

# RELATÓRIO TÉCNICO

Análise Genética das Populações de *Myrciaria dubia* (H.B.K.)  
Mc Vaugh (camu-camu) e *Ceiba pentandra* L. (samaúma)  
ocorrentes na área de Influencia da UHE Santo Antônio.

---

Etapa II: Análise Genética das Populações de *Ceiba  
pentandra* L. (samaúma).

**Centro de Pesquisa em Medicina Tropical - CEPEM  
Laboratório de Genética Molecular**

**Fevereiro/2015**

# RELATÓRIO TÉCNICO

Análise Genética das Populações de *Myrciaria dubia* (H.B.K.) Mc Vaugh (camu-camu) e *Ceiba pentandra* L. (samaúma) ocorrentes na área de Influência da UHE Santo Antônio.

---

Etapa II: Análise Genética das Populações de *Ceiba pentandra* L. (samaúma).

Equipe Técnica

Coordenadora

**Francisca de J. Holanda**

Bióloga, doutora em Genética Vegetal, professora da FIMCA

Apoio

**Elieth Afonso Mesquita**

Bióloga, doutora em Biologia Experimental, professora da UNIR

**Cooperadora** – Lab. de Bioquímica, Quimioterapia e Cultura de Célula e CIBEBI/UNIR.

**MSc. Rosimeire Cristina Dalla Martha**

Farmacêutica, doutoranda em Biologia Molecular

**Coordenadora** – Lab. de Genética de Molecular/CEPEM/RO.

**Centro de Pesquisa em Medicina Tropical - CEPEM**  
**Laboratório de Genética Molecular**

**Fevereiro/2015**

## 1. INTRODUÇÃO

Este Relatório tem por finalidade apresentar as análises de diversidade e divergência genética da população de *Ceiba pentandra* (samaúma), a partir dos parâmetros da estatística  $F$  de Wright (1951), que corresponde ao índice de fixação de uma população e das estatísticas descritivas, sendo alguns dos descritores que auxiliam na caracterização genética de uma população de acordo com os seguintes parâmetros: heterozigosidade ( $H_o$  e  $H_e$ ), coeficiente de fixação/endogamia( $f$ ), riqueza alélica, entre outras.

Finalizando assim, a proposta do projeto intitulado Análise Genética das Populações de *Myrciaria dubia* (camu-camu) e *Ceiba pentandra* (samaúma) ocorrentes na Área de Influência da UHE Santo Antônio.

## 2. OBJETIVOS

- a) Analisar a diversidade genética da população de plantas de samaúma (*C. pentandra*), coletadas às margens do rio Madeira, com marcadores microssatélites.
- b) Verificar a divergência genética da população de *C. pentandra* da área de influência da UHE Santo Antônio.

## 3. MATERIAL E MÉTODO

### 3.1 Caracterização das populações de *C. pentandra*

Este estudo foi realizado com duas populações de *C. pentandra* (samaúma), as amostras biológicas utilizadas para a extração do gDNA foram folhas frescas, coletadas ao longo das margens do no rio Madeira, identificadas, georreferenciadas e armazenadas em sacos plásticos, acondicionadas em caixa de isopor com gelo até chegar ao laboratório. A população identificada como vila de Teotônio corresponde à região a montante da barragem da Usina Hidrelétrica de Santo Antônio (UHSA), a população vila de Calama corresponde à extensão do rio a jusante da barragem. As coletas foram realizadas pela bióloga da Santo Antônio Energia.

### 3.2 Genotipagem de *C. pentandra* com marcadores SSR (Simple Sequence Repeats)

As análises genéticas da população da vila de Teotônio foram realizadas com 26 indivíduos e da vila Calama foram 9, utilizando cinco locos de marcadores microssatélites (SSR), desenvolvidos por Brondani, *et al.* (2003). As reações de genotipagem em analisador de DNA, foi realizada a partir dos seguintes parâmetros: foi ligado na extremidade 5' de cada *primers forward*, dos locos estudados, a seqüência M13 (5' TGAAAACGACGGCCAGT 3') e, uma outra seqüência M13 correspondente, marcada com fluorocromo específico (6-FAM ou HEX), seguindo o protocolo segundo Brondani *et al.* (2003), com modificações. Para a amplificação dos fragmentos marcados foi realizada uma reação de PCR nas seguintes condições: tampão 10X (TrisHCl 100 mM, KCl 500 mM, pH 8.4), dNTP a 2,0 µM, MgCl<sub>2</sub> a 25 mM, *primer Forward* M13 a 5 µM, *primer Forward* M13 marcado com fluorocromo Hex a 5 µM, *primer Reverse* a 5 µM, 1,5 U *Taq* DNA Polimerase (Invitrogen 5 U/µL) e DNA a 15 ng/µL para um volume final de 10 µL. O programa de ciclagem correspondeu a um ciclo de 94°C por quatro minutos, seguido de 25 ciclos de 94°C por 1min, temperatura de anelamento de cada *primer* T<sub>a</sub> por 1 min, e 72°C por min, em seguida um ciclo de 72°C por 7 minutos, seguido de 20 ciclos de 94°C por 10s, 50°C por 10s, e 72°C por 30s e uma extensão a 72°C por 10 min. As amplificações foram realizadas em termociclador *Verite Thermal Cycler* (AB Applied Biosystems). A genotipagem dos produtos de PCR gerados foi realizada em analisador de DNA automático. A análise dos tamanhos dos fragmentos foi realizada utilizando-se o programa *Fragment Profiler* (GE Healthcare, England), considerando o marcador de peso molecular ET-400-ROX.

Algumas amostras foram genotipadas em gel de poliacrilamida a 10%, corados com nitrato de prata, para análise dos tamanhos dos fragmentos (bandas) foram utilizados marcador de peso molecular de 50 e 100pb (Invitrogen), além dos padrões de marcadores internos.

### 3.2 Análise genética

Com os genótipos obtidos dos cinco locos analisados foi construída uma matriz de genótipos, de cada população de *C. pentandra* e seus respectivos locos, para calcular os índices genéticos das duas populações em estudo. Os parâmetros utilizados para estimar a variabilidade genética são geralmente descritas em termos de frequência alélica, heterozigosidade observada entre populações, além da proporção do número de alelos polimórficos.

A estatística  $F$  de Wright (1965) estima o coeficiente de endogamia e os parâmetros relacionados a pares de gametas, em uma população natural hierarquicamente subdividida, tendo os seguintes parâmetros:

- $F_{IT}$  - expressa a correlação entre os gametas que se unem para produzir os indivíduos em relação aos gametas da população total;
- $F_{IS}$  - expressa a média das correlações, sendo cada uma delas proveniente dos gametas que se unem em cada subpopulação em relação aos gametas desta subpopulação e;
- $F_{ST}$  - expressa a correlação entre os gametas ao acaso dentro da subpopulação em relação aos gametas da população total.

Obedecendo a seguinte equação:  $1 - F_{IT} = (1 - F_{IS})(1 - F_{ST})$

As estimativas das heterozigosidades são as bases para determinar a diversidade genética de uma população e resume-se no índice geral da variação por loco em uma população. Os parâmetros como: número de alelos por locos, porcentagem de locos polimórficos, heterozigosidade observada ( $H_o$ ) e esperada ( $H_e$ ), sob equilíbrio de Hardy-Weinberg (EHW), e o índice de fixação ( $f$ ) são estimativas que caracterizam e comparam os níveis de diversidade genética. Além disso, o índice de fixação ( $f$ ) está sendo utilizado amplamente como uma medida de desvios de panmixia do equilíbrio de Hardy-Weinberg (ZUCCHI, 2002; RODRIGUES, 2007).

As frequências alélicas foram calculadas para cada loco, com auxílio do programa GenAlEx v.6.5 (Peakall & Smouse, 2012). As estimativas de diversidade genética de número total ( $A$ ) e médio de alelos ( $\bar{A}$ ), riqueza alélica ( $A_r$ ), as heterozigosidades observada ( $H_o$ ) e esperada ( $H_e$ ), além do coeficiente de endogamia ( $f$ ), foram estimados para cada loco e para cada

população, na plataforma R com o pacote *diveRsity* (Keenan *et al.*, 2013). A significância dos valores de  $f$  foi acessada através de 1000 réplicas de *bootstrap*. O número de alelos privados ( $A_p$  - definido como um alelo encontrado em uma única população) foi calculado na plataforma R com o pacote *Poppr* (Kamvar *et al.*, 2014). Os alelos foram classificados como comum (frequência > 0,2), intermediário (> 0,05 e < 0,2) ou raro (< 0,05), similar ao esquema de Marshall & Brown (1975).

Para verificar estrutura genética entre as populações, foram estimados o  $F_{ST}$  de Weir & Cockerham (1984), bem como os análogos das estatísticas  $F$  (Wright, 1951):  $G_{ST}$  (Nei, 1973), na plataforma R com auxílio do pacote *diveRsity* (Keenan *et al.*, 2013). A significância dos estimadores de divergência genética foi acessada por meio de 1000 réplicas de *bootstrap*. Outra avaliação foi realizada por meio de análises Bayesianas com o programa *Structure* (Pritchard *et al.*, 2000) para determinar o número  $K$  de agrupamentos que melhor explica a matriz de genótipos obtida no estudo. Para cada número  $K$  de agrupamentos (com  $K$  variando de 1 a 10) foram realizadas dez simulações independentes sob o modelo de mistura com frequências alélicas correlacionadas (*Admixture model, correlated allele frequencies*) com 500.000 passos da cadeia de Markov – Monte Carlo (MCMC), após um período inicial de *burnin* de 100.000 passos. O número mais provável de  $K$  agrupamentos foi verificado de acordo com o método  $\Delta K$  de Evanno (2005), com auxílio do aplicativo on-line *Structure Harvester* (Earl & vonHoldt, 2012). A padronização dos agrupamentos foram obtidos pelas simulações independentes de cada  $K$  avaliado, utilizou-se o programa *Clumpp* (Jakobsson & Rosenberg, 2007).

Para verificar como a estrutura genética está organizada entre e dentro das populações, uma análise hierárquica da variância molecular (AMOVA, Michalakis & Excoffier, 1996) foi realizada com o programa *Arlequin v.3.5* (Excoffier & Lisher, 2010).

Para se avaliar a dispersão dos indivíduos ao longo da variação genética detectada foi efetuada uma análise de coordenadas principais (PCoA), com auxílio do programa *GenAlEx* (Peakall & Smouse, 2012).

A distância genética entre indivíduos foi avaliada por meio da construção de dendrogramas, com o método de neighbor-joining (Saitou & Nei, 1987), fundamentado nas distâncias genéticas padronizadas de Nei (1978). NJ

(Neighbour Joining ou vizinho mais próximo), esse método é baseado nas informações dos vizinhos mais próximos, o qual apresenta diferentes taxas de evolução, as árvores são construídas com sucessivos agrupamentos lineares considerando as longitudes dos ramos dos agrupamentos (Goldstein & Schlotterer, 1999). Outra análise foi realizada com base nas distâncias euclidianas de Rogers (1972). Os dendrogramas, dessas análises, foram plotados na plataforma R com o pacote Poppr (Kamvar *et al.*, 2014). A consistência dos ramos foi avaliada com 1000 réplicas de *bootstrap*.

#### 4. RESULTADOS E DISCUSSÃO

##### 4.1 Estimativas de diversidade genética

O resultado obtido com os cinco locos analisados, nas duas populações, perfaz um total de 83 alelos de 35 árvores (ver tabela 1), sendo que na população de vila Teotônio (vT) foram genotipados 51 alelos e 32 encontrados na vila Calama (vC), como mostra a tabela 2. A média de alelos foi de 16,6 por locos, sendo 10,2 e 6,4 para as respectivas populações. Foi observado polimorfismo em todos os locos, que variaram entre 9 a 12, para os locos Cp13 e Cp19, respectivamente (ver tabelas 1 e 2).

Tabela 1 – Índices de diversidade genética e coeficiente de endogamia de cinco locos microsatélites para amostras de sumaúma (*Ceiba pentandra*) provenientes de dois locais de coleta. N – número de indivíduos avaliados; A – número de alelos; Ar – número de alelos privados; Ho – heterozigosidade observada; He – heterozigosidade esperada; f – coeficiente de endogamia. \* significativo com  $p < 0,05$ .

População		Locos				
		Cp 13	Cp15	Cp 18	Cp 19	Cp 20
Vila Teotônio	N	24	26	23	23	25
	A	9	10	11	12	9
	Ar	7,38	6,55	6,96	7,87	6,46
	Ap	2	1	7	6	4
	Ho	0,792	0,385	0,652	0,652	0,440
	He	0,872	0,825	0,862	0,899	0,851
	f	0,092	0,534*	0,243*	0,274*	0,483*
	Vila Calama	N	9	9	7	7
A	7	9	4	6	6	
Ar	5,84	7,02	3,59	4,99	5,28	
Ap	-	-	-	-	1	
Ho	0,667	0,778	0,714	0,714	0,556	
He	0,802	0,846	0,684	0,796	0,802	
f	0,169	0,080	-0,045	0,103	0,308	

Tabela 2 – Índices de diversidade genética e coeficiente de endogamia para duas populações de sumaúma (*Ceiba pentandra*) avaliadas com cinco locos microssatélites. N – número de indivíduos amostrados; A – número total de alelos;  $\bar{A}$  – número médio de alelos por loco;  $A_p$  – número de alelos privados;  $H_o$  – heterozigiosidade observada;  $H_e$  – heterozigiosidade esperada;  $f$  – coeficiente de endogamia. \* significativo com  $p < 0,05$ .

População	N	A	$\bar{A}$	$A_r$	$A_p$	$H_o$	$H_e$	$f$
Vila de Teotônio	26	51	10,2	7,04	20	0,584	0,862	0,325*
Vila Calama	9	32	6,4	5,34	1	0,686	0,786	0,123

Os índices de heterozigiosidade esperada ( $H_e$ ) estimados nos cinco locos, obtidos com marcadores SSR, para a população vila Teotônio, variaram entre 0,825 (loco Cp 15) e 0,899 (loco Cp19) (ver tabela 1). Os valores das heterozigiosidades observadas ( $H_o$ ) foram inferiores a  $H_e$ , em todos os locos, variando entre 0,385 (loco Cp 15) e 0,440 (loco Cp 20), indicando carência de indivíduos heterozigotos (ver tabela 1). As estimativas de  $H_e$  para a fração amostral vila Calama variaram entre 0,684 e 0,846 nos respectivos locos, Cp 18 e Cp 15. Os valores de  $H_o$  foram todos inferiores também.

Os valores de  $H_o$  foram significativamente diferentes do esperado ( $H_e$ ) sob o Equilíbrio de Hardy-Weinberg ( $p < 0,05$ ) para todos os locos. Estes resultados estão de acordo com o estudo de Brondani *et al.* (2003), que detectou altos valores de  $H_e$ , que variaram entre 0,814 a 0,895. Considerando essas estimativas, observou alto índice do coeficiente de endogamia ( $f$ ), para a fração amostral vila Teotônio, sendo de 0,325, significativamente diferente de zero ( $p < 0,05$ ) (ver tabela 2), indicando deficiência de heterozigotos, como demonstrado pelos valores de  $H_o$ .

Em análises de genéticas de populações, o número de heterozigotos é um dado importante para estimar o índice de diversidade genética, sabendo que cada heterozigoto é caracterizado por carregar alelos diferentes, representando assim a diversidade genética presente na população, por isso, as taxas de heterozigiosidades são consideradas medidas apropriadas para quantificar a variação genética. Tanto que se o valor da heterozigiosidade observada for igual à proporção da heterozigiosidade esperada, estes índices



indicam que a população está em Equilíbrio de Hardy-Weinberg (Weir & Basten 1990; Torggler *et al.*, 1995; Weir, 1996).

No presente estudo as distribuições das freqüências dos alelos privados, entre os locos genotipados, variaram de 0,02 a 0,174 (tabela 4). Os locos CP15 e CP19 apresentaram uma distribuição menos uniforme, indicando maior freqüência de uns poucos alelos (Figura 1).

Considerando à presença dos alelos nas duas populações foi observado o número total de 9 alelos raros, 40 intermediários e 3 freqüente (ver tabelas 3 e 4). Na população como maior quantidade de amostras, vila Teotônio, foi estimado as porcentagem de alelos; 19.6 raros, 72.5 intermediários e 7.9 frequentes (ver tabela 3).

A estimativa da freqüência de um alelo particular em uma população, chamada de freqüência gênica ou alélica, é considerada fundamental em estudos evolutivos, pois a mudança genética de uma população pode ser avaliada pela mudança nas freqüências gênicas e genotípicas (Nei, 1987).

São raros estudos publicados sobre a diversidade e estrutura genética de samaúma. O estudo de Brondani *et al.* (2003), analisou 174 indivíduos adulto de samúma (*C. pentandra*) da região Amazônica com oito locos: CP1 (AF503157; GenBank), CP12 (AF503158; GenBank), CP13 (AF503159; GenBank), CP15 (AF503160; GenBank), CP18 (AF503161; GenBank); CP 19 (AF503162; GenBank), CP 20 (AF503163; GenBank) e CP 22 (AF503164; GenBank). O número total de alelos obtidos com estes locos foram de 112, com uma média de 14 alelos por locos. O presente estudo utilizou cinco (ver tabela 1 e 2) dos oito locos acima citados, desenvolvidos por Brondani *et al.* (2003), confirmando assim o alto polimorfismo dos locos, descritos pelos pesquisadores, como pode se verificar na figura 1 as freqüências alélicas obtidas com quatro *primers*, na forma de gráficos, presentes nas duas populações; vila Teotônio e vila Calama.

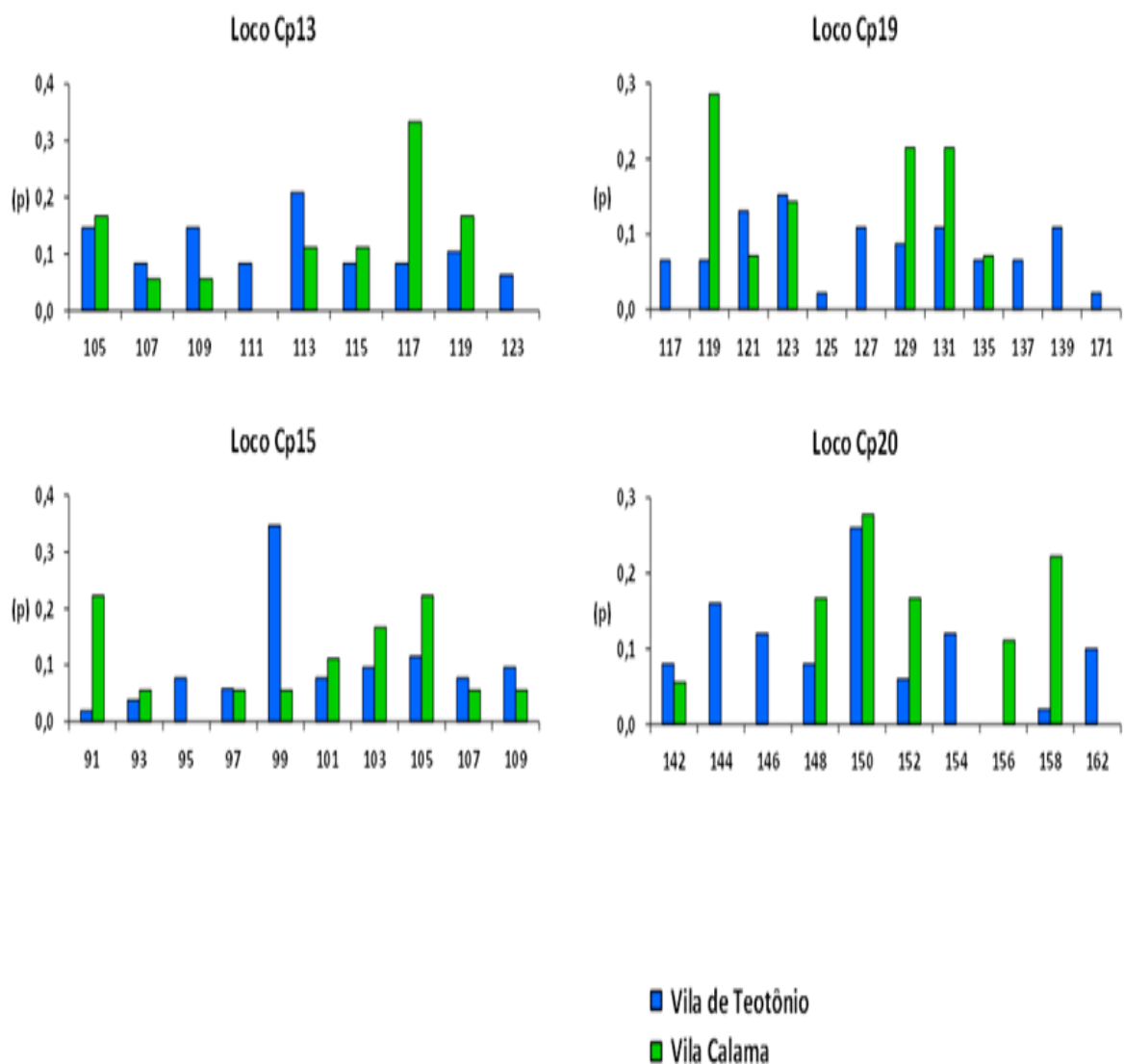


Figura 1 – Histogramas de distribuição das frequências alélicas (p), encontradas em quatro locos microsatélites (SSR) para 35 indivíduos de sumaúma (*Ceiba pentandra*), para as populações de Vila de Teotônio e Vila Calama.

Tabela 3 – Classificação do número de alelos de acordo com sua frequência e distribuição entre as populações de sumaúma (*C. pentandra*), considerando cinco locos SSR.

	Raro	Intermediário	Frequente	Total
<b>Vila de Teotônio</b>	10	37	4	51
<b>Vila Calama</b>	0	21	11	32
<b>Total</b>	9	40	3	52

Tabela 4 – Frequência dos alelos privados encontrados nas populações de sumaúma (*C. pentandra*) em dois locais de coletas.

<b>População</b>	<b>Loco</b>	<b>Alelo</b>	<b>Frequência</b>
<b>Vila Teotônio</b>	Cp13	111	0,083
		123	0,063
	Cp15	95	0,077
	Cp18	117	0,130
		120	0,043
		123	0,174
		125	0,087
		127	0,022
		138	0,022
		147	0,022
	Cp19	117	0,065
		125	0,022
		127	0,109
		137	0,065
		139	0,109
		171	0,022
	Cp20	144	0,160
146		0,120	
154		0,120	
162		0,100	
<b>Vila Calama</b>	Cp20	156	0,111

As estimativas dos alelos privados indicam, de acordo com Cole *et al.* (2007), o grau evolutivo ou de adaptação ao ambiente. Segundo Nei (1987), a estimação da frequência de um alelo particular, em estudos evolutivos de populações, é muito informativo, pois estes valores revelam as mudanças genéticas de uma população.

Portanto, os alelos privados estimados na população em estudo são representativos, considerando o potencial genético, mesmo que um alelo tenha apresentado frequência baixa, deve-se considerar a erosão genética por ação antrópica ou por fatores estocásticos.

As estatísticas descritivas em estudos de populações são relevantes, pois revelam a princípio, o polimorfismo genético de uma população. Parâmetros como: número de alelos por locos ou total, de alelos efetivos, raros, exclusivos, heterozigiosidade observada ( $H_o$ ) e esperada ( $H_e$ ), coeficiente de endogamia( $f$ ), riqueza alélica, entre outras, são alguns dos descritores que auxiliam na caracterização genética de uma população (Sánchez, 2008).

## 4.2 Análises de diferenciação genética

Os coeficientes de divergência genética global estimados para as duas populações de *C. pentandra*, com os cinco locos microssatélite, como  $H_{ST}$  – diversidade gênica total, os valores dessa diversidade variaram entre os locos, o menor de 0,059 do loco Cp 13 e o maior 0,217 para o loco Cp 18 (ver tabela 5). O índice de fixação de Nei ( $G_{ST}$ , 1973) 0,022.

Tabela 5 – Estimativas de diferenciação genética global e para cada loco entre as populações de sumaúma (*Ceiba pentandra*) de Vila de Teotônio e Vila Calama.  $H_{ST}$  – diversidade gênica total;  $D_{ST}$  – diversidade gênica entre subpopulações; Estimadores de divergência genética (índice de fixação) de Nei ( $G_{ST}$ , 1973) e Weir & Cockerham ( $F_{ST}$  (WC), 1984). \* significativo com base em 100 réplicas *bootstrap*.

Locos	$H_T$	$D_{ST}$	$G_{ST}$	$F_{ST}$ (WC)
<b>Cp13</b>	0,059	0,008	0,009	0,011
<b>Cp15</b>	0,164	0,022	0,024	0,033
<b>Cp18</b>	0,217	0,041	0,048	0,077
<b>Cp19</b>	0,067	0,008	0,008	0,003
<b>Cp20</b>	0,129	0,018	0,021	0,023
<b>Global</b>			<b>0,022</b>	<b>0,029</b>

A estimativa de divergência genética, representada pelo coeficiente de  $F_{ST}$  obtido no presente estudo foi de 0,029. Este coeficiente está relacionado com a parte da endogamia total, devido à deriva genética, independente do sistema reprodutivo.

A partir da estimativa  $F_{ST}$  pode-se obter uma estimativa indireta do fluxo gênico, representado pelo número de migrantes ( $Nm$ ), assumindo a estrutura genética populacional sob o modelo de ilhas de Wright (1951) e, a existência de equilíbrio entre migração e deriva genética, a média do número de migrantes ( $Nm$ ) encontrados entre as duas populações foi de 6,183 migrantes por geração, que segundo Govindajaru (1989), indica um alto valor de fluxo gênico. O que pode explicar esse alto valor de  $Nm$ , entre as duas populações, seria distância, que não são grande para os polinizadores, que são os insetos,

aves e morcegos (Ortega Escalona *et al.*, 1991; Joker & Salazar, 2000), da espécie *Phyllostomus hastatus* e *Phyllostomus discolor* que exercem esse papel com maior eficiência na polinização cruzada (Gribel *et al.*, 1999)

O índice de fixação  $F_{IT}$  obtido no presente estudo foi de 0,262 e representa a endogamia total da espécie ou desvio de panmixia da população total (devido ao sistema reprodutivo somado à subdivisão), para o conjunto das populações.

O índice de fixação intrapopulacional  $F_{IS}$  foi 0,228 o que significa que 22,8% dos cruzamentos ocorridos dentro das populações são endogâmicos. Este coeficiente de endogamia intrapopulacional refere-se aos desvios de panmixia, devido apenas ao sistema reprodutivo da espécie, o que reforça a informação sobre o sistema reprodutivo da samaúma, caracterizado por ser misto, preferencialmente alógama (Murawski & Hamrick, 1991; Murawski & Hamrick, 1992).

O coeficiente de endogamia é um parâmetro que afeta diretamente a diversidade genética de uma população, tendo sido introduzido por Wright (1951), e resume de forma conveniente, a estrutura de uma população (Zucchi, 2002).

Para verificar a variação genética entre e dentro das duas populações foi realizados Análise da Variação Molecular (AMOVA). A estimativa da proporção total de variação genética foi maior dentro das populações, com valor de 83,4%, o valor da variação entre as populações foi de 4% (Tabela 6). O índice de  $F_{ST}$  de 0,045 não foi significativo ( $p < 0,05$ ), indicando que as populações das duas localidades não são estruturadas, ou seja, não são geneticamente diferentes.

Tabela 6 – Análise de variância molecular (AMOVA) entre as populações de Vila de Teotônio e Vila Calama. Valor associado de  $F_{ST}$  não significativo com  $p < 0,05$ .

Fonte de Variação	G.L.	Soma de Quadrados	Componentes da Variação	Porcentagem da variação	Estatística F
Entre populações	1	1,9	0,04	<b>4,56</b>	$F_{ST} = 0,045$
Dentro de populações	68	56,75	0,834	<b>95,44</b>	
<b>Total</b>	69	58,65	0,874		

Os resultados das análises da estrutura genética das populações, com o *Structure*, não houve separação dos indivíduos em grupos, pois todos apresentavam o perfil parecido, como se pode ver na figura 2. Os outros Ks também mostram que não existe mesmo estrutura genética, pois em todos os Ks, todos os indivíduos têm perfil parecido (figura 2). Apesar dos valores de  $\Delta K$  parecerem indicar a existência de cinco agrupamentos, não existe diferenciação genética entre os locais de coleta (ver figura 2).

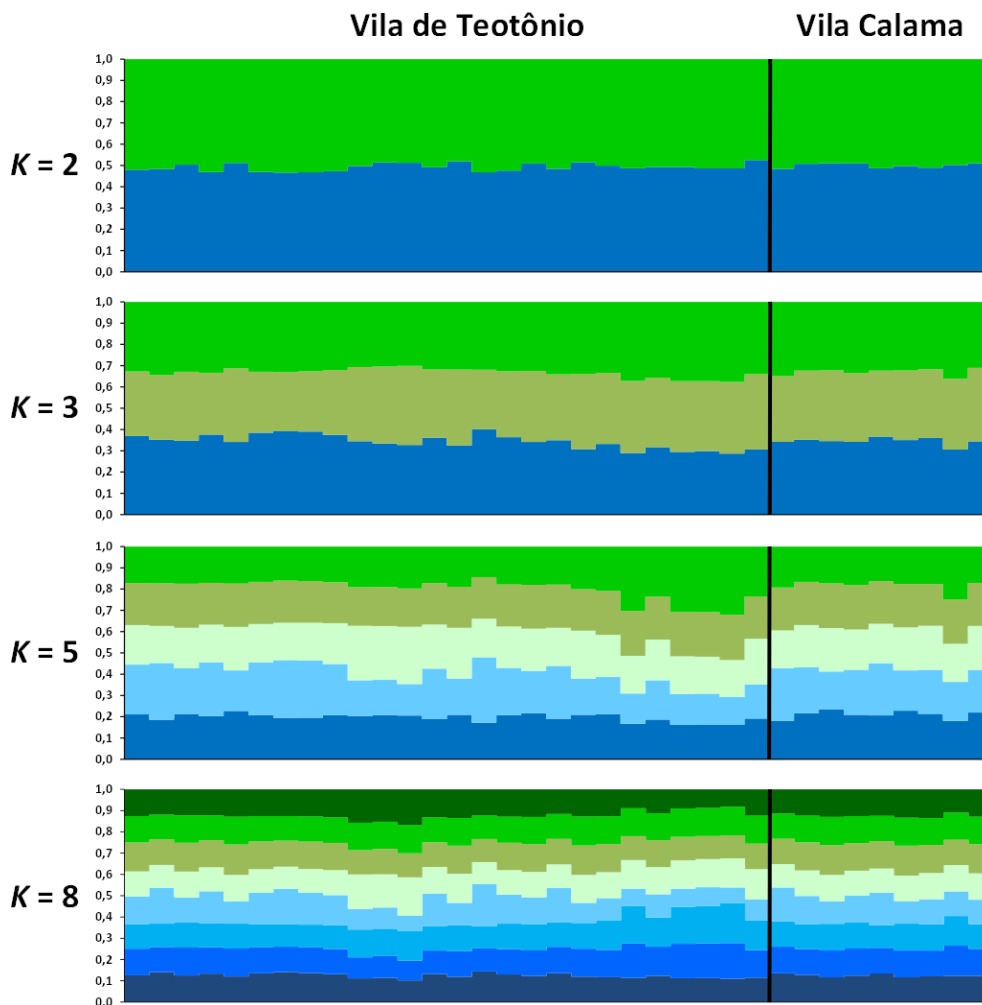


Figura 2 – Análise realizada com o programa Structure para 35 indivíduos de sumaúma (*Ceiba pentandra*), representada em gráficos de barras. A atribuição de indivíduos aos diferentes agrupamentos é representada para  $K = 2$ ,  $K = 3$ ,  $K = 5$  e  $K = 8$ . A linha preta na vertical separa os indivíduos de acordo com local de coleta especificado acima dos gráficos.

O conhecimento da variação genética entre e dentro de populações é a base para estabelecer estratégias eficazes para a conservação. Pois, mesmo a

conservação e o manejo da biodiversidade em áreas protegidas, são desafios complexos que requerem apropriação das informações básicas, tais como: a distribuição e abundância de espécies, as interações mutualísticas, sobre a biologia reprodutiva e a estrutura genética de suas populações (Frankham *et al.*, 2008).

#### 4.3 Relação das Distâncias Genéticas

Para avaliar graficamente as distâncias genéticas entre os indivíduos de *C. pentandra*, genotipados com cinco locos de SSR das duas localidades, foi realizada uma Análise de Coordenadas Principais (fig. 3). O gráfico mostra que a dispersão genética entre as amostras das duas localidades, não formaram grupos de acordo com sua população de origem, apresentam-se dispersos ao longo do gráfico (fig. 3). Estes resultados indicam que existe pouca divergência genética entre as populações amostradas.

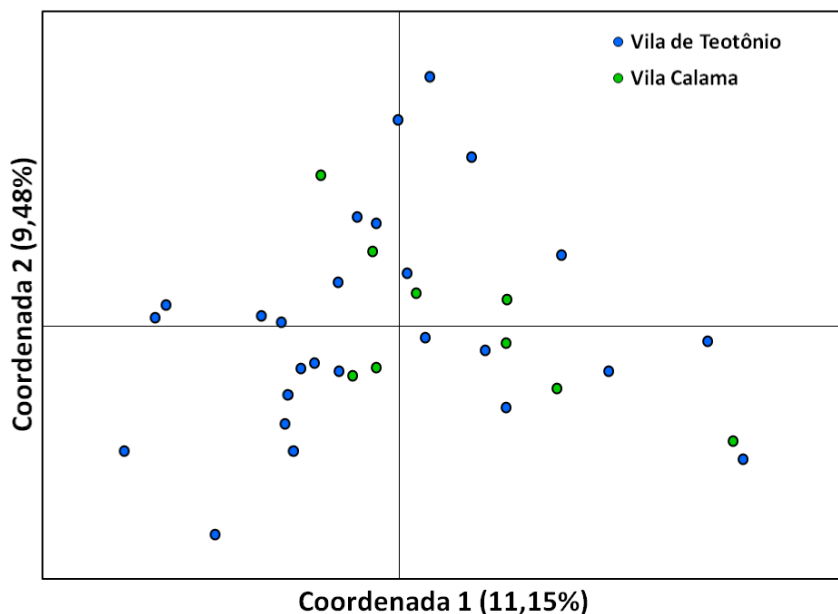


Figura 3 – Análise de Coordenadas Principais ilustrando a dispersão da distância genética de 35 indivíduos de sumaúma (*C. pentandra*) avaliados com cinco locos microssatélites.

Não existe formação de grupos de acordo com o local de coletas dos indivíduos, havendo sobreposição entre os indivíduos amostrados em diferentes locais.

As estimativas das distâncias genéticas foram realizadas segundo Nei (1978), entre as duas populações, com 1.000 reamostragem. O dendrograma plotado, entre as populações (fig.5), adotou o método neighbor-joining.

A distância genética estimada entre as duas populações, vila Teotônio e vila Calama, foi de 0,045 de dissimilaridade e 0,95 de identidade, sugerindo que essas populações não estão geneticamente estruturadas. Reforçando que os indivíduos amostrados nas duas populações são oriundos de locais diferentes do rio Madeira,

Os agrupamentos observados no dendrograma apresentam genótipos misturados; não formam *cluster* de acordo com o local de coleta, o que permite sugerir tratar-se de uma população homogenia.

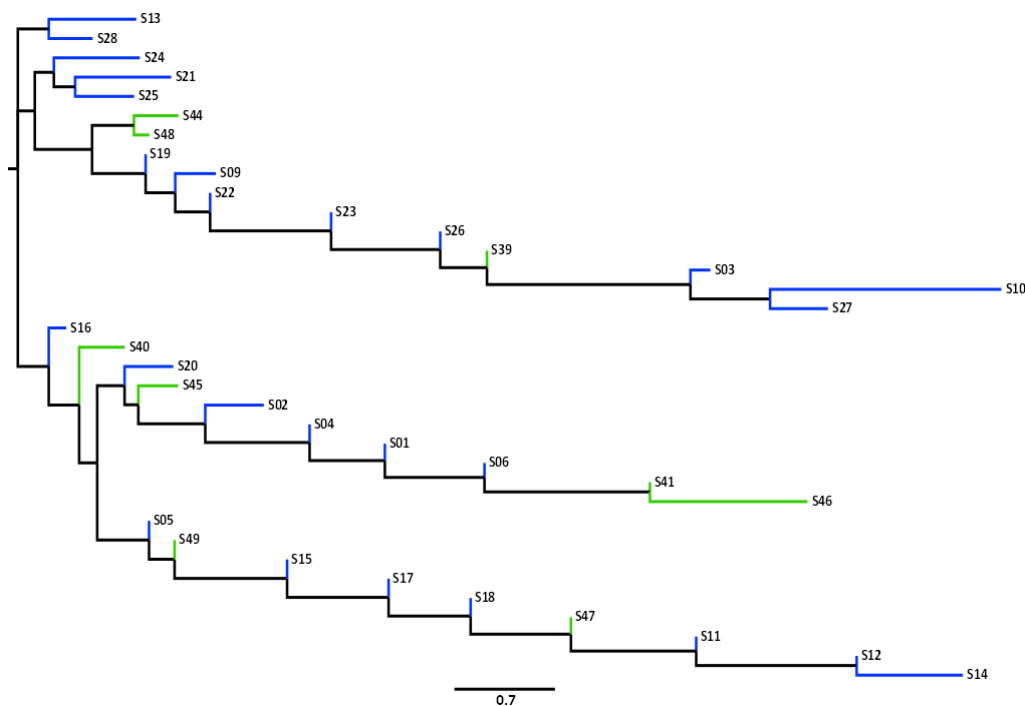


Figura 5 – Dendrograma construído com o método de neighbor-joining, com base em distâncias genéticas de Nei (1978), mostrando a relação genética de 35 indivíduos de sumaúma (*Ceiba pentandra*) avaliados com 5 locos microssatélites. Os ramos são coloridos de acordo com o local de coleta (azul – Vila de Teotônio; verde – Vila Calama).

No estudo de Brandão (2012), com a espécie *Ceibe pubiflora*, os parâmetros de distância genética indicaram que amostras de populações coletadas com distância a baixo de 400 Km quilômetros apresentaram valores indicativos de populações homogenia, ou seja, são geneticamente semelhantes.



Portanto o conhecimento da diversidade genética é primordial para a compreensão da dinâmica evolutiva das populações, sendo o fator fundamental para a adaptação às mudanças ambientais e evolutivas das espécies Frankham *et al.*, (2008).

## 5. CONSIDERAÇÕES FINAIS

Considerando as duas populações de *C. pentandra*, coletadas em dois locais diferentes, seguindo o curso rio do Madeira, pode-se concluir que os valores globais de diversidade genética foram intermediários, sendo que a maioria dos locos apresentou déficit de heterozigotos, estando abaixo do esperado, indicando que as populações não estão Equilíbrio de Hardy-Weinberg.

Quanto aos dados de divergência genética para este estudo, são indicativos de ausência de estrutura genética. O valor do  $F_{st}$ , entre as duas populações, foi baixo e não significativo, isto é, a maior parte da distribuição da variação genética encontra-se dentro das populações do que entre elas.

Considerando o valor do índice de dissimilaridade, foi baixo, concluindo-se que a distância genética estimada, entre os indivíduos das duas populações, foi de 95%, indicando que os genótipos amostrados compartilham dos recursos genéticos, tratando-se de uma população homogênea.

As informações obtidas neste estudo visam contribuir para planos efetivos de manejo e/ou melhoramento genético de samaúma (*C. pentandra*), reforçam a fragilidade desta espécie, nesta região, quanto à erosão genética, devido exploração e pressão antrópica. Os índices de diversidade genética, estimados a partir das frequências alélicas, exposto neste documento, são importantes indicadores para subsidiar medidas eficazes de conservação, *in situ* e *ex situ*, permitindo seleção de amostras representativas de populações remanescentes que conservam a variabilidade genética, imprescindível para sua sobrevivência ao longo do tempo.

## 6. BIBLIOGRAFIA

BRANDÃO, M. M., *Ceiba pubiflora* (Malvaceae): Diversidade Genética em Afloramento Calcário no Brasil Central. Tese, apresentada na Universidade Federal de Lavras, 2012.

BRONDANI, R. P. V., GAIOTTO, F. A., MISSIAGGIA, A. A., KIRST, GRIBEL, M. R., GRATTAPAGLIA, D. Microsatellite markers for *Ceiba pentandra* (Bombacaceae), an endangered tree species of the Amazon forest. *Molecular Ecology Notes* (2003) 3, 177–179

COLE DM, WHITE TL, NAIR PKR. 2007. Maintaining genetic resources of peach palm (*Bactris gasipaes* Kunth): The role of seed migration and swidden-fallow management in northeastern Peru. *Gen Resour Crop Evol* 54:189-204.

EARL, D.A.; VONHOLDT, B.M. Structure Harvester: a website and program for visualizing Structure output and implementing the Evanno method. *Cons Genet Res*, 4(2): 359-361, 2012.

EVANNO, G.; REGNAUT, S.; GOUDET, J. Detecting the number of clusters of individuals using the software Structure: a simulation study. *Mol Ecol*, 14: 2611-2620, 2005.

EXCOFFIER, L., LAVAL, G., SCHNEIDER, S. 2005. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evolution Bioinformatics Online* 1:47–50.

EXCOFFIER, L.; LISCHER, H.E.L. Arlequin suit ver. 3.5: A new series of programas to perform population genetics analyses under Linux or Windows. *Mol Ecol Resour*, 10(6): 564-567, 2010.

Forest Seed Center, 2000. 2 p. (Seed Leaflet, 22).

FRANKHAM, R., BALLOU, J. D., BRISCOE, D. A. Fundamento de Genética de Conservação. SBG. Ribeirão Preto- SP, 2008.

GOUDET, J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3).

GOVINDAJARU, R.D. 1989. Variation in gene flow levels among predominantly self-pollinated plants. *Journal of Evolutionary Biology*. 2: 173-181.

GRIBEL, P. E.; GIBBS, P. E.; QUEIRÓZ, A. L. Flowering phenology and pollination biology of *Ceiba pentandra* (Bombacaceae) in Central Amazonia.

HEDRICK, P.W. A standardized genetic differentiation measure. *Evolution*, 59: 1633–1638, 2005.

JAKOBSSON, M.; ROSENBERG, N.A. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23(14): 1801-1806, 2007.

JOKER, D.; SALAZAR, R. *Ceiba pentandra* (L.) Gaertn. Humlebaek: Danida

Journal of Tropical Ecology, v. 15, p. 247-263, 1999.

KAMVAR, Z.N.; TABIMA, J.F.; GRÜNWARD, N.J. *Poppr*: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 2:e281, 2014.

KEENAN, K.; MCGINNITY, P.; CROSS, T.F.; CROZIER, W.W.; PRODÖHL, P.A. *diveRsity*: an R package for the estimation of population genetics parameters and their associated errors, *Methods in Ecology and Evolution*, doi:10.1111/2041-210X.12067, 2013.

MARSHALL D.K.; BROWN, A.H.D. Optimum sampling strategies in genetic conservation. In: O.H. Frankel & J.G.R. Hawkes (Eds.). *Crop Genetic Resources for Today and Tomorrow*. Cambridge: Cambridge University Press, 1975. p.53-70.

MICHALAKIS, Y.; EXCOFFIER, L. A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics*, 142: 1061–1064, 1996.

MURAWSKI, D. A.; HAMRICK, J. L. Mating system and phenology of *Ceiba pentandra* (Bombacaceae) in Central Panama. *Journal of Heredity*, v. 83, n. 6,

MURAWSKI, D. A.; HAMRICK, J. L. The effect of the density of flowering individuals on the mating systems of nine tropical tree species. *Journal of Heredity*, v. 67, n. 2, p. 167-174, 1991.

NEI, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583-590.

NEI, M. Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sci USA*, 70: 3321–3323, 1973.

NEI, M. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89:583–590, 1978..

NEI, M. *Molecular evolutionary genetics*. New York: Columbia University Press, 1987.

ORTEGA ESCALONA, F.; CASTILLO MORALES, 1.; CARMONA VALDOVINOS, T. **Anatomía de la madera de veintiseis especies de la Selva Lacandona, Chiapas**. Mexico: Instituto de Ecología, Laboratorio de Ciencia y Tecnología de la Madera, 1991. 200 p. (*Angiospermas Arbóreas de México*, 3; *La Madera y su Uso*, 26). p.401-404,1992.

PEAKALL, R. & SMOUSE, P.E. 2006. *GenAIEx 6: Genetic Analysis in Excel*. Population genetic software for teaching and research. *Molecular Ecology Notes* 6: 288-295.

PEAKALL, R.; SMOUSE, P.E. GenAIEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics*, 28: 2537-2539, 2012.

PRITCHARD, J.K.; STEPHENS, M.; DONNELLY, P. Inference of population structure using multilocus genotype data. *Genetics*, 155: 945-959, 2000.

RODRIGUES, D. P. Diversidade genética e sistema de reprodução em progênies elites de pupunheira inerme (*Bactris gasipaes* Kunth) com marcadores microssatélites: implicações para o melhoramento do palmito, Tese (Doutorado em Biotecnologia) - Universidade Federal do Amazonas, Brasil, 2007.

ROGERS, J.S. 1972. Measures of genetic similarity and genetic distances. In: *Studies in genetics VII*. Publ.7213. Univ. of Texas, Austin, 145-153pp.

SAITOU, N.; NEI, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol*, 4(4): 406-425, 1987.

SÁNCHEZ, C.F.B. Diversidade Entre e Dentro de Populações Simuladas sob Deriva Genética. Dissertação desenvolvida pelo Programa de Pós-Graduação em Genética e Melhoramento. Viçosa-MG, 2008.

WEIR, B. S. & COCKERHAM, C. C. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358-1370.

WEIR, B. S. 1951. *Genetic data analysis: methods for discrete genetic data*. Sunderland: Sinauer Associates. 377p.

WEIR, B.S. & BASTEN, C.J. 1990. Sampling strategies for DNA sequence distance. *Biometrics*. 26: 551-582.

ZUCCHI, M.I. Análise da estrutura genética de *Eugenia dysenterica* DC. utilizando marcadores RAPD e SSR. 2002, 130p Tese (Doutorado em Genética e Melhoramento de Plantas) – Escola Superior de Agricultura Luiz de Queiroz, Piracicaba.