

JAY WALLACE DA SILVA E MOTA

**ANÁLISE DA DIVERSIDADE GENÉTICA DE GERMOPLASMA
DE *Theobroma cacao* L. DA AMAZÔNIA BRASILEIRA POR
MICROSSATÉLITES**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Fisiologia Vegetal, para obtenção do título de “Doctor Scientiae”

**VIÇOSA
MINAS GERAIS – BRASIL
2003**

JAY WALLACE DA SILVA E MOTA

**ANÁLISE DA DIVERSIDADE GENÉTICA DE GERMOPLASMA
DE *Theobroma cacao* L. DA AMAZÔNIA BRASILEIRA POR
MICROSSATÉLITES**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Fisiologia Vegetal, para obtenção do título de “Doctor Scientiae”

APROVADA: 30 de maio de 2003

Prof. Everaldo Gonçalves de
Barros
(Conselheiro)

Prof. Raimundo Santos Barros

Prof. Luis Orlando de Oliveira

Dr. Luis Antônio dos Santos Dias

Prof. Maurílio Alves Moreira
(Orientador)

Às pessoas mais importantes e
companheiras de todos os momentos:

Minha esposa *Santana*,

Nossos filhos: *Saulo*,

Jayana e

Diandro

Meus pais: *Wallace e*

Eudóxia

Meus sogros: *Wlá (in memoriam)*

e Flor

DEDICO

AGRADECIMENTOS

Meus mais sinceros agradecimentos a todas pessoas e instituições que de uma forma ou outra me auxiliaram no cumprimento de tão importante etapa na minha formação profissional, especialmente

Ao meu orientador, Prof. Maurílio Alves Moreira, pela amizade e orientação serena;

Aos meus conselheiros Profs. Everaldo Gonçalves de Barros, Sérgio Brommonschenkel e Dr. Cléber Novaes Bastos.

Ao Profs. Raimundo Santos Barros, Rolf Puschmann e Wagner Otoni;

Ao Dr. Antônio Figueira;

Ao Dr. Luiz Antônio dos Santos Dias;

A Lorena Sereno;

Ao grande amigo e irmão Aliomar Arapiraca;

Ao grande colega Paulo Albuquerque;

Aos colegas do Biomol: *Ivan Schuster, Ronan, Vagner, Chicão, Nilzinho, Taís, Malu, Rita, e Inês;*

Aos colegas do Seq-DNA: *Marta, Marcinho, Cynthia, Francismar, Cândida, Valéria, Marcinho (da Jussara), Márcia e Ricardo;*

Aos colegas e amigos do curso de Fisiologia Vegetal: *Jailson Lopes, Claudinéia, Marlos, Claudivan e Vespasiano;*

Às instituições: Universidade Federal de Viçosa, CNPq e CEPLAC;

Aos colegas de trabalho *Admil e Edivaldo;*

Aos queridos amigos: *Júlio, Fafá e Abelmom;*

Aos quase parentes: *Geraldinho, Fátima, Marcelo e Dudu;*

Às famílias: *Roberto, Sônia e filhas; Kleber, Natália e filhos;*

Aos funcionários e Professores da Fisiologia Vegetal e do Bioagro;

A todas as pessoas que me apoiaram nas diversas etapas da minha estada em Viçosa.

Obrigado.

CONTEÚDO

	Página
1. INTRODUÇÃO	01
2. MATERIAIS E MÉTODOS	07
2.1. Material vegetal utilizado	07
2.2. Extração de DNA	12
2.3. Detecção do polimorfismo dos microssatélites	13
2.3.1. Reações de microssatélites	13
2.3.2. Detecção dos fragmentos amplificados	14
2.4. Análises estatísticas	16
2.4.1. Modelo de mutação de microssatélites	16
2.4.2. Caracterização individual	17
2.4.3. Teste de neutralidade seletiva dos locos	18
2.4.4. Desequilíbrio genotípico	19
2.4.5. Heterozigosidade individual dos genótipos	19
2.4.6. Relação genética entre os clones	20
2.5. Variabilidade genética das subpopulações	20
2.5.1. Variabilidade genética dentro das subpopulações	20
2.5.2. Variabilidade genética entre as subpopulações	22
2.6. Distâncias genéticas entre as subpopulações	26
3. RESULTADOS E DISCUSSÃO	28
3.1. Nível de polimorfismo detectado na amostra total	28
3.2. Neutralidade seletiva e independência dos locos analisados ...	33
3.3. Poder discriminante dos microssatélites	36
3.4. Heterozigosidade individual dos clones analisados	39
3.5. Relação genética entre os clones	48
3.6. Caracterização das subpopulações analisadas	54

3.6.1. Análise das freqüências alélicas	54
3.6.2. Análise da variabilidade genéticas nas populações	57
3.6.3. Análise da variabilidade genética dentro e entre as populações	64
3.6.4. Contribuição de cada subpopulação à diversidade	70
3.7. Distâncias genéticas entre as populações	72
4. CONCLUSÕES	77
5. BIBLIOGRAFIA	79
6. ANEXOS	87

RESUMO

MOTA, Jay Wallace da Silva e, D.S., Universidade Federal de Viçosa, maio de 2003. **Análise da diversidade genética de germoplasma de *Theobroma cacao* L. da Amazônia Brasileira por microssatélites.** Orientador: Maurílio Alves Moreira. Conselheiros: Everaldo Gonçalves de Barros e Sérgio Hermínio Brommonschenkel

Uma amostra de 172 clones de cacau (*Theobroma cacao* L.), provenientes de amostras coletadas em populações espontâneas de 16 bacias hidrográficas da Amazônia Brasileira, foi caracterizada por meio de microssatélites visando identificar genótipos, estimar a diversidade genética e estruturas genéticas das populações de origem. Foram detectados 187 alelos a partir de 30 locos microssatélites, dos quais 29 apresentaram polimorfismo, seletivamente neutros e independentes, permitindo identificar 169 (98,3%) genótipos distintos e apenas três genótipos repetidos (CAB 731, CAB 732 e CAB 734), oriundos da bacia do Rio Uatumã-AM. Os genótipos agruparam-se segundo suas origens geográficas, com a dissimilaridade variando de 0,0 a 0,79 e média de 0,59. Foram identificados 68 alelos raros, dos quais 23 exclusivos, destacando-se as bacias do Acre e do Alto Amazonas com maior número de alelos privados, corroborando suas maiores riquezas alélicas. Todas as subpopulações apresentaram boa diversidade genética, exceto as de Rondônia que mostraram menor variabilidade para todos os indicadores utilizados. A diversidade genética total foi 0,658 contra uma heterozigosidade observada de 0,295, observando-se déficit de heterozigosidade em todas as amostras, caracterizado por um coeficiente de endogamia de 0,408 e taxa de fecundação cruzada de 42 %. As subpopulações apresentaram forte estruturação, com significativa diferenciação ($G_{ST}' = 0,292$ e $F_{ST} = 0,290$) e alto coeficiente de parentesco ($R_{EL} = 0,367$) dos indivíduos dentro das subpopulações. A análise das amostras agrupadas (Acre, Alto Amazonas, Rondônia e Baixo Amazonas) mostra 66,6% da variabilidade genética dentro das subpopulações, 14,9 % entre subpopulações dentro dos grupos e 18,5% entre os quatro grupos.

As populações do Alto Amazonas se destacaram em termos de contribuição relativa à diversidade e riqueza alélica total. No Acre, apesar da alta diversidade e riqueza alélica, apenas a bacia do Rio Tarauacá apresenta contribuição relativa positiva. As demais demonstram redundância na composição alélica, com contribuições relativas negativas. O dendrograma de distância genética D de Nei apresentou topologia mais robusta que $\delta\mu^2$ e D_L ($-\ln(1-F_{ST})$), confirmando o agrupamento das subpopulações do Acre e Alto Amazonas e destacando a elevada divergência da bacia do Rio Jamari em relação às demais. Os resultados realçam a importância de novas coletas nas bacias hidrográficas do Amazonas e Acre (especialmente a do rio Tarauacá) e a necessidade de agregar o máximo de novas bacias visando reunir o máximo de novos alelos, alelos raros e/ou exclusivos, de modo a enriquecer as coleções de germoplasma e a minimizar perdas por erosão genética.

ABSTRACT

MOTA, Jay Wallace da Silva e, D.S., Universidade Federal de Viçosa, May, 2003. **Genetic diversity analysis of *Theobroma cacao* L. germplasm from Brazilian Amazon with microsatellites.** Adviser: Maurílio Alves Moreira. Committee members: Everaldo Gonçalves de Barros and Sérgio Hermínio Brommonschenkel

A sample of 172 clonal genotypes of the cacao tree (*Theobroma cacao* L.) representing germplasm collections from natural subpopulations collected in 16 river basins of the Brazilian Amazon region was characterized using microsatellites to identify genotypes, assess genetic diversity and population structure. From 30 loci microsatellites analyzed 29 showed polymorphism revealing 186 alleles selectively neutral and independent, which allowed to identify 169 (98.3%) distinct genotypes and only three repeated ones (CAB 731, CAB 732 and CAB 734) from Uatumã (AM) river basin. The genotypes were clustered in accordance with their geographic origins, their dissimilarities varying from 0.0 to 0.79 and mean 0.59. Sixty eight rare alleles were identified from which 23 private ones, which places the basins from Acre and Upper Amazon with higher number of private alleles according to their greater allelic richness. All subpopulations showed relatively high genetic diversity except those from Rondonia that presented low variability to all indicators used. The Total diversity was 0.658 for an observed heterozygosity 0.295, a result of a heterozygosity deficit in every sample, characterized by an endogamy coefficient 0.408 and an outcrossing rate 42%. The subpopulations showed a strong structuring with a significant differentiation ($G_{ST}' = 0.293$ and $F_{ST} = 0.290$) and relatedness ($R_{EL} = 0.367$) of the individuals within of the subpopulations. Analyses of grouped samples (Acre, Upper Amazon, Rondonia and Low Amazon) showed 66.6% of the genetic variability within the subpopulations, 14.9% among subpopulations within the groups and 18.5% among the four groups. Upper Amazon populations produced the most relative contributions to the total diversity and allelic richness. In the Acre group, in spite of the high diversity and allelic richness, only the Tarauaca river basin presented positive relative contributions; the other Acre subpopulations showed redundant allelic

composition with negative relative contribution to total diversity and allelic richness. The Nei-based distance tree (D) showed more robust topology than $\delta\mu^2$ and $D_L [-\ln(1-F_{ST})]$ ratifying the clustering of the subpopulations from Acre and from Upper Amazon, emphasizing the high divergence between Jamari (RO) river basin and the others. The results highlight the importance of new germplasm collects in the river basins from Amazonas and Acre States (especially the Tarauaca river basin) and the need of aggregating the maximum number of new basins to assemble the maximum number of new alleles, rare and/or private alleles in order to increase the diversity of germplasm collections and minimize the loss by genetic erosion.

1. INTRODUÇÃO

O cacauzeiro (*Theobroma cacao* L.), espécie economicamente mais explorada do gênero, é um dos mais importantes cultivos perenes dos trópicos, sendo cultivado em todos os continentes, envolvendo mais de 2 milhões de produtores em mais de 50 países (Knight, 2000). Originado na região norte da América do Sul, o cacauzeiro vegeta espontaneamente no sub-bosque das florestas tropicais úmidas do hemisfério ocidental, na faixa de latitudes 18° N – 15° S, que se estende do sul do México até a bacia Amazônica. Essa dispersão, supostamente separada pelo levantamento dos Andes no início do Terciário, teria favorecido a especiação por meio do isolamento geográfico, dando origem a grupos raciais distintos denominados *Forasteiro* (*T. cacao* var. *spherocarpum*) e *Crioulo* (*T. cacao* var. *cacao*) (Cuatrecasas, 1964). Um terceiro grupo, *Trinitário*, surgiu da hibridação natural entre os dois primeiros em plantações devastadas por doenças em Trinidad (Cheesman, 1944). Hipóteses alternativas (Cheesman, 1944; Dias, 2001b) para os centros de origem e dispersão tentam explicar a diferenciação e relativo confinamento dos dois grupos separados pela Cordilheira dos Andes. Populações de Crioulo estão confinadas entre a encosta ocidental da cordilheira e a América Central, enquanto populações de Forasteiro ocorrem no Alto e no Baixo Amazonas, sendo encontrados ainda nas Guianas, Suriname e Venezuela.

Populações do grupo Forasteiro são altamente diversificadas e são freqüentemente usadas em programas de melhoramento devido ao seu vigor, precocidade e resistência a doenças. Por essas características o tipo Forasteiro é atualmente o grupo mais extensivamente cultivado, respondendo por mais de 80 % da produção mundial de cacau. Esse grupo subdivide-se em Forasteiro Baixo Amazônico e Forasteiro Alto Amazônico. Esse último apresenta maior diversidade exibindo performance agrônômica superior (Bartley, 1969, Lockwood, 1976).

A diversidade genética do cacauzeiro tem sido caracterizada através de características morfológicas (Engel, 1986; Bartley *et al.*, 1988), quantitativas (Dias, 2002; 2003a), marcadores isozimáticos (Lanaud, 1987; Sounigo *et al.* 1997; Lanaud, 1986), marcadores RAPD (Whitkus *et al.*, 1998; Lerceteau *et al.*,

1997; Figueira *et al.* 1994; N'Goran *et al.*, 1994; Laurent *et al.*, 1994; Marita *et al.*, 2001); marcadores RFLP (Laurent *et al.*,1995; Lerceteau *et al.* 1997, Motamayor *et al.*,2002) e microssatélites (Lanaud *et al.*, 1999a, Motamayor *et al.*, 2002; Sereno, 2001). Todos esses estudos têm demonstrado que as populações de cacau diferem fortemente em seus níveis de diversidade e heteroziguidade. Forasteiros Alto Amazônicos apresentam altos níveis de diversidade e níveis médios de heteroziguidade comparados aos das populações de Forasteiros Baixo Amazônicos e aos materiais nativos da Guiana Francesa, ambos caracterizados por baixas diversidade e heteroziguidade. A caracterização de coleções da CEPLAC (Bahia) com RAPD, envolvendo 270 genótipos, revela alto grau de variação entre tipos Alto Amazônicos e acessos coletados na Amazônia Brasileira. (Marita *et al.*, 2001). Esses estudos freqüentemente sugerem que efeitos fundadores e/ou áreas de refúgio desempenham importante papel na evolução da espécie.

Bem menos estudos têm se verificado sobre a estrutura genética de populações naturais de cacau. Alguns tratam como populações amostras reunindo representante dos grupos morfogeográficos ou regiões de origem como os de Lanaud, (1986, 1987), através de isoenzimas. Entretanto, Warren (1994), utilizando apenas quatro sistemas isoenzimáticos, foi capaz de detectar maior diversidade de populações de *Trinitários*, comparativamente a populações do Equador, da Colômbia e do Peru, destacando a baixa diversidade destas últimas. Novamente comparando grupos morfogeográficos, Lerceteau *et al.* (1997) detectou a baixa heteroziguidade de grupos *Forasteiros* do Peru e Equador (18,8%), em relação aos grupos *Crioulos* (0,301) e *Trinitários* (0,383%), com uma diferenciação entre os grupos de $G_{ST} = 0,161$, ficando 82 % da variabilidade dentro dos grupos. Utilizando 27 sondas RFLP, N'Goran *et al.* (2000) confirmou a maior heteroziguidade observada de *Crioulos* (0,250) e *Trinitários* (0,300), enquanto *Forasteiros Alto Amazônicos* e *Baixo Amazônicos* apresentaram apenas 0,230 e 0,126, respectivamente. A mesma tendência foi verificada por Motamayor *et al.* (2002), utilizando RFLP (25 sondas) e microssatélites (16 locos), com os grupos apresentando as seguintes heteroziguidades para locos RFLP e microssatélites, respectivamente: *Crioulos* (0,47 e 0,59), *Trinitários* (0,43 e 0,68) e *Forasteiros*

(0,18 e 0,34). Recentemente, Sereno (2001) utilizando 11 microssatélites caracterizou a estrutura de quatro populações (Alto Amazonas, Baixo Amazonas, Rondônia e Acre) constituídas por amostras agrupadas por critério geográfico, registrou heterozigosidade observada geral ($H_o=0,347$) compatível com a observada em materiais do Alto Amazonas, mas encontrou uma relativamente alta diversidade para as populações de Rondônia ($H_s = 0,555$) e Baixo Amazonas ($H_s = 0,553$), tidas como muito uniformes. As populações pouco divergiram concentrando 99,6% da variabilidade dentro das mesmas.

Maior área de ocorrência natural da espécie, a Amazônia Brasileira concentra a maior reserva da variabilidade genética do cacau, caracterizada por ampla variação de vigor, produtividade, resistência a doenças, tamanho e coloração de frutos, folhas e sementes (Barriga *et al.* 1985; Almeida *et al.* 1987, Bartley *et al.*, 1988; Almeida e Dias, 2001), componentes da produção (Dias *et al.*, 2002; 2003a), marcadores RAPD (Silva *et al.* 1998) e microssatélites (Sereno, 2001). Visando a assegurar que o máximo dessa diversidade genética seja preservada e utilizada de forma racional, a Comissão Executiva do Plano da Lavoura Cacaueira (CEPLAC) tem introduzido em suas coleções de germoplasma acessos provenientes de populações silvestres ou semi-cultivadas, coletados em expedições botânicas sistemáticas desde 1965 (Almeida *et al.*, 1995). Atualmente a Estação de Recursos Genéticos do Cacau “José Haroldo” (Marituba-PA) reúne cerca de 1800 acessos representantes de 36 bacias hidrográficas. Esse esforço envolve custos consideráveis tanto para a coleta quanto para manutenção das coleções, cujas sementes recalcitrantes não se prestam ao armazenamento prolongado. O crescimento contínuo das coleções e conseqüentes custos, associados a orçamentos escassos, tem limitado consideravelmente novas prospecções e coletas aumentando os riscos de erosão genética ante a crescente pressão antrópica na região; apenas 20 % das bacias mapeadas estão representadas nas coleções de germoplasma da CEPLAC. Por outro lado, a falta de melhor conhecimento acerca dos padrões de diversidade e diferenciação das populações naturais na região dificulta a adoção de estratégias racionais de manejo e coleta de recursos genéticos. A identificação de áreas prioritária para coletas, a eliminação de acessos duplicados, a preservação de alelos raros e/ou exclusivos, o estabelecimento

de coleções nucleares, entre outros, são medidas cuja adoção depende de uma melhor caracterização das coleções existentes, visando minimizar custos e maximizar o aproveitamento da diversidade disponível mas pouco acessível. Segundo Brown (1989), a partir da identificação de alelos comuns (raros ou dispersos), alelos raros (localizados ou dispersos) é possível estruturar coleções nucleares que representem de 70 a 80% da variação genética das coleções, utilizando-se apenas 10% de seus acessos.

O manejo racional dos recursos genéticos requer ainda conhecimentos sobre os níveis de variação existente dentro e entre as populações bem como a sua distribuição geográfica e os fatores que a influenciam. Relatos sugerindo a distribuição descontínua de pequenas concentrações de cacau ao longo de cursos d'água parecem reforçar o papel desempenhado pelo rios na divergência alopátrica das populações naturais de cacau na Amazônia, como sugerem Pound (1938), Almeida *et al.* (1987), Young (1994) e Dias *et al.* (2002; 2003a). O mesmo padrão de associação entre diversidade e cursos d'água tem sido apontado para espécies amazônicas como *Hevea brasiliensis* (Seguin *et al.*, 1999) e *Elaeis oleífera* (Barcelos, 1998), ambos citados por Motamayor (2002). A verificação de tal padrão pode fornecer subsídios muito importantes para otimizar, em bases científicas, o estabelecimento das coleções e a conservação do germoplasma.

A utilização de marcadores moleculares altamente polimórficos constitui recurso dos mais adequados à caracterização da variabilidade existentes nas coleções, visando estabelecer inferências acerca do padrão de estruturação genética das populações naturais de origem. A abundante informação gerada por marcadores moleculares é de grande utilidade no manejo da diversidade em bancos de germoplasma, facilitando o estabelecimento de coleções básicas, nucleares e de trabalho, garantindo a completa representação da diversidade e evitando a redundância dos genótipos duplicados. Além disso, as informações geradas são altamente valiosas para a identificação de progenitores divergentes de modo a maximizar heterose, visando a produção de híbridos superiores.

Vários marcadores moleculares, como já citados, têm sido freqüentemente utilizados na caracterização molecular, diversidade e estrutura

de populações de cacau. Entretanto, recentemente, os marcadores microssatélites têm sido largamente empregados, tornando-se o marcador de escolha, em função de suas vantagens em relação a outros marcadores.

Microssatélites são curtos segmentos de DNA, nos quais um grupo específico de 1-6 bases é repetido usualmente até no máximo 60 ou mais vezes. Eles são altamente polimórficos e podem ser amplificados por PCR com *primers* específicos para as regiões que flanqueiam as seqüências repetidas (Ferreira & Grattapaglia, 1998). Devido à sua excepcional variabilidade e relativa facilidade de “escorear” são considerados atualmente o mais poderoso marcador genético (Goldstein & Schlötterer, 2000). É freqüente observar-se locos com mais de 10 alelos e heterozigosidades acima de 0,6, mesmo em amostras pequenas (Deka *et al.*, 1995; Bowcock *et al.* 1994), enquanto certos locos podem ser consideravelmente mais variáveis. Além de altamente polimórficos, os microssatélites são também densamente distribuídos por todo o genoma eucariótico, tornando-os marcadores preferidos para mapeamentos de alta resolução (Goldstein & Schlöterer, 2000). Além de mapeamento, os microssatélites tem sido extensivamente usados para análises de ligação em associação com genes de susceptibilidade a doenças (Yechezkel, & Morris, 1999). São de uso quase exclusivo em testes de paternidade (Queller *et al.*, 1993) e na probabilidade de identificação em nível individual (Edwards *et al.*, 1992) e em nível de população (Paetkau *et al.* 1995). Comparações de níveis de variação entre espécies e populações na avaliação da variação genética total (Paetkau & Strobeck, 1995), estimação do tamanho efetivo da população, avaliação do grau de estruturação de populações e taxa de migração entre subpopulações (Allen *et al.*, 1995) e relações entre várias subpopulações (Bowcock *et al.* 1994; Forbes *et al.* 1995), são algumas das aplicações mais freqüentes de microssatélites no estudo de genética de populações. *Primers* específicos para cacau, recentemente desenvolvidos por Lanaud *et al.* (1999a), tem mostrado bastante polimorfismo em pequenas amostras permitindo a identificação de locos com até 13 alelos. Esses microssatélites vem sendo usados com sucesso na distinção de genótipos pertencentes a grupos morfogeográficos distintos (Motamyor, *et al.*, 2002), na construção de mapas de ligação de alta densidade (Risterucci *et al.*, 2000) e na

análise de algumas populações naturais de cacau da Amazônia (Sereno, 2001).

Desse modo, tendo por base o polimorfismo característico dos locos microssatélites, esse trabalho pretende cumprir os seguintes objetivos:

1. Caracterizar individualmente e estabelecer a relação genética entre 172 acessos clonais coletados em diferentes bacias hidrográficas da Amazônia.
2. Identificar alelos raros e/ou de ocorrência exclusiva das diferentes amostras.
3. Estabelecer inferências sobre a estrutura genética e distribuição da variabilidade dentro e entre subpopulações naturais de cacau, a partir de amostras (acessos) coletadas em diferentes bacias hidrográficas da Amazônia.
4. Estimar a contribuição relativa de cada subpopulação estudada à variabilidade total associada ao conjunto de todas as subpopulações.
5. Estimar o grau de endogamia e taxas de cruzamento dentro de cada subpopulação.
6. Estimar as distâncias genéticas entre as subpopulações e caracterizar o padrão de agrupamento associado a essas medidas.
7. Sugerir estratégias de conservação e coletas com base na distribuição da variabilidade e diferenciação das subpopulações.

2. MATERIAL E MÉTODOS

2.1. Material vegetal utilizado

Todo o material vegetal analisado no presente estudo obtido das coleções de germoplasma silvestre da Estação de Recursos Genéticos do Cacau – “José Haroldo” (ERJOH), mantida pela CEPLAC no município de Marituba-PA. Tais coleções constituem amostras de populações nativas da Amazônia, cujo critério de coleta e registro observa a dispersão do cacau associada à hidrografia da região, a qual foi subdividida em 4 bacias hidrográficas primárias, 61 secundárias e 121 terciárias (Almeida et al., 1987).

Foram utilizados 170 acessos genéticos (clones) representantes de amostras de 16 subpopulações, coletadas em diferentes micro-regiões hidrográficas da Amazônia Brasileira, conforme discriminam na Tabela 1 e Figura 1, construídas de acordo com Almeida *et al.*(1995). Acrescentam-se, ainda, os acessos alienígenas, o clone SCA-6, originário do Peru, e clone ICS-1, proveniente de cruzamentos de material Trinitário. Esses materiais, além de bastante utilizados no programa de produção de sementes da CEPLAC, são muito usados como progenitores na obtenção de progênies tolerantes à vassoura-de-bruxa (*Crinipellis perniciosa*).

O termo acesso é aqui empregado com o sentido mais comumente utilizado para qualificar amostras de germoplasma que representem a variação genética de uma população ou de um indivíduo propagado vegetativamente (clone). Esse termo é normalmente adotado para indicar que se trata de amostras obtidas através de procedimentos de coleta, as quais reproduzem as características genéticas das populações de origem. (Vilela-Morales *et al.*, 1997).

Para efeito de análise considerou-se subpopulação, a unidade amostral representativa da respectiva bacia hidrográfica, cuja agregação em escala maior foi considerada como população ou grupo representativo das respectivas regiões ou estados que a compreendem. Desse modo considerou-se como populações a reunião das amostras das subpopulações do Acre, Rondônia, Amazonas e Baixo Amazonas. Esta última, assim denominada pela inclusão de

Tabela 1. Acessos genéticos de cacau caracterizados agrupados de acordo com a localização geográfica das populações amostradas.

Nº	Acesso	Bacia Hidrográfica (subpopulação)	Local (mapa)	Local de Coleta	Grupo (população)
001	CAB 104	Rio Acre	34	Rio Acre (Ac)	Acre
002	CAB 105	Rio Acre	34	Rio Acre (AC)	Acre
003	CAB 107	Rio Acre	34	Rio Acre (AC)	Acre
004	CAB 108	Rio Acre	34	Rio Acre (AC)	Acre
005	CAB 110	Rio Acre	34	Rio Acre (AC)	Acre
006	CAB 111	Rio Acre	34	Rio Acre (AC)	Acre
007	CAB 113	Rio Acre	34	Rio Acre (AC)	Acre
008	CAB 114	Rio Acre	34	Rio Acre (AC)	Acre
009	CAB 117	Rio Acre	34	Rio Acre (AC)	Acre
010	CAB 135	Rio Acre	34	Rio Acre (AC)	Acre
011	CAB 136	Rio Acre	34	Rio Acre (AC)	Acre
012	CAB 146	Rio Acre	34	Rio Acre (AC)	Acre
013	CAB 147	Rio Acre	34	Rio Acre (AC)	Acre
014	CAB 148	Rio Acre	34	Rio Acre (AC)	Acre
015	CAB 151	Rio Acre	34	Rio Acre (AC)	Acre
016	CAB 153	Rio Acre	34	Rio Acre (AC)	Acre
017	CAB 155	Rio Acre	34	Rio Acre (AC)	Acre
018	CAB 160	Rio Acre	34	Rio Acre (AM)	Acre
019	CAB 167	Rio Acre	34	Rio Acre (AM)	Acre
020	CAB170	Rio Acre	34	Rio Acre (AM)	Acre
021	CAB 171	Rio Acre	34	Rio Acre (AM)	Acre
022	CAB 176	Rio Acre	34	Rio Acre (AM)	Acre
023	CAB 120	Rio Chandless	37	Rio Chandless - AC	Acre
024	CAB 121	Rio Chandless	37	Rio Chandless - AC	Acre
025	CAB 119	Rio Chandless	37	Rio Chandless - AC	Acre
026	CAB 066	Rio Iaco	36	Rio Caeté - AC	Acre
027	CAB 082	Rio Iaco	36	Rio Macauã - AC	Acre
028	CAB 063	Rio Iaco	36	Rio Caeté - AC	Acre
029	CAB 064	Rio Iaco	36	Rio Caeté - AC	Acre
030	CAB 067	Rio Iaco	36	Rio Caeté - AC	Acre
031	CAB 068	Rio Iaco	36	Rio Caeté - AC	Acre
032	CAB 070	Rio Iaco	36	Rio Caeté - AC	Acre
033	CAB 071	Rio Iaco	36	Rio Caeté - AC	Acre
034	CAB 072	Rio Iaco	36	Rio Caeté - AC	Acre
035	CAB 076	Rio Iaco	36	Rio Caeté - AC	Acre
036	CAB 077	Rio Iaco	36	Rio Caeté - AC	Acre
037	CAB 079	Rio Iaco	36	Rio Caeté - AC	Acre
038	CAB 080	Rio Iaco	36	Rio Caeté - AC	Acre
039	CAB 081	Rio Iaco	36	Rio Macauã - AC	Acre
040	CAB 088	Rio Iaco	36	Rio Iaco - AC	Acre
041	CAB 181	Rio Tarauacá	38	Rio Tarauacá - AC	Acre
042	CAB 200	Rio Tarauacá	38	Rio Embira - AC	Acre
043	CAB 201	Rio Tarauacá	38	Rio Embira - AC	Acre
044	CAB 202	Rio Tarauacá	38	Rio Embira - AC	Acre
045	CAB 203	Rio Tarauacá	38	Rio Embira - AC	Acre
046	CAB 182	Rio Tarauacá	38	Rio Tarauacá - AC	Acre
047	CAB 183	Rio Tarauacá	38	Rio Embira - AC	Acre
048	CAB 184	Rio Tarauacá	38	Rio Embira - AC	Acre
049	CAB 185	Rio Tarauacá	38	Rio Embira - AC	Acre
050	CAB 187	Rio Tarauacá	38	Rio Embira - AC	Acre
051	CAB 102	Rio Purus	35	Rio Purus - AM	Acre
052	CAB 191	Rio Purus	35	Rio Purus - AM	Acre
053	CAB 094	Rio Purus	35	Rio Purus - AM	Acre
054	CAB 095	Rio Purus	35	Rio Purus - AM	Acre
055	CAB 096	Rio Purus	35	Rio Purus - AM	Acre
056	CAB 098	Rio Purus	35	Rio Purus - AM	Acre
057	CAB 101	Rio Purus	35	Rio Purus - AM	Acre
058	CAB 123	Rio Purus	35	Rio Purus - AC	Acre
059	CAB 126	Rio Purus	35	Rio Purus - AC	Acre
060	CAB 127	Rio Purus	35	Rio Purus - AC	Acre
061	CAB 128	Rio Purus	35	Rio Purus - AC	Acre
062	CAB 130	Rio Purus	35	Rio Purus - AC	Acre
063	CAB 131	Rio Purus	35	Rio Purus - AC	Acre
064	CAB 132	Rio Purus	35	Rio Purus - AC	Acre
065	CAB 206	Rio Purus	35	Rio Purus - AM	Acre

Tabela 1. Acessos genéticos de cacau caracterizados agrupados de acordo com a localização geográfica das populações amostradas (continuação).

Nº	Acesso	Bacia Hidrográfica (subpopulação)	Local (mapa)	Local de Coleta	Grupo (população)
066	CJ 03	Rio Jari	8	Rio Jari/Cach Sto Antônio, Marzagão	Bx Amazonas
067	CJ 10	Rio Jari	8	Rio Jari/Cach Sto Antônio, Marzagão	Bx Amazonas
068	CJ 11	Rio Jari	8	Rio Jari/Cach Sto Antônio, Marzagão	Bx Amazonas
069	CJ 02	Rio Jari	8	Rio Jari/Cach Sto Antônio, Marzagão	Bx Amazonas
070	CJ 04	Rio Jari	8	Rio Jari/Cach Sto Antônio, Marzagão	Bx Amazonas
071	CJ 05	Rio Jari	8	Rio Jari/Cach Sto Antônio, Marzagão	Bx Amazonas
072	CJ 06	Rio Jari	8	Rio Jari/Cach Sto Antônio, Marzagão	Bx Amazonas
073	CJ 07	Rio Jari	8	Rio Jari/Cach Sto Antônio, Marzagão	Bx Amazonas
074	CJ 08	Rio Jari	8	Rio Jari/Cach Sto Antônio, Marzagão	Bx Amazonas
075	CJ 09	Rio Jari	8	Rio Jari/Cach Sto Antônio, Marzagão	Bx Amazonas
076	CAB 478	Rio Japurá	23	Alto e médio Japurá - AM	Amazonas
077	CAB 484	Rio Japurá	23	Alto e médio Japurá - AM	Amazonas
078	CAB 486	Rio Japurá	23	Alto e médio Japurá - AM	Amazonas
079	CAB 487	Rio Japurá	23	Alto e médio Japurá - AM	Amazonas
080	CAB 493	Rio Japurá	23	Alto e médio Japurá - AM	Amazonas
081	CAB 497	Rio Japurá	23	Alto e médio Japurá - AM	Amazonas
082	CAB 499	Rio Japurá	23	Alto e médio Japurá - AM	Amazonas
083	CAB 481	Rio Japurá	23	Alto e médio Japurá - AM	Amazonas
084	CAB 509	Rio Japurá	23	Alto e médio Japurá - AM	Amazonas
085	CAB 511	Rio Japurá	23	Alto e médio Japurá - AM	Amazonas
086	CAB 512	Rio Japurá	23	Alto e médio Japurá - AM	Amazonas
087	CAB 513	Rio Japurá	23	Alto e médio Japurá - AM	Amazonas
088	CAB 492	Rio Japurá	23	Alto e médio Japurá - AM	Amazonas
089	CAB 500	Rio Japurá	23	Alto e médio Japurá - AM	Amazonas
090	CAB 056	Rio Solimões	17	Manaus - AM	Amazonas
091	CAB 058	Rio Solimões	17	Manaus - AM	Amazonas
092	CAB 271	Rio Solimões	21	Tefé -AM	Amazonas
093	CAB 002	Rio Solimões	21	Ilha de Tarará, Tefé - AM	Amazonas
094	CAB 003	Rio Solimões	21	Ilha de Tarará, Tefé - AM	Amazonas
095	CAB 004	Rio Solimões	21	Ilha de Tarará, Tefé - AM	Amazonas
096	CAB 005	Rio Solimões	21	Ilha de Tarará, Tefé - AM	Amazonas
097	CAB 046	Rio Solimões	17	Manaus - AM	Amazonas
098	CAB 270	Rio Solimões	21	Tefé -AM	Amazonas
099	CAB 327	Rio Solimões	21	Tefé -AM	Amazonas
100	CAB 279	Rio Solimões	21	Tefé -AM	Amazonas
101	CAB 324	Rio Solimões	21	Tefé -AM	Amazonas
102	CAB 325	Rio Solimões	21	Tefé -AM	Amazonas
103	CAB 326	Rio Solimões	21	Tefé -AM	Amazonas
104	CAB 330	Rio Solimões	21	Tefé -AM	Amazonas
105	CAB 331	Rio Solimões	21	Tefé -AM	Amazonas
106	CAB 001	Rio Solimões	21	Ilha de Tarará, Tefé - AM	Amazonas
107	CAB 337	Rio Solimões	21	Tefé -AM	Amazonas
108	CAB 275	Rio Solimões	21	Médio Solimões nas imediações de Tefé - AM	Amazonas
109	CAB 316	Rio Solimões-B. Japurá	21	Parnás, Cubuá, Capivara, Copeá e Carapanatuba - AM	Amazonas
110	CAB 282	Rio Solimões-B. Japurá	22	Baixo Japurá - AM	Amazonas
111	CAB 266	Rio Solimões-B. Japurá	22	Baixo Japurá - AM	Amazonas
112	CAB 280	Rio Solimões-B. Japurá	22	Baixo Japurá - AM	Amazonas
113	CAB 283	Rio Solimões-B. Japurá	22	Baixo Japurá - AM	Amazonas
114	CAB 302	Rio Solimões-B. Japurá	21	Parnás, Cubuá, Capivara, Copeá E Carapanatuba - AM	Amazonas
115	CAB 304	Rio Solimões-B. Japurá	21	Parnás, Cubuá, Capivara, Copeá E Carapanatuba - AM	Amazonas
116	CAB 306	Rio Solimões-B. Japurá	21	Parnás, Cubuá, Capivara, Copeá E Carapanatuba - AM	Amazonas
117	CAB 308	Rio Solimões-B. Japurá	21	Parnás, Cubuá, Capivara, Copeá E Carapanatuba - AM	Amazonas
118	CAB 517	Rio Solimões-B. Japurá	22	Baixo Japurá - AM	Amazonas
119	CAB 313	Rio Solimões-B. Japurá	21	Parnás, Cubuá, Capivara, Copeá E Carapanatuba - AM	Amazonas
120	CAB 312	Rio Solimões-B. Japurá	21	Parnás, Cubuá, Capivara, Copeá E Carapanatuba - AM	Amazonas
121	CAB 514	Rio Solimões-B. Japurá	22	Baixo Japurá - AM	Amazonas
122	CAB 529	Rio Solimões-B. Japurá	22	Baixo Japurá - AM	Amazonas
123	CAB 730	Rio Uatamã/Balbina	16	Arredores da hid. De Balbina - AM	Amazonas
124	CAB 731	Rio Uatamã/Balbina	16	Arredores da hid. De Balbina - AM	Amazonas
125	CAB 732	Rio Uatamã/Balbina	16	Arredores da hid. De Balbina - AM	Amazonas
126	CAB 734	Rio Uatamã/Balbina	16	Arredores da hid. De Balbina - AM	Amazonas
127	CAB 006	Rio Acará	5	Rio Acará, Tomé Açu - PA	Bx Amazonas
128	CAB 020	Rio Acará	5	Rio Acará, Tomé Açu - PA	Bx Amazonas
129	CAB 036	Rio Acará	5	Rio Acará, Tomé Açu - PA	Bx Amazonas
130	CAB 016	Rio Acará	5	Rio Acará, Tomé Açu - PA	Bx Amazonas

Tabela 1. Acessos genéticos de cacau caracterizados agrupados de acordo com a localização geográfica das populações amostradas (continuação).

Nº	Acesso ¹	Bacia Hidrográfica (subpopulação)	Local ² (mapa)	Local de Coleta	Estado (população)
131	CAB 020b	Rio Acará	5	Rio Acará, Tomé Açu - PA	Bx Amazonas
132	CAB 039	Rio Acará	5	Rio Acará, Tomé Açu - PA	Bx Amazonas
133	CAB 017	Rio Maicuru	10	Rod. Alenquer-Monte Alegre, M. Alegre - PA	Bx Amazonas
134	CAB 018	Rio Maicuru	10	Rod. Alenquer-Monte Alegre, M. Alegre - PA	Bx Amazonas
135	CAB 019	Rio Maicuru	10	Rod. Alenquer-Monte Alegre, M. Alegre - PA	Bx Amazonas
136	CAB 024	Rio Maicuru	10	Rod. Alenquer-Monte Alegre, M. Alegre - PA	Bx Amazonas
137	CAB 022	Rio Maicuru	10	Rod. Alenquer-Monte Alegre, M. Alegre - PA	Bx Amazonas
138	CAB 023	Rio Maicuru	10	Rod. Alenquer-Monte Alegre, M. Alegre - PA	Bx Amazonas
139	CAB 035	Rio Maicuru	10	Stm 61/01 - PA	Bx Amazonas
140	CAB 007	Rio Maicuru	10	Alenquer - PA	Bx Amazonas
141	CAB 028	Rio Maicuru	10	Paraná-Mirim, Alenquer - PA	Bx Amazonas
142	CAB 029	Rio Maicuru	10	Paraná-Mirim, Alenquer - PA	Bx Amazonas
143	CAB 091	Rio Tocantins	6	Breu Branci, Tucuruí - PA	Bx Amazonas
144	CAB 092	Rio Tocantins	6	Breu Branci, Tucuruí - PA	Bx Amazonas
145	CAB 221	Rio Jamari	30	Ariquemes - RO	Rondônia
146	CAB 232	Rio Jamari	30	Ariquemes - RO	Rondônia
147	CAB 233	Rio Jamari	30	Ariquemes - RO	Rondônia
148	CAB 247	Rio Jamari	30	Ariquemes - RO	Rondônia
149	CAB 410	Rio Jamari	30	Ariquemes - RO	Rondônia
150	CAB 253	Rio Jamari	30	Ariquemes - RO	Rondônia
151	CAB 255	Rio Jamari	30	Ariquemes - RO	Rondônia
152	CAB 258	Rio Jamari	30	Ariquemes - RO	Rondônia
153	CAB 260	Rio Jamari	30	Ariquemes - RO	Rondônia
154	CAB 374	Rio Jamari	30	Ariquemes - RO	Rondônia
155	CAB 388	Rio Jamari	30	Ariquemes - RO	Rondônia
156	CAB 396	Rio Jamari	30	Ariquemes - RO	Rondônia
157	CAB 398	Rio Jamari	30	Ariquemes - RO	Rondônia
158	CAB 406	Rio Jamari	30	Ariquemes - RO	Rondônia
159	CAB 236	Rio Jamari	30	Ariquemes - RO	Rondônia
160	CAB 008	Rio Ji-Paraná	29	Seringal Muqui, Pdte Médice - RO	Rondônia
161	CAB 009	Rio Ji-Paraná	29	Seringal Muqui, Pdte Médice - RO	Rondônia
162	CAB 011	Rio Ji-Paraná	29	Seringal Muqui, Pdte Médice - RO	Rondônia
163	CAB 012	Rio Ji-Paraná	29	Seringal Muqui, Pdte Médice - RO	Rondônia
164	CAB 013	Rio Ji-Paraná	29	Seringal Muqui, Pdte Médice - RO	Rondônia
165	CAB 177	Rio Ji-Paraná	29	Ji-Paraná, O Preto d'oeste e Mirante da Serra - RO	Rondônia
166	CAB 178	Rio Ji-Paraná	29	Ji-Paraná, O Preto d'oeste e Mirante da Serra - RO	Rondônia
167	CAB 179	Rio Ji-Paraná	29	Ji-Paraná, O Preto d'oeste e Mirante da Serra - RO	Rondônia
168	CAB 415	Rio Ji-Paraná	29	Ji-Paraná - RO	Rondônia
169	CAB 014	Rio Xeriuni	18	Taiano - RR	Roraima
170	CAB 015	Rio Xeriuni	18	Taiano - RR	Roraima
171	SCA..6			Peru	Peru
172	ICS..1			Trinidad Tobago	Trinidad

1- CAB: Cacau da Amazônia Brasileira.

2- Local de coletas dos acessos indicados na Figura 1, conforme Almeida *et al.* (1995).

2.2. Extração de DNA

O material para a extração de DNA foi coletado em campo a partir de folhas jovens e sadias retiradas sempre de uma mesma planta de cada clone (acesso). Destas folhas, depois de lavadas com água corrente, água destilada e secas com papel absorvente, retirou-se discos foliares ($\Phi = 0,5$ cm), os quais, acondicionados em tubos tipo Eppendorf de 1,5 ml, que foram congelados em nitrogênio líquido e transportados em gelo seco para Viçosa.

A extração do DNA seguiu o protocolo de Doyle & Doyle (1990), com modificações. Os discos foliares de cada amostra foram macerados em nitrogênio líquido e transferidos para tubos Eppendorf, adicionando-se aos mesmos 800 μ L de tampão de extração [Tris-HCl 100 mM, pH 8,0, EDTA 20 mM, pH 8,0, NaCl 1,4 M, β -mecaptoetanol 0,2% (v/v) e CTAB (p/v)]. Depois de 40 minutos em banho-maria a 65 °C, os tubos foram centrifugados a 14.000 rpm, por cinco minutos, e o sobrenadante transferido para novos tubos. Ao sobrenadante foram adicionados 800 μ L de clorofórmio:álcool isoamílico (24:1), promovendo-se a mistura por inversões suaves e separando-se as fases por centrifugação a 14.000 rpm, por cinco minutos. O sobrenadante, transferido para novos tubos, recebeu isopropanol gelado, na proporção 1:1 (isopropanol:sobrenadante), o qual após suavemente misturado foi incubado por 3 h a -20 °C. Após nova centrifugação (14.000 rpm, 5 min), descartou-se o sobrenadante, lavando-se o precipitado branco duas vezes com etanol 70 % e uma vez com etanol 95 %, secando-se o precipitado à temperatura ambiente por 15 minutos. Em seguida, o precipitado foi ressuspensão em 300 μ L de TE (Tris-HCl 10 mM, EDTA 1 mM, pH 8,0), contendo RNase na concentração final de 40 μ g/mL, e incubando em banho-maria a 37 °C durante 30 minutos. Procedeu-se a nova precipitação adicionando-se ao tubo NaCl 5M na proporção 1:10 (NaCl:DNA ressuspensão) e 2/3 de isopropanol gelado, seguido de centrifugação a 14.000 rpm por 10 minutos. Após lavar duas vezes com etanol 70 % e uma vez com etanol 95 %, o precipitado foi seco ao ar e ressuspensão em 200 μ L de TE_{0,1} (Tris-HCl 10 mM, EDTA 0,1 mM, pH 7,5).

A qualidade do DNA foi avaliada em gel de agarose 0,8 % bem como por espectrofotometria, considerando a razão A_{260}/A_{280} . A concentração foi estimada a partir da absorbância a 260 nm conforme Sambrook et al. (1989),

estocando-se as amostras (estoque) a -20 °C e as amostras de trabalho (6 ng/μL) mantidas em geladeira (5 °C).

2.3. Detecção de polimorfismo dos microssatélites

2.3.1. Reações de Microssatélites

As reações dos microssatélites foram realizadas em microtubos de 0,2 mL, sendo o volume total da reação de 13 μL, contendo tampão da PCR (Tris-HCl 10 mM, pH 8, KCl 50 mM, MgCl₂ 1,0 mM), 100 μM de cada desoxirribonucleotídeos (dATP, dTTP, dGTP e dCTP), 0,2 μM de cada *primer* (*forward* e *reverse*) específico para cada microssatélite, uma unidade da enzima *Taq* polimerase e 30 ng de DNA molde.

As amplificações foram feitas em termocicladores Perkin Elmer (GeneAmp PCR System 9600) usando dois diferentes programas *touch down*, conforme a temperatura de anelamento de cada par de *primer* (Tabela 2). O programa *touch down I*, para os *primers* que anelam a 51 °C, consistiu de um ciclo de desnaturação a 94 °C durante 4 min, uma etapa de anelamento a 65 °C por 40 s, seguidos por 10 ciclos *touch down* decrescendo 1 °C por ciclo até 55 °C, a partir da qual se seguiam 30 ciclos a 55 °C por 40 s cada. Cada ciclo mantinha suas respectivas temperaturas de desnaturação (94 °C/40s) e polimerização (72 °C/1 min.). O programa *touch down II*, para os *primers* que anelam a 46 °C, diferia apenas no que diz respeito à etapa *touch down*, cujos primeiro e décimo ciclos apresentavam temperaturas de anelamento de 55 °C e 45 °C, respectivamente.

Os *primers* usados na fase de ajuste da metodologia foram cedidos pelo Dr. Antônio Figueira (USP/CENA), enquanto os *primers* utilizados na caracterização definitiva dos acessos genéticos (Tabela 2) foram sintetizados pela *Life Technologies* do Brasil, com base nas seqüências disponibilizadas pela Dra. Claire Lanaud do CIRAD-CP, Montpellier, França (Lanaud et al., 1999a). Foram testados 48 pares *primers*, entretanto somente 30 mostraram produtos de fácil detecção, dos quais, apenas os polimórficos (29) foram utilizados na análise.

2.3.2. Detecção dos fragmentos amplificados

Como se observa na Tabela 2, grande parte dos *primers* utilizados na caracterização amplificam microssatélites constituídos de repetições de dinucleotídeos, os quais, quando detectados em géis de poliacrilamida de alta resolução (desnaturantes), costumam apresentar bandas “gaguejantes” (*shadow bands*), o que dificulta ou impossibilita a genotipagem, especialmente de indivíduos heterozigotos para alelos que diferem em comprimento de apenas dois nucleotídeos. Desse modo, visando minimizar ambigüidade decorrente das bandas “gaguejantes”, decidiu-se por separar os fragmentos em gel desnaturante de poliacrilamida a 7 %, contendo 32 % de formamida/5,6 M de uréia, em TBE 1X, conforme protocolo proposto por Litt *et al.* (1993).

Foram utilizadas cubas verticais de eletroforese com placas de gel de 30 cm de altura com capacidade para 100 amostras em gel de 0,7 mm de espessura. Antes da aplicação das amostras o gel era submetido a uma pré-corrída de 30 minutos a uma potência constante, de modo a obter-se uma temperatura de aproximadamente 40 °C na placa de vidro do gel. Para tanto, mantinha-se a potência da fonte em 60 W ou 120 W, conforme se estivesse correndo um ou dois géis simultaneamente com a mesma fonte. As amostras a serem corridas eram misturadas com igual volume (4 µL) do tampão de carregamento (formamida 95%, EDTA 10 mM, azul de bromofenol 0,09 %) e, pouco antes do carregamento, as misturas eram desnaturada a 95 °C por 3-5 min e depois mantidas a 70 °C durante a aplicação das mesmas, utilizando-se bloco de aquecimento (*Chilling-Heat Block*) da Cole Parmer. Após a aplicação das amostras o gel era corrido à mesma potência da pré-corrída, ficando a temperatura da placa em torno de 50-55 °C durante a corrída. Padronizou-se paralisar a corrída quando a linha do corante (azul de bromofenol) alcançava a distância de aproximadamente 5 cm do nível do tampão inferior. Corria-se, então, metade das amostras em cada gel, repetindo-se em ambos as amostras os materiais SCA-6 e ICS-1, utilizados como padrões para, juntamente com o padrão de peso molecular, facilitar a comparação dos tamanhos dos fragmentos separados em géis diferentes.

Como a caracterização com cada microssatélite envolvia 172 genótipos, eram necessários dois géis para separar os fragmentos de cada reação.

Tabela 2. Especificações dos *primers* usados na caracterização dos genótipos

Nome do Marcador		Sequence de bases (5'-3')	Ta (°C)	Estrutura Repetida	Alelo Clonado (bp)	Cromossoma
mTcCIR1	f	GCA GGG CAG GCT CAG TGA	51	(CT) ₁₄	143	8
	r	TGG GCA ACC AGA AAA CGA T				
mTcCIR2	f	CAG GGA GCT GTG TTA TTG	51	(GA) ₃ N ₅ (AG) ₂ GG(AG) ₄	254	5
	r	AGT TAT TGT CGG CAA GGA GGA T				
mTcCIR4	f	CGA CTA AAA CCC AAA CCA	51	(TCTCTG) ₂ (TC) ₈	259	?
	r	AAT TAT TAG GCA ACC CGA ACT T				
mTcCIR6	f	TTC CCT CTA AAC TAC CCT A	46	(TG) ₇ (GA) ₁₃	231	6
	r	TAA AGC AAA GCA ATC TAA CAT A				
mTcCIR7	f	ATG CGA ATG ACA ACT GGT	51	(GA) ₁₁	160	7
	r	GCT TTC AGT CCT TTG CTT				
mTcCIR8	f	CTA GTT TCC CAT TTA CCA	46	(TC) ₅ TT(TC) ₁₇ TTT(CT) ₄	301	9
	r	TCC TCA GCA TTT TCT TTC				
mTcCIR9	f	ACC ATG CTT CCT CCT TCA	51	(CT) ₈ N ₁₅ (CT) ₅ N ₉ (TC) ₁₀	274	6
	r	ACA TTT ATA CCC CAA CCA				
mTcCIR10	f	ACA GAT GGC CTA CAC ACT	46	(TG) ₁₃	208	5
	r	CAA GCA AGC CTC ATA CTC				
mTcCIR11	f	TTT GGT GAT TAT TAG CAG	46	(TC) ₁₃	298	2
	r	GAT TCG ATT TGA TGT GAG				
mTcCIR12	f	TCT GAC CCC AAA CCT GTA	46	(CATA) ₄ N ₁₈ (TG) ₆	188	4
	r	ATT CCA GTT AAA GCA CAT				
mTcCIR13	f	CAG TCT AAC AAA GGT GAG	46	(AG) ₁₃	258	?
	r	TGC CCC ACT TGA CAA CTA				
mTcCIR17	f	AAG GAT GAA GGA TGT AAG	51	(GT) ₇ N ₄ (GA) ₁₂	271	4
	r	CCC ATA CGA GCT GTG AGT				
mTcCIR18	f	GAT AGC TAA GGG GAT TGA	51	(GA) ₁₂	345	4
	r	GGT AAT TCA ATC ATT TGA GGA TA				
mTcCIR19	f	CAC AAC CCG TGC TGA TTA	46	(CT) ₂₈	376	2
	r	GTT GTT GAG GTT GTT AGG AG				
mTcCIR25	f	CTT CGT AGT GAA TGT AGG	46	(CT) ₂₁	153	6
	r	TTA GGT AGG TAG GGT TAT CT				
mTcCIR26	f	GCA TTC ATC AAT ACA TTC	46	(TC) ₉ C(CT) ₄ TT(CT) ₁₁	298	8
	r	GCA CTC AAA GTT CAT ACT AC				
mTcCIR28	f	GAT CAA TCA GAA GCA AAC ACA	46	(TC) ₈	336	?
	r	TAA AGC AGC CTA CAA AGA AAA G				
mTcCIR32	f	GAC TTA CTC CCA TCC TAC	46	(CA) ₁₀	198	4
	r	TGA TTG GCA CAC TTT T				
mTcCIR33	f	TGG GTT GAA GAT TTG GT	51	(TG) ₁₁	285	4
	r	CAA CAA TGA AAA TAG GCA				
mTcCIR35	f	TTT CCT TGT ATT GAC CTA	46	(GT) ₁₁	235	9
	r	ATA TAA ACA CAC TTC AGA GAT				
mTcCIR40	f	AAT CCG ACA GTC TTT AAT C	51	(AC) ₁₅	286	3
	r	CCT AGG CCA GAG AAT TGA				
mTcCIR42	f	TTG CTG AAG TAT CTT TTG AC	46	(CA) ₂₁	232	5
	r	GCT CAA CCC CTA TTT G				
mTcCIR43	f	TCA TGA GAA TGC ATG TG	46	(TG) ₅ (TA)(GA) ₁₅	206	4
	r	CTG GAC ATG AAG AAG TTA T				
mTcCIR44	f	CCC ATC AAA AGT ATT AGA AG	51	(GT) ₁₀	178	2
	r	ATC AAG CAA TGG TCA AC				
mTcCIR49	f	GTA AAG CAC ATA TAC TAA ATG TCA	46	(TG) ₈	197	3
	r	TTT AAC CTT TGT AAG AAG TAT TCA				
mTcCIR54	f	AAC CTC TTG TCA CTG TA	46	(CA) ₁₅	165	1
	r	GAA GGC ATA CTT ACT ACT GT				
mTcCIR55	f	GAT ATT GTC CAA TAA TTT G	46	(CA) ₆ (GCACAC) ₅	234	7
	r	TTC CGC CTT GTT CTC				
mTcCIR57	f	TGT AGA TGT GAT TTT ATA GTT TG	46	(AC) ₁₃	253	4
	r	GGA GGG ATA AGA AGC AG				
mTcCIR58	f	TTT TTG GTG ATG GAA CTA T	51	(GT) ₄₀	266	9
	r	TGG TTA AGC AAC ACT AAA CT				
mTcCIR60	f	CGC TAC TAA CAA ACA TCA AA	51	(CT) ₇ (CA) ₂₀	207	2
	r	AGA GCA ACC ATC ACT AAT CA				

Ta: Temperatura de anelamento;

f: *forward*;r: *reverse*

Para determinar o comprimento dos fragmentos, utilizou-se o padrão de peso molecular *10 bp DNA Ladder* (Life Technologies, Cat. No. 10821-015), o qual consiste de 33 fragmentos de dsDNA, variando de 10 a 330 bp em múltiplos de 10 bp. Para os microssatélites mTcCIR-18, mTcCIR-19, mTcCIR-28 e mTcCIR-33, que apresentam alelos de comprimentos superiores a 330 bp, utilizou-se o padrão de peso molecular *25 bp DNA Ladder* (Life Technologies, Cat. No. 10597-011) que consiste de 18 fragmentos de dsDNA variando de 25 a 500 bp em múltiplos de 25 bp. Visando minimizar a interferência devido a deformações no gel (*sorrisos*), os padrões de peso molecular foram distribuídos a intervalos de 11 amostras, totalizando 8 trilhas com padrões em cada gel.

Os géis, ainda fixados em uma das placas foram mergulhados por 30 minutos em solução de Brometo de Etídio 0,5 µg/mL, lavados em seguida com água de torneira por 5 minutos e então fotografados sob luz ultravioleta (Eagle Eye II). As fotografias foram armazenadas como arquivo do tipo TIF para posterior análise do tamanho dos fragmentos com o programa ONE-Dscan (ver. 1.0).

2.4. Análises Estatísticas

2.4.1. Modelo de mutação de microssatélites

Classicamente, dois modelos extremos de mutação tem sido considerados para locos microssatélites: o modelo de *alelos infinitos* (IAM, Kimura & Crow, 1964) e o modelo de *mutações em etapas* ou *stepwise mutation model* (SMM, Kimura & Ohta, 1978). O modelo SMM baseia-se na mutação de alelos microssatélites pela perda ou ganho de uma de uma unidade simples de repetição, implicando que a mutação pode produzir estados alélicos já presentes na população. Sob o modelo IAM, uma mutação envolve qualquer número de unidades repetidas e sempre resulta em um estado alélico novo na população. A questão sobre qual modelo aplicar torna-se importante porquê a estimação de vários parâmetros populacionais (diferenciação genética, migrantes por geração, distâncias genéticas, etc) é dependente do modelo de mutação assumida para o marcador em estudo. No

caso de microssatélites essa dependência pode se tornar crítica, dado que a sensibilidade do modelo de mutação aumenta com a taxa de mutação, normalmente muito alta para microssatélites. Entretanto, não obstante os crescentes esforços no sentido de desenvolver estatísticas adequadas ao estudo de microssatélites sob o modelo SMM, vários autores vêm com reservas a aplicação das mesmas. A menos que se esteja seguro de que o sistema em estudo siga estritamente o modelo SMM e que a taxa de mutação não possa ser negligenciada quando comparada a outras forças como migração e deriva genética (Balloux *et al.*, 1999). Desse modo, salvo quando devidamente explicitado, todas as análises desenvolvidas no presente estudo foram fundamentadas no modelo de mutação de infinitos alelos (IAM).

2.4.2. Caracterização individual

Os diferentes alelos detectados por cada loco microssatélite foram codificados com o respectivo número de pares de base estimado a partir dos géis com o programa ONE-Dscan (ver. 1.0). Com a matriz obtida com os genótipos individuais de cada clone, utilizou-se o programa Popgene (ver. 1.32) para estimar os seguintes parâmetros:

Frequência alélica: estima a frequência gênica de cada loco a partir do número de cada alelo dividido pelo total de alelos encontrados no mesmo loco;

Número de alelos: conta o número de alelos com frequência maior que zero;

Heterozigosidade observada: Estima a proporção observada de heterozigotos em um dado loco;

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

onde P_{ii} é a frequência observada de genótipos homozigotos para o i -ésimo alelo.

Heterozigosidade esperada: Estima a proporção esperada de heterozigotos em um dado loco sob equilíbrio de Hardy-Weinberg.

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

onde p_i é a frequência do i -ésimo alelo de dado loco.

Heterozigosidade esperada não viesada: Estima a proporção esperada de heterozigotos em um dado loco sob equilíbrio de Hardy-Weinberg, corrigida para o tamanho das amostras, segundo Nei (1978);

$$H_{nb} = 2m(1 - \sum p_i^2) / (2m - 1)$$

onde p_i é a frequência do i -ésimo alelo de dado loco e m é o número de indivíduos da amostra.

Número efetivo de alelos: Estima o recíproco da homozigosidade esperada (Kimura & Crow, 1964);

$$n_e = 1 / (1 - H_e)$$

Índice de Fixação: Estima F_{IS} como uma medida do déficit ou excesso de heterozigosidade segundo Wright (1978):

$$F_{IS} = (H_e - H_o) / H_e$$

A partir das frequências de cada alelo calculou-se (manualmente) ainda o poder de discriminação (PD) dos locos (Jones, 1972), a probabilidade de identidade (PI) (Paetkau *et al.* 1995) e o conteúdo de informação polimórfica (PIC) (Weber, 1990), conforme as relações a seguir:

$$PD = 1 - \sum (P_x)^2 \quad PI = \sum p_i^4 + \sum_{i=1}^{l=n-1} \sum_{j=i+1}^n (2p_i p_j)^2 \quad PIC = 1 - \left(\sum_{i=1}^n p_i^2 \right) - \sum_{i=1}^{l=n-1} \sum_{j=i+1}^n 2p_i^2 p_j^2$$

Onde P_x representa a frequência de cada genótipo, p_i e p_j são as frequências do i -ésimo e j -ésimo alelos em uma dada amostra.

2.4.3. Teste de Neutralidade Seletiva dos locos

A neutralidade seletiva dos locos microssatélites foi avaliada pelo mesmo programa (Popgene, ver. 1.31) através do teste de Watterson-Ewens (Hartl & Clark, 1989), seguindo o algoritmo de Manly, 1985) para simular a distribuição da estatística F (homozigosidade esperada – $\sum p_i^2$), após comparar esta estatística com as de configurações genotípicas obtidas em 1000 hipotética para uma população obedecendo o modelo de mutação de infinito alelos (IAM). Para cada amostra obtém um valor de F , o qual é comparado ao F calculado a partir das frequências observadas da amostra. Se este última cair

dentro do intervalo de confiança gerado pelas simulações de F, aceita-se a hipótese de neutralidade seletiva para o loco em estudo.

2.4.4. Desequilíbrio genotípico

Utilizou-se teste exato de permutação (Guo & Thompson, 1992), executado pelo Programa Fstat, ver. 2.9.3 (Goudet, 2001), para verificar a significância da associação entre genótipos de cada par de locos. A estatística usada para testar as tabelas geradas a cada permutação é a razão de verossimilhança ($G = \sum \sum x_{ijkl} \ln(x_{ijkl})$), onde x_{ijkl} representa o número de indivíduos na amostra com genótipo ij no primeiro loco e kl no segundo loco. O P-valor do teste é obtido após a associação de dois locos ao acaso um número aleatório de vezes e a estatística G calculada no conjunto de dados aleatorizados é comparado com as associações observadas nos dados. O P-valor estima a proporção de vezes que a estatística dos dados simulados é maior ou igual ao observado. O teste foi aplicado na amostra total após 10.000 permutações, aplicando-se a correção de Bonferroni ajustando-se o nível nominal de significância (5%) ao número de testes aplicados (406, combinação de 29 locos dois a dois).

2.4.5. Heterozigosidade Individual dos genótipos

A heterozigosidade de cada clone foi avaliada considerando-se a identidade dos alelos (heterozigosidade observada) e o tamanho dos alelos pela medida d^2 -média, proposta por Coulson *et al* (1998), para estimar a diversidade individual sob o modelo de mutação em etapas (SMM). Ela representa a média do quadrado da diferença entre os comprimento de dois alelos de cada locos, do mesmo indivíduo, ponderada com todos os locos, conforme a relação abaixo:

$$d^2\text{-média} = (1/n) \sum^n (i_a - i_b)^2$$

onde i_a e i_b são os comprimentos dos alelos em número de repetições ou pares de base dos alelos a cada loco i e n é o número total de locos para os quais os indivíduos foi genotipado. Essa medida estima na verdade a distância genética entre os gametas que formaram o indivíduo e matematicamente, é

uma medida da variância do comprimento médio dos alelos dentro de um indivíduo.

2.4.6. Relação Genética entre os clones

A relação genética entre os genótipos individuais foi avaliada através de uma matriz de similaridade construída utilizando-se o índice de similaridade para variáveis multicategóricas ($S_i = C/(C+D)$), através do program Genes versão Windows 2001 (Cruz, 2001). O referido índice estimou a similaridade entre os genótipos a partir dos escores 0, 1 e 2, para alelo ausente, em heterozigose ou homozigose respectivamente. Com o complemento desse índice construiu-se dendrograma circular baseado no método de agrupamento hierárquico UPGMA usando o programa MEGA, ver 2,1 (Kumar *et al*, 2001). Construiu-se ainda diagrama de dispersão bidimensional dos clones através do program GENETIX, ver. 4.02, aplicando análise fatorial de correspondência a partir da matriz de escores 0, 1 e 2 de cada genótipo.

2.5. Variabilidade genética das subpopulações

Desvios do equilíbrio de Hardy-Weinberg (EHW), dentro de cada população e para a amostra inteira foram testados através do teste exato de permutação (Guo & Thompson, 1992), computado através do programa Fstat, ver. 2.9.3 (Goudet, 2001), utilizando os coeficientes de fixação F_{IS} (f de Weir & Cockerham, 1984) para o teste dentro das subpopulações e o coeficiente de fixação F_{IT} (F de Weir & Cockerhan, 1984) para o teste global. Para o teste dentro das amostras, os alelos foram permutados entre os indivíduos dentro da própria amostra; para o teste global os alelos foram permutados entre amostras. Para ambos os testes aplicou-se 10.000 permutações e o nível nominal ajustado (5 %) pela correção de Bonferroni foi de 0,00016.

2.5.1. Variabilidade genética dentro das subpopulações

A variabilidade de cada subpopulação foi caracterizada pelas freqüências alélicas, número de alelos raros, número de alelos privados, percentagem de locos polimórficos, número de alelos por loco, número total de

alelos, número efetivo de alelos, riqueza alélica padronizada, heterozigosidades observada e a diversidade não viesada ou heterozigosidade esperada no equilíbrio de Hardy-Weinberg. Estimou-se ainda o coeficiente de endogamia dentro de cada amostra, a partir do qual estimou-se a taxa de cruzamento.

A frequência de cada loco foi obtida através do programa Popgene, ver. 1.32, pela divisão do número total de cada alelo em cada população pelo número total de alelos detectados naquele loco. Alelos com frequência menor que 0,05 foi considerado raro. E privado ou exclusivo quando o alelo ocorria em apenas uma população, independente de sua frequência. Um loco foi considerado polimórfico quando a frequência do alelo mais comum é menor ou igual 0,95.

A riqueza alélica padronizada, introduzida por El-Mousadik & Petit (1996), é uma medida do número de alelos independente do tamanho da amostra permitindo que a mesma possa ser comparada entre amostras de tamanho diferentes. O princípio de sua padronização baseia-se na técnica de rarefação introduzida por Hurlbert (1971) em estudos ecológicos. O princípio é estimar o número esperado de alelos em uma sub-amostra de $2n$, dado que $2N$ tenham sido amostrados ($N \geq n$). Normalmente o n é fixado como o menor número de indivíduos genotipados para um loco em uma amostra. A riqueza alélica é então calculada pela seguinte combinação:

$$R(g) = \sum \left(1 - \frac{\binom{2N-N_i}{2n}}{\binom{2N}{2n}} \right)$$

onde N_i é o número de alelos do tipo i entre os $2N$. O termo sob a soma corresponde à probabilidade de amostrar o alelo i pelo menos uma vez em uma amostra de $2n$. Quando $2n > (2N - N_i)$ a razão é indefinida, mas a probabilidade de amostrar o alelo é considerada igual a 1. Esta estatística foi computada pelo programa Fstat, ver. 2.9.3 (Goudet, 2001).

2.5.2. Variabilidade genética entre as subpopulações

Em populações naturais, as freqüências genotípicas em cada subpopulação não necessariamente encontram-se nas proporções de Hardy-Weinberg, em função de forças não seletivas como a deriva genética e a endogamia. Ambos os fatores agem reduzindo a quantidade de heterozigosidade e aumentando a homozigosidade, efeitos que tendem a se acentuar à medida que a população se torne estruturada, quando os cruzamentos e a dispersão ocorrem somente nas vizinhanças de cada indivíduo. Wright (1951, 1969) propôs que os desvios das freqüências genotípicas em uma população subdividida podem ser medidos em termos de três parâmetros, F_{IS} , F_{IT} e F_{ST} , denominados *índices de fixação*, assim relacionados:

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

Na definição de Wright F_{IS} e F_{IT} estimam a correlação da união de gametas em relação a subpopulação e em relação à população total, respectivamente, enquanto F_{ST} é a correlação entre dois gametas retirados ao acaso de cada subpopulação.

A idéia básica na abordagem de Wright assume que as populações e subpopulações analisadas derivaram de uma população ancestral comum ao mesmo tempo e que todas as populações são igualmente relacionadas, quer haja ou não migração entre as populações. Ou seja, as populações são consideradas como uma amostra randômica de um conjunto de infinitas populações igualmente relacionadas. Nesse contexto, Weir & Cocherhan (1984) desenvolveram um sistema baseado na análise de variâncias que permite a um só tempo estimar e testar o nível de significância dos diversos índices de fixação F .

Entretanto, considerando que a maioria das populações naturais apresenta relações filogenéticas que não se adequam ao modelo proposto por Wright, Nei (1977) redefiniu os índices de fixação sem usar o conceito de correlação na união de gametas. O autor mostrou que a relação entre os diferentes índices F se mantêm em qualquer situação a despeito das relações filogenéticas, padrão de migração, número de alelos, havendo ou não seleção.

Nessa abordagem, todos os índices de fixação podem ser definidos a partir das heterozigosidades observada e esperada para a população investigada, como se verá a seguir.

No presente estudo adotou-se as duas abordagens (Wright, 1951 e Nei, 1977) para avaliar a diferenciação entre as subpopulações e grupos de subpopulações. Para tanto empregou-se o program Fstat, ver.2.9.3 (Goudet, 2001), para calcular os índices de fixação de Wright (1969), estimados pelas estatísticas globais de Nei (1987) e Weir & Cockerhan (1984).

Estatísticas de Nei

Foram estimados para cada loco e para o conjunto das amostras as estatísticas a seguir definidas:

H_o : proporção observada de indivíduos heterozigotos

$$H_o = 1 - \sum \sum P_{kij}/np,$$

Onde P_{kij} é a frequência do genótipo A_iA_i na amostra k e np é o número de amostras.

\bar{H}_s : é a heterozigosidade esperada (não viesada) dentro da amostra, nesse contexto, freqüentemente chamada de *diversidade genética*.

$$\bar{H}_s = \bar{n}[1 - \sum_i \bar{p}_i^2 - H_o/2\bar{n}]/(\bar{n}-1)$$

onde $\bar{p}_i^2 = \sum_k p_{ki}^2 / np$, np é o número de amostras e \bar{n} é a média harmônica de n_k

H_T : é a heterozigosidade total esperada, ou diversidade total do conjunto.

$$H_T = 1 - \sum_i \bar{p}_i^2 + \bar{H}_s/2\bar{n}np$$

onde $\bar{p}_i = \sum_k p_{ki} / np$,

D_{ST} : estima a quantidade de diversidade entre as amostras.

$$D_{ST} = H_T - \bar{H}_s$$

D_{ST}' : diversidade entre as amostras independente do número de amostras.

$$D_{ST}' = \frac{n}{np-1} D_{ST}$$

H_T' : diversidade total independente do tamanho da amostra

$$H_T' = \bar{H}_S + D_{ST}'$$

G_{ST} : coeficiente de diversidade relativa entre as populações – estimador de F_{ST} .

$$G_{ST} = (H_T - \bar{H}_S)/H_T = D_{ST}/H_T$$

G_{ST}' : G_{ST} corrigido independente do número de amostras.

$$G_{ST}' = (H_T' - \bar{H}_S)/H_T' = D_{ST}'/H_T'$$

G_{IS} : é um estimador do índice de fixação F_{IS} .

$$G_{IS} = (\bar{H}_S - H_O)/\bar{H}_S = 1 - H_O/\bar{H}_S$$

Estatísticas de Weir & Cockerhan:

Os estimadores F , θ e f , correspondentes aos índices de fixação de Wright, foram calculados pelo programa Fstat, ver.2.9.3 (Goudet, 2001), através da partição da variância total das frequências alélicas nos componentes relativos a variância entre as populações (σ^2_P), variância entre os indivíduos dentro da população (σ^2_I) e variância dentro dos indivíduos (σ^2_G), as quais são estimadas a partir dos respectivos quadrados médios observados para populações (QMP), indivíduos (QMI) e alelos (MSG):

$$\sigma^2_G = QMG \quad \sigma^2_I = \frac{1}{2}(QMI - QMG) \quad \sigma^2_P = (1/2n_c)(QMP - QMI)$$

obtendo-se as medidas de diferenciação a partir das relações entre os componentes da variância:

$$F = \frac{\sigma^2_P + \sigma^2_I}{\sigma^2_P + \sigma^2_I + \sigma^2_G} \quad \theta = \frac{\sigma^2_P}{\sigma^2_P + \sigma^2_I + \sigma^2_G} \quad f = \frac{F - \theta}{1 - \theta}$$

A probabilidade de estas estatísticas diferirem de zero foi determinada após 10.000 permutações e intervalos de confiança de 95%, associados a cada índice foram gerados após 10.000 bootstrap (Goudet, 2001). Para f (F_{IS}) os alelos são permutados entre indivíduos dentro das amostras; para F (F_{IT}) os alelos são permutados entre amostras. Para θ (F_{ST}), permutam-se os alelos entre amostras, quando F_{IS} não difere de zero, ou permutam-se os genótipos quando F_{IS} é significativo.

O programa calculou ainda o coeficiente de parentesco (Hamilton's, 1971) - $R_{EL} = 2\theta / (1+F)$ - estimado segundo Queller & Goodnight, 1989). Essa medida estima o parentesco médio dos indivíduos dentro das amostras em relação ao todo.

As estimativas de F_{ST} (θ) entre pares de amostras, bem como a partição da variação entre subpopulações, entre grupos de subpopulações e entre subpopulações dentro dos grupos, foram determinadas pelo programa Arlequin ver. 2000, através da análise de variância hierarquizada de Weir & Cocherhan (1984), conforme descrito acima.

A contribuição relativa de cada amostra para a variabilidade total das amostras em conjunto foi estimada para os indicadores de riqueza alélica padronizada ($R_{(g)}$) e diversidade genética (H_{nb}), conforme Petit *et al.* (1998). O procedimento consiste em se estimar os indicadores de variabilidade para o conjunto de todas as populações, exceto aquela que se quer avaliar a contribuição relativa. Conforme as relações seguintes:

$$C_{T(k)} = (X_T - X_{T/k})/X_T,$$

onde :

$C_{T(k)}$ - contribuição relativa da população k ;

X_T - representa a diversidade total ou a riqueza alélica total;

$X_{T/k}$ - representa a diversidade ou riqueza alélica total no conjunto de todas as populações exceto a população k ;

Adotando o mesmo procedimento, essa contribuição pode ser particionada em dois componentes:

$C_{S(k)}$ - contribuição da população k à diversidade total, ou à riqueza alélica total devido a sua própria diversidade;

$$C_{S(k)} = (X_k - X_{S/k})/n = (X_k - X_S)/(n-1),$$

onde:

X_k - diversidade ou riqueza alélica da população k;

X_S - diversidade ou riqueza alélica média de todas as populações.

$C_{D(k)}$ - contribuição da população k à diversidade total ou à riqueza alélica total devido a sua divergência, obtido pela diferença:

$$C_{D(k)} = C_{T(k)} - C_{S(k)}$$

2.6. Distâncias genéticas entre as subpopulações

As distâncias genéticas entre as subpopulações foram avaliadas pelos estimadores de distância D de Nei (1987), D_L de Reynolds *et al.* (1983) e $\delta\mu^2$ de Goldstein (1995), conforme as relações:

Distância de Nei

$$D = \ln I,$$

onde I é a estimativa da identidade genética entre as populações

$$I = J_{xy} / (J_x \cdot J_y)^{1/2},$$

em que

$J_x = \sum x_i^2$ probabilidade de identidade de dois alelos na população x

$J_y = \sum x_i^2$ probabilidade de identidade de dois alelos na população y

$J_{xy} = \sum x_i y_i$ probabilidade de identidade de dois alelos nas duas populações

Distância de Reynolds

$$D_L = -\ln(1-F_{ST})$$

Distância de Goldstein

$$\delta\mu^2 = \sum (\mu_{xk} - \mu_{yk})^2/L,$$

onde

$\mu_{xk} = \sum x_{ik}$ - média de unidades repetidas no loco k da população x,

$\mu_{yk} = \sum x_{jk}$ - média de unidades repetidas no loco k da população y.

Todas as distâncias foram calculadas usando o programa MICROSAT (Goldstein, 1995), obtendo-se a matriz de consenso após 100 re-Amostragens bootstraps. As matrizes geradas para cada uma das distâncias estimadas foram utilizadas para a construção de dendrogramas pelo método UPGMA. Para avaliar a consistência da topologia dos dendrogramas, utilizou-se as matrizes (100) de cada distância, obtidas pela reamostragem dos locos (bootstrap), através da rotinas NEIGHBOR e CONSENSE do pacote de programas PHYLIP 3.5c de Felsenstein (1993) .

3. RESULTADOS E DISCUSSÃO

3.1. Nível de polimorfismo detectado na amostra total

A diversidade alélica medida nos 172 acessos, caracterizados com os 30 microssatélites, revelou um perfil de loco simples, com 29 locos polimórficos e apenas um loco monomórfico (mTcCIR-28). O loco mTcCIR-28 corresponde ao *primer* APMC 526, utilizado por Sereno (2001), que também se mostrou monomórfico na caracterização de 94 clones CAB. Um exemplo do polimorfismo alélico é mostrado na Figura 2 para o *primer* mTcCIR-12. Dentre os locos polimórficos detectou-se um total de 186 alelos, variando de 2 alelos, para o *primer* mTcCIR-02, a 11 alelos, para o *primer* mTcCIR-60, com uma média de 6,4 alelos por loco, enquanto que o tamanho dos alelos variou de 129 bp (mTcCIR-01) a 454 bp (mTcCIR-18) (Tabela 3 e Figura 3). Esses resultados são comparáveis aos observados por Lanaud et al. (1999a), no que diz respeito ao tamanho dos alelos. Quanto ao número médio de alelos, mostra-se superior ao obtido por aqueles autores (5,6 alelos/loco), mas inferior ao detectado por Motamayor et al. (2002) para cacaueiros do grupo *Forastero* (8,7 alelos/loco). Há que se considerar, porém, que os autores citados não utilizaram os mesmos *primers*, sendo que o estudo de Motamayor et al. (2002) envolveu uma área de

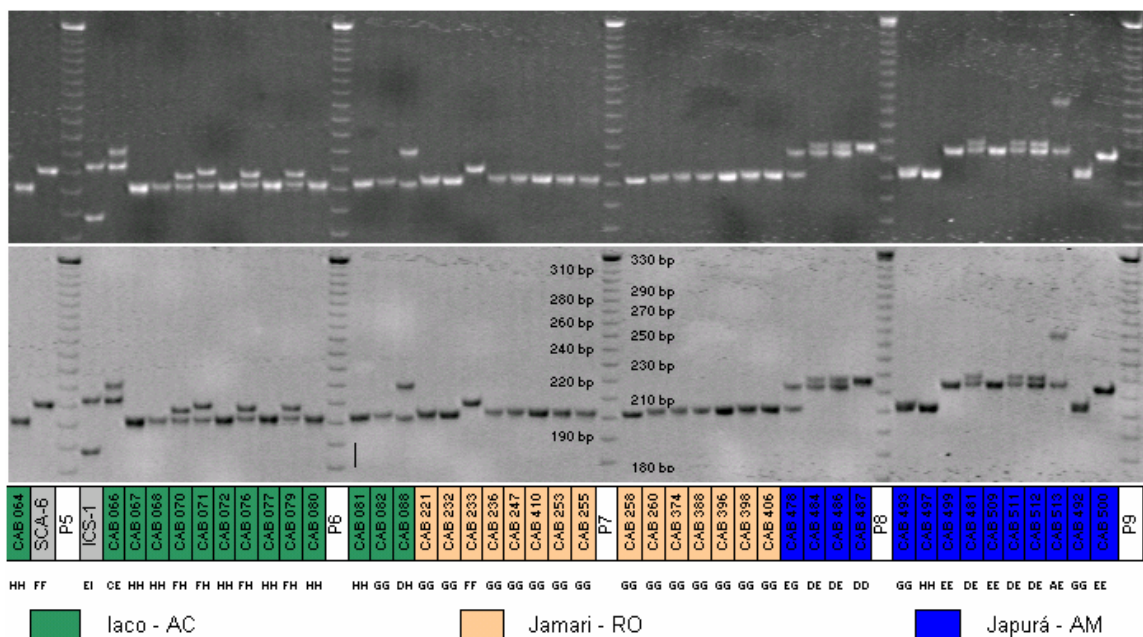


Figura 2. Exemplo de polimorfismo detectado pelo *primer* mTcCIR-12, mostrando diferentes acessos de três subpopulações, em fotografia original (superior) e negativo (inferior). As letras abaixo dos acessos codificam alelos detectados no gel. P5-P9 correspondem às trilhas com marcadores de peso molecular (10 pb).

Tabela 3. Alelos identificados (bp) na amplificação de 30 microssatélites e suas respectivas freqüências (entre parênteses) observadas em 172 acessos de cacauzeiros.

Nº Loco	Cromos-soma	Alelo Clonado (bp)	Alelos Identificados										Nº Alelos				
MtCcir 01	8	143	129	132	134	142	145	146	148								7
			(0,006)	(0,407)	(0,012)	(0,456)	(0,020)	(0,038)	(0,061)								
MtCcir 02	5	254	238	252													2
			(0,497)	(0,503)													
MtCcir 04	?	259	250	262	270												3
			(0,279)	(0,151)	(0,570)												
MtCcir 06	6	231	196	198	202	205	210	214									6
			(0,138)	(0,600)	(0,041)	(0,103)	(0,106)	(0,012)									
MtCcir 07	7	160	154	159	160	162	164	165	167	169							8
			(0,006)	(0,186)	(0,172)	(0,378)	(0,206)	(0,035)	(0,009)	(0,009)							
MtCcir 08	9	301	277	282	285	289	302										5
			(0,003)	(0,110)	(0,628)	(0,160)	(0,099)										
MtCcir 09	6	274	256	262	269	273	278	282	284	287	295						9
			(0,003)	(0,003)	(0,006)	(0,067)	(0,328)	(0,323)	(0,250)	(0,017)	(0,003)						
MtCcir 10	5	208	199	202	203	206	209	212									6
			(0,020)	(0,148)	(0,302)	(0,378)	(0,140)	(0,012)									
MtCcir 11	2	298	285	287	290	296	304	306	311								7
			(0,018)	(0,023)	(0,404)	(0,213)	(0,199)	(0,035)	(0,108)								
MtCcir 12	4	188	186	200	203	209	211	214	220	224	252						9
			(0,012)	(0,292)	(0,234)	(0,096)	(0,181)	(0,091)	(0,009)	(0,012)	(0,073)						
MtCcir 13	?	258	246	254	256	258											4
			(0,196)	(0,731)	(0,041)	(0,032)											
MtCcir 17	4	271	265	266	268	277											4
			(0,126)	(0,567)	(0,301)	(0,006)											
MtCcir 18	4	345	333	337	345	352	375	454									6
			(0,161)	(0,063)	(0,759)	(0,009)	(0,003)	(0,006)									
MtCcir 19	2	376	379	384	387	400	407										5
			(0,131)	(0,192)	(0,282)	(0,317)	(0,078)										
MtCcir 25	6	153	172	175	179	180	184	192	198	206							8
			(0,024)	(0,103)	(0,144)	(0,082)	(0,012)	(0,359)	(0,188)	(0,088)							
MtCcir 26	8	298	266	283	288	290	292	297	299	304	310						9
			(0,003)	(0,141)	(0,012)	(0,050)	(0,285)	(0,324)	(0,174)	(0,009)	(0,003)						
MtCcir 28	?	336	336														1
			(1,000)														
MtCcir 32	4	198	196	200	205												3
			(0,018)	(0,623)	(0,358)												
MtCcir 33	4	285	287	293	295	302	306	312	318	325	327	367					10
			(0,139)	(0,012)	(0,172)	(0,163)	(0,139)	(0,078)	(0,036)	(0,012)	(0,151)	(0,099)					
MtCcir 35	9	235	220	230	234												3
			(0,317)	(0,269)	(0,414)												
MtCcir 40	3	286	259	266	274	277	281	283	297	298							8
			(0,012)	(0,191)	(0,118)	(0,374)	(0,035)	(0,144)	(0,032)	(0,094)							
MtCcir 42	5	232	202	212	216	218	223	228	231	235							8
			(0,350)	(0,018)	(0,085)	(0,047)	(0,162)	(0,106)	(0,141)	(0,091)							
MtCcir 43	4	206	191	197	199	201	203	209	211	213							8
			(0,006)	(0,186)	(0,112)	(0,249)	(0,071)	(0,163)	(0,198)	(0,015)							
MtCcir 44	2	178	170	172	181	182	183	193									6
			(0,447)	(0,009)	(0,300)	(0,100)	(0,071)	(0,074)									
MtCcir 49	3	197	198	204	205	206	209	215									6
			(0,024)	(0,018)	(0,732)	(0,140)	(0,009)	(0,077)									
MtCcir 54	1	165	144	146	150	154	156	158	162	166							8
			(0,071)	(0,029)	(0,170)	(0,016)	(0,647)	(0,013)	(0,016)	(0,038)							
MtCcir 55	7	234	223	227	233												3
			(0,023)	(0,361)	(0,616)												
MtCcir 57	4	253	242	247	252	254	256	262									6
			(0,018)	(0,006)	(0,181)	(0,678)	(0,108)	(0,009)									
MtCcir 58	9	266	232	238	242	252	257	264	267	309							8
			(0,015)	(0,076)	(0,085)	(0,021)	(0,379)	(0,321)	(0,009)	(0,094)							
MtCcir 60	2	207	186	191	194	196	200	206	208	210	212	214	224				11
			(0,349)	(0,157)	(0,047)	(0,052)	(0,003)	(0,047)	(0,003)	(0,093)	(0,172)	(0,070)	(0,009)				

dispersão muito maior do que a aqui considerada, incluindo acessos de grupos morfogeográficos mais contrastantes, originários de diversos países. Quando se compara apenas os locos comuns aos estudos de Lanaud *et al.* (1999a) e Sereno (2001), obteve-se uma média de 5,8 alelos/loco, contra 6,4 e 4,2 alelos/loco dos respectivos autores. O resultado mais discrepante foi observado no loco mTcCIR-07, onde foram detectados 8 alelos, contra apenas 3 alelos detectados por Sereno (2001). Considerando que muitas populações caracterizadas são comuns aos dois estudos, o menor número de alelos por loco relatado por Sereno (2001) provavelmente resulta da menor resolução dos géis empregados pelo autor (poliacrilamida 6%, nativo) na separação das bandas (alelos). Por outro lado, os resultados ratificam o elevado polimorfismo de locos microssatélites comparativamente a outras técnicas de caracterização molecular. Caracterizações com nove isoenzimas (Lanaud, 1987) e 31 sondas de RFLP (Laurent *et al.*, 1993) em populações maiores (300 e 180 genótipos, respectivamente) revelaram 3,3 e 2,4 alelos por loco, com um máximo de 5 e 4 alelos em cada caso.

A Figura 3 mostra que a amplitude de variação no tamanho dos alelos dos diferentes microssatélites analisados permite que se compatibilize a

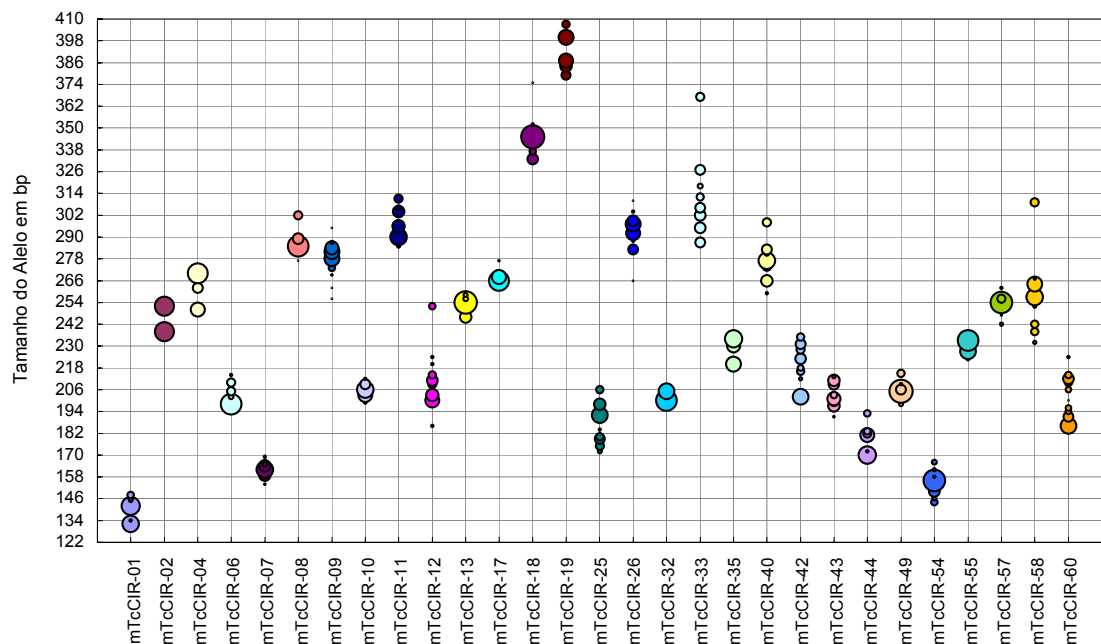


Figura 3. Perfil de distribuição dos alelos dos diferentes microssatélites, detectados nos 172 clones de cacau, em função do número de pares de bases e da frequência (diâmetro dos círculos) de cada alelo.

amplificação de conjuntos de várias combinações de *primers* em uma reação múltipla de PCR (*multiplex*), que produzam alelos de tamanhos tais que não se sobreponham, de modo que possam migrar em uma mesma canaleta do gel. Tal possibilidade poderia ser maximizada com marcação fluorescente em gel automatizado, o que resultaria em considerável economia de tempo e de recursos na caracterização de germoplasmas. A Figura 3 revela ainda que a maioria dos locos analisados apresenta um número potencialmente maior de alelos que os detectados no presente estudo. Isto fica mais claro na Figura 4, onde é mostrado que 4 dos locos mais polimórficos apresentam diferenças irregulares no tamanho dos alelos detectados. Os modelos de mutação geralmente aceitos para a evolução de microssatélites (Eisen, 1999) sugerem a possibilidade de que os pontos compreendidos nos intervalos irregulares correspondam a potenciais alelos passíveis de detecção analisando-se um número maior de indivíduos dentro das mesmas populações.

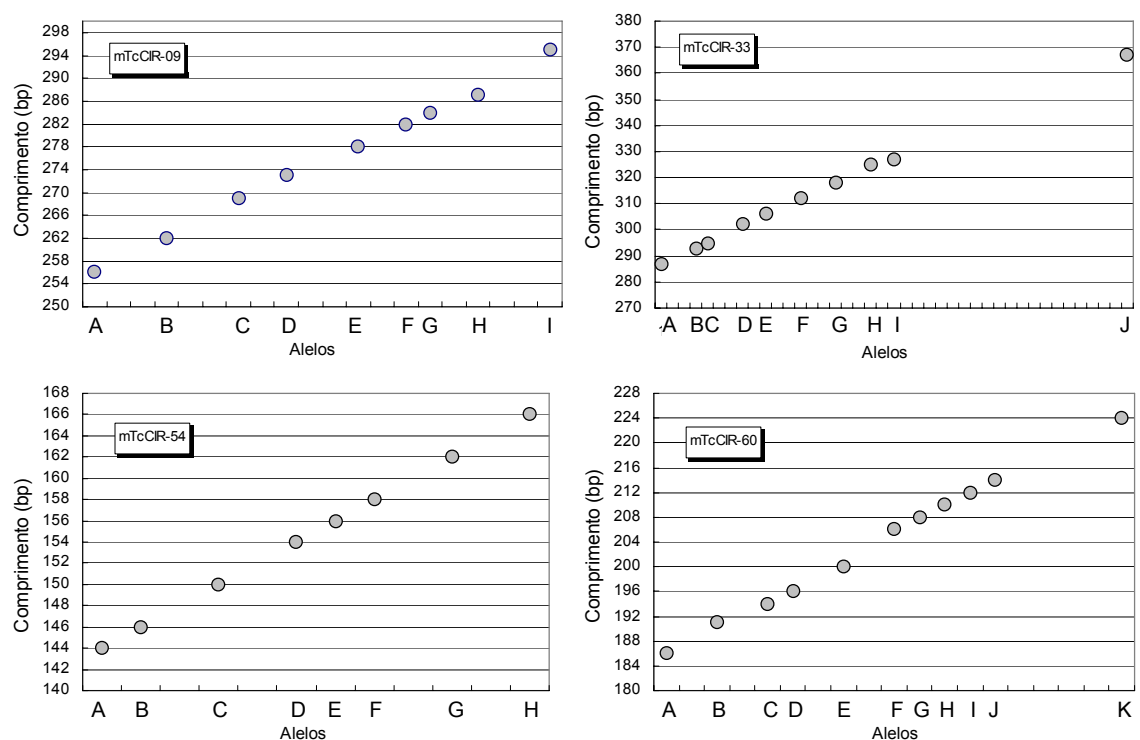


Figura 4. Exemplo da distribuição de alelos em quatro locos polimórficos, mostrando diferenças irregulares no comprimento dos fragmentos (alelos) identificados em cada um dos microssatélites. As letras nas abcissas codificam os alelos detectados em cada loco.

Os valores da diversidade (H_{nb}) variaram de 0,396 (mTcCIR-18) a 0,868 (mTcCIR-33), com uma diversidade média de 0,653 (Tabela 4). Observa-se ainda que o grau de diversidade não guarda uma relação muito estreita com o número de alelos de cada loco ($R^2 = 0,489$), refletindo a grande variação de frequência dos alelos dentro de cada loco, que, por sua vez determina os valores de heterozigidade de populações em equilíbrio de Hardy-Weinberg.

Tabela 4. Caracterização da diversidade dos microssatélites identificados nos 172 acessos analisados.

Loco	Nº Alelos	N_e	H_e	H_{nb}	H_o	F_{IS}
mTcCIR-01	7	2,64	0,6203	0,6221	0,2558	0,590
mTcCIR-02	2	2,00	0,5000	0,5016	0,5064	<u>-0,010</u>
mTcCIR-04	3	2,35	0,5746	0,5763	0,8140	<u>-0,414</u>
mTcCIR-06	6	2,48	0,5972	0,5990	0,2588	0,569
mTcCIR-07	8	3,99	0,7492	0,7513	0,2151	0,714
mTcCIR-08	5	2,26	0,5582	0,5598	0,1337	0,762
mTcCIR-09	9	3,58	0,7206	0,7227	0,2267	0,687
mTcCIR-10	6	3,62	0,7238	0,7259	0,1221	0,832
mTcCIR-11	7	3,82	0,7283	0,7405	0,2749	0,629
mTcCIR-12	9	5,10	0,8037	0,8061	0,4386	0,457
mTcCIR-13	4	1,74	0,4246	0,4258	0,1462	0,657
mTcCIR-17	4	2,34	0,5717	0,5734	0,0351	0,939
mTcCIR-18	6	1,65	0,3942	0,3963	0,2679	0,323
mTcCIR-19	5	4,16	0,7600	0,7622	0,4360	0,429
mTcCIR-25	8	4,74	0,7892	0,7915	0,4000	0,495
mTcCIR-26	9	4,18	0,7612	0,7634	0,2824	0,631
mTcCIR-32	3	1,94	0,4825	0,4839	0,2651	0,453
mTcCIR-33	10	7,40	0,8654	0,8680	0,4277	0,508
mTcCIR-35	3	2,90	0,6557	0,6577	0,3077	0,533
mTcCIR-40	8	4,50	0,7780	0,7803	0,1353	0,827
mTcCIR-42	8	5,05	0,8021	0,8044	0,1765	0,781
mTcCIR-43	8	5,55	0,8198	0,8222	0,3136	0,619
mTcCIR-44	6	3,22	0,6897	0,6917	0,3294	0,525
mTcCIR-49	6	1,78	0,4374	0,4388	0,1548	0,648
mTcCIR-54	8	2,20	0,5440	0,5458	0,1795	0,672
mTcCIR-55	3	1,96	0,4893	0,4909	0,2903	0,409
mTcCIR-57	6	1,98	0,4948	0,4963	0,1287	0,741
mTcCIR-58	8	3,71	0,7306	0,7327	0,1529	0,792
mTcCIR-60	11	5,08	0,8036	0,8059	0,2791	0,654
Média	6,41	3,38	0,6507	0,6530	0,2743	0,5673
s.d.	2,35	1,4309	0,1385	0,1391	0,1529	0,2644

Obs: n_e - número efetivo de alelos (Kimura & Crow, 1964)
 H_e - heterozigidade esperada, calculada com viés
 H_{nb} - heterozigidade esperada sem viés (Nei, 1978)
 H_o - heterozigidade observada
 F_{IS} - Índice de fixação de Wright

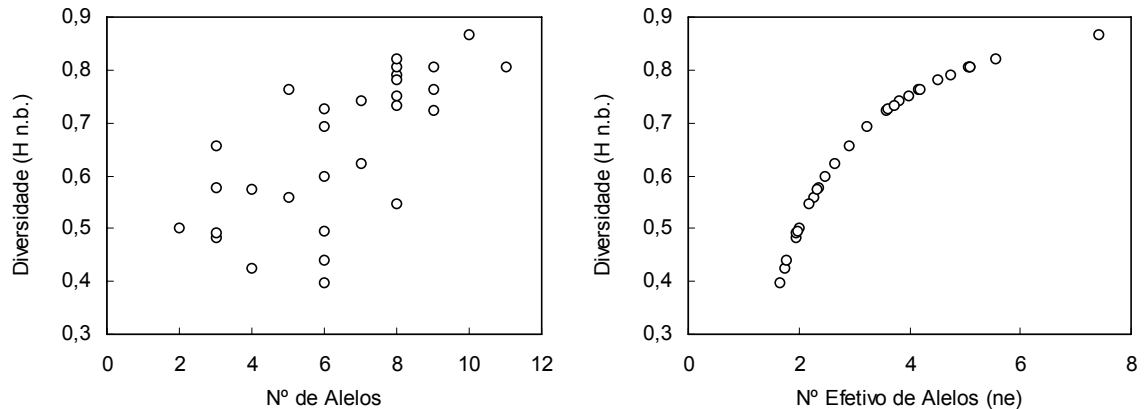


Figura 5. Diversidade (Hn.b.) dos diferentes microssatélites em função do número de alelos e do número efetivo (ne) de alelos de cada loco.

Desse modo, a diversidade é melhor refletida pelo número efetivo de alelos (ne), que representa o número de alelos igualmente freqüentes em uma população ideal, requerido para produzir a mesma homozigidade ou heterozigidade (esperadas) na população real (Hartl & Clark, 1989), conforme se observa na Figura 5.

A heterozigidade observada foi relativamente baixa na maioria dos locos analisados, variando de 0,035 (mTcCIR-17) a 0,814 (mTcCIR-04), com um valor médio de 0,274. O coeficiente de fixação F_{IS} mostra que à exceção dos locos mTcCIR-02 e mTcCIR-04, todos os demais apresentaram déficits de heterozigose. Esse resultado reflete o fato de que os indivíduos caracterizados pertencem a populações e subpopulações diferentes e geograficamente distantes, o que concorre para o aumento de homozigose (efeito de Wahlund).

3.2. Neutralidade Seletiva e Independência dos locos analisados

A hipótese da neutralidade seletiva (Kimura, 1968; King & Jukes, 1969) e a independência (equilíbrio de ligação) dos locos microssatélites são duas premissas importantes a serem consideradas antes de quaisquer análises baseadas em freqüências alélicas. A freqüência alélica de cada loco em uma população resulta da interação de forças exercidas pelos processos de deriva genética, mutações e/ou migração. Forte seleção associada a um loco pode superar estas forças e estabilizar as freqüências alélicas em todas as subpopulações levando a uma subestimação da distância genética e da sub-

estrutura da população. Do mesmo modo, diferenças na pressão seletiva entre regiões poderiam levar à fixação de alelos alternativos em diferentes subpopulações e levar a superestimativas daqueles parâmetros.

A grande maioria dos microssatélites é considerada neutra e, portanto, a frequência de seus alelos não seria influenciada pela seleção natural. Entretanto, a ligação desses marcadores a locos sob pressão de seleção é bem conhecida, como tem revelado a literatura sobre doenças genéticas humanas (Yechezkel & Morris, 1999). A localização exata dos microssatélites usados na análise dos genótipos é desconhecida, daí a necessidade de detectar possíveis efeitos da seleção antes de quaisquer análises.

Por outro lado, a análise de grande número de locos incrementa o poder de discriminação de genótipos e de estruturação da população, visto que cada loco contém uma história independente em função da intensidade das forças de deriva genética, mutação e migração. Entretanto, se os locos são geneticamente ligados (não independentes), os resultados são viesados uma vez que as informações associadas a grupos de ligação serão super-representadas no conjunto de dados combinados.

A Figura 6 representa o teste de homozigosidade para neutralidade de Ewens-Watterson (Hartl & Clark, 1989) aplicado ao conjunto de todos os genótipos (172) analisados. As curvas do gráfico representam os limites do intervalo de confiança de 95 % gerados a partir da simulação de amostras com o mesmo número de alelos e indivíduos, segundo a distribuição de amostragem de Ewens (1972) predita pelo modelo de alelos infinitos (IAM). Observa-se que à exceção dos locos mTcCIR-19, mTcCIR-33, mTcCIR-35 e mTcCIR-43, a identidade genética (F_{obs}) de todos os demais caiu dentro dos limites da região de confiança, sugerindo boa concordância com a hipótese de neutralidade ou mutação neutra. A baixa homozigosidade dos quatro locos que caíram abaixo do limite inferior do intervalo de confiança sugere que o polimorfismo nos mesmos seja mantido por algum tipo de seleção estabilizadora. Entretanto, há que se ressaltar que a identidade genética (F_{obs}), calculada a partir do somatório do quadrado das frequências alélicas de cada loco, reflete na verdade a homozigosidade esperada numa população panmítica (em equilíbrio de Hardy-Weinberg). Justamente por isso a maioria

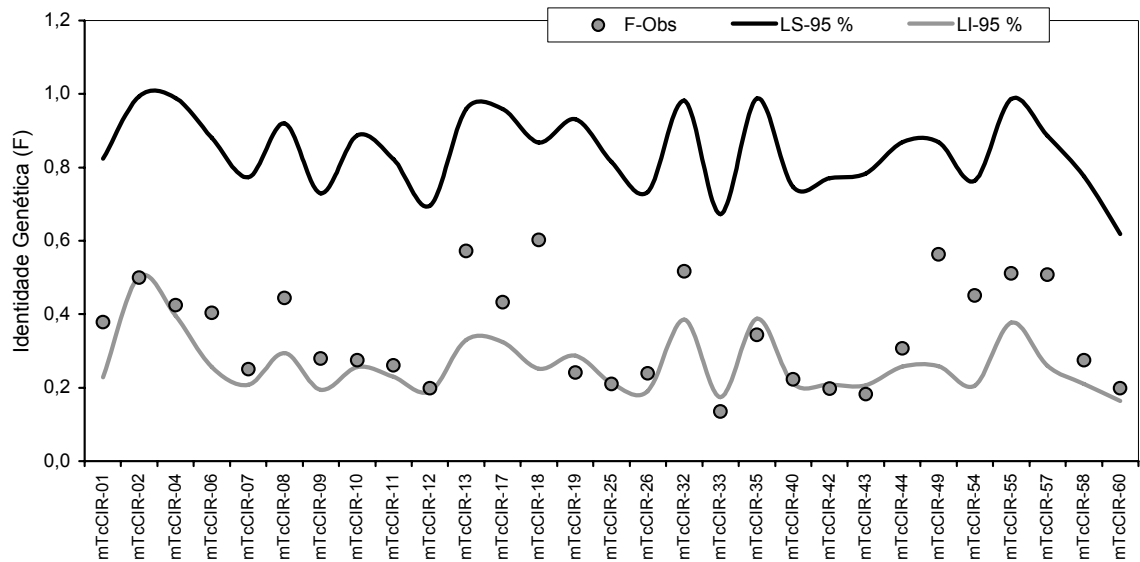


Figura 6. Identidade genética, calculada pelo teste de Waterson-Ewens para os diferentes locos de microsatélites numa amostra de 172 clones de cacau. As linhas sólidas representam os limites do intervalo de confiança a 95 % de probabilidade.

dos locos apresentaram-se mais próximos do limite inferior do intervalo de confiança, sugerindo um excesso de heterozigosidade, o que certamente deve ser levado em consideração antes de rejeitar a hipótese de neutralidade para os quatro locos que extrapolaram a região de aceitação da mesma. Se *plotadas* as homozigosidades realmente observadas em cada loco, todos se manteriam dentro dos limites da região de aceitação da neutralidade. De qualquer modo, observa-se que de maneira geral os locos utilizados se ajustam bem ao modelo de mutação neutra, não se justificando descartar quaisquer dos locos para as análises efetuadas no presente estudo. Deixou-se de testar a neutralidade dos locos dentro de cada população separadamente devido ao pequeno tamanho das amostras, em relação ao número de alelos, o que certamente introduziria estimativas viesadas do teste. Por outro lado, a despeito da fraca força do teste de homozigosidade de Watterson-Ewens para populações em não equilíbrio de Hardy-Weinberg, os resultados sugerem a realização de estudos mais acurados sobre o significado adaptativo do polimorfismo genético associado aos locos que se mostraram mais afastados da hipótese de neutralidade.

A independência dos microsatélites foi avaliada pelo teste de associação entre genótipos de todos os pares de locos analisados na amostra

total (Anexo, Tabela 1). A estatística usada para testar as tabelas geradas a cada combinação de genótipos é a razão de verossimilhança da estatística G. O P-valor estima a proporção da estatística do conjunto de dados randomizados que são maiores ou iguais ao observado. Observa-se que, de um total de 406 combinações de pares de locos, apenas uma apresentou associação significativa (mTcCIR-01 x mTcCIR-54). Tal associação, entretanto, não significa necessariamente que os locos estejam ligados, uma vez que os mesmos estão localizados em cromossomos diferentes (8 e 1, respectivamente). Como o teste de neutralidade (Figura 11) praticamente elimina a hipótese de associação dos alelos desses locos a algum gene sob pressão de seleção (seleção epistática), o desequilíbrio de ligação detectado provavelmente resulta da deriva genética e limitada migração nas subpopulações amostradas, como sugere Otha (1982). O resultado do teste de desequilíbrio genotípico mostra que de maneira geral os locos utilizados em sua grande maioria segregam independentemente. Isso sugere que os diversos locos se acham bem distribuídos pelo genoma do cacauero refletindo de forma abrangente a variabilidade genética que se pretende caracterizar nas coleções de germoplasma de cacau.

3.3. Poder discriminante dos microssatélites

O perfil dos microssatélites para cada um dos clones caracterizados com os 29 pares de *primers* (Anexo, Tabela 2) mostra que os locos utilizados foram capazes de discriminar sem ambigüidade 169 dos 172 genótipos caracterizados (98,3 %). O três genótipos indistinguíveis entre si foram os clones CAB 731, CAB 732 e CAB 734, todos pertencentes a amostra coletada nas imediações do lago da Hidrelétrica de Balbina, no Rio Uatumã (AM), o que reforça a probabilidade de identidade dos mesmos.

O alto poder de discriminação dos locos (Tabela 5) sugere que os 29 pares de *primers* poderiam discriminar quaisquer dois acessos dentro das coleções de germoplasma de cacau, principalmente levando em consideração que a amostra analisada é composta de sub-AMostras constituídas de genótipos aparentados, representativos das diversas bacias hidrográficas. O poder de

discriminação de cada loco individualmente variou de 0,394 para mTcCIR-18 a 0,865 para mTcCIR-33 e a probabilidade de identidade de quaisquer dois genótipos variou de 0,018 para o mTcCIR-33 a 0,367 para o loco mTcCIR-18. O efeito combinado de todos os locos eleva para 1,000 o poder de discriminação e reduz a probabilidade de identidade de dois genótipos para $2,8 \times 10^{-29}$, assumindo que todos os locos são independentes (Tabela 5).

Tabela 5. Nº Efetivo de alelos (n_e), Conteúdo de Informação Polimórfica (PIC), Poder Discriminante (PD) e Probabilidade de Identidade (PI), para 29 microssatélites baseados na frequência alélica detectada em 172 genótipos de cacauzeiros.

LOCO	Nº Alelos	n_e	PIC	PD	PI
mTcCIR-01	7	2,64	0,54747	0,620670	0,14389
mTcCIR-02	2	2,00	0,37499	0,499982	0,25002
mTcCIR-04	3	2,35	0,50551	0,574458	0,18109
mTcCIR-06	6	2,48	0,56531	0,597286	0,16218
mTcCIR-07	8	3,99	0,71040	0,749077	0,06296
mTcCIR-08	5	2,26	0,51927	0,558106	0,19527
mTcCIR-09	9	3,58	0,66915	0,720746	0,07798
mTcCIR-10	6	3,62	0,67721	0,723864	0,07625
mTcCIR-11	7	3,82	0,69987	0,738072	0,06861
mTcCIR-12	9	5,10	0,77714	0,804024	0,03841
mTcCIR-13	4	1,74	0,38036	0,424518	0,33118
mTcCIR-17	4	2,34	0,50063	0,571998	0,18319
mTcCIR-18	6	1,65	0,35911	0,393903	0,36735
mTcCIR-19	5	4,16	0,72033	0,759878	0,05766
mTcCIR-25	8	4,74	0,76333	0,789242	0,04442
mTcCIR-26	9	4,18	0,72267	0,760899	0,05717
mTcCIR-32	3	1,94	0,38356	0,483383	0,26689
mTcCIR-33	10	7,40	0,84967	0,864935	0,01824
mTcCIR-35	3	2,90	0,58196	0,655754	0,11851
mTcCIR-40	8	4,50	0,74996	0,777754	0,04939
mTcCIR-42	8	5,05	0,77928	0,802100	0,03916
mTcCIR-43	8	5,55	0,79477	0,819784	0,03248
mTcCIR-44	6	3,22	0,64142	0,689593	0,09635
mTcCIR-49	6	1,78	0,40897	0,437666	0,31622
mTcCIR-54	8	2,20	0,51309	0,544484	0,20749
mTcCIR-55	3	1,96	0,39025	0,489694	0,26041
mTcCIR-57	6	1,98	0,45340	0,495450	0,25457
mTcCIR-58	8	3,71	0,68964	0,730734	0,07250
mTcCIR-60	11	5,08	0,78090	0,803196	0,03873
Probabilidade de dois genótipos iguais					2,8335E-29
Poder Discriminante Combinado				1,000000	

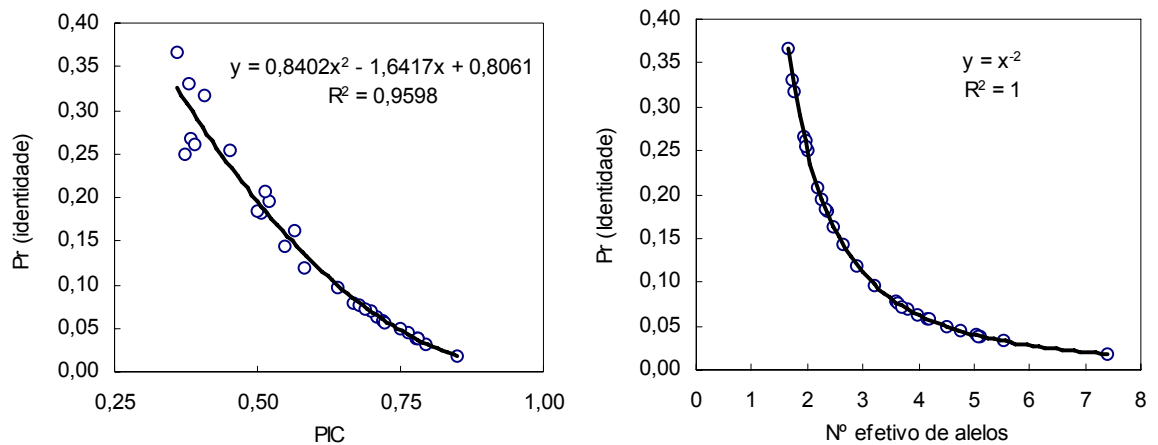


Figura 7. Probabilidade de identidade de dois genótipos em função do conteúdo de informação polimórfica (PIC) e do número efetivo de alelos.

O poder de discriminação dos microssatélites analisados reflete o elevado conteúdo de informação polimórfica (PIC) dos locos, variando de 0,850 para o mTcCIR-33 a 0,359 para o loco mTcCIR-18, com um valor médio de 0,604. Esse índice, tal como a heterozigotidade (chance de que um indivíduo tomado ao acaso seja heterozigoto) mede quão informativo é um marcador. O PIC reflete melhor o polimorfismo por eliminar do índice metade dos indivíduos heterozigotos similares, portanto não informativa. Entretanto, o número efetivo de alelos (n_e), por traduzir o polimorfismo de alelos com igual frequência, novamente mostra melhor associação tanto com o poder de discriminação (PD) dos marcadores, como com a probabilidade de identidade de dois genótipos (PI). A Figura 7 mostra que a relação $PI = f(n_e)$, por exemplo, se estabelece por meio de uma função quadrática perfeita a qual permite que se obtenha facilmente a probabilidade de identidade de dois genótipos, de um marcador, a partir da equação $PI = 1/(n_e)^2$, muito mais simples, portanto, que a fórmula tradicionalmente empregada:

$$PI = \sum p_i^4 + \sum_{i=1}^{i=n-1} \sum_{j=i+1}^n (2p_i p_j)^2$$

Considerando-se o elevado poder de discriminação dos locos mais informativos é possível reduzir bastante o número de locos utilizado na discriminação de genótipos. A Figura 8 mostra que utilizando apenas seis

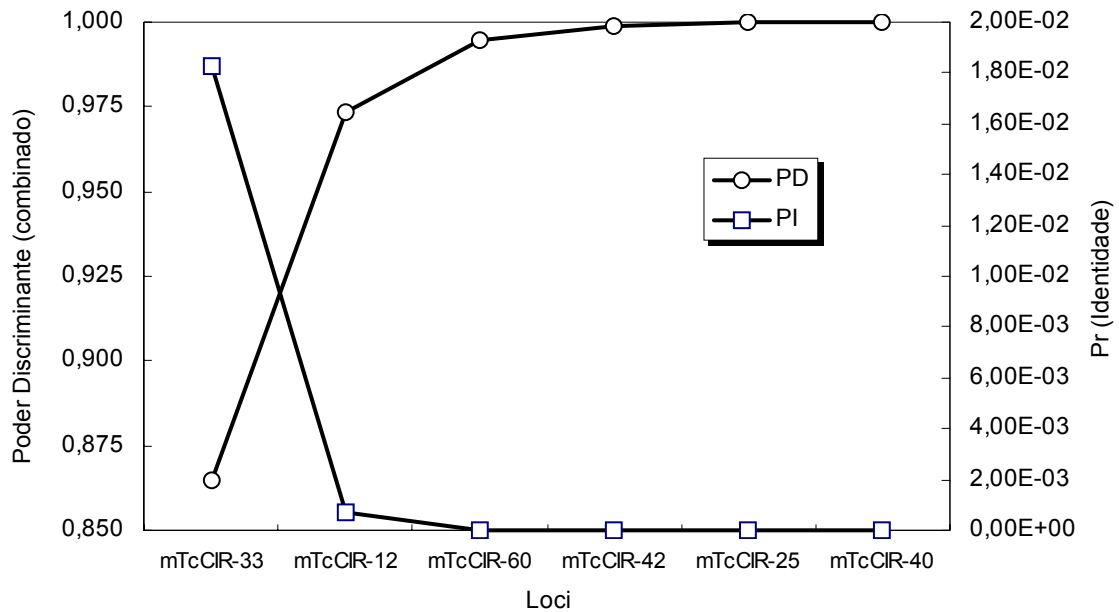


Figura 8. Poder de discriminação (PD) e Probabilidade de identidade (PI) de dois genótipos considerando-se os locos mais informativos (em ordem decrescente) para os 172 clones caracterizados.

pares dos *primers* mais informativos atinge-se uma discriminação da ordem de 0,999952 e probabilidade de identidade de dois genótipos de apenas $2,33 \times 10^{-9}$.

3.4. Heterozigidade individual dos cones analisados

Considerando as diferentes estratégias atualmente empregadas na implantação de roças de cacau na Amazônia (híbridos) e na região cacauera bahiana (clones), a caracterização do nível de heterozigidade dos acessos pode proporcionar informações importantes aos programas de melhoramento. Progenitores com baixo nível de heterozigidade produzem progênes muito homogêneas, característica desejável quando se pretende distribuir as progênes como variedades. Por outro lado, progenitores altamente heterozigotos tendem a produzir progênes muito heterogêneas, muito úteis quando se deseja selecionar indivíduos para determinados atributos.

O parentesco entre progenitores de um indivíduo pode ser visto como um contínuo, ao longo do qual podem ocorrer variações na resposta adaptativa da progênie (Hartl & Clark, 1989). Normalmente, progênes endogâmicas são menos adaptadas (“depressão endogâmica”), enquanto progênes cruzadas são mais adaptadas (“heterose”). O processo responsável por tais repostas

provavelmente inclui aumento de homoziguidade para mutações recessivas deletérias e sobredominância, embora a importância relativa de cada um desses processos ainda não esteja clara (Charlesworth & Charlesworth, 1987). Assim, a análise da heteroziguidade individual média, a partir de marcadores moleculares codominantes como microssatélites, pode ser uma ferramenta de grande utilidade, considerando que a heteroziguidade seria inversamente correlacionada com a endogamia.

Em cacau, estimativas da heteroziguidade tem sido obtidas utilizando isoenzimas e RFLP (Laurent *et al.*, 1995; Sounigo *et al.*, 2001). Entretanto os resultados são limitados tanto pelo número de locos estudados como pelo baixo grau de polimorfismo em comparação com marcadores microssatélites. Devido ao maior polimorfismo, a heteroziguidade individual medida por microssatélites deve se mostrar mais bem associada com o grau de endogamia. Por outro lado, o modelo de mutação *stepwise* de microssatélites (Eisen, 1999), permite que se acrescente ao estudo da heteroziguidade a informação histórica associada à diferença no comprimento das unidades repetidas entre os alelos de um mesmo loco, o que reflete o tempo médio de coalescência para o microssatélite estudado (Goldstein *et al.*, 1995; Slatkin, 1995). Observando esse modelo (SMM), a medida d^2 média, proposta por Coulson *et al* (1998), estimada a partir dos alelos de cada loco, em cada indivíduo, é uma medida da distância genética entre os gametas que formaram o indivíduo. Matematicamente representa a variância no comprimento médio dos alelos dentro de um indivíduo. Quando estimados a partir de uma mesma amostra de microssatélites altamente polimórficos, a heteroziguidade individual (H) e a d^2 média enfatizam diferentes extremidades do contínuo endogamia-exogamia. A heteroziguidade individual refletiria acasalamentos recentes entre parentes, enquanto a d^2 média, embora inclua a heteroziguidade, será fortemente influenciada pela variação no comprimento dos alelos, evento mais profundo no pedigree (Pemberton *et al.* 1999).

A Tabela 6 apresenta a heteroziguidade individual presente em cada material em função da identidade do alelo (H) e em função do tamanho do alelo (d^2 média), estimados a partir da caracterização do perfil de bandas de microssatélites de cada clone (Anexo, Tabela 2). Os valores de d^2 média

Tabela 6. Heterozigosidade individual dos clones em função da identidade do alelo (H) e do tamanho do alelo (d^2 média)

Clones	H	$d^2 \times 10^{-6}$	Clones	H	$d^2 \times 10^{-6}$	Clones	H	$d^2 \times 10^{-6}$
CAB 006	0,3448	101,0	CAB 233	0,2069	27,3	CAB 132	0,3929	34,8
CAB 020	0,3103	50,5	CAB 236	0,4828	112,0	CAB 191	0,2069	44,3
CAB 036	0,0690	4,4	CAB 247	0,0741	16,1	CAB 206	0,2414	32,2
CAB 016	0,0690	4,4	CAB 410	0,1429	62,5	CAB 002	0,4828	111,4
CAB 20B	0,1724	6,7	CAB 253	0,1724	30,0	CAB 003	0,4483	52,2
CAB 039	0,0690	4,4	CAB 255	0,0370	14,8	CAB 004	0,5185	223,8
CAB 104	0,1724	20,7	CAB 258	0,0357	14,3	CAB 005	0,4138	48,1
CAB 105	0,3103	37,0	CAB 260	0,0345	13,8	CAB 046	0,3448	228,0
CAB 107	0,2759	32,1	CAB 374	0,0714	17,9	CAB 056	0,2069	85,3
CAB 108	0,3448	67,1	CAB 388	0,0690	21,6	CAB 058	0,3448	191,7
CAB 110	0,3214	82,0	CAB 396	0,1379	44,3	CAB 270	0,3793	48,8
CAB 111	0,2857	41,5	CAB 398	0,2069	61,6	CAB 271	0,3793	191,6
CAB 113	0,3793	52,2	CAB 406	0,0357	14,3	CAB 275	0,1379	30,7
CAB 114	0,1786	23,1	CAB 478	0,2759	44,3	CAB 327	0,3448	60,8
CAB 117	0,3103	35,3	CAB 484	0,2759	48,4	CAB 279	0,3793	148,1
CAB 135	0,3448	54,3	CAB 486	0,3214	44,6	CAB 324	0,2143	63,4
CAB 136	0,2222	27,3	CAB 487	0,2500	31,7	CAB 325	0,2222	106,3
CAB 146	0,2069	28,8	CAB 493	0,1034	5,3	CAB 326	0,2069	41,4
CAB 147	0,3103	57,4	CAB 497	0,3103	19,2	CAB 330	0,4444	61,7
CAB 148	0,3448	45,0	CAB 499	0,3077	18,8	CAB 331	0,1923	32,0
CAB 151	0,2759	44,7	CAB 481	0,1852	24,0	CAB 001	0,4138	108,4
CAB 153	0,2414	48,1	CAB 509	0,3103	30,3	CAB 337	0,2500	27,6
CAB 155	0,3103	47,5	CAB 511	0,3103	51,4	CAB 266	0,3571	103,8
CAB 160	0,3448	48,3	CAB 512	0,3103	29,0	CAB 280	0,3704	28,4
CAB 167	0,1724	25,6	CAB 513	0,3793	94,0	CAB 283	0,4828	154,0
CAB170	0,4483	95,2	CAB 492	0,0714	4,6	CAB 282	0,3793	43,1
CAB 171	0,1429	22,7	CAB 500	0,2222	33,1	CAB 302	0,3214	110,1
CAB 176	0,3103	90,7	CAB 008	0,0690	17,2	CAB 304	0,1667	78,9
CJ 10	0,2069	18,6	CAB 009	0,3571	129,5	CAB 306	0,3462	188,8
CJ 11	0,3103	58,8	CAB 011	0,0690	19,6	CAB 308	0,1481	135,8
CJ 02	0,3103	31,5	CAB 012	0,1034	25,0	CAB 517	0,2414	61,9
CJ 03	0,5172	54,7	CAB 013	0,0690	21,6	CAB 313	0,3214	108,8
CJ 04	0,3103	33,6	CAB 177	0,2414	33,9	CAB 312	0,2963	68,2
CJ 05	0,3793	63,9	CAB 178	0,3571	69,5	CAB 316	0,4483	101,6
CJ 06	0,3571	181,4	CAB 179	0,2069	45,9	CAB 514	0,3448	113,0
CJ 07	0,4074	76,6	CAB 415	0,0690	16,0	CAB 529	0,2963	80,3
CJ 08	0,3571	40,4	CAB 017	0,2414	86,7	CAB 181	0,5172	170,7
CJ 09	0,2759	36,9	CAB 018	0,4286	45,6	CAB 182	0,3571	57,2
CAB 119	0,2069	37,5	CAB 019	0,2759	281,7	CAB 183	0,3448	521,9
CAB 120	0,4643	76,6	CAB 022	0,3333	176,7	CAB 184	0,1481	42,2
CAB 121	0,1154	9,7	CAB 023	0,3103	22,0	CAB 185	0,4828	497,3
CAB 063	0,2000	16,2	CAB 024	0,2500	38,4	CAB 187	0,3793	189,3
CAB 064	0,2500	77,4	CAB 007	0,3929	201,1	CAB 200	0,1852	42,6
CAB 066	0,4138	57,8	CAB 028	0,3793	182,6	CAB 201	0,2500	68,4
CAB 067	0,4138	60,7	CAB 029	0,4138	316,5	CAB 202	0,3103	33,5
CAB 068	0,2500	56,3	CAB 035	0,5357	76,0	CAB 203	0,2069	18,1
CAB 070	0,2414	31,2	CAB 094	0,2069	31,6	CAB 091	0,1852	53,4
CAB 071	0,2414	27,0	CAB 095	0,3448	32,2	CAB 092	0,2069	219,2
CAB 072	0,3793	41,6	CAB 096	0,2069	38,2	CAB 730	0,0345	2,2
CAB 076	0,4138	86,6	CAB 098	0,2759	27,1	CAB 731	0,0690	7,2
CAB 077	0,3103	41,1	CAB 101	0,4138	45,8	CAB 732	0,0690	7,2
CAB 079	0,3571	80,0	CAB 102	0,5172	127,5	CAB 734	0,0714	7,4
CAB 080	0,3448	53,3	CAB 123	0,3103	61,3	CAB 014	0,2759	206,7
CAB 081	0,2069	20,7	CAB 126	0,1724	30,6	CAB 015	0,2759	41,7
CAB 082	0,0400	7,8	CAB 127	0,2857	57,6	SCA-6	0,2759	42,0
CAB 088	0,3448	31,3	CAB 128	0,1724	21,3	ICS-1	0,3793	59,8
CAB 221	0,2500	39,8	CAB 130	0,2759	29,5	Média	0,2732	67,3
CAB 232	0,1429	25,9	CAB 131	0,3448	32,1	DP	0,1201	73,8

acham-se multiplicados por 10^{-6} porque foram calculados a partir do tamanho total dos alelos e não pelo número de repetições, do que resulta valores bastante elevados.

Observa-se que a heteroziguidade individual em função da identidade do alelo variou de 0,035 para os clones CAB 258, CAB 260, CAB 406 e CAB 730 a 0,536 para o clone CAB 035, com média igual a 0,273; a d^2 média variou de $2,2 \times 10^6$, para o clone CAB 730, a $521,9 \times 10^6$, para o clone CAB 183, média de $67,3 \times 10^6$. Vê-se, portanto, que não há uma correspondência direta entre as duas medidas, o que fica mais claro na Figura 9 que ilustra a extensiva variação em d^2 média dentro de cada classe de heteroziguidade individual (H).

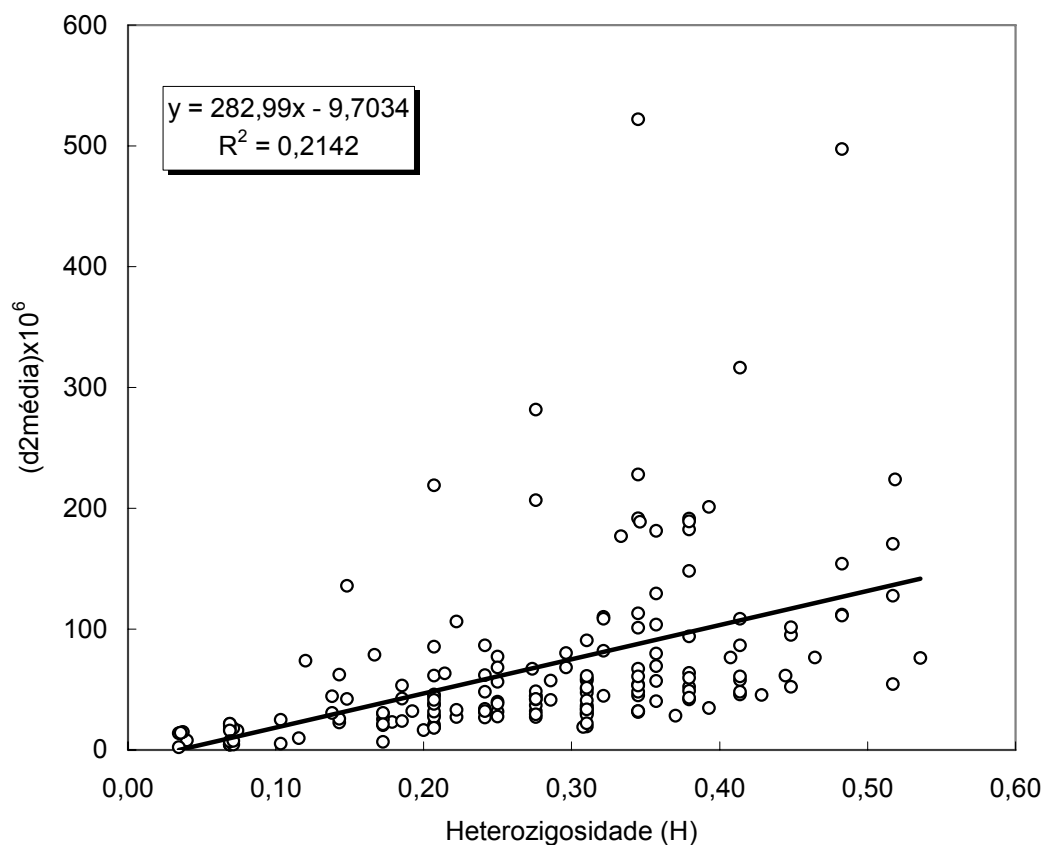


Figura 9. Relação entre a heteroziguidade individual de d^2 média de 172 clones de cacau, revelada por 29 locos microssatélites

Os gráficos que compõem a Figura 10 mostram que a distribuição de freqüências de d^2 média é bastante assimétrica em relação à da heteroziguidade individual, o que reflete a utilização de locos de diferentes graus de polimorfismo dentro da amostra. Observa-se que a heteroziguidade

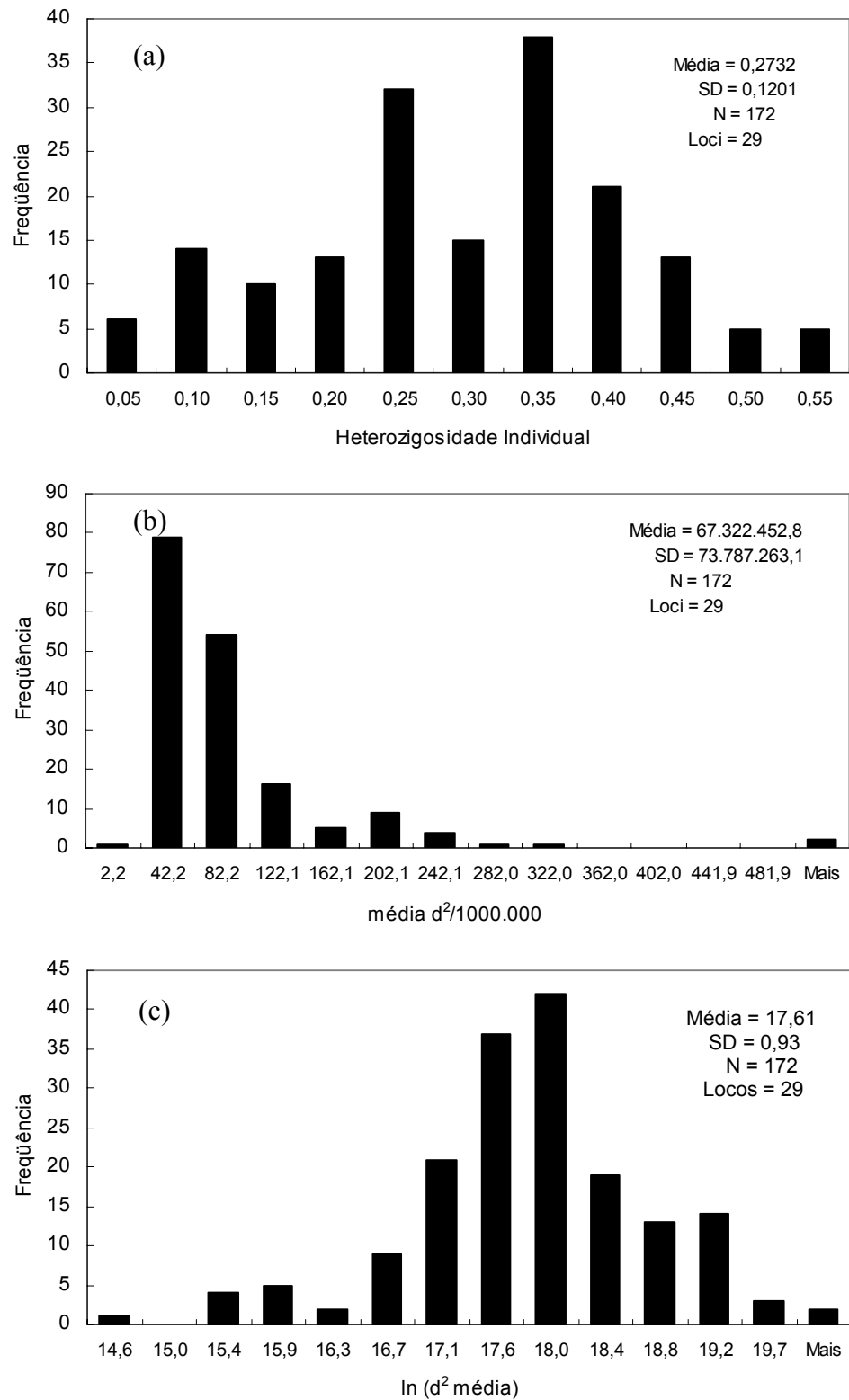


Figura 10. Distribuição de frequência da heterozigosidade individual (a), d^2 média (b) e de $\ln(d^2$ média) (c) para a amostra total de 172 clones de cacau.

individual assume uma distribuição quase normal com a maioria dos genótipos se agrupando nas classes centrais da medida, enquanto d^2 média é fortemente deslocada para a esquerda demonstrando claramente a maior amplitude de variação dessa medida. Disso resulta uma elevada variância, o que constitui uma forte limitação na utilização dessa medida em estudos de associação com características indicadoras de adaptação. Por outro lado, a simples transformação logarítmica de d^2 média permite a normalização dos dados, com redução considerável da variância, conforme mostra a Figura 10c. Entretanto, a utilização, tanto de dados transformados (Coltman *et al.*, 1998), como de não transformados (Coulson *et al.*, 1998), mostram correlação significativa com o peso ao nascer e sobrevivência neonatal de cervo (*Cervus elaphus*) e foca marinha (*Phoca vitulina*), respectivamente, quando nenhuma correlação foi observada com a heterozigidade individual em qualquer dos estudos.

Quando se considera os valores médios de H e d^2 média, em função da bacia de origem dos diferentes materiais (Figura 11), observa-se que os materiais menos heterozigotos foram os provenientes das bacias dos Rios Uatumã-AM, Jamari-RO, Ji-Paraná-RO, Acará-PA e Tocantins-PA, os quais, a exceção dos oriundos do Rio Tocantins, apresentam valores de d^2 média igualmente baixos. Por outro, desconsiderando-se os valores médios relativos às bacias do Tocantins-PA e Xeriuni-RR, sub-representadas (apenas dois indivíduos), destacam-se, em ambos os aspectos (H e d^2 média), as subpopulações representativas das bacias dos rios Tarauacá-AC, Maicuru-PA, Solimões e Solimões/Bx. Japurá-AM, em relação às demais. Isso revela que os materiais dessas regiões além de mais diversos quanto à identidade de seus alelos apresentam maior divergência entre si, como indicam os elevados valores de d^2 média. Isso sugere, que nessas bacias, ocorra com mais frequência cruzamentos entre indivíduos mais divergentes (não parentes), provavelmente devido a um maior fluxo gênico estabelecido por migrantes de populações de bacias interligadas. É o que parece explicar os altos valores de d^2 média nas amostras oriundas das bacias do Rio Solimões, Solimões/Bx Japurá-AM e Maicuru-PA. As subpopulações representadas por essas amostras encontram-se na confluência de bacias primárias com secundárias como mostra o mapa de coletas (Figura 1). Desse modo, os alelos mais

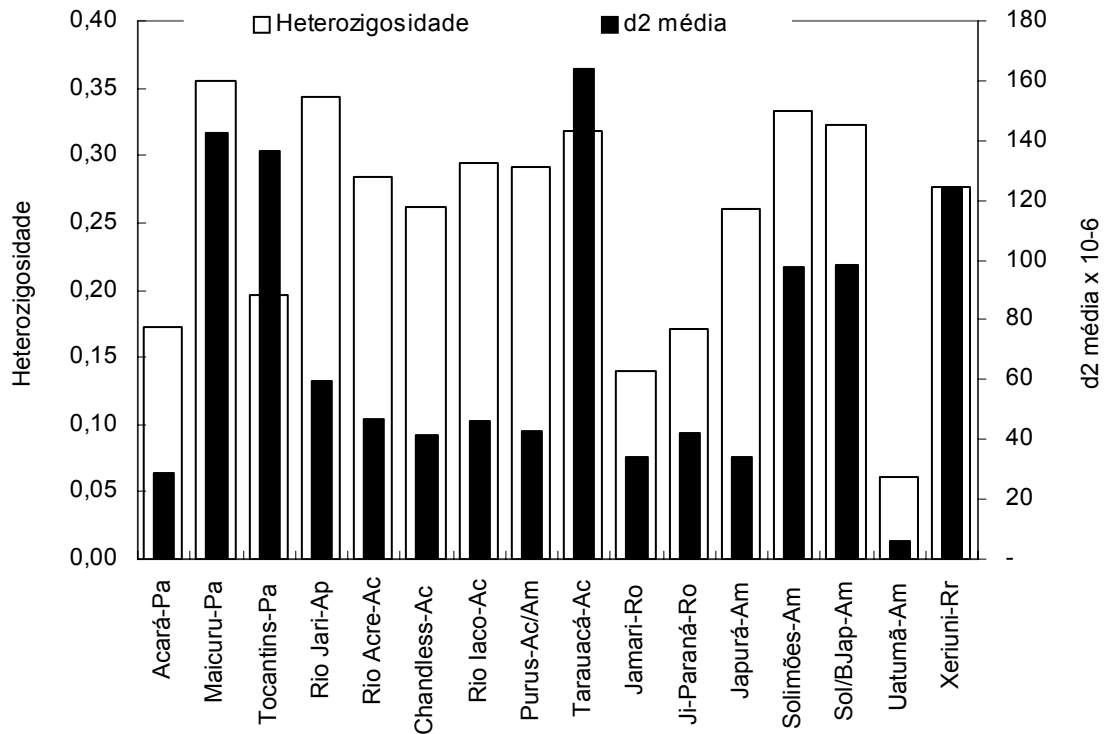


Figura 11. Valores médios da heterozigosidade individual, em função da identidade do alelo (H) e em função do tamanho do alelo (d^2 média) de acordo com as bacias de origem dos clones

divergentes detectados nessas subpopulações parecem refletir a contribuição de populações divergentes situadas à montante das respectivas bacias que se unem nas áreas amostradas. As amostras do Solimões e Solimões/Bx Japurá, por exemplo, estão à jusante de três bacias (Figura 1) importantes em termos de ocorrência de cacau silvestre: i) a bacia do Japurá, afluente da margem esquerda do Solimões, proveniente da Colômbia, um dos prováveis centros de origem da espécie (Cheesman, 1944; Dias, 2001b); ii) a bacia do Solimões, proveniente do Peru e; iii) a bacia do Rio Juruá, afluente da margem direita do Solimões, que interliga-se com a bacia do Rio Tarauacá no Acre, proveniente do Peru. Por sua vez, a amostra relativa à bacia do Maicuru foi colhida às proximidades de Alenquer, Monte Alegre e Óbidos (Figura 1), em áreas semicultivadas (Almeida *et al.*, 1985), certamente submetidas, pela ação do homem, à influência de populações do grupo *forasteiro alto amazônico*, à montante do Amazonas e de *forasteiro baixo amazônico*, à montante do Rio Maicuru. Este último, afluente da margem esquerda do Rio Amazonas, nasce

próximo à fronteira com as Guianas, outra área de ocorrência natural de cacau silvestre, tida como um dos prováveis centros de origem da espécie por Cuatrecasas (1964). Quanto à amostra relativa à bacia do Rio Tarauacá-AC, cujos altos valores de d^2 média divergem das demais subpopulações do Acre, reúne, na verdade, genótipos de duas bacias adjacentes (Rio Tarauacá e Rio Embira) e provavelmente sofre influência do fluxo gênico de populações naturais situadas à montante, no Peru.

De modo geral, os locos utilizados permitiram que se identificassem materiais bastante heterozigotos, independentemente da medida utilizada para caracterizá-los. Diversos materiais apresentaram-se mais heterogêneos que o ICS-1 que é um clone trinitário, obtido da hibridação natural de materiais de grupos bastante distintos (Crioulo x Forasteiros Amazônicos). Considerando-se a identidade dos alelos, 25 clones superam a heterozigosidade individual desse clone, enquanto 59 apresentaram d^2 média maior que o mesmo. Quanto ao material SCA-6, clone do grupo *forasteiro alto amazônico*, auto-incompatível, 96 clones apresentaram heterozigosidades individuais superiores, enquanto 94 o superaram em termos de d^2 média.

A grande variação na heterogeneidade dos clones analisados fica melhor evidenciada na Tabela 7 onde os diferentes genótipos encontram-se agrupados segundo o número absoluto de locos heterozigotos/indivíduo. Não se observou nenhum material totalmente homozigoto. Os mais homogêneos apresentaram pelo menos um loco heterozigoto ($H = 0,034$), enquanto o mais polimórfico apresentou 16 dos 29 locos heterozigotos ($H = 0,552$), com a mediana situada na classe de genótipos com 8 locos heterozigotos ($H = 0,276$). Portanto, a heterozigosidade dos materiais estudados apresenta amplitude de variação adequada a qualquer das estratégias que se deseje adotar em programas de melhoramento. A amostra reúne desde materiais altamente homogêneos (pobremente heterozigotos) até materiais altamente heterogêneos (altamente heterozigotos), cuja utilização vai depender dos atributos agronômicos superiores que cada um seja portador.

3.5. Relação Genética entre os clones

A Figura 12 representa a dispersão bi-dimensional obtida da análise fatorial de correspondência a partir dos genótipos individuais de cada clone codificados para cada alelo nos diferentes locos, conforme estejam ausentes (0), heterozigoto (1) ou homozigoto (2). Essa técnica, ainda relativamente pouco utilizada, mostra-se particularmente útil quando os objetos de análise são os vetores de frequência alélica; demonstrou-se que o traço da matriz dos auto-valores (utilizada para obtenção de combinações lineares ortogonais dos vetores das variáveis originais) corresponde a uma estimativa do F_{st} (Long, 1986); os valores de inércia ao longo de cada eixo correspondem às combinações lineares de valores de F_{st} para o loco correspondente (Guinand, 1996).

A despeito da distorção associada à representação plana do hiperespaço gerado pelos 29 locos analisados, observa-se uma nítida tendência dos genótipos se agruparem segundo suas origens geográficas, conforme já propunham Pound (1938), Bartley (1963) e Dias *et al.* (2003a), baseando-se em características distintas associadas a bacias hidrográficas específicas, e também constatado por Silva *et al.* (1998) usando marcadores RAPD. A quase totalidade dos genótipos representativos das bacias hidrográficas do estado do Acre forma uma nuvem concentrada de pontos que praticamente impede a distinção entre as mesmas, distinguindo-se, porém, completamente das demais pelos valores fortemente negativos de suas ordenadas (dimensão 2). Todas as demais amostras assumem ordenadas (dimensão 2) positivas, distinguindo-se umas das outras principalmente pela dispersão dos seus genótipos ao longo da dimensão 1.

As amostras representativas das bacias de Rondônia mostraram-se bastante dissociadas das demais, formando dois grupos discretos sem qualquer interseção entre os mesmos. A exceção de um genótipo (CAB 236), todos os demais de ambas as bacias situaram-se no terceiro quadrante do gráfico (ordenadas positivas, abscissas negativas), sendo que a discriminação entre as duas bacias se dá preponderantemente ao longo da dimensão 1. A limitada dispersão dessas amostras confirma a grande homogeneidade fenotípica das subpopulações coletadas na região da bacia do rio Jamari

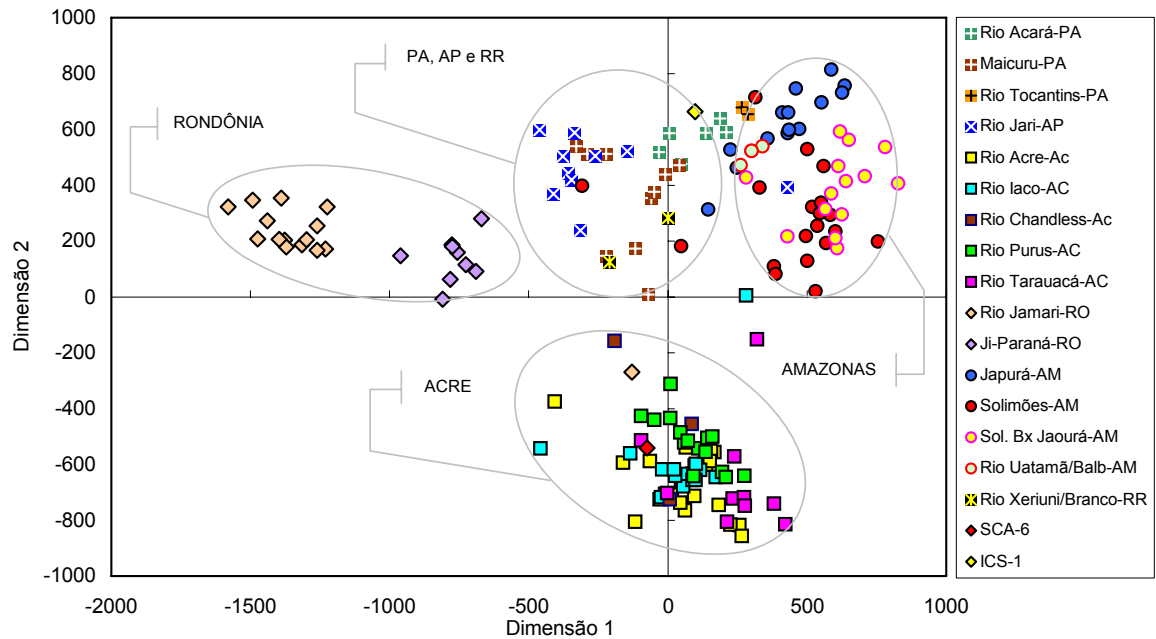


Figura 12. Dispersão dos 172 genótipos de cacau obtida por análise fatorial de correspondência utilizando os 29 locos microssatélites. Auto-valores (Fator 1 = 0,2825, perc = 6,94; Fator 2 = 0,2605, perc = 6,4).

(Ameida, 1982) e do Rio Ji-Paraná (Almeida & Almeida, 1987). Entretanto, a disjunção observada entre as amostras rechaça a possibilidade de ambas pertencerem ao mesmo complexo gênico, hipótese levantada por Almeida (2001), face às semelhanças fenotípicas e ausência de barreiras ao fluxo gênico entre as populações de terra firme, compreendidas entre os municípios de Ariquemes e Cacoal, Rondônia.

As amostras representativas das bacias do Estado do Amazonas concentraram-se predominantemente no primeiro quadrante, mostrando-se mais dispersas ao longo da dimensão 2. Dentro dessa nuvem observa-se que as amostras das bacias do Rio Solimões e as da confluência deste com o Rio Japurá (Solimões/Baixo Japurá) mostram-se mais dispersas e misturadas entre si, enquanto os genótipos das bacias do alto Japurá e do Rio Uatamã mostraram-se mais concentradas dentro de uma área relativamente distinta das demais amostras desse grupo. A grande dispersão dos pontos representativos da bacia do Rio Solimões reflete a grande área de coleta representada por essa amostra, que inclui genótipos coletados em diversas ilhas às proximidades de Tefé-AM (local 21) até às proximidades de Manaus (local 17) (Figura 1). Por

outro lado, a interpenetração dos pontos representativos dessa amostra com as do Rio Solimões/Bx. Japurá reflete a relativa proximidade dos locais de coleta 21 e 22 de grande parte dos genótipos que compõem, respectivamente, as mesmas (Figura 1).

Finalmente, os pontos relativos às amostras representativas de bacias do Pará, Amapá e Roraima mostram-se relativamente dispersos, numa faixa intermediária da dimensão 2, entre as nuvens representativas dos genótipos de Ji-Paraná-RO e as que concentram as amostras do Amazonas. Nessa nuvem observa-se que a amostra mais dispersa é a que representa a bacia do Rio Maicuru-PA, cujos pontos variam em ambas as dimensões, misturando-se com as amostras coletadas no Rio Acará-PA, Rio Jarí-AP e Rio Xeriuni-RR. Essa dispersão parece refletir os diferentes e dispersos locais de coleta dessa amostra, que inclui genótipos coletados em terra firme (Rod. Alenquer-Mte Alegre e Rod. Alenquer-Óbidos) e em faixa ribeirinha (Paraná-Mirim, Alenquer). Os materiais coletados no Rio Acará-PA mostram-se relativamente concentrados e próximos dos genótipos coletados no Rio Tocantins, ambos contíguos à nuvem correspondente às bacias do Amazonas. Entretanto, o baixo número de indivíduos nessas amostras, bem como na do Rio Xeriuni-RR, dificulta uma melhor apreciação da tendência de dispersão dos seus genótipos. A amostra do Rio Jarí-AP, representada por genótipos bastante homogêneos, apresenta-se relativamente concentrada no terceiro quadrante, a exceção do clone CJ-03 que se confunde com os genótipos que compõem a nuvem de pontos das bacias do Amazonas. É interessante notar que os coletores dessa amostra já apontavam a pouca variabilidade fenotípica dos acessos, principalmente quanto a forma dos frutos, diferindo dos demais apenas um material (CJ-08), cujos frutos destacavam-se pelo grande tamanho com cerca de 25 cm (Vello & Silva, 1968). É possível que se esteja diante de um erro de identificação entre os códigos CJ-03 e CJ-08, o qual somente poderá ser dirimido com a comparação fenotípica de ambos os materiais.

Apesar da análise de correspondência facilitar a observação do padrão geral de dispersão das amostras, a representação plana apresenta distorções consideráveis que podem dificultar a distinção entre grupos de genótipos cuja diferenciação esteja associada a fatores (eixos) sub-representados na

dispersão bi-dimensional. Tal distinção pode ser melhor avaliada por meio de medidas de similaridade entre os genótipos de múltiplos locos de cada dois indivíduos. A Figura 13 mostra a árvore construída pelo agrupamento hierárquico UPGMA a partir da matriz de dissimilaridade, obtida a partir do coeficiente de similaridade para marcadores multialélicos (Cruz, 2001). De

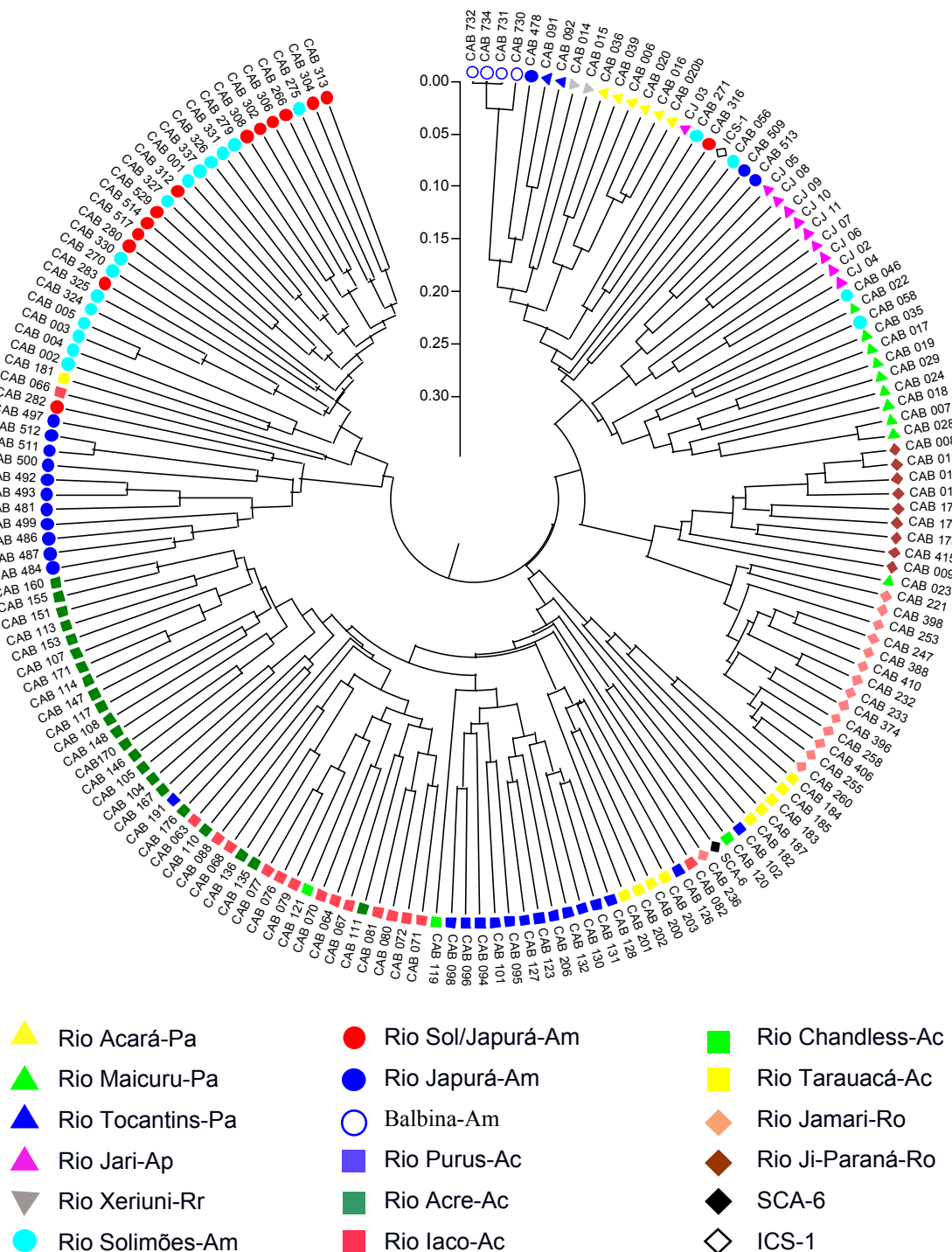


Figura 13. Dendrograma de dissimilaridade entre genótipos de cacau baseado no coeficiente de coincidência simples de 29 microssatélites, método UPGMA.

maneira geral, a topologia da árvore ratifica o agrupamento dos indivíduos segundo a sua origem geográfica, como já fora observado na dispersão fornecida pela análise de correspondência. Dos 172 clones examinados, um total de 140 (81,4%) formou grupos relativamente discretos e coincidentes com a bacia de origem das amostras. Quando se considera o estado de origem das bacias os grupos observados incluem 156 (90,7%) indivíduos. A Tabela 8 relaciona o percentual de indivíduos que agruparam dentro do seu respectivo grupo de origem, mostrando que a dissimilaridade detectada pelos 29 locos microssatélites usados foi capaz de discriminar consideravelmente a origem da grande maioria dos clones, mesmo entre bacias muito próximas, cuja distinção mostrou-se pouco evidente na análise de correspondência. Observa-se, entretanto, a baixa definição dos grupos correspondentes às bacias do Rio Chandless (33,3%) e as do Amazonas, notadamente a do Rio Solimões (53,3%), ratificando a grande dispersão observada na análise de correspondência das mesmas. Considerando-se, porém, o agrupamento formado pelos genótipos componentes das bacias Rio Solimões e Solimões/Bx. Japurá em conjunto, observa-se que o mesmo reúne 27 (93,1%) de um total de 29 clones que representam as duas bacias. Isso sugere que o fluxo gênico entre as regiões de coleta dessas amostras deve ser alto, certamente facilitada pelo complexo formado por rios, paranás, furos e canais que interligam as bacias hidrográficas do Médio Solimões e Baixo Japurá e provavelmente intensificada pela ação antrópica. Segundo Almeida et al. (1995), 74% das populações amostradas nessa região situavam-se em áreas de sítios próximas

Tabela 8. Número de clones que agruparam (Ng) dentro da respectiva bacia hidrográfica de origem em relação ao tamanho da amostra analisada (Na).

Bacia	Na	Ng	%	Bacia	Na	Ng	%
Rio Acará-PA	06	06	100,0	Rio Iaco-AC	15	13	86,7
Rio Tocantins-PA	02	02	100,0	Rio Purus-AC/Am	15	12	80,0
Rio Maicuru-PA	10	09	90,0	Rio Tarauacá-AC	10	9	90,0
Rio Jarí-AP	10	09	90,0	Rio Chandless-AC	03	1	33,3
Rio Xeriuni-RR	02	02	100,0	Rio Uatumã-AM	04	04	100,
Rio Ji-Paraná-RO	09	09	100,0	Rio Japurá-AM	14	11	78,6
Rio Jamari-RO	15	14	93,3	Rio Solimões-AM	15	08	53,3
Rio Acre-AC	22	18	81,8	Solimões/Bx Japurá-AM	14	10	71,4

de moradores, em condições de semi-cultivo. Segundo estes autores a disseminação do cacau nessa área, ao longo dos principais cursos d'água, provêm de plantios de sementes distribuídas por habitantes locais, sugerindo prática de seleção simples.

A análise de agrupamento ratifica que as amostras coletadas em Rondônia (Rio Jamari e Rio Ji-Paraná) constituem dois grupos discretos entre si, mostrando-se bem divergentes em relação a todos as demais, refutando, mais uma vez, a hipótese das mesmas constituírem um mesmo complexo gênico. A grande uniformidade dentro dessas bacias sugere a possibilidade de que o clone CAB 236 da bacia do Jamari, agrupado junto com o material do Rio Tarauacá-AC esteja erradamente codificado, considerando a forte dissimilaridade observada entre as bacias do Jamari e do Rio Tarauacá. Analogamente, a mesma consideração pode ser feita com relação ao clone CAB 023, registrado na bacia do rio Maicuru, mas agrupado junto com o, relativamente distante, material de Ji-Paraná-RO.

Os materiais representativos das populações do Pará, Amapá, Roraima, além da bacia do rio Uatumã, a despeito de formarem grupos homogêneos e quase discretos, agregam-se para formar um grupo maior tal como a nuvem observada na Figura 11. Confirma-se a relativa uniformidade da amostra do Rio Jarí, fortalecendo a suspeita de erro na identificação (codificação) do clone CJ-03, a qual pode ter sido trocada com o clone CJ-08, este sim reportado como divergente pelos coletores (Vello e Silva, 1968).

As comparações interclonais revelam que a dissimilaridade variou de 0,00 (entre os clones CAB 731, CAB 732 e CAB 734) a 0,79 entre os clones CAB 730 e CAB 187, com uma média geral de 0,5872. Observa-se, portanto, que a amostra total encerra genótipos bastante divergentes, embora deva se considerar que tal divergência é maximizada pela utilização exclusiva de *primers* polimórficos. De qualquer modo, a relativamente ampla divergência entre os clones examinados aponta para a possibilidade de obtenção de combinações híbridas com maior efeito heterótico e que ainda proporcionem maior segregação em recombinações. A obtenção de híbridos superiores, a partir de pais divergentes, resulta dos efeitos da dominância, além de efeitos epistáticos, que se manifestam na heterose (Falconer, 1981). Híbridos

resultantes do cruzamento de progenitores geneticamente próximos apresentam pouca complementaridade e baixo vigor, em função da baixa heterozigosidade alélica dos mesmos. Contrariamente, híbridos de pais divergentes tendem a diferir de forma crescente propiciando maior manifestação da heterose (Ghaderi *et al.*, 1984). No caso do cacau, são bem conhecidos os incrementos de produtividade associados ao cultivo de híbridos superiores (Dias, 2001a). Apesar da crescente utilização de clones resistentes em resposta ao avanço da vassoura-de-bruxa na Bahia, há um consenso de que a hibridação ainda será o principal método de melhoramento do cacau nas próximas décadas, nas principais regiões produtoras do mundo (Gotsch, 1997). Ensaios de campo demonstram que cacauzeiros híbridos apresentam grande adaptabilidade e rendimento superior a cultivares locais não melhorados (Dias *et al.*, 1998). A associação entre a performance heterótica de híbridos de cacau com a divergência entre seus progenitores já foi demonstrada a partir de dados fenotípicos (Dias & Kageyama, 1997) e posteriormente confirmada com marcadores RAPD (Dias *et al.*, 2003b). Desse modo, a informação fornecida pela matriz de dissimilaridade dos materiais estudados, associada a componentes da produtividade, pode proporcionar maior eficiência na escolha de progenitores divergentes, reduzindo consideravelmente o número de avaliações de cruzamentos requeridos. Por outro lado, os resultados sugerem que, embora os genótipos de cacau tendam a se agrupar segundo a sua origem geográfica, a divergência genética não guarda relação direta com a distância geográfica, conforme já observara Dias *et al.* (1997). Portanto, o empirismo baseado na distância geográfica, freqüentemente adotado na escolha aleatória de progenitores, deve ser evitado uma vez que nem sempre populações geograficamente distantes são mais divergentes que populações mais próximas.

3.6. Caracterização das populações analisadas

3.6.1. Análise das freqüências alélicas

As freqüências alélicas para cada loco dentro de cada subpopulação (Anexo, Tabela 3) mostram que a representação dos alelos presentes nas diferentes subpopulações permite distinguir todas as situações consideradas

importantes por Brown (1989) no estabelecimento de coleções nucleares. Correspondendo à categoria de alelos comuns dispersos, oito alelos (de 186) se fizeram presentes em todas as populações (mTcCIR04-270, mTcCIR06-198, mTcCIR08-285, mTcCIR13-254, mTcCIR18-345, mTcCIR49-205, mTcCIR54-156 e mTcCIR57-254), com uma frequência média de 0,70, sendo que o mTcCIR04-270 foi o alelo mais comum em todas as amostras, com uma frequência média de 0,57. Ainda nessa categoria merecem destaque alguns alelos bastante comuns que se mostraram ausentes em apenas uma subpopulação: mTcCIR17-266 (Acará-Pa), mTcCIR02-252 e mTcCIR55-233 (Ji-Paraná-RO) e mTcCIR25-192 e mTcCIR44-170 (Uatumã-AM). Na condição de alelos raros, foram detectados 68 alelos com frequência inferior a 0,05, identificando-se entre os mesmos desde alelos raros localizados, presentes em poucas amostras, até alelos raros dispersos, presentes em até 7 subpopulações (Tabela 9). Dentre esses devem ser destacados os alelos exclusivos de cada subpopulação, uma vez que se constituem alelos restritos às bacias de origem, onde podem ocorrer como alelos raros localizados ($f_i < 0,05$), ou mesmo como alelos comuns localizados ($f_i > 0,05$), conforme mostra a Tabela 10.

Tabela 9. Distribuição de alelos raros ($f_i < 0,05$) em função do número de populações em que os mesmos ocorrem.

1 sub-população		2 sub-populações		3 sub-populações		4 sub-populações		5 sub-populações		6 sub-populações		7 sub-populações	
Locus	Alelo	Locus	Alelo	Locus	Alelo	Locus	Alelo	Locus	Alelo	Locus	Alelo	Locus	Alelo
mTcCIR-01	129												
mTcCIR-08	277												
mTcCIR-09	256												
mTcCIR-09	262												
mTcCIR-09	295												
mTcCIR-10	212												
mTcCIR-13	258	mTcCIR-01	134										
mTcCIR-17	277	mTcCIR-07	154										
mTcCIR-18	375	mTcCIR-07	167										
mTcCIR-18	454	mTcCIR-07	169										
mTcCIR-26	266	mTcCIR-09	269										
mTcCIR-26	310	mTcCIR-10	199	mTcCIR-01	145								
mTcCIR-40	259	mTcCIR-11	285	mTcCIR-06	214								
mTcCIR-43	191	mTcCIR-12	186	mTcCIR-11	306	mTcCIR-07	165						
mTcCIR-54	154	mTcCIR-12	224	mTcCIR-12	220	mTcCIR-09	287						
mTcCIR-54	166	mTcCIR-25	172	mTcCIR-18	352	mTcCIR-11	287						
mTcCIR-57	247	mTcCIR-26	288	mTcCIR-33	325	mTcCIR-25	184						
mTcCIR-57	262	mTcCIR-26	304	mTcCIR-44	172	mTcCIR-33	318						
mTcCIR-58	232	mTcCIR-32	196	mTcCIR-49	204	mTcCIR-40	297						
mTcCIR-58	267	mTcCIR-33	293	mTcCIR-54	146	mTcCIR-42	212	mTcCIR-01	146				
mTcCIR-60	200	mTcCIR-49	209	mTcCIR-54	158	mTcCIR-42	218	mTcCIR-06	202				
mTcCIR-60	208	mTcCIR-54	162	mTcCIR-55	223	mTcCIR-43	213	mTcCIR-13	256				
mTcCIR-60	224	mTcCIR-57	242	mTcCIR-58	252	mTcCIR-49	198	mTcCIR-60	194	mTcCIR-40	281	mTcCIR-60	206

Tabela 10. Identificação de bacias detentoras de genótipos portadores de alelos exclusivos, com indicação do respectivo loco, tamanho do alelo (pb), clone, genótipo e estado genotípico (homozigoto ou heterozigoto) e frequências alélicas em relação à população de origem (F_i/pop) e à amostra total ($F_i/total$).

Bacia de origem	Loco	Alelo	Clones	Genótipo	Estado	F_i/pop	$F_i/total$
Acre-AC	mTcCIR-09	262	CAB 176	262::287	Hete	0,023	0,003
	mTcCIR-43	191	CAB 167	191::191	Hom	0,046	0,006
	mTcCIR-58	267	CAB 135	257::267	Hete	0,068	0,009
			CAB 136	267::267	Hom		
mTcCIR-60	208	CAB170	186::208	Hete	0,023	0,003	
Tarauaca-AC	mTcCIR-10	212	CAB 181	212::212	Hom	0,200	0,012
			CAB 187	212::212	Hom		
	mTcCIR-13	258	CAB 183	258::258	Hom	0,550	0,032
			CAB 184	258::258	Hom		
			CAB 185	254::258	Hete		
			CAB 187	258::258	Hom		
			CAB 200	254::258	Hete		
			CAB 201	258::258	Hom		
	mTcCIR-18	454	CAB 183	345::454	Hete	0,100	0,006
			CAB 185	345::454	Hete		
mTcCIR-54	154	CAB 184	154::154	Hom	0,250	0,016	
		CAB 185	154::154	Hom			
		CAB 187	154::158	Hete			
Purus-AC/Am	mTcCIR-09	256	CAB 102	256::284	Hete	0,033	0,003
	mTcCIR-17	277	CAB 102	268::277	Hete	0,033	0,003
	mTcCIR-57	247	CAB 102	247::256	Hete	0,033	0,006
Chandles-AC	mTcCIR-40	259	CAB 120	259::259	Hom	0,333	0,012
Ji-PArana-RO	mTcCIR-08	277	CAB 178	277::289	Hete	0,056	0,003
Jamari-RO	mTcCIR-57	262	CAB 221	254::262	Hete	0,100	0,009
			CAB 232	254::262	Hete		
			CAB 233	254::262	Hete		
Japura-AM	mTcCIR-09	295	CAB 500	273::295	Hete	0,036	0,003
	mTcCIR-26	310	CAB 481	292::310	Hete	0,036	0,003
	mTcCIR-54	166	CAB 484	166::166	Hom	0,462	0,038
			CAB 487	156::166	Hete		
			CAB 493	166::166	Hom		
			CAB 499	166::166	Hom		
			CAB 513	156::166	Hete		
CAB 492	166::166	Hom					
CAB 500	166::166	Hom					
Solimoes-AM	mTcCIR-18	375	CAB 325	345::375	Hete	0,029	0,003
	mTcCIR-26	266	CAB 058	266::297	Hete	0,028	0,003
	mTcCIR-58	232	CAB 046	232::252	Hete	0,079	0,015
CAB 058			232::232	Hom			
Balbina-AM	mTcCIR-60	224	CAB 731	212::224	Hete	0,375	0,009
			CAB 732	212::224	Hete		
			CAB 734	212::224	Hete		
Maicuru-PA	mTcCIR-60	200	CAB 035	191::200	Hete	0,050	0,003
	mTcCIR-01	129	CAB 007	142::129	Hete	0,100	0,006
			CAB 028	142::129	Hete		

As populações que mais mostraram alelos exclusivos foram as do Acre e as do Amazonas, destacando-se as bacias do Rio Acre e do Rio Tarauacá,

com quatro alelos privados cada, seguidas das bacias do Rio Purus, Rio Japurá e Rio Solimões, com três alelos cada. É interessante observar que embora a maioria dos alelos exclusivos apresente-se como raros, todos os alelos exclusivos da bacia do Rio Tarauacá apresentam-se como alelos comuns localizados ($f_i > 0,05$), com destaque para o alelo mTcCIR13-258 que está presente em sete genótipos ($f_i = 0,55$), tanto em estado homozigoto como heterozigoto. Situação semelhante se observa na bacia do Rio Japurá, onde o alelo mTcCIR54-166 se apresenta em sete genótipos ($f_i = 0,46$). É possível que o número de alelos raros e privados detectados guarde relação com o tamanho das amostras, o que justificaria a ausência ou baixo número de alelos exclusivos em amostras pouco representadas como Tocantins ($n=2$), Xeriuni ($n=2$), Chnadless ($n=3$), Uatumã ($n=4$) e Acará ($n=6$). A distribuição de todos os alelos raros (localizados ou dispersos) segundo a bacia de origem (Tabela 11) reforça a limitação imposta pelo tamanho das amostras. Porém, a ocorrência de um alelo privado (mTcCIR60-224) em três dos quatro clones representantes do Rio Uatumã-AM e um alelo (mTcCIR40-259) em um dos três clones coletados no Rio Chandless-AC recomenda ampliar a caracterização dessas amostras e, se for o caso proceder a novas prospecções nas referidas áreas. Por outro lado, quando se leva em conta as amostras mais bem representadas, observa-se que a maior ocorrência de alelos privados nas bacias do Alto Amazonas e do Acre sugerem uma maior diversidade nas regiões mais periféricas da Amazônia. Isso que parece reforçar as hipóteses de que a diferenciação das populações resulta do isolamento imposto por barreiras ribeirinhas (Patton *et al.*, 1994) e/ou pelo suposto grande Lago Amazônico (Klamer, 1984, citado por Dias, 2001b). Ambas as hipóteses, baseadas em evidências geológicas e biológicas, explicariam a maior diversidade da maioria das espécies na periferia (terras altas) do que na região central da Amazônia (Bush, 1994), sugerindo um fluxo gênico centrípeto, das bordas para o interior da bacia amazônica (Marroig & Cerqueira, 1997).

3.6.2. Análise da variabilidade genética nas populações

Visando evitar inferências viesadas decorrentes da inclusão de amostras muito pequenas, a partir desse ponto todas as análises se referem apenas às

subpopulações representadas por amostras que reúnam pelo menos nove indivíduos. Essas amostras quando submetidas ao teste exato de permutação (Guo & Thompson, 1992) de equilíbrio de Hardy-Weinberg (EHW) mostram que a despeito do pequeno tamanho das mesmas em relação ao número de alelos analisados, todos os locos aderiram satisfatoriamente às proporções do modelo de EHW, dentro de cada população, como demonstra a Tabela 11. Por outro lado, quando o mesmo modelo é submetido ao teste global das amostras conjunto, apenas os locos mTcCIR02 e mTcCIR04 se mantêm em equilíbrio de Hardy-Weinberg, (Tabela 11). Isso indica a estruturação da população analisada, sugerindo a ação de forças evolucionárias como deriva genética ou endogamia. Há que se ressaltar, porém, que o teste aplicado é bastante conservador, em decorrência da correção de Bonferroni para testes múltiplos, a qual consiste em dividir o nível nominal de probabilidade do teste (5%) pelo número de testes (n° amostras x n° locos). Assim, o nível nominal ajustado de probabilidade assume valores muito baixos para a rejeição da hipótese de nulidade, à medida que aumenta o número de testes. De qualquer modo, os resultados do teste exato mostram-se mais confiáveis que testes baseados em χ^2 ou razão de verossimilhança (G2-valor), quando a análise envolve amostras pequenas e locos multialélicos, onde é muito comum a ocorrência de classes genotípicas de baixa frequência.

A Tabela 12 apresenta diversos indicadores de diversidade caracterizados nas diferentes subpopulações com os 29 locos microssatélites. A percentagem de locos polimórficos foi relativamente alta, com a maioria das amostras apresentando polimorfismo em mais de 90% dos locos. As exceções foram as bacias do Rio Ji-Paraná-RO e do Maicuru-PA que apresentaram apenas 79,3% e 89,7% de locos polimórficos, respectivamente. As amostras com maior número de alelos foram as representantes das bacias do Amazonas e do Acre, contrastando com a pobre diversidade alélica das bacias de Rondônia. A subpopulação de composição alélica mais rica (Rio Solimões, com 124 alelos) apresenta, excluídos os 29 alelos automaticamente presentes devido aos 29 locos, 61,3% $[(124-29)/(184-29)]$ da diversidade alélica total (%DA) encontrada. Em contraste, a amostra da bacia do Rio Ji-Paraná-RO, com apenas 67 alelos, apresentou apenas 24,5% da diversidade alélica total.

Essa tendência, logicamente é acompanhada pelo número médio de alelos por loco (n_a). Entretanto, quando se analisa o número efetivo de alelos (n_e), destacam-se as amostras do Rio Tarauacá-AC, Rio Solimões e Solimões/Bx. Japurá-AM (n_e : 2,4 a 2,5) seguidas das demais subpopulações do Acre, Rio Japurá-AM, Maicuru-PA e Jarí-AP (n_e : 2,0 a 2,2) e por último as amostras de Rondônia (n_e : 1,6).

Tabela 12. Estatísticas de diversidade obtidas de onze subpopulações de cacaueiros pela análise de 29 locos microssatélites, mostrando o tamanho da amostra (n), percentagem de locos polimórficos (P_{95}), total de alelos (N_a), percentagem da diversidade alélica ($\%Da$), número de alelos por loco (n_a), número efetivo de alelos (n_e), riqueza alélica ($R_{(16)}$), heterozigosidades observada (H_o), e esperada sem viés (H_{nb}) (Nei, 1978), coeficiente de endogamia (f) e taxa de cruzamento (t).

Subpopulação	n	$P_{95\%}^1$	N_a	$\%Da^2$	n_a	n_e	$R_{(16)}$	H_o	H_{nb}	f^3	t
Acre-AC	22	93,1	100	45,8	3,4	2,1	2,807	0,285	0,454	0,378*	0,451
Iaco-AC	15	100,0	100	45,8	3,4	2,0	2,967	0,298	0,459	0,359*	0,472
Purus-AC/Am	15	96,6	103	47,7	3,6	2,1	3,094	0,291	0,464	0,380*	0,449
Tarauacá-AC	10	100,0	104	48,4	3,6	2,5	3,415	0,325	0,580	0,453*	0,376
Jari-AP	10	96,6	85	36,1	2,9	2,1	2,828	0,341	0,499	0,329*	0,505
Maicuru-PA	10	89,7	86	36,8	3,0	2,2	2,874	0,352	0,479	0,276*	0,567
Jamari-RO	15	100,0	78	31,6	2,7	1,6	2,313	0,140	0,312	0,559*	0,283
Ji-Paraná-RO	09	79,3	67	24,5	2,3	1,6	2,273	0,170	0,321	0,486*	0,346
Japurá-AM	14	96,6	102	47,1	3,5	2,2	3,170	0,261	0,506	0,494*	0,339
Solimões-AM	19	100,0	124	61,3	4,3	2,5	3,595	0,334	0,581	0,432*	0,397
S.B.Japurá-AM	14	96,6	101	46,5	3,5	2,4	3,199	0,327	0,525	0,387*	0,442
Amostra Total	144	100	184	100,0	6,34	3,4	4,228	0,285	0,658	0,408*	0,420

1 Considerado polimórfico quando a frequência do alelo mais comum não exceder 95%;

2 % de alelos que excede o número automaticamente esperado pelo número de locos (29);

3 Nível de significância para $f > 0$ determinado após 10 000 permutações e intervalo de confiança de 95% após 10 000 bootstrap (Goudet, 2001).

A utilização de todos esse índices, baseados no número de alelos e percentagem de locos polimórficos, é de grande importância para propósitos de conservação. O grande problema desses indicadores de diversidade é que dependem fortemente do tamanho das amostras analisadas e, portanto, não podem ser comparados com outros estudos. Daí a importância da utilização de medidas de riqueza alélicas padronizadas como a proposta por El Mousadik & Petit (1996). Essa medida, baseada no método de rarefação de Hurlbert (1971), é usada para estimar a riqueza alélica de um dado loco para um fixado

tamanho de amostra. Sob esse método, a riqueza alélica $R_{(g)}$ corresponde ao número de diferentes alelos encontrados quando g genes (tamanho especificado da amostra) são amostrados no loco em questão. No presente estudo, a riqueza alélica $R_{(16)}$ foi padronizada a partir de uma amostra de 16 genes, que corresponde ao número (8) de indivíduos completamente caracterizados na menor amostra (Ji-Paraná-RO). Em cada população a probabilidade de amostrar cada um dos diferentes alelos de um dado loco é computada como se apenas 16 genes (8 indivíduos) tivessem sido escoreados. $R_{(16)}$, portanto, resulta da soma destas probabilidades que corresponde ao número esperado de diferentes alelos em uma amostra observada de 16 genes. Desse modo, para cada população o número médio de alelos por loco para amostras de tamanho uniforme de 16 genes pode ser comparado com os resultados observados nas amostras de tamanhos diferentes. Vê-se agora que as subpopulações do Amazonas, juntamente com a do Rio Tarauacá-AC se destacaram mais em relação às demais, ao mesmo tempo em que as amostras do Baixo Amazonas (Jarí e Maicuru) se equipararam às demais bacias do Acre, ratificando-se a composição mais pobre das subpopulações de Rondônia (Jamari e Ji-Paraná). Observa-se que esse resultado guarda estreita relação com o obtido com o número efetivo de alelos ($R^2 = 0,93$) e com a diversidade ($R^2 = 0,91$) mostrando que a variação de tamanhos entre as amostras analisadas não compromete as inferências baseadas nesses índices de diversidade. A diversidade genética (H_{nb}) variou de 0,31 (Jamari-RO) a 0,58 (Solimões-AM) para uma diversidade global de 0,658. Novamente se destacam as amostras do Amazonas (H_{nb} : 0,51-0,58), juntamente com a de Tarauacá-AC (H_{nb} : 0,58), contrastando com as baixas diversidades genéticas das amostras de Rondônia (H_{nb} : 0,31-0,32). Nesse ponto merece destaque a relativamente alta diversidade apresentada pelas subpopulações do Baixo Amazonas (Maicuru-PA: 0,48 e Jarí-AP: 0,50), em comparação às demais amostras do Acre (H_{nb} : 0,45-0,46), geralmente tidas com mais polimórficas. Com respeito à heterozigiosidade observada (H_o), confirmam-se os baixos valores das subpopulações de Jamari-RO (0,14) e Ji-Paraná-RO (0,17), enquanto os maiores valores estiveram associados às subpopulações do baixo Amazonas (Maicuru, 0,35 e Jarí, 0,34), mas ainda comparáveis às amostras do Rio

Tarauacá-AC (0,32) e às amostras do Amazonas (0,33), a exceção da bacia do Rio Japurá-AM que apresentou valor mais baixo (0,261).

Praticamente todos os indicadores de diversidade apresentados na Tabela 12 ratificam a grande homogeneidade morfológica e baixa variabilidade fenotípica das subpopulações de Rondônia, relatadas por Almeida & Almeida (1987). De igual modo, os indicadores de diversidade das amostras representativas de subpopulações do Acre mostram-se coerentes com a grande diversidade fenotípica para caracteres morfológicos, agronômicos e moleculares reportados para acessos coletados nessa região (Almeida *et al.*, 1987; Barriga *et al.*, 1985; Pires *et al.* 2000; Figueira *et al.* 2000). A alta diversidade das amostras proveniente do baixo Amazonas confirma semelhante observação de Sereno (2001), embora seu estudo não incluía nenhuma população do Amapá. Esses resultados parecem conflitar com a suposta grande homogeneidade de populações típicas do grupo Forasteiros Baixo Amazônicos (Cheesman, 1944), corroborada por caracterizações moleculares (Lanaud, 1987; Lanaud *et al.* 1999b; N'Goran *et al.* 2000). Entretanto, as amostras relativas à bacia do Maicuru-PA, coletadas em plantios formados a partir de sementes de cacauzeiros silvestres, apresentam bom nível de variabilidade, quanto a produção, forma e peso de frutos, tamanho e cor de sementes (inclusive brancas) (Machado, 1977, citado por Almeida *et al.* 1995). A variabilidade em frutos se traduz nas amplitudes de peso de sementes úmidas (100 a 200g), número de sementes (26 a 52) e peso individual de sementes úmidas (2,61 a 4,25g) (Bartley, 1977, citado por Almeida *et al.*, 1995). A subpopulação da região do Rio Jarí-AP, embora considerada uniforme pelos coletores, apresentava considerável variabilidade de frutos, cujas formas apresentavam-se como cilíndricas, arredondadas ou amelonadas, com tipos semelhantes ao cacau “Comum”, Maranhão Liso” ou “Maranhão” da Bahia, até tipos bem diferentes semelhantes ao Scavina; a casca dos frutos variavam ainda quanto a rugosidade, coloração e espessura e as sementes apresentavam formas (achatadas, alongadas, arredondadas ou cilíndricas) e colorações (violeta claro ou violeta escuro) diversas (Vello e Silva, 1968). Além dessas evidências, não se pode desconsiderar a provável intensificação do fluxo gênico através da ação antrópica nessas regiões. Já no início do século

passado, suspeitava-se que todo o cacau existente entre o estuário amazônico e Óbidos e Santarém, no Pará, não era mais silvestre (Huber, 1901). Suspeita semelhante foi levantada por Bartley *et al.* (1988), em relação às populações existentes no município de Alenquer (Pa), na zona compreendida pela amostra que representa a bacia do Maicuru-PA.

O coeficiente de endogamia ou de fixação (f) apresentou valores positivos e significativamente diferentes de zero em todas as amostras, refletindo o déficit de heterozigotos observado nas mesmas e indicando desvios de panmixia provavelmente resultantes de endogamia ou deriva genética. Os altos e significativos valores do coeficiente de endogamia demonstram o quão conservador foi o resultado do teste de aderência ao equilíbrio de Hardy-Weinberg (Tabela 11), o qual não foi capaz de detectar desvios de panmixia em quaisquer dos locos examinados. A taxa estimada de fecundação cruzada para o conjunto das onze subpopulações foi de 42%, variando de 28,3% (Jamari-RO) a 56,7% (Maicuru-PA). Esses valores sugerem predomínio de auto-fecundação e, portanto, de materiais auto-compatíveis, o que parece estranho para uma espécie considerada predominantemente de fecundação cruzada devido à sua morfologia floral e à ocorrência de um sistema de auto-incompatibilidade (Cope, 1976). A sub-estimação das taxas de fecundação cruzada não deve ser descartada, uma vez que são estimadas a partir do coeficiente de endogamia, cuja estimativa pode envolver erros de amostragem, como ainda ser superestimada em decorrência da homoplasia que não permite distinção de alelos idênticos por descendência dos idênticos por estado (Slatkin, 1995). Além disso, não se pode descartar a possibilidade de a homozigosidade estar superestimada em decorrência presença de alelos nulos. Entretanto, déficits de heterozigotos, associados a altos valores de F_{IS} , em populações do alto e do baixo Amazonas, têm sido detectados por RFLP (N’Goran *et al.*, 2000). Por outro lado, Lanaud *et al.* (1987) têm reportado taxas de autofecundação de até 96 % em clones tidos como auto-incompatíveis. Além disso, a limitada distância de vôo (menos de 60 m) da mosquinha (*Forcipomya* sp), principal agente polinizador do cacauzeiro (Posnette, 1950), restringe o fluxo de pólen a árvores vizinhas, favorecendo a formação de grupos de famílias. Esse fluxo restrito, segundo Dias *et al.* (2003a),

intensificaria o efeito da vizinhança, aumentando a endogamia devido ao cruzamento entre parentes.

3.6.3. Análise da variabilidade dentro e entre as populações

A análise a seguir fundamenta-se nos coeficientes de fixação de Wright (1951), estimados segundo os modelos propostos por Nei (1973, 1987) (Tabela 13) e por Weir & Cockerhan (1984) (Tabela 14). Independentemente da abordagem adotada, verifica-se que embora as populações se mostrem fortemente diferenciadas, a variabilidade genética predomina dentro das populações. De acordo com as estatísticas de Nei, no conjunto das onze subpopulações, para uma diversidade total independente do número de amostras (H_T') de 0,676, a diversidade dentro das amostras foi $H_S = 0,479$, enquanto a diferenciação relativa entre as subpopulações foi $G_{ST}' = 0,292$ (Tabela 13). Portanto, segundo esses estimadores, a variabilidade dentro das subpopulações responde por 70,8% da variabilidade total, enquanto a divergência entre as mesmas representa apenas 29,2%. A divergência estimada por $F_{ST} = 0,290$ (Tabela 14), significativamente diferente de zero, ratifica a distribuição da variabilidade, ao mesmo tempo em que mostra uma alta e significativa divergência entre as subpopulações das bacias hidrográficas analisadas.

Os valores de $G_{IS} = 0,407$ (Tabela 13) de $F_{IS} = 0,408$ (Tabela 14) mostram-se concordantes e elevados, ratificando o alto déficit de heterozigose ($H_O = 0,284 \times H_S = 0,479$), conforme já havia sido caracterizado no item anterior. A comparação dos índices de fixação dentro das subpopulações ($F_{IS} = 0,408$) e para toda a amostra ($F_{IT} = 0,580$) sugere que o alto nível de endogamia reflete mais o sistema de reprodução que o efeito da deriva genética. Isso, entretanto, deve ser considerado com reservas, ante a possibilidade de superestimação dos coeficientes de endogamia, conforme já discutido anteriormente. Por outro lado, esse resultado é reforçado pelo alto valor assumido pelo coeficiente de parentesco $R_{EL} = 0,367$ (Tabela 14). Esse coeficiente proposto por Hamilton (1971), calculado pela relação $R_{EL} = 2F_{ST}/(1+F_{IT})$, é uma medida do parentesco dos indivíduos dentro das amostras quando comparados com o conjunto das subpopulações.

Tabela 13. Indicadores de diversidade de Nei, estimados para o conjunto das onze subpopulações, a partir de 29 locos microsatélites.

Locos	H _O	H _S	H _T	H _T '	G _{ST}	G _{ST} '	G _{IS}
mTcCIR-01	0,284	0,364	0,646	0,674	0,437	0,461	0,218
mTcCIR-02	0,508	0,389	0,499	0,510	0,220	0,237	-0,305
mTcCIR-04	0,804	0,513	0,581	0,588	0,118	0,128	-0,568
mTcCIR-06	0,288	0,470	0,621	0,636	0,243	0,261	0,388
mTcCIR-07	0,219	0,518	0,754	0,778	0,313	0,334	0,577
mTcCIR-08	0,147	0,450	0,624	0,642	0,280	0,299	0,673
mTcCIR-09	0,203	0,430	0,710	0,738	0,394	0,417	0,529
mTcCIR-10	0,138	0,595	0,732	0,746	0,188	0,203	0,768
mTcCIR-11	0,313	0,569	0,739	0,756	0,230	0,247	0,450
mTcCIR-12	0,472	0,606	0,805	0,825	0,247	0,265	0,221
mTcCIR-13	0,187	0,304	0,502	0,521	0,394	0,417	0,386
mTcCIR-17	0,026	0,438	0,558	0,570	0,215	0,232	0,940
mTcCIR-18	0,232	0,275	0,380	0,390	0,275	0,295	0,158
mTcCIR-19	0,418	0,592	0,750	0,766	0,211	0,227	0,294
mTcCIR-25	0,386	0,677	0,782	0,793	0,134	0,146	0,429
mTcCIR-26	0,293	0,529	0,780	0,805	0,321	0,342	0,447
mTcCIR-32	0,285	0,317	0,477	0,493	0,334	0,356	0,104
mTcCIR-33	0,470	0,689	0,868	0,886	0,206	0,222	0,318
mTcCIR-35	0,323	0,446	0,659	0,680	0,323	0,345	0,275
mTcCIR-40	0,147	0,570	0,785	0,806	0,273	0,292	0,743
mTcCIR-42	0,226	0,492	0,804	0,835	0,389	0,412	0,540
mTcCIR-43	0,315	0,627	0,822	0,841	0,237	0,255	0,497
mTcCIR-44	0,357	0,483	0,693	0,714	0,303	0,324	0,261
mTcCIR-49	0,132	0,356	0,443	0,451	0,195	0,211	0,630
mTcCIR-54	0,173	0,465	0,562	0,572	0,173	0,188	0,628
mTcCIR-55	0,308	0,335	0,525	0,544	0,362	0,384	0,081
mTcCIR-57	0,136	0,409	0,477	0,484	0,142	0,154	0,667
mTcCIR-58	0,141	0,473	0,720	0,745	0,343	0,365	0,702
mTcCIR-60	0,308	0,507	0,791	0,820	0,359	0,381	0,392
Geral	0,284	0,479	0,658	0,676	0,272	0,292	0,407

H_O: Heterozigosidade observada; H_S: Diversidade genética dentro das subpopulações; H_T: Diversidade genética total; H_T': Diversidade genética total independente do número de amostras; G_{ST}: Coeficiente de diferenciação relativa entre as amostras; G_{ST}': Coeficiente de diferenciação relativa independente do número de amostras.

Tabela 14. Estimativas dos coeficientes de fixação de Wright, segundo Weir & Cockerhan (1984) e coeficiente de parentesco segundo Queller & Goodnight (1989).

Locos	σ_A^1	σ_B	σ_W	F_{IT}	F_{ST}	F_{IS}	R_{EL}
mTcCIR-01	0,292	0,093	0,275	0,584	0,443	0,253	0,560
mTcCIR-02	0,113	-0,144	0,540	-0,059	0,222	-0,362	0,473
mTcCIR-04	0,071	-0,303	0,810	-0,402	0,123	-0,599	0,411
mTcCIR-06	0,189	0,187	0,265	0,586	0,294	0,414	0,371
mTcCIR-07	0,244	0,304	0,222	0,712	0,317	0,578	0,371
mTcCIR-08	0,194	0,289	0,131	0,787	0,316	0,689	0,353
mTcCIR-09	0,293	0,227	0,235	0,688	0,388	0,491	0,459
mTcCIR-10	0,149	0,456	0,131	0,822	0,203	0,777	0,222
mTcCIR-11	0,201	0,240	0,303	0,593	0,270	0,443	0,339
mTcCIR-12	0,216	0,133	0,467	0,428	0,265	0,222	0,371
mTcCIR-13	0,204	0,108	0,164	0,655	0,427	0,397	0,516
mTcCIR-17	0,125	0,408	0,033	0,942	0,221	0,925	0,227
mTcCIR-18	0,127	0,032	0,253	0,386	0,308	0,113	0,444
mTcCIR-19	0,174	0,174	0,431	0,446	0,223	0,287	0,308
mTcCIR-25	0,112	0,299	0,397	0,508	0,138	0,429	0,183
mTcCIR-26	0,263	0,247	0,285	0,642	0,331	0,465	0,403
mTcCIR-32	0,197	0,025	0,264	0,457	0,405	0,087	0,556
mTcCIR-33	0,198	0,231	0,453	0,487	0,225	0,338	0,302
mTcCIR-35	0,237	0,125	0,327	0,526	0,344	0,277	0,451
mTcCIR-40	0,220	0,449	0,152	0,814	0,268	0,747	0,295
mTcCIR-42	0,353	0,273	0,199	0,759	0,428	0,579	0,487
mTcCIR-43	0,211	0,302	0,333	0,606	0,249	0,475	0,310
mTcCIR-44	0,228	0,140	0,364	0,503	0,312	0,277	0,415
mTcCIR-49	0,106	0,214	0,160	0,667	0,221	0,572	0,265
mTcCIR-54	0,093	0,305	0,187	0,680	0,159	0,620	0,189
mTcCIR-55	0,197	0,027	0,299	0,427	0,376	0,082	0,527
mTcCIR-57	0,100	0,285	0,132	0,745	0,193	0,684	0,222
mTcCIR-58	0,266	0,346	0,152	0,801	0,348	0,694	0,387
mTcCIR-60	0,315	0,218	0,288	0,650	0,384	0,431	0,465
Geral²				0,580± 0,091	0,290± 0,031	0,408± 0,122	0,367± 0,038

1- F_{IT} (F): Deficit global de heterozigotos (índice de fixação global); F_{ST} (θ): deficit de heterozigotos entre populações (coeficiente de diferenciação); F_{IT} (f): deficit de heterozigotos dentro das populações (índice de fixação dentro das populações); R_{EL} : coeficiente de parentesco dos indivíduos dentro das populações; σ_A , σ_B e σ_W : componentes da variância entre amostras, entre indivíduos dentro das amostras e entre indivíduos.

2- Intervalo de confiança de 95% após 10.000 bootstrap

Os elevados coeficientes de divergência obtidos ($G'_{ST} = 0,292$ e $F_{ST} = 0,290$), de acordo com classificação proposta por Wright (1965), citado por Nei (1972), sugerem uma diferenciação genética muito elevada entre as subpopulações analisadas e, consideravelmente superior à encontrada ($F_{ST} = 0,042$ e $G_{ST} = 0,080$) por Sereno (2001) para populações de cacau da Amazônia, por meio de microssatélites. Entretanto, Alves (2003), utilizando os mesmos marcadores, encontrou um coeficiente de divergência $F_{ST} = 0,307$, entre populações naturais de cupuaçu (*T. grandiflorum*) da Amazônia, espécie caracterizada pelo autor como perfeitamente alógama (taxa de fertilização cruzada = 1).

A inclusão de subpopulações do alto Japurá (Am), mais próximas ao provável centro de origem da espécie (Cheesman, 1944; Dias, 2001b), bem como a maior representatividade de subpopulações mais uniformes de Rondônia (Jamari e Ji-Paraná) e do Baixo Amazonas (Jari-AP), provavelmente contribuiu para a maior diferenciação genética observada no presente estudo em relação ao de Sereno (2001). Concorre ainda para a discrepância observada, a comparação, pela autora, de populações formadas por grupos de subpopulações pouco representadas, agregando-se ainda o fato de que a população denominada Alta Amazonas incluía a subpopulação da bacia do Rio Purus. Os genótipos representantes dessa bacia hidrográfica foram colhidos predominantemente na região fronteira do Acre com o Amazonas, muito próximo das bacias do Rio Acre e do Rio Iaco, que, aliás, são tributários do Rio Purus (Figura 1). A grande similaridade dos genótipos provenientes da bacia do rio Purus com os procedentes das demais bacias do Acre ficou evidenciado tanto na análise de correspondência (Figura 12), como na análise de agrupamento a partir da matriz de dissimilaridade entre os genótipos (Figura 13).

Observando-se o critério da proximidade geográfica, reforçado pelos resultados das análises de correspondência (Figura 12) e agrupamento (Figura 13) procedeu-se à análise comparativa das populações formadas a partir da reunião das diversas subpopulações em quatro grupos distintos: i) Alto Amazonas (Rio Solimões-AM, Rio Solimões/Bx. Japurá-AM e Rio Japurá-AM); ii) Baixo Amazonas (Rio Maicuru-PA e Rio Jari-AP); iii) Acre (Rio Acre-AC, Rio

Iaco-AC, Rio Purus-AC/Am e Rio Tarauacá-AC), e iv) Rondônia (Rio Jamari-RO e Rio Ji-Paraná-RO).

Os indicadores de diversidade para os referidos grupos (Tabela 15) confirmam a maior diversidade genética dentro do grupo Alto Amazonas ($H_S = 0,551$), seguida dos grupos do Baixo Amazonas ($H_O = 0,498$) e do Acre ($H_S = 0,483$). Mais uma vez se constata que a maior heterozigosidade observada no Baixo Amazonas ($H_O = 0,348$) não se traduz em maior riqueza alélica ($r(g) = 2,851$), conforme se observa nos grupos do Alto Amazonas ($r(g) = 3,321$) e do Acre ($r(g) = 3,071$). O grupo de Rondônia ratifica sua baixa diversidade sob quaisquer dos indicadores analisados, contrastando com os elevados valores e diversidade reportados por Sereno (2001) para o mesmo grupo ($H_S = 0,5096$). Se por um lado é pouco diverso, o grupo de Rondônia é o que apresenta populações mais divergentes ($F_{ST} = 0,349$), o que se reflete no maior parentesco entre os indivíduos que compõem suas subpopulações ($R_{EL} = 0,412$). Deve-se destacar, porém que o ambos os indicadores (F_{ST} e R_{EL}) não se mostraram diferentes de zero pelo P-valor obtido no teste exato com 10.000 permutações. De qualquer forma, considerando os valores de F_{ST} nos demais grupos, observa-se que a elevada divergência entre as duas subpopulações de Rondônia ($F_{ST} = 0,349$) contribuiu fortemente para o alto coeficiente de diferenciação obtido quando da análise das subpopulações isoladamente ($F_{ST} = 0,290$).

Tabela 15. Estatísticas de diversidade obtidas grupos de populações de cacauzeiros pela análise de 29 locos microssatélites, mostrando o tamanho da amostra (n), riqueza alélica ($R_{(16)}$), heterozigosidades observada (H_O), e diversidade intragrupo (H_S), coeficiente de endogamia (F_{IS}), Coeficiente de divergência (F_{ST}) e Coeficiente de parentesco (R_{EL}).

Grupo	n	$R_{(16)}$	H_O	H_S	F_{IS}	F_{ST}	R_{EL}
Baixo Amazonas	20	2,851	0,348	0,498	0,301	0,122	0,177
Rondônia	23	2,293	0,151	0,322	0,530	0,349	0,412
Acre	62	3,071	0,296	0,483	0,388	0,148	0,201
Alto Amazonas	47	3,321	0,310	0,551	0,437	0,135	0,179
P-valor		0,0089	0,0007	0,0156	0,0001	0,2168	0,2502

¹ P-valor para teste bi-caudal obtido após 10.000 permutações.

A análise hierárquica dos quatro grupos de populações através da análise de variância (Tabela 16) fornece valores significativos para os coeficientes de diferenciação entre os grupos ($F_{CT} = 0,1852$), entre as subpopulações dentro dos grupos ($F_{SC} = 0,1824$) e entre as subpopulações ($F_{ST} = 0,3338$). Esses valores confirmam a maior variabilidade dentro das subpopulações (66,6%), seguida da variabilidade entre grupos (18,5%), ficando a variação entre subpopulações dentro dos grupos com 14,9%.

Tabela 16. Análise de variância para os grupos de subpopulações segundo Weir & Cockerhan (1984), seguindo Excoffier *et al.* (1992)

Fonte de Variação	GL	Soma de quadrados	Componentes da variância	% de variação
Entre grupos (populações)	3	494,303	1,70549 Va	18,52
Entre subpopulações dentro dos grupos	7	309,891	1,36828 Vb	14,86
Dentro de Subpopulações	295	1809,470	6,13380 Vc	66,62
Total	305	2.613,663	9,20756	

Índices de Fixação

Fst:	0,3338	(P-valor p/ 10023 permutações = 0,0000)
Fsc:	0,1824	(P-valor p/ 10023 permutações = 0,0000)
Fct:	0,1852	(P-valor p/ 10023 permutações = 0,0000)

A elevada divergência entre as populações estudadas é confirmada quando se comparam os valores de F_{ST} entre pares de subpopulações (Tabela 17). As comparações duas a duas, todas significativamente diferentes de zero, mostram-se coerentes com os resultados obtidos na análise de correspondência dos genótipos individualmente (Figura 11), confirmando a menor divergência entre as bacias de cada grupo. Este resultado reforça a hipótese de que a diferenciação das populações naturais de cacau da Amazônia guarda relação com suas bacias hidrográficas. Por outro lado, a despeito da significativa diferenciação entre as subpopulações, de maneira geral a partição da variabilidade genética é nitidamente maior dentro do que

entre as amostras analisadas, como já observado em outros estudos com cacau, envolvendo microssatélites (Sereno, 2001), RAPD e RFLP (Russel, *et al.* 1993; Lerceteau *et al.*, 1997), isozimas (Ronning & Schnell, 1994; Sounigo *et a.*, 1997) e descritores morfológicos/quantitativos (Dias *et al.*, 2002; 2003a). Esses resultados sugerem que o acervo genético das coleções *ex situ* pode ser enriquecido com novas coletas que enfatizem a maior representatividade dentro das subpopulações com maior riqueza alélica (Alto Amazonas e Acre), agregando-se, porém o máximo de novas bacias visando tirar proveito da relativa divergência observada entre os grupos e entre as subpopulações.

Tabela 17. Coeficiente de diferenciação (F_{ST}) entre pares de subpopulações significativamente (5%) diferentes de zero.

Subpopulações	1	2	3	4	5	6	7	8	9	10	11
1-Rio Acre-AC	0,000										
2-Rio Jari-AP	0,339	0,000									
3-Rio Iaco-AC	0,141	0,318	0,000								
4-Jamari-RO	0,459	0,395	0,461	0,000							
5-Japurá-AM	0,352	0,262	0,313	0,445	0,000						
6-Ji-Paraná-RO	0,403	0,256	0,366	0,383	0,400	0,000					
7-Maicuru-PA	0,363	0,148	0,342	0,415	0,281	0,222	0,000				
8-Rio Purus-AC	0,148	0,306	0,150	0,465	0,346	0,376	0,287	0,000			
9-Solimões-AM	0,277	0,221	0,259	0,422	0,214	0,341	0,181	0,238	0,000		
10-SBJapurá-AM	0,316	0,266	0,308	0,483	0,213	0,397	0,244	0,301	0,037	0,000	
11-Tarauacá-AC	0,206	0,282	0,176	0,437	0,292	0,359	0,268	0,161	0,211	0,255	0,000

3.6.4. Contribuição de cada subpopulação à diversidade

Para efeito de conservação genética, tão importante quanto a distribuição da diversidade é a contribuição de cada população à diversidade total, conforme enfatizam Petit *et al.* (1998). Esses autores demonstram que a contribuição de uma dada população à diversidade total (diversidade de Nei, riqueza alélica, etc.) pode ser medida como a diferença entre a diversidade conjunta de todas as populações menos a diversidade de todas excluindo a população em questão. Esta contribuição pode ser dividida em dois componentes; um relacionado à diversidade da população e outro devido à sua

divergência em relação às demais. Adotando o mesmo procedimento, a decomposição da diversidade total e da riqueza alélica (padronizada para 16 genes), mostra que as contribuições relativas de algumas populações se alteram dependendo do parâmetro avaliado (Figura 14). Sob a ótica da diversidade genética, a subpopulação de Jamari-RO equipara-se com as populações do Amazonas e a de Tarauacá-AC. Quando se analisa a riqueza alélica total, a contribuição relativa é bem reduzida em relação às outras quatro, que continuam se destacando das demais. Observa-se que algumas subpopulações contribuem negativamente para a diversidade da população total. Isso decorre, ou porque a diversidade da subpopulação é menor que a diversidade média, ou porque a população é muito divergente. No geral, observa-se que três subpopulações do Acre (Acre, Iaco e Purus) e a de Ji-Paraná-RO apresentam contribuições relativas negativas para ambos os indicadores de diversidade total. Apenas Tarauacá-AC e as três subpopulações do Alto Amazonas apresentam contribuições positivas devido às suas próprias diversidades e riquezas alélicas, mas apenas a do Japurá apresenta essa contribuição maior que a dada pela diferenciação. É interessante notar que, embora apresentem riqueza alélica e diversidade elevadas, as populações do Acre (exceto Tarauacá), juntamente com a de Maicuru-PA, apresentam baixas contribuições à variabilidade total. Isso parece resultar da redundância dos alelos presentes nessas subpopulações, em relação ao *pool* de alelos da amostra total, como sugere a contribuição negativa de suas divergências.

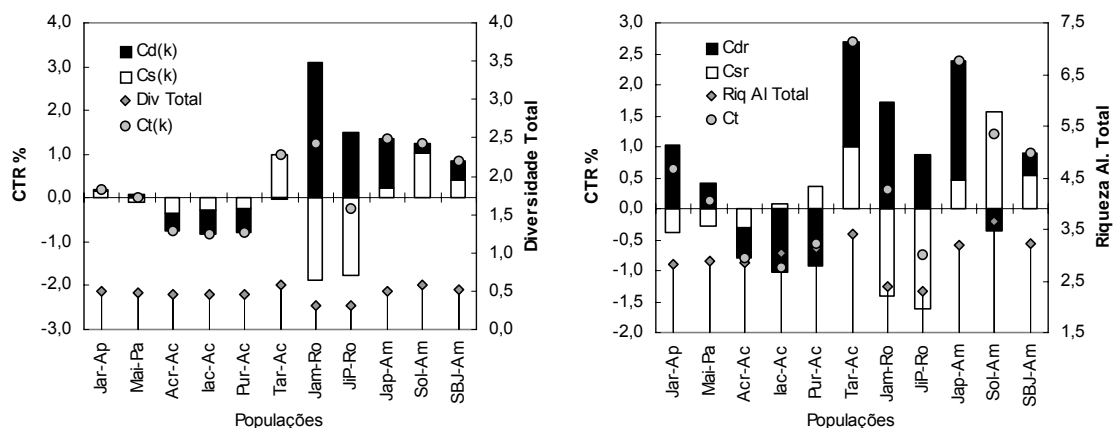


Figura 14. Contribuição relativa (CTR %), dos componentes diversidade e diferenciação à diversidade total (A) ou à riqueza alélica total, de cada sub-população.

Considerando que para fins de conservação, a preservação de alelos é mais importante que a manutenção de frequências alélicas (El Mousadik & Petit, 1996), os resultados recomendam novas coletas dentro das subpopulações que apresentaram maiores contribuições positivas à riqueza alélica total, devido às suas próprias riquezas alélicas (Csr) - Tarauacá-AC, Japurá, Solimões e Solimões/Bx Japurá (Am). Entretanto, a forte contribuição de várias subpopulações devido suas diferenciações (Cdr) recomenda que novas bacias sejam prospectadas visando incrementar a riqueza alélica das coleções de germoplasma.

3.7. Distâncias genéticas entre as populações

Embora sem a pretensão de definir o modelo que melhor se adequa ao estudo de microssatélites em cacau, avaliou-se a distância genética entre as subpopulações utilizando estimadores D (distância padrão de Nei, 1978), baseado no modelo IAM e $(\delta\mu)^2$ de Goldstein *et al.* (1995), baseado no modelo SMM (Tabela 18). E, tendo em vista a forte diferenciação entre as subpopulações revelada pelo F_{ST} , computou-se também a distância D_L de Reynolds *et al.* (1983), uma linearização de F_{ST} ($-\ln(1-F_{ST})$) (Tabela 19).

De maneira geral verifica-se boa concordância entre os três estimadores de distância genética, embora apresentam ligeiras discordâncias quando se analisa valores extremos de divergência entre os pares de subpopulações. As populações mais próximas foram Solimões-AM e Solimões/Bx Japurá-AM, com $D = 0,045$, $\delta\mu^2 = 2,2$ e $D_L = 0,052$. No outro extremo, enquanto D (1,0) e D_L (0,687) apontam como mais divergentes as amostras do Rio Jamari-RO e as do Solimões/Bx Japurá-AM, $\delta\mu^2$ (176,9) mostra maior distância entre Jamari-RO e Japurá-AM. De certo modo coerente com estreita distância apresentada pelas populações do Solimões e Solimões/Bx Japurá-AM. No geral, os indicadores de distância mostram a subpopulação do Rio Jamari-RO como a que mais divergiu das demais e confirmam a menor distância entre as bacias agrupadas pelo critério geográfico, como já se havia percebido nas análises de correspondência (Figura 12) e de dissimilaridade individual dos clones (Figura 13). Esse resultado mostra-se coerente com grande divergência, detectada por

Dias *et al.* (2002; 2003a), entre população de Rondônia (Ji-Paraná) e populações dos rios Amazonas e Japurá.

Tabela 18. Distâncias D (Nei, 1978), abaixo da diagonal e $\delta\mu^2$ (Goldstein *et al.*, 1995), acima da diagonal, para 29 locos microssatélites de 11 subpopulações naturais de cacauzeiros.

Subpopulações	1	2	3	4	5	6	7	8	9	10	11
1-Rio Acre-AC		64,4	4,6	91,8	62,5	33,8	52,8	8,3	36,0	54,3	22,5
2-Rio Jari-AP	0,555		59,7	109,2	29,9	72,5	14,4	74,2	29,9	26,6	61,1
3-Rio Iaco-AC	0,130	0,504		111,4	47,2	34,5	46,4	4,7	32,6	50,0	15,3
4-Jamari-RO	0,734	0,475	0,730		176,9	75,5	123,2	122,6	123,9	128,7	137,2
5-Japurá-AM	0,633	0,456	0,514	0,832		70,4	22,4	58,7	17,2	18,1	52,4
6-Ji-Paraná-RO	0,637	0,264	0,527	0,296	0,782		41,5	32,6	37,2	58,2	39,7
7-Maicuru-PA	0,623	0,160	0,553	0,545	0,485	0,216		48,1	15,7	17,6	31,1
8-Rio Purus-AC	0,141	0,491	0,148	0,795	0,633	0,600	0,430		31,3	51,6	10,4
9-Solimões-AM	0,466	0,391	0,434	0,878	0,369	0,664	0,289	0,381		2,2	31,4
10-SBJapurá-AM	0,518	0,462	0,501	1,082	0,313	0,791	0,383	0,480	0,045		48,5
11-Tarauacá-AC	0,315	0,601	0,273	0,786	0,662	0,652	0,545	0,261	0,450	0,530	

Tabela 19. Distâncias $D_L = -\ln(1-F_{st})$ (Reynolds *et al.*, 1983) para 29 locos microssatélites de 11 subpopulações naturais de cacauzeiros.

Subpopulações	1	2	3	4	5	6	7	8	9	10	11
1-Rio Acre-AC											
2-Rio Jari-AP	0,408										
3-Rio Iaco-AC	0,150	0,379									
4-Jamari-RO	0,611	0,496	0,625								
5-Japurá-AM	0,432	0,333	0,373	0,631							
6-Ji-Paraná-RO	0,548	0,324	0,501	0,464	0,580						
7-Maicuru-PA	0,449	0,171	0,412	0,549	0,355	0,288					
8-Rio Purus-AC	0,160	0,374	0,164	0,649	0,423	0,537	0,354				
9-Solimões-AM	0,318	0,260	0,293	0,558	0,245	0,457	0,218	0,271			
10-SBJapurá-AM	0,375	0,325	0,358	0,687	0,245	0,569	0,297	0,350	0,052		
11-Tarauacá-AC	0,263	0,354	0,232	0,583	0,369	0,497	0,346	0,225	0,253	0,317	

Os dendrogramas correspondentes às matrizes de distância (Figura 15), embora apresentem diferenças topológicas, confirmam a tendência de as subpopulações de bacias geograficamente mais próximas formarem grupos discretos. Embora o pequeno número de bootstraps (100) não sustente inferências seguras sobre as topologias obtidas, verifica-se uma certa

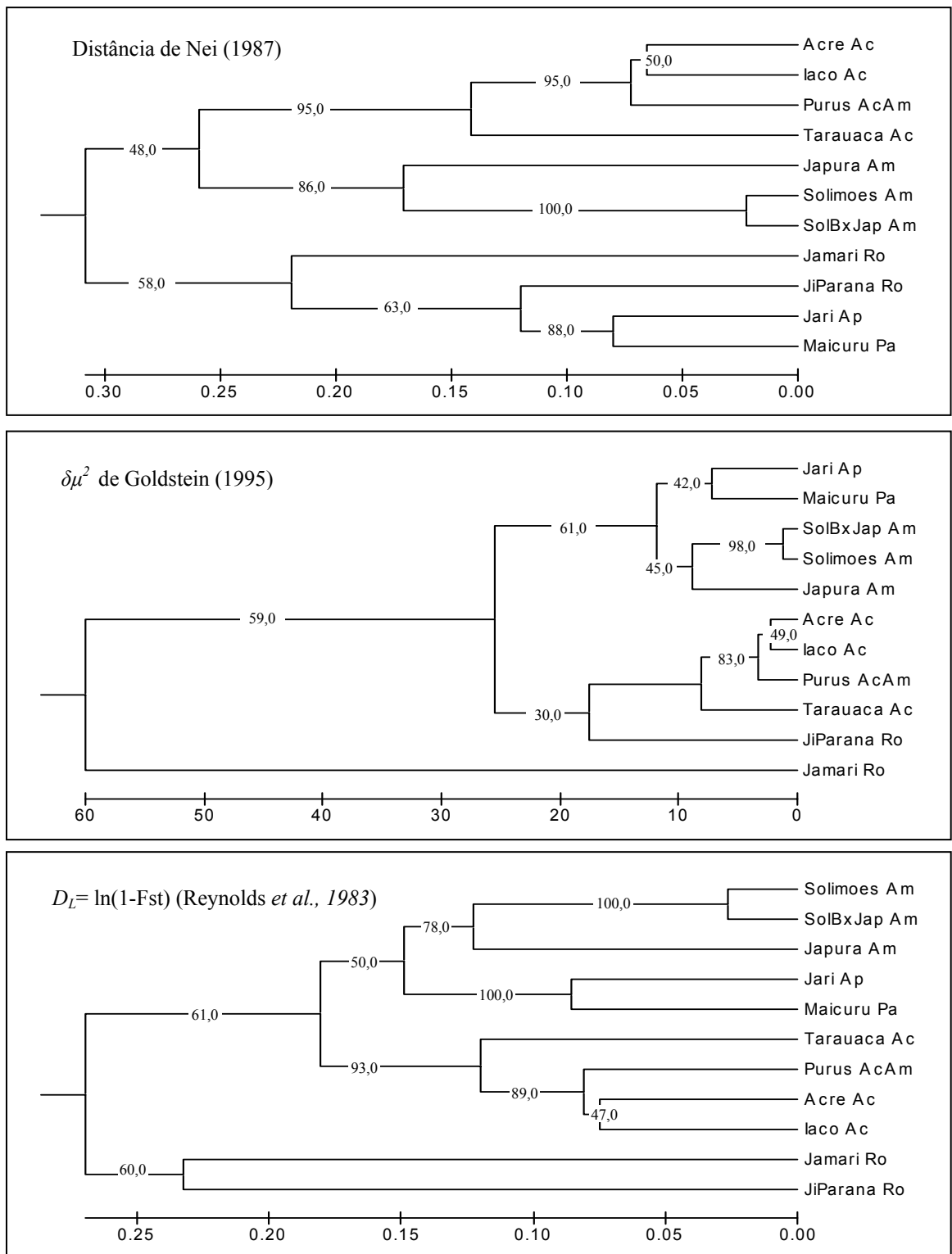


Figura 15. Dendrograma (UPGMA) de distâncias genéticas de sub-populações de cacau obtidas com 29 locos microssatélites. O número em cada nó representa o valor do bootstrap para 100 repetições, para os grupos à direita.

concordância na identificação dos grupos formados pelas subpopulações do Acre e do Alto Amazonas. As maiores discordâncias aparecem quanto ao agrupamento das subpopulações de Rondônia, embora todos destaquem a grande divergência da subpopulação do Jamari em relação às demais.

Segundo Felsenstein (1985), somente agrupamentos com valores de bootstraps maiores ou iguais a 95% podem ser considerados significativos. Por esse critério, sem esquecer as limitações estatísticas (pequeno número de bootstraps), o estimador de distância genética D (Nei, 1978) foi o que proporcionou topologia mais robusta. O dendrograma associado a D , bastante coerente com a dispersão bidimensional observada na Figura 12, permitiu identificar pelo menos três agrupamento significativamente distintos: i) Rio Acre, Iaco e Purus, com 95%; ii) Tarauacá (singular), com 95 % e , iii) Solimões e Solimões/Bx Japurá, com 100%. A distância D_L (Reynolds *et al.*, 1983), embora apresente topologia bastante coerente com as análises de correspondência (Figura 12) e com a grande divergência apresentada pela subpopulações de Rondônia, somente permitiu confirmar dois grupos com segurança estatística: i) Solimões e Solimões/Bx Japurá, 100% e, ii) Jarí e Maicuru, 100%. O estimador $\delta\mu^2$ (Goldstein *et al.*, 1995), de topologia menos robusta, somente ratificou o grupo formado pelas subpopulações do Solimões e do Solimões/Bx Japurá (98%). Esse resultado parece confirmar as limitações freqüentemente associadas ao uso de estimadores baseados em SMM para estudos de distância genética por microssatélites. A elevada variância de $\delta\mu^2$, associada à irregularidade no padrão mutacional dos microssatélites são apontados como grande limitações ao uso desse estimador distância (Nei & Kumar, 2000).

O padrão de agrupamento observado, tanto de indivíduos (Figuras 12 e 13), como das subpopulações (Figura 15), associado à distribuição da diversidade dentro e entre as subpopulações e populações agrupadas, enfatiza o papel desempenhado pelos rios na divergência alopátrica das populações naturais de cacau da Amazônia, já reconhecido por diversos autores (Pound, 1938; Almeida *et al.*, 1987; Young, 1994; Dias *et al.*, 2003a; Sereno, 2001). Semelhante associação entre diversidade e cursos d'água tem sido reportado para outras espécies amazônicas como *Hevea brasiliensis* (Seguin *et al.*, 1999)

e *Elaeis oleifera* (Barcelos, 1998) (ambos citados por Motamayor *et al.*, 2002). Esse padrão de diferenciação tem sido frequentemente associado à teoria do refúgio (Haffer, 1969; Simpson & Haffer, 1978). Essa teoria propõe que fortes alterações climáticas impostas pela glaciação no Pleistoceno provocaram ciclos de fragmentação e coalescência de áreas da floresta, constituindo refúgios isolados onde as espécies associadas evoluíam alopaticamente sob seleção natural e deriva genética. Desse modo, populações isoladas de cacauzeiros em tais refúgios, provavelmente constituindo florestas de galeria ao longo de cursos d'água dispersos, podem ter atravessado períodos climáticos adversos, durante o Quaternário, evoluindo alopaticamente antes da subsequente expansão da floresta, conforme sugerem Lanaud (1987) e Young (1994). A mesma teoria tem sido argüida por Lachenaud (1997) para a evolução de populações de cacau nativo na Guiana Francesa.

A forte divergência revelada pelo coeficiente de diferenciação (Tabela 17), mesmo entre amostras coletadas em bacias adjacentes, parece reforçar o papel dos refúgios na evolução das populações nativas de cacau na Amazônia. Entretanto, a alta divergência entre as populações de Rondônia (Jamari e Ji-Paraná) parece resultar de eventos mais recentes. As amostras coletadas naquelas áreas, além de muito próximas, são em sua maioria provenientes de florestas de terra firme, constituindo vastas comunidade sem qualquer barreira geográfica aparente, levando Almeida (1982) a sugerir tratar-se de um só complexo gênico. Entretanto, a julgar pela alta uniformidade interna dessas subpopulações, é provável que a sua divergência resulte da intervenção humana através de cultivos e seleção. Concentrações espontâneas de cacau em áreas de até 10 ha na região de Mirante da Serra, incluída na amostra de Ji-Paraná-RO, são indícios da participação humana através de cultivos (Almeida, 2001). Por outro lado, vestígios de ocupação humana na forma de fragmentos de cerâmica e ferramentas etc., têm sido registrados às proximidades de algumas populações silvestres amostradas na bacia do Jamari-RO, sugerindo intervenção de comunidades indígenas primitivas na dispersão e estabelecimento de populações de cacau nessa região (Almeida, 1982).

4. CONCLUSÕES

A utilização de locos polimórficos de microssatélites mostrou-se bastante informativa para a identificação molecular individual de genótipos de cacau (*Theobroma cacao* L.), bem como na caracterização da variabilidade genética de populações naturais de cacau.

Dos 30 locos avaliados, 29 apresentaram-se suficientemente polimórficos (6,4 alelos/loco), seletivamente neutros e independentes, permitindo identificar genótipos com elevado poder discriminação ($PD = 1$) e baixíssima probabilidade de identidade ($PI = 2,8 \cdot 10^{-29}$). Tal informatividade permitiu identificar como distintos 169 (98,3%) dos 172 clones amostrados em populações espontâneas de 16 bacias hidrográficas da Amazônia Brasileira. Apenas os clones CAB 731, CAB 732 e CAB 734, oriundos da bacia do rio Uatumã-AM, mostraram-se indistintos, sugerindo tratar-se de genótipos repetidos. Assim caracterizados, o genótipos agruparam-se segundo suas origens geográficas, com a dissimilaridade variando de 0,0 a 0,79 e média de 0,59.

Foram identificados 68 alelos raros, dos quais 23 exclusivos, destacando-se as bacias do Acre e do Alto Amazonas com maior número de alelos privados, corroborando suas maiores riquezas alélicas. Todas as subpopulações apresentaram boa diversidade genética, exceto as de Rondônia que mostraram menor variabilidade para todos os indicadores utilizados. A diversidade genética total foi 0,658 contra uma heterozigosidade observada de 0,295, observando-se déficit de heterozigosidade em todas as amostras, caracterizado por um coeficiente de endogamia de 0,408 e taxa de fecundação cruzada de 42 %. As subpopulações apresentaram forte estruturação, com significativa diferenciação ($G_{ST}' = 0,292$ e $F_{ST} = 0,290$) e alto coeficiente de parentesco ($R_{EL} = 0,367$) dos indivíduos dentro das subpopulações. A análise das amostras agrupadas (Acre, Alto Amazonas, Rondônia e Baixo Amazonas) mostra 66,6% da variabilidade genética dentro das subpopulações, 14,9 % entre subpopulações dentro dos grupos e 18,5% entre os quatro grupos. As populações do Alto Amazonas se destacaram em termos de contribuição relativa à diversidade e riqueza alélica total. No acre, a despeito da alta

diversidade e riqueza alélica, apenas a bacia do Rio Tarauacá apresenta contribuição relativa positiva. As demais demonstram redundância na composição alélica, com contribuições relativas negativas. O dendrograma de distância genética D de Nei apresentou topologia mais robusta que $\delta\mu^2$ e D_L ($-\ln(1-F_{ST})$), confirmando o agrupamento das subpopulações do Acre e Alto Amazonas, destacando a elevada divergência da bacia do Rio Jamari em relação às demais.

A grande variabilidade e riqueza alélica dentro das subpopulações, bem como a grande ocorrência de alelos raros e exclusivos (raros ou dispersos), associado à considerável divergência entre as amostras estudadas, permitirá nortear estratégias de novas coletas visando incrementar a variabilidade genética das coleções de germoplasma à disposição de melhoristas. Os resultados realçam a importância de novas coletas visando aumentar a representatividade das subpopulações com maior riqueza alélica e, ao mesmo tempo agregar o máximo de novas bacias visando recolher o máximo alelos raros e/ou exclusivos, antes que a ação antrópica destrutiva os alcancem. Por outro lado, o padrão de diversidade associado às bacias e a grande ocorrência de alelos comuns (frequência $> 0,05$) abre perspectivas para a constituição de coleções nucleares maximizando a utilização dos recursos genéticos e minimizando os custos de manutenção dos bancos de germoplasma.

5. BIBLIOGRAFIA

- Almeida, CMVC. 1982. Relatório da expedição botânica para a coleta de cacau nativo na região de Ariquemes, Rondônia, 1982. Belém, PA., Brasil, CEPLAC/DEPEA, 21 p.
- Almeida, CMVC. 2001. Ecologia de Populações Naturais. In: *Melhoramento Genético do Cacaueiro*, ed. LAS Dias. FUNAPE, Viçosa, MG, UFV, 2001. Capítulo 4, p 129-162.
- Almeida, CMVC e Almeida, CFG. 1987. Coleta de cacau silvestre no Estado de Rondônia, Brasil. *Revista Theobroma* (Brasil). 17:65-92.
- Almeida, CMVC e Dias, LAS. 2001. Recursos Genéticos .In: *Melhoramento Genético do Cacaueiro*.(Ed. LAS Dias), FUNAPE, Viçosa, MG, 2001.Capítulo 5, p.163-216.
- Almeida, CMVC; Barriga, JP; Machado, PFR and Bartley, BGD. 1987. Evolução do programa de conservação dos recursos genéticos de cacau na Amazônia brasileira. Boletim Técnico nº 5. Ministério da Agricultura. Comissão Executiva do Plano da Lavoura Cacaueira, DEPEA, 108p.
- Almeida, CMVC; Machado, PFR; Barriga, JP e Silca FCO. 1995. *Coleta de cacau na Amazônia Brasileira: uma abordagem histórica e analítica*. CEPLAC/SUPOC, Porto Velho. 83p.
- Allen, PJ; Amos, W; Pomeroy, PP and Twiss, SD. 1995. Microsatellite variation in gey seals (*Halichoerus grypus*) shows evidence fo genetic differentiation between two British breeding colonies. *Molecular Ecology* 4:653-62.
- Arlequin ver. 2.000. A software for population genetics data analysis.URL: <http://anthro.unige.ch/arlequin>.
- Balloux F, Brüner H, Lugon-Moulin N, Hausser J and Goudet J, 2000. Microsatellites can be misleading: an empirical and simulation study. *Evolution*. 54:1414-22
- Barriga, JP; Machado, PFR; Almeida, CMVC e Almeida, CFG. 1985. A preservação e utilização dos recursos genéticos de cacau na Amazônia brasileira. In: *International Cocoa Research Conference, 9 Proceedings*. Cocoa Producers' Alliance, London, p.73-79.
- Bartley, BGD. 1963. Exploration for *Theobroma* in the Amazon valley. *Genética Agrária* 17:345-9.
- Bartley, BGD. 1969. Twenty years of cacao breeding at the imperial College of Tropical Agriculture, Trinidad. In: *International Cocoa Research Conference, 2 Proceedings*. CEPLAC, Salvador, Itabuna, p. 29-34.
- Bartley, BGD; Machado, PFR; Ahnert, D; Barriga, JP & Almeida, CMVC. 1988. Descrição de populações de cacau da Amazônia brasileira I. Observações preliminares sobre populações de Alenquer, Pará. In: *International Cocoa Research Conference, 10, Proceedings*. Cocoa Producers' Alliance, London, p. 665-72.

- Bowcock, AM, Ruiz-Linares, A; Tomfohrde, J; Minch, E; Kidd, JR and Cavalli-Sforza, LL. 1994. High resolution of human evolution with polymorphic microsatellites. *Nature* 368:455-457.
- Botstein, D; White, RL; Skolnick, M; Davis, RW. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* 32:314-31.
- Brown, AHD. 1989a. The case for core collections. In: *The use of plant genetic resources*. (Eds.) Brown, AHD; Frankel, O; Marshall, DR and Williams, JT. Cambridge: Cambridge University Press, 1989. p.136-56.
- Brown, AHD. 1989b. Core collections: a practical approach to genetics resources managements. *Genome* 31:818-24.
- Bush, MB. 1994. Amazonian speciation: a necessarily complex model. *Journal of Biogeography* 21:-5-17.
- Charlesworth, D and Charlesworth, B. 1987. Inbreeding depression and its evolutionary consequences. *Annual Review of Ecology and Systematics* 18:237-68.
- Cheesman, EE. 1944. Notes on the nomenclature, classification and possible relationships of cacao populations. *Tropical Agriculture* 21:144-59.
- Cope, FW. 1962. The mechanism of pollen incompatibility in *Theobroma cacao*. *Heredity* 17:157-82.
- Coltman, DW; Bowen, WD and Wright, JM. 1998. Birth weight and neonatal survival of harbour seal pups are positively correlated with genetic variation measured by microsatellites. *Proc. of the Royal Soc. of London B* 265:803-9.
- Coulson, T; Pemberton, J; Albon, S; Beaumont, M; Marshall, T; Slate, J. 1998. Microsatellites reveal heterosis in red deer. *Proceedings of the Royal Society of London, B* 265:489-95.
- Cruz, CD. 2001. *Programa Genes: versão Windows; aplicativo computacional em genética e estatística*. Viçosa; UFV. 2001. 648p. il.
- Cuatrecasas, J. 1964. Cacao and its allies: a taxonomic revision of the genus *Theobroma*. *Conmt. Nat. Herb. (USA)* 35:379-614.
- Deka, R; Jin, L; Shriver, MD; Yu LM; Decroo, S; Hubdrieser, J; Bunker, CH; Ferrell, RE and Chakraborty, R. 1995. Population genetic of dinucleotide (dC-dA)_n-(dG-dT) polymorphisms in world populations. *Am J Hum Genet* 56:461-74.
- Dias, LAS & Kageyama, PY. 1997. Temporal stability of multivariate genetic divergence in caçao (*Theobroma cacao* L.) in Southern Bahia conditions. *Euphytica* 93:181-7.
- Dias, LAS; Souza, CAS; Augusto, SG; Siqueira, PR & Miller, MW. 1998. Performance and temporal stability analyses of cacao cultivars in Linhares, Brasil. *Plant Re. Develop.* 5:343-55.
- Dias, LAS. 2001a. Contribuições do melhoramento. In: *Melhoramento Genético do Cacaueiro*.(Ed. LAS Dias), FUNPAPE, Viçosa, MG, 2001. Capítulo 12, p.493-529.

- Dias, LAS. 2001b. Origem e dispersão de *Theobroma cacao* L.: novo cenário. In: *Melhoramento Genético do Cacaueiro*.(Ed. LAS Dias), FUNPAPE, Viçosa, MG, 2001. Capítulo 3, p.81-127.
- Dias, LAS & Almeida, CMVC. 2002. Cocoa germplasm: how and how much to collect. *INGENIC Newsletter* 6:8-11.
- Dias, LAS; Barriga, JP; Kageyama, PY & Almeida, CMVC. 2003a. Variation and its distribution within and between wild cacao populations from Brazilian Amazon. *Brazilian Archives of Biology and Technology* 46(4) (no prelo).
- Dias, LAS; Kageyama, PY & Castro, GCT. 1997. Divergência fenética multivariada na preservação de germoplasma de cacau (*Theobroma cacao* L.). *Agrotropica* 9:29-40.
- Dias, LAS; Marita, J; Cruz, CD; Barros, EG and Salomão, TMF. 2003b. Genetic distance and its association with heterosis in cacao. *Brazilian Archives of Biology and Technology*, 46(4) (no prelo).
- Doyle, JJ and Doyle, JL. 1990. Isoaltion of plant DNA from fresh tissue. *Focus* 12:13-5.
- Edwards, A; Hammond, HA; Jin, L; Caskey, CT and Chakraborty, R. 1992. Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics* 12:241-53.
- Eisen, JA. 1999. Mechanistic basis for microsatellite instability. In *Microsatellites Evolution and Applications* (Ed. Goldstein, D.B. and Schlötterer, C.), pp 34-48. Oxford University Press, New York.
- El Mousadik, A & Petit, J. 1996. High level of genetic differentiation for allelic richness among population of the argan tree (*Argania spinosa* L.) Skeels endemic of Morocco. *Theoretical and Applied Genetics* 92:832-839.
- Engels, JMN; Bartley, BDG; Enriquez, GA. 1980. Cacao descriptors , their states and *modus operandi*. *Turrialba* 30:209-218.
- Excoffier, L; Smouse, P & Quattro, J. 1992. Analysis of molecular variance inferred from metric distance among DNA haplotyps: Application of human mitochondrial DNA restriction data. *Genetics* 131:479-491.
- Ewens, WJ. 1972. The samplig theory of selectively neutral alleles. *Theoretical Population Biology* 3:87-112
- Falconer, DS. 1981. *Introduction to quantitative genetics*. Longman, London
- Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39:783-791.
- Felsenstein, J. 1993. PHYLIP version 3.50. Seattle, WA, USA, PO Box 35760
- Ferreira, ME & Grattapaglia, D. Introdução ao uso de marcadores moleculares em análise genética. 3^a. ed. EMBRAPA-CENARGEN, Brasília, 1998. pp 220.
- Figueira, A. Pires, JL; Cascardo, JCM; Cardoso, RD; Nascimento, CS and Lambert, SV. 2000. Cocoa butter quality can be improved by genetic breeding for fatty acid and triacylglycerol composition and hardness. In:

- Proceedings 12th International Cocoa Research Conference*, Salvador, Bahia, Brazil, Cocoa Producers' Alliance, Lagos, Nigeria, pp. 309-407.
- Figueira, A; Janick, J; Levy, M & Goldsbrough, PB. 1994. Reexamining the classification of *Theobroma cacao* L. using molecular markers. *J. Amer. Soc. Hort. Sci.* 119:1073-82.
- Forbes, SH; Hogg, JT; Buchanan, FC; Crawford, AM and Allendorf, FW. 1995. Microsatellite evolution in congeneric mammals: Domestic and Bighorn sheep. *Molecular Ecology* 12:1106-13.
- GENETIX, ver. 4.01 logiciel sous WindowsTM pour la génétique des populations. Laboratoire Génome et Populations, CNRS UPR 9060, Université de Montpellier II, Montpellier (France).
- Ghaderi, A; Adams, MW & Nassib, AM. 1984. Relationship between genetic distance and heterosis for field and morphological traits in dry edible bean and faba bean. *Crop Sci.* 24:37-42.
- Goldstein, B and Schlötterer, C. 2000. *Microsatellites: Evolution and Applications*. University Press, Oxford.
- Goldstein, DB; Ruiz Linares, A; Cavalli-Sforza, LL and Feldman, MW. 1995. An evaluation of genetic distances for use with microsatellite loci. *Genetics* 139:463-71.
- Goudet, J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from <http://www.unil.ch/izea/software/fstat.html>. Updated from Goudet (1995)
- Gotsch, N. 1997. Cocoa biotechnology: status, constraints and future prospects. *Biotechnologies Adv.* 15:333-52.
- Guinand B, 1996. Use of a multivariate model using allele frequency distributions to analyse patterns of genetic differentiation among populations. *Biol. J. Linn. Soc.* 58:173-95.
- Guo, SW and Thompson, EA. 1992. Performing the exact test of Hardy-Weinberg proportion for multiples allele. *Biometrics* 48:361-372.
- Haffer, J. 1969. Speciation in Amazonian forest birds. *Science* 165:131-7.
- Hamilton WD, 1971. Selection of selfish and altruistic behaviour in some extreme models. In: *Man and Beast: Comparative Social Behavior*. Eisenberg and Dillon (Eds.) Smithsonian Inst. Press, Wash, DC. pp. 57-91
- Hartl, DL & Clark, AG. 1989. *Principles of population genetics*. 2nd ed. Sinauer Associates, Sunderland, MA.
- Hurlbert, SH. 1971. The nonconcept of species diversity: a critique and alternative parameters. *Ecology* 52:577-86.
- Kimura, M and Crow, JF. 1964. The number of alleles that can be maintained in a finite population. *Genetics* 49:725-38.
- Kimura, M. 1968. Evolutionary rate at the molecular level. *Nature* 217:624-6.
- Kimura, M and Ohta, T. 1978. Stepwise mutation model and distribution of allelic frequencies in a finite population. *Proceedings of the National Academy of Sciences of the USA* 75:2868-72.

- King, JL & Jukes, TH. 1969. Non-Darwinian evolution. *Science* 164:788-98.
- Knight, C. 2000. Cocoa review: supply and demand trends. American Cocoa Research Institute. <http://www.acri-cocoa.org/acri/index.cfm?item=publications>.
- Kumar,S; Tamura, K; Jakobsen, IB and Nei, M.2001. MEGA2: Molecular Evolutionary Genetics Analysis software, Bioinformatics (submitted).
- Lanaud, C.1986. Utilisation des marqueurs enzymatiques pour l'étude génétique du cacaoyer: *Theobroma cacao* L. I. Contrôle génétique et "linkage" de neuf marqueurs enzymatiques. *Café, Cacao Thé* 30:259-267.
- Lanaud, C. 1987. Nouvelles données sur la biologie du cacaoyer (*T. cacao* L.): diversité des populations, système d'incompatibilité, haploïdes spontanés. Les conséquences pour l'amélioration génétique de cette espèce. Tese de PhD, Université de Paris XI.
- Lanaud, C; Risterucci, AM; Pieretti, I; Falque, M; Bouet, A; Lagoda, P.JL. 1999a. Isolation and characterization of microsatellites in *Theobroma cacao* L. *Molecular Ecology* 8:2141-2143.
- Lanaud, C; Motamayor, JC et Sounigo. 1999b. *Le cacaoyer. Diversité génétique des plants tropicales cultivés*. Eds. P Hamon; X Perrier & JC Glaszmann. CIRAD, Montpellier, France, pp: 141-169.
- Lanaud, C; Sounigo, O; Amefia, YK; Paulin, D; Lachenaud, P and Clement, D. 1997. Nouvelles données sur le fonctionnement du système d'incompatibilité du cacaoyer et ses conséquences pour la sélection. *Café, Cacao, Thé* 31:267-277.
- Lachenaud, P. 1997. Genetic taxonomic structuring of *Theobroma cacao* L. species – fresh hypothesis. *INGENIC Newsletter* 3:10-11.
- Laurent, V; N'Goran, JAK; Eskes, AB and Lanaud, C. 1995. Estimation of the level of heterozygosity and genetic diversity with RFLP markers. *INGENIC Newsletter* 1:5-7.
- Laurent, V; Risterucci, AM; Lanaud, C. 1993. Genetic diversity in cocoa revealed by cDNA probes. *Theoretical and Applied Genetics* 88:193-8.
- Laurent, V; Risterucci, AM; Lanaud, C.1994. RFLP study of genetic diversity of *Theobroma cacao* L. *Angew Botany* 68:36-39.
- Lerceteau, E; Robert, T; Pétiard, V & Cruzillat, D. 1997. Evaluation of the extent of genetic variability among *Theobroma cacao* accessions using RAPD and RFLP markers. *Theoretical and Applied Genetics* 95:10-9.
- Litt, M; Hauge, X & Sharma, V. 1993. Shadow bands seen when typing polymorphic dinucleotide repeats: some causes and cures. *Biotechniques* 15:280-4.
- Lockwood, G. 1976. A comparison of the growth and yield during a 20 year period of amelonado and upper amazon hybrid cocoa in Ghana. *Euphytica*. 25:647-58.
- Long JC, 1986. The allelic correlation of Gaij- and Kalam-speaking people. I. The estimation and interpretation of Wright's F-Statistics. *Genetics* 112:629-647

- Marita, JM; Nienhuis, J; Pires, JL and Aitken, WM. 2001. Analysis of genetic diversity in *Theobroma cacao* with emphasis on witches' broom disease resistance. *Crop Science* 41:1305-16.
- Manly, BFJ. 1985. *The statistic of natural selection on animal populations*. Chapman and Hall, London.
- Mayr, HL. 1977. *Populações, espécies e evolução*. USP, São Paulo.
- Marroig, G & Cerqueira, R. 1997. Plio-Pleistocene South América history and the Amazonas Lagoon hypothesis: a piece in the puzzle of the Amazonian diversification. *Journal of Comparative Biology* 2:103-19.
- Vilela-Morales, EA; Valois, ACC & Nass, LL. 1997. *Recursos genéticos vegetales*. Brasília: EMBRAPA-SPI/Embrapa-Cenargen, 78p.
- Motamayor, JC; Risterucci, AM; Lopez, PA; Ortiz, CF; Morena, A and Lanaud, C. 2002. Cacao domestication I: the origin of the cacao cultivated by the Mayas. *Heredity* 89:380-6.
- Nei, M. 1972. Genetic distance between populations. *Amer. Nat.* 106:283-92.
- Nei, M. 1973. Analysis of gene diversity in subdivided population. *Proc. Natl. Acad. Sci. USA* 70:3321-3.
- Nei, M. 1977. *F*-statistic and analysis of gene diversity in subdivided populations. *Ann. Hum. Genetics* 41:225-33.
- Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89:583-90.
- Nei, M. 1987. *Molecular Evolutionary Genetics*. Columbia University Press, New York.
- Nei, M & Kumar, S. 2000. *Molecular Evolution and Phylogenetics*. Oxford University Press, NY.
- N'Goran, JAK; Laurent, V; Risterucci, AM and Lanaud, C. 1994. Comparative genetic diversity studies of *Theobroma cacao* L. using RFLP and RAPD markers. *Heredity* 73:589-97.
- N'Goran, JAK; Laurent, V; Risterucci, AM and Lanaud, C. 2000. The genetic structure of cocoa populations (*Theobroma cacao* L.) revealed by RFLP analysis. *Euphytica*, 115:83-90.
- One-Dimensional Gel Analyse. ONE-Dscan, ver. 1.0, Copyright 1994/1995, Scanalytic, a dimensional CST Inc.
- Ohta, T. 1982. Linkage disequilibrium due to random genetic drift in finite subdivided populations. *Proc. Natl. Acad. Sci. USA* 79:1940-4.
- Paetkau, D and Strobeck, I. 1995. The molecular basis and evolutionary history of a microsatellite null allele in bears. *Molecular Ecology* 4:519-520.
- Paetkau, D; Calvert, W; Stirling, I and Strobeck, I. 1995. Microsatellite analysis of population structure in Canadian polar bears. *Molecular Ecology* 4:347-354.
- Patton, JL; Silva, MNF & Malcolm, JR. 1994. Gene genealogy and differentiation among arboreal spiny rats (*Rodentia: Echimyidae*) of the

- Amazon basin: a test of the riverine barrier hypothesis. *Evolution* 48:1314-23.
- Pemberton, JM; Coltman, DW; Coulson, TN and Slate J. 1999. Using microsatellites to measure the fitness consequences of inbreeding and outbreeding. In *Microsatellites Evolution and Applications* (Ed. Goldstein, D.B. and Schlötterer, C.), Oxford University Press, New York. p 151-64.
- Petit, R; El Mousadik, A & Pons, O. 1998. Identifying populations for conservation on the basis of genetic markers. *Conservation Biology* 12:844-855.
- Pires, JL; Monteiro, WR; Pinto, LRM; Figueira, A; Yamada, MM and Ahnert, D. 2000. A proposal for cocoa breeding. In: *Proceedings of 12th International Cocoa Research Conference*, Salvador, Bahia, Brazil, Cocoa Producers' Alliance, Lagos, Nigeria. Pp. 287-292.
- Popgene, Population Genetic Analysis, ver. 1.32.
ftp://ftp.microsoft.com/SoftLib/MSLFILES/HPGL.EXE.
- Posnette, AF. 1950. The pollination of cacao in the Gold Coast. *Journal of Horticultural Science* 25:155-63
- Pound, FJ. 1938. *Cacao and witches' broom disease (Marasmius perniciosus) of South America, with notes on their species of Theobroma; report on a visit to Ecuador the Amazon Valley and Colombia April 1937 – April 1938*. Yuille's Printerie, Trinidad.
- Queller, DC; Strassmann, JE. And Hughes, CR. 1993. Microsatellites and Kinships. *Tree* 8:285-288
- Queller DC and Goodnight KF, 1989. Estimating relatedness using genetic markers. *Evolution* 42:258-275.
- Reynolds, J; Weir, BS & Cockerhan, CC. 1993. Estimation of co-ancestry coefficient: basis for a short-term genetic distance. *Genetics* 105:767-9.
- Risterucci, AM; Grivet, L; N'Goran, JAK; Pieretti, I; Flament, MH and Lanaud, C. 2000. A high-density linkage map of *Theobroma cacao* L. *Theor. Appl. Gene.* 101:948-55.
- Ronning, CM & Schnell, RJ, 1994. Allozyme diversity in a germplasm collection of *Theobroma cacao* L. *The Journal of Heredity* 85:291-5.
- Russel, JR; Hosein, F; Johnson, E; Waugh, R & Powell, W. 1993. Genetic differentiation of cocoa *Theobroma cacao* L.). *Molecular Ecology* 2:89-97.
- Sambrook, J; Fritsch, EF & Maniatis, T. 1989. *Molecular cloning: laboratory manual*, 2nd edition CSHL, Cold Spring Harbor, NY.
- Sereno, ML. 2001. Estimación de la diversidad genética de poblaciones silvestres de *Theobroma cacao* L. amazónico brasileño, mediante microsatélites. Piracicaba, 2001. 70p. Tese de Mestrado – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo.
- Silva, FCO; Neto, EF; Kodama, KR and Figueira, A. 1998. Avaliação das relações genéticas entre genótipos de cacaueteiro (*Theobroma cacao* L.) contrastantes para reação à vassoura-de-bruxa, através de marcadores RAPD. *Genetics and Molecular Biology*, 21:105 (supplement).

- Simpson, BB & Haffer, J. 1978. Speciation patterns in the Amazonian forest biota. *Ann. Rev. Ecol. Systemat* 9:497-518.
- Slatkin, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, 139:457-62.
- Sounigo, O; Umaharan, R & Sankar, A. 1997. Assessing genetic diversity in the International Cocoa Genebank of Trinidad (ICGT) using isozyme electrophoresis. *INGENIC Newsletter* 3:11-6.
- Sounigo, O; Umaharan, R; Sankar, A and Ramdahin, S. 2001. Estimation of the level of heterozygosity of 570 cocoa clones using isozyme electrophoresis. *INGENIC Newsletter* 6:15-18.
- Vello, F e Silva, LF. 1968. Relatório de viagem à região amazônica. Ilhéus, BA., Brasil, CEPLAC/CEPEC. Comunicado Técnico nº 22. 19p.
- Warren, JM. 1994. Isozyme variation in a number of populations of *Theobroma cacao* L. obtained through various sampling regimes. *Euphytica* 72:121-126.
- Weber, JL. 1990. The informativeness of human (DC-DA)N(DG-DT)N polymorphisms. *Genomics* 7:524-539.
- Weir, BS and Cockerham, CC. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38:1358-70.
- Weir, BS. 1990. Genetic data analysis. Sinauer Associate Sunderland, Mass.
- Whytkus, R; de La Cruz, M; Mota-Bravo, L. e Gómez-Pompa, A 1998. Genetic diversity and relationship of cacao (*Theobroma cacao* L.) in Southern México. *Theoretical and Applied Genetics* 96:621-7.
- Wright S, 1969. Evolution and the genetics of populations. Vol. 2. The theory of gene frequencies. University of Chicago Press.
- Wright S. 1951. The genetical structure of populations. *Ann. Eugen.* 15:323-54.
- Wright, S. 1978. Variability within and among natural populations. Vol. 4. The Univ. of Chicago Press, Chicago.
- Yechezkel, K and Morris, S. 1999. Functional roles of microsatellites and minisatellites. In *Microsatellites Evolution and Applications* (Ed. Goldstein, D.B. and Schlötterer, C.), pp 10-23, Oxford University Press, New York.
- Young, AM. 1994. *The Chocolate Tree: a natural history of cacao*. Smithsonian Institution Press. Washington.

ANEXOS

Tabela 1. P-valor para o desequilíbrio genotípico baseado em 8120 permutações. P-valor para o nível nominal de 5 % = 0,000123.

Locos	P-valor	Locos	P-valor	Locos	P-valor	Locos	P-valor
mTc-01 X mTc-02	0,1919	mTc-02 X mTc-55	0,1320	mTc-06 X mTc-55	0,0865	mTc-09 X mTc-10	0,5500
mTc-01 X mTc-04	0,1964	mTc-02 X mTc-57	0,7142	mTc-06 X mTc-57	0,3547	mTc-09 X mTc-11	0,1069
mTc-01 X mTc-06	0,0467	mTc-02 X mTc-58	0,2536	mTc-06 X mTc-58	0,0048	mTc-09 X mTc-12	0,0092
mTc-01 X mTc-07	0,0372	mTc-02 X mTc-60	0,1574	mTc-06 X mTc-60	0,0053	mTc-09 X mTc-13	0,8120
mTc-01 X mTc-08	0,4740	mTc-04 X mTc-06	0,1301	mTc-07 X mTc-08	0,0032	mTc-09 X mTc-17	0,1972
mTc-01 X mTc-09	0,1852	mTc-04 X mTc-07	0,0720	mTc-07 X mTc-09	0,0281	mTc-09 X mTc-18	0,0009
mTc-01 X mTc-10	0,0479	mTc-04 X mTc-08	0,5693	mTc-07 X mTc-10	0,0430	mTc-09 X mTc-19	0,0015
mTc-01 X mTc-11	0,0094	mTc-04 X mTc-09	0,0438	mTc-07 X mTc-11	0,0325	mTc-09 X mTc-25	0,0188
mTc-01 X mTc-12	0,0751	mTc-04 X mTc-10	0,0642	mTc-07 X mTc-12	0,1089	mTc-09 X mTc-26	0,1496
mTc-01 X mTc-13	0,2516	mTc-04 X mTc-11	0,1344	mTc-07 X mTc-13	0,2353	mTc-09 X mTc-32	0,0888
mTc-01 X mTc-17	0,6005	mTc-04 X mTc-12	0,0406	mTc-07 X mTc-17	0,0511	mTc-09 X mTc-33	0,0165
mTc-01 X mTc-18	0,0174	mTc-04 X mTc-13	0,6900	mTc-07 X mTc-18	0,2232	mTc-09 X mTc-35	0,3370
mTc-01 X mTc-19	0,1531	mTc-04 X mTc-17	0,0842	mTc-07 X mTc-19	0,0861	mTc-09 X mTc-40	0,1847
mTc-01 X mTc-25	0,1919	mTc-04 X mTc-18	0,6994	mTc-07 X mTc-25	0,0025	mTc-09 X mTc-42	0,3734
mTc-01 X mTc-26	0,1203	mTc-04 X mTc-19	0,2181	mTc-07 X mTc-26	0,0191	mTc-09 X mTc-43	0,0791
mTc-01 X mTc-32	0,7336	mTc-04 X mTc-25	0,0647	mTc-07 X mTc-32	0,2867	mTc-09 X mTc-44	0,3888
mTc-01 X mTc-33	0,6817	mTc-04 X mTc-26	0,1552	mTc-07 X mTc-33	0,0218	mTc-09 X mTc-49	0,0809
mTc-01 X mTc-35	0,0557	mTc-04 X mTc-32	0,3362	mTc-07 X mTc-35	0,2128	mTc-09 X mTc-54	0,1132
mTc-01 X mTc-40	0,0748	mTc-04 X mTc-33	0,4197	mTc-07 X mTc-40	0,3945	mTc-09 X mTc-55	0,0888
mTc-01 X mTc-42	0,0004	mTc-04 X mTc-35	0,6543	mTc-07 X mTc-42	0,1799	mTc-09 X mTc-57	0,1833
mTc-01 X mTc-43	0,0730	mTc-04 X mTc-40	0,1208	mTc-07 X mTc-43	0,0133	mTc-09 X mTc-58	0,0436
mTc-01 X mTc-44	0,0414	mTc-04 X mTc-42	0,6867	mTc-07 X mTc-44	0,0152	mTc-09 X mTc-60	0,0438
mTc-01 X mTc-49	0,0218	mTc-04 X mTc-43	0,1452	mTc-07 X mTc-49	0,6356	mTc-10 X mTc-11	0,1521
mTc-01 X mTc-54	0,0001	mTc-04 X mTc-44	0,1051	mTc-07 X mTc-54	0,1659	mTc-10 X mTc-12	0,0594
mTc-01 X mTc-55	0,3299	mTc-04 X mTc-49	0,0278	mTc-07 X mTc-55	0,0043	mTc-10 X mTc-13	0,9127
mTc-01 X mTc-57	0,1647	mTc-04 X mTc-54	0,0340	mTc-07 X mTc-57	0,2341	mTc-10 X mTc-17	0,5368
mTc-01 X mTc-58	0,2523	mTc-04 X mTc-55	0,0749	mTc-07 X mTc-58	0,0626	mTc-10 X mTc-18	0,3315
mTc-01 X mTc-60	0,0220	mTc-04 X mTc-57	0,6278	mTc-07 X mTc-60	0,0805	mTc-10 X mTc-19	0,1560
mTc-02 X mTc-04	0,9483	mTc-04 X mTc-58	0,0824	mTc-08 X mTc-09	0,0248	mTc-10 X mTc-25	0,0163
mTc-02 X mTc-06	0,8436	mTc-04 X mTc-60	0,0590	mTc-08 X mTc-10	0,1058	mTc-10 X mTc-26	0,5932
mTc-02 X mTc-07	0,0301	mTc-06 X mTc-07	0,1659	mTc-08 X mTc-11	0,2999	mTc-10 X mTc-32	0,0175
mTc-02 X mTc-08	0,8011	mTc-06 X mTc-08	0,1275	mTc-08 X mTc-12	0,0744	mTc-10 X mTc-33	0,2966
mTc-02 X mTc-09	0,1339	mTc-06 X mTc-09	0,0849	mTc-08 X mTc-13	0,2982	mTc-10 X mTc-35	0,0732
mTc-02 X mTc-10	0,8350	mTc-06 X mTc-10	0,0117	mTc-08 X mTc-17	0,6786	mTc-10 X mTc-40	0,0254
mTc-02 X mTc-11	0,1546	mTc-06 X mTc-11	0,0160	mTc-08 X mTc-18	0,2440	mTc-10 X mTc-42	0,8936
mTc-02 X mTc-12	0,9909	mTc-06 X mTc-12	0,0052	mTc-08 X mTc-19	0,0712	mTc-10 X mTc-43	0,3664
mTc-02 X mTc-13	0,1203	mTc-06 X mTc-13	0,0190	mTc-08 X mTc-25	0,3716	mTc-10 X mTc-44	0,0147
mTc-02 X mTc-17	0,3809	mTc-06 X mTc-17	0,6132	mTc-08 X mTc-26	0,0087	mTc-10 X mTc-49	0,0911
mTc-02 X mTc-18	0,4990	mTc-06 X mTc-18	0,2814	mTc-08 X mTc-32	0,5083	mTc-10 X mTc-54	0,6374
mTc-02 X mTc-19	0,1363	mTc-06 X mTc-19	0,1886	mTc-08 X mTc-33	0,5865	mTc-10 X mTc-55	0,1739
mTc-02 X mTc-25	0,8370	mTc-06 X mTc-25	0,0006	mTc-08 X mTc-35	0,1414	mTc-10 X mTc-57	0,1457
mTc-02 X mTc-26	0,0355	mTc-06 X mTc-26	0,5257	mTc-08 X mTc-40	0,4617	mTc-10 X mTc-58	0,0287
mTc-02 X mTc-32	0,7573	mTc-06 X mTc-32	0,2583	mTc-08 X mTc-42	0,0665	mTc-10 X mTc-60	0,0044
mTc-02 X mTc-33	0,0791	mTc-06 X mTc-33	0,0225	mTc-08 X mTc-43	0,0925	mTc-11 X mTc-12	0,2680
mTc-02 X mTc-35	0,7840	mTc-06 X mTc-35	0,6762	mTc-08 X mTc-44	0,4112	mTc-11 X mTc-13	0,0004
mTc-02 X mTc-40	0,8052	mTc-06 X mTc-40	0,1921	mTc-08 X mTc-49	0,3395	mTc-11 X mTc-17	0,0038
mTc-02 X mTc-42	0,0140	mTc-06 X mTc-42	0,0767	mTc-08 X mTc-54	0,0027	mTc-11 X mTc-18	0,1147
mTc-02 X mTc-43	0,1193	mTc-06 X mTc-43	0,0371	mTc-08 X mTc-55	0,3652	mTc-11 X mTc-19	0,0643
mTc-02 X mTc-44	0,0506	mTc-06 X mTc-44	0,0814	mTc-08 X mTc-57	0,2738	mTc-11 X mTc-25	0,0010
mTc-02 X mTc-49	0,8044	mTc-06 X mTc-49	0,7016	mTc-08 X mTc-58	0,7595	mTc-11 X mTc-26	0,0347
mTc-02 X mTc-54	0,5530	mTc-06 X mTc-54	0,1964	mTc-08 X mTc-60	0,4801	mTc-11 X mTc-32	0,0111

Tabela 1. P-valor para o desequilíbrio genotípico baseado em 8120 permutações. P-valor para o nível nominal de 5 % = 0,000123 (continuação).

Locos	P-valor	Locos	P-valor	Locos	P-valor	Locos	P-valor
mTc-11 X mTc-33	0,1254	mTc-17 X mTc-25	0,5493	mTc-25 X mTc-42	0,4463	mTc-35 X mTc-55	0,0804
mTc-11 X mTc-35	0,3881	mTc-17 X mTc-26	0,2394	mTc-25 X mTc-43	0,1282	mTc-35 X mTc-57	0,7132
mTc-11 X mTc-40	0,0014	mTc-17 X mTc-32	0,4773	mTc-25 X mTc-44	0,2035	mTc-35 X mTc-58	0,8062
mTc-11 X mTc-42	0,3951	mTc-17 X mTc-33	0,3759	mTc-25 X mTc-49	0,3782	mTc-35 X mTc-60	0,1037
mTc-11 X mTc-43	0,0003	mTc-17 X mTc-35	0,5230	mTc-25 X mTc-54	0,2905	mTc-40 X mTc-42	0,4778
mTc-11 X mTc-44	0,1233	mTc-17 X mTc-40	0,4458	mTc-25 X mTc-55	0,0129	mTc-40 X mTc-43	0,3886
mTc-11 X mTc-49	0,5462	mTc-17 X mTc-42	0,6796	mTc-25 X mTc-57	0,0415	mTc-40 X mTc-44	0,1802
mTc-11 X mTc-54	0,4429	mTc-17 X mTc-43	0,5850	mTc-25 X mTc-58	0,2421	mTc-40 X mTc-49	0,1601
mTc-11 X mTc-55	0,0624	mTc-17 X mTc-44	0,0911	mTc-25 X mTc-60	0,1044	mTc-40 X mTc-54	0,7472
mTc-11 X mTc-57	0,0404	mTc-17 X mTc-49	0,0071	mTc-26 X mTc-32	0,4293	mTc-40 X mTc-55	0,4690
mTc-11 X mTc-58	0,0722	mTc-17 X mTc-54	0,1516	mTc-26 X mTc-33	0,0020	mTc-40 X mTc-57	0,0436
mTc-11 X mTc-60	0,1219	mTc-17 X mTc-55	0,2009	mTc-26 X mTc-35	0,1207	mTc-40 X mTc-58	0,1154
mTc-12 X mTc-13	0,0954	mTc-17 X mTc-57	0,1357	mTc-26 X mTc-40	0,2840	mTc-40 X mTc-60	0,0041
mTc-12 X mTc-17	0,1220	mTc-17 X mTc-58	0,4601	mTc-26 X mTc-42	0,3633	mTc-42 X mTc-43	0,0120
mTc-12 X mTc-18	0,0698	mTc-17 X mTc-60	0,0555	mTc-26 X mTc-43	0,0286	mTc-42 X mTc-44	0,0120
mTc-12 X mTc-19	0,5650	mTc-18 X mTc-19	0,0617	mTc-26 X mTc-44	0,1096	mTc-42 X mTc-49	0,0610
mTc-12 X mTc-25	0,2395	mTc-18 X mTc-25	0,0201	mTc-26 X mTc-49	0,0390	mTc-42 X mTc-54	0,0122
mTc-12 X mTc-26	0,3197	mTc-18 X mTc-26	0,0558	mTc-26 X mTc-54	0,1312	mTc-42 X mTc-55	0,0134
mTc-12 X mTc-32	0,0197	mTc-18 X mTc-32	0,3463	mTc-26 X mTc-55	0,4007	mTc-42 X mTc-57	0,1523
mTc-12 X mTc-33	0,2532	mTc-18 X mTc-33	0,3408	mTc-26 X mTc-57	0,1232	mTc-42 X mTc-58	0,3445
mTc-12 X mTc-35	0,2879	mTc-18 X mTc-35	0,2209	mTc-26 X mTc-58	0,2659	mTc-42 X mTc-60	0,1344
mTc-12 X mTc-40	0,0319	mTc-18 X mTc-40	0,1136	mTc-26 X mTc-60	0,3611	mTc-43 X mTc-44	0,0055
mTc-12 X mTc-42	0,4315	mTc-18 X mTc-42	0,0367	mTc-32 X mTc-33	0,7900	mTc-43 X mTc-49	0,2250
mTc-12 X mTc-43	0,0092	mTc-18 X mTc-43	0,0014	mTc-32 X mTc-35	0,0004	mTc-43 X mTc-54	0,5712
mTc-12 X mTc-44	0,0283	mTc-18 X mTc-44	0,1934	mTc-32 X mTc-40	0,6216	mTc-43 X mTc-55	0,2847
mTc-12 X mTc-49	0,1096	mTc-18 X mTc-49	0,3038	mTc-32 X mTc-42	0,0087	mTc-43 X mTc-57	0,0503
mTc-12 X mTc-54	0,0438	mTc-18 X mTc-54	0,7158	mTc-32 X mTc-43	0,0437	mTc-43 X mTc-58	0,0169
mTc-12 X mTc-55	0,0132	mTc-18 X mTc-55	0,7706	mTc-32 X mTc-44	0,0134	mTc-43 X mTc-60	0,0967
mTc-12 X mTc-57	0,5229	mTc-18 X mTc-57	0,0400	mTc-32 X mTc-49	0,2637	mTc-44 X mTc-49	0,0866
mTc-12 X mTc-58	0,0897	mTc-18 X mTc-58	0,2602	mTc-32 X mTc-54	0,7920	mTc-44 X mTc-54	0,0620
mTc-12 X mTc-60	0,0177	mTc-18 X mTc-60	0,0078	mTc-32 X mTc-55	0,1342	mTc-44 X mTc-55	0,1156
mTc-13 X mTc-17	0,2866	mTc-19 X mTc-25	0,0209	mTc-32 X mTc-57	0,3182	mTc-44 X mTc-57	0,5260
mTc-13 X mTc-18	0,0520	mTc-19 X mTc-26	0,5202	mTc-32 X mTc-58	0,0406	mTc-44 X mTc-58	0,1025
mTc-13 X mTc-19	0,1326	mTc-19 X mTc-32	0,3230	mTc-32 X mTc-60	0,0213	mTc-44 X mTc-60	0,0039
mTc-13 X mTc-25	0,1100	mTc-19 X mTc-33	0,2565	mTc-33 X mTc-35	0,5670	mTc-49 X mTc-54	0,1137
mTc-13 X mTc-26	0,0003	mTc-19 X mTc-35	0,1591	mTc-33 X mTc-40	0,3220	mTc-49 X mTc-55	0,3314
mTc-13 X mTc-32	0,6410	mTc-19 X mTc-40	0,2019	mTc-33 X mTc-42	0,4968	mTc-49 X mTc-57	0,1062
mTc-13 X mTc-33	0,4298	mTc-19 X mTc-42	0,1224	mTc-33 X mTc-43	0,0318	mTc-49 X mTc-58	0,3337
mTc-13 X mTc-35	0,3282	mTc-19 X mTc-43	0,1744	mTc-33 X mTc-44	0,5153	mTc-49 X mTc-60	0,8365
mTc-13 X mTc-40	0,8402	mTc-19 X mTc-44	0,0744	mTc-33 X mTc-49	0,1799	mTc-54 X mTc-55	0,9341
mTc-13 X mTc-42	0,7281	mTc-19 X mTc-49	0,6825	mTc-33 X mTc-54	0,0096	mTc-54 X mTc-57	0,3437
mTc-13 X mTc-43	0,0169	mTc-19 X mTc-54	0,1823	mTc-33 X mTc-55	0,5581	mTc-54 X mTc-58	0,3890
mTc-13 X mTc-44	0,0623	mTc-19 X mTc-55	0,0053	mTc-33 X mTc-57	0,5621	mTc-54 X mTc-60	0,1159
mTc-13 X mTc-49	0,6746	mTc-19 X mTc-57	0,5270	mTc-33 X mTc-58	0,0794	mTc-55 X mTc-57	0,3079
mTc-13 X mTc-54	0,9751	mTc-19 X mTc-58	0,1729	mTc-33 X mTc-60	0,3658	mTc-55 X mTc-58	0,4411
mTc-13 X mTc-55	0,1898	mTc-19 X mTc-60	0,0123	mTc-35 X mTc-40	0,4727	mTc-55 X mTc-60	0,0535
mTc-13 X mTc-57	0,4237	mTc-25 X mTc-26	0,0091	mTc-35 X mTc-42	0,2504	mTc-57 X mTc-58	0,5750
mTc-13 X mTc-58	0,5647	mTc-25 X mTc-32	0,2636	mTc-35 X mTc-43	0,0571	mTc-57 X mTc-60	0,5786
mTc-13 X mTc-60	0,1879	mTc-25 X mTc-33	0,4509	mTc-35 X mTc-44	0,0022	mTc-58 X mTc-60	0,0541
mTc-17 X mTc-18	0,0059	mTc-25 X mTc-35	0,6990	mTc-35 X mTc-49	0,4147		
mTc-17 X mTc-19	0,0977	mTc-25 X mTc-40	0,3836	mTc-35 X mTc-54	0,6402		

Tabela 2: Perfil de microssatélites, em pares, de base de 29 loci em 172 clones de cacau.

Clones	IntTCR 01	IntTCR 02	IntTCR 04	IntTCR 06	IntTCR 07	IntTCR 08	IntTCR 09	IntTCR 10	IntTCR 11	IntTCR 12	IntTCR 13	IntTCR 17	IntTCR 18	IntTCR 19	IntTCR 25	IntTCR 26	IntTCR 32	IntTCR 33	IntTCR 35	IntTCR 40	IntTCR 42	IntTCR 43	IntTCR 44	IntTCR 49	IntTCR 54	IntTCR 55	IntTCR 57	IntTCR 68	IntTCR 80
	CAB 006	142	238	250	198	159	285	273	206	311	200	254	265	333	379	192	297	200	327	234	277	231	201	170	205	156	227	254	238
CAB 020	142	238	250	198	159	285	278	203	311	214	254	265	337	379	175	297	200	367	234	277	235	203	170	205	156	227	254	257	210
CAB 036	142	252	270	198	162	285	278	203	311	214	254	265	345	400	192	299	205	367	234	277	235	203	170	205	156	233	254	257	210
CAB 016	142	252	262	198	159	285	278	203	311	214	254	265	337	387	192	299	200	327	234	277	231	201	170	205	156	233	252	257	210
CAB 20b	142	252	270	198	159	285	278	203	311	214	254	265	345	400	192	297	205	367	234	277	235	201	170	205	156	233	254	257	210
CAB 039	142	252	270	198	159	285	278	203	311	252	246	265	337	379	192	297	200	327	234	277	212	201	170	205	156	233	252	257	210
CAB 104	132	238	250	198	164	289	284	202	290	203	246	266	337	379	192	299	200	306	230	274	202	199	181	215	156	233	254	264	186
CAB 105	142	238	250	198	164	285	284	202	290	203	246	266	337	379	179	299	200	287	220	266	202	199	181	205	150	233	256	264	186
CAB 107	132	238	250	198	164	285	278	203	290	203	246	266	337	384	179	297	200	306	234	298	202	209	181	205	156	227	256	264	186
CAB 108	132	238	250	198	162	285	284	203	290	200	246	266	337	384	179	292	200	302	230	274	202	209	170	205	156	233	256	264	186
CAB 110	132	..	250	198	162	285	282	203	296	200	246	268	345	379	175	283	200	287	220	266	202	199	181	205	156	227	254	264	186
CAB 111	142	238	250	198	164	285	282	202	290	200	254	266	345	379	179	297	200	302	230	266	202	199	181	205	156	233	254	264	186
CAB 113	132	238	250	198	164	285	278	202	290	203	254	266	345	384	179	297	200	302	230	283	202	199	181	215	156	233	254	264	186
CAB 114	132	238	250	198	164	285	278	203	290	203	254	266	345	384	192	292	200	302	230	298	202	209	181	205	156	233	254	264	186
CAB 117	132	238	250	198	162	285	278	206	290	200	254	265	337	384	192	297	200	306	230	274	202	199	170	205	156	233	256	264	186
CAB 135	132	252	270	198	162	285	278	209	290	200	254	266	345	387	192	297	200	306	230	274	202	209	182	205	156	233	256	264	196
CAB 136	132	238	250	198	162	285	284	199	290	211	254	266	337	384	180	292	200	287	230	283	202	211	181	205	156	233	252	257	186
CAB 146	132	..	250	198	162	285	284	202	287	200	254	266	345	384	192	292	200	..	230	283	202	211	181	205	156	227	256	267	186
CAB 147	132	238	250	198	164	285	282	203	290	203	254	265	345	384	192	297	200	302	220	283	202	199	181	205	150	233	254	264	196
CAB 148	132	252	270	198	162	285	284	206	290	200	254	266	345	384	198	292	200	302	220	298	202	211	182	215	156	233	254	264	196
CAB 151	132	238	250	198	164	285	278	203	290	203	254	266	345	384	184	292	200	295	220	266	202	199	181	205	156	233	254	264	186
CAB 153	132	238	250	198	162	285	278	203	290	209	254	266	345	400	192	292	200	295	230	266	202	211	181	215	156	233	254	264	186
CAB 155	142	238	250	198	162	285	278	203	290	200	254	266	337	384	198	292	200	306	230	298	202	209	181	215	156	233	256	264	186
CAB 160	142	252	270	198	162	285	278	206	290	203	254	266	345	384	198	292	200	306	230	274	202	209	181	215	150	233	256	264	206
CAB 167	132	238	250	198	162	289	282	203	290	209	254	266	345	400	179	297	200	295	230	274	202	191	181	205	150	227	256	264	196
CAB170	132	238	250	198	164	285	282	203	290	203	254	265	345	384	192	297	200	302	230	266	202	199	170	205	150	233	256	264	186
CAB 171	132	252	270	198	164	285	284	203	311	209	254	265	345	400	198	297	200	325	230	266	202	209	182	215	156	233	256	264	208
CAB 176	132	238	250	198	164	285	282	203	290	203	254	266	345	400	192	297	200	306	230	298	..	209	181	205	150	233	256	264	186
CJ 10	142	252	262	198	159	282	278	206	296	209	246	266	345	407	198	297	200	306	230	266	202	211	181	205	150	227	256	264	196
CJ 11	145	238	262	198	160	285	278	199	296	214	246	266	345	379	175	297	200	312	220	277	223	201	170	205	156	227	254	257	212
CJ 02	146	238	270	198	160	285	278	203	296	214	246	266	345	400	192	297	205	312	234	277	235	201	193	205	156	227	254	257	212
CJ 03	145	238	262	196	160	285	282	199	296	209	246	266	345	379	192	297	200	312	234	277	223	201	183	205	156	233	242	309	194
CJ 04	142	252	262	196	159	285	282	199	304	214	254	266	333	387	179	292	200	327	220	277	212	203	170	205	144	233	252	309	214
CJ 05	132	238	262	196	159	285	282	206	296	200	246	266	345	407	192	292	205	327	234	277	231	209	182	205	156	233	252	257	191
CJ 06	146	238	270	196	160	285	282	203	290	211	246	266	345	400	192	297	200	367	234	298	235	197	193	205	156	..	242	257	194
CJ 07	142	..	262	196	160	289	282	203	296	200	246	266	345	379	175	297	200	306	..	298	223	201	170	215	150	233	242	257	194
CJ 08	142	..	270	198	160	289	282	206	296	211	254	266	345	400	192	299	200	306	..	298	223	201	193	215	150	233	254	257	214
CJ 09	142	..	262	196	159	285	278	206	296	209	246	266	345	379	175	297	200	367	234	281	223	197	170	205	156	227	242	257	191
CAB 119	145	238	262	198	159	285	278	206	296	211	254	266	333	400	175	297	200	367	234	298	223	203	170	205	156	227	252	309	194
CAB 120	132	238	270	205	162	285	284	203	290	200	254	266	345	387	179	304	200	318	220	281	228	197	170	205	144	233	254	257	186
CAB 121	142	238	250	198	162	285	282	203	311	200	254	268	345	384	175	283	..	295	230	259	202	197	172	205	156	233	252	242	191
CAB 063	132	238	270	198	162	285	284	202	290	200	254	268	345	384	179	297	200	287	220	266	202	197	181	252	264	186
CAB 064	146	238	270	198	167	285	284	206	290	200	254	268	345	384	179	297	..	302	220	266	228	209	181	254	264	186
CAB 066	142	238	250	205	162	285	278	206	304	211	254	266	333	387	179	292	200	306	220	281	202	201	170	215	150	233	252	252	191
	142	252	270	205	162	285	278	206	304	220	254	266	345	387	184	297	200	306	230	281	202	203	182	215	150	233	254	257	214

Tabela 2: Perfil de microssatélites, em pares, de base de 29 loci em 172 clones de cacau (continuação).

Clones	Loci																												
	mTcCR 01	mTcCR 02	mTcCR 04	mTcCR 06	mTcCR 07	mTcCR 08	mTcCR 09	mTcCR 10	mTcCR 11	mTcCR 12	mTcCR 13	mTcCR 17	mTcCR 18	mTcCR 19	mTcCR 25	mTcCR 26	mTcCR 32	mTcCR 33	mTcCR 35	mTcCR 40	mTcCR 42	mTcCR 43	mTcCR 44	mTcCR 49	mTcCR 54	mTcCR 55	mTcCR 57	mTcCR 58	mTcCR 60
CAB 067	148	238	250	202	164	285	284	203	290	200	254	266	345	384	192	283	200	287	230	266	202	197	181	205	150	233	252	257	186
CAB 068	148	252	270	202	165	285	284	203	290	200	254	266	345	387	198	299	200	302	230	266	202	197	181	205	156	233	252	257	186
CAB 070	132	238	270	198	164	282	284	202	290	200	254	268	345	387	192	299	200	318	230	283	202	211	181	205	156	233	252	257	186
CAB 071	132	238	270	202	162	285	284	203	296	200	254	266	345	387	179	299	200	306	230	266	202	209	181	205	150	227	254	257	186
CAB 072	132	238	250	202	162	285	284	202	285	200	254	266	345	384	179	297	200	306	220	266	202	197	181	205	150	233	254	257	186
CAB 076	132	238	250	202	162	285	284	209	290	200	254	268	345	384	179	299	200	302	220	266	202	197	170	205	156	227	252	242	186
CAB 077	132	238	250	202	162	285	284	209	296	209	254	268	345	387	198	299	205	302	230	298	202	197	182	205	156	233	252	252	186
CAB 079	132	238	270	198	162	285	284	203	290	200	254	268	345	387	179	299	200	302	220	266	202	197	181	205	156	233	252	257	186
CAB 080	148	252	270	202	164	285	284	209	290	200	254	266	345	387	192	299	205	318	230	266	202	197	181	205	156	233	254	257	186
CAB 081	132	238	270	198	164	285	284	203	290	200	254	266	337	387	192	299	200	302	220	266	223	197	181	205	150	233	254	257	186
CAB 082	148	238	270	198	164	282	284	203	296	203	246	266	345	387	192	283	200	295	266	202	211	181	205	156	233	254	257	186	
CAB 088	132	238	270	198	164	282	284	202	287	200	254	266	345	387	179	299	200	302	230	283	202	209	170	205	156	227	252	264	186
CAB 221	142	238	250	198	160	285	278	203	290	203	246	266	345	384	198	283	200	287	234	297	223	197	170	206	156	227	254	309	191
CAB 232	142	238	250	198	160	302	278	206	290	203	246	268	345	387	175	299	200	287	234	228	197	170	206	156	227	254	309	191	
CAB 233	142	238	250	198	160	302	278	206	296	209	246	268	345	387	192	299	200	287	234	283	228	197	170	206	156	227	254	309	191
CAB 236	142	238	250	198	164	285	282	209	290	203	246	268	345	387	192	297	200	302	220	274	202	197	170	205	150	227	252	242	186
CAB 247	142	238	250	198	160	302	278	206	290	203	246	268	345	387	192	283	200	287	234	298	228	197	170	206	156	227	252	309	191
CAB 410	142	238	270	198	160	302	278	206	296	203	246	268	345	387	175	283	200	287	234	298	228	197	170	206	150	227	252	309	191
CAB 253	142	238	270	198	160	302	278	206	290	203	246	266	345	387	175	283	200	302	234	297	216	197	170	205	150	233	254	309	191
CAB 255	142	238	270	198	160	282	278	206	296	203	246	268	345	384	175	283	200	287	234	283	202	170	206	156	227	254	309	191	
CAB 258	142	238	270	198	160	282	278	206	296	203	246	268	345	387	175	283	200	287	234	283	202	170	206	156	227	254	309	191	
CAB 260	142	238	250	198	160	302	278	206	296	203	246	268	345	387	175	283	200	287	234	283	228	197	170	206	156	227	254	309	191
CAB 374	142	238	250	198	160	302	278	206	290	203	246	268	345	384	175	283	200	287	234	297	223	197	170	206	156	227	254	309	191
CAB 388	142	238	250	198	160	285	278	203	290	203	246	268	345	387	192	283	200	287	234	297	223	197	170	206	156	227	254	309	191
CAB 396	142	238	250	198	160	302	278	206	290	203	246	266	345	384	175	283	200	287	234	298	223	197	170	206	156	227	254	309	191
CAB 398	142	238	250	198	160	302	278	206	290	203	246	268	345	387	175	283	200	287	234	298	223	197	170	205	150	227	254	309	191
CAB 406	142	238	270	198	160	302	278	206	296	203	246	268	345	384	175	283	200	287	234	223	199	170	205	156	227	254	309	191	
CAB 478	142	252	270	198	162	285	278	206	304	211	254	268	345	379	198	297	205	327	234	277	231	201	193	205	166	233	254	257	194
CAB 484	142	252	270	210	164	285	273	202	296	211	254	268	345	384	196	297	205	327	234	266	216	201	183	205	166	233	254	238	214
CAB 486	142	238	250	205	159	285	273	203	296	211	254	268	345	384	206	297	205	327	230	266	216	201	172	205	156	227	254	238	191
CAB 487	142	252	270	210	159	282	273	206	296	214	254	268	345	384	206	297	205	327	230	266	216	201	183	205	156	233	254	238	191
CAB 493	142	252	270	210	159	282	282	203	287	203	254	266	345	384	206	297	200	327	230	274	231	201	183	205	166	233	254	238	214
CAB 497	142	238	262	198	162	282	273	206	290	200	246	268	337	379	175	292	200	327	230	266	231	211	181	205	156	233	254	238	212
CAB 499	132	252	270	198	162	282	278	206	290	200	246	268	345	387	175	297	200	327	234	266	231	211	183	205	156	233	254	238	212
CAB 481	142	252	270	210	164	282	282	206	296	211	254	268	345	384	198	297	205	327	234	266	231	201	183	205	166	233	254	264	212
CAB 509	142	252	270	205	164	282	282	203	311	211	254	268	333	384	202	292	205	327	234	277	235	203	181	205	156	227	242	238	212
CAB 511	142	252	270	205	169	282	287	206	296	214	254	266	345	387	172	292	205	302	230	266	231	203	183	205	162	233	254	264	212
CAB 512	142	238	262	205	159	282	273	206	296	211	254	268	333	387	172	292	205	293	230	266	231	203	183	205	162	233	254	264	212
CAB 513	142	252	270	210	169	282	287	206	296	214	254	268	345	387	172	292	205	302	230	266	231	203	183	205	162	233	254	264	212
CAB 492	132	252	270	205	162	282	278	206	304	252	254	268	333	387	198	297	205	327	234	274	216	201	170	205	156	227	254	257	210
CAB 500	142	238	262	205	162	282	273	206	304	211	254	268	345	387	206	297	200	327	230	266	231	201	183	205	166	233	254	238	214
CAB 008	142	238	250	198	160	285	282	202	296	200	246	266	345	387	192	283	200	302	220	298	223	197	170	205	156	227	254	257	186
CAB 009	142	238	250	198	160	285	282	202	296	200	246	266	345	387	192	283	200	302	230	298	223	197	170	204	156	227	254	257	186
CAB 011	142	238	270	198	162	285	282	203	296	200	246	266	345	387	192	283	200	287	234	277	223	197	170	205	156	227	254	257	186

Tabela 2: Perfil de microssatélites, em pares, de base de 29 loci em 172 clones de cacau (continuação).

Clones	mtcCR 01	mtcCR 02	mtcCR 04	mtcCR 06	mtcCR 07	mtcCR 08	mtcCR 09	mtcCR 10	mtcCR 11	mtcCR 12	mtcCR 13	mtcCR 17	mtcCR 18	mtcCR 19	mtcCR 25	mtcCR 26	mtcCR 32	mtcCR 33	mtcCR 35	mtcCR 40	mtcCR 42	mtcCR 43	mtcCR 44	mtcCR 49	mtcCR 54	mtcCR 55	mtcCR 57	mtcCR 58	mtcCR 60																																																											
CAB 012	142 238 250 198 160 285 282 202 296 200 246 266 345 387 192 283 200 287 220 277 223 197 170 205 156 227 254 257 186	142 238 270 198 160 285 282 202 296 200 246 266 345 387 192 283 200 302 230 277 223 197 170 205 156 227 254 257 186	142 238 250 198 160 282 282 202 296 200 246 266 345 387 192 283 200 287 234 277 228 197 170 205 156 227 254 257 186	142 238 270 198 160 282 282 202 296 200 246 266 345 387 192 283 200 302 234 277 228 197 170 205 156 227 254 257 186	142 238 250 198 160 302 282 203 296 200 246 265 345 387 192 283 200 287 220 277 212 197 170 205 156 227 254 264 186	142 238 270 205 160 302 282 203 304 200 256 265 345 400 192 283 200 287 234 277 212 199 170 205 156 227 254 264 186	142 238 250 205 160 277 282 206 290 200 246 266 345 387 175 283 200 287 220 277 223 197 170 205 ... 227 254 257 186	142 238 270 205 160 289 282 206 304 209 256 266 345 400 175 299 200 287 234 277 223 199 170 205 ... 227 254 257 206	142 238 250 198 160 302 282 203 290 200 254 265 345 400 192 283 200 287 230 277 223 197 170 205 156 227 254 257 186	142 238 270 198 160 302 282 203 304 200 254 265 345 400 192 299 200 295 234 277 223 197 170 205 156 227 254 257 206	142 238 250 198 160 302 282 206 290 200 246 266 345 400 175 283 200 287 234 277 223 197 170 205 150 227 252 257 186	142 238 270 198 160 302 282 206 290 200 246 266 345 400 175 283 200 295 234 277 223 197 170 205 150 227 252 257 186	142 238 250 198 159 289 282 209 296 211 254 266 345 400 192 288 205 367 220 277 223 201 170 204 156 227 254 257 191	142 252 270 205 162 289 282 209 296 252 254 266 345 400 192 288 205 367 220 277 235 201 170 204 156 233 252 257 191	142 238 250 196 159 285 282 209 290 200 246 266 345 400 192 290 ... 295 230 277 223 197 170 205 156 233 252 257 194	142 252 270 198 162 289 282 209 290 200 254 266 345 400 192 290 ... 306 234 277 231 209 170 205 156 233 252 257 206	142 252 250 198 160 289 282 206 296 211 254 266 345 400 175 290 205 295 220 277 223 201 170 205 156 227 254 257 191	142 252 270 205 160 289 282 206 311 252 254 266 345 400 192 290 205 367 234 277 235 201 170 205 156 227 254 257 191	142 238 250 196 160 282 282 206 311 ... 246 266 345 400 192 290 200 306 220 277 223 201 170 205 156 227 254 ... 191	142 238 270 196 160 282 282 206 311 ... 256 266 345 400 192 290 205 367 234 277 231 203 170 205 156 233 254 ... 206	142 238 262 198 160 282 282 206 304 252 246 265 345 400 206 283 200 287 234 277 223 197 170 205 156 227 254 257 194	142 252 270 198 164 282 282 206 311 252 256 265 345 400 206 290 200 295 234 277 231 197 170 205 156 233 254 257 194	142 252 250 196 159 285 282 209 296 211 254 265 345 400 192 ... 200 295 234 277 235 201 170 206 156 227 252 257 191	142 252 270 198 159 289 282 209 311 211 254 265 345 400 192 ... 200 295 234 277 235 201 183 206 156 233 252 257 206	132 238 250 210 164 285 284 206 290 203 254 266 345 387 179 288 200 306 220 277 202 213 181 205 150 233 254 242 186	132 252 270 198 164 289 284 206 290 203 254 266 345 400 179 288 200 306 230 277 202 213 181 205 156 233 254 242 186	134 238 270 198 164 285 282 206 290 203 254 266 345 387 179 290 200 287 220 266 228 211 181 205 156 233 252 257 186	132 252 270 202 164 289 284 206 290 209 254 266 345 400 192 290 200 302 220 266 228 211 181 205 156 233 252 257 196	132 238 270 198 164 285 284 206 290 203 256 266 345 400 179 290 200 287 220 277 202 209 170 205 156 233 256 238 186	132 252 270 198 164 289 284 206 304 203 256 266 345 400 192 290 200 306 220 277 202 209 183 205 156 233 256 238 186	134 238 270 205 164 285 284 209 290 203 256 266 345 400 179 290 200 295 230 277 202 209 170 205 156 233 256 242 186	132 252 270 205 164 289 284 209 304 214 254 266 345 400 179 299 200 295 230 277 202 209 183 205 156 233 256 242 186	132 238 270 198 164 285 282 206 290 203 254 266 345 379 179 290 200 293 220 277 202 211 181 206 156 233 256 257 186	132 252 270 198 164 289 284 209 290 209 254 266 345 400 192 290 205 302 230 277 202 211 183 215 156 233 256 257 186	132 238 250 198 160 285 256 206 290 186 256 268 345 387 198 292 200 287 230 274 202 201 172 209 156 227 247 257 186	132 252 270 214 164 285 284 206 296 203 256 277 345 407 198 292 200 318 230 274 202 203 183 215 156 233 256 257 186	132 238 270 198 162 289 284 203 290 203 254 266 345 400 179 297 200 295 220 277 202 199 170 205 150 233 254 257 186	132 252 270 205 162 289 284 203 304 209 254 266 345 400 179 297 200 312 220 277 228 211 182 205 150 233 254 264 186	132 238 250 198 162 289 284 202 290 209 254 266 345 400 192 297 200 295 220 266 202 211 170 205 144 233 254 242 196	132 252 270 198 162 289 284 202 290 209 254 266 345 400 192 297 205 295 220 277 202 211 182 205 143 233 254 242 196	132 238 250 198 162 285 284 203 290 203 254 266 345 400 192 297 200 302 220 277 202 197 ... 205 156 233 254 257 186	132 252 270 198 162 289 284 203 290 214 254 266 345 400 192 299 200 312 230 277 228 197 ... 205 156 233 254 257 186	132 238 270 198 162 285 284 209 290 203 254 266 345 400 192 299 200 293 220 274 202 211 181 206 156 233 254 264 186	132 252 270 198 162 289 284 209 290 203 254 266 345 400 198 299 200 312 220 277 202 211 181 206 156 233 254 264 186	132 238 270 198 164 285 284 203 290 200 254 266 345 400 179 292 205 312 220 277 202 197 170 204 156 233 254 264 186	132 252 270 198 164 289 284 203 290 209 254 266 345 400 192 299 205 312 220 277 202 211 182 206 156 233 254 264 186	132 238 270 198 162 285 273 203 290 203 254 265 345 400 192 297 200 312 220 277 202 211 170 204 150 233 254 264 186	132 252 270 198 162 302 284 203 290 203 254 265 345 400 198 299 205 312 230 277 202 211 181 206 156 233 254 264 186	134 238 270 ... 164 285 284 202 290 203 254 266 345 379 192 297 200 295 220 277 202 211 181 206 156 227 252 264 186	132 252 270 ... 167 285 284 203 290 214 254 266 345 400 198 297 205 295 230 277 202 211 181 206 156 233 254 264 186	132 238 250 198 162 289 278 206 290 203 254 266 345 379 192 297 200 287 230 277 202 211 181 205 150 233 254 242 186	145 252 270 198 162 289 278 206 290 209 254 266 345 379 192 297 200 287 230 277 202 211 181 206 150 233 254 264 186	132 238 270 210 162 289 284 209 290 203 254 266 345 400 179 292 200 295 220 277 202 197 181 206 156 233 254 257 186	132 252 270 210 162 289 284 209 304 203 254 266 345 400 179 297 200 306 230 277 202 211 181 206 156 233 254 257 196	132 252 250 210 162 285 273 206 304 203 254 265 333 379 172 292 200 295 220 277 216 201 193 206 150 233 252 257 186	132 252 270 210 162 285 278 206 304 211 254 265 345 407 179 292 205 312 234 277 235 211 193 215 156 233 252 257 212	132 238 250 196 164 285 282 209 304 200 254 265 333 407 198 297 200 295 220 283 216 201 170 206 156 227 252 264 210	132 252 270 210 164 285 282 209 304 200 254 265 345 407 206 299 205 295 234 283 218 211 182 206 156 233 252 264 212	142 252 250 196 162 285 273 209 290 200 254 268 - 384 172 292 ... 295 220 283 218 201 182 206 146 233 252 264 194	132 252 270 210 162 285 287 209 304 252 254 268 - 407 198 292 ... 295 234 283 235 211 193 206 150 233 252 264 212	132 238 250 196 164 285 282 209 304 200 254 265 333 407 198 297 200 295 220 274 218 201 170 206 156 227 254 264 210	132 252 270 205 164 285 282 209 304 200 254 265 345 407 206 299 205 295 234 274 218 201 182 206 156 233 254 264 212	142 238 250 196 159 285 282 209 290 200 254 266 345 384 192 283 200 306 234 277 216 199 170 206 ... 227 254 257 194	129 238 270 198 162 289 282 209 311 200 254 266 345 400 192 299 200 367 234 277 216 199 170 206 ... 233 254 257 212	142 238 262 196 159 285 282 209 290 200 254 266 345 384 192 283 200 306 234 277 216 199 170 205 156 227 254 257 194	129 238 270 198 162 289 282 209 311 200 254 266 345 400 192 299 200 367 234 277 216 199 170 205 156 233 254 257 212	142 238 250 198 159 285 282 209 290 211 246 266 345 384 179 292 205 295 234 277 216 201 170 205 156 223 254 257 191	142 252 270 205 162 285 282 209 290 252 256 266 345 407 192 292 205 367 234 297 235 201 170 205 156 233 254 257 191	142 238 250 196 159 285 282 206 296 200 254 265 333 400 179 292 200 287 234 277 223 199 170 205 156 227 254 ... 191	142 252 270 198 159 285 282 206 304 214 254 265 345 407 198 292 205 306 234 277 231 201 182 205 156 233 254 ... 200	142 238 250 198 160 285 278 209 304 200 254 266 345 384 179 292 200 295 220 261 228 203 170 205 146 227 254 232 206	142 238 270 205 164 285 278 209 311 200 254 266 345 400 179 292 205 367 234 281 228 203 170 205 146 233 254 252 206	142 252 262 196 162 285 282 202 290 211 254 266 333 379 180 292 205 295 234 277 216 201 170 205 158 227 254 238 186	142 252 270 202 162 285 282 202 290 211 254 266 345 379 180 292 205 327 234 277 216 201 193 205 158 227 254 238 212	142 238 250 198 159 289 278 206 290 200 246 266 345 407 175 266 200 306 220 277 231 199 170 205 156 227 254 232 191	142 238 270 202 169 289 278 206 290 200 256 266 345 407 175 297 205 367 234 277 231 201 170 205 156 233 254 232 191	142 238 250 210 162 285 278 206 290 200 246 266 345 379 180 292 205 295 220 281 235 201 181 205 150 227 254 257 194	146 252 270 210 165 285 282 206 304 211 256 266 345 379 180 292 205 312 220 281 235 201 181 205 156 233 254 257 194	142 252 262 198 159 282 269 206 290 214 254 266 333 379 192 292 205 327 234 277 212 203 170 205 146 233 254 257 212	142 252 270 198 164 282 282 206 311 252 254 266 345 407 192 292 205 367 234 277 228 203 193 205 156 233 254 257 212	142 238 262 196 154 285 278 206 304 200 254 266 333 379 179 292 205 306 220 283 216 201 181 205 144 233 254 264 210	132 252 270 196 162 285 278 206 311 211 254 266 345 400 179 292 205 327 220 283 216 201 181 205 156 233 254 264 210	142 238 250 196 162 285 273 206 304 200 254 268 333 379 180 292 200 295 234 283 235 201 170 198 146 233 254 264 210	142 252 270 210 162 285 278 206 304 252 254 268 333 400 180 292 205 295 234 283 235 211 182 205 150 233 254 264 210	132 252 250 210 162 289 273 206 304 200 254 268 - 379 172 292 205 295 220 274 218 201 181 205 150 233 254 257 210	132 252 270 210 162 289 273 206 304 211 254 268 - 400 198 292 205 295 220 274 218 201 181 205 156 233 254 257 210	132 252 250 210 162 289 278 206 304 200 254 266 345 379 198 ... 200 295 234 274 ... 211 181 205 150 233 254 257 210	132 252 270 210 162 289 287 206 304 200 254 266 375 400 198 ... 205 327 234 274 ... 211 181 205 150 233 254 257 210

Tabela 2: Perfil de microssatélites, em pares, de base de 29 loci em 172 clones de cacau (continuação).

Clones	Loci																												
	mTcCR 01	mTcCR 02	mTcCR 04	mTcCR 06	mTcCR 07	mTcCR 08	mTcCR 09	mTcCR 10	mTcCR 11	mTcCR 12	mTcCR 13	mTcCR 17	mTcCR 18	mTcCR 19	mTcCR 25	mTcCR 26	mTcCR 32	mTcCR 33	mTcCR 35	mTcCR 40	mTcCR 42	mTcCR 43	mTcCR 44	mTcCR 49	mTcCR 54	mTcCR 55	mTcCR 57	mTcCR 58	mTcCR 60
CAB 326	132	252	250	196	164	285	278	206	304	200	254	268	333	400	198	292	205	306	220	277	216	201	181	198	156	233	252	264	212
	132	252	270	196	164	285	282	206	304	200	254	268	333	400	198	292	205	327	234	277	216	211	181	205	156	233	252	264	212
CAB 330	132	238	262	196	162	285	282	206	304	200	254	266	333	384	180	283	205	327	220	283	218	201	170	215	252	257	212
	132	252	270	210	162	285	284	206	304	211	254	266	333	400	192	297	205	327	234	283	218	213	182	215	254	257	212
CAB 331	142	..	250	196	162	285	278	206	290	200	254	266	333	400	198	290	205	..	234	283	235	209	181	205	146	233	..	242	212
	142	..	270	196	162	285	282	206	304	211	254	266	333	400	198	290	205	..	234	283	235	209	181	215	146	233	..	242	212
CAB 001	132	252	250	196	162	285	282	206	304	211	254	266	333	400	198	292	205	295	220	277	231	201	182	205	150	227	254	242	212
	132	252	270	214	162	285	282	206	304	200	254	266	333	400	206	299	205	306	234	277	235	211	193	205	156	233	254	242	212
CAB 337	132	..	250	196	162	285	278	206	304	200	254	266	333	400	198	292	205	295	220	277	216	201	181	205	156	227	252	242	214
	132	..	270	196	162	285	278	206	304	211	254	266	333	400	206	292	205	295	220	277	216	201	181	205	156	233	254	242	214
CAB 266	142	252	250	198	162	285	282	203	304	200	254	268	333	387	179	292	205	295	220	274	231	201	170	205	156	..	254	264	194
	132	252	270	210	165	285	282	206	304	211	254	268	333	400	179	292	205	327	220	274	231	201	193	205	156	..	254	264	214
CAB 275	132	252	250	196	162	285	282	203	306	200	254	266	333	384	179	292	205	295	220	283	216	211	181	205	156	227	254	257	194
	132	252	270	205	162	285	282	206	306	200	254	268	333	384	179	292	205	295	220	283	216	211	181	205	156	227	254	257	214
CAB 280	132	..	262	205	154	285	273	206	304	211	254	266	345	379	180	292	205	287	220	283	218	..	181	205	150	233	252	264	212
	132	..	270	205	162	289	282	206	304	211	254	266	345	387	198	299	205	295	220	283	218	..	181	205	156	233	254	264	212
CAB 283	132	238	250	196	162	289	273	206	296	200	254	266	333	379	198	292	200	306	234	274	216	201	170	205	150	227	254	257	210
	132	252	270	196	165	289	282	206	306	252	254	266	345	384	198	292	200	327	234	274	216	211	182	205	156	233	254	264	210
CAB 282	142	238	262	210	162	285	273	206	296	211	254	265	333	384	192	290	200	327	230	266	231	211	182	205	156	233	254	238	214
	132	252	270	210	162	285	282	206	304	211	254	266	345	400	206	292	205	327	230	266	231	211	193	205	156	233	254	238	214
CAB 302	142	..	250	196	162	285	273	206	304	200	254	268	333	379	180	292	205	287	230	283	218	211	193	205	146	233	254	257	212
	132	..	270	205	162	285	282	206	304	200	254	268	333	387	198	292	205	327	230	283	231	211	193	205	158	233	254	257	212
CAB 304	142	..	262	196	162	285	282	206	304	211	254	..	333	407	180	292	205	..	220	274	235	211	181	198	254	264	212
	132	..	270	196	162	285	282	206	304	252	254	..	333	407	180	292	205	..	220	274	235	211	181	205	254	264	212
CAB 306	132	238	250	196	162	285	282	206	290	200	254	268	333	379	180	292	200	306	230	283	216	201	182	254	264	212
	132	252	270	196	162	285	282	206	306	252	254	268	333	400	198	292	205	327	230	283	216	201	193	254	264	212
CAB 308	142	252	250	196	165	285	282	206	304	200	254	268	333	400	180	292	205	306	234	283	231	211	182	205	254	264	214
	142	252	270	196	165	285	282	206	304	252	254	268	333	400	180	292	205	327	234	283	231	211	193	205	254	264	214
CAB 517	142	238	262	196	162	285	282	206	306	211	254	268	333	379	198	292	205	295	220	277	218	211	170	198	144	233	252	264	210
	142	252	270	196	162	285	282	206	306	211	254	268	333	407	198	292	205	306	234	277	235	211	182	198	144	233	252	264	210
CAB 313	142	252	250	205	159	285	282	203	290	211	254	266	333	400	180	292	200	306	220	274	231	211	181	205	156	..	252	257	212
	132	252	270	205	162	285	282	203	304	252	254	266	333	400	192	292	205	327	220	274	231	211	181	205	156	..	252	264	212
CAB 312	142	252	250	205	165	285	282	206	296	200	254	266	333	379	179	292	..	306	220	283	231	201	193	205	156	..	252	264	212
	132	252	270	205	165	285	282	206	306	200	254	266	345	400	179	292	..	325	234	283	231	211	193	205	156	..	252	264	212
CAB 316	142	252	262	198	159	285	269	206	304	203	254	266	337	400	192	292	205	325	234	277	228	203	170	198	144	227	252	257	196
	142	252	270	198	162	285	284	206	311	211	254	266	345	400	192	292	205	367	234	277	231	203	182	205	156	233	252	257	214
CAB 514	142	238	262	210	159	285	273	206	306	211	254	266	333	384	198	292	205	295	220	281	231	201	193	205	144	233	254	238	210
	132	252	270	210	165	285	278	206	306	252	254	268	333	400	198	292	205	295	220	281	231	211	193	205	156	233	254	264	210
CAB 529	142	252	250	..	159	285	278	206	306	211	254	266	333	384	180	292	205	..	220	274	231	201	170	205	156	233	254	257	210
	132	252	270	..	165	285	278	206	306	211	254	266	352	407	180	292	205	..	234	274	231	201	193	205	156	233	254	257	214
CAB 181	132	238	250	198	162	285	282	212	304	203	254	266	333	387	192	292	205	302	220	283	223	209	170	198	150	227	252	252	186
	148	252	270	205	162	285	282	212	311	252	254	266	345	407	198	292	205	306	220	283	228	209	182	209	150	233	252	252	212
CAB 182	148	..	250	198	162	289	278	209	296	203	254	268	345	400	192	297	200	302	220	274	202	197	170	206	144	223	254	252	186
	148	..	270	198	162	289	278	209	296	209	254	268	345	400	198	297	205	302	230	274	228	209	182	206	144	227	254	257	186
CAB 183	148	238	262	198	162	285	278	209	287	186	258	265	345	387	192	292	200	318	230	274	202	209	181	205	144	227	254	238	186
	148	252	262	198	162	289	278	209	287	224	258	265	454	407	192	297	200	318	230	283	228	209	181	205	144	227	254	242	206
CAB 184	148	238	262	198	162	289	282	209	285	209	258	266	345	387	198	297	220	274	202	199	170	205	154	227	254	242	186
	148	252	262	198	162	289	282	209	285	209	258	266	345	407	198	297	220	274	202	211	170	205	154	227	254	242	206
CAB 185	148	238	262	198	162	285	278	209</																					

Tabela 3. Frequências alélicas nas amostras (N) representativas de dezesseis sub-populações de cacauzeiros de diferentes bacias hidrográficas, estimadas a partir de 29 locos polimórficos de microssatélites (Continuação).

Locus Alelos	Acará (Pa)	Rio Acre (Ac)	Rio Jari (Ap)	Chandless (Ac)	Rio Iaco (Ac)	Jamari (Ro)	Japurá (Am)	Ji-Paraná (Ro)	Maicuru (Pa)	Rio Purus (Ac/Am)	Solimões (Am)	Sol/Bx-Jap (Am)	Tarauacá (Ac)	Tocantins (Pa)	Uatumã (Am)	Xeruni (Rr)
mTc-25																
(N)	6	22	10	3	15	13	14	9	10	15	19	14	10	2	4	2
172	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,1786	0,0000	0,0000	0,0000	0,0789	0,0000	0,0000	0,0000	0,0000	0,0000
175	0,0833	0,0455	0,3500	0,1667	0,0000	0,5769	0,0714	0,2222	0,0500	0,0000	0,0526	0,0000	0,0000	0,0000	0,0000	0,0000
179	0,0000	0,1818	0,0500	0,5000	0,3667	0,0000	0,0000	0,0000	0,1000	0,4000	0,1842	0,1429	0,0000	0,0000	0,0000	0,2500
180	0,0000	0,0682	0,0000	0,1667	0,0000	0,0000	0,0714	0,0000	0,0000	0,0000	0,1842	0,3571	0,1500	0,2500	0,0000	0,0000
184	0,0000	0,0227	0,0000	0,0000	0,0333	0,0000	0,0357	0,0000	0,0000	0,0000	0,0000	0,0000	0,0500	0,0000	0,0000	0,0000
192	0,8333	0,3182	0,5500	0,1667	0,4667	0,3077	0,0714	0,6667	0,7000	0,4333	0,0789	0,1429	0,4500	0,5000	0,0000	0,7500
198	0,0833	0,3636	0,0000	0,0000	0,1333	0,0769	0,1429	0,0556	0,0500	0,1667	0,3158	0,3214	0,3500	0,2500	0,0000	0,0000
206	0,0000	0,0000	0,0500	0,0000	0,0000	0,0385	0,4286	0,0556	0,1000	0,0000	0,1053	0,0357	0,0000	0,0000	1,0000	0,0000
mTc-26																
(N)	6	22	10	3	15	15	14	9	9	15	18	14	10	2	4	2
266	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0278	0,0000	0,0000	0,0000	0,0000	0,0000
283	0,0000	0,0227	0,0000	0,1667	0,1000	0,7667	0,0000	0,8333	0,1667	0,0000	0,0278	0,0000	0,0000	0,0000	0,0000	0,0000
288	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,1111	0,0667	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
290	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,3889	0,2333	0,0556	0,0357	0,0000	0,0000	0,0000	0,0000
292	0,0000	0,3636	0,1000	0,0000	0,0333	0,0000	0,2143	0,0000	0,2222	0,1333	0,6944	0,9286	0,3500	1,0000	0,2500	0,0000
297	0,6667	0,4545	0,6500	0,3333	0,2333	0,0667	0,7500	0,0000	0,0000	0,3667	0,1111	0,0000	0,5000	0,0000	0,7500	1,0000
299	0,3333	0,1591	0,2500	0,1667	0,6000	0,1667	0,0000	0,1667	0,1111	0,2000	0,0833	0,0357	0,1500	0,0000	0,0000	0,0000
304	0,0000	0,0000	0,0000	0,3333	0,0333	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
310	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0357	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
mTc-32																
(N)	6	22	10	2	14	15	14	9	9	15	18	13	9	2	4	2
196	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,2778	0,0000	0,0000	0,2500
200	0,5833	0,9773	0,7000	1,0000	0,8571	0,9667	0,3929	1,0000	0,5556	0,8000	0,1944	0,1923	0,4444	0,0000	0,0000	0,0000
205	0,4167	0,0227	0,3000	0,0000	0,1429	0,0333	0,6071	0,0000	0,4444	0,2000	0,8056	0,8077	0,2778	1,0000	1,0000	0,7500
mTc-33																
(N)	6	21	10	3	15	15	14	9	10	15	18	12	9	2	3	2
287	0,0000	0,1190	0,0000	0,3333	0,0333	0,6000	0,0000	0,6111	0,1000	0,1667	0,0000	0,0833	0,0000	0,0000	0,0000	0,0000
293	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0714	0,0000	0,0000	0,0667	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
295	0,0000	0,0952	0,0000	0,1667	0,1000	0,0000	0,0000	0,1111	0,3000	0,2667	0,5556	0,2083	0,2222	0,2500	0,0000	0,2500
302	0,0000	0,2857	0,0500	0,0000	0,4667	0,3333	0,1071	0,2778	0,0000	0,1000	0,0000	0,0000	0,3333	0,0000	0,0000	0,0000
306	0,0000	0,4762	0,1000	0,0000	0,1333	0,0000	0,0000	0,0000	0,2500	0,1333	0,1111	0,2500	0,0556	0,0000	0,0000	0,0000
312	0,0000	0,0000	0,4000	0,1667	0,1333	0,0333	0,0357	0,0000	0,0000	0,2333	0,0556	0,0000	0,1111	0,0000	0,0000	0,0000
318	0,0000	0,0000	0,0000	0,3333	0,1333	0,0000	0,0000	0,0000	0,0000	0,0333	0,0000	0,0000	0,2778	0,0000	0,0000	0,0000
325	0,0000	0,0238	0,0000	0,0000	0,0000	0,0333	0,0000	0,0000	0,0000	0,0000	0,0000	0,0833	0,0000	0,0000	0,0000	0,0000
327	0,4167	0,0000	0,1000	0,0000	0,0000	0,0000	0,7857	0,0000	0,0000	0,0000	0,1944	0,3333	0,0000	0,0000	1,0000	0,0000
367	0,5833	0,0000	0,3500	0,0000	0,0000	0,0000	0,0000	0,0000	0,3500	0,0000	0,0833	0,0417	0,0000	0,7500	0,0000	0,7500

Tabela 3. Frequências alélicas nas amostras (N) representativas de dezesseis sub-populações de cacauzeiros de diferentes bacias hidrográficas, estimadas a partir de 29 locos polimórficos de microssatélites (Continuação).

Locus Alelos	Acará (Pa)	Rio Acre (Ac)	Rio Jari (Ap)	Chandless (Ac)	Rio Iaco (Ac)	Jamari (Ro)	Japurá (Am)	Ji-Paraná (Ro)	Maicuru (Pa)	Rio Purus (Ac/Am)	Solimões (Am)	Sol/Bx-Jap (Am)	Tarauacá (Ac)	Tocantins (Pa)	Uatumã (Am)	Xeruni (Rr)
mTc-35																
(N)	6	22	9	3	14	15	13	9	10	15	19	14	10	2	4	2
220	0,0000	0,2727	0,1111	0,5000	0,5000	0,0333	0,0385	0,2222	0,2000	0,6000	0,5000	0,4643	0,7500	0,0000	0,0000	0,2500
230	0,0000	0,6818	0,0000	0,3333	0,5000	0,0000	0,6154	0,1667	0,0500	0,4000	0,0000	0,2143	0,2500	0,0000	0,0000	0,0000
234	1,0000	0,0455	0,8889	0,1667	0,0000	0,9667	0,3462	0,6111	0,7500	0,0000	0,5000	0,3214	0,0000	1,0000	1,0000	0,7500
mTc-40																
(N)	6	22	10	3	15	13	14	9	10	15	19	14	10	2	4	2
259	0,0000	0,0000	0,0000	0,3333	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
266	0,0000	0,3182	0,0000	0,3333	0,6667	0,0385	0,6429	0,0000	0,0000	0,1000	0,0000	0,0714	0,2500	0,0000	0,0000	0,0000
274	0,0000	0,2045	0,0000	0,0000	0,0000	0,0385	0,0714	0,0000	0,0000	0,1000	0,1579	0,3571	0,4500	0,0000	0,0000	0,0000
277	1,0000	0,0455	0,6000	0,0000	0,0000	0,0385	0,2143	0,8333	0,9500	0,8000	0,3684	0,1429	0,0000	1,0000	1,0000	1,0000
281	0,0000	0,0000	0,0500	0,3333	0,0667	0,0000	0,0357	0,0000	0,0000	0,0000	0,1053	0,0714	0,0000	0,0000	0,0000	0,0000
283	0,0000	0,1364	0,0000	0,0000	0,1667	0,3077	0,0357	0,0000	0,0000	0,0000	0,3684	0,3571	0,2500	0,0000	0,0000	0,0000
297	0,0000	0,0227	0,0000	0,0000	0,0000	0,3077	0,0000	0,0000	0,0500	0,0000	0,0000	0,0000	0,0500	0,0000	0,0000	0,0000
298	0,0000	0,2727	0,3500	0,0000	0,1000	0,2692	0,0000	0,1667	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
mTc-42																
(N)	6	21	10	3	15	15	14	9	10	15	18	14	10	2	4	2
202	0,0000	1,0000	0,0000	0,6667	0,8667	0,1000	0,0000	0,0000	0,0000	0,8667	0,0000	0,0000	0,8000	0,0000	0,0000	0,0000
212	0,1667	0,0000	0,0500	0,0000	0,0000	0,0000	0,0000	0,1111	0,0000	0,0000	0,0278	0,0000	0,0000	0,0000	0,0000	0,0000
216	0,0000	0,0000	0,0000	0,0000	0,0000	0,0333	0,2857	0,0000	0,2500	0,0000	0,3056	0,1429	0,0000	0,0000	0,0000	0,0000
218	0,0000	0,0000	0,0000	0,0000	0,0000	0,0333	0,0714	0,0000	0,0000	0,0000	0,2500	0,1429	0,0000	0,0000	0,0000	0,0000
223	0,0000	0,0000	0,6500	0,0000	0,0667	0,5333	0,0357	0,7778	0,3000	0,0000	0,0000	0,0000	0,0500	0,0000	0,0000	0,5000
228	0,0000	0,0000	0,0000	0,3333	0,0667	0,3000	0,0000	0,1111	0,0000	0,1333	0,0833	0,0357	0,1500	0,0000	1,0000	0,5000
231	0,5000	0,0000	0,0500	0,0000	0,0000	0,0000	0,5357	0,0000	0,2000	0,0000	0,0833	0,5357	0,0000	1,0000	0,0000	0,0000
235	0,3333	0,0000	0,2500	0,0000	0,0000	0,0000	0,0714	0,0000	0,2500	0,0000	0,2500	0,1429	0,0000	0,0000	0,0000	0,0000
mTc-43																
(N)	6	22	10	3	15	14	13	9	10	15	19	13	10	2	4	2
191	0,0000	0,0455	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
197	0,0000	0,0000	0,3000	0,5000	0,3000	0,8214	0,0000	0,7778	0,1500	0,1333	0,0000	0,0000	0,0500	0,0000	0,0000	0,0000
199	0,0000	0,3636	0,0000	0,0000	0,0000	0,1429	0,0000	0,2222	0,2500	0,0333	0,0263	0,0000	0,2500	0,0000	0,0000	0,5000
201	0,8333	0,0000	0,5000	0,0000	0,0333	0,0000	0,5000	0,0000	0,5000	0,0333	0,4737	0,3462	0,0000	0,5000	1,0000	0,5000
203	0,1667	0,0000	0,1500	0,0000	0,0333	0,0000	0,2308	0,0000	0,0500	0,0333	0,1053	0,0769	0,0000	0,5000	0,0000	0,0000
209	0,0000	0,3864	0,0500	0,3333	0,4000	0,0000	0,1923	0,0000	0,0500	0,1333	0,0526	0,0000	0,4500	0,0000	0,0000	0,0000
211	0,0000	0,2045	0,0000	0,1667	0,2000	0,0000	0,0769	0,0000	0,0000	0,5667	0,3158	0,5769	0,2500	0,0000	0,0000	0,0000
213	0,0000	0,0000	0,0000	0,0000	0,0333	0,0357	0,0000	0,0000	0,0000	0,0667	0,0263	0,0000	0,0000	0,0000	0,0000	0,0000

Tabela 3. Frequências alélicas nas amostras (N) representativas de dezesseis sub-populações de cacauzeiros de diferentes bacias hidrográficas, estimadas a partir de 29 locos polimórficos de microsátélites (Continuação).

Locus Alelos	Acará (Pa)	Rio Acre (Ac)	Rio Jari (Ap)	Chandless (Ac)	Rio Iaco (Ac)	Jamari (Ro)	Japurá (Am)	Ji-Paraná (Ro)	Maicuru (Pa)	Rio Purus (Ac/Am)	Solimões (Am)	Sol/Bx-Jap (Am)	Tarauacá (Ac)	Tocantins (Pa)	Uatumã (Am)	Xeriuni (Rr)
mTc-58																
(N)	6	22	10	3	15	15	14	9	8	15	19	14	10	2	4	2
232	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0789	0,0000	0,0000	0,0000	0,0000	0,0000
238	0,1667	0,0000	0,0000	0,0000	0,0000	0,0000	0,5714	0,0000	0,0000	0,0667	0,0526	0,1071	0,0500	0,0000	0,0000	0,0000
242	0,0000	0,0227	0,0000	0,1667	0,0333	0,0667	0,0000	0,0000	0,0000	0,2333	0,1579	0,0000	0,3500	0,0000	0,0000	0,5000
252	0,0000	0,0000	0,0000	0,0000	0,1000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0263	0,0000	0,1500	0,0000	0,0000	0,0000
257	0,8333	0,1136	0,8000	0,3333	0,5667	0,0000	0,1786	0,8889	1,0000	0,3667	0,3684	0,2857	0,1500	1,0000	0,0000	0,5000
264	0,0000	0,7955	0,0000	0,5000	0,3000	0,0000	0,2500	0,1111	0,0000	0,3333	0,3158	0,6071	0,3000	0,0000	1,0000	0,0000
267	0,0000	0,0682	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
309	0,0000	0,0000	0,2000	0,0000	0,0000	0,9333	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
mTc-60																
(N)	6	22	10	3	15	15	14	9	10	15	19	14	10	2	4	2
186	0,0000	0,6818	0,0000	0,6667	0,9000	0,0333	0,0000	0,8889	0,0000	0,8667	0,0526	0,0000	0,7000	0,0000	0,0000	0,0000
191	0,0000	0,0000	0,1500	0,3333	0,0333	0,9333	0,2143	0,0000	0,4500	0,0000	0,0526	0,0000	0,0000	0,0000	0,0000	0,7500
194	0,0000	0,0000	0,2000	0,0000	0,0000	0,0000	0,0714	0,0000	0,2500	0,0000	0,1053	0,0357	0,0000	0,0000	0,0000	0,0000
196	0,0000	0,2500	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,1333	0,0000	0,0357	0,0000	0,0000	0,0000	0,0000
200	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0500	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
206	0,0000	0,0455	0,0000	0,0000	0,0333	0,0000	0,0000	0,1111	0,1500	0,0000	0,0526	0,0000	0,2500	0,0000	0,0000	0,2500
208	0,0000	0,0227	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
210	1,0000	0,0000	0,0500	0,0000	0,0000	0,0000	0,0714	0,0000	0,0000	0,0000	0,2632	0,2500	0,0000	0,0000	0,0000	0,0000
212	0,0000	0,0000	0,3500	0,0000	0,0000	0,0333	0,3571	0,0000	0,1000	0,0000	0,3947	0,4286	0,0500	1,0000	0,6250	0,0000
214	0,0000	0,0000	0,2500	0,0000	0,0333	0,0000	0,2857	0,0000	0,0000	0,0000	0,0789	0,2500	0,0000	0,0000	0,0000	0,0000
224	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,3750	0,0000