Revista Ciência Agronômica, v. 53, e20207540, 2022
Centro de Ciências Agrárias - Universidade Federal do Ceará, Fortaleza, CE
www.ccarevista.ufc.br

Scientific Article
ISSN 1806-6690

Proposal of descriptors for the characterisation of *Jatropha curcas*[1]

Proposição de descritores para caracterização do pinhão-manso

Marcus Vinícius Magro dos Reis[2], Pedro Corrêa Damasceno Junior[3]*, Luiz Antônio dos Santos Dias[4], Rosemar Antoniassi[5], Tatiana de Oliveira Pinto[2], Antônio Carlos de Souza Abboud[3]

ASBSTRACT - Despite the importance of the physic nut (*Jatropha curcas*) in biodiesel production, knowledge of its genetic variability is still in the early stages, with the development of descriptors bring essential to achieving this goal. The aim of this study, therefore, was to propose and develop a list of standardised and tested descriptors, and to indicate those that make the greatest contribution to characterising the physic nut. Fifty-three morphological and reproductive descriptors of a quantitative, multi-category or binary nature were proposed. Descriptive estimates were obtained from among the descriptors. Using PCA, highly correlated descriptors were discarded. Then, the descriptors with the greatest relative contribution to estimating genetic diversity among the plants under analysis were identified. Reproductive descriptors had the highest coefficients of variation. Of 38 quantitative descriptors, 19 were sufficient to discriminate between the plants. This covered production, the crown of the plant, inflorescences, and fruit maturation. Descriptors related to grain yield, the uniformity of maturation and the crown showed the greatest relative contributions to diversity. Of the 13 multi-category descriptors, only five vegetative descriptors were discriminative. None of the proposed binary descriptors showed anyvariation.

Key words: Biodiesel. Germplasm. *Jatropha curcas*. Genetic gain.

RESUMO - Apesar da importância do pinhão-manso na produção de biodiesel, conhecimentos sobre a sua variabilidade genética ainda é incipiente. A idealização de descritores é essencial para se alcançar tal intento. Portanto, objetivou-se propor e desenvolver uma lista de descritores normatizados e testados, e indicar os de maior contribuição para atividades de caracterização em pinhão-manso. Propuseram-se 53 descritores morfológicos e reprodutivos de naturezas quantitativa, multicategórica e binária. Obtiveram-se estimativas descritivas entre os descritores. Procedeu-se, via PCA, o descarte dos descritores altamente correlacionados. Em seguida identificaram-se os descritores de maior contribuição relativa para a estimação da diversidade genética entre as plantas analisadas. Descritores reprodutivos apresentaram os maiores coeficientes de variação. De 38 descritores quantitativos, 19 foram suficientes para discriminar as plantas. Estes abrangeram a produção, a copa de planta, a inflorescência, e a maturação dos frutos. Descritores relacionados à produção de grãos, a uniformidade de maturação e a copa de planta apresentaram as maiores contribuições relativas para a diversidade. Dos 13 descritores multicategóricos, apenas cinco, vegetativos, foram discriminativos. Nenhum descritor binário proposto variou.

Palavras-chave: Biodiesel. Germoplasma. Jatropha curcas. Melhoramento genético.

DOI: 10.5935/1806-6690.20220026
Editor do artigo: Prof. Alek Sandro Dutra - alekdutra@ufc.br
*Author for correspondence
Received for publication in 18/08/2020; approved in 24/09/2021
[1]Parte da Dissertação do primeiro autor, apresentada ao Curso de Pós-Graduação em Fitotecnia - Universidade Federal Rural do Rio de Janeiro (UFRRJ)
[2]Departamento de Fitotecnia, Programa de Pós-graduação em Fitotecnia, Instituto de Agronomia, Universidade Federal Rural do Rio de Janeiro (UFRRJ), Seropédica-RJ, Brasil, marcus.ufrrj@gamail.com (ORCID ID 0000-0001-5505-0230), tatianaoliveira91@hotmail.com (ORCID ID 0000-0002-1039-9233)
[3]Departamento de Fitotecnia, Instituto de Agronomia, Universidade Federal Rural do Rio de Janeiro (UFRRJ), Seropédica-RJ, Brasil, damascenojunior2009@gmail.com (ORCID ID 0000-0001-8879-4850), abboud@ufrrj.br (ORCID ID 0000-0001-6054-071X)
[4]Departamento de Agronomia, Universidade Federal de Viçosa (UFV), Campus Universitário, Viçosa-MG, Brasil, lasdias@ufv.br (ORCID ID 0000-0002-1828-6938)
[5]Empresa Brasileira de Pesquisa Agropecuária, Embrapa Agroindústria de Alimentos, Rio de Janeiro-RJ, Brasil, rosemar.antoniassi@embrapa.br (ORCID ID 0000-0002-0795-9877)

INTRODUCTION

The physic nut (*Jatropha curcas* L.) is a monoecious species of family Euphorbiaceae distributed, according to Kumar and Tewary (2015), over a wide area of Central and South America. The species is characterised as perennial, shrub-like and fast-growing, with a high oil content in the seeds. Due to its potential for replacing fossil diesel (GUDETA, 2016), the physic nut has been tested as a source of raw material for biodiesel production (PRASADI *et al.*, 2012). Despite this, the productive potential of the crop is little known, and its germplasm is still little used. For Montes and Melchinger (2016), the physic nut is a non-domesticated species. According to Pereira *et al.* (2018), the lack of uniformity of the genotypes planted in Brazil seriously hampers expansion of the crop in the country. On a global scale, cultivation of the physic nut has been greatly affected by the scarcity of improved genotypes (EDRISI *et al.*, 2015).

Correct exploitation of the physic nut germplasm is mainly aimed at understanding the genetic variability that exists in the species. Without doubt, studies related to estimating genetic diversity resulting from the application of morphological and agronomic descriptors, substantially contribute to the planning of strategies that lead to maximising genetic gains for grain production in breeding programs.

One of the greatest obstacles to the efficient use of genotypes preserved in germplasm collections and germplasm banks, for example, are inadequate characterisation and evaluation. According to Bioversity International (2007), curators and breeders cannot efficiently exploit plant genetic resources when descriptors are omitted and/or described in an inappropriate and non-standardised way. In this regard, the development of descriptors that are easy to apply, efficient in their description of genotypes, and that include important agronomic aspects, is undoubtedly the greatest bottleneck in efficient exploitation of the genetic resources of the physic nut. According to Pinto *et al.* (2018), the oil content of the seeds of the physic nut may vary between genotypes, or even within genotypes during the production period. This fact shows the importance of efficient characterisation during the plant selection process.

Up to now, the literature has reported several proposals for characterising the physic nut (ALBUQUERQUE *et al.*, 2017; LAVIOLA *et al.*, 2011; OLIVEIRA *et al.*, 2016; PINTO *et al.*, 2018; SUNIL *et al.*, 2013). However, some bottlenecks are still found, such as descriptors that discriminate genotypes in terms of fruit maturation. It is also true that standardised and tested lists of descriptors allow researchers to expand the potential use of germplasm, as well as exchanges between institutions, and a proper comparison of research results published in the literature.

Given the above, the aim of the present study was to develop and propose a list of standardised quantitative, multi-category and binary descriptors for the physic nut, as recommended by Bioversity International (2007), which would consider the morphological and reproductive aspects of the plants. In addition, the aim was to discuss their developmentand importance to the crop, as well as their variation and relative contribution to estimating diversity. Finally, the aim was to decide which descriptors were most important in discriminating the genotypes, considering the plant population used.

MATERIAL AND METHODS

Location of the experiment and plant material

Data were obtained from 50 three-year-old plants, open-pollinated physic nut (*Jatropha curcas*) genotypes, randomly selected from the physic nut germplasm collection of the Department of Crop Science of the Federal Rural University of Rio de Janeiro, in Seropédica (22°45’ S; 43° 41’ W), Rio de Janeiro, Brazil.

Design, development and standardisation of the morpho-agronomic descriptors

Initially, the descriptors were developed based on the descriptor lists for soya (INTERNATIONAL BOARD FOR PLANT GENETIC RESOURCES, 1984), cotton (INTERNATIONAL BOARD FOR PLANT GENETIC RESOURCES, 1985) and castor bean (MILANI, 2008). In addition, specific descriptors for the physic nut available in the literature (LAVIOLA *et al.*, 2011) were also tested and, when necessary, adapted for inclusion in the list presented here. Moreover, a novel descriptor was proposed for the uniformity of fruit maturation (UFM).

Fifty-three descriptors were proposed, each capable of adaptation and application to the physic nut. These were defined to make up the preliminary list of descriptors for the species. The entire list was developed following the Guidelines for the Development of Crop Descriptor Lists (BIOVERSITY INTERNATIONAL, 2007). Each descriptor comprised three essential elements: a name, able to describe the attribute simply and concisely; a state, able to express the character of the observed attribute; and a clear and easy-to-understand method of measurement.

The evaluations were divided into two steps based on the stages of plant development, one for the application of descriptors referring to the vegetative aspects, and the other for the reproductive aspects. Based on this division, the descriptors were standardised and classified into quantitative, multi-category and binary variables. In all, 2,650 registrations were made, i.e. 53 descriptors were applied to 50 plants.

Statistical analysis

By means of the registrations, the mean value, variance, coefficient of variation, standard deviation, and minimum and maximum values were noted for each quantitative descriptor. The most discriminating quantitative descriptors were identified by principal component analysis, as per Jolliffe (1973). The variables referring to eigenvectors showing the greatest absolute coefficients associated with principal components with eigenvalues estimated up to 0.70, corresponded to those variables (descriptors) that contributed least to explain the total variance, and as suchare less discriminative. For this analysis, the data were normalised dividing the observed value by the standard deviation of the corresponding variable. After identifying the most discriminating descriptors, the relative importance of each was evaluated as per Singh (1981) in relation to its ability to contribute to an analysis of genetic diversity.

All the multi-category and binary descriptors were registered based on their respective percentages in each of the developed classes.

The statistical analysis was carried out using the Rv3.6.2 software (R DEVELOPMENT CORE TEAM, 2020).

RESULTS AND DISCUSSION

Design and development of the descriptors

Overall, of the 53 descriptors that were developed, 38 quantitative descriptors (19 vegetative and 19 reproductive), 13 multi-category (8 vegetative and 5 reproductive) and 2 binary (1 vegetative and 1 reproductive) were developed and tested (Tables 1 and 2, respectively).

All the quantitative descriptors were registered using their exact value. Of these, only the reproductive variable, number of locules per fruit (NLF), had a discrete quantitative character, and was therefore transformed into a multi-category variable (unilocular, bilocular, trilocular and tetra-locular) (Table 2). According to the Guidelines for the Development of Crop Descriptor Lists (BIOVERSITY INTERNATIONAL, 2007), quantitative variables should preferably be registered using their exact value; however, discrete quantitative variables are better suited for transformation than are continuous variables. Despite its discrete character, the number of lobes per leaf (NLL) was registered using its exact value due to its high variability (Table 1). Sunil *et al*. (2013) chose to define the same descriptor as a discrete quantitative, dividing it into three categories: 0 to 2 lobes, 3 to 5 lobes, and greater than 6 lobes.

For some of the quantitative descriptors, a minimum number of samples was defined for registration. Among the vegetative descriptors, the following were considered: length of the primary branches (LPB), NLL, internode length of the primary branches (ILPB), length of branches with an inflorescence (LBI), internode length of branches with an inflorescence (ILBI), leaf insertion angle (LIA), mean petiole length (MPL), leaf length (LL) and leaf width (LW) (Table 1). Among the reproductive descriptors, the minimum number of samples was considered for the length of the inflorescence peduncle (LIP), fruit length (FL), fruit width (FW), mean fruit weigth (MFW), seed length (SL) and seed width (SW) (Tables 2).

Each of the multi-category variables was registered with a maximum of five states (or classes). According to Bioversity International (2007), a high number of classes tends to affect the description criteria of the evaluator. For the binary variables, '0' was registered for the absence of a trait and '1' for its presence.

Table 1 - Preliminary list of vegetative descriptors developed for the physic nut (*Jatropha curcas*). The list includes the type of variable, descriptor name, methodology and, for multi-category and binary descriptors, the state. The list is shown in alphabetical order

Quantitative descriptors:
CPA - Crown projection area: evaluated from the end of the first reproductive cycle of the plant. The CPA (in square metres) is determined using the following estimator: $CPA=[(CPAn+ CPAs)\times(CPAe+CPAw)]/2$
CPAe – Eastern projection of the crown: evaluated from the end of the first reproductive cycle of the plant. The CPAe (in meters) is measured from the central region of the base of the stem to the maximum projection of the branch in an easterly direction.
CPAn - Northern projection of the crown: evaluated from the end of the first reproductive cycle of the plant. The CPAn (in meters) is measured from the central region of the base of the stem to the maximum projection of the branch in a northerly direction.

Continuation Table 1

CPAs - Southern projection of the crown: evaluated from the end of the first reproductive cycle of the plant. The CPAs (in meters) is measured from the central region of the base of the stem to the maximum projection of the branch in a southerly direction.
CPAw - Western projection of the crown: evaluated from the end of the first reproductive cycle of the plant. The CPAw (in meters) is measured from the central region of the base of the stem to the maximum projection of the branch in a westerly direction.
DPH - Height of developed plants: assessed from the end of the first reproductive cycle of the plant. The DPH (in centimetres) is measured from the base of the stem (ground) to the upper end of the tallest branch.
ILBI - Internode length of branches with an inflorescence: assessed from the first reproductive cycle of the plant. The ILBI (in centimetres) is determined from the mean length of the internodes measured on the three longest branches containing an inflorescence. Four internodes are measured on the median part of each branch with an inflorescence. For plants with erect architecture and no branches, the ILBI is given by the mean value of the four internodes on the median part of the stem, providing it contains an inflorescence.
ILPB - Length of the internode of primary branches: assessed from the first reproductive cycle of the plant. The ILPB (in centimetres) is determinedby the mean length of the internodes measured on the three longest primary branches on the plant. Six internodes are measured on the median part of each primary branch. For plants with erect architecture and no branches, the ILPB is given by the mean value of the measurements of six internodes on the middle part of the stem.
LBI - Length of branches with an inflorescence: evaluated from the first reproductive cycle of the plant. The LBI (in centimetres) is determined by the mean value of the length of the three longest branches containing an inflorescence, measured from the base of the branch close to the ground to its upper end. For plants with erect architecture and no branches, the LBI is given by the height of the plant, measured from the base of the stem close to the ground to the upper end of the single branch, providing the branch contains an inflorescence.
LIA - Leaf insertion angle: evaluated at any stage during the development cycle of the plant. The LIA is determined by the mean value in degrees of the measurements of 10 fully developed leaves, measured between the leaf insertion (petiole) and the stem.
LL - Leaf Length: assessed at any stage during the development cycle of the plant. LL (in centimetres) is determined from the mean length of 10 fully developed leaves per plant.
LLWR - Leaf length-width ratio: assessed at any stage during the development cycle of the plant. The LLWR is determined from the ratio of leaf length (LL) to leaf width (LW).
LPB - Length of the primary branches: Evaluated from the first reproductive cycle of the plant. The LPB (in centimetres) is determined from the mean length of the three longest primary branches, measured from the base of the branch close to the ground and its upper end. For plants with erect architecture and no branches, the LPB is given by the height of the plant, measured from the base of the stem close to the ground to its upper end.
LW - Leaf Width: assessed at any stage during the development cycle of the plant. The LW (in centimetres) is determined from the mean width of 10 fully developed leaves per plant.
MPL - Mean length of the petiole: assessed at any stage during the development cycle of the plant. The MPL (in centimetres) is determined from the mean value of the length of the petiole in 10 fully developed leaves per plant.
NBI - Number of branches with an inflorescence: evaluated from the first reproductive cycle of the plant. The NBI is determined by quantifying the total number of branches that have at least one inflorescence.
NLL - Number of lobes per leaf: assessed at any stage during the development cycle of the plant. The NLL is obtained by quantifying the lobes in 10 fully developed leaves per plant.
STD - Stem diameter: Assessed from the first reproductive cycle of the plant. The STD (in centimetres) is measured in the region of the collar at a maximum height of 10 cm from the base of the stem in the ground.
TNB - Total number of branches: assessed at any stage during the development cycle of the plant. The TNB is determined by quantifying the total number of branches on the plant.
Multi-category descriptors:
ARC - Plant architecture: observed from the end of the first reproductive cycle of the plant. ARC can assume the following states: 1- Upright; 2- Closed; 3- Open. Plants with Erect ARC have no branches in the lower third of the plant. Plants with Closed ARC have branches in the lower third of the plant.

Continuation Table 1

BRA - Type of branching: observed from the end of the first reproductive cycle of the plant. BRA can assume the following states: 0- No branching; 1- Monopodial; 2- Dichasial; 3- Sympodial.
BRC – Colour of the branches: observed from the end of the first reproductive cycle of the plant. The BRC is observed disregarding the presence of waxiness on the branch. BRC can assume the following states: 1- Green 2- Dark green; 3- Gray; 4- Purple; 5- Blue.
DLC - Colour of developed leaves: observed at any stage during the development cycle of the plant. The DLC is observed on the upper (adaxial) face of the leaf blade in a representative sample of 10 fully developed leaves per plant. DLC can assume the following states: 1- Light green; 2- Green; 3- Dark green; 4- Purple; 5- Red.
LVC – Colour of the leaf vein: observed at any stage during the development cycle of the plant. The LVC is observed on the upper surface (adaxial) of the leaf blade in a sample of 10 fully developed leaves per plant. LVC can assume the following states: 1- Green; 2- Purple; 3- Red.
PEC – Colour of the petiole: observed at any stage during the development cycle of the plant. The PEC is observed in a representative sample of 10 fully developed leaves per plant. PEC can assume the following states: 1- Green; 2- Greenish purple; 3- Purple; 4- Red.
STC – Colour of the stem: observed from the end of the first reproductive cycle of the plant. The STC is observed in the median region, ignoring the presence of wax. STC can assume the following states: 1- Green; 2- Dark green; 3- Grey; 4- Purple; 5- Blue.
YLC – Colour of the young leaves: observed at any stage during the development cycle of the plant. The YLC is observed on the upper (adaxial) face of the leaf blade in a representative sample of 10 young leaves per plant. YLC can assume the following states: 1- Light green; 2- Green; 3- Dark green; 4- Purple; 5- Red.
Binary descriptor:
WAX - Waxiness on the plant: observed from the end of the first reproductive cycle of the plant. The plant is classified according to the presence or absence of waxiness on the stem and other branches. WAX can assume the following states: 0- Absence; 1- Presence.

Table 2 - Preliminary list of reproductive descriptors developed for the psychic nut (*Jatropha curcas*). The list includes the type of variable, descriptor name, methodology and, for multi-category and binary descriptors, the state. The list is shown in alphabetical order

Quantitative descriptors:
DSW - Mean dry-seed weight per plant: evaluated from the first reproductive cycle of the plant. The DSW (in grams) is determined from the mean weight of all the seeds relative to the total number of seeds produced, weighed at a low moisture content of 12% after drying for approximately 72 hours at a temperature of 37 °C in an oven. The DSW can also be determined by the mean weight of 100 seeds, with 12% moisture, after drying for approximately 72 hours at a temperature of 37 °C in an oven.
FL – Fruit length: assessed from the first reproductive cycle of the plant. FL (in centimetres) is determined from the mean length of a minimum of 10 and a maximum of 50 yellow (ripe) fruit from different, randomly defined branches.
FLWR - Fruit length-width ratio: evaluated from the first reproductive cycle of the plant. The FLWR is determined from the ratio between the length of the fruit (FL) and the width of the fruit (FW).
FW - Fruit width: evaluated from the first reproductive cycle of the plant. The FW (in centimetres) is determined from the mean width of a minimum of 10 and a maximum of 50 yellow (ripe) fruit from different randomly defined branches on the plant.
GPP - Grain production per plant: evaluated from the first reproductive cycle of the plant. The GPP (in g plant-1) is obtained by multiplying the number of seeds per plant (NSP) by the weight of dry seeds (DSW) per plant.
HFI - Height of the first inflorescence: assessed as soon as the reproductive cycle of the plant begins. The HFI (in centimetres) is given by measuring from the base of the stem close to the ground to the lowest inflorescence (first inflorescence) on the plant.
LIP - Length of the inflorescence peduncle: evaluated from the first fruiting. The LIP (in centimetres) is determined from the mean length of the peduncle of 10 inflorescences per plant.
MFW - Mean fruit weight: evaluated from the first reproductive cycle of the plant. The MFW (in grams) is determined from the mean weight of a minimum of 10 and a maximum of 50 yellow fruit produced by the plant.

Continuation Table 2

NFP - Number of fruit per plant: evaluated from the first reproductive cycle of the plant. The NFP is determined by quantifying the total number of fruit produced by the plant.
NSF - Number of seeds per fruit: evaluated from the first reproductive cycle of the plant. The NSF is determined from the ratio between the number of seeds from the fruit used in measuring the NFP (number of fruit per plant) and the value of the NFP itself.
NSP - Number of seeds per plant: evaluated from the first reproductive cycle of the plant. The NSP is obtained by quantifying the total number of seeds produced by the plant.
OPP - Oil production per plant: evaluated from the first reproductive cycle of the plant. The OPP (in $g.plant^{-1}$) is obtained by multiplying the value for grain production per plant (GPP) by the seed oil content on a dry basis (SOC).
PFF – Period of fruit formation: evaluated from the first reproductive cycle of the plant. The PFF is obtained by quantifying the days elapsed between floral anthesis and the formation of the first completely yellow (ripe) fruit on the plant.
SL - Mean seed length: evaluated from the first reproductive cycle of the plant. The SL (in centimetres) is determined from the mean length of a minimum of 30 and a maximum of 50 seeds per plant.
SLWR - Seed length-width ratio: evaluated from the first reproductive cycle of the plant. The SLWR is obtained from the ratio between the length of the seed (SL) to the width of the seed (SW).
SOC - Seed oil content on a dry basis: evaluated from the first reproductive cycle of the plant. The seeds of fruit at the yellow stage are removed and dried until reaching 12% moisture. The oil is extracted in a Soxhlet extractor for 16 hours, using petroleum ether at 30ºC-60ºC as a solvent. The solvent is removed under a stream of nitrogen. The SOC (as a percentage) is obtained via lipid determination, based on method 945.38 of the Official Methods of Analysis of the Association of Official Analytical Chemists (2010).
SW - Seed Width: evaluated from the first reproductive cycle of the plant. The SW (in centimetres) is determined from the mean width of a minimum of 30 and a maximum of 50 seeds per plant.
TNI - Total number of inflorescences: evaluated from the first reproductive cycle. The TNI is obtained by quantifying (counting) all inflorescences on the plant throughout its reproductive cycle.
UFM – Uniformity of fruit maturation: evaluated from the first reproductive cycle of the plant. The UFM is determined by the mean total number of yellow (ripe) fruit in relation to the total number of fruit [green, greenish yellow (ripe) and black (dry)] counted per bunch per plant and expressed as a percentage.
Multi-category descriptors:
CSE – Colour of the seeds: observed at any stage during the development cycle of the plant. The SC should be observed in a representative sample of a minimum of 30 and a maximum of 50 seeds per plant. CSE should assume the following states: 1- Brown; 2- Black; 3- Purple; 4- Red.
FS – Shape of the fruit: observed from the first reproductive cycle of the plant. The FS should be observed in a minimum of 10 and a maximum of 50 yellow (ripe) fruit. FS should assume the following states: 1- Elliptical; 2- Oval; 3- Triangular.
NLF - Number of locules per fruit: evaluated from the first reproductive cycle of the plant. The NLF is observed in a minimum of 10 and a maximum of 50 yellow (ripe) fruit, and should assume the following states: 1- Unilocular; 2- Bilocular; 3- Trilocular; 4- Tetra-locular. The plant will assume the state that largely prevails after analysis.
SPT - Seed Pattern: observed at any stage during the development cycle of the plant. The SPT should be observed in a representative sample of a minimum of 30 and a maximum of 50 seeds per plant. SPT should assume the following states: 1- Single colour; 2- Brindled; 3- Spotted; 4- Speckled.
STX - Seed Texture: observed at any stage during the development cycle of the plant. The STX should be observed in a representative sample of a minimum of 30 and a maximum of 50 seeds per plant. The STX should assume the following states: 1- Smooth; 2- Rough; 3- Wrinkled.
Binary Descriptors:
CAR - Seed caruncle: observed at any stage during the development cycle of the plant. The CAR should be observed in a representative sample of a minimum of 30 and a maximum of 50 seeds per plant. CAR should assume the following states: 0- Absence; 1- Presence.

The descriptors related to the crown were adapted from Laviola *et al.* (2011), namely: northern (CPAn), southern (CPAs), eastern (CPAe) and western (CPAw) crown projection (Table 1). Based on these, it was proposed to estimate the area of the crown using the descriptor, crown projection area (CPA) (Table 1). From the internodes on the branches, it was found that for any one plant, the length of the primary branches was visibly different to that of the branches with inflorescences. It was therefore proposed to measure the branches using two descriptors: ILPB and ILBI, respectively (Table 1). It should be noted that the 'internode length' descriptor comes from the list of castor bean descriptors (MILANI, 2008). The descriptor 'number of branches with an inflorescence' (NBI) was adapted from the descriptor 'number of secondary branches' from Laviola *et al.* (2011), albeit considering branches with at least one inflorescence in the present study. The descriptors, height of developed plants (DPH), stem diameter (STD), total number of branches (TNB) and the leaf length-width ratio (LLWR) (Table 1), were also taken from Milani (2008).

The multi-category vegetative descriptors 'plant architecture' (ARC) and 'type of branching' (BRA) (Table 1) were adapted from the castor bean descriptors 'plant architecture' and 'branch type' from Milani (2008) and Milani *et al.* (2006), respectively. To facilitate the use of ARC in the physic nut, its states were described as 'upright', 'closed' and 'open'. Instead of 'semi-erect', we opted for 'closed'. The BRA descriptor had the categories, 'trifurcated', 'cup' and 'universal', replaced by 'no branching', 'monopodial', 'dichasial' and 'sympodial', common terms in botany. Sunil *et al.* (2013) also proposed multi-category descriptors to assess plant architecture in the physic nut. The vegetative descriptors, colour of the stem (STC), branches (BRC), young leaves (YLC), developed leaves (DLC), leaf vein (LVC) and petiole (PEC) (Table 1), were also registered as multi-category and taken from the list of soya (INTERNATIONAL BOARD FOR PLANT GENETIC RESOURCES, 1984) and cotton descriptors (INTERNATIONAL BOARD FOR PLANT GENETIC RESOURCES, 1985). For these, only the states related to the colour seen on the stem, leaves and petioles of the plants under analysis were changed. For the descriptor 'waxiness on the plant' (WAX) (Table 1) referring to the vegetative binary variable, the descriptor used for the castor bean was adopted (MILANI, 2008), recording only the presence or absence of waxiness.

Regarding the reproductive descriptors, it was decided to work with LIP as a quantitative variable (Table 2), whereas Laviola *et al.* (2011) considered the same variable as binary in the physic nut. Among the reproductive descriptors, the UFM descriptor (uniformity of fruit maturation) was proposed as original (Table 2). The lack of fruiting uniformity is one of the biggest problems in cultivating the physic nut, and can be considered one of the main limitations to the expansion of the crop in Brazil. It is worth noting that Albuquerque *et al.* (2017) mentioned a lack of fruiting synchrony in the physic nut, however, did not propose to quantify the trait.

The proposal of descriptors directly related to the production of grain and oil in the physic nut is of paramount importance for the selection of genotypes of high agronomic performance. As such, descriptors were proposed related to inflorescence (height of the first inflorescence-HFI and total number of inflorescences-TNI), fruiting (period of fruit formation-PFF, number of fruit per plant-NFP, mean fruit weight-MFW, FL, FW, fruit length-width ratio-FLWR, and number of seeds per fruit-NSF), and seeds (number of seeds per plant-NSP, mean dry-seed weight per plant-DSW, SL, SW and seed length-width ratio-SLWR) (Table 2).

In the present study, oil production per plant (OPP) is estimated from the product of grain production per plant (GPP) and the oil content of the seeds on a dry basis (SOC). In turn, GPP is estimated by the product of NSP and DSW (Table 2). It is important to note that standardisation of the DSW (Table 2) is of paramount importance, so that production (GPP and OPP) is not over- or underestimated. It is also worth pointing out that SOC must be estimated on a dry basis, so that comparisons including yield (SOC) and oil production (OPP) can safely be carried out. In order to determine the period taken for fruit production, the PFF descriptor (period of fruit formation) was proposed (Table 2). This was adapted from the 'planting cycle' descriptor used by Milani (2008) for the castor bean. Similarly, FL, FW, SL, SW, NFP, MFW, FLWR, NSF, NSP and SLWR) (Table 2) were taken from Milani (2008).

Among the multi-category reproductive descriptors (Table 2), both FS (fruit shape) and NLF were adapted from Laviola *et al.* (2011). Seeking greater practicality for FS, the originally used categories were changed: 'spherical ellipsoid', 'lanceolate ellipsoid' and 'ovoid ellipsoid' becoming 'elliptical', 'oval' and 'triangular', respectively. The colour (CSE), texture (STX) and pattern (SPT) of the seeds were taken from descriptor lists for soya (INTERNATIONAL BOARD FOR PLANT GENETIC RESOURCES, 1984) and cotton (INTERNATIONAL BOARD FOR PLANT GENETIC RESOURCES, 1985). Sunil *et al.* (2013) classified the surface of the physic nut seed as rough, smooth or shiny; in the present work, we chose to classify STX as smooth, rough or wrinkled.

The only reproductive binary descriptor, seed caruncle (CAR), was adapted from the castor bean descriptor 'caruncle type' (MILANI, 2008). However, due the lack of variability, CAR was registered as present or absent only.

Variation and relative importance of the descriptors

Of the 38 proposed quantitative descriptors, only nine had coefficients of variation greater than 30.0%, namely descriptors related to the crown projection (CPA) (Table 3), the uniformity of fruit maturation (UFM) (Table 4), the inflorescences (NBI-Table 3, LIP and TNI-Table 4) and production (OPP, GPP, NSP and NFP) (Table 4). Of these, maturation and production stand out with the highest coefficients. The lowest estimates for the coefficient of variation were for plant height (DPH), stem diameter (STD), leaves (LIA, LL, LW and LLWR) (Table 3), oil content of the seeds (SOC), and the morphological attributes of the fruit (MFW, FL, FW, FLWR, NSF, PFF) and seeds (DSW, SL, SW and SLWR) (Table 4).

Table 3 - Descriptive statistics of 19 quantitative vegetative variables measured in genotypes from the *Jatropha curcas* germplasm collection of UFRRJ

Variable	Mean	Minimum	Maximum	CV%	Variance	SD
DPH	246.18	210.00	305.00	9.10	502.15	22.41
CPA	4.42	1.89	9.15	33.80	2.23	1.49
CPAn	1.60	0.96	2.35	17.71	0.08	0.28
CPAs	1.38	0.80	2.00	20.03	0.07	0.28
CPAe	1.41	0.91	2.30	22.80	0.10	0.32
CPAw	1.35	0.81	2.31	24.58	0.11	0.33
STD	11.28	8.70	13.80	9.26	1.09	1.04
TNB	29.32	13.00	44.00	24.98	53.65	7.32
NBI	22.62	8.00	38.00	31.86	51.93	7.20
LPB	182.56	141.00	255.00	16.03	856.21	29.26
ILPB	1.47	1.00	2.00	17.12	0.06	0.25
LBI	74.52	45.00	124.00	19.78	217.17	14.74
ILBI	1.76	1.08	2.72	18.92	0.11	0.33
NLL	4.63	3.00	5.00	16.84	0.61	0.78
LIA	133.20	112.00	153.00	6.48	74.43	8.63
MPL	14.57	8.30	19.60	14.70	4.59	2.14
LL	12.74	8.80	14.40	7.53	0.92	0.96
LW	12.96	8.60	15.40	8.49	1.21	1.10
LLWR	0.98	0.86	1.10	4.36	0.01	0.04

Legend: CV%-Coefficient of phenotypic variation (%); SD-Standard Deviation; DPH-Height of developed plants (cm); CPA-Crown projection area (m^2); CPAn-Northern projection of the crown (m); CPAs-Southern projection of the crown (m); CPAe–Eastern projection of the crown (m); CPAw-Western projection of the crown (m); STD-Stem diameter (cm); TNB-Total number of branches; NBI-Number of branches with an inflorescence; LPB-Length of the primary branches (cm); ILPB-Internode length of the primary branches (cm); LBI-Length of branches with an inflorescence (cm); ILBI-Internode length of branches with an inflorescence (cm); NLL-Number of lobes per leaf ; LIA-leaf insertion angle (degrees); MPL-Mean petiole length (cm); LL-Leaf length (cm); LW-Leaf width (cm); LLWR-Leaf length-width ratio

Table 4 - Descriptive statistics of 17 quantitative reproductive variables measured in genotypes from the *Jatropha curcas* germplasm collection of UFRRJ

Variable	Mean	Minimum	Maximum	CV%	Variance	SD
HFI	92.63	30.00	145.00	27.19	634.37	25.19
LIP	2.49	1.50	7.40	30.76	0.58	0.76
TNI	123.05	19.00	353.00	38.62	2.258.87	47.53
UFM	47.63	8.00	100.00	52.82	633.15	25.16
PFF	81.68	74.00	105.00	6.95	32.26	5.68

Continuation Table 4

NFP	266.52	11.00	756.00	58.97	24.698.42	157.16
MFW	12.49	9.29	14.62	7.31	0.83	0.91
FL	3.09	2.86	3.75	4.31	0.02	0.13
FW	2.70	2.53	2.93	3.43	0.01	0.09
FLWR	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
NSP	676.08	30.00	1.973.00	58.23	154.965.40	393.66
DSW	0.73	0.64	0.79	5.17	0.00	0.04
SL	1.79	1.08	1.95	6.99	0.01	0.13
SW	1.09	0.92	1.15	4.70	0.00	0.04
SLWR	1.65	1.57	1.99	4.21	0.00	0.07
GPP	487.02	23.85	1.402.04	59.10	82.841.16	287.82
SOC	37.42	26.60	40.35	6.12	5.24	2.30
OPP	210.36	71.32	561.35	46.69	9.647.34	98.22

Legend: CV%-Coefficient of phenotypic variation (%); DP-Standard Deviation; HFI-Height of the first inflorescence (cm); LIP-length of the inflorescence peduncle (cm); TNI-Total number of inflorescences; UFM-Uniformity of fruit maturation (%); PFF-Length of fruit formation (days); NFP-Number of fruit per plant; MFW-Mean fruit weight (g); FL-Fruit length (cm); FW-Fruit width (cm); FLWR-Fruit length-width ratio; NSF-Number of seeds per fruit; NSP-Number of seeds per plant; DSW-Dry-seed weight per plant (g); SL-Seed length (cm); SW-Seed width (cm); SLWR-Seed length-width ratio; GPP-Grain production per plant (g.plant-1); SOC-Seed oil content (%); OPP-Oil production per plant (g.plant-1)

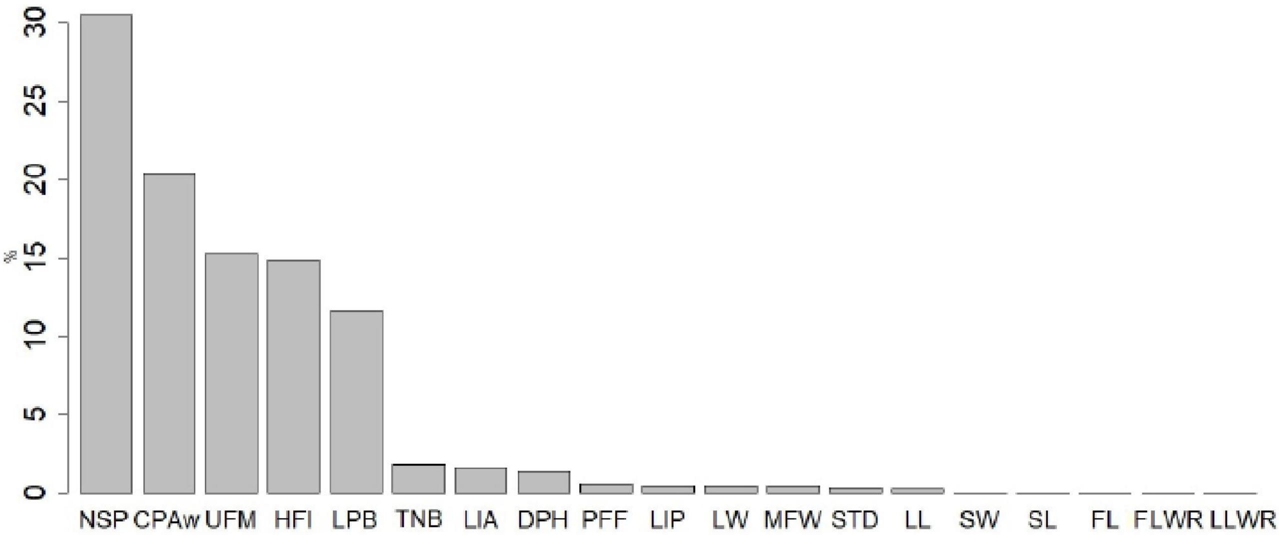
For the multi-category descriptors, only five out of the proposed 13 (38.46%) showed any variation, namely: ARC (16% of the plants closed, 85% open), BRA (8% of the plants monopodial, 8% dichasial, 84% sympodial), YLC (8% of the plants light green, 56% green, 36% purple), DLC (2% of the plants light green, 14% green, 84% dark green) and PEC (82% of the plants green, 18% greenish purple). The other multi-category descriptors showed no variation: STC (100% of the plants grey), BRC (100% of the plants grey), LVC (100% of the plants green) (vegetative descriptors); FS (100% of the plants elliptical), NLF (100% of the plants trilocular), CSE (100% of the plants black), STX (100% of the plants smooth) and SPT (100% of the plants single colour) (reproductive descriptors). For the binary descriptors proposed here, none showed any variation: WAX (present in 100% of the plants) and CAR (present in 100% of the plants). From the above it can be seen that only the vegetative descriptors showed any variation.

Of the 38 proposed quantitative descriptors, 19 (50%) were sufficient to discriminate the genotypes under analysis as per the methodology proposed by Jolliffe (1973). These included different parts of the plant, such as the crown (DPH and CPAw), branches (STD, TNB and LPB), inflorescences (HFI and LIP), fruit (MFW, FL, FLWR, UFM and PFF), seeds (SL and SW), leaf (LW, LL, LLWR and LIA), and grain production, represented by NSP. In addition to the variation in the descriptor, aspects related to the correlation between variables are also considered in the proposal by Jolliffe (1973). Among the descriptors cited as the most discriminating, nine were vegetative and 10 were reproductive.

Based on Figure 1, of the 19 most discriminating descriptors, NSP (30.46%), CPAw (20.40%), UFM (15.25%), HFI (14.90%) and LPB (11.64%) stood out for their relative importance in discriminating the genotypes. These represent variables related to grain production (NSP and HFI), aspects related to the crown (CPAw and LPB), and fruit maturation (UFM). It is worth mentioning that the relevance of the descriptors depends on the population under analysis.

Considering only Brazilian genotypes, Oliveira *et al.* (2016) reported that grain production for the crop showed a high rate of phenotypic plasticity. Laviola *et al.* (2011) argued that grain production was among the variables with the greatest contribution to estimating diversity in the physic nut. For production, estimates based on SOC and OPP are also extremely important in the physic nut. According to Pinto *et al.* (2018), seed weight in the physic nut does not show a high correlation with oil yield. In the present study, the SOC ranged from 26.60% to 40.35%, with a mean value of 37.42% (Table 4). Freitas *et al.* (2016) found an oil yield of between 30% to 39.60% in Brazilian germplasm. Whereas Pinto *et al.* (2018) and Singh *et al.* (2016) found a maximum seed oil concentration of 43.89% and 37.49%, respectively. In the present study, oil production per plant (OPP) ranged between 71.32 and 561.35 g (Table 4). Singh *et al.* (2016) found a variation of between 70.0 to 470.0 grams of oil per plant.

Figure 1 - Relative importance of 19 quantitative descriptors applied in characterising genotypes of the physic nut



Legend: NSP: Number of seeds per plant; CPAw: Western projection of the crown; UFM: Uniformity of fruit maturation; HFI: Height of the first inflorescence; LPB: Length of the primary branches; TNB: Total number of branches; LIA: Leaf insertion angle; DPH: Height of developed plants; PFF: Period of fruit formation; LIP: Length of the inflorescence peduncle; LW: Leaf width; MFW: Mean fruit weight; STD: Stem diameter; LL: Leaf length; SW: Seed width; SL: Seed length; FL: Fruit length; FLWR: fruit length-width ratio; LLWR: leaf length-width ratio

Singh *et al.* (2016) also obtained high estimates of phenotypic variance for the crown and seed production per plant in the physic nut. These results corroborate the present study, where the crown projection area (CPA) varied considerably (1.89 to 9.15 m^2) (Table 3). Basu, Gunupuru and Sahoo (2017) found low variation in the crown in Indian genotypes. Oliveira *et al.* (2016) also pointed out reduced phenotypic plasticity for this trait. The contradictory results show that characterising the crown of plants of the physic nut still needs to be improved, so that any existing variability can be satisfactorily captured.

As mentioned above, fruit maturation proved to be a discriminative variable of high relative importance. Albuquerque *et al.* (2017) reported that all the genotypes under analysis showed uneven fruiting. However, in the present study, the UFM showed a high variation (CV = 52.82%), with plants showing a UFM of between 8% to 100% of ripe fruit on each bunch (Table 4). These results demonstrate the possibility of successfully selecting physic nut plants for uniform fruit maturation per bunch.

CONCLUSIONS

1. Of the 38 proposed quantitative descriptors, 19 were sufficient to discriminate the genotypes: DPH and CPAw (crown projection), STD, TNB and LPB (branches), HFI and LIP (inflorescences), MFW, FL, FLWR, UFM and PFF (fruit and their maturation), SL and SW (seeds), LW, LL, LLWR and LIA (leaves) and NSP (production);
2. Descriptors related to grain yield (NSP and HFI), the crown (CPAw and LPB) and fruit maturation (UFM) showed the greatest relative contributions to estimating genetic diversity;
3. Only five out of 13 of the proposed multi-category descriptors (ARC, BRA, DLC, PEC and YLC) were effective in discriminating genotypes. All were vegetative;
4. Neither of the proposed binary descriptors (WAX and CAR) was able to detect any variation between the genotypes.

ACKNOWLEDGEMENTS

The authors would like to thank PETROBRAS Biocombustíveis and FAPERJ for their financial assistance, and CNPq for their grant of a scholarship.

REFERENCES

ALBUQUERQUE, N. *et al*. Characterization of *Jatropha curcas* accessions based in plant growth characteristics and oil quality. **Industrial Crops & Products**, v. 109, p. 693-698, 2017.

ASSOCIATION OF OFFICIAL ANALYTICAL CHEMISTS. **Official methods of analysis of AOAC International**. 18. ed. Maryland: AOAC International, 2010. 1298p.

BASU, A.; GUNUPURU, L. R.; SAHOO, L. Morphometric characterization of *Jatropha curcas* germplasm of North-East

India. **African Journal of Biotechnology**, v. 16, p. 648-656, 2017.

BIOVERSITY INTERNATIONAL. **Guidelines for the development of crop descriptor lists**. Rome: Bioversity International, 2007. 72p. (Bioversity Technical Bulletin Series).

EDRISI, S. A. *et al*. Jatropha curcas L.: a crucified plant waiting for resurgence. **Renewable and Sustainable Energy Reviews**, v. 41, p. 855-862, 2015.

FREITAS, R. G. *et al*. Diversity and genetic parameter estimates for yield and its components in *Jatropha curcas* L.. **Genetics and Molecular Research**, v. 15, p. 1-10, 2016.

GUDETA, T. B.Chemical composition, bio-diesel potential and uses of *Jatropha curcas* L. (Euphorbiaceae). **American Journal of Agriculture and Forestry**, v. 4, n. 2, p. 35-48, 2016.

INTERNATIONAL BOARD FOR PLANT GENETIC RESOURCES. **Cotton descriptors (Revised)**. Rome: IBPGR,1985. 25p.

INTERNATIONAL BOARD FOR PLANT GENETIC RESOURCES. **Descriptors for soybean**. Rome: IBPGR, 1984. 38p.

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

KUMAR, A.; TEWARY, S. K. Origin, distribution, ethnobotany and pharmacology of *Jatropha curcas*. **Research Journal of Medicinal Plant**, v. 9, n. 2, p. 48-59, 2015.

LAVIOLA, B. G. *et al.* Caracterização morfo-agronômica do banco de germoplasma de pinhão-manso na fase jovem. **Bioscience Journal**, v. 27, n. 3, p. 371-379, 2011.

MILANI, M. **Descritores de mamona utilizados pela Embrapa Algodão**.Campina Grande, PB: Embrapa Algodão, 2008. 39p. (Documentos, 192).

MILANI, M. *et al.* **Caracterização taxonômica de acessos de mamona (*Ricinus communis* L.) do banco ativo de germoplasma da Embrapa Algodão**. Campina Grande, PB: Embrapa Algodão, 2006. 17p. (Boletim de pesquisa e desenvolvimento, 67).

MONTES, J. M.; MELCHINGER, A. E. Domestication and breeding of *Jatropha curcas* L. **Trends in Plant Science**, v. 21, n. 12, p. 1045-1057, 2016.

OLIVEIRA, J. P. de M. *et al*. Phenotypic characterization of physic nut populations. **African Journal of Agricultural Research**, v. 11, n. 45, p. 4559-4566, 2016.

PEREIRA, I. R. *et al*. Trends and gaps in the global scientific literature about *Jatropha curcas* L. (Euphorbiaceae), a tropical plant of economic importance. **Semina: Ciências Agrárias**, v. 39, n. 1, p. 7-18, 2018.

PINTO, M. S. *et al*. Diversity between *Jatropha curcas* L. accessions based on oil characteristics and X-ray digital images analysis from it seeds. **Crop Breeding and Applied Biotechnology**, v. 18, p.292-300, 2018.

PRASADI, L. *et al*. Experimental assessment of toxic phorbol ester in oil, biodiesel and seed cake of Jatropha curcas and use of biodiesel in diesel engine. **Applied Energy**, v. 93, p. 245-250, 2012.

R DEVELOPMENTCORE TEAM. **R**: a language and environment for statistical computing. Viena, Austria: R Foundation for Statistical Computing, 2020. Available at: https://www.R-project.orf/. Accessed: May 29, 2020

SINGH, D. The relative importance of characters affecting genetic divergence. **Indian Journal of Genetic and Plant Breeding**, v. 41, p. 237-245, 1981.

SINGH, S. *et al*. Genetic variability, character association and divergence studies in *Jatropha curcas* for improvement in oil yield. **Trees**, v. 30, 2016.

SUNIL, N. *et al*. Minimal descriptors for characterization avaluation of *Jatropha curcas* L. germplasm for utilization in crop improvement. **Biomass and Bioenergy**, v. 48, p. 239-249, 2013.