1,045	1,013	1,032	1,017	1,041	0,994
6	1,096	1,010	1,039	0,946	1,004	0,908	1,016	1,023
7	1,029	0,943	1,025	1,023	0,967	0,928	1,027	1,063
8	0,801	1,048	0,997	1,029	1,011	1,025	1,048	1,029
9	1,080	0,791	0,987	1,006	1,031	0,999	1,010	1,0008
10	0,937	0,845	0,989	1,029	1,024	1,046	1,041	1,017

A ineficiência da estatística  $H'$  para diferenciar as populações de deriva genética pode ser causada pela especialidade da mesma, pois esta é adequada para análise de

amostras coletadas ao acaso em grandes populações, considerando a riqueza das espécies representada pelo número de genótipos (Brower & Zar, 1984).

Neste trabalho, com dados simulados sob um modelo destinado a avaliação dos efeitos da deriva genética, por meio do uso da estatística de Shannon-Wiener os valores obtidos foram pequenos. Entretanto, na literatura encontram-se estimativas bem elevadas. Para ilustrar serão descritos alguns trabalhos práticos cujos valores de  $H'$  foram utilizados para definir o nível de diversidade genética das populações.

Mantovani (1996), encontrou índices de diversidade de Shannon- Wiener de 2,78 e 2,28 em duas populações do rio Cará Cará, sendo o índice de diversidade total  $H'=2,97$ . Em outro trabalho analisando a diversidade e riqueza em espécies de aranhas em duas ambientes do Parque Metropolitano de Pituçu, a diversidade de Shannon-Wiener foi  $H'=3,3741$  e  $H'=2,8705$ , apontando diferenças bastante significativas, entre os ambientes (Alves et al., 2005).

Ávila et al., (2008) encontraram índice de diversidade  $H'=1,6933$  para espécies de himenópteros, seguindo padrão encontrado em outras áreas de cerrado. Almeida (2002), em estudo de Apoidea, obteve, em área de cerradão, índice de diversidade (Shannon-Wiener) de 2,3987. Outras áreas de cerrado apresentaram os seguintes resultados para índice de diversidade, Corumbataí  $H'= 1,70$ ; Paraopeba  $H'=1,92$ ; Cajuru  $H'=1,51$ ; Cassilândia  $H'=0,8920$  (Campos, 1989).

O índice de Shannon-Wiener considera que quanto mais próxima de zero, mais baixa é a diversidade. Nas populações de *E. erythropappus* analisadas em Aiuruoca e São Tomé das Letras, o índice de Shannon  $H'$  foi 0,48 e 0,45, e nas subpopulações I e II de Aiuruoca os valores foram 0,48 e 0,49. Em outras espécies arbóreas como *Populus tremuloides* valores próximos de diversidade foram encontrados de 0,58-0,69 (Yeh, 1997), e para *Aspidosperma polyneuron*, Torezan et al., (2005) obtiveram  $H'= 0,41$  para indivíduos adultos.

### 4.3 Diversidade Genética em Nível Interpopulacional

#### 4.3.1 Estatística F de Wright (1951)

A análise da estrutura genética das populações tem sido baseada em princípios subjacentes à estatística F de Wright (Wright, 1951 e 1978). Os valores médios, máximos e mínimos da estatística F são apresentados na Tabela 5, os quais são utilizados para medir a extensão de endogamia dentro das subpopulações ( $F_{IS}$ ), o nível da diferenciação genética entre as subpopulações ( $F_{ST}$ ) e a redução média em heterozigiosidade de um indivíduo em relação a toda população ( $F_{IT}$ ).

**Tabela 5-** Valores máximos e mínimos dos índices de fixação dentro das populações ( $F_{IS}$ ), de diversidade genética entre populações ( $F_{ST}$ ) e de fixação para o conjunto das populações ( $F_{IT}$ ), conforme Wright (1951)

Tamanho		$F_{IS}$	Loco/Geração	$F_{ST}$	Loco/Geração	$F_{IT}$	Loco/Geração
N=20	Máximo	0,031	3/1	0,401	4/1	0,2643	2/10
	Mínimo	-0,0777	4/6	0,0253	5/2	0,0065	6/1
N=50	Máximo	0,0157	4/4	0,0848	7/8	0,075	3/9
	Mínimo	-0,0279	3/4	0,0064	10/1	-0,0105	3/2
N=100	Máximo	0,0181	10/5	0,1086	5/10	0,1153	8/10
	Mínimo	-0,0269	7/1	0,0054	6/1	-0,038	7/1
N=200	Máximo	0,005	6/8	0,066	9/8	0,0623	6/8
	Mínimo	-0,0232	9/8	0,0041	6/1	-0,0086	1/1

$F_{IS}$  Coeficiente de endogamia devido a acasalamentos endogâmicos preferenciais;  $F_{ST}$  grau de diferenciação entre populações;  $F_{IT}$  redução da heterozigiosidade devido a endogamia dentro de populações e deriva genética entre populações.

O índice de fixação/endogamia é um dos parâmetros mais importantes em genética de populações, por medir o balanço entre homozigotos e heterozigotos nas populações. A explicação para populações ou espécies que contém maior número de locos em homozigose e menor em heterozigose, deve estar associada ao sistema reprodutivo e, ou, à deriva genética. O principal efeito da endogamia é diminuir a

heterozigotidade da população, quando comparada a heterozigotidade esperada (Kageyama, 2003). Em termos biológicos, é possível quantificá-la pela comparação das proporções de heterozigotos observados e heterozigotos esperados. Dentro deste conceito, a estimativa do coeficiente  $F$  possibilita mensurar a deficiência ou excesso de heterozigotos nas populações. A redução dos indivíduos de uma população causa o aumento das taxas de endogamia aleatória e maiores efeitos da deriva genética, levando a perdas de alelos importantes.

As estimativas dos parâmetros  $F$  de Wright revelaram valores semelhantes, para os índices de fixação dentro de populações ( $F_{IS}$ ), para todos os quatro tamanhos populacionais, indicando baixo coeficiente de endogamia sugerindo uma dispersão dos indivíduos apropriada para regular a taxa de incidência de cruzamento entre os indivíduos relacionados (Chesser, 1991).

Oliveira et al., (2006) obtiveram valores médios para o coeficiente de endogamia de ( $F_{IS}=0,305$ ) sendo o mais elevado de ( $F_{IS}=0,788$ ) revelando portanto que maior parte da diversidade genética encontra-se dentro das populações. Considerando os altos índices encontrados em tais literaturas, os índices aqui obtidos, evidenciam uma condição de panmixia das populações no geral, ou seja, indica acasalamento ao acaso. A magnitude dos valores encontrados aqui concordam com o  $F_{IS}=0,05$  estimados por Wadt & Kageyama (2004), que consideraram uma condição de panmixia de suas populações.

Diversos autores consideram valores de  $F_{ST}$  superiores 0,5 como presença de alta diferenciação interpopulacional. As populações com valores mais altos de diferenciação entre populações  $F_{ST}$  são as formadas por vinte indivíduos (Tabela 5). Estes índices  $F_{ST}$ , foram encontrados a partir da primeira geração de acasalamento ao acaso. Os valores médios de  $F_{ST}$  para os 10 locos demonstram o efeito da deriva genética em alterar as frequências alélicas, apresentando incremento do valor  $F_{ST}$  entre populações com o passar das gerações (Tabela 6), ademais se pode observar que valores de  $F_{ST}$  acima de 0,100 foram alcançados para  $N=20$  após a 5ª geração de acasalamento ao acaso.

O parâmetro  $F_{ST}$  estima o coeficiente de endogamia de uma subpopulação em relação à população total, medindo o grau de diferenciação entre as subpopulações. Pode também ser um índice de fixação, que representa a redução na heterozigotidade em uma subpopulação devido à deriva genética.

Os resultados demonstram maior contribuição da diferenciação interpopulacional à diversidade total que a diferenciação intrapopulacional. Brondani et al., (2005)

também observaram, com marcadores SSR, valores elevados para a diferenciação interpopulacional ( $F_{ST} = 0,847$ ), sendo  $F_{IT} = 0,968$  e  $F_{IS} = 0,794$ . Já Karasawa et al., (2007) obtiveram valores menores para a diferenciação interpopulacional ( $F_{ST}=0,491$ ) em 11 populações de *O. glumaepatula*, sendo que as estimativas da endogamia total ( $F_{IT}=0,888$ ) foram consideradas decorrentes principalmente do sistema reprodutivo, em função da alta endogamia intrapopulacional observada ( $F_{IS}=0,780$ ).

**Tabela 6-** Valores médios de  $F_{ST}$ , para os 10 locos, ao longo das gerações, nos quatro tamanhos populacionais.

Tamanho	Gerações									
	1	2	3	4	5	6	7	8	9	10
N=20	0,025	0,059	0,065	0,078	0,113	0,122	0,142	0,153	0,170	0,187
N=50	0,006	0,019	0,037	0,055	0,070	0,083	0,092	0,099	0,099	0,099
N=100	0,009	0,016	0,031	0,045	0,057	0,072	0,069	0,072	0,065	0,070
N=200	0,004	0,014	0,026	0,038	0,046	0,049	0,054	0,058	0,056	0,050

Uma importante consequência em populações estruturadas em subpopulações é a redução na proporção média de genótipos heterozigotos relativos ao esperado na população única de acasalamento ao acaso (Hartl & Clark, 1997). Oliveira et al., (2006) cita que os valores  $F_{IT}$  negativos indicam excesso de heterozigotos e valores altos e positivos indicam excesso de homozigotos, que podem ser decorrente de endogamia advinda de autofecundação e cruzamento entre aparentados. A Tabela 7 apresenta o número de índices  $F_{IT}$  negativos para cada subpopulação, nas respectivas gerações de ocorrência. Observou-se, portanto, que na subpopulações de 20 indivíduos, a maior quantidade do índice  $F_{IT}$  negativos está concentrada nas primeiras gerações de acasalamentos ao acaso e nas subpopulações de 50 e 100 e 200 indivíduos os índices  $F_{IT}$  negativos ainda estão distribuídos nas 10 gerações, com redução gradual.

**Tabela 7-** Número de índices  $F_{IT}$  negativos para cada subpopulação.

<b>Geração</b>	<b>N=20</b>	<b>N=50</b>	<b>N=100</b>	<b>N=200</b>
1	6	8	10	10
2	3	5	8	10
3	0	4	7	9
4	0	2	6	6
5	0	4	7	8
6	0	5	5	10
7	1	2	5	6
8	0	3	4	4
9	0	0	4	5
10	0	0	4	3

Portanto, com base nos índices negativos de  $F_{IT}$ , nas subpopulações de 20 indivíduos, pode-se inferir que, há uma redução brusca do número de heterozigotos com o passar das gerações, enquanto que nas subpopulações de 50, 100 e 200 indivíduos isto não ocorre. As subpopulações com amostras contendo 20 indivíduos sofrem maior efeito da deriva genética que subpopulações com tamanho maior, tornando a estatística F de Wright apropriada para determinar o efeito da deriva genética sobre populações subdivididas.

Com relação ao efeito de deriva genética, Oliveira et al., (2006), utilizando índices de medição de fixação alélica de três populações, obtiveram baixos valores para uma das populações. Os autores discutem que tais índices observados na população podem ser, provavelmente, atribuídos à ocorrência da deriva genética por efeito de gargalo, devido à redução aleatória no tamanho da população, associado ao aumento intenso do parentesco na mesma. Outra causa pode ser que o efeito de fundação tenha ocorrido na ocasião de povoamento da área. A fundação da população a partir de poucos indivíduos provoca aumento no índice de fixação.

### 4.3.2 Heterozigosidade de Weir (1996)

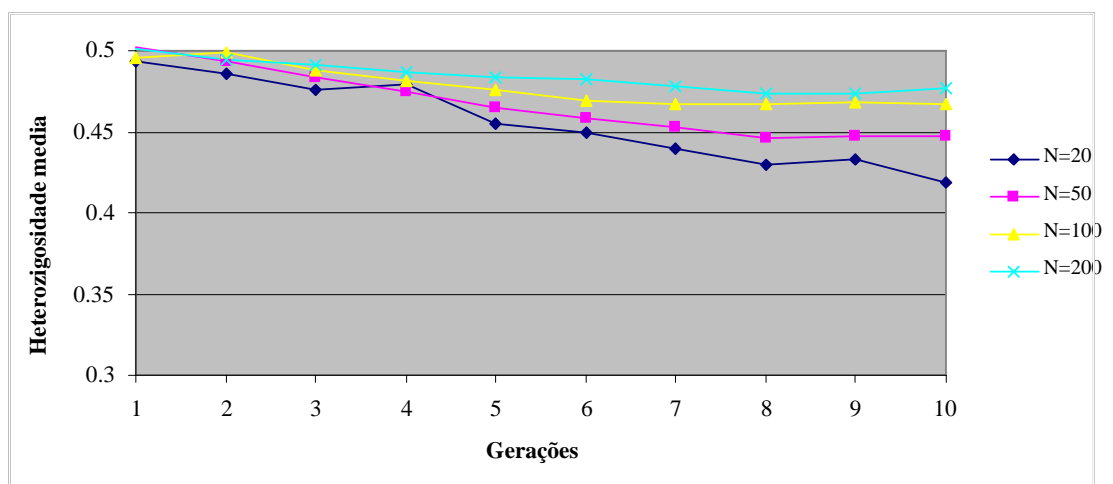
De acordo com Weir (1996) a frequência de heterozigotos é uma maneira de representar a existência de variação genética, desde que carreguem diferentes alelos. No entanto, há situações em que a variação é resultante da presença contínua de diferentes homozigotos, a exemplo das populações de autofecundação. Nessa condição, o termo diversidade gênica torna-se mais apropriado.

O método proposto por Weir (1996), trabalhando com a heterozigosidade média ( $\bar{H}$ ), permite analisar a diversidade genética entre as subpopulações de todos os locos nas 10 gerações. Os resultados mostram que a subpopulação de 20 indivíduos apresenta maior variação nas estimativas de ( $\bar{H}$ ), mostrando perda mais acelerada de indivíduos heterozigóticos após dez gerações de acasalamento, em relação às subpopulações de 50, 100 e 200 indivíduos (Tabela 8), demonstrando que o efeito da deriva genética é mais forte nas populações com um menor número de indivíduos (Figura 13). Entretanto, Weir (1996) trabalha com a frequência observada dos alelos de forma que, analisando as variâncias entre as frequências médias observadas, de todos os locos, nas 10 gerações, a subpopulação de 20 indivíduos apresenta maior variância em relação às subpopulações de 50, 100 e 200 indivíduos. Infere-se, com base nos dados de variância da frequência de heterozigotos observada, maior variabilidade entre subpopulações de 20 indivíduos do que entre subpopulações de 200 indivíduos (Tabela 8).

Os desvios das proporções do EHW indicam que os efeitos da seleção, mistura de populações e acasalamentos não aleatórios ocorrem na população, e é este o primeiro passo nas investigações sobre estruturação genética de populações (Raymond & Rousset, 1995).

**Tabela 8-** Heterozigose média ( $\bar{H}$ ), ao longo de 10 gerações de acasalamentos ao acaso, para 4 diferentes tamanhos populacionais.

Geração	Tamanho			
	N=20	N=50	N=100	N=200
1	0,49385	0,4958	0,50212	0,5006
2	0,48535	0,49884	0,49372	0,49476
3	0,4762	0,4878	0,48309	0,490905
4	0,47915	0,48146	0,47479	0,486335
5	0,4554	0,47528	0,46455	0,48304
6	0,4495	0,46976	0,45802	0,482305
7	0,4394	0,46652	0,45293	0,477605
8	0,43005	0,4671	0,44635	0,47318
9	0,43295	0,46866	0,4474	0,47407
10	0,4182	0,46712	0,447	0,477175
<b>Máximo</b>	0,545	0,55	0,551	0,539
<b>Mínimo</b>	0,29	0,366	0,207	0,3605
<b>Variância</b>	0,00069	0,00042	0,00015	0,00008



**Figura 13-** Heterozigosidade média ( $\bar{H}$ ), ao longo de 10 gerações de acasalamentos ao acaso, para 4 diferentes tamanhos populacionais.

São apresentados os quadros de análise de variância (Heterozigosidade de Weir, 1996), para as populações no estudo (20, 50, 100 e 200 indivíduos) em diferentes gerações de acasalamento ao acaso.

**Tabela 9-** Quadrados médios obtidos da análise de variância da heterozigosidade observada, para os diferentes tamanhos populacionais nas gerações (1ª, 3ª, 5ª, 7ª e 10ª) de acasalamento ao acaso.

**Tabela 9a-** Tamanho populacional: N=20

FV	GL	Gerações				
		1	3	5	7	10
Populações (Pop.)	99	0,116**	0,192**	0,378 <sup>ns</sup>	0,813 <sup>ns</sup>	0,530 <sup>ns</sup>
Indivíduos/Pop.	1900	0,292**	0,246 <sup>ns</sup>	0,238 <sup>ns</sup>	0,946 <sup>ns</sup>	0,220 <sup>ns</sup>
Locos	9	0,007 <sup>ns</sup>	0,034 <sup>ns</sup>	0,778**	2,173**	2,672**
Locos x Popul.	891	0,191 <sup>ns</sup>	0,221 <sup>ns</sup>	0,240 <sup>ns</sup>	1,278**	0,387**
Locos x Ind/Pop	17100	0,249	0,251	0,248	0,813	0,235
Total	19999					

ns: não significativo, \*\* significativo a 1 %, \* significativo a 5 %.

**Tabela 9b-** Tamanho populacional: N=50

FV	GL	Gerações				
		1	3	5	7	10
Populações	99	1,551**	1,569**	1,798**	1,534**	2,528**
Indivíduos/Pop.	4900	1,122**	1,090**	1,054**	1,012**	0,991 <sup>ns</sup>
Locos	9	0,598 <sup>ns</sup>	0,234 <sup>ns</sup>	0,602 <sup>ns</sup>	0,344 <sup>ns</sup>	0,405 <sup>ns</sup>
Locos x Popul.	891	0,357 <sup>ns</sup>	0,421 <sup>ns</sup>	0,434 <sup>ns</sup>	0,702 <sup>ns</sup>	0,629 <sup>ns</sup>
Locos x Ind/Pop	44100	1,551	1,569	1,798	1,534	2,528
Total	49999					

ns: não significativo, \*\* significativo a 1 %, \* significativo a 5 %.

**Tabela 9c**-Tamanho populacional: N=100

FV	GL	Gerações				
		1	3	5	7	10
Populações	99	0,457**	0,333**	1,088**	2,569**	4,416**
Indivíduos/Pop.	9900	0,263**	0,257**	0,255**	0,254**	0,257**
Locos	9	0,151**	0,171**	0,163**	0,050 <sup>ns</sup>	0,234**
Locos x Popul.	891	0,066 <sup>ns</sup>	0,078 <sup>ns</sup>	0,088 <sup>ns</sup>	0,120 <sup>ns</sup>	0,138 <sup>ns</sup>
Locos x Ind/Pop	89100	0,250	0,250	0,248	0,245	0,242
Total	99999					

ns: não significativo, \*\* significativo a 1 %, \* significativo a 5 %.

**Tabela 9d**-Tamanho populacional: N=200

FV	GL	Gerações				
		1	3	5	7	10
Populações	99	0,760**	0,722**	0,793**	2,379**	2,717**
Indivíduos/Pop.	19900	0,267**	0,257**	0,268**	0,271**	0,259**
Locos	9	0,212**	0,039**	0,111**	0,114**	0,296**
Locos x Popul.	891	0,053 <sup>ns</sup>	0,075 <sup>ns</sup>	0,091 <sup>ns</sup>	0,105 <sup>ns</sup>	0,114 <sup>ns</sup>
Locos x Ind/Pop	17910	0,248	0,249	0,248	0,246	0,247
Total	199999					

ns: não significativo, \*\* significativo a 1 %, \* significativo a 5 %.

Infere-se pelo resultado da análise de variância que existe diferença significativa entre populações, pelo teste F a 1% de significância nos quatro tamanhos populacionais, assim pode ser observado o efeito da deriva sobre as populações com o passar das gerações de acasalamento, originando uma perda da diversidade mais acelerada nas populações de menor tamanho. Baixos tamanhos efetivos populacionais são provavelmente uma das mais importantes causas da perda de variabilidade genética, em função da deriva genética (Frankham, 1996).

Infere-se pelo resultado da análise de variância, a partir das conclusões obtidas pelo teste F, que a variabilidade significativa entre as populações de 20 indivíduos perde-se a partir da terceira geração, enquanto que nas subpopulações de 50, 100 e 200 indivíduos há diferenciação significativa. O componente locos (estimativa dentro), nas

populações de 20 indivíduos, retira toda a variância do componente populações (estimativa entre), ficando, portanto pouca variância para o componente populações. Provavelmente, a metodologia de Heterozigosidade de Weir (1996) baseada em componentes de variância entre e dentro de populações, não seja uma boa estatística para analisar a magnitude da diversidade genética.

Nybom (2004) compilou 307 trabalhos, publicados de 1993 a 2003, usando marcadores moleculares para avaliação da diversidade genética entre e dentro de populações em espécies silvestres de angiospermas e gimnospermas. Ele explica que um aumento da população analisada leva a um aumento no número de alelos computados, estando assim correlacionado com a diversidade dentro de população. O autor ainda verificou que o número de locos microssatélites estudados teve efeito negativo em  $H_e$ , possivelmente, porque os pesquisadores tendem a investigar mais locos nos táxon, onde cada loco exibe pouco polimorfismo. O número de alelos tem um efeito positivo tanto em  $H_e$  quanto em  $H_o$ , possivelmente porque locos altamente heterozigotos permitem a detecção de um número maior de alelos por loco.

Estes índices de diversidade não podem ser comparáveis entre estimativas obtidas via marcadores isoenzimáticos e microssatélites, pois o poder da técnica de microssatélites em detectar diferentes alelos nos locos, é muito superior às isoenzimas (Kageyama, 2003).

#### **4.3.3 Análise de Variância Molecular (AMOVA)**

Os resultados da análise de variância molecular (AMOVA) mostram variação genética significativa entre populações e variações genéticas dentro, revelando uma concordância com os resultados obtidos pelas demais estatísticas estudadas. Pode ser visualizado, na Tabela 10, que a estimativa máxima de  $\phi_{ST}$  foi de 23,30% nas subpopulações de vinte indivíduos, na décima geração de acasalamentos ao acaso. Esta estimativa indica que após 10 gerações o 76,70 % da variabilidade genética, está distribuída dentro populações. Wright indica que a grande maioria da variabilidade é devida a diferenças de indivíduos dentro de populações.

Nos demais tamanhos de subpopulações 50, 100 e 200 indivíduos, as estimativas máximas da variabilidade genética estimada entre populações ( $\phi_{ST}$ ) foram 14,30; 8,57 e

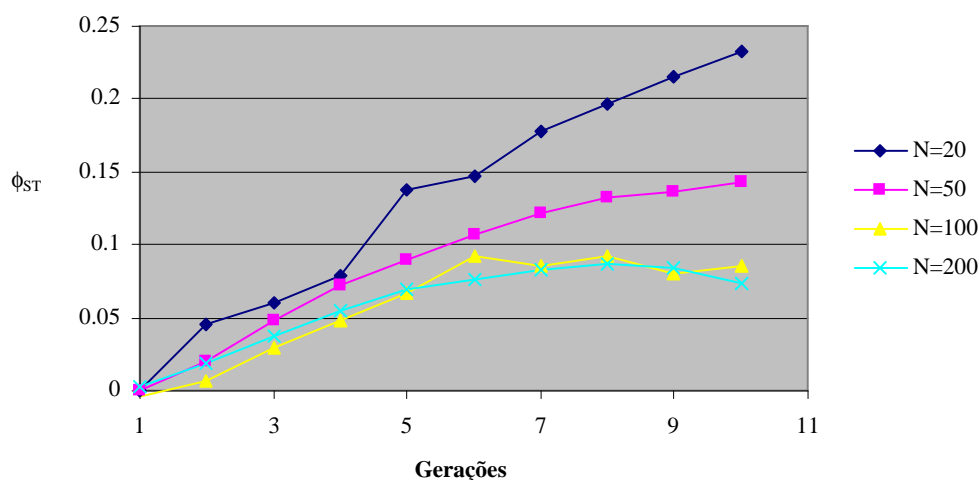
7,41% e, respectivamente (Figura 14). Estes resultados confirmam, portanto, que quanto menor for os tamanhos das subpopulações, maior será a perda da diversidade genética dentro das mesmas, fazendo perceptível o efeito de deriva genética. Também é factível que a magnitude da diferenciação genética, entre populações, ( $\phi_{ST}$ ) tem aumento ao longo de dez gerações de acasalamentos ao acaso para todos os tamanhos populacionais.

Para Zucchi et al., (2003), o alto nível de diferenciação genética entre populações ( $\phi_{ST}=0,2703$ ) é decorrente de efeitos da deriva genética, ocasionados pela ação antrópica e pela predominância de fluxo gênico a pequenas distâncias.

**Tabela 10-** Estimativas de  $\phi_{ST}$  ao longo de 10 gerações de acasalamentos ao acaso, para 4 diferentes tamanhos populacionais.

<b>Gerações</b>	<b>N=20</b>	<b>N=50</b>	<b>N=100</b>	<b>N=200</b>
1	-0,00051	0,000025	-0,00363	0,002236
2	0,045485	0,019592	0,007042	0,018242
3	0,059789	0,048161	0,02918	0,037481
4	0,079017	0,072403	0,048635	0,054417
5	0,137252	0,089267	0,067189	0,06954
6	0,146596	0,106531	0,091874	0,07658
7	0,177276	0,122094	0,085471	0,082484
8	0,196214	0,132319	0,09173	0,086652
9	0,215046	0,136018	0,08034	0,084474
10	0,233048	0,143088	0,085774	0,074156

**Figura 14-** Tendência das estimativas de diversidade genética entre populações ( $\phi_{ST}$ ) ao longo de 10 gerações.



Apresenta-se o quadro de AMOVA e os cálculos das estimativas dos componentes de variância e  $\phi_{ST}$ , na 10ª geração de acasalamento ao acaso, da subpopulação de 20 indivíduos. Este procedimento foi feito para todos os tamanhos de subpopulações, nas dez gerações de acasalamentos ao acaso e amostra a obtenção da estimativa  $\phi_{ST}$  em uma das situações avaliadas.

**Tabela 11-** Resumo da análise de variância molecular (AMOVA), na 10ª geração de acasalamento ao acaso, da subpopulação de 20 indivíduos

<b>FV</b>	<b>GL</b>	<b>QM</b>
Entre	99	1273,834
Dentro	1900	179,990
Total	1999	234,162

<b>Componente de variância</b>	<b>Estimativa</b>	<b>Porcentagem</b>
Entre	54,692	23,304
Dentro	179,990	76,695
Total	234,682	100,0
Estatística $\phi_{ST}$	0,233	

Em estudos de diversidade genética em populações naturais de algumas espécies vegetais, alguns autores como Wallace (2002), Wadt & Kageyama (2004) e Shah et al., (2008) encontraram estimativas  $\phi_{ST}$  de: 21; 28,10 e 56,85%, respectivamente. Estes consideraram estes valores como altos o suficiente para inferir a ocorrência de diferenciação entre populações. As estimativas  $\phi_{ST}$  obtidas, neste trabalho, mais condizentes, com os altos valores encontrados na literatura foram às obtidas em populações com 20 indivíduos a partir da 9ª geração de acasalamentos ao acaso.

Quando o tamanho da população é pequeno Rutten et al., (2004) indica, mesmo em cruzamentos totalmente ao acaso, uma diminuição de heterozigose pode ocorrer. Segundo a AMOVA (Weir & Cockerham, 1984; Excoffier et al., 1992), grande parte da variabilidade genética está concentrada dentro dos plantéis (87,25%), com menor porcentagem entre eles (12,75%). Demonstrando o efeito da subdivisão populacional, que representa a redução da heterozigosidade em uma subpopulação em razão do efeito da deriva genética.

Altos níveis de diferenciação genética foram registrados por Padilha & Damasceno (2008) entre procedência de açazeiro que compõem a coleção da Embrapa Amazônia Oriental, com 0,301 para o marcador dominante e 0,242 para o co-dominante. Para os dois marcadores, a AMOVA apresentou grande variabilidade dentro das procedências (acima de 69%). De modo geral, o alto nível de diferenciação detectado entre procedências pode ter sido influenciado também pela ocorrência de altas taxas de autofecundação e de cruzamentos entre indivíduos aparentados dentro das mesmas originando um efeito de deriva genética.

Um teste de AMOVA foi realizado para investigar a diferenciação genética entre populações de *Panthera onça* em diferentes populações. O teste indicou um valor de  $F_{st}=0,053$  ( $P=0,000$ ) entre todas as populações. Os resultados obtidos entre Ivinhema e Morro do Diabo ( $F_{st}=0,039$ ;  $P=0,018$ ), Ivinhema e Porto Primavera ( $F_{st}=0,021$ ;  $P=0,009$ ), Morro do Diabo e Porto Primavera ( $F_{st}=0,050$ ;  $P=0,000$ ) e Porto primavera e Iguazu ( $F_{st}=0,044$ ;  $P=0,009$ ) indicaram um índice de fixação baixo, porém significativo. A maior diferenciação ocorreu entre Iguazu e Morro do Diabo ( $F_{st}=0,178$ ;  $P=0,000$ ) seguido de Iguazu e Ivinhema ( $F_{st}=0,124$ ;  $P=0,000$ ). Os resultados indicam a existência de uma diferenciação genética entre estes fragmentos, o que pode ser uma consequência de deriva genética recente em populações isoladas muito pequenas (Haag et al., 2007).

#### 4.3.4 Estatística de Nei (1973)

Para os marcadores empregados, calcularam-se as frequências alélicas para cada população amostrada. A diferenciação genética entre as populações foi analisada através da análise da diversidade gênica em populações subdivididas de Nei (1973, 1977, 1987), que permite a comparação dos níveis de heterozigosidade entre e dentro das populações, bem como a obtenção de uma estimativa de divergência, a partir de uma base diferente daquela que é fundamentada nas estimativas de  $F_{ST}$  e  $\phi_{ST}$ . Esta abordagem deriva-se da fórmula de diversidade gênica média total dada por  $H_{IT} = H_{IS} + D_{ST}$ , onde  $H_{IS}$  é a diversidade gênica média dentro das populações, e  $D_{ST}$  é a diversidade gênica média entre as populações, com a diferenciação genética relativa entre populações sendo obtida por  $G_{ST} = D_{ST}/H_{IT}$ . Os valores médios de  $H_{IT}$  e  $H_{IS}$  sobre todos os locos foram usados para calcular  $G_{ST}$  (Tabela 12).

**Tabela 12-** Valores médios de  $G_{ST}$ , para os 10 locos, ao longo das dez gerações, em 4 tamanhos populacionais

Tamanho	Gerações									
	1	2	3	4	5	6	7	8	9	10
N=20	0,0124	0,0250	0,0382	0,0514	0,0659	0,0766	0,0841	0,0905	0,0974	0,1083
N=50	0,0032	0,0108	0,0218	0,0345	0,0444	0,0480	0,0482	0,0443	0,0408	0,0397
N=100	0,0043	0,0022	0,0057	0,0038	0,0071	0,0048	0,0080	0,0077	0,0101	0,0120
N=200	0,0020	0,0064	0,0095	0,0082	0,0076	0,0079	0,0091	0,0100	0,0125	0,0139

Estes testes baseiam-se no princípio de que populações que experimentaram uma recente redução de seu tamanho efetivo populacional exibem uma correlacionada redução do número de alelos e da diversidade gênica ( $H_e$ , ou heterozigosidade esperada pelas proporções de EHW) nos locos polimórficos (Piry et al., 1999).

Observou-se uma concordância entre as estatísticas  $G_{ST}$ ,  $F_{ST}$  e  $\phi_{ST}$ , que apresentam valores inferiores de diferenciação genética nas subpopulações de tamanho maior (50, 100, 200 indivíduos) com relação às subpopulações de tamanhos inferiores

(20 indivíduos) de forma que o grau de diferenciação genética entre as subpopulações de tamanho (50, 100 e 200 indivíduos) é baixo (Tabela 12). Segundo Wright (1978), valores de  $F_{ST}$ , que correspondente ao  $G_{ST}$ , de 0,00 a 0,05 indicam baixa diferenciação genética, valores de 0,05 a 0,15 indicam média diferenciação genética, e valores de 0,15 a 0,25 indicam alta diferenciação genética.

Pode-se estabelecer que o valor  $G_{ST}$  mostrado pelas populações formadas por 20 indivíduos na primeira geração (0,0124) é equivalente ao valor  $G_{ST}$  na nona geração das populações formadas por 200 indivíduos (0,0125). Assim, populações que estiveram submetidas a um efeito de diminuição drástica do tamanho populacional podem entrar em processo de deriva genética que se traduz a uma redução da heterozigose (Avice, 1994).

Hamrick & Godt (1990) realizaram revisão em 653 trabalhos em que foi estudada a diversidade genética por isoenzimas envolvendo 449 espécies de plantas. Neste estudo, os autores encontraram diversidade genética média em nível populacional de 0,160 para gimnospermas, 0,149 para espécies arbóreas, 0,159 para espécies de ampla distribuição, 0,109 para espécies tropicais, 0,124 para espécies polinizadas por animais e 0,123 para espécies com dispersão pelo vento. Valores inferiores a estes foram encontrados em todas as populações em estudo.

Wallace (2002) estudou os efeitos da fragmentação com influências na variação genética em dez populações de *Platanthera leucophaea* (Orchidaceae), para 14 locos polimórficos tendo encontrado estimativas  $G_{ST}$  de 0,26. Shah et al., (2008) encontraram estimativas  $G_{ST}$  de 0,584, e, consideraram estas como altas o suficiente para inferir a ocorrência de diferenciação entre populações.

A Tabela 13 apresenta os resultados das estatísticas (Nei 1987) de diversidade em populações subdivididas. Observa-se que a diversidade gênica total ( $H_T$ ), que é essencialmente a heterozigosidade esperada, foi similar para todas as populações na primeira geração e apresentou perda com o transcorrer das mesmas com uma maior perda nas subpopulações com 20 indivíduos. Por sua vez, a diversidade gênica dentro das populações ( $H_S$ ) para as subpopulações de 20 indivíduos na primeira geração, foi responsável por 98,78% da diversidade gênica total, enquanto para a décima geração foi unicamente de 80,76%. As demais subpopulações tem um comportamento similar com menor perda de diversidade, indicando a existência de maior variabilidade ocorrendo dentro das populações do que entre as mesmas. Isto também confirma o fato de que, com

o passar do tempo, a deriva genética age de forma que a divergência gênica é aumentada entre subpopulações e é reduzida dentro das mesmas. Dentro de subpopulações a homoziguidade é aumentada após muitas gerações de acasalamentos ao acaso.

**Tabela 13-** Estimativas das estatísticas (Nei 1972) de 4 populações para dez gerações

Geração	Tamanho											
	N=20			N=50			N=100			N=200		
	H <sub>IT</sub>	H <sub>IS</sub>	D <sub>ST</sub>	H <sub>IT</sub>	H <sub>IS</sub>	D <sub>ST</sub>	H <sub>IT</sub>	H <sub>IS</sub>	D <sub>ST</sub>	H <sub>IT</sub>	H <sub>IS</sub>	D <sub>ST</sub>
1	0,494	0,488	0,006	0,498	0,496	0,002	0,498	0,496	0,002	0,499	0,498	0,001
2	0,488	0,476	0,012	0,494	0,489	0,005	0,499	0,498	0,001	0,497	0,493	0,003
3	0,482	0,463	0,018	0,489	0,479	0,011	0,497	0,494	0,003	0,495	0,491	0,005
4	0,475	0,451	0,024	0,484	0,467	0,017	0,498	0,496	0,002	0,496	0,492	0,004
5	0,468	0,438	0,031	0,479	0,458	0,021	0,497	0,493	0,004	0,496	0,493	0,004
6	0,463	0,428	0,035	0,477	0,454	0,023	0,498	0,495	0,002	0,496	0,492	0,004
7	0,459	0,420	0,038	0,476	0,453	0,023	0,496	0,492	0,004	0,496	0,491	0,005
8	0,455	0,415	0,041	0,477	0,456	0,021	0,496	0,492	0,004	0,495	0,490	0,005
9	0,452	0,409	0,044	0,479	0,459	0,020	0,495	0,490	0,005	0,494	0,488	0,006
10	0,447	0,399	0,048	0,479	0,460	0,019	0,494	0,488	0,006	0,493	0,486	0,007

H<sub>IT</sub> = diversidade gênica total; H<sub>IS</sub> = diversidade gênica dentro das amostras; D<sub>ST</sub> = diversidade gênica entre as amostras, dependente do número de amostras; D<sub>ST</sub> = diversidade gênica entre as amostras.

Ridley (2006) cita que Dobzhansky e Pavlovsky (1957), trabalhando com *Drosophila pseudoobscura*, obtiveram 10 populações com 4 mil membros iniciais (populações grandes) e 10 com 20 membros iniciais (populações pequenas) e observaram as mudanças nas frequências de duas variantes cromossômicas durante 18 meses. O efeito médio foi o mesmo nas populações grandes e pequenas, mas a variabilidade foi significativamente maior entre as populações pequenas. Este mesmo autor cita que a taxa de mudança da frequência gênica por deriva genética depende do tamanho da população e que efeitos da amostragem aleatória são mais importantes em populações pequenas.

Alzira et al., (2006) estimaram os valores médios do índice de fixação dentro das populações (F<sub>IS</sub>), divergência genética entre populações (F<sub>ST</sub>) e índice de fixação para o conjunto das populações (F<sub>IT</sub>) conforme Nei (1977). Os valores médios obtidos para estes índices mostram que a maior parte da diversidade genética encontra-se dentro das

populações (0,788), embora exista grande variação entre as populações ( $F_{ST}=0,212$ ). Tais índices revelam que as populações podem estar isoladas. Já a diversidade observada dentro das populações pode ser considerada alta (0,788). A diversidade genética mantida dentro e entre populações ocorre em função de eventos históricos e processos evolutivos recentes (Lee et al., 2002).

## 5. CONCLUSÕES

A alteração da frequência gênica pelo efeito da deriva genética depende do tamanho da população. Quanto menor a população, mais evidentes são os efeitos da amostragem aleatória, por deriva genética. Em subpopulações com 20 indivíduos esses efeitos são mais evidentes do que nas subpopulações com 200 indivíduos.

Nas subpopulações com 20 indivíduos houve fixação de alelos em dez gerações de acasalamento ao acaso enquanto as subpopulações com mais de 50 indivíduos só apresentam fixação após 20 gerações.

Os índices de fixação podem ser considerados indicadores de diferenciação genética, pois permitem uma comparação dos efeitos de uma população estruturada em subpopulações.

A Heterozigosidade observada  $H_o$  média é um bom indicador do efeito da deriva genética dentro das subpopulações, demonstrando redução da frequência de heterozigotos.

Em todas as interações avaliadas a estatística  $H'$  foi ineficiente em detectar a diferenciação das populações proporcionada pela deriva genética em 10 gerações.

A diversidade genética medida pelas estatísticas  $G_{ST}$ ,  $F_{ST}$  e  $\phi_{ST}$  aumentaram sua magnitude com o avanço das gerações, constituindo-se em bons indicadores da diferenciação entre populações.

O método de Weir pode ser empregado para representar efetivamente a existência do efeito de deriva genética, por ser fundamentada na quantidade observada de heterozigotos na população.

A variância entre subpopulações de tamanhos pequenos é maior que a variância entre subpopulações de tamanhos maiores.

## 6. REFERÊNCIAS BIBLIOGRÁFICAS

- ABJAUD, R. E.; MARCOS DE SOUZA, A.; MOURA, M.; BOTREL, M.; MENDONÇA E.; CARVALHO, D. Genetic diversity in natural populations of candeia (*Eremanthus erythropappus* (DC.) MacLeish). **Scientia Forestalis**. n. 70, p. 97-106. , abril 2006
- ALMEIDA, D. Espécies de abelhas (Hymenoptera, Apoidea) e tipificação dos méis por elas produzidos em área de cerrado no município de Pirassununga, Estado de São Paulo. 2002. 103 f. Dissertação (Mestrado em Entomologia) - Escola Superior de Agricultura "Luiz de Queiroz", Piracicaba, 2002.
- ALTER, S.E.; RYNES, E. & PALUMBI, S.R. DNA evidence for historic population size and past ecosystem impacts of gray whales. **Proceeding National Academy Press**, n.38, 104: 15162–15167, 2007.
- ALVES, A.; LIMA, M.; DIAS, M.; FERREIRA, G.; SOUTO, L. Estudo das comunidades de aranhas (Arachnida: Araneae) em ambiente de Mata Atlântica no Parque Metropolitano de Pituáçu – PMP, Salvador, Bahia. **Biota Neotropica**, Vol. 5 (número 1a). 2005. ISSN 1676-0603.
- ALZIRA, C.; FERREIRA, E.; GONZAGA S.; CARACIOLO, R. Diversidade e estrutura genética em populações de *Caesalpinia echinata* (Lam.) na Estação Ecológica do Tapacurá, PE. **Scientia Forestalis**. n. 70, p. 77-83, abril 2006.

- ÁVILA, M.; MARCHINI, L. Análise faunística de himenópteros visitantes florais em fragmento de Cerradão em Itirapina, SP. **Ciência Florestal**, Santa Maria, v. 18, n. 2, p. 271-279, abr.-junho, 2008. ISSN 0103-9954
- AVISE, J.C. Molecular markers, natural history e evolution. New York: Chapman e Hall, Inc. 511p., 1994.
- BARRET, S. C. H.; KOHN, J. R. Genetic and evolutionary consequences of small population size in plants: implications for conservation. In: FALK, D. A.; OLSINGER, K. E. (Ed.). **Genetic and conservation of rare plants**. New York: Oxford University Press, 1991. p.3-30.
- BARKER, J.S.F.; TAN, S.G.; MOORE, S.S.; MUKHERJEE, T.K.; MATHESON, J-L.; SELVARAJ, O.S. Genetic variation within and relationship among populations of Asian goats (*Capra hircus*). **Journal of Animal Breeding Genetics**, v.118, p.213-233. 2001.
- BAWA, K.S. Mating systems, genetic differentiation and speciation in tropical rain forest plants. **Biotropica** 24:250-255. 1992.
- BRONDANI, R.P.V.; ZUCCHI, M.I.; BRONDANI, C.; RANGEL, P.H.N.; BORBA, T.C.O.; RANGEL, P.N.; MAGALHÃES, M.R.; VENCOSKY, R. Genetic structure of wild rice *Oryza glumaepatula* populations in three Brazilian biomes using microsatellite markers. **Genética** 125:115-123. 2005
- BROWER, E.J.; H.J. ZAR. Field and laboratory methods for general ecology. 2ª ed. Iowa, Wm. C. **Brown (Publishers)**, 226p. 1984.
- BROWN, A. Measuring genetic variability in plant populations. In: Isozymes in plant genetics and breeding. Part A. Tanksley S. e T. Orton (Eds). **Elsevier**. p. 219-287. 1983.

- CAMPOS, M. J. O. Estudo das interações entre a comunidade de Apoidea, na procura de recursos alimentares, e a vegetação de cerrado da Reserva de Corumbataí, SP. 1989. 123 f. Tese (Doutorado em Ecologia) - Universidade Federal de São Carlos, São Carlos, 1989.
- CARLINI, G.; VENCovsky, R.; COELHO, A.S.G. Variance additivity of genetic populational parameter estimates obtained through bootstrapping. **Scientia Agricola**. v. 60, n. 1, p. 97-103, 2003.
- CHARLESWORTH, D. & CHARLESWORTH, B. Inbreeding depression and its evolutionary consequences. **Annual Review of Ecology and Systematics** 18: 237-268. 1987.
- CHESSER, R. K. Influence of gene flow and breeding tactics on gene diversity within populations. **Genetics** 127: 372-388. 1991.
- COCKERHAM, C.C. Analysis of gene frequencies. **Genetics**. n.74, p. 679-700, 1973.
- CORNUET, J.M. & LUIKART, G. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. **Genetics**, 144(4): 2001-2014. 1996.
- COLE, T.C. Genetic Variation in rare and common plants. *Annu. Rev. Evol. Syst.* 34: 213-237. 2003.
- CRUZ, C.D. Programa Genes: análise multivariada e simulação. Viçosa: UFV, 2006a. 175p.
- CRUZ, COSME DAMIÃO, Programa GENES: diversidade genética. Viçosa, Ed. UFV, 2008.

- EGUIARTE, L.E.; PEREZ NASSER, L.; PIÑERO, D. Genetic structure, outcrossing rate and heterosis in *Astrocaryum mexicanum* (tropical palm): implications for evolution and conservation. *Heredity* 69:217-228. 1992.
- ELLSTRAND, N. C.; ELAM, D. R. Population genetic consequences of small population size: implications for plant conservation. *Annual Review of Ecology and Systematics*, Palo Alto, v. 24, p. 217-242, 1993.
- EPPERSON, B.K. Spatial patterns of genetic variation within plant populations. In: BROWN, A.H.D.; CLEGG, M.T.; KAHLER, A.L.; WEIR, B.S., ed. *Plant population genetics, breeding and genetic resources*. Sunderland: Sinauer, 1990. p.229-253
- EXCOFFIER, L., SMOUSE, P.E., and QUATTRO, J.M. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. **Genetics** 131: 479-491, 1992.
- FALCONER, D. S; MACKAY, T. F. *Introduction to quantitative genetics*. 4 ed. Harlow: **Longman Scientific**, 464 p., 1997.
- FARINA, L.; YOSHIO, P.; MAGNO, A. Estrutura genética em populações fragmentadas de *Chorisia speciosa* St. Hil (Bombacaceae). **Scientia Forestalis**. n. 65, p. 70-79, jun. 2004.
- FEDOROV, A. The structure of the tropical rain forest and speciation in the humid tropics. **Journal of Ecology**. 54:1-11. 1966.
- FERREIRA, D. F. *Estatística Básica*. Lavras: UFLA, 664 p., 2005.
- FRANKHAM, R.; BALLOUD, J.; BRISCOE, D. *Fundamentos de genética da conservação*. Ribeirão Preto, SP: SBG (Sociedade Brasileira de Genética), 280 p.: il. Título em inglês: *A primer of conservation genetics*. Traduzido para o português por Mercival Roberto e Izeni Pires Farias. ISBN: 978-85-89265-08-9, 2008.

FRANKHAM, R. Relationship of genetic variation to populations size in wildlife. **Conservation Biology** 10: 1500-1508, 1996.

FUTUYMA, D. **Biologia evolutiva**. Ribeirao Preto: FUNPEC. 2003.

GUSSON, E.; SEBBENN, M. A.; KAGEYAMA, P. Diversidade e estrutura genética espacial em duas populações de *Eschweilera ovata*. **Scientia Forestalis**. n. 67, p.123-135, abr. 2005.

HAAG, T.; SANTOS, A.; SANA, D.; CULLEN, L.; MORATO, R.; CRAWSHAW, P.; SALZANO, F.; EIZIRIK, E. Variabilidade genética e diferenciação entre populações de onça-pintada (*Panthera onca*) em fragmentos de Mata Atlântica de Interior. Resumos do 53º Congresso Brasileiro de Genética, Brasil. 2007. ISBN 978-85-89109-06-2.

HAMRICK, J.L.; GODT, M.J.W. Alloenzyme diversity in plant species. In: BROWN, A.H.D.; CLEGG, M.T.; KAHLER, A.L. (Ed). Plant population genetics, breeding and genetic resources. Sunderland: Sinauer Associates, 1990. p.145-162.

HAMRICK, J.L.; MURAWSKI, E.D.A. Levels of allozyme diversity in populations of uncommon neotropical tree species. **Journal of tropical ecology**, v.7, p.395-399, 1991.

HARTL, D.L.; CLARK, A.G. Principles of population genetics. Sunderland: **Sinauer Associates**, 3ª ed. 542p. 1997.

HEDRICK, P. W. A new approach to measuring genetic similarity. **Evolution**. v. 25, p. 276-280, 1971.

HELDRIK, P. W.; MILLER, P. S. Conservation genetics: techniques and fundamentals. **Ecology Applied**, v. 2, p. 30-46, 1992.

- HOLCOMB, J.; TOLBERT, D. M.; JAIN, K. A diversity analysis of genetic resources in rice. **Euphytica**. v. 26, p. 441-449, 1977.
- JAMHOUR, J.; LAROCA, S., A community of wild bees (Hym., Apoidea) of Pato Branco (PR-Brazil): diversity, phenology, floral resources, and biogeographical aspects. In: **Acta Biol. Par.**, Curitiba, 33 (1, 2, 3, 4): 27-119. 2004.
- KAGEYAMA, P. Diversidade genética em espécies tropicais de diferentes estágios sucessionais por marcadores genéticos. **Scientia Forestalis**. v. 64, p. 93-107, 2003.
- KARASAWA, M.A.G.; VENCOVSKY, R.; SILVA, C.M.; ZUCCHI, M.I.; OLIVEIRA, G.C.X.; VEASEY, E.A. 2007. Genetic structure of Brazilian *Oryza glumaepatula* Steud. (Poaceae) populations studied with microsatellite markers. *Gen Mol Biol* 30(2):400-410.
- KERR, W. E.; WRIGHT, S. Experimental studies of the distribution of gene frequencies in very small populations of *Drosophila melanogaster*: III. Aristapedia and spineless. **Evolution**, v.8, p. 293-302, 1954.
- LEE, S.L.; NG, K.K.S.; NORWATI, A.; SALWANA, M.H.S.; LEE, C.T.; NORWATI, M. Population genetics of *Intsia palembanica* (Leguminosae) and genetic conservation of Virgin Jungle Reserves in Peninsular Malaysia. *American Journal of botany*, St. Louis, v.89, p.447-459, 2002.
- LEWIS, P. O.; ZAYKIN, D. GDA - Genetic Data Analysis: version 1.1 for Windows 95/NT. 2002. Disponível em: <http://hydrodictyon.eeb.uconn.edu/people/plewis/software.php>.
- LIU, K.; MUSE, S. V. PowerMarker: An integrated analysis environment for genetic marker analysis. **Bioinformatics**. v. 21, p. 2128-2129, 2005.
- LOVELESS, M. D. & J. L. HAMRICK. Distribución de la variación genética en especies arbóreas tropicales. *Revista de Biología Tropical*. 35 (Supl. 1):165-175. 1987.

- LUIKART, G.; BIJU-DUVAL, M-P.; ERTUGRUL, O; ZAGDSUREN, Y.; MAUDET, C.; TABERLET, P. Power of 22 microsatellite markers in fluorescent multiplexes for parentage testing in goats (*Capra hircus*). **Animal Genetics**, v.30, p.431-438, 1999.
- LYNCH, M.; MILLIGAN, B.G. Analysis of population genetic structure with RAPD markers. **Molecular Ecology**, v.3, p. 91-99, 1994.
- MADSEN, T.; R. SHINE, M. OLSSON & H. WITZELL. Restoration of an inbred adder population. **Nature** 402, 34-35. 1999.
- MANTOVANI, W. Methods for Assessment of Terrestrial Phanerogams Biodiversity. In: BICUDO, C. E. M.; MENEZES, N. A. (Eds.) Biodiversity in Brazil: a first approach. São Paulo: CNPq, p.119-44. 1996.
- MARTINS, P. S. Estrutura populacional, fluxo gênico e conservação *in situ*. IPEF, Piracicaba, v. 35, p. 71-78, abr. 1987.
- MORAES, P.L.R. & CARVALHO, R. Taxas de cruzamento em uma população natural de *Cryptocarya moschata* Nees (Lauraceae). **Biota Neotropica** 2: 10p. (<http://www.biotaneotropica.cria.org.br>). 2003.
- MOREIRA, R. Avaliação dos efeitos da deriva genética em duas gerações de amostragens nas populações de milho BR-105 e BR-106/ Rosângela Maria Pinto, Piracicaba, 1999.
- NEI, M. Analysis of gene diversity in subdivided populations. Proceedings of the National Academy of Sciences of the United States of America. Washington, v. 70, p. 3321-3323, 1973.
- NEI, M. Estimation of average heterozygosity and genetic distance from a small number of individuals. **Genetics**. Pittsburgh, v. 89, p. 583-590, 1978.

- NEI, M. *F*-statistics and analysis of gene diversity in subdivided populations. *Ann. Hum. Genet.* 41:225-233. 1977.
- NEI, M. *Molecular evolutionary genetics*. Columbia University Press, New York. 1987.
- NEIGEL, J.E. Is  $F_{ST}$  obsolete? ***Conservation Genetics***, 3: 167-173, 2002.
- NYBOM, H. Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. ***Molecular Ecology***. v.13, p.1143-1155. 2004.
- OLIVEIRA, C.A.M. de; SILVA, E.F. da; MOLICA, S.G.; FERREIRA, R.L.C.; LIRA, D.A.S. de e JÚNIOR, J.A.B.B. de. Diversity and genetic structure in populations of *Caesalpinia echinata* (Lam.) in the Ecological Station of Tapacurá, Pernambuco, Brazil. ***Scientia Forestalis***, 70: 77-83, 2006.
- PADILHA, M.; DAMASCENO, K. Diferenciação Genética Entre Procedências De Açazeiro Por Marcadores RAPD e SSR. ***Revista Brasileira Fruticultura***, Jaboticabal - SP, v. 30, n. 2, p. 438-443, Junho 2008.
- PICOLI, E. A. T. Influence of RAPD number of markers and sample size on Eucalyptus genetic distance and diversity. ***Crop Breeding and Applied Biotechnology***. v.4, p. 384-390, 2004.
- PIRY, S.; LUIKART, G.; CORNUET, J. A computer program for detecting recent reductions in the effective population size using allele frequency data. *J. Hered.* 90:502-503. 1999.
- RAYMOND, M.; ROUSSET, F. Population genetics software for exact tests and ecumenicism. ***Journal of Heredity***. v.86, p. 248-249, 1995.
- RIDLEY, MARK. *Evolução*; tradução Henrique Ferreira, Luciane Passaglia, Rivo Fisher. 3 ed. Porto Alegre: Artmed, 752 p, 2006.

- ROBINSON, I.P. Aloenzimas na genética de populações de plantas. In Eletroforese de isoenzimas e proteínas afins: fundamentos e aplicações em plantas e microrganismos (A.C. Alfenas, ed.). Editora UFV, Viçosa, p.329-380, 1998.
- ROUSSET, F.; Raymond, M. Testing heterozygote excess and deficiency. *Genetics*. v. 140, p. 1413-1419, 1995.
- RUTTEN, M.J.M.; KOMEN, H.; DEERENBERG, R.M.; SIWEK, M.; BOVENHUIS, H. Genetic characterization of four strains of Nile tilapia (*Oreochromis niloticus* L.) using microsatellite markers. **Animal Genetics**, v.35, p.93-97, 2004.
- SHAH, A.; LI, D.; GAO, L.; LI, H. e MÖLLER, M. Genetic diversity within and among populations of the endangered species *Taxus fuana* (Taxaceae) from Pakistan and implications for its conservation. **Biochemical Systematics and Ecology**, 36: 183-193, 2008.
- TAKAHATA, N.; PALUMBI, S. Extranuclear differentiation and gene flow in the finite island model. *Genetics* 109: 441-457. 1985
- TEIXEIRA, M.; KAGEYAMA, P.; SEBBENN, A. Mating system in small fragmented populations and isolated trees of *Hymenaea stigonocarpa*. **Scientia Forestalis**. n. 74, p. 75-86, junho 2007.
- TORGGLER, M.G.F.; CONTEL, E.P.B.; TORGGLER, S.P. Isoenzimas: variabilidade genética em plantas. Ribeirão Preto: SBG, 1995, 186p.
- TOREZAN, J. M. D. et al., Genetic variability of pré and post-fragmentation cohorts of *Aspidosperma polyneuron* Muell. Arg. (Apocynaceae). *Brazilian Archives of Biology and Technology*, v.48, n.2, p.171-180, 2005.
- YEEH, Y.; KANG, S.S.; CHUNG, M.G. Evaluation of the natural monument populations of *Camellia japonica* (Thearaceae) in Korea based on allozyme

- studies. Botanical Bulletin of Academia Sinica, Shanghai, v.37, n.1, p.141-146, 1996.
- YEH, F. C. POPGENE, the user-friendly shareware for population genetic analysis. Molecular Biology and Biotechnology Centre, University of Alberta, Canada. 1997.
- WADT, L.H.O. de; KAGEYAMA, P. Y. Estrutura genética e sistema de acasalamento de *Piper hispidinervum*. **Pesquisa Agropecuária Brasileira**, n.2, 39: 151-157, 2004.
- WAHLUND, S. Zusammensetzung von Populationen und korrelationenscheinungen von Standpunkt der Vererbungslehre aus betrachtet. Hereditas 11: 65-106, 1928, cf. Li, C.C. Population genetics. Univ. Chicago Press, Chicago, 7a. reimpressão, 1972.
- WALLACE, L.E. Examining the effects of fragmentation on genetic variation in *Platanthera leucophaea* (Orchidaceae): Inferences from allozyme and random amplified polymorphic DNA markers. Plant Species Biology, 17: 37-49, 2002.
- WEIR, B. S. Genetic Data Analysis II. Sinauer Associates, Inc., MA, USA, 1996.
- WEIR, B. S.; COCKERHAM, C. C; Estimating F-statistics for the analysis of population structure. **Evolution**. v. 38, p. 1358-1370, 1984.
- WRIGHT, S. Evolution and genetics of populations. University of Chicago Press, Chicago, 511 p. 1978.
- WRIGHT, S. The genetical structure of populations. **Annual of Eugenics**. v. 15, p. 313-354, 1951.
- WRIGHT, S. The roles of mutation, inbreeding, crossbreeding, and selection in evolution. Proceedings of the 6<sup>th</sup> International Congress of Genetics 1:356-366. 1932.

WRIGHT, S.; KERR, W. E. Experimental studies of the distribution of gene frequencies in very small populations of *Drosophila melanogaster*: II. Bar. **Evolution**, v.8, p. 225-240, 1954.

ZUCCHI, M. I.; BRONDANI, R. V.; PINHEIRO, J. B.; COELHO, A. S. G.; CHAVES, L. J.; VENCOVSKY, R. Genetic structure and gene flow in *Eugenia dysenterica* DC in the brazilian cerrado utilizing SSR markers. **Genetics and Molecular Biology**, Ribeirão preto, v. 26, n. 4, p. 449-458, 2003.