

UFRRJ

INSTITUTO DE AGRONOMIA

CURSO DE PÓS-GRADUAÇÃO EM FITOTECNIA

DISSERTAÇÃO

**Lista Preliminar de Descritores, Caracterização
Morfoagronômica e Diversidade Genética em Germoplasma
de Pinhão-manso (*Jatropha curcas* L.)**

Marcus Vinicius Magro Reis

2013



**UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO DE AGRONOMIA
CURSO DE PÓS-GRADUAÇÃO EM FITOTECNIA**

**LISTA PRELIMINAR DE DESCRITORES, CARACTERIZAÇÃO
MORFOAGRONÔMICA E DIVERSIDADE GENÉTICA EM
GERMOPLASMA DE PINHÃO-MANSO
(*Jatropha curcas* L.)**

MARCUS VINICIUS MAGRO REIS

Sob a Orientação do Professor
Pedro Corrêa Damasceno Junior

Dissertação submetida como requisito
parcial para obtenção do grau de
**Mestre em Ciências no Curso de
Pós-Graduação em Fitotecnia.**

Seropédica, RJ
Julho de 2013

633.85

R3751

T

Reis, Marcus Vinicius Magro, 1984-
Lista preliminar de descritores,
caracterização morfoagronômica e
diversidade genética em germoplasma de
pinhão-manso (*Jatropha curcas* L.) / Marcus
Vinicius Magro Reis. - 2013.
118 f.: il.

Orientador: Pedro Corrêa Damasceno
Junior.

Dissertação (mestrado) - Universidade
Federal Rural do Rio de Janeiro, Curso de
Pós-Graduação em Fitotecnia, 2013.

Bibliografia: f. 74-88.

1. Pinhão-manso - Genética - Teses. 2.
Pinhão-manso - Genética - Métodos
estatísticos - Teses. 3. Pinhão-manso -
Melhoramento genético - Teses. 4.
Pinhão-manso - Recursos do germoplasma -
Teses. 5. Biodiesel - Teses. I. Damasceno
Junior, Pedro Corrêa, 1973-. II.
Universidade Federal Rural do Rio de
Janeiro. Curso de Pós-Graduação em
Fitotecnia. III. Título.

**UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO DE AGRONOMIA
CURSO DE PÓS-GRADUAÇÃO EM FITOTECNIA**

MARCUS VINICIUS MAGRO REIS

Dissertação submetida como requisito parcial para obtenção do grau de **Mestre em Ciências**
no Curso de Pós-Graduação em Fitotecnia

DISSERTAÇÃO APROVADA EM: 29/07/2013

Pedro Corrêa Damasceno Junior. D.Sc. Genética e Melhoramento de Plantas. UFRRJ.
(Orientador)

Rosana Rodrigues. D.Sc. Produção Vegetal / Genética e Melhoramento de Plantas. UENF.

Luiz Beja Moreira. D.Sc. Fitotecnia. UFRRJ

RESUMO

REIS, Marcus Vinicius Magro. **Lista Preliminar de Descritores, Caracterização Morfoagronômica e Diversidade Genética em Germoplasma de Pinhão-mansô (*Jatropha curcas* L.)**. 2013. 109p. Dissertação (Mestrado em Fitotecnia). Instituto de Agronomia, Departamento de Fitotecnia, Universidade Federal Rural do Rio de Janeiro, Seropédica, RJ, 2013.

A crescente demanda por fontes energéticas de menor impacto ambiental têm promovido um grande interesse acerca de potenciais espécies oleaginosas, alternativas à exploração de combustíveis fósseis. Sob este aspecto, a euforbiácea *Jatropha curcas* L., popularmente conhecida como pinhão-mansô, se destaca pelo óleo de excelente qualidade para produção de biodiesel. Apesar de promissora, a carência por variedades geneticamente melhoradas, aliada à desuniformidade dos materiais hoje cultivados, ainda se apresentam como os principais obstáculos ao plantio comercial da espécie. Nesse contexto, o conhecimento da diversidade genética entre um grupo de genitores se faz essencial ao uso racional de seus recursos genéticos. Entretanto, informações sobre a diversidade genética existente no germoplasma de *J. curcas* ainda se mostram incipientes. O presente trabalho teve como principais objetivos desenvolver e validar uma lista preliminar de descritores, bem como estimar a diversidade genética entre 60 genótipos da Coleção de germoplasma de *J. curcas* da Universidade Federal Rural do Rio de Janeiro (UFRRJ), através da aplicação de diversos métodos estatísticos multivariados, e também conhecer o potencial produtivo desta aos três anos de idade. A lista preliminar de descritores permitiu um melhor aproveitamento dos recursos genéticos da coleção, otimizando sua fenotipagem. A técnica de componentes principais foi eficiente em descartar 20 caracteres redundantes, de menor contribuição para a variabilidade observada, e em identificar os descritores relativos ao tempo de formação do fruto (TFF), número de sementes (NSE) e ao comprimento da folha (CFO) e do pedúnculo da inflorescência (CPI) como os mais importantes para a variação genética da coleção. A análise de trilha descreveu importantes correlações entre TFF e os caracteres referentes ao número de sementes por fruto (NSF) e projeção da copa no sentido Oeste (PCOO) frente ao número de sementes gerado. Esta ainda apontou um moderado efeito direto negativo, de -0,50, de TFF sobre NSE, retratando a possibilidade de, na coleção da UFRRJ, TFF ser utilizado como base para a seleção indireta visando uma maior produção de grãos em um menor tempo de colheita. As diferentes técnicas estatísticas multivariadas revelaram que a distância de Chebyshev e o algoritmo de Gower, conjugados ao método de agrupamento UPGMA, demonstraram, dentre outras metodologias, maior robustez na estimação da diversidade genética detectada. De modo geral, a coleção da UFRRJ apresentou uma considerável diversidade genética para os descritores avaliados. Os acessos 860, 345, 346 e 062 se apresentaram como os mais promissores a serem envolvidos em futuros cruzamentos com vistas à obtenção de híbridos de alta produção comercial e não-tóxicos. A avaliação do potencial produtivo retratou que a produção se concentrou ao longo de janeiro de 2012, com exceção do acesso 062, destacando um pico de produção para todos os genótipos. Devido ao caráter de agricultura familiar, a concentração da produção em um único mês, e também a possibilidade de redução de seu custo, fazem destes resultados um importante atrativo à adoção do pinhão-mansô pelo pequeno agricultor. O Programa de melhoramento de *J. curcas* da UFRRJ, a partir da clonagem dos acessos 860, 345 e 346 (mais produtivos), poderia, a curto prazo, maximizar a uniformidade e a precocidade de produção na região metropolitana do Rio de Janeiro.

Palavras-chave: Análise de trilha, análise multivariada, biodiesel, PCA, recursos genéticos.

ABSTRACT

REIS, Marcus Vinicius Magro. **Preliminary List of Descriptors, Morphoagronomic Characterization and Genetic Diversity in Germplasm of Physic-nut (*Jatropha curcas* L.)**. 2013. 109p. Dissertation (Master Science in Fitotecnia). Instituto de Agronomia, Departamento de Fitotecnia, Universidade Federal Rural do Rio de Janeiro, Seropédica, RJ, 2013.

The growing demand for energy sources with lower environmental impact has promoted a great interest about potential oil varieties, exploring alternatives to fossil fuels. In this scenario, the Euphorbiaceae *Jatropha curcas* L., commonly known as physic-nut, is distinguished by the excellent quality oil for biodiesel production. Although promising, the lack of genetically improved varieties, coupled with the uneven material now grown, still stand as major obstacles to the commercial crop of this species. In this context, genetic diversity among a group of parents is essential to make rational use of their genetic resources. However, information about the genetic diversity in germplasm of *J. curcas* is still incipient. The present work had as main objectives to develop and validate a preliminary list of descriptors as well as to estimate the genetic diversity among 60 *J. curcas* genotypes from Universidade Federal Rural do Rio de Janeiro (UFRRJ) germplasm collection through the application of various multivariate statistical methods, and also identify its yield potential at three years old. The preliminary descriptors list allowed a better use of the collection's genetic resources, optimizing their characterization. The principal components technique was efficient in discard 20 redundant characters, of minor contribution to the observed variability, and in identifying the descriptors related to the fruit ripening time (TFF), number of seeds (NSE) and leaf (CFO) and stalk of the inflorescence (CPI) length as the most important for the genetic variation of the collection. The breakdown of correlations in direct and indirect effects, by path-analysis, described important correlations between TFF and the characters related to the number of seeds per fruit (NSF) and the projection of the crown towards West (PCOo) over the number of seeds generated. This also showed a moderate negative direct effect of -0.50 from TFF on NSE, portraying the possibility, in UFRRJ's collection, of TFF be used as a basis for indirect selection towards a higher grain yield in a shorter harvesting. The different multivariate techniques revealed that the Chebyshev's distance and the Gower's algorithm, conjugated to the UPGMA method, showed, among other methodologies, greater robustness in the estimation of detected genetic diversity. In general UFRRJ's germplasm collection presented a reasonable genetic diversity for the evaluated descriptors. Accessions 860, 345 346 and 062 were presented as the most promising to be involved in future crosses in order to obtain non-toxic and high production commercial hybrid. The yield potential evaluation portrayed that production was concentrated through January 2012, with the exception of accession 062, highlighting a yield peak for all genotypes. Due to the nature of family farming, the concentration of production in a single month, and also the possibility of reducing costs, makes these results an important attraction to the adoption of physic-nut by the small farmer. The UFRRJ's *J. curcas* Improvement Program could, through the cloning of accessions 860, 345 and 346 (high yielders), in the short term, maximize the uniformity and the precocity production in the metropolitan region of Rio de Janeiro.

Keywords: Biodiesel, genetic resources, multivariate analysis, PCA, path-analysis.

ÍNDICE DE FIGURAS

- Figura 1.** Dendrogramas gerados através do método de agrupamento UPGMA, a partir do algoritmo de Gower, para as análises conjunta de variáveis quantitativas e multicategóricas (a) e individual de variáveis multicategóricas (b), e em função das distâncias Euclidiana (c) e de Chebyshev (d), para a análise individual de variáveis quantitativas. 62
- Figura 2.** Produção semanal de 60 genótipos da Coleção de germoplasma de *Jatropha curcas* da UFRRJ entre os meses de dezembro de 2011 e maio de 2012. 69
- Figura 3.** Temperatura média e precipitação acumulada na região de Seropédica, RJ, onde está localizada a Coleção de germoplasma de *Jatropha curcas* avaliada, entre os meses de setembro de 2011 e maio de 2012. Fonte: INMET/SIPA/UFRRJ. 70

ÍNDICE DE TABELAS

Tabela 1. Precipitação acumulada e valores médios mensais de umidade relativa, velocidade do vento e temperatura em Seropédica, RJ, no período de setembro de 2011 a maio de 2012.	26
Tabela 2. Procedência, número e identificação dos 60 genótipos de <i>Jatropha curcas</i> da Coleção de germoplasma da UFRRJ definidos para caracterização.	27
Tabela 3. Lista preliminar de descritores morfoagronômicos desenvolvidos para <i>Jatropha curcas</i> : tipo de variável, nome do descritor, metodologia e estado.....	28
Tabela 4. Classificação do grau de multicolinearidade (MONTGOMERY & PECK, 1981).	37
Tabela 5. Tipo de análise, distâncias, respectivos estimadores e agrupamentos utilizados para a estimação da diversidade genética a partir de 60 genótipos da Coleção de germoplasma de <i>Jatropha curcas</i> da UFRRJ.	39
Tabela 6. Classificação dos níveis de distorção e stress (KRUSKAL, 1964).	40
Tabela 7. Relação de variáveis e respectivos descritores elaborados para a caracterização, estudo da diversidade genética e potencial produtivo de 60 genótipos da Coleção de germoplasma de <i>Jatropha curcas</i> da UFRRJ.	42
Tabela 8. Estatística descritiva de 36 variáveis (19 vegetativas e 17 reprodutivas) mensuradas em 60 genótipos da Coleção de germoplasma de <i>Jatropha curcas</i> da UFRRJ.	46
Tabela 9. Resumo das estimativas da variância (autovalores) associadas aos componentes principais, contribuição relativa, acumulada e caracteres passíveis de descarte (redundantes), a partir de 36 variáveis quantitativas avaliadas em 60 genótipos da Coleção de germoplasma de <i>Jatropha curcas</i> da UFRRJ.	48
Tabela 10. Diagnóstico de multicolinearidade e principais correlações apontadas pelo teste de Montgomery e Peck (1981), a partir de 36 variáveis quantitativas descritas em 60 genótipos de <i>Jatropha curcas</i> da UFRRJ.	50
Tabela 11. Estimativas da variância (autovalores), coeficientes de ponderação associados (autovetores) e contribuição relativa e acumulada a partir de 16 variáveis quantitativas avaliadas em 60 genótipos da Coleção de germoplasma de <i>Jatropha curcas</i> da UFRRJ.	51
Tabela 12. Coeficientes de correlação fenotípica entre as 16 variáveis quantitativas de maior importância para a variabilidade genética encontrada em 60 genótipos da Coleção de germoplasma de <i>Jatropha curcas</i> da UFRRJ.	53
Tabela 13. Resumo das estimativas de efeitos diretos e indiretos de 15 variáveis quantitativas sobre a característica Número de sementes por planta (NSE) em 60 genótipos de <i>Jatropha curcas</i> da UFRRJ.	56
Tabela 14. Consistência do padrão de agrupamento proporcionado pelas técnicas do Vizinheiro mais próximo e UPGMA, em relação ao coeficiente geral de Gower, para a análise conjunta	

de todas as variáveis e individual de variáveis multicategóricas, e às distâncias Euclidiana, de Chebyshev, Manhattan e de Canberra, para variáveis quantitativas, em função do coeficiente de correlação cofenética, aglomerativo e níveis de distorção e stress, em 60 genótipos da Coleção de germoplasma de *Jatropha curcas* da UFRRJ. 58

Tabela 15. Correlação entre as matrizes de distância Euclidiana, de Chebyshev, Manhattan e de Canberra, estimadas a partir de 16 descritores quantitativos, em 60 genótipos de *Jatropha curcas* da UFRRJ..... 60

Tabela 16. Ordem decrescente de produção e respectivas procedências dos 60 genótipos de *Jatropha curcas* da Coleção de germoplasma da UFRRJ definidos para avaliação, entre dezembro de 2011 e maio de 2012. 66

Tabela 17. Coeficientes de correlação fenotípica entre oito descritores relativos a frutos e sementes, avaliados de dezembro de 2011 a maio de 2012, em 60 genótipos de *Jatropha curcas* da Coleção de germoplasma da UFRRJ. 68

Tabela 18. Comparação entre os valores médios da produção de grãos mensal no período de dezembro de 2011 a maio de 2012, em 60 genótipos da Coleção de germoplasma de *Jatropha curcas* da UFRRJ..... 72

SUMÁRIO

1. INTRODUÇÃO	1
2. REVISÃO DE LITERATURA.....	3
2.1. Aspectos Gerais de <i>Jatropha curcas</i> L.....	3
2.1.1. Descrição botânica e morfológica	3
2.1.2. Centro de origem e de diversidade	4
2.1.3. Aspectos agrônômicos de cultivo	5
2.2. Recursos Genéticos Vegetais	7
2.2.1. Conservação de recursos genéticos vegetais	7
2.2.2. Tipos de coleções.....	8
2.2.3. Formas de conservação.....	10
2.2.4. Caracterização e avaliação.....	11
2.2.4.1. Importância para a conservação e o melhoramento de plantas	11
2.2.4.2. Tipos de caracterização	13
2.2.5. Descritores mínimos	15
2.2.5.1. Definição e importância para a caracterização	15
2.2.5.2. Órgão regulamentador: <i>Bioversity International</i>	16
2.2.5.3. Regras básicas para elaboração de listas de descritores padronizadas	16
2.3. Diversidade Genética.....	18
2.3.1. Importância para o melhoramento genético vegetal.....	18
2.3.2. Estimação da diversidade genética	18
2.3.2.1. Tipos de variáveis utilizadas na estimação da divergência genética.....	19
2.3.2.2. Medidas de similaridade e dissimilaridade.....	19
2.3.2.3. Análises de agrupamento.....	21
2.3.2.4. Análises de componentes principais e variáveis canônicas.....	22
2.3.2.5. Análise de Trilha	23
2.4. Melhoramento Genético de <i>Jatropha curcas</i> L.....	23
3. MATERIAL E MÉTODOS	26
3.1. Localização do Experimento e Condições Edafoclimáticas.....	26
3.2. Germoplasma e Condições Experimentais	26
3.3. Desenvolvimento, Validação e Normatização de Descritores Morfoagronômicos.....	27
3.4. Análises Estatísticas	36
3.4.1. Importância e descarte de variáveis via análise de componentes principais	36
3.4.2. Correlação fenotípica e análise de trilha.....	38
3.4.3. Estimação da diversidade genética e consistência do padrão de agrupamento	38

3.4.4. Avaliação do potencial produtivo.....	40
4. RESULTADOS E DISCUSSÃO	42
4.1. Desenvolvimento, Validação e Normatização de Descritores Morfoagronômicos.....	42
4.2. Análises Estatísticas	45
4.2.1. Importância e descarte de variáveis via análise de componentes principais	45
4.2.2. Correlação fenotípica e análise de trilha.....	52
4.2.3. Estimção da diversidade genética e consistência do padrão de agrupamento	57
4.2.4. Avaliação do potencial produtivo.....	66
5. CONCLUSÕES.....	73
6. REFERÊNCIAS BIBLIOGRÁFICAS	74
7. ANEXOS	89

1. INTRODUÇÃO

A carência mundial por fontes de energia renováveis e menos poluidoras, capazes de garantir uma base energética mais segura e sustentável, se tornou um dos principais desafios deste século. Nesse contexto, sabe-se que o uso de óleos vegetais constitui uma das alternativas mais promissoras para a obtenção de biocombustíveis. Entre as diversas culturas oleaginosas estudadas para este fim, o pinhão-manso (*Jatropha curcas* L.), espécie monóica pertencente à família Euphorbiaceae, a mesma da mamona (*Ricinus* sp.) e da mandioca (*Manihot* sp.), destaca-se por possuir uma excelente qualidade de óleo para produção de biodiesel (HELLER, 1996; DRUMOND et al., 2010).

A espécie é caracterizada por ser perene, arbustiva, de rápido crescimento e possuir um alto teor de óleo nas sementes. Segundo Sunil et al. (2008), este teor pode variar de 22 a 42%. O pinhão-manso produz, em média, 700 a 1500 kg.ha⁻¹ de grãos em plantio de sequeiro e de 1500 a 2500 kg.ha⁻¹ sob condições irrigadas, a partir do quinto ano de cultivo (WANI et al., 2006). Tais características, aliadas ao fato de não ser utilizada na alimentação humana e apresentar manejo compatível ao perfil da agricultura familiar, fizeram o governo brasileiro incluir a espécie no Programa Nacional de Produção e Uso de Biodiesel (PNPB). Lançado em 2004, o PNPB subsidia o desenvolvimento de tecnologias de caráter sustentável, visando melhorar a qualidade e o rendimento de óleos vegetais, através de políticas de enfoque social e fortalecimento da agricultura familiar (ARRUDA et al., 2004; RODRIGUES, 2005).

Apesar de demonstrar grande potencial, a espécie *J. curcas* ainda se encontra em fase de domesticação para o cultivo comercial, processo que se iniciou somente nos últimos 30 anos (SATURNINO et al., 2005). De acordo com Durães et al. (2009), a oleaginosa vem sendo implantada em várias regiões do Brasil, porém, os genótipos utilizados são geneticamente desconhecidos e desuniformes, uma vez que ainda não há variedades melhoradas sobre as quais se tenham informações e garantias técnicas sobre a potencialidade de produção e adaptação para as diversas áreas produtoras. Nesse sentido, o programa de melhoramento genético de pinhão-manso deve buscar a obtenção de cultivares com características que atendam tanto as exigências dos produtores quanto as do mercado, como alta produtividade de grãos e óleo, rusticidade, floração sincronizada, ausência de curcina e ésteres de forbol (sem toxicidade), resistência a estresses bióticos e abióticos, além de ampla faixa de adaptação edafoclimática (LAVIOLA, 2010; JOHNSON et al., 2011).

O conhecimento da diversidade genética entre um grupo de genitores se faz essencial para o uso racional dos recursos genéticos de uma espécie (LOARCE et al., 1996). No entanto, informações acerca da diversidade genética existente no germoplasma de *J. curcas* ainda se mostram incipientes e pouco conclusivas. Segundo Rao et al. (2008), por ser uma espécie de grande adaptabilidade edafoclimática e ampla distribuição geográfica, acredita-se que haja uma considerável variabilidade genética a ser explorada em seu germoplasma. Nessa perspectiva, o levantamento de dados para descrição e diferenciação entre acessos, em coleções de germoplasma, se torna uma atividade essencial para a detecção de genes úteis para o melhoramento da cultura. Para tal, estudos sobre a variação genética podem ser realizados empregando-se descritores botânicos, morfológicos, agrônômicos e moleculares (MOREIRA et al., 1994).

Os maiores obstáculos encontrados para o uso eficiente de acessos conservados em coleções e bancos de germoplasma devem-se a uma série de dificuldades e deficiências relativas ao seu gerenciamento, como a falta de documentação e a caracterização/avaliação inadequada do material vegetal existente. Tais fatos limitam a ação do melhorista, visto que a variabilidade genética só pode ser utilizada de maneira adequada quando devidamente avaliada e quantificada (CHIORATO, 2004). Sob este aspecto, a inexistência de uma lista oficial de descritores mínimos para a espécie *J. curcas*, sem dúvida, acarreta maior

dificuldade na avaliação de acessos desta em coleções e bancos de germoplasma. Desse modo, tanto a formulação de uma lista de descritores normatizada quanto o emprego correto de técnicas estatísticas multivariadas, para uma estimação mais fiel da diversidade genética, se tornam imprescindíveis no intuito de organizar e otimizar processos que envolvam a conservação e o uso de recursos genéticos de *J. curcas* em coleções de germoplasma.

Perante o exposto, o presente trabalho teve como objetivos gerais desenvolver uma lista preliminar de descritores; validar esta lista; e estimar a diversidade genética existente entre 60 acessos de *J. curcas* da Coleção de germoplasma da Universidade Federal Rural do Rio de Janeiro (UFRRJ). Para tal, fundamentou-se, como objetivos específicos, realizar a caracterização morfoagronômica dos 60 acessos; propor, com base em análises de robustez estatística de técnicas multivariadas, as metodologias que mais se adequam à estimação da diversidade genética da coleção; bem como conhecer o potencial produtivo desta aos três anos de idade, para, assim, possibilitar a seleção de genótipos divergentes e produtivos a serem utilizados no Programa de melhoramento genético de *J. curcas* da UFRRJ.

2. REVISÃO DE LITERATURA

2.1. Aspectos Gerais de *Jatropha curcas* L.

2.1.1. Descrição botânica e morfológica

O pinhão-manso é um arbusto caducifólio pertencente à família Euphorbiaceae, do gênero *Jatropha* e espécie *Jatropha curcas* (BÁRTOLI, 2008). O gênero *Jatropha* possui cerca de 175 espécies registradas, entre as quais *Jatropha curcas*, considerada sua forma mais primitiva (DEHGAN & WEBSTER, 1979). A espécie é caracterizada por ser de rápido crescimento, cuja altura normal é de dois a três metros, podendo alcançar até cinco metros em condições especiais. Seu caule é liso, com aproximadamente 20 cm de diâmetro, de lenho mole e medula desenvolvida, porém, pouco resistente. A cultura possui floema com longos canais, que se estendem até as raízes, nos quais circula o látex, líquido leitoso e cáustico que corre em abundância quando a planta é ferida (CORTESÃO, 1956). A espécie segue a arquitetura clássica das Euforbiáceas, onde a primeira inflorescência é apical, originando a emissão de novos ramos que passam a ser axilares até o surgimento de novas inflorescências que, por sua vez, impedem novamente o crescimento apical, surgindo os próximos ramos e, assim, sucessivamente (ALBUQUERQUE et al., 2008).

De acordo com a descrição de Bártoli (2008), as raízes da cultura são curtas e pouco ramificadas. Cada plântula possui, normalmente, cinco raízes, sendo uma central e quatro periféricas (duas secundárias e duas terciárias). Seu tronco, ou fuste, é dividido desde a base, em longos ramos, com inúmeras cicatrizes produzidas pela queda das folhas na estação seca, as quais ressurgem logo após as primeiras chuvas (CORTESÃO, 1956).

Quando jovens, as folhas do pinhão-manso possuem coloração vermelha e roxa, e, à medida que se expandem, tornam-se verdes, largas, esparsas e brilhantes (ALVES et al., 2008). São alternas, pecioladas, com nervuras esbranquiçadas e salientes em sua face inferior, e possuem forma de palma, com três a cinco lóbulos (CORTESÃO, 1956). As folhas são decíduas e contam com limbo menor que o pecíolo, este de até 18 cm de comprimento (HELLER, 1996). Entretanto, Bártoli (2008) descreve folhas com limbo de 7 a 16 cm de comprimento e largura, maiores que o pecíolo, este, de até 15 cm de comprimento.

Segundo Alves et al. (2008), sua floração é monóica, possuindo flores masculinas em maior número nas extremidades das ramificações e femininas ao longo das ramificações. A inflorescência é do tipo panícula em cimeira definida, isto é, possui eixo principal com eixos secundários, formando pequenas flores pediceladas (6 a 8 mm) de coloração amarelo-esverdeada (HELLER, 1996; BÁRTOLI, 2008).

Seu fruto é indeiscente, ou seja, não liberam a semente até se tornarem maduros, e possuem forma capsular e ovóide, com diâmetro de 1,5 a 3,0 cm, na maioria das vezes trilocular, com uma semente em cada cavidade, e formada por um pericarpo ou casca dura e lenhosa. Apresenta coloração inicialmente verde, passando para amarelo, castanho, e por fim, preto, quando atinge seu estado de maturação (CORTESÃO, 1956; BÁRTOLI, 2008). De acordo com Cortesão (1956), são constituídos por 53 a 62% de sementes, e 38 a 47% de casca. Todavia, Drumond et al. (2008) declaram que o fruto do pinhão é formado por 64 a 75% de sementes, e 15 a 36% de casca.

As sementes são relativamente grandes e, quando secas, medem de 1,2 a 2,0 cm de comprimento, e de 1,0 a 1,3 cm de largura. Possuem albúmen abundante, branco e oleaginoso, contendo embrião provido de dois largos cotilédones achatados. Ainda apresentam tegumento rígido, quebradiço e de textura resinosa (HELLER, 1996). Para Drumond et al. (2008), a semente do pinhão, que pesa de 0,50 a 0,70 g, pode ter, dependendo das condições ecológicas e tratos culturais, 34 a 45% de casca, e 55 a 66% de amêndoa. Conforme Peixoto (1973), a semente da oleaginosa possui 37,5% de óleo. Entretanto, Sunil et al. (2008), estudando

genótipos indianos de pinhão-manso, alegam que o teor de óleo em suas sementes pode variar entre 22 e 42%. Em contraponto, ao avaliarem acessos do Banco de Germoplasma de *J. curcas* do Instituto Agronômico de Campinas (IAC), Ferrari et al. (2009) identificaram genótipos com teor de óleo variando entre 11 e 39%. De acordo com Akbar et al. (2009), o óleo de *J. curcas* pode ser classificado como oléico-linoléico.

2.1.2. Centro de origem e de diversidade

Na concepção do geneticista russo Nicolay Ivanovich Vavilov (1887-1943), o centro de origem de uma cultura equivalia à região onde o ancestral silvestre de uma determinada espécie exibia a maior diversidade genética para um número seletivo de características, diminuindo sua variabilidade à medida que se deslocava para a periferia da distribuição. No entanto, o conhecimento moderno raramente valida tal proposição, pois nem sempre o centro de origem da espécie estudada coincide com a região em que esta demonstra maior diversidade genética.

Segundo Harlan (1975), algumas culturas possuem centros de origem, outras não, bem como podem ter sido domesticadas somente uma vez, ou diversas vezes ao longo dos anos, por diferentes povos, em épocas e áreas distintas. Apoiado nestas afirmações, Harlan criou a teoria dos *Centres and Non-centres*, na qual o fator “espaço” e a atividade humana eram os principais responsáveis pelo padrão de distribuição geográfico das espécies cultivadas. Para este autor, cada espécie deve ser estudada separadamente, visto que sua dispersão geográfica é altamente dependente das diferentes formas de uso que as diversas civilizações faziam destas, suas preferências e manejo cultural. Hawkes (1983) também questionou a teoria de Vavilov, afirmando que a agricultura não teria se originado uma única vez, mas, possivelmente, de forma simultânea em várias regiões do globo terrestre. Para tal, Hawkes propôs a existência dos “Centros nucleares e Regiões de grande diversidade”, onde os *centros nucleares* teriam sido as áreas onde, de fato, a agricultura teria se originado, com evidências arqueológicas da presença dos ancestrais silvestres das espécies, enquanto que as *regiões de grande diversidade* corresponderiam às áreas para onde a agricultura teria se dispersado, junto com o processo de migração do homem, sem vestígios arqueológicos da presença do ancestral silvestre da espécie. Sendo assim, a definição original de Vavilov para *centro de origem* foi complementada com a introdução do novo termo *centro de diversidade*. Hoje, tem-se por centro de origem a região onde o ancestral silvestre de uma cultura se distribuiu em seu estado nativo, e entende-se o termo centro de diversidade como zona geográfica onde se encontra um número elevado de espécies de um gênero, ou de gêneros de uma família, contrastando com sua menor frequência em outras áreas (VALOIS et al., 1996).

A origem da espécie *Jatropha curcas* ainda é confusa e indefinida, e embora seja amplamente encontrada ao redor das regiões intertropicais, ocorrendo em maior escala nas zonas tropicais e em número reduzido em localidades temperadas, a cultura segue despertando discussões e controvérsias em relação à sua origem. Contudo, relatos citam que a espécie pode ter sido difundida mundialmente pelos portugueses, e apontam as Américas Central e do Sul como prováveis centros de origem (HELLER, 1996; ARRUDA et al., 2004). Para Bártoli (2008), o pinhão-manso é oriundo do México, e tem a América do Sul, Índia, Ásia e África como centros de diversidade. Já Openshaw (2000) afirma que a cultura está distribuída, principalmente, nas Américas, África e parte da Ásia, sendo largamente cultivada no México, Nicarágua, Tailândia e partes da Índia. Conforme Dias et al. (2012), a ocorrência de genótipos não-tóxicos e o conhecimento da existência dessa espécie pela civilização mexicana Olmeca, que viveu há 3500-5000 anos atrás, faz do México seu centro de origem e de domesticação. No Brasil, sua distribuição geográfica é bastante vasta, podendo ser encontrado do Nordeste até os Estados de São Paulo e Paraná, sobretudo ao longo de Goiás e Norte de Minas Gerais

(CORTESÃO, 1956; PEIXOTO, 1973; ARRUDA et al., 2004). De acordo com Dias et al. (2012), a presença de genótipos com alto teor de óleo no Norte do Estado de Minas Gerais caracteriza esta região como um dos seus centros secundários de diversidade.

2.1.3. Aspectos agronômicos de cultivo

De acordo com Drumond et al. (2010), a produtividade do pinhão-mansinho pode ser muito variável, dependendo da região, do método de cultivo e dos tratamentos culturais, assim como da regularidade pluviométrica e da fertilidade do solo. Também não se pode deixar de mencionar que uma das causas da produção inconstante entre plantas diz respeito à variabilidade genética existente na espécie.

Em relação às necessidades hídricas, Foidl et al. (1996) descreveram que a cultura se desenvolve bem com médias anuais de precipitação variando de 250 a 3000 mm, podendo ser encontrada desde o nível do mar até 1800 m de altitude. Para Heller (1996), seu ótimo desenvolvimento ocorre em regiões com temperaturas médias anuais entre 20 e 28 °C, e baixas altitudes (0-500 m). Silva et al. (2009c) apontam que o estresse hídrico pode afetar o desenvolvimento do pinhão-mansinho, comprometendo sua altura e diâmetro caulinar.

Embora seja caracterizada como uma planta pouco exigente em relação ao clima e adaptada a áreas degradadas de baixa fertilidade, o pinhão-mansinho deve ser, preferencialmente, cultivado em solos profundos, bem estruturados, não compactados, drenados e pouco argilosos (ARRUDA et al., 2004). Para Foidl et al. (1996), a espécie se desenvolve bem em quase todos os solos, com exceção do Vertissolo, por preferir textura mais arenosa. Segundo Vale et al. (2006a), o aumento da densidade do solo pode provocar diminuição da altura da planta, diâmetro caulinar, número de folhas e peso seco da parte aérea e das raízes. Estas, por sua vez, não conseguem se desenvolver de forma satisfatória em solos ácidos, com pH abaixo de 4,5. Nessa situação se faz necessária a realização de calagem com base na análise química do solo, a qual indicará a quantidade de calcário, gesso, macro e micronutrientes necessários para satisfazer as exigências da cultura (PEIXOTO, 1973).

A salinização do solo progride quando a quantidade de sais nele acumulada, pela água de irrigação, é maior que a quantidade removida pela água de drenagem (PIZARRO, 1985). Nesse cenário, Andréo-Souza et al. (2010) constataram que sua semente sofre atraso no processo germinativo quando submetida à condição de estresse salino durante a fase de embebição. Estudos de Nery et al. (2009) concluíram que a altura da planta, o diâmetro caulinar, o número de folhas e a área foliar do pinhão são afetados, linearmente, com decréscimos de 3,78, 7,35, 9,75 e 17,74%, respectivamente, por aumento unitário da condutividade elétrica da água, em ambiente protegido. Ainda nessas condições, Vale et al. (2006b) comprovaram que a cultura é sensível à presença de sal na água de irrigação em sua fase de desenvolvimento inicial. Entretanto, Silva et al. (2009b) observaram que o pinhão-mansinho é capaz de se ajustar osmoticamente em presença de salinidade através da redução intensa no potencial osmótico e aumento do estado hídrico das folhas em concentrações elevadas de NaCl. Segundo estes autores, o bom nível de hidratação se dá, principalmente, por meio do acúmulo dos próprios íons salinos Na^+ e Cl^- , que contribuem para a retenção de água no tecido. Nessa ocasião, os íons inorgânicos são sequestrados no vacúolo, enquanto os solutos orgânicos são compartimentalizados no citoplasma, balanceando o baixo potencial osmótico nos vacúolos (TAIZ & ZEIGHER, 2004).

Quando plantado no princípio da estação chuvosa, o pinhão-mansinho pode iniciar sua produção de frutos já no primeiro ano de cultivo, embora atinja seu ápice produtivo somente a partir do quarto ano, depois de estabelecido. De acordo com Juhász et al. (2009), seu florescimento inicia-se após um período de repouso vegetativo, normalmente após o inverno, no Brasil. Todavia, em regiões permanentemente úmidas, Heller (1996) registrou que sua

floração pode ser observada ao longo de todo ano. Dependendo do espaçamento, sua produtividade pode passar dos 6000 kg.ha⁻¹ de sementes, chegando a produzir mais de 2000 kg.ha⁻¹ de óleo (LAVIOLA & DIAS, 2008). No entanto, para se obter alta produção de frutos, a planta exige solos férteis e com boas condições físicas. Ainda segundo Laviola e Dias (2008), a colheita dos frutos de pinhão extrai elevada quantidade de nutrientes do solo, podendo causar seu empobrecimento ao longo dos anos de cultivo, demandando reposição adequada por adubação. Logo, a correção da acidez e fertilidade é decisiva para se obter sucesso e lucro em seu plantio. De acordo com os estudos de Silva et al. (2009a), sobre deficiência nutricional no pinhão-mansão, as omissões de macro e micronutrientes podem provocar sintomas visuais comuns a outras espécies. Openshaw (2000) relatou que a cultura necessita do uso de fertilizantes, e que em situações de altos níveis de nitrogênio, ocorre aumento de produtividade. Já Arruda et al. (2004), indicaram solos com altos teores de matéria orgânica, sendo recomendado o uso de adubação verde com leguminosas, que irão fixar o nitrogênio atmosférico e transferir aos solos, por decomposição orgânica, os nutrientes essenciais à planta, como fósforo, cálcio e enxofre.

No que diz respeito ao plantio e propagação da espécie, a escolha do melhor método deverá ser em função da finalidade à qual será utilizada. Geralmente, para cercas-vivas, utilizam-se estacas, devido ao seu crescimento mais rápido, porém, sua longevidade e resistência à seca serão bem menores. Em cultivos comerciais, com fins para produção de óleo vegetal, recomenda-se o uso de sementes, com o intuito de gerar plantas mais vigorosas e de maior longevidade (HELLER, 1996; ARRUDA et al., 2004). Para Heller (1996), o bom desenvolvimento da planta é assegurado com o plantio no início do período chuvoso. Ainda segundo este autor, o espaçamento ideal para plantio convencional é de 2 x 2 m (2500 plantas.ha⁻¹), e para o consórcio com culturas anuais, durante os primeiros anos, é de 2,5 x 2,5 m (1600 plantas.ha⁻¹) ou 3 x 3m (1111 plantas.ha⁻¹).

Os tratos culturais empregados no cultivo do pinhão-mansão são semelhantes aos realizados na cultura da mamona (DRUMOND et al., 2008). De acordo com Arruda et al. (2004), a cultura é atacada por algumas pragas e doenças, e práticas, como o controle de espontâneas, devem ser realizadas sempre que necessárias, principalmente em estádios iniciais de crescimento, devido à alta competição. Entre as pragas mais observadas no cultivo do pinhão, estão os pulgões, cochonilhas, formigas saúvas e o ácaro-branco. Já entre as doenças, se destacam os fungos causadores da ferrugem (*Phakopsora jatrophiicola*), da antracnose (*Colletotrichum gloeosporioides*) e da gomose (*Phytophthora* sp.), além da bactéria *Xanthomonas axonopodis* pv. *malvacearum*, causadora da mancha angular do pinhão (HELLER, 1996; BÁRTOLI, 2008; DRUMOND et al., 2008). A podridão do colo das plantas, também conhecida como *seca descendente*, causada pelo fungo *Lasiodiplodia theobromae*, é uma das doenças mais importantes no cultivo da espécie. Segundo Freire e Parente (2006), a doença se caracteriza pela seca nas extremidades superiores dos ramos, onde progride até chegar ao caule e todo órgão ser afetado, podendo levar a planta à morte. Conforme estes autores, a doença também atinge a espécie *J. podagrica*, relativa ao gênero *Jatropha*, em que a podridão na base do caule se estende até às raízes, fazendo com que a planta perca sustentação e seja facilmente retirada do solo.

Um importante aspecto observado na espécie *Jatropha curcas* se refere à toxicidade presente na torta, subproduto gerado a partir do processo de extração do óleo de suas sementes. De acordo com Alves et al. (2008), esta é tóxica devido à presença de alcalóides conhecidos como ésteres de forbol (ésteres diterpenos), responsáveis por atividades carcinogênicas e ação inflamatória em animais e ao homem. Por muito tempo, tinha-se que a atividade tóxica da espécie era causada pela ação da lectina curcina, substância abundante em suas sementes (FERRARI et al., 2009). Contudo, Evans (1986) concluiu que o caráter tóxico do pinhão-mansão é dado, principalmente, pelos ésteres de forbol, amplamente encontrados em

espécies vegetais das famílias Euphorbiaceae e Thymelaeaceae. Makkar et al. (1998), avaliando genótipos mexicanos considerados tóxicos e de baixa toxicidade de *J. curcas*, verificaram que os níveis de outros componentes, também julgados tóxicos, presentes na espécie, como as lectinas, apresentavam valores similares entre as duas variedades, todavia, o teor de ésteres de forbol diferia drasticamente entre elas. Corroborando tal resultado, Ferrari et al. (2009) observaram níveis de ésteres de forbol variando entre 1,41 e 8,97 mg.g⁻¹, em acessos do Banco de Germoplasma de *J. curcas* do IAC.

2.2. Recursos Genéticos Vegetais

2.2.1. Conservação de recursos genéticos vegetais

Segundo Hidalgo (1991), os recursos genéticos vegetais correspondem à soma de todas as combinações de genes resultantes da evolução de uma espécie, englobando desde espécies silvestres com potencial agrícola até genes clonados. Para o *International Board of Plant Genetic Resources* (IBPGR, 1991), o emprego do termo *recursos genéticos* implica que o material possui, ou possa vir a apresentar, valor econômico e potencial de utilização atual ou futura, sendo de vital importância aquele que contribui para a segurança alimentar mundial. Dessa maneira, pode-se inferir que os recursos fitogenéticos constituem o principal pilar de subsistência da humanidade, uma vez que auxiliam na resolução de problemas crônicos da sociedade, como a fome e a pobreza, e suprem as necessidades básicas do homem (JARAMILLO & BAENA, 2000). No entanto, sua perda, devido ao uso irracional, é considerada um processo irreversível que interfere diretamente na segurança alimentar do planeta. De acordo com Martins (2001), um dos problemas mais graves enfrentados pela agricultura moderna é a vulnerabilidade dos agroecossistemas às adversidades ambientais e desastres naturais causada pela erosão genética.

Dado o grande valor dos recursos fitogenéticos para o bem estar das gerações contemporâneas e futuras, sua conservação compõe uma das estratégias mais eficientes para posterior exploração sustentável da variabilidade genética vegetal. Na acepção da Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA), conservação, no sentido amplo, corresponde ao conjunto de atividades e políticas que asseguram a contínua disponibilidade e existência de um recurso, enquanto que, em sentido mais restrito, é o armazenamento e a guarda do germoplasma em condições ideais, permitindo a manutenção de sua integridade (VALOIS et al., 1996). Definida pelo IBPGR (1991), conservação de recursos genéticos vegetais refere-se à manutenção das populações em seu habitat natural (*in situ*) ou de espécimes/acessos/entradas destas populações em bancos de germoplasma, fora de seu habitat natural (*ex situ*). A conservação ainda pode ser dada por meio da combinação dos métodos *in situ* e *ex situ*, de maneira complementar, onde a escolha de um ou vários métodos depende das necessidades, possibilidades e da espécie em causa. Para alcançar sucesso, a conservação requer planejamento, continuidade, conhecimento sobre as espécies alvo e apoio institucional contínuo, para gerir de maneira sustentável os recursos econômicos, humanos e técnicos necessários para manter as coleções (JARAMILLO & BAENA, 2000).

Tendo como principal objetivo a conservação de genes ou de genótipos de plantas fora de suas comunidades naturais de ocorrência, a conservação *ex situ* faz parte de um importante conjunto de atividades que compõem a gestão dos recursos fitogenéticos, sendo considerada complementar à conservação *in situ*, pois não é possível conservar *ex situ* todas as espécies (PARODA & ARORA, 1991; JARAMILLO & BAENA, 2000). Por ser aplicada tanto às espécies silvestres quanto às já domesticadas, a conservação *ex situ* atinge um amplo espectro taxonômico, conservando a variabilidade genética gerada ao longo do processo evolutivo de domesticação (HIDALGO, 1991). De forma geral, o modelo *ex situ* pode abrigar uma gama de materiais de importância agrícola, cada qual com uma possibilidade distinta de uso, mas

tendo como principal meta a conservação da variabilidade genética. Espécies silvestres, por exemplo, são utilizadas como fontes de genes de interesse para o melhoramento, como resistência a pragas e doenças (HARLAN, 1975). Já variedades tradicionais, cultivares modernas e produtos da biotecnologia, como plantas transgênicas, fragmentos de DNA e genes clonados, são também conservados *ex situ* como uma importante estratégia para garantia da segurança alimentar, através de sua exploração racional (JARAMILLO & BAENA, 2000). Segundo Paroda e Arora (1991), a conservação *ex situ* inclui os jardins botânicos, herbários, coleções *in vitro*, em câmaras frias, nitrogênio líquido (crioconservação) e bancos de germoplasma a campo e de sementes, sendo este último o mais comum. Entretanto, estes autores alertam para o risco potencial que bancos de germoplasma a campo possuem de perder seus acessos, dada sua maior propensão ao ataque de pragas e doenças, desastres naturais e outros estresses bióticos e abióticos. Outra desvantagem de sua utilização, observada por estes autores, reside na grande demanda por espaço e mão de obra necessária para manutenção de uma pequena proporção de diversidade genética. Atento a esta desvantagem, Jaramillo e Baena (2000) recomendam a conservação em bancos de germoplasma a campo somente às espécies que se reproduzem, exclusivamente, de forma vegetativa e àquelas que possuem sementes recalcitrantes.

Em contraponto ao modelo *ex situ*, a conservação *in situ* refere-se à ação de conservar o vegetal em seu habitat natural (VALOIS et al., 1996). De acordo com Paroda e Arora (1991), o modelo *in situ* apresenta vantagem sobre a conservação *ex situ*, pois a influência direta das forças da seleção natural sobre o material, em seu próprio local de ocorrência, permite uma evolução contínua, aumentando sua adaptação ao ambiente. Para Hawkes (1982), a conservação *in situ* é extremamente importante para parentais silvestres de culturas comerciais, espécies perenes e florestais, uma vez que o modelo *ex situ* se mostra limitado e inconveniente para a conservação eficiente destas. Uma definição mais abrangente de conservação *in situ* é dada pelo IBPGR (1991), que relaciona o modelo com a conservação do ecossistema e do habitat natural e, no caso das espécies domesticadas, com o local onde estas desenvolveram suas características adaptativas, envolvendo os sistemas tradicionais agrícolas. Portanto, dentro da definição de conservação *in situ*, podem-se citar dois conceitos: conservação genética em reservas e conservação *on farm*. O primeiro inclui o manejo e o monitoramento dos recursos genéticos de populações silvestres dentro de áreas definidas para conservação ativa a longo prazo, como em parques nacionais, reservas naturais, estações ecológicas e santuários da vida silvestre. Já a conservação *on farm* corresponde ao cultivo e manejo contínuo de populações de plantas realizado por produtores de comunidades locais e povos indígenas (VALOIS et al., 1996; VETELÄINEN et al., 2009).

Maxted et al. (1997) definiram a conservação *on farm* como: “*manejo sustentável da diversidade genética de variedades locais (land races), associado às espécies silvestres e espontâneas, realizado por agricultores, em sistemas agroflorestais ou de agricultura tradicional*”. Dessa forma, pode-se concluir que o conceito de conservação *on farm* baseia-se não somente na conservação do germoplasma existente, mas também nas condições que permitem o desenvolvimento de novos germoplasmas (BITOCCHI et al., 2009). A variabilidade genética mantida nesse tipo de conservação, onde há uma grande diversidade inter e intra-específica, denota a importância dessas comunidades em manter e amplificar a variação genética (PERONI & HANAZAKI, 2002).

2.2.2. Tipos de coleções

De acordo com Goedert (2007), a palavra *germoplasma* foi criada pelo cientista alemão August Weisman, em 1883, pela combinação do prefixo “germ”, que significa “do qual algo nasce”, e “plasma”, que exprime “material que dele se forma”. Segundo esta autora,

o termo *germoplasma* é aplicado ao elemento dos recursos genéticos que maneja a variabilidade dentro de cada espécie (intra-específica), com fins de utilização para a pesquisa em melhoramento genético e biotecnologia. Allard (1960) define germoplasma como a base física do cabedal genético que reúne o conjunto de materiais hereditários de uma espécie. No entanto, Giacometti e Ferreira (1988) elucidam que genes e germoplasma são materiais praticamente idênticos, porém, o germoplasma governa o processo da hereditariedade, enquanto os genes constituem os elementos desse processo.

Para Chang (1987), as principais atividades da agricultura moderna estão relacionadas ao uso racional dos germoplasmas vegetais, uma vez que estes estabelecem a base central de variabilidade genética que subsidia a pesquisa, o melhoramento genético vegetal e a produção agrícola. Tendo em vista a importância da manutenção desta diversidade genética para o bem estar das futuras gerações, amostras representativas da variabilidade genética de espécies alvo de conservação são agrupadas em coleções. Segundo Jaramillo e Baena (2000), *coleções de germoplasma* são conjuntos de espécimes/acessos/entradas representativos de uma variabilidade genética objeto de conservação, e/ou utilização, podendo conter de dezenas até milhares de amostras, em ambientes e condições apropriadas. Nesse aspecto, estes autores notificam que a maneira mais prática de salvaguardar o material genético, sob condições adequadas, é em *bancos de germoplasma*. Nestes são realizados trabalhos de aquisição, multiplicação, documentação, regeneração, caracterização, avaliação, distribuição e intercâmbio de germoplasma visando sua conservação. Desse modo, as coleções de germoplasma são classificadas quanto ao material que as compõem e quanto ao tempo e condições que as conservam, em *coleção base*, *coleção ativa*, *coleção nuclear* e *coleção de trabalho*.

A coleção base agrupa a maior variabilidade genética possível das espécies alvo, incluindo parentes silvestres, variedades tradicionais, cultivares e germoplasma de elite, em amostras de sementes ortodoxas ou material vegetativo (VILELA-MORALES & VALOIS, 1996). De acordo com Towil e Roos (1989), é estabelecida para conservação de germoplasma a longo prazo e como precaução contra possíveis perdas, e não para distribuição ou intercâmbio. Para Valois et al. (1996), pela variabilidade que contém e pela função que cumpre, a coleção base é vista como uma estratégia de segurança alimentar para o país.

Oriunda da coleção base, a coleção ativa promove a conservação, a curto e médio prazo, de espécimes na forma de sementes, material vegetativo em campo ou *in vitro*, para gestão e distribuição (JARAMILLO & BAENA, 2000). Dessa forma, é rotineiramente usada para propósitos de pesquisa, caracterização, avaliação e utilização de materiais, que são multiplicados de acordo com a demanda pelo germoplasma, por parte de pesquisadores, e regenerados periodicamente. Seu caráter dinâmico é dado pelo fato que acessos entram e saem de seu inventário conforme decisões gerenciais (VALOIS et al., 1996).

Conhecida também como *core collection*, a coleção nuclear reúne a maior variabilidade genética de uma espécie no menor número possível de amostras, na forma de sementes ou material vegetativo, ou seja, nas mesmas condições de uma coleção ativa (VAN HINTUM et al., 2000). O conceito de coleção nuclear é aplicado em coleções de germoplasma com 10 a 15% do tamanho da coleção original, representando 70 a 80% da variabilidade genética disponível na espécie de interesse e nos parentes silvestres. O resto da coleção permanece reservado, como fonte de genes para futuras utilizações (VALOIS et al., 1996). Segundo Frankel et al. (1995) e Hodgkin et al. (1995), a coleção nuclear permite detectar acessos duplicados na coleção base e definir prioridades para caracterizar e avaliar as amostras. Cordeiro e Abadie (2007) ressaltam que a coleção nuclear estabelece uma base de informação mais completa sobre um conjunto de acessos, levando efetivamente a uma ampliação do uso do germoplasma.

Já a coleção de trabalho, considerada “coleção do melhorista”, é classificada como um modelo de conservação a curto prazo, na forma de sementes ou plantas em campo, com o objetivo de fornecer material para o melhorista ou para instituições de pesquisa que fazem melhoramento. Geralmente, de tamanho limitado e composta por germoplasma de elite, a coleção de trabalho é mantida para propósitos específicos do melhorista (VALOIS et al., 1996; JARAMILLO & BAENA, 2000).

2.2.3. Formas de conservação

A conservação dos recursos fitogenéticos visa assegurar a existência destes, portando suas características originais, em condições adequadas. Sendo assim, não se limita à obtenção e posse do material, este deve estar, também, geneticamente íntegro e disponível para exploração. Tal segurança pode ser obtida a partir do conhecimento sobre a reação da espécie objeto à conservação e sua forma de reprodução. Estas características irão determinar em quais condições ela permanecerá viável. Nesse sentido, as alternativas mais eficientes de conservação de germoplasma se dão sob a forma de sementes e material vegetativo, em campo ou *in vitro* (JARAMILLO & BAENA, 2000).

A manutenção da diversidade genética através do armazenamento de sementes é a forma de conservação *ex situ* mais utilizada e conveniente (ENGELS & VISSER, 2003). Para total garantia, sua conservação deve ser realizada via controle dos níveis de temperatura e umidade durante o armazenamento, para que se possa inibir ou reduzir o metabolismo da amostra (RAO & JACKSON, 1996). Entretanto, somente sementes ortodoxas garantem sua viabilidade quando armazenadas dessa forma, uma vez que sementes recalcitrantes perdem rapidamente sua viabilidade quando expostas a ambientes de baixa umidade relativa, necessitando de metodologias específicas para armazenamento, mesmo para curto período de tempo (SAAD & RAO, 2001).

Conforme Engels e Visser (2003), espécies que possuem sementes recalcitrantes devem ser conservadas de forma vegetativa, em bancos de germoplasma a campo ou *in vitro*. Além destas, fazem parte do grupo de espécies recomendadas para conservação em bancos de germoplasma a campo, as florestais, perenes, silvestres, e as que raramente produzem sementes. Apesar de se mostrar uma forma de conservação importante para estas espécies, Engelmann e Engels (2002) alertam para os possíveis transtornos da conservação a campo, como o alto custo de mão de obra para manejo e manutenção, a ocorrência de pragas e doenças, e a vulnerabilidade frente estresses ambientais, como déficit hídrico e alagamentos.

Quando uma forma de conservação de germoplasma se mostra suscetível à perda, como ocorre em bancos a campo, métodos alternativos e/ou complementares devem ser sempre considerados e explorados (ENGELS & VISSER, 2003). Nesse contexto, a cultura de tecidos permite conservar *in vitro* uma ampla variedade de espécies em diversos tipos de amostras, como plantas completas, rebentos, ápices caulinares, meristemas, embriões, protoplastos, pólen e DNA (JARAMILLO & BAENA, 2000). De acordo com Engels e Visser (2003), a conservação *in vitro* baseia-se em controlar o crescimento normal de explantes viáveis através de modificações em seu meio de cultura e condições de armazenamento. Para tal, Scowcroft (1984) enuncia duas formas básicas para manutenção do germoplasma *in vitro*: via “crescimento lento” ou por crioconservação. Condições para promoção de “crescimento lento” são geradas a partir de alterações no meio de cultura, adicionando inibidores de crescimento e diminuindo a oferta de nutrientes necessários para o desenvolvimento do explante, por exemplo, e/ou modificando as condições em que se mantém, por meio de redução de temperatura e iluminação (ENGELMANN, 1997). Porém, a falta de protocolos adequados para a grande maioria das espécies, adicionada à necessidade de repicagem periódica do material, pode aumentar seu risco de contaminação, tornando esse método pouco

prático e laborioso, se prestando mais a pequenas coleções (PARODA & ARORA, 1991). Visando contornar possíveis desvantagens do cultivo *in vitro* via “crescimento lento”, a crioconservação se apresenta como uma técnica ainda recente, mas de boas perspectivas, pois permite manter qualquer espécie que tolere e sobreviva ao congelamento por períodos indefinidos (JARAMILLO & BAENA, 2000). Esta prática possibilita a manutenção da estabilidade e viabilidade, genética e fisiológica, do material através de armazenamento sob temperaturas ultra-reduzidas em nitrogênio líquido (-196 °C). Segundo Engelmann (1997), processos metabólicos e divisões celulares são “virtualmente” paralisados nessas condições, permitindo o armazenamento sem qualquer tipo de alteração sobre o vegetal.

Paroda e Arora (1991) afirmam que a crioconservação pode ser aplicada com sucesso em amostras de protoplastos, pólen, DNA e embriões somáticos e zigóticos. Dentre os diferentes tipos de materiais já testados, a conservação de pólen e fragmentos de DNA tem ganhado grande valor. Roberts (1975) afirma que o armazenamento de pólen se faz essencial para espécies que o apresentam em baixa disponibilidade e curto período de viabilidade. Por ser facilmente coletado e criopreservado em grande quantidade, utilizando relativamente uma pequena porção de espaço para isto, Alexander e Ganeshan (1993) reiteram sua importância para a polinização entre plantas de floração assíncronas. No que tange à conservação de DNA, seu armazenamento é, a princípio, simples e barato de ser realizado (SAAD & RAO, 2001). O progresso científico alcançado pela engenharia genética ao longo dos últimos anos já permite a conservação de DNA extraído de cloroplastos, mitocôndrias e núcleos celulares, tornando real a produção de plantas transgênicas, graças à possibilidade de transferência de genes entre diferentes organismos (ENGELS & VISSER, 2003). Tais avanços levaram à criação de uma “biblioteca internacional de DNA”, que visa guardar coleções de fragmentos de DNA genômico (ADAMS, 1997). No entanto, Adams et al. (1992) e Ramanatha Rao (1998), esclarecem que estratégias e procedimentos de utilização desse material armazenado na forma de DNA devem ser desenvolvidos, uma vez que ainda não é possível a regeneração de um indivíduo por completo a partir destes fragmentos.

2.2.4. Caracterização e avaliação

2.2.4.1. Importância para a conservação e o melhoramento de plantas

Recursos genéticos vegetais são coletados não apenas para conservação, mas também para serem explorados em programas de melhoramento genético. Entretanto, o emprego racional desses recursos só se faz possível mediante o conhecimento mais apurado sobre suas características e possíveis usos. De acordo com Jaramillo e Baena (2000), a informação que permite conhecer e determinar a utilidade de um germoplasma provém de registrar e analisar um conjunto de dados sobre este, em diversas etapas da conservação, mas principalmente durante a caracterização e a avaliação. Conforme estes autores, caracterização e avaliação são atividades complementares que consistem em descrever os atributos de amostras de uma mesma espécie com o objetivo de diferenciá-las, determinando suas estruturas, variabilidade genética e relações entre-si.

Segundo Faleiro (2007), a caracterização pode ser definida como a atividade de descrição e registro de características morfológicas, agronômicas, ecogeográficas e moleculares de um indivíduo ou acesso. Em concordância com este autor, Valois et al. (1996) adicionam à definição de caracterização a descrição e o registro de características citogenéticas e bioquímicas do indivíduo. Posterior à caracterização, a avaliação consiste em descrever características agronômicas, geralmente quantitativas e de baixa herdabilidade, como a produção de grãos e a resistência a estresses bióticos e abióticos de acessos, no maior número possível de diferentes ambientes, com vistas à identificação de genes úteis para o melhoramento de culturas (JARAMILLO & BAENA, 2000).

Do ponto de vista histórico, domesticação e melhoramento de plantas são processos indissociáveis, uma vez que o melhoramento começou a ser praticado somente a partir da domesticação de culturas consideradas fundamentais para sobrevivência do homem, a cerca de 10000 anos atrás (NASS et al., 2007). Nesse início, ambas as atividades dependiam diretamente de sua capacidade de observação e intuição, para, desta maneira, proceder a seleção de plantas que mais lhe trariam proveito, principalmente no âmbito alimentar e medicinal. Paterniani (1988) ressalta que das mais de 300000 espécies descritas, o homem chegou a utilizar cerca de 3000 para sua alimentação. Para tal, a observação, em especial, se tornou uma importante aliada para os “primeiros melhoristas” estabelecerem bases de informações para a caracterização de plantas e, com isto, descrever diferenças e semelhanças entre elas.

A partir da caracterização, o ainda incipiente processo de melhoramento genético vegetal conseguiu unir subsídios suficientes para aperfeiçoar a domesticação de materiais com características interessantes ao homem. Nesse sentido, homem e natureza não visam os mesmos fenótipos, pois a seleção natural age em prol dos indivíduos mais adaptados a uma determinada condição ambiental, enquanto a seleção artificial, praticada pelo homem, prioriza características importantes para suas necessidades. Para Nass et al. (2007), a seleção feita pelo homem pode caminhar no sentido oposto ao da seleção natural, em que algumas espécies chegam a atingir um grau de domesticação tão elevado que, hoje, já não sobrevivem naturalmente sem a interferência humana, como, por exemplo, o milho (*Zea mays* L.). Dessa forma, pode-se inferir que a caracterização atuou promovendo uma maior uniformidade do material genético das espécies selecionadas ao longo do processo da domesticação, selecionando plantas com caracteres mais convenientes ao homem e excluindo da seleção, plantas portadoras de características indesejadas.

Apesar da grande importância da caracterização para o início da domesticação de plantas, o melhoramento genético ainda era tratado como uma “arte empírica”, pois a opção por um determinado tipo de planta era baseada somente em observações fenotípicas das características de interesse (NASS et al., 2007). Apenas com o surgimento de trabalhos clássicos no campo da genética, como a formulação das “Leis da hereditariedade” por Gregor Mendel (1866) e a “Teoria das linhas puras” de Wilhelm Johannsen (1903), que o melhoramento genético vegetal passou a ser realizado com maior embasamento teórico, sendo então considerado uma nova ciência. Dado o maior conhecimento científico obtido por meio destes novos estudos, entre as etapas de conservação de recursos genéticos vegetais, a prática de caracterização se tornou a principal atividade a ser realizada em programas de melhoramento genético vegetal, sendo, hoje em dia, determinante para o sucesso destes (VALLS, 2007).

Segundo Saad e Rao (2001), até que sejam devidamente caracterizados e avaliados, a simples coleta e introdução de novos materiais vegetais, em bancos ou coleções de germoplasma, não garantem sua pronta utilidade em programas de melhoramento genético. Para tanto, a caracterização deve ser empregada a fim de permitir ganhos genéticos mais promissores ao longo do melhoramento da espécie alvo e potencializar o uso destes recursos por agricultores (COELHO et al., 2007). Sob a ótica do melhoramento genético vegetal, Jaramillo e Baena (2000) definem caracterização como a descrição e registro de características de alta herdabilidade, as quais podem ser facilmente observadas a olho nu, e que são pouco influenciadas pelo ambiente.

Em estudo sobre gerenciamento de coleções de germoplasma vegetal, Engels e Visser (2003) revelam que uma caracterização bem executada pode contribuir diretamente para uma administração mais racional dos recursos fitogenéticos, fornecendo ao curador subsídios importantes para identificação de duplicatas; para tomada de decisões sobre quais acessos devem ser regenerados; e, principalmente, para estimação da variabilidade genética da

coleção. Nesse sentido, Reif et al. (2005) e Coelho et al. (2007) reiteram que a quantificação da diversidade genética, baseada nos dados gerados pela caracterização, auxilia no planejamento de estratégias mais eficazes que venham maximizar os ganhos genéticos no melhoramento através da identificação de combinações parentais adequadas à obtenção de híbridos.

A caracterização também exerce papel valioso sob o enfoque da conservação de recursos genéticos vegetais, visto que está intimamente ligada à manutenção e potencial exploração destes, provendo informações mais profundas sobre a variabilidade genética das coleções de germoplasma (CHIORATO, 2004). Para Perssoni (2007), este conhecimento também proporciona contribuições importantes à conservação e ao entendimento do processo evolutivo das espécies.

De acordo com Jaramillo e Baena (2000), assim como qualquer plano estratégico, a conservação de recursos fitogenéticos implica em planejamento e tomada de decisões baseadas em informações prévias. Nesse ponto, a caracterização pode contribuir significativamente para a discriminação da variabilidade genética do germoplasma, auxiliando no estabelecimento de coleções bases e nucleares, estrategicamente importantes para segurança alimentar de uma nação, e exploração racional dos recursos genéticos das espécies alvos.

A informação registrada através da caracterização deve ser prontamente repassada a um sistema de documentação capaz de determinar padrões de variação, bem como identificar os acessos da coleção (SAAD & RAO, 2001). Para Engels e Visser (2003), estas informações facilitam a utilização das coleções, pois permitem a detecção de materiais mal identificados e eventuais erros cometidos ao longo das atividades procedidas no banco, estabelecendo maior segurança sobre as práticas referentes à conservação e utilização. Dessa maneira, Valls (2007) conclui que o incremento das atividades de caracterização deve se impor entre as prioridades na estratégia de abordagem e manejo de recursos genéticos vegetais. Superadas as deficiências de informação, o germoplasma disponível será mais útil para programas de melhoramento e sua própria conservação.

2.2.4.2. Tipos de caracterização

Estudos que envolvem a diversidade genética em plantas perenes têm sido realizados frequentemente com base em descritores morfológicos, botânicos e agrônômicos (DIAS et al., 1997). Entretanto, Ramos et al. (2006) atestam que a caracterização de uma coleção de germoplasma também pode ser realizada baseada em aspectos reprodutivos, citogenéticos e moleculares.

Segundo Mariot et al. (2008), a caracterização morfológica é a primeira atividade a ser efetuada dentre as etapas que envolvem a identificação de acessos em bancos de germoplasma. Para Ramalho et al. (2000), a caracterização via descritores morfológicos é dada através da descrição de caracteres altamente herdáveis, na qual o genótipo do vegetal estudado pode ser facilmente avaliado por meio de seu fenótipo. Nesse sentido, Jarvis et al. (2000) declaram que o emprego de critérios morfoagronômicos para seleção de genótipos potencialmente úteis ao melhoramento é comumente usado como forma de identificar a diversidade genética entre o germoplasma da espécie objeto através da análise de sua variabilidade fenotípica. Ainda sob esta perspectiva, estudos de enfoque botânico são especialmente necessários para a determinação correta de aspectos morfológicos da espécie alvo. Nesse quesito, as dificuldades inerentes à complexa taxonomia das plantas exigem alta especialização por parte de, ao menos, um membro da equipe envolvida no processo, principalmente quando o germoplasma avaliado inclui espécies silvestres (VALLS, 2007).

De acordo com Juhász et al. (2009), os procedimentos adotados na execução de programas de melhoramento genético de espécies perenes dependem, necessariamente, de seu modo de reprodução, de informações sobre sua biologia e morfologia floral, e de aspectos relativos à polinização e fertilização. A caracterização reprodutiva, geralmente baseada na citogenética (através da contagem de cromossomos, análise do comportamento meiótico, exame do saco embrionário e análise da viabilidade de pólen), também pode requerer cruzamentos controlados e estudos de progênies, dependendo da disponibilidade de laboratórios adequados, em termos de mão de obra capacitada, equipamentos e reagentes (VALLS, 2007). Segundo este autor, o conhecimento da forma de reprodução de uma determinada espécie já se faz importante durante as etapas iniciais de caracterização, visto que seu desconhecimento pode levar à rápida descaracterização dos acessos na própria coleção ativa, uma vez que as sementes eventualmente colhidas podem ser resultantes de cruzamentos espontâneos entre genótipos distintos, nada tendo a ver com a expressão das características anotadas nos genitores maternos. No entanto, também pode ser observada a ocorrência de casos de autogamia e apomixia, por exemplo, em plantas inicialmente consideradas alógamas. Silva et al. (2001) afirmam que o conhecimento da estrutura floral numa dada cultura é vital para que o melhorista desenvolva técnicas de emasculação e polinização adequadas. Dessa maneira, apesar de onerosa, a caracterização reprodutiva pode poupar anos de insucesso na condução de trabalhos com o germoplasma e fornecer informações de pronto uso em programas de melhoramento genético (VALLS, 2007).

Assim como a citogenética, a caracterização molecular também é uma importante ferramenta auxiliar à caracterização morfoagronômica em estudos de diversidade genética entre plantas. Apesar de permitir grandes avanços nos processos de conhecimento e organização de coleções de germoplasma vegetal, a caracterização morfoagronômica tem sido complementada por outros critérios, visto que somente o uso de descritores morfoagronômicos pode apresentar algumas limitações, como: exigência de um grande número de descritores para possibilitar uma classificação adequada; ser intensivo em tempo e recursos para execução; e pelo fato de algumas características serem muito plásticas, sendo fortemente influenciadas por variações ambientais (FERREIRA et al., 2007b). Buscando superar tal limitação, novos métodos de análise foram introduzidos em auxílio à caracterização morfológica, como a caracterização baseada em técnicas de marcadores moleculares (KARP et al., 1997; FERREIRA et al., 2007b).

Segundo Ferreira e Grattapaglia (1998), marcadores moleculares são características de DNA, herdadas geneticamente, que diferenciam dois ou mais indivíduos. A alta resolução de técnicas moleculares permite identificar diferenças genéticas entre indivíduos onde não se observa distinção fenotípica aparente, tornando ilimitado o potencial de mensuração de polimorfismo entre acessos de uma coleção (FERREIRA et al., 2007b). Para estes autores, uma importante qualidade destes marcadores é a investigação direta dos genótipos, que permite a detecção de variação no DNA, excluindo, portanto, influências ambientais. De acordo com Karp et al. (1997), a análise baseada em marcadores moleculares também pode ser realizada em diferentes estádios de desenvolvimento da planta, sem apresentar risco sobre a interpretação dos resultados. Ferreira et al. (2007b) ressaltam que diversas técnicas de marcadores moleculares têm sido desenvolvidas, mas nenhuma deve ser considerada universal. Dentre os marcadores moleculares mais utilizados, destacam-se o RFLP (*Restriction Fragment Length Polymorphism*), os Minissatélites ou VNTR (*Variable Number of Tandem Repeats*), o RAPD (*Random Amplified Polymorphic DNA*), o SCAR (*Sequence Characterized Amplified Regions*), os Microssatélites ou SSR (*Simple Sequence Repeats*), o ISSR (*Inter Simple Sequence Repeats*), o AFLP (*Amplified Fragment Length Polymorphism*) e o SNP (*Single Nucleotide Polymorphism*) (FERREIRA & GRATTAPAGLIA, 1998; FALEIRO, 2007). Para Faleiro (2007), os SNPs, que se baseiam nas posições de um único

nucleotídeo para a identificação de mutações e polimorfismos, apresentam vantagens em relação aos outros marcadores, pois permitem a detecção de um grande número de polimorfismos entre alelos de um determinado gene. Ainda segundo Faleiro (2007), outro marcador genético-molecular importante são os isoenzimáticos, referentes às isoenzimas, grupo de múltiplas formas moleculares de uma enzima, resultante de variações alélicas dos genes codificadores. As principais vantagens desta técnica são o baixo custo e a facilidade de sua metodologia, bem como a obtenção de marcadores genéticos co-dominantes, isto é, que permitem a diferenciação dos locos em homozigose de locos em heterozigose. Entretanto, este método apresenta desvantagens, como o baixo número de sistemas enzimáticos polimórficos e influência das condições ambientais e dos tecidos vegetais nas atividades enzimáticas (ALFENAS, 1998; FALEIRO, 2007). Nesse contexto, em detrimento à técnica isoenzimática, outros marcadores co-dominantes podem ser preferidos, como os microssatélites ou até mesmo o RFLP. Sendo assim, a escolha do método de caracterização depende da espécie estudada, da questão a ser esclarecida, bem como do financiamento e das tecnologias disponíveis (REDONDO, 2011). No entanto, analisando sob esta perspectiva, Karp et al. (1997) advertem que a caracterização morfológica não deve ser substituída por quaisquer técnica molecular. Ainda sob este foco, Saad e Rao (2001) e Ferreira et al. (2007b) atentam que dados morfológicos e moleculares devem ser usados de forma complementar sempre que possível.

2.2.5. Descritores mínimos

2.2.5.1. Definição e importância para a caracterização

De acordo com o *Bioversity International* (2007), uma das principais causas da utilização irracional do germoplasma de diversas espécies é a carência de informações relativas à sua identificação, como a falta, ou inadequada, caracterização e avaliação, e a omissão de dados de passaporte. Segundo esta Organização Científica Internacional, curadores e melhoristas não conseguem explorar os recursos fitogenéticos de maneira eficiente quando tais informações são omitidas e/ou descritas de forma imprópria e não-padronizada. Tal fato é justificado mediante a importância destes dados para o estabelecimento de coleções, em que a identificação de duplicatas e materiais ainda não identificados, ou mal descritos, se faz essencial para a avaliação da variabilidade genética presente no germoplasma estudado.

A quantificação da diversidade genética existente em um dado germoplasma vegetal pode ser fundamentada sobre aspectos morfológicos e agrônômicos (SILVA et al., 2011). Nesse sentido, a descrição de características capazes de distinguir acessos dentro de uma coleção de germoplasma deve ser realizada de maneira eficaz, possibilitando traçar a melhor estratégia de condução e seleção de populações para o sucesso do programa de melhoramento genético da espécie alvo. Dessa maneira, Saad e Rao (2001) revelam que o processo de caracterização do germoplasma deve ser iniciado com o uso de uma lista de descritores apropriada para a espécie objeto de estudo.

Um descritor é definido como um atributo, ou característica, passível de ser mensurada, que pode ser observada entre os diferentes acessos de uma coleção de germoplasma (BIOVERSITY INTERNATIONAL, 2007). Jaramillo e Baena (2000) declaram que estes devem ser específicos para cada espécie, além de expressar o atributo de maneira uniforme e precisa. Dessa forma, a *lista de descritores* de uma espécie é constituída pelo conjunto de descritores individuais capazes de distinguir um genótipo do outro dentro de uma coleção de germoplasma (PAINTING et al., 1993; BIOVERSITY INTERNATIONAL, 2007).

2.2.5.2. Órgão regulamentador: *Bioversity International*

Para que um sistema de dados se torne amplamente difundido e aplicável, suas informações devem ser standardizadas em termos de unidades de medida e terminologias (PARODA & ARORA, 1991). Com a finalidade de simplificar operações relativas à conservação de recursos genéticos vegetais, como o registro, a modificação/atualização e a recuperação de informações, o uso de listas de descritores padronizadas pode reduzir a ocorrência de erros e assegurar uma maior uniformidade para avaliação de dados intra e entre diferentes bancos de germoplasma (PAINTING et al., 1993).

Objetivando a documentação de informações mais precisas e seguras sobre a origem, caracterização e desempenho agrônômico, com vistas à conservação de recursos genéticos, o *Bioversity International*, Organização Científica Internacional (autônoma), atua, mundialmente, promovendo listas de descritores normatizadas, manuais e guias para a elaboração destas a melhoristas e curadores de coleções, como o *Guidelines for the development of crop descriptor lists* (BIOVERSITY INTERNATIONAL, 2007). Esta Organização tem como principal objetivo assegurar o bem estar das futuras gerações por meio da regulamentação e aperfeiçoamento de métodos e políticas relativas à conservação da biodiversidade em florestas e propriedades rurais (BIOVERSITY & THE CHRISTENSEN FUND, 2009). Para tal, o *Bioversity International* opera através de quatro programas (*Diversity for Livelihoods*; *Understanding and Managing Biodiversity*; *Global Partnerships*; e *Commodities for Livelihoods*) financiados por mais de 150 doadores ao redor do mundo, como governos federais, empresas privadas e organizações internacionais, entre elas, o Fundo para a Agricultura e Alimentação das Nações Unidas (FAO) (BIOVERSITY INTERNATIONAL, 2007).

A adoção das listas de descritores formuladas pelo *Bioversity International* permite uma maior padronização no processamento e administração dos dados do germoplasma avaliado (PARODA & ARORA, 1991; SAAD & RAO, 2001; BIOVERSITY INTERNATIONAL, 2007). Sob esta perspectiva, tanto o *Bioversity International* quanto seus predecessores (*IBPGR - International Board for Plant Genetic Resources* e *IPGRI - International Plant Genetic Resources Institute*) já publicaram listas de descritores para, aproximadamente, 100 culturas, nas quais destacam-se os descritores de espécies oleaginosas, como a soja (*Glycine max* - 1984) e o girassol (*Helianthus annuus* L. - 1985) e, também, de grandes culturas, como o milho (*Zea mays* L.- 1991) e o arroz (*Oryza sativa* - 2006) (SAAD & RAO, 2001; BIOVERSITY INTERNATIONAL, 2007).

Apesar da grande quantidade de listas já publicadas, ainda há uma grande demanda para o desenvolvimento de descritores para espécies florestais e variedades de importância local e regional (BIOVERSITY INTERNATIONAL, 2007). Sendo assim, os guias para elaboração de listas de descritores, formulados pelo *Bioversity International*, fornecem aos curadores e melhoristas informações sobre a melhor estrutura, metodologia e elementos essenciais para a construção de listas padronizadas para a espécie de interesse. De acordo com Painting et al. (1993), tal procedimento permite uma melhor comunicação entre pesquisadores e instituições, resultando em um maior aproveitamento dos recursos genéticos conservados. O intercâmbio de informações entre diferentes instituições só é realizado de forma eficiente através de um sistema de documentação compatível e padrão, que possa facilitar a comunicação e o entendimento por ambas as partes (BIOVERSITY INTERNATIONAL, 2007).

2.2.5.3. Regras básicas para elaboração de listas de descritores padronizadas

Baseado no guia previamente citado, *Guidelines for the development of crop descriptor lists*, a elaboração de listas de descritores deve seguir algumas regras básicas, como

a utilização de descritores simples e auto-explicativos, que permitam uma fácil interpretação por parte do leitor. Para tal, cada descritor deve possuir três elementos essenciais para simplificar sua aplicação prática no campo: um *nome*, que deve descrever o atributo a ser caracterizado de forma simples, compacta e não ambígua; um *estado*, que deve ser capaz de expressar a característica do atributo observado, e, se possível, fazer uso de materiais de referência, como desenhos, esquemas gráficos e fotografias, para melhor entendimento; e um *método* de mensuração e coleta que, assim como os elementos já mencionados, deve ser claro e de fácil compreensão, visando otimizar sua aplicação prática. Sendo assim, estes três elementos devem evitar a ambiguidade e prezar, sempre, pela simplicidade e clareza. Dessa forma, a lista elaborada alcançará um maior nível de padronização e reprodução, contemplando um maior número de profissionais para seu uso (BIOVERSITY INTERNATIONAL, 2007).

Atenção especial também deve ser dada ao registro e à representação de dados qualitativos e quantitativos. Descritores de caracteres qualitativos devem ser auto-explicativos e descritos em escala nominal, ordinal ou binária. A escala nominal é usada para atributos definidos por texto, como coloração de frutos, onde os códigos numéricos, utilizados para identificação da característica, não seguem uma ordem ou *ranking* especial. Ao contrário da escala nominal, a ordinal estabelece uma ordem para enumerar os diferentes estados expressos pela característica a ser descrita. Dessa forma, os códigos numéricos são expostos em ordem crescente, distinguindo o estado da característica de menor ao de maior intensidade, por exemplo. Já a escala binária é utilizada para caracteres qualitativos de apenas duas categorias, denotando sua ausência ou presença (sim/não; presença/ausência) como, por exemplo, em presença ou ausência de pilosidade em folhas. Dessa maneira, é identificado por “0” para *não* ou *ausência*, e “1” para *sim* ou *presença* (BIOVERSITY INTERNATIONAL, 2007).

De acordo com Jaramillo e Baena (2000), características quantitativas são poligênicas, altamente influenciadas pelo efeito ambiental e de distribuição contínua. Entretanto, caracteres quantitativos podem ser registrados tanto através da escala discreta (para características que podem ser facilmente quantificadas, como número de plantas) quanto pela contínua (para atributos mensuráveis, como altura, peso, comprimento e diâmetro de planta). Sendo assim, a escala discreta é empregada em casos onde o maior objetivo é a simples identificação da variação apresentada pelo atributo, uma vez que este tipo de escala assume somente um número finito de valores para registro, não requerendo grande precisão para sua avaliação. Nesses casos, a divisão em “classes” de distribuição discreta basta. No entanto, o *Biodiversity International* sugere que o registro de variáveis quantitativas seja dado, de preferência, pelo valor exato expresso pela característica avaliada, recomendando o uso do Sistema Internacional de Medidas (*Système International d’Unités* - SI). Desse modo, do ponto de vista estatístico, o registro do valor exato expresso pelo *estado* da característica confere maior precisão e fidelidade à descrição, permitindo, quando julgada conveniente, sua conversão para escala discreta e, assim, ser também avaliada como um dado multicategórico, através da definição de classes.

Quando seguidas estas recomendações, a lista de descritores recém elaborada pode funcionar como uma importante ferramenta de auxílio ao curador e ao melhorista, promovendo uma documentação mais consistente e segura do germoplasma conservado e garantindo a exploração contínua deste material pela comunidade envolvida na conservação de recursos genéticos vegetais (BIOVERSITY INTERNATIONAL, 2007).

2.3. Diversidade Genética

2.3.1. Importância para o melhoramento genético vegetal

Tendo em vista a importância da quantificação da variabilidade genética para a identificação de genótipos portadores de características desejáveis ao melhoramento, a estimação da divergência genética entre um grupo de acessos em coleções de germoplasma pode fornecer materiais genéticos superiores para programas de melhoramento, gerando, no futuro, linhagens, compostos e até híbridos de alta performance agronômica.

Segundo Cruz e Carneiro (2003), a análise da diversidade genética permite a diferenciação de acessos em coleções de germoplasma, promovendo a identificação de genitores adequados à obtenção de híbridos de maior efeito heterótico, capazes de proporcionar maior segregação em recombinações e, conseqüentemente, maior possibilidade de recuperação de genótipos superiores. Conforme Wesp et al. (2008), tal fenômeno é conhecido como *segregação transgressiva*.

Além de promover a identificação de genitores divergentes e possuidores de características complementares, a diversidade genética ainda pode ser utilizada como uma importante ferramenta para a melhor compreensão do processo evolutivo das plantas (MIRANDA, 1998). Para Falconer (1981), a expectativa de que pais divergentes proporcionem boas combinações híbridas procede do fato de que a heterose manifestada em híbridos é dada em função dos efeitos da dominância dos genes para o caráter em questão, e do quadrado da diferença das frequências gênicas de seus genitores, somados aos efeitos epistáticos, geralmente negligenciados. De acordo com Mather e Jinks (1971), *heterose* (ou vigor híbrido) pode ser definida como o desempenho superior de híbridos da F₁ em relação aos seus genitores. Faleiro (2007) declara que o vigor híbrido ocorre quando o comportamento do híbrido está situado acima da média de seus genitores. Dessa maneira, é possível que genótipos geneticamente dissimilares, porém, detentores de fenótipos similares e superiores para o caráter de interesse, tenham locos distintos para o controle deste, que podem vir a ser reunidos em um terceiro genótipo, pela ação de genes complementares, dando origem a um segregante transgressivo, indivíduo com valores fenotípicos superiores aos dos genitores para o caráter sob seleção (CARVALHO et al., 2001; RODRIGUES, 2008).

2.3.2. Estimação da diversidade genética

Segundo Miranda (1998) e Cruz e Carneiro (2003), a diversidade genética entre genótipos pode ser estimada, geralmente, a partir de características morfológicas, agronômicas e moleculares, através da utilização de métodos biométricos baseados na quantificação da heterose ou por meio de processos preditivos.

Dentre os métodos relativos à natureza quantitativa de avaliação da diversidade, estão as análises dialélicas. Este método é caracterizado pela avaliação de p genitores e de todas (ou amostras de) as suas combinações híbridas, resultando num total de $p(p-1)/2$ híbridos a serem avaliados (CRUZ & CARNEIRO, 2003). Conforme estes autores, entre os métodos preditivos da heterose citam-se aqueles que tomam as diferenças morfológicas, fisiológicas ou moleculares como base para quantificá-las em alguma medida de dissimilaridade capaz de expressar o grau de diversidade genética entre os genitores. Desse modo, a predição da divergência genética pode ser dada através da aplicação de técnicas estatísticas multivariadas, entre as quais, destacam-se as análises de agrupamento (métodos aglomerativos), de componentes principais e de variáveis canônicas (CRUZ & REGAZZI, 2001). Para tal, a análise multivariada utiliza-se de coeficientes de distância genética entre os genótipos. Sob esta perspectiva, parte-se da hipótese de que quanto maior a distância genética entre dois genótipos, maiores também são as chances de se obter combinações mais promissoras. Sendo assim, linhagens que possuem um grande número de alelos em comum para uma determinada

característica são designadas como similares. Contrariamente, linhagens divergentes, que apresentam diferenças na frequência dos locos considerados, são ditas mais apropriadas para cruzamentos e formação de populações segregantes (CHIORATO, 2004).

A avaliação simultânea de vários descritores via análises de agrupamento, componentes principais e variáveis canônicas são opções vantajosas na otimização do uso e da avaliação de coleções de germoplasma, sendo de grande utilidade na predição de combinações heteróticas, onde centenas de cruzamentos indesejáveis podem ser evitados (DIAS, 1994). Entretanto, considerando as vantagens e desvantagens de cada método, Cruz e Regazzi (2001) alertam que a escolha da metodologia mais adequada deve ser dada de acordo com a cultura estudada, precisão desejada pelo pesquisador e pela facilidade da análise e forma como os dados foram obtidos.

2.3.2.1. Tipos de variáveis utilizadas na estimação da divergência genética

Segundo Cruz e Carneiro (2003), estudos de diversidade genética, de maneira geral, têm sido realizados através de medidas de distância (de similaridade ou de dissimilaridade), obtidas a partir de variáveis quantitativas, multicategóricas e binárias.

Uma variável quantitativa é toda aquela que apresenta como possíveis realizações valores oriundos de uma contagem, apuração ou medição (AZAMBUJA, 2005). Dessa forma, as variáveis quantitativas podem ser discretas ou contínuas. Uma variável quantitativa é considerada discreta quando o conjunto de valores que ela assume for finito (enumerável). Por outro lado, uma variável quantitativa é dita contínua quando o conjunto de valores por ela assumidos for infinito (não-enumerável) (BUSSAB et al., 1990; AZAMBUJA, 2005; BIOVERSITY INTERNATIONAL, 2007).

De acordo com Azambuja (2005), uma variável é considerada qualitativa quando apresenta como possíveis valores uma qualidade (ou atributo) do indivíduo estudado. Conforme Bussab et al. (1990), elas se dividem em variáveis qualitativas nominais, cujos valores não possuem ordenação, e ordinais, quando estes apresentam uma ordenação. Para Cruz e Carneiro (2003), quando uma variável qualitativa pode ser ordenada, esta pode ser considerada multicategórica. Nesse sentido, Sneath e Sokal (1973) observam que, quando ordenados, estes valores podem ser analisados como variáveis quantitativas discretas, estabelecendo-se uma escala de classes.

Para Gower (1971), variáveis binárias (ou dicotômicas) são empregadas para avaliar caracteres qualitativos de apenas duas categorias, demonstrando sua ausência ou presença. Cruz e Carneiro (2003) atentam para a grande importância do emprego das variáveis binárias na avaliação da diversidade genética entre genótipos a partir de dados obtidos de estudos moleculares, onde é julgada a presença ou a ausência de marcas (bandas), codificadas por “0”, para ausência, e “1” para presença.

2.3.2.2. Medidas de similaridade e dissimilaridade

De acordo com Stähelin et al. (2011), a estatística multivariada pode auxiliar melhoristas de plantas na identificação de genótipos divergentes a serem utilizados na obtenção de populações segregantes de maior variabilidade. Nesse sentido, Moreira et al. (1994) afirmam que a divergência genética está diretamente relacionada ao grau de distanciamento entre indivíduos/populações quanto ao conjunto de caracteres que lhes são peculiares. Dessa maneira, a aplicação de técnicas estatísticas multivariadas, como as análises de agrupamento, para predição da divergência genética entre um grupo de genitores, requer o emprego de algum coeficiente, ou tipo de medida, que meça a distância entre estes, quantificando o quanto são parecidos (BUSSAB et al., 1990; AZAMBUJA, 2005). Segundo Valentin (1995), o coeficiente capaz de quantificar o grau de associação entre um par de

objetos, ou de descritores, é chamado de *medida de semelhança*. Para Bussab et al. (1990), tal medida é denominada *coeficiente de parença* e, tecnicamente, pode ser dividida em duas categorias: *medidas de similaridade* e *medidas de dissimilaridade*.

Nas medidas de similaridade, quanto maior o valor observado, mais parecidos serão os objetos (mais semelhantes), ao passo que, nas medidas de dissimilaridade, quanto maior for o valor apresentado, menos parecidos serão os objetos (mais dissimilares) (BUSSAB et al., 1990; CHIORATO, 2004). Para Azambuja (2005), quando o maior valor calculado para um coeficiente for empregado para indicar que dois objetos são parecidos, trata-se da utilização de um *coeficiente de similaridade*, em contrapartida, se o menor valor relativo a um coeficiente indicar que dois objetos são parecidos, refere-se ao uso de um *coeficiente de dissimilaridade*. Um exemplo de medida de similaridade é o coeficiente de correlação, pois quanto maior seu valor, maior também será a associação entre os objetos, e consequentemente, a semelhança entre eles. Por outro lado, a distância Euclidiana, um dos coeficientes mais utilizados, é um exemplo de medida de dissimilaridade, visto que quanto menor o seu valor, mais próximos os objetos estão uns dos outros (BUSSAB et al., 1990; AZAMBUJA, 2005).

As medidas de distância consideram que, se dois indivíduos são similares, estes estão próximos um do outro, ou seja, são comuns ao conjunto de variáveis e vice-versa. No entanto, a utilização de diferentes medidas de distância pode gerar resultados distintos de aglomeração. Sob esta perspectiva, Azambuja (2005) ressalta que certos coeficientes se adaptam melhor a determinados tipos de variáveis e situações, em que, geralmente, coeficientes de dissimilaridade se mostram mais adequados às variáveis quantitativas, enquanto coeficientes de similaridade são mais apropriados às variáveis qualitativas. Desse modo, Cruz e Carneiro (2003) ressaltam que a utilização de medidas de dissimilaridade para variáveis quantitativas se faz de grande valia em estudos de diversidade genética. Nesse sentido, estes autores atentam que a estruturação da matriz de dissimilaridade é usualmente dada pela distância Euclidiana e pela distância generalizada de Mahalanobis (D^2). De acordo com Cruz e Regazzi (2001), a distância de Mahalanobis oferece vantagem em relação à distância Euclidiana por levar em consideração a existência das variâncias e covariâncias residuais entre os caracteres avaliados, porém, necessita do emprego de delineamentos estatísticos. Contudo, deve-se enfatizar que a distância Euclidiana é uma análise baseada no Teorema de Pitágoras, juntamente com a aplicação de múltiplos eixos ortogonais, e, quando obtida de variáveis padronizadas ou estandardizadas, torna-se mais eficaz (CHIORATO, 2004).

No âmbito do melhoramento genético vegetal, um dos principais aspectos a serem observados no emprego de métodos de agrupamento é a escolha da medida de distância mais adequada para estimar a divergência genética entre um grupo de genitores, visto que cada técnica de agrupamento tem como base uma medida de distância apropriada para o método escolhido, dessa forma, as diferentes técnicas de agrupamento nem sempre resultam no mesmo padrão de aglomeração (JACKSON et al., 1989). Segundo estes autores, várias discrepâncias podem ser observadas em relação aos agrupamentos e às inferências referentes à quantificação da variabilidade entre acessos de um banco de germoplasma. Sob este aspecto, deve-se enfatizar a existência de outras medidas, como as distâncias de Minkowski, Chebyshev (*Maximum*), Manhattan (*City-block*) e de Canberra, recomendadas para variáveis quantitativas, o coeficiente de concordância simples ou de coincidência (*Simple Matching* ou *Matching Coefficient*), de Jaccard e o de Sørensen-Dice (Nei e Li), indicados para variáveis qualitativas binárias, além do coeficiente de Cole-Rodger, para variáveis multicategóricas, entre outras, enaltecendo, sempre, a importância de se testar medidas diferentes para comparação de resultados (BUSSAB et al., 1990; MALHOTRA, 2001; CRUZ & CARNEIRO, 2003; AZAMBUJA, 2005). Com o objetivo de avaliar as alterações provocadas por diferentes medidas de similaridade, Duarte et al. (1999) testaram os coeficientes de

Anderberg, Simple Matching, Rogers e Tanimoto, Russel e Rao, Ochiai, Jaccard, Sørensen-Dice e Ochiai II no agrupamento de 27 cultivares de feijão, analisados por marcadores moleculares RAPD, e concluíram que, mesmo provocando poucas mudanças na estrutura dos grupos mais diferenciados, estes coeficientes alteraram algumas relações entre cultivares com alta similaridade genética.

O uso de técnicas estatísticas multivariadas tem colaborado para o aumento de estudos referentes à quantificação da divergência genética em coleções de germoplasma (SANTOS, 2010). Apesar da extensa lista de medidas de distância disponível para a aplicação destas técnicas, grande parte dos trabalhos publicados sobre o tema pecam por não justificarem a escolha do coeficiente utilizado para estimação (DUARTE et al., 1999). Considerando que tanto o agrupamento quanto a ordenação dos resultados podem ser diretamente influenciados por esta escolha, Jackson et al. (1989) e Duarte et al. (1999) reiteram a importância de se conhecer melhor os coeficientes disponíveis para utilização, para que, dessa forma, os mais adequados sejam aplicados. Nesse contexto, com o intuito de facilitar o emprego de medidas de distância através da criação de um coeficiente geral único, capaz de abranger o critério de aplicação de todos os outros já existentes, Gower (1971) propôs uma técnica que permite a análise simultânea de dados quantitativos e qualitativos. Este método permite que os valores da matriz de distância fiquem compreendidos entre 0 e 1, através da padronização das variáveis (BUSSAB et al., 1990). Tal técnica foi adotada com sucesso por Das et al. (2008), onde, aliada ao método hierárquico de agrupamento UPGMA, utilizaram o coeficiente geral de Gower para estimar a divergência genética entre 18 acessos indianos de *J. curcas*. Também por meio do método de agrupamento UPGMA, Santos (2010) empregou o algoritmo proposto por Gower para avaliar a variabilidade genética entre acessos de maracujazeiro a partir de dados quantitativos, moleculares e multicategóricos.

2.3.2.3. Análises de agrupamento

As análises de agrupamento (*Cluster analysis*) têm por finalidade separar um grupo original de observações em vários subgrupos, de tal forma que se obtenha homogeneidade dentro dos grupos e heterogeneidade entre estes (BUSSAB et al., 1990). Segundo Cruz e Carneiro (2003), dentre os métodos de agrupamento mais empregados por melhoristas, estão os de otimização (Tocher) e os hierárquicos (Vizinho mais próximo; Vizinho mais distante; UPGMA; e Ward). No entanto, outros métodos hierárquicos também podem ser utilizados, tais como: UPGMC, ou Centróides (LANCE & WILLIAMS, 1967), WPGMA, ou Mcquitty (SOKAL & MICHENER, 1958) e WPGMC, ou Mediana (GOWER, 1967). É importante notar que estes vários métodos de agrupamento se diferenciam pelo tipo de resultado e pelas diferentes maneiras de declarar graficamente a proximidade entre um indivíduo e um grupo já formado ou entre dois grupos quaisquer, para os quais não se conhece, *a priori*, o número de grupos a serem estabelecidos (CRUZ & CARNEIRO, 2003).

O método de otimização de Tocher, citado por Rao (1952), é caracterizado pela formação de grupos a partir da adequação de algum critério de agrupamento. Dessa forma, o método tem como objetivo alcançar uma partição dos indivíduos que otimize, maximizando ou minimizando, alguma medida de distância pré-estabelecida, diferindo dos métodos hierárquicos, pelo fato de os grupos formados serem mutuamente exclusivos (CRUZ & REGAZZI, 2001; CRUZ & CARNEIRO, 2003). Este método foi empregado para estudar a variabilidade genética entre genótipos de *J. curcas* por Abreu et al. (2009), Laviola et al. (2010) e Rosado et al. (2010).

Nos métodos hierárquicos, os genótipos são agrupados por um processo que se repete em vários níveis, até que seja estabelecido um dendrograma que permite a identificação de grupos homogêneos, sem maior preocupação com a obtenção de um número ótimo de grupos.

Para tal, Cruz e Carneiro (2003) apresentam quatro diferentes formas de representar a estrutura de agrupamento com base na distância entre os pares de genótipos: a) utilizando a menor distância existente entre um par de genótipos, pelo método do *Vizinho mais próximo* (ou ligação simples); b) fazendo uso da maior distância existente entre um par de genótipos, pelo método do *Vizinho mais distante* (ou ligação completa); c) pela média das distâncias entre todos os pares de genótipos para formação de cada grupo, através do método da *Distância média* (UPGMA - *Unweighted pair-group method using an arithmetic average*); e d) considerando os indivíduos que proporcionam a menor soma de quadrados dos desvios para a formação inicial do grupo, por meio do método de *Ward*. Sendo assim, cabe ao pesquisador adotar o método que melhor represente a estrutura de agrupamento esperada com base em seu conjunto de dados. Baseados em caracteres vegetativos e reprodutivos, Rao et al. (2008) utilizaram o método de Ward para conhecer a diversidade genética entre genótipos indianos de *J. curcas*. Chiorato (2004) deu preferência ao método do *Vizinho mais próximo* para identificar a diversidade genética entre acessos de feijoeiro do Banco de Germoplasma do IAC.

Embora sejam amplamente empregadas para a identificação da diversidade genética em populações e/ou entre acessos em coleções de germoplasma, Chiorato (2004) lembra que as análises de agrupamento requerem a pré-definição de medidas de distância para estimação da divergência genética e o estabelecimento da formação dos grupos. Para Bussab et al. (1990) e Azambuja (2005), a definição da medida de distância deve ser dada com muito cuidado, pois esta deve ser capaz de ressaltar as qualidades específicas explicitadas nos objetivos do estudo. Tendo a medida de distância escolhida, torna-se possível a construção da matriz de distância, de onde surgirão os grupos. Diante do exposto, Cruz e Carneiro (2003) e Azambuja (2005) enumeram as seguintes etapas a serem seguidas para proceder-se com a análise de agrupamento: a) seleção dos genótipos a serem analisados; b) obtenção e sistematização dos dados; c) definição da medida de distância a ser estimada (de similaridade ou de dissimilaridade); d) escolha do método de agrupamento a ser aplicado; e e) validação, avaliação e interpretação dos resultados obtidos.

2.3.2.4. Análises de componentes principais e variáveis canônicas

Por permitir a simplificação de um conjunto de dados, a partir do resumo das informações originalmente contidas em um grupo de n variáveis, em poucos componentes (que apresentam as propriedades de reterem o máximo da variação originalmente disponível e serem independentes entre-si), a análise de componentes principais (*Principal components analysis* - PCA) tem sido largamente empregada no melhoramento genético vegetal (CRUZ & CARNEIRO, 2003). Segundo estes autores, a análise de componentes principais baseia-se apenas nas informações individuais de cada acesso, não necessitando de dados com repetições.

Assim como a técnica de componentes principais, a análise de variáveis canônicas também permite a simplificação no conjunto de dados estudado, resumindo suas informações em poucas variáveis capazes de reterem o máximo da variação originalmente disponível e serem independentes entre si. No entanto, em harmonia com Amaral Júnior (1999), Cruz e Carneiro (2003) atentam para o fato que esta análise baseia-se em informações intra e entre acessos (ou entre indivíduos de cada acesso), havendo necessidade de dados com repetições. Dessa maneira, quando há repetição dos dados, independente do tipo de variável, o uso das variáveis canônicas permite, com elevada robustez, extrair informações sobre a diversidade genética presente. Por outro lado, quando não existem repetições, a análise de componentes principais apresenta-se como a alternativa analítica mais conveniente (AMARAL JÚNIOR, 1999).

2.3.2.5. Análise de Trilha

O entendimento das relações entre um conjunto de caracteres é de grande valia para o programa de melhoramento, uma vez que a obtenção de ganhos genéticos e a definição dos melhores genótipos são, em grande parte, direcionadas a um grupo de características agronômicas desejáveis (FERREIRA et al., 2007a). Nesse sentido, o estudo das correlações constitui-se em uma importante estratégia de análise. Segundo Carvalho et al. (1999), seu conhecimento possibilita a seleção indireta de uma característica desejável (ex.: produção) através da seleção direta sobre uma característica de fácil mensuração, e de preferência, de alta herdabilidade. Para Ramalho et al. (1993), a correlação permite ao pesquisador conhecer as modificações que ocorrem em um determinado caráter em função da seleção praticada em outro correlacionado a ele.

Embora seja intensamente empregada como um parâmetro para a seleção indireta, Coimbra et al. (2005) alertam para o fato de que a correlação não é uma medida real de causa e efeito. Dessa forma, a interpretação direta de suas magnitudes pode resultar em equívocos na estratégia de seleção, visto que uma alta correlação entre duas variáveis pode ser resultado do(s) efeito(s) de outra(s) sobre estas. Com o intuito de superar esta limitação, e entender melhor as causas envolvidas entre as associações, Wright (1921) propôs uma técnica denominada análise de trilha (*Path analysis*), capaz de desdobrar as correlações estimadas em efeitos diretos e indiretos de caracteres sobre uma variável básica (também chamada de *principal* ou *dependente*), ampliando a compreensão das relações de causa e efeito entre as mesmas. Portanto, esta metodologia pode ser aplicada a partir de correlações fenotípicas, genotípicas ou ambientais.

De acordo com Cruz e Carneiro (2003), a análise de trilha permite, essencialmente, avaliar se a relação entre dois caracteres é de causa e efeito ou determinada pela influência de outras variáveis. Sob esta perspectiva, para que a avaliação dos efeitos diretos e indiretos promova uma interpretação confiável e apropriada, é de fundamental importância que se teste o grau de colinearidade entre as variáveis independentes, visto que a estimação dos coeficientes de trilha pode ser adversamente afetada pelos efeitos da multicolinearidade (COIMBRA et al., 2005). Em tais condições, Cruz e Carneiro (2003) e Coimbra et al. (2005) afirmam que estas estimativas podem assumir valores absurdos ou sem nenhuma coerência com o fenômeno biológico estudado.

2.4. Melhoramento Genético de *Jatropha curcas* L.

Ainda não se conhecem variedades melhoradas ou cultivares de pinhão-mansão, no entanto, grande parte dos países interessados na cultura está prospectando sua diversidade genética e realizando coleta de materiais ao redor do mundo, visando enriquecer sua coleção de germoplasma para fins carburantes, alimentares e/ou medicinais. Apesar da carência de informações técnicas relativas ao seu plantio e procedência genética, a espécie vem sendo difundida e implantada em diversas regiões do Brasil (LAVIOLA, 2010).

Para Marques e Ferrari (2008), a chave para o sucesso de um programa de melhoramento genético está em uma adequada variabilidade genética e na avaliação de acessos divergentes com características desejáveis, como alta produtividade de sementes, resistência a pragas e doenças, e uniformidade e precocidade de maturação. De modo a auxiliar futuros programas de melhoramento genético da espécie em questão, Paiva Neto et al. (2009) constataram que alguns fatores característicos desta merecem atenção, como a necessidade do aumento do número de flores femininas por inflorescência, a elevação do teor de óleo na semente, maior sincronia na diferenciação floral e consequente antese por inflorescência. No entanto, a caracterização molecular do Banco de Germoplasma de *J.*

curcas da EMBRAPA revelou que os materiais coletados em diversos locais do território brasileiro possuem base genética estreita, provavelmente causada por ancestralidade comum em sua difusão em séculos passados (ROSADO et al., 2010). Visando alcançar um maior ganho genético a longo prazo, Laviola (2010) alerta que, talvez, seja necessária a introdução de materiais oriundos dos prováveis centro de origem e/ou de diversidade da cultura.

O uso da variabilidade genética pré-existente de *Jatropha curcas* está vinculado ao estudo de sua extensão no germoplasma disponível utilizando-se descritores morfológicos e, também, marcadores moleculares, tais como o RAPD e o AFLP (BASHA & SUJATHA, 2007). Santos et al. (2010a) e Redondo (2011) aplicaram o marcador molecular AFLP para analisar a diversidade genética entre genótipos de pinhão-manso de diferentes lugares do Brasil. Não obstante, outros marcadores moleculares podem ser igualmente utilizados para detectar a variabilidade genética em *J. curcas*, como os Microssatélites, empregados por Rosado et al. (2010), e o ISSR, por Chen et al. (2011). Com este mesmo fim, Basha e Sujatha (2007) realizaram a caracterização molecular de genótipos indianos de *J. curcas* via marcadores RAPD.

Em estudo sobre a divergência genética em *J. curcas*, via marcadores moleculares RAPD e SSR, Rosado et al. (2010) revelaram a existência de uma estreita base genética entre os 192 genótipos caracterizados. Tal resultado corrobora com os trabalhos de Basha e Sujatha (2007) e Rao et al. (2008), que, além de observarem baixa diversidade genética, identificaram a ocorrência de genótipos de diferentes procedências reunidas em um mesmo grupo, concluindo que a diferença geográfica nem sempre implica na ocorrência de divergência genética entre o material avaliado. Todavia, ao estudar a variabilidade genética entre 10 genótipos de pinhão-manso na fase juvenil, por meio de descritores morfológicos, Abreu et al. (2009) observaram que a divergência genética entre os grupos formados pode estar relacionada com as origens geográficas das plantas. Santos et al. (2010a) utilizaram marcadores moleculares AFLP para identificar a similaridade genética entre acessos de *J. curcas* e também constataram que a variabilidade genética entre estes era geneticamente dada em função da origem. Tamanha discordância entre os resultados apresentados pode ser atribuída à falta de padronização e detalhamento nas caracterizações procedidas até o momento, bem como ao emprego de metodologias estatísticas inadequadas para estimação da real divergência genética existente no material estudado.

Grande parte dos estudos de avaliação do germoplasma disponível de pinhão-manso tem sido realizada utilizando materiais coletados de plantas de diferentes regiões, idades e modo de propagação (sementes e estacas). Comparações de características relacionadas à produtividade destes genótipos podem gerar conclusões errôneas em relação à superioridade dos materiais identificados, pois são atributos altamente influenciados pelo ambiente. Desse modo, tais avaliações devem ser realizadas em germoplasmas submetidos às mesmas condições edafoclimáticas e agronômicas. Posteriormente, a superioridade dos acessos identificados pode ser confirmada através de ensaios de validação em diferentes localidades (JARAMILLO & BAENA, 2000).

Atualmente, diversos centros de pesquisa têm realizado programas de melhoramento genético para a espécie, como a EMBRAPA, dentro do Projeto BRJATROPHA, e a Universidade Federal do Ceará (UFC), Universidade Federal Rural do Semiárido (UFERSA), Universidade Federal da Bahia (UFBA), Empresa de Pesquisa Agropecuária de Minas Gerais (EPAMIG), Universidade Federal de Campina Grande (UFGC), e o Instituto Agrônomo de Pernambuco (IPA), juntamente à Universidade Federal Rural do Rio de Janeiro, através do Projeto PetroBio Girassol (P BIO Girassol), financiado pela PETROBRAS. Estes projetos visam o desenvolvimento e a validação de tecnologias de cultivo no que se refere ao melhoramento genético, à produção de sementes e mudas, nutrição mineral, espaçamentos, controle de pragas e doenças, e manejo da cultura.

Perante a necessidade de desenvolvimento de variedades superiores de *Jatropha curcas*, com vistas à produção de biodiesel, Marques e Ferrari (2008) ressaltam a importância da integração das técnicas de biotecnologia ao programa de melhoramento genético convencional, visto que juntas poderão auxiliar no desenvolvimento de cultivares para a espécie. Sendo uma cultura perene e em fase de domesticação para o cultivo comercial, estima-se que ainda serão necessários mais alguns anos para que se obtenham as primeiras cultivares melhoradas, bem como informações cientificamente embasadas sobre o sistema de produção para as diversas regiões do Brasil (DURÃES et al., 2009; LAVIOLA, 2010).

3. MATERIAL E MÉTODOS

3.1. Localização do Experimento e Condições Edafoclimáticas

O presente trabalho foi conduzido no campo experimental do Departamento de Fitotecnia, do Instituto de Agronomia da Universidade Federal Rural do Rio de Janeiro (DFITO/IA/UFRRJ), situada no município de Seropédica (22°45' S; 43° 41' W) (RAMOS et al., 1973), região metropolitana do Estado do Rio de Janeiro, entre os dias 01/09/2011 e 28/05/2012. A área experimental cultivada se encontra ao nível do mar e o solo à classe dos Planossolos Háplicos, de textura arenosa de baixada. Segundo a classificação de Köppen, o clima da região enquadra-se no tipo Aw, sendo caracterizado pela alternância entre a estação chuvosa no verão e seca durante o inverno, com precipitação e temperatura anual média na ordem de 1213 mm e 24,5 °C (CARVALHO et al., 2006). Ao longo dos nove meses de avaliação, o município de Seropédica, RJ, apresentou uma temperatura média de 23,87 °C e uma precipitação mensal de 91,40 mm (Tabela 1).

Tabela 1. Precipitação acumulada, valores médios mensais de umidade relativa, velocidade do vento e temperatura, e amplitude térmica em Seropédica, RJ, no período de setembro de 2011 a maio de 2012.

Mês e ano	Precipitação (mm)	Umidade (%)	Velocidade do vento (m/s ²)	Temperatura (°C)	Amplitude térmica (°C)
Setembro de 2011	20,60	75,05	2,70	21,05	10,50-36,30
Outubro de 2011	104,60	78,72	2,56	22,99	12,70-38,50
Novembro de 2011	103,60	79,63	2,40	22,59	13,40-35,60
Dezembro de 2011	132,80	80,33	2,26	24,87	18,60-37,30
Janeiro de 2012	298,00	81,53	2,25	24,98	18,10-35,50
Fevereiro de 2012	42,60	72,58	2,36	27,21	19,40-38,00
Março de 2012	0,00	80,31	2,20	25,48	18,80-38,50
Abril de 2012	64,60	84,77	1,90	24,27	17,80-35,30
Mai de 2012	55,80	83,91	1,75	21,42	13,50-33,60
Média	91,40*	79,65	2,27	23,87	-

* Calculada a partir de dados relativos à precipitação acumulada mensal. Fonte: INMET/SIPA/UFRRJ.

3.2. Germoplasma e Condições Experimentais

Os dados foram coletados em 60 acessos pertencentes à Coleção de germoplasma de *Jatropha curcas* do DFITO/IA/UFRRJ. Atualmente, esta coleção de germoplasma conta com 286 plantas, de 13 diferentes procedências, plantados em abril de 2009, em espaçamento 4 x 2 m, sendo todos oriundos de polinização aberta. Todas as 286 plantas foram dispostas em linhas, conforme procedência e sem repetições, não havendo delineamento estatístico definido, visto que cada procedência possuía um número distinto de genótipos. Desse modo, cada planta foi tratada como um acesso. Devido ao elevado número de descritores, foram caracterizados 60 acessos, de 12 diferentes procedências, conhecidamente mais produtivos aos três anos de idade, presentes nesta (Tabela 2).

Tabela 2. Procedência, identificação e número de acessos por procedência dos 60 acessos de *Jatropha curcas* da Coleção de germoplasma da UFRRJ definidos para caracterização.

Procedência	Identificação dos acessos UFRRJ	Número de acessos
Janaúba - MG	860 / 869 / 874 / 872 / 346 / 858 / 875 / 878 / 863 / 890 / 864 / 861 / 385 / 876 / 059 / 856 / 857	17
Dourados - MS	825 / 807 / 900 / 991 / 911 / 337 / 393 / 908 / 932 / 955 / 905 / 910 / 824 / 820	14
Petrolina - PE	002 / 376 / 912 / 828 / 931 / 831 / 849 / 024	8
CNPS - MA	371 / 356 / 058 / 331	4
Lavras - Lote 000 - MG	399 / 332 / 005 / 330	4
Lavras - Lote 210 - MG	357 / 345 / 328	3
Rio das Flores - RJ	321 / 343	2
Rio Verde - GO	383 / 053	2
Barra dos Bugres - MT	003 / 344	2
Bom Jardim - RJ	355 / 386	2
Guapimirim - RJ	341	1
Jaíba - MG	062	1

O preparo do campo experimental para o plantio iniciou-se em novembro de 2008, quando foi realizada a limpeza da área. Para tal, utilizou-se roçadeira e aração, seguida de gradagem, para incorporação dos restos vegetais, visando à melhoria das propriedades químicas, físicas e biológicas do solo. Com base em análise de solo, realizada no Laboratório de Fertilidade da UFRRJ (LABFER), foram efetuadas a calagem e a adubação química da área. Para esta análise, foram coletadas amostras de solo nas profundidades de 0-10 cm e de 10-20 cm. Dessa forma, em fevereiro de 2009, foram procedidas nova aração e gradagem para incorporação do calcário. Em março de 2009, com o auxílio de um encanteirador, foram formados canteiros de 1m de largura. Ainda em março de 2009, a área foi adubada com 200 g de NPK, formulação 4-31-4 e 0,4% de Zn (por cova), sendo 18% de P₂O₅ solúvel em ácido cítrico.

As mudas (com idade de 6-7 meses) foram transferidas para o campo em abril de 2009, onde foram transplantadas para covas com dimensões de 40 x 40 x 40 cm. Uma nova adubação foi realizada na primeira quinzena de novembro de 2010, quando foram adicionados 100 g.planta⁻¹ da formulação 10-10-10 de NPK.

3.3. Desenvolvimento, Validação e Normatização de Descritores Morfoagronômicos

A partir dos 60 acessos definidos, uma lista preliminar de descritores morfoagronômicos para *Jatropha curcas* foi elaborada tendo como principal base as recomendações sugeridas pelo *Guidelines for the development of crop descriptor lists* (BIOVERSITY INTERNATIONAL, 2007). Conforme recomendado, cada descritor foi composto por três elementos essenciais para aplicação prática deste no campo, sendo eles: um *nome*, capaz de descrever o atributo de forma simples e compacta; um *estado*, capaz de expressar a característica do atributo observado; e um *método* de mensuração e coleta que, assim como os elementos já mencionados, deve ser claro e de fácil entendimento.

Para validação, os descritores desenvolvidos foram aplicados de acordo com as diferentes fases de desenvolvimento da planta. Desse modo, as avaliações foram divididas em duas etapas: uma para a aplicação dos descritores referentes aos aspectos vegetativos e outra relativa aos descritores de seus aspectos reprodutivos.

Baseada nesta divisão, os descritores foram normatizados de acordo com o tipo de variável, classificando-os em variáveis quantitativas, multicategóricas e binárias (Tabela 3).

Tabela 3. Lista preliminar de descritores morfoagronômicos desenvolvidos para *Jatropha curcas*: tipo de variável, nome do descritor, metodologia e estado.

Tipo de variável	Nome do descritor	Metodologia	Estado
<i>1. Descritores vegetativos</i>			
Quantitativas	1.1. Altura de plantas desenvolvidas (APD)	Avaliada a partir do final do 1º (primeiro) ciclo reprodutivo da planta. A altura de plantas desenvolvidas (em cm) deve ser medida entre a base do caule (solo) e a extremidade superior do ramo mais alto.	-
	1.2. Projeção da copa no sentido Norte (PCOn)	Avaliada a partir do final do 1º (primeiro) ciclo reprodutivo da planta. A projeção da copa (em m) deve ser mensurada a partir da região central da base do caule até a projeção máxima do ramo no sentido Norte.	-
	1.3. Projeção da copa no sentido Sul (PCOs)	Avaliada a partir do final do 1º (primeiro) ciclo reprodutivo da planta. A projeção da copa (em m) deve ser mensurada a partir da região central da base do caule até a projeção máxima do ramo no sentido Sul.	-
	1.4. Projeção da copa no sentido Leste (PCOI)	Avaliada a partir do final do 1º (primeiro) ciclo reprodutivo da planta. A projeção da copa (em m) deve ser mensurada a partir da região central da base do caule até a projeção máxima do ramo no sentido Leste.	-
	1.5. Projeção da copa no sentido Oeste (PCOo)	Avaliada a partir do final do 1º (primeiro) ciclo reprodutivo da planta. A projeção da copa (em m) deve ser mensurada a partir da região central da base do caule até a projeção máxima do ramo no sentido Oeste.	-
	1.6. Área de projeção da copa (PCO)	Avaliada a partir do final do 1º (primeiro) ciclo reprodutivo da planta. A área de projeção da copa (em m ²) deve ser obtida através da seguinte fórmula: $[(PCOn + PCOs) \times (PCOI + PCOo)] / 2$, onde os descritores PCOn, PCOs, PCOI e PCOo, previamente citados, corresponderão aos valores de projeção nos sentidos Norte, Sul, Leste e Oeste, respectivamente. Para tal, cada valor de projeção (em m) deve ser mensurado a partir da região central da base do caule até a projeção máxima do ramo em cada sentido.	-
	1.7. Diâmetro do caule (DCA)	Avaliado em qualquer fase do ciclo de desenvolvimento da planta. O diâmetro do caule (em cm) deve ser medido na região do colo da planta a uma altura máxima de 10 cm da base do caule (solo).	-
	1.8. Número de ramos totais (NRT)	Avaliado em qualquer fase do ciclo de desenvolvimento da planta. O número de ramos totais deve ser obtido pela	-

	quantificação do número total de ramos da planta.	
1.9. Número de ramos com inflorescência (NRI)	Avaliado a partir do 1º (primeiro) ciclo reprodutivo da planta. O número de ramos com inflorescência deve ser obtido pela quantificação do número total de ramos que culminam em uma inflorescência.	-
1.10. Comprimento de ramos primários (CRP)	Avaliado a partir do 1º (primeiro) ciclo reprodutivo da planta. O comprimento de ramos primários deve ser obtido pela média das medidas efetuadas entre a base do ramo (extremidade inferior) e sua extremidade superior nos 3 ramos primários de maior comprimento. A média do comprimento de ramos primários por planta (em cm) deve ser dada pela divisão da soma de todas as medidas aferidas pelo número total de ramos primários avaliados. Para plantas de arquitetura ereta e sem ramificações, o comprimento de ramos primários deve ser dado pela própria altura da planta, mensurada entre a base do caule (solo), nesse caso, considerado o único ramo primário, e sua extremidade superior.	-
1.11. Comprimento do entrenó de ramos primários (CERp)	Avaliado a partir do 1º (primeiro) ciclo reprodutivo da planta. O comprimento do entrenó de ramos primários deve ser obtido pela média das medidas efetuadas nos 3 ramos primários de maior comprimento, avaliando-se 6 entrenós na porção mediana de cada ramo primário. A média do comprimento do entrenó de ramos primários por planta (em cm) deve ser dada pela divisão da soma de todos os entrenós mensurados pelo número total de entrenós avaliados. Para plantas de arquitetura ereta e sem ramificações, o comprimento do entrenó de ramos primários deve ser dado pela própria média das medidas aferidas dos 6 entrenós da porção mediana do caule da planta, nesse caso, considerado o único ramo primário.	-
1.12. Comprimento de ramos com inflorescência (CRI)	Avaliado a partir do 1º (primeiro) ciclo reprodutivo da planta. O comprimento de ramos com inflorescência deve ser obtido pela média das medidas efetuadas entre a base do ramo (extremidade inferior) e sua extremidade superior nos 3 ramos de maior comprimento com inflorescência. A média do comprimento de ramos com inflorescência por planta (em cm) deve ser dada pela divisão da soma de todas as medidas aferidas pelo número total de ramos com inflorescência avaliados. Para plantas de arquitetura ereta e sem ramificações, o comprimento de ramos com inflorescência deve ser dado pela própria altura	-

	da planta, mensurada entre a base do caule (solo), nesse caso, considerado o único ramo com inflorescência, e sua extremidade superior.	
1.13. Comprimento do entrenó de ramos com inflorescência (CERi)	Avaliado a partir do 1º (primeiro) ciclo reprodutivo da planta. O comprimento do entrenó de ramos com inflorescência deve ser obtido pela média das medidas efetuadas nos 3 ramos de maior comprimento com inflorescência, avaliando-se 4 entrenós na porção mediana de cada ramo com inflorescência. A média do comprimento do entrenó de ramos com inflorescência por planta (em cm) será dada pela divisão da soma de todos os entrenós mensurados pelo número total de entrenós avaliados. Para plantas de arquitetura ereta e sem ramificações, o comprimento do entrenó de ramos com inflorescência será dado pela própria média das medidas aferidas dos 4 entrenós da porção mediana do caule da planta, nesse caso, considerado o único ramo com inflorescência.	-
1.14. Número de lóbulos foliar (NLF)	Avaliado em qualquer fase do ciclo de desenvolvimento da planta. O número de lóbulos foliar deve ser obtido pela quantificação dos lóbulos em 10 folhas completamente desenvolvidas por planta.	-
1.15. Ângulo de inserção da folha no caule (AIFc)	Avaliado em qualquer fase do ciclo de desenvolvimento da planta. O ângulo de inserção da folha no caule deve ser obtido pela média das medidas entre a inserção da folha (pecíolo) e o caule em 10 folhas completamente desenvolvidas por planta. A média do ângulo de inserção da folha no caule (em graus) deve ser dada pela divisão do valor da soma de todos os ângulos aferidos pelo número total de folhas avaliadas.	-
1.16. Comprimento do pecíolo (CMP)	Avaliado em qualquer fase do ciclo de desenvolvimento da planta. O comprimento do pecíolo deve ser obtido pela média do comprimento do pecíolo em 10 folhas completamente desenvolvidas por planta. A média do comprimento do pecíolo (em cm) deve ser dada pela divisão do valor da soma do comprimento de todos os pecíolos mensurados pelo número total de folhas avaliadas.	-
1.17. Comprimento da folha (CFO)	Avaliado em qualquer fase do ciclo de desenvolvimento da planta. O comprimento da folha deve ser obtido pela média do comprimento em 10 folhas completamente desenvolvidas por planta. A média do comprimento das folhas (em cm) deve ser dada pela divisão do valor da soma do comprimento de todas as	-

		folhas mensuradas pelo número total de folhas avaliadas.	
	1.18. Largura da folha (LFO)	Avaliada em qualquer fase do ciclo de desenvolvimento da planta. A largura da folha deve ser obtida pela média da largura em 10 folhas completamente desenvolvidas por planta. A média da largura das folhas (em cm) deve ser dada pela divisão do valor da soma da largura de todas as folhas mensuradas pelo número total de folhas avaliadas.	-
	1.19. Relação entre o comprimento e a largura da folha (CLFo)	Avaliada em qualquer fase do ciclo de desenvolvimento da planta. Relação obtida através da razão entre o valor do <i>comprimento da folha</i> pelo valor da <i>largura da folha</i> (ambos citados anteriormente).	-
	1.20. Arquitetura da planta (ARQ)	Observada a partir do final do 1º (primeiro) ciclo reprodutivo da planta.	1- Ereta; 2- Fechada; 3- Aberta
	1.21. Tipo de ramificação (RAM)	Observada a partir do final do 1º (primeiro) ciclo reprodutivo da planta.	0- Ausência de ramificação; 1- Monopodial; 2- Dicasial; 3- Simpodial
	1.22. Coloração do caule (CCA)	Observada a partir do final do 1º (primeiro) ciclo reprodutivo da planta. A coloração do caule deve ser observada na região mediana deste, desconsiderando a presença de cerosidade.	1- Verde; 2- Verde escuro; 3- Cinza; 4- Roxo; 5- Azul
	1.23. Coloração do ramo (CRA)	Observada a partir do final do 1º (primeiro) ciclo reprodutivo da planta. A coloração do ramo deve ser observada desconsiderando a presença de cerosidade na ramificação.	1- Verde 2- Verde escuro; 3- Cinza; 4- Roxo; 5- Azul
Multicategóricas	1.24. Coloração de folhas jovens (CFJ)	Observada em qualquer fase do ciclo de desenvolvimento da planta. A coloração de folhas jovens deve ser observada na face superior (adaxial) do limbo foliar em uma amostra representativa de 10 folhas jovens por planta.	1- Verde claro; 2- Verde; 3- Verde escuro; 4- Roxa; 5- Vermelha
	1.25. Coloração de folhas desenvolvidas (CFD)	Observada em qualquer fase do ciclo de desenvolvimento da planta. A coloração de folhas desenvolvidas deve ser observada na face superior (adaxial) do limbo foliar em uma amostra representativa de 10 folhas completamente desenvolvidas por planta.	1- Verde claro; 2- Verde; 3- Verde escuro; 4- Roxa; 5- Vermelha
	1.26. Coloração da nervura foliar (CNF)	Observada em qualquer fase do ciclo de desenvolvimento da planta. A coloração da nervura foliar deve ser observada na face superior (adaxial) do limbo em uma amostra representativa de 10 folhas completamente desenvolvidas por planta.	1- Verde; 2- Roxa; 3- Vermelha
	1.27. Coloração do pecíolo (CPE)	Observada em qualquer fase do ciclo de desenvolvimento da planta. A coloração do pecíolo deve ser observada em uma amostra representativa de 10 folhas completamente	1- Verde; 2- Roxo-esverdeado; 3- Roxo; 4- Vermelho

		desenvolvidas por planta.	
Binárias	1.28. Cerosidade (CER)	Observada a partir do final do 1º (primeiro) ciclo reprodutivo da planta. A planta deve ser classificada quanto à presença ou ausência de cerosidade no caule e demais ramificações.	0- Ausência; 1- Presença
	1.29. Pilosidade da folha (PFO)	Observada em qualquer fase do ciclo de desenvolvimento da planta. A pilosidade da folha deve ser observada quanto a presença ou a ausência de pêlos na porção adaxial e abaxial em uma amostra representativa de 10 folhas completamente desenvolvidas por planta.	0- Ausência; 1- Presença
<i>2. Descritores reprodutivos</i>			
Quantitativas	2.1. Altura da primeira inflorescência (API)	Avaliada assim que se inicia o ciclo reprodutivo da planta. A altura da primeira inflorescência (em cm) deve ser dada pela medida compreendida entre a base do caule (solo) e a inflorescência mais baixa (primeira inflorescência) da planta.	-
	2.2. Comprimento do pedúnculo da inflorescência (CPI)	Avaliado a partir da 1º (primeira) frutificação. O comprimento do pedúnculo (em cm) deve ser obtido pela média do comprimento do pedúnculo em 10 inflorescências por planta. A média do comprimento do pedúnculo deve ser dada pela divisão do valor da soma do comprimento de todos os pedúnculos mensurados pelo número total de inflorescências avaliadas.	-
	2.3. Número total de inflorescências (NTI)	Avaliado a partir do 1º (primeiro) ciclo reprodutivo. O número total de inflorescências por planta deve ser obtido pela quantificação de todas as inflorescências na planta. Para tal, deve-se realizar a contagem de todas as inflorescências da planta ao longo de seu ciclo reprodutivo.	-
	2.4. Uniformidade de maturação do cacho por planta (UMC)	Avaliado a partir do 1º (primeiro) ciclo reprodutivo da planta. A uniformidade de maturação do cacho deve ser obtida pela média do número total de frutos amarelos (maduros) em relação ao número total de frutos em 5 cachos por planta. A média da uniformidade de maturação do cacho (em %) deve ser dada pela divisão do valor da soma de todos os frutos amarelos (maduros) do cacho pelo número total de frutos verdes, amarelos (maduros) e pretos (secos) presentes no cacho. A uniformidade de maturação do cacho por planta deve ser dada pela divisão do valor da soma da média de uniformidade de maturação de todos os cachos quantificados pelo número total de cachos avaliados.	-
	2.5. Tempo de formação do fruto (TFF)	Avaliado a partir do 1º (primeiro) ciclo reprodutivo da planta. O tempo de formação do fruto deve ser obtido pela quantificação dos dias decorridos entre a antese floral e a	-

	formação do primeiro fruto amarelo (maduro) da planta.	
2.6. Número de frutos por planta (NFR)	Avaliado a partir do 1º (primeiro) ciclo reprodutivo da planta. O número de frutos por planta deve ser obtido pela quantificação do número total de frutos produzidos pela planta.	-
2.7. Peso de frutos por planta (PFR)	Avaliado a partir do 1º (primeiro) ciclo reprodutivo da planta. O peso de frutos por planta deve ser obtido pela média do peso de todos os frutos em relação ao número total de frutos produzidos. A média do peso de frutos por planta (em g) deve ser dada pela divisão do valor da soma do peso de todos os frutos, pesados na fase amarela de maturação, pelo número total de frutos produzidos.	-
2.8. Comprimento do fruto (CFR)	Avaliado a partir do 1º (primeiro) ciclo reprodutivo da planta. O comprimento do fruto deve ser obtido pela média do comprimento de 50 frutos amarelos (maduros) de diferentes ramos da planta. A média do comprimento de frutos (em cm) deve ser dada pela divisão do valor da soma do comprimento de todos os frutos mensurados pelo número total de frutos avaliados.	-
2.9. Largura do fruto (LFR)	Avaliada a partir do 1º (primeiro) ciclo reprodutivo da planta. A largura média do fruto deve ser obtida pela média da largura de 50 frutos amarelos (maduros) de diferentes ramos da planta. A média da largura de frutos (em cm) deve ser dada pela divisão do valor da soma da largura de todos os frutos mensurados pelo número total de frutos avaliados.	-
2.10. Relação entre o comprimento e a largura do fruto (CLFr)	Avaliada a partir do 1º (primeiro) ciclo reprodutivo da planta. Relação obtida através da razão entre o valor do <i>comprimento do fruto</i> pelo valor da <i>largura do fruto</i> (ambos citados anteriormente).	-
2.11. Número de sementes por fruto (NSF)	Avaliado a partir do 1º (primeiro) ciclo reprodutivo da planta. O número de sementes por fruto deve ser obtido pela razão entre o valor do <i>número de sementes por planta</i> (citado posteriormente) pelo <i>número de frutos por planta</i> (citado anteriormente).	-
2.12. Número de sementes por planta (NSE)	Avaliado a partir do 1º (primeiro) ciclo reprodutivo da planta. O número de sementes por planta deve ser obtido pela quantificação do número total de sementes produzidas pela planta.	-
2.13. Peso médio da semente seca (PMSs)	Avaliado a partir do 1º (primeiro) ciclo reprodutivo da planta. O peso médio da semente seca deve ser obtido pela média do	-

		peso de todas as sementes em relação ao número total de sementes produzidas. A média do peso de sementes secas por planta (em g) deve ser dada pela divisão do valor da soma do peso de todas as sementes, pesadas com baixo teor de umidade (após secagem durante 3 dias a uma temperatura de 35°C em estufa), pelo número total de sementes produzidas.	
	2.14. Comprimento da semente (CMS)	Avaliado a partir do 1º (primeiro) ciclo reprodutivo da planta. O comprimento da semente deve ser obtido pela média do comprimento de 50 sementes por planta. A média do comprimento de sementes (em cm) deve ser dada pela divisão do valor da soma do comprimento de todas as sementes mensuradas pelo número total de sementes avaliadas.	-
	2.15. Largura da semente (LSE)	Avaliada a partir do 1º (primeiro) ciclo reprodutivo da planta. A largura da semente deve ser obtida pela média da largura de 50 sementes por planta. A média da largura de sementes (em cm) deve ser dada pela divisão do valor da soma da largura de todas as sementes mensuradas pelo número total de sementes avaliadas.	-
	2.16. Relação entre o comprimento de sementes e largura de sementes (CLSe)	Avaliada a partir do 1º (primeiro) ciclo reprodutivo da planta. Relação obtida através da razão entre o valor do <i>comprimento da semente</i> pelo valor da <i>largura da semente</i> (ambos citados anteriormente).	-
	2.17. Produção de grãos por planta (PGP)	Avaliada a partir do 1º (primeiro) ciclo reprodutivo da planta. A produção de grãos por planta (em g.planta ⁻¹) deve ser obtida pela multiplicação entre o valor do <i>número de sementes por planta</i> pelo valor do <i>peso médio da semente seca</i> (ambos citados anteriormente).	-
Multicategóricas	2.18. Formato do fruto (FFR)	Observado a partir do 1º (primeiro) ciclo reprodutivo da planta. O formato do fruto deve ser observado em 50 frutos amarelos (maduros).	1- Elíptico; 2- Oval; 3- Triangular
	2.19. Número de lóculos por fruto (NLFr)	Avaliado a partir do 1º (primeiro) ciclo reprodutivo da planta. O número de lóculos por fruto deve ser observado em 50 frutos amarelos (maduros).	1- Unilocular; 2- Bilocular; 3- Trilocular; 4- Tetralocular
	2.20. Coloração da semente (CSE)	Observada em qualquer fase do ciclo de desenvolvimento da planta. A coloração da semente deve ser observada em uma amostra representativa de 50 sementes por planta.	1- Marrom; 2- Preta; 3- Roxa; 4- Vermelha
	2.21. Textura da semente (TSE)	Observada em qualquer fase do ciclo de desenvolvimento da planta. A textura da semente deve ser observada em uma amostra representativa de 50 sementes por planta.	1- Lisa; 2- Áspera; 3- Rugosa

	2.22. Padrão da semente (PSE)	Observada em qualquer fase do ciclo de desenvolvimento da planta. O padrão da semente deve ser observado em uma amostra representativa de 50 sementes por planta.	1- Cor única; 2- Pintada; 3- Pontuada; 4- Rajada
Binárias	2.23. Carúncula da semente (CAR)	Observada em qualquer fase do ciclo de desenvolvimento da planta. A carúncula da semente deve ser observada em uma amostra representativa de 50 sementes por planta.	0- Ausência; 1- Presença

3.4. Análises Estatísticas

3.4.1. Importância e descarte de variáveis via análise de componentes principais

Objetivando o reconhecimento dos descritores que mais e menos contribuíram para a variabilidade investigada, uma análise preliminar desta foi realizada através da técnica de componentes principais. Primeiramente, a partir do registro das 36 variáveis quantitativas, uma análise descritiva auxiliou esta identificação. Nesta, seus valores máximos, mínimos e médios, bem como sua variância, desvio-padrão e coeficiente de variação, ajudaram na detecção dos caracteres que mais e menos variaram.

Posteriormente à análise descritiva, prosseguiu-se com a análise de componentes principais. Para tal, seguindo as recomendações de Cruz e Carneiro (2003), foi realizada a padronização dos dados originais, em virtude das diferentes unidades de medida aqui utilizadas. Na padronização, a média original X_{ij} , obtida pelo descritor j do acesso i , foi dividida pelo desvio-padrão (S_j) do correspondente descritor j , gerando a média reduzida Z_{ij} , de variância unitária:

$$Z_{ij} = \frac{X_{ij}}{S_j}$$

Conforme Strapasson et al. (2000), cada componente é formado por uma combinação linear de todas as variáveis originais, sendo independentes entre-si e construídos de forma a explicar o máximo da variabilidade total destas. Desse modo, a partir da matriz de correlação R , a análise de componentes principais consiste em transformar um conjunto de v variáveis Z_1, Z_2, \dots, Z_v , pertencentes a um n conjunto de indivíduos (ou populações), em um novo conjunto de variáveis Y_1 (CP 1), Y_2 (CP 2), ..., Y_v (CP V) (CRUZ & CARNEIRO, 2003). Sendo assim, o número de componentes principais se dá de acordo com o número de variáveis registradas. Segundo Cruz e Carneiro (2003), se Y_{il} é um componente principal, então, este é uma combinação linear das variáveis Z_i 's:

$$Y_{il} = a_1 Z_{i1} + a_2 Z_{i2} + \dots + a_v Z_{iv}$$

Logo, se $Y_{i2}, Y_{i3}, \dots, Y_{iv}$ são componentes principais, outras combinações lineares também serão formadas, assim como Y_{i1} , onde, entre todos os componentes, este apresenta a maior variância, Y_{i2} a segunda maior e assim sucessivamente. Dito isto, tem-se, na presente dissertação, que a técnica de componentes principais foi aplicada duas vezes. Em primeiro momento, após proceder-se com a padronização, para o descarte dos descritores menos discriminantes, e em seguida, dentre os restantes, na identificação dos mais importantes para a variabilidade genética encontrada.

Segundo Jolliffe (1972 e 1973), quando se dispõe de um elevado número de descritores, existe a possibilidade de que muitos deles sejam redundantes, logo, de baixo valor informativo. Para Paiva et al. (2010), além da difícil interpretação, esta situação não aumenta a precisão da análise, tornando-a ainda mais complexa. Portanto, segundo estes autores, variáveis de baixa discriminação podem ser eliminadas. De acordo com Bertini et al. (2010), estudos com variáveis padronizadas, onde autovalores são obtidos a partir da matriz de correlação R , têm-se mostrado comuns em descartar o descritor de maior coeficiente (em valor absoluto) a partir do último componente, até aquele cujo autovalor não exceda 0,70, seguindo, assim, as recomendações de Jolliffe (1972 e 1973). Em consonância, Cruz e Carneiro (2003) observam que variáveis relativas a autovetores de maiores coeficientes, associadas aos componentes de autovalor até 0,70, devem ser as menos importantes para

explicar a variância total, sendo, portanto, passíveis de descarte. Diante do exposto, optou-se, nos dados aqui trabalhados, por descartar as variáveis de baixa discriminação através do critério de Jolliffe (1972 e 1973). No entanto, somada a este critério, também utilizou-se das recomendações de Cruz e Carneiro (2003) e Chiorato (2004), avaliando, por meio de correlações simples, a importância agrônoma da característica para o programa de melhoramento.

Baseado no princípio de que a variância (ou importância) decresce do primeiro para o último componente (STRAPASSON et al., 2000), empregou-se, na presente dissertação, para identificar os descritores de maior importância, o critério de Cruz e Carneiro (2003) e Strapasson et al. (2000). Para Cruz e Carneiro (2003), os descritores de maior contribuição dizem respeito aos autovetores (coeficientes de ponderação) de elevado valor, normalmente associados a autovalores responsáveis por, no mínimo, 80% da variação explicada. Sob este aspecto, Strapasson et al. (2000) afirmam que o primeiro componente também é definido como o de maior importância, uma vez que abriga grande parte da variância encontrada nos dados originais.

Concomitante à análise de componentes principais, efetuou-se o diagnóstico de multicolinearidade. Proposto por Montgomery e Peck (1981), este teste visa a detecção da severidade provocada pelos efeitos da colinearidade sobre a matriz de correlação. Segundo Cruz e Carneiro (2003), tais efeitos podem ser advindos do grande número de variáveis avaliadas e/ou pela presença de variáveis altamente correlacionadas, caracterizando, assim, uma dependência linear entre elas. De acordo com Montgomery e Peck (1981), o grau de multicolinearidade é estabelecido através do Número de condição (NC), ou seja, entre a razão do maior e menor autovalor (λ_{max} e λ_{min}) da matriz de correlação genética. De modo prático, conforme estes autores, quando o número de condição é menor que 100, a multicolinearidade é fraca; quando se encontra entre 100 e 1000, seu grau é de moderado a forte; e quando maior que 1000, sua classificação é severa (Tabela 4). Para Carvalho et al. (1999), somente o grau fraco de multicolinearidade não constitui problema sério à análise. Contudo, Cruz e Carneiro (2003) advertem que, quando elevada, a multicolinearidade pode levar a estimativas de coeficientes de correlação e regressão instáveis, distorcendo interpretações sobre análises de trilha, regressão e índices de seleção. Desse modo, um alto grau de severidade pode levar à formação de matrizes singulares ou mal condicionadas (PAIVA et al., 2010).

Tabela 4. Classificação do grau de multicolinearidade (MONTGOMERY & PECK, 1981).

Número de condição (NC)	Multicolinearidade
$NC \leq 100$	Fraca (não constitui problema sério)
$100 < NC < 1000$	Moderada a forte
$NC \geq 1000$	Severa

Tendo em vista que variáveis de alta correlação são potencialmente descartáveis, uma vez que já são representadas por outros caracteres (STRAPASSON et al., 2000), o diagnóstico de multicolinearidade, no presente estudo, também foi empregado como um parâmetro de eliminação. Assim, paralelo ao critério de Jolliffe (1972 e 1973), via componentes principais, descritores menos discriminantes também tiveram suas correlações avaliadas pelo teste de Montgomery e Peck (1981).

Por fim, o descarte de variáveis qualitativas foi dado a partir do registro de suas classes, sejam elas binárias ou multicategóricas. Dessa maneira, descritores que não obtiveram qualquer tipo de variação, ao longo do processo de caracterização, foram descartados.

3.4.2. Correlação fenotípica e análise de trilha

Após descarte e conhecimento das características de maior importância, um estudo das correlações foi realizado via análise de trilha (WRIGHT, 1921).

Primeiramente, conforme Cruz e Carneiro (2003), estas foram submetidas a análises de correlações simples, de Pearson.

A partir da matriz de correlação gerada, realizou-se, novamente, o diagnóstico de multicolinearidade (MONTGOMERY & PECK, 1981) e, em seguida, a decomposição dos efeitos diretos e indiretos de 15 variáveis explicativas, a saber: “projeção da copa no sentido Oeste”, “diâmetro do caule”, “comprimento de ramos com inflorescência”, “comprimento do entrenó de ramos com inflorescência”, “número de lóbulos foliar”, “ângulo de inserção da folha no caule”, “comprimento de folha”, “relação entre o comprimento e a largura de folha”, “altura da primeira inflorescência”, “comprimento do pedúnculo da inflorescência”, “tempo de formação do fruto”, “largura do fruto”, “número de sementes por fruto”, “comprimento da semente” e “largura da semente”, sobre a variável principal “número de sementes por planta”, através da resolução do sistema linear $\hat{\beta} = (X'X)^{-1} X'Y$, conforme proposto por Cruz e Carneiro (2003) e Rocha et al. (2009):

$$\hat{\beta} = \begin{bmatrix} p_1 \\ p_2 \\ \dots \\ p_n \end{bmatrix}; \quad X'X = \begin{bmatrix} 1 & r_{12} & \dots & r_{1n} \\ r_{12} & 1 & \dots & r_{2n} \\ \dots & \dots & \dots & \dots \\ r_{1n} & r_{2n} & \dots & 1 \end{bmatrix}; \quad \text{e} \quad X'Y = \begin{bmatrix} r_{1y} \\ r_{2y} \\ \dots \\ r_{ny} \end{bmatrix}$$

Onde,

$\hat{\beta}$ é um vetor coluna correspondendo às estimativas dos coeficientes de trilha;

$X'X$ é uma matriz não singular das correlações entre as variáveis explicativas; e

$X'Y$ é um vetor coluna das correlações entre a variável principal e as variáveis explicativas.

3.4.3. Estimação da diversidade genética e consistência do padrão de agrupamento

De forma semelhante à análise de trilha, baseado nos descritores que mais contribuíram para a variabilidade observada, estimou-se a diversidade genética da coleção em questão. Para tal, adotando as sugestões de Bussab et al. (1990), Malhotra (2001) e Azambuja (2005), nas quais destacam que, com base na consistência do padrão de agrupamento gerado, para melhor comparação dos grupos formados, se teste as diferentes técnicas multivariadas disponíveis, avaliou-se a combinação de cinco medidas de distâncias junto a dois métodos de agrupamento. Desse modo, testou-se a combinação do coeficiente geral de Gower (Algoritmo de Gower) e das distâncias Euclidiana, de Chebyshev (*Maximum*), Manhattan (*City-block*) e de Canberra com os métodos de agrupamento hierárquico UPGMA (*Unweighted pair-group method using an arithmetic average*) e Vizinheiro mais próximo (Ligação simples) (Tabela 5). O coeficiente geral de Gower foi empregado na análise individual de variáveis multicategóricas e na análise conjunta destas com as variáveis quantitativas, ao passo que as distâncias Euclidiana, de Chebyshev, Manhattan e de Canberra prestaram-se à análise individual de variáveis quantitativas.

Tabela 5. Tipo de análise, distâncias, respectivos estimadores e agrupamentos utilizados para a estimação da diversidade genética a partir de 60 genótipos da Coleção de germoplasma de *Jatropha curcas* da UFRRJ.

Tipo de análise (variáveis)	Distância	Estimador
Multicategóricas e Multicategóricas + Quantitativas	Algoritmo de Gower*	$S_{ijk} = \frac{\sum_{k=1}^p W_{ijk} \cdot S_{ijk}}{\sum_{k=1}^p W_{ijk}}$
Quantitativas	Distância Euclidiana**	$d_e = \left\{ \sum_{i=1}^p (x_i - y_i)^2 \right\}^{\frac{1}{2}}$
	Distância de Chebyshev**	$d_{ch} = \max_i x_i - y_i $
	Distância de Manhattan**	$d_m = \sum_{i=1}^p x_i - y_i $
	Distância de Canberra**	$d_{ca} = \sum_{i=1}^p \frac{ x_i - y_i }{x_i + y_i}$
Agrupamento	Formação do grupo	Referência
UPGMA	Utiliza a média das distâncias entre todos os pares de genótipos	Sneath e Sokal (1973)
Vizinho mais próximo	Utiliza a menor distância existente entre um par de genótipos	Lukaszewicz (1951)

* Adaptado de Moura et al. (2010), onde $k = 0$ o número de variáveis ($k = 1, 2, \dots, p$); i e $j =$ dois indivíduos que representem o acesso; $W_{ijk} =$ peso dado à comparação ijk , atribuindo valor 1 para comparações válidas e valor 0 para comparações inválidas (quando o valor da variável está ausente em um ou ambos indivíduos); $S_{ijk} =$ contribuição da variável k na similaridade entre os indivíduos i e j , com valores entre 0 e 1. Para uma variável qualitativa (nominal), se o valor da variável k é o mesmo para ambos os indivíduos, i e j , então $S_{ijk} = 1$, caso contrário, é igual a 0; para uma variável quantitativa (contínua) $S_{ijk} = 1 - |x_{ik} - x_{jk}| / R_k$, onde x_{ik} e x_{jk} são os valores da variável k para os indivíduos i e j , respectivamente, e R_k é o intervalo (valor máximo menos valor mínimo), da variável k na amostra. A divisão por R_k elimina as diferenças entre escalas das variáveis, produzindo um valor dentro do intervalo [0, 1] e pesos iguais; ** Adaptado de Webb (2002).

Visando o reconhecimento do grau de similaridade entre as distâncias relativas às variáveis quantitativas, estas matrizes foram submetidas a uma análise de correlação simples (Pearson), para efeito de comparação.

O número ótimo de grupos foi concebido pela metodologia de corte no dendrograma proposta por Kelley et al. (1996) (*Kelley-Gardner-Sutcliffe penalty function*), dada a partir da média das distâncias da matriz de agrupamento formada (matriz cofenética).

A consistência do padrão de agrupamento foi verificada através do grau de ajuste entre a representação gráfica da matriz de distância original (matriz fenética) e a matriz de agrupamento obtida (matriz cofenética), via coeficiente de correlação cofenética (SOKAL & ROHLF, 1962), utilizando o Teste de Mantel com 1000 permutações (MANTEL, 1967); pela qualidade das bifurcações construídas, através do coeficiente aglomerativo (VALE, 2005); e por meio dos níveis de distorção e stress (KRUSKAL, 1964; CRUZ & CARNEIRO, 2003).

Para Cruz e Carneiro (2003), o coeficiente de correlação cofenética é uma medida de concordância entre os valores originais de dissimilaridade e aqueles representados no dendrograma, fazendo uso, somente, dos valores encontrados acima da diagonal das referidas matrizes (fenética e cofenética).

Segundo Vale (2005), o coeficiente aglomerativo é capaz de medir a qualidade do agrupamento formado. Para tal, para cada genótipo i , $d(i)$ é sua dissimilaridade, em relação ao primeiro agrupamento em que foi inserido, dividido pela dissimilaridade na etapa final do algoritmo. Seus valores variam entre 0 e 1, em que, quanto mais pertos de 0, pior é a estrutura

do grupo, e quanto mais próximos de 1, mais definida é esta estrutura. O coeficiente aglomerativo é definido por:

$$Ca = \frac{1}{n} \sum_i^n 1 - d(i)$$

Onde,

n é o número total de genótipos avaliados.

Conforme Cruz e Carneiro (2003), os níveis de distorção e stress são igualmente aplicados como medida de adequação da representação de medidas de dissimilaridade via gráficos de dispersão 2D e 3D. O grau de distorção é definido por $(1 - \alpha) \times 100$, em que:

$$\alpha = \frac{\sum_{i < j} \sum d^2_{gij}}{\sum_{i < j} \sum d^2_{oij}}$$

A representação estatística do nível de stress (S) foi sugerida por Kruskal (1964). De acordo com este autor, tal parâmetro também tem por finalidade avaliar a adequação da projeção gráfica em relação às medidas de distâncias originais:

$$S = \sqrt{\frac{\sum_{i < j} \sum (d_{oij} - d_{gij})^2}{\sum_{i < j} \sum d^2_{oij}}} \times 100$$

Onde,

d^2_{gij} e d^2_{oij} são as distâncias gráficas (da matriz cofenética) e originais (da matriz fenética), respectivamente, de todos os pares de genótipos i e j .

Segundo Kruskal (1964), a adequação dos níveis de distorção e stress pode ser classificada em: insatisfatório, regular, bom, excelente e perfeito (Tabela 6).

Tabela 6. Classificação dos níveis de distorção e stress (KRUSKAL, 1964).

Nível (%)	Adequação
0	Perfeito
5	Excelente
10	Bom
20	Regular
40	Insatisfatório

3.4.4. Avaliação do potencial produtivo

A produção foi avaliada com base no registro dos descritores NSE e PMSs, relativos ao número e ao peso médio da semente seca por planta, coletados, semanalmente e por genótipo, entre os meses de dezembro de 2011 e maio de 2012, período relativo à frutificação das plantas de pinhão-manso nas condições climáticas de Seropédica, RJ. Dessa forma, avaliou-se a produção de grãos por planta ($PGP = NSE \times PMSs$) durante seis meses, totalizando 24 semanas de colheita.

Com o intuito de se entender melhor como as características referentes aos frutos e sementes poderiam influenciar no número de grãos gerados, avaliou-se, via correlação fenotípica, as associações entre seus descritores: peso do fruto (PFR), comprimento do fruto (CFR), largura do fruto (LFR), peso médio da semente seca (PMSs), comprimento da semente (CMS), largura da semente (LSE), estes mensurados a partir de paquímetro e balança digital, e tempo de formação do fruto (TFF) e produção de grãos por planta (PGP). Nesta etapa, similarmente às análises anteriormente empregadas, os referidos descritores também foram submetidos ao teste de multicolinearidade (MONTGOMERY & PECK, 1981).

Fundamentada nos valores semanais de produção, traçou-se um gráfico relativo ao comportamento produtivo dos 60 genótipos em questão, durante o período citado. A partir deste gráfico, identificou-se a época de maior produtividade, bem como os respectivos picos de produção.

A análise estatística para a distinção dos meses de colheita foi realizada por meio da comparação de médias duas a duas, através do teste *t*, comparando os seis meses entre-si. Para tal, os valores também foram transformados pela expressão $\sqrt{(x + 1)}$.

Com exceção da análise de trilha e dos testes *t* e de multicolinearidade, realizadas no Programa GENES (CRUZ, 2008), todas as demais análises estatísticas foram conduzidas no Programa R, versão 1.7.1 (R DEVELOPMENT CORE TEAM, 2011), utilizando-se os pacotes *cluster* (MAECHLER et al., 2011), *vegan* (OKSANEN et al., 2011), *ade4* (DRAY & DUFOUR, 2007) e *maptree* (WHITE, 2010).

4. RESULTADOS E DISCUSSÃO

4.1. Desenvolvimento, Validação e Normatização de Descritores Morfoagronômicos

Mediante a falta de descritores oficiais para *Jatropha curcas*, descritores de culturas afins foram empregados como referenciais na elaboração da lista para a caracterização da espécie. Dessa maneira, buscou-se, sempre que possível, descritores desenvolvidos pelo *Bioversity International* que poderiam ser adaptados para a espécie em questão. Assim, foram seguidas como referência as listas de descritores do gênero *Gossypium* (algodão) e das espécies *Glycine max* (soja) e *Ricinus communis* (mamona), por serem oleaginosas e, a última, também ser uma Euforbiácea. Somadas a estas, descritores individuais não-oficiais para *Jatropha curcas*, encontrados isoladamente em trabalhos sobre a caracterização da espécie, também foram testados e, quando necessário, adaptados para a composição da presente lista. De posse destes, 52 características, passíveis de serem adaptadas e aplicadas em *Jatropha curcas*, foram definidas para compor a lista preliminar de descritores da espécie.

Primeiramente, a partir dos descritores selecionados como referência, juntamente às recomendações propostas pelo *Guidelines for the development of crop descriptor lists* (BIOVERSITY INTERNATIONAL, 2007), aplicou-se a referida lista em função da separação das características relativas a aspectos vegetativos dos caracteres reprodutivos, respeitando, assim, as diferentes fases de desenvolvimento da planta. Nesta etapa, dos 52 descritores aplicados, contabilizou-se 29 para aspectos vegetativos e 23 para reprodutivos. Tal metodologia permitiu uma melhor abordagem prática no campo, uma vez que concentrou os esforços em dois momentos distintos ao longo do processo de caracterização, promovendo, em consequência, uma maior organização e otimização do rendimento da equipe e suas ferramentas de trabalho em relação ao tempo disponível para a descrição.

Posteriormente, para a normatização da lista, concomitante à divisão das etapas vegetativa e reprodutiva, a classificação dos 52 caracteres, fundamentada no tipo de variável, também auxiliou no processo de padronização dos descritores. Deste modo, foram estabelecidos três diferentes padrões para a normatização da lista: para variáveis quantitativas, multicategóricas e binárias, nas quais foram contabilizadas 36 variáveis quantitativas, 13 multicategóricas e 3 binárias (Tabela 7).

Tabela 7. Relação de variáveis e respectivos descritores elaborados para a caracterização, estudo da diversidade genética e potencial produtivo de 60 genótipos da Coleção de germoplasma de *Jatropha curcas* da UFRRJ.

Variáveis	Descritores		Número de variáveis
	Vegetativos	Reprodutivos	
Multicategóricas	ARQ / RAM / CCA CRA / CFJ / CFD CNF / CPE	FFR / NLFr / CSE TSE / PSE	13
Binárias	CER / PFO	CAR	3
Quantitativas	APD / PCOn / PCOs PCOI / PCOo / PCO DCA / NRT / NRI / CRP CERp / CRI / CERi NLF / AIFc / CMP CFO / LFO / CLFo	API / CPI / NTI UMC / TFF / NFP PMF / CFR / LFR CLFr / NSF / NSE PMS / CMS / LSE CLSe / PGP	36
Número de descritores	29	23	52

Legenda: ARQ - Arquitetura da planta; RAM - Tipo de ramificação; CCA - Coloração do caule; CRA - Coloração do ramo; CFJ - Coloração de folhas jovens; CFD - Coloração de folhas desenvolvidas; CNF - Coloração da nervura foliar; CPE - Coloração do pecíolo; FFR - Formato do fruto; NLFr - Número de lóculos por fruto; CSE - Coloração da semente; TSE - Textura da semente; PSE - Padrão da semente; CER - Cerosidade; PFO - Pilosidade da folha; CAR - Carúncula da semente; APD - Altura de plantas desenvolvidas; PCO - Área de projeção da copa; PCOn - Projeção da copa no sentido Norte; PCOs - Projeção da copa no sentido Sul; PCOI -

Projeção da copa no sentido Leste; PCOo - Projeção da copa no sentido Oeste; DCA - Diâmetro do caule; NRT - Número de ramos totais; NRI - Número de ramos com inflorescência; CRP - Comprimento de ramos primários; CERp - Comprimento do entrenó de ramos primários; CRI - Comprimento de ramos com inflorescência; CERi - Comprimento do entrenó de ramos com inflorescência; NLF - Número de lóbulos foliar; AIFc - Ângulo de inserção da folha no caule; CMP - Comprimento do pecíolo; CFO - Comprimento da folha; LFO - Largura da folha; CLFo - Relação entre o comprimento e a largura da folha; API - Altura da primeira inflorescência; CPI - Comprimento do pedúnculo da inflorescência; NTI - Número total de inflorescências; UMC - Uniformidade de maturação do cacho por planta; TFF - Tempo de formação do fruto; NFR - Número de frutos por planta; PFR - Peso de frutos por planta; CFR - Comprimento do fruto; LFR - Largura do fruto; CLFr - Relação entre o comprimento e a largura do fruto; NSF - Número de sementes por fruto; NSE - Número de sementes por planta; PMSs - Peso médio da semente seca; CMS - Comprimento da semente; LSE - Largura da semente; CLSe - Relação entre o comprimento e a largura da semente; e PGP - Produção de grãos por planta.

Variáveis quantitativas, de acordo com o já mencionado guia (BIOVERSITY INTERNATIONAL, 2007), devem, preferencialmente, ser registradas através de seu valor exato, recomendando-se o uso do Sistema Internacional de Medidas. No entanto, este guia demonstra que, quando conveniente, tanto variáveis quantitativas discretas quanto contínuas podem ser transformadas em variáveis multicategóricas. Apesar da possibilidade, variáveis quantitativas discretas são mais apropriadas à transformação, visto que estas podem ter seus *estados* facilmente enumerados em classes e dispostos em escala numérica para anotação, sem a necessidade da construção de “intervalos de classes” e o risco de sofrerem maiores distorções. Dessa forma, com exceção do descritor número de lóculos por fruto (NLFr), variável quantitativa discreta posteriormente convertida em variável multicategórica, não foram estabelecidas classes multicategóricas para as demais variáveis quantitativas relacionadas, visando, portanto, uma maior fidelidade ao valor expresso. Nesse sentido, embora igualmente passível de conversão, o descritor número de lóbulos foliar (NLF) não foi tratado como uma variável multicategórica, pois, na aplicação prática deste, além da dificuldade inerente à descrição, constatou-se a possibilidade de se trabalhar com um grande número de classes para sua classificação, dada à variabilidade de *estados* encontrados para este descritor entre as plantas estudadas. Como tal fato não é recomendado para o registro de caracteres multicategóricos, uma vez que um elevado número de categorias pode dificultar a descrição do atributo, julgou-se mais adequado não atribuir classes a este descritor. Já para variáveis multicategóricas, de caráter nominal ou ordinal, diferentes classes foram estabelecidas de modo a se adequar a melhor forma de representação do *estado* a ser descrito. Sob esta perspectiva, preferiu-se estabelecer um número máximo de até 5 diferentes categorias para classificação. Conforme o guia referido e já comentado, um excessivo número de classes categóricas pode comprometer o critério de descrição do avaliador. Variáveis binárias foram codificadas pelo emprego de “0” para denotar ausência e “1” para presença da característica.

Perante as observações enunciadas, os descritores vegetativos APD, PCOn, PCOs, PCOI, PCOo, PCO, DCA, NRT, NRI, CRP, CERp, CRI, CERi, NLF, AIFc, CMP, CFO, LFO e CLFo, expressos por variáveis quantitativas, tiveram o valor exato da característica registrado. Dentre estes, os descritores CRP, CERp, CRI, CERi, NLF, AIFc, CMP, CFO, e LFO tiveram discriminados em sua metodologia uma quantidade ótima de unidades amostrais a serem avaliadas, com o objetivo de se definir uma amostra significativa padrão para caracterização. Este procedimento permite que o valor descrito possa ser comparado de forma igual entre diferentes avaliadores, aumentando sua compatibilidade e poder discriminatório, uma vez que amplia seu potencial de uso e intercâmbio entre diferentes instituições. Ainda em relação às variáveis quantitativas, os descritores PCOn, PCOs, PCOI e PCOo foram adaptados dos descritores relativos à projeção da copa entre linhas e na linha de plantio, utilizados por Laviola et al. (2011), em trabalho sobre a caracterização de germoplasma de pinhão-manso em fase jovem. Nos presentes descritores, deu-se preferência em registrar os valores de

projeção nos quatro sentidos cardeais, visando, posteriormente, investigar suas correlações com outras características da lista desenvolvida. Adicionada a estes, também criou-se um descritor específico para descrever a área de projeção da copa (PCO), utilizando para tal os valores observados pelos descritores PCOn, PCOs, PCOl e PCOo na fórmula $[(PCOn + PCOs) \times (PCOl + PCOo)] / 2$. Assim como os descritores relativos à projeção da copa, o descritor NRI foi igualmente adaptado do descritor número de ramos secundários, de Laviola et al. (2011). No entanto, na presente descrição, levando em consideração a emissão de inflorescências e o tipo de ramificação da espécie, preferiu-se adotar a caracterização de ramos com inflorescência, visto que estas podem ser emitidas não somente por ramos secundários. Os descritores CERp e CERi foram adaptados do descritor “comprimento dos internódios”, da lista de descritores de mamona utilizados pela EMBRAPA (MILANI, 2008). Na lista citada, este descritor preconiza a medição do comprimento de entrenós somente no caule principal, mas como foram observadas diferenças visuais claras desta característica em relação aos diferentes ramos da planta, adequou-se a caracterização para ramos primários e com inflorescência. Demais variáveis quantitativas referentes a aspectos vegetativos tiveram seus descritores (APD, DCA, NRT e CLFo) igualmente transcritos da supracitada lista de Milani (2008) para *Ricinus communis*.

Os descritores ARQ, RAM, CCA, CRA, CFJ, CFD, CNF e CPE representaram as características vegetativas relativas às variáveis multicategóricas. Entre elas, os descritores ARQ e RAM foram adaptados dos descritores de mamona “arquitetura da planta” e “tipo de ramificação”, respectivamente, de Milani (2008) e Milani et al. (2006). Visando facilitar o entendimento e o emprego destes descritores por parte do avaliador, seus *estados* foram modificados por termos mais formais, comumente utilizados pela botânica e organografia. Desse modo, a arquitetura da planta foi classificada em *ereta*, *fechada* e *aberta*, em oposição ao termo “semi-ereta” originalmente adotado, ao passo que o tipo de ramificação teve as categorias “trifurcada”, “taça” e “universal” trocadas por *ausência de ramificação*, *monopodial*, *dicasial* e *simpodial*, visto que os termos originais empregados poderiam levantar dúvidas quanto à descrição destas características (VIDAL & VIDAL, 2006). Já os descritores CCA, CRA, CFJ, CFD, CNF e CPE, que dizem respeito à coloração, foram transpostos da lista de descritores de soja (IBPGR, 1984) e algodão (IBPGR, 1985), propostas pelo IBPGR. Nesse sentido, só foram alterados os *estados* relativos às cores observadas para *Jatropha curcas* e, a estas, adicionada o *estado* expresso pela cor *azul* para CCA e CRA. Embora não registrada oficialmente na literatura, relatos informais afirmam a existência de tal coloração para a espécie, desse modo, julgou-se conveniente adotar tal classificação. Para os descritores CER e PFO, referentes a variáveis binárias, foram adotados os descritores correspondentes aos da mamona (MILANI, 2008), relatando somente a presença ou ausência da característica.

Assim como os descritores vegetativos CRP, CERp, CRI, CERi, NLF, AIFc, CMP, CFO, e LFO, os descritores reprodutivos CPI, UMC, CFR, LFR, CMS e LSE, igualmente relativos a variáveis quantitativas, tiveram discriminados em sua metodologia uma amostra ótima a ser avaliada, contemplando alcançar maior compatibilidade e uso mediante diferentes avaliadores. Entre estes, o descritor CPI foi adaptado de Laviola et al. (2011), onde o tamanho do pedúnculo da inflorescência de *Jatropha curcas* foi relatado de forma qualitativa. Dessa maneira, adequou-se este de acordo com o descritor de mamona “comprimento dos pedúnculos”, por Milani (2008), em que foi avaliado como variável quantitativa, em escala contínua, por mensuração. Como uma das características alvo de melhoramento genético em *Jatropha curcas* diz respeito à uniformidade de maturação dos frutos, e esta não foi relacionada entre os descritores selecionados para composição da presente lista, propôs-se o descritor UMC para caracterizar a uniformidade dos cachos de frutos emitidos. Nesse contexto, os descritores reprodutivos API, NTI, TFF, NFR, PFR, CLFr, NSF, NSE, PMSs,

CLSe e PGP são, também, referentes a características importantes para o melhoramento genético da espécie estudada. Dentre estes, os descritores API e PGP foram inteiramente transcritos destas mesmas características estudadas por Laviola et al. (2011), onde a altura da primeira inflorescência (API) foi obtida através da medida compreendida entre o solo e a primeira inflorescência da planta. Na presente lista, a produção de grãos por planta (PGP), produto entre o número de sementes e o peso médio da semente seca por planta, foi calculada baseada no peso médio de sementes secas, denotado pelo descritor PMSs, visto que, na contabilização prática da produção, dificilmente será utilizada a semente úmida para extração de óleo. Objetivando estudar futuramente a previsibilidade de produção, foi proposto o descritor NTI, que, com base no número de inflorescências emitidas pela planta, pode vir a auxiliar na estimação de um parâmetro para tal. Com o intuito de se identificar a uniformidade e o período levado para produção, aspectos que podem refletir sobre o tempo exigido para a colheita, adaptou-se o descritor TFF do descritor “ciclo de plantio”, utilizado por Milani (2008) para mamona. Da mesma forma, demais variáveis quantitativas representadas por aspectos reprodutivos tiveram seus descritores (CFR, LFR, CMS, LSE, NFR, PFR, CLFr, NSF, NSE e CLSe) transpostos da referida lista de *Ricinus communis*.

Representando os caracteres reprodutivos relativos a variáveis multicategóricas estão os descritores FFR, NLFr, CSE, TSE e PSE, onde, com exceção dos descritores NLFr e FFR, este último adaptado de Laviola et al. (2011), foram transcritos das listas de descritores de soja e algodão organizadas pelo IBPGR. Nesse caso, somente os *estados* relativos às características da semente foram modificados por novos *estados*, condizentes aos encontrados em *Jatropha curcas*. Para o descritor FFR, as categorias originalmente utilizadas, “elipsóide esférico”, “elipsóide lanceolado” e “elipsóide ovóide” foram trocadas por *elíptico*, *oval* e *triangular*, visando facilitar o entendimento para a aplicação prática do descritor. Como discutido previamente, o descritor NLFr, inicialmente tratado como uma variável quantitativa, foi avaliado como multicategórica, através da definição das classes *unilocular*, *bilocular*, *trilocular* e *tetralocular*. A variável binária referente ao descritor CAR foi adaptada do descritor de mamona “tipo de carúncula”, o qual foi caracterizado pelas categorias “protuberante” e “não-protuberante”, por Milani (2008). Todavia, por tais *estados* não terem sido observados ao longo da aplicação do descritor, decidiu-se relatar somente a presença ou ausência desta característica.

4.2. Análises Estatísticas

4.2.1. Importância e descarte de variáveis via análise de componentes principais

Visando reconhecer a importância das 36 variáveis quantitativas sobre a diversidade observada, uma análise multivariada desta foi realizada através da técnica de componentes principais. Para tal, estas foram submetidas, primeiramente, a uma análise estatística descritiva, onde seus valores máximos, mínimos e médios, bem como seus coeficientes de variação, auxiliaram na identificação das que mais e menos variaram (Tabela 8).

Tabela 8. Estatística descritiva de 36 variáveis (19 vegetativas e 17 reprodutivas) mensuradas em 60 genótipos da Coleção de germoplasma de *Jatropha curcas* da UFRRJ.

Descritores	Variáveis	Média	Mínimo	Máximo	CV	Variância	Desvio padrão
Vegetativos	APD	241,92	179,00	305,00	10,05	591,23	24,31
	PCO	4,23	1,52	9,15	35,39	2,24	1,50
	PCOn	1,60	0,96	2,35	17,71	0,08	0,28
	PCOs	1,38	0,80	2,00	20,03	0,07	0,28
	PCOI	1,41	0,91	2,30	22,80	0,10	0,32
	PCOo	1,35	0,81	2,31	24,58	0,11	0,33
	DCA	11,16	8,60	13,80	10,47	1,37	1,17
	NRT	28,07	12,00	44,00	28,67	64,74	8,05
	NRI	22,62	8,00	38,00	31,86	51,93	7,20
	CRP	177,92	121,00	255,00	16,74	887,37	29,79
	CERp	1,47	1,00	2,00	17,00	0,06	0,25
	CRI	74,52	45,00	124,00	19,78	217,17	14,74
	CERi	1,76	1,08	2,72	18,92	0,11	0,33
	NLF	4,63	3,00	5,00	16,84	0,61	0,78
	AIFc	133,20	112,00	153,00	6,48	74,43	8,63
	CMP	14,57	8,30	19,60	14,70	4,59	2,14
	CFO	12,74	8,80	14,40	7,53	0,92	0,96
	LFO	12,96	8,60	15,40	8,49	1,21	1,10
	CLFo	0,98	0,86	1,10	4,36	0,01	0,04
Reprodutivos	API	92,63	30,00	145,00	27,19	634,37	25,19
	CPI	2,49	1,50	7,40	30,76	0,58	0,76
	NTI	123,05	19,00	353,00	38,62	2258,87	47,53
	UMC	47,63	8,00	100,00	52,82	633,15	25,16
	TFF	83,85	74,00	113,00	9,58	64,47	8,03
	NFR	266,52	11,00	756,00	58,97	24698,42	157,16
	PFR	12,49	9,29	14,62	7,31	0,83	0,91
	CFR	3,09	2,86	3,75	4,31	0,02	0,13
	LFR	2,70	2,53	2,93	3,43	0,01	0,09
	CLFr	1,15	1,10	1,41	3,81	0,00	0,04
	NSF	1,67	1,00	3,00	4,35	0,33	0,57
	NSE	676,08	30,00	1973,00	58,23	154965,40	393,66
	PMSs	0,72	0,63	0,80	5,62	0,00	0,04
	CMS	1,80	1,73	1,92	2,53	0,00	0,04
	LSE	1,09	0,92	1,15	3,32	0,00	0,04
	CLSe	1,65	1,57	1,99	4,21	0,00	0,07
	PGP	487,02	23,85	1402,04	59,10	82841,16	287,82

Legenda: APD - Altura de plantas desenvolvidas (cm); PCO - Área de projeção da copa (m²); PCOn - Projeção da copa no sentido Norte (m); PCOs - Projeção da copa no sentido Sul (m); PCOI - Projeção da copa no sentido Leste (m); PCOo - Projeção da copa no sentido Oeste (m); DCA - Diâmetro do caule (cm); NRT - Número de ramos totais; NRI - Número de ramos com inflorescência; CRP - Comprimento de ramos primários (cm); CERp - Comprimento do entrenó de ramos primários (cm); CRI - Comprimento de ramos com inflorescência (cm); CERi - Comprimento do entrenó de ramos com inflorescência (cm); NLF - Número de lóbulos foliar; AIFc - Ângulo de inserção da folha no caule (graus); CMP - Comprimento do pecíolo (cm); CFO - Comprimento da folha (cm); LFO - Largura da folha (cm); CLFo - Relação entre o comprimento e a largura da folha; API - Altura da primeira inflorescência (cm); CPI - Comprimento do pedúnculo da inflorescência (cm); NTI - Número total de inflorescências; UMC - Uniformidade de maturação do cacho por planta (%); TFF - Tempo de formação do fruto (dias); NFR - Número de frutos por planta; PFR - Peso de frutos por planta (g); CFR - Comprimento do fruto (cm); LFR - Largura do fruto (cm); CLFr - Relação entre o comprimento e a largura do fruto; NSF - Número de sementes por fruto; NSL - Número de sementes por loco; NSE - Número de sementes por planta; PMSs - Peso médio da semente seca (g); CMS - Comprimento da semente (cm); LSE - Largura da semente (cm); CLSe - Relação entre o comprimento e a largura da semente; PGP - Produção de grãos por planta (g.planta⁻¹); e CV - Coeficiente de variação fenotípica (%).

De acordo com os valores apresentados na Tabela 8, pode-se observar que os descritores vegetativos PCO e NRI, juntamente aos reprodutivos CPI, NTI, UMC, NFR, NSE e PGP demonstraram um alto coeficiente de variação fenotípica (acima de 30%),

evidenciando uma grande variabilidade para estas características, apontada pelos coeficientes de 35,39, 31,86, 30,76, 38,62, 52,82, 58,97, 58,23 e 59,10%, respectivamente. Tendo em vista que caracteres associados ao aumento e à uniformidade de produção constituem um dos principais focos do melhoramento genético em *Jatropha curcas*, tais resultados podem ser considerados de grande valia para programas de melhoramento da espécie, uma vez que possibilitam a seleção de genótipos superiores para características fortemente relacionadas à produtividade, como os descritores NRI, NTI, UMC, NFR, NSE e a própria produção de grãos por planta (PGP). Os caracteres que apresentaram menor variação foram os descritores vegetativos AIFc, CFO, LFO e CLFo, referentes às folhas, e TFF, PFR, CFR, LFR, CLFr, NSF, PMSs, CMS, LSE e CLSe, de caráter reprodutivo e relativos a atributos de frutos e sementes, todos representados por coeficientes abaixo de 10%. Em concordância aos valores aqui encontrados, Mohapatra e Panda (2010), em trabalho sobre a variabilidade genética em genótipos indianos de *J. curcas*, aos 4 anos de idade, também observaram baixos coeficientes de variação para os descritores LFR, CMS e LSE, com valores de 3,58, 9,27 e 6,67%, respectivamente, e altos valores para NRI (34,15%), NTI (38,11%) e NFR (40,76%). Igualmente entre genótipos indianos de *J. curcas*, porém, aos 2,5 anos de idade, uma alta variabilidade fenotípica para a produção de grãos foi identificada por Rao et al. (2008), expressa pelo valor máximo de 55,26%.

Baseado nos dados aqui registrados, e sustentado pelo descritor PGP, obtido através do produto entre o número de sementes por planta (NSE) e o peso médio da semente seca (PMSs), é possível inferir que, no presente caso, o sucesso no processo de seleção de acessos superiores para o aumento da produção em *Jatropha curcas* pode ser alcançado tanto pela seleção de plantas com alta variabilidade para a produção de grãos quanto pelo aumento da média do número de sementes por planta, dada à grande variabilidade genética destas variáveis na coleção de germoplasma em questão. Tendo em vista os valores expressos na Tabela 8, pode-se notar que a média geral do número de sementes foi de 676,08, no entanto, a Coleção de germoplasma de *Jatropha curcas* da UFRRJ possui genótipos com amplitude de NSE variando entre 30,00 (mínima) e 1973,00 sementes (máxima). Dessa maneira, a seleção truncada diretamente em NSE pode contribuir para o aumento da média desta variável em gerações subsequentes e, por conseguinte, para o incremento da produção de grãos por planta. Ainda em relação a atributos intrínsecos às sementes, Santoso (2011) relatou valores de comprimento e largura oscilando entre 1,78 e 1,85 cm e 0,80 e 0,86 cm, respectivamente, em genótipos de *J. curcas* da Indonésia, ao passo que os valores encontrados no presente trabalho variaram de 1,73 a 1,92 cm de comprimento e 0,92 a 1,15 cm de largura, denotando, assim, a existência de genótipos portadores de sementes de maior tamanho na coleção pesquisada, com valor médio de peso de 0,72 g e amplitude de 0,63 a 0,80 g. Tal fato se mostra de grande valor para a seleção de genitores superiores de *J. curcas* no que diz respeito à elevação do teor de óleo em suas sementes, haja visto que Rocha et al. (2009) relataram que esta característica está diretamente correlacionada ao peso de sementes e amêndoas, com estimativas, respectivamente, de correlação de 0,93 e 0,15 sobre o rendimento de óleo.

A análise de componentes principais permitiu identificar as variáveis que contribuíram em maior e menor escala para a variabilidade encontrada. De acordo com Santos et al. (1995), esta técnica permite transformar um grupo original de variáveis em um conjunto secundário, de dimensão equivalente, porém, independentes entre-si, portando a máxima informação necessária para explicar a variação observada. Desse modo, a análise de componentes principais possibilita reduzir o conjunto original de variáveis, sendo, por isso, frequentemente utilizada na estimação da diversidade genética, via dispersão gráfica dos componentes de maior variação. Nesse sentido, Morrison (1976) adverte que os dois primeiros componentes devem reter 75% de toda variação pronunciada. Complementando esta afirmativa, Cruz e Carneiro (2003) recomendam que, no mínimo, 80% da variação seja absorvida até os três

primeiros componentes principais, tornando, assim, mais segura a representação gráfica em planos bi ou tridimensionais.

Conforme exposto na Tabela 9, observa-se, no presente trabalho, que foram necessários 11 componentes principais para explicar os 80,11% da variância total obtida, onde os 3 primeiros componentes foram responsáveis por, somente, 44,26% desta variação. Sendo assim, Cruz e Carneiro (2003) observam que uma representação gráfica consistente, a respeito da diversidade apresentada, seria altamente dependente da plotagem de gráficos bidimensionais, em combinação, do primeiro (CP 1) até o décimo primeiro componente principal (CP 11). Contudo, devido ao elevado número de gráficos necessários para realizar a análise, tal procedimento dificultaria tecer maiores inferências acerca da diversidade genética expressa pelos acessos. Resultado análogo foi descrito por Chiorato (2004), que fez uso de 10 componentes principais ao ressaltar 77,38% da variabilidade acumulada em 993 genótipos de feijão comum (*Phaseolus vulgaris*), tornando, do mesmo modo, impraticável sua representação gráfica. Em discordância ao presente trabalho, Carvalho (2010), estudando o comportamento genético de pinhão-manso, conseguiu reunir 80,94% da variação observada em apenas dois componentes principais, destacando, nestes, a importância dos caracteres número de cachos e ramos secundários, respectivamente.

Tabela 9. Resumo das estimativas da variância (autovalores) associadas aos componentes principais, contribuição relativa, acumulada e caracteres passíveis de descarte (redundantes), a partir de 36 variáveis quantitativas avaliadas em 60 genótipos da Coleção de germoplasma de *Jatropha curcas* da UFRRJ.

Componentes principais	λ_j - Autovalores	λ_j - Contribuição (%)		Caracteres passíveis de descarte
		Relativa	Acumulada	
CP 1	2,90	23,40	23,40*	-
CP 2	2,08	12,07	35,47*	-
CP 3	1,78	8,79	44,26*	-
CP 4	1,59	6,99	51,25*	-
CP 5	1,48	6,10	57,35*	-
CP 6	1,36	5,18	62,53*	-
CP 7	1,22	4,17	66,70*	-
CP 8	1,16	3,75	70,45*	-
CP 9	1,12	3,47	73,92*	-
CP 10	1,10	3,37	77,29*	-
CP 11	1,01	2,82	80,11*	-
CP 12	0,98	2,64	82,75	-
CP 13	0,92	2,38	85,13	-
CP 14	0,88	2,14	87,27	-
CP 15	0,81	1,84	89,11	-
CP 16	0,77	1,66	90,77	-
CP 17	0,67**	1,24	92,01	CERp
CP 18	0,65**	1,17	93,18	CRP
CP 19	0,60**	1,02	94,20	NRT
CP 20	0,57**	0,91	95,11	UMC
CP 21	0,56**	0,89	96,00	CMP
CP 22	0,55**	0,83	96,83	PCOI
CP 23	0,48**	0,64	97,47	PMSs
CP 24	0,45**	0,57	98,04	APD
CP 25	0,45**	0,57	98,61	PCOs
CP 26	0,39**	0,43	99,04	PFR
CP 27	0,36**	0,36	99,40	PCOn
CP 28	0,30**	0,25	99,65	NRI
CP 29	0,28**	0,21	99,86	CLFr
CP 30	0,18**	0,09	99,95	NTI
CP 31	0,08**	0,02	99,97	NFR
CP 32	0,07**	0,01	99,98	PCO
CP 33	0,06**	0,01	99,99	LFO

Continuação da Tabela 9

CP 34	0,03**	0,00	100,00	CLSe
CP 35	0,02**	0,00	100,00	PGP
CP 36	0,01**	0,00	100,00	CFR

Legenda: CERp - Comprimento do entrenó de ramos primários; CRP - Comprimento de ramos primários; NRT - Número de ramos totais; UMC - Uniformidade de maturação do cacho por planta; CMP - Comprimento do pecíolo; PCOl - Projeção da copa no sentido Leste; PMSs - Peso médio da semente seca; APD - Altura de plantas desenvolvidas; PCOs - Projeção da copa no sentido Sul; PFR - Peso de frutos por planta; PCOn - Projeção da copa no sentido Norte; NRI - Número de ramos com inflorescências; CLFr - Relação entre o comprimento e a largura do fruto; NTI - Número total de inflorescências; NFR - Número de frutos por planta; PCO - Área de projeção da copa; LFO - Largura da folha; CLSe - Relação entre o comprimento e a largura da semente; PGP - Produção de grãos por planta; e CFR - Comprimento do fruto. *, ** Valores em negrito relativos aos critérios de Cruz e Carneiro (2003), para a representação gráfica, e Jolliffe (1972 e 1973), para o descarte de caracteres, respectivamente.

Segundo Jolliffe (1972 e 1973), quando se dispõe de um grande número de descritores, como no presente estudo, muitos destes podem vir a ser redundantes, portanto, passíveis de descarte, dado o baixo valor informativo adicionado à avaliação. Assim, este autor recomenda que o número de variáveis descartadas seja igual ao número de componentes cuja variância (autovalor) não ultrapasse o valor de 0,70. Desse modo, a partir dos autovalores discriminados na Tabela 9, nota-se que, dos 36 componentes principais gerados, 20 apresentaram variância inferior a 0,70, compreendendo do décimo sétimo (CP 17) ao trigésimo sexto componente (CP 36). Tendo por base os maiores autovetores associados aos componentes principais de autovalores inferiores a 0,70, os descritores CERp, CRP, NRT, UMC, CMP, PCOl, PMSs, APD, PCOs, PFR, PCOn, NRI, CLFr, NTI, NFR, PCO, LFO, CLSe, PGP e CFR, respectivamente associados aos componentes principais 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 e 36, foram indicados à eliminação (Tabela 9). Dessa forma, pode-se inferir que os descritores vegetativos demonstraram uma importância equivalente à conferida pelos reprodutivos, dada à quantidade de variáveis quantitativas inicialmente propostas (19 vegetativas e 17 reprodutivas) e o número recomendado para exclusão (11 vegetativas e 9 reprodutivas), para a variabilidade genética encontrada. Ao descartar variáveis via componentes principais em *J. curcas*, Carvalho (2010) também desconsiderou o descritor “número de ramificações”.

De acordo com Daher et al. (1997) e Paiva et al. (2010), variáveis menos discriminantes são recomendadas ao descarte por já estarem significativamente correlacionadas a demais características, ou pela sua própria invariância, ao passo que descritores de maior poder discriminatório apresentam menor correlação entre-si. No presente trabalho, tal fato pôde ser comprovado através do diagnóstico de multicolinearidade, proposto por Montgomery e Peck (1981), onde o número de condição (NC) gerado, da ordem de 69200,28, expôs a existência de um severo grau de colinearidade, caracterizando, assim, uma forte correlação entre algumas variáveis, como, por exemplo, entre NFR e PGP, de 0,98 (Tabela 10). Dessa maneira, pode-se observar que, embora não descartável, o descritor NSE está fortemente correlacionado aos caracteres número de frutos (0,98) e produção de grãos por planta (0,99), indicados à exclusão, visto que, este último, é obtido através do produto entre NSE e PMSs. Entretanto, Cruz e Carneiro (2003) ressaltam que, estimativas de correlação, subestimadas e/ou superestimadas, ocasionadas por multicolinearidade severa, podem ser advindas de um grande número de variáveis trabalhadas. Portanto, tal efeito pode ser justificado pelo extenso grupo de descritores aqui envolvidos, uma vez que 36 variáveis foram investigadas.

Tabela 10. Diagnóstico de multicolinearidade e principais correlações apontadas pelo teste de Montgomery e Peck (1981), a partir de 36 variáveis quantitativas descritas em 60 genótipos de *Jatropha curcas* da UFRRJ.

Variáveis	Correlação	Significância - Alfa (%)
NFR x NSE	0,98	0,00*
NFR x PGP	0,98	0,00*
NSE x PGP	0,99	0,00*
Diagnóstico de multicolinearidade		
Número de condição (NC)		69200,28
Grau		Severa

Legenda: NFR - Número de frutos por planta; NSE - Número de sementes por planta; e PGP - Produção de grãos por planta. * Significativo a 1% de probabilidade pelo teste *t*.

Objetivando contornar possíveis complicações geradas pelo elevado grau de colinearidade, Carvalho et al. (1999) atentam para a identificação e exclusão da(s) provável (eis) “variável(eis)-problema”, como aquela(s) relacionada(s) ao maior elemento do autovetor associado aos autovalores de menor magnitude. Nesse sentido, Cruz e Carneiro (2003) afirmam que autovalores próximos de zero podem indicar a existência de dependência linear, isto é, estimativas de correlação próximas a 1 (Tabela 10). Perante o exposto, percebe-se que o reconhecimento destas variáveis se dá igualmente ao critério de Jolliffe (1972 e 1973), instituído à identificação de descritores redundantes via análise de componentes principais. Desse modo, no presente estudo, as “variáveis-problema”, responsáveis pelos efeitos adversos da multicolinearidade, corresponderam às mesmas variáveis passíveis de descarte, elucidadas pelos componentes principais indicados na Tabela 9. Portanto, procurando-se evitar maiores distorções sobre as estimativas de correlação, decidiu-se excluir as 20 variáveis de menor peso para a avaliação e, a fim de se conhecer, dentre os 16 descritores restantes, os de maior importância, proceder com uma nova análise de componentes principais. A exclusão de variáveis, visando alcançar uma maior consistência sobre o padrão de agrupamento gerado, também já foi testada com sucesso por Cargnelutti Filho e Guadagnin (2011), em trabalho sobre a diversidade genética em milho (*Zea mays* L.).

Quando novamente submetida ao Teste de Montgomery e Peck (1981), a matriz de correlação, dispondo agora das 16 variáveis pré-selecionadas, demonstrou um fraco grau de severidade, ilustrado por um número de condição de 13,67, concordando, assim, com a hipótese fornecida por Carvalho et al. (1999). Para Cruz e Carneiro (2003), coeficientes de ponderação (autovetores) de elevado valor, associados a autovalores responsáveis por 80% da variação explicada, referem-se aos caracteres de maior contribuição para a diversidade encontrada. Tendo por base os autovetores apontados na Tabela 11, nota-se que os descritores TFF, CFO, CPI, DCA, CLFo, PCOo, LFR, API e LSE, relativos aos componentes principais compreendidos entre os autovalores 1,64 (CP 1) e 0,88 (CP 9), de variância acumulada até 81,04%, representaram, dentre os 16 descritores considerados, as características de maior peso para a variabilidade genética da coleção estudada. Nesse sentido, Strapasson et al. (2000) afirmam que o primeiro componente principal também é definido como o de maior importância, uma vez que retém os caracteres responsáveis por maior parte da variação total observada. Desse modo, junto as 9 características apontadas, pode-se incluir o descritor NSE, de autovetor de valor -0,45, por, junto ao descritor TFF (0,49), também associado ao CP 1, ser responsável por grande parcela da variação anotada.

Tabela 11. Estimativas da variância (autovalores), coeficientes de ponderação associados (autovetores) e contribuição relativa e acumulada a partir de 16 variáveis quantitativas avaliadas em 60 genótipos da Coleção de germoplasma de *Jatropha curcas* da UFRRJ.

Descritores	CP1	CP2	CP3	CP4	CP5	CP6	CP7	CP8
PCOo	-0,29	0,13	-0,06	0,12	-0,24	0,56	-0,21	0,01
DCA	-0,11	-0,02	-0,29	0,48	0,19	-0,27	0,01	-0,34
CRI	0,15	-0,06	0,47	0,20	-0,06	-0,12	0,19	-0,41
CERi	-0,28	-0,41	-0,18	0,09	0,01	-0,33	-0,19	0,05
NLF	-0,22	-0,01	-0,07	-0,44	-0,18	-0,33	-0,45	-0,24
AIFc	-0,13	-0,34	0,03	0,08	0,35	0,18	-0,37	0,16
CFO	-0,14	0,51	0,02	-0,08	0,05	-0,30	0,21	-0,05
CLFo	0,18	0,22	0,13	-0,25	0,46	0,01	-0,15	0,34
API	0,08	0,15	-0,06	0,28	-0,25	-0,43	-0,09	0,64
CPI	0,18	0,23	-0,54	0,05	0,07	0,18	0,09	-0,01
TFF	0,49	-0,14	-0,24	0,01	-0,16	-0,10	-0,19	-0,09
LFR	-0,01	0,41	0,16	0,06	-0,34	0,03	-0,50	-0,15
NSF	-0,37	0,17	0,06	-0,19	0,39	-0,14	0,02	-0,07
NSE	-0,45	0,10	-0,28	0,09	-0,14	0,09	0,23	0,06
CMS	0,15	0,28	-0,08	0,37	0,39	0,02	-0,35	-0,19
LSE	-0,19	0,00	0,42	0,42	-0,01	0,01	-0,06	0,19
λ_j - Autovalores	1,64	1,34	1,27	1,24	1,19	1,07	1,02	0,97
λ_j relativa (%)	16,76	11,31	10,13	9,62	8,92	7,19	6,47	5,84
λ_j acumulada (%)	16,76	28,07	38,20	47,82	56,74	63,93	70,40	76,24
Descritores	CP9	CP10	CP11	CP12	CP13	CP14	CP15	CP16
PCOo	-0,09	0,26	-0,22	0,40	-0,26	0,34	0,01	-0,01
DCA	-0,42	0,03	0,16	-0,11	-0,28	0,19	-0,33	0,12
CRI	-0,10	0,32	-0,60	-0,00	0,08	-0,14	0,04	-0,01
CERi	0,31	0,24	-0,14	-0,25	-0,05	0,31	0,47	-0,09
NLF	0,16	-0,02	-0,17	0,16	-0,18	-0,25	-0,35	0,24
AIFc	-0,33	-0,47	-0,39	-0,02	0,18	-0,12	-0,01	-0,08
CFO	0,03	-0,51	-0,30	0,10	-0,23	0,26	0,24	-0,19
CLFo	-0,10	0,37	-0,13	-0,31	-0,45	0,01	-0,09	-0,09
API	-0,22	0,16	-0,14	0,26	0,17	-0,08	-0,02	0,16
CPI	0,31	0,03	-0,43	-0,28	0,31	0,12	-0,25	0,22
TFF	0,04	-0,01	-0,04	0,18	-0,05	0,08	-0,18	-0,72
LFR	-0,20	-0,01	0,13	-0,50	0,26	0,07	0,13	-0,10
NSF	-0,05	0,29	0,10	0,23	0,55	0,19	-0,19	-0,29
NSE	-0,05	0,11	-0,08	-0,19	-0,11	-0,62	0,02	-0,39
CMS	0,30	0,05	0,12	0,31	-0,01	-0,35	0,34	0,06
LSE	0,52	-0,15	0,04	-0,12	-0,08	0,06	-0,47	-0,14
λ_j - Autovalores	0,88	0,83	0,78	0,69	0,68	0,58	0,51	0,44
λ_j relativa (%)	4,80	4,32	3,77	2,98	2,88	2,12	1,64	1,25
λ_j acumulada (%)	81,04	85,36	89,13	92,11	94,99	97,11	98,75	100,00

Legenda: AIFc - Ângulo de inserção da folha no caule; API - Altura da primeira inflorescência; CERi - Comprimento do entrenó de ramos com inflorescência; CFO - Comprimento da folha; CLFo - Relação entre o comprimento e a largura da folha; CMS - Comprimento da semente; CPI - Comprimento do pedúnculo da inflorescência; CRI - Comprimento de ramos com inflorescência; DCA - Diâmetro do caule; LFR - Largura do fruto; LSE - Largura da semente; NLF - Número de lóbulos foliar; NSE - Número de sementes por planta; NSF - Número de sementes por fruto; PCOo - Projeção da copa no sentido Oeste; e TFF - Tempo de formação do fruto. Descritores e valores em negrito relativos aos critérios de importância e contribuição de Strapasson et al. (2000) e Cruz e Carneiro (2003).

Sendo assim, tem-se que, no presente trabalho, a primeira análise multivariada, acerca dos componentes principais, possibilitou a identificação de 20 descritores passíveis de descarte, reduzindo o conjunto original de 36 para 16 variáveis quantitativas discriminatórias (PCOo, DCA, CRI, CERi, NLF, AIFc, CFO, CLFo, API, CPI, TFF, LFR, NSF, NSE, CMS e LSE), ao passo que a segunda aplicação desta técnica permitiu a detecção dos 10 caracteres, dentre os 16 restantes, que mais contribuíram para a variabilidade observada (TFF, NSE, CFO, CPI, DCA, CLFo, PCOo, LFR, API e LSE, respectivamente). No entanto, como demonstrado na Tabela 11, tal metodologia não proporcionou, novamente, o acúmulo da máxima variação até os 3 primeiros componentes (80%), como proposto por Cruz e Carneiro

(2003). Dessa forma, de maneira semelhante à primeira, optou-se por não representar graficamente a diversidade discriminada, bem como não dar continuidade aos descartes nesta segunda análise, julgando-se mais prudente, para posteriores estimativas, considerar os 16 descritores previamente selecionados na primeira, visto que, se considerado novo descarte, seguindo o critério de Jolliffe (1972 e 1973), variáveis importantes do ponto de vista agrônomo da cultura, como o número de sementes por planta (NSE), o número de sementes por fruto (NSF) e o tempo de formação do fruto (TFF), de interesse ao melhoramento genético da espécie, poderiam ser excluídas, tornando a metodologia de descarte inconsistente. Sob esta perspectiva, Cruz e Carneiro (2003) enaltecem que a tomada de decisão sobre qual variável descartar deve ser baseada, não somente, por critérios estatísticos, mas, também, pela manutenção de caracteres de interesse ao melhoramento genético da espécie, como, por exemplo, os de importância reprodutiva e agrônoma. Dessa maneira, assim como Chiorato (2004), descritores de interesse agrônomo para o melhoramento da espécie, além dos componentes principais, tiveram suas correlações consideradas como critério auxiliar ao método de descarte.

Ainda relativo ao descarte, o registro de caracteres qualitativos também auxiliou no reconhecimento das variáveis binárias e multicategóricas de menor peso para a variabilidade genética da coleção. Por meio deste, verificou-se que as três variáveis binárias caracterizadas não obtiveram variação: cerosidade (CER) e pilosidade da folha (PFO), referentes a características vegetativas, e carúncula da semente (CAR), atributo reprodutivo, bem como as variáveis multicategóricas coloração do caule (CCA), coloração do ramo (CRA) e coloração da nervura foliar (CNF), de caráter vegetativo, e número de lóculos por fruto (NLFr), coloração da semente (CSE), textura da semente (TSE) e padrão da semente (PSE), relacionadas a aspectos reprodutivos. Dessa forma, por não contribuírem efetivamente para a variação observada, assim como as 20 variáveis quantitativas já descartadas, tais descritores também foram desconsiderados nas posteriores análises, pertinentes à estimativa da diversidade genética da coleção em questão.

Finalizando esta etapa, conclui-se que, dos 52 descritores aplicados, apenas os 22 que mais contribuíram para a variação genética persistiram, totalizando 13 caracteres relacionados a aspectos vegetativos e 9 a reprodutivos, onde 16 destes dizem respeito a variáveis quantitativas e 6 a multicategóricas.

4.2.2. Correlação fenotípica e análise de trilha

A Tabela 12 descreve as estimativas de correlação fenotípica, mensuradas a partir dos 16 caracteres quantitativos selecionados. Através desta, observa-se que as correlações dos descritores relativos às sementes (NSE e NSF), sobre a característica tempo de formação de fruto (TFF), representaram os coeficientes de maior magnitude, descrevendo um valor de -0,51 para NSE x TFF e de -0,57 para NSF x TFF.

Tabela 12. Coeficientes de correlação fenotípica entre as 16 variáveis quantitativas de maior importância para a variabilidade genética encontrada em 60 genótipos da Coleção de germoplasma de *Jatropha curcas* da UFRRJ.

Variável	PCOo	DCA	CRI	CERi	NLF	AIFc	CFO	CLFo	API	CPI	TFF	LFR	NSF	CMS	LSE	NSE
PCOo	1,00	0,00	-0,13	0,03	0,07	0,07	-0,02	-0,18	-0,05	0,01	-0,30*	0,22	0,10	-0,03	0,14	0,42**
DCA		1,00	-0,00	0,23	-0,13	0,11	0,05	-0,19	0,09	0,06	-0,03	-0,04	0,09	0,26*	-0,00	0,24
CRI			1,00	-0,12	-0,16	-0,08	-0,05	-0,00	-0,01	-0,21	0,07	0,05	-0,13	0,01	0,18	-0,28*
CERi				1,00	0,30*	0,23	-0,22	-0,24	0,02	-0,10	-0,14	-0,25	0,15	-0,15	0,16	0,26*
NLF					1,00	-0,00	0,15	-0,07	-0,10	-0,15	-0,07	0,15	0,24	-0,18	-0,14	0,14
AIFc						1,00	-0,20	0,02	-0,13	-0,15	-0,13	-0,18	0,08	0,00	0,09	-0,02
CFO							1,00	0,07	0,11	0,10	-0,29*	0,18	0,26*	0,09	0,06	0,19
CLFo								1,00	0,01	0,07	0,03	0,00	0,13	0,21	-0,14	-0,29*
API									1,00	0,03	0,16	0,16	-0,14	0,04	0,09	0,02
CPI										1,00	0,31*	-0,03	-0,14	0,27*	-0,28*	0,08
TFF											1,00	-0,03	-0,57**	0,15	-0,36**	-0,51**
LFR												1,00	-0,03	0,13	0,10	-0,01
NSF													1,00	0,04	0,04	0,29*
CMS														1,00	0,14	-0,18
LSE															1,00	0,08
NSE																1,00

Legenda: PCOo - Projeção da copa no sentido Oeste; DCA - Diâmetro do caule; CRI - Comprimento de ramos com inflorescência; CERi - Comprimento do entrenó de ramos com inflorescência; NLF - Número de lóbulos foliar; AIFc - Ângulo de inserção da folha no caule; CFO - Comprimento da folha; CLFo - Relação entre o comprimento e a largura da folha; API - Altura da primeira inflorescência; CPI - Comprimento do pedúnculo da inflorescência; TFF - Tempo de formação do fruto; LFR - Largura do fruto; NSF - Número de sementes por fruto; CMS - Comprimento da semente; LSE - Largura da semente; e NSE - Número de sementes por planta. *, ** Significativo a 5% e 1% de probabilidade pelo teste *t*, respectivamente.

Do ponto de vista agrônomo, é de se notar associações de grande importância para o melhoramento genético de *Jatropha curcas*, como, principalmente, as significativas correlações dos descritores vegetativos PCOo e CRI, em conjunto com os reprodutivos TFF e NSF, sobre o número de sementes produzido (Tabela 12). De acordo com Rao et al. (2008), tanto o diâmetro quanto o volume da copa constituem, junto às características “peso de sementes” e “número de ramos”, os principais componentes da produtividade do pinhão-manso.

Como já esperado, pode-se perceber valores de correlação positivos de PCOo e NSF com NSE, seguidamente de 0,42 e 0,29, visto que plantas com maior área de projeção de copa possuem, igualmente, maior número de ramos, contribuindo, possivelmente, para o aumento de sua produção. Já a característica comprimento de ramos com inflorescência demonstrou negativo efeito sobre NSE (-0,28), validando a afirmação de Saturnino et al. (2005). Segundo estes autores, o pinhão-manso produz inflorescências em gemas terminais de ramos crescidos no ano corrente, fazendo com que o aumento de sua produção de frutos seja altamente dependente do maior número de ramos na planta. Dessa forma, visando o aumento de produção, é preferível investir na seleção direcionada ao maior número de ramos do que ao maior comprimento de ramos com inflorescência. A mesma observação pode ser feita em relação ao descritor NSF: quanto maior o número de sementes por fruto, maior também será a quantidade de sementes produzidas por planta.

O número de sementes por fruto ainda obteve correlação significativa com o comprimento da folha (0,26), e este, por sua vez, relação negativa com TFF (-0,29), relatando que quanto maior o campo foliar verde, menor deverá ser o tempo de formação do fruto. Assim como CFO, o descritor PCOo também registrou significativo coeficiente negativo, de magnitude de -0,30, sobre TFF, indicando que, nas condições expressas pela Coleção de germoplasma da UFRRJ, plantas portadoras de ampla área de copa projetada podem ser utilizadas como critério de seleção indireta na diminuição do tempo gasto para a colheita. Tal resultado, juntamente aos valores de correlação negativos, obtidos por TFF sobre NSF (-0,57), LSE (-0,36) e NSE (-0,51), podem ser considerados de grande valia para o melhoramento genético da espécie, representando um significativo passo em direção à obtenção de variedades mais produtivas e uniformes quanto ao período de colheita, uma vez que plantas que demoraram menos para frutificar alcançaram um maior número de sementes por fruto e, em consequência, maior produção de sementes no geral, assim como sementes de maior largura, fato este que, talvez, possa conferir maior teor de óleo para extração.

Resultados de menor expressão foram observados nas correlações de DCA com CMS (0,26), já esperada, de NLF com CERi (0,30), CLFo e NSE (-0,29) e de CPI em relação aos descritores CMS (0,27), LSE (-0,28) e TFF (0,31). Nesse contexto, percebeu-se, de maneira pontual, que o acesso 062, procedente de Jaíba - MG, caracterizado por folhas desenvolvidas de coloração verde claro e um longo pedúnculo de inflorescência, apresentou, fielmente, estas 3 últimas associações: excessivo tempo para frutificar e frutos/sementes mais delgados(as) e compridos(as), quando comparado à grande maioria dos genótipos da coleção. Genótipos possuidores destas mesmas características, como o acesso 062, também foram descritos por Laviola et al. (2011), em trabalho sobre a caracterização morfoagronômica da Coleção de germoplasma de *J. curcas* da EMBRAPA Agroenergia.

Resultado inesperado foi encontrado na correlação positiva entre CERi e NSE, de 0,26, visto que, como já comentado, quanto maior o comprimento de ramos com inflorescência, menor é o número de sementes gerado. Desse modo, um comprimento de entrenó mais intervalado deveria ser responsável por uma diminuição no número de sementes produzidas pela planta. No entanto, a decomposição das estimativas de correlação, realizada pela análise de trilha, em efeitos diretos e indiretos (sobre a variável principal NSE), demonstrou, por meio da Tabela 13, que o comprimento do entrenó de ramos com

inflorescência apresentou um baixo efeito direto sobre o número de sementes produzido (0,09), relatando que, apesar do coeficiente significativo (0,26), apontado pela correlação fenotípica, esta variável proporcionou pouca influência direta sobre NSE, assim, não sendo determinante para o número de sementes gerado. Ainda sob este aspecto, também é possível notar a negativa participação direta, de -0,13, de CRI sobre NSE, corroborando com o coeficiente de correlação fenotípica encontrado (-0,28), igualmente em NSE, e com as afirmações de Saturnino et al. (2005) e Rao et al. (2008), anteriormente discutidas.

Tabela 13. Resumo das estimativas de efeitos diretos e indiretos de 15 variáveis quantitativas sobre a característica Número de sementes por planta (NSE) em 60 genótipos de *Jatropha curcas* da UFRRJ.

Variável	Efeitos indiretos															Efeitos diretos NSE
	PCOo	DCA	CRI	CERi	NLF	AIFc	CFO	CLFo	API	CPI	TFF	LFR	NSF	CMS	LSE	
PCOo	-	0,00	0,02	0,00	0,00	-0,01	-0,00	0,03	-0,00	0,00	0,15	-0,01	-0,00	0,00	-0,01	0,24
DCA	0,00	-	0,00	0,02	-0,01	-0,01	0,00	0,03	0,01	0,01	0,02	0,00	-0,00	-0,03	0,00	0,19
CRI	-0,03	-0,00	-	-0,01	-0,01	0,01	-0,00	0,00	-0,00	-0,04	-0,03	-0,00	0,00	-0,00	-0,01	-0,13
CERi	0,00	0,04	0,02	-	0,02	-0,02	-0,01	0,04	0,00	-0,02	0,07	0,02	-0,00	0,02	-0,01	0,09
NLF	0,02	-0,03	0,02	0,03	-	0,00	0,00	0,01	-0,01	-0,03	0,04	-0,01	-0,00	0,02	0,01	0,07
AIFc	0,02	0,02	0,01	0,02	-0,00	-	-0,00	-0,00	-0,01	-0,03	0,07	0,01	-0,00	-0,00	-0,00	-0,11
CFO	-0,00	0,01	0,01	-0,02	0,01	0,02	-	-0,01	0,01	0,02	0,14	-0,01	-0,00	-0,01	-0,00	0,03
CLFo	-0,04	-0,04	0,00	-0,02	-0,00	-0,00	0,01	-	0,00	0,01	-0,01	-0,00	-0,00	-0,03	0,01	-0,16
API	-0,01	0,02	0,00	0,00	-0,01	0,01	0,00	-0,00	-	0,01	-0,08	-0,01	0,00	-0,00	-0,00	0,10
CPI	0,00	0,01	0,03	-0,01	-0,01	0,02	0,00	-0,01	0,00	-	-0,15	0,00	0,00	-0,03	0,01	0,21
TFF	-0,07	-0,01	-0,01	-0,01	-0,00	-0,01	-0,01	-0,00	0,01	0,07	-	0,00	0,01	-0,02	0,02	-0,50
LFR	0,05	-0,01	-0,01	-0,02	0,01	0,02	0,00	-0,00	0,01	-0,01	0,02	-	0,00	-0,02	-0,00	-0,06
NSF	0,02	0,02	0,02	0,01	0,02	-0,01	0,01	-0,02	-0,01	-0,03	0,29	0,00	-	-0,00	-0,00	-0,01
CMS	-0,01	0,05	-0,00	-0,01	-0,01	-0,00	0,00	-0,03	0,00	0,06	-0,08	-0,01	-0,00	-	-0,01	-0,13
LSE	0,03	-0,00	-0,02	0,01	-0,01	-0,01	0,00	0,02	0,01	-0,06	0,18	-0,01	-0,00	-0,02	-	-0,05

Legenda: PCOo - Projeção da copa no sentido Oeste; DCA - Diâmetro do caule; CRI - Comprimento de ramos com inflorescência; CERi - Comprimento do entrenó de ramos com inflorescência; NLF - Número de lóbulos foliar; AIFc - Ângulo de inserção da folha no caule; CFO - Comprimento da folha; CLFo - Relação entre o comprimento e a largura da folha; API - Altura da primeira inflorescência; CPI - Comprimento do pedúnculo da inflorescência; TFF - Tempo de formação do fruto; LFR - Largura do fruto; NSF - Número de sementes por fruto; CMS - Comprimento da semente; LSE - Largura da semente; e NSE - Número de sementes por planta. Coeficiente de determinação = 0,56; Efeito da variável residual = 0,66.

Através dos coeficientes de trilha, ilustrados na Tabela 13, pode-se observar que, com exceção dos descritores TFF (-0,50) e PCOo (0,24), que relataram um moderado efeito direto sobre o número de sementes produzido, todas as outras características aqui estudadas exerceram baixo efeito direto sobre a variável principal. Do mesmo modo, a decomposição das estimativas de correlação demonstrou que, igualmente ao descritor PCOo, os descritores CFO, CPI, NSF e LSE foram os únicos que contribuíram de forma indireta para o número de sementes encontrado, com valores de efeito indireto de 0,15, 0,14, -0,15, 0,29 e 0,18, todos via TFF, respectivamente, sobre NSE. Nesse sentido, cabe ressaltar que, apesar de não ter apresentado valor de correlação significativo com NSE, o descritor CPI, juntamente aos descritores DCA, AIFc, API e CMS, demonstraram efeito direto sobre esta variável, de 0,21, 0,19, -0,11, 0,10 e -0,13, seguidamente. De forma oposta, o número de sementes por fruto, que obteve significativa correlação positiva com NSE, descreveu um negativo e baixo efeito direto junto a esta variável, de -0,01. Enquanto que, em consonância ao coeficiente de correlação obtido (-0,29), o descritor CLFo exerceu um negativo efeito direto sobre NSE, de -0,16.

Por fim, deve-se ressaltar, mais uma vez, a importância dos descritores TFF e PCOo, relativos ao tempo de formação do fruto e à projeção da copa no sentido Oeste, para os resultados aqui alcançados, que, através das estimativas de efeito direto, da ordem de -0,50 e 0,24, respectivamente, sobre NSE, reiteraram a possibilidade de, na Coleção de germoplasma da UFRRJ, serem empregados como critério de seleção indireta para a obtenção de variedades de menor tempo de colheita e, simultaneamente, maior produção de grãos, principais alvos do melhoramento genético de *Jatropha curcas*. Em contraste aos resultados aqui descritos, Rao et al. (2008) observaram indesejáveis efeitos diretos positivos, de 0,26 e 0,43, para caracteres referentes ao período exigido para frutificação (tempo requerido da floração à frutificação e dias contados para maturação) sobre a produção de grãos por planta. Resultados semelhantes aos observados no presente trabalho, relativos à projeção da copa, também foram notados por Laviola et al. (2011), onde os descritores de projeção de copa na linha e entre linhas apresentaram efeitos diretos de 0,25 e 0,21 sobre o rendimento de grãos.

4.2.3. Estimação da diversidade genética e consistência do padrão de agrupamento

De posse destas informações, utilizou-se os 22 descritores remanescentes (16 quantitativos e 6 multicategóricos) para estimar a diversidade genética entre os 60 genótipos de *Jatropha curcas*, previamente definidos a partir da Coleção de germoplasma da UFRRJ. Para tal, seguindo as recomendações de Azambuja (2005), que ressalta a importância de se testar as diversas medidas de distância e métodos de agrupamento disponíveis, para que, com base na consistência do padrão aglomerativo obtido, possa melhor comparar os grupos formados, avaliou-se a combinação do coeficiente geral de Gower e das distâncias Euclidiana, de Chebyshev, Manhattan e de Canberra com os métodos de agrupamento hierárquico UPGMA e Vizinho mais próximo. Dessa maneira, aplicou-se o coeficiente geral de Gower, para a análise individual de variáveis multicategóricas e sua análise conjunta entre estas e as variáveis quantitativas, ao passo que empregou-se as demais distâncias à análise individual destas últimas.

A Tabela 14 apresenta os resultados referentes à consistência dos agrupamentos obtidos a partir das combinações entre as diferentes matrizes de distância e técnicas de agrupamento testadas. Através desta, observa-se que, em relação aos métodos de agrupamento aplicados, o padrão de grupos gerados pela média das distâncias entre todos os pares de genótipos, dado pela técnica UPGMA, apresentou maior consistência quando comparada à metodologia do Vizinho mais próximo, que faz uso do menor elemento do conjunto de distâncias dos pares de indivíduos para a formação dos grupos. Nesse sentido, pode-se notar

que os coeficientes de correlação cofenética (CC), aglomerativo (CA) e os níveis de distorção (DT) e stress (ST), empregados como parâmetros na avaliação da consistência das técnicas testadas, apresentaram, sem exceção, valores superiores quando as matrizes de distância estimadas foram agrupadas com base na metodologia conferida pelo método UPGMA (Tabela 14).

Tabela 14. Consistência do padrão de agrupamento proporcionado pelas técnicas do Vizinho mais próximo e UPGMA, em relação ao coeficiente geral de Gower, para a análise conjunta de todas as variáveis e individual de variáveis multicategóricas, e às distâncias Euclidiana, de Chebyshev, Manhattan e de Canberra, para variáveis quantitativas, em função do coeficiente de correlação cofenética, aglomerativo e níveis de distorção e stress, em 60 genótipos da Coleção de germoplasma de *Jatropha curcas* da UFRRJ.

Variáveis	Medidas	Agrupamento								Nº de descritores
		Vizinho mais próximo				UPGMA				
		CC	CA	DT	ST	CC	CA	DT	ST	
Conjunta	Gower	73	60	36	40	81	64	0	17	22
Multicategóricas	Gower	81	94	38	50	84	95	0	29	6
	Euclidiana	81	58	28	31	82	61	0	15	
Quantitativas	Chebyshev	74	67	29	36	85	70	0	18	16
	Manhattan	75	49	32	36	80	58	0	16	
	Canberra	79	45	35	39	80	55	0	18	

Legenda: CA - Coeficiente aglomerativo (%); CC - Coeficiente de correlação cofenética (%); DT - Distorção (%); e ST - Stress (%).

Segundo Sokal e Rohlf (1962), quanto mais perto da unidade for o coeficiente de correlação cofenética (CC), maior também será o ajuste entre a representação gráfica da matriz de distância genética original (matriz fenética) e a matriz de agrupamento obtida (matriz cofenética). Com base nos resultados apontados pela Tabela 14, aferiu-se, no presente estudo, que a análise individual de variáveis quantitativas, estimada pela distância de Chebyshev junto ao método de agrupamento UPGMA, apresentou a melhor adequação entre estas matrizes, com um CC de 85%, seguida, com um CC de 84%, da combinação entre o algoritmo de Gower e, de maneira semelhante, da técnica UPGMA, para a análise individual de variáveis multicategóricas. De forma oposta, o ensaio entre a matriz gerada pelo coeficiente geral de Gower junto ao método do Vizinho mais próximo, para a análise conjunta de variáveis quantitativas e multicategóricas, reproduziu a menor consistência entre as demais combinações testadas, com um CC de 73%. Estes resultados estão de acordo com as considerações de Sokal e Rohlf (1962), nas quais afirmam que, dentre os métodos hierárquicos de agrupamento, o UPGMA é o que apresenta coeficientes de correlação cofenética máximos. Tal eficiência é dada, possivelmente, à própria metodologia da técnica UPGMA, na qual os grupos são formados em função de médias aritméticas da medida de dissimilaridade empregada, ou seja, não utiliza valores extremos, como os métodos do Vizinho mais próximo e do Vizinho mais distante, que fazem referência, respectivamente, ao menor e ao maior elemento das matrizes de dissimilaridade, fato este que pode resultar em uma maior deturpação entre a matriz fenética e a matriz gerada pelo método de agrupamento (cofenética).

Carvalho (2010), ao estimar a divergência genética entre progênies de polinização aberta de *J. curcas*, por meio da distância generalizada de Mahalanobis e a técnica de agrupamento UPGMA, relatou um coeficiente cofenético de 64%, ilustrando, deste modo, um baixo ajuste entre as matrizes original e a cofenética. Em estudo molecular sobre a diversidade genética entre genótipos de *J. curcas*, Santos et al. (2010a) observaram um valor de 91% para o coeficiente cofenético expresso pela matriz de distância do coeficiente de Jaccard junto ao método UPGMA. Há de se destacar, nos resultados aqui discutidos, a eficiência demonstrada pelo algoritmo de Gower, quando combinado ao método UPGMA, na

análise simultânea de variáveis multicategóricas e quantitativas, apontada pelo CC de 81%, evidenciando, portanto, a viabilidade no emprego desta técnica para a análise conjunta de dados provenientes de diferentes tipos de variáveis. Tal resultado também foi ressaltado por Santos (2010), com a obtenção de um CC de 90%, na avaliação simultânea de dados quantitativos, moleculares e multicategóricos para a estimação da diversidade genética entre genótipos de maracujazeiro.

Proposto por Vale (2005), o coeficiente aglomerativo (CA), empregado de maneira auxiliar na aferição da qualidade dos grupos formados, ainda é pouco utilizado em estudos de diversidade genética. Contudo, assim como o coeficiente cofenético, tal parâmetro também pode ser considerado de grande valia ao se inferir sobre a consistência das técnicas aplicadas, uma vez que, junto ao próprio coeficiente de correlação cofenética e aos níveis de distorção e stress, ajuda a ampliar a base de dados a serem considerados para a definição da melhor reprodução dos resultados atingidos. Através da Tabela 14, é possível observar que a combinação entre o coeficiente geral de Gower e o método aglomerativo UPGMA, para a análise individual de caracteres multicategóricos, foi o procedimento que proporcionou o maior coeficiente aglomerativo, de 95%, em relação às outras metodologias testadas, demonstrando, dessa forma, uma boa definição para as bifurcações geradas. No entanto, esta combinação obteve um nível de stress maior que as outras avaliadas (29%), o que relata um ajuste variando entre o regular e o insatisfatório, segundo Kruskal (1964), e fora do padrão estipulado por Cruz e Carneiro (2003), onde somente valores abaixo de 20% representariam uma boa adequação entre a matriz originalmente estimada e a gerada pela técnica de agrupamento. Opostamente ao stress, os níveis de distorção foram nulos para todas as distâncias combinadas ao método UPGMA, denotando uma excelente adequação entre as matrizes fenética e cofenética estimadas, como sugerido por Kruskal (1964). Em consonância ao presente trabalho, valores de stress acima do recomendado por Kruskal (1964) e Cruz e Carneiro (2003), de 36,78%, aliado a baixos níveis de distorção (13,53%), também foram observados por Carvalho (2010), ao estimar a diversidade genética de *J. curcas* com base em 8 caracteres morfológicos.

De modo geral, percebe-se, nos resultados aqui obtidos, que a combinação entre as matrizes de distância de Chebyshev e o procedimento aglomerativo UPGMA, para variáveis quantitativas, reportou a maior consistência entre os agrupamentos produzidos, juntamente à conjugação entre o algoritmo de Gower e o mesmo método de agrupamento, utilizados para a análise individual de caracteres multicategóricos. Entretanto, sustentado pelo coeficiente cofenético, um dos parâmetros mais empregados na avaliação da consistência em estudos de diversidade, pode-se inferir, por meio da Tabela 14, que todas as combinações aqui testadas, utilizando o método UPGMA, reproduziram um bom ajuste na discriminação dos grupos gerados (CC > 80%). Através desta, também é possível notar, a partir dos valores de CC e CA, que a consistência do padrão de agrupamento tende a diminuir à medida que mais descritores passam a ser avaliados. Efeito semelhante aos valores aqui descritos já foi explanado por Cargnelutti Filho e Guadagnin (2011), em estudo sobre a consistência do padrão de agrupamento em cultivares de milho. Tais resultados ressaltam a importância de sempre verificar as diferentes metodologias multivariadas a serem combinadas, bem como conhecer e analisar as propriedades intrínsecas às distâncias disponíveis, como já enfatizado por Bussab et al. (1990), Malhotra (2001) e Azambuja (2005).

Como retratado neste trabalho, a maior consistência, dada através da distância de Chebyshev em relação à Euclidiana, no âmbito do melhoramento, pode ser interpretada como uma importante consideração para posteriores avaliações, visto que esta última é a medida comumente mais utilizada em estudos de diversidade genética. No que diz respeito somente às distâncias relativas a variáveis quantitativas, utilizando o método UPGMA, observa-se que o aumento da consistência da técnica deu-se na crescente ordem: Canberra, Manhattan,

Euclidiana e Chebyshev (Tabela 14). Apesar do alto coeficiente de correlação obtido entre as matrizes de distância Euclidiana e de Manhattan, de 0,95, tal relação não se refletiu sobre os valores de consistência alcançados por estas, visto que os resultados referentes à distância Euclidiana se mostraram melhores que os atingidos pela medida de Manhattan, tanto quando combinada à técnica UPGMA quanto a do Vizinho mais próximo, como demonstrado na Tabela 15.

Tabela 15. Correlação entre as matrizes de distância Euclidiana, de Chebyshev, Manhattan e de Canberra, estimadas a partir de 16 descritores quantitativos, em 60 genótipos de *Jatropha curcas* da UFRRJ.

Matrizes	Euclidiana	Chebyshev	Manhattan	Canberra
Euclidiana	1,00	0,85	0,95	0,73
Chebyshev		1,00	0,68	0,43
Manhattan			1,00	0,82
Canberra				1,00

Em trabalho comparando diferentes medidas de distâncias e técnicas de agrupamento, Cargnelutti Filho e Guadagnin (2011) observaram que a distância Euclidiana proporciona maior consistência no padrão de agrupamento em relação à distância de Manhattan, sendo que a robustez do método aumenta na seguinte ordem: Ward, Vizinho mais distante, Vizinho mais próximo e UPGMA. Ao avaliar a consistência dos agrupamentos formados a partir de cultivares de feijão, Cargnelutti Filho et al. (2010) constataram que os grupos obtidos através da distância Euclidiana, com o método UPGMA, demonstram maior consistência quando comparado às suas derivações (Euclidiana média, padronizada, média padronizada, quadrado da distância Euclidiana e quadrado da distância Euclidiana padronizada) e às distâncias generalizadas de Mahalanobis e Mahalanobis padronizada. Sabendo-se que a distância de Chebyshev (*Maximum*) se vale da máxima diferença entre as variáveis multidimensionais de dois elementos (AZAMBUJA, 2005), os resultados aqui apresentados se tornam ainda mais importantes. Isto porque, em estudos de diversidade, onde preconiza-se a identificação dos genitores geneticamente mais distantes (para um melhor aproveitamento da heterose), as metodologias propostas pelas distâncias aqui utilizadas, quando não testadas suas consistências, poderão compor uma matriz de valores sub ou superestimados, aumentando o risco de se proporcionar agrupamentos distorcidos e distantes da realidade descrita pelos genótipos, de acordo com a técnica aglomerativa aplicada. No entanto, em função dos resultados aqui demonstrados (Tabela 14), sugere-se que o emprego da técnica delineada por Chebyshev, assim como a distância Euclidiana, também pode alcançar resultados robustos, quando avaliada com consistência, na busca dos genitores mais apropriados à obtenção de híbridos de maior efeito heterótico.

Fundamentada na consistência dos agrupamentos aqui conferidos, decidiu-se utilizar a combinação de todas as medidas de distância testadas, com exceção das distâncias de Manhattan e Canberra (de menor consistência), junto à técnica de agrupamento UPGMA, a fim de se representar graficamente a diversidade genética entre os 60 acessos de *Jatropha curcas* descritos, da Coleção de germoplasma da UFRRJ.

Aplicando a metodologia de corte proposta por Kelley et al. (1996), baseada na média das distâncias da matriz cofenética para a delimitação dos grupos formados, nota-se, sem exceções, a partir da Figura 1, que a localização geográfica não foi determinante para a disposição final dos agrupamentos gerados, uma vez que o número de grupos, para cada combinação testada, não se igualou, em nenhum momento, ao total das 12 procedências avaliadas. Tal resultado se mostra de acordo com os estudos de Rosado et al. (2010), que avaliaram, via marcadores RAPD e SSR, 192 genótipos de *J. curcas* coletados de diversas localidades do território brasileiro. Para estes autores, o descompasso entre os diferentes grupos formados e suas reais origens geográficas, no Brasil, pode ser consequência direta da

interferência humana sobre seu processo dispersivo, via propagação por sementes e/ou estacas. Basha e Sujatha (2007) e Rao et al. (2008), ao estimarem a diversidade genética de *J. curcas* de diferentes regiões da Índia, também descreveram comportamento semelhante, detectando a ocorrência de genótipos de procedências distintas reunidas sob um mesmo grupo. Entretanto, em dissonância com o padrão aqui registrado, Abreu et al. (2009) relatam que as diferentes localidades podem, sim, ser preponderantes para a organização dos agrupamentos formados. Em trabalho sobre a variabilidade genética do Banco ativo de germoplasma de *J. curcas* da Universidade Federal de Mato Grosso do Sul (UFMS), estes autores, utilizando o método de Tocher, observaram que a separação dos 2 grupos gerados (a partir da caracterização de 10 genótipos aos 3 meses de idade) se deu segundo sua origem geográfica, alocando, no primeiro, genótipos originários dos Estados de Minas Gerais e Goiás e, no segundo, acessos de regiões localizadas mais ao Sul do país, como o Paraguai e os Estados de São Paulo e Paraná. Em concordância com Abreu et al. (2009), Santos et al. (2010a) constataram que a variabilidade genética entre 50 genótipos de *J. curcas*, de 12 diferentes localidades do Brasil e do mundo, também era geneticamente estruturada em função de suas origens.

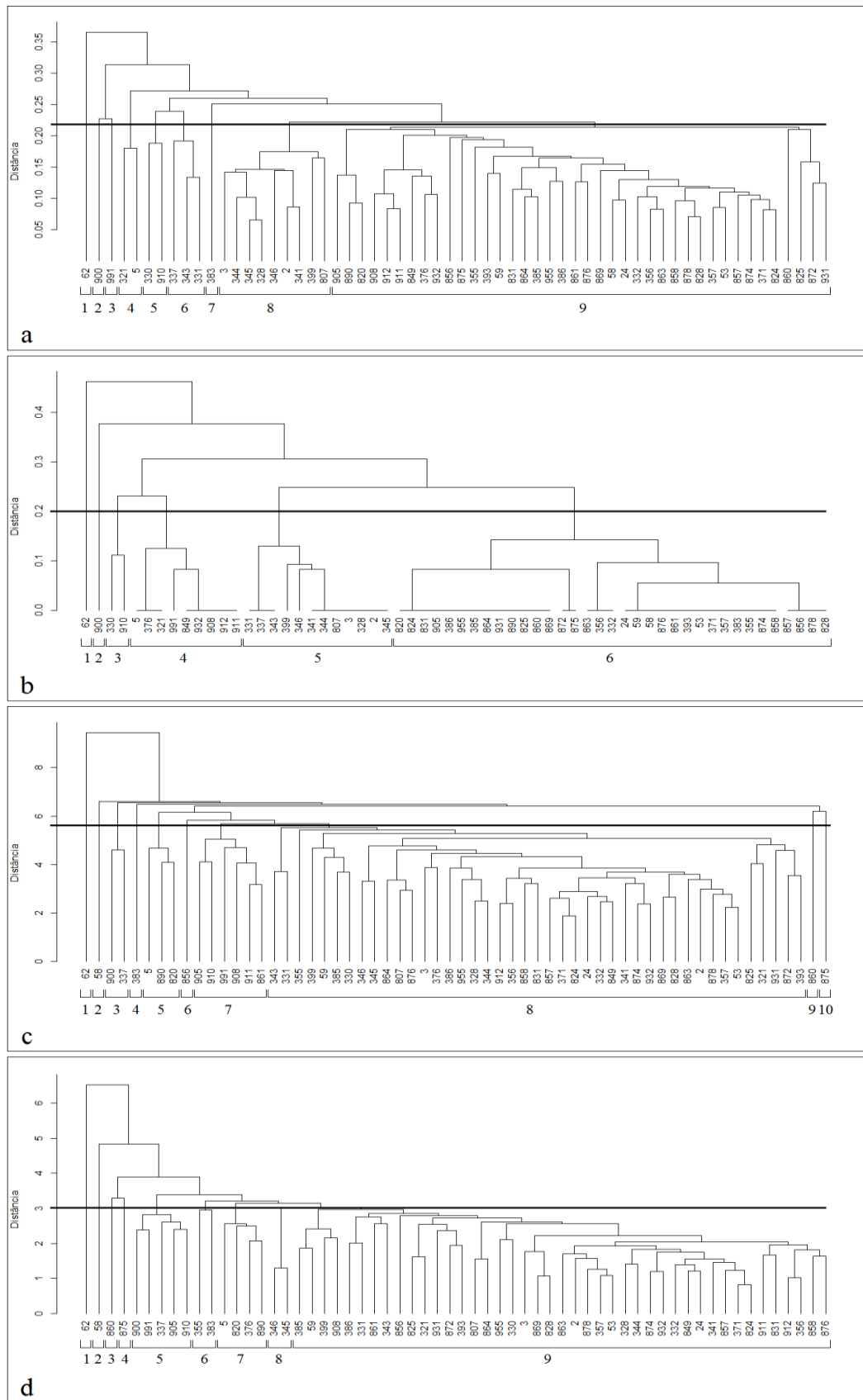


Figura 1. Dendrogramas gerados através do método de agrupamento UPGMA, a partir do algoritmo de Gower, para as análises conjunta de variáveis quantitativas e multicategóricas (a) e individual de variáveis multicategóricas (b), e em função das distâncias Euclidiana (c) e de Chebyshev (d), para a análise individual de variáveis quantitativas.

No que tange à composição da matriz conjunta (variáveis quantitativas avaliadas simultaneamente às multicategóricas), inferi-se que os descritores quantitativos obtiveram maior importância relativa, quando comparados aos multicategóricos, para a constituição desta (Figura 1). Tal inferência pode ser justificada pelo semelhante número de grupos conferidos entre a análise simultânea, através do coeficiente geral de Gower (nove grupos) (Figura 1a), e a individual de variáveis quantitativas, por meio das distâncias Euclidiana e de Chebyshev, seguidamente de 10 e 9 grupos (Figura 1c e 1d), em relação aos seis agrupamentos obtidos pelo algoritmo de Gower na análise individual de variáveis multicategóricas (Figura 1b). Entretanto, ressalvas podem ser consideradas aos acessos 062 e 900, formadores dos grupos individuais 1 e 2, e dos acessos 330 e 910, do grupo 5, onde os caracteres multicategóricos podem ter exercido maior peso sobre a organização final do dendrograma (Figura 1a). Isto porque, em termos de alocação, estes mesmos genótipos compuseram, de modo idêntico, os grupos singulares 1 e 2 e o grupo 3 na análise individual de variáveis multicategóricas (Figura 1b). Nesse sentido, deve-se ressaltar que tanto o acesso 062, do Estado de Minas Gerais, quanto o genótipo de numeração 900, de Dourados (Mato Grosso do Sul), são detentores de atributos morfológicos únicos em relação aos demais acessos da referida coleção.

O acesso 062, dentre os 60 genótipos avaliados, foi o único responsável por frutos de formato oval e folhas desenvolvidas de coloração verde claro, além de contar com sementes e pedúnculo da inflorescência de maior comprimento. Em trabalho sobre a caracterização morfoagronômica de um Banco ativo de germoplasma de pinhão-manso, Laviola et al. (2011) descreveram plantas com estas mesmas características, observando, nestas, a ausência de ésteres de forbol em seus grãos. Nesse sentido, o resultado apresentado se faz de grande importância para o melhoramento genético de *J. curcas*, visto que genótipos como o acesso 062 podem ser fonte de variabilidade genética para o futuro desenvolvimento de cultivares comerciais não-tóxicas. Já o acesso 900 foi o único a apresentar ramificação do tipo monopodial, quando a grande maioria das plantas descreveu ramificação simpodial. A partir destes resultados, pode-se inferir que tanto os descritores vegetativos relativos ao tipo de ramificação (RAM) e coloração de folhas desenvolvidas (CFD), quanto ao descritor reprodutivo referente ao formato do fruto (FFR), bem como ao comprimento da semente (CMS) e pedúnculo da inflorescência (CPI), podem ser empregados como importantes marcadores morfoagronômicos para a discriminação da diversidade genética na espécie *J. curcas*.

As Figuras 1c e 1d dizem respeito, respectivamente, às análises individuais de variáveis quantitativas via distância Euclidiana e de Chebyshev. Nestas, é possível perceber um padrão de agrupamento similar, na qual a distância Euclidiana proporcionou um dendrograma portador de 10 grupos, ao passo que a técnica de Chebyshev apresentou a formação de 9. Tal semelhança pode ser explicada através do alto valor de correlação entre estas matrizes, da ordem de 85%, dessa forma, denotando certa correspondência para os valores de distância aferidos (Tabela 15).

Assim como os dendrogramas relativos à análise conjunta e individual de variáveis multicategóricas, os dendrogramas gerados a partir das distâncias Euclidiana e de Chebyshev também demonstraram agrupamentos idênticos. Sob este aspecto, nota-se que os grupos individuais 1, 2, 9 e 10, pertinentes à análise através da distância Euclidiana (Figura 1c), foram compostos pelos mesmos acessos (062, 058, 860 e 875, seguidamente) correspondentes aos agrupamentos 1, 2, 3 e 4, dispostos no dendrograma obtido pela medida de Chebyshev (Figura 1d). No entanto, apesar da grande similaridade, algumas modificações foram verificadas. Por meio dos dendrogramas apontados nas Figuras 1c e 1d, observa-se que os acessos 905, 910, 991, 908, 911 e 861, formadores do grupo 7 na análise realizada pela

distância Euclidiana, tiveram os acessos 908, 911 e 861 deslocados para o agrupamento 9, juntos à maioria, quando avaliados pela distância de Chebyshev. Ainda aqui, é possível notar que o grupo 5, dos mesmos acessos 905, 910 e 991, dado pela análise de Chebyshev, recebeu os acessos 900 e 337, antes reunidos sob o grupo 3 na avaliação promovida pela distância Euclidiana, se tornando, assim, o grupo com os genótipos de menor produção. Do mesmo modo, o grupo 5 da análise via distância Euclidiana, formado pelos acessos 005, 890 e 820, teve sua organização modificada, com a entrada do acesso 376, quando posteriormente avaliado pela distância de Chebyshev, constituindo, nesta, o grupo 7.

Contrapondo estas mudanças mais evidentes, apurou-se que, tanto para a análise via distância Euclidiana quanto pela medida de Chebyshev, o acesso 860, caracterizado como o mais produtivo dentre os 60 genótipos avaliados, foi discriminado em grupos unitários, formando o grupo 9 quando avaliado pela distância Euclidiana, e o 3 pela de Chebyshev. Já os acessos 345 e 346, concomitantemente o segundo e o terceiro de maior produção, foram ordenados de modo distinto para estas duas análises, sendo colocados em conjunto com a grande parte dos genótipos, pela distância Euclidiana (grupo 8), e separados dos demais, porém em dupla, através da técnica de Chebyshev (grupo 8), como demonstrado nas Figuras 1c e 1d, respectivamente.

Tendo em vista as mudanças anteriormente descritas, observa-se que a distância de Chebyshev, quando comparada à Euclidiana, expressou melhor a produção de grãos dos genótipos estudados, sendo capaz de ressaltar tanto o grupo de menor quanto o de maior produção. Dessa forma, baseado nas modificações impostas pelas distâncias Euclidiana e de Chebyshev, sugere-se, no presente trabalho, que, mesmo apresentando matrizes de grande compatibilidade, a definição da medida a ser utilizada pode resultar na alteração da estrutura de determinados grupos, bem como modificar o ordenamento entre genótipos de maior similaridade. Este comportamento faz com que os descritores número de sementes (NSE) e produção de grãos por planta (PGP), altamente correlacionados, também sejam considerados potenciais marcadores morfoagronômicos para o estudo da diversidade genética em *J. curcas*.

A partir dos índices de consistência alcançados pelo método de agrupamento UPGMA, percebe-se, através dos dendrogramas da Figura 1, uma razoável variabilidade genética para os caracteres aqui descritos. Esta inferência pode ser constatada, sobretudo, através do número de grupos formados nas duas combinações de padrão mais consistentes (9 e 6 grupos, seguidamente, para as análises individuais de variáveis quantitativas, via distância de Chebyshev, e multicategóricas, por meio do algoritmo de Gower), como exposto nas Figuras 1d e 1b. No entanto, as análises de agrupamento também demonstraram a presença de um extenso grupo reunindo grande parte dos genótipos caracterizados, retratando uma limitada diversidade nestes. Na análise conjunta, observa-se 66,67% dos genótipos no grupo 9, na individual de variáveis multicategóricas este valor é de 58,3% (grupo 6), já na análise individual de variáveis quantitativas este número corresponde a 71,67%, nos grupos 8 e 9, respectivamente, através da distância Euclidiana e de Chebyshev (Figura 1). A fim de se ampliar esta base, almejando uma maior diversidade para futuros cruzamentos, a serem realizados pelo Programa de melhoramento da espécie, recomenda-se a inserção de novos materiais genéticos, procedentes de mais localidades e, especialmente, de seus centros de origem e de diversidade.

O cenário aqui expresso delineou-se em consonância às análises descritivas e de componentes principais previamente efetuadas, onde, mesmo após o descarte de 20 variáveis quantitativas, verificou-se que a variação encontrada ainda se apresentava de forma diluída, isto é, homoganeamente distribuída ao longo dos 16 componentes principais formados (Tabela 11). Quando avaliados conjuntamente pelo coeficiente geral de Gower, os 60 acessos de *Jatropha curcas*, da Coleção de germoplasma da UFRRJ, continuaram a apresentar 9 grupos, demonstrando que a variabilidade para os descritores de caráter quantitativo se

mostrou maior que a relativa às características qualitativas, como já discutido, para a diversidade genética observada.

Como o Programa de melhoramento genético de *J. curcas* tem como principal objetivo a obtenção de cultivares produtivas e não-tóxicas, portadoras de características desejáveis para o manejo e sustentabilidade da cadeia produtiva de biodiesel, os resultados aqui apresentados sugerem que os acessos 062 (formador de grupo individual em todas as análises e portador de características relativas à ausência de ésteres de forbol em suas sementes), junto aos acessos 860, 345 e 346 (mais produtivos), se mostraram como os mais promissores a serem envolvidos em futuros cruzamentos. Com base nestes resultados, pode-se inferir que existe, sim, variabilidade genética para os descritores aqui avaliados, demonstrando, para o futuro, a possibilidade de desenvolvimento de cultivares comerciais não-tóxicas e de alta produção, a serem avaliadas para as diferentes regiões de potencial produtor do Brasil.

Também utilizando a técnica de agrupamento UPGMA, Freitas (2010), ao estimar a diversidade genética em 78 genótipos de *J. curcas* (aos 14 meses de idade), via descritores morfoagronômicos, relatou a formação de 15 grupos distintos, onde o grupo 1 reuniu 64,10% das plantas caracterizadas. Estimando a variabilidade genética entre 20 genótipos de pinhão-manso, Carvalho (2010) descreveu a obtenção de 4 grupos de dissimilaridade, onde os grupos 2 e 3 foram responsáveis, cada um, pelo agrupamento de 9 genótipos. Para Rosado et al. (2010), em estudo molecular acerca da diversidade genética entre 192 acessos de *J. curcas* do Banco de germoplasma da EMBRAPA Agroenergia, a base genética dos genótipos nacionais pode ser considerada estreita, uma vez que, como os resultados anteriormente citados, um único grupo concentrou grande parte dos genótipos descritos, expressando, através deste, uma limitada diversidade entre eles. Segundo estes autores, a baixa variabilidade do germoplasma brasileiro pode estar relacionada à sua origem a partir de um ancestral comum, bem como à intensa seleção sobre o material hoje cultivado, desde sua introdução no país. Como já salientado por Basha e Sujatha (2007) e Rao et al. (2008), reduzidos níveis de diversidade também foram observados no germoplasma indiano de *J. curcas*. Do mesmo modo, Sun et al. (2008) e Zhang et al. (2011) demonstraram, por meio de marcadores AFLP, uma limitada variabilidade genética para os genótipos chineses. Em contraste a estes resultados, Santos et al. (2010a) alegam ter observado uma alta variabilidade genética, via marcadores AFLP, em 50 acessos de *J. curcas* da Coleção de germoplasma da EMBRAPA Semiárido, com um agrupamento organizado em função de suas origens geográficas e uma variação da ordem de 0,27 entre as diferentes populações avaliadas. Entretanto, vale dizer que, talvez por terem atingido um coeficiente cofenético de 91%, estes autores não adotaram qualquer metodologia de corte para a discriminação dos agrupamentos formados.

Apesar da variabilidade aqui encontrada, de modo geral, os estudos acerca do germoplasma da espécie *Jatropha curcas* parecem convergir para a discriminação de uma modesta diversidade genética, sendo esta provocada por: um ancestral comum, referente às diversas localidades nas quais a espécie vêm sendo explorada; e, principalmente, por sua dispersão via propagação vegetativa (ROSADO et al. 2010); pela ocorrência de apomixia (BHATTACHARYA et al. 2005); pelo pequeno número de plantas introduzidas nas regiões amostradas; pelo emprego de poucos marcadores (NOOR CAMELLIA et al. 2012); ou até mesmo por ocorrer, majoritariamente, de forma isolada, fazendo com que a autofecundação prevaleça sobre sua inclinação à reprodução cruzada e, por conseguinte, aumente o número de locos em homozigose, como explanado por Santos et al. (2010a). Estes autores descreveram um baixo fluxo gênico ($N_m = 0,66$) entre as diferentes populações de *J. curcas* por eles avaliadas, atribuindo tal feito a um resultado direto do isolamento espacial entre estas. De modo semelhante ao germoplasma chinês, genótipos brasileiros de *J. curcas* podem estar se reproduzindo preferencialmente por autogamia, dado o alto número de locos em homozigose detectado ao longo destes (ROSADO et al. 2010). Contudo, há de se concordar com as

suposições de Rao et al. (2008) e Rosado et al. (2010), onde declaram que, por ser uma espécie de grande adaptabilidade edafoclimática, e também ampla distribuição geográfica, acredita-se, ainda, que haja uma considerável variabilidade genética a ser explorada em seu germoplasma.

4.2.4. Avaliação do potencial produtivo

Tendo estimado a diversidade genética, bem como os efeitos diretos e indiretos das correlações entre os descritores propostos, partiu-se, por fim, para uma investigação mais específica sobre o comportamento produtivo dos 60 acessos em questão, da Coleção de germoplasma de *Jatropha curcas* da UFRRJ.

Na presente avaliação, a Tabela 16 demonstra, em ordem decrescente, os valores da produção de grãos alcançada entre os meses de dezembro de 2011 e maio de 2012, relativo à frutificação das plantas de pinhão-mansão nas condições climáticas de Seropédica, RJ, aos três anos de idade.

Tabela 16. Ordem decrescente de produção e respectivas procedências dos 60 genótipos de *Jatropha curcas* da Coleção de germoplasma da UFRRJ definidos para avaliação, entre dezembro de 2011 e maio de 2012.

	Acesso	Procedência	Produção*	Acesso	Procedência	Produção*	
1	860	Janaúba - MG	1402,04	31	332	Lavras - lote 000 - MG	459,67
2	345	Lavras - lote 210 - MG	1391,20	32	857	Janaúba - MG	455,38
3	346	Janaúba - MG	1121,36	33	062	Jaíba - MG	446,68
4	858	Janaúba - MG	903,31	34	058	CNPS - MA	424,99
5	878	Janaúba - MG	884,83	35	003	Barra dos Bugres - MT	424,98
6	002	Petrolina - PE	824,26	36	955	Dourados - MS	417,00
7	383	Rio Verde - GO	775,18	37	872	Janaúba - MG	415,31
8	355	Bom Jardim - RJ	768,28	38	385	Janaúba - MG	411,95
9	828	Petrolina - PE	768,27	39	911	Dourados - MS	381,09
10	876	Janaúba - MG	736,66	40	059	Janaúba - MG	363,45
11	932	Dourados - MS	717,41	41	005	Lavras - lote 000 - MG	357,48
12	357	Lavras - lote 210 - MG	697,07	42	875	Janaúba - MG	348,25
13	869	Janaúba - MG	653,93	43	393	Dourados - MS	341,05
14	825	Dourados - MS	642,81	44	849	Petrolina - PE	340,84
15	376	Petrolina - PE	641,03	45	344	Barra dos Bugres - MT	307,74
16	356	CNPS - MA	639,71	46	864	Janaúba - MG	283,94
17	321	Rio das Flores - RJ	606,73	47	861	Janaúba - MG	272,13
18	931	Petrolina - PE	585,66	48	341	Guapimirim - RJ	264,65
19	831	Petrolina - PE	559,43	49	900	Dourados - MS	253,37
20	053	Rio Verde - GO	558,29	50	343	Rio das Flores - RJ	210,49
21	328	Lavras - lote 210 - MG	549,45	51	856	Janaúba - MG	182,51
22	908	Dourados - MS	548,95	52	330	Lavras - lote 000 - MG	164,64
23	912	Petrolina - PE	527,54	53	820	Dourados - MS	162,05
24	807	Dourados - MS	525,90	54	331	CNPS - MA	149,11
25	874	Janaúba - MG	500,28	55	386	Bom Jardim - RJ	142,02
26	371	CNPS - MA	484,27	56	337	Dourados - MS	120,45
27	399	Lavras - lote 000 - MG	482,48	57	905	Dourados - MS	88,62
28	863	Janaúba - MG	466,33	58	991	Dourados - MS	66,73
29	024	Petrolina - PE	462,87	59	890	Janaúba - MG	55,47
30	824	Dourados - MS	459,72	60	910	Dourados - MS	23,85

* Produção de grãos em g.planta⁻¹.

Assim como os 16 descritores já avaliados por meio das análises de agrupamento, a produção, quando examinada de maneira individual, também não sofreu influência das diferentes localidades dos genótipos estudados. Tal suposição pode ser conferida através dos genótipos de procedências distintas, presentes tanto entre as 10 plantas mais produtivas quanto nas 10 de menor produção. Ainda assim, nota-se que dos 10 acessos mais produtivos,

seis são provenientes do Estado de Minas Gerais, mas de localidades diferentes, sendo os acessos 860 (1402,04 g.planta⁻¹), 346 (1121,36 g.planta⁻¹), 858 (903,31 g.planta⁻¹), 878 (884,83 g.planta⁻¹) e 876 (736,66 g.planta⁻¹) oriundos de Janaúba e o acesso 345 (1391,20 g.planta⁻¹) de Lavras. A mesma relação ocorreu entre os 10 genótipos menos produtivos, onde a metade destes (acessos 820, 337, 905, 991 e 910) é procedente de Dourados, do Estado de Mato Grosso do Sul (Tabela 16).

Sobre o uso das origens geográficas como indicador de diversidade genética, Cruz e Carneiro (2003) alertam que as intensas trocas de sementes e estruturas propagativas entre instituições de diferentes regiões acabam por causar a perda de individualidade e ocorrência de tipos particulares. Nesse sentido, é possível observar que, apesar de Janaúba (MG) ser a procedência de grande parte dos 10 acessos mais produtivos, os acessos 890, 856, 861, 864, 875, 059, 385, 872 e 857, também de Janaúba, figuraram entre os 30 genótipos menos produtivos da coleção, assim como os acessos 932, 825, 908, 807 e 824, de Dourados (MS), localidade de 5 entre os 10 de menor produção, também se encontraram entre as 30 plantas mais produtivas da referida coleção, indicando, desse modo, a grande similaridade existente em meio a este material genético quanto à produção (Tabela 16). Outro exemplo deste desempenho, como visto na Tabela 16, foi dado pelos acessos 355 e 386, ambos de Bom Jardim (RJ), onde o primeiro, produzindo 768,28 g.planta⁻¹, foi o 8º genótipo de maior produção, ao passo que o acesso 386, com 142,02 g.planta⁻¹, listou a 55ª produção.

A despeito da recente intensificação sobre a exploração comercial desta espécie, os dados literários acerca da produtividade do pinhão-manso ainda seguem contraditórios. Sem especificar a idade das plantas estudadas, Martin e Mayeux (1984), citados por Heller (1996), contabilizaram uma produção variando entre 3000 e 5000 g.planta⁻¹ de grãos em genótipos de *J. curcas* de Madagascar. Sukarin et al. (1987), também citados por Heller (1996), registraram genótipos tailandeses atingindo 63,8 g.planta⁻¹ de sementes no primeiro ano de produção. Shabanimofrad et al. (2013), estudando 48 genótipos de *J. curcas* da Malásia, conferiram uma variação entre 107,25 e 745,72 g.planta⁻¹ de grãos. No entanto, como exposto na Tabela 16, tal observação não foi corroborada, visto que, no presente estudo, a produção se mostrou bem superior, variando entre 23,85 e 1402,04 g.planta⁻¹, onde os acessos mais produtivos, 860, 345 e 346, ultrapassaram a marca dos 1000 g.planta⁻¹, apresentando uma grande variabilidade genética para produção.

Conhecida a produção, a próxima etapa focou-se em entender, via correlação fenotípica, como as características inerentes aos frutos e sementes - peso do fruto (PFR), comprimento do fruto (CFR), largura do fruto (LFR), tempo de formação do fruto (TFF), peso médio da semente seca (PMSs), comprimento da semente (CMS) e largura da semente (LSE) - poderiam influenciar no número de grãos gerados em *Jatropha curcas*. Sendo assim, da mesma forma que os 16 descritores anteriormente empregados, para a estimação da diversidade genética, foram submetidos ao teste de multicolinearidade (item 4.2.1. deste trabalho), os descritores PFR, CFR, LFR, TFF, PMSs, CMS, LSE, e a própria produção de grãos por planta (PGP), agora analisados à parte, também foram diagnosticados pelo Teste de Montgomery e Peck (1981). Através deste, o número de condição, de 16,52, indicou um fraco grau de severidade ($NC < 100$), retratando um baixo grau de colinearidade e, portanto, um bom acondicionamento da matriz de correlação, não apresentando maiores problemas sobre a interpretação desta.

Através da Tabela 17, nota-se, como já previsto, que apenas as características peso médio da semente seca e tempo de formação do fruto obtiveram moderadas correlações significativas sobre a produção de grãos por planta. O descritor PMSs obteve um positivo valor de 0,30, uma vez que o descritor PGP é inerente ao produto entre o número e o peso médio da semente seca por planta. Corroborando à correlação de -0,51, sobre o número de sementes por planta (NSE), demonstrado no item 4.2.2. do presente trabalho, o descritor TFF

também apresentou um negativo valor sobre PGP, de -0,52. Igualmente, as significativas correlações positivas entre CFR junto a LFR, TFF e CMS, de PFR com LSE, e de PMSs com CMS, bem como a negativa correlação entre CFR e LSE, já eram esperadas.

Tabela 17. Coeficientes de correlação fenotípica entre oito descritores relativos a frutos e sementes, avaliados de dezembro de 2011 a maio de 2012, em 60 genótipos de *Jatropha curcas* da Coleção de germoplasma da UFRRJ.

Descritores	PFR	CFR	LFR	TFF	PMSs	CMS	LSE	PGP
PFR	1,00	-0,20	0,13	-0,16	0,18	0,09	0,34**	-0,21
CFR		1,00	0,54**	0,42**	0,20	0,48**	-0,37**	-0,13
LFR			1,00	-0,04	0,16	0,13	0,09	0,01
TFF				1,00	-0,17	0,16	-0,35**	-0,52**
PMSs					1,00	0,42**	0,20	0,30*
CMS						1,00	0,13	-0,14
LSE							1,00	0,10
PGP								1,00

Legenda: CFR - Comprimento do fruto; CMS - Comprimento da semente; LFR - Largura do fruto; LSE - Largura da semente; PFR - Peso do fruto; PGP - Produção de grãos por planta; PMSs - Peso médio da semente seca; e TFF - Tempo de formação do fruto. *, ** Significativo a 5% e 1% de probabilidade pelo teste *t*, respectivamente.

Com base nestas correlações, é importante ressaltar que, embora não significativas, as estimativas negativas entre a produção de grãos e os descritores PFR, CFR e CMS podem vir a comprometer a produção (Tabela 17). Tal comportamento pode ser explicado pelo fato de que a produção de frutos do pinhão-manso tende a aumentar de acordo com o seu número de ramos, visto que suas inflorescências são geradas nas gemas terminais destes (SATURNINO et al., 2005). Isto é, plantas de maior produção de grãos possuem também um maior número de frutos. Dessa maneira, genótipos de *Jatropha curcas* de alta produção optam por investir em um maior número de frutos de menor peso e tamanho, em detrimento a alocar seus nutrientes na promoção de um maior enchimento destes.

A partir dos dados referentes ao número de sementes (NSE) e peso médio da semente seca por planta (PMSs), coletados semanalmente entre os meses de dezembro de 2011 e maio de 2012, foi possível traçar um comportamento relativo à produção (PGP = NSE x PMSs) dos 60 genótipos ao longo do período de colheita, identificando, assim, as datas de maior e menor produtividade de grãos, como demonstrado na Figura 2.

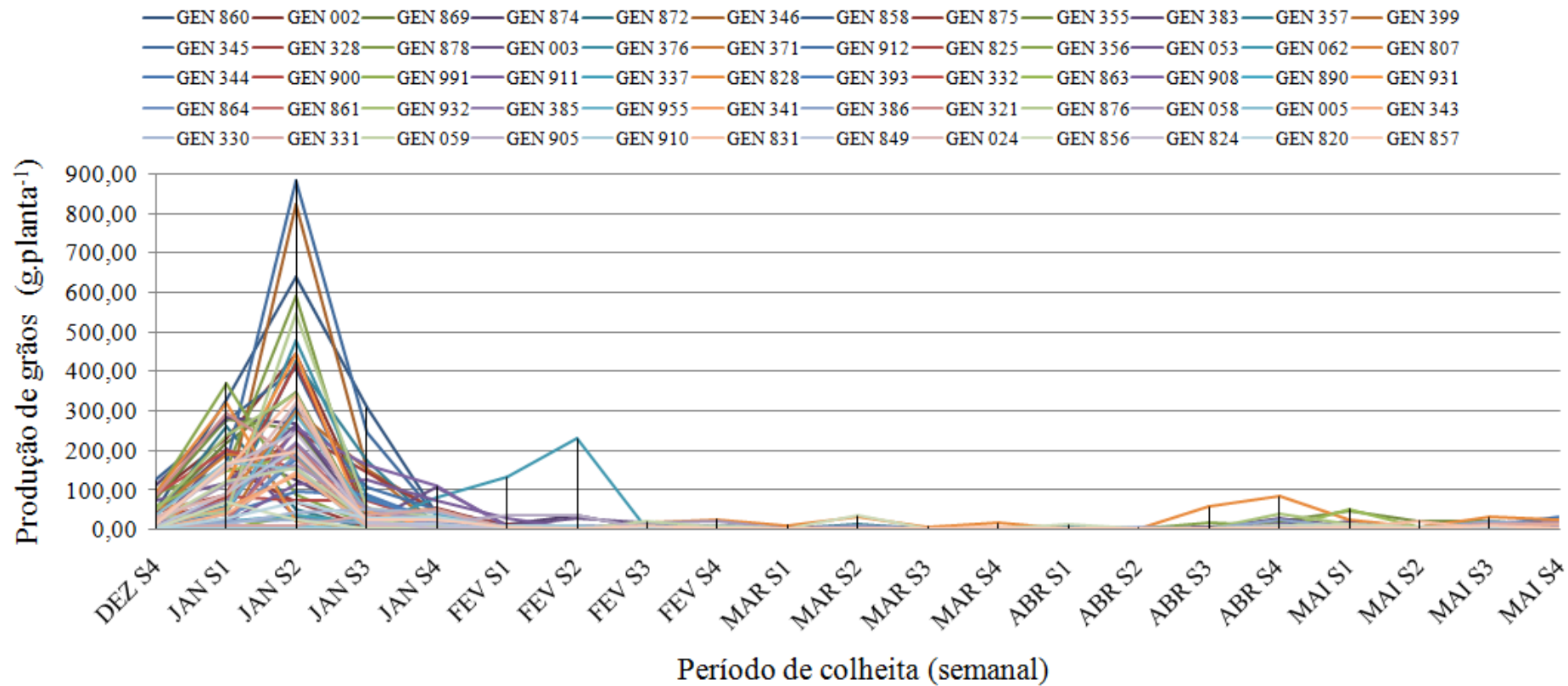


Figura 2. Produção semanal de 60 genótipos da Coleção de germoplasma de *Jatropha curcas* da UFRRJ entre os meses de dezembro de 2011 e maio de 2012.

Por meio da Figura 2, observa-se, com exceção do acesso 062, que demonstrou pico de produção em fevereiro de 2012 (229,50 g.planta⁻¹), que a produção de grãos dos genótipos aqui caracterizados se concentrou ao longo da segunda semana de janeiro de 2012. O acesso 860, o primeiro da lista dos mais produtivos, alcançou uma produção de 637,56 g.planta⁻¹ nesta mesma semana. Nesse contexto, é importante destacar que, exceto o acesso 828, todos os outros genótipos apresentaram apenas um pico de produção. Este genótipo foi o único a produzir grãos ao longo de todos os 6 meses de avaliação, obtendo dois picos de produção, um na primeira semana de janeiro, com 321,37 g.planta⁻¹, e o segundo na quarta semana de abril de 2012, com 82,81 g.planta⁻¹ (Figura 2). Tal comportamento sinaliza para a possibilidade de, futuramente, se selecionar genótipos para duas diferentes épocas de colheita, sendo a primeira em janeiro (fevereiro para o acesso 062) e a segunda de forma mais tardia, em abril, nas condições edafoclimáticas do município de Seropédica, RJ. Nesta região, localizada na baixada fluminense, o período expresso entre os meses de outubro e março (primavera e verão) são caracterizados pela estação chuvosa, enquanto a temporada compreendida entre abril e setembro (outono e inverno) é marcada pela seca (Figura 3). Dessa maneira, entende-se que a coleção aqui estudada apresentou maior produção durante a época mais chuvosa, mais especificamente ao longo do mês de janeiro de 2012. Todavia, baseado no segundo pico de produção do acesso 828, pode-se inferir que este genótipo também deve se mostrar mais responsivo sob manejo irrigado, visto que, além de ter produzido grãos durante os 6 meses de avaliação, este também apresentou uma sensível resposta positiva à produção em abril de 2012, período caracterizado pela volta da chuva, em contraste às demais plantas estudadas (Figuras 2 e 3). Sob este aspecto, é importante ressaltar que, apesar de distintos, estes dois picos de produção tiveram emissão floral em períodos anteriores e com regimes de precipitação em comum, sendo outubro de 2011 para o primeiro pico (janeiro de 2012), e em janeiro de 2012 para o segundo (abril de 2012), caracterizados por épocas chuvosas.

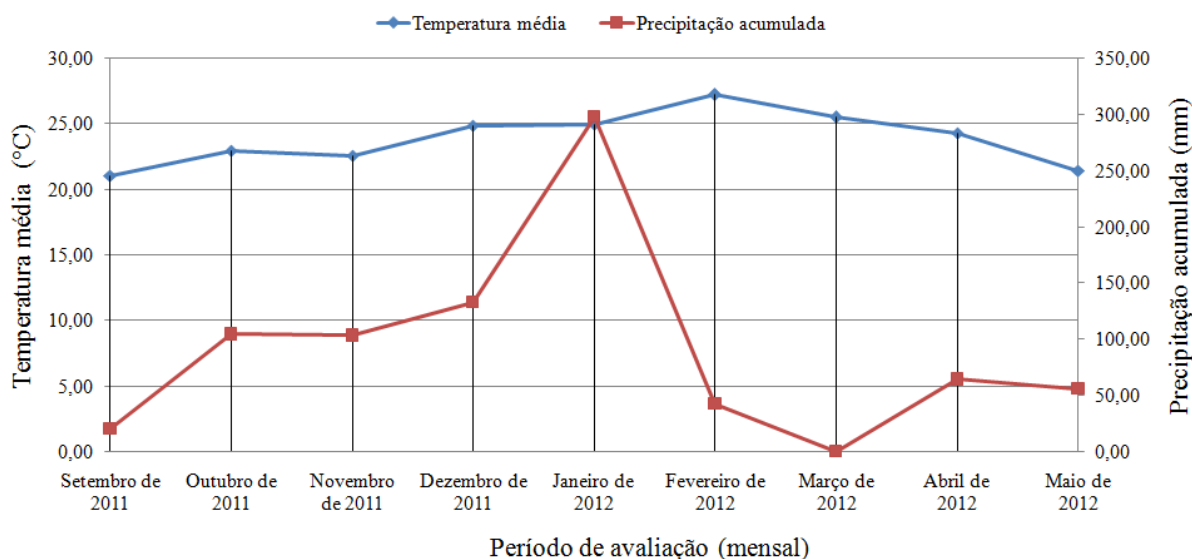


Figura 3. Temperatura média e precipitação acumulada na região de Seropédica, RJ, onde está localizada a Coleção de germoplasma de *Jatropha curcas* avaliada, entre os meses de setembro de 2011 e maio de 2012. Fonte: INMET/SIPA/UFRRJ.

Em consonância ao padrão obtido no presente trabalho, Santos et al. (2010b), estudando a fenologia e o crescimento de genótipos de pinhão-manso cultivados no Estado de Alagoas, verificaram dois períodos distintos de frutificação, sendo o primeiro de menor intensidade (entre os meses de dezembro de 2006 e fevereiro de 2007) e o segundo de maior

magnitude (entre abril e agosto de 2007), respectivamente nas estações de seca e chuva daquela região. Dado o caráter de agricultura familiar atribuído ao pinhão-manso, a possibilidade de se obter dois períodos de colheita, como aqui descrito, também se faz de grande relevância para o pequeno agricultor, possibilitando-o diversificar sua produção, implantando a prática do consórcio com culturas de ciclo curto, por exemplo, e planejar, de acordo com estas épocas, a melhor maneira de adequar sua colheita.

Dentre os genótipos mais produtivos, ressalta-se que os acessos 860, 345 e 346 (o primeiro, segundo e terceiro de maior produção) atingiram, respectivamente, 45,47%, 63,43% e 76,36% de sua produção total ao longo da segunda semana de janeiro de 2012. Ainda sob esta perspectiva, deve-se enfatizar que o acesso 345, de produção total de 1391,20 g.planta⁻¹, cessou sua produção já no mês de fevereiro de 2012. Já os acessos 890 (Janaúba - MG), 905 e 910 (ambos de Dourados - MS), três dos menos produtivos (55,47; 88,62; e 23,85 g.planta⁻¹, respectivamente), produziram somente entre janeiro e a quarta semana de fevereiro de 2012, também não voltando mais a produzir (Figura 2). Perante a crescente demanda por grãos em *Jatropha curcas*, os programas de melhoramento genético desta devem adotar estratégias emergenciais e paliativas, que busquem maximizar sua produção já em curto prazo. Dessa forma, uma abordagem rápida, capaz de atender, provisoriamente, tal demanda, seria a utilização de genótipos conhecidamente mais produtivos, como os acessos 860, 345 e 346, aqui discutidos, em técnicas de propagação vegetativa. Nesse contexto, tanto a propagação vegetativa convencional quanto o cultivo *in vitro* possibilitariam obter resultados de forma mais rápida. A partir da clonagem dos melhores indivíduos, dois dos grandes “gargalos” da cadeia produtiva e do melhoramento genético do pinhão-manso já poderiam ser amenizados: a redução do tempo para a primeira produção comercial (precocidade de produção) e a questão da uniformidade dos plantios.

Quando comparada as médias da produção semanal pelos 60 genótipos, entre os 6 meses avaliados, observou-se, na Tabela 18, que o mês de janeiro de 2012 foi o único a diferir estatisticamente da produção obtida por todos os demais meses, a 5% de significância pelo teste *t*. Apesar da ausência de chuva em março de 2012, verificou-se que, estatisticamente, este período não diferiu de abril do mesmo ano e dezembro de 2011. No entanto, cabe dizer que esta resposta está diretamente ligada ao período de emissão floral das plantas, que, em consequência, irá refletir no tempo de formação dos frutos e na época em que estes serão gerados. Nesse caso, a produção de março e abril de 2012 foi correspondente à floração de dezembro de 2011 e janeiro de 2012, respectivamente, marcada por períodos chuvosos. Por sua vez, os meses de dezembro de 2011 e abril de 2012 só diferiram de forma significativa do mês de janeiro. Já a produção alcançada em fevereiro de 2012 diferiu significativamente dos meses de janeiro e março, não sendo considerada distinta de abril e maio de 2012 e dezembro de 2011 (Tabela 18). Silva et al. (2012), avaliando a concentração do ciclo de produção no primeiro ano de cultivo, via podas de formação e produção, observaram que o período de colheita em plantas de pinhão-manso não podadas perdurou por 120 dias, enquanto que os tratamentos referentes aos genótipos podados não ultrapassaram os 48 dias de duração, para a poda de formação, e 90 dias, para a poda de produção, na região do semiárido de Minas Gerais.

Tabela 18. Comparação entre os valores médios da produção de grãos mensal no período de dezembro de 2011 a maio de 2012, em 60 genótipos da Coleção de germoplasma de *Jatropha curcas* da UFRRJ.

Meses	Produção - Média original	Produção - Média transformada*
Dezembro de 2011	432,26 bc	11,40 bc
Janeiro de 2012	6194,61 a	73,28 a
Fevereiro de 2012	253,23 b	16,57 b
Março de 2012	43,93 c	7,17 c
Abril de 2012	110,49 bc	10,11 bc
Maio de 2012	270,37 b	17,28 b
CV (%)	251,11	123,56

Legenda: CV - Coeficiente de variação experimental. Médias seguidas pela mesma letra não diferem estatisticamente entre-si pelo teste *t* a 5% de probabilidade. * Dados transformados pela expressão $\sqrt{(x + 1)}$.

Ao avaliar o padrão de florescimento em regiões equatoriais constantemente úmidas, Heller (1996) verificou que o pinhão-manso pode emitir flores durante o ano todo. Já em condições de déficit hídrico, em regiões de clima quente da China, Meng et al. (2009) descreveram genótipos de *J. curcas* que florescem apenas uma vez ao ano. Para Santos et al. (2010b), as respostas anteriormente descritas reiteram a grande influência do clima sobre a fenologia reprodutiva da cultura. Segundo Juhász et al. (2009), no Brasil, o florescimento do pinhão-manso se dá após um período de temperatura e precipitação reduzidas. Esta afirmação de Juhász et al. (2009) se mostra em consonância ao observado no presente trabalho, onde, nas condições climáticas de Seropédica, RJ, a floração se iniciou na primavera, junto à volta do período chuvoso e logo após a seca do inverno (Figura 3).

De acordo com Laviola et al. (2011), a emissão de inflorescências e a maturação de frutos de *J. curcas* ainda seguem muito desuniformes, fazendo com que os produtores realizem diversas colheitas durante a etapa produtiva, onerando o custo sobre a mão de obra e tornando seu cultivo economicamente inviável. No entanto, a partir dos resultados conferidos no presente estudo, inferi-se que, a despeito da marcante desuniformidade de floração/maturação da espécie, a Coleção de germoplasma de *Jatropha curcas* da UFRRJ já acena para a possibilidade de seleção de genótipos com a capacidade de se concentrar a produção no mês de janeiro, sem fazer uso de manejos alternativos, como a poda (Tabela 18 e Figuras 2 e 3). Como o extenso período de colheita e a heterogeneidade do florescimento ainda são dois dos principais alvos a serem trabalhados em *Jatropha curcas*, este resultado pode ser julgado de grande importância para o programa de melhoramento genético da espécie. De forma semelhante aos dois períodos de colheita já citados, tal fato também pode ser considerado essencial para o pequeno agricultor familiar, uma vez que concentrando a produção, um menor plantel de mão de obra será necessário para efetuar a colheita, reduzindo, assim, o custo de produção da cultura.

5. CONCLUSÕES

A partir dos resultados aqui expressos, pode-se concluir que:

❖ O desenvolvimento e a aplicação da lista preliminar de descritores colaborou tanto para o estabelecimento de um novo conjunto de dados, a serem utilizados no Programa de melhoramento de *Jatropha curcas*, quanto na identificação de potenciais marcadores morfoagronômicos para a discriminação da diversidade genética da coleção da UFRRJ, promovendo um melhor aproveitamento de seus recursos genéticos e a otimização da fenotipagem.

❖ A técnica de componentes principais se mostrou eficiente no reconhecimento dos descritores de maior e menor importância para a diversidade genética encontrada, descartando 20 variáveis quantitativas de baixa contribuição e apontando, dentre as restantes, os 10 descritores que mais colaboraram para a variabilidade genética da coleção.

❖ O estudo das correlações e análise de trilha descreveram importantes associações entre os descritores NSF, TFF, PCOo e CRI junto ao descritor NSE. Tais resultados corroboraram a possibilidade de, na Coleção de germoplasma da UFRRJ, estes descritores serem utilizados como base para a seleção indireta, visando à obtenção de variedades de menor tempo de colheita e, simultaneamente, maior produção de grãos.

❖ A distância de Chebyshev e o algoritmo de Gower, aliados ao método de agrupamento UPGMA, apresentaram, em detrimento às demais combinações testadas, maior robustez para a discriminação da diversidade genética da coleção estudada, confirmando a importância de se testar as diferentes combinações de distâncias e técnicas de agrupamento disponíveis, bem como as distintas metodologias de verificar sua consistência. De modo geral, a Coleção de germoplasma de *Jatropha curcas* da UFRRJ demonstrou uma considerável diversidade genética para os descritores aqui aplicados, na qual os acessos 860, 345, 346 e 062 se apresentaram como os mais promissores a serem envolvidos em futuros cruzamentos visando à obtenção de cultivares comerciais produtivas e não-tóxicas.

❖ As correlações estimadas somente entre os descritores referentes a frutos e sementes demonstraram que quanto maior o comprimento destes, e mais pesado o fruto, menor é a produção de grãos por planta. A avaliação do potencial produtivo da coleção em questão revelou que sua produção se concentrou ao longo da segunda semana de janeiro de 2012. Os acessos UFRRJ 860, 345 e 346 foram, seguidamente, os três genótipos de maior produção, enquanto os acessos 905, 991, 890 e 910 obtiveram o menor número de grãos apurados. Tal avaliação também demonstrou que as maiores épocas de floração e colheita, nas condições edafoclimáticas de Seropédica, RJ, se deram, respectivamente, durante a primavera (outubro e novembro de 2011) e verão (janeiro de 2012), na estação chuvosa da região.

❖ Tanto a concentração da produção em um único mês (janeiro) quanto a futura possibilidade de se obter dois diferentes períodos de colheita (janeiro e abril) fazem dos resultados aqui apresentados importantes atrativos à adoção do pinhão-manso pelo pequeno agricultor familiar, permitindo a diversificação e a redução dos custos de sua produção. O Programa de melhoramento de *Jatropha curcas* da UFRRJ poderia proporcionar, a curto prazo, a maximização da uniformidade e a precocidade de produção, através da clonagem dos acessos 860, 345 e 346, amenizando, na região metropolitana do Estado do Rio de Janeiro, os principais obstáculos da cadeia produtiva desta cultura.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- ABREU, F. B.; RESENDE, M. D. V. de; ANSELMO, J. L.; SATURNINO, H. M.; BRENHA, J. A. M.; FREITAS, F. B. de. Variabilidade genética entre acessos de pinhão-manso na fase juvenil. **Magistra**, Cruz das Almas-BA, v. 21, n. 1, p. 036-040, 2009. ISSN 0102-5333.
- ADAMS, R. P. Conservation of DNA: DNA banking. *In*: CALLOW, J. A.; FORD-LLOYD, B. V. and NEWBURY, H. J. (eds.). **Biotechnology and plant genetic resources conservation and use**. CAB International Publishing, Wallingford, UK, 1997. p. 163-174.
- ADAMS, R. P.; DO, N.; GE-LIN, C. Preservation of DNA in plant specimens from tropical species by desiccation. *In*: ADAMS, R. P. and ADAMS, J. E. (eds.). **Conservation of plant genes, DNA banking and in vitro biotechnology**. Academic Press Inc., San Diego, USA, 1992. p. 153-181.
- AKBAR, E.; YAAKOB, Z.; KAMARUDIN, SK.; ISMAIL, M.; SALIMON, J. Characteristic and composition of *Jatropha Curcas* oil seed from Malaysia and its potential as biodiesel feedstock. **European Journal of Scientific Research**, v. 29, n. 3, p. 396-403, 2009.
- ALBUQUERQUE, F. A. de; OLIVEIRA, M. I. P. de; LUCENA, A. M. A. de; BARTOLOMEU, C. R. C.; BELTRÃO, N. E. de M. Crescimento e desenvolvimento do pinhão-manso: 1º ano agrícola. **Documentos**, 197. Campina Grande, PB: Embrapa Algodão, 2008. 21 p. ISSN 0103-0205.
- ALEXANDER, M. P.; GANESHAN, S. Pollen storage. *In*: CHADHA, K. L. and PAREEK, O. P. (eds.). **Advances in horticulture**, volume I - fruit crops, part I. Malhotra Publishing House, New Delhi, India, 1993. p. 481-496.
- ALFENAS, A. C. **Eletroforese de isoenzimas e proteínas afins: fundamentos e aplicações em plantas e microrganismos**. Viçosa: Editora UFV, 1998. 574 p.
- ALLARD, R. W. **Principles of plant breeding**. London: J. Wiley, 1960. 485 p.
- ALVES, J. M. A.; SOUSA, A. de A.; SILVA, S. R. G. da; LOPES, G. N.; SMIDERLE, O. J.; UCHÔA, S. C. P. Pinhão-manso: uma alternativa para produção de biodiesel na agricultura familiar da Amazônia brasileira. **Agro@mbiente**, Boa Vista, v. 2, n. 1, p. 57-68, 2008. ISSN 1982-8470.
- AMARAL JÚNIOR, A. T. do. Divergência genética entre acessos de moranga do banco de germoplasma de hortaliças da Universidade Federal de Viçosa. Palestra. Suplemento. **Horticultura Brasileira**, Brasília, v. 17, p. 03-06, 1999. ISSN 0102-0536.
- ANDRÉO-SOUZA, Y.; PEREIRA, A. L.; SILVA, F. F. S. da; RIEBEIRO REIS, R. C.; EVANGELISTA, M. R. V.; CASTRO, R. D. de; DANTAS, B. F. Efeito da salinidade na germinação de sementes e no crescimento inicial de mudas de pinhão-manso. **Revista Brasileira de Sementes**, Londrina, v. 32, n. 2, p. 083-092, 2010. ISSN 0101-3122.
- ARRUDA, F. P.; BELTRÃO, N. E. M.; ANDRADE, A. P.; PEREIRA, W. E.; SEVERINO, L. S. Cultivo de pinhão-manso (*Jatropha curcas*) como alternativa para o semi-árido

nordestino. **Revista Brasileira de Oleaginosas e Fibrosas**, Campina Grande, v. 8, n. 1, p. 789-799, 2004. ISSN 1980-4830.

AZAMBUJA, S. de. **Estudo e implementação da análise de agrupamento em ambientes virtuais de aprendizagem**. 2005. 212 p. Dissertação (Mestrado em Informática) - Universidade Federal do Rio de Janeiro, Instituto de Matemática / Núcleo de Computação Eletrônica, 2005.

BÁRTOLI, J. A. A. **Physic-nut (*Jatropha curcas*) cultivation, Honduras handbook**. Agricultural Communication Center of the Honduran Foundation for Agricultural Research (FHIA). La Lima, Cortés, Honduras, May 2008. 40 p.

BASHA, S. D.; SUJATHA, M. Inter and intra-population variability of *Jatropha curcas* (L.) characterized by RAPD and ISSR markers and development of population-specific SCAR markers, **Euphytica**, Wageningen, v. 156, p. 375-386, 2007. ISSN 0014-2336.

BERTINI, C. A. de M.; PINHEIRO, E. A. R.; NÓBREGA, G. N.; DUARTE, J. M. de L. Desempenho agrônomo e divergência genética de genótipos de coentro. **Revista Ciência Agrônômica**, Ceará, v. 41, n. 3, p. 409-416, 2010. ISSN 1806-6690.

BHATTACHARYA, A.; DATTA, K.; DATTA, S. K. Floral biology, floral resource constraints and pollination limitation in *Jatropha curcas* L. **Pakistan Journal of Biological Sciences**, v. 8, p. 456-460, 2005. ISSN 1812-5735.

BIOVERSITY AND THE CHRISTENSEN FUND. **Descriptors for farmers' knowledge of plants**. Bioversity International, Rome, Italy and The Christensen Fund, Palo Alto, California, USA, 2009. 13 p. ISBN 978-92-9043-801-4.

BIOVERSITY INTERNATIONAL. Guidelines for the development of crop descriptor lists. **Bioversity Technical Bulletin Series**. Bioversity International, Rome, Italy, 2007. xii+72 p. ISBN 978-92-9043-792-1.

BITOCCHI, E.; NANNI, L.; ROSSI, M.; RAU, D.; BELLUCCI, E.; GIARDINI, A.; BUONAMICI, A.; VENDRAMIN, G. G.; PAPA, R. Introgression from modern hybrid varieties into landrace populations of maize (*Zea mays* ssp. *mays* L.) in central Italy. **Molecular Ecology**, v. 18, n. 4, p. 603-621, 2009. ISBN 0962-1083.

BUSSAB, W. O.; MIAZAKI, E. S.; ANDRADE, D. F. Introdução à análise de agrupamentos. In: SIMPÓSIO NACIONAL DE PROBABILIDADE E ESTATÍSTICA, 9, 1990. São Paulo. **Anais...** Associação Brasileira de Estatística, ABE, 1990. 87 p.

CARGNELUTTI FILHO, A.; GUADAGNIN, J. P. Consistência do padrão de agrupamento de cultivares de milho. **Ciência Rural**, Santa Maria, v. 41, n. 9, p. 1503-1508, 2011. ISSN 0103-8478.

CARGNELUTTI FILHO, A.; RIBEIRO, N. D.; BURIN, C. Consistência do padrão de agrupamento de cultivares de feijão conforme medidas de dissimilaridade e métodos de agrupamento. **Pesquisa Agropecuária Brasileira**, v. 45, n. 3, p. 236-243, 2010. ISSN 1678-3921.

CARVALHO, D. S. **Comportamento genético de progênies de meio-irmãos de pinhão manso no Recôncavo Baiano**. 2010. 51 p. Dissertação (Mestrado em Ciências Agrárias) - Universidade Federal do Recôncavo da Bahia, Centro de Ciências Agrárias, Ambientais e Biológicas, 2010.

CARVALHO, C. G. P.; OLIVEIRA, V. R.; CRUZ, C. D.; CASALI, V. W. D. Análise de trilha sob multicolinearidade em pimentão. **Pesquisa Agropecuária Brasileira**, Brasília, v. 34, n. 4, p. 603-613, abr.1999. ISSN 1678-3921.

CARVALHO, D. F.; SILVA, L. D. B.; FOLEGATTI, M. V.; COSTA, J. R.; CRUZ, F. A. Avaliação da evapotranspiração de referência na região de Seropédica-RJ, utilizando lisímetro de pesagem. **Revista Brasileira de Agrometeorologia**, v. 14, n. 2, p. 108-116, 2006. ISSN 2175-7666.

CARVALHO, F. I. F. de; SILVA, S. A.; KUREK, A. J.; MARCHIORI, V. S. **Estimativas e implicações da herdabilidade como estratégia de seleção**. Pelotas: Editora da UFPel, 2001. 99 p. ISBN 85-7192-162-8.

CHANG, T. T. Saving crop germplasm. **Span** (Feb. issue), England, v. 30, n. 2, p. 62-63, 1987.

CHEN, K.; REN, P.; YING, C.; JIANG, Q.; JIA, X. Genetic relationships among *Jatropha curcas* L. clones from Panzhihua, China as revealed by RAPD and ISSR. **African Journal of Agricultural Research**, v. 6, n. 11, p. 2582-2585, 2011. ISSN 1991-637X.

CHIORATO, A. F. **Divergência genética em acessos de feijoeiro (*Phaseolus vulgaris* L.) do Banco de Germoplasma do Instituto Agrônomo-IAC**. 2004. 85 p. Dissertação (Mestrado em Agricultura Tropical e Subtropical) - Instituto Agrônomo de Campinas, 2004.

COELHO, C. M. M.; COIMBRA, J. L. M.; SOUZA, C. A. de; BOGO, A.; GUIDOLIN, A. F. Diversidade genética em acessos de feijão (*Phaseolus vulgaris* L.). **Ciência Rural**, Santa Maria, v. 37, n. 5, p. 1241-1247, 2007. ISSN 0103-8478.

COIMBRA, J. L. M.; BENIN, G.; VIEIRA, E. A.; OLIVEIRA, C. de; CARVALHO, F. I. F.; GUIDOLIN, A. F.; SOARES, A. P. Consequências da multicolinearidade sobre análise de trilha em canola. **Ciência Rural**, Santa Maria, v. 35, n. 2, p. 347-352, mar-abr. 2005. ISSN 0103-8478.

CORDEIRO, C. M. T.; ABADIE, T. Coleções nucleares. *In*: NASS, L. L. (Editor técnico). **Recursos genéticos vegetais**. Brasília, DF: Embrapa Recursos Genéticos e Biotecnologia, 2007. p. 575-602. ISBN 978-85-87697-34-9.

CORTESÃO, M. **Culturas tropicais: plantas oleaginosas, coqueiro, rícino, purgueira e aleurites**. Lisboa: Livraria Clássica, 1956. 231 p.

CRUZ, C. D. **Programa Genes - Diversidade Genética**. 1. ed. Viçosa: Editora UFV, 2008. 278 p.

CRUZ, C. D.; CARNEIRO, P. C. S. **Modelos biométricos aplicados ao melhoramento genético**. Volume 2. Viçosa: Editora UFV, 2003. 585 p.

CRUZ, C. D.; REGAZZI, A. J. **Modelos biométricos aplicados ao melhoramento genético**. 2. ed. Viçosa: Editora UFV, 2001. 390 p.

DAHER, R. F.; MORAES, C. F.; CRUZ, C. D.; PEREIRA, A. V.; XAVIER, D. F. Seleção de caracteres morfológicos discriminantes em capim-elefante (*Pennisetum purpureum* Schum.). **Revista Brasileira de Zootecnia**, v. 26, n. 2, p. 247-254, 1997. ISSN 1806-9290.

DAS, S.; MOHAPATRA, A. K.; PATNAIK, R. K.; SWAIN, D.; GANTAYAT, B. P. Clustering of *Jatropha curcas* provenances. **Agricultural Science Digest**, v. 28, n. 4, p. 298-300, 2008. ISSN 0253-150X.

DEHGAN, B.; WEBSTER, G. L. **Morphology and infrageneric relationships of the genus *Jatropha* (Euphorbiaceae)**. University of California Publications in Botany, University of California Press, 1979. 73 p.

DIAS, L. A. dos S. **Divergência genética e fenética multivariada na predição de híbridos e preservação de germoplasma de cacau (*Theobroma cacao* L.)**. 1994. 94 p. Tese (Doutorado) - Escola Superior de Agricultura "Luiz de Queiroz" - ESALQ, Piracicaba, 1994.

DIAS, L. A. dos S.; KAGEYAMA, P. Y.; CASTRO, G. C. T. Divergência genética multivariada na preservação de germoplasma de cacau (*Theobroma cacao* L.). **Agrotropica**, v. 9, n. 1, p. 29-40, 1997. ISSN 0103-3816.

DIAS, L. A. dos S.; MISSIO, R. F.; DIAS, D. C. F. S. Antiquity, botany, origin and domestication of *Jatropha curcas* (Euphorbiaceae), a plant species with potential for biodiesel production. **Genetics and Molecular Research**, v. 11, n. 3, p. 2719-2728, 2012. ISSN 1676-5680.

DRAY, S.; DUFOUR, A. B. The ade4 package: implementing the duality diagram for ecologists. **Journal of Statistical Software**, v. 22, n. 4, p. 1-20, 2007. ISSN 1548-7660.

DRUMOND, M. A.; ARRUDA, F. P. de; ANJOS, J. B. dos. Pinhão-Manso - *Jatropha curcas* L. **Documentos**, 212. Petrolina, PE: Embrapa Semi-Árido, dezembro 2008. 15 p. ISSN 1808-9992.

DRUMOND, M. A.; SANTOS, C. A. F.; OLIVEIRA, V. R. de; MARTINS, J. C.; ANJOS, J. B. dos; EVANGELISTA, M. R. V. Desempenho agrônômico de genótipos de pinhão-manso no semi-árido pernambucano. **Ciência Rural** [online], v. 40, n. 1, p. 44-47, 2010. Epub 20-Nov-2009. ISSN 0103-8478.

DUARTE, J. M.; SANTOS, J. B. dos; MELO, L. C. Comparison of similarity coefficients based on RAPD markers in the common bean. **Genetics and Molecular Biology**, v. 22, n. 3, p. 427-432, 1999. ISSN 1415-4757.

DURÃES, F. O. M.; LAVIOLA, B. G.; SUNDFELD, E.; MENDONÇA, S.; BHERING, L. L. Pesquisa, desenvolvimento e inovação em pinhão-manso para produção de biocombustíveis. **Documentos**, 01. Brasília, DF: Embrapa Agroenergia, novembro 2009. 17 p. ISSN 2177-4439.

ENGELMANN, F. *In vitro* conservation methods. In: CALLOW, J. A.; FORD-LLOYD, B. V. and NEWBURY, H. J. (eds.). **Biotechnology and plant genetic resources: conservation and use**. CAB International Publishing, Wallingford, UK, 1997. p. 119-162.

ENGELMANN, F.; ENGELS, J. M. M. Technologies and strategies for *ex situ* conservation. In: ENGELS, J. M. M.; RAMANATHA RAO, V.; BROWN, A. H. D. and JACKSON, M. T. (eds.). **Managing plant genetic diversity**. CAB International Publishing, Wallingford, UK, 2002. p. 89-103.

ENGELS, J. M. M.; VISSER, L. A guide to effective management of germplasm collections. **IPGRI Handbooks for Genebanks**, No. 6. IPGRI, Rome, Italy, 2003. 172 p. ISBN 92-9043-582-8.

EVANS, F. J. **Environmental hazards of diterpene esters from plants**. Boca Raton, Flórida: CRC Press, 1986.

FALCONER, D. S. **Introdução à genética quantitativa**. Viçosa. Universidade Federal de Viçosa, MG, 1981. 279 p.

FALEIRO, F. G. **Marcadores genético-moleculares aplicados a programas de conservação e uso de recursos genéticos**. Planaltina, DF. Embrapa Cerrados, 2007. 102 p. ISBN 978-85-7075-035-8.

FERRARI, R. A.; CASARINI, M. B.; MARQUES, D. de A.; SIQUEIRA, W. J. Avaliação da composição química e de constituinte tóxico em acessos de pinhão-manso de diferentes origens. **Brazilian Journal of Food Technology**, Campinas, v. 12, n. 4, p. 309-314, out-dez. 2009. ISSN 1981-6723.

FERREIRA, F. M.; BARROS, W. S.; SILVA, F. L. da.; BARBOSA, M. H. P.; CRUZ, C. D.; BASTOS, I. T. Relações fenotípicas e genotípicas entre componentes de produção em cana-de-açúcar. **Bragantia**, Campinas, v. 66, n. 4, p. 527-533, 2007a. ISSN 0006-8705.

FERREIRA, M. E.; GRATTAPAGLIA, D. **Introdução ao uso de marcadores moleculares em análise genética**. 3ª ed. Brasília: DF, Embrapa Cenargen, 1998. 220 p.

FERREIRA, M. E.; MORETZSOHN, M. de C.; BUSO, G. S. C. Fundamentos de caracterização molecular de germoplasma vegetal. In: NASS, L. L. (Editor técnico). **Recursos genéticos vegetais**. Brasília, DF: Embrapa Recursos Genéticos e Biotecnologia, 2007b. p. 377-415. ISBN 978-85-87697-34-9.

FOIDL, N.; FOIDL, G.; SANCHEZ, M.; MITTELBAACH, M.; HACKEL, S. *Jatropha curcas* L. as a source for the production of biofuel in Nicaragua. **Bioresource Technology**, Great Britain, v. 58, n. 1, p. 77-82, 1996. ISSN 0960-8524.

FRANKEL, O. H.; BROWN, A. H. D.; BURDON, J. J. **Conservation of plant biodiversity**. United Kingdom. Cambridge University Press, 1995. 299 p.

FREIRE, F. das C. O.; PARENTE, G. B. As Doenças das Jatrofas (*Jatropha curcas* L. e *J. podagrica* Hook.) no Estado do Ceará. **Comunicado Técnico** [online], 120. Fortaleza, CE: Embrapa Agroindústria Tropical, dezembro 2006. 4 p. ISSN 1679-6535.

FREITAS, R. G. de. **Caracterização de germoplasma de pinhão manso (*Jatropha curcas* L.) por descritores morfo-agronômicos**. 2010. 62 p. Dissertação (Mestrado) - Universidade Federal de Viçosa - UFV, Viçosa, 2010.

GIACOMETTI, D. C.; FERREIRA, F. R. Organização e uso de germoplasma de fruteiras. *In*: CONGRESSO BRASILEIRO DE FRUTICULTURA, 9, 1987. Campinas. **Anais...** Campinas: SBF, 1988. p. 11-17.

GOEDERT, C. O. de. Histórico e avanços em recursos genéticos no Brasil. *In*: NASS, L. L. (Editor técnico). **Recursos genéticos vegetais**. Brasília, DF: Embrapa Recursos Genéticos e Biotecnologia, 2007. p. 23-59. ISBN 978-85-87697-34-9.

GOWER, J. C. A comparison of some methods of cluster analysis. **Biometrics**, Arlington, v. 23, n. 4, p. 623-637, 1967. ISSN 1541-0420.

GOWER, J. C. A general coefficient of similarity and some of its properties. **Biometrics**, Arlington, v. 27, n. 4, p. 857-874, 1971. ISSN 1541-0420.

HARLAN, J. R. **Crops and Man**. American Society of Agronomy, Madison, Wisconsin, 1975. 295 p.

HAWKES, J. G. Genetic conservation of recalcitrant species - an overview. *In*: WITHERS, L. A. and WILLIAMS, J. T. (eds.). **Crop genetic resources - The conservation of difficult material**. 1980. **Proceedings...** Int. Workshop, Univ. of Reading, U.K., Paris, International Union of Biological Sciences, Series B 42, 1982. p. 83-92.

HAWKES, J. G. **The diversity of crop plants**. Harvard University Press, Cambridge, Mass, 1983. 184 p. ISBN 0-674-2186-X.

HELLER, J. **Physic nut. *Jatropha curcas* L. Promoting the conservation and use of underutilized and neglected crops**. Rome: Institute of Plant Genetics and Crop Plant Research, 1996. 66 p.

HIDALGO, R. Conservation *ex situ*. *In*: CASTILLO, R.; ESTRELLA, J. e TAPIA, C. (eds.). **Técnicas para a gestão e uso dos recursos fitogenéticos**. Instituto Nacional de Investigaciones Agropecuárias, Equador, 1991. p. 71-87.

HODGKIN, T.; BROWN, A. H. D.; VAN HINTUM, T. J. L.; VILELA-MORALES, E. A. **Core collections of plant genetic resources**. United Kingdom, John Wiley & Sons, 1995. 269 p.

IBPGR (comp.). **Elsevier's dictionary of plant genetic resources**. Netherlands, Elsevier Science Publishers B.V., 1991. 187 p. ISSN 0922-3444.

IBPGR. **Cotton descriptors** (Revised). Rome: International Board for Plant Genetic Resources, 1985. 25 p.

IBPGR. **Descriptors for soyabean**. Rome: International Board for Plant Genetic Resources, 1984. 38 p.

INSTITUTO NACIONAL DE METEOROLOGIA (INMET). **Dados relativos às Normais Climatológicas**. Disponível em:< <http://www.inmet.gov.br/portal/>>. Último acesso em 20 de março de 2013.

JACCARD, P. Étude comparative de la distribution florale dans une portion des Alpes et des Jura. **Bulletin Societe' Vandoise des sciences naturelles**, v. 37, p. 547-579, 1901.

JACKSON, A. A.; SOMERS, K. M.; HARVEY, H. H. Similarity coefficients: measures for co-occurrence and association or simply measures of occurrence? **American Naturalist**, v. 133, n. 3, p. 436-453, 1989. ISSN 1537-5323.

JARAMILLO, S.; BAENA, M. **Material de apoio a la capacitación em conservación ex situ de recursos fitogenéticos**. Instituto Internacional para os Recursos Fitogenéticos. Cali, Colômbia, 2000. 221 p.

JARVIS, D. I.; MYEZ, L.; KLEMICK, H.; GUARINO, L.; SMALE, M.; BROWN, A. H. D.; SADIKI, M.; STHAPIT, B.; HODGKIN, T. **A training guide for in situ conservation on-farm**, Version 1. IPGRI, Rome, Italy, 2000. 161 p. ISBN 92-9043-452-X.

JOHNSON, T. S.; ESWARAN, N.; SUJATHA, M. Molecular approaches to improvement of *Jatropha curcas* Linn. as a sustainable energy crop. **Plant Cell Reports**, v. 30, p. 1573-1591, 2011. ISSN 1432-203X.

JOLLIFFE, I. T. Discarding variables in a principal component analysis. I. Artificial data. **Applied Statistics**, v. 21, n. 2, p. 160-173, 1972.

JOLLIFFE, I. T. Discarding variables in a principal component analysis. II. Real data. **Applied Statistics**, v. 22, n. 1, p. 21-31, 1973.

JUHÁSZ, A. C. P.; PIMENTA, S.; SOARES, B. O.; MORAIS, D. de L. B.; RABELLO, H. de O. Biologia floral e polinização artificial de pinhão-mansão no norte de Minas Gerais. **Pesquisa Agropecuária Brasileira**, Brasília, v. 44, n. 9, p. 1073-1077, setembro 2009. ISSN 1678-3921.

KARP, A.; KRESOVICH, S.; BHAT, K. V.; AYAD, W. G.; HODGKIN, T. Molecular tools in plant genetic resources conservation: a guide to the technologies. **IPGRI Technical Bulletin No. 2**. International Plant Genetic Resources Institute, Rome, Italy, 1997. 47 p. ISBN 92-9043-323-X.

KELLEY, L. A.; GARDNER, S. P.; SUTCLIFFE, M. J. An automated approach for clustering an ensemble of NMR-derived protein structures into conformationally related subfamilies. **Protein Engineering**, v. 9, n. 11, p. 1063-1065, 1996. ISSN 1741-0134.

KRUSKAL, J. B. Multidimensional-scaling by optimizing goodness of fit to a non-metric hypothesis. **Psychometrica**, Williamsburg, v. 29, n. 1, p. 1-27, 1964.

LANCE, G. N.; WILLIAMS, W. T. A general theory of classificatory sorting strategies. I. Hierarchical systems. **Computer Journal**, v. 9, n. 4, p. 373-380, 1967. ISSN 1460-2067.

LAVIOLA, B. G. **Diretrizes para o melhoramento de pinhão-manso visando à produção de biocombustíveis.** 2010. Artigo em Hypertexto. Disponível em: <http://www.infobibos.com/Artigos/2010_1/DiretrizesPinhaoManso/index.htm>. Acesso em: 19/09/2011.

LAVIOLA, B. G.; BHERING, L. L.; MENDONÇA, S.; ROSADO, T. B.; ALBRECHT, J. C. Caracterização morfo-agronômica do banco de germoplasma de pinhão-manso na fase jovem. **Bioscience Journal**, Uberlândia, v. 27, n. 3, p. 371-379, 2011. ISSN 1981-3163.

LAVIOLA, B. G.; DIAS, L. A. S. Teor e acúmulo de nutrientes em folhas e frutos de pinhão-manso. **Revista Brasileira de Ciência do Solo**, Viçosa, v. 32, n. 5, p. 1969-1975, 2008. ISSN 0100-0683.

LAVIOLA, B. G.; ROSADO, T. B.; BHERING, L. L.; KOBAYASHI, A. K.; RESENDE, M. D. V. de. Genetic parameters and variability in physic nut accessions during early developmental stages. **Pesquisa Agropecuária Brasileira** [online], v. 45, n. 10, p. 1117-1123, 2010. ISSN 0100-204X.

LOARCE, Y.; GALLEGO, R.; FERRER, E. A comparative analysis of the genetic relationship between rye cultivars using RFLP and RAPD markers. **Euphytica**, Wageningen, v. 88, n. 2, p. 107-115, 1996. ISSN 0014-2336.

LUKASZEWICZ, J. Sur la liaison et la division des points d'un ensemble fini. **Colloquium Mathematicae**, v. 2, p. 282-285, 1951. ISSN 0010-1354.

MAECHLER, M.; ROUSSEUW, P.; STRUYF, A.; HUBERT, M.; HORNIK, K. **cluster: Cluster Analysis Basics and Extensions.** R package version 1.14.1, 2011.

MAKKAR, H. P. S.; BECKER, K.; SCHMOOK, B. Edible provenances of *Jatropha curcas* from Quintana Roo state of Mexico and effects of roasting on antinutrient and toxic factor in seeds. **Plant Foods for Human Nutrition**, Dordrecht, v. 52, n. 1, p. 31-36, 1998. ISSN 0921-9668.

MALHOTRA, N. K. **Pesquisa de marketing: uma orientação aplicada.** 3ª edição. Porto Alegre, Bookman, 2001. 719 p.

MANTEL, N. The detection of disease clustering and generalized regression approach. **Cancer Research**, Birmingham, v. 27, n. 2, p. 209-220, 1967. ISSN 0008-5472.

MARIOT, M. P.; BARBIERI, R. L.; SINIGAGLIA, C.; RIBEIRO, M, V. Variabilidade em matrizes de acessos de espinheira-santa. **Ciência Rural**, Santa Maria, v. 38, n. 2, p. 351-357, mar-abr. 2008. ISSN 0103-8478.

MARQUES, D. de A.; FERRARI, R. A. O papel das novas biotecnologias no melhoramento genético do pinhão-manso. **Biológico**, São Paulo, v. 70, n. 2, p. 65-67, jul-dez. 2008.

MARTINS, S. R. **Agricultura, ambiente e sustentabilidade: seus limites para América Latina.** EMATER, 2001. 84 p. CD-ROM.

MARTIN, G.; MAYEUX, A. Réflexions sur les cultures oléagineuses énergétiques. II.- Le Pourghère (*Jatropha curcas* L.): un carburant possible. **Oléagineux**, v. 39, n. 5, p. 283-287, 1984.

MATHER, K.; JINKS, J. L. **Biometrical genetics**. Second revised edition. Ithaca, New York, Cornell University Press, 1971. 382 p.

MAXTED, N.; FORD-LLOYD, B. V.; HAWKES, J. G. Complementary conservation strategies. *In*: MAXTED, N.; FORD-LLOYD, B. V. and HAWKES, J. G. (eds.). **Plant genetic conservation: the in situ approach**. Chapman & Hall, London, United Kingdom, 1997. p. 20-55.

MENG, Y.; LI, C.; FRANCIS, G.; MAKAR, H. P. S. Current situation and prospects of *Jatropha curcas* as a multipurpose tree in China. **Agroforestry Systems**, v. 76, n. 2, p. 487-497, 2009. ISSN 1572-9680.

MILANI, M. Descritores de mamona utilizados pela Embrapa Algodão. **Documentos**, 192. Campina Grande, PB: Embrapa Algodão, 2008. 39 p. ISSN 0103-0205.

MILANI, M.; MIGUEL JÚNIOR, S. R.; OLIVEIRA, E. C. S. de; NÓBREGA, M. B. de M.; COSTA, M. N. da. Caracterização taxonômica de acessos de mamona (*Ricinus communis* L.) do banco ativo de germoplasma da Embrapa Algodão. **Boletim de pesquisa e desenvolvimento**, 67. Campina Grande, PB: Embrapa Algodão, 2006. 17 p. ISSN 0103-0841.

MIRANDA, G. V. **Diversidade genética e desempenho de cultivares de soja como progenitores**. 1998. 117 p. Tese (Doutorado em Genética e Melhoramento de Plantas) - Universidade Federal de Viçosa, Viçosa, 1998.

MOHAPATRA, S.; PANDA, P. K. Genetic variability on growth, phenological and seed characteristics of *Jatropha curcas* L. **Notulae Scientia Biologicae**, v. 2, n. 2, p. 127-132, 2010. ISSN 2067-3205.

MONTGOMERY, D. C.; PECK, E. A. **Introduction to linear regression analysis**. New York: John Wiley & Sons, 1981. 504 p.

MOREIRA, J. de A. N.; SANTOS, J. W. dos; OLIVEIRA, S. R. de M. **Abordagens e metodologias para avaliação de germoplasma**. Brasília: Embrapa SPI, 1994. 115 p. ISBN 85-85760-01-X.

MORRISON, D. F. **Multivariate statistical methods**. 2ª edição. Singapore: McGraw-Hill, 1976. 415 p.

MOURA, M. da C. C. L.; GONÇALVES, L. S. A.; SUDRÉ, C. P.; RODRIGUES, R.; AMARAL JÚNIOR, A. T. do; PEREIRA, T. N. S. Algoritmo de Gower na estimativa da divergência genética em germoplasma de pimenta. **Horticultura Brasileira**, v. 28, n. 2, p. 155-161, 2010. ISSN 0102-0536.

NASS, L. L.; NISHIKAWA, M. A. N.; FÁVERO, A. P.; LOPES, M. A. Pré-melhoramento de germoplasma vegetal. *In*: NASS, L. L. (Editor técnico). **Recursos genéticos vegetais**.

Brasília, DF: Embrapa Recursos Genéticos e Biotecnologia, 2007. p. 683-710. ISBN 978-85-87697-34-9.

NERY, A. R.; RODRIGUES, L. N.; SILVA, M. B. R.; FERNANDES, P. D.; CHAVES, L. H. G.; NETO, J. D.; GHEYI, H. R. Crescimento do pinhão-mansão irrigado com águas salinas em ambiente protegido. **Revista Brasileira de Engenharia Agrícola e Ambiental**, Campina Grande, PB, UAEEA/UFCG, v. 13, n. 5, p. 551-558, 2009. ISSN 1415-4366.

NOOR CAMELLIA, N. A.; THOHIRAH LEE, A.; ABDULLAH, N. A. P. Genetic relationships and diversity of *Jatropha curcas* accessions in Malaysia. **African Journal of Biotechnology**, v. 11, n. 13, p. 3048-3054, 2012. ISSN 1684-5315.

OKSANEN, J.; BLANCHET, F. G.; KINDT, R.; LEGENDRE, P.; MINCHIN, P. R.; O'HARA, R. B.; SIMPSON, G. L.; SOLYMOS, P.; STEVENS, M. H. H.; WAGNER, H. **vegan: Community Ecology Package**. R package version 2.0-2, 2011. Disponível em: URL <http://CRAN.R-project.org/package=vegan>.

OPENSHAW, K. A review of *Jatropha curcas*: an oil plant of unfulfilled promise. **Biomass and Bioenergy**, Silver Spring, v. 19, n. 1, p. 01-15, 2000. ISSN 0961-9534.

PAINTING, K. A.; PERRY M. C.; DENNING, R. A.; AYAD, W. G. **Guía para La Documentación de Recursos Genéticos**. Roma, Consejo Internacional de Recursos Fitogenéticos, 1993. 310 p. ISBN 92-9043-215-2.

PAIVA, A. L. da C.; TEIXEIRA, R. B.; YAMAKI, M.; MENEZES, G. R. de O.; LEITE, C. D. S.; TORRES, R. de A. Análise de componentes principais em características de produção de aves de postura. **Revista Brasileira de Zootecnia**, v. 39, n. 2, p. 285-288, 2010. ISSN 1806-9290.

PAIVA NETO, V. B. de; BRENHA, J. A. M.; FREITAS, F. B. de; ZUFFO, M. C. R.; ALVAREZ, R. de C. F. Aspectos da biologia reprodutiva de *Jatropha curcas* L. **Ciência e agrotecnologia**, Lavras, v. 34, n. 3, p. 558-563, mai-jun. 2010. ISSN 1413-7054.

PARODA, R. S.; ARORA, R. K. **Plant genetics resources conservation and management**. International Board for Plant Genetic Resources Regional Office for South and Southeast Asia, New Delhi, India, 1991. 309 p.

PATERNIANI, E. Diversidade genética em plantas. *In*: ENCONTRO SOBRE RECURSOS GENÉTICOS, 1, 1988. Jaboticabal, SP. **Anais...** Jaboticabal: FCAV, 1988. p. 75-77.

PEIXOTO, A. R. **Plantas oleaginosas arbóreas**. São Paulo: Nobel, 1973. 284 p.

PERONI, N.; HANAZAKI, N. Current and lost diversity of cultivated varieties, especially cassava, under swidden cultivation systems in the Brazilian Atlantic Forest. **Agriculture, Ecosystems and Environment**, Amsterdam, v. 92, n. 2, p. 171-183, 2002. ISSN 0167-8809.

PERSSONI, L. A. **Estratégias de análise da diversidade em germoplasma de cajueiro (*Anacardium spp* L.)**. 2007. 159 p. Tese (Doutorado em Genética e Melhoramento) - Universidade Federal de Viçosa, 2007.

PIZARRO, F. **Drenaje agrícola y recuperacion de suelos salinos**. Segunda edição. Madri: Editorial Española S.A., 1985. 542 p.

R DEVELOPMENT CORE TEAM. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria, 2011. ISBN 3-900051-07-0. Disponível em: URL <http://www.r-project.org/>.

RAMALHO, M. A. P.; SANTOS, J. B. dos; PINTO, C. A. B. P. **Genética na Agropecuária**. Lavras: UFLA, 2000. 472 p.

RAMALHO, M. A. P.; SANTOS, J. B. dos; ZIMMERMANN, M. J. de O. **Genética quantitativa em plantas autógamas: Aplicações ao melhoramento do feijoeiro**. Goiânia: ED. da UFG, 1993. 271 p.

RAMANATHA RAO, V. Complementary conservation strategy. *In*: FOURTH SOUTH ASIA COORDINATORS MEETING, 1-3 September, 1998. (In press). **Proceedings...** Nepal Agricultural Research Council. Khumaltar, Kathmandu, 1998.

RAMOS, D. P.; CASTRO, A. F.; CAMARGO, M. N. Levantamento detalhado de solos da área da Universidade Federal Rural do Rio de Janeiro. **Pesquisa Agropecuária Brasileira**, Série Agronomia, Brasília, v. 8, n. 1, p. 1-27, 1973. ISSN 1678-3921.

RAMOS, S. R. R.; FREIRE FILHO, F. R.; MEIRELLES, A. C. de; BARROS, G. B.; AZEVEDO, J. N. de; SANTOS, E. P. A. dos; ROCHA, M. de M.; SANTOS, J. O.; GALVÃO, J. C.; OLIVEIRA, C. R. R.; SOBRAL, P. V. C.; RIBEIRO, V. Q.; WETZEL, M. M. V. S. Banco de germoplasma de *Vigna* sp. da Embrapa Meio-Norte: Status e prioridades para manejo. **Magistra**, Cruz das Almas-BA, v. 18, n. especial, 2006. ISSN 0102-5333.

RAO, C. R. **Advanced statistical methods in biometric research**. New York, John Wiley, 1952.

RAO, G. R.; KORWAR, G. R.; SHANKER, A. K.; RAMAKRISHNA, Y. S. Genetic associations, variability and diversity in seed characters, growth, reproductive phenology and yield in *Jatropha curcas* (L.) accessions. **Trees: Structure and Function**, Berlin, v. 22, p. 697-709, 2008. ISSN 1432-2285.

RAO, N. K.; JACKSON, M. T. Seed production environment and storage longevity of japonica rices (*Oryza sativa* L.). **Seed Science Research**, v. 6, n. 1, p. 17-21, 1996. ISSN 0960-2585.

REDONDO, M. L. C. **Caracterização molecular por AFLP (Amplified Fragment Length Polymorphism) de acessos de pinhão-manso (*Jatropha curcas* L.)**. 2011. 90 p. Dissertação (Mestrado) - Centro de Energia Nuclear, Agricultura da Universidade de São Paulo, Piracicaba, 2011.

REIF, J. C.; MELCHINGER, A. E.; FRISCH, M. Genetical and mathematical properties of similarity and dissimilarity coefficients applied in plant breeding and seed bank management. **Crop Science**, Madison, v. 45, n. 1, p. 1-7, 2005. ISSN 0011-183X.

ROBERTS, E. H. Problems of long-term storage of seed and pollen for genetic resources conservation. *In*: FRANKEL, O. H. and HAWKES, J. G. (eds.). **Crop genetic resources for today and tomorrow**. Cambridge University Press, Cambridge, 1975. p. 269-295.

ROCHA, R. B.; MARCOLAN, A. L.; RAMALHO, A. R.; VIEIRA JÚNIOR, J. R.; FERNANDES, C. de F.; SPINELLI, V. M.; RIBEIRO, R. da S.; SILVA, F. C. G. da.; MILITÃO, J. S. L. T. Caracterização dos componentes de rendimento de óleo do pinhão manso (*Jatropha curcas* L.). **Circular Técnica**, 111. Porto Velho, RO: Embrapa Rondônia, 2009. 6 p. ISSN 0103-9334.

RODRIGUES, J. I. da S. **Mapeamento de QTLs associados a conteúdo de proteína, óleo e componentes de produção em soja**. 2008. 54 p. Dissertação (Mestrado em Genética e Melhoramento) - Universidade Federal de Viçosa, 2008.

RODRIGUES, R. Plano Nacional de Agroenergia 2006-2011. **Informe do Ministério da Agricultura, Pecuária e Abastecimento**, Brasília, v. 1, n. 1, 2005. 120 p.

ROSADO, T. B.; LAVIOLA, B. G.; FARIA, D. A.; PAPPAS, M. C. R.; BHERING, L. L.; QUIRINO, B.; GRATTAPAGLIA, D. Molecular markers reveal limited genetic diversity in a large germplasm collection of the biofuel crop *Jatropha curcas* L. in Brazil. **Crop Science**, v. 50, p. 2372-2382, 2010. ISSN 0011-183X.

SAAD, M. S.; RAMANATHA RAO, V. **Establishment and Management of Field Genebank, a Training Manual**. IPGRI-APO (Asia, Pacific and Oceania), Serdang, 2001. 121 p. ISBN 92-9043-464-3.

SANTOS, D. B. dos. **Procedimentos multivariados no agrupamento de genótipos de maracujazeiro com base em matriz de distância conjunta e em separado para características quantitativas e categóricas**. 2010. 59 p. Dissertação (Mestrado em Fitotecnia) - Universidade Federal do Recôncavo da Bahia, 2010.

SANTOS, C. A. F.; DRUMOND, M. A.; RODRIGUES, M. A.; EVANGELISTA, M. R. V. Genetic similarity of *Jatropha curcas* accessions based on AFLP markers. **Crop Breeding and Applied Biotechnology** [online], v. 10, n. 4, p. 364-369, 2010a. ISSN 1984-7033.

SANTOS, C. M. dos.; ENDRES, L.; WANDERLEY FILHO, H. C. de Lins.; ROLIM, E. V.; FERREIRA, V. M. Fenologia e crescimento do pinhão-manso cultivado na zona da mata do estado de Alagoas, Brasil. **Scientia Agraria**, Curitiba, v. 11, n. 3, p. 201-209, May/Jun, 2010b. ISSN 1983-2443.

SANTOS, C. A. F.; OLIVEIRA, C. A. V.; MENEZES, E. A. Seleção de descritores na caracterização e avaliação preliminar de germoplasma de guandu. **Pesquisa Agropecuária Brasileira**, Brasília, v. 30, n. 7, p. 971-975, 1995. ISSN 1678-3921.

SANTOSO, B. B. Variation in oil contents, and seed and seedling characteristics of *Jatropha curcas* of West Nusa Tenggara selected genotypes and their first improved population. **Nusantara Bioscience**, v. 3, n. 3, p. 130-135, 2011. ISSN 2087-3948

SATURNINO, H. M.; PACHECO, D. D.; KAKIDA, J.; TOMINAGA, N.; GONÇALVES, N. P. Cultura do pinhão-mansão (*Jatropha curcas* L.). **Informe Agropecuário**, Belo Horizonte, v. 26, n. 229, p. 44–78, 2005. ISSN 0100-3364.

SCOWCROFT, W. R. Genetic variability in tissue culture: Impact on germplasm conservation and utilization. **IBPGR Technical Report**. International Board for Plant Genetic Resources, Rome, Italy, 1984. 41 p.

SHABANIMOFRAD, M.; RAFII, M. Y.; MEGAT WAHAB, P. E.; BIABANI, A. R.; LATIF, M. A. Phenotypic, genotypic and genetic divergence found in 48 newly collected Malaysian accessions of *Jatropha curcas* L. **Industrial Crops and Products**, v. 42, p. 543-551, 2013.

SILVA, R. M.; BANDEL, G.; FARALDO, M. I. F.; MARTINS, P. S. Biologia reprodutiva de etnovarietades de mandioca. **Scientia Agricola**, Piracicaba, v. 58, n. 1, p. 101-107, 2001. ISSN 0103-9016.

SILVA, V. A.; MORAIS, D. de L. B.; KAKIDA, J.; FERREIRA, E. A.; SILVA, V. de F. Concentração do ciclo de produção de pinhão-mansão por meio de podas de formação ou de produção. **Pesquisa Agropecuária Brasileira**, Brasília, v. 47, n. 1, p. 134-137, jan. 2012. ISSN 1678-3921.

SILVA, M. B. R.; NETO, J. D.; FERNANDES, P. D.; FARIAS, M. S. S. de. Cultivo de pinhão-mansão sob condições de estresse hídrico e salino, em ambiente protegido. **Revista de Biologia e Ciências da Terra**, Campina Grande, v. 9, n. 2, 2º Semestre 2009c. ISSN 1519-5228.

SILVA, G. C.; OLIVEIRA, F. J. de; ANUNCIÇÃO FILHO, C. J. da; SIMÕES NETO, D. E.; MELO, L. J. O. T. de. Divergência genética entre genótipos de cana-de-açúcar. **Revista Brasileira de Ciências Agrárias**, Recife, PE, UFRPE, v. 6, n. 1, p. 52-58, jan-mar. 2011. ISSN 1981-0997.

SILVA, E. N. da; SILVEIRA, J. A. G.; RODRIGUES, C. R. F.; LIMA, C. S. de; VIÉGAS, R. A. Contribuição de solutos orgânicos e inorgânicos no ajustamento osmótico de pinhão-mansão submetido à salinidade. **Pesquisa Agropecuária Brasileira**, Brasília, v. 44, n. 5, p. 437-445, maio 2009b. ISSN 1678-3921.

SILVA, E. de B.; TANURE, L. P. P.; SANTOS, S. R.; JÚNIOR, P. S. de R. Sintomas visuais de deficiências nutricionais em pinhão-mansão. **Pesquisa Agropecuária Brasileira**, Brasília, v. 44, n. 4, p. 392-397, abril 2009a. ISSN 1678-3921.

SNEATH, P. H.; SOKAL, R. R. **Numerical taxonomy: The principles and practice of numerical classification**. San Francisco: W. H. Freeman, 1973. 573 p.

SOKAL, R. R.; MICHENER, C. D. A statistical method for evaluating systematic relationships. **University of Kansas Science Bulletin**, v. 38, p. 1409-1438, 1958.

SOKAL, R. R.; ROHLF, F. J. The comparison of dendrograms by objective methods. **Taxon**, v. 11, p. 33-40, 1962.

STÄHELIN, D.; VALENTINI, G.; ANDRADE, L. R. B. de; VERISSIMO, M. A. A.; BERTOLDO, J. G.; GUIDOLIN, A. F.; COIMBRA, J. L. M. *Screening* multivariado entre acessos e cultivares de feijão do grupo preto para utilização em blocos de cruzamento. **Revista Biotemas**, v. 24, n. 1, p. 95-103, março 2011. ISSN 2175-7925.

STRAPASSON, E.; VENCOSKY, R.; BATISTA, L. A. R. Seleção de Descritores na Caracterização de Germoplasma de *Paspalum sp.* por meio de Componentes Principais. **Revista Brasileira de Zootecnia**, v. 29, n. 2, p. 373-381, 2000. ISSN 1806-9290.

SUKARIN, W.; YAMADA, Y.; SAKAGUCHI, S. Characteristics of physic nut, *Jatropha curcas* L. as a new biomass crop in the Tropics. **Japan Agriculture Research Quarterly**, Japan, v. 20, n. 4, p. 302-303, 1987. ISSN 0919-8822.

SUN, Q. B.; LI, L. F.; LI, Y.; WU, G. J.; GE, X. J. SSR and AFLP markers reveal low genetic diversity in the biofuel plant *Jatropha curcas* in China. **Crop Science**, Madison, v. 48, p. 1865-1871, 2008. ISSN 0011-183X.

SUNIL, N.; VARAPRASAD, K. S.; SIVARAJ, N.; SURESH KUMAR, T.; BABU ABRAHAM; PRASAD, R. B. N. Assessing *Jatropha curcas* L. germplasm in situ - a case study. **Biomass and Bioenergy**, v. 32, p. 198-202, 2008. ISSN 0961-9534.

TAIZ, L.; ZEIGHER, E. **Plant physiology**. 3ªed. Porto Alegre: ARTMED, 2004. 719 p.

TOWIL, L. E.; ROOS, E. E. Techniques for preserving of plant germplasm. *In*: KNUTSON, L. and STONER, A. K. (eds.). **Biotic diversity and germplasm preservation, global imperatives**. Kluwer Academic Publishers, Netherlands, 1989. p. 379-403.

VALE, M. N. **Agrupamentos de Dados: Avaliação de Métodos e Desenvolvimento de Aplicativo para Análise de Grupos**. 2005. 120 p. Dissertação (Mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, RJ, 2005.

VALE, L. S.; SEVERINO, L. S.; BELTRÃO, N. E. de M. Crescimento do pinhão-mansão em solo compactado. *In*: CONGRESSO BRASILEIRO DE TECNOLOGIA DE BIODIESEL, 1, 2006. Brasília, D. F. **Anais...** Brasília-D. F: MCT/ABIPTI, 2006a. p. 78-91.

VALE, L. S.; SEVERINO, L. S.; BELTRÃO, N. E. de M. Efeito da salinidade da água sobre o pinhão-mansão. *In*: CONGRESSO BRASILEIRO DE TECNOLOGIA DE BIODIESEL, 1, 2006. Brasília, D. F. **Anais...** Brasília-D. F: MCT/ABIPTI, 2006b. p. 87-90.

VALENTIN, J. L. Tópicos em Tratamento de Dados Biológicos. *In*: PERES-NETO, P. R.; VALENTIN, J. L. e FERNANDEZ, F. A. S. (eds.). **Oecologia brasiliensis**, 2. Programa de Pós-Graduação em Ecologia, Instituto de Biologia, UFRJ, Rio de Janeiro, 1995. p. 27-55.

VALLS, J. F. M. Caracterização de recursos genéticos vegetais. *In*: NASS, L. L. (Editor técnico). **Recursos genéticos vegetais**. Brasília, DF: Embrapa Recursos Genéticos e Biotecnologia, 2007. p. 281-300. ISBN 978-85-87697-34-9.

VALOIS, A. C. C.; SALOMÃO, A. N.; ALLEM, A. C. Glossário de recursos genéticos vegetais. **Documentos**, 22. Brasília, DF: Embrapa Cenargen, 1996. 62 p.

VAN HINTUM, T. J. L.; BROWN, A. H. D.; SPILLANE, C.; HODGKIN, T. Core collections of plant genetic resources. **IPGRI Technical Bulletin No. 3**. International Plant Genetic Resources Institute, Rome, Italy, 2000. 48 p. ISBN 92-9043-454-6.

VETELÄINEN, M.; NEGRI, V.; MAXTED, N. European landraces on farm conservation, management and use. **Bioversity Technical Bulletin No. 15**. Bioversity International, Rome, Italy, 2009. 344 p. ISBN 978-92-9043-805-2.

VIDAL, W. N.; VIDAL, M. R. S. **Botânica - Organografia; quadros sinóticos ilustrados de fanerógamos**. 4ª ed. rev. ampl. Viçosa: Editora UFV, 2006. 124 p. ISBN 85-7269-054-9.

VILELA-MORALES, E. A.; VALOIS, A. C. C. Princípios genéticos para recursos genéticos. *In*: DIÁLOGO XLV: CONSERVAÇÃO DO GERMOPLASMA VEGETAL, 1994. **Anais...** Instituto Interamericano de Cooperação para a Agricultura, Uruguay, 1996. p. 35-48.

WANI, S. P.; OSMAN, M.; D'SILVA, E.; SREEDEVI, T. K. Improved livelihoods and environmental protection through biodiesel plantations in Asia. **Asian Biotechnology and Development Review**, v. 8, n. 2, p. 11-29, 2006. ISSN 0972-7566.

WEBB, A. R. **Statistical Pattern Recognition**, Second Edition. John Wiley & Sons, Ltd. 2002. 496 p. ISBN 0-470-84513-9.

WHITE, D. **maptree: Mapping, pruning, and graphing tree models**. R package version 1.4-6, 2010. Disponível em: URL <http://CRAN.R-project.org/package=maptree>.

WESP, C. L.; MARTINELI, J. A.; CHAVES M. S.; GRAICHEN, F. A. S.; FEDERIZZI, L. C. Herança da resistência quantitativa à ferrugem da folha em linhagens recombinantes de aveia branca. **Tropical Plant Pathology**, Brasília, v. 33, n. 2, p. 138-147, 2008. ISSN 1982-5676.

WRIGHT, S. Correlation and causation. **Journal of Agricultural Research**, Washington, v. 20, n. 7, p. 557-585, 1921.

ZHANG, Z.; GUO, X.; LIU, B.; TANG, L.; CHEN, F. Genetic diversity and genetic relationship of *Jatropha curcas* between China and Southeast Asian revealed by amplified fragment length polymorphisms. **African Journal of Biotechnology**, v. 10, n. 15, p. 2825-2832, 2011. ISSN 1684-5315.

7. ANEXOS

Anexo A. Lista de descritores morfoagronômicos para caracterização e avaliação de <i>Jatropha curcas</i>	90
Anexo B. Descrição de classes observadas em descritores relativos a variáveis multicatóricas e binárias em 60 genótipos da Coleção de germoplasma de <i>Jatropha curcas</i> da UFRRJ.	98
Anexo C. Estimativas da variância (autovalores), respectivos coeficientes de ponderação (autovetores) associados aos componentes principais (CP) e contribuição acumulada a partir de 36 variáveis quantitativas avaliadas em 60 genótipos de <i>Jatropha curcas</i> da Coleção de germoplasma da UFRRJ.....	99
Anexo D. Estimativas dos efeitos diretos e indiretos de 15 variáveis explicativas sobre a característica Número de sementes por planta (NSE) em 60 genótipos de <i>Jatropha curcas</i> da UFRRJ.....	101
Anexo E. Fotos da Coleção de germoplasma de <i>Jatropha curcas</i> da UFRRJ, localizada no campo experimental do DFITO/IA, em Seropédica, RJ, aos três anos de idade.....	106
Anexo F. Fotos das 16 variáveis qualitativas (13 multicatóricas e 3 binárias) caracterizadas em 60 genótipos da Coleção de germoplasma de <i>Jatropha curcas</i> da UFRRJ.....	1067
Anexo G. Fotos ilustrativas para auxiliar na mensuração de 18 variáveis quantitativas, elaboradas a partir da Coleção de germoplasma de <i>Jatropha curcas</i> da UFRRJ.....	1068
Anexo H. Croqui da Coleção de germoplasma de <i>Jatropha curcas</i> da UFRRJ e respectiva diagramação dos acessos definidos para caracterização.....	109

Anexo A. Lista preliminar de descritores morfoagronômicos para *Jatropha curcas*.

1. Descritores relacionados a aspectos vegetativos

1.1. Altura de plantas desenvolvidas (APD)

Avaliada a partir do final do 1º (primeiro) ciclo reprodutivo da planta. A altura de plantas desenvolvidas (em cm) deve ser medida entre a base do caule (solo) e a extremidade superior do ramo mais alto.

1.2. Projeção da copa no sentido Norte (PCOn)

Avaliada a partir do final do 1º (primeiro) ciclo reprodutivo da planta. A projeção da copa (em m) deve ser mensurada a partir da região central da base do caule até a projeção máxima do ramo no sentido Norte.

1.3. Projeção da copa no sentido Sul (PCOs)

Avaliada a partir do final do 1º (primeiro) ciclo reprodutivo da planta. A projeção da copa (em m) deve ser mensurada a partir da região central da base do caule até a projeção máxima do ramo no sentido Sul.

1.4. Projeção da copa no sentido Leste (PCOl)

Avaliada a partir do final do 1º (primeiro) ciclo reprodutivo da planta. A projeção da copa (em m) deve ser mensurada a partir da região central da base do caule até a projeção máxima do ramo no sentido Leste.

1.5. Projeção da copa no sentido Oeste (PCOo)

Avaliada a partir do final do 1º (primeiro) ciclo reprodutivo da planta. A projeção da copa (em m) deve ser mensurada a partir da região central da base do caule até a projeção máxima do ramo no sentido Oeste.

1.6. Área de projeção da copa (PCO)

Avaliada a partir do final do 1º (primeiro) ciclo reprodutivo da planta. A área de projeção da copa (em m²) deve ser obtida através da seguinte fórmula: $[(PCOn + PCOs) \times (PCOl + PCOo)] / 2$, onde os descritores PCOn, PCOs, PCOl e PCOo, previamente citados, corresponderão aos valores de projeção nos sentidos Norte, Sul, Leste e Oeste, respectivamente. Para tal, cada valor de projeção (em m) deve ser mensurado a partir da região central da base do caule até a projeção máxima do ramo em cada sentido.

1.7. Diâmetro do caule (DCA)

Avaliado em qualquer fase do ciclo de desenvolvimento da planta. O diâmetro do caule (em cm) deve ser medido na região do colo da planta a uma altura máxima de 10 cm da base do caule (solo).

1.8. Número de ramos totais (NRT)

Avaliado em qualquer fase do ciclo de desenvolvimento da planta. O número de ramos totais deve ser obtido pela quantificação do número total de ramos da planta.

1.9. Número de ramos com inflorescência (NRI)

Avaliado a partir do 1º (primeiro) ciclo reprodutivo da planta. O número de ramos com inflorescência deve ser obtido pela quantificação do número total de ramos que culminam em uma inflorescência.

1.10. Comprimento de ramos primários (CRP)

Avaliado a partir do 1º (primeiro) ciclo reprodutivo da planta. O comprimento de ramos primários deve ser obtido pela média das medidas efetuadas entre a base do ramo (extremidade inferior) e sua extremidade superior nos 3 ramos primários de maior comprimento. A média do comprimento de ramos primários por planta (em cm) deve ser dada pela divisão da soma de todas as medidas aferidas pelo número total de ramos primários avaliados. Para plantas de arquitetura ereta e sem ramificações, o comprimento de ramos primários deve ser dado pela própria altura da planta, mensurada entre a base do caule (solo), nesse caso, considerado o único ramo primário, e sua extremidade superior.

1.11. Comprimento do entrenó de ramos primários (CERp)

Avaliado a partir do 1º (primeiro) ciclo reprodutivo da planta. O comprimento do entrenó de ramos primários deve ser obtido pela média das medidas efetuadas nos 3 ramos primários de maior comprimento, avaliando-se 6 entrenós na porção mediana de cada ramo primário. A média do comprimento do entrenó de ramos primários por planta (em cm) deve ser dada pela divisão da soma de todos os entrenós mensurados pelo número total de entrenós avaliados. Para plantas de arquitetura ereta e sem ramificações, o comprimento do entrenó de ramos primários deve ser dado pela própria média das medidas aferidas dos 6 entrenós da porção mediana do caule da planta, nesse caso, considerado o único ramo primário.

1.12. Comprimento de ramos com inflorescência (CRI)

Avaliado a partir do 1º (primeiro) ciclo reprodutivo da planta. O comprimento de ramos com inflorescência deve ser obtido pela média das medidas efetuadas entre a base do ramo (extremidade inferior) e sua extremidade superior nos 3 ramos de maior comprimento com inflorescência. A média do comprimento de ramos com inflorescência por planta (em cm) deve ser dada pela divisão da soma de todas as medidas aferidas pelo número total de ramos com inflorescência avaliados. Para plantas de arquitetura ereta e sem ramificações, o comprimento de ramos com inflorescência deve ser dado pela própria altura da planta, mensurada entre a base do caule (solo), nesse caso, considerado o único ramo com inflorescência, e sua extremidade superior.

1.13. Comprimento do entrenó de ramos com inflorescência (CERi)

Avaliado a partir do 1º (primeiro) ciclo reprodutivo da planta. O comprimento do entrenó de ramos com inflorescência deve ser obtido pela média das medidas efetuadas nos 3 ramos de maior comprimento com inflorescência, avaliando-se 4 entrenós na porção mediana de cada ramo com inflorescência. A média do comprimento do entrenó de ramos com inflorescência por planta (em cm) será dada pela divisão da soma de todos os entrenós mensurados pelo número total de entrenós avaliados. Para plantas de arquitetura ereta e sem ramificações, o comprimento do entrenó de ramos com inflorescência será dado pela própria média das medidas aferidas dos 4 entrenós da porção mediana do caule da planta, nesse caso, considerado o único ramo com inflorescência.

1.14. Número de lóbulos foliar (NLF)

Avaliado em qualquer fase do ciclo de desenvolvimento da planta. O número de lóbulos foliar deve ser obtido pela quantificação dos lóbulos em 10 folhas completamente desenvolvidas por planta.

1.15. Ângulo de inserção da folha no caule (AIFc)

Avaliado em qualquer fase do ciclo de desenvolvimento da planta. O ângulo de inserção da folha no caule deve ser obtido pela média das medidas entre a inserção da folha (pecíolo) e o caule em 10 folhas completamente desenvolvidas por planta. A média do ângulo de inserção da folha no caule (em graus) deve ser dada pela divisão do valor da soma de todos os ângulos aferidos pelo número total de folhas avaliadas.

1.16. Comprimento do pecíolo (CMP)

Avaliado em qualquer fase do ciclo de desenvolvimento da planta. O comprimento do pecíolo deve ser obtido pela média do comprimento do pecíolo em 10 folhas completamente desenvolvidas por planta. A média do comprimento do pecíolo (em cm) deve ser dada pela divisão do valor da soma do comprimento de todos os pecíolos mensurados pelo número total de folhas avaliadas.

1.17. Comprimento da folha (CFO)

Avaliado em qualquer fase do ciclo de desenvolvimento da planta. O comprimento da folha deve ser obtido pela média do comprimento em 10 folhas completamente desenvolvidas por planta. A média do comprimento das folhas (em cm) deve ser dada pela divisão do valor da soma do comprimento de todas as folhas mensuradas pelo número total de folhas avaliadas.

1.18. Largura da folha (LFO)

Avaliada em qualquer fase do ciclo de desenvolvimento da planta. A largura da folha deve ser obtida pela média da largura em 10 folhas completamente desenvolvidas por planta. A média da largura das folhas (em cm) deve ser dada pela divisão do valor da soma da largura de todas as folhas mensuradas pelo número total de folhas avaliadas.

1.19. Relação entre o comprimento e a largura da folha (CLFo)

Avaliada em qualquer fase do ciclo de desenvolvimento da planta. Relação obtida através da razão entre o valor do *comprimento da folha* pelo valor da *largura da folha* (ambos citados anteriormente).

1.20. Arquitetura da planta (ARQ)

Observada a partir do final do 1º (primeiro) ciclo reprodutivo da planta. A arquitetura da planta deve ser classificada em:

- 1- Ereta
- 2- Fechada
- 3- Aberta

1.21. Tipo de ramificação (RAM)

Observada a partir do final do 1º (primeiro) ciclo reprodutivo da planta. O tipo de ramificação da planta deve ser classificado em:

- 0- Ausência de ramificação
- 1- Monopodial
- 2- Dicasial
- 3- Simpodial

1.22. Coloração do caule (CCA)

Observada a partir do final do 1º (primeiro) ciclo reprodutivo da planta. A coloração do caule deve ser observada na região mediana deste, desconsiderando a presença de cerosidade e classificando sua coloração em:

- 1- Verde
- 2- Verde escuro
- 3- Cinza
- 4- Roxo
- 5- Azul

1.23. Coloração do ramo (CRA)

Observada a partir do final do 1º (primeiro) ciclo reprodutivo da planta. A coloração do ramo deve ser observada desconsiderando a presença de cerosidade na ramificação, classificando sua coloração em:

- 1- Verde
- 2- Verde escuro
- 3- Cinza
- 4- Roxo
- 5- Azul

1.24. Coloração de folhas jovens (CFJ)

Observada em qualquer fase do ciclo de desenvolvimento da planta. A coloração de folhas jovens deve ser observada na face superior (adaxial) do limbo foliar em uma amostra representativa de 10 folhas jovens por planta, classificando-as em:

- 1- Verde claro
- 2- Verde
- 3- Verde escuro
- 4- Roxa
- 5- Vermelha

1.25. Coloração de folhas desenvolvidas (CFD)

Observada em qualquer fase do ciclo de desenvolvimento da planta. A coloração de folhas desenvolvidas deve ser observada na face superior (adaxial) do limbo foliar em uma amostra representativa de 10 folhas completamente desenvolvidas por planta, classificando-as em:

- 1- Verde claro
- 2- Verde
- 3- Verde escuro
- 4- Roxa
- 5- Vermelha

1.26. Coloração da nervura foliar (CNF)

Observada em qualquer fase do ciclo de desenvolvimento da planta. A coloração da nervura foliar deve ser observada na face superior (adaxial) do limbo em uma amostra representativa de 10 folhas completamente desenvolvidas por planta, classificando-as em:

- 1- Verde
- 2- Roxa
- 3- Vermelha

1.27. Coloração do pecíolo (CPE)

Observada em qualquer fase do ciclo de desenvolvimento da planta. A coloração do pecíolo deve ser observada em uma amostra representativa de 10 folhas completamente desenvolvidas por planta, classificando-os em:

- 1- Verde
- 2- Roxo-esverdeado
- 3- Roxo
- 4- Vermelho

1.28. Cerosidade (CER)

Observada a partir do final do 1º (primeiro) ciclo reprodutivo da planta. A planta deve ser classificada quanto à presença ou ausência de cerosidade no caule e demais ramificações.

- 0- Ausência
- 1- Presença

1.29. Pilosidade da folha (PFO)

Observada em qualquer fase do ciclo de desenvolvimento da planta. A pilosidade da folha deve ser observada quanto a presença ou a ausência de pêlos na porção adaxial e abaxial em uma amostra representativa de 10 folhas completamente desenvolvidas por planta, classificando-as em:

- 0- Ausência
- 1- Presença

2. Descritores relacionados a aspectos reprodutivos:

2.1. Altura da primeira inflorescência (API)

Avaliada assim que se inicia o ciclo reprodutivo da planta. A altura da primeira inflorescência (em cm) deve ser dada pela medida compreendida entre a base do caule (solo) e a inflorescência mais baixa (primeira inflorescência) da planta.

2.2. Comprimento do pedúnculo da inflorescência (CPI)

Avaliado a partir da 1º (primeira) frutificação. O comprimento do pedúnculo (em cm) deve ser obtido pela média do comprimento do pedúnculo em 10 inflorescências por planta. A média do comprimento do pedúnculo deve ser dada pela divisão do valor da soma do comprimento de todos os pedúnculos mensurados pelo número total de inflorescências avaliadas.

2.3. Número total de inflorescências (NTI)

Avaliado a partir do 1º (primeiro) ciclo reprodutivo. O número total de inflorescências por planta deve ser obtido pela quantificação de todas as inflorescências na planta. Para tal, deve-se realizar a contagem de todas as inflorescências da planta ao longo de seu ciclo reprodutivo.

2.4. Uniformidade de maturação do cacho por planta (UMC)

Avaliado a partir do 1º (primeiro) ciclo reprodutivo da planta. A uniformidade de maturação do cacho deve ser obtida pela média do número total de frutos amarelos (maduros) em relação ao número total de frutos em 5 cachos por planta. A média da uniformidade de maturação do cacho (em %) deve ser dada pela divisão do valor da soma de todos os frutos amarelos (maduros) do cacho pelo número total de frutos verdes, amarelos (maduros) e pretos (secos) presentes no cacho. A uniformidade de maturação do cacho por planta deve ser dada pela

divisão do valor da soma da média de uniformidade de maturação de todos os cachos quantificados pelo número total de cachos avaliados.

2.5. Tempo de formação do fruto (TFF)

Avaliado a partir do 1º (primeiro) ciclo reprodutivo da planta. O tempo de formação do fruto deve ser obtido pela quantificação dos dias decorridos entre a antese floral e a formação do primeiro fruto amarelo (maduro) da planta.

2.6. Número de frutos por planta (NFR)

Avaliado a partir do 1º (primeiro) ciclo reprodutivo da planta. O número de frutos por planta deve ser obtido pela quantificação do número total de frutos produzidos pela planta.

2.7. Peso de frutos por planta (PFR)

Avaliado a partir do 1º (primeiro) ciclo reprodutivo da planta. O peso de frutos por planta deve ser obtido pela média do peso de todos os frutos em relação ao número total de frutos produzidos. A média do peso de frutos por planta (em g) deve ser dada pela divisão do valor da soma do peso de todos os frutos, pesados na fase amarela de maturação, pelo número total de frutos produzidos.

2.8. Comprimento do fruto (CFR)

Avaliado a partir do 1º (primeiro) ciclo reprodutivo da planta. O comprimento do fruto deve ser obtido pela média do comprimento de 50 frutos amarelos (maduros) de diferentes ramos da planta. A média do comprimento de frutos (em cm) deve ser dada pela divisão do valor da soma do comprimento de todos os frutos mensurados pelo número total de frutos avaliados.

2.9. Largura do fruto (LFR)

Avaliada a partir do 1º (primeiro) ciclo reprodutivo da planta. A largura média do fruto deve ser obtida pela média da largura de 50 frutos amarelos (maduros) de diferentes ramos da planta. A média da largura de frutos (em cm) deve ser dada pela divisão do valor da soma da largura de todos os frutos mensurados pelo número total de frutos avaliados.

2.10. Relação entre o comprimento e a largura do fruto (CLFr)

Avaliada a partir do 1º (primeiro) ciclo reprodutivo da planta. Relação obtida através da razão entre o valor do *comprimento do fruto* pelo valor da *largura do fruto* (ambos citados anteriormente).

2.11. Número de sementes por fruto (NSF)

Avaliado a partir do 1º (primeiro) ciclo reprodutivo da planta. O número de sementes por fruto deve ser obtido pela razão entre o valor do *número de sementes por planta* (citado posteriormente) pelo *número de frutos por planta* (citado anteriormente).

2.12. Número de sementes por planta (NSE)

Avaliado a partir do 1º (primeiro) ciclo reprodutivo da planta. O número de sementes por planta deve ser obtido pela quantificação do número total de sementes produzidas pela planta.

2.13. Peso médio da semente seca (PMSs)

Avaliado a partir do 1º (primeiro) ciclo reprodutivo da planta. O peso médio da semente seca deve ser obtido pela média do peso de todas as sementes em relação ao número total de sementes produzidas. A média do peso de sementes secas por planta (em g) deve ser dada pela divisão do valor da soma do peso de todas as sementes, pesadas com baixo teor de umidade

(após secagem durante 3 dias a uma temperatura de 35°C em estufa), pelo número total de sementes produzidas.

2.14. Comprimento da semente (CMS)

Avaliado a partir do 1º (primeiro) ciclo reprodutivo da planta. O comprimento da semente deve ser obtido pela média do comprimento de 50 sementes por planta. A média do comprimento de sementes (em cm) deve ser dada pela divisão do valor da soma do comprimento de todas as sementes mensuradas pelo número total de sementes avaliadas.

2.15. Largura da semente (LSE)

Avaliada a partir do 1º (primeiro) ciclo reprodutivo da planta. A largura da semente deve ser obtida pela média da largura de 50 sementes por planta. A média da largura de sementes (em cm) deve ser dada pela divisão do valor da soma da largura de todas as sementes mensuradas pelo número total de sementes avaliadas.

2.16. Relação entre o comprimento de sementes e largura de sementes (CLSe)

Avaliada a partir do 1º (primeiro) ciclo reprodutivo da planta. Relação obtida através da razão entre o valor do *comprimento da semente* pelo valor da *largura da semente* (ambos citados anteriormente).

2.17. Produção de grãos por planta (PGP)

Avaliada a partir do 1º (primeiro) ciclo reprodutivo da planta. A produção de grãos por planta (em g.planta⁻¹) deve ser obtida pela multiplicação entre o valor do *número de sementes por planta* pelo valor do *peso médio da semente seca* (ambos citados anteriormente).

2.18. Formato do fruto (FFR)

Observado a partir do 1º (primeiro) ciclo reprodutivo da planta. O formato do fruto deve ser observado em 50 frutos amarelos (maduros), classificando-os em:

- 1- Elíptico
- 2- Oval
- 3- Triangular

2.19. Número de lóculos por fruto (NLFr)

Avaliado a partir do 1º (primeiro) ciclo reprodutivo da planta. O número de lóculos por fruto deve ser observado em 50 frutos amarelos (maduros), classificando-os em:

- 1- Unilocular
- 2- Bilocular
- 3- Trilocular
- 4- Tetralocular

2.20. Coloração da semente (CSE)

Observada em qualquer fase do ciclo de desenvolvimento da planta. A coloração da semente deve ser observada em uma amostra representativa de 50 sementes por planta, classificando-as em:

- 1- Marrom
- 2- Preta
- 3- Roxa
- 4- Vermelha

2.21. Textura da semente (TSE)

Observada em qualquer fase do ciclo de desenvolvimento da planta. A textura da semente deve ser observada em uma amostra representativa de 50 sementes por planta, classificando-as em:

- 1- Lisa
- 2- Áspera
- 3- Rugosa

2.22. Padrão da semente (PSE)

Observada em qualquer fase do ciclo de desenvolvimento da planta. O padrão da semente deve ser observado em uma amostra representativa de 50 sementes por planta, classificando-as em:

- 1- Cor única
- 2- Pintada
- 3- Pontuada
- 4- Rajada

2.23. Carúncula da semente (CAR)

Observada em qualquer fase do ciclo de desenvolvimento da planta. A carúncula da semente deve ser observada em uma amostra representativa de 50 sementes por planta, classificando-as quanto sua presença ou ausência em:

- 0- Ausência
- 1- Presença

Anexo B. Descrição de classes observadas em descritores relativos a variáveis multicategóricas e binárias em 60 genótipos da Coleção de germoplasma de *Jatropha curcas* da UFRRJ.

Acessos	Descritores vegetativos										Descritores reprodutivos					
	Multicategóricas					Binárias					Multicategóricas				Binárias	
	ARQ	RAM	CCA	CRA	CFJ	CFD	CNF	CPE	CER	PFO	FFR	NLFr	CSE	TSE	PSE	CAR
860	3	3	3	3	4	3	1	1	1	0	1	3	2	1	1	1
002	3	3	3	3	2	3	1	2	1	0	1	3	2	1	1	1
869	3	3	3	3	4	3	1	1	1	0	1	3	2	1	1	1
874	3	3	3	3	2	3	1	1	1	0	1	3	2	1	1	1
872	3	3	3	3	4	2	1	1	1	0	1	3	2	1	1	1
346	3	3	3	3	2	2	1	2	1	0	1	3	2	1	1	1
858	3	3	3	3	2	3	1	1	1	0	1	3	2	1	1	1
875	3	3	3	3	4	2	1	1	1	0	1	3	2	1	1	1
355	3	3	3	3	2	3	1	1	1	0	1	3	2	1	1	1
383	3	3	3	3	2	3	1	1	1	0	1	3	2	1	1	1
357	3	3	3	3	2	3	1	1	1	0	1	3	2	1	1	1
399	3	2	3	3	2	3	1	2	1	0	1	3	2	1	1	1
345	3	3	3	3	2	3	1	2	1	0	1	3	2	1	1	1
328	3	3	3	3	2	3	1	2	1	0	1	3	2	1	1	1
878	3	3	3	3	1	3	1	1	1	0	1	3	2	1	1	1
003	3	3	3	3	2	3	1	2	1	0	1	3	2	1	1	1
376	2	3	3	3	2	3	1	1	1	0	1	3	2	1	1	1
371	3	3	3	3	2	3	1	1	1	0	1	3	2	1	1	1
912	2	3	3	3	4	3	1	1	1	0	1	3	2	1	1	1
825	3	3	3	3	4	3	1	1	1	0	1	3	2	1	1	1
356	3	3	3	3	2	2	1	1	1	0	1	3	2	1	1	1
053	3	3	3	3	2	3	1	1	1	0	1	3	2	1	1	1
062	3	3	3	3	4	1	1	1	1	0	2	3	2	1	1	1
807	3	3	3	3	2	3	1	2	1	0	1	3	2	1	1	1
344	3	3	3	3	2	3	1	2	1	0	1	3	2	1	1	1
900	3	1	3	3	4	2	1	1	1	0	1	3	2	1	1	1
991	2	3	3	3	4	2	1	1	1	0	1	3	2	1	1	1
911	2	3	3	3	4	3	1	1	1	0	1	3	2	1	1	1
337	3	3	3	3	4	3	1	2	1	0	1	3	2	1	1	1
828	3	3	3	3	1	3	1	1	1	0	1	3	2	1	1	1
393	3	3	3	3	2	3	1	1	1	0	1	3	2	1	1	1
332	3	3	3	3	2	2	1	1	1	0	1	3	2	1	1	1
863	3	3	3	3	2	2	1	1	1	0	1	3	2	1	1	1
908	2	3	3	3	4	3	1	1	1	0	1	3	2	1	1	1
890	3	3	3	3	4	3	1	1	1	0	1	3	2	1	1	1
931	3	3	3	3	4	3	1	1	1	0	1	3	2	1	1	1
864	3	3	3	3	4	3	1	1	1	0	1	3	2	1	1	1
861	3	3	3	3	2	3	1	1	1	0	1	3	2	1	1	1
932	2	3	3	3	4	3	1	1	1	0	1	3	2	1	1	1
385	3	3	3	3	4	3	1	1	1	0	1	3	2	1	1	1
955	3	3	3	3	4	3	1	1	1	0	1	3	2	1	1	1
341	3	3	3	3	2	3	1	2	1	0	1	3	2	1	1	1
386	3	3	3	3	4	3	1	1	1	0	1	3	2	1	1	1
321	2	3	3	3	2	3	1	1	1	0	1	3	2	1	1	1
876	3	3	3	3	2	3	1	1	1	0	1	3	2	1	1	1
058	3	3	3	3	2	3	1	1	1	0	1	3	2	1	1	1
005	2	3	3	3	2	3	1	1	1	0	1	3	2	1	1	1
343	3	3	3	3	4	3	1	2	1	0	1	3	2	1	1	1
330	2	3	3	3	2	3	1	2	1	0	1	3	2	1	1	1
331	3	3	3	3	4	3	1	2	1	0	1	3	2	1	1	1
059	3	3	3	3	2	3	1	1	1	0	1	3	2	1	1	1
905	3	3	3	3	4	3	1	1	1	0	1	3	2	1	1	1
910	2	3	3	3	4	3	1	2	1	0	1	3	2	1	1	1
831	3	3	3	3	4	3	1	1	1	0	1	3	2	1	1	1
849	2	3	3	3	4	3	1	1	1	0	1	3	2	1	1	1
024	3	3	3	3	2	3	1	1	1	0	1	3	2	1	1	1
856	3	3	3	3	1	3	1	1	1	0	1	3	2	1	1	1
824	3	3	3	3	4	3	1	1	1	0	1	3	2	1	1	1
820	3	3	3	3	4	3	1	1	1	0	1	3	2	1	1	1
857	3	3	3	3	1	3	1	1	1	0	1	3	2	1	1	1

Legenda: ARQ - Arquitetura da planta (1- ereta; 2- fechada; 3- aberta); RAM - Tipo de ramificação (0- ausência de ramificação; 1- monopodial; 2- dicasial; 3- simpodial); CCA - Coloração do caule (1- verde; 2- verde escuro; 3- cinza; 4- roxo; 5- azul); CRA - Coloração do ramo (1- verde; 2- verde escuro; 3- cinza; 4- roxo; 5- azul); CFJ - Coloração de folhas jovens (1- verde claro; 2- verde; 3- verde escuro; 4- roxa; 5- vermelha); CFD - Coloração de folhas desenvolvidas (1- verde claro; 2- verde; 3- verde escuro; 4- roxa; 5- vermelha); CNF - Coloração da nervura foliar (1- verde; 2- roxa; 3- vermelha); CPE - Coloração do pecíolo (1- verde; 2- roxo-esverdeado; 3- roxo; 4- vermelho); CER - Cerosidade (0- ausência; 1- presença); PFO - Pilosidade da folha (0- ausência; 1- presença); FFR - Formato do fruto (1- elíptico; 2- oval; 3- triangular); NLFr - Número de lóculos por fruto (1- unilocular; 2- bilocular; 3- trilocular; 4- tetralocular); CSE - Coloração da semente (1- marrom; 2- preta; 3- roxa; 4- vermelha); TSE - Textura da semente (1- lisa; 2- áspera; 3- rugosa); PSE - Padrão da semente (1- cor única; 2- pintada; 3- pontuada; 4- rajada); e CAR - Carúncula da semente (0- ausência; 1- presença).

Anexo C. Estimativas da variância (autovalores), respectivos coeficientes de ponderação (autovetores) associados aos componentes principais (CP) e contribuição acumulada a partir de 36 variáveis quantitativas avaliadas em 60 genótipos de *Jatropha curcas* da Coleção de germoplasma da UFRRJ.

	CP1	CP2	CP3	CP4	CP5	CP6	CP7	CP8	CP9	CP10	CP11	CP12	CP13	CP14	CP15	CP16	CP17	CP18
APD	0,20	-0,02	0,05	-0,20	0,20	-0,04	0,16	-0,14	0,26	-0,25	0,15	0,03	0,21	0,30	0,13	0,05	-0,08	-0,15
PCO	0,28	-0,12	0,01	0,27	0,03	0,09	-0,03	-0,13	-0,01	0,10	0,14	-0,10	-0,01	-0,06	-0,00	0,02	-0,03	-0,06
PCOn	0,25	-0,06	0,03	0,16	0,06	0,16	-0,07	-0,08	0,25	0,04	0,22	-0,06	-0,02	-0,04	0,18	0,11	0,20	-0,06
PCOs	0,23	-0,16	-0,02	0,22	0,01	0,02	-0,06	-0,15	-0,01	-0,03	0,24	-0,12	-0,02	0,15	-0,24	-0,14	-0,01	-0,08
PCOI	0,23	-0,13	-0,06	0,19	0,00	0,10	-0,09	-0,16	-0,19	-0,02	0,09	0,01	-0,24	-0,08	-0,08	-0,26	-0,17	0,12
PCOo	0,23	-0,05	0,06	0,27	0,05	0,02	0,12	0,00	0,08	0,25	-0,02	-0,18	0,25	-0,11	0,08	0,26	-0,10	-0,19
DCA	0,10	-0,05	-0,02	-0,30	0,21	-0,20	-0,07	-0,32	0,23	0,05	-0,25	0,22	-0,09	-0,04	-0,20	0,15	0,24	-0,00
NRT	0,27	-0,10	0,11	-0,10	0,12	-0,02	-0,02	0,04	0,03	0,14	-0,09	0,11	0,11	-0,13	0,10	0,01	-0,08	-0,06
NRI	0,28	-0,03	0,08	-0,14	0,09	-0,04	0,05	0,03	-0,01	-0,04	-0,22	0,09	0,25	-0,06	-0,03	0,01	-0,12	0,21
CRP	0,14	0,08	0,03	-0,38	0,04	0,08	0,03	0,04	0,02	-0,09	0,25	-0,12	-0,39	0,22	0,06	-0,08	-0,18	-0,46
CERp	-0,05	-0,06	0,03	-0,22	0,00	0,21	0,35	-0,43	-0,20	-0,15	0,06	-0,09	0,24	0,03	0,01	-0,35	0,37	0,09
CRI	-0,09	0,08	-0,04	0,18	0,11	0,06	0,06	-0,54	-0,15	-0,16	-0,06	0,28	-0,11	-0,36	0,16	0,03	-0,35	-0,14
CERi	0,07	0,16	0,16	-0,33	-0,13	-0,06	0,13	-0,10	-0,21	0,27	0,05	-0,12	-0,13	-0,08	-0,21	0,34	-0,24	-0,08
NLF	0,05	0,03	-0,08	-0,15	-0,27	0,25	0,38	0,15	-0,10	0,34	0,06	0,19	0,11	-0,15	-0,19	-0,26	-0,16	-0,07
AIFc	0,02	0,09	0,15	-0,16	0,12	0,18	-0,17	-0,07	0,04	0,38	-0,29	-0,36	-0,10	-0,00	0,52	-0,33	-0,01	0,08
CMP	0,06	0,01	-0,38	-0,12	0,02	0,07	-0,19	-0,10	0,20	0,16	-0,09	-0,13	0,21	-0,19	-0,21	-0,00	0,06	-0,16
CFO	0,07	-0,02	-0,48	-0,06	-0,10	0,01	-0,10	0,09	-0,06	-0,14	-0,14	0,00	0,12	0,00	0,13	-0,10	-0,07	-0,27
LFO	0,10	0,04	-0,45	-0,04	-0,18	-0,20	-0,07	-0,06	0,02	0,05	-0,06	0,13	0,12	0,18	0,18	-0,11	-0,10	-0,08
CLFo	-0,09	-0,12	0,03	-0,03	0,15	0,39	-0,05	0,26	-0,10	-0,38	-0,20	-0,25	0,02	0,29	-0,15	0,07	0,11	-0,32
API	-0,02	0,02	-0,11	-0,13	0,16	0,33	0,28	0,12	0,15	-0,24	0,21	-0,32	0,06	-0,49	0,18	-0,05	-0,22	0,20
CPI	0,01	-0,36	0,00	-0,06	-0,11	-0,16	-0,03	-0,07	-0,15	-0,14	-0,09	-0,30	0,05	0,30	-0,03	0,05	-0,32	0,18
NTI	0,24	0,01	0,23	0,03	0,15	0,12	0,10	0,06	0,12	-0,07	-0,25	0,03	0,15	0,10	-0,29	-0,12	-0,14	-0,03
UMC	0,02	0,00	0,26	0,05	0,22	-0,31	-0,07	0,29	-0,12	0,02	0,16	0,29	0,22	-0,03	0,05	-0,40	-0,03	-0,20
TFF	-0,22	-0,22	0,07	-0,04	-0,03	-0,19	0,12	-0,12	0,09	0,18	0,24	-0,07	0,00	-0,08	0,03	0,03	0,22	-0,22
NFR	0,30	0,00	0,06	-0,02	-0,20	-0,14	-0,00	0,04	-0,09	-0,11	-0,12	0,00	-0,15	-0,14	0,03	-0,02	0,18	-0,01
PFR	-0,03	0,19	-0,20	-0,01	0,39	-0,01	-0,11	0,01	0,18	0,07	0,08	-0,14	-0,22	0,01	-0,38	-0,26	-0,12	0,23
CFR	-0,03	-0,40	-0,13	0,04	0,06	-0,02	0,26	0,06	0,02	0,08	-0,22	-0,05	-0,11	0,08	-0,02	0,00	-0,01	0,01
LFR	0,03	-0,06	-0,20	0,20	0,15	0,03	0,55	0,15	0,15	0,05	-0,22	0,07	-0,30	0,16	0,09	0,03	0,05	0,01
CLFr	-0,07	-0,41	0,03	-0,14	-0,06	-0,05	-0,20	-0,07	-0,11	0,05	-0,04	-0,13	0,14	-0,06	-0,10	-0,03	-0,07	-0,00
NSF	0,17	0,07	-0,09	-0,19	0,04	0,39	-0,07	0,12	-0,09	-0,16	0,18	0,20	0,06	0,08	0,20	0,25	-0,05	0,35
NSE	0,29	0,03	0,03	-0,03	-0,22	-0,16	-0,00	0,05	-0,03	-0,12	-0,08	-0,02	-0,20	-0,13	-0,00	-0,08	0,20	0,04
PMSs	0,16	-0,09	-0,27	-0,01	0,20	0,02	0,04	0,12	-0,25	0,21	0,33	-0,05	-0,08	-0,06	-0,03	0,13	0,21	0,17
CMS	-0,00	-0,25	-0,08	-0,10	0,40	-0,05	-0,08	0,07	-0,36	0,10	-0,05	0,19	-0,15	-0,02	0,10	0,08	0,07	-0,08
LSE	0,06	0,27	-0,06	0,11	0,28	-0,17	0,06	-0,03	-0,46	0,01	-0,12	-0,15	0,07	0,15	0,00	0,10	0,11	-0,11
CLSE	-0,05	-0,39	0,00	-0,16	0,01	0,12	-0,11	0,08	0,17	0,05	0,07	0,25	-0,17	-0,15	0,07	-0,04	-0,08	0,04
PGP	0,30	0,02	0,01	-0,04	-0,20	-0,16	0,00	0,05	-0,06	-0,10	-0,05	-0,03	-0,19	-0,14	-0,00	-0,07	0,20	0,04
λ_j	2,90	2,08	1,78	1,59	1,48	1,36	1,22	1,16	1,12	1,10	1,01	0,98	0,92	0,88	0,81	0,77	0,67	0,65
$\lambda_j(\%)$	23,40	35,47	44,26	51,25	57,35	62,53	66,70	70,45	73,92	77,29	80,11	82,75	85,13	87,27	89,11	90,77	92,01	93,18
	CP19	CP20	CP21	CP22	CP23	CP24	CP25	CP26	CP27	CP28	CP29	CP30	CP31	CP32	CP33	CP34	CP35	CP36
APD	-0,11	0,04	0,32	0,13	-0,12	-0,42	-0,14	0,02	0,31	-0,05	-0,24	0,02	0,03	0,02	0,00	0,00	0,00	0,00
PCO	0,07	0,01	0,04	0,11	0,08	-0,01	-0,02	0,03	-0,00	0,05	-0,02	0,02	-0,42	-0,72	-0,17	-0,01	0,05	-0,01
PCOn	-0,06	-0,11	-0,30	0,08	0,16	0,35	0,15	0,26	0,45	0,19	-0,10	-0,05	0,15	0,19	0,01	0,01	-0,01	0,01
PCOs	-0,04	-0,25	0,26	-0,40	0,18	-0,21	0,41	-0,04	-0,22	0,11	0,14	0,03	0,06	0,19	0,04	-0,00	-0,01	-0,00

PCOI	0,04	0,09	-0,02	0,54	0,16	-0,06	-0,13	-0,39	-0,02	-0,11	-0,13	-0,09	0,13	0,24	0,05	0,00	-0,02	0,00
PCOo	0,22	0,22	0,16	0,14	-0,19	0,05	-0,22	0,24	-0,36	0,12	0,18	0,14	0,15	0,24	0,06	-0,00	-0,00	-0,00
DCA	-0,03	-0,19	0,22	0,33	-0,05	0,25	0,20	-0,09	-0,13	0,08	0,20	0,16	-0,03	-0,04	0,01	-0,02	0,00	-0,00
NRT	-0,50	-0,00	-0,13	-0,20	0,36	0,05	-0,15	-0,03	-0,07	-0,54	0,12	0,14	0,00	-0,00	0,01	-0,02	-0,00	-0,00
NRI	-0,11	0,26	-0,27	-0,14	0,13	-0,09	0,03	-0,18	-0,17	0,61	-0,23	0,09	-0,00	0,00	-0,01	-0,00	0,02	-0,00
CRP	0,02	0,20	-0,29	-0,03	-0,12	0,11	-0,01	-0,12	-0,13	0,10	0,28	0,08	-0,03	-0,03	0,01	-0,00	0,00	0,00
CERp	0,23	0,10	-0,07	-0,06	0,19	0,12	-0,15	0,15	-0,10	-0,09	0,05	0,08	0,01	0,00	0,01	-0,01	0,01	-0,00
CRI	-0,10	-0,14	-0,06	-0,24	-0,30	0,03	-0,07	0,04	0,06	0,07	-0,02	0,02	0,03	0,02	-0,02	0,01	-0,00	0,00
CERi	0,26	-0,26	0,14	-0,08	0,36	0,03	-0,16	0,04	0,12	0,05	-0,22	-0,11	0,05	0,02	-0,02	0,01	0,00	0,00
NLF	-0,28	-0,00	0,08	0,23	-0,18	-0,04	0,30	0,20	0,13	0,06	0,02	0,10	-0,03	0,00	-0,01	-0,00	-0,00	0,00
AIFc	-0,00	-0,17	0,18	-0,05	-0,11	-0,05	0,09	-0,08	-0,01	0,07	-0,07	-0,03	0,00	0,01	-0,01	0,01	-0,01	-0,00
CMP	0,26	-0,17	-0,45	-0,00	-0,10	-0,39	0,05	-0,07	0,03	-0,19	-0,06	0,06	0,00	0,01	0,01	-0,00	0,01	0,00
CFO	-0,02	-0,04	0,20	-0,01	0,16	0,25	-0,03	0,04	-0,07	0,07	-0,16	-0,12	0,08	-0,17	0,56	0,00	0,17	0,00
LFO	0,04	0,06	0,09	-0,04	0,12	0,17	-0,03	0,04	-0,03	0,02	0,02	-0,15	-0,09	0,17	-0,64	0,02	-0,20	-0,00
CLFo	-0,12	-0,14	0,16	0,07	0,05	0,05	-0,05	-0,04	-0,04	0,07	-0,14	0,12	-0,02	0,09	-0,34	0,01	-0,09	0,00
API	0,09	0,01	0,04	0,08	0,10	0,03	0,26	-0,04	0,02	-0,05	0,17	-0,11	-0,03	-0,01	-0,00	0,00	-0,01	0,00
CPI	-0,03	-0,20	-0,21	0,10	-0,23	0,21	0,15	0,19	-0,07	-0,17	-0,26	0,32	-0,00	-0,00	-0,02	0,01	0,00	0,00
NTI	0,19	0,07	-0,03	-0,15	-0,24	0,28	0,15	-0,09	0,05	-0,24	-0,05	-0,56	0,02	-0,04	-0,01	0,02	-0,03	-0,00
UMC	0,24	-0,43	-0,09	0,08	-0,01	0,06	-0,17	0,03	-0,03	0,10	-0,01	0,13	0,01	-0,00	-0,02	0,01	0,01	0,00
TFF	-0,30	-0,07	-0,13	0,08	-0,10	-0,01	-0,02	-0,03	-0,39	0,03	-0,37	-0,42	-0,07	0,03	-0,01	-0,01	-0,02	-0,00
NFR	0,05	-0,00	0,02	-0,07	-0,13	-0,06	-0,06	0,09	0,07	-0,04	-0,06	0,00	-0,70	0,39	0,19	-0,04	0,00	-0,00
PFR	-0,18	0,06	0,05	-0,06	-0,01	0,07	-0,32	0,42	-0,07	0,08	-0,11	-0,00	-0,08	0,05	0,01	-0,00	0,01	0,00
CFR	-0,02	-0,17	-0,08	-0,04	0,01	-0,11	-0,23	-0,02	0,16	0,15	0,27	-0,14	-0,00	0,00	0,02	0,00	-0,01	0,65
LFR	0,09	-0,21	-0,11	-0,06	0,08	-0,05	-0,05	-0,09	-0,05	-0,00	-0,08	0,08	0,00	0,00	-0,00	-0,01	0,01	-0,51
CLFr	-0,10	-0,01	-0,00	0,01	-0,06	-0,08	-0,22	0,06	0,24	0,18	0,38	-0,23	0,00	0,01	0,03	-0,02	0,01	-0,56
NSF	-0,01	-0,37	-0,08	0,13	-0,06	-0,10	-0,05	0,15	-0,33	0,00	0,17	-0,26	-0,07	0,01	0,01	-0,01	-0,01	-0,00
NSE	-0,04	-0,02	0,01	-0,04	0,14	-0,11	-0,07	0,15	-0,05	-0,03	-0,01	-0,07	0,33	-0,08	-0,26	0,21	0,64	0,01
PMSs	-0,02	0,00	-0,14	-0,25	-0,37	0,24	-0,13	-0,40	0,15	-0,03	-0,09	0,13	0,03	0,01	-0,03	0,03	0,04	-0,00
CMS	0,15	0,27	-0,01	0,01	0,07	-0,18	0,29	0,29	0,02	-0,09	-0,09	-0,07	0,01	-0,00	0,02	0,43	-0,11	-0,01
LSE	-0,16	0,01	-0,10	0,14	0,00	-0,08	0,15	0,08	0,08	0,01	0,08	-0,13	0,04	0,01	-0,05	-0,57	0,14	0,01
CLSe	0,27	0,16	0,09	-0,13	0,01	-0,12	0,06	0,11	-0,04	-0,07	-0,16	0,04	0,05	0,02	-0,08	-0,64	0,17	0,01
PGP	-0,02	-0,03	0,02	-0,06	-0,15	-0,10	-0,09	0,14	-0,03	0,00	-0,05	-0,00	0,34	-0,24	0,08	-0,17	-0,67	-0,01
λ_j	0,60	0,57	0,56	0,55	0,48	0,45	0,45	0,39	0,36	0,30	0,28	0,18	0,08	0,07	0,06	0,03	0,02	0,01
$\lambda_j(\%)$	94,20	95,11	96,00	96,83	97,47	98,04	98,61	99,04	99,40	99,65	99,86	99,95	99,97	99,98	99,99	100,0	100,0	100,0

Legenda: Consultar Tabela 3 e/ou 7; λ_j - Variância (autovalor); e $\lambda_j (\%)$ - Variância acumulada. Descritores e valores em negrito referentes ao critério de descarte de variáveis redundantes de Jolliffe (1972 e 1973).

Anexo D. Estimativas dos efeitos diretos e indiretos de 15 variáveis explicativas sobre a característica Número de sementes por planta (NSE) em 60 genótipos de *Jatropha curcas* da UFRRJ.

Variável	Vias de associação	Coeficientes de trilha		Coeficiente de correlação
		Efeito direto	Efeito indireto	
PCOo	Efeito direto sobre NSE	0,23726346		
	Efeito indireto via DCA		0,00074681	
	Efeito indireto via CRI		0,01781248	
	Efeito indireto via CERi		0,00245857	
	Efeito indireto via NLF		0,00484241	
	Efeito indireto via AIFc		-0,0071629	
	Efeito indireto via CFO		-0,00049175	
	Efeito indireto via CLFo		0,02938467	
	Efeito indireto via API		-0,00505661	
	Efeito indireto via CPI		0,00254111	
	Efeito indireto via TFF		0,15183531	
	Efeito indireto via LFR		-0,01361095	
	Efeito indireto via NSF		-0,00151282	
	Efeito indireto via CMS		0,00370662	
	Efeito indireto via LSE		-0,00695542	
	Efeito total			0,415801
DCA	Efeito direto sobre NSE	0,19534654		
	Efeito indireto via PCOo		0,00090706	
	Efeito indireto via CRI		0,00071699	
	Efeito indireto via CERi		0,02122419	
	Efeito indireto via NLF		-0,0093289	
	Efeito indireto via AIFc		-0,01197886	
	Efeito indireto via CFO		0,00136563	
	Efeito indireto via CLFo		0,03101271	
	Efeito indireto via API		0,00879937	
	Efeito indireto via CPI		0,01297298	
	Efeito indireto via TFF		0,01708342	
	Efeito indireto via LFR		0,00243574	
	Efeito indireto via NSF		-0,00125716	
	Efeito indireto via CMS		-0,03366124	
	Efeito indireto via LSE		0,00027753	
	Efeito total			0,235916
CRI	Efeito direto sobre NSE	-0,13540876		
	Efeito indireto via PCOo		-0,03121106	
	Efeito indireto via DCA		-0,00103436	
	Efeito indireto via CERi		-0,0113296	
	Efeito indireto via NLF		-0,0114082	
	Efeito indireto via AIFc		0,00879671	
	Efeito indireto via CFO		-0,00138005	
	Efeito indireto via CLFo		0,0007402	
	Efeito indireto via API		-0,00102575	
	Efeito indireto via CPI		-0,0453395	
	Efeito indireto via TFF		-0,03505174	
	Efeito indireto via LFR		-0,00325239	
	Efeito indireto via NSF		0,0018652	
	Efeito indireto via CMS		-0,00141702	
	Efeito indireto via LSE		-0,00922666	
	Efeito total			-0,275683
CERi	Efeito direto sobre NSE	0,09262947		
	Efeito indireto via PCOo		0,00629745	
	Efeito indireto via DCA		0,04475975	
	Efeito indireto via CRI		0,01656198	
	Efeito indireto via NLF		0,02108183	
	Efeito indireto via AIFc		-0,0251943	

	Efeito indireto via CFO		-0,00572931
	Efeito indireto via CLFo		0,04008452
	Efeito indireto via API		0,00224268
	Efeito indireto via CPI		-0,02224483
	Efeito indireto via TFF		0,06947584
	Efeito indireto via LFR		0,01572899
	Efeito indireto via NSF		-0,00219943
	Efeito indireto via CMS		0,01954904
	Efeito indireto via LSE		-0,00807568
	Efeito total		0,264968
NLF	Efeito direto sobre NSE	0,071299	
	Efeito indireto via PCOo		0,01611422
	Efeito indireto via DCA		-0,02555953
	Efeito indireto via CRI		0,02166608
	Efeito indireto via CERi		0,02738887
	Efeito indireto via AIFc		0,00045801
	Efeito indireto via CFO		0,00400598
	Efeito indireto via CLFo		0,01143297
	Efeito indireto via API		-0,00910605
	Efeito indireto via CPI		-0,03218155
	Efeito indireto via TFF		0,03732255
	Efeito indireto via LFR		-0,00973199
	Efeito indireto via NSF		-0,00352678
	Efeito indireto via CMS		0,02383556
	Efeito indireto via LSE		0,00727766
	Efeito total		0,140695
AIFc	Efeito direto sobre NSE	-0,10832855	
	Efeito indireto via PCOo		0,01568833
	Efeito indireto via DCA		0,02160123
	Efeito indireto via CRI		0,01099573
	Efeito indireto via CERi		0,02154311
	Efeito indireto via NLF		-0,00030145
	Efeito indireto via CFO		-0,00518684
	Efeito indireto via CLFo		-0,00342413
	Efeito indireto via API		-0,01238735
	Efeito indireto via CPI		-0,03200905
	Efeito indireto via TFF		0,06658639
	Efeito indireto via LFR		0,01108276
	Efeito indireto via NSF		-0,00113775
	Efeito indireto via CMS		-0,0005934
	Efeito indireto via LSE		-0,00442803
	Efeito total		-0,020299
CFO	Efeito direto sobre NSE	0,02589366	
	Efeito indireto via PCOo		-0,00450587
	Efeito indireto via DCA		0,01030258
	Efeito indireto via CRI		0,00721688
	Efeito indireto via CERi		-0,02049547
	Efeito indireto via NLF		0,0110306
	Efeito indireto via AIFc		0,02169962
	Efeito indireto via CLFo		-0,01204494
	Efeito indireto via API		0,01032794
	Efeito indireto via CPI		0,02145258
	Efeito indireto via TFF		0,14412695
	Efeito indireto via LFR		-0,01145092
	Efeito indireto via NSF		-0,00373781
	Efeito indireto via CMS		-0,01126049
	Efeito indireto via LSE		-0,00294229
	Efeito total		0,185613
CLFo	Efeito direto sobre NSE	-0,16459787	
	Efeito indireto via PCOo		-0,04235722

Efeito indireto via DCA		-0,03680622
Efeito indireto via CRI		0,00060893
Efeito indireto via CERi		-0,02255805
Efeito indireto via NLF		-0,00495243
Efeito indireto via AIFc		-0,00225356
Efeito indireto via CFO		0,00189485
Efeito indireto via API		0,00097257
Efeito indireto via CPI		0,01515749
Efeito indireto via TFF		-0,01340066
Efeito indireto via LFR		-0,00002722
Efeito indireto via NSF		-0,00189817
Efeito indireto via CMS		-0,02674583
Efeito indireto via LSE		0,00713439
Efeito total		-0,289829
<hr/>		
Efeito direto sobre NSE	0,09530346	
Efeito indireto via PCOo		-0,01258872
Efeito indireto via DCA		0,01803635
Efeito indireto via CRI		0,0014574
Efeito indireto via CERi		0,00217976
Efeito indireto via NLF		-0,00681248
Efeito indireto via AIFc		0,01408033
Efeito indireto via CFO		0,00280607
Efeito indireto via CLFo		-0,00167972
Efeito indireto via CPI		0,00757563
Efeito indireto via TFF		-0,07927383
Efeito indireto via LFR		-0,00991834
Efeito indireto via NSF		0,00202192
Efeito indireto via CMS		-0,00463217
Efeito indireto via LSE		-0,00451464
Efeito total		0,024041
<hr/>		
Efeito direto sobre NSE	0,21482008	
Efeito indireto via PCOo		0,00280659
Efeito indireto via DCA		0,01179698
Efeito indireto via CRI		0,0285791
Efeito indireto via CERi		-0,00959187
Efeito indireto via NLF		-0,01068109
Efeito indireto via AIFc		0,01614139
Efeito indireto via CFO		0,00258582
Efeito indireto via CLFo		-0,01161386
Efeito indireto via API		0,00336088
Efeito indireto via TFF		-0,15380682
Efeito indireto via LFR		0,00198978
Efeito indireto via NSF		0,00199014
Efeito indireto via CMS		-0,03523083
Efeito indireto via LSE		0,01459973
Efeito total		0,077746
<hr/>		
Efeito direto sobre NSE	-0,50051033	
Efeito indireto via PCOo		-0,07197648
Efeito indireto via DCA		-0,00666757
Efeito indireto via CRI		-0,00948295
Efeito indireto via CERi		-0,0128579
Efeito indireto via NLF		-0,0053167
Efeito indireto via AIFc		0,01441171
Efeito indireto via CFO		-0,00745634
Efeito indireto via CLFo		-0,00440694
Efeito indireto via API		0,01509473
Efeito indireto via CPI		0,06601421
Efeito indireto via LFR		0,00221699
Efeito indireto via NSF		0,0083772
Efeito indireto via CMS		-0,01966447

	Efeito indireto via LSE		0,01862583
	Efeito total		-0,513599
LFR	Efeito direto sobre NSE	-0,06314962	
	Efeito indireto via PCOo		0,05113858
	Efeito indireto via DCA		-0,00753471
	Efeito indireto via CRI		-0,00697396
	Efeito indireto via CERi		-0,02307168
	Efeito indireto via NLF		0,01098789
	Efeito indireto via AIFc		0,01901166
	Efeito indireto via CFO		0,0046953
	Efeito indireto via CLFo		-0,00007094
	Efeito indireto via API		0,01496846
	Efeito indireto via CPI		-0,00676877
	Efeito indireto via TFF		0,01757142
	Efeito indireto via NSF		0,00051598
	Efeito indireto via CMS		-0,01630039
	Efeito indireto via LSE		-0,00497621
	Efeito total		-0,009957
NSF	Efeito direto sobre NSE	-0,01467989	
	Efeito indireto via PCOo		0,02445095
	Efeito indireto via DCA		0,01672909
	Efeito indireto via CRI		0,01720477
	Efeito indireto via CERi		0,0138783
	Efeito indireto via NLF		0,0171293
	Efeito indireto via AIFc		-0,0083959
	Efeito indireto via CFO		0,00659307
	Efeito indireto via CLFo		-0,02128316
	Efeito indireto via API		-0,01312653
	Efeito indireto via CPI		-0,02912294
	Efeito indireto via TFF		0,28562022
	Efeito indireto via LFR		0,00221965
	Efeito indireto via CMS		-0,00465762
	Efeito indireto via LSE		-0,0020633
	Efeito total		0,290496
CMS	Efeito direto sobre NSE	-0,12984735	
	Efeito indireto via PCOo		-0,00677292
	Efeito indireto via DCA		0,05064105
	Efeito indireto via CRI		-0,00147772
	Efeito indireto via CERi		-0,01394574
	Efeito indireto via NLF		-0,01308807
	Efeito indireto via AIFc		-0,00049506
	Efeito indireto via CFO		0,00224552
	Efeito indireto via CLFo		-0,0339037
	Efeito indireto via API		0,00339986
	Efeito indireto via CPI		0,05828606
	Efeito indireto via TFF		-0,07579879
	Efeito indireto via LFR		-0,00792749
	Efeito indireto via NSF		-0,00052657
	Efeito indireto via LSE		-0,00700208
	Efeito total		-0,176213
LSE	Efeito direto sobre NSE	-0,05128003	
	Efeito indireto via PCOo		0,03218147
	Efeito indireto via DCA		-0,00105722
	Efeito indireto via CRI		-0,02436369
	Efeito indireto via CERi		0,01458747
	Efeito indireto via NLF		-0,01011875
	Efeito indireto via AIFc		-0,00935417
	Efeito indireto via CFO		0,0014857
	Efeito indireto via CLFo		0,02289984
	Efeito indireto via API		0,00839042

Efeito indireto via CPI	-0,06116056	
Efeito indireto via TFF	0,18179436	
Efeito indireto via LFR	-0,00612804	
Efeito indireto via NSF	-0,00059066	
Efeito indireto via CMS	-0,01773014	
<hr/>		
Efeito total		0,079556

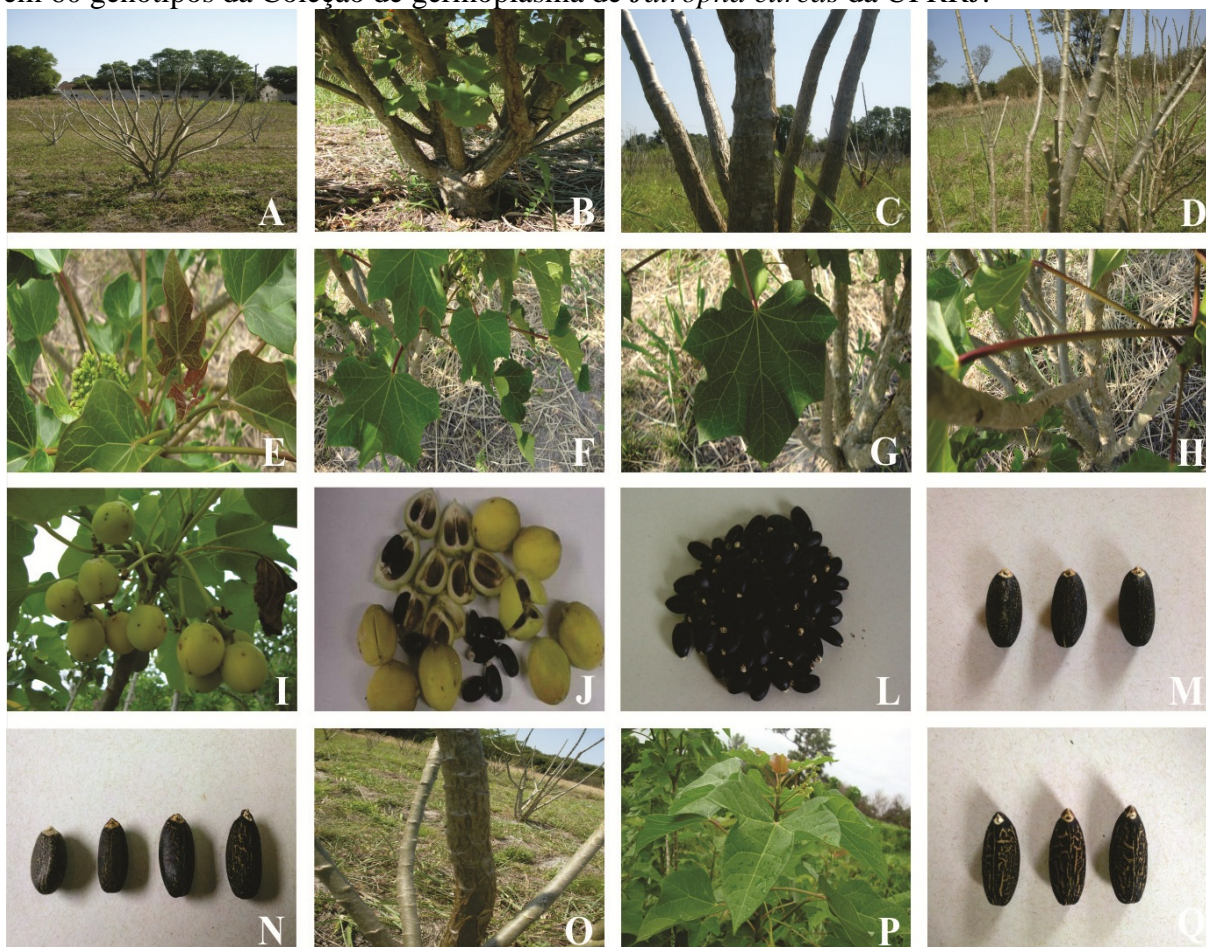
Legenda: Consultar Tabela 3 e/ou 7. Coeficiente de determinação = 0,56; e Efeito da variável residual = 0,66.

Anexo E. Fotos da Coleção de germoplasma de *Jatropha curcas* da UFRRJ, localizada no campo experimental do DFITO/IA, em Seropédica, RJ, aos três anos de idade.



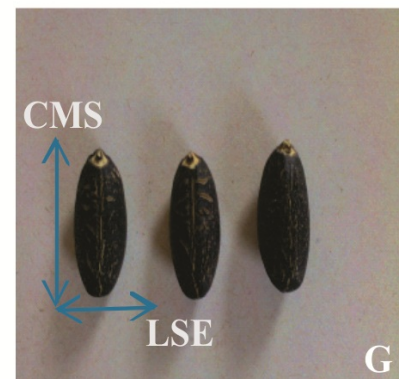
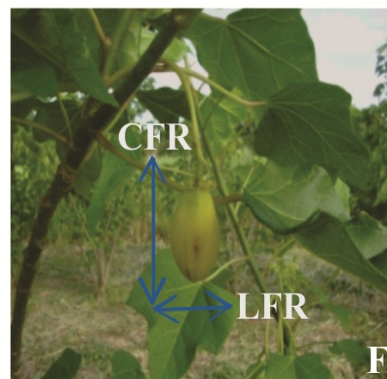
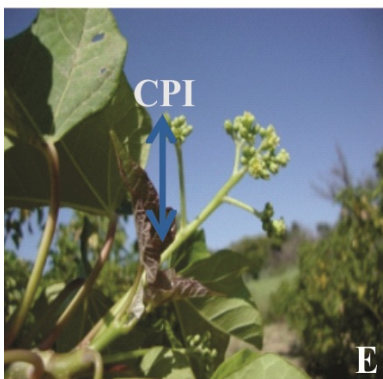
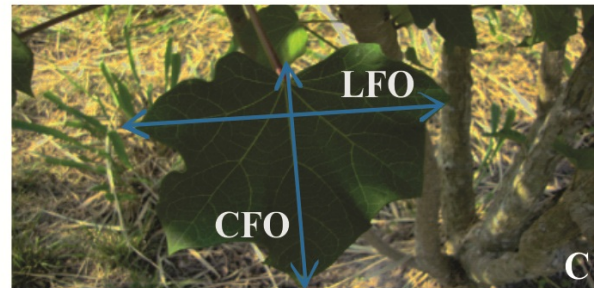
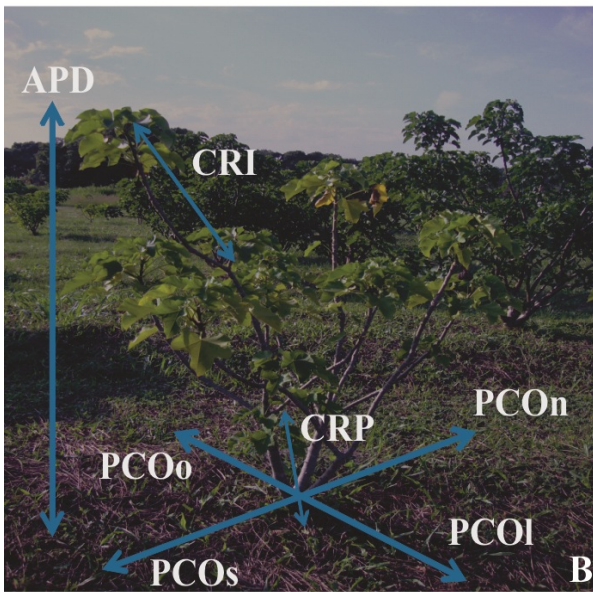
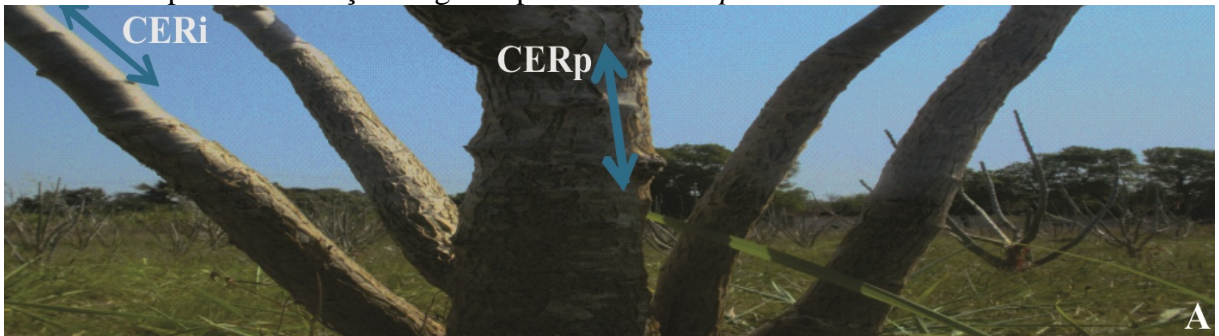
Legenda: A - Sentido Leste; B - Sentido Norte; C - Sentido Sul; e D - Sentido Oeste. Fonte: Reis (2011).

Anexo F. Fotos das 16 variáveis qualitativas (13 multicategóricas e 3 binárias) caracterizadas em 60 genótipos da Coleção de germoplasma de *Jatropha curcas* da UFRRJ.



Legenda: Fotos A, B, C, D, E, F, G e H relativas aos descritores vegetativos multicategóricos ARQ, RAM, CCA, CRA, CFJ, CFD, CNF e CPE; fotos I, J, L, M e N referentes aos descritores reprodutivos multicategóricos FFR, NLFr, CSE, TSE e PSE; fotos O e P relativas aos descritores vegetativos binários CER e PFO; e foto Q referente ao descritor reprodutivo binário CAR. Fonte: Reis (2012).

Anexo G. Fotos ilustrativas para auxiliar na mensuração de 18 variáveis quantitativas, elaboradas a partir da Coleção de germoplasma de *Jatropha curcas* da UFRRJ.















Legenda: Ilustrações A, B, C e D relativas à mensuração dos descritores vegetativos CERi, CERp (A), APD, CRI, CRP, PCOn, PCOI, PCOs, PCOo (B), CFO, LFO (C), CMP e AIFc (D); ilustrações E, F e G referentes aos descritores reprodutivos CPI (E), CFR, LFR (F), CMS e LSE (G). Fonte: Reis (2012).

Anexo H. Croqui da Coleção de germoplasma de *Jatropha curcas* da UFRRJ e respectiva diagramação dos acessos definidos para caracterização.

Bordadura	080	079	078	057	972	X*	956	955	954	Dreno 01	X*	900	991	910	911	912	930	931	932	Dreno 02	002	003	004	005	006	007	008	Dreno 03
	309	X*	X*	970	969	968	049	952	951		897	898	899	907	908	909	927	928	929		345	346	347	348	071	350	001	
	306	077	076	056	966	X*	950	949	948		894	895	045	904	905	906	043	925	926		339	340	341	342	343	344	X*	
	075	074	073	X*	X*	055	X*	946	048		891	892	893	901	902	903	041	042	923		332	333	334	335	336	337	338	
	990	989	072	X*	X*	081	047	046	942		888	889	890	998	999	035	039	040	920		068	069	070	328	329	330	331	
	987	986	985	959	958	054	941	940	939		885	044	X*	995	X*	997	X*	038	917		397	398	399	400	322	323	324	
	X*	983	051	978	X*	976	938	937	936		882	883	884	992	993	994	036	X*	914		390	391	392	393	394	395	X*	
	050	980	979	975	X*	974	935	934	933		825	826	021	852	853	854	X*	033	034									
											822	823	824	849	850	851	876	877	878									
											819	820	821	X*	031	848	873	874	875		383	384	385	386	387	388	067	
											816	019	020	X*	029	030	870	871	872		376	377	378	379	380	065	066	
											813	017	018	028	X*	842	867	868	869		369	370	371	372	373	063	064	
											810	811	812	025	026	027	864	865	032		362	363	053	X*	X*	X*	368	
									807	808	809	834	X*	037	861	862	863	355	356	357	358	359	360	361				
									804	805	806	831	024	833	858	859	860	062	321	351	052	353	354	X*				
									016	802	803	828	022	023	855	856	857	058	059	315	316	060	318	061				

* Correspondente a plantas não mais existentes na Coleção.

Legenda	Procedências	Número de acessos	Grupos de acessos definidos para caracterização																
	Janaúba - MG	17	860	869	874	872	346	858	875	878	863	890	864	861	385	876	059	856	857
	Dourados - MS	14	825	807	900	991	911	337	393	908	932	955	905	910	824	820			
	Petrolina - PE	8	002	376	912	828	931	831	849	024									
	CNPS - MA	4	371	356	058	331													
	Lavras - Lote 000 - MG	4	399	332	005	330													
	Lavras - Lote 210 - MG	3	357	345	328														
	Rio das Flores - RJ	2	321	343															
	Rio Verde - GO	2	383	053															
	Barra dos Bugres - MT	2	003	344															
	Bom Jardim - RJ	2	355	386															
	Guapimirim - RJ	1	341																
	Jafba - MG	1	062																