

Diversidade genética estimada com marcadores entre sequências simples repetidas em cultivos comerciais de Cupuaçuzeiro

Genetic diversity estimated using inter-simple sequence repeat markers in commercial crops of cupuassu tree

Bruna Mezzalira da Silva^I Ana Aparecida Bandini Rossi^{II}
Juliana de Freitas Encinas Dardengo^{III} Vitor Arreguy Amado Correa de Araujo^{III}
Fernanda Saragosa Rossi^{IV} Luiz Orlando de Oliveira^V Wellington Ronildo Clarindo^{VI}

RESUMO

Quinze primers ISSR (entre sequências simples repetidas) foram utilizados para avaliar a diversidade genética entre e dentro de pomares comerciais de *Theobroma grandiflorum* (Willd. ex Spreng.) K. Schum. Para isso, foram analisados sessenta indivíduos, distribuídos nos três cultivos. Um total de 102 bandas foi amplificado, com uma porcentagem de 52,0% de polimorfismo em nível de espécie e média de 6,8 alelos por primer ISSR. A média do Índice de Conteúdo Polimórfico (PIC) foi de 0,55. Em relação aos índices de diversidade gênica de Nei (H) e de Shannon (I), os cultivos analisados apresentaram os valores: SAR H = 0,114 e I = 0,177; SSL H = 0,108 e I = 0,162 e SEC H = 0,104 e I = 0,156, considerados valores de moderados a baixos. A AMOVA revelou 34,91% da variância total entre os cultivos e 65,09% dentro deles. Os marcadores moleculares ISSR revelaram que há diversidade genética dentro de cada cultivo comercial estudado, portanto é possível selecionar genótipos superiores que poderão ser utilizados para originar cultivos mais uniformes. Esse resultado tem sido considerado de grande relevância, por fornecer ferramentas para a implementação de programas de melhoramento e delineamento de estratégias de conservação ex situ e in situ.

Palavras-chave: *Theobroma grandiflorum*, variabilidade genética, marcadores moleculares.

ABSTRACT

Fifteen ISSR (inter-simple sequence repeat) primers were used to evaluate the genetic diversity among and within

commercial crops of *T. grandiflorum* (Willd. ex Spreng.) K. Schum. For this, 60 specimens were analyzed, distributed in three crops. A total of 102 bands were amplified, with a polymorphism percentage of 52.0% at species level and an average of 6.8 alleles per ISSR primer. The average for polymorphism information content (PIC) index was 0.55. In relation to the genetic diversity index of Nei (H), and Shannon (I), crops analyzed showed the following values: SAR H = 0,114 e I = 0,177; SSL H = 0,108 e I = 0,162 e SEC H = 0,104 e I = 0,156, considered moderate to low values. AMOVA showed 34.91% of total variance among the crops, and 65.09% within them. The ISSR molecular markers revealed that there is genetic diversity within each commercial crops studied, thus is possible to select superior genotypes that can be used to give more uniform crops. This result has been considered of great relevance, to provide tools for breeding implementation programs and design conservation strategies ex situ and in situ.

Key words: *Theobroma grandiflorum*, genetic variability, molecular markers.

INTRODUÇÃO

Theobroma grandiflorum, conhecido popularmente como cupuaçu, é uma espécie frutífera

^IPrograma de Pós-graduação em Genética e Melhoramento de Plantas, Universidade do Estado de Mato Grosso (UNEMAT), Alta Floresta, MT, Brasil.

^{II}Faculdade de Ciências Biológicas e Agrárias, UNEMAT. 78580-000, Alta Floresta, MT, Brasil. E-mail: anabanrossi@gmail.com. Autor para correspondência.

^{III}Programa de Pós-graduação em Biodiversidade e Biotecnologia pela rede BIONORTE-MT, UNEMAT, Alta Floresta, MT, Brasil.

^{IV}Programa de Pós-graduação em Genética e Melhoramento de Plantas, Universidade Estadual do Norte Fluminense (UENF), Campos dos Goytacazes, RJ, Brasil.

^VDepartamento de Bioquímica e Biologia Celular, Universidade Federal de Viçosa (UFV), Viçosa, MG, Brasil.

^{VI}Departamento de Biologia, Universidade Federal do Espírito Santo (UFE), Alegre, ES, Brasil.

nativa da Amazônia brasileira e, nas últimas décadas, transformou-se em um cultivo importante para a agricultura do Norte do Brasil (ALVES et al., 2003).

O cupuaçuzeiro é uma árvore que, em condições naturais, pode atingir mais de 30m de altura. Nas áreas cultivadas, sua altura varia de 6 a 10m, apresenta dois picos de floração: um menor entre julho e agosto; e um principal que se dá entre outubro e novembro. O fruto do cupuaçuzeiro também chamado de “cupu” possui de 12cm a 25cm de comprimento e 10cm a 12cm de diâmetro, apresentando em média peso de 1,0kg, sendo 30% de polpa e 15% a 20% de sementes (em média, 35 unidades) (FRAIFE FILHO, 1984).

O cupuaçu apresenta grandes perspectivas agroindustriais. Alguns anos atrás, sua produção era considerada semiextrativista, hoje, desponta como uma das melhores e mais promissoras frutíferas da Amazônia, apresentando-se como uma alternativa rentável e economicamente viável. Vários produtos são fabricados a partir da polpa, como sucos, sorvetes, cremes, bombons, doces, licores e compotas, além de produtos de beleza. A partir da semente, obtém-se um chocolate de excelente qualidade, conhecido como cupulate. (CAVALCANTE, 1991).

No norte do estado de Mato Grosso, a contínua expansão do comércio de polpa de cupuaçu tem gerado uma demanda crescente pelo produto, garantindo ao produtor a segurança de venda da produção agrícola dessa espécie e, portanto, aumentando a renda familiar dos agricultores da região.

O uso de marcadores moleculares tem tido cada vez mais destaque como uma ferramenta importante para o melhoramento genético de plantas, visando ao mapeamento de genes, à análise de diversidade genética, aos diagnósticos de doenças e aos estudos taxonômicos e evolutivos (WÜNSCH & HORMAZA, 2007).

Os marcadores ISSR (*inter-simple sequence repeat*) são amplamente utilizados em estudos de diversidade e variabilidade genética, por não necessitarem de informação prévia da sequência de DNA, ter baixos custos de desenvolvimento e os procedimentos laboratoriais podem ser transferidos para qualquer espécie de planta (BARTH et al., 2002),

Considerando a importância do cultivo do cupuaçu para a região norte do Mato Grosso e visando a conservação da variabilidade genética por meio de programas de preservação, o presente estudo empregou os marcadores moleculares ISSR para avaliar a diversidade genética entre e dentro de cultivos comerciais de cupuaçu no município de Alta Floresta-MT.

MATERIAL E MÉTODOS

Três cultivos comerciais de *T. grandiflorum* (Willd. ex Spreng.) K. Schum, localizados no norte do Estado de Mato Grosso, no Município de Alta Floresta, foram amostrados: Setor Industrial (SSL); Estrada 5ª leste (SAR); e Comunidade Sol Nascente (SEC) (Figura 1). Os cultivos foram implantados pelos produtores rurais com o propósito de aumentar a renda familiar, através da comercialização do fruto *in natura* e processado em suas diversas formas.

Para a implantação dos cultivos de cupuaçuzeiro, os agricultores adquiriram mudas distribuídas pela CEPLAC do município de Alta Floresta-MT nos anos de 2003 e 2004. As mudas foram obtidas a partir de sementes coletadas no pomar da CEPLAC/Alta Floresta-MT, que, por sua vez, foi implantado através da introdução de genótipos provenientes do Banco de Germoplasma da CEPLAC, em Belém-PA.

As três áreas estudadas estão distantes entre si aproximadamente 30km, não havendo troca de germoplasma entre os produtores e tampouco novos cultivos originados a partir dos cultivos estudados.

Em cada cultivo, foram coletadas folhas jovens de 20 indivíduos, as quais foram utilizadas para extração de DNA total. O DNA foi extraído com base no protocolo de CTAB (DOYLE & DOYLE, 1987), com modificações para *T. grandiflorum*, aumentando a concentração de CTAB para 5% e utilizando concentração de PVP de 2% e de b-mercaptoetanol de 2,5% no tampão de extração. A qualidade e a concentração do DNA foram confirmadas por eletroforese em gel de agarose 1%, com marcador DNA l.

As análises de ISSR foram conduzidas por meio de 15 *primers*, selecionados com base no padrão de amplificação inicial com 40 marcadores de ISSR.

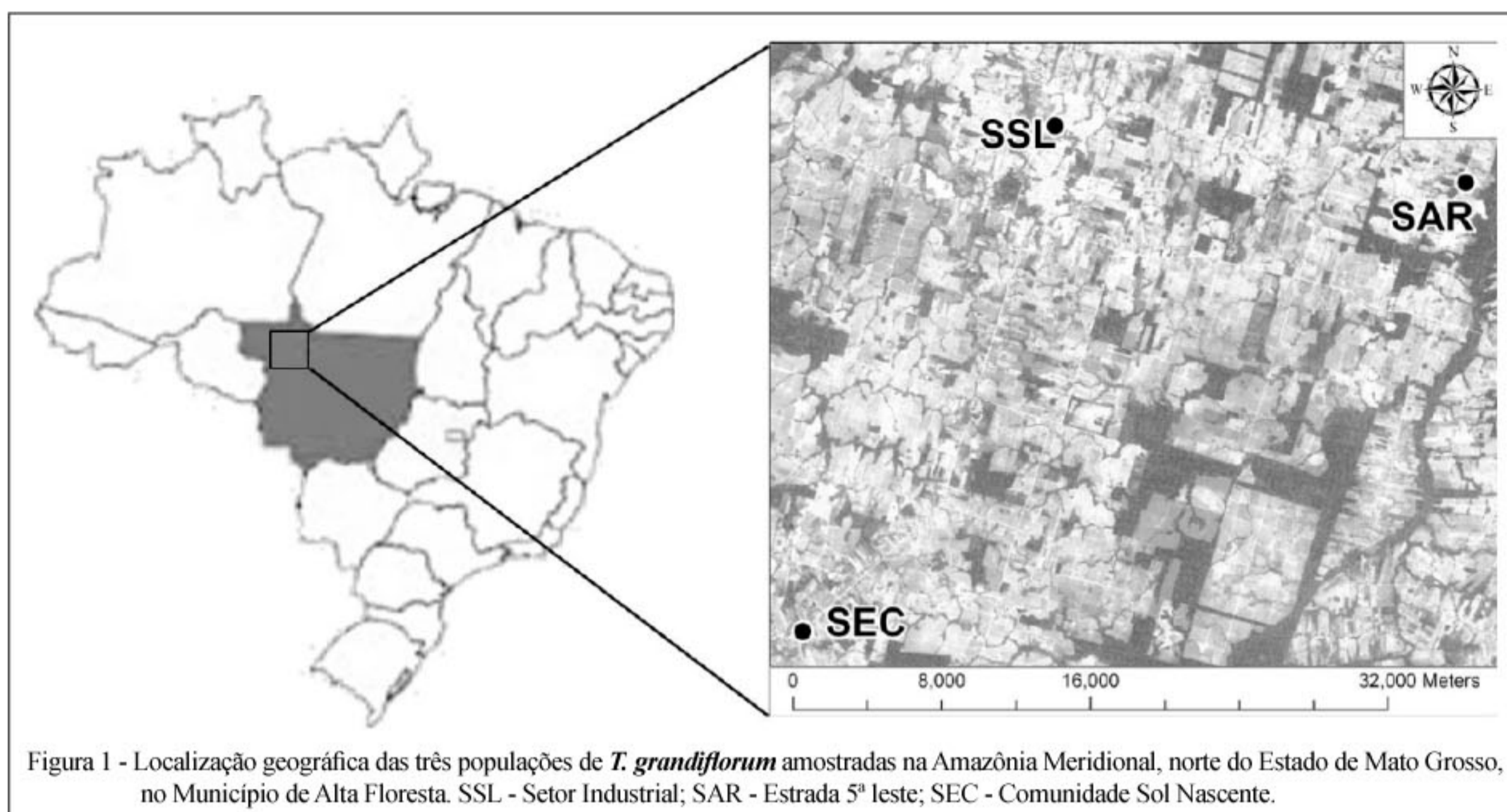


Figura 1 - Localização geográfica das três populações de *T. grandiflorum* amostradas na Amazônia Meridional, norte do Estado de Mato Grosso, no Município de Alta Floresta. SSL - Setor Industrial; SAR - Estrada 5ª leste; SEC - Comunidade Sol Nascente.

As condições de amplificação por meio da reação em cadeia da polimerase (PCR) foram de acordo com o descrito por CHARTERS & WILKINSON (2000). Os produtos de amplificação foram separados por eletroforese em gel de agarose 1,5% em tampão de corrida TBE 1X. A coloração do gel foi feita com brometo de etídeo ($0,6\text{ng mL}^{-1}$).

Os fragmentos de ISSR foram tratados como marcadores dominantes e julgados como caracteres binários: presente (1) ou ausente (0). A matriz binária de presença/ausência foi analisada utilizando o programa POPGENE 1.31 (YEH et al., 1999) para estimar a porcentagem de locos polimórficos (P), o índice de diversidade de Shannon (I) e a diversidade gênica de Nei (H) (NEI, 1978). A análise de variância molecular (AMOVA) foi utilizada para revelar a distribuição da diversidade genética dentro e entre os cultivos; ela foi realizada de acordo com EXCOFFIER et al. (1992), com o auxílio do programa ARLEQUIN 3.01 (EXCOFFIER et al., 2006).

A diversidade genética do loco ou PIC (Índice de Conteúdo Polimórfico) é uma estimativa utilizada para a avaliação do poder discriminatório de um loco. A informatividade do loco p_i é a frequência do alelo p no loco p_i , calculado pela equação: $\text{PIC} = 1 - \sum p_i^2$, e a

informatividade do primer p_{ij} é a frequência do alelo p do loco i , no primer j , sendo calculada pela equação: $\text{PIC}_{\text{primer}} = 1 - \sum_i \cdot \sum_j p_{ij}^2$ (REZENDE et al., 2009).

A análise de coordenadas principais (PCoA) em nível de indivíduo foi realizada de acordo com Nei e do coeficiente de Li (1979) por meio do programa NTSYSpc 2.2 (ROHLF, 2005).

O programa “Structure” (PRITCHARD et al., 2000), baseado em estatística bayesiana, foi utilizado para inferir o número de grupos (K), nos quais os genótipos encontram-se estruturados. Foram realizadas 20 corridas para cada valor de K, 200.000 interações iniciais (*burn-ins*) e 500.000 simulações de Monte Carlo, via Cadeias de Markov (MCMC). Para definição do K, mais provável em relação aos propostos, foram utilizados os critérios descritos por PRITCHARD & WEN (2004) e EVANO et al. (2005).

RESULTADOS E DISCUSSÃO

Os 15 primers de ISSR selecionados produziram um total de 102 bandas, sendo 52% polimórficas. O cultivo SAR revelou maior porcentagem de polimorfismo (40,2%) em comparação com os cultivos SEC e SSL, que apresentaram 32,65 e 31,4% de polimorfismo, respectivamente (Tabela 1).

Tabela 1 - Diversidade genética dentro de populações de *T. grandiflorum*.

Populações	N	P(%)	H	I
SAR	20	40,2	0,11	0,17
SEC	20	32,6	0,10	0,15
SSL	20	31,8	0,10	0,16
Média		34,9	0,10	0,16
Espécie	60	52,0	0,15	0,24

N, tamanho da amostra; P, % de polimorfismo; H, diversidade gênica de NEI (1978); I, índice de diversidade gênica de Shannon.

Neste estudo, os valores de PIC variaram de 0,00 do *primer* DiAG3'G até 0,69 do *primer* DiAC3'C, com média de 0,55. O número de locos por *primer* variou de 3 (em DiTG3'G) a 11 (em DiAC3'G), com média de 6,8 alelos por *primer* ISSR, valor similar ao encontrado por LANAUD et al. (1999), que, analisando 24 genótipos de *Theobroma cacao*, obtiveram uma média de 5,6 alelos por locos. Os *primers* DiAG3'YT, DiAC3'C e DiAC3'G tiveram um valor de PIC equivalente 0,55; 0,69 e 0,67, respectivamente, portanto, esses *primers* foram considerados muito informativos para o cupuaçu. Os resultados estão de acordo com a classificação descrita por BOTSTEIN et al. (1980), os quais relataram que marcadores com valores de PIC superiores a 0,5 são considerados muito informativos.

Em relação aos índices de diversidade gênica de Nei (H) e de Shannon (I), os cultivos analisados apresentaram os valores: SAR H = 0,114 e I = 0,177; SSL H = 0,108 e I = 0,162 e SEC H = 0,104 e I = 0,156 (Tabela 1), considerados valores de moderados a baixos, uma vez que os índices variam de 0 a 1, sendo 0 considerado diversidade gênica nula e 1 diversidade gênica máxima (GIUSTINA et al., 2014). Tais resultados podem ser justificados por se tratar de cultivos comerciais, em que a uniformidade da cultura é primordial.

Para a espécie, encontrou-se uma diversidade gênica de Nei igual a 0,160 e índice de diversidade gênica de Shannon de 0,244 (Tabela 1), valores próximos aos encontrados por RIVAS et al. (2013), em estudo com *Theobroma subincanum* (H= 0,151 e I= 0,219), e por GIUSTINA et al., 2014 ao estudar *Theobroma speciosum* (H= 0,167 e I= 0,254).

O resultado da AMOVA mostrou que a maior diversidade genética ocorre dentro dos pomares (65,09%), sugerindo que há variabilidade a ser explorada em cada um dos referidos cultivos (Tabela 2). Estudos realizados com espécies de plantas tropicais também obtiveram o mesmo padrão de distribuição da variabilidade (BERTONI et al., 2010).

Pesquisas com marcadores RAPD e ISSR têm comprovado que espécies com fecundação cruzada comumente são caracterizadas por baixos níveis de diversidade entre populações e alta diferenciação dentro de populações (SUN et al., 2006).

A análise das coordenadas principais (Figura 2) indicou que os indivíduos estão separados em três grupos, de acordo com sua procedência. O agrupamento dos indivíduos por localidade geográfica indica que, apesar de terem sido originados a partir do mesmo banco de germoplasma, os pomares possivelmente são provenientes de matrizes diferentes. A maior parte da variação entre indivíduos é explicada pela primeira coordenada principal (18,1%), que, em conjunto com a segunda e terceira coordenadas principais, representam cerca de 37% da variação total.

O programa *Structure*, baseado em estatística bayesiana, foi utilizado para inferir o número de grupos (k). Segundo PRITCHARD & WEN (2004), nesse programa, o número de grupos formados pode ser pré-determinado, mas são os dados que definem o K (número de grupos), cujo valor mais confiável é estimado pelo menor número com valor negativo de Ln e pelo menor desvio padrão encontrado durante a análise estatística.

Tabela 2 - Análise de variância molecular (AMOVA) das três populações de *T. grandiflorum*, estudadas a partir de 15 marcadores de ISSR.

Fonte de Variação	GL*	SQ*	CV*	VT (%)*	Valor de p*
Entre Populações	2	130.933	2.99421	34.91	<0,000
Dentro de Populações	57	318.200	5.58246	65.09	
Total	59	449.133	8.57667		

Grau de Liberdade (GL), Soma dos Quadrados (SQ), Componente de Variância (CV), Variância Total (VT) e P são as probabilidades de ter um componente de variância maior que os valores observados ao acaso. As probabilidades foram calculadas por 1023 permutações ao acaso. $F_{ST} = 0.34911$.

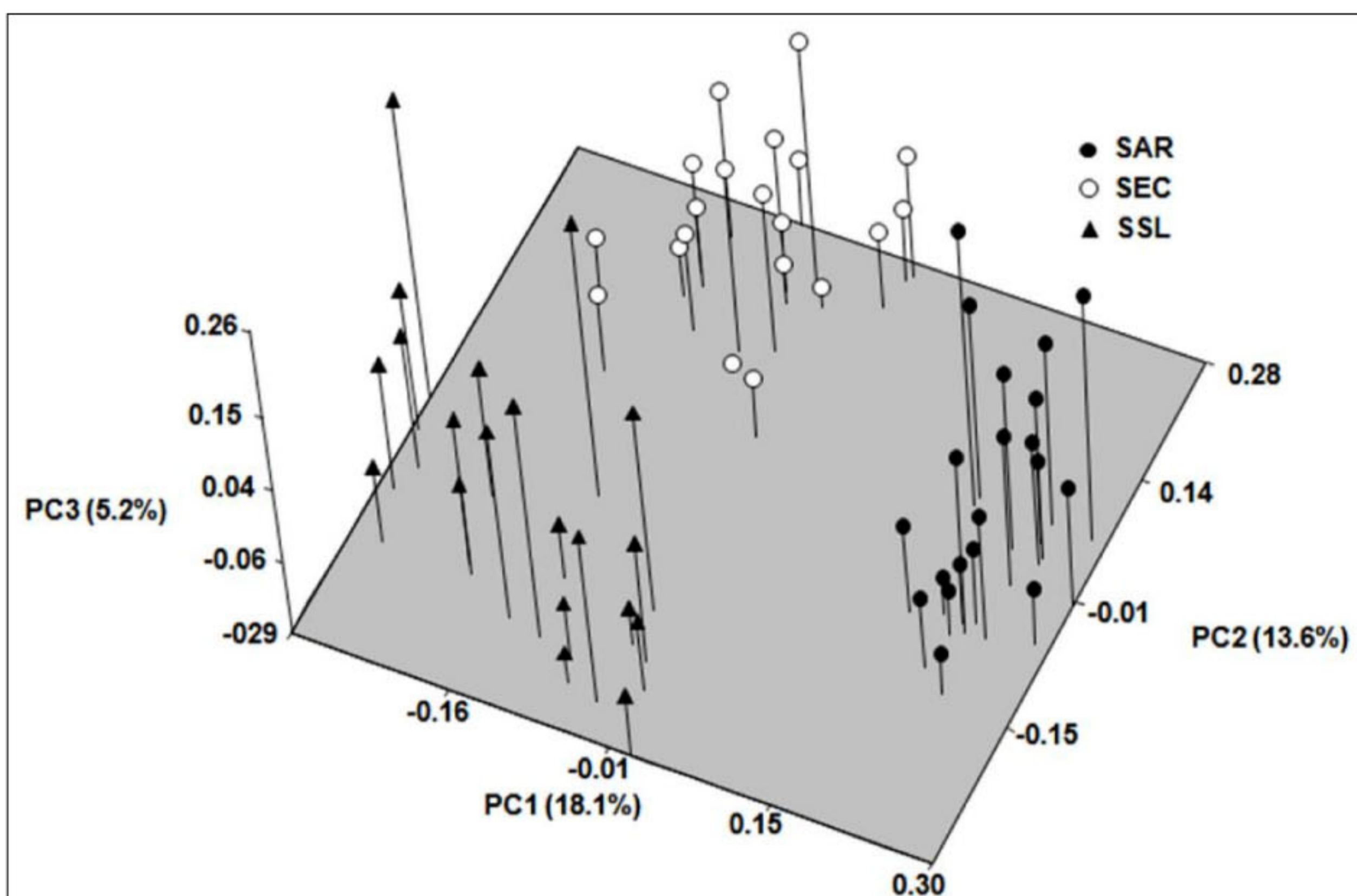


Figura 2 - Representação gráfica das três primeiras coordenadas principais para cada indivíduo de *Theobroma grandiflorum*, amostrado nas três populações estudadas.

O melhor k encontrado relacionando os três cultivos foi representado por três grupos ($k=3$), pôde-se verificar que os indivíduos estão claramente alocados. Comparando os dados da análise bayesiana determinada pelo *Structure* com o agrupamento da análise das coordenadas principais, pode-se notar que houve uma total correspondência entre os grupos (Figura 3).

Apesar de serem cultivos comerciais, há variabilidade genética dentro de cada cultivo estudado, revelando que eles podem servir como fontes de genes em programas de melhoramento genético do cupuaçu.

CONCLUSÃO

A caracterização molecular revelou que há diversidade genética entre e dentro de cada cultivo comercial estudado. O nível de resolução genética e confiabilidade obtida por meio da análise com marcadores moleculares ISSR possibilitou a discriminação de genótipos diferentes geneticamente, que poderão ser utilizados para o manejo dos recursos genéticos da espécie e, uma vez selecionados, poderão originar cultivos mais uniformes, sendo importante para programas de melhoramento que visam à seleção de genótipos superiores e à uniformidade da cultura.

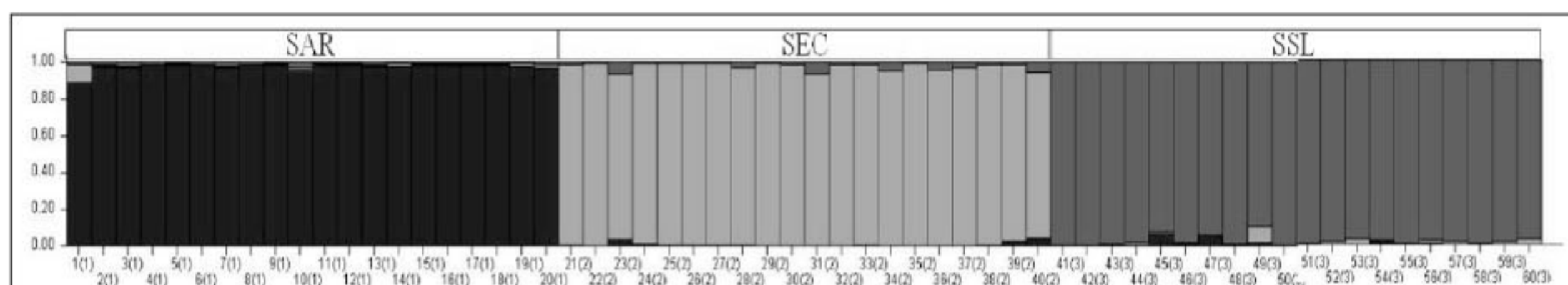


Figura 3 - Representação dos 60 indivíduos de *Theobroma grandiflorum* dos três cultivos comerciais estudados: SSL - Setor Industrial; SAR - Estrada 5ª leste; SEC - Comunidade Sol Nascente, em grupos segundo dados moleculares com 15 ISSRs, utilizando o programa "Structure". Os indivíduos estão representados por barras verticais com coloração de acordo com o grupo ao qual pertencem (três grupos, $K = 3$).

AGRADECIMENTOS

À Fundação de amparo à pesquisa do Estado de Mato Grosso (FAPEMAT), pelo apoio financeiro ao Projeto: "Diversidade, conservação e uso de recursos genéticos de quatro espécies amazônicas de *Theobroma* com potencial econômico para a região"; (PROCESSO nº. 461849/2009). Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela bolsa concedida à Bruna Mezzalira da Silva.

REFERÊNCIAS

- ALVES, R.M. et al. Seleção de descritores botânico-agronômicos para caracterização de germoplasma de cupuaçuzeiro. **Pesquisa agropecuária brasileira**, v. 38, n.7, p.807-818, 2003. Disponível em: <<http://dx.doi.org/10.1590/S0100-204X2003000700004>>. Acesso em: 15 mar. 2013. doi: 10.1590/S0100-204X2003000700004.
- BARTH, S. et al. Genetic diversity in *Arabidopsis thaliana* L. Heynh. Investigated by cleaved amplified polymorphic sequence (CAPS) and inter-simple sequence repeat (ISSR) markers. **Molecular Ecology**, v.11, p.495-505, jul. 2002.
- BERTONI, B.W. et al. Genetic diversity in natural populations of *Jacaranda decurrens* Cham. Determined using RAPD and AFLP markers. **Genetics and Molecular Biology**, n.33, p.532-538, 2010. Disponível em: <<http://dx.doi.org/10.1590/S1415-47572010005000068>>. Acesso em: 20 mar. 2013. doi: 10.1590/S1415-47572010005000068.
- BOTSTEIN, D. et al. Construction of a genetic linkage map in man using restriction fragment length polymorphism. **American Journal of Human Genetics**, v.32, p.314-331, 1980. Disponível em: <<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1686077/pdf/ajhg00189-0020.pdf>>. Acesso em: 27 mar. 2013. doi: PMC1686077/pdf/ajhg00189-0020.
- CAVALCANTE, P.B. **Frutas comestíveis da Amazônia**. Belém: CEJUP/Museu Paraense Emílio Goeldi, 1991. 150p.
- CHARTERS, Y.M.; WILKINSON, M.J. The use of self-pollinated progenies as 'in-groups' for the genetic characterization of cocoa germplasm. **Theoretical and Applied Genetics**, n.100, p.160-166, 2000. Disponível em: <<http://link.springer.com/content/pdf/10.1007%2FPL00002903>>. Acesso em: 27 mar. 2013. doi: 10.1007%2FPL00002903.
- DOYLE, J.J.; DOYLE, J.L. Isolation of plant DNA from fresh tissue. **Focus**, n.12, p.013-15, 1987.
- EVANO, G. et al. Detecting the number of clusters of individuals using the software structure: a simulation study. **Molecular Ecology**, v.14, p.2611-2620, 2005.
- EXCOFFIER, L. et al. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. **Genetics**, v.131, p.479-491, 1992. Disponível em: <<http://www.genetics.org/content/131/2/479.full.pdf+html>>. Acesso em: 27 mar. 2013. doi: 131/2/479.
- EXCOFFIER, L. et al. **ARLEQUIN: an integrated software package for population genetics data analysis**. Bern: University of Bern, Zoological Institute, 2006. Version 3.01
- FRAIFEFILHO, G.A. **Alternativas para diversificação agroeconômica da região Sudeste da Bahia**. Ilhéus: CEPLAC, 1984. 120 p.
- GIUSTINA, L.D. et al. Population structure and genetic diversity in natural populations of *Theobroma speciosum* Willd. Ex Spreng (*Malvaceae*). **Genetics and Molecular Research**, v.13, n.5, p.47-53, 2014.
- LANAUD, C. et al. Isolation and characterization of microsatellites in *Theobroma cacao* L. **Molecular Ecology**, v.8, p.2141-2143, 1999. Disponível em: <<http://onlinelibrary.wiley.com/doi/10.1046/j.1365-294x.1999.00802.x/pdf>>. Acesso em: 24 mar. 2013. doi: 10.1046/j.1365-294x.1999.00802.
- NEI, M. Estimation of average heterozygosity and genetic distance from a small number of individuals. **Genetics**, v.89, p.583-590, 1978. Disponível em: <<http://www.genetics.org/content/89/3/583.full.pdf+html>>. Acesso em: 27 mar. 2013. doi: 89/3/583.
- PRITCHARD, J.K. et al. Inference of population structure using multilocus genotype data. **Genetics**, v.155, p.945-959, 2000.
- PRITCHARD, J.K.; WEN W. **Documentation for structure software: Version 2.1**, 2004. Disponível em: <<http://pritch.bsd.uchicago.edu>>. Acesso em: 15 abr. 2013.
- REZENDE, R.K. et al. Divergência genética entre cultivares de gérbera utilizando marcadores RAPD. **Ciência Rural**, v.39, n.8, p.61-72 2009. Disponível em: <<http://dx.doi.org/10.1590/S0103-84782009005000176>>. Acesso em: 27 mar. 2013. doi: 10.1590/S0103-84782009005000176.
- RIVAS, L.H. et al. Genetic diversity in natural populations of *Theobroma subincanum* Mart. in the Brazilian Amazon. **Genetics and Molecular Research**, v.12, n.4, p. 34-41, 2013.
- ROHLF, F.J. **NTSYS-pc**. Numerical taxonomy and multivariate analysis system. Setauket: Exeter Software, 2005. Version 2.2.
- SUN, K. et al. Genetic variation in *Hippoph aerhamnoides* ssp. *sinensis* (*Elaeagnaceae*) revealed by RAPD markers. **Biochemical Genetics**, n.44 p.186-197, 2006. Disponível em: <<http://evolution.ibcas.ac.cn/PDF/SunK%20et%20al-2006.pdf>>. Acesso em: 17 mar. 2013. doi: 0.1007/s10528-006-9025-2].
- YEH, F.C. et al. **POPGENE**. Microsoft Windows-based freeware for population genetic analysis. Release 1.31. Edmonton: University of Alberta, 1999. 149p.
- WÜNSCH, A.; HORMAZA J.I. Characterization of variability and genetic similarity of European pear using microsatellite loci developed in apple. **Scientia Horticulturae**, v.113, p.37-43, 2007.